The extremal collapse threshold and the third law of black hole thermodynamics

Ryan Unger

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Mathematics
Adviser: Mihalis Dafermos

May 2024

Abstract

In this dissertation, we investigate extremal black holes in general relativity. Extremal black holes are exceptional solutions of Einstein's equations which have absolute zero temperature in the celebrated thermodynamic analogy of black hole mechanics.

Our first main result is a definitive disproof of the "third law of black hole thermodynamics." We construct examples of black hole formation from regular, one-ended asymptotically flat Cauchy data for the Einstein–Maxwell-charged scalar field system which are exactly isometric to extremal Reissner–Nordström after a finite advanced time along the event horizon. Moreover, in each of these examples the apparent horizon of the black hole coincides with that of a Schwarzschild solution at earlier advanced times. We also prove similar black hole formation results for very slowly rotating Kerr black holes in vacuum.

Our second main result is a proof that extremal black holes arise on the threshold of gravitational collapse. More precisely, we construct smooth one-parameter families of smooth, spherically symmetric solutions to the Einstein–Maxwell–Vlasov system which interpolate between dispersion and collapse and for which the critical solution is an extremal Reissner–Nordström black hole. We call this critical phenomenon extremal critical collapse and the present work constitutes the first rigorous result on the black hole formation threshold in general relativity.

The above mentioned results constitute Part I of this dissertation and were all obtained in joint work with Christoph Kehle.

In Part II of this dissertation, we study extensions of the celebrated positive mass theorem to a very general class of initial data, including extremal black holes. These results were obtained in collaboration with Dan A. Lee, Martin Lesourd, and Shing-Tung Yau. We provide a resolution of the spacetime positive mass theorem on manifolds with boundary, a resolution of the remaining cases of Schoen and Yau's Liouville conjecture for locally conformally flat manifolds, and demonstrate a novel scalar curvature shielding phenomenon for the ADM mass.

Acknowledgements

I am indebted to a great number of people who have helped me on my journey to completing this dissertation. It is not possible to do everyone justice and I apologize in advance for the many inevitable omissions.

The first person I would like to thank is Alex Freire. I met Alex on the first day of my undergraduate studies at the University of Tennessee and after talking to him for just a few moments it was clear that I should become a mathematician. Theodora Bourni and Mat Langford joined UT in my junior year and nothing in this dissertation would have been possible without the countless hours spent discussing geometry and analysis with them. I am also greatly indebted to Cristian Batista, Jochen Denzler, Maik Lang, Tadele Mengesha, Tim Schulze, Carl Sundberg, Morwen Thistlethwaite, and Stella Thistlethwaite for all their teaching and support during my years at Tennessee.

I would like to express my sincere gratitude towards my PhD advisor, Mihalis Dafermos. Mihalis has an incredibly infectious enthusiasm for mathematics and a strong aesthetic sense that has definitively shaped my identity as a scientist. Aside from being a mere mathematical advisor, Mihalis has been an invaluable source of support, wisdom, and friendship. He has also taught me much about Greek and Turkish culture and has been an exceptional tour guide. The many conversations we had during my time in Princeton and Cambridge have been truly invaluable and I am lucky to have been his student.

I would like to thank Igor Rodnianski for many fascinating and inspiring conversations over the years, Frans Pretorius for his interest in my work and many helpful conversations, and Alice Chang and Fernando Codá Marques for their support during my time at Princeton. I am furthermore grateful to Frans, Igor, and Mihalis for agreeing to be on my dissertation committee.

I owe many thanks to my collaborators Christoph Kehle, Dan Lee, Martin Lesourd, and Shing-Tung Yau for all they have taught me. I wish to especially thank Christoph for encouraging me to pursue general relativity, his relentless and invaluable optimism, and being such a reliable friend and collaborator, and Martin for keeping me sane during the pandemic and for being with me every step of the way.

I spent the fourth year of my PhD at the University of Cambridge. I am very grateful to Harvey Reall, Claude Warnick, and István Kádár for many interesting and insightful conversations, to DPMMS for its hospitality, and to King's College for exceptionally warm accommodation.

I would like to thank Maxime van de Moortel for his friendship, carrot cake, and kindly agreeing to be a reader of my dissertation. Elena Giorgi, Dan Lee, Jonathan Luk, and Sung-Jin Oh deserve special thanks for their invaluable assistance with my job search and I apologize in advance to Jonathan and Sung-Jin for being saddled with me for the next four years. The work of Georgios Moschidis has been a great inspiration for me and I am truly thankful for all the conversations we have had over the years. I would also like to thank Rita Teixeira da Costa for her sage advice.

I was very lucky to befriend Thomas Massoni and Érico Melo Silva in graduate school. I wish to thank them for the countless conversations we had about math, life, food, tea, coffee, video games, books, and pens. I especially wish to thank Thomas for his constant support and bearing the brunt of my complaints not directed at Christoph. I was also very lucky to have had great housemates and grilling partners Douglas Stryker, Rohil Prasad, and Vikram Giri during the pandemic and Lorenzo Sarnataro in the watermelon house.

As a student of a Cretan, graduate school would not have been complete without the preparation and consumption of large quantities of meat. I am grateful to John Anderson and Federico Pasqualotto for their friendship. John taught me how to grill (a truly life-changing skill) and Federico's grandmother has provided countless delicious mozzarella sticks. I am indebted to the very exclusive butcher shops Costco Wholesale in Lawrence Township and Al Amin in Cambridge for providing high quality and affordable steak and lamb. I am grateful to Grigorios Fournodavlos for his help in making Easter 2022 an unforgettable experience, Warren Li for the gaiwans, and Igor Medvedev for starting a new tradition with his desserts. I am indebted to Nikiforos Aerakis, Gialaftis, Zeki Müren, and many others for providing the soundtrack of my PhD.

I would also like to acknowledge Yannis Angelopoulos, Marios Apetroaie, Gabriele Benomio, David Bick, Serban Cicortas, Lili Fehértói-Nagy, Dimitris Fetsios, Fernando Figueroa Zamora, Dejan Gajic, André Guerra, Otte Heinävaara, Jeff Jauregui, Leonhard Kehrberger, Noah Kravitz, Hamed Masaood, Georgios Mavrogiannis, Paul Minter, Alex Mramor, Haydee Pacheco, Lindsey Reinhard, Seung-Yeon Ryoo, Shikhin Sethi, Jaydeep Singh, Anna Skorobogatova, Andreas Stavrou, Martin Taylor, Giuseppe Tinaglia, Tin-Yau Tsang, Stefan Tudose, Tuomas Tuukkanen, Renato Velozo Ruiz, Katy Woo, and Nina Zubrilina for all the conversations and good times we shared.

I would like to express thanks to Jill LeClair and Ankit Tak for making my time in Fine Hall go very smoothly. I am very grateful to Princeton University for awarding me the Porter Ogden Jacobus Fellowship in my final year.

Finally, I would like to thank my parents and family for all of their support and patience during my studies. This dissertation is dedicated to my grandmother Georgia Strednansky, an amazing woman who passed away during my PhD.

To Babi.

Contents

Part I

The extremal collapse threshold and the third law

Chapter 1

Overview of Part I

One of the most spectacular predictions of general relativity is the existence and formation of *black holes*, which are regions of spacetime where gravity is so strong that not even light can escape from within. Matter and energy, evolving under the *Einstein field equations* [Ein15]

$$\mathrm{Ric}(g) - \tfrac{1}{2}R(g)g = 2\mathbf{T}, \tag{1.0.1}$$

can undergo *gravitational collapse*, becoming so dense that a black hole forms dynamically. The Einstein equations relate the geometry of *spacetime*, a (3+1)-dimensional Lorentzian manifold $(\mathcal{M}^4, g)$, to its matter content, which is represented by the *energy-momentum tensor* $\mathbf{T}$.

The first solution of (1.0.1) containing a black hole was discovered by Schwarzschild [Sch16] almost immediately after Einstein's formulation of the field equations. Building on earlier work of Lemaître [Lem33], Oppenheimer and Snyder [OS39] produced the first actual example of gravitational collapse by showing that a homogeneous dust cloud can collapse to form a Schwarzschild black hole in finite time. However, the significance of this work was not understood for another 20 years, until the celebrated work of Penrose [Pen65]. Now black holes are one of the central objects of study in astrophysics and high energy physics.

The most important explicit black hole solutions of the Einstein equations are the Reissner–Nordström and Kerr families. Each of these black holes is characterized by a mass M and either a charge $|e| \le M$ for Reissner–Nordström or a specific angular momentum $|a| \le M$ for Kerr. As with every black hole, the event horizon $\mathcal{H}^+$ of a Reisser–Nordström or Kerr black hole is a null hypersurface. For each member of these families, there is a Killing vector field K normal and tangent to $\mathcal{H}^+$—the event horizons are *Killing horizons*. This Killing field satisfies

$$\nabla_K K = \kappa K$$

along $\mathcal{H}^+$, where $\kappa \geq 0$ is a constant known as the *surface gravity*, which can be explicitly calculated in terms of M and e or a. If $|e| < M$ or $|a| < M$, $\kappa > 0$ and the black hole is called *subextremal*. If $|e| = M$ or $|a| = M$, $\kappa = 0$ and the black hole is called *extremal*. As we will see, extremal black holes have exceptional properties and play a fundamental role in the structure of the moduli space of solutions of the Einstein equations.

Part I of this dissertation describes the beginning of a research program, which is joint work with Christoph Kehle, to study the dynamical formation of extremal black holes and develop a picture of phase space around extremal black holes in gravitational collapse. More precisely, we pose and solve the following three problems:

1. We prove that a subextremal black hole can become extremal in finite time in the gravitational collapse of charged matter, which definitively disproves the so-called *third law of black hole thermodynamics*. See already Section 1.1. This result was unexpected because the third law was widely believed to be true, with a supposed proof by Israel [Isr86] and several other supporting numerical and heuristic studies in the literature.

2. We prove that any sufficiently slowly rotating Kerr (including Schwarzschild) black hole can form in vacuum gravitational collapse in finite time. See already Section 1.1.5. We hope that the techniques developed here will allow us to prove that extremal Kerr can form in vacuum gravitational collapse and disprove the third law *in vacuum*.

3. We prove that there exist extremal black holes on the threshold between collapsing and dispersing charged matter, without the use of infinitesimally thin shells or other singular matter. See already Section 1.2. This is a novel critical phenomenon which we call *extremal critical collapse*.

1.1 The third law of black hole thermodynamics

1.1.1 Retiring the third law

Following pioneering work of Christodoulou [Chr70] and Hawking [Haw71] on energy extraction from rotating black holes and Bekenstein's proposal of a black hole entropy [Bek72], Bardeen, Carter, and

Hawking [BCH73] proposed—via analogy to classical thermodynamics—the celebrated *four laws of black hole thermodynamics*. The analogy asserts that the entropy S of a black hole is proportional to the *surface area* A of the horizon and the temperature T is proportional to the *surface gravity* κ of the horizon. While these identifications were later vindicated by the discovery of Hawking radiation [Haw75], the laws proposed by Bardeen–Carter–Hawking are a set of mathematical statements about classical general relativity.

Law	Classical thermodynamics	Black hole dynamics
Zeroth	T constant in equilibrium	κ constant on stationary horizon
First	$dE = TdS + \cdots$	$dM = \kappa dA + \cdots$
Second	$dS \geq 0$	$dA \geq 0$
Third	$T \not\to 0$ in finite process	$\kappa \not\to 0$ in finite advanced time

Table 1.1: The four laws of black hole thermodynamics. Extremal black holes have *absolute zero temperature* in this analogy. These "laws" are to be thought of as conjectures, and laws 0, 1, and 2 have been proved in the form stated here [Haw72a; BCH73]. In the physics literature, these "laws" (sometimes suitably modified) are interpreted as fundamental meta-theorems which are to be true in any reasonable physical theory.

In analogy to Nernst's "unattainability law" in classical thermodynamics, we have:

Conjecture (The third law of black hole thermodynamics)**.** *A subextremal black hole cannot become extremal in finite time by any continuous process, no matter how idealized, in which the spacetime and matter fields remain regular and obey the weak energy condition.*

This version is distilled from the literature, particularly from the work of Israel [Isr86; Isr92] who added explicit mention of regularity and the weak energy condition to avoid previously known examples [DI67; Kuc68; Bou73; FH79; SI80; Pró83] which would otherwise violate the third law. In this dissertation (taken from the work [KU22]), we show that the third law is fundamentally flawed in a manner that does not appear to be salvageable by further reformulation. Indeed, we construct counterexamples in the Einstein–Maxwell-charged scalar field model in spherical symmetry, a model which satisfies the dominant energy condition, arising from arbitrarily regular initial data on a one-ended asymptotically flat hypersurface.

Theorem 1.1.1. *Subextremal black holes can become extremal in finite time, evolving from regular initial data. In fact, there exist regular one-ended Cauchy data for the Einstein–Maxwell-charged scalar field system which undergo gravitational collapse and form an exactly Schwarzschild apparent horizon, only for the spacetime to form an exactly extremal Reissner–Nordström event horizon at a later advanced time.*

In particular, the "third law of black hole thermodynamics" is false.

For the more precise version of the theorem, see Theorem 1.1.11 below. For a Penrose diagram of the counterexample, see Fig. 1.1 below.

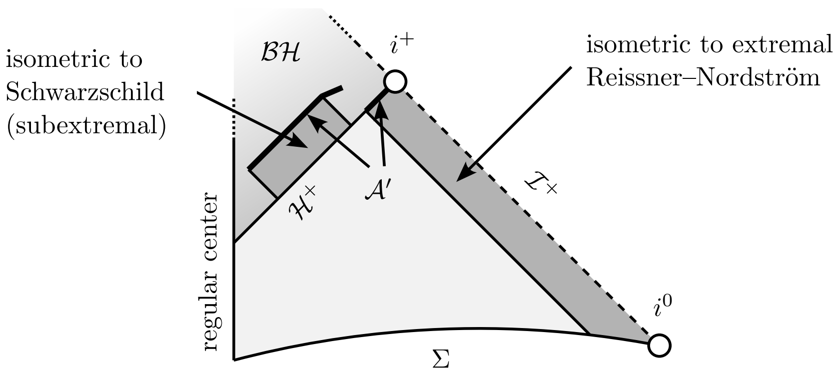

Figure 1.1: Penrose diagram of our counterexample to the third law arising from regular initial data on Σ. The northwest edge of the Schwarzschild region is exactly isometric to a section of the $r = 2M$ hypersurface in Schwarzschild. The outermost apparent horizon $\mathcal{A}'$ is initially indistinguishable from Schwarzschild and then jumps out in finite time to be exactly isometric to the event horizon of extremal Reissner–Nordström. For speculations about the future boundary of the interior, see already Section 1.1.4.1. The behavior of our solutions can be modified to be subextremal near i^0, see already Remark 1.1.2.

Our result also clarifies some issues raised by Israel in [Isr86; Isr92] who seemingly associated a disconnected outermost apparent horizon with a severe lack of regularity of the spacetime metric and/or matter fields. We stress that our examples are regular despite the disconnectedness of the apparent horizon. We note moreover that Israel seemed to associate extremization with the black hole "losing its trapped surfaces." This confusion appears to be related to his implicit assumption that the apparent horizon is connected. Since the Einstein–Maxwell-charged scalar field matter manifestly obeys the dominant energy condition, trapped surfaces are not lost in any sense, nonetheless, the black hole becomes extremal in finite time. In the examples we construct, there exists an open set of trapped spheres inside the black hole region, which persist for all advanced time until they encounter the Cauchy horizon or a curvature singularity inside the black hole. However, there is a neighborhood of the event horizon which does not contain any (strictly) trapped surfaces. For an extended discussion of these issues, see already Section 1.1.3.

Remark 1.1.2. Note that in discussions of the third law, the focus is typically on dynamics near the event horizon and apparent horizon, in late advanced time. Our counterexamples depicted in Fig. 1.1 are isometric to extremal Reissner–Nordström for all sufficiently late advanced times and all retarded times to the past of the event horizon, in particular near spatial infinity i^0. However, by using a scattering argument as in [Keh22b], one can easily modify our examples so as to be subextremal in a neighborhood of i^0, if desired.

The Einstein–Maxwell-charged scalar field (EMCSF) system featured in Theorem 1.1.1 reads

$$R_{\mu\nu}(g) - \tfrac{1}{2}R(g)g_{\mu\nu} = 2\left(T^{\mathrm{EM}}_{\mu\nu} + T^{\mathrm{CSF}}_{\mu\nu}\right), \tag{1.1.1}$$

$$\nabla^\mu F_{\mu\nu} = 2\mathfrak{e}\,\mathrm{Im}(\phi\overline{D_\nu\phi}), \tag{1.1.2}$$

$$g^{\mu\nu}D_\mu D_\nu\phi = 0, \tag{1.1.3}$$

for a quintuplet $(\mathcal{M}, g, F, A, \phi)$, where $(\mathcal{M}, g)$ is a (3+1)-dimensional Lorentzian manifold, ϕ is a complex-valued scalar field, A is a real-valued 1-form, $F = dA$ is a real-valued 2-form, $D = d + i\mathfrak{e}A$ is the gauge covariant derivative, $\mathfrak{e} \in \mathbb{R}\setminus\{0\}$ is a fixed coupling constant representing the charge of the scalar field, and the energy momentum tensors are defined by

$$T^{\mathrm{EM}}_{\mu\nu} \doteq g^{\alpha\beta}F_{\alpha\nu}F_{\beta\mu} - \tfrac{1}{4}F^{\alpha\beta}F_{\alpha\beta}g_{\mu\nu}, \tag{1.1.4}$$

$$T^{\mathrm{CSF}}_{\mu\nu} \doteq \mathrm{Re}(D_\mu\phi\overline{D_\nu\phi}) - \tfrac{1}{2}g_{\mu\nu}g^{\alpha\beta}D_\alpha\phi\overline{D_\beta\phi}. \tag{1.1.5}$$

We refer to Section 2.2 for the form of the EMCSF system in spherical symmetry.

Remark 1.1.3. Theorem 1.1.1 also holds for the *Einstein–Maxwell–charged Klein–Gordon* system in which the wave equation (1.1.3) is replaced by the Klein–Gordon equation

$$g^{\mu\nu}D_\mu D_\nu\phi = \mathfrak{m}^2\phi,$$

where $\mathfrak{m} \in \mathbb{R}_{>0}$ represents the mass of the scalar field and satisfies $\mathfrak{m}M \ll \mathfrak{e}M$. Here M denotes the mass of the black hole to be formed.

We emphasize that not only are our data in the above examples regular, but the spacetimes arise from gravitational collapse, i.e., the initial data surface is one-ended, has a regular center, lies entirely in the domain of outer communication, and the black hole forms strictly to the future of initial data. In particular, in contrast to what has been suggested numerically [TA14; CIP21], there is no upper bound (strictly less than unity) on the charge to mass ratio of a black hole which can be achieved in gravitational collapse for this model.

1.1.2 Gravitational collapse to Reissner–Nordström black holes with prescribed parameters

For appropriate matter models, the Einstein equations (1.0.1) are well-posed (see [Fou52; CG69] for the vacuum case) as a Cauchy problem for suitable initial data posed on a 3-manifold Σ, which will then be isometrically embedded as a spacelike hypersurface in a Lorentzian manifold $(\mathcal{M}, g)$. The textbook explicit black hole solutions such as the Schwarzschild spacetime do not contain *one-ended* Cauchy surfaces $\Sigma \cong \mathbb{R}^3$ but are instead foliated by *two-ended* hypersurfaces $\Sigma \cong \mathbb{R} \times S^2$. Thus, a natural and physically relevant problem is to construct regular asymptotically flat data on $\Sigma \cong \mathbb{R}^3$ which evolve to a black hole spacetime.

Our counterexample of the third law, Theorem 1.1.1, is preceded by a more general construction, presented as Theorem 1.1.4 below. We construct regular one-ended Cauchy data for the Einstein–Maxwell-charged scalar field system in spherical symmetry whose black hole exterior evolves (in fact is eventually isometric) to a Schwarzschild black hole with prescribed mass $M > 0$ or to a *subextremal* or *extremal* Reissner–Nordström black hole with prescribed mass $M > 0$ and prescribed charge to mass ratio $\mathfrak{q} \doteq e/M \in [-1, 1]$.

Theorem 1.1.4 (Exact Reissner–Nordström arising from gravitational collapse)**.** *For any regularity index $k \in \mathbb{N}$ and charge to mass ratio $\mathfrak{q} \in [-1, 1]$, there exist spherically symmetric, asymptotically flat Cauchy data for the Einstein–Maxwell-charged scalar field system, with $\Sigma \cong \mathbb{R}^3$ and a regular center, such that the maximal future globally hyperbolic development $(\mathcal{M}^4, g)$ has the following properties:*

- *All dynamical quantities are at least C^k-regular.*
- *Null infinity $\mathcal{I}^+$ is complete.*
- *The black hole region is non-empty, $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+) \neq \emptyset$.*
- *The Cauchy surface Σ lies in the causal past of future null infinity, $\Sigma \subset J^-(\mathcal{I}^+)$. In particular, Σ does not intersect the event horizon $\mathcal{H}^+ \doteq \partial(\mathcal{BH})$. Furthermore, Σ contains no trapped or antitrapped surfaces.*
- *For sufficiently late advanced times $v \geq v_0$, the domain of outer communication, including the event horizon, is isometric to that of a Reissner–Nordström solution with charge to mass ratio*

$\mathfrak{q}$. *For $v \geq v_0$, the event horizon of the spacetime can be identified with the event horizon of Reissner–Nordström.*

For the Penrose diagram of this spacetime, see Fig. 1.2 below. The construction of Cauchy data on $\Sigma \cong \mathbb{R}^3$ in Theorem 1.1.4 will follow from the characteristic gluing statement Theorem 4.4.1, which we present in Section 4.4 below. The detailed proof of Theorem 1.1.4 is given in Section 5.6.1.

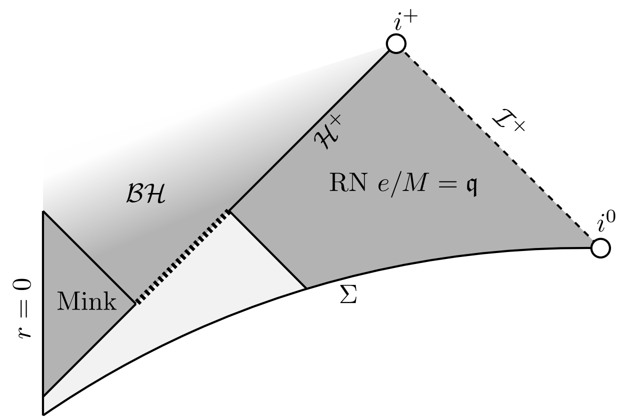

Figure 1.2: Penrose diagram for Theorem 1.1.4. The textured line segment is where the data constructed in Theorem 4.4.1 live.

The key step in the proof of Theorem 1.1.4 is a novel *characteristic/null gluing* result which we report as Theorem 4.4.1 below. The study of the characteristic gluing problem for the Einstein vacuum equations (outside of spherical symmetry) was recently initiated by Aretakis, Czimek, and Rodnianski [ACR21; ACR23b; ACR23a] in the perturbative regime around Minkowski space. Our setup is directly inspired by their work. In contrast, however, our null gluing construction (while in spherical symmetry) necessarily exploits the *large data regime* in order to glue a cone of Minkowski space to a black hole event horizon along a null hypersurface within the EMCSF model. The characteristic gluing problem will be extensively discussed in Chapter 4.

Note that in the case $|\mathfrak{q}| = 1$, this does not yet furnish a counterexample to the third law of black hole thermodynamics, as the spacetime does not necessarily contain a subextremal apparent horizon. For the counterexample we must defer to Theorem 1.1.11 in Section 1.1.4 below.

However, in our proof of Theorem 1.1.4, forming an extremal black hole with $|\mathfrak{q}| = 1$ is no different from any subextremal charge to mass ratio $|\mathfrak{q}| < 1$ (see already Section 1.1.4.2). In particular, in contrast with what has been suggested by numerical simulations [TA14; CIP21], there is no universal upper bound (strictly less than unity) for $|\mathfrak{q}|$. Given that we have now proved that extremal Reissner–Nordström can arise in gravitational collapse, it would be interesting to rethink the numerical approach to this problem and develop a scheme to construct such solutions numerically. Because our construction is fundamentally teleological (see already Section 2.4), it might be challenging to directly find suitable data on Σ by trial and error.

The formation of black holes is a very well studied problem in spherical symmetry. We mention here only the Einstein-scalar field model, for which Christodoulou [Chr91b] first showed that concentration of the scalar field can lead to formation of a black hole. This result played a decisive role in Christodoulou's proof of weak cosmic censorship in spherical symmetry [Chr99b]. Dafermos constructed solutions of the Einstein-scalar field system which collapse to the future but are complete and regular to the past [Daf09]. For work on other matter models, see for example [And14; AL22a]. In this dissertation, we also consider black hole formation in vacuum, to be discussed in Section 1.1.5 below.

Remark 1.1.5. Our derivation of Theorem 1.1.4 from Theorem 4.4.1 is completely soft and does not make use of spherical symmetry. Therefore, if versions of the main gluing theorems were known for the Einstein vacuum equations (for example, gluing a Minkowski cone to an extremal Kerr event horizon, or more generally a Schwarzschild exterior sphere to an extremal Kerr event horizon), then our procedure would yield vacuum spacetimes arising from gravitational collapse which are eventually isometric to extremal Kerr. Furthermore, such a construction would also yield a disproof of the third law in vacuum. See already Section 1.1.5.1 and Section 4.5 for the case of very slowly rotating Kerr.

Remark 1.1.6. By the very nature of the gluing procedure, our constructions have finite regularity (C^k for arbitrarily large k). It would be mathematically interesting to create such examples with C^∞ regularity. See already Remark 4.2.1.

Remark 1.1.7. The existence of dynamical spacetimes satisfying the dominant energy condition which are extremal at spacelike infinity i^0 does not contradict the positive mass theorem "with charge" [GHHP83; CRT06] because the matter itself carries charge. Concretely, condition (27) in [GHHP83] is false for various charged matter models, in particular the Einstein–Maxwell-charged scalar field model with small (or zero) mass.

1.1.3 Commentary on the formulation of the third law

In this section we give more details on the background and history of the third law of black hole thermodynamics, in particular the motivation for Israel's formulation in [Isr86; Isr92].

While the *zeroth*, *first*, and *second* laws of black hole thermodynamics are by now well understood in the literature (see e.g. [Wal01]), the validity of the third law has been a source of debate up until today. In the original form of Bardeen–Carter–Hawking (BCH), in analogy to Nernst's version of the third law of classical thermodynamics [Ner26][1], it reads:

It is impossible by any procedure, no matter how idealized, to reduce κ to zero by a finite sequence of operations.

There were two main motivations in [BCH73] for proposing a third law of black hole thermodynamics (aside from the aesthetic desire of having the full analogy between black hole and thermodynamics as laid out as in Table 1.1):

1. Heuristic perturbative calculations seemed to indicate that something like the third law could be true. Our disproof of the third law shows that such calculations do not accurately represent the fully dynamical regime of the Einstein equations.

2. It was thought that if one could charge/spin a black hole up to extremality, then one could go further and create a naked singularity, which would violate the weak cosmic censorship conjecture. This was a rather severe misunderstanding of the geometry of maximally extended superextremal Reissner–Nordström and Kerr solutions. We will return to this point in Remark 1.2.11 below.

A number of arguably pathological (e.g. singular or energy condition violating) examples of extremal black hole formation were put forth in [Kuc68; DI67; Bou73; FH79; SI80; Pró83], which Israel [Isr86; Isr92] took into account to make the third law more precise:

A nonextremal black hole cannot become extremal (i.e., lose its trapped surfaces) at a finite advanced time in any continuous process in which the stress-energy tensor of accreted matter stays bounded and satisfies the weak energy condition in a neighborhood of the outer apparent horizon.

The parenthetical comment "(i.e., lose its trapped surfaces)" is an extra source of confusion which will be specifically addressed in Section 1.1.3.3. We will now discuss the papers [Kuc68; DI67; Pró83; Bou73; FH79; SI80; Isr86; Isr92] and where the issues lie.

1.1.3.1 The singular thin charged shell model

It has been known since the 60's that an extremal black hole can be formed instantly by collapsing an infinitesimally thin shell of charged massive dust [Kuc68; DI67; Bou73; Pró83]. Later, Farrugia and

[1] For a discussion of various versions of the third law of classical thermodynamics, see [Wal97].

Hajicek [FH79] showed how to “turn a subextremal Reissner–Nordström spacetime into an extremal one” by firing an appropriately charged singular massive shell into the black hole.

The resulting spacetime metric is not C^2-regular and the energy-momentum tensor is concentrated along a timelike hypersurface (the shell). The Penrose diagram of the spacetime they construct is similar to our Fig. 1.1 (see [FH79, p. 296 Fig. 2]). In particular, we note the presence of a disconnected outermost apparent horizon in their example. Israel seemed to associate the disconnectedness of the apparent horizon with a singularity of the matter and/or spacetime: “Violations can also be produced by any process that induces discontinuous behavior of the apparent horizon—for example, absorption of an infinitely thin massive shell, which will force this horizon to jump outward.”

On the basis of this, he dismissed this example in his formulation of the third law by explicitly requiring regularity of the energy-momentum tensor. We note, however, that Farrugia and Hajicek suggest that their construction can in principle be desingularized—we do not know if this point was ever addressed again, because if true, it would seem to provide an alternative route to constructing a counterexample apart from our own.

As is clear in Fig. 1.1, the outermost apparent horizon is disconnected in our counterexamples to the third law. As we will discuss in Section 1.1.3.3, disconnectedness of the outermost apparent horizon has nothing to do with regularity—it is an intrinsic feature of extremization. Therefore, dismissing the charged thin shell on the basis of “undesirable” behavior of the apparent horizon was unwarranted. We will return to the thin charged shell in Section 1.2.1 below.

1.1.3.2 The charged null dust model

An interesting example motivating explicit mention of the weak energy condition in the third law was provided by Sullivan and Israel [SI80] in spherical symmetry, with the charged null dust matter model. This matter model allows for *dynamical* violations of the weak energy condition—even if the initial data satisfies the weak energy condition, the solution might violate it in the future. Sullivan and Israel showed that extremization is impossible in this model without such a violation, which can also be seen from Penrose diagrams. They interpreted this result as further evidence that the third law holds as long as the weak energy condition is demanded near the apparent horizon.

We note, however, that Ori has proposed a different interpretation of the model studied by Sullivan and Israel which does not violate the weak energy condition [Ori91]. This version is arguably more physically correct, and it is a pity that Ori's work was seemingly ignored in the literature. We will return to Ori's dust model in Chapter 7 below and prove that it arises as a limit of smooth solutions to the Einstein–Maxwell–Vlasov system in Chapter 8.

1.1.3.3 "Losing trapped surfaces" and connectedness of the outermost apparent horizon

We will now clarify the issue of "losing trapped surfaces" appearing prominently in [Isr86; Isr92] and the implicit assumption of connectedness of the outermost apparent horizon.

The black hole region in a subextremal Reissner–Nordström or Kerr spacetime is foliated by trapped spheres. Conversely, extremal Reissner–Nordström and Kerr black holes have no trapped surfaces, but the event horizon is a marginally trapped tube in both cases. As $|\mathfrak{q}| \to 1$ (where we take $\mathfrak{q} \doteq e/M$ for Reissner–Nordström and $\mathfrak{q} \doteq a/M$ for Kerr), $r_- \to r_+$, and one might be inclined to think that extremizing involves "squeezing" away the trapped region inside the black hole. However, it is an immediate consequence of Raychaudhuri's equation [HE73; Wal84] that trapped surfaces persist in evolution as long as the spacetime satisfies the weak energy condition. Since the typical explicit extremal black holes have no trapped surfaces (in particular none near the event horizon), one might wonder if Raychaudhuri's equation alone could be used to "prove" the third law.

This is what Israel attempted to do in [Isr86; Isr92]. We will formalize his observation in Definition 1.1.8 and Proposition 1.1.10 below. In order to reconstruct Israel's argument mathematically, let us formulate the following definition. For precise definitions relating to spherical symmetry, see already Chapter 2.

Definition 1.1.8. Let H be a connected dynamical apparent horizon, i.e., a connected, achronal curve in the $(1+1)$-dimensional reduction $(\mathcal{Q}, g_{\mathcal{Q}})$ of a spherically symmetric spacetime $(\mathcal{M}, g)$, along which $\partial_v r$ vanishes identically. We say that H *becomes extremal in finite time in the sense of Israel* if

1. H is not completely contained in a null cone.
2. Let $\tau \mapsto H(\tau)$ be a parametrization of H. Then there exists a $\tau_0 \in \mathbb{R}$ so that for all $\tau \geq \tau_0$, $\tau \mapsto H(\tau)$ is a future-directed constant u curve.
3. There exists a $\tau_1 > \tau_0$ and a neighborhood $\mathcal{N}$ of $H_{\tau \geq \tau_1}$ such that $\mathcal{N} \setminus H_{\tau \geq \tau_0}$ contains only strictly untrapped spheres ($\partial_v r > 0$).

Remark 1.1.9. The outermost apparent horizon $\mathcal{A}'$ (see already Section 2.4), if connected, is an example of a connected dynamical apparent horizon.

As a simple consequence of Raychaudhuri's equation in a spacetime satisfying the weak energy condition [HE73; Wal84], we have:

Proposition 1.1.10 (Israel's observation)**.** *Let $(\mathcal{M}, g)$ be a spherically symmetric black hole spacetime. If the spacetime satisfies the weak energy condition, has a nonempty trapped region, and a* <u>*connected*</u> *outermost apparent horizon $\mathcal{A}'$ as defined in [Kom13], then the outermost apparent horizon $\mathcal{A}'$ does not become extremal in finite time in the sense of Israel.*

However, it is clear that in view of our main theorem, the correct reading of this proposition is the *contrapositive*, namely that violations of the third law necessarily have a disconnected apparent horizon. This effect has nothing to do with singularities of spacetime or the matter model (and there was never actually any *a priori* reason to believe that the outermost apparent horizon was connected). This situation is depicted in Fig. 1.3.

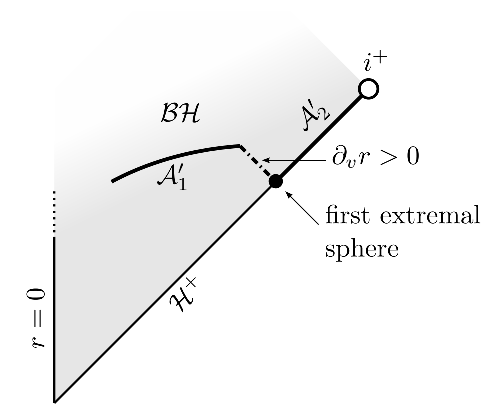

Figure 1.3: Illustration of the contrapositive of Proposition 1.1.10. The outermost apparent horizon $\mathcal{A}' = \mathcal{A}_1' \cup \mathcal{A}_2'$ becomes disconnected when a black hole with trapped surfaces "becomes extremal," while the spacetime and matter fields remain regular. The trapped region begins to the north of $\mathcal{A}_1'$ and persists for all advanced time.

1.1.3.4 Aside: Extremal horizons with nearby trapped surfaces

Though not directly relevant for the considerations of this dissertation, we would like to point out that there is another issue with the attempt to characterize extremality by the lack of trapped surfaces near the horizon, i.e., by the third property of Definition 1.1.8. In fact, it would appear that the property of having no trapped surfaces in the interior near the horizon is actually stronger than being extremal.

For a spacetime $(\mathcal{M}, g)$ with Killing field K, a Killing horizon H is said to be *extremal* if the surface gravity κ, defined by $\nabla_K K = \kappa K$ on H, vanishes identically. Equivalently, extremality means that $g(K, K)$ vanishes to at least second order along null geodesics crossing H transversely. If K is timelike to the past of H and $g(K, K)$ vanishes to an *even* order on H, then K passes from timelike, to null, then back to timelike across H, and there are no strictly trapped surfaces near the horizon. This is precisely the situation for extremal Reissner–Nordström and Kerr black holes, where $g(K, K)$ vanishes to second order on the event horizon.

However, there exist spacetimes for which $g(K, K)$ vanishes to an *odd* order (at least three), in which case there may be trapped surfaces just behind the horizon. Indeed, in Proposition 5.A.1 of Section 5.A we construct an example of a stationary spacetime containing an extremal Killing horizon, with trapped surfaces just behind the horizon, and satisfying the dominant energy condition. In this case $g(K, K)$ is exactly cubic in an ingoing null coordinate system. It would be interesting to construct such a spacetime with a specific matter model, or an extremal black hole with this behavior.

While extremal Kerr, Reissner–Nordström, and other known examples are extremal in the sense of Definition 1.1.8, it is far from obvious that all *hairy* (i.e., carrying non-EM matter fields) extremal black holes should be free of trapped surfaces. In view of our example in Section 5.A, any mechanism which enforces this must necessarily be global in nature and/or depend on particular properties of the matter model in question.

One could define the notion of a *nondegenerate* extremal Killing horizon, i.e., the Killing field K has the property that $g(K, K)$ vanishes only to second order, which would then be compatible with Definition 1.1.8. See already Remark 5.A.2.

For more discussion about possible definitions of extremality, see for instance [BF08; Boo16; MRT13].

1.1.4 Detailed description of our counterexample to the third law

With this discussion out of the way, we present now a detailed version of our counterexample to the third law, which satisfies all of Israel's requirements. It is essentially a corollary of the more general version of our main gluing result Theorem 4.4.1 with a Schwarzschild exterior sphere in place of a Minkowski sphere (see already Section 5.4) and will be given in Section 5.6.2. For an illustration of the spacetime, we refer the reader back to Fig. 1.1.

Theorem 1.1.11 (Detailed version of Theorem 1.1.1)**.** *For any regularity index $k \in \mathbb{N}$, there exist spherically symmetric, asymptotically flat Cauchy data for the Einstein–Maxwell-charged scalar field system, with $\Sigma \cong \mathbb{R}^3$ and a regular center, such that the maximal future globally hyperbolic development $(\mathcal{M}^4, g)$ has the following properties:*

- *The spacetime satisfies all the conclusions of Theorem 1.1.4 with $\mathfrak{q} = 1$, including C^k-regularity of all dynamical quantities.*

- *The black hole region contains an isometrically embedded portion of a Schwarzschild exterior horizon neighborhood. In particular, there is a portion of a null cone behind the event horizon of $(\mathcal{M}, g)$ which can be identified with a portion of the apparent horizon of Schwarzschild.*

- *The "Schwarzschild horizon" piece is a part of the outermost apparent horizon $\mathcal{A}'$ of the spacetime. The set $\mathcal{A}'$ is disconnected and agrees with the event horizon $\mathcal{H}^+$ to the future of the first marginally trapped sphere on the event horizon.*

- *There is a neighborhood of the event horizon that contains no trapped surfaces. Nonetheless, the black hole region contains trapped surfaces. In fact, there are trapped surfaces at arbitrarily late advanced time in the interior of the black hole.*

To reiterate, the scalar field collapses to form an *exact* Schwarzschild spacetime, including the horizon, only to collapse further to form an *exact* extremal Reissner–Norström for all late advanced time. The spacetime is regular (for any fixed $k \geq 1$, one can construct an example which is C^k) and the matter model satisfies the dominant energy condition.

1.1.4.1 Future boundary of the interior in third law violating solutions

The future boundary of the black hole region of dynamical black holes formed from gravitational collapse in the EMCSF system is known to be intricate (see e.g. [Daf03; Kom13; Van18b]). We refer to [Kom13] for a detailed description of the most general possible structure of the interior, but see already Fig. 2.1 for a summary of the most salient features. In this subsection we will first discuss the future boundary of the black hole interior in Theorem 1.1.11. Further, we will present additional corollaries of our characteristic gluing method which provide examples of gravitational collapse to black holes with a piece of null boundary (a "Cauchy horizon") and a construction of spacetimes for which a Cauchy horizon closes off the interior region.

For our main counterexample to the third law in Theorem 1.1.11, we obtain that the regular center Γ extends into the black hole region. Regarding the future boundary of the spacetime, we do not know whether there exists a piece of possibly singular null boundary emanating from i^+ as in the subextremal case [Daf03; Van18b] or whether a spacelike singularity emanates from i^+. Note that the result of [GL19], which shows the existence of a Cauchy horizon emanating from i^+, does not apply directly since their analysis requires $|\mathfrak{e}|M \leq 0.1$, whereas our construction requires $|\mathfrak{e}|M$ large. Nevertheless, one may speculate that a piece of Cauchy horizon occurs (for which the linear analysis of [Gaj17a; Gaj17b] would be relevant), which could eventually turn into a spacelike singularity. (Note that one can readily set up the data such that the future boundary of the interior in Theorem 1.1.11 has a piece of spacelike singularity. See however already Section 4.6.2.)

1.1.4.2 Exceptionality and stability of third law violating solutions

The third law is manifestly concerned with exceptional behavior, which is why the phrases "no matter how idealized" [BCH73] or "in any continuous process" [Isr86] are specifically included in formulations of the third law. Indeed, keeping a horizon at exactly constant temperature (or equivalently constant surface gravity), any temperature, is of course exceptional. (Exactly stationary behavior on the horizon for all late advanced times is itself an infinite codimension phenomenon in the moduli space of solutions.) In view of our construction, the case of gravitational collapse to zero temperature in finite time is no more exceptional than any other fixed temperature.

We would also like to address the interesting question of whether creating *asymptotically* extremal black holes should be viewed any differently from the subextremal case. Indeed, any mechanism which forms a black hole with *exactly specified* parameters is inherently unstable, because a small perturbation can just change the parameters. As an example of this, we note the codimension-3 nonlinear stability of the Schwarzschild family by Dafermos–Holzegel–Rodnianski–Taylor [DHRT]. In order to preserve the final black hole parameters, only a codimension-3 submanifold of the moduli space of data is admissible in their theorem.

The stability problem for extremal black holes is exceptional because they suffer from a linear instability known as the *Aretakis instability* [Are11a; Are11b; Are15; Ape22]. This instability is weak, and a restricted form of nonlinear stability is nevertheless conjectured to hold with the same codimensionality as in the subextremal case. See [DHRT, Section IV.2] for conjectures about stability of extremal black holes, [Ang16; AAG20] for stability results on a nonlinear model problem, and numerical work [MRT13; LMRT13] which is consistent with the above conjecture. The Aretakis instability should not be thought of as a manifestation of the third law and understanding its ramifications in the full nonlinear theory is a fundamental open problem in general relativity.

Therefore, asymptotic stability for any fixed parameter ratio (up to and including extremality) should be formulated as a positive codimension statement. In our spherically symmetric setting, we are led to conjecture that for every solution constructed in Theorem 1.1.4, there exists a codimension-1 family of perturbations which asymptote to a Reissner–Nordström black hole with the same final parameter ratio. Since the conjectured codimension is the same for every ratio, we are then led to conclude that asymptotically extremal black holes are not qualitatively rarer than any fixed positive temperature.

We will return to the issue of the stability of extremal black holes in Section 1.2.5 below.

1.1.5 Conjecture: the third law is false in vacuum

In light of Theorem 1.1.1, we are motivated to make the following conjecture:

Conjecture 1.1.12. *There exist regular one-ended Cauchy data for the Einstein vacuum equations*

$$\mathrm{Ric}(g) = 0 \tag{1.1.6}$$

which undergo gravitational collapse and form an exactly Schwarzschild apparent horizon, only for the spacetime to form an exactly extremal Kerr event horizon at a later advanced time. In particular, already in vacuum, the "third law of black hole thermodynamics" is false.

Remark 1.1.13. It is not possible to have a solution of the pure Einstein–Maxwell equations which behaves like one of the solutions in Theorem 1.1.1. This is because the vacuum Maxwell equation $d{\star}F = 0$ always gives rise to a conserved electric charge $(4\pi)^{-1}\int_S \star F$, even outside of spherical symmetry. On Schwarzschild, this charge is zero, and on Reissner–Nordström, it equals the charge parameter e.

Remark 1.1.14. Similarly, if a vacuum spacetime has an axial Killing field Z, then the Komar angular momentum $(16\pi)^{-1}\int_S \star dZ^\flat$ is conserved. Therefore, Conjecture 1.1.12 cannot be proved in axisymmetry.

1.1.5.1 Gravitational collapse to very slowly rotating Kerr black holes with prescribed parameters in vacuum

The black holes in Theorem 1.1.1 are constructed in two stages: First the scalar field is used to form an exact Schwarzschild apparent horizon in finite time, which is then charged up to extremality by exploiting the coupling of the scalar field with the electromagnetic field.

In [KU23], we showed how to generalize the first step of forming an exact Schwarzschild black hole in vacuum. In fact, we can form any sufficiently slowly rotating Kerr black hole:

Theorem 1.1.15 (Gravitational collapse with prescribed M and $0 \le |a| \ll M$ in vacuum)**.** *There exists a constant* $0 < \mathfrak{a}_0 \ll 1$ *such that for any mass* $M > 0$ *and specific angular momentum parameter* a *satisfying* $0 \le |a|/M \le \mathfrak{a}_0$*, there exist one-ended asymptotically flat Cauchy data* $(g_0, k_0) \in H^{7/2-}_{\mathrm{loc}} \times H^{5/2-}_{\mathrm{loc}}$ *for the Einstein vacuum equations* (1.1.6) *on* $\Sigma \cong \mathbb{R}^3$*, satisfying the* *constraint equations, such that the maximal future globally hyperbolic development* $(\mathcal{M}^4, g)$ *contains a black hole* $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+)$ *and has the following properties:*

- *The Cauchy surface* Σ *lies in the causal past of future null infinity,* $\Sigma \subset J^-(\mathcal{I}^+)$*. In particular,* Σ *does not intersect the event horizon* $\mathcal{H}^+ \doteq \partial(\mathcal{BH})$ *or contain trapped surfaces.*
- $(\mathcal{M}, g)$ *contains trapped surfaces.*
- *For sufficiently late advanced times* $v \geq v_0$*, the domain of outer communication, including the event horizon* $\mathcal{H}^+$*, is isometric to that of a Kerr solution with parameters* M *and* a*. For* $v \geq v_0$*, the event horizon of the spacetime can be identified with the event horizon of Kerr.*

For the relevant Penrose diagram, consult Fig. 1.2 below.

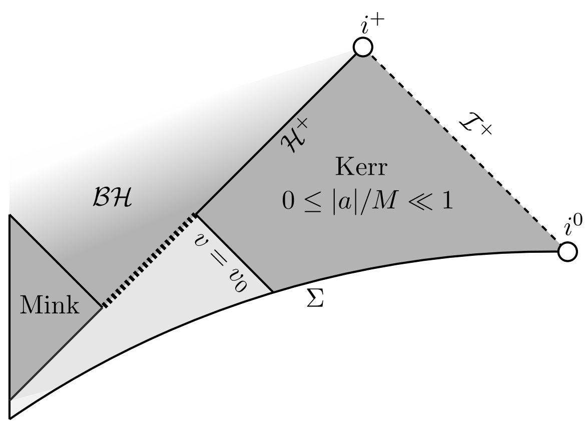

Figure 1.4: Penrose diagram for Theorem 1.1.15. The textured line segment is where the gluing data constructed in Theorem 4.5.2 live.

Remark 1.1.16. It is a classical result that the Einstein equations are well posed in $H^{7/2-}_{\mathrm{loc}} \times H^{5/2-}_{\mathrm{loc}}$, see [HKM76] and also [PR07; Chr13].

The proof is again a characteristic gluing construction, but now for the Einstein vacuum equations. See already Section 4.5.

Outside of spherical symmetry (for the Einstein vacuum equations), formation of black holes was studied by Christodoulou in the seminal monograph [Chr09]. Christodoulou constructed *characteristic data* for the Einstein vacuum equations containing no trapped surfaces, but whose evolution contains trapped surfaces in the future. Li and Yu [LY15] showed how to combine Christodoulou's construction with the spacelike gluing technique of Corvino and Schoen [CS06] to construct asymptotically flat *Cauchy data* containing no trapped surfaces, but whose evolution contains trapped surfaces in the future. Later, Li and Mei [LM20] observed that the Corvino–Schoen gluing can be done “behind the event horizon,” which yields a genuine construction of gravitational collapse in vacuum arising from one-ended asymptotically flat Cauchy data.[2]

The constructions of the above type rely on the observation that if an additional restriction is imposed on the seed data in [Chr09], then the resulting spacetime has a region of controlled size which is close to Schwarzschild. The Corvino–Schoen gluing then selects *very slowly rotating* Kerr parameters for the exterior region. Our Theorem 1.1.15 can be compared with the main theorem of Li and Mei in [LM20]. In particular, we also prove trapped surface formation starting from Cauchy data outside of the black hole region.

Our proof of trapped surface formation starting from Cauchy data is fundamentally different from [LM20] because it does not appeal to Christodoulou’s trapped surface formation mechanism [Chr09]. In fact, the only aspect of the evolution problem we require is Cauchy stability. Furthermore, we can directly prescribe the (very slowly rotating) Kerr parameters of the black hole to be formed. In particular, we may take $a = 0$, which guarantees the existence of a spacelike singularity, see already Corollary 4.6.5. However, our data is of limited regularity (but still in a well-posed class). Nevertheless, by appealing to Cauchy stability once again, Theorem 1.1.15 has the further corollary of showing the existence of an *open set* of vacuum Cauchy data not containing trapped surfaces, but which lead to trapped surface formation in evolution. This method of obtaining trapped surface formation *softly* is new and fundamentally different from Christodoulou’s method in [Chr09], which revolves around a semiglobal evolution problem.

Remark 1.1.17. In [KU23] (and this dissertation), the Cauchy data $(\bar{g}, \bar{k})$ are constructed with regularity $H^{7/2-}_{\mathrm{loc}} \times H^{5/2-}_{\mathrm{loc}}$, which is well above the threshold for classical existence and uniqueness for the Einstein vacuum equations [HKM76; PR07; Chr13]. This limited regularity is because the characteristic gluing results [ACR21; CR22] which we use as a black box are limited to C^2 regularity of transverse derivatives in the non-bifurcate case. Using the more recent *spacelike* gluing results of Mao–Oh–Tao [MOT23], it is possible to construct suitable Cauchy data in $H^s_{\mathrm{loc}} \times H^{s-1}_{\mathrm{loc}}$ for any s.

1.1.5.2 The Thorne bound

When a black hole forms in nature, it is typically surrounded by a so-called *accretion disk*, consisting of gas, dust, and other diffuse material. The matter in the disk is susceptible to friction, which raises the temperature and causes emission of electromagnetic radiation [PT74]. A famous heuristic argument in astrophysics, called the *Thorne bound* [Tho74], would imply that Conjecture 1.1.12 is

[2]Here, gravitational collapse refers to a solution of Einstein’s equations containing a black hole, such that the Cauchy hypersurface on which data are posed does not intersect the black hole region.

false in the presence of an accretion disk because of the back-reaction of photons. It would be very interesting to investigate this argument further in light of our recent results.

1.2 Extremal black hole formation as a critical phenomenon

1.2.1 Statement of the main result

In contrast to the examples of gravitational collapse presented in the previous section, small initial data for the Einstein equations (with reasonable matter content) tend to *disperse* without a black hole forming. It is a fundamental problem in classical general relativity to understand how these different classes of spacetimes—collapsing and dispersing—fit together in the moduli space of solutions. The interface between collapse and dispersion is known as the *black hole formation threshold* and families of solutions crossing this threshold are said to exhibit *critical collapse*. Spacetimes lying on the threshold are called *critical solutions*.

Critical collapse has been extensively studied numerically, starting with the influential work of Choptuik [Cho93] on the spherically symmetric Einstein-scalar field model, in a regime where the critical solutions are believed to be naked singularities. The Einstein–Vlasov system is believed to have static "star-like" critical solutions [RRS98; OC02], but critical collapse involving naked singularities has so far not been observed. These numerical studies on critical collapse (see also the survey [GM07]) have yet to be made rigorous.

At first glance, the Reissner–Nordström family of metrics (indexed by the mass $M > 0$ and charge e) appears to exhibit a type of critical behavior: the solution contains a black hole when $|e| < M$ (subextremal) or $|e| = M$ (extremal) and does not contain a black hole when $|e| > M$ (superextremal). However, the Reissner–Nordström black holes are eternal and arise from two-ended Cauchy data, while the superextremal variants contain an eternal "naked singularity" that has historically caused much confusion. Moreover, it was long thought that extremal black holes could not form dynamically (a consideration closely related to the third law of black hole thermodynamics). Were this true, it would seem to rule out extremal Reissner–Nordström as the late-time behavior of any critical solution. As discussed in the previous section, Kehle and the present author disproved the third law in the Einstein–Maxwell-charged scalar field model and showed that an exactly extremal Reissner–Nordström domain of outer communication can indeed form in gravitational collapse.

We now continue our investigation of extremal black hole formation by showing that extremal Reissner–Nordström *does* arise as a critical solution in gravitational collapse for the Einstein–

Maxwell–Vlasov model, giving rise to a new phenomenon that we call *extremal critical collapse.*

Theorem 1.2.1. *There exist extremal black holes on the threshold between collapsing and dispersing smooth configurations of charged matter. More precisely, for any mass $M > 0$, fundamental charge $\mathfrak{e} \neq 0$, and particle mass $0 \leq \mathfrak{m} \leq \mathfrak{m}_0$, where $0 < \mathfrak{m}_0 \ll 1$ depends only on M and $\mathfrak{e}$, there exists a smooth one-parameter family of smooth, spherically symmetric, one-ended asymptotically flat Cauchy data $\{\Psi_\lambda\}_{\lambda\in[0,1]}$ for the Einstein–Maxwell–Vlasov system for particles of fundamental charge $\mathfrak{e}$ and mass $\mathfrak{m}$, such that the maximal globally hyperbolic development of Ψ_λ, denoted by $\mathcal{D}_\lambda$, has the following properties.*

1. *$\mathcal{D}_0$ is isometric to Minkowski space and there exists $\lambda_* \in (0,1)$ such that for $\lambda < \lambda_*$, $\mathcal{D}_\lambda$ is future causally geodesically complete and disperses towards Minkowski space. In particular, $\mathcal{D}_\lambda$ does not contain a black hole or naked singularity. If $\lambda < \lambda_*$ is sufficiently close to λ_*, then for sufficiently large advanced times and sufficiently small retarded times, $\mathcal{D}_\lambda$ is isometric to an appropriate causal diamond in a superextremal Reissner–Nordström solution.*

2. *$\mathcal{D}_{\lambda_*}$ contains a nonempty black hole region $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+)$ and for sufficiently large advanced times, the domain of outer communication, including the event horizon $\mathcal{H}^+ \doteq \partial(\mathcal{BH})$, is isometric to that of an extremal Reissner–Nordström solution of mass M. The spacetime contains no trapped surfaces.*

3. *For $\lambda > \lambda_*$, $\mathcal{D}_\lambda$ contains a nonempty black hole region $\mathcal{BH}$ and for sufficiently large advanced times, the domain of outer communication, including the event horizon $\mathcal{H}^+$, is isometric to that of a subextremal Reissner–Nordström solution. The spacetime contains an open set of trapped surfaces.*

In addition, for every $\lambda \in [0,1]$, $\mathcal{D}_\lambda$ is past causally geodesically complete, possesses complete null infinities $\mathcal{I}^+$ and $\mathcal{I}^-$, and is isometric to Minkowski space near the center $\{r = 0\}$ for all time.

In the proof of Theorem 1.2.1, we construct one-parameter families of charged Vlasov beams coming in from past timelike infinity (if $\mathfrak{m} > 0$, cf. Fig. 1.5) or from past null infinity (if $\mathfrak{m} = 0$, cf. Fig. 1.6). In the dispersive case $\lambda < \lambda_*$, the area-radius r of the beam grows linearly in time as the matter expands towards the future. Moreover, the macroscopic observables of the Vlasov matter (the particle current N and energy momentum tensor T) decay at the sharp t^{-3} rate in the massive case and at the sharp t^{-2} rate in the massless case (with faster decay for certain null components), see already (8.2.10)–(8.2.14) in Proposition 8.2.3. In fact, this same dispersive behavior occurs in the *past* for every $\lambda \in [0,1]$.

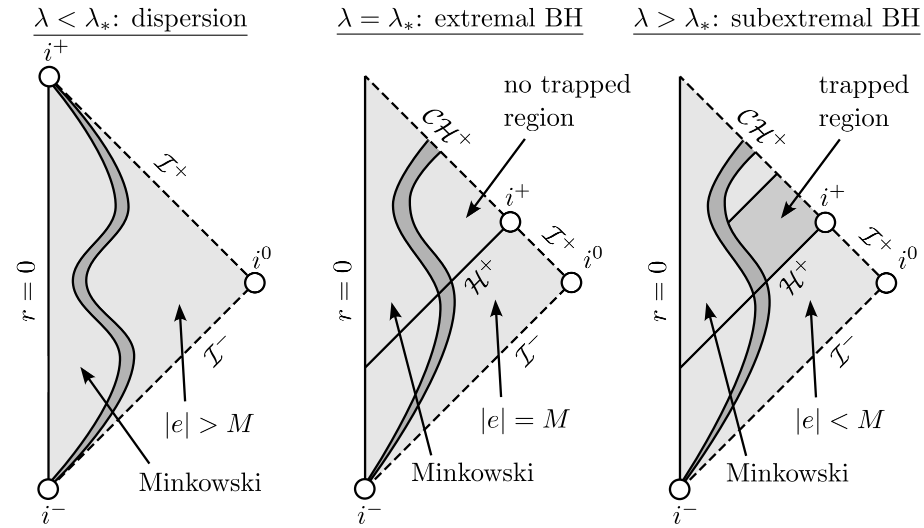

Figure 1.5: Penrose diagrams of the one-parameter family $\{\mathcal{D}_\lambda\}$ from Theorem 1.2.1 in the case of massive particles. The dark gray region depicts the physical space support of the Vlasov matter beam. The region of spacetime to the left of the beam is exactly Minkowski space and the region to the right of the beam is exactly Reissner–Nordström with the parameter ratio as depicted. In every case, the beam "bounces" before it hits the center $\{r = 0\}$ due to the repulsive effects of angular momentum and the electromagnetic field. When $\lambda < \lambda_*$, the beam bounces before a black hole is formed. The leftmost figure, with superextremal exterior parameters, represents the case of λ close to λ_*. For λ close to zero, the exterior parameters can be subextremal (but nevertheless no black hole forms). We note already that the beams actually have more structure than is depicted here in these "zoomed out" pictures. See already Fig. 8.2.

As a direct consequence of Theorem 1.2.1, we obtain

Corollary 1.2.2. *The very "black hole-ness" of an extremal black hole arising in gravitational collapse can be unstable: There exist one-ended asymptotically flat Cauchy data for the Einstein–Maxwell–Vlasov system, leading to the formation of an extremal black hole, such that an arbitrarily small smooth perturbation of the data leads to a future causally geodesically complete, dispersive spacetime.*

This is in stark contrast to the subextremal case, where formation of trapped surfaces behind the event horizon—and hence *stable* geodesic incompleteness [Pen65]—is expected. Despite this inherent instability of the critical solution, we expect extremal critical collapse itself to be a stable phenomenon: We conjecture that there exists a teleologically determined "hypersurface" $\mathfrak{B}_{\mathrm{crit}}$ in moduli space which consists of asymptotically extremal black holes, contains $\mathcal{D}_{\lambda_*}$, and locally delimits the boundary in moduli space between future complete and collapsing spacetimes. This "codimension-one" property is expected to hold for other variants of critical collapse and will be discussed in detail in Section 1.2.5.

We further expect extremal critical collapse to be a more general phenomenon: we conjecture it

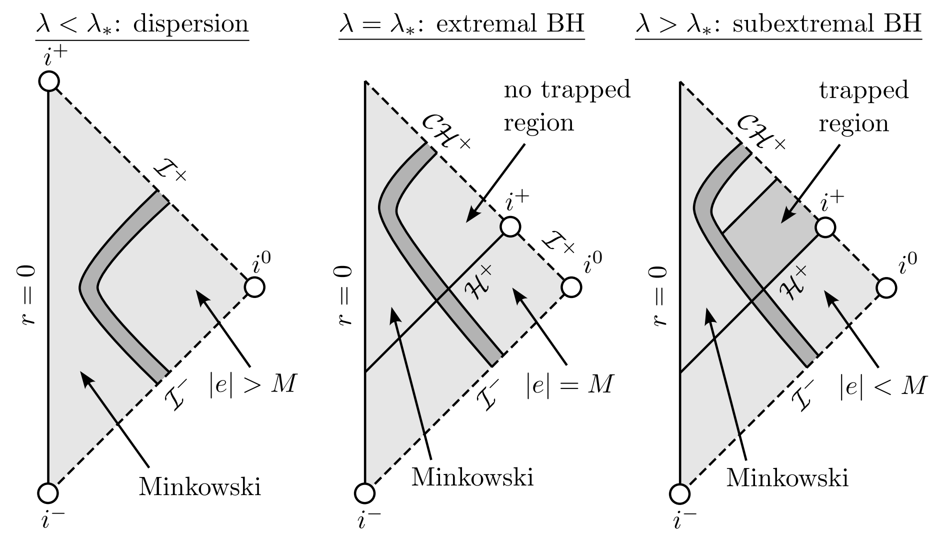

Figure 1.6: Penrose diagrams of the one-parameter family $\{\mathcal{D}_\lambda\}$ from Theorem 1.2.1 in the massless case. In the ingoing (resp., outgoing) phases, the massless beams are entirely contained in slabs of finite advanced (resp., retarded) time. Therefore, for sufficiently early advanced times and sufficiently late retarded times, the solutions are vacuum and isometric to Minkowski space.

to occur in the spherically symmetric Einstein–Maxwell-charged scalar field model and also for the Einstein vacuum equations, where extremal Kerr is the model critical solution. In this paper, we also prove (see already Theorem 7.3.2 in Section 7.3) that extremal critical collapse already occurs in the simpler—but singular—*bouncing charged null dust model*, which was first introduced by Ori in [Ori91]. The proof of Theorem 1.2.1, which will be outlined in Section 8.1, can be viewed as a global-in-time desingularization of these extremal critical collapse families in dust.

Besides the Einstein–Maxwell–Vlasov and bouncing charged null dust models, it turns out that the *thin charged shell model* [Isr66; DI67] also exhibits extremal critical collapse: Prószyński observed in [Pró83] that if a thin charged shell is injected into Minkowski space (so the interior of the shell is always flat), the parameters can be continuously varied so that the exterior of the shell goes from forming a subextremal Reissner–Nordström black hole, to forming an extremal Reissner–Nordström black hole, to forming no black hole or naked singularity at all: the shell "bounces" off to future timelike infinity. Because the thin shell model is quite singular (the energy-momentum tensor is merely a distribution and the metric can fail to be C^1 across the shell), it seems to have been discounted as a serious matter model. We refer to the previous discussion in Section 1.1.3.1 in reference to the thin charged shell counterexample to the third law by Farrugia and Hajicek [FH79]. Theorem 1.2.1 can be viewed as a vindication of [Pró83], since our smooth Einstein–Maxwell–massive Vlasov solutions exhibit all of the qualitative features of Prószyński's dust shells. In particular, Fig. 8.1 below is strikingly similar to Fig. 3 in [Pró83].

1.2.2 The Einstein–Maxwell–Vlasov system

In Theorem 1.2.1, we consider the *Einstein–Maxwell–Vlasov* system, which models a distribution of collisionless, self-gravitating charged particles with *mass* $\mathfrak{m} \geq 0$ and *fundamental charge* $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$. The model consists of a quadruple $(\mathcal{M}^4, g, F, f)$, where $(\mathcal{M}^4, g)$ is a spacetime, F is a closed 2-form representing the electromagnetic field strength, and $f = f(x,p)$, called the *distribution function* of the Vlasov matter, is a smooth nonnegative function defined on the *mass shell*

$$P^{\mathfrak{m}} \doteq \{(x,p) \in T\mathcal{M} : p \text{ is future-directed causal and } g(p,p) = -\mathfrak{m}^2\}.$$

The equations of motion are

$$R_{\mu\nu} - \tfrac{1}{2} R g_{\mu\nu} = 2\left(T^{\mathrm{EM}}_{\mu\nu} + T_{\mu\nu}\right), \tag{1.2.1}$$

$$\nabla_\mu F^{\mu\nu} = -\mathfrak{e} N^\nu, \tag{1.2.2}$$

$$Xf = 0, \tag{1.2.3}$$

where $T^{\mathrm{EM}}_{\mu\nu} \doteq F_\mu{}^\alpha F_{\nu\alpha} - \frac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta}$ is the energy-momentum tensor of the electromagnetic field, N and T are the number current and energy-momentum tensor of the Vlasov matter, defined by

$$N^\mu[f](x) \doteq \int_{P^{\mathfrak{m}}_x} p^\mu f(x,p)\, d\mu^{\mathfrak{m}}_x(p), \quad T^{\mu\nu}[f](x) \doteq \int_{P^{\mathfrak{m}}_x} p^\mu p^\nu f(x,p)\, d\mu^{\mathfrak{m}}_x(p), \tag{1.2.4}$$

and $X \in \Gamma(TT\mathcal{M})$ is the *electromagnetic geodesic spray* vector field, defined relative to canonical coordinates (x^μ, p^μ) on $T\mathcal{M}$ by

$$X \doteq p^\mu \frac{\partial}{\partial x^\mu} - \left(\Gamma^\mu_{\alpha\beta} p^\alpha p^\beta - \mathfrak{e} F^\mu{}_\alpha p^\alpha\right) \frac{\partial}{\partial p^\mu}. \tag{1.2.5}$$

For the definition of the family of measures $d\mu^{\mathfrak{m}}_x$ on $P^{\mathfrak{m}}$ and a proof of the consistency of the system (1.2.1)–(1.2.3), we refer to Section 2.3.1.

The integral curves of the vector field X consist of curves of the form $s \mapsto (\gamma(s), p(s)) \in T\mathcal{M}$, where $p = d\gamma/ds$ and p satisfies the *Lorentz force equation*

$$\frac{Dp^\mu}{ds} = \mathfrak{e} F^\mu{}_\nu p^\nu.$$

We refer to such curves γ as *electromagnetic geodesics.* The vector field X is tangent to $P^{\mathfrak{m}}$ for any $\mathfrak{m} \geq 0$, and the Vlasov equation (1.2.3) implies that f is conserved along electromagnetic geodesics.

Since $f \geq 0$, N is a future-directed causal vector field on $\mathcal{M}$ and the model satisfies the dominant energy condition.

When $\mathfrak{m} > 0$, the system (1.2.1)–(1.2.3) is locally well-posed outside of symmetry, which can be seen as a special case of results in [BC73] or by applying the general methods of [Rin13]. Well-posedness when $\mathfrak{m} = 0$ is conditional and is a delicate issue that we will return to in Section 3.2.1. We emphasize at this point that Theorem 1.2.1 produces examples of extremal critical collapse for any sufficiently small positive particle mass, where well-posedness is unconditional and valid outside of spherical symmetry.

1.2.3 The problem of critical collapse

We would like to place Theorem 1.2.1 into the larger picture of *critical collapse*, the general study of the black hole formation threshold. In particular, we conjecture that our examples in Theorem 1.2.1 have a suitable *codimension-one* property as is expected to hold for other, so far only numerically observed, critical phenomena in gravitational collapse.

In order to discuss the general concept of critical collapse, it is very helpful to have a notion of "phase space" or *moduli space* for initial data (or maximal Cauchy developments) for the Einstein equations. Consider, formally, the set $\mathfrak{M}$ of one-ended asymptotically flat Cauchy data for the Einstein equations with a fixed matter model (or vacuum) and perhaps with an additional symmetry assumption. We will be intentionally vague about what regularity elements of $\mathfrak{M}$ have, what decay conditions to impose, or what topology to endow $\mathfrak{M}$ with. We will also not discuss gauge conditions, which could be viewed as taking specific quotients of $\mathfrak{M}$. These questions are related to several fundamental issues in general relativity, see for instance [Chr94; Chr99b; Chr02; DS18; LO19; Keh22a; RSR23; Keh23; KM24; Sin24].[3] Indeed, it seems likely that there is no single "correct" definition—it is doubtful that a single moduli space will capture every interesting phenomenon.

Nevertheless, we will pretend in this section that a "reasonable" definition of $\mathfrak{M}$ exists. At the very least, $\mathfrak{M}$ ought to consist of initial data possessing a well-posed initial value problem. For each element $\Psi = (\bar{g}, \bar{k}, \dots) \in \mathfrak{M}$ (where $\bar{g}$ is a Riemannian metric on $\mathbb{R}^3$, $\bar{k}$ the induced second fundamental form, and $\dots$ denotes possible matter fields), we have a unique maximal globally hyperbolic development $\mathcal{D} = (\mathcal{M}, g, \dots)$ of Ψ, where $\mathcal{M} \cong \mathbb{R}^4$ [Fou52; CG69; Sbi16].[4] We assume

[3] In particular, it would actually be most natural to define $\mathfrak{M}$ in terms of (perhaps a quotient space of) *scattering data* on past infinity (past null infinity $\mathcal{I}^-$ in the case of massless fields). However, since a nonlinear scattering theory for the full Einstein equations has not yet been developed in any regime, we limit ourselves to the Cauchy problem for now.

[4] By an abuse of terminology, we will interchangeably refer to either Ψ or its development $\mathcal{D}$, which is of course only unique up to isometry.

that $(\mathcal{M}, g)$ is asymptotically flat. In particular, we assume that we have a well-defined notion of future null infinity $\mathcal{I}^+$ and past null infinity $\mathcal{I}^-$.

Remark 1.2.3. In the proof of Theorem 1, we define a "naive moduli space" $\mathfrak{M}_\infty$ consisting of all smooth solutions of the Einstein–Maxwell–Vlasov constraint equations on $\mathbb{R}^3$, equipped with the C^∞_{loc} topology, and with no identifications made. See already Definition 8.9.8. This topology is inadequate for addressing asymptotic stability questions but since our families are electrovacuum outside a fixed large compact set anyway, they will be continuous in any "reasonable" topology that respects asymptotic flatness.

Let $\mathfrak{C} \subset \mathfrak{M}$ denote the subset of initial data with *future causally geodesically complete* developments. We also highlight the special class $\mathfrak{D} \subset \mathfrak{C}$ of initial data with *dispersive* developments, i.e, those solutions whose geometry asymptotically converges to Minkowski space in the far future and matter fields decay suitably.[5] Nontrivial stationary states, if they exist, lie in $\mathfrak{C} \setminus \mathfrak{D}$ since they do not decay.[6] Let $\mathfrak{B} \subset \mathfrak{M}$ denote the set of initial data leading to the formation of a nonempty *black hole region*, i.e., $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+) \neq \emptyset$. The question of *critical collapse* is concerned with the study of phase transitions between $\mathfrak{C}$, $\mathfrak{D}$, and $\mathfrak{B}$, that is, the structure of the boundaries $\partial\mathfrak{C}$, $\partial\mathfrak{D}$, and $\partial\mathfrak{B}$, how they interact, and characterizing solutions lying on the threshold.

A natural way of exploring this phase transition is by studying continuous paths of initial data interpolating between future complete and black hole forming solutions.

Definition 1.2.4. An *interpolating family* is a continuous one-parameter family $\{\Psi_\lambda\}_{\lambda\in[-1,1]} \subset \mathfrak{M}$ such that $\Psi_0 \in \mathfrak{C}$ and $\Psi_1 \in \mathfrak{B}$. Given such a family, we may define the *critical parameter* λ_* and the *critical solution* $\mathcal{D}_{\lambda_*}$ (the development of Ψ_{λ_*}) by

$$\lambda_* \doteq \sup\{\lambda \in [0,1] : \Psi_\lambda \in \mathfrak{C}\}.$$

The prototypical critical collapse scenario consists of a spherically symmetric self-gravitating massless scalar field pulse with fixed profile and "total energy" $\sim \lambda$. At $\lambda = 0$, the solution is Minkowski space and for λ very close to 0, the solution disperses and is future complete [Chr86]. As λ approaches 1, a trapped surface forms in evolution, signaling the formation of a black hole [Chr91c]. This is precisely the scenario first studied numerically by Christodoulou in his thesis [Chr71] and then later by Choptuik in the influential work [Cho93]. Based on numerical evidence,

[5] Again, we are being intentionally vague here.

[6] According to a famous theorem of Lichnerowicz, the Einstein vacuum equations do not admit nontrivial asymptotically flat stationary solutions on $\mathbb{R}^3 \times \mathbb{R}$ (with an everywhere timelike Killing field) [Lic55]. On the other hand, the Einstein–Vlasov and Einstein–Maxwell–Vlasov systems, for example, have many asymptotically flat stationary solutions [RR93; Tha19; Tha20].

it is believed that the critical solutions for these types of families are *naked singularities* that form a codimension-one "submanifold" in moduli space. For discussion of Choptuik's results we refer to the survey [GM07].

Remark 1.2.5. A codimension-one submanifold of naked singularities is nongeneric and therefore compatible with the weak cosmic censorship conjecture, which has been proved in this model by Christodoulou [Chr99b].

Remark 1.2.6. A rigorous understanding of Choptuik's critical collapse scenario would in particular give a construction of naked singularities in the Einstein-scalar field system starting from *smooth* initial data, in contrast to Christodoulou's examples in [Chr94]. It already follows from work of Christodoulou [Chr91c] that a critical solution cannot be a black hole in this model and from work of Luk and Oh that a critical solution cannot "scatter in BV norm" [LO15]. This leaves the possibility of either a first singularity along the center not hidden behind an event horizon[7] or a solution in $\mathfrak{C}\setminus\mathfrak{D}$ which "blows up at infinity." Ruling out this latter case is an interesting open problem.

When massive fields are introduced, such as in the spherically symmetric Einstein–massive Klein–Gordon or Einstein–massive Vlasov systems, then static "star-like" critical solutions can be observed numerically [BCG97; RRS98; OC02; AR06; AAR21]. These static solutions are nonsingular and lie in $\mathfrak{C}\setminus\mathfrak{D}$. It is interesting to note that while Einstein–Klein–Gordon also displays Choptuik-like naked singularity critical solutions, there is no numerical evidence for the existence of naked singularities in the Einstein–Vlasov system. We again refer to [GM07] for references and would also like to point out the new development [Bau+23] on numerical critical collapse in vacuum.

1.2.4 Extremal critical collapse

So far, all numerically observed critical solutions are believed to be either naked singularities or complete and nondispersive. It follows at once from Penrose's incompleteness theorem [Pen65] and Cauchy stability that a critical solution cannot contain a trapped surface. While a generic black hole is expected to contain trapped surfaces,[8] members of the extremal Kerr–Newman family do not. In view of this, we raise the question of whether extremal black holes can arise on the black hole formation threshold:

Definition 1.2.7. An interpolating family $\{\Psi_\lambda\}_{\lambda\in[0,1]}$ exhibits *extremal critical collapse* if the critical solution $\mathcal{D}_{\lambda_*}$ asymptotically settles down to an extremal black hole.

[7] See [Kom13, Page 10] for a catalog of the possible Penrose diagrams in this case.

[8] By the celebrated redshift effect, one expects a spacetime asymptoting to a subextremal Kerr–Newman black hole to contain trapped surfaces asymptoting to future timelike infinity i^+. See [Daf05c; DR05b; DL17; Van18a; AH23].

Our main result, Theorem 1.2.1, proves that the Einstein–Maxwell–Vlasov system exhibits extremal critical collapse, with critical solution $\mathcal{D}_{\lambda_*}$ exactly isometric to extremal Reissner–Nordström in the domain of outer communication at late advanced times. As shown by Prószyński [Pró83] and the present authors in Theorem 7.3.2, the fundamentally singular thin charged shell and charged null dust models, respectively, exhibit extremal critical collapse, also with extremal Reissner–Nordström as the critical solution. We expect this phenomenon to also occur in the spherically symmetric Einstein–Maxwell-charged scalar field system and even for the Einstein vacuum equations, where the critical solution is expected to be based on the extremal Kerr solution. Note that we only require the *asymptotic geometry* of the critical solution to be an extremal black hole in Definition 1.2.7, which is a much weaker condition than being exactly extremal as in Theorem 1.2.1.

Remark 1.2.8. Because black holes in the spherically symmetric Einstein-scalar field model always contain trapped surfaces [Chr91c], this model *does not* exhibit extremal critical collapse. In particular, since the presence of a trapped surface in this model already implies completeness of null infinity and the existence of a black hole [Daf05b], $\mathfrak{B}$ *is open in the spherically symmetric Einstein-scalar field model.* Moreover, black holes in this model always settle down to (subextremal) Schwarzschild [Chr87].

Remark 1.2.9. We reiterate the points of Remarks 1.1.13 and 1.1.14. It is not possible for a Kerr solution with nonzero angular momentum (i.e., not Schwarzschild) to appear as the asymptotic state in axisymmetric vacuum gravitational collapse. This is because the Komar angular momentum is independent of the sphere S, which is nullhomologous. Similarly, it is not possible for a Kerr–Newman solution with nonzero charge (i.e., not Kerr) to appear as the asymptotic state in gravitational collapse for the Einstein–Maxwell system. This is because the charge is independent of the sphere S, which is nullhomologous. The presence of charged matter is essential in Theorem 1.2.1.

Remark 1.2.10 (Stationary solutions and the extremal limit). In the 1960s and '70s, it was suggested that astrophysical black holes could form through quasistationary accretion processes. In a landmark work, Bardeen and Wagoner [Bar70; BW71] numerically studied axisymmetric stationary states of the Einstein-dust system (modeling accretion disks) and found that a "black hole limit" was only possible in the "extremal limit" of the dust configuration.[9] In this limit, the exterior metric of the disk converges, in a certain sense, to the metric of the domain of outer communication of extremal

[9]Recall that the classical Buchdahl inequality states that a spherically symmetric stationary fluid ball is always "far away" from being a black hole in the sense that $\frac{2m}{r} < \frac{8}{9}$, which quantitatively forbids (even marginally) trapped surfaces [Buc59]. This bound is relaxed outside of spherical symmetry or in the presence of charge. In particular, the sharp charged Buchdahl inequality in [And09] is consistent with becoming arbitrarily close to extremality and forming a marginally trapped surface.

Kerr.

However, the event horizon of a stationary black hole is necessarily a Killing horizon and therefore an exactly stationary black hole solution cannot admit a one-ended asymptotically flat Cauchy hypersurface.[10] It follows that a sequence of one-ended stationary states cannot actually smoothly converge to a black hole spacetime up to and including the event horizon, and that the black hole threshold cannot be directly probed by studying limits of stationary states—black hole formation is a fundamentally dynamical process.

Nevertheless, there is a substantial body of numerical and heuristic literature exploring "extremal black hole limits" of stationary solutions in dust models [NM95; Mei06; Mei+08; MH11; KLM11] and using Einstein–Yang–Mills–Higgs magnetic monopoles [LW99; LW00]; see also references therein. In particular, we refer the reader to [MH11] for a cogent explanation of the exact nature of the convergence of these stationary states to extremal Reissner–Nordström/Kerr exteriors and throats. It would be interesting to see if perturbing these "near-extremal" non-black hole stationary states can provide another route to extremal critical collapse (and also perhaps to new examples of third law violating solutions), but this seems to be a difficult and fully dynamical problem as stationarity necessarily has to be broken in order for a black hole to form.

Remark 1.2.11 (Overcharging and overspinning). Extremal critical collapse should not be confused with the attempt to overcharge or overspin a black hole, i.e., the attempt to destroy the event horizon and create a "superextremal naked singularity" by throwing charged or spinning matter into an extremal or near-extremal black hole. The fear of forming such a naked singularity provided some impetus for the original formulation of the third law in [BCH73][11] and many arguments for and against have appeared in the literature, see [Wal74; Hub99; JS09; SW17] and references therein. Overcharging has been definitively disproved in spherical symmetry for the class of "weakly tame" matter models [Daf05b; Kom13], which includes the Einstein–Maxwell-charged scalar field and Einstein–Maxwell–Vlasov systems considered in this dissertation. We expect overcharging and overspinning to be definitively disproved with a positive resolution of the black hole stability problem for extremal black holes, to be discussed in Section 1.2.5 below.

[10] The original dust disk configuration is one-ended.

[11] With this in mind, the formulation of the third law in [BCH73] can be thought of as simply outright forbidding the formation of extremal black holes. The formulation in Israel's work [Isr86] is more refined and specifically refers to subextremal black holes "becoming" extremal in a dynamical process. In any case, both formulations are false as demonstrated in this dissertation.

1.2.5 Stability of extremal critical collapse

Before discussing the stability of our interpolating families in Theorem 1.2.1, we must first address the expected notion of stability for the domain of outer communication of the extremal Reissner–Nordström solution.

Firstly, since the asymptotic parameter ratio of the black hole is inherently unstable, we can at most expect a *positive codimension* stability statement for extremal Reissner–Nordström. This should be compared with the codimension-three nonlinear stability theorem of the Schwarzschild solution by Dafermos, Holzegel, Rodnianski, and Taylor [DHRT]: Only a codimension-three "submanifold" of moduli space can be expected to asymptote to Schwarzschild, which has codimension three in the Kerr family (parametrized by the mass and specific angular momentum vector). In the case of Reissner–Nordström, the set of extremal solutions has *codimension one* in the full family. Indeed, any fixed parameter ratio subfamily of the Reissner–Nordström family has codimension one. See already Remark 1.2.17.

Secondly, and far less trivially, the stability problem for extremal black holes is complicated by the absence of the celebrated *redshift effect*, which acts as a stabilizing mechanism for the event horizon of subextremal black holes. The event horizon of extremal Reissner–Nordström (and axisymmetric extremal black holes in general) suffers from a linear instability known as the *Aretakis instability* [Are11a; Are11b; Are15; Ape22], which causes ingoing translation invariant null derivatives of solutions to the linear wave equation to (generically) either not decay, or to blow up polynomially along the event horizon as $v \to \infty$. Weissenbacher has recently shown that a similar instability (non-decay of the first derivative of the energy-momentum tensor) occurs for the linear massless Vlasov equation on extremal Reissner–Nordström [Wei23].

However, the Aretakis instability is weak and does not preclude asymptotic stability and decay *away from the event horizon.* Including the horizon, we expect a degenerate type of stability, with decay in directions tangent to it, and possible non-decay and growth transverse to it (so-called *horizon hair*). This behavior has been shown rigorously for a semilinear model problem on a fixed background [Ang16; AAG20] and numerically for the coupled spherically symmetric nonlinear Einstein–Maxwell-(massless and neutral) scalar field system [MRT13].

To further complicate matters, the massive and massless Vlasov equations behave fundamentally differently and we state two separate conjectures. In these statements, we consider characteristic data posed on a bifurcate null hypersurface $C_{\mathrm{out}} \cup \underline{C}_{\mathrm{in}}$, where C_{out} is complete and $\underline{C}_{\mathrm{in}}$ penetrates the event horizon in the case of trivial data. Solutions of the linear massless Vlasov equation decay exponentially on subextremal Reissner–Nordström black holes [Big23; Wei23] and Velozo Ruiz has proved nonlinear asymptotic stability of Schwarzschild for the spherically symmetric Einstein–massless Vlasov system [Vel23]. Based on this, [MRT13; Ang16; AAG20], and [DHRT, Conjecture IV.2], we make the

Conjecture 1.2.12. *The extremal Reissner–Nordström solution is nonlinearly asymptotically stable to spherically symmetric perturbations in the Einstein–Maxwell–massless Vlasov model in the following sense: Given sufficiently small characteristic data posed on a bifurcate null hypersurface* $C_{\mathrm{out}} \cup \underline{C}_{\mathrm{in}}$ *and lying on a "codimension-one submanifold"* $\mathfrak{M}_{\mathrm{stab}}$ *(which contains the trivial solution) of the moduli space of such initial data, the maximal Cauchy development contains a black hole which asymptotically settles down to the domain of outer communication of an extremal Reissner–Nordström solution, away from the event horizon* $\mathcal{H}^+$*. Moreover, along the horizon, the solution decays towards extremal Reissner–Nordström in tangential directions, with possibly growing "Vlasov hair" transverse to the horizon.*

Remark 1.2.13. There exist nontrivial spherically symmetric static solutions of the Einstein–massless Vlasov system containing a black hole which are isometric to a Schwarzschild solution in a neighborhood of the event horizon [And21].[12] However, these are not small perturbations of Schwarzschild as the structure of trapping for null geodesics is significantly modified in the construction. Their existence is therefore consistent with [Vel23] and Conjecture 1.2.12.

The massive Vlasov equation admits many nontrivial stationary states on black hole backgrounds, which is an obstruction to decay and we do not expect a general asymptotic stability statement to hold, even in the subextremal case. In fact, it has been shown that there exist spherically symmetric static solutions of Einstein–massive Vlasov bifurcating off of Schwarzschild [Rei94; Jab21]. We refer to [Vel23] for a characterization of the "largest" region of phase space on which one can expect decay for the massive Vlasov energy-momentum tensor on a Schwarzschild background. However, one might still hope for *orbital* stability of the exterior, with a non-decaying Vlasov matter *atmosphere*, and that the horizon itself decays to that of extremal Reissner–Nordström:

Conjecture 1.2.14. *The extremal Reissner–Nordström solution is nonlinearly orbitally stable to spherically symmetric perturbations in the Einstein–Maxwell–massive Vlasov model in the following sense: Given sufficiently small characteristic data posed on a bifurcate null hypersurface* $C_{\mathrm{out}} \cup \underline{C}_{\mathrm{in}}$ *and lying on a "codimension-one submanifold"* $\mathfrak{M}_{\mathrm{stab}}$ *of the moduli space of such initial data, the*

[12] A similar construction can presumably be performed for the Einstein–Maxwell–massless Vlasov system and Reissner–Nordström black holes which makes this relevant to the current discussion.

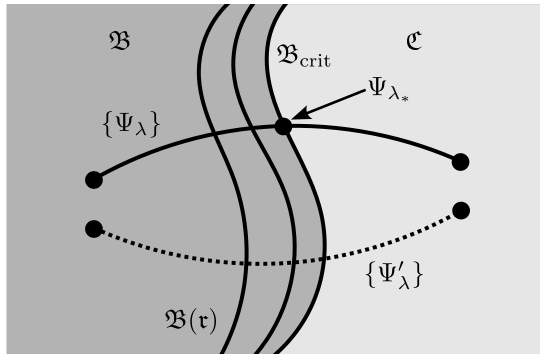

Figure 1.7: A cartoon depiction of the conjectured structure of a neighborhood of moduli space near an interpolating family $\{\Psi_\lambda\}$ from Theorem 1.2.1. We have suppressed infinitely many dimensions and emphasize the codimension-one property of the critical "submanifold" $\mathfrak{B}_{\text{crit}}$ which consists of asymptotically extremal black holes in accordance with Conjectures 1.2.12 and 1.2.14. The interpolating family $\{\Psi'_\lambda\}$ is a small perturbation of $\{\Psi_\lambda\}$ which also crosses $\mathfrak{B}_{\text{crit}}$ and exhibits extremal critical collapse. Locally, $\mathfrak{B}$ is foliated by "hypersurfaces" $\mathfrak{B}(\mathfrak{r})$ consisting of black hole spacetimes with asymptotic parameter ratio $\mathfrak{r}$ close to 1.

maximal Cauchy development contains a black hole which remains close to an extremal Reissner–Nordström solution in the domain of outer communication and asymptotically settles down to extremal Reissner–Nordström tangentially along the horizon, with possibly growing "Vlasov hair" transverse to the horizon.

Remark 1.2.15. We emphasize that this type of nonlinear orbital stability for massive Vlasov has not yet been proven even in the subextremal case, where we do not expect horizon hair to occur.

With the conjectured description of the stability properties of the *exterior* of the critical solution at hand, we are now ready to state our conjecture for the *global* stability of the extremal critical collapse families in Theorem 1.2.1. Refer to Fig. 1.7 for a schematic depiction of this conjecture.

Conjecture 1.2.16. *Extremal critical collapse is stable in the following sense: Consider the moduli space $\mathfrak{M}$ of the spherically symmetric Einstein–Maxwell–Vlasov system for particles of mass $\mathfrak{m}$. Let $\{\Psi_\lambda\}$ be one of the interpolating families given by Theorem 1.2.1. Then there exists a "codimension-one submanifold" $\mathfrak{B}_{\text{crit}}$ of $\mathfrak{M}$ such that $\Psi_0 \in \mathfrak{B}_{\text{crit}} \subset \mathfrak{B}$, which has the following properties:*

1. *$\mathfrak{B}_{\text{crit}}$ is critical in the sense that $\mathfrak{B}$ and $\mathfrak{C}$ locally lie on opposite sides of $\mathfrak{B}_{\text{crit}}$.*

2. *If $\mathfrak{m} = 0$ and $\Psi \in \mathfrak{B}_{\text{crit}}$, the domain of outer communication of the maximal Cauchy development of Ψ asymptotically settles down to an extremal Reissner–Nordström black hole as in Conjecture 1.2.12.*

3. *If $\mathfrak{m} > 0$ and $\Psi \in \mathfrak{B}_{\text{crit}}$, the domain of outer communication of the maximal Cauchy development of Ψ remains close to an extremal Reissner–Nordström black hole and the event*

horizon asymptotically settles down to an extremal Reissner–Nordström event horizon as in Conjecture 1.2.14.

Therefore, any nearby interpolating family $\{\Psi'_\lambda\}$ *also intersects* $\mathfrak{B}_{\mathrm{crit}}$ *and exhibits extremal critical collapse.*

Remark 1.2.17. We further conjecture that given $\mathfrak{r} \in [1-\varepsilon, 1]$ for some $\varepsilon > 0$, there exists a one-parameter family of disjoint "codimension-one submanifolds" $\mathfrak{B}(\mathfrak{r}) \subset \mathfrak{B}$, varying "continuously" in $\mathfrak{r}$, such that $\mathfrak{B}(1) = \mathfrak{B}_{\mathrm{crit}}$ and if $\Psi \in \mathfrak{B}(\mathfrak{r})$, then the maximal Cauchy development of Ψ contains a black hole which asymptotes to a Reissner–Nordström black hole with parameter ratio $\mathfrak{r} = e_f/M_f$, where M_f is the final renormalized Hawking mass and e_f is the final charge. One can then interpret equation (8.9.5) below as saying that the families $\{\Psi_\lambda\}$ in Theorem 1.2.1 intersect the foliation $\{\mathfrak{B}(\mathfrak{r})\}$ *transversally*, as depicted in Fig. 1.7.

While one should think that $\mathfrak{B}_{\mathrm{crit}}$ in Conjecture 1.2.16 corresponds to $\mathfrak{M}_{\mathrm{stab}}$ in Conjectures 1.2.12 and 1.2.14, Part 1 of Conjecture 1.2.16 is also a highly nontrivial statement about the *interiors* of the black holes arising from $\mathfrak{B}_{\mathrm{crit}}$. In particular, by the incompleteness theorem, it would imply that there are no trapped surfaces in the maximal developments of any member of $\mathfrak{B}_{\mathrm{crit}}$; see [GL19, Remark 1.8] and the following remark.

Remark 1.2.18. Conjecture 1.2.16 implies that $\mathfrak{B}$ is *locally closed* near Ψ_0: there exists an open set $U \subset \mathfrak{M}$ containing Ψ_0 such that $\mathfrak{B} \cap U$ is closed in U. This property is not expected to hold near other types of critical solutions, such as naked singularties or star-like solutions.

Remark 1.2.19. Using arguments from [LO19, Appendix A], one can show the following statement in the spherically symmetric Einstein–Maxwell-(neutral and massless) scalar field model: If the maximal Cauchy development of a partial Cauchy hypersurface[13] with $\partial_u r < 0$ contains a black hole with asymptotically extremal parameter ratio, then the development does not contain trapped symmetry spheres. The argument uses crucially the constancy of charge and absence of T_{uv} in this model.

1.2.6 Extremal critical collapse of a charged scalar field and in vacuum

It is natural to conjecture the analog of Theorem 1.2.1 for a massless charged scalar field in spherical symmetry:

[13]By this, we mean an asymptotically flat spacelike hypersurface which terminates at a symmetric sphere with positive area-radius. If the charge is nonzero and nondynamical (as in the neutral scalar field model), one cannot have a regular center.

Conjecture 1.2.20. *Extremal critical collapse occurs in the spherically symmetric Einstein–Maxwell-charged scalar field model and there exist critical solutions which are isometric to extremal Reissner–Nordström in the domain of outer communication after sufficiently large advanced time.*

In [KU22] (see Section 4.6.2 below), we showed that a black hole with an extremal Reissner–Nordström domain of outer communication and containing no trapped surfaces can arise from regular one-ended Cauchy data in the spherically symmetric charged scalar field model (see Corollary 4.6.3). The proof is based on a *characteristic gluing* argument, in which we glue a late ingoing cone in the interior of extremal Reissner–Nordström to an ingoing cone in Minkowski space. The desired properties of the spacetime are obtained softly by Cauchy stability arguments. In particular, the method is inadequate to address whether the solution constructed in Corollary 4.6.3 is critical.

It is also natural to conjecture the analog of Theorem 1.2.1 for the Einstein vacuum equations,

$$\mathrm{Ric}(g) = 0, \tag{1.2.6}$$

where the role of extremal Reissner–Nordström is played by the rotating *extremal Kerr* solution.[14]

Conjecture 1.2.21. *Extremal critical collapse occurs in vacuum gravitational collapse and there exist critical solutions which are isometric to extremal Kerr in the domain of outer communication after sufficiently large advanced time.*

In [KU23] (refer to Section 1.1.5 in this dissertation), the present authors constructed examples of vacuum gravitational collapse which are isometric to Kerr black holes with prescribed mass M and specific angular momentum a, where M and a are any Kerr parameters satisfying $0 \le |a|/M \le \mathfrak{a}_0$ for some small positive constant $\mathfrak{a}_0$. The proof does not work for large values of a and whether extremal Kerr black holes can form in gravitational collapse remains open.

If extremal critical collapse involving the Kerr solution does occur, then one may also ask about stability as in Section 1.2.5. In this case, the question hangs on the stability properties of extremal Kerr, which are more delicate than for extremal Reissner–Nordström. While extremal Kerr is mode-stable [TdC20], axisymmetric scalar perturbations have been shown to exhibit the same non-decay and growth hierarchy as general scalar perturbations of extremal Reissner–Nordström [Are12; Are15]. In light of the newly discovered *azimuthal instabilities* of extremal Kerr by Gajic [Gaj23], in which growth of scalar perturbations already occurs at first order of differentiability, the full (in)stability picture of extremal Kerr may be one of spectacular complexity!

[14]Recall also Remark 1.2.9: replacing "vacuum" with "electrovacuum" and "Kerr" with "Kerr–Newman with nonzero charge" in Conjecture 1.2.21 is not possible!

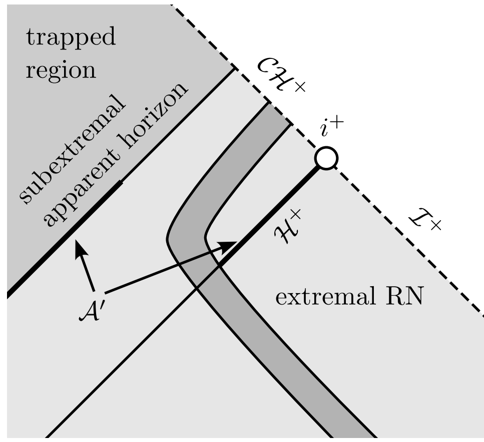

Figure 1.8: Penrose diagram of a counterexample to the third law of black hole thermodynamics in the Einstein–Maxwell–Vlasov model from Theorem 1.2.22. The broken curve $\mathcal{A}'$ is the outermost apparent horizon of the spacetime. This view is zoomed in on the Vlasov beam that charges up the subextremal black hole to extremality. We refer to Fig. 8.6 in Section 8.11.1 for diagrams of the entire spacetime.

1.2.7 Event horizon jumping at extremality

The techniques used to prove Theorem 1.2.1 can also be immediately used to disprove the third law in the Einstein–Maxwell–Vlasov model, which complements our previous disproof in the Einstein–Maxwell-charged scalar field model [KU22]. The present method has the advantage of constructing counterexamples which are past causally geodesically complete, like the spacetimes in Theorem 1.2.1.

Theorem 1.2.22. *There exist smooth solutions of the Einstein–Maxwell–Vlasov system for either massless or massive particles that violate the third law of black hole thermodynamics: a subextremal Reissner–Nordström apparent horizon can evolve into an extremal Reissner–Nordström event horizon in finite advanced time due to the incidence of charged Vlasov matter.*

More precisely, there exist smooth, spherically symmetric, one-ended asymptotically flat Cauchy data for the Einstein–Maxwell–Vlasov system for either massive or massless particles such that the maximal globally hyperbolic development $\mathcal{D}$ has the following properties.

1. *$\mathcal{D}$ contains a nonempty black hole region and for sufficiently large advanced times, the domain of outer communication, including the event horizon $\mathcal{H}^+$, is isometric to that of an extremal Reissner–Nordström solution.*

2. *$\mathcal{D}$ contains a causal diamond which is isometric to a causal diamond in a subextremal Reissner–Nordström black hole, including an appropriate portion of the subextremal apparent horizon. This subextremal region contains an open set of trapped surfaces.*

3. The outermost apparent horizon $\mathcal{A}'$ of $\mathcal{D}$ has at least two connected components. One component of $\mathcal{A}'$ coincides in part with the subextremal apparent horizon and the last component (with respect to v) coincides with the extremal event horizon.

4. $\mathcal{D}$ is past causally goedesically complete, possesses complete null infinities $\mathcal{I}^+$ and $\mathcal{I}^-$, and is isometric to Minkowski space near the center $\{r = 0\}$ for all time.

Refer to Fig. 1.8 for a Penrose diagram of one of these solutions. Note the disconnectedness of the outermost apparent horizon $\mathcal{A}'$, which is necessary in third law violating spacetimes—refer back to the discussion in Section 1.1.3.3. It is striking that the Vlasov beams we construct in the proof of Theorem 1.2.22 do not even touch the subextremal apparent horizon, which should be compared with the hypothetical situation depicted in Fig. 1 of [Isr86]. As with Theorem 1.2.1, Theorem 1.2.22 is proved by desingularizing suitable bouncing charged null dust spacetimes which we construct in Section 7.4.

It is now very natural to ask if some critical behavior can be seen in the examples from Theorem 1.2.22. They are clearly not candidates for critical collapse because they contain trapped surfaces. Nevertheless, by tuning the final charge to mass ratio of the outermost beam in Theorem 1.2.22 (subextremal to superextremal as in Theorem 1.2.1), we construct one-parameter families of solutions satisfying the following

Theorem 1.2.23. *There exist smooth one-parameter families of smooth, spherically symmetric, one-ended asymptotically flat Cauchy data $\{\Psi_\lambda\}_{\lambda\in[-1,1]}$ for the Einstein–Maxwell–Vlasov system for either massive or massless particles with the following properties. Let $\mathcal{D}_\lambda$ be a choice of maximal globally hyperbolic development*[15] *of Ψ_λ for which the double null gauge (u, v) is continuously synchronized as a function of λ. (See already Assumption 8.11.1 and Remark 8.11.2 for the definition of continuous synchronization.) Then the following holds:*

1. For $\lambda \neq 0$, $\mathcal{D}_\lambda$ contains a black hole whose domain of outer communication is isometric to that of a subextremal Reissner–Nordström black hole with mass M_λ and charge $|e_\lambda| < M_\lambda$ for sufficiently large advanced times.

2. $\mathcal{D}_0$ contains a black hole whose domain of outer communication is isometric to that of an extremal Reissner–Nordström black hole with mass M_0 and charge $|e_0| = M_0$ for sufficiently large advanced times.

[15] Typically, one refers to "the" maximal globally hyperbolic development [CG69], which is an equivalence class of isometric developments. In this statement, however, it is crucial that the development comes equipped with a fixed coordinate system.

3. *The location of the event horizon is* <u>*discontinuous*</u> *as a function of* λ*: Let* $u_{\lambda,\mathcal{H}^+}$ *denote the retarded time coordinate of the event horizon* $\mathcal{H}^+_\lambda$ *of* $\mathcal{D}_\lambda$ *with respect to the continuously synchronized gauge* (u,v)*. Then* $\lambda \mapsto u_{\lambda,\mathcal{H}^+}$ *is* <u>*continuous from the left*</u> *but* <u>*discontinuous from the right*</u>*, and*

$$\lim_{\lambda\to 0^+} u_{\lambda,\mathcal{H}^+} > \lim_{\lambda\to 0^-} u_{\lambda,\mathcal{H}^+}. \tag{1.2.7}$$

4. *The functions* $\lambda \mapsto M_\lambda$ *and* $\lambda \mapsto e_\lambda$ *are* <u>*continuous from the left*</u> *but* <u>*discontinuous from the right*</u>*, and*

$$\lim_{\lambda\to 0^+} M_\lambda < \lim_{\lambda\to 0^-} M_\lambda, \quad \lim_{\lambda\to 0^+} |e_\lambda| < \lim_{\lambda\to 0^-} |e_\lambda|, \tag{1.2.8}$$

$$\lim_{\lambda\to 0^+} \frac{|e_\lambda|}{M_\lambda} < \lim_{\lambda\to 0^-} \frac{|e_\lambda|}{M_\lambda} = 1, \quad \lim_{\lambda\to 0^+} r_{\lambda,\mathcal{H}^+} < \lim_{\lambda\to 0^-} r_{\lambda,\mathcal{H}^+}, \tag{1.2.9}$$

where $r_{\lambda,\mathcal{H}^+} \doteq M_\lambda + \sqrt{M_\lambda^2 - e_\lambda^2}$*.*

In addition, for every $\lambda \in [-1,1]$*,* $\mathcal{D}_\lambda$ *is past causally geodesically complete, possesses complete null infinities* $\mathcal{I}^+$ *and* $\mathcal{I}^-$*, and is isometric to Minkowski space near the center* $\{r=0\}$ *for all time.*

From the perspective of the dynamical extremal black hole $\mathcal{D}_0$, an arbitrarily small perturbation to $\mathcal{D}_\lambda$ with $\lambda > 0$ causes the event horizon to *jump* by a definite amount in u (i.e., not $o(1)$ in λ) and the parameter ratio to drop by a definite amount. The proof of Theorem 1.2.23 relies crucially on the absence of trapped surfaces in a double null neighborhood of the horizon in the solutions of Theorem 1.2.22, cf. Fig. 1.8. In the asymptotically subextremal case, trapped surfaces are expected to asymptote towards future timelike infinity i^+. In this case, we prove in Proposition 8.11.4 below that the *location of the event horizon is continuous* as a function of initial data, under very general assumptions in spherical symmetry. Therefore, (1.2.7) is a characteristic feature of extremal black hole formation.

We expect this "local critical behavior" to be *stable* in the sense of Section 1.2.5 and to play a key role in the general stability problem for extremal black holes.

Remark 1.2.24. By a suitable modification of the characteristic gluing techniques in [KU22], Theorem 1.2.23 can be proved for the spherically symmetric Einstein–Maxwell-charged scalar field model, but past completeness of the solutions does not follow immediately from our methods. It is also natural to conjecture analogs for Theorem 1.2.23 in (electro)vacuum; see in particular [DHRT, Section IV.2].[16]

[16] In fact, the statement of Theorem 1.2.23 is not actually reliant on the black holes forming in gravitational collapse and can be made sense of in terms of characteristic data as in Conjectures 1.2.12 and 1.2.14. In this case, one can study the local critical behavior of extremal Reissner–Nordström in electrovacuum since Remark 1.2.9 no longer applies. Indeed, this is precisely the context of the discussion in [DHRT, Section IV.2].

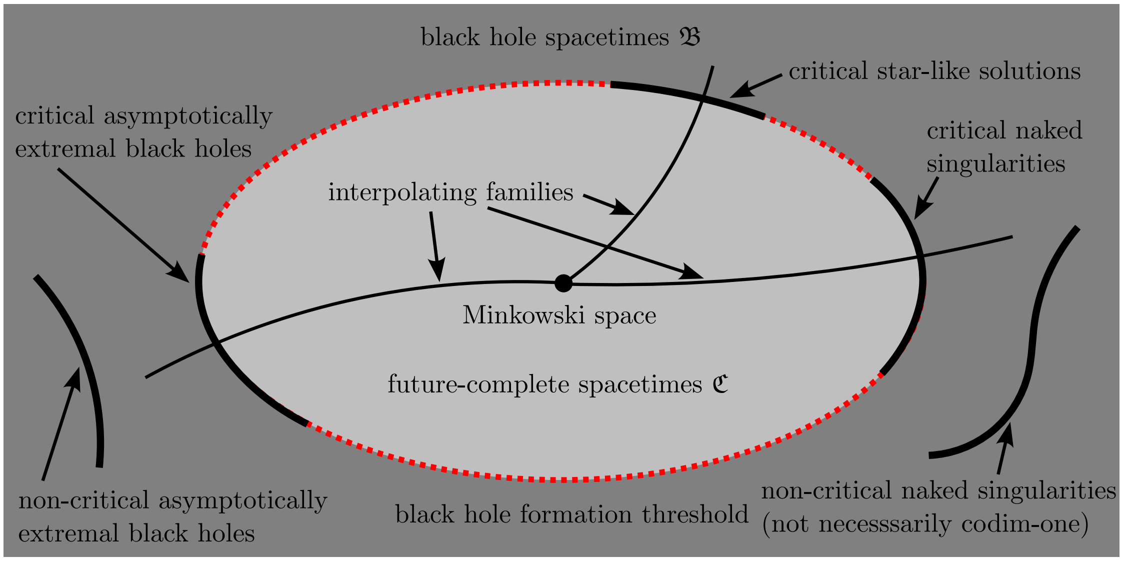

Figure 1.9: A cartoon picture of the conjectured structure of the black hole formation threshold in moduli space. The dashed red line represents the black hole formation threshold which is crossed by three interpolating families of data. On the threshold we have highlighted three distinct regimes of critical collapse: naked singularities, star-like solutions, and asymptotically extremal black holes. These thick black lines represent codimension-one submanifolds. Near the critical extremal black hole threshold, the structure is represented in more detail by Fig. 1.7. The figure also schematically depicts that not all extremal black holes are critical—there are codimension-one submanifolds consisting of asympotically extremal black holes which do not like in the closure of $\mathfrak{C}$, such as those with trapped surfaces far behind the horizon. When crossing such a hypersurface in moduli space, one might be subjected to the discontinuities of Theorem 1.2.23. Furthermore, there might be non-critical naked singularities, but there is no a priori reason to believe that these constitute codimension-one submanifolds.

1.2.8 The conjectural picture of moduli space

We conclude this introduction with a conjectural picture of the qualitative structure of the black hole formation threshold and related phenomena. Refer to Fig. 1.9. The reader is warned that Fig. 1.9 is to be taken with a large grain of salt—the author has taken some artistic liberties in representing the relative sizes, locations, number of connected components, and shapes of the highlighted areas.

Not every part of this picture will be present for every model, of course. For the spherically symmetric scalar field model, we have already remarked that extremal critical collapse cannot happen (Remark 1.2.8). It is not known (or even suggested numerically) whether naked singularities occur in spherically symmetric Einstein–Vlasov models. Therefore, while we have proved in this dissertation that interpolating families exhibiting extremal critical collapse do occur for the spherically symmetric Einstein–Maxwell–Vlasov model, we do not yet have evidence for whether Fig. 1.9 accurately reflects the rest of the moduli space. However, Fig. 1.9 does seem to be a reasonable conjectural representation for the spherically symmetric Einstein–Maxwell–charged Kein–Gordon model or the

Einstein–Maxwell–Vlasov system outside of symmetry.

1.3 Outline of the contents of Part I

Chapter 2: In this chapter, we set up the Einstein–Maxwell-charged scalar field and Einstein–Maxwell–Vlasov systems in spherical symmetry. In Section 2.1, we lay the foundations for the study of spherically symmetric charged spacetimes and electromagnetic geodesics. In Section 2.2 we define the spherically symmetric Einstein–Maxwell-charged scalar field system and in Section 2.3 the spherically symmetric Einstein–Maxwell–Vlasov system. In Section 2.4, we recall Kommemi's a priori characterization of the spacetime boundary which is valid for both of these models. Finally, in Section 2.5 we show that there are no nonspherically symmetric trapped or antitrapped surfaces in a spacetime if there are no spherically symmetric trapped or antitrapped surfaces.

Chapter 3: In this chapter, we study the characteristic intial value problem for the spherically symmetric Einstein–Maxwell-charged scalar field and Einstein–Maxwell–Vlasov systems. We handle the case of a charged scalar field in Section 3.1 and charged Vlasov in Section 3.2. In Section 3.2, we also prove the generalized extension principle for the charged Vlasov model and set up the spacelike/characteristic initial value problem which will be utilized in the construction of extremal critical collapse in Chapter 8. Finally, in Section 3.3 we give a detailed proof of local well-posedness for Einstein–Maxwell–Vlasov.

Chapter 4: In this chapter, we give an overview of the characteristic gluing method and applications to black hole formation. After providing a general outline of the problem in Section 4.1, we explain Aretakis' work on characteristic gluing for the linear wave equation in Section 4.2. In Section 4.3, we recall Aretakis, Czimek, and Rodnianski's work on characteristic gluing for the Einstein vacuum equations near Minkowski space. In Section 4.4, we explain our event horizon gluing result for the spherically symmetric charged scalar field model and give an outline of the proof. In Section 4.5, we explain our event horizon gluing result for slowly rotating Kerr in vacuum. Finally, in Section 4.6 we give several additional applications of our characteristic gluing methods.

Chapter 5: In this chapter, we prove characteristic gluing results in the spherically symmetric charged scalar field model and disprove the third law of black hole thermodynamics for this model. We give a precise definition of *sphere data*, *cone data*, and *characteristic gluing* in Section 5.1 and Section 5.2 and define the reference sphere data in Minkowski, Schwarzschild, and Reissner–

Nordström in Section 5.3. In Section 5.4, we give the precise statements of our gluing theorems, which are then proved in Section 5.5. In Section 5.6, we construct the glued spacetimes and disprove the third law. Finally, in the appendix Section 5.A we show that there is no local mechanism that forces an extremal Killing horizon to not have trapped surfaces behind it (by constructing an explicit example).

Chapter 6: In this chapter, we prove characteristic gluing of Minkowski space to Schwarzschild and very slowly rotating Kerr in vacuum. We recall the setup of the Einstein vacuum equations in double null gauge in Section 6.1 and Section 6.2. In Section 6.3 we define sphere data outside of symmetry and define the reference Minkowski, Schwarzschild, and Kerr sphere data in Section 6.4. In Section 6.5, we recall the near-Minkowski obstruction free gluing of Czimek and Rodnianski [CR22]. In Section 6.6, we prove our vacuum gluing results and then construct the spacetimes in Section 6.7.

Chapter 7: Before turning to the proof of extremal critical collapse in the Einstein–Maxwell–Vlasov model in Chapter 8, we show in this chapter that a singular toy model—Ori's bouncing charged null dust model—exhibits extremal critical collapse. We first recall the definition of the model in Section 7.1. We then introduce a radial parametrization of bouncing charged null dust spacetimes in Section 7.2 in which we teleologically prescribe a regular, spacelike, totally geodesic bounce hypersurface. These spacetimes consist of an explicit ingoing charged Vaidya metric pasted along the radially parametrized bounce hypersurface to an outgoing charged Vaidya metric through a physically motivated surgery procedure. In Sections 7.3 and 7.4, we use the radial parametrization to construct new examples of bouncing charged null dust spacetimes. In Section 7.3, we show that Ori's model exhibits extremal critical collapse (Theorem 7.3.2) and in Section 7.4, we show that the third law of black hole thermodynamics is false in Ori's model (Theorem 7.4.1). We then discuss the fundamental flaws of Ori's model in Section 7.5: the ill-posedness across the bounce hypersurface, the singular nature of the solutions, and the ill-posedness near the center. In Section 7.6, we conclude the chapter with the formal radial charged null dust system in double null gauge which will be important for the setup of our initial data in Chapter 8.

Chapter 8: This chapter is devoted to the construction of extremal critical collapse for the Einstein–Maxwell–Vlasov system, Theorem 1.2.1. The proof relies crucially on a very specific teleological choice of Cauchy data which aims at *desingularizing* the dust examples of extremal critical collapse in Theorem 7.3.2, globally in time. In Section 8.1, we give a detailed guide to the proof of Theorem 1.2.1. In Section 8.2, we define the hierarchy of beam parameters and state the key ingredient in the proof of Theorem 1.2.1: the existence and global structure of outgoing charged Vlasov beams (Proposition 8.2.3). These beams arise from data posed on a Cauchy hypersurface which is analogous to the "bounce hypersurface" associated to Ori's model in Chapter 7. In Section 8.3, we solve the constraint equations and prove estimates for the solution along the initial data hypersurface. Section 8.4 is devoted to estimates for the "near region" establishing the bouncing character of our Vlasov beams. To overcome certain difficulties associated with low momenta, our construction features an "auxiliary beam" which is treated in Section 8.5. Section 8.6 is concerned with the "far region" and in Section 8.7 we prove the sharp dispersive estimates in the case of massive particles. In Section 8.8 we conclude the proof of Proposition 8.2.3. This proposition is then used to prove Theorem 1.2.1 in Section 8.9. Finally, in Section 8.10 we show that in a certain hydrodynamic limit of our parameters, the family of solutions constructed in Theorem 1.2.1 converge in a weak* sense to the family constructed in the charged null dust model in Chapter 7. This result rigorously justifies Ori's bouncing charged null dust construction from [Ori91]. Section 8.11. In this final section, we disprove the third law of black hole thermodynamics for the Einstein–Maxwell–Vlasov model (Theorem 1.2.22) in Section 8.11.1. Section 8.11.2 is concerned with the phenomenon of event horizon jumping at extremality. We first show in Proposition 8.11.4 that for a general class of (so-called *weakly tame*) spherically symmetric Einstein-matter systems the retarded time coordinate of the event horizon is lower semicontinuous as a function of initial data. Secondly, we show by example (Theorem 1.2.23) that event horizon jumping can occur in the Einstein–Maxwell–Vlasov system for extremal horizons, which proves the sharpness of semicontinuity in Proposition 8.11.4.

Chapter 2

Spherically symmetric charged matter models

In this chapter, we introduce basic definitions and properties of Einstein-matter systems and electromagnetic fields in spherical symmetry and formulate the Einstein–Maxwell-charged scalar field and Einstein–Maxwell–Vlasov models.

2.1 The geometry of spherically symmetric charged spacetimes

2.1.1 Double null gauge

Let $(\mathcal{M}, g)$ be a smooth, connected, time-oriented, four-dimensional Lorentzian manifold. We say that $(\mathcal{M}, g)$ is *spherically symmetric* with (possibly empty) *center* of symmetry $\Gamma \subset \mathcal{M}$ if $\mathcal{M} \setminus \Gamma$ splits diffeomorphically as $\mathring{\mathcal{Q}} \times S^2$ with metric

$$g = g_{\mathcal{Q}} + r^2\gamma,$$

where $(\mathcal{Q}, g_{\mathcal{Q}})$, $\mathcal{Q} = \mathring{\mathcal{Q}} \cup \Gamma$, is a (1+1)-dimensional Lorentzian spacetime with boundary Γ (possibly empty), $\gamma \doteq d\vartheta^2 + \sin^2\vartheta\, d\varphi^2$ is the round metric on the unit sphere, and r is a nonnegative function on $\mathcal{Q}$ which can be geometrically interpreted as the area-radius of the orbits of the isometric SO(3) action on $(\mathcal{M}, g)$. In a mild abuse of notation, we denote by Γ both the center of symmetry in $\mathcal{M}$ and its projection to $\mathcal{Q}$. Moreover, if Γ is non-empty, we assume that the SO(3) action fixes Γ and that Γ consists of one timelike geodesic along which $r = 0$. We further assume that $(\mathcal{Q}, g_{\mathcal{Q}})$ admits a *global double-null coordinate system* (u, v) such that the metric g takes the form

$$g = -\Omega^2\, dudv + r^2\gamma \tag{2.1.1}$$

for a positive function $\Omega^2 \doteq -2g_{\mathcal{Q}}(\partial_u, \partial_v)$ on $\mathcal{Q}$ and such that ∂_u and ∂_v are future-directed. Our conventions are so that $u = t - r$ and $v = t + r$ give a double null coordinate system on $(3+1)$-dimensional Minkowski space, with $r = \frac{1}{2}(v - u)$ and $\Omega^2 \equiv 1$. We will also use the notation $\not{g} \doteq r^2\gamma$. The constant u and v curves are null in $(\mathcal{Q}, g_{\mathcal{Q}})$ and correspond to null hypersurfaces "upstairs" in the full spacetime $(\mathcal{M}, g)$. We further assume that along the center Γ, the coordinate v is outgoing and u is ingoing, i.e., $\partial_v r|_\Gamma > 0$, $\partial_u r|_\Gamma < 0$. We will often refer interchangeably to $(\mathcal{M}, g)$ and the reduced spacetime $(\mathcal{Q}, r, \Omega^2)$.

Recall the *Hawking mass* $m : \mathcal{M} \to \mathbb{R}$, which is defined by

$$m \doteq \frac{r}{2}(1 - g(\nabla r, \nabla r))$$

and can be viewed as a function on $\mathcal{Q}$ according to

$$m = \frac{r}{2}\left(1 + \frac{4\partial_u r \partial_v r}{\Omega^2}\right). \tag{2.1.2}$$

We will frequently use the formula

$$\Omega^2 = \frac{4(-\partial_u r)\partial_v r}{1 - \frac{2m}{r}} \tag{2.1.3}$$

to estimate Ω^2 when $1 - \frac{2m}{r} > 0$.

The isometric action of SO(3) on $(\mathcal{M}, g)$ extends to the tangent bundle $T\mathcal{M}$ as follows: Let $\varrho : \mathrm{SO}(3) \to \mathrm{Diff}(\mathcal{M})$ be the representation of SO(3) given by the spherically symmetric ansatz, so that the group action is given by

$$R \cdot x \doteq \varrho(R)(x)$$

for $R \in \mathrm{SO}(3)$ and $x \in \mathcal{M}$. For $(x, p) \in T\mathcal{M}$, we define

$$R \cdot (x, p) \doteq (\varrho(R)(x), \varrho(R)_* p), \tag{2.1.4}$$

where of course $\varrho(R)_* p$ lies in $T_{\varrho(R)(x)}\mathcal{M}$.

Finally, we note that the double null coordinates (u, v) above are not uniquely defined and for any strictly increasing smooth functions $U, V : \mathbb{R} \to \mathbb{R}$, we obtain new global double null coordinates $(\tilde{u}, \tilde{v}) = (U(u), V(v))$ such that $g = -\tilde{\Omega}^2\, d\tilde{u}\, d\tilde{v} + \not{g}$, where $\tilde{\Omega}^2(\tilde{u}, \tilde{v}) = (U'V')^{-1}\Omega^2(U^{-1}(\tilde{u}), V^{-1}(\tilde{v}))$ and $r(\tilde{u}, \tilde{v}) = r(U^{-1}(\tilde{u}), V^{-1}(\tilde{v}))$.

2.1.2 Canonical coordinates on the tangent bundle

Given local coordinates $(\vartheta^1, \vartheta^2)$ on a (proper open subset of) S^2, the quadruple $(u, v, \vartheta^1, \vartheta^2)$ defines a local coordinate system on the spherically symmetric spacetime $(\mathcal{M}, g)$. Given $p \in T_x\mathcal{M}$, we may expand

$$p = p^u \partial_u|_x + p^v \partial_v|_x + p^1 \partial_{\vartheta^1}|_x + p^2 \partial_{\vartheta^2}|_x.$$

The octuple $(u, v, \vartheta^1, \vartheta^2, p^u, p^v, p^1, p^2)$ defines a local coordinate system on $T\mathcal{M}$, and is called a *canonical coordinate system* on $T\mathcal{M}$ dual to $(u, v, \vartheta^1, \vartheta^2)$. One is to think of p as the "momentum coordinate" and x as the "position coordinate." The tangent bundle of $\mathcal{Q}$ trivializes globally as $\mathcal{Q} \times \mathbb{R}^2$, with coordinates p^u and p^v on the second factor. We let

$$\pi : T\mathcal{Q} \to \mathcal{Q}$$

denote the canonical projection.

On a spherically symmetric spacetime, we define the *angular momentum* function by

$$\begin{aligned} \ell : T\mathcal{M} &\to [0, \infty) \\ (x, p) &\mapsto \sqrt{r^2 \not{g}_{AB} p^A p^B}, \end{aligned}$$

where summation over $A, B \in \{1, 2\}$ is implied. This function is independent of the angular coordinate system chosen and is itself spherically symmetric as a function on $T\mathcal{M}$.

Given a double null gauge (u, v), it will be convenient to define a "coordinate time" function

$$\tau \doteq \tfrac{1}{2}(v + u).$$

Associated to this time function is the τ-momentum coordinate

$$p^\tau \doteq \tfrac{1}{2}(p^v + p^u).$$

2.1.3 The Einstein equations and helpful identities in double null gauge

A tensor field on a spherically symmetric spacetime is said to be spherically symmetric if it is itself invariant under the SO(3)-action of the spacetime. If $(\mathcal{M}, g)$ is a spherically symmetric spacetime satisfying the Einstein equations (1.0.1), then the energy-momentum tensor $\mathbf{T}$ is a spherically symmetric, symmetric $(0,2)$-tensor field. We may decompose

$$\mathbf{T} = \mathbf{T}_{uu}\, du^2 + \mathbf{T}_{uv}(du \otimes dv + dv \otimes du) + \mathbf{T}_{vv}\, dv^2 + \mathbf{S}\not{g},$$

where

$$\mathbf{S} \doteq \frac{1}{2}\not{g}^{AB}\mathbf{T}_{AB} = \frac{1}{2}\operatorname{tr}_g \mathbf{T} + \frac{2}{\Omega^2}\mathbf{T}_{uv}.$$

It will be convenient to work with the contravariant energy momentum tensor, which takes the form

$$\mathbf{T}^{\sharp\sharp} = \mathbf{T}^{uu}\partial_u \otimes \partial_u + \mathbf{T}^{uv}(\partial_u \otimes \partial_v + \partial_v \otimes \partial_u) + \mathbf{T}^{vv}\partial_v \otimes \partial_v + \mathbf{S}\not{g}^{-1},$$

where

$$\mathbf{T}_{uu} = \tfrac{1}{4}\Omega^4\mathbf{T}^{vv}, \quad \mathbf{T}_{uv} = \tfrac{1}{4}\Omega^4\mathbf{T}^{uv}, \quad \mathbf{T}_{vv} = \tfrac{1}{4}\Omega^4\mathbf{T}^{uu}.$$

The Christoffel symbols involving null coordinates are given by

$$\begin{aligned}
\Gamma^u_{uu} &= \partial_u \log \Omega^2, & \Gamma^v_{vv} &= \partial_v \log \Omega^2,\\
\Gamma^u_{AB} &= \frac{2\partial_v r}{\Omega^2 r}\not{g}_{AB}, & \Gamma^v_{AB} &= \frac{2\partial_u r}{\Omega^2 r}\not{g}_{AB},\\
\Gamma^A_{Bu} &= \frac{\partial_u r}{r}\delta^A_B, & \Gamma^A_{Bv} &= \frac{\partial_v r}{r}\delta^A_B,
\end{aligned}$$

and the totally spatial Christoffel symbols Γ^A_{BC} are the same as for γ in the coordinates $(\vartheta^1, \vartheta^2)$.

For a spherically symmetric metric g written in double null gauge (2.1.1), the Einstein equations (1.0.1) separate into the wave equations for the geometry,

$$\partial_u\partial_v r = -\frac{\Omega^2}{2r^2}m + \tfrac{1}{4}r\Omega^4\mathbf{T}^{uv}, \tag{2.1.5}$$

$$\partial_u\partial_v \log \Omega^2 = \frac{\Omega^2 m}{r^3} - \tfrac{1}{2}\Omega^4\mathbf{T}^{uv} - \Omega^2\mathbf{S}, \tag{2.1.6}$$

and Raychaudhuri's equations

$$\partial_u \left(\frac{\partial_u r}{\Omega^2} \right) = -\tfrac{1}{4} r \Omega^2 \mathbf{T}^{vv}, \tag{2.1.7}$$

$$\partial_v \left(\frac{\partial_v r}{\Omega^2} \right) = -\tfrac{1}{4} r \Omega^2 \mathbf{T}^{uu}. \tag{2.1.8}$$

The Hawking mass (2.1.2) satisfies the equations

$$\partial_u m = \tfrac{1}{2} r^2 \Omega^2 (\mathbf{T}^{uv} \partial_u r - \mathbf{T}^{vv} \partial_v r), \tag{2.1.9}$$

$$\partial_v m = \tfrac{1}{2} r^2 \Omega^2 (\mathbf{T}^{uv} \partial_v r - \mathbf{T}^{uu} \partial_u r). \tag{2.1.10}$$

If X is a spherically symmetric vector field, then

$$X = X^u \partial_u + X^v \partial_v$$

and X satisfies $\operatorname{div}_g X = 0$ if and only if

$$\partial_u (r^2 \Omega^2 X^u) + \partial_v (r^2 \Omega^2 X^v) = 0. \tag{2.1.11}$$

The contracted Bianchi identity,

$$\operatorname{div}_g \mathbf{T} = 0,$$

which follows from the Einstein equations (1.0.1), implies the following pair of identities:

$$\partial_u (r^2 \Omega^4 \mathbf{T}^{uu}) + \partial_v (r^2 \Omega^4 \mathbf{T}^{uv}) = \partial_v \log \Omega^2 \, r^2 \Omega^4 \mathbf{T}^{uv} - 4r \partial_v r \Omega^2 \mathbf{S}, \tag{2.1.12}$$

$$\partial_v (r^2 \Omega^4 \mathbf{T}^{vv}) + \partial_u (r^2 \Omega^4 \mathbf{T}^{uv}) = \partial_u \log \Omega^2 \, r^2 \Omega^4 \mathbf{T}^{uv} - 4r \partial_u r \Omega^2 \mathbf{S}. \tag{2.1.13}$$

If α is a spherically symmetric two-form which annihilates TS^2, then it may be written as

$$\alpha = -\frac{\Omega^2 f}{2r^2} \, du \wedge dv,$$

where $f : \mathcal{Q} \to \mathbb{R}$ is a smooth function. We then have

$$\nabla^\mu \alpha_{u\mu} = \frac{\partial_u f}{r^2}, \quad \nabla^\mu \alpha_{v\mu} = -\frac{\partial_v f}{r^2}. \tag{2.1.14}$$

2.1.4 Spherically symmetric electromagnetic fields

We will additionally assume that our spherically symmetric spacetime $(\mathcal{M}, g)$ carries a spherically symmetric electromagnetic field with no magnetic charge. The electromagnetic field is represented by a closed two-form F, which takes the Coulomb form

$$F = -\frac{\Omega^2 Q}{2r^2}\, du \wedge dv, \tag{2.1.15}$$

for a function $Q : \mathcal{Q} \to \mathbb{R}$. The number $Q(u, v)$ is the total electric charge enclosed in the (u, v)-symmetry sphere $S_{u,v} \subset \mathcal{M}$, which can be seen from the gauge-invariant formula

$$Q(u, v) = \frac{1}{4\pi} \int_{S_{u,v}} \star F, \tag{2.1.16}$$

where $\star$ is the Hodge star operator and we orient $\mathcal{M}$ by $du \wedge dv \wedge d\vartheta \wedge d\varphi$.

The electromagnetic energy momentum tensor is defined by

$$T^{\mathrm{EM}}_{\mu\nu} \doteq F_{\mu}{}^{\alpha} F_{\nu\alpha} - \tfrac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta}, \tag{2.1.17}$$

and relative to a double null gauge is given by

$$T^{\mathrm{EM}} = \frac{\Omega^2 Q^2}{4r^4}(du \otimes dv + dv \otimes du) + \frac{Q^2}{2r^4} \not{g}$$

in spherical symmetry. If F satisfies Maxwell's equation

$$\nabla^{\alpha} F_{\mu\alpha} = J_{\mu}$$

for a charge current J, then the divergence of the electromagnetic energy momentum tensor satisfies

$$\nabla^{\mu} T^{\mathrm{EM}}_{\mu\nu} = -F_{\nu\alpha} J^{\alpha}. \tag{2.1.18}$$

In spherical symmetry, Maxwell's equations read (see (2.1.14))

$$\partial_u Q = -\tfrac{1}{2} r^2 \Omega^2 J^v, \quad \partial_v Q = \tfrac{1}{2} r^2 \Omega^2 J^u.$$

Finally, we will utilize the *renormalized Hawking mass*

$$\varpi \doteq m + \frac{Q^2}{2r} \tag{2.1.19}$$

to account for the contribution of the electromagnetic field to the Hawking mass m.

2.1.5 The Lorentz force

We next briefly recall the *Lorentz force law* for the motion of a charged particle. Let $(\mathcal{M}, g, F)$ be a charged spacetime, where F is a closed 2-form representing the electromagnetic field. If γ is the worldline of a particle of mass $\mathfrak{m} > 0$ and charge $\mathfrak{e} \in \mathbb{R}$, then it satisfies the Lorentz force equation

$$\mathfrak{m}\frac{Du^\mu}{d\tau} = \mathfrak{e}F^\mu{}_\nu u^\nu,$$

where τ is proper time and $u \doteq d\gamma/d\tau$ (so that $g(u,u) = -1$). Defining the momentum of γ by $p \doteq \mathfrak{m}u$ and rescaling proper time to $s = \mathfrak{m}^{-1}\tau$ (so that $p = d\gamma/ds$), we can rewrite the Lorentz force equation as

$$\frac{Dp^\mu}{ds} = \mathfrak{e}F^\mu{}_\nu p^\nu. \tag{2.1.20}$$

This equation, which we call the *electromagnetic geodesic equation*, makes sense for null curves as well, and can be taken as the equation of motion for all charged particles [Ori91].

Remark 2.1.1. Because the Lorentz force equation (2.1.20) is not quadratic in p, s is *not* an affine parameter, which has fundamental repercussions for the dynamics of the electromagnetic geodesic flow. In particular, trajectories of the Lorentz force with parallel, but not equal, initial velocities will in general be distinct, even up to reparametrization.

The electromagnetic geodesic flow has two fundamental conserved quantities which will feature prominently in this work.

Lemma 2.1.2. *Let $\gamma : I \to \mathcal{M}$ be a causal electromagnetic geodesic in a spherically symmetric charged spacetime, where $I \subset \mathbb{R}$ is an interval. Let $p = d\gamma/ds$. Then the rest mass*

$$\mathfrak{m}[\gamma] \doteq \sqrt{-g_\gamma(p,p)}$$

and the angular momentum

$$\ell[\gamma] \doteq \ell(\gamma, p)$$

are conserved quantities along γ.

Proof. To show that $\mathfrak{m}$ is constant, we compute

$$\frac{d}{ds}g(p,p) = 2g\left(\frac{Dp}{ds},p\right) = 2\mathfrak{e}F(p,p) = 0,$$

where the final equality follows from the antisymmetry of F.

Since ℓ is independent of the coordinates chosen on S^2, we may assume that $(\vartheta^1,\vartheta^2)$ are normal coordinates on S^2 at the point $(\gamma^1(s_0),\gamma^2(s_0))$. To show that ℓ is constant, we then compute

$$\frac{d}{ds}r^4\gamma_{AB}p^Ap^B\bigg|_{s=s_0} = 4r^3\frac{dr}{ds}\gamma_{AB}\dot\gamma^A\dot\gamma^B - 2r^4\gamma_{AB}\left(-2\Gamma^A_{Cu}\dot\gamma^u - 2\Gamma^A_{Cv}\dot\gamma^v\right)\dot\gamma^B\dot\gamma^C = 0,$$

where we used the formulas for the Christoffel symbols in spherical symmetry from Section 2.1.3. □

For an electromagnetic geodesic $\gamma(s)$ with angular momentum $\ell = \ell[\gamma]$ and mass $\mathfrak{m} = \mathfrak{m}[\gamma]$, the Lorentz force law can be written as

$$\frac{d}{ds}p^u = -\partial_u\log\Omega^2(p^u)^2 - \frac{2\partial_v r}{r\Omega^2}\frac{\ell^2}{r^2} - \mathfrak{e}\frac{Q}{r^2}p^u, \tag{2.1.21}$$

$$\frac{d}{ds}p^v = -\partial_v\log\Omega^2(p^v)^2 - \frac{2\partial_u r}{r\Omega^2}\frac{\ell^2}{r^2} + \mathfrak{e}\frac{Q}{r^2}p^v, \tag{2.1.22}$$

$$\Omega^2p^up^v = \frac{\ell^2}{r^2} + \mathfrak{m}^2, \tag{2.1.23}$$

where $p^u \doteq d\gamma^u/ds$, $p^v \doteq d\gamma^v/ds$, and the third equation, known as the *mass shell relation*, is directly equivalent to the definition of mass and angular momentum. In this work, it will not be necessary to follow the angular position of the electromagnetic geodesics in the $(3+1)$-dimensional spacetime. It is very convenient to rewrite (2.1.21) and (2.1.22) as

$$\frac{d}{ds}(\Omega^2p^u) = \left(\partial_v\log\Omega^2 - \frac{2\partial_v r}{r}\right)\frac{\ell^2}{r^2} - \mathfrak{e}\frac{Q}{r^2}(\Omega^2p^u), \tag{2.1.24}$$

$$\frac{d}{ds}(\Omega^2p^v) = \left(\partial_u\log\Omega^2 - \frac{2\partial_u r}{r}\right)\frac{\ell^2}{r^2} + \mathfrak{e}\frac{Q}{r^2}(\Omega^2p^v). \tag{2.1.25}$$

2.2 The Einstein–Maxwell-charged scalar field system

In this section we introduce the spherically symmetric Einstein–Maxwell-charged scalar field (EM-CSF) system, which will feature in Chapter 5. The Einstein–Maxwell-charged scalar field system

$$\begin{aligned} R_{\mu\nu}(g) - \tfrac{1}{2}R(g)g_{\mu\nu} &= 2\left(T^{\mathrm{EM}}_{\mu\nu} + T^{\mathrm{CSF}}_{\mu\nu}\right), \\ \nabla^\mu F_{\mu\nu} &= 2\mathfrak{e}\,\mathrm{Im}(\phi\overline{D_\nu\phi}), \\ g^{\mu\nu}D_\mu D_\nu\phi &= 0, \end{aligned}$$

is invariant with respect to the following gauge transformations

$$\phi \mapsto e^{-i\mathfrak{e}\chi}\phi, \quad A \mapsto A + d\chi \tag{2.2.1}$$

for real-valued functions $\chi = \chi(u,v)$, where $A = A_u du + A_v dv$ is the potential 1-form and $\mathfrak{e}$ is a dimensionful coupling constant representing the charge of the scalar field. More abstractly, the Einstein–Maxwell-scalar field system is a $U(1)$-gauge theory and we refer to [Kom13] for more details. In order to break the symmetry we will use the global electromagnetic gauge

$$A_v = 0 \tag{2.2.2}$$

when discussing this model.

Definition 2.2.1. The *spherically symmetric Einstein–Maxwell-charged scalar field* model with coupling constant $\mathfrak{e} \in \mathbb{R}$ consists of a spherically symmetric charged spacetime $(\mathcal{Q}, r, \Omega^2, Q)$ and a smooth complex-valued scalar field $\phi : \mathcal{Q} \to \mathbb{C}$. The system satisfies the wave equations

$$\partial_u\partial_v\phi = -\frac{\partial_u\phi\partial_v r}{r} - \frac{\partial_u r\partial_v\phi}{r} + \frac{i\mathfrak{e}\Omega^2 Q}{4r^2}\phi - i\mathfrak{e}A_u\frac{\partial_v r}{r}\phi - i\mathfrak{e}A_u\partial_v\phi, \tag{2.2.3}$$

$$\partial_u\partial_v r = -\frac{\Omega^2}{4r} - \frac{\partial_u r\partial_v r}{r} + \frac{\Omega^2}{4r^3}Q^2, \tag{2.2.4}$$

$$\partial_u\partial_v\log(\Omega^2) = \frac{\Omega^2}{2r^2} + 2\frac{\partial_u r\partial_v r}{r^2} - \frac{\Omega^2}{r^4}Q^2 - 2\mathrm{Re}(D_u\phi\overline{\partial_v\phi}), \tag{2.2.5}$$

the Raychaudhuri equations,

$$\partial_u\left(\frac{\partial_u r}{\Omega^2}\right) = -\frac{r}{\Omega^2}|D_u\phi|^2, \tag{2.2.6}$$

$$\partial_v\left(\frac{\partial_v r}{\Omega^2}\right) = -\frac{r}{\Omega^2}|\partial_v\phi|^2, \tag{2.2.7}$$

and the Maxwell equations,

$$\partial_u Q = -\mathfrak{e}r^2 \mathrm{Im}(\phi\overline{D_u\phi}), \tag{2.2.8}$$

$$\partial_v Q = \mathfrak{e}r^2 \mathrm{Im}(\phi\overline{\partial_v\phi}), \tag{2.2.9}$$

$$\partial_v A_u = -\frac{Q\Omega^2}{2r^2} \tag{2.2.10}$$

From these equations we easily derive

$$\partial_v(r\partial_u r) = -\frac{\Omega^2}{4}\left(1-\frac{Q^2}{r^2}\right), \tag{2.2.11}$$

$$\partial_v\partial_u(r\phi) = -\frac{\Omega^2 m}{2r^2}\phi + i\frac{\mathfrak{e}\Omega^2 Q}{4r}\phi + \frac{\Omega^2 Q^2}{4r^3}\phi - i\mathfrak{e}A_u\partial_v(r\phi), \tag{2.2.12}$$

as well as

$$\partial_v m = 2\Omega^{-2}r^2(-\partial_u r)|\partial_v\phi|^2 + \frac{1}{2}\frac{Q^2}{r^2}\partial_v r, \tag{2.2.13}$$

which will be useful later.

2.3 The Einstein–Maxwell–Vlasov system

In this section, we introduce the main matter model considered in Chapter 8, the *Einstein–Maxwell–Vlasov* system. This model describes an ensemble of collisionless self-gravitating charged particles which are either massive or massless. We begin by defining the general system outside of symmetry in Section 2.3.1 and then specialize to the spherically symmetric case in Section 2.3.2. In Sections 3.2.1 and 3.2.2, we formulate the fundamental local theory for this model, local existence and a robust continuation criterion known as the *generalized extension principle*. Finally, in Section 3.2.3, we define a procedure for solving the constraint equations for the spherically symmetric Einstein–Maxwell–Vlasov system.

2.3.1 The general system

2.3.1.1 The volume form on the mass shell

Let $(\mathcal{M}^4, g)$ be a spacetime. For $x \in \mathcal{M}$ and $\mathfrak{m} \geq 0$, we define the (future) *mass shell* at x by

$$P_x^{\mathfrak{m}} \doteq \{p \in T_x\mathcal{M} : p \text{ is future directed and } g_x(p,p) = -\mathfrak{m}^2, p \neq 0\}$$

and the associated smooth fiber bundle

$$P^{\mathfrak{m}} \doteq \bigcup_{x\in\mathcal{M}} P^{\mathfrak{m}}_x,$$

with projection maps $\pi_{\mathfrak{m}} : P^{\mathfrak{m}} \to \mathcal{M}$.

Fix $x \in \mathcal{M}$ and let (x^μ) be normal coordinates at x so that

$$g_x = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2.$$

Let p^μ be dual coordinates to x^μ on $T\mathcal{M}$, defined by $p^\mu = dx^\mu(p)$ for $p \in T_x\mathcal{M}$. For $\mathfrak{m} \geq 0$, the restrictions of p^1, p^2, p^3 to $P^{\mathfrak{m}}_x$, denoted by $\bar{p}^1, \bar{p}^2, \bar{p}^3$, define a global coordinate system on $P^{\mathfrak{m}}_x$, with p^0 determined by

$$p^0 = \sqrt{\mathfrak{m}^2 + |\bar{p}^1|^2 + |\bar{p}^2|^2 + |\bar{p}^3|^2}. \tag{2.3.1}$$

Definition 2.3.1. Let $\mathfrak{m} \geq 0$ and $x \in \mathcal{M}$. The *canonical volume form* $\mu^{\mathfrak{m}}_x \in \Omega^3(P^{\mathfrak{m}}_x)$ is defined by

$$\mu^{\mathfrak{m}}_x = (p^0)^{-1}\, d\bar{p}^1 \wedge d\bar{p}^2 \wedge d\bar{p}^3,$$

in normal coordinates at x, where p^0 is given by (2.3.1).

One can show that this form is independent of the choice of normal coordinates. When $\mathfrak{m} > 0$, $P^{\mathfrak{m}}_x$ is a spacelike hypersurface in $T_x\mathcal{M}$ if it is equipped with the Lorentzian metric g_x. In this case, $\mu^{\mathfrak{m}}_x = \mathfrak{m}^{-1}\omega^{\mathfrak{m}}_x$, where $\omega^{\mathfrak{m}}_x$ is the induced Riemannian volume form on $P^{\mathfrak{m}}_x$. The factor of $\mathfrak{m}^{-1}$ is needed to account for the degeneration of $\omega^{\mathfrak{m}}_x$ as $\mathfrak{m} \to 0$, since P^0_x is a null hypersurface. For more information about the volume form on the mass shell, see [SW77; Rin13; SZ14]. The canonical volume form is uniquely characterized by the following property, which can be found in [SW77, Corollary 5.6.2].

Lemma 2.3.2. *The form $\mu^{\mathfrak{m}}_x$ defined above does not depend on the choice of local coordinates on $\mathcal{M}$. Moreover, it is uniquely characterized by the following property. Let $H(p) \doteq \frac{1}{2}g_x(p,p)$. If α is a 3-form in $T_x\mathcal{M}$ defined along $P^{\mathfrak{m}}_x$ such that*

$$dH \wedge \alpha = \sqrt{-\det g(x)}\, dp^0 \wedge dp^1 \wedge dp^2 \wedge dp^3,$$

then $i^*_{\mathfrak{m}}\alpha = \mu^{\mathfrak{m}}_x$, where $i_{\mathfrak{m}} : P^{\mathfrak{m}}_x \to T_x\mathcal{M}$ denotes the inclusion map.

We denote the integration measure associated to $\mu^{\mathfrak{m}}_x$ by $d\mu^{\mathfrak{m}}_x$. A *distribution function* is a nonnegative function $f \in C^\infty(P^\mathfrak{m})$ which decays sufficiently quickly on the fibers so that the relevant integrals are well-defined and are smooth functions of x. Given a distribution function f we may now define the *number current* N and the *energy momentum tensor* T of f by

$$N^\mu(x) \doteq \int_{P^\mathfrak{m}_x} p^\mu f(x,p)\, d\mu^\mathfrak{m}_x, \quad T^{\mu\nu}(x) \doteq \int_{P^\mathfrak{m}_x} p^\mu p^\nu f(x,p)\, d\mu^\mathfrak{m}_x. \tag{2.3.2}$$

These are readily verified to be tensor fields on $\mathcal{M}$. Taking divergences, we have [Rin13, Appendix D]

$$\nabla_\mu N^\mu = \int_{P^\mathfrak{m}_x} X_0(f)\, d\mu^\mathfrak{m}_x, \quad \nabla_\mu T^{\mu\nu} = \int_{P^\mathfrak{m}_x} p^\nu X_0(f)\, d\mu^\mathfrak{m}_x, \tag{2.3.3}$$

where $X_0 \doteq p^\mu \partial_{x^\mu} - \Gamma^\mu_{\nu\rho} p^\nu p^\rho \partial_{p^\mu} \in \Gamma(TT\mathcal{M})$ is the *geodesic spray* vector field.

2.3.1.2 The equations

Definition 2.3.3. The *Einstein–Maxwell–Vlasov* system for particles of *mass* $\mathfrak{m} \in \mathbb{R}_{\geq 0}$ and *fundamental charge* $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$ consists of a charged spacetime $(\mathcal{M}, g, F)$ and a distribution function $f : P^\mathfrak{m} \to [0,\infty)$ satisfying the following equations:

$$R_{\mu\nu} - \tfrac{1}{2} R g_{\mu\nu} = 2\left(T^{\mathrm{EM}}_{\mu\nu} + T_{\mu\nu}\right), \tag{2.3.4}$$

$$\nabla^\alpha F_{\mu\alpha} = \mathfrak{e} N_\mu, \tag{2.3.5}$$

$$Xf = 0, \tag{2.3.6}$$

where T^{EM} is the electromagnetic energy momentum tensor defined in (2.1.17), $T_{\mu\nu}$ and N_μ are the Vlasov energy-momentum tensor and number current, respectively, defined in (2.3.2), and

$$X \doteq p^\mu \frac{\partial}{\partial x^\mu} - \left(\Gamma^\mu_{\alpha\beta} p^\alpha p^\beta - \mathfrak{e} F^\mu{}_\alpha p^\alpha\right) \frac{\partial}{\partial p^\mu} \in \Gamma(TT\mathcal{M}) \tag{2.3.7}$$

is the *electromagnetic geodesic spray* vector field.

The vector field X is easily seen to be tangent to the mass shell $P^\mathfrak{m}$, which means that the Vlasov equation (2.3.6) is indeed a transport equation on $P^\mathfrak{m}$. The vector field $F^\mu{}_\alpha p^\alpha \partial_{p^\mu}$ is itself tangent to $P^\mathfrak{m}$ and we have the integration by parts formulas

$$\int_{P^\mathfrak{m}_x} F^\mu{}_\alpha p^\alpha \partial_{p^\mu} f\, d\mu^\mathfrak{m}_x = 0, \quad \int_{P^\mathfrak{m}_x} F^\mu{}_\alpha p^\nu p^\alpha \partial_{p^\mu} f\, d\mu^\mathfrak{m}_x = -F^\nu{}_\alpha N^\alpha,$$

which are easily verified in normal coordinates. Combined with (2.3.3) and the transport equation

(2.3.6), we obtain the fundamental conservation laws

$$\nabla_\mu N^\mu = 0, \tag{2.3.8}$$

$$\nabla^\mu T_{\mu\nu} = \mathfrak{e} N^\alpha F_{\nu\alpha}. \tag{2.3.9}$$

We now see that the Einstein–Maxwell–Vlasov system is consistent: (2.3.8) implies that Maxwell's equation (2.3.5) is consistent with antisymmetry of F and (2.3.9) implies (using also (2.1.18)) the contracted Bianchi identity

$$\nabla^\mu \left(T^{\mathrm{EM}}_{\mu\nu} + T_{\mu\nu}\right) = 0. \tag{2.3.10}$$

for the total energy-momentum tensor of the system.

2.3.1.3 Relation to the relativistic Maxwell–Vlasov system

The system (1.2.1)–(1.2.3) includes gravity and thus generalizes the special relativistic *Maxwell–Vlasov system* which is typically written in the form[1] (cf. [Gla96])

$$\begin{aligned}
&\partial_t f_{\mathrm{K}} + \hat{v}\cdot\partial_x f_{\mathrm{K}} + \mathfrak{e}(E + \hat{v}\times B)\cdot\partial_v f_{\mathrm{K}} = 0,\\
&\partial_t E - \nabla\times B = -j_{\mathrm{K}}, \quad \partial_t B + \nabla\times E = 0,\\
&\nabla\cdot E = \rho_{\mathrm{K}}, \quad \nabla\cdot B = 0,
\end{aligned}$$

where $f_{\mathrm{K}}(t,x,v) \geq 0$ is the distribution function, $(t,x,v) \in \mathbb{R}\times\mathbb{R}^3\times\mathbb{R}^3$, E is the electric field, B is the magnetic field, $\hat{v} \doteq (\mathfrak{m}^2 + |v|^2)^{-1/2} v$ is the "relativistic velocity" and has modulus smaller than unity, and

$$\rho_{\mathrm{K}}(t,x) \doteq \mathfrak{e}\int_{\mathbb{R}^3} f(t,x,v)\,dv, \quad j_{\mathrm{K}}(t,x) \doteq \mathfrak{e}\int_{\mathbb{R}^3} \hat{v} f(t,x,v)\,dv.$$

This system is equivalent to the covariant equations (2.3.5) and (2.3.6) in Minkowski space under the identifications $(\sqrt{\mathfrak{m}^2+|v|^2}, v) = p$, $f_{\mathrm{K}}(t,x,v) = f(t,x,\sqrt{\mathfrak{m}^2+|v|^2}, v)$, $E_i = F_{i0}$, $B_i = \frac{1}{2}\varepsilon_i{}^{jk}F_{jk}$, $\rho_{\mathrm{K}} = \mathfrak{e}N^0$, and $(j_{\mathrm{K}})^i = \mathfrak{e}N^i$.

[1] The subscript K stands for "kinetic theory literature."

2.3.2 Spherically symmetric definitions and equations

Let $(\mathcal{Q}, r, \Omega^2)$ be the $(1+1)$-dimensional reduced spacetime associated to a spherically symmetric spacetime. For $\mathfrak{m} \geq 0$, we define the *reduced mass shell* by

$$P^{\mathfrak{m}}_{\mathrm{red}} \doteq \left\{(u, v, p^u, p^v) \in T\mathcal{Q} : \Omega^2(u,v) p^u p^v \geq \mathfrak{m}^2, p^\tau > 0\right\}, \tag{2.3.11}$$

where the second condition in the definition forbids $\Omega^2 p^u p^v = 0$ in the $\mathfrak{m} = 0$ case. The angular momentum function from Section 2.1.2 can be defined on the reduced mass shell by

$$\begin{aligned} \ell : P^{\mathfrak{m}}_{\mathrm{red}} &\to \mathbb{R}_{\geq 0} \\ (u, v, p^u, p^v) &\mapsto r\sqrt{\Omega^2 p^u p^v - \mathfrak{m}^2}. \end{aligned}$$

Note that $\ell > 0$ on P^0_{red}. The definition of ℓ can be rewritten as the fundamental *mass shell relation*

$$\Omega^2 p^u p^v = \frac{\ell^2}{r^2} + \mathfrak{m}^2. \tag{2.3.12}$$

Definition 2.3.4. A *spherically symmetric distribution function* of *massive* (if $\mathfrak{m} > 0$) or *massless* (if $\mathfrak{m} = 0$) *particles* is a smooth function

$$f : P^{\mathfrak{m}}_{\mathrm{red}} \to \mathbb{R}_{\geq 0}.$$

We say that f has *locally compact support in* p if for every compact set $K \subset \mathcal{Q}$ there exists a compact set $K' \subset \mathbb{R}^2$ such that $\mathrm{spt}(f) \cap P^{\mathfrak{m}}_{\mathrm{red}}|_K \subset K \times K'$. We say that f has *locally positive angular momentum* if for every $K \subset \mathcal{Q}$ compact there exists a constant $c_K > 0$ such that $\ell \geq c_K$ on $\mathrm{spt}(f) \cap P^{\mathfrak{m}}_{\mathrm{red}}|_K$.

In order to define appropriate moments of a distribution function f on a spherically symmetric spacetime $(\mathcal{Q}, r, \Omega^2)$, we require that f decays in the momentum variables p^u and p^v. For $\sigma > 0$, $k \geq 0$ an integer, and $K \subset \mathcal{Q}$ compact, we define the norm

$$\|f\|_{C^k_\sigma(P^{\mathfrak{m}}_{\mathrm{red}}|_K)} \doteq \sum_{0 \leq i_1 + i_2 \leq k} \sup_{P^{\mathfrak{m}}_{\mathrm{red}}|_K} \langle p^\tau \rangle^{\sigma + i_2} |\partial_x^{i_1} \partial_p^{i_2} f|, \tag{2.3.13}$$

where $\partial_x^{i_1}\partial_p^{i_2} f$ ranges over all expressions involving i_1 derivatives in the (u,v)-variables and i_2 derivatives in the (p^u, p^v)-variables. If the norm (2.3.13) is finite for all compact sets $K \subset \mathcal{Q}$, we say that

$f \in C^\infty_{\sigma,\mathrm{loc}}(P^{\mathfrak{m}}_{\mathrm{red}})$. If f has locally compact support in p, then it lies in $C^\infty_{\sigma,\mathrm{loc}}(P^{\mathfrak{m}}_{\mathrm{red}})$.

Remark 2.3.5. Our well posedness theory for the Einstein–Maxwell–Vlasov system requires that p-derivatives of f decay faster, which is the reason for the i_2 weight in (2.3.13).

Given a spherically symmetric spacetime $(\mathcal{Q}, r, \Omega^2)$ with distribution function f, we define the *Vlasov number current* by

$$N^u(u,v) \doteq \pi\Omega^2 \int_{\Omega^2 p^u p^v \geq \mathfrak{m}^2} p^u f(u,v,p^u,p^v)\, dp^u dp^v, \tag{2.3.14}$$

$$N^v(u,v) \doteq \pi\Omega^2 \int_{\Omega^2 p^u p^v \geq \mathfrak{m}^2} p^v f(u,v,p^u,p^v)\, dp^u dp^v \tag{2.3.15}$$

and the *Vlasov energy momentum tensor* by

$$T^{uu}(u,v) \doteq \pi\Omega^2 \int_{\Omega^2 p^u p^v \geq \mathfrak{m}^2} (p^u)^2 f(u,v,p^u,p^v)\, dp^u dp^v, \tag{2.3.16}$$

$$T^{uv}(u,v) \doteq \pi\Omega^2 \int_{\Omega^2 p^u p^v \geq \mathfrak{m}^2} p^u p^v f(u,v,p^u,p^v)\, dp^u dp^v, \tag{2.3.17}$$

$$T^{vv}(u,v) \doteq \pi\Omega^2 \int_{\Omega^2 p^u p^v \geq \mathfrak{m}^2} (p^v)^2 f(u,v,p^u,p^v)\, dp^u dp^v, \tag{2.3.18}$$

$$S(u,v) \doteq \frac{\pi}{2}\Omega^2 \int_{\Omega^2 p^u p^v \geq \mathfrak{m}^2} (\Omega^2 p^u p^v - \mathfrak{m}^2) f(u,v,p^u,p^v)\, dp^u dp^v \leq \frac{\Omega^2}{2} T^{uv}(u,v). \tag{2.3.19}$$

If $f \in C^\infty_{\sigma,\mathrm{loc}}(P^{\mathfrak{m}}_{\mathrm{red}})$ with $\sigma > 4$, then these moments are well defined smooth functions of u and v.

Definition 2.3.6. The *spherically symmetric Einstein–Maxwell–Vlasov* model for *particles of mass* $\mathfrak{m} \in \mathbb{R}_{\geq 0}$ and *fundamental charge* $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$ consists of a smooth spherically symmetric charged spacetime $(\mathcal{Q}, r, \Omega^2, Q)$ and a smooth distribution function $f \in C^\infty_{\sigma,\mathrm{loc}}(P^{\mathfrak{m}}_{\mathrm{red}})$ for a decay rate $\sigma > 4$ fixed. When $\mathfrak{m} = 0$, we require that f also has locally positive angular momentum. To emphasize that the distribution functions we consider have these regularity properties in p, we say that such a solution has *admissible momentum*.

The system satisfies the wave equations

$$\partial_u \partial_v r = -\frac{\Omega^2}{2r^2}\left(m - \frac{Q^2}{2r}\right) + \tfrac{1}{4} r \Omega^4 T^{uv}, \tag{2.3.20}$$

$$\partial_u \partial_v \log \Omega^2 = \frac{\Omega^2 m}{r^3} - \frac{\Omega^2 Q^2}{r^4} - \tfrac{1}{2}\Omega^4 T^{uv} - \Omega^2 S, \tag{2.3.21}$$

the Raychaudhuri equations

$$\partial_u\left(\frac{\partial_u r}{\Omega^2}\right) = -\tfrac{1}{4}r\Omega^2 T^{vv}, \tag{2.3.22}$$

$$\partial_v\left(\frac{\partial_v r}{\Omega^2}\right) = -\tfrac{1}{4}r\Omega^2 T^{uu}, \tag{2.3.23}$$

and the Maxwell equations

$$\partial_u Q = -\tfrac{1}{2}\mathfrak{e}r^2\Omega^2 N^v, \tag{2.3.24}$$

$$\partial_v Q = +\tfrac{1}{2}\mathfrak{e}r^2\Omega^2 N^u, \tag{2.3.25}$$

where N^u, N^v, T^{uu}, T^{uv}, T^{vv}, and S are defined by equations (2.3.14)–(2.3.19). Finally, f satisfies the *spherically symmetric Vlasov equation*

$$Xf = 0, \tag{2.3.26}$$

where $X \in \Gamma(TP^{\mathfrak{m}}_{\mathrm{red}})$ is the *reduced electromagnetic geodesic spray*

$$X \doteq p^u\partial_u + p^v\partial_v - \left(\partial_u \log\Omega^2 (p^u)^2 + \frac{2\partial_v r}{r\Omega^2}(\Omega^2 p^u p^v - \mathfrak{m}^2) + \mathfrak{e}\frac{Q}{r^2}p^u\right)\partial_{p^u} \\ - \left(\partial_v \log\Omega^2 (p^v)^2 + \frac{2\partial_u r}{r\Omega^2}(\Omega^2 p^u p^v - \mathfrak{m}^2) - \mathfrak{e}\frac{Q}{r^2}p^v\right)\partial_{p^v}. \tag{2.3.27}$$

Remark 2.3.7. Since $f \geq 0$ for a solution of the Einstein–Maxwell–Vlasov system, the components N^u and N^v of the number current are nonnegative. It follows from the Maxwell equations (2.3.24) and (2.3.25) that $\mathfrak{e}Q$ is decreasing in u and increasing in v, unconditionally. This monotonicity property is a fundamental feature of the spherically symmetric Einstein–Maxwell–Vlasov system and will be exploited several times in this work.

Remark 2.3.8. Both the electromagnetic energy-momentum tensor T^{EM} and the Vlasov energy-momentum tensor T of the Einstein–Maxwell–Vlasov system satisfy the dominant energy condition.

Remark 2.3.9. As an abuse of notation, we have denoted the spray (2.3.7) on $T\mathcal{M}$ and the spray (2.3.27) on $T\mathcal{Q}$ by the same letter X. It will always be clear from the context which vector field we are referring to. They are related by the pushforward along the natural projection map $P^{\mathfrak{m}} \to P^{\mathfrak{m}}_{\mathrm{red}}$.

For a solution of the Einstein–Maxwell–Vlasov system, the Hawking mass m satisfies the equations

$$\partial_u m = \tfrac{1}{2} r^2\Omega^2 (T^{uv}\partial_u r - T^{vv}\partial_v r) + \frac{Q^2}{2r^2}\partial_u r, \tag{2.3.28}$$

$$\partial_v m = \tfrac{1}{2} r^2\Omega^2 (T^{uv}\partial_v r - T^{uu}\partial_u r) + \frac{Q^2}{2r^2}\partial_v r, \tag{2.3.29}$$

which can be derived from (2.1.9) and (2.1.10). The modified Hawking mass ϖ can then be seen to satisfy

$$\partial_u \varpi = \tfrac{1}{2} r^2\Omega^2 (T^{uv}\partial_u r - T^{vv}\partial_v r) - \tfrac{1}{2}\mathfrak{e} r\Omega^2 Q N^v, \tag{2.3.30}$$

$$\partial_v \varpi = \tfrac{1}{2} r^2\Omega^2 (T^{uv}\partial_v r - T^{uu}\partial_u r) + \tfrac{1}{2}\mathfrak{e} r\Omega^2 Q N^u. \tag{2.3.31}$$

The particle current N satisfies the conservation law

$$\partial_u(r^2\Omega^2 N^u) + \partial_v(r^2\Omega^2 N^v) = 0 \tag{2.3.32}$$

by (2.1.11) and (2.3.8). Alternatively, it can be directly derived from the spherically symmetric Vlasov equation, which we will do the proof of Proposition 3.2.3 in Section 3.3. Finally, for the Einstein–Maxwell–Vlasov system, the Bianchi identities (2.1.12) and (2.1.13) read

$$\partial_u(r^2\Omega^4 T^{uu}) + \partial_v(r^2\Omega^4 T^{uv}) = \partial_v \log\Omega^2\, r^2\Omega^4 T^{uv} - 4r\partial_v r\Omega^2 S - \mathfrak{e}\Omega^4 Q N^u, \tag{2.3.33}$$

$$\partial_v(r^2\Omega^4 T^{vv}) + \partial_u(r^2\Omega^4 T^{uv}) = \partial_u \log\Omega^2\, r^2\Omega^4 T^{uv} - 4r\partial_u r\Omega^2 S + \mathfrak{e}\Omega^4 Q N^v. \tag{2.3.34}$$

Again, this follows either from (2.3.10) or directly from the spherically symmetric equations.

2.3.2.1 The spherically symmetric reduction

Proposition 2.3.10. *Let $(\mathcal{Q}, r, \Omega^2, Q, f_{\mathrm{sph}})$ be a solution of the spherically symmetric Einstein–Maxwell–Vlasov system as defined by Definition 2.3.6. Then $(\mathcal{M}, g, F, f)$ solves the Einstein–*

Maxwell–Vlasov system if we lift the solution according to

$$\mathcal{M} \doteq \mathcal{Q} \times S^2,$$

$$g \doteq -\frac{\Omega^2}{2}(du \otimes dv + dv \otimes du) + r^2\gamma, \tag{2.3.35}$$

$$F \doteq -\frac{Q}{2r^2} du \wedge dv, \tag{2.3.36}$$

$$f(u, v, \vartheta^1, \vartheta^2, p^u, p^v, p^1, p^2) \doteq f_{\mathrm{sph}}(u, v, p^u, p^v), \tag{2.3.37}$$

where $(\vartheta^1, \vartheta^2)$ *is a local coordinate system on* S^2*.*

Note that f in (2.3.37) is SO(3)-invariant as a function on $T\mathcal{M}$ as defined in Section 2.1.1.

Proof. As the equations (2.3.20)–(2.3.25) are equivalent to the Einstein equations and Maxwell equations, we must only check that f, defined by (2.3.37), satisfies $Xf = 0$, where X is given by (2.3.7), and that the spherically symmetric formulas (2.3.14)–(2.3.19) appropriately reconstruct the $(3+1)$-dimensional number current and energy-momentum tensor.

Let $\gamma(s)$ be an electromagnetic geodesic. Then $Xf = 0$ at $(\gamma(s_0), \dot\gamma(s_0)) \in P^{\mathfrak{m}}$ is equivalent to

$$\frac{d}{ds}\bigg|_{s=s_0} f(\gamma(s), \dot\gamma(s)) = 0. \tag{2.3.38}$$

Using the chain rule

$$\frac{d}{ds}\bigg|_{s=s_0} f_{\mathrm{sph}}(\gamma^u(s), \gamma^v(s), p^u(s), p^v(s)) = p^u \partial_u f_{\mathrm{sph}} + p^v \partial_v f_{\mathrm{sph}} + \frac{dp^u}{ds}\partial_{p^u} f_{\mathrm{sph}} + \frac{dp^v}{ds}\partial_{p^v} f_{\mathrm{sph}},$$

the spherically symmetric Lorentz force equations (2.1.21) and (2.1.22), the mass shell relation (2.1.23), and the spherically symmetric Vlasov equation (2.3.26), we see that (2.3.38) holds.

To compute the energy-momentum tensor in spherical symmetry, we use Lemma 2.3.2. Let $x \in \mathcal{M}$ and take $(\vartheta^1, \vartheta^2)$ to be local coordinates on S^2 which are normal at the spherical component of x, so that $\gamma_{AB} = \delta_{AB}$. We have

$$dH = -\frac{\Omega^2}{2} p^v dp^u - \frac{\Omega^2}{2} p^u dp^v + r^2 p^1 dp^1 + r^2 p^2 dp^2.$$

Therefore, if we define

$$\alpha \doteq -r^2 (p^v)^{-1} dp^v \wedge dp^1 \wedge dp^2,$$

then

$$dH \wedge \alpha = \tfrac{1}{2}\Omega^2 r^2\, dp^u \wedge dp^v \wedge dp^1 \wedge dp^2.$$

Therefore, by Lemma 2.3.2,

$$d\mu_x^{\mathfrak{m}} = r^2 (p^v)^{-1}\, dp^v dp^1 dp^2$$

as measures on $(0,\infty)\times\mathbb{R}^2_{(p^1,p^2)}$. The remaining momentum variable p^u is obtained from p^v, p^1, and p^2 via the mass shell relation (2.3.12). If we set $\tan\beta = p^2/p^1$ and use that $\ell^2 = r^4((p^1)^2+(p^2)^2)$, then

$$d\mu_x^{\mathfrak{m}} = r^{-2}(p^v)^{-1} dp^v\, \ell\, d\ell\, d\beta.$$

For any weight function $w = w(p^u, p^v)$, we therefore have

$$\int_{(0,\infty)\times\mathbb{R}^2} wf\, d\mu_x^{\mathfrak{m}} = r^{-2}\int_0^{2\pi}\int_0^\infty\int_0^\infty w\left(\frac{\ell^2+\mathfrak{m}^2r^2}{r^2\Omega^2p^v}, p^v\right) f_{\mathrm{sph}}\left(u, v, \frac{\ell^2+\mathfrak{m}^2 r^2}{r^2\Omega^2 p^v}, p^v\right)\frac{dp^v}{p^v}\,\ell\, d\ell\, d\beta. \tag{2.3.39}$$

Integrating out β and applying the coordinate transformation $(p^v, \ell)\mapsto (p^u, p^v)$ reproduces the formulas (2.3.14)–(2.3.19) for N and T. □

Remark 2.3.11. Other works on the spherically symmetric Einstein–Vlasov system in double null gauge, such as [DR16; Mos18; Mos23], represent the distribution function f differently, opting to (at least implicitly) eliminate either p^u or p^v in terms of ℓ using the mass shell relation (2.3.12). This leads to different formulas for N^μ and $T^{\mu\nu}$, as these will then involve an integral over ℓ, as in (2.3.39). To make this precise, we can define the *outgoing representation*[2] of the spherically symmetric f by

$$f_\nearrow(u, v, p^v, \ell) \doteq f\left(u, v, \frac{\ell^2 + r^2\mathfrak{m}^2}{r^2\Omega^2 p^v}, p^v\right), \tag{2.3.40}$$

and (2.3.39) implies, for instance,

$$T^{uv} = \frac{2\pi}{r^4\Omega^2}\int_0^\infty\int_0^\infty \frac{\ell^2 + r^2\mathfrak{m}^2}{p^v} f_\nearrow(u, v, p^v, \ell)\, dp^v\, \ell\, d\ell.$$

The outgoing representation may be taken as an alternative *definition* of the spherically symmetric Vlasov system. We have chosen the formulation here in terms of p^u and p^v because of its explicit symmetry, which is key for constructing time-symmetric initial data in the proof of Theorem 1.2.1. We have also chosen to always write N^μ and $T^{\mu\nu}$ in contravariant form, so that N^u is associated with p^u, etc. This causes extra factors of $g_{uv} = -\frac{1}{2}\Omega^2$ to appear in various formulas, compared to

[2] Of course, we may also define an *ingoing representation* $f_\nwarrow(u, v, p^u, \ell)$ by interchanging p^u and p^v.

[DR16; Mos18; Mos23].

2.3.2.2 Previous work on the Einstein–Maxwell–Vlasov system

Besides the general local existence result of [BC73], the Einstein–Maxwell–Vlasov model does not seem to have been extensively studied. Dispersion for small data solutions of the Einstein–Maxwell–massive Vlasov model (stability of Minkowski space) in spherical symmetry was proved by Noundjeu [Nou05]. See also [NN04] for local well-posedness in Schwarzschild coordinates and [NNR04] for the existence of nontrivial solutions of the constraints. Many static solutions are known to exist for the massive system, first studied numerically by Andréasson–Eklund–Rein [AER09] and proved rigorously by Thaller [Tha19] in spherical symmetry. Thaller has also shown that stationary and axisymmetric (but not spherically symmetric) solutions exist [Tha20].

2.4 Maximal future developments of asymptotically flat data and the a priori characterization of the boundary

The theorems and corollaries in this dissertation are stated in the framework of the Cauchy problem for the Einstein–Maxwell-charged scalar field and Einstein–Maxwell–Vlasov systems. Cauchy data for these systems consist of the usual Cauchy data (Σ, g_0, k_0) for the Einstein equations, where Σ is a 3-manifold, g_0 a Riemannian metric on Σ, and k_0 a symmetric 2-tensor field, together with initial data for the matter fields, namely initial electric and magnetic fields, E_0 and B_0, and finally the scalar field ϕ_0 and its "time derivative" ϕ_1 or the distribution function f_0. (See e.g. [Cho09, Section VI.10] for a treatment of the Einstein–Maxwell Cauchy problem or [Rin13] for the Einstein–Vlasov Cauchy problem.) Associated to a Cauchy data set is a unique maximal future globally hyperbolic development $(\mathcal{M}^4, g, F, A, \phi)$ or $(\mathcal{M}^4, g, F, f)$ [Fou52; CG69]. If the Cauchy data are moreover spherically symmetric, then the maximal development will be spherically symmetric by uniqueness.

In the context of gluing constructions, we will not, however, actually construct our spacetimes by directly evolving Cauchy data. Rather, we construct the spacetimes *teleologically* by gluing together explicit spacetimes with the help of our characteristic gluing results and Proposition 5.2.4. In each case, a Cauchy hypersurface Σ is then found, within the spacetime, whose future domain of dependence contains the physically relevant region, and contains no antitrapped spheres. At this point, all attention is restricted to this future domain of dependence. *A posteriori*, by the existence and uniqueness theory for the maximal globally hyperbolic development, the spacetime will then be contained in the maximal development of the induced data on the Cauchy hypersurface Σ.

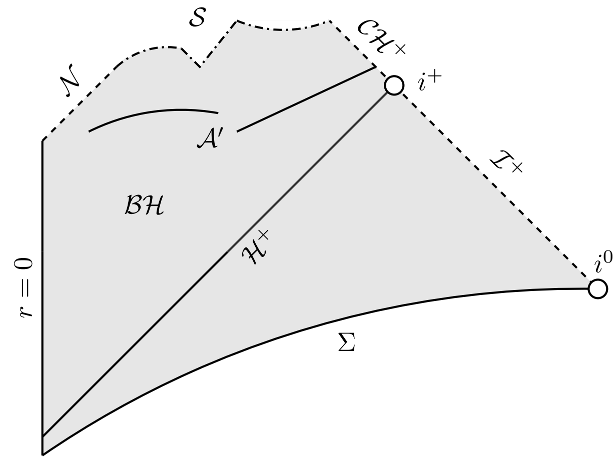

Figure 2.1: General structure of the MFGHD of asymptotically flat Cauchy data Σ in the EMCSF system in spherical symmetry [Kom13]. What is depicted is the quotient manifold $\mathcal{Q}$ as a bounded subset of $\mathbb{R}^{1+1}_{u,v}$ with boundary suitably labeled. Note that various components of the diagram can be empty.

Since all of the examples constructed in this dissertation are maximal globally hyperbolic developments of asymptotically flat, spherically symmetric Cauchy data with no antitrapped spheres of symmetry for Einstein-matter system satisfying the generalized extension principle, we can make use of a general characterization of the boundary of spacetime in this context appearing in [Kom13].[3] In particular one can rigorously associate a global Penrose diagram, and unambiguously identify a nonempty null boundary component *future null infinity* $\mathcal{I}^+$, *domain of outer communication* $J^-(\mathcal{I}^+)$, (possibly empty) *black hole region* $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+)$, (possibly empty) *event horizon* $\mathcal{H}^+ \doteq \partial(\mathcal{BH})$, (possibly empty) *Cauchy horizon* $\mathcal{CH}^+$, (possibly empty) $r = 0$ singularity $\mathcal{S}$, and (possibly empty) null boundary component $\mathcal{N}$ emanating from a (possibly absent) "locally naked" singularity at the center. The Penrose diagram $\mathcal{Q} \subset \mathbb{R}^{1+1}_{u,v}$ can be viewed as a global double null chart for the spacetime, with v the "outgoing" null coordinate and u the "ingoing" coordinate. See Fig. 2.1.[4]

For use in the statement and proofs of many of the results in this dissertation, we recall that the *apparent horizon* is defined by

$$\mathcal{A} \doteq \{\partial_v r = 0\} \subset \mathcal{BH}.$$

[3]The generalized extension principle is proved for EMCSF in [Kom13] and for Einstein–Maxwell–Vlasov in Section 3.2.2 below.

[4]Note that the above general boundary decomposition in particular proves that one cannot form a globally naked singularity once a marginally trapped surface has developed in the spacetime, which already rules out naked singularity formation by supercharging a black hole in spherical symmetry, see [Kom13, Section 1.9]. It is thus not at all surprising that ongoing numerical searches for these continue to be futile.

Since $\mathcal{A}$ might have a complicated structure (in particular, it might have nonempty interior), we define an appropriate notion of boundary as follows. The *outermost apparent horizon* $\mathcal{A}'$ consists of those points $p \in \mathcal{A}$ whose past-directed ingoing null segment lies in the strictly untrapped region $\{\partial_v r > 0\}$ and eventually exits the black hole region, i.e., enters $J^-(\mathcal{I}^+)$. $\mathcal{A}'$ is a possibly disconnected achronal curve in the $(1+1)$-dimensional reduction $\mathcal{Q}$ of $\mathcal{M}$. Note, as depicted in Fig. 2.1, that $\mathcal{A}'$ does not necessarily asymptote to future timelike infinity i^+.

For definiteness, we will make extensive use of these notions in our theorems and corollaries. However, our notation and usage should be sufficiently familiar to readers acquainted with standard concepts in general relativity so that they may read our diagrams and understand our theorems without specific reference to [Kom13].

We also note that when referring to spherically symmetric subsets of $(\mathcal{M}, g)$, such as the event horizon $\mathcal{H}^+$, we may view them as objects in $\mathcal{M}$ or in the reduced space $\mathcal{Q}$. The context will make it clear which point of view we are taking.

Remark 2.4.1. In the following section, we show by a barrier argument that since $\partial_u r < 0$ in a spacetime satisfying the hypotheses of [Kom13], there are also no nonspherically symmetric antitrapped surfaces.

2.5 General trapped and antitrapped surfaces in spherically symmetric spacetimes

In this section we infer the absence of nonspherically symmetric trapped or antitrapped surfaces from the absence of spherically symmetric trapped or antitrapped surfaces. This technical result will be used later in Chapter 5.

Our definition of trapped surface is completely standard, see Definition 2.5.6 below. (Note that we assume trapped surfaces to be closed and strictly trapped.) Our definition of antitrapped is as in [Chr93; Kom13], i.e., an antitrapped surface is closed and past weakly outer trapped, see Definition 2.5.7 below.

Proposition 2.5.1. *Let $(\mathcal{M}^4, g)$ be a spherically symmetric spacetime as defined in Section 2.1.1. Then there are no trapped surfaces contained in the sets*

$$A \doteq \{p \in \mathcal{M} : \partial_u r \geq 0\}, \tag{2.5.1}$$

$$B \doteq \{p \in \mathcal{M} : \partial_v r \geq 0\}. \tag{2.5.2}$$

Remark 2.5.2. Note that there could be trapped surfaces contained in $A \cup B$. There might also be trapped surfaces which merely intersect A or B.

Proposition 2.5.3. *Let* $(\mathcal{M}^4, g, F, A, \phi)$ *be a spherically symmetric spacetime arising as the maximal future globally hyperbolic development from one-ended asymptotically flat Cauchy data for the EMCSF system with no antitrapped spheres of symmetry as in [Kom13]. Then:*

1. *If* S *is a trapped surface in* $\mathcal{M}$*, then* $S \cap J^-(\mathcal{I}^+) = \emptyset$.
2. $\mathcal{M}$ *does not contain any antitrapped surfaces.*

Remark 2.5.4. Under stronger assumptions on $\mathcal{I}^+$, the first part of the previous proposition would follow from a classical result of Hawking [Haw72b; HE73, Proposition 9.2.1].

For the proofs, we recall some facts from Lorentzian geometry [Gal00]. Let H be a null hypersurface in a spacetime $(\mathcal{M}^4, g)$, i.e., H is a 3-dimensional submanifold of $\mathcal{M}$ and admits a future-directed normal vector field L which is null and whose integral curves can be reparametrized to be null geodesics. We say that L is a *(future-directed) null generator* of H.

The *second fundamental form of* H *with respect to* L is given by

$$B^L(X, Y) = g(\nabla_X L, Y) \tag{2.5.3}$$

for $X, Y \in TH$. If e_1 and e_2 are an orthonormal pair of spacelike vectors at $p \in H$, we define the *null expansion of* H *with respect to* L by

$$\theta^L = B^L(e_1, e_1) + B^L(e_2, e_2) \tag{2.5.4}$$

at p, and this definition is independent of the pair e_1 and e_2. If $\tilde{L}$ is another future-directed null generator of H, then there is a positive function f on H such that $\tilde{L} = fL$. In this case, we have

$$\theta^{\tilde{L}} = f\theta^L. \tag{2.5.5}$$

Lemma 2.5.5 (Comparison principle for null hypersurfaces)**.** *Let* H_1 *and* H_2 *be null hypersurfaces in* $(\mathcal{M}^4, g)$*, with* H_1 *to the future of* H_2 *and generated by* L_1 *and* L_2*, respectively. If* H_1 *and* H_2 *are tangent at a point* p*, and* $L_1(p) = L_2(p)$*, then*

$$\theta^{L_1}_{H_1}(p) \geq \theta^{L_2}_{H_2}(p). \tag{2.5.6}$$

Proof. By (2.5.5), it suffices to prove (2.5.6) with respect to some choice of null generators of H_1 and H_2 which agree at p. Let (t, x, y, z) be normal coordinates for g based at p so that ∂_t is future-directed and $\{\frac{1}{2}(\partial_t + \partial_x), \partial_y, \partial_z\}$ spans $T_pH_1 = T_pH_2$. We introduce approximate null coordinates $u = t - x$ and $v = t + x$, so that

$$\partial_u = \tfrac{1}{2}(\partial_t - \partial_x), \quad \partial_v = \tfrac{1}{2}(\partial_t + \partial_x).$$

Note that ∂_u and ∂_v are only guaranteed to be null at p.

By the implicit function theorem, there exist functions $f_1(v, y, z)$ and $f_2(v, y, z)$ defined near p, so that, upon defining

$$\zeta_1(u, v, y, z) \doteq f_1(v, y, z) - u, \quad \zeta_2(u, v, y, z) \doteq f_2(v, y, z) - u,$$

we have $H_i = \{\zeta_i = 0\}$ for $i = 1, 2$. Note that $f_1(p) = f_2(p) = 0$ and that p is a critical point for f_1 and f_2. The vector fields $Z_i = \operatorname{grad} \zeta_i$ are null on H_i and define there future-directed null generators. In particular, we have $Z_1(p) = Z_2(p) = \partial_v|_p$.

We first show that $f_1 \geq f_2$ near p. If a point $q = (u, v, y, z)$ lies to the past of H_1, then $\zeta_1(q) \geq 0$. If $q \in H_2$, then $\zeta_2(q) = 0$, so combining these inequalities yields

$$f_1(v, y, z) = \zeta_1(q) + u \geq \zeta_2(q) + u = f_2(v, y, z),$$

as claimed.

We now show that

$$B^{Z_1}_{H_1}(\partial_y, \partial_y)(p) \geq B^{Z_2}_{H_2}(\partial_y, \partial_y)(p), \tag{2.5.7}$$

the corresponding statement and proof for ∂_z being the same. By (2.5.4) this will complete the proof. Since $f_1 \geq f_2$ near p, p is a local minimum for $f_1 - f_2$. It follows that

$$\partial_y^2(f_1 - f_2)(p) \geq 0 \tag{2.5.8}$$

by the second derivative test. Since we are working in a normal coordinate system,

$$B^{Z_i}_{H_i}(\partial_y, \partial_y)(p) = g(\nabla_{\partial_y} \nabla \zeta_i, \partial_y)(p) = \partial_y^2 f_i(p),$$

whence (2.5.8) proves (2.5.7), which completes the proof. □

Definition 2.5.6. A closed spacelike 2-surface S in a spacetime $(\mathcal{M}^4, g)$ is always the intersection of two locally defined null hypersurfaces. We say that S is *trapped* if both of these hypersurfaces have negative future null expansion along S.

Proof of Proposition 2.5.1. We show that there is no trapped surface $S \subset B$. The argument for $S \subset A$ is analogous after noting that $A \cap \Gamma = \emptyset$ by our definition of spherical symmetry and convention for u.

Let $S \subset \{\partial_v r \geq 0\}$ be a closed 2-surface. Let $\pi : \mathcal{M} \to \mathcal{Q}$ be the projection of the spherically symmetric spacetime to its Penrose diagram. Then $\pi(S)$ is a compact subset of $\mathcal{Q}$ and hence u attains a minimum u_0 on $\pi(S)$.

Therefore, there exists a symmetry sphere S_{u_0,v_0} on which $\partial_v r \geq 0$ such that S lies to the future of C_{u_0} and is tangent to this cone at a point $p \in S_{u_0,v_0}$. Note that $p \notin \Gamma$ because C_{u_0} is not regular there. The condition $\partial_v r \geq 0$ means C_{u_0} has nonnegative future expansion. By Lemma 2.5.5, one of the two null hypersurfaces emanating from S also has nonnegative future expansion, so S is not trapped. □

Definition 2.5.7. Let $(\mathcal{M}^4, g)$ be a spacetime satisfying the hypotheses of Proposition 2.5.3. A closed spacelike 2-surface S which bounds a compact spacelike hypersurface Ω is said to be *antitrapped* if its future-directed inward null expansion is nonnegative. Here the (locally defined) inward null hypersurface H_{in} emanating from S is chosen to be the one which smoothly extends the boundary of the causal past of Ω.

Proof of Proposition 2.5.3. 1. Since $r \to \infty$ at $\mathcal{I}^+$ [Kom13], Raychaudhuri's equation (2.2.7) implies $\partial_v r > 0$ in $J^-(\mathcal{I}^+)$. Let S be a closed 2-surface such that $S \cap J^-(\mathcal{I}^+) \neq \emptyset$. Let $\pi : \mathcal{M} \to \mathcal{Q}$ be the projection to the Penrose diagram. Then u attains a minimum u_0 on $\pi(S)$. By the causal properties of $J^-(\mathcal{I}^+)$, there exists a symmetry sphere $S_{u_0,v_0} \subset J^-(\mathcal{I}^+)$ such that S lies to the future of C_{u_0} and is tangent to the cone at $p \in S_{u_0,v_0}$. Arguing as in the proof of Proposition 2.5.1, we see that one of the null hypersurfaces emanating from S has positive future expansion, so S is not trapped.

2. Let $\pi : \mathcal{M} \to \mathcal{Q}$ be again the projection. Then v attains a maximum v_0 on $\pi(S)$ and again there exists a non-central symmetry sphere S_{u_0,v_0} such that $\partial_u r(u_0, v_0) < 0$, S lies to the past of C_{v_0}, and is tangent to the cone at a point $p \in S_{u_0,v_0}$. Now C_{v_0} is tangent to H_{in} at p and lies to the future, so by Lemma 2.5.5, H_{in} has negative null expansion at p. Therefore, S is not antitrapped. □

Chapter 3

The characteristic initial value problem in spherical symmetry

3.1 The Einstein–Maxwell-charged scalar field system

In this section, we give a detailed explanation of the setup and characteristic initial value problem for the Einstein equations with charged scalar fields in spherical symmetry, with a view towards the characteristic gluing problem. See [Kom13] for more details on the EMCSF system.

3.1.1 Bifurcate characteristic data

We first define precisely what we mean by a C^k solution. For now, we may restrict attention to solutions away from the center.

Definition 3.1.1. Let $k \in \mathbb{N}$. A C^k *solution* for the Einstein–Maxwell-charged scalar field system in the EM gauge (2.2.2) consists of a domain $\mathcal{Q} \subset \mathbb{R}^{1+1}_{u,v}$ and functions $r \in C^{k+1}(\mathcal{Q})$ and $\Omega^2, \phi, Q, A_u \in C^k(\mathcal{Q})$, such that $r > 0$, $\Omega^2 > 0$, ϕ is complex-valued, $\partial_v^{k+1} A_u \in C^0(\mathcal{Q})$, and the functions satisfy[1] equations (2.2.3)–(2.2.7).

Next, we formulate the characteristic initial value problem for this class of solutions. Let $\mathbb{R}^{1+1}_{u,v}$ denote the standard (1+1)-dimensional Minkowski space. We introduce the *bifurcate null hypersurface*

[1] Note that the wave equations (2.2.3) and (2.2.5) can readily be interpreted for $k = 1$.

$C \cup \underline{C} \subset \mathbb{R}^{1+1}_{u,v}$, where

$$C \doteq C_{-1} \doteq \{u = -1\} \cap \{v \geq 0\}$$

$$\underline{C} \doteq \underline{C}_0 \doteq \{v = 0\} \cap \{u \geq -1\}.$$

The special point $(-1, 0)$ is called the *bifurcation sphere*. We pose data for ϕ, Q, r, Ω^2 and A_u for the Einstein–Maxwell-charged-scalar field system on $C \cup \underline{C}$.

Definition 3.1.2. Let $k \in \mathbb{N}$. A C^k *bifurcate characteristic initial data set* on $C \cup \underline{C}$ for the Einstein–Maxwell-charged scalar field system in the EM gauge (2.2.2) consists of continuous functions $r > 0$, $\Omega^2 > 0$, ϕ (complex-valued), Q, and A_u on $C \cup \underline{C}$. It is required that $r \in C^{k+1}$, $\Omega^2 \in C^k$, $\phi \in C^k$, $Q \in C^k$, and $A_u \in C^k$ on $C \cup \underline{C}$.[2] Finally, the data are required to satisfy equations (2.2.8)–(2.2.7), which implies also $\partial_v^{k+1} A_u \in C^0(C)$.

Given characteristic initial data on a portion of $C \cup \underline{C}$ containing the bifurcation sphere, we can solve in a full double null neighborhood to the future. The proof is a standard iteration argument.

Proposition 3.1.3. *Given a C^k bifurcate characteristic initial data set for the EMCSF system on*

$$(\{u = -1\} \times \{0 \leq v \leq v_0\}) \cup (\{-1 \leq u \leq u_0\} \times \{v = 0\}) \subset C \cup \underline{C},$$

where $u_0 > -1$ and $v_0 > 0$, there exists a number $\delta > 0$ and a unique spherically symmetric C^k solution of the EMCSF system on

$$(\{-1 \leq u \leq -1 + \delta\} \times \{0 \leq v \leq v_0\}) \cup (\{-1 \leq u \leq u_0\} \times \{0 \leq v \leq \delta\})$$

which extends the initial data on $C \cup \underline{C}$.

3.2 Einstein–Maxwell–Vlasov

3.2.1 Local well-posedness in spherical symmetry

Electromagnetic geodesics, in contrast to ordinary geodesics, can have limit points in $\mathcal{M}$. By standard ODE theory, this can only occur if $p(s) \to 0$ as $s \to \pm\infty$.[3] On a fixed spherically symmetric

[2] By "C^k on $C \cup \underline{C}$" is meant that v derivatives are continuous on C and u derivatives are continuous on $\underline{C}$.

[3] This does not occur for ordinary geodesics because of the following homogeneity property: If $s \mapsto \gamma(s)$ is a geodesic, then so is $s \mapsto \gamma(as)$ for any $a > 0$. See [ONe83, Lemma 5.8], where homogeneity is used to identify radial geodesics emanating from the same point with parallel velocity.

background, one can show that an electromagnetic geodesic γ cannot have a limit point if either $\mathfrak{m}[\gamma] > 0$ or $\ell[\gamma] > 0$. However, even in the massless case, an electromagnetic geodesic with initially positive momentum will still have positive momentum for a short (coordinate) time. Therefore, one can show that local well-posedness in double null gauge holds in the case of massive particles or massless particles with momentum initially supported away from zero.

Remark 3.2.1. Bigorgne has shown that the relativistic Maxwell–massless Vlasov system is classically ill-posed if the initial data are allowed to be supported near zero momentum [Big22]. We expect a similar result to hold for the Einstein–Maxwell–massless Vlasov system.

We now state our fundamental local well-posedness result for the spherically symmetric Einstein–Maxwell–Vlasov system. We formulate this in terms of the *characteristic initial value problem*, though the techniques used apply to the Cauchy problem as well. Note that we work in function spaces that allow for noncompact support in the momentum variables, although this is not needed for the applications in this dissertation (but is useful in the context of cosmic censorship [DR16]). The proof of local existence is deferred to Section 3.3.

Given $U_0 < U_1$ and $V_0 < V_1$, let

$$\mathcal{C}(U_0, U_1, V_0, V_1) \doteq (\{U_0\} \times [V_0, V_1]) \cup ([U_0, U_1] \times \{V_0\}),$$
$$\mathcal{R}(U_0, U_1, V_0, V_1) \doteq [U_0, U_1] \times [V_0, V_1].$$

We will consistently omit (U_0, U_1, V_0, V_1) from the notation for these sets when the meaning is clear. A function $\phi : \mathcal{C} \to \mathbb{R}$ is said to be *smooth* if it is continuous and $\phi|_{\{U_0\}\times[V_0,V_1]}$ and $\phi|_{[U_0,U_1]\times\{V_0\}}$ are C^∞ single-variable functions. This definition extends naturally to functions $f : P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{C}} \to \mathbb{R}_{\geq 0}$.

Definition 3.2.2. A smooth *(bifurcate) characteristic initial data set* for the spherically symmetric Einstein–Maxwell–Vlasov system with parameters $\mathfrak{m}, \mathfrak{e}$, and σ consists of smooth functions $\mathring{r}, \mathring{\Omega}^2, \mathring{Q} : \mathcal{C} \to \mathbb{R}$ with $\mathring{r}$ and $\mathring{\Omega}^2$ positive, and a smooth function $\mathring{f} : P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{C}} \to \mathbb{R}_{\geq 0}$, where $P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{C}}$ is defined using $\mathring{\Omega}^2$. Moreover, we assume that the norms

$$\|\mathring{f}\|_{C^k_\sigma(P|_{\mathcal{C}})} \doteq \sum_{0 \leq i_1 + i_2 \leq k} \left(\sup_{P^\kappa|_{\{U_0\}\times[V_0,V_1]}} \langle p^\tau \rangle^{\sigma + i_2} |\partial_v^{i_1} \partial_p^{i_2} \mathring{f}| + \sup_{P^\kappa|_{[U_0,U_1]\times\{V_0\}}} \langle p^\tau \rangle^{\sigma + i_2} |\partial_u^{i_1} \partial_p^{i_2} \mathring{f}| \right)$$

are finite for every $k \geq 0$. In the case $\mathfrak{m} = 0$, we also assume that $\mathring{f}$ has locally positive angular momentum. Finally, we assume that Raychaudhuri's equations (2.3.22) and (2.3.23), together with Maxwell's equations (2.3.24) and (2.3.25) are satisfied to all orders in directions tangent to $\mathcal{C}$.

Proposition 3.2.3. *For any* $\mathfrak{m} \geq 0$, $\mathfrak{e} \in \mathbb{R}$, $\sigma > 4$, $B > 0$, *and* $c_\ell > 0$, *there exists a constant* $\varepsilon_{\mathrm{loc}} > 0$ *with the following property. Let* $(\mathring{r}, \mathring{\Omega}^2, \mathring{Q}, \mathring{f})$ *be a characteristic initial data set for the spherically symmetric Einstein–Maxwell–Vlasov system on* $\mathcal{C}(U_0, U_1, V_0, V_1)$. *If* $U_1 - U_0 < \varepsilon_{\mathrm{loc}}$, $V_1 - V_0 < \varepsilon_{\mathrm{loc}}$,

$$\|\log \mathring{r}\|_{C^2(\mathcal{C})} + \|\log \mathring{\Omega}^2\|_{C^2(\mathcal{C})} + \|\mathring{Q}\|_{C^1(\mathcal{C})} + \|\mathring{f}\|_{C^1_\sigma(P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{C}})} \leq B,$$

and in the case $\mathfrak{m} = 0$, $\ell \geq c_\ell$ *on* $\mathrm{spt}(\mathring{f})$, *then there exists a unique smooth solution* (r, Ω^2, Q, f) *of the spherically symmetric Einstein–Maxwell–Vlasov system on* $\mathcal{R}(U_0, U_1, V_0, V_1)$ *which extends the initial data. If* $\mathring{f}$ *has locally compact support in* p, *then so does* f. *Moreover, the norms*

$$\|\log r\|_{C^k(\mathcal{R})}, \|\log \Omega^2\|_{C^k(\mathcal{R})}, \|Q\|_{C^k(\mathcal{R})}, \|f\|_{C^k_\sigma(P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{R}})}$$

are finite for any k *and can be bounded in terms of appropriate higher order initial data norms.*

The proof of the proposition is given in Section 3.3.2.

3.2.2 The generalized extension principle

Recall that a spherically symmetric Einstein-matter model is said to satisfy the *generalized extension principle* if any "first singularity" either emanates from a point on the spacetime boundary with $r = 0$, or its causal past has infinite spacetime volume. This property has been shown to hold for the Einstein-massless scalar field system by Christodoulou [Chr93], for the Einstein–massive Vlasov system by Dafermos and Rendall [DR16], and for the Einstein–Maxwell–charged Klein–Gordon system by Kommemi [Kom13]. We now extend the generalized extension principle of Dafermos–Rendall to the Einstein–Maxwell–Vlasov system:

Proposition 3.2.4 (The generalized extension principle)**.** *Let* $(\mathcal{Q}, r, \Omega^2, Q, f)$ *be a smooth solution of the spherically symmetric Einstein–Maxwell–Vlasov system with admissible momentum, which is defined on an open set* $\mathcal{Q} \subset \mathbb{R}^2_{u,v}$. *If* $\mathcal{Q}$ *contains the set* $\mathcal{R}' \doteq \mathcal{R}(U_0, U_1, V_0, V_1) \setminus \{(U_1, V_1)\}$ *and the following two conditions are satisfied:*

1. $\mathcal{R}'$ *has finite Lorentzian volume, i.e.,*

$$\iint_{\mathcal{R}'} \Omega^2 \, dudv < \infty, \tag{3.2.1}$$

2. the area-radius is bounded above and below, i.e.,

$$\sup_{\mathcal{R}'} |\log r| < \infty, \tag{3.2.2}$$

then the solution extends smoothly, with admissible momentum, to a neighborhood of (U_1, V_1).

Therefore, since this system satisfies the dominant energy condition (Remark 2.3.8), the Einstein–Maxwell–Vlasov system is *strongly tame* in Kommemi's terminology [Kom13], under the admissible momentum assumption. This is an important "validation" of the Einstein–Maxwell–Vlasov model over the charged null dust model and means the model enjoys Kommemi's a priori boundary characterization [Kom13], which will be used in Section 8.11.2.1 below.

Proposition 3.2.4 is also used crucially in the proof of Theorem 1.2.1 because it provides a continuation criterion at zeroth order. This allows us to avoid commutation when treating the singular "main beam" in the construction of bouncing charged Vlasov beams.

We now give the proof of Proposition 3.2.4, assuming the "fundamental local spacetime estimate" to be stated and proved in Section 3.2.2.2 below. The proof of the local estimate is based on a streamlining of the ideas already present in [DR16] together with the monotonicity of charge inherent to the Einstein–Maxwell–Vlasov system and a quantitative lower bound on the "coordinate time momentum" $p^u + p^v$ obtained from the mass shell relation.

Proof of Proposition 3.2.4. By Lemma 3.2.5 below, (3.2.1) and (3.2.2) imply that

$$B \doteq \|\log r\|_{C^2(\mathcal{R}')} + \|\log \Omega^2\|_{C^2(\mathcal{R}')} + \|Q\|_{C^1(\mathcal{R}')} + \|f\|_{C^1_\sigma(P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{R}'})} < \infty.$$

Let $U_1' > U_1$ and $V_1' > V_1$ be such that the segments $[U_1, U_1'] \times \{V_0\}$ and $\{U_0\} \times [V_1, V_1']$ lie inside of $\mathcal{Q}$ and let $c_\ell > 0$ be a lower bound for ℓ on $\mathrm{spt}(f) \cap P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{C}(U_0,U_1',V_0,V_1')}$ if $\mathfrak{m} = 0$. Let $\varepsilon_{\mathrm{loc}} > 0$ be the local existence time for the spherically symmetric Einstein–Maxwell–Vlasov system with parameters $(\mathfrak{m}, \mathfrak{e}, \sigma, 2B, c_\ell)$ given by Proposition 3.2.3. Fix $(\tilde{U}, \tilde{V}) \in \mathcal{R}'$ with $U_1 - \tilde{U} < \varepsilon_{\mathrm{loc}}$ and $V_1 - \tilde{V} < \varepsilon_{\mathrm{loc}}$.

Observe that if $U_2 > U_1$ is sufficiently close to U_1 and $V_2 > V_1$ is sufficiently close to V_1, then

$$B \doteq \|\log r\|_{C^2(\tilde{\mathcal{C}})} + \|\log \Omega^2\|_{C^2(\tilde{\mathcal{C}})} + \|Q\|_{C^1(\tilde{\mathcal{C}})} + \|f\|_{C^1_\sigma(P^{\mathfrak{m}}_{\mathrm{red}}|_{\tilde{\mathcal{C}}})} \le 2B$$

and $\ell \ge c_\ell$ on $\mathrm{spt}(f) \cap P^{\mathfrak{m}}_{\mathrm{red}}|_{\tilde{\mathcal{C}}}$ if $\mathfrak{m} = 0$, where $\tilde{\mathcal{C}} \doteq \mathcal{C}(\tilde{U}, U_2, \tilde{V}, V_2)$. Indeed, this is clear for $\log r$, $\log \Omega^2$, and Q by smoothness of these functions on $\mathcal{Q}$. For f, we can also easily show this using the mean value theorem and the finiteness of $\|f\|_{C^2_\sigma(P^{\mathfrak{m}}_{\mathrm{red}}|_K)}$ on compact sets $K \subset \mathcal{Q}$. For ℓ, this follows immediately from conservation of angular momentum and the domain of dependence property if $U_2 \le U_1'$ and $V_2 \le V_1'$.

Therefore, by Proposition 3.2.3, the solution extends smoothly, with admissible momentum, to the rectangle $\mathcal{R}(\tilde{U}, U_2, \tilde{V}, V_2)$, which contains (U_1, V_1). This completes the proof. □

3.2.2.1 Horizontal lifts and the commuted Vlasov equation

Local well-posedness for (r, Ω^2, Q, f) takes place in the space $C^2 \times C^2 \times C^1 \times C^1$, since the Christoffel symbols and electromagnetic field need to be Lipschitz regular to obtain a unique classical solution of the Vlasov equation (2.3.26). In order to estimate $\partial_u^2\Omega^2$ and $\partial_v^2\Omega^2$, one has to commute the wave equation for Ω^2, (2.3.21), with ∂_u and ∂_v. This commuted equation contains terms such as $\partial_u T^{uv}$, which can only be estimated by first estimating $\partial_u f$ and $\partial_v f$. On the other hand, naively commuting the spherically symmetric Vlasov equation, (2.3.26), with spatial derivatives introduces highest order nonlinear error terms such as $\partial_u^2\Omega^2\,\partial_{p^u} f$.[4] Therefore, it would appear that the system does not close at this level of regularity.

However, as was observed by Dafermos and Rendall in [DR05a] in the case of Einstein–Vlasov (see also the erratum of [RR92]), the *horizontal lifts*

$$\hat{\partial}_u f \doteq \partial_u f - p^u \partial_u \log \Omega^2 \partial_{p^u} f,$$
$$\hat{\partial}_v f \doteq \partial_v f - p^v \partial_v \log \Omega^2 \partial_{p^v} f$$

of $\partial_u f$ and $\partial_v f$ with respect to the Levi–Civita connection satisfy a better system of equations without these highest order errors. In the case of Einstein–Maxwell–Vlasov, we directly commute (2.3.26) with $\{\hat{\partial}_u, \hat{\partial}_v, \partial_{p^u}, \partial_{p^v}\}$ to obtain

$$\begin{aligned} X(\hat{\partial}_u f) &= p^u \partial_u \log \Omega^2 \hat{\partial}_u f + \left(p^u \partial_u \log \Omega^2 \partial_{p^u} \zeta^u - \partial_u \zeta^u - \partial_u \log \Omega^2 \zeta^u\right) \partial_{p^u} f \\ &\quad + \left(\partial_u \partial_v \log \Omega^2 (p^u)^2 - \partial_u \zeta^v + p^u \partial_u \log \Omega^2 \partial_{p^u} \zeta^v\right) \partial_{p^v} f, \end{aligned} \tag{3.2.3}$$

$$\begin{aligned} X(\hat{\partial}_v f) &= p^v \partial_v \log \Omega^2 \hat{\partial}_v f + \left(\partial_u \partial_v \log \Omega^2 (p^v)^2 - \partial_v \zeta^u + p^v \partial_v \log \Omega^2 \partial_{p^v} \zeta^u\right) \partial_{p^u} f \\ &\quad + \left(p^v \partial_v \log \Omega^2 \partial_{p^v} \zeta^v - \partial_v \zeta^v - \partial_v \log \Omega^2 \zeta^v\right) \partial_{p^v} f, \end{aligned} \tag{3.2.4}$$

$$X(\partial_{p^u} f) = -\hat{\partial}_u f + (3p^u \partial_u \log \Omega^2 - \partial_{p^u} \zeta^u) \partial_{p^u} f - \partial_{p^u} \zeta^v \partial_{p^v} f, \tag{3.2.5}$$

$$X(\partial_{p^v} f) = -\hat{\partial}_v f - \partial_{p^v} \zeta^u \partial_{p^u} f + (3p^v \partial_v \log \Omega^2 - \partial_{p^v} \zeta^v) \partial_{p^v} f, \tag{3.2.6}$$

[4] This is clearly not an issue for local well-posedness since the "time interval" of the solution can be taken sufficiently small to absorb (the time integral of) this term.

where

$$\zeta^u \doteq \frac{2\partial_v r}{r\Omega^2}(\Omega^2 p^u p^v - \mathfrak{m}^2) + \mathfrak{e}\frac{Q}{r^2}p^u, \quad \zeta^v \doteq \frac{2\partial_u r}{r\Omega^2}(\Omega^2 p^u p^v - \mathfrak{m}^2) - \mathfrak{e}\frac{Q}{r^2}p^v.$$

Upon using the wave equation (2.3.21), we see that the right-hand sides of (3.2.3)–(3.2.6) do not contain second derivatives of Ω^2.

3.2.2.2 The fundamental local spacetime estimate

Lemma 3.2.5. *For any $\mathfrak{m} \geq 0$, $\mathfrak{e} \in \mathbb{R}$, $\sigma > 4$, and $C_0 > 0$, there exists a constant $C_* < \infty$ with the following property. Let (r, Ω^2, Q, f) be a solution of the spherically symmetric Einstein–Maxwell–Vlasov system with admissible momentum for particles of charge $\mathfrak{e}$, mass $\mathfrak{m}$, and momentum decay rate σ defined on $\mathcal{R}' \doteq \mathcal{R}(U_0, U_1, V_0, V_1) \setminus \{(U, V)\}$. Assume $U_1 - U_0 \leq C_0$, $V_1 - V_0 \leq C_0$, the initial data estimates*

$$\|\log r\|_{C^2(\mathcal{C})} + \|\log \Omega^2\|_{C^2(\mathcal{C})} + \|Q\|_{C^1(\mathcal{C})} + \|f\|_{C^1_\sigma(P^{\mathfrak{m}}_{\mathrm{red}}|_\mathcal{C})} \leq C_0, \tag{3.2.7}$$

the global estimates

$$\iint_{\mathcal{R}'} \Omega^2 \, dudv \leq C_0, \tag{3.2.8}$$

$$\sup_{\mathcal{R}'} |\log r| \leq C_0, \tag{3.2.9}$$

and in the case $\mathfrak{m} = 0$, assume also that

$$\inf_{\mathrm{spt}(f) \cap P^{\mathfrak{m}}_{\mathrm{red}}|_\mathcal{C}} \ell \geq C_0^{-1}. \tag{3.2.10}$$

Then we have the estimate

$$\|\log r\|_{C^2(\mathcal{R}')} + \|\log \Omega^2\|_{C^2(\mathcal{R}')} + \|Q\|_{C^1(\mathcal{R}')} + \|f\|_{C^1_\sigma(P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{R}'})} \leq C_*.$$

Proof. In this proof, we use the notation $A \lesssim 1$ to mean that $A \leq C$, where C is a constant depending only on $\mathfrak{m}$, $\mathfrak{e}$, and σ, and C_0. When writing area integrals, we will also make no distinction between $\mathcal{R}$ and $\mathcal{R}'$, though the integrands are strictly speaking only assumed to be defined on $\mathcal{R}'$.

From (3.2.7) and the monotonicity properties of Maxwell's equations (2.3.24) and (2.3.25), it follows that

$$\sup_{\mathcal{R}'} |Q| \lesssim 1. \tag{3.2.11}$$

Rewriting (2.3.20), we obtain

$$r^2\Omega^4 T^{uv} = 2\partial_u\partial_v r^2 + \Omega^2\left(1 - \frac{Q^2}{r^2}\right). \tag{3.2.12}$$

Integrating over $\mathcal{R}'$ and using (3.2.7), (3.2.8), and (3.2.9) yields

$$\iint_{\mathcal{R}} \Omega^4 T^{uv}\, dudv \lesssim \iint_{\mathcal{R}} r^2\Omega^4 T^{uv}\, dudv \lesssim \iint_{\mathcal{R}} \partial_u\partial_v r^2\, dudv + \iint_{\mathcal{R}} \Omega^2\, dudv \lesssim 1. \tag{3.2.13}$$

Rewriting (3.2.12) slightly, we obtain

$$\partial_u(r\partial_v r) = -\frac{\Omega^2}{4}\left(1 - \frac{Q^2}{r^2}\right) + \tfrac{1}{4}r^2\Omega^4 T^{uv}. \tag{3.2.14}$$

Integrating this in u and using (3.2.7), (3.2.9), and (3.2.11), we have

$$\sup_{[U_0,U_1]\times\{v\}} |r\partial_v r| \lesssim 1 + \int_{U_0}^{U_1} \Omega^2(u,v)\, du + \int_{U_0}^{U_1} \Omega^4 T^{uv}(u,v)\, du$$

for any $v \in [0, V]$. Integrating this estimate in v and using (3.2.9), (3.2.8), and (3.2.13) yields

$$\int_{V_0}^{V_1} \sup_{[U_0,U_1]\times\{v\}} |\partial_v r|\, dv \lesssim \int_{V_0}^{V_1} \sup_{[U_0,U_1]\times\{v\}} |r\partial_v r|\, dv \lesssim 1. \tag{3.2.15}$$

By Raychaudhuri's equation (2.3.22), $\partial_u r$ changes signs at most once on each ingoing null cone. Therefore, by the fundamental theorem of calculus and (3.2.9),

$$\sup_{v\in[V_0,V_1]} \int_{U_0}^{U_1} |\partial_u r|(u,v)\, du \leq 2\left(\sup_{\mathcal{R}'} r - \inf_{\mathcal{R}'} r\right) \lesssim 1. \tag{3.2.16}$$

Combining (3.2.15) and (3.2.16) yields

$$\iint_{\mathcal{R}} |\partial_u r\partial_v r|\, dudv \leq \left(\int_{V_0}^{V_1} \sup_{[U_0,U_1]\times\{v\}} |\partial_v r|\, dv\right)\left(\sup_{v\in[V_0,V_1]} \int_{U_0}^{U_1} |\partial_u r|(u,v)\, du\right) \lesssim 1. \tag{3.2.17}$$

Using the definition of the Hawking mass (2.1.2), (3.2.9), and (3.2.8), we readily infer

$$\iint_{\mathcal{R}} \Omega^2 |m|\, dudv \lesssim 1 \tag{3.2.18}$$

By the wave equation (2.3.21), the fundamental theorem of calculus, and the estimates (2.3.19), (3.2.8), (3.2.9), (3.2.11), (3.2.13), and (3.2.18), we have

$$\begin{aligned}\sup_{\mathcal{R}'}|\log\Omega^2| &\lesssim 1+\left|\iint_{\mathcal{R}}\partial_u\partial_v\log\Omega^2\,dudv\right|\\ &\lesssim 1+\iint_{\mathcal{R}}\left(\Omega^2+\Omega^2|m|+\Omega^4T^{uv}\right)dudv\\ &\lesssim 1.\end{aligned}$$

We now prove estimates for the electromagnetic geodesic flow. Let $\gamma:[0,S]\to\mathcal{R}'$ be an electromagnetic geodesic such that $(\gamma(0),p(0))\in\operatorname{spt}(f)\cap P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{C}}$. We aim to prove that

$$\left|\log\left(\frac{\Omega^2p^u(s)}{\Omega^2p^u(0)}\right)\right|+\left|\log\left(\frac{\Omega^2p^v(s)}{\Omega^2p^v(0)}\right)\right|\lesssim 1 \tag{3.2.19}$$

for $s\in[0,S]$, where the implied constant does not depend on γ.

It suffices to prove this estimate for p^u, as the proof of the estimate for p^v differs only in notation. Following [Mos18] (see also [Daf06]), we write an integral formula for $\log(\Omega^2p^u)$, which can then be estimated using our previous area estimates. Rewriting the mass shell relation (2.3.12) as

$$\frac{\ell^2}{r^2}=\left(\frac{\ell^2}{\ell^2+\mathfrak{m}^2r^2}\right)\Omega^2p^up^v,$$

we deduce from the Lorentz force equation (2.1.24) that

$$\frac{d}{ds}\log(\Omega^2p^u)=\left(\partial_v\log\Omega^2-\frac{2\partial_vr}{r}\right)\left(\frac{\ell^2}{\ell^2+\mathfrak{m}^2r^2}\right)p^v-\mathfrak{e}\frac{Q}{r^2}.$$

Integrating in s and changing variables yields

$$\log\left(\frac{\Omega^2p^u(s)}{\Omega^2p^u(0)}\right)=\int_{\gamma([0,s])}\left(\partial_v\log\Omega^2-\frac{2\partial_vr}{r}\right)\left(\frac{\ell^2}{\ell^2+\mathfrak{m}^2r^2}\right)dv-\int_0^s\mathfrak{e}\frac{Q}{r^2}\bigg|_{\gamma(s')}ds'.$$

We use the fundamental theorem of calculus on the first integral to obtain

$$\begin{aligned}&\int_{\gamma([0,s])}\left(\partial_v\log\Omega^2-\frac{2\partial_vr}{r}\right)\left(\frac{\ell^2}{\ell^2+\mathfrak{m}^2r^2}\right)dv\\ &\qquad=\int_{\gamma^v(0)}^{\gamma^v(s)}\int_0^{\gamma^u(s_v)}\partial_u\left[\left(\partial_v\log\Omega^2-\frac{2\partial_vr}{r}\right)\left(\frac{\ell^2}{\ell^2+\mathfrak{m}^2r^2}\right)\right]dudv\\ &\qquad\qquad+\int_{\{0\}\times[0,\gamma^v(s)]}\left(\partial_v\log\Omega^2-\frac{2\partial_vr}{r}\right)\left(\frac{\ell^2}{\ell^2+\mathfrak{m}^2r^2}\right)dv,\end{aligned}$$

where $s_v\in[0,S]$ is defined by $\gamma^v(s_v)=v$. Using now the wave equations (2.3.20) and (2.3.21), we arrive at

$$\log\left(\frac{\Omega^2 p^u(s)}{\Omega^2 p^u(0)}\right) = \int_{\gamma^v(0)}^{\gamma^v(s)}\int_0^{\gamma^u(s_v)}\left(\frac{3\Omega^2 m}{r^3} - \frac{3\Omega^2 Q^2}{2r^4} - \frac{\Omega^2}{2r^2} - \Omega^4 T^{uv} - \Omega^2 S\right)\left(\frac{\ell^2}{\ell^2+\mathfrak{m}^2 r^2}\right) dudv$$
$$+ \int_{\gamma^v(0)}^{\gamma^v(s)}\int_0^{\gamma^u(s_v)}\left(2\partial_u r\partial_v r - \partial_u r\partial_v \log\Omega^2\right)\frac{2\mathfrak{m}^2\ell^2}{(\ell^2+\mathfrak{m}^2 r^2)^2}\, dudv$$
$$+ \int_{\{0\}\times[0,\gamma^v(s)]}\left(\partial_v \log\Omega^2 - \frac{2\partial_v r}{r}\right) dv - \int_0^s \mathfrak{e}\frac{Q}{r^2}\bigg|_{\gamma(s')} ds'. \tag{3.2.20}$$

We now bound each of the four terms in (3.2.20). Using (3.2.8), (3.2.9), (3.2.11), (3.2.13), and (3.2.18), the first double integral in (3.2.20) is readily seen to be uniformly bounded. To estimate the second double integral, we integrate the wave equation (2.3.21) in u to obtain

$$\sup_{[0,U]\times\{v\}} |\partial_v \log\Omega^2| \lesssim 1 + \int_0^U \left(\Omega^2|m| + \Omega^2 + \Omega^4 T^{uv}\right)(u,v)\, du$$

for any $v \in [0,V]$. Integrating this estimate in v and using the previous area estimates yields

$$\int_0^V \sup_{[0,U]\times\{v\}} |\partial_v \log\Omega^2|\, dv \lesssim 1$$

which when combined with (3.2.16) gives

$$\iint_{\mathcal{R}} |\partial_u r||\partial_v \log\Omega^2|\, dudv \lesssim 1. \tag{3.2.21}$$

Combined with (3.2.17), we now readily see that the second double integral in (3.2.20) is uniformly bounded. The integral along initial data is clearly also bounded by assumption.

Using (3.2.9) and (3.2.11), we estimate

$$\int_0^s \mathfrak{e}\frac{|Q|}{r^2}\, ds \lesssim S. \tag{3.2.22}$$

To estimate S, define the function $\tau(s) \doteq \tau|_{\gamma(s)}$, which is strictly increasing and satisfies $0 \le \tau(s) \lesssim 1$ for every $s \in [0,S]$. Using the mass shell relation (2.3.12), we have

$$\sqrt{\frac{4}{\Omega^2}\left(\frac{\ell^2}{r^2}+\mathfrak{m}^2\right)} \le p^\tau = \frac{d\tau}{ds}. \tag{3.2.23}$$

Since either $\mathfrak{m} > 0$ or (3.2.10) holds, it follows that $d\tau/ds$ is uniformly bounded away from zero, which implies $S \lesssim 1$ for any electromagnetic geodesic in the support of f. Combined with (3.2.22), this uniformly bounds the final term in (3.2.20) and completes the proof of (3.2.19).

Since f is constant along $(\gamma(s), p(s))$, we therefore have

$$\langle p^\tau(s)\rangle^\sigma f(\gamma(s), p(s)) \lesssim \langle p^\tau(0)\rangle^\sigma f(\gamma(0), p(0)),$$

which implies

$$\|f\|_{C^0_\sigma(P^{\mathrm{m}}_{\mathrm{red}}|_{\mathcal{R}'})} + \sup_{\mathcal{R}'}(N^u + N^v + T^{uu} + T^{uv} + T^{vv} + S) \lesssim 1.$$

By (2.3.24) and (2.3.25),

$$\sup_{\mathcal{R}'}(|\partial_u Q| + |\partial_v Q|) \lesssim 1.$$

By integrating (3.2.14) and also using that $\partial_u(r\partial_v r) = \partial_v(r\partial_u r)$, we now readily estimate

$$\sup_{\mathcal{R}'}(|\partial_u r| + |\partial_v r|) \lesssim 1.$$

As this bounds the Hawking mass pointwise, we can now estimate

$$\sup_{\mathcal{R}'}\left(|\partial_u \log\Omega^2| + |\partial_v \log\Omega^2| + |\partial_u\partial_v \log\Omega^2|\right) \lesssim 1$$

using (2.3.21). It then follows immediately from (2.3.20), (2.3.22), and (2.3.23) that

$$\sup_{\mathcal{R}'}\left(|\partial_u^2 r| + |\partial_u\partial_v r| + |\partial_v^2 r|\right) \lesssim 1.$$

Along an electromagnetic geodesic γ lying in the support of f, we have that

$$\begin{aligned}
|X(\hat\partial_u f)| &\lesssim p^\tau|\hat\partial_u f| + (p^\tau)^2|X(\partial_{p^u} f)| + (p^\tau)^2|X(\partial_{p^v} f)|,\\
|X(\hat\partial_v f)| &\lesssim p^\tau|\hat\partial_v f| + (p^\tau)^2|X(\partial_{p^u} f)| + (p^\tau)^2|X(\partial_{p^v} f)|,\\
|X(\partial_{p^u} f)| &\lesssim |\hat\partial_u f| + p^\tau|X(\partial_{p^u} f)| + p^\tau|X(\partial_{p^v} f)|,\\
|X(\partial_{p^v} f)| &\lesssim |\hat\partial_v f| + p^\tau|X(\partial_{p^u} f)| + p^\tau|X(\partial_{p^v} f)|
\end{aligned}$$

by (3.2.3)–(3.2.6) and all of the estimates obtained so far. It follows that, defining

$$\mathcal{F}(s) \doteq \Big((p^\tau)^\sigma\hat\partial_u f, (p^\tau)^\sigma\hat\partial_v f, (p^\tau)^{\sigma+1}\partial_{p^u} f, (p^\tau)^{\sigma+1}\partial_{p^v} f\Big)(s)$$

along $\gamma(s)$, we have

$$\left|\frac{d}{ds}\mathcal{F}\right| \lesssim p^\tau |\mathcal{F}|.$$

By Grönwall's inequality and (3.2.23), it follows that $|\mathcal{F}| \lesssim 1$ along γ. Recovering $\partial_u f$ and $\partial_v f$ from $\hat{\partial}_u f$ and $\hat{\partial}_v f$, we now readily bound

$$\|f\|_{C^1_\sigma(P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{R}'})} + \sup_{\mathcal{R}'} (|\partial_u T^{uv}| + |\partial_v T^{uv}| + |\partial_u S| + |\partial_v S|) \lesssim 1.$$

Commuting the wave equation (2.3.21) with ∂_u and ∂_v, we obtain the final estimates

$$\sup_{\mathcal{R}'} \left(|\partial_u^2 \log \Omega^2| + |\partial_v^2 \log \Omega^2|\right) \lesssim 1,$$

which completes the proof. □

3.2.2.3 Local existence in characteristic slabs

The spacetime estimate Lemma 3.2.5 can also be used to improve Proposition 3.2.3 to local existence in a full double null neighborhood of a bifurcate characteristic hypersurface of arbitrary length:

Proposition 3.2.6. *For any $\mathfrak{m} \geq 0$, $\mathfrak{e} \in \mathbb{R}$, $\sigma > 4$, $B > 0$, and $c_\ell > 0$, there exists a constant $\varepsilon_{\mathrm{slab}} > 0$ with the following property. Let $(\mathring{r}, \mathring{\Omega}^2, \mathring{Q}, \mathring{f})$ be a characteristic initial data set for the spherically symmetric Einstein–Maxwell–Vlasov system on $\mathcal{C}(U_0, U_1, V_0, V_1)$ with admissible momentum. If*

$$\|\log \mathring{r}\|_{C^2(\mathcal{C})} + \|\log \mathring{\Omega}^2\|_{C^2(\mathcal{C})} + \|\mathring{Q}\|_{C^1(\mathcal{C})} + \|\mathring{f}\|_{C^1_\sigma(P^{\mathfrak{m}}_{\mathrm{red}}|_{\mathcal{C}})} \leq B$$

and either $\mathfrak{m} > 0$ or $\mathfrak{m} = 0$ and $\ell \geq c_\ell$ on $\mathrm{spt}(\mathring{f})$, then there exists a unique smooth solution (r, Ω^2, Q, f) of the spherically symmetric Einstein–Maxwell–Vlasov system with admissible momentum on

$$\mathcal{R}(U_0, U_0 + \min\{\varepsilon_{\mathrm{slab}}, U_1 - U_0\}, V_0, V_1) \cup \mathcal{R}(U_0, U_1, V_0, V_0 + \min\{\varepsilon_{\mathrm{slab}}, V_1 - V_0\})$$

which extends the initial data.

Proof. We prove existence in the slab which is thin in the u-direction, the proof in the other slab being identical. Let $C_0 = 10B$ and let C_* be the constant obtained from the fundamental local spacetime estimate Lemma 3.2.5 with this choice. Let $\mathcal{A} \subset [V_0, V_1]$ denote the set of $\tilde{V}$ such that the solution exists on $\mathcal{R}(U_0, U_1', V_0, \tilde{V})$, where $U_1' \doteq U_0 + \min\{\varepsilon_{\text{slab}}, U_1 - U_0\}$ and satisfies the estimates

$$\sup_{\mathcal{R}(U_0,U_1',V_0,\tilde{V})} |\log r| + \sup_{\mathcal{R}(U_0,U_1',V_0,\tilde{V})} |\log \Omega^2| \leq C_0. \tag{3.2.24}$$

We will show that if $\varepsilon_{\text{slab}} = \min\{\varepsilon_{\text{loc}}(C_*), B(U_1 - U_0)^{-1}C_*^{-1}\}$, then $\mathcal{A}$ is nonempty, closed, and open.

Nonemptyness follows from Proposition 3.2.3 and closedness by continuity of the bootstrap assumptions. Let now $\tilde{V} \in \mathcal{A}$. To improve the bootstrap assumptions, we note that $|\partial_u \partial_v \log r| \leq C_*$ on $\mathcal{R}(U_0, U_1', V_0, \tilde{V})$ by Lemma 3.2.5, whence $|\log r| \leq \varepsilon_{\text{slab}}(U_1 - U_0)C_* + 3B \leq \frac{1}{2}C_0$ by the fundamental theorem of calculus. A similar argument applies for $\log \Omega^2$. Therefore, by applying Proposition 3.2.3 again, a simple continuity argument shows that $\tilde{V} + \eta \in \mathcal{A}$ for $\eta > 0$ sufficiently small. □

3.2.3 Time-symmetric seed data and their normalized developments

In the proof of Theorem 1.2.1, we will pose data for the Einstein–Maxwell–Vlasov system on a mixed spacelike-null hypersurface, with the Vlasov field f supported initially on the spacelike hypersurface and away from the center. The initial data is given by a compactly supported distribution function $\mathring{f}$ on the spacelike hypersurface, a numerical parameter that fixes the location of the initial outgoing null cone, together with the mass and charge of the particles. As we will only consider *time-symmetric* initial configurations, these data are sufficient to uniquely determine the solution.

Definition 3.2.7. A *time-symmetric seed data set* $\mathcal{S} \doteq (\mathring{f}, r_2, \mathfrak{m}, \mathfrak{e})$ for the spherically symmetric Einstein–Maxwell–Vlasov system consists of a real numbers $r_2 \in \mathbb{R}_{>0}$, $\mathfrak{m} \in \mathbb{R}_{\geq 0}$, and $\mathfrak{e} \in \mathbb{R}$, together with a compactly supported nonnegative function $\mathring{f} \in C^\infty\big((0,\infty)_r \times (0,\infty)_{p^u} \times (0,\infty)_{p^v}\big)$ which is symmetric in the second and third variables, $\mathring{f}(\,\cdot\,, p^u, p^v) = \mathring{f}(\,\cdot\,, p^v, p^u)$, and satisfies $\operatorname{spt}\big(\mathring{f}(\,\cdot\,, p^u, p^v)\big) \subset (0, r_2]$ for every $p^u, p^v \in (0,\infty)$.

Given a seed data set $\mathcal{S} = (\mathring{f}, r_2, \mathfrak{m}, \mathfrak{e})$ and $r \in [0, r_2]$, we define

$$\mathring{\varrho}(r) \doteq \pi \int_0^\infty \int_0^\infty \mathring{f}(v, p^u, p^v)\, dp^u dp^v, \tag{3.2.25}$$

$$\mathring{\mathcal{N}}^u(r) \doteq \mathring{\mathcal{N}}^v(r) \doteq \pi \int_0^\infty \int_0^\infty p^u \mathring{f}(r, p^u, p^v)\, dp^u dp^v, \tag{3.2.26}$$

$$\mathring{\mathcal{T}}^{uu}(r) \doteq \mathring{\mathcal{T}}^{vv}(r) \doteq \pi \int_0^\infty \int_0^\infty (p^u)^2 \mathring{f}(r, p^u, p^v)\, dp^u dp^v, \tag{3.2.27}$$

$$\mathring{\mathcal{T}}^{uv}(r) \doteq \pi \int_0^\infty \int_0^\infty p^u p^v \mathring{f}(r, p^u, p^v)\, dp^u dp^v. \tag{3.2.28}$$

Remark 3.2.8. These formulas are missing a factor of Ω^2 compared to (2.3.14)–(2.3.19). This is

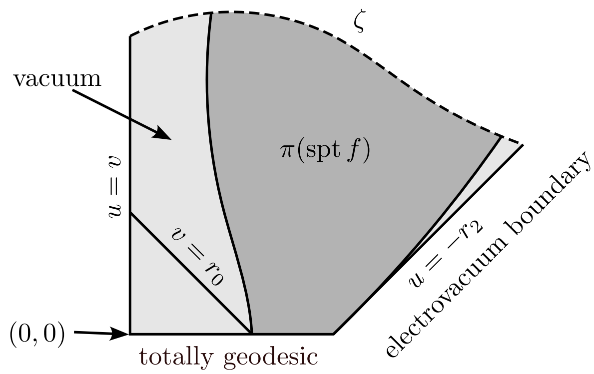

Figure 3.1: Penrose diagram of a normalized development $\mathcal{U}$ of a time symmetric seed $\mathcal{S}$. The spacelike hypersurface $\{\tau = 0\}$ is totally geodesic, i.e., time symmetric, and the outgoing cone $\{u = -r_2\}$ has $f = 0$. To the left of the support of f, the spacetime is vacuum: both the Hawking mass m and charge Q vanish identically. For the significance of the cone $\{v = r_0\}$, see already Remark 3.2.13.

because Ω^2 is not explicitly known on the initial data hypersurface and is accounted for by extra factors of Ω^2 in the constraint system (3.2.29)–(3.2.30) below.

Let the functions $\mathring{m} = \mathring{m}(r)$ and $\mathring{Q} = \mathring{Q}(r)$ be the unique solutions of the first order system

$$\frac{d}{dr}\mathring{m} = \frac{r^2}{4}\left(1 - \frac{2\mathring{m}}{r}\right)^{-2}\left(\mathring{\mathcal{T}}^{uu} + 2\mathring{\mathcal{T}}^{uv} + \mathring{\mathcal{T}}^{vv}\right) + \frac{\mathring{Q}^2}{2r^2}, \tag{3.2.29}$$

$$\frac{d}{dr}\mathring{Q} = \frac{1}{2}\mathfrak{e}r^2\left(1 - \frac{2\mathring{m}}{r}\right)^{-2}\left(\mathring{\mathcal{N}}^u + \mathring{\mathcal{N}}^v\right) \tag{3.2.30}$$

with initial conditions $\mathring{m}(0) = 0$ and $\mathring{Q}(0) = 0$. If

$$\sup_{r\in[0,r_2]} \frac{2\mathring{m}}{r} < 1,$$

then $\mathring{m}$ and $\mathring{Q}$ exist on the entire interval $[0, r_2]$ and we say that $\mathcal{S}$ is *untrapped.* We also define

$$\mathring{\Omega}^2 \doteq \left(1 - \frac{2\mathring{m}}{r}\right)^{-1}, \quad \mathring{\varpi} \doteq \mathring{m} + \frac{\mathring{Q}^2}{2r}.$$

Finally, we say that $\mathcal{S}$ is *consistent with particles of mass* $\mathfrak{m}$ if $\mathring{\Omega}^2(r)p^u p^v \geq \mathfrak{m}^2$ for every $(r, p^u, p^v) \in \operatorname{spt} \mathring{f}$.

Remark 3.2.9. We have not attempted to formulate the most general notion of seed data for the spherically symmetric Einstein–Maxwell–Vlasov Cauchy problem here as it is not needed for our purposes.

Associated with time-symmetric seed data as in Definition 3.2.7, we will introduce normalized developments of such data in the following. For $r_2 > 0$, let

$$\mathcal{C}_{r_2} \doteq \{\tau \geq 0\} \cap \{v \geq u\} \cap \{u \geq -r_2\} \subset \mathbb{R}^2_{u,v}$$

and let $\mathfrak{U}_{r_2}$ denote the collection of connected relatively open subsets $\mathcal{U} \subset \{v \geq u\} \subset \mathbb{R}^2_{u,v}$ for which there exists a (possibly empty) achronal curve $\zeta \subset \mathcal{C}_{r_2}$, extending from the center $\{u = v\}$ and reaching the cone $\{u = -r_2\}$, such that $\mathcal{U} = \mathcal{C}_{r_2} \cap \{u+v < \zeta^u + \zeta^v\}$ and $\{\tau = 0\} \cap \{0 \leq v \leq r_2\} \subset \mathcal{U}$. See Fig. 3.1.

We also define the cones

$$C_{u_0} \doteq \mathcal{U} \cap \{u = u_0\}, \quad \underline{C}_{v_0} \doteq \mathcal{U} \cap \{v = v_0\}.$$

Definition 3.2.10. Let $\mathcal{S} = (\mathring{f}, r_2, \mathfrak{m}, \mathfrak{e})$ be an untrapped time-symmetric seed data set which is consistent with particles of mass $\mathfrak{m}$. A *normalized development* of $\mathcal{S}$ consists of a domain $\mathcal{U} \in \mathfrak{U}_{r_2}$ and a spherically symmetric solution (r, Ω^2, Q, f) of the Einstein–Maxwell–Vlasov system for particles of mass $\mathfrak{m}$ and fundamental charge $\mathfrak{e}$ defined over $\mathcal{U} \setminus \{u = v\}$ such that the following holds.

1. For every $(v, p^u, p^v) \in (0, r_2] \times (0, \infty) \times (0, \infty)$,

$$r(-v, v) = v, \tag{3.2.31}$$

$$\partial_v r(-v, v) = \tfrac{1}{2}, \tag{3.2.32}$$

$$\partial_u r(-v, v) = -\partial_v r(-v, v), \tag{3.2.33}$$

$$\Omega^2(-v, v) = \mathring{\Omega}^2(v), \tag{3.2.34}$$

$$\partial_v \Omega^2(-v, v) = \frac{1}{2}\left(\frac{d}{dr}\mathring{\Omega}^2\right)(v), \tag{3.2.35}$$

$$\partial_u \log \Omega^2(-v, v) = -\partial_v \log \Omega^2(-v, v), \tag{3.2.36}$$

$$f(-v, v, p^u, p^v) = \mathring{f}(v, p^u, p^v). \tag{3.2.37}$$

2. Along the initial outgoing null cone C_{-r_2},

$$r(-r_2, v) = \tfrac{1}{2}r_2 + \tfrac{1}{2}v, \tag{3.2.38}$$

$$\Omega^2 = \mathring{\Omega}^2(r_2), \tag{3.2.39}$$

$$Q = \mathring{Q}(r_2), \tag{3.2.40}$$

$$f = 0. \tag{3.2.41}$$

3. The functions r, Ω^2, Q, and m extend smoothly to the center $\Gamma \doteq \mathcal{U} \cap \{u = v\}$ and satisfy there the boundary conditions

$$r = m = Q = 0, \tag{3.2.42}$$

$$\partial_u r < 0, \quad \partial_v r > 0. \tag{3.2.43}$$

4. Let $\gamma : [0, S) \to \mathcal{U} \setminus \Gamma$ be a future-directed electromagnetic geodesic such that $r(\gamma(s)) \to 0$ as $s \to S$.[5] Then $(\gamma^u(s), \gamma^v(s), p^u(s), p^v(s))$ attains a limit on Γ, say (u_*, v_*, p^u_*, p^v_*), and there exists a unique electromagnetic geodesic $\gamma' : (S, S+\varepsilon) \to \mathcal{U} \setminus \Gamma$ for some $\varepsilon > 0$ such that $(\gamma'^u(s), \gamma'^v(s), p'^u(s), p'^v(s)) \to (u_*, v_*, p^u_*, p^v_*)$ as $s \to S$. We then require that

$$\lim_{s \nearrow S} f(\gamma^u(s), \gamma^v(s), p^u(s), p^v(s)) = \lim_{s \searrow S} f(\gamma'^u(s), \gamma'^v(s), p'^u(s), p'^v(s)).$$

We use the adjective "normalized" to emphasize the choice of a development with double null gauge anchored to the data as in points 1. and 2. above.

Remark 3.2.11. The "time-symmetric" aspect of the development is captured by the first equalities in (3.2.27) and (3.2.28), and the equations (3.2.33) and (3.2.36). One can moreover easily verify, using (3.2.33), (3.2.36), and the formulas for the Christoffel symbols in Section 2.1.3, that $\{\tau = 0\} \cap \mathcal{U}$ is a totally geodesic spacelike hypersurface with respect to the $(3+1)$-dimensional metric (2.1.1).

[5] Such a curve necessarily has $\ell = 0$.

For a normalized development of seed data, we clearly have

$$\begin{aligned} N^u &= \mathring{\Omega}^2 \mathring{\mathcal{N}}^u, & N^v &= \mathring{\Omega}^2 \mathring{\mathcal{N}}^v, \\ T^{uu} &= \mathring{\Omega}^2 \mathring{\mathcal{T}}^{uu}, & T^{vv} &= \mathring{\Omega}^2 \mathring{\mathcal{T}}^{vv}, \\ T^{uv} &= \mathring{\Omega}^2 \mathring{\mathcal{T}}^{uv}, & S &= \frac{\mathring{\Omega}^4}{2} \mathring{\mathcal{T}}^{uv} - \mathring{\Omega}^2 \mathfrak{m}^2 \mathring{\varrho}, \\ Q &= \mathring{Q}, & \varpi &= \mathring{\varpi} \end{aligned}$$

along $\{\tau = 0\} \cap \mathcal{U}$.

Proposition 3.2.12. *Let $\mathcal{S}$ be an untrapped time-symmetric seed data set which is consistent with particles of mass $\mathfrak{m}$. Then there exists a $\delta > 0$ and a unique normalized development (r, Ω^2, Q, f) of $\mathcal{S}$ defined on $\{0 \le \tau < \delta\} \cap \mathcal{C}_{r_2}$.*

Proof. Using essentially the same methods as the proof of Proposition 3.2.3 in Section 3.3, we obtain a unique local smooth solution (r, Ω^2, Q, f) to the system of equations (2.3.20), (2.3.21), (2.3.24), and (2.3.26), with initial data given by (3.2.31)–(3.2.41). It remains to show that the constraints (2.3.22), (2.3.23), and (2.3.25) hold.

By the same calculation as in the proof of Proposition 3.2.3, equation (2.3.26) implies the conservation law (2.3.32) for N. Let $v \in (0, r_2)$. By integration of (2.3.24),

$$Q(u,v) = \mathring{Q}(v) - \int_{-v}^{u} \tfrac{1}{2} \mathfrak{e} r^2 \Omega^2 N^v \, du'$$

for $u \ge -v$. Differentiating in v, using (3.2.30), (2.3.32), and the fundamental theorem of calculus yields (2.3.25) at (u, v).

To prove that (2.3.22) and (2.3.23) hold, we argue as in the proof of Proposition 3.2.3. Therefore, it suffices to show that (2.3.23) holds on initial data (the corresponding argument for (2.3.22) being the same). By (3.2.38) and (3.2.41), (2.3.23) clearly holds on the initial outgoing cone. By taking the absolute v-derivative of $\partial_v r(-v, v) = \frac{1}{2}$, we obtain $\partial_v^2 r(-v, v) = \partial_u \partial_v r(-v, v)$. Therefore, using

(2.3.20), (3.2.29), and (3.2.35), we readily compute

$$\begin{aligned}
\partial_v^2 r - \partial_v r \partial_v \log \Omega^2 + \frac{1}{4} r\Omega^4 T^{vv} &= \partial_u\partial_v r - \frac{1}{4\Omega^2}\frac{d}{dr}\left(1-\frac{2\mathring{m}}{r}\right)^{-1} + \frac{1}{4} r\Omega^4 T^{vv} \\
&= -\frac{\Omega^2 m}{2r^2} + \frac{\Omega^2 Q^2}{4r^3} + \frac{1}{4} r\Omega^4 T^{uv} \\
&\quad - \frac{1}{4\Omega^2}\left(-\frac{2\Omega^4\mathring{m}}{r^2} + \frac{r\Omega^6}{2}\left(T^{uu}+2T^{uv}+T^{vv}\right) + \frac{\Omega^4 Q^2}{r^3}\right) + \frac{1}{4} r\Omega^4 T^{vv} \\
&= 0,
\end{aligned}$$

where every function is being evaluated at $(-v, v)$. This is equivalent to (2.3.23) and completes the proof. □

Remark 3.2.13. Let $r_0 \in (0, r_2)$ be such that $\mathring{f}(r, p^u, p^v) = 0$ if $r \in (0, r_0]$. Since $\mathring{f}$ is assumed to be compactly supported, such an r_0 necessarily exists. Then if $(\mathcal{U}, r, \Omega^2, Q, f)$ is a normalized development of $\mathcal{S}$, the portion of the triangle $\{v \le r_0\}$ inside of $\mathcal{U}$ is identically Minkowskian in the sense that

$$r = \tfrac{1}{2}(v-u),$$
$$\Omega^2 = 1,$$
$$Q = f = 0$$

on $\mathcal{U} \cap \{v \le r_0\}$. In fact, we may therefore assume that any normalized development of $\mathcal{S}$ contains the full corner $\mathcal{C}_{r_2} \cap \{v \le r_0\}$.

Remark 3.2.14. One can verify that a normalized development as in Definition 3.2.10 defines a solution of the constraint equations associated to the $(3+1)$-dimensional Einstein–Maxwell–Vlasov system after applying the correspondence of Proposition 2.3.10. In particular, the lift of $\tau = 0$ will be totally geodesic in the $(3+1)$-dimensional spacetime.

Remark 3.2.15. One can "maximalize" Proposition 3.2.12 to show the existence of a maximal globally hyperbolic development of $\mathcal{S}$, but this requires treating the local existence and uniqueness problem for the spherically symmetric Einstein–Maxwell–Vlasov system at the center of symmetry, which we do not address here.[6] Indeed, since our charged Vlasov beams spacetimes will always be vacuum near $r = 0$, existence and uniqueness near the center will be completely trivial in our specific construction and is established in the following lemma.

[6]In the case $\mathfrak{m} > 0$ one could directly appeal to [BC73] to get local well-posedness near $r = 0$.

Lemma 3.2.16. *Let $u_0 \le v_0 < v_1$, $r_0 \ge 0$, $\lambda_0 > 0$, and $\alpha : [u_0, v_0] \to \mathbb{R}_{>0}$ and $\beta : [v_0, v_1] : \mathbb{R}_{>0}$ be smooth functions satisfying the relations*

$$\alpha(u_0) = \beta(v_0), \quad \alpha(v_0) = 4\lambda_0^2, \quad r_0 = \frac{1}{4\lambda_0}\int_{u_0}^{v_0} \alpha(u')\,du'.$$

Then there exists a unique smooth solution (r, Ω^2, Q, f) of the spherically symmetric Einstein–Maxwell–Vlasov system on

$$[u_0, v_0] \times [v_0, v_1] \cup (\{v \ge u\} \cap \{u \ge v_0\} \cap \{v \le v_1\})$$

with Q and f identically vanishing, satisfying the boundary conditions of Definition 3.2.10 along $\{u = v\}$, together with

$$r(u_0, v_0) = r_0, \quad \partial_v r(u_0, v_0) = \lambda_0, \quad \Omega^2|_{[u_0,v_0]\times\{v_0\}} = \alpha, \quad \Omega^2|_{\{u_0\}\times[v_0,v_1]} = \beta.$$

The solution is given by the explicit formulas

$$r(u, v) = r_0 + \frac{\lambda_0}{\beta(v_0)}\int_{v_0}^{v} \beta(v')\,dv' - \frac{1}{4\lambda_0}\int_{u_0}^{u} \alpha(u')\,du', \quad \Omega^2(u, v) = \frac{\alpha(u)\beta(v)}{\beta(v_0)}$$

for $(u, v) \in [u_0, v_0] \times [v_0, v_1]$ and

$$r(u, v) = \frac{\lambda_0}{\beta(v_0)}\int_{u}^{v} \beta(v')\,dv', \quad \Omega^2(u, v) = \frac{4\lambda_0}{\beta(v_0)^2}\beta(u)\beta(v)$$

for $(u, v) \in \{v \ge u\} \cap \{u \ge v_0\} \cap \{v \le v_1\}$.

Remark 3.2.17. From the last formula, it follows that

$$\partial_u \Omega^2(u, u) = \partial_v \Omega^2(u, u). \tag{3.2.44}$$

3.3 The characteristic initial value problem for spherically symmetric nonlinear wave-transport systems

In this section, we prove local well-posedness for the spherically symmetric Einstein–Maxwell–Vlasov system in small characteristic rectangles away from the center. In fact, we consider the general system of equations

$$\partial_u\partial_v\Psi = F(\Psi,\partial\Psi,Q,M[f],M^{uv}[f]), \tag{3.3.1}$$

$$\partial_u Q = K(\Psi, M^v[f]), \tag{3.3.2}$$

$$X(p^u,p^v,\Psi,\partial\Psi,Q)f = 0, \tag{3.3.3}$$

where $\Psi:\mathbb{R}^2_{u,v}\to\mathbb{R}^N$ is a vector-valued function taking the role of the "wave-type" variables r and Ω^2, $Q:\mathbb{R}^2_{u,v}\to\mathbb{R}$ is the charge, $f:T\mathbb{R}^2_{u,v}\to\mathbb{R}_{\geq 0}$ is the distribution function,

$$M[f]\doteq\int_0^\infty\int_0^\infty f\,dp^u dp^v,\quad M^v[f]\doteq\int_0^\infty\int_0^\infty p^v f\,dp^u dp^v,\quad M^{uv}[f]\doteq\int_0^\infty\int_0^\infty p^u p^v f\,dp^u dp^v$$

are moments of f, $F=(F_1,\dots,F_N)$ and K are smooth functions of their variables, and X is a vector field on $\mathbb{R}^2_{u,v}$ of the form

$$X(p^u,p^v,\Psi,\partial\Psi,Q) = p^u\partial_u + p^v\partial_v + \xi^u(p^u,p^v,\Psi,\partial\Psi,Q)\partial_{p^u} + \xi^v(p^u,p^v,\Psi,\partial\Psi,Q)\partial_{p^v},$$

where ξ^u and ξ^v are smooth functions of their variables. Letting $a\in\{u,v\}$, we can write X using Einstein notation as

$$X = p^a\partial_a + \xi^a\partial_{p^a}.$$

We assume that there exist functions $G_k:\mathbb{R}_{\geq 0}\to\mathbb{R}_{\geq 0}$ for $k\geq 0$ such that

$$|D^{i_1}_{\Psi,\partial\Psi,Q}\partial_p^{i_2}\xi^u(p^u,p^v,\Psi,\partial\Psi,Q)| + |D^{i_1}_{\Psi,\partial\Psi,Q}\partial_p^{i_2}\xi^v(p^u,p^v,\Psi,\partial\Psi,Q)| \leq G_k(M)\langle p^\tau\rangle^{2-i_2} \tag{3.3.4}$$

if $|\Psi|+|\partial\Psi|+|Q|\leq M$ and $i_1+i_2=k$, where $\langle s\rangle\doteq\sqrt{1+s^2}$ and $p^\tau\doteq\frac{1}{2}(p^u+p^v)$. Here $D^{i_1}_{\Psi,\partial\Psi,Q}\partial_p^{i_2}$ denotes any expression involving i_1 derivatives in the $(\Psi,\partial\Psi,Q)$-variables and i_2 derivatives in the (p^u,p^v)-variables. We also assume that there exists a constant $\mathfrak{m}\geq 0$ such that

$$\xi^u p^v + \xi^v p^u = 0 \tag{3.3.5}$$

whenever $p^u p^v = \mathfrak{m}^2$. These structural assumptions are verified for a renormalized version of the spherically symmetric Einstein–Maxwell–Vlasov system.

Given $U_0 < U_1$ and $V_0 < V_1$, let

$$\mathcal{C}(U_0,U_1,V_0,V_1) \doteq (\{U_0\}\times[V_0,V_1])\cup([U_0,U_1]\times\{V_0\}),$$
$$\mathcal{R}(U_0,U_1,V_0,V_1) \doteq [U_0,U_1]\times[V_0,V_1].$$

We will consistently omit (U_0,U_1,V_0,V_1) from the notation for these sets. We also define

$$P^{\mathfrak{m}} \doteq \{(u,v,p^u,p^v)\in T\mathbb{R}^2 : p^u\geq 0, p^v\geq 0, p^\tau>0, p^up^v\geq \mathfrak{m}^2\},$$
$$H^{\kappa} \doteq \{(u,v,p^u,p^v)\in T\mathbb{R}^2 : p^u\geq 0, p^v\geq 0, p^\tau>\kappa\}$$

and set $P \doteq P^0$. A function $\phi:\mathcal{C}\to\mathbb{R}$ is said to be *smooth* if it is continuous and $\phi|_{\{U_0\}\times[V_0,V_1]}$ and $\phi|_{[U_0,U_1]\times\{V_0\}}$ are C^∞ single-variable functions. This definition extends naturally to functions $f : P^{\mathfrak{m}}|_{\mathcal{C}}\to\mathbb{R}$. A smooth *characteristic initial data set* for the system (3.3.1)–(3.3.3) consists of a triple $(\mathring{\Psi},\mathring{Q},\mathring{f})$ and numbers $\kappa>0$, $\sigma>4$, where $\mathring{\Psi}:\mathcal{C}\to\mathbb{R}^N$, $\mathring{Q}:\{U_0\}\times[V_0,V_1]\to\mathbb{R}$, and $\mathring{f}:P^{\mathfrak{m}}|_{\mathcal{C}}\to\mathbb{R}_{\geq 0}$ are smooth. We additionally assume that $\operatorname{spt}(\mathring{f})\subset H^\kappa|_{\mathcal{C}}$ for some $\kappa>0$ (which is only an extra assumption when $\mathfrak{m}=0$) and that

$$\|\mathring{f}\|_{C^k_\sigma(P|_{\mathcal{C}})} \doteq \sum_{0\leq i_1+i_2\leq k}\left(\sup_{P^{\mathfrak{m}}|_{\{U_0\}\times[V_0,V_1]}}\langle p^\tau\rangle^{\sigma+i_2}|\partial_v^{i_1}\partial_p^{i_2}\mathring{f}| + \sup_{P^{\mathfrak{m}}|_{[U_0,U_1]\times\{V_0\}}}\langle p^\tau\rangle^{\sigma+i_2}|\partial_u^{i_1}\partial_p^{i_2}\mathring{f}|\right) \tag{3.3.6}$$

is finite for every $k\geq 0$. For $f:P^{\mathfrak{m}}|_{\mathcal{R}}\to\mathbb{R}_{\geq 0}$ and $k\geq 0$, we define the norms

$$\|f\|_{C^k_\sigma(P|_{\mathcal{R}})} \doteq \sum_{0\leq i_1+i_2\leq k}\sup_{P^{\mathfrak{m}}|_{\mathcal{R}}}\langle p^\tau\rangle^{\sigma+i_2}|\partial_x^{i_1}\partial_p^{i_2}f|,$$

where $\partial_x^{i_1}$ denotes i_1 derivatives in the (u,v)-variables.

Proposition 3.3.1. *For any $B>0$, $\kappa>0$, and $\sigma>4$ there exists a constant $\varepsilon>0$ (depending also on F, K, and X) with the following property. Let $(\mathring{\Psi},\mathring{Q},\mathring{f})$ be a smooth characteristic initial data set for the system* (3.3.1)–(3.3.3) *on $\mathcal{C}(U_0,U_1,V_0,V_1)$. If $U_1-U_0<\varepsilon$, $V_1-V_0<\varepsilon$, and*

$$\|\mathring{\Psi}\|_{C^2(\mathcal{C})}+\|\mathring{Q}\|_{C^1(\mathcal{C})}+\|\mathring{f}\|_{C^1_\sigma(P|_{\mathcal{C}})}\leq B, \tag{3.3.7}$$

then there exists a unique smooth solution (Ψ,Q,f) of (3.3.1)–(3.3.3) *on $\mathcal{R}(U_0,U_1,V_0,V_1)$ which extends the initial data. Moreover, the distribution function f is supported in $H^{\kappa/2}$ and for any*

$k \geq 0$, the norm

$$\|\Psi\|_{C^k(\mathcal{R})} + \|Q\|_{C^k(\mathcal{R})} + \|f\|_{C^k_\sigma(P|_\mathcal{R})}$$

is finite and can be bounded in terms of initial data norms.

Remark 3.3.2. While we assume that the initial data $(\mathring{\Psi}, \mathring{Q}, \mathring{f})$ are smooth (and that $\mathring{f}$ satisfies the nontrivial bound (3.3.6) at any order), the existence time ε in the proposition depends only on the estimate (3.3.7).

Remark 3.3.3. If we assume that $\mathring{f}$ has compact support in the momentum variables, then (3.3.6) is automatic by smoothness.

3.3.1 Proof of Proposition 3.3.1

In this section, we assume the hypotheses and setup of Proposition 3.3.1. We also set $U_0 = V_0 = 0$ and define $\tau \doteq \frac{1}{2}(v+u)$. Therefore, $0 \leq \tau \leq \varepsilon$ on $\mathcal{R}$.

We will construct the solution (Ψ, Q, f) as the limit of an iteration scheme.

Lemma 3.3.4. *There exist sequences of constants $\{\tilde{C}_k\}$ and $\{C_k\}$ such that the following holds. For any ε sufficiently small and every $n \geq 1$, there exist functions $(\Psi_n, Q_n, f_n) \in C^\infty(\mathcal{R}) \times C^\infty(\mathcal{R}) \times C^\infty(P^\mathfrak{m}|_\mathcal{R})$ solving the iterative system*

$$\partial_u\partial_v\Psi_n = F(\Psi_{n-1}, \partial\Psi_{n-1}, Q_{n-1}, M[f_{n-1}], M^{uv}[f_{n-1}]), \tag{3.3.8}$$

$$\partial_u Q_n = K(\Psi_{n-1}, M^v[f_{n-1}]), \tag{3.3.9}$$

$$X(p^u, p^v, \Psi_{n-1}, \partial\Psi_{n-1}, Q_{n-1}) f_n = 0, \tag{3.3.10}$$

where we set (Ψ_0, Q_0, f_0) to be identically zero, with initial conditions

$$\Psi_n|_\mathcal{C} = \mathring{\Psi}, \quad Q_n|_\mathcal{C} = \mathring{Q}, \quad f_n|_{P|_\mathcal{C}} = \mathring{f}. \tag{3.3.11}$$

Moreover, $\mathrm{spt}(f_n) \subset H^{\kappa/2}$ and these functions satisfy the bounds

$$\|\Psi\|_{C^k(\mathcal{R})} \leq \tilde{C}_k e^{C_k\tau}, \tag{3.3.12}$$

$$\|Q_n\|_{C^k(\mathcal{R})} \leq \tilde{C}_{k+1} e^{C_{k+1}\tau}, \tag{3.3.13}$$

$$\|f_n\|_{C^k_\sigma(P|_\mathcal{R})} \leq \tilde{C}_{k+1} e^{C_{k+1}\tau}. \tag{3.3.14}$$

It is convenient to set

$$\begin{aligned}
F_n &\doteq F(\Psi_n, \partial\Psi_n, Q_n, M[f_n], M^{uv}[f_n]),\\
K_n &\doteq K(\Psi_n, M^v[f_n]),\\
X_n &\doteq X(p^u, p^v, \Psi_n, \partial\Psi_n, Q_n)
\end{aligned}$$

for $n \geq 0$. We first require a preliminary lemma about integral curves of the vector field X_n.

Lemma 3.3.5. *For $n \geq 0$, let $\Gamma_n^{\mathfrak{m},\kappa}$ denote the set of maximal integral curves $\tilde{\gamma} = (\gamma, p) : I \to T\mathcal{R}$ (where I is a closed interval containing 0) of the vector field X_n subject to the condition that $\tilde{\gamma}(0) \in H^\kappa \cap P^{\mathfrak{m}}$. Assume that Ψ_n satisfies (3.3.12) for $k = 1$ and Q_n satisfies (3.3.13) for $k = 0$. Then for ε sufficiently small (depending in particular on κ) and any $\tilde{\gamma} \in \Gamma_n$, γ is a future-directed causal curve in $\mathcal{R}$ connecting $\mathcal{C}$ with the future boundary of $\mathcal{R}$, $\tilde{\gamma}(s) \in P^{\mathfrak{m}} \cap H^{\kappa/2}$ for every $s \in I$, and*

$$\frac{1}{2}p^\tau(0) \leq p^\tau(s) \leq 2p^\tau(0) \tag{3.3.15}$$

for every $s \in I$.

Proof. Let $\tilde{\gamma} = (\gamma, p) \in \Gamma_n$ and set $\tau_0 \doteq \gamma^\tau(0)$. By definition,

$$\frac{d\gamma^a}{ds} = p^a, \quad \frac{dp^a}{ds} = \xi^a_n$$

for $a \in \{u, v\}$, where $\xi^a_n \doteq \xi^a(p^u, p^v, \Psi_n, \partial\Psi_n, Q_n)$. Observe that X_n is tangent to the the boundary of $P^{\mathfrak{m}}$, $\{(u, v, p^u, p^v) \in T\mathcal{R} : p^u p^v = \mathfrak{m}^2\}$, by (3.3.5), so $\tilde{\gamma}$ remains within $P^{\mathfrak{m}}$. Reparametrizing $\tilde{\gamma}$ by τ gives

$$\frac{d}{d\tau}(p^\tau)^2 = 2(\xi^u_n + \xi^v_n).$$

Using (3.3.4), the assumptions on Ψ_n and Q_n, and Grönwall's inequality, we have

$$p^\tau(\tau)^2 \leq e^{O(\varepsilon)}\left(p^\tau(\tau_0)^2 + O(\varepsilon)\right) \leq 2p^\tau(\tau_0)^2$$

for τ in the domain of $\tilde{\gamma}$ and for ε sufficiently small. This proves the second inequality in (3.3.15). To prove the first inequality, we observe that by the estimate we have just proved,

$$|(p^\tau(\tau))^2 - (p^\tau(\tau_0))^2| \lesssim \varepsilon \sup_{\tilde{\gamma}} |\xi^u_n + \xi^v_n| \lesssim \varepsilon p^\tau(\tau_0)^2$$

Choosing ε perhaps even smaller proves (3.3.15) and completes the proof of the lemma. □

Proof of Lemma 3.3.4. The proof is by induction on n and induction on k for each fixed n. As the existence and estimates for the base case $n=0$ are trivial, we assume the existence of $(\Psi_{n-1}, Q_{n-1}, f_{n-1})$ satisfying (3.3.12)–(3.3.14), where the constants are still to be determined. We will choose the constants to satisfy $\tilde{C}_k \le C_k \le \tilde{C}_{k+1}$, which we use without comment in the sequel. By (3.3.12)–(3.3.14) for $(\Psi_{n-1}, Q_{n-1}, f_{n-1})$ and iterating the chain rule, it is easy to see that

$$|\partial^k \mathcal{M}_{n-1}| \le C(\tilde{C}_{k+1}) e^{C_{k+1}\tau} \quad (k \ge 0), \tag{3.3.16}$$

$$|F_{n-1}| \le C(C_1), \tag{3.3.17}$$

$$|\partial^k F_{n-1}| \le C(\tilde{C}_{k+1}) e^{C_{k+1}\tau} \quad (k \ge 1), \tag{3.3.18}$$

$$|K_{n-1}| \le C(C_1), \tag{3.3.19}$$

$$|\partial^k K_{n-1}| \le C(\tilde{C}_{k+1}) e^{C_{k+1}\tau} \quad (k \ge 1), \tag{3.3.20}$$

where $\mathcal{M}_{n-1} \in \{M[f_{n-1}], M^v[f_{n-1}], M^{uv}[f_{n-1}]\}$. We also define the number

$$B' \doteq |F(\mathring{\Psi}, \partial\mathring{\Psi}, \mathring{Q}, M[\mathring{f}], M^{uv}[\mathring{f}])|_{(0,0)} + |K(\mathring{\Psi}, M^v[\mathring{f}])|_{(0,0)}.$$

Step 1. The function Ψ_n is defined by the explicit formula

$$\Psi_n(u,v) \doteq \int_0^u \int_0^v F_{n-1}(u',v')\, dv' du' + \mathring{\Psi}(u,0) + \mathring{\Psi}(0,v) - \mathring{\Psi}(0,0). \tag{3.3.21}$$

It follows by inspection of this representation formula and (3.3.17) that

$$\|\Psi_n\|_{C^1(\mathcal{R})} \le 10B \tag{3.3.22}$$

if ε is sufficiently small depending on C_1. We estimate k-th order derivatives ($k \ge 2$) of the form $\partial_u^k \Psi_n$, $\partial_x^{k-2}\partial_u\partial_v\Psi_n$, and $\partial_v^k\Psi_n$ separately. For the first type, we simply differentiate the representation formula (3.3.21) k times to obtain

$$|\partial_u^k \Psi_n| \le (\text{data}) + \int_0^v |\partial_u^{k-1} F_{n-1}|\, dv' \le C + \frac{C(\tilde{C}_k)}{C_k} e^{C_k\tau} \le \tilde{C}_k e^{C_k\tau}$$

for appropriate choices of $\tilde{C}_k$ and C_k. For mixed derivatives, we differentiate the wave equation to obtain

$$|\partial_x^{k-2}\partial_u\partial_v\Psi_{n-1}| \leq \sup_{\{0\}\times[0,V_0]} |\partial_x^{k-2}F_{n-1}| + \int_0^u |\partial_u\partial^{k-2}F_{n-1}|\,du' \leq C(C_{k-1}) + \frac{C(\tilde{C}_k)}{C_k}e^{C_k\tau} \leq \tilde{C}_k e^{C_k\tau}$$

for appropriate choices of $\tilde{C}_k$ and C_k. The estimate for $\partial_v^k\Psi_n$ is similar to $\partial_u^k\Psi_n$. For later use, we derive a slightly improved estimate for $\partial_x^2\Psi_n$. Using the mean value theorem and (3.3.18), we have

$$|F_{n-1} - B'| \leq C(C_2)\tau.$$

We also have

$$|\partial_u^2\Psi_n - (\text{data})| + |\partial_v^2\Psi_n - (\text{data})| \leq C(C_2)\tau.$$

Therefore, for ε sufficiently small depending on C_2, combined with (3.3.22), we infer

$$\|\Psi_n\|_{C^2(\mathcal{R})} \leq 20(B + B'). \tag{3.3.23}$$

This completes Step 1.

Step 2. The function Q_n is defined by the explicit formula

$$Q_n(u,v) = \int_0^u K_{n-1}(u',v)\,du' + \mathring{Q}(0,v).$$

It follows by inspection of this representation formula and (3.3.19) that

$$\|Q\|_{C^0(\mathcal{R})} \leq 10B$$

for ε sufficiently small depending on C_1. For $k \geq 1$, we estimate

$$|\partial_v^k Q_n| \leq (\text{data}) + \int_0^u |\partial_v^k K_{n-1}|\,du' \leq C + \frac{C(\tilde{C}_{k+1})}{C_{k+1}}e^{C_{k+1}\tau} \leq \tilde{C}_{k+1}e^{C_{k+1}\tau},$$

$$|\partial^{k-1}\partial_u Q_n| \leq \sup_{\{0\}\times[0,V_0]} |\partial^{k-1}K_{n-1}| + \int_0^u |\partial^k K_{n-1}|\,du' \leq C(C_k) + \frac{C(\tilde{C}_{k+1})}{C_{k+1}}e^{C_{k+1}\tau} \leq \tilde{C}_{k+1}e^{C_{k+1}\tau}$$

for appropriate chices of $\tilde{C}_{k+1}$ and C_{k+1}. Arguing as in Step 1, for ε sufficiently small depending on C_2, we also infer

$$\|Q_n\|_{C^1(\mathcal{R})} \leq 20(B + B'). \tag{3.3.24}$$

This completes Step 2.

Step 3. For f_n we do not have an explicit representation formula and must instead infer its existence from general properties of flows of vector fields. A slight technical issue is that the "initial data hypersurface" $P|_{\mathcal{C}}$ is not smooth because of the corner in $\mathcal{C}$. Therefore, in order to prove the existence of a smooth f_n, we will construct it as a smooth limit of smooth solutions to the X_{n-1} transport equation, corresponding to initial data where we smooth out the corner in $\mathcal{C}$. To carry out this idea, we first extend $\mathring{f}$ to $C^\infty_\sigma(P^{\mathfrak{m}}|_{\mathcal{R}})$ according to

$$\mathcal{E}\mathring{f}(u,v,p^u,p^v) \doteq \mathring{f}(u,0,p^u,p^v) + \mathring{f}(0,v,p^u,p^v) - \mathring{f}(0,0,p^u,p^v)$$

and set, for $j \geq 1$,

$$\mathcal{S}_j \doteq \{(u,v) \in \mathcal{R} : uv = 2^{-j}\},$$
$$\mathcal{R}_j \doteq \{(u,v) \in \mathcal{R} : uv \geq 2^{-j}.\}$$

For any $j \geq j_0$ sufficiently large that $\mathcal{S}_j \neq \emptyset$, there exists a unique function $f_{n,j} \in C^\infty(P^{\mathfrak{m}}|_{\mathcal{R}_j})$ such that

$$X_{n-1} f_{n,j} = 0 \tag{3.3.25}$$

in $\mathcal{R}_j$, with initial data $f_{n,j} = \mathcal{E}\mathring{f}$ on $P^{\mathfrak{m}}|_{\mathcal{S}_j}$. The existence follows immediately from the flowout theorem (see [Lee13, Theorem 9.20]), the fact that X_{n-1} is transverse to $P^{\mathfrak{m}}|_{\mathcal{S}_j}$, and Lemma 3.3.5. It also follows from Lemma 3.3.5 that $\operatorname{spt}(f_{n,j}) \subset H^{\kappa/2}$. We will use the fact that $p^\tau \geq \kappa/2$ for any $\tilde{\gamma} \in \Gamma^{\mathfrak{m},\kappa}_{n-1}$ often and without further comment in the sequel.

We now claim that we can choose $\tilde{C}_k$ and C_k such that

$$\|f_{n,j}\|_{C^k_\sigma(P|_{\mathcal{R}_j})} \leq \tilde{C}_{k+1} e^{C_{k+1}\tau} \tag{3.3.26}$$

for every $n \geq 0$, $j \geq j_0$, and $k \geq 0$. Let $\mathcal{F}_{n,j,k}$ denote the vector with $\binom{k+3}{k}$ components of the form

$$(p^\tau)^{\sigma+i_2} \partial_x^{i_1} \partial_p^{i_2} f_{n,j}, \tag{3.3.27}$$

where $i_1 + i_2 = k$. We will show inductively that $\tilde{C}_{k+1}$ and C_{k+1} can be chosen so that

$$\sup_{P^{\mathfrak{m}}|_{\mathcal{R}_j}} |\mathcal{F}_{n,j,k}| \leq \tilde{C}_{k+1} e^{C_{k+1}\tau}, \tag{3.3.28}$$

which would imply (3.3.26).

Orders $k = 0$ and $k = 1$ are slightly anomalous in our scheme and we handle them first. We require the estimate

$$\langle p^\tau \rangle^{-2} |\partial_x^{\leq 1} \xi_{n-1}^a| + \langle p^\tau \rangle^{-1} |\partial_p \xi_{n-1}^a| \lesssim 1, \tag{3.3.29}$$

which follows from (3.3.4), (3.3.23), and (3.3.24). Using (3.3.25), we compute

$$X_{n-1}((p^\tau)^\sigma f_{n,j}) = \frac{\sigma}{2}(p^\tau)^{\sigma-1}(\xi_{n-1}^u + \xi_{n-1}^v) f_{n,j}.$$

From (3.3.29) we then infer

$$|X_{n-1}((p^\tau)^\sigma f_{n,j})| \lesssim (p^\tau)^{\sigma+1} f_{n,j}.$$

Let $\tilde{\gamma} \in \Gamma_{n-1}^{\mathfrak{m},\kappa}$ be parametrized by coordinate time τ. Then along $\tilde{\gamma}$ we have

$$\left| \frac{d}{d\tau} \mathcal{F}_{n,j,0} \right| \lesssim |\mathcal{F}_{n,j,0}|,$$

whence by Grönwall's inequality

$$|\mathcal{F}_{n,j,0}| \lesssim \|\mathcal{E} \mathring{f}\|_{C^0_\sigma(P|_\mathcal{R})} \lesssim \|\mathring{f}\|_{C^0_\sigma(P|_\mathcal{C})}.$$

Next, we compute

$$|X_{n-1}((p^\tau)^\sigma \partial_x f_{n,j})| \lesssim (p^\tau)^{\sigma+1} |\partial_x f_{n,j}| + (p^\tau)^\sigma |[X_{n-1}, \partial_x] f_{n,j}|.$$

The commutator is estimated using (3.3.29) by

$$|[X_{n-1}, \partial_x] f_{n,j}| \lesssim |\partial_x \xi_{n-1}^a| |\partial_p f_{n,j}| \lesssim (p^\tau)^2 |\partial_p f_{n,j}|.$$

The p derivative of $f_{n,j}$ satisfies

$$|X_{n-1}((p^\tau)^{\sigma+1} \partial_p f_{n,j})| \lesssim (p^\tau)^{\sigma+2} |\partial_p f_{n,j}| + (p^\tau)^{\sigma+1} |[X_{n-1}, \partial_p] f_{n,j}|,$$

where the commutator is now estimated by

$$|[X_{n-1}, \partial_p] f_{n,j}| \lesssim |\partial_x f_{n,j}| + |\partial_p \xi_{n-1}^a| |\partial_x f_{n,j}| \lesssim |\partial_x f_{n,j}| + p^\tau |\partial_p f_{n,j}|.$$

Putting these estimates together, we find that

$$\left|\frac{d}{d\tau}\mathcal{F}_{n,j,1}\right| \lesssim |\mathcal{F}_{n,j,1}|,$$

whence again by Grönwall's inequality we conclude a uniform bound

$$|\mathcal{F}_{n,j,1}| \lesssim \|\mathring{f}\|_{C^1_\sigma(P|_\mathcal{C})}. \tag{3.3.30}$$

Having now established cases $k = 0$ and $k = 1$ of (3.3.28), we now assume (3.3.28) up to order $k - 1$. Let φ_{i_1,i_2} be a component of $\mathcal{F}_{n,j,k}$. We adopt the convention that if either i_1 or i_2 are negative, then φ_{i_1,i_2} is interpreted as identically zero. Using (3.3.4) and (3.3.29), we estimate

$$\begin{aligned}|X_{n-1}(\varphi_{i_1,i_2})| &\lesssim (p^\tau)^{\sigma+i_2-1}|\xi^u_{n-1}+\xi^v_{n-1}||\partial_x^{i_1}\partial_p^{i_2}f_{n,j}| + (p^\tau)^{\sigma+i_2}|[X_{n-1},\partial^k]f_{n,j}| \\ &\lesssim p^\tau|\varphi_{i_1,i_2}| + (p^\tau)^{\sigma+i_2}|[p^a\partial_a,\partial_x^{i_1}\partial_p^{i_2}]f_{n,j}| + (p^\tau)^{\sigma+i_2}|[\xi^a_{n-1}\partial_{p^a},\partial_x^{i_1}\partial_p^{i_2}]f_{n,j}|.\end{aligned} \tag{3.3.31}$$

The first commutator, $[p^a\partial_a,\partial_x^{i_1}\partial_p^{i_2}]f_{n,j}$, vanishes unless $i_1 \geq 1$ and therefore consists of terms of the form $\partial_x^{i_1+1}\partial_p^{i_2-1}f_{n,j}$, which implies

$$(p^\tau)^{\sigma+i_2}|[p^a\partial_a,\partial_x^{i_1}\partial_p^{i_2}]f_{n,j}| \lesssim p^\tau|\varphi_{i_1+1,i_2-1}| \lesssim p^\tau|\mathcal{F}_{n,j,k}|. \tag{3.3.32}$$

The second commutator can be estimated by

$$|[\xi^a_{n-1}\partial_{p^a},\partial_x^{i_1}\partial_p^{i_2}]f| \lesssim \sum_{\substack{1\leq j_1+j_2\\ j_1\leq i_1, j_2\leq i_2}} |\partial_x^{j_1}\partial_p^{j_2}\xi^a_{n-1}||\partial_x^{i_1-j_1}\partial_p^{i_2+1-j_2}f_{n,j}|.$$

By inspection, $\partial_x^{j_1}\partial_p^{j_2}\xi^a_{n-1}$ is linear in $\partial_x^{j_1+1}\Psi_{n-1}$, which is the worst behaved term in our inductive hierarchy. Therefore, using again (3.3.4), we may estimate

$$|\partial_x^{j_1}\partial_p^{j_2}\xi^a_{n-1}| \lesssim \left(C(C_k) + \tilde{C}_{j_1+1}e^{C_{j_1+1}\tau}\right)(p^\tau)^{2-j_2}$$

and therefore infer

$$(p^\tau)^{\sigma+i_2}|[\xi^a_{n-1}\partial_{p^a},\partial_x^{i_1}\partial_p^{i_2}]f_{n,j}| \lesssim p^\tau \sum_{\substack{1\leq j_1+j_2\\ j_1\leq i_1, j_2\leq i_2}} \left(C(C_k) + \tilde{C}_{j_1+1}e^{C_{j_1+1}\tau}\right)|\mathcal{F}_{n,j,k+1-j_1-j_2}|.$$

If $j_1 < k$, then $\tilde{C}_{j_1+1}e^{C_{j_1+1}\tau} \leq C(C_k)$. If $j_1 = k$, then $j_2 = 0$ and

$$\left(C(C_k) + \tilde{C}_{j_1+1}e^{C_{j_1+1}\tau}\right)|\mathcal{F}_{n,j,k+1-j_1-j_2}| = \left(C(C_k) + \tilde{C}_{k+1}e^{C_{k+1}\tau}\right)|\mathcal{F}_{n,j,1}| \lesssim C(C_k)\tilde{C}_{k+1}e^{C_{k+1}\tau}, \tag{3.3.33}$$

where we have used (3.3.30). Therefore, the sum in (3.3.1) can be estimated by

$$\lesssim C(C_k)\left(|\mathcal{F}_{n,j,k}| + \tilde{C}_{k+1}e^{C_{k+1}\tau}\right).$$

Putting (3.3.31), (3.3.32), (3.3.1), and (3.3.33) together, we arrive at

$$|X_{n-1}\mathcal{F}_{n,j,k}| \lesssim C(C_k)p^\tau\left(|\mathcal{F}_{n,j,k}| + \tilde{C}_{k+1}e^{C_{k+1}\tau}\right).$$

As before, a simple Grönwall argument now establishes (3.3.26) for appropriate choices of $\tilde{C}_{k+1}$ and C_{k+1}.

Having now established the boundedness of the sequence $f_{n,j}$, we may take the limit $j \to \infty$ (after perhaps passing to a subsequence). This shows the existence of a function $f_n \in C^\infty(P^{\mathfrak{m}}|_{\mathcal{R}})$ with $\mathrm{spt}(f_n) \subset H^{\kappa/2}$, satisfying the estimates (3.3.14), and attaining $\mathring{f}$ on $P^{\mathfrak{m}}|_{\mathcal{C}}$. Finally, uniqueness of f_n is immediate since it is constant along the integral curves of X_{n-1}. □

Proof of Proposition 3.3.1. We prove first that the sequence iterative sequence (Ψ_n, Q_n, f_n) constructed in Lemma 3.3.4 is Cauchy in $C^2 \times C^1 \times C^1_\sigma$. We claim that if ε is sufficiently small, then the following estimate holds for every $n \geq 2$:

$$\begin{aligned}&\|\Psi_n - \Psi_{n-1}\|_{C^2(\mathcal{R})} + \|Q_n - Q_{n-1}\|_{C^1(\mathcal{R})} + \|f_n - f_{n-1}\|_{C^1_\sigma(P|_{\mathcal{R}})}\\ &\qquad \leq \frac{1}{2}\left(\|\Psi_{n-1} - \Psi_{n-2}\|_{C^2(\mathcal{R})} + \|Q_{n-1} - Q_{n-2}\|_{C^1(\mathcal{R})} + \|f_{n-1} - f_{n-2}\|_{C^1_\sigma(P|_{\mathcal{R}})}\right).\end{aligned} \tag{3.3.34}$$

Using the mean value theorem and the boundedness of the iterative sequence, we immediately estimate

$$\|F_{n-1}-F_{n-2}\|_{C^1(\mathcal{R})} \lesssim \|\Psi_{n-1}-\Psi_{n-2}\|_{C^2(\mathcal{R})}+\|Q_{n-1}-Q_{n-2}\|_{C^1(\mathcal{R})}+\|f_{n-1}-f_{n-2}\|_{C^1_\sigma(P|_{\mathcal{R}})}. \tag{3.3.35}$$

Using the formula

$$(\Psi_n - \Psi_{n-1})(u,v) = \int_0^u\int_0^v (F_{n-1} - F_{n-2})\,dv'du'$$

we readily infer that for ε sufficiently small, $\Psi_n - \Psi_{n-1}$ and $\partial_a^i(\Psi_n - \Psi_{n-1})$ for $a \in \{u, v\}$ and

$i \in \{1,2\}$ are bounded by an arbitrarily small multiple of the right-hand side of (3.3.34). To estimate the mixed second partial derivative, we simply use the fundamental theorem of calculus to bound

$$\|\Psi_{n-1} - \Psi_{n-2}\|_{C^1(\mathcal{R})} + \|Q_{n-1} - Q_{n-2}\|_{C^0(\mathcal{R})} + \|\mathcal{M}_{n-1} - \mathcal{M}_{n-2}\|_{C^0(\mathcal{R})}$$

by an arbitrarily small multiple of the right-hand side of (3.3.34) and then use the mean value theorem to estimate

$$\begin{aligned}|\partial_u \partial_v (\Psi_n - \Psi_{n-1})| &\leq \|F_{n-1} - F_{n-2}\|_{C^0(\mathcal{R})} \\ &\lesssim \|\Psi_{n-1} - \Psi_{n-2}\|_{C^1(\mathcal{R})} + \|Q_{n-1} - Q_{n-2}\|_{C^0(\mathcal{R})} + \|\mathcal{M}_{n-1} - \mathcal{M}_{n-2}\|_{C^0(\mathcal{R})}.\end{aligned}$$

The argument for bounding $\|Q_n - Q_{n-1}\|_{C^1(\mathcal{R})}$ is essentially the same and is omitted.

To estimate $f_n - f_{n-1}$, we examine the quantity

$$\mathcal{F}_n \doteq (p^\tau)^\sigma \left(f_n - f_{n-1}, \partial_x(f_n - f_{n-1}), p^\tau \partial_p (f_n - f_{n-1})\right)$$

along integral curves of X_{n-1}. First, note that $\mathcal{F}_n \neq 0$ only along curves in $\Gamma^{\mathfrak{m},\kappa}_{n-1} \cup \Gamma^{\mathfrak{m},\kappa}_{n-2}$. By Lemma 3.3.5 (and choosing ε perhaps smaller), any value $\tilde{\gamma}_{n-2}(s)$ for a curve $\tilde{\gamma}_{n-2} \in \Gamma^{\mathfrak{m},\kappa}_{n-2}$ can be realized as an initial value $\tilde{\gamma}_{n-1}(0)$ for a curve $\tilde{\gamma}_{n-1} \in \Gamma^{\mathfrak{m},\kappa/2}_{n-1}$. Therefore, it suffices to observe that the following estimate holds along any curve $\tilde{\gamma}_{n-1} \in \Gamma^{\mathfrak{m},\kappa/2}_{n-1}$:

$$\left|\frac{d}{d\tau}\mathcal{F}_n\right| \lesssim |\mathcal{F}_n| + \|\Psi_{n-1} - \Psi_{n-2}\|_{C^2(\mathcal{R})} + \|Q_{n-1} - Q_{n-2}\|_{C^1(\mathcal{R})} + \|f_{n-1} - f_{n-2}\|_{C^1_\sigma(P|_{\mathcal{R}})},$$

which is obtained by simply differentiating $\mathcal{F}_n$ and using the estimates proved in Lemma 3.3.4. Therefore, since $\mathcal{F}_n$ vanishes along $P^{\mathfrak{m}}|_{\mathcal{C}}$, Grönwall's inequality implies

$$|\mathcal{F}_n| \lesssim \varepsilon \left(\|\Psi_{n-1} - \Psi_{n-2}\|_{C^2(\mathcal{R})} + \|Q_{n-1} - Q_{n-2}\|_{C^1(\mathcal{R})} + \|f_{n-1} - f_{n-2}\|_{C^1_\sigma(P|_{\mathcal{R}})}\right).$$

After choosing ε sufficiently small, the proof of (3.3.35) is complete.

Therefore, (Ψ_n, Q_n, f_n) converges to a solution (Ψ, Q, f) of the system (3.3.1)–(3.3.3) in $C^2 \times C^1 \times C^1$, which is moreover C^∞ smooth by the higher order estimates proved in Lemma 3.3.4. Uniqueness of the solution can be proved along the same lines as the proof of the estimate (3.3.34) and is omitted. □

3.3.2 Proof of local well posedness for the Einstein–Maxwell–Vlasov system

In this section, we prove Proposition 3.2.3, local well-posedness for the Einstein–Maxwell–Vlasov system in small characteristic rectangles. The proof has 3 steps: In the first step, we solve the wave equations (2.3.20) and (2.3.21), the *ingoing* Maxwell equation (2.3.24), and the Maxwell–Vlasov equation (2.3.26) using Proposition 3.3.1. In order to directly quote Proposition 3.3.1, we in fact consider a *renormalized* system that fixes the location of the mass shell to the fixed family of hyperbolas $\tilde{p}^u\tilde{p}^v = \mathfrak{m}^2$. In step 2, we show that the *outgoing* Maxwell equation holds as a result of conservation law (2.3.32), which follows from the Maxwell–Vlasov equation (2.3.26). Finally, in step 3, we show that Raychaudhuri's equations (2.3.22) and (2.3.23) hold, using now the Bianchi identities (2.3.33) and (2.3.34), which again follow from (2.3.26). Steps 2 and 3 may be thought of as *propagation of constraints*, as they require the relevant equations to hold on initial data.

Proof of Proposition 3.2.3. Step 1. Consider the wave-transport system

$$\partial_u\partial_v r = -\frac{\Omega^2}{4r} - \frac{\partial_u r\partial_v r}{r} + \frac{\pi r\Omega^2}{4}\int_{\tilde{p}^u\tilde{p}^v\geq\mathfrak{m}^2} \tilde{p}^u\tilde{p}^v\tilde{f}(u,v,\tilde{p}^u,\tilde{p}^v)\,d\tilde{p}^u d\tilde{p}^v, \tag{3.3.36}$$

$$\begin{aligned}\partial_u\partial_v \log\Omega^2 = {} & \frac{\Omega^2}{2r^2} + \frac{2\partial_u r\partial_v r}{r^2} + \frac{\pi\Omega^2}{2}\int_{\tilde{p}^u\tilde{p}^v\geq\mathfrak{m}^2} \tilde{p}^u\tilde{p}^v\tilde{f}(u,v,\tilde{p}^u,\tilde{p}^v)\,d\tilde{p}^u d\tilde{p}^v \\ & - \frac{\pi\Omega^2}{2}\int_{\tilde{p}^u\tilde{p}^v\geq\mathfrak{m}^2} (\tilde{p}^u\tilde{p}^v-\mathfrak{m}^2)\tilde{f}(u,v,\tilde{p}^u,\tilde{p}^v)\,d\tilde{p}^u d\tilde{p}^v,\end{aligned} \tag{3.3.37}$$

$$\partial_u Q = -\frac{\pi}{2}\mathfrak{e}r^2\Omega\int_{\tilde{p}^u\tilde{p}^v\geq\mathfrak{m}^2} \tilde{p}^u\tilde{f}(u,v,\tilde{p}^u,\tilde{p}^v)\,d\tilde{p}^u d\tilde{p}^v, \tag{3.3.38}$$

$$\tilde{X}\tilde{f} = 0, \tag{3.3.39}$$

where

$$\begin{aligned}\tilde{X} \doteq {} & \tilde{p}^u\partial_u + \tilde{p}^v\partial_v - \left(\partial_u\log\Omega(\tilde{p}^u)^2 - \partial_v\log\Omega\tilde{p}^u\tilde{p}^v + \frac{2\partial_v r}{r}(\tilde{p}^u\tilde{p}^v-\mathfrak{m}^2) + \mathfrak{e}\frac{\Omega Q}{r^2}\tilde{p}^u\right)\partial_{\tilde{p}^u} \\ & - \left(\partial_v\log\Omega(\tilde{p}^v)^2 - \partial_u\log\Omega\tilde{p}^u\tilde{p}^v + \frac{2\partial_u r}{r}(\tilde{p}^u\tilde{p}^v-\mathfrak{m}^2) - \mathfrak{e}\frac{\Omega Q}{r^2}\tilde{p}^v\right)\partial_{\tilde{p}^v},\end{aligned}$$

with initial data $\mathring{\Psi} = (\log\mathring{r}, \log\mathring{\Omega}^2)$, $\mathring{Q}$, and $\mathring{\tilde{f}}(u,v,\tilde{p}^u,\tilde{p}^v) = \mathring{f}(u,v,\mathring{\Omega}^{-1}\tilde{p}^u,\mathring{\Omega}^{-1}\tilde{p}^v)$. Since $\log\mathring{\Omega}^2$ is bounded on $\mathcal{C}$ and either $\mathfrak{m}>0$ or $\ell\geq c_\ell$ on $\operatorname{spt}(\mathring{f})$, there exists a $\kappa>0$ such that $\operatorname{spt}(\mathring{\tilde{f}})\subset H^\kappa$. Furthermore, the structural conditions (3.3.4) and (3.3.5) are easily verified for $\tilde{X}$, so Proposition 3.3.1 produces a unique local smooth solution to (3.3.36)–(3.3.39) if $\varepsilon_{\mathrm{loc}}$ is chosen sufficiently small. Making the change of variables $\tilde{p}^u\mapsto\Omega p^u$ and $\tilde{p}^v\mapsto\Omega p^v$, defining $f(u,v,p^u,p^v) = \tilde{f}(u,v,\Omega p^u,\Omega p^v)$, and observing that

$$Xf = \Omega^{-1}\tilde{X}\tilde{f} = 0,$$

we have obtained a unique local smooth solution (r, Ω^2, Q, f) for the system (2.3.20), (2.3.21), (2.3.24), and (2.3.26) on $\mathcal{R}$ which extends the initial data.

Step 2. We first aim to prove (2.3.32) using only (2.3.26). At this point, one could apply Proposition 2.3.10 and (2.1.11) to derive (2.3.32), but we give now a direct proof.

First, using the definitions (2.3.14) and (2.3.15), we have

$$\begin{aligned}\partial_u\left(\frac{r^2\Omega^2}{\pi}N^u\right)+\partial_v\left(\frac{r^2\Omega^2}{\pi}N^v\right) &= \partial_u\int_0^\infty\int_{\mathfrak{m}^2/(\Omega^2p^v)}^\infty r^2\Omega^4p^uf\,dp^udp^v+\partial_v\int_0^\infty\int_{\mathfrak{m}^2/(\Omega^2p^u)}^\infty r^2\Omega^4p^vf\,dp^vdp^u\\ &= \int_{\Omega^2p^up^v\geq\mathfrak{m}^2}\left(2r\partial_ur\Omega^4p^uf+2r^2\Omega^4\partial_u\log\Omega^2p^uf+r^2\Omega^4p^u\partial_uf\right)dp^udp^v\\ &\quad+\int_{\Omega^2p^up^v\geq\mathfrak{m}^2}\left(2r\partial_vr\Omega^4p^vf+2r^2\Omega^4\partial_v\log\Omega^2p^vf+r^2\Omega^4p^v\partial_vf\right)dp^udp^v\\ &\quad+\int_0^\infty\frac{\mathfrak{m}^4r^2}{(p^v)^2}\partial_u\log\Omega^2f\,dp^v+\int_0^\infty\frac{\mathfrak{m}^4r^2}{(p^u)^2}\partial_v\log\Omega^2f\,dp^u. \qquad (3.3.40)\end{aligned}$$

Adding both terms involving spatial derivatives of f and using (2.3.26) yields

$$\begin{aligned}\int_{\Omega^2p^up^v\geq\mathfrak{m}^2}r^2\Omega^4\left(p^u\partial_uf+p^v\partial_vf\right)dp^udp^v &= \int_0^\infty\int_{\mathfrak{m}^2/(\Omega^2p^v)}^\infty r^2\Omega^4\Xi^u\partial_{p^u}f\,dp^udp^v\\ &\quad+\int_0^\infty\int_{\mathfrak{m}^2/(\Omega^2p^u)}^\infty r^2\Omega^4\Xi^v\partial_{p^v}f\,dp^vdp^u, \qquad (3.3.41)\end{aligned}$$

where

$$\begin{aligned}\Xi^u &\doteq \partial_u\log\Omega^2(p^u)^2+\frac{2\partial_vr}{r\Omega^2}(\Omega^2p^up^v-\mathfrak{m}^2)+\mathfrak{e}\frac{Q}{r^2}p^u,\\ \Xi^v &\doteq \partial_v\log\Omega^2(p^v)^2+\frac{2\partial_ur}{r\Omega^2}(\Omega^2p^up^v-\mathfrak{m}^2)-\mathfrak{e}\frac{Q}{r^2}p^v.\end{aligned}$$

Integrating the first term on the right-hand side of (3.3.41) by parts, we find

$$\begin{aligned}\int_0^\infty\int_{\mathfrak{m}^2/(\Omega^2p^v)}^\infty r^2\Omega^4\Xi^u\partial_{p^u}f\,dp^udp^v &= -\int_{\Omega^2p^up^v\geq\mathfrak{m}^2}\left(2r^2\Omega^4\partial_u\log\Omega^2p^u+2r\partial_vr\Omega^4p^v+\mathfrak{e}\Omega^4Q\right)f\,dp^udp^v\\ &\quad-\int_0^\infty\left(\frac{\mathfrak{m}^4r^2}{(p^v)^2}\partial_u\log\Omega^2+\mathfrak{e}\frac{\mathfrak{m}^2\Omega^2}{p^v}Q\right)f\,dp^v, \qquad (3.3.42)\end{aligned}$$

where f is evaluated at $p^u = \mathfrak{m}^2/(\Omega^2 p^v)$, and for the second term,

$$\int_0^\infty \int_{\mathfrak{m}^2/(\Omega^2 p^u)}^\infty r^2\Omega^4 \Xi^v \partial_{p^v} f\, dp^v dp^u = -\int_{\Omega^2 p^u p^v \geq \mathfrak{m}^2} \big(2r^2\Omega^4 \partial_v \log\Omega^2 p^v + 2r\partial_u r \Omega^4 p^u - \mathfrak{e}\Omega^4 Q\big)\, f\, dp^u dp^v$$
$$- \int_0^\infty \left(\frac{\mathfrak{m}^4 r^2}{(p^u)^2}\partial_v \log\Omega^2 - \mathfrak{e}\frac{\mathfrak{m}^2\Omega^2}{p^u} Q\right) f\, dp^u, \tag{3.3.43}$$

where f is evaluated at $p^v = \mathfrak{m}^2/(\Omega^2 p^u)$. Combining (3.3.40)–(3.3.43) yields (2.3.32) after noting that

$$\int_0^\infty \mathfrak{e}\frac{\mathfrak{m}^2\Omega^2}{p^v} Q f\, dp^v = \int_0^\infty \mathfrak{e}\frac{\mathfrak{m}^2\Omega^2}{p^u} Q f\, dp^u.$$

We can now derive the ingoing Maxwell equation (2.3.25). By (2.3.24), we have

$$Q(u,v) = Q(U_0,v) - \int_{U_0}^u \tfrac{1}{2}\mathfrak{e} r^2\Omega^2 N^v\, du'$$

on $\mathcal{R}$. We then derive

$$\partial_v Q(u,v) = \partial_v Q(U_0,v) - \int_{U_0}^u \partial_v(\tfrac{1}{2}\mathfrak{e} r^2\Omega^2 N^v)\, du'$$
$$= \partial_v Q(U_0,v) + \int_{U_0}^u \partial_u(\tfrac{1}{2}\mathfrak{e} r^2\Omega^2 N^u)\, du' = \tfrac{1}{2}\mathfrak{e} r^2\Omega^2 N^v(u,v),$$

where in the final equality we used the fundamental theorem of calculus and the assumption that (2.3.25) holds on $\{U_0\}\times[V_0,V_1]$.

Step 3. By a lengthy calculation which is very similar to the one performed in step 2, one may use (2.3.26) to derive the Bianchi identities in the form (2.3.33) and (2.3.34). Using the Maxwell equations (2.3.24) and (2.3.25), this implies the Bianchi identities in the form (2.1.12) and (2.1.13), where

$$\mathbf{T}^{uu} = T^{uu}, \quad \mathbf{T}^{uv} = T^{uv} + \frac{Q^2}{\Omega^2 r^4}, \quad \mathbf{T}^{vv} = T^{vv}, \quad \mathbf{S} = S + \frac{Q^2}{2r^4}.$$

By another lengthy calculation, using now also the wave equations (2.3.20) and (2.3.21), one can derive the pair of identities

$$\partial_v\left(r\partial_u^2 r - r\partial_u r\partial_u \log\Omega^2 + \frac{1}{4}r^2\Omega^4\mathbf{T}^{vv}\right) = 0,$$
$$\partial_u\left(r\partial_v^2 r - r\partial_v r\partial_v \log\Omega^2 + \frac{1}{4}r^2\Omega^4\mathbf{T}^{uu}\right) = 0.$$

These identities, together with the assumption that (2.3.22) holds on $[U_0,U_1]\times\{V_0\}$ and (2.3.23) holds on $\{U_0\} \times [V_0, V_1]$, prove that (2.3.22) and (2.3.23) hold throughout $\mathcal{R}$. $\square$

Chapter 4

Formation of black holes via characteristic gluing

In this chapter, we give a brief outline of the characteristic gluing constructions used in Chapter 5 and Chapter 6.

4.1 The problem of characteristic gluing

Characteristic gluing is a powerful new method for constructing solutions of the *Einstein field equations* by gluing together two existing solutions along a null hypersurface. The setup of characteristic gluing is depicted in Fig. 4.1 below and we will repeatedly refer to this diagram for definiteness.

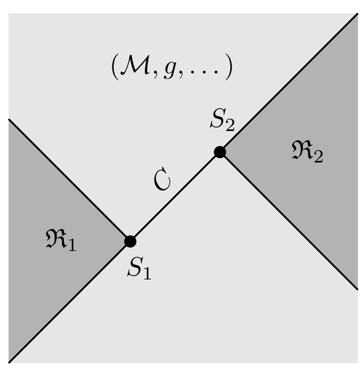

Figure 4.1: Penrose diagram depicting the setup of characteristic gluing. The null hypersurface C is declared to be "outgoing."

In Fig. 4.1, the two dark gray regions $\mathfrak{R}_1$ and $\mathfrak{R}_2$ carry Lorentzian metrics and matter fields which satisfy (1.0.1). The goal is to embed these regions into a spacetime $(\mathcal{M}, g, \dots)$ which satisfies

(1.0.1) globally, in the configuration depicted in Fig. 4.1. The characteristic gluing problem reduces to constructing characteristic data along a null hypersurface C going between spheres $S_1 \subset \mathfrak{R}_1$ and $S_2 \subset \mathfrak{R}_2$, so that after constructing the light gray regions in Fig. 4.1 by solving a characteristic initial value problem, the resulting spacetime is of the desired global regularity.

Characteristic gluing is a useful tool for constructing spacetimes that share features of two existing solutions, and therefore display interesting behavior. Moreover, the causal structure of the glued spacetime can be immediately read off from the construction, which makes it well suited to constructing examples of black hole formation.

4.2 Characteristic gluing for the linear wave equation

The study of the characteristic gluing problem was initiated by Aretakis for the linear scalar wave equation

$$\Box_g \phi = 0 \tag{4.2.1}$$

on general spacetimes $(\mathcal{M}^{3+1}, g)$ in [Are17]. Aretakis showed that there is always a finite-dimensional (but possibly trivial) space of obstructions to the characteristic gluing problem. More precisely, he showed that there are at most finitely many (possibly none) *conserved charges* that are computed from the given solutions at S_1 and S_2 in Fig. 4.1 that determine whether characteristic gluing can be performed. These charges are conserved along C for any solution of (4.2.1). This gives a definitive answer[1] to the characteristic gluing problem for (4.2.1): There is a precise characterization of which solutions can be glued—the matching of all conserved charges is both necessary and sufficient.

The gluing problem along characteristic hypersurfaces for hyperbolic equations and associated null constraints already appears for the linear wave equation on Minkowski space. On $\mathbb{R}^{3+1}$, let $u = \frac{1}{2}(t-r)$, $v = \frac{1}{2}(t+r)$, and let ϕ be a spherically symmetric solution to the wave equation, i.e.

$$\partial_u \partial_v (r\phi) = 0. \tag{4.2.2}$$

Let $C \cup \underline{C}$ be a spherically symmetric bifurcate null hypersurface, that is, $C = \{u = u_0\} \cap \{v \geq v_0\}$ and $\underline{C} = \{u_1 \geq u \geq u_0\} \cap \{v = v_0\}$. The wave equation (4.2.2) implies that $\partial_u(r\phi)$ *is conserved along the outgoing cone* C. This implies that $\partial_u \phi$ cannot be freely prescribed along C, but is in fact determined by $\partial_u \phi$ on the bifurcation sphere $C \cap \underline{C}$. Indeed, as we have seen, the characteristic initial value problem is well posed with just ϕ itself prescribed along $C \cup \underline{C}$—the full 1-jet of ϕ can

[1] [Are17] only deals with C^1 characteristic gluing, but is definitive in this regularity class.

then be recovered from (4.2.2). For general spacetimes, the question of null gluing for the linear wave equation was studied by Aretakis [Are17].

For a general wave equation, ingoing derivatives satisfy transport equations along outgoing null cones. The general C^k characteristic gluing problem is to be given two spheres S_1 and S_2 along an outgoing null cone C, and k ingoing and outgoing derivatives of ϕ at S_1 and S_2. One then seeks to prescribe ϕ along the part of C between S_1 and S_2 so that the outgoing derivatives agree with the given ones and the solutions of the transport equations for the ingoing derivatives have the specified initial and final values. In general, the linear characteristic gluing problem is obstructed due to the presence of *conserved charges* stemming from *conservation laws* along C.

Remark 4.2.1. Even in the absence of conservation laws at any order, C^∞ gluing of transverse derivatives may be obstructed in linear theory. This can be seen already for the $(1+1)$-dimensional wave equation $\partial_u\partial_v\phi = f(u,v)\phi$ for generic $f \in C^\infty(\mathbb{R}^{1+1})$. For such an f, there are no conservation laws at any order and by imposing trivial data at S_1 and very rapidly growing (in k) ∂_u^k-derivatives at S_2, one can show that C^∞ gluing cannot be achieved. Note that in $1+1$ dimensions, S_1 and S_2 are points.

4.3 Characteristic gluing for the Einstein vacuum equations near Minkowski space

The characteristic gluing problem for the Einstein equations amounts to the following:

Question 4.3.1. *Which spheres S_1 and S_2 in which vacuum spacetimes can be characteristically glued as in Fig. 4.1 as a solution of the Einstein vacuum equations? Are there any nontrivial obstructions? If so, can they be characterized geometrically?*

For example, a genuine obstruction arises from *Raychaudhuri's equation* (see already (2.3.23)), which implies that S_2 cannot be strictly outer untrapped if S_1 is (marginally) outer trapped. Another genuine obstruction arises from the rigidity of the positive mass theorem, which implies that if $\mathfrak{R}_2$ is Minkowski space, then $\mathfrak{R}_1$ is either Minkowski space or must be singular or incomplete in some sense.

Characteristic gluing for the Einstein vacuum equations (1.1.6) was initiated by Aretakis, Czimek, and Rodnianski in a fundamental series of papers [ACR21; ACR23b; ACR23a]. They study the perturbative regime around Minkowski space, that is, when both spheres S_1 and S_2 in Fig. 4.1 are close to symmetry spheres in Minkowski space. Their proof uses the inverse function theorem to reduce the nonlinear problem to a linear characteristic gluing problem for the linearized Einstein equations in double null gauge around Minkowski space, in the formalism of Dafermos–Holzegel–Rodnianski [DHR]. In the course of their argument, they discover that the linearized Einstein equations around Minkowski space in double null gauge possess *infinitely many conserved charges*. However, it turns out that all but ten of these charges are due to *gauge invariance* of the Einstein equations (cf. the *pure gauge solutions* of [DHR]). The remaining charges, which we define precisely in Definition 6.5.2 below, are genuine obstructions to the linear characteristic gluing problem. The inverse function theorem then gives nonlinear gluing close to Minkowski space, provided that, *a posteriori*, the 10 transported gauge-invariant charges at S_2 agree with the prescribed charges on S_2 (which is in fact also perturbed to deal with the gauge-dependent charges).

The conserved charges of Aretakis–Czimek–Rodnianski can be identified with the ADM energy, linear momentum, angular momentum, and center of mass. This identification is used in [ACR23a] to give a new proof of the spacelike gluing results of [Cor00; CS06; CS16] using characteristic gluing.

Later, Czimek and Rodnianski [CR22] made the fundamental observation that the linear conservation laws can be violated at the nonlinear level by certain explicit "high frequency" seed data for the characteristic initial value problem.[2] They then use these high frequency perturbations to *adjust* the linearly conserved charges in the full nonlinear theory, so that the main theorem of [ACR21] applies. The result, which we state as Theorem 6.5.3 in Section 6.5.2 below, is that two spheres close to Minkowski space can be glued if the differences of the conserved charges satisfy a certain *coercivity* condition. Roughly, the assumption is that the change in the Hawking mass be larger than the changes in the other conserved charges and that the change in angular momentum is itself much smaller than the distance of S_1 and S_2 to spheres in Minkowski space. Their result has the remarkable corollary of obstruction-free spacelike gluing of asymptotically flat Cauchy data to Kerr in the far region.

We note at this point that the analysis of [ACR21; ACR23b; ACR23a; CR22] is limited to C^2 regularity in the ingoing direction u, but allows for arbitrarily high regularity in v and angular directions.[3] It is not clear whether their analysis (especially [CR22]) can be generalized to higher order transverse derivatives.

The linearized characteristic gluing problem for (1.1.6) was redone in Bondi gauge and extended to incorporate a cosmological constant and different topologies of the cross sections of the null

[2]The high frequency perturbations are becoming singular in the Minkowski limit and hence do not linearize in a regular fashion.

[3]In this paragraph we are referring to the results in [ACR21; ACR23b; ACR23a; CR22] that have to do with characteristic gluing as is depicted in Fig. 4.1. Aretakis–Czimek–Rodnianski also consider another type of characteristic gluing, *bifurcate characteristic gluing*, which works to arbitrarily high order of differentiability.

hypersurface C by Chruściel, Cong, and Gray [CC22; CCG24]. This work also addresses linearized characteristic gluing of higher order transverse derivatives.

Question 4.3.2. *Is there a general geometric characterization of conservation laws associated to the linearized Einstein equations around a fixed background? Is there always a finite number of conservation laws? Is the generic spacetime free of conservation laws at the linear level?*

One might also wonder if there is a precise connection between the conservation laws observed in the null setting with the cokernel of the linearized constraint map studied in the spacelike gluing problem [Cor00; CS06; CD03]. We refer the reader to [CC23] for more discussion about these issues.

Remark 4.3.3. More recently, Mao, Oh, and Tao [MOT23] have developed a different approach to obstruction-free gluing in the asymptotically flat regime using an explicit solution operator for the linearized spacelike constraints. This method gives solutions to the spacelike constraints of arbitrarily high regularity.

4.4 Event horizon gluing for the Einstein–Maxwell-charged scalar field system in spherical symmetry

In the context of black hole formation, we are however interested in a different regime of gluing. We wish to glue two specific null cones: a light cone in Minkowski space and a Reissner–Nordström event horizon, as a solution of the EMCSF null constraint system. On the one hand, this is a genuine "large data" gluing problem, as these cones are very dissimilar in a gauge invariant sense and there is no known spacetime around which one could reasonably linearize the equations. On the other hand, we study our problem in spherical symmetry, which makes it considerably more tractable. We refer to Section 5.2 below for a precise definition of characteristic gluing in spherical symmetry.

We will now state the rough version of our main null gluing theorem, which concerns gluing a null cone in Minkowski space to a Reissner–Nordström event horizon.

Theorem 4.4.1 (Rough version)**.** *Let $k \in \mathbb{N}$ be a regularity index, $\mathfrak{q} \in [-1, 1]$ a charge to mass ratio, and $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$ a fixed coupling constant. For any M sufficiently large depending on k, $\mathfrak{q}$, and $\mathfrak{e}$, there exist spherically symmetric characteristic data for the Einstein–Maxwell-charged scalar field system with coupling constant $\mathfrak{e}$ gluing a Minkowski null cone of radius $\frac{1}{2}M$ to a Reissner–Nordström event horizon with mass M and charge $e = \mathfrak{q}M$ up to order k.*

We also refer to Fig. 4.2 for an illustration of our construction.

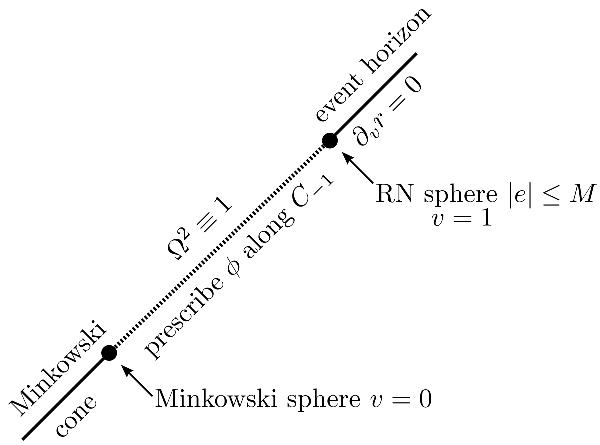

Figure 4.2: Setup of Theorem 4.4.1.

For the precise version of Theorem 4.4.1 we refer to Theorem 5.4.1 and Theorem 5.4.2 in Section 5.4. In fact, more generally, we can replace the Minkowski sphere with certain Schwarzschild exterior spheres at $v = 0$, which is important for constructing counterexamples to the third law of black hole thermodynamics (see already Section 1.1.3). Furthermore, when $\mathfrak{q} = 0$ we may take the scalar field to be real-valued, in which case the EMCSF system collapses to the Einstein-scalar field system.

Remark 4.4.2. For the proofs of Corollary 4.6.1 and Corollary 4.6.3 below, we will use versions of Theorem 4.4.1 where the top sphere is not located on a horizon. See Theorem 5.4.4 and Theorem 5.4.7 in Section 5.4 below.

Remark 4.4.3. With our methods one can also construct characteristic data which are exactly Minkowski initially and then settle down, but only asymptotically, to a Schwarzschild or (sub-)extremal Reissner–Nordström event horizon of prescribed mass and charge. The rate of decay can be chosen to be $|\partial_v \phi| \approx v^{-p}$, $p > \frac{1}{2}$, in a standard Eddington–Finkelstein gauge for Schwarzschild or subextremal Reissner–Nordström black holes. This provides examples of "global" characteristic data settling down at certain prescribed rates as assumed in [Van18b; GL19; KV21].

4.4.1 Outline of the proof of Theorem 4.4.1

In the Einstein–Maxwell-charged scalar field model in spherical symmetry, the spacetime metric is written in double null gauge as

$$g = -\Omega^2 du dv + r^2 g_{S^2},$$

where Ω^2 is the lapse and r the area-radius. We also have a complex-valued scalar field ϕ and a real-valued charge Q, which is related to the only nonzero component of the electromagnetic tensor

F. We choose an electromagnetic gauge in which $A = A_u\, du$, where A is a gauge potential for F. The dynamical variables to be glued along an outgoing cone (which we will call $C_{-1} \doteq \{u = -1\}$) are $(r, \Omega^2, \phi, Q, A_u)$. The charge Q solves first order equations in u and v, A_u is computed from Q via $F = dA$, and the variables r, Ω^2, and ϕ solve coupled nonlinear wave equations involving also Q and A_u. See already equations (2.2.3)–(2.2.10). Since the value of Ω^2 along any given null cone (or bifurcate null hypersurface) can be adjusted by reparametrizing the double null gauge, we impose that $\Omega^2 \equiv 1$.

We first consider *Raychaudhuri's equation* (see already (2.2.7)), which reads in the gauge $\Omega^2 \equiv 1$

$$\partial_v^2 r = -r|\partial_v \phi|^2. \tag{4.4.1}$$

This equation gives a nonlinear constraint on C_{-1} and completely determines r on C_{-1} given r and $\partial_v r$ at one point of C_{-1} and ϕ along C_{-1}. Thus, in the gauge $\Omega^2 \equiv 1$ along C_{-1}, up to specifying the dynamical quantities at a sphere, the free data in this problem is exactly ϕ on C_{-1}: All other dynamical quantities and their derivatives (both in the u and v coordinates) along C_{-1} can be obtained from ϕ and the equations (2.2.3)–(2.2.7).

We will choose ϕ to be compactly supported on the textured segment in Fig. 4.2 and set

$$\partial_u \phi(0) = \cdots = \partial_u^k \phi(0) = 0,$$

where k is the order at which we wish to glue. A first attempt to solve the gluing problem would be to set (r, Ω^2, Q, A_u) and derivatives to have their "Minkowski values" at the sphere $v = 0$ and then prescribe $\phi(v)$ so that the dynamical variables reach their "Reissner–Nordström values" at $v = 1$. However, specifying a "Minkowski value" for $\partial_v r$ is essentially another gauge choice, and the gauge invariance of the equations enables a much more convenient strategy.

Given that ϕ vanishes to order k at $v = 0$, to know that the sphere $v = 0$ is a sphere in Minkowski space to order k, we merely need to know that $r(0) > 0$ and that the charge Q and the Hawking mass (see already (2.1.2)) both vanish. See already Lemma 5.5.1. This reduces to the statement that in the gauge $\Omega^2 \equiv 1$,

$$\partial_u r(0) \partial_v r(0) = -\tfrac{1}{4}.$$

Since r solves a wave equation (see already (2.2.4)), $\partial_u r$ solves a first order equation in v, so it is determined on C_{-1} by $\partial_u r(0)$ alone. Given ϕ, we solve Raychaudhuri's equation (4.4.1) *backwards*, i.e., we teleologically normalize r at the final sphere by setting

$$r(1) = r_+ \doteq \left(1 + \sqrt{1-\mathfrak{q}^2}\right) M$$
$$\partial_v r(1) = 0$$

and then set

$$\partial_u r(0) \doteq \frac{-\frac{1}{4}}{\partial_v r(0)}.$$

Therefore the only "constraint" is that $\partial_v r(0) > 0$, which will be automatically satisfied by the monotonicity property of Raychaudhuri's equation as long as $r > 0$.

The charge Q is determined by *Maxwell's equation* (see already (2.2.9))

$$\partial_v Q = \mathfrak{e} r^2 \mathrm{Im}(\phi \overline{\partial_v \phi}).$$

Integrating this forwards in v yields the charge condition

$$\int_0^1 \mathfrak{e} r^2 \mathrm{Im}(\phi \overline{\partial_v \phi})\, dv = \mathfrak{q} M. \tag{4.4.2}$$

At this point we note that if $r(0) \geq \frac{1}{2}M$, then the left-hand side of this equation is $\approx M^2 \int \mathrm{Im}(\phi\overline{\partial_v\phi})$. So by modulating $\int \mathrm{Im}(\phi\overline{\partial_v\phi})$, we can hope to satisfy this equation just on the basis of scaling ϕ itself.

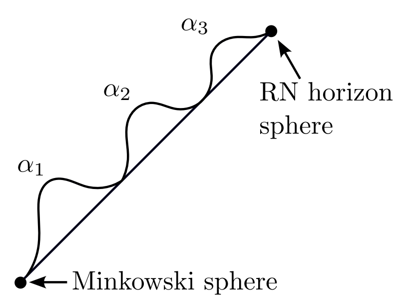

Figure 4.3: Schematic illustration of the pulses.

Our ansatz for the scalar field will be

$$\phi_\alpha = \sum_{1 \leq j \leq 2k+1} \alpha_j \phi_j,$$

where $\alpha = (\alpha_1, \ldots, \alpha_{2k+1}) \in \mathbb{R}^{2k+1}$ and the ϕ_j's are smooth compactly supported complex-valued functions with disjoint supports. We assume $\mathfrak{q} \neq 0$ now, the $\mathfrak{q} = 0$ case being in fact much easier.

The charge condition (4.4.2) is examined on every ray $\mathbb{R}_+\hat{\alpha} \in \mathbb{R}^{2k+1}$, $\hat{\alpha} \in S^{2k}$. We show that for a given choice of baseline profiles ϕ_j, there is a smooth starshaped hypersurface $\mathfrak{Q}^{2k} \subset \mathbb{R}^{2k+1}$ which is isotopic to the unit sphere S^{2k} and invariant under the antipodal map $\alpha \mapsto -\alpha$ such that (4.4.2) holds for every $\alpha \in \mathfrak{Q}^{2k}$.

The condition that M is large depending on k, $\mathfrak{q}$, and $\mathfrak{e}$ in Theorem 4.4.1 comes from natural conditions that arise when attempting to construct the hypersurface $\mathfrak{Q}^{2k}$. The charge condition (4.4.2) implies $|\mathfrak{e}|M^2|\alpha|^2 \approx |\mathfrak{q}|M$ on $\mathfrak{Q}^{2k}$. However, to keep $r \geq \frac{1}{2}M$ on C_{-1}, we find the condition $|\alpha| \lesssim 1$, see already Lemma 5.5.6. These conditions are consistent only if $|\mathfrak{e}|M \gtrsim |\mathfrak{q}|$. Furthermore, this condition is crucially used to propagate the condition $\partial_u r < 0$, see already Lemma 5.5.10.

The remaining equations ($2k$ real equations since the scalar field is complex)

$$\partial_u^i \phi_\alpha(1) = 0 \quad 1 \leq i \leq k \tag{4.4.3}$$

can naturally be viewed as *odd* equations as a function of α. So when restricted to $\alpha \in \mathfrak{Q}^{2k}$, we can use the classical Borsuk–Ulam theorem to find a simultaneous solution. Once we have an $\alpha \in \mathfrak{Q}^{2k}$ such that (4.4.3) is satisfied, ϕ_α will glue all relevant quantities to k-th order, as desired.

Theorem 4.4.4 (Borsuk–Ulam [Bor33])**.** *If $f : S^k \to \mathbb{R}^k$ is a continuous odd function, i.e., $f(-x) = -f(x)$ for every $x \in S^k$, then f has a root.*

For a nice proof using only basic degree theory and transversality arguments, see Nirenberg's lecture notes [Nir01].

4.5 Event horizon gluing in vacuum: the very slowly rotating case

The spherically symmetric gluing outlined in Section 4.4.1 can be summarized as follows:

(i) The Hawking mass is glued by judiciously initiating the transport equations at S_1 *or* S_2 and directly exploiting gauge freedom in the form of boosting the double null gauge by hand.

(ii) The charge of the Maxwell field is glued by exploiting a monotonicity property of Maxwell's equation specific to spherical symmetry.

(iii) Transverse derivatives of the scalar field are glued by exploiting a parity symmetry of the Einstein–Maxwell-charged scalar field system specific to spherical symmetry and invoking the

Borsuk–Ulam theorem.

The argument for vacuum gluing has two crucial ingredients: the implementation of idea (i) above in the context of the Einstein vacuum equations, and the obstruction-free characteristic gluing of Czimek–Rodnianski [CR22], which replaces the Borsuk–Ulam argument in vacuum. See already Section 4.5.1 for the outline of our proof.

Our first theorem shows that the characteristic gluing of Minkowski space to (positive mass) Schwarzschild solutions is completely unobstructed, provided that we aim to glue to a symmetry sphere in Schwarzschild. In the statements of our theorems, refer to Fig. 4.1.

Theorem 4.5.1. *Let $M > 0$. Let S_2 be any non-antitrapped symmetry sphere in the Schwarzschild solution of mass M. Then S_2 can be characteristically glued to a sphere S_1 as depicted in Fig. 4.1, to order C^2 as a solution of the Einstein vacuum equations* (1.1.6)*, where S_1 is a spacelike sphere in Minkowski space which is arbitrarily close to a round symmetry sphere.*

For the precise statement of this theorem, see already Theorem 6.6.11 in Section 6.6.2 below. Our method also immediately generalizes to very slowly rotating Kerr, and gives the following particularly clean statement about event horizon gluing:

Theorem 4.5.2. *There exists a constant $0 < \mathfrak{a}_0 \ll 1$ such that if S_2 is a spacelike section of the event horizon of a Kerr black hole with mass $M > 0$ and specific angular momentum a satisfying $0 \le |a|/M \le \mathfrak{a}_0$, then S_2 can be characteristically glued to a sphere S_1 as depicted in Fig. 4.1, to order C^2 as a solution of the Einstein vacuum equations* (1.1.6)*, where S_1 is a spacelike sphere in Minkowski space which is close to a round symmetry sphere.*

This theorem is a special case of Theorem 6.6.13 in Section 6.6.3 below.

Remark 4.5.3. There is an apparent asymmetry in the statements of Theorem 4.5.1 and Theorem 4.5.2 about the allowable S_2's. In fact, in Theorem 6.6.13 below, we show that *any* Kerr coordinate sphere can be connected to a sphere in Schwarzschild with smaller mass, but the maximum value of allowed angular momentum depends on the sphere in a non-explicit way that we prefer to explain later, see already Section 6.4.

Remark 4.5.4. In Theorem 4.5.1, the bottom sphere S_1 can be made arbitrarily close to an exact symmetry sphere in Minkowski, whereas in Theorem 4.5.2, the closeness to an exact symmetry sphere is limited by the size of a/M.

Remark 4.5.5. The C^2 regularity of the spacetime metric in Theorem 4.5.1 and Theorem 4.5.2 is due to limited regularity in the direction transverse to C. The regularity of the metric in directions tangent to C can be made arbitrarily high (but finite).

Theorem 4.5.1 above can be viewed as a generalization to vacuum of the first step in the proof of Theorem 1.1.1. We conjecture that the second step, making the black hole extremal, can also be generalized to vacuum:

Conjecture 4.5.6. *The Schwarzschild symmetry sphere of mass M_i and radius R_i can be characteristically glued to any non-antitrapped Kerr coordinate sphere with radius $R_f \gg R_i$ in a Kerr solution with mass $M_f \gg M_i$ and specific angular momentum $0 \le |a_f| \le M_f$.*

If this conjecture holds, Schwarzschild can be spun up to extremality, which would likely lead to a proof of Conjecture 1.1.12.

Remark 4.5.7. By using negatively charged pulses in [KU22], we can design characteristic data that also "discharges" the black hole. It would be very interesting to find a mechanism that can both "spin up" and "spin down" a Kerr black hole, or move the rotation axis without changing the angular momentum much.

4.5.1 Outline of the proof

In this section we give a very brief outline of the proof of Theorem 4.5.1. The gluing is performed in two stages and should be thought of as being performed backwards in time.

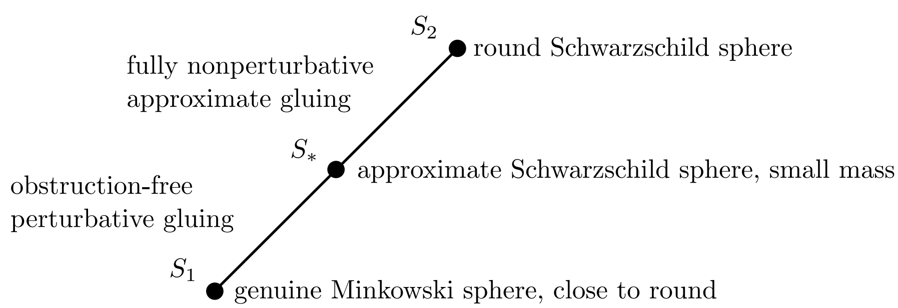

Figure 4.4: Two-step process for the proof of Theorem 4.5.1.

First, a fully nonperturbative mechanism is used to connect the exact Schwarzschild sphere S_2 of mass M and radius R to a sphere data set S_* which is very close to a Schwarzschild sphere of mass $0 < M_* \ll R$ and radius $\approx R$. See already Proposition 6.6.1 below. In the second stage of the gluing, we use the main theorem of [CR22], which we state below as Theorem 6.5.3, as a black box. In order to satisfy the necessary coercivity conditions required for obstruction-free characteristic gluing, we choose $0 < \varepsilon_\sharp \ll M_* \ll R$, where $\varepsilon_\sharp$ measures the closeness of S_* to the (M_*, R)-Schwarzschild sphere.

The nonperturbative mechanism which glues S_* to S_2 involves the injection of two pulses[4] of gravitational waves (described mathematically by the shear $\hat{\chi}$) along $C = [0,1]_v \times S^2$, of amplitude $\delta^{1/2} = O(\varepsilon_\sharp)$, together with a choice of outgoing null expansion $\operatorname{tr}\chi$ at S_2 such that $|\operatorname{tr}\chi| \lesssim \delta$. In order to fix the Hawking mass of S_2 to be M, the ingoing null expansion $\operatorname{tr}\underline{\chi}$ is then chosen to be $\approx -\delta^{-1}$ at S_2.[5] The Hawking mass of S_* is fixed to be arbitrarily close to M_* by tuning $\hat{\chi}$ and using the monotonicity of Raychaudhuri's equation; see already Lemma 6.6.8.[6] We then step through the null structure equations and Bianchi identities as in [Chr09, Chapter 2] and establish a δ-weighted hierarchy for the sphere data at S_*; see already Lemma 6.6.10. Finally, we boost the cone by δ, which brings S_2 to a reference Schwarzschild sphere and S_* within $\varepsilon_\sharp$ of a reference Schwarzschild sphere with mass M_*. This construction may be thought of as a direct adaptation, in vacuum, of the idea used to prove Schwarzschild event horizon gluing in spherical symmetry for the Einstein-scalar field system in Section 5.5.1.

4.5.2 Relation to Christodoulou's short pulse method

After the boost, one can interpret the above approximate nonperturbative gluing mechanism as a "short pulse" data set defined on $[0,\delta] \times S^2$ as in [Chr09], but fired backwards. In this context, our equation (6.6.9) below should be compared with the condition (4.1) in [LY15] (see also [LM20; AL22b]). This condition guarantees that certain components of the sphere data at S_* are *a posteriori* closer to spherical symmetry.

4.6 Further applications of characteristic gluing

4.6.1 Gravitational collapse with a piece of smooth Cauchy horizon

Another corollary of our method is the construction of regular one-ended Cauchy data which evolve to a subextremal or extremal black hole for which there exists a piece of Cauchy horizon emanating from i^+. We refer to Fig. 4.5 for the Penrose diagram of the spacetime constructed in Corollary 4.6.1. The proof of Corollary 4.6.1 is given in Section 5.6.3.

[4] By the Poincaré–Hopf theorem, the shear $\hat{\chi}$ vanishes somewhere on each sphere. This means that in any characteristic gluing problem in vacuum where the null expansion $\operatorname{tr}\chi$ has to change everywhere, the zero of $\hat{\chi}$ has to move along the sphere as v increases. In our setting, we choose the first pulse to be supported away from the north pole and the second pulse to be supported away from the south pole.

[5] $\operatorname{tr}\underline{\chi} < 0$ on S_2 means that S_2 is not antitrapped.

[6] In [KU22], we use a similar monotonicity to glue the charge.

Corollary 4.6.1 (Gravitational collapse to RN with a smooth Cauchy horizon)**.** *For any regularity index $k \in \mathbb{N}$ and nonzero charge to mass ratio $\mathfrak{q} \in [-1,1] \setminus \{0\}$, there exist spherically symmetric, asymptotically flat Cauchy data for the Einstein–Maxwell-charged scalar field system in spherical symmetry, with $\Sigma \cong \mathbb{R}^3$ and a regular center, such that the maximal future globally hyperbolic development $(\mathcal{M}^4, g)$ has the following properties:*

- *The spacetime satisfies all the conclusions of Theorem 1.1.4 with $\mathfrak{q} \neq 0$, including C^k-regularity of all dynamical quantities.*
- *The black hole region contains an isometrically embedded portion of a Reissner–Nordström Cauchy horizon neighborhood with charge to mass ratio $\mathfrak{q}$.*

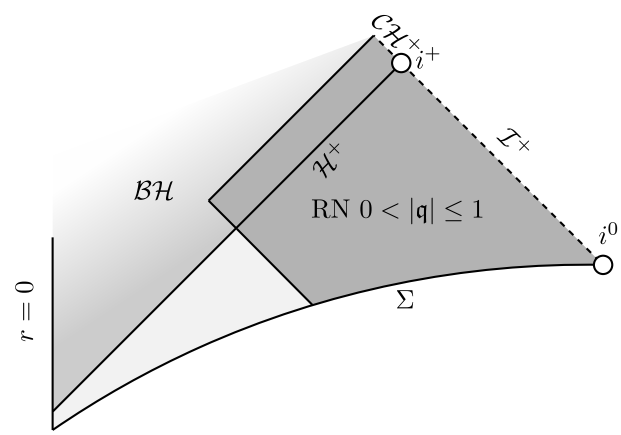

Figure 4.5: Penrose diagram depicting Corollary 4.6.1: Gravitational collapse to Reissner–Nordström with nonempty piece of Cauchy horizon $\mathcal{CH}^+$.

Remark 4.6.2. When $|\mathfrak{q}| = 1$, the spacetime constructed in Corollary 4.6.1 does not contain trapped symmetry spheres in the dark shaded region in Fig. 4.5. By a slight modification of the argument in Proposition 2.5.1 below, this implies no trapped surfaces *intersect* the dark shaded region. In particular, the trapped region (in the sense of [HE73, p. 319]) of the spacetime (if nonempty) avoids a whole double null neighborhood of the event horizon. Nevertheless, the event horizon agrees with the outermost apparent horizon for late advanced times.

4.6.2 Black hole interiors for which the Cauchy horizon closes off spacetime

Our horizon gluing method can also be extended to glue Reissner–Nordström interior spheres to a regular center along an ingoing cone, see already Theorem 5.4.7. Using this, we construct asymptotically flat Cauchy data for which the future boundary of the black hole region $\mathcal{BH}$ is a Cauchy horizon $\mathcal{CH}^+$ which closes off spacetime. We refer to Fig. 4.6 for the Penrose diagram of the spacetime constructed in Corollary 4.6.3. The proof of Corollary 4.6.3 is given in Section 5.6.4.

Corollary 4.6.3 (Cauchy horizon that closes off the spacetime)**.** *For any regularity index $k \in \mathbb{N}$, and nonzero charge to mass ratio $\mathfrak{q} \in [-1, 1] \setminus \{0\}$, there exist spherically symmetric, asymptotically flat Cauchy data for the Einstein–Maxwell-charged scalar field system, with $\Sigma \cong \mathbb{R}^3$ and a regular center, such that the maximal future globally hyperbolic development $(\mathcal{M}^4, g)$ has the following properties:*

- *All dynamical quantities are at least C^k-regular.*
- *The black hole region is non-empty, $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+) \neq \emptyset$.*
- *The future boundary of $\mathcal{BH}$ is a C^k-regular Cauchy horizon $\mathcal{CH}^+$ which closes off spacetime.*
- *The black hole exterior is isometric to a Reissner–Nordström exterior with charge to mass ratio $\mathfrak{q}$. In particular, null infinity $\mathcal{I}^+$ is complete.*
- *The spacetime does not contain antitrapped surfaces.*
- *When $|\mathfrak{q}| = 1$, the spacetime does not contain trapped surfaces.*

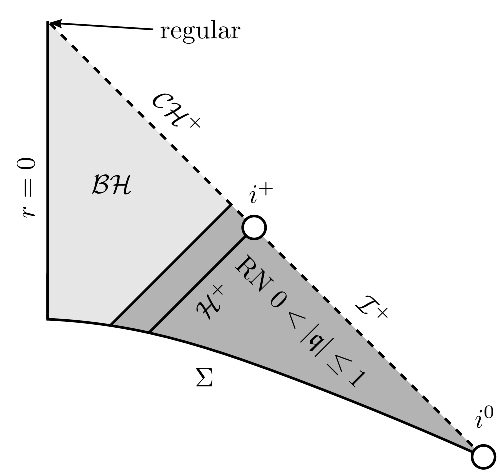

Figure 4.6: Penrose diagram depicting Corollary 4.6.3: The Cauchy horizon is regular and closes off the spacetime in a regular fashion.

Remark 4.6.4. In contrast to our previous constructions, the Cauchy surface Σ in Corollary 4.6.3 could contain trapped surfaces and Σ intersects the black hole region. It would be interesting to construct a spacetime as in Corollary 4.6.3 which depicts genuine gravitational collapse, i.e., for which $\Sigma \subset J^-(\mathcal{I}^+)$.

In the subextremal case, the behavior exhibited by our construction can be seen as exceptional as one generically expects a Cauchy horizon which forms in gravitational collapse to be a weak null singularity [Daf03; Van18b; LO19]. In particular, in the case where the Cauchy horizon $\mathcal{CH}^+$ is weakly singular, Van de Moortel [Van23] showed that the Cauchy horizon $\mathcal{CH}^+$ cannot close off spacetime in the sense of Fig. 4.6. Thus, our construction in Corollary 4.6.3 makes [Van23] sharp in the sense that the singularity assumption of $\mathcal{CH}^+$ in [Van23] is needed. Restricted to the extremal case, however, on the basis of a more regular Cauchy horizon as in [GL19], one may speculate that there exists a set of data (open as a subset of the positive codimension set of data settling down to ERN) for which the Cauchy horizon closes off spacetime as depicted in Fig. 4.6.

4.6.3 Vacuum gravitational collapse with a spacelike singularity

By performing characteristic gluing as in Theorem 4.5.1 of Minkowski space to spheres in a Schwarzschild solution lying just inside the horizon and using Cauchy stability, we also obtain:

Corollary 4.6.5 (Gravitational collapse with a spacelike singularity)**.** *There exist one-ended asymptotically flat Cauchy data* $(g_0, k_0) \in H^{7/2-}_{\mathrm{loc}} \times H^{5/2-}_{\mathrm{loc}}$ *for the Einstein vacuum equations* (1.1.6) *on* $\Sigma \cong \mathbb{R}^3$*, satisfying the constraint equations, such that the maximal future globally hyperbolic development* $(\mathcal{M}^4, g)$ *contains a black hole* $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+)$ *and has the following properties:*

- *The Cauchy surface* Σ *lies in the causal past of future null infinity,* $\Sigma \subset J^-(\mathcal{I}^+)$*. In particular,* Σ *does not intersect the event horizon* $\mathcal{H}^+ \doteq \partial(\mathcal{BH})$ *or contain trapped surfaces.*

- *For sufficiently late advanced times* $v \geq v_0$*, the domain of outer communication, together with a full double null slab lying in the interior of the black hole, is isometric to a portion of a Schwarzschild solution as depicted in Fig. 4.7. The double null slab terminates in the future at a spacelike singularity, isometric to the "*$r = 0$*" singularity in Schwarzschild.*

We will sketch the proof of this result in Section 6.7 below.

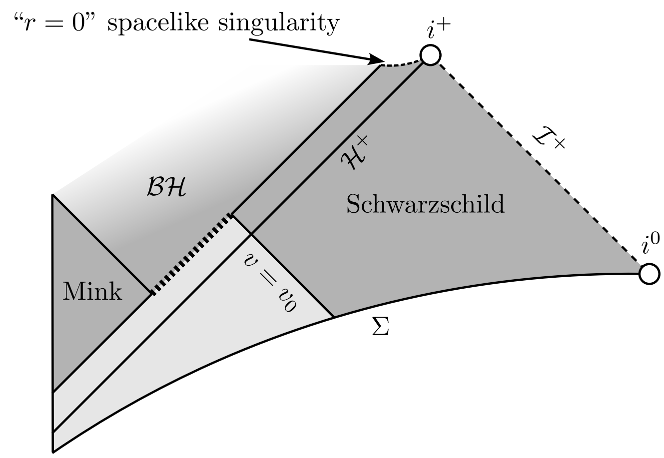

Figure 4.7: Example of gravitational collapse with a piece of a spacelike singularity emanating from timelike infinity i^+. Characteristic gluing is performed along the textured line segment, where the top gluing sphere has radius very close to the Schwarzschild radius of the black hole to be formed.

Chapter 5

Characteristic gluing for the charged scalar field model and the third law of black hole thermodynamics

In this chapter, we establish the characteristic gluing method for the spherically symmetric Einstein–Maxwell-charged scalar field model and use this to construct examples of gravitational collapse to Reissner–Nordström black holes with specified parameters.

5.1 Sphere data and cone data

5.1.1 Determining transversal derivatives from tangential data

We use here the notation of Section 3.1. According to Proposition 3.1.3, characteristic data on the bifurcate null hypersurface $C \cup \underline{C}$ extends (locally) to a solution of the system (2.2.3)–(2.2.7). We can then use the equations to compute all the partial derivatives of the solution along $C \cup \underline{C}$. We now describe a procedure for determining all u-derivatives on C just in terms of r, Ω^2, ϕ, Q, and A_u (as functions of v) and their u derivatives at the bifurcation sphere.

Proposition 5.1.1. *Let* $(r, \Omega^2, \phi, Q, A_u)$ *be a* C^k *<u>bifurcate</u> characteristic initial data set as in Definition 3.1.2. Then the EMCSF system can be used to determine as many u-derivatives of* $r, \Omega^2, \phi, Q,$

and A_u on C as is consistent with Definition 3.1.1, explicitly from the data on $C \cup \underline{C}$.

Proof. Since $(r, \Omega^2, \phi, Q, A_u)$ are all given on $\underline{C}$, we can compute as many u-derivatives of these quantities at the bifurcation sphere $(-1, 0)$ as the regularity k allows. We describe an inductive procedure for computing u-derivatives of $(r, \Omega^2, \phi, Q, A_u)$ on C, starting with $\partial_u r$. Since $\partial_u r(-1, 0)$ is known, and the wave equation (2.2.4) can be written as

$$\left(\partial_v + \frac{\partial_v r}{r}\right)\partial_u r = -\frac{\Omega^2}{4r} + \frac{\Omega^2}{4r^3}Q^2,$$

where everything on the right-hand side is already known, $\partial_u r(-1, v)$ can be found by solving this ODE. In the same manner, $\partial_u \phi(-1, v)$ and then $\partial_u \log(\Omega^2)(-1, v)$ can be found. To find $\partial_u Q(-1, v)$, differentiate (2.2.8) in v and then integrate. (Alternatively, differentiate (2.2.9) in u.) Finally, $\partial_u A_u(-1, v)$ is found by differentiating (2.2.10) in v and then integrating.

Proceeding in this way, by commuting all the equations with ∂_u^i, every partial derivative of $(r, \Omega^2, \phi, Q, A_u)$ which is consistent with the initial C^k regularity can be found. We finally note that $\partial_u^{k+1} r(-1, v)$ is found from differentiating (2.2.6) an appropriate number of times, since the wave equation it satisfies is not consistent with the level of regularity of the rest of the dynamical variables. □

Remark 5.1.2. Both Proposition 3.1.3 and Proposition 5.1.1 exploit the *null condition* satisfied by the EMCSF system in double null gauge. For a general nonlinear wave equation, the solution may not exist in a full double null neighborhood of the initial bifurcate null hypersurface as in Proposition 3.1.3. Indeed, the null condition means the transport equations for transversal derivatives in Proposition 5.1.1 are linear and hence do not blow up in finite time.

5.1.2 Sphere data

In order to define a notion of characteristic gluing later, we introduce a notion of *sphere data* inspired by [ACR21; ACR23b]. Given a C^k solution of the EMCSF system in spherical symmetry, for every $(u_0, v_0) \in \mathcal{Q}$ one can extract a list of numbers corresponding to $r(u_0, v_0)$, $\Omega^2(u_0, v_0)$, $\phi(u_0, v_0)$, $Q(u_0, v_0)$, $\partial_u r(u_0, v_0)$ etc. Our definition of sphere data formalizes this (long) list of numbers and incorporates the constraints (2.2.8)–(2.2.7), so we may refer to the data induced by a C^k solution on a sphere without reference to an actual solution of the equations themselves.

Definition 5.1.3. Let $k \geq 1$. A *sphere data set* with *regularity index* k for the Einstein–Maxwell-charged scalar field in the EM gauge (2.2.2) is the following list of numbers[1]:

1. $\varrho > 0$, $\varrho_u^1, \dots, \varrho_u^{k+1}, \varrho_v^1, \dots, \varrho_v^{k+1} \in \mathbb{R}$
2. $\omega > 0$, $\omega_u^1, \dots, \omega_u^k, \omega_v^1, \dots, \omega_v^k \in \mathbb{R}$
3. $\varphi, \varphi_u^1, \dots, \varphi_u^k, \varphi_v^1, \dots, \varphi_v^k \in \mathbb{C}$
4. $q, q_u^1, \dots, q_u^k, q_v^1, \dots, q_v^k \in \mathbb{R}$
5. $a, a_u^1, \dots, a_u^k, a_v^1, \dots, a_v^k, a_v^{k+1} \in \mathbb{R}$

subject to the following conditions:

(i) ϱ_u^{i+2} can be expressed as a rational function of ϱ_u^{j+1}, ω_u^{j+1}, φ_u^{j+1}, and a_u^j for $0 \le j \le i$ by formally differentiating (2.2.6),

(ii) ϱ_v^{i+2} can be expressed as a rational function of ϱ_v^{j+1}, ω_v^{j+1}, and φ_v^{j+1} for $0 \le j \le i$ by formally differentiating (2.2.7),

(iii) q_u^{i+1} can be expressed as a polynomial of ϱ_u^j, φ_u^j, and a_u^j for $0 \le j \le i$ by formally differentiating (2.2.8),

(iv) q_v^{i+1} can be expressed as a polynomial of ϱ_u^j, and φ_u^j for $0 \le j \le i$ by formally differentiating (2.2.9), and

(v) a_v^{i+1} can be expressed as a rational function of ϱ_v^j, ω_v^j, and q_v^j for $0 \le j \le i$ by formally differentiating (2.2.10),

where we have adopted the convention that $\varrho_u^0 = \varrho$, etc. We denote by $\mathcal{D}_k$ the set of such sphere data sets with regularity index k.

Gauge freedom is a very important aspect of the study of the EMCSF system. Our next definition records the gauge freedom present in sphere data. We need to consider both double null gauge transformations

$$u = f(U), \quad v = g(V),$$

where f and g are increasing functions on $\mathbb{R}$ and EM gauge transformations (2.2.1)

$$\phi \mapsto e^{-i\mathfrak{e}\chi}\phi, \quad A \mapsto A + d\chi,$$

where χ is a function of u alone, i.e. $\partial_v \chi = 0$, in order to satisfy (2.2.2).

[1]One should think that formally $r(u_0, v_0) = \varrho$, $\phi(u_0, v_0) = \varphi$, $\Omega^2(u_0, v_0) = \omega$, $\partial_v^i r(u_0, v_0) = \varrho_v^i$, etc.

Definition 5.1.4. We define the *full gauge group* of the Einstein–Maxwell-charged scalar field system in spherically symmetric double null gauge with the EM gauge condition (2.2.2) as

$$\mathcal{G} \doteq \{(f,g) : f,g \in \mathrm{Diff}_+(\mathbb{R}), f(0)=g(0)=0\} \times C^\infty(\mathbb{R}),$$

with the group multiplication given by[2]

$$((f_2,g_2),\chi_2)\cdot((f_1,g_1),\chi_1) = ((f_2\circ f_1, g_2\circ g_1), \chi_2\circ f_1^{-1}+\chi_1).$$

The gauge group defines an action on sphere data as follows. Given sphere data $D \in \mathcal{D}_k$, assign functions $r(u,v)$, $\Omega^2(u,v)$, $\phi(u,v)$, $Q(u,v)$, and $A_u(u,v)$ whose jets agree with the sphere data D. For $\tau = ((f,g),\chi) \in \mathcal{G}$, let

$$\tilde{r}(u,v) = r(f(u),g(v)) \tag{5.1.1}$$

$$\tilde{\Omega}^2(u,v) = f'(u)g'(v)\Omega^2(f(u),g(v)) \tag{5.1.2}$$

$$\tilde{\phi}(u,v) = e^{-i\mathfrak{e}\chi(f(u))}\phi(f(u),g(v)) \tag{5.1.3}$$

$$\tilde{Q}(u,v) = Q(f(u),g(v)) \tag{5.1.4}$$

$$\tilde{A}_u(u,v) = f'(u)A_u(f(u),g(v)) + f'(u)\chi'(f(u)). \tag{5.1.5}$$

The components of τD are then defined by formally differentiating equations (5.1.1)–(5.1.5) and evaluating at $u=v=0$. For example, $\tau(\varrho)=\varrho$, $\tau(\varrho_v^1) = g'(0)\varrho_v^1$, and $\tau(\varphi_u^1) = (1-i\mathfrak{e}\chi'(0))e^{-i\mathfrak{e}\chi(0)}\varphi$.

If one is given a bifurcate characteristic initial data set (r,Ω^2,ϕ,Q,A_u), the lapse Ω^2 can be set to unity on $C\cup\underline{C}$ by reparametrizing u and v. In the sphere data setting, we have an analogous notion:

Definition 5.1.5. A sphere data set $D\in\mathcal{D}_k$ is said to be *lapse normalized* if $\omega=1$ and $\omega_u^i=\omega_v^i=0$ for $1\le i\le k$. Every sphere data set is gauge equivalent to a lapse normalized sphere data set.

5.1.3 Cone data and seed data

In the previous subsection, we saw how a C^k solution (r,Ω^2,ϕ,Q,A_u) on $\mathcal{Q}$ gives rise to a continuous map $\mathcal{Q}\to\mathcal{D}_k$. For the purpose of characteristic gluing, it is convenient to consider one-parameter families of sphere data which are to be thought of as being induced by constant u cones in $\mathcal{Q}$.

[2]One can view this as a left semidirect product.

More precisely, if we consider a null cone $C \subset \mathcal{Q}$, parametrized by $v \in [v_1, v_2]$, then a solution of the EMCSF system induces a continuous map $D : [v_1, v_2] \to \mathcal{D}_k$ by sending each v to its associated sphere data $D(v)$. In fact, this map can be produced by knowing only $D(v_1)$ and the values of $(r, \Omega^2, \phi, Q, A_u)$ on C. Arguing as in Proposition 5.1.1 with $D(v_1)$ taking the role of the bifurcation sphere gives:

Proposition 5.1.6. *Let $k \in \mathbb{N}$, $v_1 < v_2 \in \mathbb{R}$, $r, A_u \in C^{k+1}([v_1, v_2])$, and $\Omega^2, \phi, Q \in C^k([v_1, v_2])$ which satisfy the constraints* (2.2.9), (2.2.10), *and* (2.2.7) *on $[v_1, v_2]$. Let $D_1 \in \mathcal{D}_k$ such that all v-components of D_1 agree with the corresponding v-derivatives of $(r, \Omega^2, \phi, Q, A_u)$ at v_1. Then there exists a unique continuous function $D : [v_1, v_2] \to \mathcal{D}_k$ such that $D(v_1) = D_1$ and upon identification of the formal symbols $\varrho(D(v))$, $\varrho_u^1(D(v))$, etc., with the dynamical variables $(r, \Omega^2, \phi, Q, A_u)$ and their u- and v-derivatives, satisfies the EMCSF system and agrees with $(r, \Omega^2, \phi, Q, A_u)$ in the v-components for every $v \in [v_1, v_2]$.*

Definition 5.1.7. Let $k \in \mathbb{N}$ and $v_1 < v_2 \in \mathbb{R}$. A C^k *cone data set* for the Einstein–Maxwell-charged scalar field in spherical symmetry is a continuous function $D : [v_1, v_2] \to \mathcal{D}_k$ satisfying the conclusion of Proposition 5.1.6, i.e., formally satisfying the EMCSF system.

We now discuss a procedure for generating solutions of the "tangential" constraint equations, (2.2.9), (2.2.10), and (2.2.7), which were required to be satisfied in the previous proposition.

Proposition 5.1.8 (Seed data)**.** *Let $k \in \mathbb{N}$, $v_1 < v_2 \in \mathbb{R}$, and $D_1 \in \mathcal{D}_k$ be lapse normalized. For any $\phi \in C^k([v_1, v_2])$ such that $\partial_v^i \phi(v_1) = \varphi_v^i(D_1)$ for $0 \le i \le k$, there exist unique functions $r, A_u \in C^{k+1}([v_1, v_2])$ and $Q \in C^k([v_1, v_2])$ such that $(r, \Omega^2, \phi, Q, A_u)$ satisfies the hypotheses of Proposition 5.1.6 with $\Omega^2(v) = 1$ for every $v \in [v_1, v_2]$.*

Proof. When $\Omega^2 \equiv 1$, Raychaudhuri's equation (2.2.7) reduces to

$$\partial_v^2 r = -r|\partial_v \phi|^2,$$

which is a second order ODE for $r(v)$. Setting $r(v_1) = \varrho(D_1)$ and $\partial_v r(v_1) = \varrho_v^1(D_1)$, we obtain a unique solution $r \in C^{k+1}([v_1, v_2])$. The charge is obtained by integrating Maxwell's equation (2.2.9):

$$Q(v) = q(D_1) + \int_0^v \mathfrak{e} r^2(v') \mathrm{Im}(\phi(v') \overline{\partial_v \phi(v')})\, dv'.$$

Finally, the gauge potential is obtained by integrating (2.2.10):

$$A_u(v) = a(D_1) - \int_0^v \frac{Q(v')}{2r^2(v')}\,dv'.$$

The v-derivatives of $(r, \Omega^2, \phi, Q, A_u)$ agree with the v-components of D_1 by virtue of the definitions.

□

5.2 Characteristic gluing in spherical symmetry

In this section we give precise statements of our main theorems. In order to do this, we carefully define the notion of *characteristic gluing*.

Definition 5.2.1 (Characteristic gluing)**.** Let $k \in \mathbb{N}$. Let $D_1, D_2 \in \mathcal{D}_k$ be sphere data sets. We say that D_1 can be *characteristically glued to D_2 to order k* in the Einstein–Maxwell-charged scalar field system in spherical symmetry if there exist $v_1 < v_2$ and a C^k cone data set $D : [v_1, v_2] \to \mathcal{D}_k$ such that $D(v_1)$ is gauge equivalent to D_1 and $D(v_2)$ is gauge equivalent to D_2.

Remark 5.2.2. It is clear that if D_1 and D_2 can be characteristically glued and $\tau_1, \tau_2 \in \mathcal{G}$, then $\tau_1 D_1$ and $\tau_2 D_2$ can be characteristically glued.

Remark 5.2.3. Definition 5.2.1 on characteristic gluing along an outgoing cone has a natural analog defining characteristic gluing along an ingoing cone by parametrizing the cone data with u and letting v denote the transverse null coordinate, but keeping the definition of sphere data unchanged.

By Proposition 5.1.8, characteristic gluing is equivalent to choosing an appropriate seed ϕ in the following sense. By applying a gauge transformation to D_1, we may assume it to be lapse normalized. Then cone data sets with $\Omega^2 \equiv 1$ agreeing with D_1 at v_1 are parametrized precisely by functions $\phi \in C^k([v_1, v_2]; \mathbb{C})$ with the correct v-jet at v_1. Therefore, characteristic gluing reduces to finding ϕ so that the final data set $D(v_2)$ produced by Proposition 5.1.6 is gauge equivalent to D_2.

5.2.1 Spacetime gluing from characteristic gluing

If the two sphere data sets in Definition 5.2.1 come from spheres in two spherically symmetric EMCSF spacetimes, we can use local well posedness for the EMCSF characteristic initial value problem, Proposition 3.1.3, to glue parts of the spacetimes themselves. This principle underlies all of our constructions in Section 5.6.

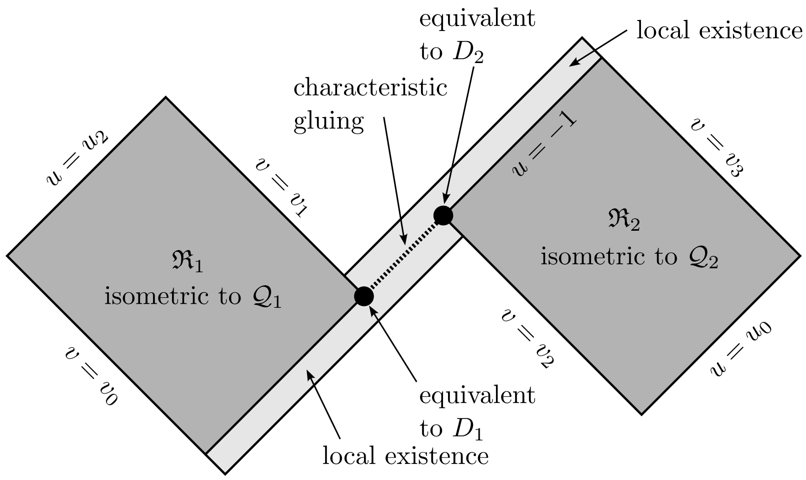

Figure 5.1: Spacetime gluing obtained from characteristic gluing. The two spacetimes (dark gray) are glued along the cone $u = -1$. Note that the dark gray regions are causally disconnected except for the cone $u = -1$. Such a spacetime exists if and only if D_1 and D_2 can be characteristically glued.

Proposition 5.2.4. *Let $(\mathcal{Q}_1, r_1, \Omega_1^2, \phi_1, Q_1, A_{u\,1})$ and $(\mathcal{Q}_2, r_2, \Omega_2^2, \phi_2, Q_2, A_{u\,2})$ be two C^k solutions of the EMCSF system in spherical symmetry, where each $\mathcal{Q}_i$ is a double null rectangle, i.e.,*

$$\mathcal{Q}_1 = [u_{0,1}, u_{1,1}] \times [v_{0,1}, v_{1,1}]$$

$$\mathcal{Q}_2 = [u_{0,2}, u_{1,2}] \times [v_{0,2}, v_{1,2}].$$

Let D_1 be the sphere data induced by the first solution on $(u_{0,1}, v_{1,1})$ and D_2 be the sphere data induced by the second solution on $(u_{1,2}, v_{0,2})$. If D_1 can be characteristically glued to D_2 to order k, then there exists a spherically symmetric C^k solution $(\mathcal{Q}, r, \Omega^2, \phi, Q, A_u)$ of the EMCSF system with the following property: There exists a global double null gauge (u, v) on $\mathcal{Q}$ containing double null rectangles

$$\mathfrak{R}_1 = [-1, u_2] \times [v_0, v_1],$$

$$\mathfrak{R}_2 = [u_0, -1] \times [v_2, v_3],$$

such that the restricted solutions $(\mathfrak{R}_i, r, \Omega^2, \phi, Q, A_u)$ are isometric and gauge equivalent to the solutions $(\mathcal{Q}_i, r_i, \Omega_i^2, \phi_i, Q_i, A_{u\,i})$ for $i = 1, 2$, the sphere data induced on $(-1, v_1)$ is equal to D_1 to k-th order, and the sphere data induced on $(-1, v_2)$ is gauge equivalent to D_2 to k-th order.

Proof. In this proof, we will refer to spherically symmetric solutions of the EMCSF system by their domains alone.

By Definition 5.2.1, since D_1 and D_2 can be characteristically glued, we obtain $v_1 < v_2$, functions r, Ω^2, ϕ, Q, and A_u on $[v_1, v_2]$, and a gauge transformation $\tau \in \mathcal{D}_k$ which acts on D_2. We now build the spacetime out of two pieces which will then be pasted along $u = -1$ and match to order C^k. See Fig. 5.2.

First, we prepare the given spacetimes. We relabel the double null gauge on $\mathcal{Q}_1$ by changing the southeast edge to be $u = -1$ and the northeast edge to be $v = v_1$. This also determines u_2 and v_0 and we apply no further gauge transformation to $\mathcal{Q}_1$. We denote this region by $\mathfrak{R}_1$.

Next, the gauge transformation τ is extended and applied to $\mathcal{Q}_2$. We relabel the double null gauge to have $u = -1$ on the northwest edge and $v = v_2$ on the southwest edge. We denote this region by $\mathfrak{R}_2$.

We now construct the left half of Fig. 5.2 as follows. Extend the cone $u = -1$ in $\mathfrak{R}_1$ until $v = v_3$, and extend the functions $(r, \Omega^2, \phi, Q, A_u)$ on $u = -1$ by taking them from the definition of characteristic gluing for $v \in [v_1, v_2]$, and then from the induced data on $u = -1$ in $\mathfrak{R}_2$ for $v \in [v_2, v_3]$. We now appeal to local existence, Proposition 3.1.3, the EMCSF system in spherical symmetry to construct the solution in a thin slab $\mathcal{S}_1$ to the future of

$$(\{u = -1\} \times [v_1, v_3]) \cup ([-1, u_2] \times \{v = v_1\}).$$

This completes the construction of $\mathfrak{R}_1 \cup \mathcal{S}_1$.

The region $\mathfrak{R}_2 \cup \mathcal{S}_2$ is constructed similarly, with the cone $u = -1$ now being extended backwards, first using the characteristic gluing data and then using the tangential data induced by $\mathfrak{R}_1$ on $u = -1$. Again, Proposition 3.1.3 is used to construct the thin strip $\mathcal{S}_2$.

Finally, the spacetime is constructed by taking $\mathcal{Q} \doteq (\mathfrak{R}_1 \cup \mathcal{S}_1) \cup (\mathfrak{R}_2 \cup \mathcal{S}_1)$ and pasting r, Ω^2, ϕ, Q, and A_u. From the construction, it is clear that the dynamical variables, together with all v-derivatives consistent with C^k regularity are continuous on $\mathcal{Q}$. To show that all u-derivatives are continuous across $u = -1$, we observe that all transverse quantities are initialized consistently to k-th order at $(-1, v_1)$ and that the tangential data agrees by construction. Now Proposition 5.1.1 implies that the transverse derivatives through order k are equal on $u = -1$ in both $\mathfrak{R}_1 \cup \mathcal{S}_1$ and $\mathfrak{R}_2 \cup \mathcal{S}_2$. This completes the proof. □

Remark 5.2.5. If the characteristic gluing hypothesis is C^k but no better and the original solutions $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are more regular than C^k, then one expects $(k+1)$-th derivatives of dynamical quantities to jump across any of the null hypersurfaces bordering the light gray regions in Fig. 5.1.

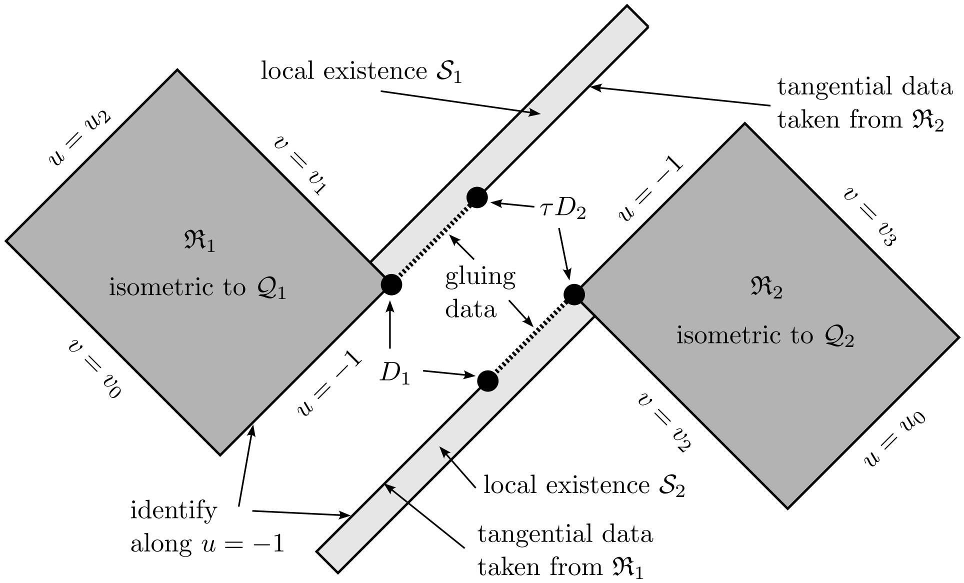

Figure 5.2: Proof of Proposition 5.2.4.

5.3 Sphere data in Minkowski, Schwarzschild, and Reissner–Nordström

Before stating our main gluing results, we need to precisely define the terms *Minkowski sphere*, *Schwarzschild event horizon sphere*, and *Reissner–Nordström event horizon sphere*.

Definition 5.3.1 (Minkowski sphere data)**.** Let $k \in \mathbb{N}$ and $R > 0$. The unique lapse normalized sphere data set satisfying

- $\varrho = R$,
- $\varrho_u^1 = -\frac{1}{2}$,
- $\varrho_v^1 = \frac{1}{2}$, and
- all other components zero,

is called the *Minkowski sphere data of radius* R and is denoted by $D^{\mathrm{M}}_{R,k}$.

Definition 5.3.2 (Schwarzschild sphere data)**.** Let $k \in \mathbb{N}$, $R > 0$, and $0 \leq 2M \leq R$. The unique lapse normalized sphere data set satisfying

- $\varrho = R$,
- $\varrho_u^1 = -\frac{1}{2}$

- $\varrho_v^1 = \frac{1}{2}(1 - 2M/R)$, and
- all other components zero,

is called the *Schwarzschild sphere data of mass M and radius R* and is denoted by $D^{\mathrm{S}}_{M,R,k}$. Note that $D^{\mathrm{S}}_{0,R,k} = D^{\mathrm{M}}_{R,k}$.

Definition 5.3.3 (Reissner–Nordström horizon sphere data)**.** Let $k \in \mathbb{N}$, $M > 0$, and $0 \le |e| \le M$. The unique lapse normalized sphere data set satisfying

- $\varrho = r_+ \doteq M + \sqrt{M^2 - e^2}$,
- $\varrho_u^1 = -\frac{1}{2}$,
- $\varrho_v^1 = 0$,
- $q = e$, and
- all other components zero,

is called the *Reissner–Nordström horizon sphere data with parameters M and e* and is denoted by $D^{\mathrm{RN}\mathcal{H}}_{M,e,k}$. Note that $D^{\mathrm{RN}\mathcal{H}}_{M,0,k} = D^{\mathrm{S}}_{M,2M,k}$.

We will also define sphere data for general Reissner–Nordström spheres. To do so, we extend the Hawking mass (2.1.2) to a function on sphere data sets $D \in \mathcal{D}_k$ by setting

$$m(D) \doteq \frac{\varrho}{2}\left(1 + \frac{4\varrho_u^1 \varrho_v^1}{\omega}\right).$$

We also define the *modified Hawking mass* of a spherically symmetric spacetime with charge by

$$\varpi \doteq m + \frac{Q^2}{2r}$$

and extend it to sphere data sets by

$$\varpi(D) \doteq m(D) + \frac{q^2}{2\varrho}.$$

In a Reissner–Nordström spacetime of mass M and charge e, any sphere data set D associated to a symmetry sphere has $\varpi(D) = M$. Note that given $\varrho > 0, \varrho_u^1 < 0, \omega$, and q, ϱ_v^1 is determined uniquely by $\varpi(D)$.

Recall that the horizons of Reisser–Nordström with parameters $|e| \leq M$ are located at

$$r_\pm = M \pm \sqrt{M^2 - e^2}.$$

Definition 5.3.4 (Reissner–Nordström sphere data)**.** Let $k \in \mathbb{N}$, $e \in \mathbb{R}$, and $R > 0$ satisfy $M > e^2/(2R)$. A lapse normalized sphere data set satisfying

- $\varrho = R$,
- $q = e$,
- $\varpi = M$,
- $|\varrho_v^1| = \frac{1}{2}$, or $|\varrho_u^1| = \frac{1}{2}$, or $\varrho_v^1 = \varrho_u^1 = 0$,
- all other components zero

is called a *Reissner–Nordström sphere data set of modified Hawking mass M, charge e, and radius* R and is denoted by $D^{\mathrm{RN}}_{M,e,R,k}$.

Remark 5.3.5. A Reissner–Nordström sphere data set of modified Hawking mass M, charge e, and radius R, $D^{\mathrm{RN}}_{M,e,R,k}$, gives rise to unique sphere data if either $\varrho_v^1 = \varrho_u^1 = 0$, or one additionally specifies $\mathrm{sgn}(\varrho_v^1) \in \{+,-\}$ or $\mathrm{sgn}(\varrho_u^1) \in \{+,-\}$.

5.4 Main gluing theorems

With the previous definitions of Section 5.2 and Section 5.3 at hand, we are now in a position to state our main gluing results.

Our first gluing theorem concerns gluing a sphere in Minkowski space to a Schwarzschild event horizon with a *real* scalar field. When the scalar field ϕ in the EMCSF system is real-valued, Maxwell's equation decouples from the rest of the system and the charge Q is constant throughout the spacetime. Since Q must vanish on any sphere in Minkowski space, it vanishes everywhere and the EMCSF system reduces to the Einstein-scalar field system.

Theorem 5.4.1. *For any $k \in \mathbb{N}$ and $0 < R_i < 2M_f$, the Minkowski sphere of radius R_i, $D^{\mathrm{M}}_{R_i,k}$, can be characteristically glued to the Schwarzschild event horizon sphere with mass M_f, $D^{\mathrm{S}}_{M_f,k}$, to order C^k within the Einstein-scalar field model in spherical symmetry.*

The proof of Theorem 5.4.1 is given in Section 5.5.1. We have separated out Minkowski to Schwarzschild gluing as a special case because it is simpler and highlights our topological argument. We will actually use this special case as the first step to produce our counterexample to the third law in Section 5.6.2.

Our second gluing theorem concerns gluing a sphere in the domain of outer communication of a Schwarzschild spacetime to a Reissner–Nordström event horizon with specified mass and charge to mass ratio.

Theorem 5.4.2. *For any $k \in \mathbb{N}$, $\mathfrak{q} \in [-1, 1]$, and $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$, there exists a number $M_0(k, \mathfrak{q}, \mathfrak{e}) \geq 0$ such that if $M_f > M_0$, $0 \leq M_i \leq \frac{1}{8} M_f$, and $2M_i < R_i \leq \frac{1}{2} M_f$, then the Schwarzschild sphere of mass M_i and radius R_i, $D^{\mathrm{S}}_{M_i,R_i,k}$, can be characteristically glued to the Reissner–Nordström event horizon with mass M_f and charge to mass ratio $\mathfrak{q}$, $D^{\mathrm{RN}\mathcal{H}}_{M_f,\mathfrak{q}M_f,k}$, to order C^k within the Einstein–Maxwell-charged scalar field model with coupling constant $\mathfrak{e}$. The associated characteristic data can be chosen to have no spherically symmetric antitrapped surfaces, i.e. $\partial_u r < 0$ everywhere.*

The proof of Theorem 5.4.2 is given in Section 5.5.2.

Remark 5.4.3. The data constructed in the proof of Theorem 5.4.1 will automatically not contain spherically symmetric antitrapped surfaces because of a special monotonicity property in the absence of charge. Namely,

$$\partial_v (r \partial_u r) = -\frac{\Omega^2}{4}, \tag{5.4.1}$$

so $r\partial_u r$ is decreasing. In particular, since $r\partial_u r$ is negative in Minkowski space, the sign will propagate in view of (5.4.1) for the Einstein-scalar field model.

Our next gluing theorem supersedes Theorem 5.4.1 and Theorem 5.4.2 by relaxing the requirement that the final sphere lie on the event horizon. The proof is slightly more involved than Theorem 5.4.2 but has the same basic structure and is given in Section 5.5.3 below.

Theorem 5.4.4. *For any $k \in \mathbb{N}, \mathfrak{q} \in \mathbb{R}$, $\mathfrak{e} \in \mathbb{R}\setminus\{0\}$ and $\mathfrak{r} > 0$, there exists a number $M_0(k, \mathfrak{q}, \mathfrak{e}, \mathfrak{r}) > 0$ such that if $M_f > M_0$ and*

$$R_f \geq \frac{M_f}{2}(1 + \mathfrak{r})\mathfrak{q}^2, \tag{5.4.2}$$

then there exists $R_i \in (0, R_f)$ such that the Minkowski sphere of radius R_i, $D^{\mathrm{M}}_{R_i,k}$, can be characteristically glued to the Reissner–Nordström sphere with modified Hawking mass M_f, charge $\mathfrak{q}M_f$, and radius R_f, $D^{\mathrm{RN}}_{M_f,\mathfrak{q}M_f,R_f,k}$ with $\varrho^1_u < 0$, to order C^k within the Einstein–Maxwell-charged scalar field system with coupling constant $\mathfrak{e}$. The associated characteristic data can be chosen to have no spherically symmetric antitrapped surfaces, i.e., $\partial_u r < 0$ everywhere.

Remark 5.4.5. Reissner–Nordström spheres with modified Hawking mass M, charge $\mathfrak{q}M$ and radius $R \leq \frac{M}{2}\mathfrak{q}^2$ have non-positive Hawking mass, $m \leq 0$. In this sense, the assumption $\mathfrak{r} > 0$ in Theorem 5.4.4 is necessary. Indeed, one immediately sees that (5.4.2) implies

$$m \geq \frac{\mathfrak{r}}{1+\mathfrak{r}} M_f,$$

so that $\mathfrak{r} > 0$ ensures $m > 0$.

Remark 5.4.6. Theorem 5.4.4 also allows for gluing of Minkowski space to Reissner–Nordström Cauchy horizons located at $r = r_-$. This is achieved by setting $\mathfrak{r} = \mathfrak{q}^2/4$ in Theorem 5.4.4, see already the proof of Corollary 5.6.7.

While all the above theorems are stated as gluing results along outgoing cones, by mapping $u \mapsto -v$ and $v \mapsto -u$, they also hold true for gluing along ingoing cones, recall Remark 5.2.3. In particular, restating Theorem 5.4.4 for gluing along ingoing cones gives

Theorem 5.4.7. *For any $k \in \mathbb{N}, \mathfrak{q} \in \mathbb{R}$, $\mathfrak{e} \in \mathbb{R}\backslash\{0\}$ and $\mathfrak{r} > 0$, there exists a number $M_0(k, \mathfrak{q}, \mathfrak{e}, \mathfrak{r}) > 0$ such that if $M_f > M_0$ and*

$$R_f \geq \frac{M_f}{2}(1+\mathfrak{r})\mathfrak{q}^2, \tag{5.4.3}$$

then there exists $R_i \in (0, R_f)$ such that the Reissner–Nordström sphere with modified Hawking mass M_f, charge $\mathfrak{q}M_f$, and radius R_f, $D^{\mathrm{RN}}_{M_f,\mathfrak{q}M_f,R_f,k}$ with $\varrho^1_v > 0$, can be characteristically glued along an ingoing cone to the Minkowski sphere of radius R_i, $D^{\mathrm{M}}_{R_i,k}$, to order C^k within the Einstein–Maxwell-charged scalar field system with coupling constant $\mathfrak{e}$. The associated characteristic data can be chosen to have no spherically symmetric trapped surfaces, i.e., $\partial_v r > 0$ everywhere.

5.5 Proofs of the main gluing theorems

We begin with two lemmas which identify the orbits of Schwarzschild and Reissner–Nordström sphere data under the action of the full gauge group. This essentially amounts to a version of Birkhoff's theorem for sphere data.

Lemma 5.5.1 (Schwarzschild exterior sphere identification)**.** *If $D \in \mathcal{D}_k$ satisfies*

- $\varrho = R > 0$,
- $\varrho^1_u < 0$,
- $\varrho^1_v > 0$,

- $\frac{1}{2}\varrho(1 + 4\varrho_u^1\varrho_v^1) = M$,

- $q = 0$, *and*

- $\varphi_u^i = \varphi_v^i = 0$ *for* $0 \le i \le k$,

then $R > 2M$ *and* D *is equivalent to* $D^{\mathrm{S}}_{M,R,k}$ *up to a gauge transformation.*

Proof. First, we observe that by the relations obtained from Maxwell's equations, $q_u^i = q_v^i = 0$ for $1 \le i \le k$. Since $\varphi_u^i = \varphi_v^i = 0$, we can perform an EM gauge transformation to make $a_u^i = 0$ for $0 \le i \le k$. Also, $a_v^i = 0$ for $1 \le i \le k$ from $F = d(A_u\,du)$. Next, we can normalize the lapse. Finally, $R > 2M$ follows from the definitions and $\varrho_u^1\varrho_v^1 < 0$. □

Lemma 5.5.2 (Reissner–Nordström horizon sphere identification)**.** *If* $D \in \mathcal{D}_k$ *satisfies*

- $\varrho = (1 + \sqrt{1 - \mathfrak{q}^2})M$ *for* $\mathfrak{q} \in [-1, 1]$ *and* $M > 0$,

- $\varrho_u^1 < 0$,

- $\varrho_v^1 = 0$,

- $q = \mathfrak{q}M$, *and*

- $\varphi_u^i = \varphi_v^i = 0$ *for* $0 \le i \le k$,

then D *is equivalent to* $D^{\mathrm{RN}}_{M,\mathfrak{q}M,k}$ *up to a gauge transformation.*

Proof. As before, the charge vanishes to all orders and we normalize the gauge potential and lapse. We then use the additional double null gauge freedom $u \mapsto \lambda u$, $v \mapsto \lambda^{-1}v$ to make $\varrho_u^1 = -\frac{1}{2}$. □

Remark 5.5.3. Without the condition $\varrho_u^1 < 0$ in the previous lemma, the sphere data in the extremal case could also arise from the *Bertotti–Robinson* universe.

With these lemmas and Remark 5.2.2 in mind, we follow the strategy discussed in Section 5.2. We fix the interval $[0, 1]$, set $\Omega^2 \equiv 1$, and solve Raychaudhuri's equation, Maxwell's equation, and the transport equation for transverse derivatives of ϕ with appropriate initial and final values. We do **not** have to track transverse derivatives of $\partial_u r$, Ω^2, Q, or A_u, because these will be "gauged away" at the end of the proof.

5.5.1 Proof of Theorem 5.4.1

In this subsection we prove Theorem 5.4.1. We first note that if the scalar field is chosen to be real-valued, the Einstein–Maxwell-charged scalar field system collapses to the Einstein-scalar field system. If the initial data has no charge ($Q(0) = 0$), then this is equivalent to setting $\mathfrak{e} = 0$ and A_u and all its derivatives to be identically zero.

We will first set up our scalar field ansatz as a collection of pulses. To do so, let

$$0 = v_0 < v_1 < \cdots < v_k < v_{k+1} = 1$$

be an arbitrary partition of $[0, 1]$. For each $1 \leq j \leq k + 1$, fix a nontrivial bump function

$$\chi_j \in C_c^\infty((v_{j-1}, v_j); \mathbb{R}).$$

In the rest of this section, the functions $\chi_1, \dots, \chi_{k+1}$ are fixed and our constructions depend on these choices.

Let $\alpha = (\alpha_1, \dots, \alpha_{k+1}) \in \mathbb{R}^{k+1}$ and set

$$\phi_\alpha(v) \doteq \phi(v; \alpha) \doteq \sum_{1 \leq j \leq k+1} \alpha_j \chi_j(v). \tag{5.5.1}$$

We set $\Omega^2(v; \alpha) \equiv 1$ along $[0, 1]$ and define $r(v; \alpha)$ as the unique solution of Raychauduri's equation (2.2.7) with this scalar field ansatz,

$$\partial_v^2 r(v; \alpha) = -r(v; \alpha)(\partial_v \phi_\alpha(v))^2, \tag{5.5.2}$$

with prescribed "final values"

$$r(1; \alpha) = 2M_f$$
$$\partial_v r(1; \alpha) = 0.$$

Let $0 < \varepsilon < 2M_f - R_i$. By Cauchy stability and monotonicity properties of Raychaudhuri's equation (5.5.2), there exists a $\delta > 0$ such that for every $0 < |\alpha| \leq \delta$,

$$\sup_{[0,1]} |r(\cdot;\alpha) - 2M_f| \leq \varepsilon,$$

$$\inf_{[0,1]} \partial_v r(\cdot;\alpha) \geq 0,$$

$$\partial_v r(0;\alpha) > 0.$$

The final inequality follows from the fact that $\alpha \neq 0$.

We now consider the sphere $S^k_\delta \doteq \{\alpha \in \mathbb{R}^{k+1} : |\alpha| = \delta\}$. For each $\alpha \in S^k_\delta$, define $D_\alpha(0) \in \mathcal{D}_k$ by setting

- $\varrho = r(0;\alpha) > 0$,
- $\varrho^1_v = \partial_v r(0;\alpha) > 0$,
- $\varrho^1_u = -\frac{1}{4}(\varrho^1_v)^{-1}$,
- $\omega = 1$, and
- all other components to zero.

By Lemma 5.5.1, $D_\alpha(0)$ is equivalent to $D^{\mathrm{M}}_{r(0;\alpha),k}$ up to a gauge transformation.

For each $\alpha \in S^k_\delta$, we now apply Proposition 5.1.6 and Proposition 5.1.8 to uniquely determine cone data

$$D_\alpha : [0,1] \to \mathcal{D}_k,$$

with initialization $D_\alpha(0)$ above and seed data ϕ_α given by (5.5.1). By standard ODE theory, $D_\alpha(v)$ is jointly continuous in v and α. Note that $\varrho(D_\alpha(v)) = r(v;\alpha)$ and $\varphi(D_\alpha(v)) = \phi(v;\alpha)$ by definition. We now use the notation

$$\partial^i_u \phi(v;\alpha) \doteq \varphi^i_u(D_\alpha(v))$$

for $i = 1, \ldots, k$ to denote the transverse derivatives of the scalar field obtained by Proposition 5.1.6.

By construction, the data set $D_\alpha(1)$ satisfies

- $\varrho = 2M_f$,
- $\varrho^1_u < 0$,
- $\varrho^1_v = 0$,
- $\omega = 1$, and

- $\varphi_v^i = 0$ for $0 \leq i \leq k$.

The second property follows from the initialization of ϱ_u^1 in $D_\alpha(0)$ and the monotonicity of

$$(r\partial_u r)(v;\alpha) \doteq \varrho(D_\alpha(v))\varrho_u^1(D_\alpha(v))$$

in the Einstein-scalar field system discussed in Remark 5.4.3.

In order to glue to Schwarzschild at $v = 1$, by Lemma 5.5.2, it suffices to find an $\alpha_* \in S_\delta^k$ for which additionally

$$\partial_u\phi(1;\alpha_*) = \cdots = \partial_u^k\phi(1;\alpha_*) = 0.$$

The following discrete symmetry of the Einstein-scalar field system plays a decisive role in finding α_*.

A function $f(v;\alpha)$ is *even in* α if $f(v;-\alpha) = f(v;\alpha)$ and *odd in* α if $f(v;-\alpha) = -f(v;\alpha)$.

Lemma 5.5.4. *As functions on* $[0,1] \times S_\delta^k$*, the metric coefficients* $r(v;\alpha)$*,* $\Omega^2(v;\alpha)$ *and all their ingoing and outgoing derivatives are even functions of* α*. The scalar field* $\phi(v;\alpha)$ *and all its ingoing and outgoing derivatives are odd functions of* α*. In particular, the map*

$$\begin{aligned} F : S_\delta^k &\to \mathbb{R}^k \\ \alpha &\mapsto \big(\partial_u\phi(1;\alpha), \ldots, \partial_u^k\phi(1;\alpha)\big) \end{aligned} \tag{5.5.3}$$

is continuous and odd.

Proof. The scalar field itself is odd by the definition (5.5.1). Since Raychaudhuri's equation (5.5.2) involves the square of $\partial_v\phi(v;\alpha)$, $r(v;\alpha)$ will be automatically even. Next, $\partial_u r(v;\alpha)$ is found by integrating the wave equation for the radius (2.2.4), forwards in v with initial value determined by $D_\alpha(0)$. Since ϕ enters into this equation with an even power (namely zero), $\partial_u r(v;\alpha)$ will also be even. The wave equation for $r\phi$ in the Einstein-scalar field model can be derived from (2.2.12) and reads

$$\partial_u\partial_v(r\phi) = -\frac{\Omega^2 m}{2r^3}\, r\phi,$$

and the right-hand side is odd in α (the Hawking mass is constructed from metric coefficients so is also even). Recall from Proposition 5.1.6 that this wave equation is used to compute $\varphi_u^i(D_\alpha(v))$. By inspection $\partial_u(r\phi)$ is odd, whence $\partial_u\phi(v;\alpha)$ is also odd. The proof now follows by inductively following the procedure of Proposition 5.1.1, taking note of the fact that the transport equations for ingoing derivatives of r and Ω^2 only involve even powers of ϕ and its derivatives, whereas the transport equations for ingoing derivatives of ϕ only involve odd powers.

The claim about the map F follows from the oddness of ingoing derivatives of ϕ and the continuity of all dynamical quantities in α, per standard ODE theory. $\square$

We now complete the proof of Theorem 5.4.1. By the Borsuk–Ulam theorem stated as Theorem 4.4.4, $F(\alpha_*) = 0$ for some $\alpha_* \in S^k_\delta$, where F is as in (5.5.3). By Lemma 5.5.2, $D_{\alpha_*}(1)$ is gauge equivalent to $D^{\mathrm{S}}_{M_f,k}$.

So far we have glued $D^{\mathrm{M}}_{r(0;\alpha),k}$ to $D^{\mathrm{S}}_{M_f,k}$, and since $r(0;\alpha) > R_i$, we extend the data trivially in order to glue $D^{\mathrm{M}}_{R_i,k}$ to $D^{\mathrm{S}}_{M_f,k}$, which concludes the proof of Theorem 5.4.1. $\square$

5.5.2 Proof of Theorem 5.4.2

In this subsection we prove Theorem 5.4.2. We assume that $\mathfrak{q} \neq 0$, the $\mathfrak{q} = 0$ version of this result being essentially a repeat of the arguments in the previous section combined with the new initialization of $\partial_u r(0;\alpha)$ in (5.5.17) below.

In this subsection we adopt the notational convention that $A \lesssim B$ means $A \le CB$, where C is a constant that depends only on k and the baseline scalar field profile, but not on $\mathfrak{q}$, $\mathfrak{e}$, M_i, M_f, or α. The notation $A \approx B$ means $A \lesssim B$ and $B \lesssim A$.

Let

$$0 = v_0 < v_1 < \cdots < v_{2k} < v_{2k+1} = 1$$

be an arbitrary partition of $[0, 1]$. For each $1 \le j \le 2k + 1$, fix a nontrivial bump function

$$\chi_j \in C^\infty_c((v_{j-1}, v_j); \mathbb{R}).$$

In the rest of this section, the functions $\chi_1, \dots, \chi_{2k+1}$ are fixed and our constructions depend on these choices.

For $\alpha = (\alpha_1, \dots, \alpha_{2k+1}) \in \mathbb{R}^{2k+1}$, set

$$\phi_\alpha(v) \doteq \phi(v;\alpha) \doteq \sum_{1\le j\le 2k+1} \alpha_j \chi_j(v) e^{-iv}. \tag{5.5.4}$$

Remark 5.5.5. If $\mathfrak{e} > 0$, this choice of ϕ will make $Q \ge 0$, with is consistent with $\mathfrak{q} > 0$. If $\mathfrak{e} > 0$ and $\mathfrak{q} < 0$, then we replace $-iv$ in the exponential with $+iv$. Similarly, the cases $\mathfrak{e} < 0$, $\mathfrak{q} > 0$ and $\mathfrak{e} < 0$, $\mathfrak{q} < 0$ can be handled. Therefore, we assume without loss of generality that $\mathfrak{e} > 0$, $\mathfrak{q} > 0$.

For $\hat{\alpha} \in S^{2k}$ (the unit sphere in $\mathbb{R}^{2k+1}$) and $\beta \geq 0$, it is convenient to define $r(v;\beta,\hat{\alpha}) = r(v;\beta\hat{\alpha})$, etc. We again set $\Omega^2(v;\alpha) \equiv 1$ and study the equations (2.2.7) and (2.2.9) for $v \in [0,1]$ with the ϕ_α ansatz:

$$\partial_v^2 r(v;\alpha) = -|\alpha|^2 r(v;\alpha)|\partial_v\phi_{\hat{\alpha}}(v)|^2, \tag{5.5.5}$$

$$\partial_v Q(v;\alpha) = \mathfrak{e}|\alpha|^2 r(v;\alpha)^2 \mathrm{Im}(\phi_{\hat{\alpha}}(v)\overline{\partial_v\phi_{\hat{\alpha}}(v)}). \tag{5.5.6}$$

In addition, we again define r at $v=1$ by

$$r(1;\alpha) = r_+,$$
$$\partial_v r(1;\alpha) = 0,$$

and Q at $v = 0$ by

$$Q(0;\alpha) = 0, \tag{5.5.7}$$

which together with (5.5.5) and (5.5.6) uniquely determine r and Q on $[0,1]$. Note that we will initialize $\partial_u r$ only later in (5.5.17).

We first note that basic calculations yield

$$|\partial_v\phi_{\hat{\alpha}}|^2 = \sum_{1\leq j\leq 2k+1} \hat{\alpha}_j^2\left(\chi_j^2 + \chi_j'^2\right)$$

and

$$\mathrm{Im}(\phi_{\hat{\alpha}}\overline{\partial_v\phi_{\hat{\alpha}}}) = \sum_{1\leq j\leq 2k+1} \hat{\alpha}_j^2\chi_j^2.$$

Therefore,

$$\int_0^1 |\partial_v\phi_{\hat{\alpha}}|^2\,dv \approx \int_0^1 \mathrm{Im}(\phi_{\hat{\alpha}}\overline{\partial_v\phi_{\hat{\alpha}}})\,dv \approx 1$$

for any $\hat{\alpha} \in S^{2k}$.

Lemma 5.5.6. *There exists a constant $0 < c \lesssim 1$ such that if $0 < \beta \leq c$, then for any $\hat{\alpha} \in S^{2k}$, $r(\cdot;\beta\hat{\alpha})$ satisfies*

$$r(v;\beta\hat{\alpha}) \geq \tfrac{1}{2}r_+ \tag{5.5.8}$$

$$\partial_v r(v;\beta\hat{\alpha}) \geq 0 \tag{5.5.9}$$

for $v \in [0,1]$*, where*

$$r_+ \doteq \left(1+\sqrt{1-\mathfrak{q}^2}\right) M_f.$$

Furthermore,

$$\partial_v r(0;\beta\hat{\alpha}) > 0. \tag{5.5.10}$$

Proof. This is a simple bootstrap argument in v. Assume that on $[v_0,1]\subset[0,1]$, we have

$$\inf_{[v_0,1]} r \geq 0$$
$$\inf_{[v_0,1]} \partial_v r \geq 0.$$

This is clear for v_0 close to 1 by Cauchy stability. From Raychaudhuri's equation (5.5.5), $r \geq 0$ implies $\partial_v r$ is monotone decreasing, hence is bounded above by $\partial_v r(v_0)$, which can be estimated by

$$\partial_v r(v_0) = \int_{v_0}^1 \beta^2 r |\partial_v \phi_{\hat{\alpha}}|^2\, dv \lesssim \beta^2 r_+, \tag{5.5.11}$$

since $r \leq r_+$ on $[v_0,1]$. It follows that

$$r(v_0) = r_+ - \int_{v_0}^1 \partial_v r\, dv \geq r_+ - C\beta^2 r_+ \tag{5.5.12}$$

for some $C \lesssim 1$. Choosing $\beta > 0$ sufficiently small shows $r(v_0) \geq \frac{1}{2} r_+$ which improves the bootstrap assumptions and proves the desired estimate (5.5.8). Finally, note that (5.5.10) holds true as $\partial_v r$ is monotone decreasing and r is not constant ($\beta > 0$ and the scalar field is not identically zero). □

Lemma 5.5.7. *By potentially making the constant* c *from Lemma 5.5.6 smaller, we have that for any* $0 < \beta \leq c$ *and* $\hat{\alpha} \in S^{2k}$*, the following estimate holds*

$$\frac{\partial}{\partial \beta} Q(1;\beta,\hat{\alpha}) > 0.$$

Proof. Integrating Maxwell's equation (5.5.6) and using (5.5.7), we find

$$Q(1;\beta,\hat{\alpha}) = \int_0^1 \mathfrak{e}\beta^2 r^2 \mathrm{Im}(\phi_{\hat{\alpha}} \overline{\partial_v \phi_{\hat{\alpha}}})\, dv.$$

A direct computation yields

$$\partial_\beta Q(1;\beta,\hat{\alpha}) = 2\mathfrak{e}\beta \int_0^1 (r^2 + \beta r \partial_\beta r)\mathrm{Im}(\phi_{\hat{\alpha}}\overline{\partial_v \phi_{\hat{\alpha}}})\, dv.$$

Note that $\mathrm{Im}(\phi_{\hat{\alpha}}\overline{\partial_v \phi_{\hat{\alpha}}}) \geq 0$ pointwise and is not identically zero. Since $0 < \beta \leq c$, we use Lemma 5.5.6 to estimate

$$r^2 + \beta r \partial_\beta r \geq \tfrac{1}{4} r_+^2 - C\beta r_+^2 = r_+^2(\tfrac{1}{4} - C\beta),$$

where we also used $|\partial_\beta r| \lesssim r_+$ which follows directly from differentiating (5.5.5) with respect to $\beta = |\alpha|$. Therefore, by choosing c even smaller, we obtain $\partial_\beta Q(1;\beta,\hat{\alpha}) > 0$. □

Lemma 5.5.8. *If $\mathfrak{e}M_f/\mathfrak{q}$ is sufficiently large depending only on k and the choice of profiles, then there is a smooth function $\beta_Q : S^{2k} \to (0,\infty)$ so that $Q(1;\beta_Q(\hat{\alpha}),\hat{\alpha}) = \mathfrak{q}M_f$ for every $\hat{\alpha} \in S^{2k}$, which also satisfies*

$$\beta_Q(\hat{\alpha}) \approx \frac{\sqrt{\mathfrak{q}M_f}}{\sqrt{\mathfrak{e}}r_+} \tag{5.5.13}$$

$$\beta_Q(-\hat{\alpha}) = \beta_Q(\hat{\alpha}) \tag{5.5.14}$$

for every $\hat{\alpha} \in S^{2k}$.

Proof. As in the proof of Lemma 5.5.7 we have

$$Q(1;\beta,\hat{\alpha}) = \mathfrak{e}\beta^2 \int_0^1 r^2 \mathrm{Im}(\phi_{\hat{\alpha}}\overline{\partial_v \phi_{\hat{\alpha}}}).$$

If β is sufficiently small so that Lemma 5.5.6 and Lemma 5.5.7 apply, we estimate

$$Q(1;\beta,\hat{\alpha}) \approx \mathfrak{e}\beta^2 r_+^2.$$

For $\mathfrak{e}M_f/\mathfrak{q}$ sufficiently large as in the assumption, we apply now the intermediate value theorem, to obtain a $\beta_Q(\hat{\alpha})$ satisfying $0 < \beta_Q(\hat{\alpha}) \leq c$ such that

$$Q(1;\beta_Q,\hat{\alpha}) = \mathfrak{q}M_f. \tag{5.5.15}$$

Note that $\beta_Q(\hat{\alpha})$ is unique since $Q(1;\cdot,\hat{\alpha})$ is strictly increasing as shown in Lemma 5.5.7. Moreover, since $Q(1;\cdot,\cdot)$ is smooth (note that $\hat{\alpha} \in S^{2k}$ and $\beta > 0$ enter as smooth parameters in (5.5.6) which defines Q), a direct application of the implicit function theorem using that $\partial_\beta Q(1;\cdot,\hat{\alpha}) \neq 0$ shows that $\beta_Q : S^{2k} \to (0, \infty)$ is smooth.

Moreover, by (5.5.2) and (5.5.15), β_Q satisfies

$$\mathfrak{e}\beta_Q^2 r_+^2 \approx \mathfrak{q} M_f$$

which shows (5.5.13). Finally, note that $Q(1; \beta, -\hat{\alpha}) = Q(1; \beta, \hat{\alpha})$, from which (5.5.14) follows. □

Lemma 5.5.9. *Let $\mathfrak{e}M_f/\mathfrak{q}$ be sufficiently large (depending only on k and the choice of profiles) so that Lemma 5.5.8 applies. Then*

$$p_Q : S^{2k} \to \mathfrak{Q}^{2k}$$
$$\hat{\alpha} \mapsto \beta_Q(\hat{\alpha})\hat{\alpha}$$

is a diffeomorphism, where

$$\mathfrak{Q}^{2k} \doteq \{\beta_Q(\hat{\alpha})\hat{\alpha} \colon \hat{\alpha} \in S^{2k}\} \subset \mathbb{R}^{2k+1}$$

is the radial graph of β_Q. Moreover, $\mathfrak{Q}^{2k}$ is invariant under the antipodal map $A(\alpha) = -\alpha$ and p_Q commutes with the antipodal map.

Proof. By definition of $\mathfrak{Q}^{2k}$ and the facts that β_Q is smooth, positive, and invariant under the antipodal map as proved in Lemma 5.5.8, the stated properties of $\mathfrak{Q}^{2k}$ and p_Q follow readily. □

Having identified the set $\mathfrak{Q}^{2k}$ which guarantees gluing of the charge Q, for the rest of the section we will always take $\alpha \in \mathfrak{Q}^{2k}$. Recall from (5.5.13) that for every $\alpha \in \mathfrak{Q}^{2k}$:

$$|\alpha| \approx \frac{\sqrt{\mathfrak{q}M_f}}{\sqrt{\mathfrak{e}}r_+}. \tag{5.5.16}$$

Before proceeding to choose sphere data, we will need to examine the equation for $\partial_u r$ because this will place a further restriction on α which must be taken into account before setting up the topological argument. We continue by using the definition of the Hawking mass m in (2.1.2), to impose the condition

$$m(0; \alpha) = M_i$$

by initializing

$$\partial_u r(0; \alpha) = -\left(1 - \frac{2M_i}{r(0; \alpha)}\right) \frac{1}{4\partial_v r(0; \alpha)}. \tag{5.5.17}$$

The transverse derivative $\partial_u r(v;\alpha)$ is now determined by solving (2.2.4),

$$\partial_v \partial_u r(v;\alpha) = -\frac{1}{4r(v;\alpha)^2} - \frac{\partial_u r(v;\alpha)\partial_v r(v;\alpha)}{r(v;\alpha)^2} + \frac{Q(v;\alpha)^2}{4r(v;\alpha)^3}, \tag{5.5.18}$$

with initialization (5.5.17).

Note that (5.5.17) is well-defined by (5.5.10) and (5.5.8) from Lemma 5.5.6. Furthermore,

$$1 - \frac{2M_i}{r(0;\alpha)} \geq 1 - \frac{4M_i}{M_f} > 0,$$

so

$$\partial_u r(0;\alpha) < 0. \tag{5.5.19}$$

Having initialized $\partial_u r$ at $v = 0$, we determine $\partial_u r(v;\alpha)$ using (5.5.18), and we will now show that for $\mathfrak{e}M_f/\mathfrak{q}$ sufficiently large, $\partial_u r(v;\alpha) < 0$ for all $v \in [0,1]$.

Lemma 5.5.10. *If $\mathfrak{e}M_f/\mathfrak{q}$ is sufficiently large depending only on k and the choice of profiles and if $0 \leq M_i \leq \frac{1}{8}M_f$, then*

$$\sup_{v\in[0,1]} \partial_u r(v;\alpha) < 0 \tag{5.5.20}$$

for every $\alpha \in \mathfrak{Q}^{2k}$.

Proof. Since $r > 0$ on $[0,1]$, it suffices to show that

$$\sup_{[0,1]} r\partial_u r < 0.$$

First, by (2.2.11),

$$|\partial_v(r\partial_u r)| = \left|\frac{1}{4}\left(1 - \frac{Q^2}{r^2}\right)\right| \lesssim 1, \tag{5.5.21}$$

as

$$Q(v;\alpha) \leq Q(1;\alpha) = \mathfrak{q}M_f \lesssim r(v;\alpha),$$

where we used (5.5.8). Integrating (5.5.21), we have

$$\sup_{v\in[0,1]} r(v)\partial_u r(v) \leq r(0)\partial_u r(0) + C_1, \tag{5.5.22}$$

where $C_1 \lesssim 1$ is a constant. Analogously to (5.5.11), we estimate

$$\partial_v r(0;\alpha) \lesssim |\alpha|^2 r_+ \lesssim \frac{\mathfrak{q}}{\mathfrak{e}},$$

where we used (5.5.16). Now, using (5.5.17),

$$\begin{aligned}
-r(0)\partial_u r(0) &= \frac{r(0) - 2M_i}{4\partial_v r(0)} \\
&\gtrsim \frac{\mathfrak{e}}{\mathfrak{q}}(\tfrac{1}{2}M_f - 2M_i) \\
&\gtrsim \frac{\mathfrak{e}}{\mathfrak{q}}M_f.
\end{aligned}$$

Therefore, we improve (5.5.22) to

$$\sup_{v\in[0,1]} r(v)\partial_u r(v) \leq -C_2\frac{\mathfrak{e}}{\mathfrak{q}}M_f + C_1$$

for some $C_2 \lesssim 1$. Thus, if $\mathfrak{e}M_f/\mathfrak{q}$ is sufficiently large we obtain (5.5.20). □

To continue the proof of Theorem 5.4.2, we now put our construction into the framework of the sphere data in Section 5.3. For each $\alpha \in \mathfrak{Q}^{2k}$, define $D_\alpha(0) \in \mathcal{D}_k$ by setting

- $\varrho = r(0;\alpha) \geq \frac{1}{2}r_+$ (see (5.5.8)),
- $\varrho_v^1 = \partial_v r(0;\alpha) > 0$ (see (5.5.10)),
- $\varrho_u^1 = \partial_u r(0;\alpha) < 0$ (see (5.5.17) and (5.5.19)),
- $\omega = 1$, and
- all other components to zero.

By Lemma 5.5.1, $D_\alpha(0)$ is equivalent to $D^{\mathrm{S}}_{M_i,r(0;\alpha),k}$ up to a gauge transformation.

For each $\alpha \in \mathfrak{Q}^{2k}$, we now apply Proposition 5.1.6 and Proposition 5.1.8 to uniquely determine cone data

$$D_\alpha : [0,1] \to \mathcal{D}_k,$$

with initialization $D_\alpha(0)$ above and seed data ϕ_α given by (5.5.4). By standard ODE theory, $D_\alpha(v)$ is jointly continuous in v and α. Note that $\varrho(D_\alpha(v)) = r(v;\alpha)$, $\varphi(D_\alpha(v)) = \phi(v;\alpha)$, and $q(D_\alpha(v)) = Q(v;\alpha)$ by definition. As in the proof of Theorem 5.4.1, we use the notation

$$\partial_u^i \phi(v;\alpha) \doteq \varphi_u^i(D_\alpha(v))$$

for $i = 1, \dots, k$ to denote the transverse derivatives of the scalar field obtained by Proposition 5.1.6. Note also that

$$\partial_u r(v;\alpha) = \varrho_u^1(D_\alpha(v)),$$

where $\partial_u r(v;\alpha)$ is as in (2.2.4) above.

By construction, the data set $D_\alpha(1)$ satisfies

- $\varrho = 2M_f$,
- $\varrho_u^1 < 0$ (see Lemma 5.5.10),
- $\varrho_v^1 = 0$,
- $\omega = 1$,
- $q = \mathfrak{q}M_f$ (definition of $\mathfrak{Q}^{2k}$), and
- $\varphi_v^i = 0$ for $0 \le i \le k$.

In order to glue to the appropriate Reissner–Nordström event horizon sphere, by Lemma 5.5.2, it suffices to find an $\alpha_* \in \mathfrak{Q}^{2k}$ for which additionally

$$\partial_u \phi(1;\alpha_*) = \cdots = \partial_u^k \phi(1;\alpha_*) = 0.$$

Analogously to Lemma 5.5.4 we first establish

Lemma 5.5.11. *The metric coefficients $r(v;\alpha)$, $\Omega^2(v;\alpha)$, the electromagnetic quantities $Q(v;\alpha)$, $A_u(v;\alpha)$, and all their ingoing and outgoing derivatives are even functions of α. The scalar field $\phi(v;\alpha)$ and all its ingoing and outgoing derivatives are odd functions of α.*

Proof. The proof is essentially the same as Lemma 5.5.4, noting that equations (2.2.8), (2.2.9), and (2.2.10) are also even in ϕ. □

We now complete the proof of Theorem 5.4.2. Recall from Lemma 5.5.9 that $p_Q : S^{2k} \to \mathfrak{Q}^{2k}$ is a diffeomorphism which commutes with the antipodal map. We now argue similarly to Section 5.5.1. By Lemma 5.5.11, the function

$$\begin{aligned} F : \mathfrak{Q}^{2k} &\to \mathbb{C}^k \\ \alpha &\mapsto \big(\partial_u \phi(1;\alpha), \dots, \partial_u^k \phi(1;\alpha)\big) \end{aligned}$$

is continuous and odd. Therefore, the Borsuk–Ulam theorem, stated as Theorem 4.4.4, applied to

$$(\mathrm{Re}F^1, \mathrm{Im}F^1, \ldots, \mathrm{Re}F^k, \mathrm{Im}F^k) \circ p_Q : S^{2k} \to \mathbb{R}^{2k},$$

where F^i is the ith component of F, shows that there is an $\alpha_* \in \mathfrak{Q}^{2k}$ such that $F(\alpha_*) = 0$. By Lemma 5.5.2, $D_{\alpha_*}(1)$ is gauge equivalent to $D^{\mathrm{RN}\mathcal{H}}_{M_f,\mathfrak{q}M_f,k}$ which concludes the gluing construction. Since we have already established that $\partial_u r < 0$ for all $v \in [0,1]$ in Lemma 5.5.10, this concludes the proof of Theorem 5.4.2. □

5.5.3 Proof of Theorem 5.4.4

In this section we extend our characteristic gluing result Theorem 5.4.2 to allow for sphere data at the final sphere which is not necessarily located on a horizon. Recall Definition 5.3.4 for the definition of general Reissner–Nordström sphere data. As the steps in the proof below are direct generalizations of the proof of Theorem 5.4.2, our presentation here will have fewer details.

Proof of Theorem 5.4.4. We only consider the case $\mathfrak{q} \neq 0$, the case $\mathfrak{q} = 0$ being strictly easier and requiring only "gluing 3" below. Without loss of generality, we may also assume $R_f \le 3M_f$ as for $r \ge 3M_f$ we can extend trivially with Reissner–Nordström data satisfying $\partial_v r > 0$ and $\partial_u r < 0$. In the following proof, we use the convention that all constants appearing in $\lesssim, \gtrsim$ and $\approx$ to also depend on $\mathfrak{q}, \mathfrak{r}$ and $\mathfrak{e}$. The theorem is proved as a consequence of the following three intermediate gluings:

1. $D^{\mathrm{M}}_{R_i,k}$ is glued to $D^{\mathrm{RN}}_{M',Q_f,R_1,k}$ with a complex scalar field,
2. $D^{\mathrm{RN}}_{M',Q_f,R_1,k}$ is glued to $D^{\mathrm{RN}}_{M',Q_f,R_2,k}$ trivially (i.e., with identically vanishing scalar field), and
3. $D^{\mathrm{RN}}_{M',Q_f,R_2,k}$ is glued to $D^{\mathrm{RN}}_{M_f,Q_f,R_f,k}$ with a real scalar field,

where $R_i \doteq R_f - M_f^{3/4}$, $0 < M' < M_f$ is an intermediate modified Hawking mass, $Q_f \doteq \mathfrak{q}M_f$, R_1, R_2 are intermediate radii which satisfy $R_i < R_1 < R_2$.

Gluing 1. In the interval $v \in [0,1]$ we impose the ansatz (5.5.4). At $v = 0$, we set

$$r(0) = R_i, \quad m(0) = Q(0) = 0, \quad \partial_u r(0) = -\frac{1}{2M_f^{1/2}}, \quad \partial_v r(0) = \frac{M_f^{1/2}}{2}. \tag{5.5.23}$$

The pulse parameters α_* which achieve gluing of transverse derivatives of ϕ are determined by the procedure of Section 5.5.2, with charge condition $Q(1;\alpha) = Q_f$. As in Section 5.5.2 we find that the gluing can be performed with parameters satisfying $|\alpha_*|^2 \lesssim M_f^{-1}$. Using this estimate on α_*, we obtain from Raychaudhuri's equation (2.2.7) and (5.5.23) that $\frac{1}{2}M_f^{1/2} \geq \partial_v r \geq \frac{1}{4}M_f^{1/2}$ for every $v \in [0,1]$ by choosing $M_0(k,\mathfrak{q},\mathfrak{e},\mathfrak{r})$ sufficiently large. This also implies $R_i \leq r \leq R_i + \frac{1}{2}M_f^{1/2}$. Using $r \geq R_f - M_f^{3/4}$ and the estimate analogous to (5.5.21) we infer $|r\partial_u r - r(0)\partial_u r(0)| \lesssim 1$ for every $v \in [0,1]$, i.e., $0 < -\partial_u r \leq M_f^{-1/2}$. We now estimate the Hawking mass at $v = 1$ by integrating (2.2.13),

$$m(1) = \int_0^1 2r^2(-\partial_u r)|\partial_v \phi|^2\,dv + \int_0^1 \frac{Q^2}{2r^2}\partial_v r\,dv \lesssim M_f^2 M_f^{-1/2} M_f^{-1} + M_f^{1/2} \lesssim M_f^{1/2}. \qquad (5.5.24)$$

Setting $R_1 = r(1)$ and $M' = m(1) + Q_f^2/(2R_1)$, we have shown that $R_i < R_1 \leq R_i + \frac{1}{2}M_f^{1/2}$. The condition (5.4.2) shows that $Q_f^2/(2R_f) \leq M_f/(1+\mathfrak{r})$. In particular, since $R_1 \geq R_f - M_f^{3/4}$ we estimate

$$M' = m(1) + \frac{Q_f^2}{2R_1} \leq \frac{1}{2}\left(1 + \frac{1}{1+\mathfrak{r}}\right)M_f = \frac{2+\mathfrak{r}}{2+2\mathfrak{r}}M_f \qquad (5.5.25)$$

by possibly taking $M_0(k,\mathfrak{q},\mathfrak{e},\mathfrak{r})$ larger. This completes the first gluing step.

Gluing 3. It is more convenient to now carry out the third gluing step and simply ensure that $R_2 > R_1$. We use a collection of $k+1$ real-valued pulses as in (5.5.1) on $v \in [0,1]$. We impose

$$r(1) = R_f, \quad \partial_u r(1) = -M_f, \quad Q(1) = Q_f, \quad \varpi(1) = M_f. \qquad (5.5.26)$$

This uniquely determines $\partial_v r(1)$ which can have either sign but satisfies $|\partial_v r(1)| \lesssim M_f^{-1}$. We also note that as long as $|\alpha|^2 \leq M_f^{-3/2}$, we have $|\partial_v r| \lesssim M_f^{-1/2}$ and thus $|r - R_f| \lesssim M_f^{-1/2}$ on $[0,1]$. This also gives $-\partial_u r \approx M_f$. Using

$$\varpi(1) - \varpi(0) = \int_0^1 2r^2(-\partial_u r)|\partial_v \phi|^2\,dv$$

and (5.5.25), we write the mass condition $\varpi(1) = M_f$ and $\varpi(0) = M'$ as a sphere of α's ($|\alpha|^2 \approx M_f^{-2}$) for which we will apply the Borsuk–Ulam argument. We use here that $M_f - M' = M_f\mathfrak{r}/(2+2\mathfrak{r})$. With $|\alpha_*|^2 \approx M_f^{-2}$ we have the improved estimate $|\partial_v r| \lesssim M_f^{-1}$ for $v \in [0,1]$ and thus, $|r(0) - R_f| \lesssim M_f^{-1}$. Taking now $M_0(k,\mathfrak{q},\mathfrak{e},\mathfrak{r})$ sufficiently large makes $R_2 \doteq r(0) > R_1$.

Gluing 2. By the previous constructions, we have $R_1 < R_2$, $\varpi(0) = \varpi(1) = M'$, $Q(0) = Q(1) = Q_f$, $\partial_v r(0) > 0$, and $\partial_u r(0) < 0$. Now $D^{\mathrm{RN}}_{M',Q_f,R_k}$ can be trivially glued to $D^{\mathrm{RN}}_{M',Q_f,R_2,k}$ by choosing $\phi \equiv 0$, and we must merely ensure that $\partial_u r < 0$ along the way. Since $\partial_v r > 0$ by Raychauduri's equation, this amounts to proving $\frac{2m}{r} < 1$. Indeed,

$$m(v) \le m(0) + \int_0^1 \frac{Q_f^2}{2r^2}\partial_v r\, dv = m(0) + \int_{R_1}^{R_2} \frac{Q_f^2}{2r^2}\, dr \lesssim m(0) + (R_2 - R_1) \lesssim M_f^{1/2} + M_f^{3/4},$$

where we used (5.5.24). In particular, by choosing $M_0(k, \mathfrak{q}, \mathfrak{e}, \mathfrak{r})$ larger, we can make $m(v)/M_f$ arbitrarily small and thus $\partial_u r < 0$ throughout gluing 2. □

5.6 Constructing the spacetimes and Cauchy data

In this final section we will prove our main result Theorem 1.1.11 as well as Theorem 1.1.4, Corollary 4.6.1, and Corollary 4.6.3.

5.6.1 Construction of gravitational collapse to Reissner–Nordström

We now state a more precise version of Theorem 1.1.4 as follows.

Corollary 5.6.1. *For any $k \in \mathbb{N}$, $\mathfrak{q} \in [-1, 1] \setminus \{0\}$, and $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$, let $M_0(k, \mathfrak{q}, \mathfrak{e})$ be as in Theorem 5.4.2. Then for any $M \ge M_0$ there exist asymptotically flat, spherically symmetric Cauchy data $(\Sigma, g_0, k_0, E_0, B_0, \phi_0, \phi_1)$ for the EMCSF system, with $\Sigma \cong \mathbb{R}^3$ and a regular center, such that the maximal future globally hyperbolic development $(\mathcal{M}^4, g, F, A, \phi)$ has the following properties:*

- *All dynamical quantities are at least C^k-regular.*
- *Null infinity $\mathcal{I}^+$ is complete.*
- *The black hole region is nonempty, $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+) \neq \emptyset$.*
- *The Cauchy surface Σ lies in the domain of outer communication $J^-(\mathcal{I}^+)$. In particular, it does not intersect the event horizon $\mathcal{H}^+ \doteq \partial(\mathcal{BH})$.*
- *The initial data hypersurface does not contain trapped surfaces.*
- *The spacetime does not contain antitrapped surfaces.*
- *For sufficiently late advanced times $v \ge v_0$, the domain of outer communication, including the event horizon, is isometric to that of a Reissner–Nordström solution with mass M charge to mass ratio $\mathfrak{q}$. For $v \ge v_0$, the event horizon of the spacetime can be identified with the event horizon of Reissner–Nordström.*

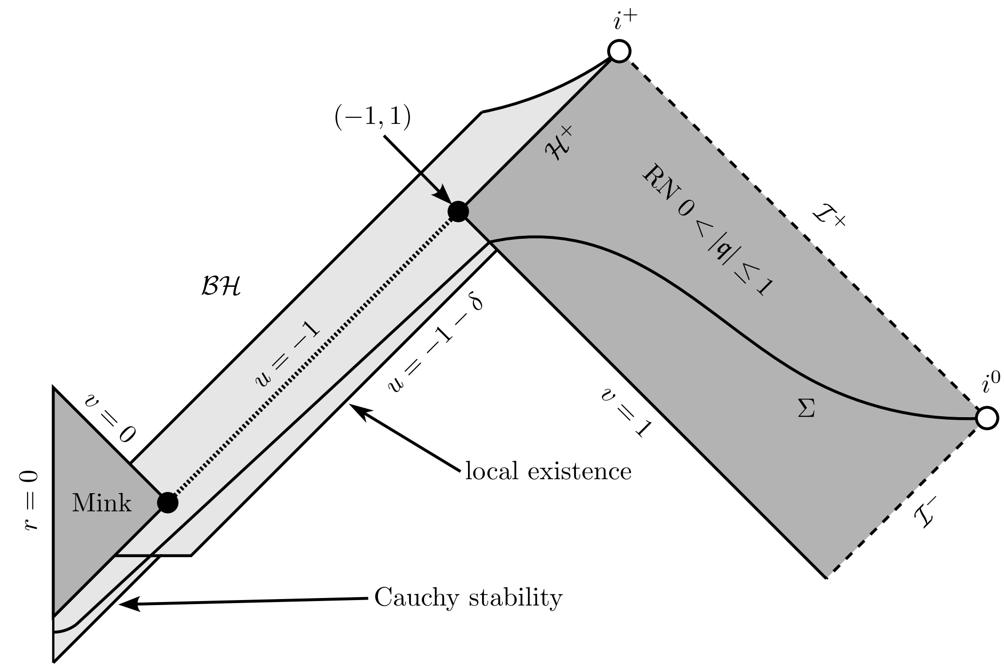

Figure 5.3: Penrose diagram for the proof of Corollary 5.6.1.

Remark 5.6.2. A similar statement can be made with $\mathfrak{q} = 0$ for the Einstein-scalar field model, using instead Theorem 5.4.1. In that case, there will also be no assumption made on the mass.

Proof. We refer the reader to Fig. 5.3 for a visual guide to the proof. Using Theorem 5.4.2 with regularity index $k+1$ (see footnote below) and Proposition 5.2.4, a portion of Minkowski space

$$t + r \leq \tfrac{1}{2}M,$$
$$t - r \geq -\tfrac{1}{2}M,$$

can be glued to a Reissner–Nordström solution with parameters M and $\mathfrak{q}M$. Note that as depicted, one can solve for a complete future neighborhood of the event horizon, which might not be a complete double null neighborhood.

Since we are in spherical symmetry, standard techniques (see [Chr93, Section 5] or [LOY18, Section 3]) allow the "local existence" region emanating from the Reissner–Nordström portion of the spacetime to be extended all the way up to the center.[3] (In this figure, this region is denoted "Cauchy stability" for reasons that will become clear below.)

We now identify a spacelike curve Σ connecting spacelike infinity i^0 in the exactly Reissner–Nordström region to the center, to the past of the cone $u = -1$. The curve Σ can be chosen so the

[3]The wave equation in spherical symmetry loses one derivative at the center when compared to characteristic data. Therefore, to obtain a globally C^k solution, we take C^{k+1} characteristic data.

induced data on it is asymptotically flat near i^0. For example, it may be taken to be a constant t curve near i^0 in standard coordinates. Furthermore, by having Σ hug the gluing region closely enough, we are guaranteed to have no spherically symmetric antitrapped surfaces on Σ.

Completeness of null infinity $\mathcal{I}^+$ is inherited from the exact Reissner–Nordström solution. By inspecting Fig. 5.3, we see that the null hypersurface C_{-1} is the event horizon $\mathcal{H}^+ = \partial J^-(\mathcal{I}^+)$ of the spacetime and that Σ can be arranged to lie in the domain of outer communication $J^-(\mathcal{I}^+)$. The statement about trapped surfaces follows from Proposition 2.5.3 below.

We now consider the (unique) maximal future globally hyperbolic development $(\mathcal{M}^4, g, F, A, \phi)$ of the induced data $(\Sigma, g_0, k_0, E_0, B_0, \phi_0, \phi_1)$ on Σ. By uniqueness of the MFGHD, it contains the domain of dependence of Σ in the gluing spacetime (and thus all shaded regions to the future of Σ in Fig. 5.3). Therefore, by construction, $(\mathcal{M}^4, g, F, A, \phi)$ has all the properties listed in the statement of Corollary 5.6.1. Note that the property of having no antitrapped symmetry spheres is propagated to the whole development by Raychaudhuri's equation (2.2.6). By Proposition 2.5.3, the spacetime does not contain any nonspherically symmetric antitrapped surfaces either. This concludes the proof. $\square$

The above proof made use of spherical symmetry in the local existence region and the region up to the center. In view of potentially extending our work to the Einstein vacuum equations in the future, we give a second construction of these regions which does not invoke spherical symmetry. First, the "local existence region" can be constructed outside of spherical symmetry by the well-known theorem of Luk [Luk12]. Once such a region has been constructed, we can use the fact that it lies "outside" of a Minkowski region to construct the rest of the spacetime, up to the center, by Cauchy stability:

Lemma 5.6.3. *Let B_{r_0} and B_{r_1} denote the (open) balls of radii $r_0 > 0$ and $r_1 > r_0$ in $\mathbb{R}^3$, respectively. Consider on B_{r_1} data for the Einstein–Maxwell-charged scalar field system corresponding to Minkowski space, $(\delta, 0, 0, 0, 0, 0)$. Let $D \doteq (g_0, k_0, E_0, B_0, \phi_0, \phi_1)$ be a C^k (for $k \in \mathbb{N}$ sufficiently large and not assumed to be spherically symmetric) initial data set for the Einstein–Maxwell-charged scalar field system defined on B_{r_1} which agrees with the Minkowski data set on B_{r_0}. Then the maximal globally hyperbolic development of D contains the Minkowski cone over B_{r_0} "in its interior" in the following sense:*

There exists an $\varepsilon > 0$ and a development (g, F, A, ϕ) of the data D on $K_{r_0+\varepsilon} \doteq \{t + r < r_0 + \varepsilon\} \cap \{t \geq 0\} \subset \mathbb{R}^{3+1}$ so that the development of the Minkowski portion of the data is defined on $K_{r_0} \doteq \{t + r < r_0\} \cap \{t \geq 0\}$ and is the Minkowski metric in those coordinates.

Proof. Since this is a standard Cauchy stability argument we merely sketch the proof. For $0 < \varepsilon < \frac{r_1 - r_0}{2}$, let θ_ε be a cutoff function which is equal to one on $B_{r_0+\varepsilon}$ and vanishes outside $B_{r_0+2\varepsilon}$. On B_{r_1}, we consider the "initial data set"

$$D_\varepsilon \doteq (\theta_\varepsilon g_0 + (1-\theta_\varepsilon)\delta, \theta_\varepsilon k_0, \theta_\varepsilon E_0, \theta_\varepsilon B_0, \theta_\varepsilon \phi_0, \theta_\varepsilon \phi_1).$$

This does not solve the constraints everywhere, but it does solve them on $B_{r_0+\varepsilon}$, where it equals D. We assume that $k \geq 5$ and show that D_ε is $O(\varepsilon)$-close to the Minkowski data set in H^4. Then Cauchy stability for the reduced Einstein equations (in harmonic coordinates) will show that a solution to the reduced equations with data D_ε exists on $K_{r_0+2\varepsilon}$ for ε sufficiently small. By domain of dependence arguments, a genuine solution will then exist on a smaller domain which still contains the entirety of $\overline{K_{r_0}}$ in its interior.

To show that D_ε is close to Minkowski data we must check it componentwise. For brevity, we only check $\theta_\varepsilon k_0$. Note first that

$$\|\theta_\varepsilon k_0\|_{H^4} \lesssim \|\theta_\varepsilon k_0\|_{C^4}.$$

Now since k_0 vanishes on B_{r_0} and is at least C^5, Taylor's theorem implies

$$|\partial_r^i \nabla\!\!\!/^j k_0| \lesssim \max\{0, r - r_0\}^{5-i-j},$$

if $0 \leq i + j \leq 5$. In the region where either θ_ε or $\partial_r \theta_i$ are nonvanishing, $\max\{0, r - r_0\} \lesssim \varepsilon$. It follows that

$$\|\theta_\varepsilon k_0\|_{H^4} \lesssim \sum_{0 \leq i+j \leq 4} \sup_{B_{r_1}} |\partial_r^i \nabla\!\!\!/^j (\theta_\varepsilon k_0)| \lesssim \varepsilon,$$

which proves the claim and hence the lemma. □

5.6.2 Construction of counterexample to the third law

In this section we prove Theorem 1.1.11 with an analogous approach as in the proof of Corollary 5.6.1. We first restate the result in more detail.

Theorem 5.6.4. *For any $k \in \mathbb{N}$ and $\mathfrak{e} \in \mathbb{R}\backslash\{0\}$, there exist asymptotically flat, spherically symmetric Cauchy data $(\Sigma, g_0, k_0, E_0, B_0, \phi_0, \phi_1)$, with $\Sigma \cong \mathbb{R}^3$ and a regular center, for the EMCSF system such that the maximal future globally hyperbolic development $(\mathcal{M}^4, g, F, A, \phi)$ has the following properties:*

- *All dynamical quantities are at least C^k-regular.*

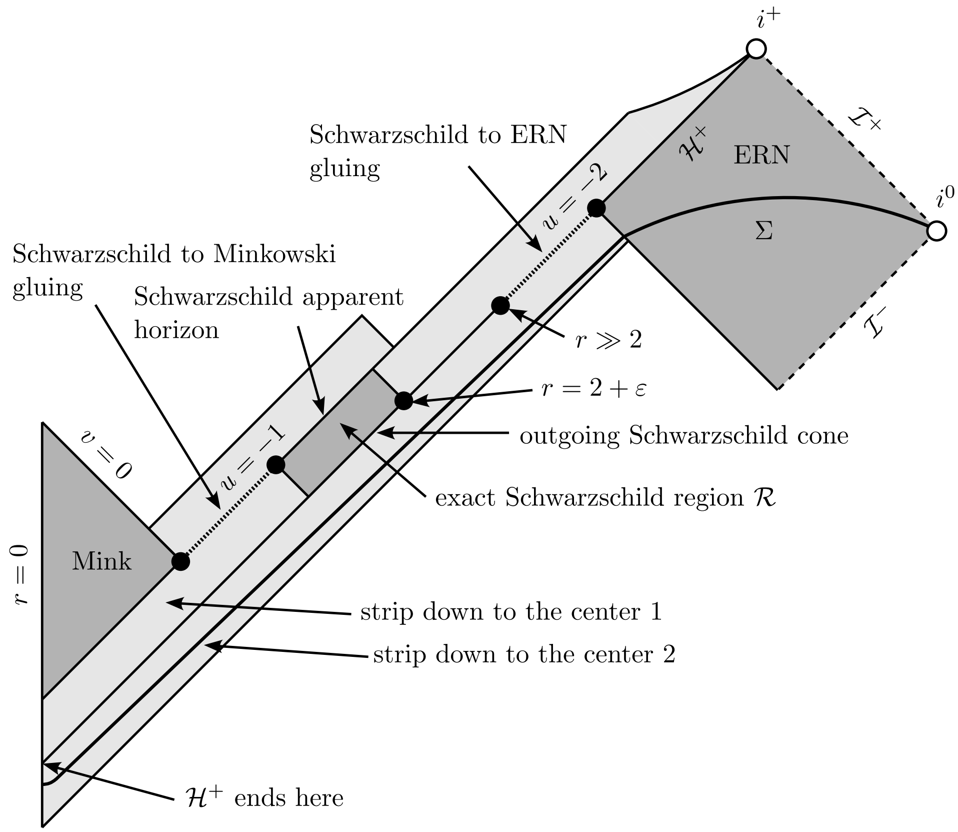

Figure 5.4: Penrose diagram for the proof of Theorem 5.6.4.

- *The spacetime and Cauchy data satisfy all the conclusions of Corollary 5.6.1 with* $\mathfrak{q} = 1$ *and final mass* $M_f \geq M_0(1, \mathfrak{e}, k) + 8$.

- *The spacetime contains a double null rectangle of the form* $\mathfrak{R} \doteq \{-2 \leq u \leq -1\} \cap \{1 \leq v \leq 2\}$ *which is isometric to a double null rectangle in a Schwarzschild spacetime of mass* 1.

- *The cone* $\{u = -1\} \cap \mathfrak{R}$ *lies in the outermost apparent horizon* $\mathcal{A}'$ *of the spacetime and is isometric to an appropriate portion of the* $r = 2$ *hypersurface in the Schwarzschild spacetime of mass* 1.

- *The outermost apparent horizon* $\mathcal{A}'$ *is disconnected.*

- *The spacetime contains trapped surfaces in the black hole region, for all arbitrarily late advanced time. More precisely, for every symmetry sphere* $S_{u,v} \subset \mathcal{H}^+$, $J^+(S_{u,v})$ *contains a trapped sphere.*

- *There exists a neighborhood* $\mathcal{U}$ *of* $\mathcal{H}^+$ *in* $\mathcal{M}$ *such that there are no trapped surfaces* $S \subset \mathcal{U}$.

Proof. We refer to Fig. 5.4 for a Penrose diagram illustrating the proof. The proof begins as the proof of Corollary 5.6.1 (recall also Proposition 5.2.4), by gluing a Minkowski cone to a Schwarzschild event horizon of unit mass along $\{u=-1\}$. Then, attach a double null rectangle $\mathfrak{R}$ of Schwarzschild along the hypersurface $r=2$, as in Corollary 5.6.1, but stop after a finite advanced time $v=2$. Now place $u=-2$ so that

$$\sup_{\{u=-2\}\cap\mathfrak{R}} r = 2+\varepsilon \le 3.$$

For ε sufficiently small, the first strip down to the center can be constructed as in the proof of Corollary 5.6.1. Now let $M_f \ge M_0+8$ and extend the cone $u=-2$ to the future with trivial scalar field until $r=\frac{1}{2}(M_0+8)\gg 3$. Then using Theorem 5.4.2, extremal Reissner–Nordström of mass M_f can be attached. We again solve backward up to the center as in Corollary 5.6.1 and have now constructed the spacetime depicted in Fig. 5.4.

As in the proof of Corollary 5.6.1, we again find an asymptotically flat spacelike curve Σ connecting i^0 with the center and lying entirely in $J^-(\mathcal{I}^+)$. The maximal future globally hyperbolic development $(\mathcal{M},g,F,A,\phi)$ of the induced data on Σ contains the domain of dependence of Σ in the spacetime constructed above (and thus all shaded regions to the future of Σ in Fig. 5.4) and satisfies all the conclusions of Corollary 5.6.1 with $\mathfrak{q}=1$ and final mass $M_f \ge M_0(1,\mathfrak{e},k)+8$. By construction, $\mathcal{M}$ contains the double null rectangle $\mathfrak{R}$ which satisfies the stated properties. Further, the cone $\{u=-1\}\cap\mathfrak{R}$ lies in the apparent horizon $\mathcal{A}$ of $(\mathcal{M},g)$ and $\{u=-1\}\cap\mathfrak{R}$ is isometric to an appropriate portion of the $r=2$ hypersurface in the Schwarzschild spacetime of mass 1.

We readily see that $(\mathcal{M},g)$ contains trapped surfaces in any (future) neighborhood of $\{u=-1\}\cap\mathfrak{R}$ as $\partial_v r=0$ along $\{u=-1\}\cap\mathfrak{R}$ and (2.2.11) evaluated on $\{u=-1\}\cap\mathfrak{R}$ gives

$$\partial_u(r\partial_v r) = -\frac{\Omega^2}{4}.$$

To prove that trapped surfaces exist for arbitrarily late advanced time, we invoke the general boundary characterization of [Kom13]. If the $r=0$ singularity $\mathcal{S}$ is empty, then the outgoing cone starting from one of these trapped spheres terminates on the Cauchy horizon $\mathcal{CH}^+$ and the claim is clearly true by Raychaudhuri's equation (2.2.7). If $\mathcal{S}$ is nonempty, then every ougoing null cone which terminates on $\mathcal{S}$ is eventually trapped since r extends continuously by zero on $\mathcal{S}$. Furthermore, $\mathcal{S}$ terminates at the Cauchy horizon $\mathcal{CH}^+$ or future timelike infinity i^+, so the claim is also true in this case.

We now show that there exists a neighborhood $\mathcal{U}$ of $\mathcal{H}^+$ in $\mathcal{M}$ which does not contain *spherically* *symmetric* trapped surfaces. It suffices to show that there is a neighborhood $\mathcal{V}$ of $\mathcal{H}^+$ in $\mathcal{Q}$ such that $\partial_v r > 0$ on $\mathcal{V} \setminus \mathcal{H}^+$, where we use the same symbol for the event horizon in $\mathcal{M}$ and $\mathcal{Q}$. Let $p \in \mathcal{H}^+$ be any sphere after the final gluing sphere, see Fig. 5.4. Then $r(p) = Q(p) = M_f$, $\partial_v r(p) = 0$, and $\phi(p) = 0$. Reparametrize the double null gauge so that $\Omega \equiv 1$ on the ingoing cone $\underline{C}$ passing through p. By the wave equation for the radius (2.2.4),

$$\partial_u \partial_v r(p) = -\frac{1}{4M_f} + \frac{M_f^2}{4M_f^3} = 0.$$

Differentiating (2.2.4) in u, we find

$$\partial_u^2 \partial_v r = \frac{\partial_u r}{4r^2} - \partial_u(\partial_u \log r)\partial_v r - (\partial_u \log r)\partial_u \partial_v r - \frac{3Q^2 \partial_u r}{4r^4} + \frac{Q\partial_u Q}{2r^3}.$$

Evaluating at p, we find $\partial_u Q(p) = 0$ by Maxwell's equation (2.2.8), so we have

$$\partial_u^2 \partial_v r(p) = \frac{\partial_u r(p)}{4M_f^2} - \frac{3M_f^2 \partial_u r(p)}{4M_f^4} = -\frac{2\partial_u r(p)}{M_f^2} > 0.$$

Therefore, $\partial_v r$ becomes immediately positive for all points along $\underline{C}$ sufficiently close to the event horizon but not on it (see also Fig. 1.3).[4]

By the monotonicity of Raychaudhuri's equation (2.2.7) and since $p \in \mathcal{H}^+$ after the final gluing sphere was arbitrary, this shows that there exists a neighborhood $\mathcal{V}$ of $\mathcal{H}^+$ contained in $\mathcal{Q}$ that does not contain trapped symmetry spheres except for $\mathcal{H}^+$ itself. That there are also no nonspherically symmetric trapped surfaces in $\mathcal{U} \doteq \mathcal{V} \times S^2$ now follows immediately from Proposition 2.5.1 below.

The claim about the disconnectedness of the outermost apparent horizon $\mathcal{A}'$ now follows from the fact that $\mathcal{A}' \cap \mathcal{H}^+$ is one connected component of $\mathcal{A}'$ which does not contain $\{u = -1\} \cap \mathfrak{R} \subset \mathcal{A}'$. This concludes the proof. □

5.6.3 Construction of collapse to Reissner–Nordström with piece of Cauchy horizon

In this section, we show that a mild modification of the proof of Corollary 5.6.1 allows us to construct examples of gravitational collapse such that the black hole region admits a piece of future boundary which is a Cauchy horizon which is isometric to a subextremal or extremal Reissner–Nordström Cauchy horizon.

[4] This calculation is related to the discussion in Section 1.1.3.4 above and Section 5.A below. In fact, we have effectively just proved the claim in Remark 5.A.2.

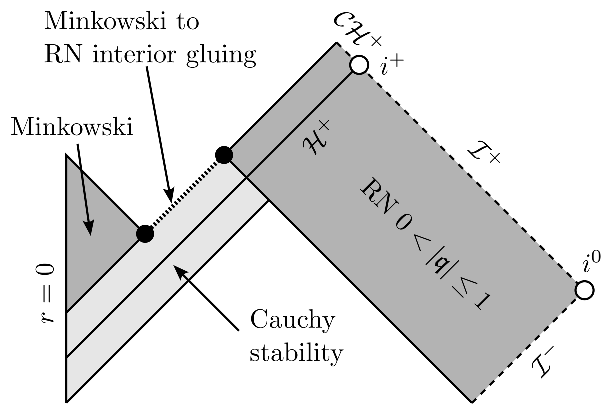

Figure 5.5: Penrose diagram depicting the proof of Corollary 5.6.5.

Corollary 5.6.5. *For any $k \in \mathbb{N}$, $\mathfrak{q} \in [-1,1] \setminus \{0\}$, and $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$, let $M_0(k, \mathfrak{q}, \mathfrak{e}, 1/2)$ be as in Theorem 5.4.4. Then for any $M \geq M_0$ there exist asymptotically flat, spherically symmetric Cauchy data $(\Sigma, g_0, k_0, E_0, B_0, \phi_0, \phi_1)$, with $\Sigma \cong \mathbb{R}^3$ and a regular center, for the EMCSF system such that the maximal future globally hyperbolic development $(\mathcal{M}^4, g, F, A, \phi)$ has the following properties:*

- *All dynamical quantities are at least C^k-regular.*
- *The spacetime and Cauchy data satisfy all the conclusions of Corollary 5.6.1.*
- *The black hole region contains an isometrically embedded portion of a Reissner–Nordström Cauchy horizon neighborhood with parameters M and $\mathfrak{q}M$, in particular $\mathcal{CH}^+ \neq \emptyset$.*

Proof. The proof is completely analogous to the proof of Corollary 5.6.1. We apply the gluing construction of Theorem 5.4.4 to glue a sphere in Minkowski space to a Reissner–Nordström interior sphere with radius $R_f < r_+$ and $r_+ - R_f$ small. Indeed, this can be achieved by setting $\mathfrak{r} = \frac{1}{2}$ in Theorem 5.4.4 as then $\frac{1}{2} M_f (1 + \mathfrak{r}) \mathfrak{q}^2 \leq \frac{3}{4} M_f < M_f \leq r_+$. We then apply the local existence and Cauchy stability argument as in the proof of Corollary 5.6.1. We note that the u-width of the local existence and Cauchy stability argument remains uniform as $R_f \to r_+$ so by choosing R_f sufficiently close to r_+, we guarantee that we find a Cauchy hypersurface Σ which does not intersect the event horizon. We refer to Fig. 5.5 for the Penrose diagram explaining the proof. □

Remark 5.6.6. As in Remark 5.6.2, we note that a similar statement with a piece of Schwarzschild interior including the $\{r = 0\}$ singularity can be made with $\mathfrak{q} = 0$.

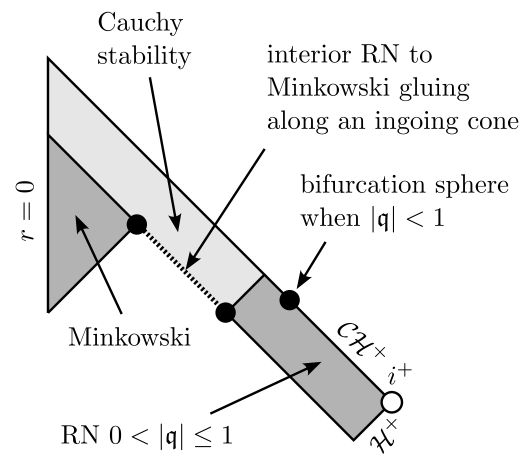

Figure 5.6: Penrose diagram depicting the proof of Corollary 5.6.7.

5.6.4 Construction of black hole interior for which the Cauchy horizon closes off spacetime

We now give our construction of a spacetime for which the Cauchy horizon closes off the black hole region.

Corollary 5.6.7. *For any $k \in \mathbb{N}$, $\mathfrak{q} \in [-1, 1] \setminus \{0\}$, $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$, let $\tilde{M}_0(k, \mathfrak{q}, \mathfrak{e}, \mathfrak{q}^2/4)$ be as in Theorem 5.4.7. Then for any $M \geq \tilde{M}_0$ there exist asymptotically flat, spherically symmetric Cauchy data $(\Sigma, g_0, k_0, E_0, B_0, \phi_0, \phi_1)$, with $\Sigma \cong \mathbb{R}^3$ and a regular center, for the EMCSF system such that the maximal future globally hyperbolic development $(\mathcal{M}^4, g, F, A, \phi)$ has the following properties:*

- *All dynamical quantities are at least C^k-regular.*
- *The spacetime does not contain antitrapped surfaces.*
- *The black hole region is nonempty, $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+) \neq \emptyset$.*
- *The future boundary of the black hole region is a C^k-regular Cauchy horizon $\mathcal{CH}^+$ which closes off spacetime, i.e., $\mathcal{N} \cup \mathcal{S} = \emptyset$ in Fig. 2.1.*
- *The exterior region is isometric to a Reissner–Nordström exterior with mass M and charge $\mathfrak{q}M$. In particular, future null infinity $\mathcal{I}^+$ is complete.*

Proof. Analogous to the proof of Corollary 5.6.5 we glue a Reissner–Nordström interior sphere with $R_f < r_-$ and $r_- - R_f$ small to a sphere in Minkowski space along an *ingoing* cone using Theorem 5.4.7. We can choose R_f arbitrarily close to r_- in Theorem 5.4.7 by setting $\mathfrak{r} = \mathfrak{q}^2/4$.

Indeed, in this case

$$
\begin{aligned}
r_- - \frac{M_f}{2}\left(1+\frac{\mathfrak{q}^2}{4}\right)\mathfrak{q}^2 = M_f\left(1-\sqrt{1-\mathfrak{q}^2}\right) - \frac{M_f}{2}\left(1+\frac{\mathfrak{q}^2}{4}\right)\mathfrak{q}^2 \\
= M_f\left(1-\sqrt{1-\mathfrak{q}^2}-\frac{\mathfrak{q}^2}{2}-\frac{\mathfrak{q}^4}{8}\right) \geq M_f\frac{\mathfrak{q}^6}{16},
\end{aligned}
$$

where in the last step we used the Taylor expansion of $\sqrt{1-\mathfrak{q}^2}$ around $\mathfrak{q}=0$. The rest of the proof is now analogous to Corollary 5.6.5 and can be read off from Fig. 5.6. We note that an isometric copy of the Reissner–Nordström exterior can be attached to the past of $\mathcal{H}^+$ in Fig. 5.6. □

5.A An isolated extremal horizon with nearby trapped surfaces

In this appendix we show that, in the context of the dominant energy condition, there is no local mechanism forcing a stationary extremal Killing horizon to have no trapped surfaces "just inside" of the horizon. We also refer back to Section 1.1.3.4.

Proposition 5.A.1. *There exists a C^∞ spherically symmetric spacetime $(\mathcal{M}^4, g)$ with a complete null hypersurface $H \subset \mathcal{M}$ and a Killing vector field T with the following properties. The Killing field T is spherically symmetric, timelike in $I^-(H)$, spacelike in $I^+(H)$, null and tangent along H, where it also satisfies $\nabla_T T = 0$, i.e., its integral curves are affinely parametrized null generators of H. Furthermore, $(\mathcal{M}, g)$ contains no antitrapped symmetry spheres, i.e., $\partial_u r < 0$, and satisfies the dominant energy condition. Therefore, H is an extremal Killing horizon and $I^+(H)$ is foliated by trapped symmetry spheres.*

We recall that a spacetime $(\mathcal{M}, g)$ satisfies the *dominant energy condition* if for all future directed causal vectors $X \in T\mathcal{M}$, $-G(\cdot, X)^\sharp$ is future directed causal or zero. Here G denotes the Einstein tensor of g,

$$G(g) \doteq \mathrm{Ric}(g) - \tfrac{1}{2}R(g)g.$$

Proof. The spacetime is given by the spherically symmetric ansatz

$$
\begin{aligned}
\mathcal{M} &= \mathcal{Q} \times S^2 \\
g &= g_{\mathcal{Q}} + r^2 g_{S^2},
\end{aligned}
$$

where

$$\mathcal{Q} = \{(t,u) \in \mathbb{R}^2 : t \in \mathbb{R}, -\varepsilon < u < \varepsilon\}$$

for ε to be chosen later, and $r = r(u)$. Let $f = f(u)$ and set

$$g_{\mathcal{Q}} = f dt^2 - 2dtdu.$$

The vector field $\underline{L} = \partial_u$ is geodesic and null and we declare it to be future directed. The Killing vector field $T = \partial_t$ satisfies $g(T,T) = f$. Letting $f(u) = u^2 F(u)$ for a smooth function $F(u)$ makes $H = \{u = 0\}$ an extremal Killing horizon and ∂_t is future directed where it is causal. The conjugate null vector to $\underline{L}$ is $L = \partial_t + \frac{1}{2} f \partial_u$ such that $g(\underline{L}, L) = -1$. The symmetry spheres $S_{t_0,u_0} = \{t = t_0\} \cap \{u = u_0\}$ are trapped if

$$Lr < 0$$
$$\underline{L}r < 0,$$

which can be more simply written as

$$f(u) r'(u) < 0$$
$$r'(u) < 0.$$

From this we see that $r'(u) < 0$ implies no antitrapped spheres of symmetry and $f(u) < 0$ for $u < 0$ and $f(u) > 0$ for $u > 0$ implies the symmetry spheres to the past (respectively, future) of H are untrapped (respectively, trapped). This also makes T timelike to the past of H. Since we require $f(u) = u^2 F(u)$ but also that f changes sign, we in fact have $f(u) = u^3 \tilde{F}(u)$.

We will now see which restrictions on f, r, and ε enforce the dominant energy condition. The Einstein tensor of g is given by

$$G = -\theta g_{\mathcal{Q}} - \frac{2r''}{r} du^2 + \zeta r^2 g_{S^2}, \tag{5.A.1}$$

where

$$\theta \doteq \frac{1 + (u')^2 f + r f r' + 2 f r r''}{r^2}, \quad \zeta \doteq -\tfrac{1}{2} f'' - \frac{r'}{r} f' - \frac{r''}{r} f.$$

For $f(u) = u^3 \tilde{F}(u)$ and $r(u)$ fixed and $\varepsilon > 0$ sufficiently small, we have $\theta(u) > 0$ and $|\zeta(u)| \ll \theta(u)$ for $|u| < \varepsilon$.

Let X be a future causal vector, that is

$$g_{\mathcal{Q}}(X,X) + r^2 g_{S^2}(X,X) \leq 0, \quad g_{\mathcal{Q}}(\underline{L}+L, X) < 0. \tag{5.A.2}$$

To show that $-G(\cdot, X)^\sharp$ is causal or zero, it suffices to show that

$$g_{\mu\nu} G^\mu{}_\rho G^\nu{}_\sigma X^\rho X^\sigma \leq 0. \tag{5.A.3}$$

To simplify the calculation, we assume r'' vanishes identically and then the left-hand side of (5.A.3), using (5.A.1) and (5.A.2), can be estimated as

$$g_{\mu\nu} G^\mu{}_\rho G^\nu{}_\sigma X^\rho X^\sigma = \theta^2 g_{\mathcal{Q}}(X,X) + \zeta^2 r^2 g_{S^2}(X,X) \leq (\zeta^2 - \theta^2) r^2 g_{S^2}(X,X).$$

Since $\zeta^2 - \theta^2 \leq 0$, this proves that $-G(\cdot, X)^\sharp$ is causal. To show that $-G(\cdot, X)^\sharp$ is future directed we compute using (5.A.2)

$$g(\underline{L}+L, -G(\cdot, X)^\sharp) = -G(\underline{L}+L, X) = \theta g_{\mathcal{Q}}(\underline{L}+L, X) < 0.$$

Finally, an explicit example of a metric satisfying all of our conditions is

$$g = u^3 dt^2 - 2dtdu + (1-u)^2 g_{S^2}. \qquad \square$$

Remark 5.A.2. Extremal Reissner–Nordström has $f(u) \sim -u^2$. One might say that an extremal horizon constructed in the above manner with $f(u)$ vanishing faster than u^2 is a *degenerate extremal horizon*.

Chapter 6

Characteristic gluing in vacuum and formation of very slowly rotating Kerr black holes

In this chapter, we prove that Minkowski space can be characteristically glued to any slowly rotating Kerr event horizon, see Section 1.1.5.1 and Section 4.5.

6.1 Spacetimes in double null gauge

We first briefly recall the basic notion of *double null gauge* for the Einstein vacuum euqations outside of spherical symmetry [Chr91a; Chr09].

6.1.1 Double null gauge

Let $\mathcal{W} \subset \mathbb{R}^2_{u,v}$ be a domain and define $\mathcal{M}^{3+1} \doteq \mathcal{W} \times S^2$. Denote $S_{u,v} \doteq \{(u,v)\} \times S^2 \subset \mathcal{M}$. The distinguished foliation of $\mathcal{M}$ by these spheres carries a tangent bundle TS and cotangent bundle $T^*S \doteq (TS)^*$. An S-*tensor (field)* is a section of a vector bundle consisting of tensor products of TS and T^*S. Let $\not{g}$ be a positive-definite $(0,2)$ S-tensor field, let Ω^2 be a positive function on $\mathcal{M}$, and b be an S-vector field. Under these assumptions, the formula

$$g = -4\Omega^2\, du\, dv + \not{g}_{AB}(d\vartheta^A - b^A\, dv)(d\vartheta^B - b^B\, dv)$$

defines a Lorentzian metric on $\mathcal{M}$, where $(\vartheta^1, \vartheta^2)$ are arbitrary local coordinates on S^2 and $\not{g}_{AB}$ (resp., b^A) are the components of $\not{g}$ (resp., b) relative to this coordinate basis. The coordinate functions u and v satisfy the eikonal equation, i.e.,

$$g^{\mu\nu}\partial_\mu u \partial_\nu u = 0 \quad \text{and} \quad g^{\mu\nu}\partial_\mu v \partial_\nu v = 0.$$

Consequently, the hypersurfaces $C_u \doteq \{u = \text{const.}\}$ and $\underline{C}_v \doteq \{v = \text{const.}\}$ are null hypersurfaces. We time orient $(\mathcal{M}, g)$ by declaring $\partial_u + \partial_v + b^A \partial_{\vartheta^A}$ to be future-directed.

The vector fields

$$L' \doteq -2(du)^\sharp \quad \text{and} \quad \underline{L}' \doteq -2(dv)^\sharp$$

are future-directed null geodesic vector fields. We set

$$L \doteq \Omega^2 L' \quad \text{and} \quad \underline{L} \doteq \Omega^2 \underline{L}',$$

which then satisfy

$$Lu = 0, \quad Lv = 1,$$
$$\underline{L}u = 1, \quad \underline{L}v = 0.$$

Finally, we set

$$e_4 \doteq \Omega L', \quad e_3 \doteq \Omega \underline{L}'.$$

Given arbitrary coordinates ϑ^A on S^2 and defining $e_A \doteq \partial_{\vartheta^A}$, the quadruple $\{e_1, e_2, e_3, e_4\}$ is called a (normalized) *null frame*, which satisfies

$$g(e_A, e_3) = g(e_A, e_4) = 0, \quad g(e_3, e_4) = -2, \quad g(e_3, e_3) = g(e_4, e_4) = 0.$$

6.1.2 Algebra and calculus of S-tensors

Let $(\mathcal{M}, g)$ be a spacetime equipped with a double null gauge as above. For vector fields on $\mathcal{M}$, we define the orthogonal projection to S vector fields by

$$\Pi : T\mathcal{M} \to TS, \quad \Pi X \doteq X + \tfrac{1}{2} g(X, e_3) e_4 + \tfrac{1}{2} g(X, e_4) e_3$$

which we extend componentwise to contravariant tensors of higher rank. We note that $\Pi \circ i = \mathrm{id}$ on TS, where $i\colon TS \subset T\mathcal{M}$ is the natural inclusion. By duality, this defines a "promotion" operator $\Pi^*\colon T^*S \to T^*\mathcal{M}$ which extends componentwise to covariant S-tensors and satisfies $i^* \circ \Pi^* = \mathrm{id}$ on T^*S.

We now define projected Lie derivatives $\not{\mathcal{L}}_L$ and $\not{\mathcal{L}}_{\underline{L}}$ on S-tensors. If X is an S-vector field, then

$$\not{\mathcal{L}}_L X \doteq \mathcal{L}_L X, \quad \not{\mathcal{L}}_{\underline{L}} X \doteq \mathcal{L}_{\underline{L}} X$$

are already S-vector fields. If ξ is an S-1-form, then

$$\not{\mathcal{L}}_L \xi \doteq i^* \mathcal{L}_L(\Pi^* \xi) = \mathcal{L}_L(\xi \circ \Pi)|_{TS}, \quad \not{\mathcal{L}}_{\underline{L}} \xi \doteq i^* \mathcal{L}_{\underline{L}}(\Pi^* \xi) = \mathcal{L}_{\underline{L}}(\xi \circ \Pi)|_{TS},$$

where we have explicitly written the "promotion" operation which will be consistently omitted in the sequel. The operation is extended to general S-tensor fields via the Leibniz rule. As a shorthand, we write

$$D \doteq \not{\mathcal{L}}_L, \quad \underline{D} \doteq \not{\mathcal{L}}_{\underline{L}}.$$

The symbol $\not\nabla$ acts on functions and S-vector fields as the induced covariant derivative on the spheres and is extended to general S-tensors by the Leibniz rule.

We will frequently make use of the following notation: Let ξ, η be S-1-forms and θ, ϕ symmetric covariant S-2-tensor fields. We then define

$$\begin{aligned}
(\xi \mathbin{\hat{\otimes}} \eta)_{AB} &\doteq \xi_A \eta_B + \xi_B \eta_A - (\xi \cdot \eta) \not{g}_{AB} \\
(\not\nabla \mathbin{\hat{\otimes}} \xi)_{AB} &\doteq \not\nabla_A \xi_B + \not\nabla_B \xi_A - (\mathrm{d}\not{\mathrm{i}}\mathrm{v}\, \xi) \not{g}_{AB} \\
\mathrm{d}\not{\mathrm{i}}\mathrm{v}\, \xi &\doteq \not{g}^{AB} \not\nabla_A \xi_B \\
\mathrm{r}\not{\mathrm{o}}\mathrm{t}\, \xi &\doteq \not{\epsilon}^{AB} \not\nabla_A \xi_B \\
({}^*\xi)_A &\doteq \not{\epsilon}_{AB} \not{g}^{BC} \xi_C \\
\hat{\theta}_{AB} &\doteq \theta_{AB} - \tfrac{1}{2} \operatorname{tr} \theta \, \not{g}_{AB} \\
\theta \wedge \phi &\doteq \not{\epsilon}^{AB} \not{g}^{CD} \theta_{AC} \phi_{BD},
\end{aligned}$$

where $\not{\epsilon}$ is the induced volume form on $S_{u,v}$. The notation $\not{g}^{AB}$ denotes the inverse of the induced metric $\not{g}_{AB}$. Indices of S-tensors are raised and lowered with $\not{g}$ and $\not{g}^{-1}$.

6.1.3 Ricci and curvature components

The Ricci components are given by the null second fundamental forms

$$\chi_{AB} \doteq g(\nabla_A e_4, e_B), \quad \underline{\chi}_{AB} \doteq g(\nabla_A e_3, e_B),$$

the torsions

$$\eta_A \doteq -\tfrac{1}{2} g(\nabla_3 e_A, e_4), \quad \underline{\eta}_A \doteq -\tfrac{1}{2} g(\nabla_4 e_A, e_3),$$

and

$$\omega \doteq D \log \Omega, \quad \underline{\omega} \doteq \underline{D} \log \Omega.$$

The 1-form

$$\zeta \doteq \eta - \nabla\!\!\!/ \log \Omega$$

is also referred to as the torsion. Note that

$$\eta + \underline{\eta} = 2 \nabla\!\!\!/ \log \Omega. \tag{6.1.1}$$

The null curvature components are given by

$$\begin{aligned} \alpha_{AB} &\doteq R(e_A, e_4, e_B, e_4), & \underline{\alpha}_{AB} &\doteq R(e_A, e_3, e_B, e_3), \\ \beta_A &\doteq \tfrac{1}{2} R(e_A, e_4, e_3, e_4), & \underline{\beta}_A &\doteq \tfrac{1}{2} R(e_A, e_3, e_3, e_4), \\ \rho &\doteq \tfrac{1}{4} R(e_4, e_3, e_4, e_3), & \sigma &\doteq \tfrac{1}{4} {}^*R(e_4, e_3, e_4, e_3), \end{aligned}$$

where $R(W, Z, X, Y) = g(R(X, Y)Z, W)$ is the Riemann tensor.

6.1.4 Normalized sphere data determined by a geometric sphere

For the notion of sphere data used here, see Section 6.3.1.1.

Lemma 6.1.1. *Let* $(\mathcal{M}^4, g)$ *be a spacetime satisfying the Einstein vacuum equations* (1.1.6) *and* $i : S^2 \to \mathcal{M}$ *an embedding with spacelike image* $S \doteq i(S^2)$*. Let* $\underline{L}$ *be a null vector field along* S *which is normal to* S*. Then for any* $m \geq 0$ *there exists a unique associated* $C^2_u C^{2+m}_v$ *sphere data set* $x[g, i, \underline{L}]$ *such that* $g\!\!\!/ = i^* g$*,* $\Omega^2 = 1$*, and* $\omega = D\omega = \cdots = D^{m+1}\omega = \underline{\omega} = \underline{D\omega} = 0$*.*

The sphere data $x[g, i, \underline{L}]$ *depends smoothly on* $(g, i, \underline{L})$ *in the natural way.*

We say that $x[g, i, \underline{L}]$ is *generated* by $(g, i, \underline{L})$. If $\psi : S^2 \to S^2$ is a diffeomorphism, then $x[g, i \circ \psi, \underline{L}]$

is related to $x[g, i, \underline{L}]$ by a sphere diffeomorphism as in Definition 6.3.4.

Proof. The geometric sphere S is identified with the round sphere by i, which endows S with a choice of round metric γ. The dual null vector field L is uniquely determined by the requirement that $L \perp TS$ and $g(L, \underline{L}) = -2$. Let $C \cup \underline{C}$ be the (locally defined) bifurcate null hypersurface passing through S such that L is tangent to C and $\underline{C}$ is tangent to $\underline{L}$. Let $\Omega^2 = 1$ identically on $C \cup \underline{C}$. Given this data, there is a unique double null foliation with respect to g covering a neighborhood of S in $\mathcal{M}$. Now $x[g, i, \underline{L}]$ is constructed by computing the corresponding quantities in this double null foliation and taking the values at S. □

6.2 The Einstein equations in double null gauge

We now assume that the spacetime metric g satisfies the Einstein vacuum equations (1.1.6). In double null gauge, the Einstein equations are equivalent to the *null structure equations* (with the Ricci coefficients on the left-hand side) and the *Bianchi equations* (with the curvature components on the left-hand side). The Einstein equations (1.1.6) imply

$$\operatorname{tr}\alpha = 0, \qquad \operatorname{tr}\underline{\alpha} = 0. \tag{6.2.1}$$

6.2.1 The null structure equations

First variation formulas:

$$D\not{g} = 2\Omega\chi = 2\Omega\hat{\chi} + \Omega \operatorname{tr}\chi \not{g} \tag{6.2.2}$$

$$\underline{D}\not{g} = 2\Omega\underline{\chi} = 2\Omega\hat{\underline{\chi}} + 2\Omega \operatorname{tr}\underline{\chi} \not{g} \tag{6.2.3}$$

Raychaudhuri's equations:

$$D \operatorname{tr}\chi + \tfrac{1}{2}\Omega(\operatorname{tr}\chi)^2 - \omega \operatorname{tr}\chi = -\Omega|\hat{\chi}|^2 \tag{6.2.4}$$

$$\underline{D} \operatorname{tr}\underline{\chi} + \tfrac{1}{2}\Omega(\operatorname{tr}\underline{\chi})^2 - \underline{\omega} \operatorname{tr}\underline{\chi} = -\Omega|\hat{\underline{\chi}}|^2 \tag{6.2.5}$$

Transport equations for Ricci components:

$$D\hat{\chi} = \Omega|\hat{\chi}|^2 \not{g} + \omega\hat{\chi} - \Omega\alpha \tag{6.2.6}$$

$$\underline{D}\hat{\underline{\chi}} = \Omega|\hat{\underline{\chi}}|^2 \not{g} + \underline{\omega}\hat{\underline{\chi}} - \Omega\underline{\alpha} \tag{6.2.7}$$

$$D\eta = \Omega(\chi \cdot \underline{\eta} - \beta) \tag{6.2.8}$$

$$\underline{D}\underline{\eta} = \Omega(\underline{\chi} \cdot \eta + \underline{\beta}) \tag{6.2.9}$$

$$D\underline{\omega} = \Omega^2(2\eta \cdot \underline{\eta} - |\eta|^2 - \rho) \tag{6.2.10}$$

$$\underline{D}\omega = \Omega^2(2\eta \cdot \underline{\eta} - |\eta|^2 - \rho) \tag{6.2.11}$$

$$D\underline{\eta} = -\Omega(\chi \cdot \underline{\eta} - \beta) + 2\nabla\!\!\!/\,\omega \tag{6.2.12}$$

$$\underline{D}\eta = -\Omega(\underline{\chi} \cdot \eta - \beta) + 2\nabla\!\!\!/\,\omega \tag{6.2.13}$$

$$D(\Omega \operatorname{tr} \underline{\chi}) = 2\Omega^2 \operatorname{d\!\!/iv} \underline{\eta} + 2\Omega^2|\underline{\eta}|^2 - \Omega^2(\hat{\chi}, \hat{\underline{\chi}}) - \tfrac{1}{2}\Omega^2 \operatorname{tr}\chi \operatorname{tr}\underline{\chi} + 2\Omega^2\rho \tag{6.2.14}$$

$$\underline{D}(\Omega \operatorname{tr} \chi) = 2\Omega^2 \operatorname{d\!\!/iv} \eta + 2\Omega^2|\eta|^2 - \Omega^2(\hat{\chi}, \hat{\underline{\chi}}) - \tfrac{1}{2}\Omega^2 \operatorname{tr}\chi \operatorname{tr}\underline{\chi} + 2\Omega^2\rho \tag{6.2.15}$$

$$D(\Omega\hat{\underline{\chi}}) = \Omega^2\left((\hat{\chi}, \hat{\underline{\chi}}) + \tfrac{1}{2}\operatorname{tr}\chi\,\hat{\underline{\chi}} + \nabla\!\!\!/\,\hat{\otimes}\,\underline{\eta} + \underline{\eta}\,\hat{\otimes}\,\underline{\eta} - \tfrac{1}{2}\operatorname{tr}\underline{\chi}\,\hat{\chi}\right) \tag{6.2.16}$$

$$\underline{D}(\Omega\hat{\chi}) = \Omega^2\left((\hat{\chi}, \hat{\underline{\chi}}) + \tfrac{1}{2}\operatorname{tr}\underline{\chi}\,\hat{\chi} + \nabla\!\!\!/\,\hat{\otimes}\,\eta + \eta\,\hat{\otimes}\,\eta - \tfrac{1}{2}\operatorname{tr}\chi\,\hat{\underline{\chi}}\right) \tag{6.2.17}$$

Gauss equation:

$$K + \tfrac{1}{4}\operatorname{tr}\chi \operatorname{tr}\underline{\chi} - \tfrac{1}{2}(\hat{\chi}, \hat{\underline{\chi}}) = -\rho \tag{6.2.18}$$

Codazzi equations:

$$\operatorname{d\!\!/iv} \hat{\chi} - \tfrac{1}{2}\nabla\!\!\!/ \operatorname{tr}\chi + \hat{\chi} \cdot \zeta - \tfrac{1}{2}\operatorname{tr}\chi\,\zeta = -\beta \tag{6.2.19}$$

$$\operatorname{d\!\!/iv} \hat{\underline{\chi}} - \tfrac{1}{2}\nabla\!\!\!/ \operatorname{tr}\underline{\chi} - \hat{\underline{\chi}} \cdot \zeta + \tfrac{1}{2}\operatorname{tr}\underline{\chi}\,\zeta = \underline{\beta} \tag{6.2.20}$$

Curl equations:

$$\operatorname{r\!\!/ot}\eta = -\operatorname{r\!\!/ot}\underline{\eta} = \operatorname{r\!\!/ot}\zeta = -\tfrac{1}{2}\hat{\chi} \wedge \hat{\underline{\chi}} - \sigma \tag{6.2.21}$$

We also require

$$D\underline{D}\underline{\omega} = -12\Omega^2(\eta - \nabla\!\!\!/ \log\Omega, \nabla\!\!\!/\underline{\omega}) + 2\Omega^2\underline{\omega}\left((\eta, -3\eta + 4\nabla\!\!\!/ \log\Omega) - \rho\right) \tag{6.2.22}$$
$$+ 4\Omega^3\underline{\chi}(\eta, \nabla\!\!\!/ \log\Omega) + \Omega^3(\underline{\beta}, 7\eta - 3\nabla\!\!\!/ \log\Omega) + \tfrac{3}{2}\Omega^3\rho \operatorname{tr}\underline{\chi} + \Omega^3 \operatorname{d\!\!/iv}\underline{\beta} + \tfrac{1}{2}\Omega^3(\hat{\chi}, \underline{\alpha}).$$

6.2.2 The Bianchi identities

In this chapter, we only need the following two Bianchi identities:

$$\hat{D}\underline{\alpha} + \left(2\omega - \tfrac{1}{2}\Omega \operatorname{tr}\chi\right)\underline{\alpha} = \Omega\left(-\nabla\!\!\!/\, \hat{\otimes}\, \underline{\beta} - (4\underline{\eta} - \zeta)\,\hat{\otimes}\,\zeta - 3\underline{\hat{\chi}}\rho + 3^*\underline{\hat{\chi}}\sigma\right), \tag{6.2.23}$$

$$D\underline{\beta} + \left(\tfrac{1}{2}\Omega \operatorname{tr}\chi - \Omega\hat{\chi} + \omega\right)\underline{\beta} = \Omega\left(-\nabla\!\!\!/\,\rho + {}^*\nabla\!\!\!/\,\sigma - 3\underline{\eta}\rho + 3^*\underline{\eta}\sigma + 2\underline{\hat{\chi}}\cdot\beta\right). \tag{6.2.24}$$

6.3 Characteristic initial data and characteristic gluing

In this section, we give a brief review of the characteristic gluing problem for the Einstein vacuum equations (1.1.6) in double null gauge [ACR21; ACR23b; ACR23a; CR22]. We follow the conventions of [CR22] unless otherwise stated.

6.3.1 Sphere data, null data, and seed data

The terminology used in this chapter is in agreement with [CR22], which we will be using as a black box, and therefore differs slightly from Chapter 5. We hope this facilitates the reader in understanding exactly how the main notions and results from [CR22] are being used here.

6.3.1.1 Sphere data

Given a solution $(\mathcal{M}^4, g)$ of the Einstein vacuum equations (1.1.6) and a sphere S in a double null foliation, the 2-jet of g can be determined from knowledge of the metric coefficients, Ricci coefficients, and curvature components. However, the equations themselves, such as the Codazzi equation (6.2.19) allow some of these degrees of freedom to be computed from the others, just in terms of derivatives tangent to S. This leads to the following definition:

Definition 6.3.1 (C^2 sphere data, [ACR23b, Definition 2.4])**.** Let S be a 2-sphere. Sphere data x on S consists of a choice of round metric γ on S and the following tuple of S-tensors

$$x = (\Omega, g\!\!\!/, \Omega \operatorname{tr}\chi, \hat{\chi}, \Omega \operatorname{tr}\underline{\chi}, \underline{\hat{\chi}}, \eta, \omega, D\omega, \underline{\omega}, \underline{D\omega}, \alpha, \underline{\alpha}), \tag{6.3.1}$$

where Ω is a positive function, $g\!\!\!/$ a Riemannian metric, $\Omega \operatorname{tr}\chi, \hat{\chi}, \Omega \operatorname{tr}\underline{\chi}, \omega, D\omega, \underline{\omega}, \underline{D\omega}$ are scalar functions, η is a vector field, $\hat{\chi}, \underline{\hat{\chi}}$, α and $\underline{\alpha}$ are symmetric $g\!\!\!/$-traceless 2-tensors.

Definition 6.3.2 ($C^2_u C^{2+m}_v$ sphere data, [ACR23b, Definition 2.28])**.** Let S be a 2-sphere and $m \geq 0$ an integer. Higher order sphere data x on S consists of a choice of round metric γ on S, the tuple

(6.3.1), together with

$$(\hat{D}\alpha, \ldots, \hat{D}^m\alpha, D^2\omega, \ldots, D^{m+1}\omega), \tag{6.3.2}$$

where $\hat{D}^j\alpha$ are symmetric $\not{g}$-traceless 2-tensors and $D^j\omega$ scalar functions. We will write $x = (x^{\mathrm{low}}, x^{\mathrm{high}})$, where x^{low} is a C^2 sphere data set and x^{high} denotes the tuple (6.3.2).

When sphere data is obtained from a geometric sphere in a vacuum spacetime, one has to make a choice of normal null vector fields L and $\underline{L}$. See Lemma 6.1.1 below, for instance. As is well known, the null pair $\{L, \underline{L}\}$ can be "boosted" by the transformation

$$\tilde{L} \doteq \frac{1}{\lambda}L, \quad \underline{\tilde{L}} \doteq \lambda\underline{L}, \tag{6.3.3}$$

where $\lambda \in \mathbb{R}_+$. This boost freedom was also quite useful in Chapter 5.

Definition 6.3.3 (Boosted sphere data)**.** Let x be a sphere data set as in Definition 6.3.2 and $\lambda \in \mathbb{R}_+$. Then the λ-*boosted sphere data set* is the $C^2_uC^{2+m}_v$ sphere data set given by

$$\mathfrak{b}_\lambda(x^{\mathrm{low}}) \doteq (\Omega, \not{g}, \lambda^{-1}\Omega\operatorname{tr}\chi, \lambda^{-1}\hat{\chi}, \lambda\Omega\operatorname{tr}\underline{\chi}, \lambda\underline{\hat{\chi}}, \eta, \lambda^{-1}\omega, \lambda^{-2}D\omega, \lambda\underline{\omega}, \lambda^2\underline{D\omega}, \lambda^{-2}\alpha, \lambda^2\underline{\alpha}),$$
$$\mathfrak{b}_\lambda(x^{\mathrm{high}}) \doteq (\lambda^{-3}\hat{D}\alpha, \ldots, \lambda^{-2-m}\hat{D}^m\alpha, \lambda^{-3}D\omega, \ldots, \lambda^{-2-m}D^{m+1}\omega).$$

This is the effect that the boost (6.3.3) has on the metric coefficients, Ricci coefficients, and curvature components in double null gauge.

There is a norm $\|x\|_{\mathcal{X}^m}$ defined on $C^2_uC^{2+m}_v$ sphere data sets employed in [ACR23b; CR22], which is just a sum of high order (in the angular variable θ) Sobolev norms of the sphere data components [ACR23b, Definition 2.5]. We will show very strong pointwise smallness for arbitrary numbers of angular derivatives later and thus will not need the exact form of these norms in order to apply the result of [CR22].

Definition 6.3.4 (Sphere diffeomorphisms)**.** Given a diffeomorphism $\psi : S^2 \to S^2$, we let ψ act on $C^2_uC^{2+m}_v$ data sets by pullback on each component.

6.3.1.2 Null data

Definition 6.3.5 (Ingoing and outgoing null data [ACR23b, Definition 2.6])**.** Let $v_1 < v_2$. An *outgoing null data set* is given by an assignment $v \mapsto x(v)$, where $x(v)$ is a C^2 sphere data set. We may say that the null data lives on the null hypersurface $C \doteq C^{[v_1,v_2]} \doteq [v_1, v_2] \times S^2$. An *ingoing null* *data set* is defined in the same way, but with the formal variable v replaced with u and η replaced by $\underline{\eta}$ in (6.3.1).

Higher order null data is defined in the obvious way, with $x(v)$ being $C^2_u C^{2+m}_v$ sphere data. Null data on its own is not assumed to satisfy the null structure equations and Bianchi identities.

There are several norms on null data that are employed in [ACR21; ACR23b; ACR23a; CR22]. These include the standard norm $\mathcal{X}$ defined on ingoing and outgoing null data the high regularity norm $\mathcal{X}^+$ defined on ingoing null data, and the high frequency norm $\mathcal{X}^{\text{h.f.}}$ defined on outgoing null data using in obstruction-free characteristic gluing. As we will not need the precise forms of these norms in the present work, we refer the reader to [ACR23b, Definition 2.7] for details.

6.3.1.3 Christodoulou seed data

We will employ the following method, originating in the work of Christodoulou [Chr09], for producing solutions of the null structure and Bianchi equations along a null hypersurface C.

For definiteness, we seek a solution of the null constraints on the null cone segment $C \doteq C^{[0,1]} \doteq [0,1] \times S^2$. The coordinate along $[0,1]$ is called v and we set $L = \partial_v$. On S^2, we have the round metric

$$\gamma \doteq d\vartheta^2 + \sin^2\vartheta\, d\varphi^2,$$

where (ϑ, φ) are standard spherical polar coordinates. We interpret γ as a symmetric S-tensor on C (see Section 2.3.1 for this terminology) by imposing $\gamma(\partial_v, \cdot) = 0$. We set $S_v \doteq \{v\} \times S^2$.

Lemma 6.3.6. *Let $\hat{\not{g}}$ be a smooth S-(0,2)-tensor field on C which induces a Riemannian metric on the sections of C satisfying*

$$\operatorname{tr}_{\hat{\not{g}}} D\hat{\not{g}} = 0, \tag{6.3.4}$$

where $D\hat{\not{g}} \doteq \not{\mathcal{L}}_L \hat{\not{g}}$ as in Section 2.3.1.[1] *Let $\not{g}_1$ be a Riemannian metric on S_1 which is conformal to $\hat{\not{g}}(1)$, $\operatorname{tr}\chi_1$ and $\operatorname{tr}\underline{\chi}_1$ be functions on S_1, η_1 be a 1-form on S_1, and $\hat{\underline{\chi}}_1$ and $\underline{\alpha}_1$ be $\not{g}_1$-traceless symmetric 2-tensors on S_1.*

Then there exist uniquely determined $\not{g}_1$-traceless symmetric 2-tensors $\hat{\chi}_1$ and $\alpha_1, \hat{D}\alpha_1 \ldots, \hat{D}^m\alpha_1$

[1] Concretely, this means $\hat{\not{g}} = \hat{\not{g}}(v)$ is a smooth 1-parameter family of Riemannian metrics on S^2. We identify S^2 with $S_v \subset C$. Since $b = 0$, $\not{\mathcal{L}}_L \hat{\not{g}} = \mathcal{L}_L \hat{\not{g}}$ and relative to any angular coordinates ϑ^A defined on S_1 extended to C according to $L\vartheta^A = 0$, $(D\hat{\not{g}})_{AB} = \partial_v(\hat{\not{g}}_{AB})$.

on S_1, $v_0 \in (-1,1)$, and null data

$$x^{\text{low}}(v) = (\Omega, \not{g}, \Omega \operatorname{tr}\chi, \hat{\chi}, \Omega \operatorname{tr}\underline{\chi}, \hat{\underline{\chi}}, \eta, \omega, D\omega, \underline{\omega}, \underline{D\omega}, \alpha, \underline{\alpha}),$$
$$x^{\text{high}}(v) = (\hat{D}\alpha, \ldots, \hat{D}^m\alpha, D^2\omega, \ldots, D^{m+1}\omega),$$

defined for $v \in (v_0, 1] \cap [0,1]$, satisfying the null structure equations and Bianchi identities along C, such that

$$x^{\text{low}}(1) = (1, \not{g}_1, \operatorname{tr}\chi_1, \hat{\chi}_1, \operatorname{tr}\underline{\chi}_1, \hat{\underline{\chi}}_1, \eta_1, 0, 0, 0, 0, \alpha_1, \underline{\alpha}_1) \tag{6.3.5}$$
$$x^{\text{high}}(1) = (\hat{D}\alpha_1, \ldots, \hat{D}^{m+1}\alpha_1, 0, \ldots, 0), \tag{6.3.6}$$

For every $v \in (v_0, 1] \cap [0,1]$, $\not{g}(v)$ is conformal to $\hat{\not{g}}(v)$ and $\Omega^2(v) = 1$ identically, so $D^j\omega(v) = 0$ identically for $j = 0, \ldots, m+1$. The number v_0 is either strictly negative (in which case x exists on all of C), or is nonnegative and satisfies

$$\lim_{v \searrow v_0} \inf_{S_v} |\not{g}(v)|_{\hat{\not{g}}} = 0. \tag{6.3.7}$$

A *conformal class of Riemannian metrics* on C is the equivalence class $\mathfrak{K}$ of symmetric S-$(0,2)$-tensors on C which are positive definite on each S_v with the equivalence relation $\not{g}', \not{g}'' \in \mathfrak{K}$ if there exists a smooth positive function ψ on C such that $\not{g}' = \psi^2 \not{g}''$.

Lemma 6.3.6 shows that the free data[2] for the characteristic data (in the gauge $\Omega^2 = 1$) are given by

$$\mathfrak{K}, \not{g}(1), \operatorname{tr}\chi(1), \operatorname{tr}\underline{\chi}(1), \eta(1), \hat{\underline{\chi}}(1), \text{and } \underline{\alpha}(1),$$

subject to the condition that $\not{g}(1)$ is compatible with $\mathfrak{K}$ and that $\hat{\underline{\chi}}(1)$ and $\underline{\alpha}(1)$ are traceless, which is a notion that depends only on $\mathfrak{K}$. The desired induced metric $\not{g}$ will be a representative of $\mathfrak{K}$. One often writes $\mathfrak{K} = [\not{g}]$, so the prescription of $\mathfrak{K}$ is the prescription of the conformal class of the induced metric $\not{g}$ which is to be found.

The condition (6.3.4) on the representative $\hat{\not{g}}$ of $\mathfrak{K}$ can be imposed without loss of generality, i.e., $\mathfrak{K}$ always contains a representative satisfying (6.3.4). Indeed, let $\tilde{\not{g}} \in [\not{g}]$ and let $\psi \doteq \exp(\int_v^1 \frac{1}{4} \operatorname{tr}_{\tilde{\not{g}}} D\tilde{\not{g}}\, dv')$. Then $\hat{\not{g}} \doteq \psi^2 \tilde{\not{g}}$ satisfies (6.3.4).

Remark 6.3.7. In Lemma 6.6.2 below, we will directly construct a specific $\hat{\not{g}}$ satsifying the volume form condition $D(d\mu_{\hat{\not{g}}}) = 0$, which easily implies (6.3.4) by the first variation formula for area.

[2] That is, the quantities that may be freely prescribed.

Outline of the proof of Lemma 6.3.6. Let ϕ_1 be the positive function on S_1 satisfying $\not{g}_1 = \phi_1^2 \hat{\not{g}}_1$. We make the ansatz

$$\not{g} = \phi^2 \hat{\not{g}}, \tag{6.3.8}$$

on C, where ϕ is now a positive function on C agreeing with ϕ_1 on S_1.

We define

$$e \doteq \tfrac{1}{8} \hat{\not{g}}^{AB} \hat{\not{g}}^{CD} \partial_v \hat{\not{g}}_{AC} \partial_v \hat{\not{g}}_{BD} \tag{6.3.9}$$

relative to any Lie-transported angular coordinate system on the spheres. We set

$$\partial_v \phi_1 \doteq 2\phi_1 \operatorname{tr}\chi_1 \tag{6.3.10}$$

and let ϕ be the unique solution of the second order ODE

$$\partial_v^2 \phi + e\phi = 0, \tag{6.3.11}$$

with initial conditions $(\phi_1, \partial_v \phi_1)$. If ϕ remains strictly positive on all of C, then we let v_0 be any strictly negative number. If however ϕ has a zero on C, then we take v_0 to be the supremum of $v \in [0,1]$ for which $\inf_{S_v} \phi \leq 0$. This definition gives (6.3.7).

We now set

$$\hat{\chi} \doteq \tfrac{1}{2} \phi^2 D\hat{\not{g}} \quad \text{and} \tag{6.3.12}$$

$$\operatorname{tr}\chi \doteq \tfrac{1}{2} \partial_v \log \phi \tag{6.3.13}$$

along C and observe that this choice of $\operatorname{tr}\chi$ is consistent with (6.3.5). By (6.3.4), the shear defined by (6.3.12) is $\not{g}$-traceless. From (6.3.8), we have

$$D\not{g} = \phi^2 D\hat{\not{g}} + 2\phi \partial_v \phi \hat{\not{g}} = \phi^2 D\hat{\not{g}} + 2\partial_v \log \phi \, \not{g},$$

and by comparing with the first variation formula (6.2.2) written in the form

$$D\not{g} = 2\hat{\chi} + \operatorname{tr}\chi \, \not{g},$$

we conclude that (6.3.12) and (6.3.13) are consistent with the first variation formula.

From (6.3.12), we have

$$e = \tfrac{1}{2}|\hat{\chi}|^2, \tag{6.3.14}$$

so the ODE (6.3.11) is seen to be equivalent to Raychaudhuri's equation (6.2.4).

From here, the full null data along C can be determined by stepping through the null structure and Bianchi equations in the right order, as in [Chr09]. We will outline this procedure in the proof of Lemma 6.6.10 below. □

6.4 Reference sphere data for the Kerr family

Definition 6.4.1 (The Kerr family of metrics)**.** Let $\mathcal{M}_* \doteq (-\infty, \infty)_v \times (0, \infty)_r \times S^2$, where S^2 carries standard spherical polar coordinates ϑ and φ. The *Kerr family* of metrics is the smooth two-parameter family of Lorentzian metrics

$$g_{M,a} = -\left(1 - \frac{2Mr}{\Sigma}\right) dv^2 + 2\,dv\,dr - \frac{4Mar\sin^2\vartheta}{\Sigma} dv\,d\varphi - 2a\sin^2\vartheta\,dr\,d\varphi + \Sigma\,d\vartheta^2 + \rho^2\sin^2\vartheta\,d\varphi^2 \tag{6.4.1}$$

on $\mathcal{M}_*$, where $M \geq 0$ is the mass, $a \in \mathbb{R}$ is the specific angular momentum,

$$\Sigma \doteq r^2 + a^2\cos^2\vartheta, \quad \text{and}$$
$$\rho^2 \doteq r^2 + a^2 + \frac{2Ma^2r\sin^2\vartheta}{\Sigma}.$$

When $a = 0$, $g_{M,a}$ reduces to the *Schwarzschild metric*

$$g_M = -\left(1 - \frac{2M}{r}\right) dv^2 + 2\,dv\,dr + r^2\gamma, \tag{6.4.2}$$

where $\gamma \doteq d\vartheta^2 + \sin^2\vartheta\,d\varphi^2$. When $M = 0$, $g_{M,a}$ reduces to the *Minkowski metric*

$$m \doteq -dv^2 + 2\,dv\,dr + r^2\gamma.$$

The metrics $g_{M,a}$ solve the Einstein vacuum equations (1.1.6). The spacetime $(\mathcal{M}_*, g_{M,a})$ is time-oriented by ∂_v for $r \gg 1$. The vector field ∂_v is Killing—the Kerr family is stationary. If $|a| \leq M$ and $M > 0$, these metrics describe black hole spacetimes. For $0 \leq |a| < M$, the black hole is said to be *subextremal*, and for $0 < |a| = M$, *extremal*.

Remark 6.4.2. In the context of the Schwarzschild solution, the coordinates $(v, r, \vartheta, \varphi)$ are called *ingoing Eddington–Finkelstein coordinates.* Indeed, defining

$$t \doteq v - r - 2M \log|r - 2M|$$

brings g_M into the familiar form

$$g_M = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2\gamma$$

and $(t, r, \vartheta, \varphi)$ are called *Schwarzschild coordinates.* In the context of the Kerr solution, the coordinates $(v, r, \vartheta, \varphi)$ are called *Kerr-star coordinates.* For the relation to the perhaps more familiar *Boyer–Lindquist coordinates*, see [ONe95]. The advantage of defining the Kerr family $g_{M,a}$ directly in these coordinates is that we may view it as a smooth two-parameter family of Lorentzian metrics on the *fixed* smooth manifold $\mathcal{M}_*$, even across the horizons located at $r_\pm = M \pm \sqrt{M^2 - a^2}$ when $M > 0$.

Remark 6.4.3. The spacetimes $(\mathcal{M}_*, g_{M,a})$ defined here do not cover the entire maximal analytic extensions of the Minkowski, Schwarzschild, and Kerr solutions. Most importantly, $(\mathcal{M}_*, g_{M,a})$ includes the portion of the future event horizon $\mathcal{H}^+ \doteq \{r = r_+\}$ strictly to the future of the bifurcation sphere.

We will now define the reference sphere data for the Kerr family. We will use the notion of sphere data $x[g, i, \underline{L}]$ generated by a Lorentzian metric g on a smooth manifold $\mathcal{M}$, an embedding $i : S^2 \to \mathcal{M}$, and a choice of null vector field $\underline{L}$ defined along and orthogonal to $i(S^2)$, which is defined in Lemma 6.1.1 below. Note that

$$Y \doteq -\partial_r$$

is a future-directed null vector field for $(\mathcal{M}_*, g_{M,a})$. We also define the family of embeddings

$$\begin{aligned} i_R : S^2 &\to \mathcal{M}_* \\ (\vartheta, \varphi) &\mapsto (0, R, \vartheta, \varphi) \end{aligned}$$

for $R > 0$.

Definition 6.4.4 (Reference sphere data)**.** Let $M \geq 0$, $a \in \mathbb{R}$, $R > 0$, and $m \geq 0$ be an integer.

The *reference Kerr sphere data set* of mass M, specific angular momentum a, and radius[3] R is the $C^2_u C^{2+m}_v$ sphere data set given by

$$\mathfrak{k}_{M,a,R} \doteq x[g_{M,a}, i_R, Y]. \tag{6.4.3}$$

The *reference Schwarzschild data sets* are defined by

$$\mathfrak{s}_{M,R} \doteq \mathfrak{k}_{M,0,R} \tag{6.4.4}$$

and the *reference Minkowski data sets* are defined by

$$\mathfrak{m}_R \doteq \mathfrak{s}_{0,R}. \tag{6.4.5}$$

We will colloquially refer to $\mathfrak{k}_{M,a,R}$ as a "Kerr coordinate sphere" and $\mathfrak{s}_{M,R}$ (resp., $\mathfrak{m}_R$) as a "(round) Schwarzschild symmetry sphere" (resp., "(round) Minkowski symmetry sphere").

In the notation of Section 6.3.1.1, one can show that

$$\mathfrak{s}^{\mathrm{low}}_{M,R} = \left(1, R^2\gamma, \frac{2}{R}\left(1 - \frac{2M}{R}\right), 0, -\frac{2}{R}, 0, \ldots, 0\right), \tag{6.4.6}$$

$$\mathfrak{s}^{\mathrm{high}}_{M,R} = 0. \tag{6.4.7}$$

A similarly simple expression is neither available nor needed for Kerr. Indeed, we have the

Lemma 6.4.5. *For any integer $m \geq 0$, $\mathfrak{k}_{M,a,R}$ is a smooth three-parameter family of $C^2_u C^{2+m}_v$ sphere data sets. In particular,*

$$\lim_{a\to 0} \|\mathfrak{k}_{M,a,R} - \mathfrak{s}_{M,R}\|_{\mathcal{X}^m} = 0. \tag{6.4.8}$$

Proof. The metrics $g_{M,a}$ are defined on the fixed smooth manifold $\mathcal{M}_*$. By inspection of (6.4.1), $g_{M,a}$ varies smoothly in M and a. Therefore, the smooth dependence of $\mathfrak{k}_{M,a,R}$ on the parameters and (6.4.8) follow from the smooth dependence of the sphere data generated by $(g, i, \underline{L})$ on g, i, and $\underline{L}$; see Lemma 6.1.1. □

We conclude this section with several remarks.

Remark 6.4.6. As was already mentioned, the Kerr family is stationary. Defining $i_R(\vartheta, \varphi) = (v, R, \vartheta, \varphi)$ for any $v \in \mathbb{R}$ leads to the same sphere data.

[3]We use the term radius because it is associated to the Kerr coordinate r, but the spheres are not round!

Remark 6.4.7. We always take the Kerr axis to point along the poles of the fixed identification of the Kerr coordinate spheres with the usual unit sphere.

Remark 6.4.8. The induced metric $\not{g}_{M,a,R}$ in $\mathfrak{k}_{M,a,R}$ is not conformal to the round metric γ (defined relative to the Kerr angular coordinates). For this reason we have slightly modified the setup in Section 6.3.1.3 by imposing (6.3.4) instead of simply $d\mu_{\hat{\not{g}}} = d\mu_\gamma$ as in [Chr09, Chapter 2]. See already Lemma 6.6.2 below.

Remark 6.4.9. The induced metric $\not{g}_{M,a,R}$ is given in Kerr angular coordinates by

$$\not{g}_{M,a,R} = \Sigma\, d\vartheta^2 + \rho^2 \sin^2\vartheta\, d\varphi^2. \tag{6.4.9}$$

To show that this extends smoothly over the poles relative to the smooth structure defined by the Kerr angular coordinates, we note the identity

$$\Sigma\, d\vartheta^2 + \rho^2 \sin^2\vartheta\, d\varphi^2 = \Sigma(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2) + a^2\left(1 + \frac{2Mr}{\Sigma}\right)\sin^4\vartheta\, d\varphi^2. \tag{6.4.10}$$

Now $\sin^2\vartheta\, d\varphi$ is a globally defined smooth 1-form on S^2 since it is the γ-dual of the globally defined vector field ∂_φ, so the right-hand side of (6.4.10) can be extended smoothly over the poles.

6.5 Perturbative characteristic gluing

Since the characteristic gluing results of [ACR21; ACR23b; ACR23a; CR22] pass through linear theory, the conserved charges in Minkowski space play an important role. In Section 6.5.1, we give the definition of conserved charges. In Section 6.5.2, we state the main result of [CR22] in the form which we will directly apply it.

6.5.1 Conserved charges

Definition 6.5.1 (Spherical harmonics)**.** For $\ell \in \mathbb{N}_0$ and $m = -\ell, \dots, \ell$, let Y^ℓ_m denote the standard real-valued spherical harmonics on the unit sphere (S^2, γ). We also define the electric and magnetic 1-form spherical harmonics by

$$E^\ell_m \doteq -\frac{1}{\sqrt{\ell(\ell+1)}}\not{\nabla} Y^\ell_m \quad \text{and} \quad H^\ell_m \doteq \frac{1}{\sqrt{\ell(\ell+1)}}{}^*\not{\nabla} Y^\ell_m$$

for $\ell \geq 1$ and $|m| \leq \ell$. By a standard abuse of notation, we will use the same symbol for the vector-valued spherical harmonics, with the understanding that γ is used to raise the index.

Definition 6.5.2 (Linearly conserved charges)**.** Let x be C^2 sphere data and define the 1-form $\mathbf{B}$ and scalar function $\mathbf{m}$ by

$$\mathbf{B} \doteq \frac{\phi^3}{2\Omega^2}\left(\nabla\!\!\!/(\Omega \operatorname{tr}\chi) + \Omega \operatorname{tr}\chi(\eta - 2\nabla\!\!\!/ \log \Omega)\right) \text{ and}$$
$$\mathbf{m} \doteq \phi^3\left(K + \tfrac{1}{4}\operatorname{tr}\chi \operatorname{tr}\underline{\chi}\right) - \phi \operatorname{d\!\!\!/iv} \mathbf{B}.$$

The conformal factor ϕ is defined as the unique positive function on S^2 such that $d\mu_{g\!\!\!/} = \phi^2 d\mu_\gamma$, where γ is the distinguished choice of round metric on S. Then the charges $\mathbf{E}, \mathbf{P}, \mathbf{L}$, and $\mathbf{G}$ (where the latter three are vectors in $\mathbb{R}^3$ indexed by $m \in \{-1, 0, 1\}$) are defined by

$$\mathbf{E} \doteq \mathbf{m}^{\ell=0}$$
$$\mathbf{P} \doteq \mathbf{m}^{\ell=1}$$
$$\mathbf{L} \doteq \mathbf{B}^{\ell=1,H}$$
$$\mathbf{G} \doteq \mathbf{B}^{\ell=1,E}.$$

Here the modes are defined by

$$f^{\ell=0} \doteq \int_{S^2} f\, Y^0_0 \, d\mu_\gamma, \qquad (f^{\ell=1})^m \doteq \int_{S^2} f\, Y^1_m \, d\mu_\gamma,$$
$$(X^{\ell=1,E})^m \doteq \int_{S^2} \gamma(X, E^1_m)\, d\mu_\gamma, \qquad (X^{\ell=1,H})^m \doteq \int_{S^2} \gamma(X, H^1_m)\, d\mu_\gamma.$$

6.5.2 Czimek–Rodnianski obstruction-free perturbative characteristic gluing

The following theorem is a combination of [ACR23b, Theorem 3.2], [CR22, Theorem 2.9], and Remark (5) after Theorem 2.9 in [CR22].

Theorem 6.5.3 (Czimek–Rodnianski obstruction-free characteristic gluing)**.** *For any $C_{\mathbf{E}} > 0$ and integer $m \geq 0$, there exist constants $C_* > 0$ and $\varepsilon_0 > 0$ such that the following holds. Let $\underline{x}$ be ingoing null data on an ingoing cone $\underline{C} = [-\frac{1}{100}, \frac{1}{100}]_u \times S^2$ solving the null structure equations and Bianchi identities, and x_2 be $C^2_u C^{2+m}_v$ sphere data. Let x_1 be the sphere data in $\underline{x}$ corresponding to*

$u = 0$. *Let*

$$(\Delta\mathbf{E}, \Delta\mathbf{P}, \Delta\mathbf{L}, \Delta\mathbf{G}) \doteq (\mathbf{E}, \mathbf{P}, \mathbf{L}, \mathbf{G})(x_2) - (\mathbf{E}, \mathbf{P}, \mathbf{L}, \mathbf{G})(x_1)$$

be the difference of the conserved charges of x_2 *and* x_1. *If the data sets satisfy the smallness condition*

$$\|\underline{x} - \mathfrak{m}\|_{\mathcal{X}^+(\underline{C})} + \|x_2 - \mathfrak{m}_2\|_{\mathcal{X}^m} < \varepsilon_\flat \tag{6.5.1}$$

for some $0 < \varepsilon_\flat < \varepsilon_0$, *where* $\mathfrak{m}$ *is reference Minkowski null data*[4] *and* $\mathfrak{m}_2$ *is reference Minkowski sphere data, and the following "coercivity" conditions on the charge differences*

$$\Delta\mathbf{E} > C_{\mathbf{E}}\varepsilon_\flat, \tag{6.5.2}$$

$$|\Delta\mathbf{L}| < \varepsilon_\flat^2, \text{ and} \tag{6.5.3}$$

$$|\Delta\mathbf{P}| + |\Delta\mathbf{G}| < C_*\Delta\mathbf{E}, \tag{6.5.4}$$

then there is a solution $x \in \mathcal{X}(C)$ *of the null structure equations along a null hypersurface* $C = [1,2]_v \times S^2$ *such that*

$$x(1) = x_1', \quad x(2) = x_2, \tag{6.5.5}$$

and

$$\|x - \mathfrak{m}\|_{\mathcal{X}^{\text{h.f.}}(C)} + \|x_1' - \mathfrak{m}_1\|_{\mathcal{X}} \lesssim \|\underline{x} - \mathfrak{m}\|_{\mathcal{X}^+(\underline{C})} + \|x_2 - \mathfrak{m}_2\|_{\mathcal{X}^m}.$$

The sphere data x_1' *is obtained by applying a sphere diffeomorphism and a transversal sphere perturbation to* x_1 *inside of* $\underline{C}$. *See [ACR23b; CR22] for the precise definitions of these terms.*

Remark 6.5.4. The matching condition (6.5.5) is to order C^{2+m} in directions tangent to the cone.

Since all hypotheses in Theorem 6.5.3 are open conditions, we immediately have:

Corollary 6.5.5. *If the sphere data set* x_2 *satisfies the hypotheses of Theorem 6.5.3, there exists an* $\varepsilon_* > 0$ *such that if* $\tilde{x}_2$ *is another sphere data set such that*

$$\|\tilde{x}_2 - x_2\|_{\mathcal{X}^m} < \varepsilon$$

for some $0 \leq \varepsilon < \varepsilon_*$, *then the conclusion of the theorem holds for* $\tilde{x}_2$ *in place of* x_2.

[4] That is, reference Minkowski sphere data defined along the ingoing cone $\underline{C}$. See [ACR23b] for details.

6.6 Proofs of the main gluing theorems

6.6.1 Gluing an almost-Schwarzschild sphere to a round Schwarzschild sphere with a larger mass

In this subsection, we prove the main technical lemma of the chapter. In essence, we show how to decrease the mass of a Schwarzschild sphere (going backwards in time) by an arbitrary amount, with an arbitrary small error.

Proposition 6.6.1. *Given any $0 \leq M_* < M$, $R > 0$, integer $m \geq 0$, and any $\varepsilon_\sharp > 0$, there exists a $\delta > 0$ and null data x on $C_1^{[0,1]} = [0,1] \times S^2$ solving the null structure equations and Bianchi identities such that*

$$\mathfrak{b}_\delta(x(1)) = \mathfrak{s}_{M,R} \tag{6.6.1}$$

and

$$\|\mathfrak{b}_\delta(x(0)) - \mathfrak{s}_{M_*,R}\|_{\mathcal{X}^m} < \varepsilon_\sharp, \tag{6.6.2}$$

where $\mathfrak{b}$ is the boost operation defined in Definition 6.3.3 and $\mathcal{X}^m$ is the sphere data norm appearing in Theorem 6.5.3.

The proof of the proposition is given at the end of this subsection. We first give a general construction of seed data $\hat{\not{g}}$ compatible with the hypotheses of Lemma 6.3.6.

Lemma 6.6.2. *Let C be as in Section 6.3.1.3. Let $\tilde{\gamma}$ be a Riemannian metric on S^2. There exists an explicitly defined smooth assignment $\tilde{\gamma} \mapsto \mathfrak{h}$, where $\mathfrak{h}$ is a traceless $(1,1)$-S-tensor field along C, such that for any $\lambda \in \mathbb{R}$,*

$$\hat{\not{g}}_{AB} \doteq \tilde{\gamma}_{AC} \exp(\lambda\mathfrak{h})^C{}_B \tag{6.6.3}$$

defines a Riemannian metric for each v, satisfies condition (6.3.4), and $\hat{\not{g}}(1) = \tilde{\gamma}$. Here $\tilde{\gamma}$ is defined along C according to $D\tilde{\gamma} = 0$. We have

$$\partial_v \hat{\not{g}}_{AB} = \lambda \hat{\not{g}}_{AC} \partial_v \mathfrak{h}^C{}_B \tag{6.6.4}$$

and the inverse metric is given by

$$\hat{\not{g}}^{AB} = \left(\hat{\not{g}}^{-1}\right)^{AB} = (\tilde{\gamma}^{-1})^{AC} \exp(-\lambda\mathfrak{h})^B{}_C = (\tilde{\gamma}^{-1})^{BC} \exp(-\lambda\mathfrak{h})^A{}_C. \tag{6.6.5}$$

Proof. We first fix some cutoff functions. Let $\chi \in C_c^\infty(0, \frac{1}{2})$ be nonnegative and not identically zero.

Let $\chi_1 \doteq \chi(v)$ and $\chi_2(v) \doteq \chi(v - \frac{1}{2})$. Let p_1 be the north pole of S^2, p_2 the south pole, and set $U_i \doteq S^2 \setminus \{p_i\}$ for $i = 1, 2$. Let $f_1 \in C_c^\infty(U_1)$ and $f_2 \in C_c^\infty(U_2)$ be such that $f_1^2 + f_2^2 = 1$ on S^2.

Let $(\vartheta_1^1, \vartheta_1^2)$ be a coordinate chart covering U_1, $(\vartheta_2^1, \vartheta_2^2)$ be a coordinate chart covering U_2, and set

$$\mathring{h}_i \doteq d\vartheta_i^1 \otimes \frac{\partial}{\partial \vartheta_i^1} - d\vartheta_i^2 \otimes \frac{\partial}{\partial \vartheta_i^2} \text{ on } U_i.$$

As matrices, these tensor fields are given by $\operatorname{diag}(1, -1)$ in the respective coordinate systems.

We now claim that the symmetric $(0, 2)$-tensor fields

$$h_{i\,AB} \doteq \tfrac{1}{2}\left(\tilde{\gamma}_{AC}\mathring{h}_i{}^C{}_B + \tilde{\gamma}_{BC}\mathring{h}_i{}^C{}_A\right)$$

are nowhere vanishing on their respective domains of definition. This follows from the fact that

$$h_{i\,AA} = \tilde{\gamma}\left(\frac{\partial}{\partial \vartheta_i^A}, \frac{\partial}{\partial \vartheta_i^A}\right),$$

where no summation is implied. Since $\tilde{\gamma}$ is positive definite, we must at each point have both $h_{i\,11}$ and $h_{i\,22}$ nonvanishing, so h_i is always nonvanishing. Let $h_i^\sharp$ be the $(1, 1)$-tensor field obtained by dualizing h_i with $\tilde{\gamma}$. We then finally define

$$\mathfrak{h} \doteq \chi_1 f_1 |h_1|_{\tilde{\gamma}}^{-1} h_1^\sharp + \chi_2 f_2 |h_2|_{\tilde{\gamma}}^{-1} h_2^\sharp. \tag{6.6.6}$$

It is clear that $\operatorname{tr} \mathfrak{h} = 0$ and that $\mathfrak{h}^\flat$ is symmetric, where $\flat$ is taken relative to $\tilde{\gamma}$.

We now show that (6.6.3) defines a Riemannian metric. Viewing $\mathfrak{h}$ as an endomorphism $TS^2 \to TS^2$, the power series

$$\exp(\lambda\mathfrak{h}) = \sum_{n=0}^{\infty} \frac{1}{n!}(\lambda\mathfrak{h})^n \tag{6.6.7}$$

converges and defines a smooth family of endomorphisms.

To verify that $\hat{\not{g}}$ is symmetric, we examine (6.6.7) term by term:

$$\begin{aligned}
\tilde{\gamma}_{AC}\left((\lambda\mathfrak{h})^n\right)^C{}_B &= \lambda^n \tilde{\gamma}_{AD_1}\mathfrak{h}^{D_1}{}_{D_2}\mathfrak{h}^{D_2}{}_{D_3}\cdots\mathfrak{h}^{D_{n-1}}{}_{D_n}\mathfrak{h}^{D_n}{}_B \\
&= \lambda^n \tilde{\gamma}_{D_1D_2}\mathfrak{h}^{D_1}{}_{A}\mathfrak{h}^{D_2}{}_{D_3}\cdots\mathfrak{h}^{D_{n-1}}{}_{D_n}\mathfrak{h}^{D_n}{}_B \\
&= \cdots \\
&= \lambda^n \tilde{\gamma}_{D_{n-1}D_n}\mathfrak{h}^{D_1}{}_{D_2}\mathfrak{h}^{D_2}{}_{D_3}\cdots\mathfrak{h}^{D_{n-1}}{}_{A}\mathfrak{h}^{D_n}{}_B \\
&= \lambda^n \tilde{\gamma}_{D_{n-1}D_n}\mathfrak{h}^{D_1}{}_{D_2}\mathfrak{h}^{D_2}{}_{D_3}\cdots\mathfrak{h}^{D_{n-1}}{}_{B}\mathfrak{h}^{D_n}{}_A \\
&= \tilde{\gamma}_{BC}\left((\lambda\mathfrak{h})^n\right)^C{}_A,
\end{aligned}$$

where we used the symmetry of $\mathfrak{h}^\flat$ repeatedly. That $\hat{\not{g}}$ is positive definite follows easily from the fact that at the origin of a normal coordinate system for $\tilde{\gamma}$, $\hat{\not{g}}_{AB}$ is the matrix exponential of a symmetric matrix, and hence positive definite.

To show that (6.3.4) is satisfied, we use Jacobi's formula to calculate

$$\det\hat{\not{g}} = \det\tilde{\gamma}\,\exp(\lambda\operatorname{tr}\mathfrak{h}) = \det\tilde{\gamma}$$

relative to any coordinate system, where we used $\operatorname{tr}\mathfrak{h} = 0$. We conclude that the volume form of $\hat{\not{g}}$ satisfies

$$d\mu_{\hat{\not{g}}} = d\mu_{\tilde{\gamma}}. \tag{6.6.8}$$

Observe that since $D(d\mu_{\tilde{\gamma}}) = 0$ by definition of $\tilde{\gamma}$ along C, (6.6.8) implies

$$0 = D(d\mu_{\hat{\not{g}}}) = \tfrac{1}{2}\operatorname{tr}_{\hat{\not{g}}}(D\hat{\not{g}})\,d\mu_{\hat{\not{g}}},$$

so $D\hat{\not{g}}$ is $\hat{\not{g}}$-traceless.

To prove (6.6.4), we use the fact that $\mathfrak{h}(v)$ and $\mathfrak{h}(v')$ commute for any v and v' sufficiently close to simply differentiate (6.6.3):

$$D\hat{\not{g}}_{AB} = \partial_v\hat{\not{g}}_{AB} = \gamma_{AC}\exp(\lambda\mathfrak{h})^C{}_D\lambda\partial_v\mathfrak{h}^D{}_B = \lambda\hat{\not{g}}_{AC}\partial_v\mathfrak{h}^C{}_B.$$

The formula (6.6.5) is immediately seen to hold. □

Remark 6.6.3. By the Poincaré–Hopf theorem, the shear $\hat{\chi}$ must vanish at some point on each $S_v \subset C$. Equivalently, any $\mathfrak{h}$ for which (6.6.3) satisfies condition (6.3.4), must vanish at some point on each S_v. In order to satisfy (6.6.9) below, this zero cannot stay along the same generator of C. The simplest solution to this problem is the two-pulse configuration above.

With this general construction out of the way, we begin the proof of Proposition 6.6.1 in earnest. We specialize now to the case of $\tilde{\gamma} = \gamma$, the round metric on the unit sphere.

Convention. We now introduce a parameter $\delta > 0$ satisfying $0 < \delta < \delta_0$, where $\delta_0 > 0$ is a sufficiently small fixed parameter only depending on M_*, M, R and the fixed choices of χ, U_1, U_2, f_1, and f_2. We will further use in this section the notation that implicit constants in $\lesssim, \gtrsim$, and $\approx$ may depend M_*, M, R and χ, U_1, U_2, f_1, and f_2. We also use the notation $\lesssim_j, \gtrsim_j$, and $\approx_j$ if the implicit constants in $\lesssim, \gtrsim$, and $\approx$ depend on an additional parameter j.

Lemma 6.6.4. *The geometric quantity* e*, defined in* (6.3.14)*, satisfies*

$$\not\nabla \int_0^1 e\, dv = 0 \tag{6.6.9}$$

and

$$|\not\nabla^j e| \lesssim_j \lambda^2. \tag{6.6.10}$$

for $j \geq 0$.

Proof. We have

$$e = \tfrac{1}{8}\lambda^2 \partial_v \mathfrak{h}^A{}_B \partial_v \mathfrak{h}^B{}_A = \tfrac{1}{8}\lambda^2 \left((\chi_1')^2 f_1^2 + (\chi_2')^2 f_2^2\right)$$

by (6.3.9), (6.6.5), and (6.6.6). Therefore, we have

$$\not\nabla^j e = \tfrac{1}{8}\lambda^2 \left((\chi_1')^2 \not\nabla^j f_1^2 + (\chi_2')^2 \not\nabla^j f_2^2\right),$$

which immediately proves (6.6.10). To prove (6.6.9), we note that

$$\int_0^1 \left((\chi_1')^2 f_1^2 + (\chi_2')^2 f_2^2\right) dv = \left(f_1^2 + f_2^2\right) \int_0^1 (\chi')^2\, dv = \int_0^1 (\chi')^2\, dv,$$

which is independent of the angle on S^2. □

Along $[0,1] \times S^2$ we impose the gauge condition $\Omega^2 = 1$ and at $v = 1$ we impose

$$\operatorname{tr}\chi(1) = \delta \frac{2}{R}\left(1 - \frac{2M}{R}\right). \tag{6.6.11}$$

The conformal factor ϕ solves Raychaudhuri's equation (6.3.11) with final values (see (6.3.13) and

(6.6.11))

$$\phi(1) = R_f$$
$$\partial_v\phi(1) = \delta\left(1 - \frac{2M}{R}\right)$$

Lemma 6.6.5. *If* $0 < \delta \le \delta_0$ *and* $0 \le \lambda \le \delta^{1/4}$*, then*

$$|\nabla\!\!\!/^j(\phi - R)| + |\nabla\!\!\!/^j \partial_v\phi| \lesssim_j \delta + \lambda^2 \tag{6.6.12}$$

uniformly on $[0,1] \times S^2$ *for every integer* $j \ge 0$.

Proof. Integrating (6.3.11), we obtain

$$\partial_v\phi(v) = \delta\left(1 - \frac{2M}{R}\right) + \int_v^1 \phi\, e\, dv'. \tag{6.6.13}$$

Assuming $|\phi| \le 10R_f$ in the context of a simple bootstrap argument, we see that (6.6.13) and (6.6.10) imply

$$|\partial_v\phi| \lesssim \delta + \lambda^2, \tag{6.6.14}$$

which implies

$$|\phi - R| \lesssim \delta + \lambda^2. \tag{6.6.15}$$

Since $\lambda^2 \le \delta^{1/2}$, taking $\delta_0 > 0$ sufficiently small closes the boostrap and (6.6.14) and (6.6.15) hold on $[0,1]$. Commuting (6.3.11) repeatedly with $\nabla\!\!\!/$ and arguing inductively using (6.6.14) and (6.6.15) as the base cases, we easily obtain (6.6.12). □

Lemma 6.6.6. *Fix an angle* $\theta_0 \in S^2$*. For* $0 < \delta \le \delta_0$*, the function*

$$\lambda \longmapsto \operatorname{tr}\chi(0, \theta_0; \lambda)$$

is monotonically increasing.

Proof. Since $\Omega = 1$ identically, Raychaudhuri's equation (6.2.4) becomes

$$\partial_v \operatorname{tr}\chi = -2\lambda^2 e_1 - \tfrac{1}{2}(\operatorname{tr}\chi)^2, \tag{6.6.16}$$

where $e_1 \doteq \frac{1}{8}\partial_v \mathfrak{h}^A{}_B \partial_v \mathfrak{h}^B{}_A$. Taking the ∂_λ derivative of (6.6.16) gives

$$\partial_v\left(\partial_\lambda \operatorname{tr}\chi\right) = -4\lambda e_1 - \operatorname{tr}\chi(\partial_\lambda \operatorname{tr}\chi)$$

This is at once solved for

$$\partial_\lambda \operatorname{tr}\chi(v) = 4\lambda e^{\int_v^1 \operatorname{tr}\chi\, dv'} \int_v^1 e^{-\int_{v'}^1 \operatorname{tr}\chi\, dv''} e_1\, dv',$$

which is strictly positive at $v = 0$ for $\lambda > 0$. □

Lemma 6.6.7. *Let $0 < \delta \le \delta_0$ and $0 \le \lambda \le \delta^{1/4}$. Then $\operatorname{tr}\chi$ is monotonically decreasing along each generator and*

$$\inf_{[0,\frac{1}{2}]\times S^2} \operatorname{tr}\chi \gtrsim \delta + \lambda^2. \tag{6.6.17}$$

Proof. Monotonicity of $\operatorname{tr}\chi$ follows at once from Raychaudhuri's equation (6.6.16). We can immediately integrate (6.6.16) to obtain

$$\operatorname{tr}\chi(v) = \delta\frac{2}{R}\left(1 - \frac{2M}{R}\right) + 2\lambda^2 e^{\frac{1}{2}\int_v^1 \operatorname{tr}\chi\, dv'} \int_{v'}^1 e^{-\frac{1}{2}\int_{v'}^1 \operatorname{tr}\chi\, dv''} e_1 dv', \tag{6.6.18}$$

By (6.6.12), $\operatorname{tr}\chi \lesssim \delta + \lambda^2 \lesssim \delta^{1/2} \le 1$, so (6.6.18) implies (6.6.17). □

Lemma 6.6.8. *Fix an angle $\theta_0 \in S^2$. For $0 < \delta \le \delta_0$, there exists a unique $\lambda_0 = \lambda_0(\delta) \in (0, \delta^{1/4})$ (depending also on θ_0) such that*

$$\operatorname{tr}\chi(0, \theta_0; \lambda_0) = \delta\frac{2}{R}\left(1 - \frac{2M_*}{R}\right), \tag{6.6.19}$$

which also satisfies

$$\lambda_0(\delta) \approx \delta^{1/2}. \tag{6.6.20}$$

Proof. Let

$$c \doteq \frac{4}{R}(M - M_*) \quad \text{and} \quad C \doteq 2\int_0^1 e_1(v, \vartheta_0)\, dv.$$

Then the condition (6.6.19) becomes (see (6.6.11) and (6.6.16))

$$c\delta = C\lambda^2 + \tfrac{1}{2}\int_0^1 (\operatorname{tr}\chi)^2\, dv. \tag{6.6.21}$$

Combining (6.6.12) with (6.6.17) shows immediately that (6.6.21) can be achieved by a $\lambda_0(\delta)$ satisfying (6.6.20). ☐

From now on, we always take $\lambda = \lambda_0(\delta)$ as constructed in Lemma 6.6.8. With this λ, our main estimates (6.6.12) are improved to:

$$|\nabla\!\!\!/^j e| + |\nabla\!\!\!/^j (\phi - R)| + |\nabla\!\!\!/^j \operatorname{tr}\chi| + \delta^{1/2}|\nabla\!\!\!/^j \hat{\chi}| \lesssim_j \delta \tag{6.6.22}$$

for any $j \geq 0$ and uniformly on $[0,1] \times S^2$. Importantly, we also have

Lemma 6.6.9. *Let $0 < \delta \leq \delta_0$. Then,*

$$|\nabla\!\!\!/^j \operatorname{tr}\chi(0)| \lesssim_j \delta^2 \tag{6.6.23}$$

at $v = 0$ for $j \geq 1$. Hence,

$$\left|\nabla\!\!\!/^j \left(\operatorname{tr}\chi(0) - \delta\frac{2}{R}\left(1 - \frac{2M_*}{R}\right)\right)\right| \lesssim_j \delta^2 \tag{6.6.24}$$

for all $j \geq 0$ at $v = 0$.

Proof. Applying $\nabla\!\!\!/^j$ to (6.6.16), integrating in v, and applying (6.6.9) yields

$$\nabla\!\!\!/^j \operatorname{tr}\chi(0) = -\tfrac{1}{2}\int_0^1 \nabla\!\!\!/^j (\operatorname{tr}\chi)^2\, dv.$$

We arrive at (6.6.23) after applying (6.6.22). This also proves (6.6.24) for $j \geq 1$. For $j = 0$, we integrate (6.6.23) along geodesics emanating from θ_0 and use (6.6.19). ☐

The remaining sphere data at $v = 1$ is now specified as follows:

$$\begin{aligned} \operatorname{tr}\underline{\chi}(1) &= -\frac{1}{\delta}\frac{2}{R} \\ \underline{\hat{\chi}}(1) &= 0 \\ \eta(1) &= 0 \\ \underline{\omega}(1) = \underline{D\omega}(1) &= 0 \\ \alpha(1) = \underline{\alpha}(1) &= 0. \end{aligned}$$

Combining everything and using the null structure and Bianchi equations to solve the rest of the system, we have

Lemma 6.6.10. *For $0 < \delta \le \delta_0$ we have* <u>*at $v = 0$*</u>

$$\begin{aligned}\delta^{-1}|\nabla\!\!\!/^j(g\!\!\!/ - R^2\gamma)| + \delta^{-2}\left|\nabla\!\!\!/^j\left(\operatorname{tr}\chi(0) - \delta\frac{2}{R}\left(1 - \frac{2M_*}{R}\right)\right)\right| + \delta^{-1}\left|\nabla\!\!\!/^j\left(K - \frac{1}{R^2}\right)\right| \\ + \delta^{-1/2}|\nabla\!\!\!/^j\eta| + \left|\nabla\!\!\!/^j\left(\operatorname{tr}\underline{\chi} + \frac{1}{\delta}\frac{2}{R}\right)\right| + \delta^{1/2}|\nabla\!\!\!/^j\hat{\underline{\chi}}| + \delta^{-3/2}|\nabla\!\!\!/^j\beta| \\ + \delta^{-1/2}\left|\nabla\!\!\!/^j\left(\rho + \frac{2M_*}{R^3}\right)\right| + \delta^{-1/2}|\nabla\!\!\!/^j\sigma| + |\nabla\!\!\!/^j\underline{\beta}| + \delta^{1/2}|\nabla\!\!\!/^j\underline{\alpha}| + |\nabla\!\!\!/^j\underline{\omega}| + \delta|\nabla\!\!\!/^j\underline{D\omega}| \lesssim_j 1 \qquad (6.6.25)\end{aligned}$$

for every $j \ge 0$ and

$$\hat{\chi}(0) = 0, \qquad \alpha(0) = 0. \qquad (6.6.26)$$

The terms in (6.6.25) *are displayed in the order in which they are estimated.*

Proof. The proof follows the procedure of [Chr09, Chapter 2], which we now outline. The first term is estimated using (6.6.22). The second term was estimated in (6.6.24). The third term is estimated using the formula

$$K_{g\!\!\!/} = \phi^{-2}(K_{\hat{g\!\!\!/}} - \Delta\!\!\!/_{\hat{g\!\!\!/}}\log\phi).$$

Note that the first and third terms are estimated by $\delta^{1/2}$ on the whole cone, but are improved at $v = 0$. To estimate the fourth term, the transport equation (6.2.8) combined with the Codazzi equation (6.2.19) and (6.1.1) yields

$$\partial_v\eta_A + \operatorname{tr}\chi\,\eta_A = (\operatorname{d\!\!\!/iv}\chi)_A - \nabla\!\!\!/_A\operatorname{tr}\chi.$$

Now $|\eta|$ can be estimated using Grönwall's inequality and (6.6.22). To estimate the fifth term, the transport equation (6.2.14) is combined with the Gauss equation (6.2.18) to give

$$\partial_v\operatorname{tr}\underline{\chi} + \operatorname{tr}\chi\ \operatorname{tr}\underline{\chi} = -2K - 2\operatorname{d\!\!\!/iv}\eta + |\eta|^2.$$

Grönwall gives $|\operatorname{tr}\underline{\chi}| \lesssim \delta^{-1}$, which then easily implies the desired estimate by Grönwall applied to

$$(\partial_v + \operatorname{tr}\chi)\left(\operatorname{tr}\underline{\chi} + \frac{1}{\delta}\frac{2}{R}\right) = -2K - 2\operatorname{d\!\!\!/iv}\eta + |\eta|^2 + \frac{1}{\delta}\frac{2}{R}\operatorname{tr}\underline{\chi}.$$

To estimate the sixth term, apply Grönwall directly to (6.2.16). The first variation formula (6.3.12), the second variation formula (6.2.6), and (6.6.4) imply (6.6.26). The seventh term in (6.6.25) is estimated directly from the Codazzi equation (6.2.19). The eighth term is estimated directly from the Gauss equation (6.2.18). The ninth term is estimated directly from the curl equation (6.2.21).

The tenth term is estimated using Grönwall and the Bianchi equation (6.2.24). The eleventh term is estimated by using the Einstein equations (6.2.1) and the first variation formula (6.2.2) to compute

$$0 = D \operatorname{tr} \underline{\alpha} = \operatorname{tr} D\underline{\alpha} - 2\Omega(\chi, \underline{\alpha}).$$

Combined with the Bianchi identity (6.2.23), this yields

$$\partial_v \underline{\alpha}_{AB} - (\hat{\chi}, \underline{\alpha})\not{g}_{AB} - \tfrac{1}{2} \operatorname{tr} \chi \underline{\alpha}_{AB} = -(\not{\nabla} \hat{\otimes} \underline{\beta})_{AB} + 5(\eta \hat{\otimes} \underline{\beta})_{AB} - 3\underline{\hat{\chi}}_{AB}\rho + 3^* \underline{\hat{\chi}}_{AB}\sigma,$$

from which the desired estimate follows by Grönwall. The twelfth term is estimated by integrating (6.2.10) and the thirteeth term is estimated by integrating (6.2.22). □

We are now ready to prove the main result of this subsection.

Proof of Proposition 6.6.1. Let $x^{\text{low}}(v)$, $v \in [0,1]$, be the null data constructed above. We have defined

$$x^{\text{low}}(1) = \left(1, R^2\gamma, \delta \frac{2}{R}\left(1 - \frac{2M}{R}\right), 0, -\frac{1}{\delta}\frac{2}{R}, 0, \dots, 0\right)$$

and we set $x^{\text{high}}(1) \doteq 0$. Immediately from the definition of the boost $\mathfrak{b}_\delta$ in Definition 6.3.3, we have (6.6.1).

Since $\Omega^2 = 1$ along C and $\hat{\chi}$ is compactly supported away from $v = 0$, we have $x^{\text{high}}(0) = 0$. The boost $\mathfrak{b}_\delta$ changes every positive power of δ on the left-hand side of (6.6.25) into a negative power, so that

$$\|\mathfrak{b}_\delta(x(0)) - \mathfrak{s}_{M_*,R}\|_{C^j} \lesssim_j \delta^{1/2}$$

for any $j \geq 0$. Therefore, taking j sufficiently large, we have

$$\|\mathfrak{b}_\delta(x(0)) - \mathfrak{s}_{M_*,R}\|_{\mathcal{X}^m} \lesssim \|\mathfrak{b}_\delta(x(0)) - \mathfrak{s}_{M_*,R}\|_{C^j} \lesssim \delta^{1/2},$$

where $\mathcal{X}^m$ is the sphere data norm appearing in Theorem 6.5.3. Now (6.6.2) follows follows by taking δ sufficiently small. □

6.6.2 Gluing Minkowski space to any round Schwarzschild sphere

Theorem 6.6.11. *Let $M > 0$, $R > 0$, and $k \in \mathbb{N}$. For any $\varepsilon > 0$ there exists a solution x of the null constraints on a null cone $C^{[0,1]}$ such that $x(1)$ equals $\mathfrak{s}_{M,R}$ after a boost and $x(0)$ can be realized as a sphere in Minkowski space in the following sense: There exists a C^k spacelike 2-sphere*

S in Minkowski space and a choice of L and $\underline{L}$ on S such that the induced $C^2_u C^k_v$ sphere data on this sphere equals $x(0)$ after a boost and a sphere diffeomorphism.

Remark 6.6.12. The sphere S can be made arbitrarily close to round and the sphere diffeomorphism can be made arbitrarily close to the identity.

Proof. By scaling, it suffices to prove the theorem when $R = 2$. We first use Proposition 6.6.1 to connect $\mathfrak{b}_\delta(\mathfrak{s}_{M,R})$ to the sphere data set $\mathfrak{b}_\delta(x(0))$ with $M_* = \varepsilon_\sharp^{1/4} \ll 1$. We now aim to use Theorem 6.5.3 to connect $x_2 \doteq \mathfrak{b}_\delta(x(0))$ to a sphere in Minkowski space. Let $\underline{x}$ be the usual ingoing Minkowski null data passing through the unit sphere at $u = 0$.[5]

By a direct computation, $\|\mathfrak{s}_{M_*,2} - \mathfrak{m}_2\|_{\mathcal{X}^m} \approx M_*$. It follows that

$$\|x_2 - \mathfrak{m}_2\|_{\mathcal{X}^m} < C_1 M_* \tag{6.6.27}$$

if $\varepsilon_\sharp$ is sufficiently small, where C_1 does not depend on $\varepsilon_\sharp$.

We must estimate the conserved charge deviation vector

$$(\Delta\mathbf{E}, \Delta\mathbf{P}, \Delta\mathbf{L}, \Delta\mathbf{G}) = (\mathbf{E}, \mathbf{P}, \mathbf{L}, \mathbf{G})(\mathfrak{b}_\delta(x(0))).$$

By (6.6.2),

$$|\mathbf{B}| + |\phi \, \mathrm{d}\!\!/\mathrm{iv}\, \mathbf{B}| \lesssim \varepsilon_\sharp.$$

We then compute

$$\begin{aligned} \phi^3 \left(K + \tfrac{1}{4} \operatorname{tr}\chi \operatorname{tr}\underline{\chi}\right) &= 2^3 \left(\frac{1}{2^2} + \frac{1}{4}\left(\frac{2}{2}\left(1 - \frac{2M_*}{2}\right)\right)\left(-\frac{2}{2}\right)\right) + O(\varepsilon_\sharp) \\ &= 2M_* + O(\varepsilon_\sharp), \end{aligned}$$

where $O(\varepsilon_\sharp)$ denotes a function all of whose angular derivatives are $\lesssim \varepsilon_\sharp$. It follows that for $\varepsilon_\sharp$ sufficiently small,

$$\Delta\mathbf{E} \geq \tfrac{3}{2} M_* > M_* \tag{6.6.28}$$

and

$$|\Delta\mathbf{P}| + |\Delta\mathbf{L}| + |\Delta\mathbf{G}| \lesssim \varepsilon_\sharp. \tag{6.6.29}$$

Let and C_* and ε_0 as in Theorem 6.5.3 for the choice $C_{\mathbf{E}} = C_1^{-1}$. For $0 < \varepsilon_\sharp < (\varepsilon_0/C_1)^4$, set

[5] Note that Theorem 6.5.3 was formulated for $C = [1,2]_v \times S^2$, but we are applying it on $C = [0,1]_v \times S^2$ here, which is merely a change of notation. We refer the reader back to Fig. 4.4 for the setup of this proof.

$\varepsilon_\flat = C_1 M_*$. Then (6.5.1) and (6.5.2) are satisfied,

$$|\Delta\mathbf{L}| < C\varepsilon_\sharp = CM_*^4 \leq (C_1 M_*)^2$$

if $\varepsilon_\sharp$ is sufficiently small, so (6.5.3) is satisfied, and finally

$$|\Delta\mathbf{P}| + |\Delta\mathbf{G}| < C\varepsilon_\sharp = CM_*^3 \cdot M_* \leq C_* \Delta\mathbf{E} \tag{6.6.30}$$

if $\varepsilon_\sharp$ is sufficiently small, so (6.5.4) is also satisfied.

By applying Theorem 6.5.3, we obtain a null data set for which the bottom sphere x_1' is a sphere diffeomorphism of a genuine Minkowski sphere data set and satisfies

$$\|x_1' - \mathfrak{m}_1\|_{\mathcal{X}} \lesssim \|x_2 - \mathfrak{m}_2\|_{\mathcal{X}} \lesssim \varepsilon_\sharp^{1/4}, \tag{6.6.31}$$

which can be made arbitrarily small and hence completes the proof of the theorem. □

6.6.3 Gluing Minkowski space to any Kerr coordinate sphere in very slowly rotating Kerr

In this section, we perform Kerr gluing for small angular momentum essentially as a corollary of the Schwarzschild work.

Theorem 6.6.13. *For any $k \in \mathbb{N}$, there exists a function $\mathfrak{a}_0 : (0,\infty)^2 \to (0,\infty)$ with the following property. Let $M > 0$ and $R > 0$. If $0 \leq |a| \leq \mathfrak{a}_0(M,R)M$, there exists a solution x of the null constraints on a null cone $C_0^{[0,1]}$ such that $x(1)$ equals $\mathfrak{k}_{M,a,R}$ after a boost and $x(0)$ can be realized as a sphere in Minkowski space in the following sense: There exists a C^k spacelike 2-sphere S in Minkowski space and a choice of L and $\underline{L}$ on S such that the induced $C_u^2 C_v^k$ sphere data on this sphere equals $x(0)$ after a boost and a sphere diffeomorphism.*

Proof. Again, it suffices to prove the theorem for $R = 2$ and $M > 0$ fixed but otherwise arbitrary. Let $x(v)$ and δ be the associated null data set and boost parameter constructed in Proposition 6.6.1, where M and R are as in the statement of the present theorem and $\varepsilon_\sharp$ is sufficiently small that the argument of Theorem 6.6.11 applies.

Let $\tilde{\gamma} \doteq 2^{-2}\not{g}_{M,a,2}$ and define $\hat{\not{g}}$ by (6.6.3) with $\lambda = \lambda_0(\delta)$ from Lemma 6.6.8. By (6.4.8), Cauchy stability for the proof of Proposition 6.6.1, and Corollary 6.5.5, we conclude that $\mathfrak{k}_{M,a,2}$ can be glued to Minkowski space as in Fig. 4.1 for a sufficiently small. □

6.7 Gravitational collapse to a Kerr black hole of prescribed mass and angular momentum

In this section we give the proof of Theorem 1.1.15 and the sketch of the proof of Corollary 4.6.5. Recall the fractional Sobolev spaces H^s, $s \in \mathbb{R}$, and their local versions H^s_{loc}. Recall also the notation $f \in H^{s-}_{\mathrm{loc}}$ which means $f \in H^{s'}_{\mathrm{loc}}$ for every $s' < s$.

We refer the reader back to Fig. 1.2 for the Penrose diagram associated to the following result.

Corollary 6.7.1. *There exists a constant $\mathfrak{a}_0 > 0$ such that the following holds. For any mass $M > 0$ and specific angular momentum a satisfying $a/M \in [-\mathfrak{a}_0, \mathfrak{a}_0]$, there exist one-ended asymptotically flat Cauchy data $(g_0, k_0) \in H^{7/2-} \times H^{5/2-}$ for the Einstein vacuum equations* (1.1.6) *on $\Sigma \cong \mathbb{R}^3$, satisfying the constraint equations*

$$R_{g_0} + (\operatorname{tr}_{g_0} k_0)^2 - |k_0|^2_{g_0} = 0 \text{ and} \tag{6.7.1}$$

$$\operatorname{div}_{g_0} k_0 - {}^{g_0}\nabla \operatorname{tr}_{g_0} k_0 = 0, \tag{6.7.2}$$

such that the maximal future globally hyperbolic development $(\mathcal{M}^4, g)$ has the following properties:

- *Null infinity $\mathcal{I}^+$ is complete.*
- *The black hole region is non-empty, $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+) \neq \emptyset$.*
- *The Cauchy surface Σ lies in the causal past of future null infinity, $\Sigma \subset J^-(\mathcal{I}^+)$. In particular, Σ does not intersect the event horizon $\mathcal{H}^+ \doteq \partial(\mathcal{BH})$ or contain trapped surfaces.*
- *$(\mathcal{M}, g)$ contains trapped surfaces.*
- *For sufficiently late advanced times $v \geq v_0$, the domain of outer communication, including the event horizon, is isometric to that of a Kerr solution with parameters M and a. For $v \geq v_0$, the event horizon of the spacetime can be identified with the event horizon of Kerr.*

Remark 6.7.2. The spacetime metric g is in fact C^2 everywhere away from the region labeled "Cauchy stablity" in Fig. 6.1 below. Near the set $\dot{\mathcal{H}}^+_-$ (see [HE73, p. 187] for notation), the spacetime metric might fail to be C^2, but is consistent with the regularity of solutions constructed in [HKM76] with $s = \frac{7}{2}-$. See also [Chr13] for the notion of the maximal globally hyperbolic development in low regularity.

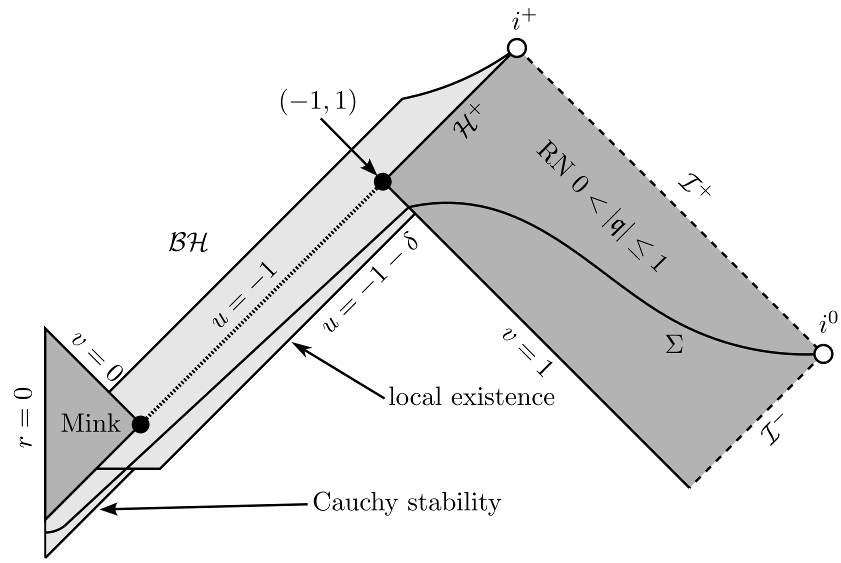

Figure 6.1: Penrose diagram for the proof of Theorem 1.1.15. The diagram does not faithfully represent the geometry of the spacetime near the "bottom" of the event horizon $\dot{\mathcal{H}}^+_-$ (i.e., the locus where the null geodesic generators "end"). The event horizon does not necessarily end in a point since the distinguished Minkowski sphere is not necessarily round.

Proof. We refer the reader to Fig. 6.1 for the anatomy of the proof, which is essentially the same as Corollary 5.6.1. The region to the left of $\mathcal{H}^+$ is constructed using our gluing theorem, Theorem 6.6.13, and local existence (in this case we appeal to [Luk12]). The region to the right of the horizon, save for the part labeled "Cauchy stability" in Fig. 6.1, is constructed in the same manner. These two regions can now be pasted along $u=0$ and the resulting spacetime will be C^2.

We can now use a Cauchy stability argument to construct the remainder of the spacetime. A very similar argument in carried out in Lemma 5.6.3, but the lower regularity of our gluing result in the present proof forces us to use slightly more technology here. As in Lemma 5.6.3, we take the induced data (g_*,k_*) on a suitably chosen spacelike hypersurface Σ_* passing through the bottom gluing sphere. See Fig. 6.1. This data lies in the regularity class $H^{7/2-}_{\mathrm{loc}}\times H^{5/2-}_{\mathrm{loc}}$ by part (i) of Lemma 6.7.3 below and satisfies the constraint equations. The cutoff argument presented in Lemma 5.6.3 goes through using (ii) of Lemma 6.7.3 and the low regularity well-posedness theory in [HKM76]. Note that for simplicity we have applied well-posedness in the class $H^{3-}_{\mathrm{loc}}\times H^{2-}_{\mathrm{loc}}$ because of a loss of half a derivative in our Hardy inequality argument below, but since (g_*,k_*) actually lies in the better space $H^{7/2-}_{\mathrm{loc}}\times H^{5/2-}_{\mathrm{loc}}$, the spacetime metric has regularity consistent with $H^{7/2-}_{\mathrm{loc}}\times H^{5/2-}_{\mathrm{loc}}$ initial data by propagation of regularity.

To show that $(\mathcal{M},g)$ contains trapped surfaces, it suffices to observe that $\underline{D}(\Omega\operatorname{tr}\chi)<0$ on

$\mathfrak{k}_{M_f,a_f,r_+}$ for $\mathfrak{a}_0$ sufficiently small by (6.4.8) and (6.2.15).[6]

Having constructed the spacetime, we can finally extract a Cauchy hypersurface Σ, which completes the proof. ☐

Lemma 6.7.3. *Let f and g be functions defined on $B_2 \subset \mathbb{R}^3$, the ball of radius two, such that $f \in C^2(B_2)$, $f|_{B_1} \in C^3(\overline{B_1})$, $f|_{B_2\setminus\overline{B_1}} \in C^3(\overline{B_2} \setminus B_1)$, $g \in C^1(B_2)$, $g|_{B_1} \in C^2(\overline{B_1})$, and $g|_{B_2\setminus\overline{B_1}} \in C^2(\overline{B_2} \setminus B_1)$. Then:*

(i) $(f, g) \in H^{7/2-} \times H^{5/2-}(B_2)$.

(ii) Suppose that $f = g = 0$ identically on B_1. For $0 < \varepsilon < \frac{1}{2}$, let θ_ε be a cutoff function which is equal to one on $B_{1+\varepsilon}$ and zero outside of $B_{2+\varepsilon}$. Then $f_\varepsilon \doteq \theta_\varepsilon f$ and $g_\varepsilon \doteq \theta_\varepsilon g$ satisfy

$$\lim_{\varepsilon\to 0} \|(f_\varepsilon, g_\varepsilon)\|_{H^s\times H^{s-1}(B_2)} = 0 \tag{6.7.3}$$

for any $s < 3$.

Proof. The proof of (i) follows in a straightforward manner from the physical space characterization of fractional Sobolev spaces (such as in [DPV12]) and is effectively an elaboration of the fact that the characteristic function of B_1 lies in $H^{1/2-}$.

Proof of (ii): Using Taylor's theorem as in Lemma 5.6.3, we see that $\|(f_\varepsilon, g_\varepsilon)\|_{H^2\times H^1(B_2)} \to 0$ as $\varepsilon \to 0$. By iterating Hardy's inequality, we see that

$$\int_{B_2} f^2|\partial^3\theta_\varepsilon|^2\,dx \lesssim \int_{B_{1+2\varepsilon}\setminus B_{1+\varepsilon}} \frac{f^2}{\varepsilon^6}\,dx \lesssim \int_{B_2} f^2 + |\partial f|^2 + |\partial^2 f|^2 + |\partial^3 f|^2\,dx,$$

so $(f_\varepsilon, g_\varepsilon)$ is bounded in $H^3 \times H^2$. We now obtain (6.7.3) by interpolation. ☐

Remark 6.7.4. In fact (6.7.3) holds for $s < \frac{7}{2}$, but this requires a fractional Hardy inequality.

We now sketch the proof of Corollary 4.6.5 and refer the reader back to Fig. 4.7 for the associated Penrose diagram.

Sketch of the proof of Corollary 4.6.5. Using Theorem 6.6.11, we glue Minkowski space to a round Schwarzschild sphere of mass 1 and radius $R = 2 - \varepsilon$ for $0 \le \varepsilon \ll 1$. As $\varepsilon \to 0$ (perhaps only along a subsequence $\varepsilon_j \to 0$), the gluing data converge to the horizon gluing data used in the proof of Corollary 6.7.1, in an appropriate norm. It then follows by Cauchy stability that the spacetimes

[6] For convenience, we have deduced the presence of trapped surfaces in very slowly rotating Kerr perturbatively from Schwarzschild. However, it is well known that any subextremal Kerr black hole contains trapped surfaces right behind the event horizon, and one may invoke that fact instead since the spacetime metric constructed here is C^2 across the event horizon $\mathcal{H}^+$.

constructed by solving backwards as in the proof of Corollary 6.7.1 contain the full event horizon, for ε sufficiently small. $\square$

Chapter 7

Revisiting the charged Vaidya metric

In this chapter, we introduce Ori's *bouncing charged null dust model* [Ori91]. We then show that Ori's model exhibits extremal critical collapse and can be used to construct counterexamples to the third law of black hole thermodynamics. Later, in Chapter 8, we will then show that these dust solutions can be (in an appropriate sense) globally desingularized by passing to smooth solutions of the Einstein–Maxwell–Vlasov system. The constructions in this section are crucial to motivate the choice of initial data in the proof of Theorem 1.2.1 in Chapter 8.

7.1 Ori's bouncing charged null dust model

We begin by recalling the general notion of charged null dust from [Ori91]:

Definition 7.1.1. The *Einstein–Maxwell–charged null dust model* for particles of *fundamental charge* $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$ consists of a charged spacetime $(\mathcal{M}, g, F)$, a future-directed null vector field k representing the momentum of the dust particles, and a nonnegative function ρ which describes the energy density of the dust. The equations of motion are

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = 2\left(T^{\mathrm{EM}}_{\mu\nu} + T_{\mu\nu}\right), \tag{7.1.1}$$

$$\nabla^\alpha F_{\mu\alpha} = \mathfrak{e}\rho k_\mu, \tag{7.1.2}$$

$$k^\nu \nabla_\nu k^\mu = \mathfrak{e}F^\mu{}_\nu k^\nu, \tag{7.1.3}$$

$$\nabla_\mu(\rho k^\mu) = 0, \tag{7.1.4}$$

where $T^{\mathrm{EM}}_{\mu\nu}$ was defined in (2.1.17) and $T_{\mu\nu} \doteq \rho k_\mu k_\nu$ is the energy-momentum tensor of a pressureless perfect fluid. By the forced Euler equation (7.1.3), the integral curves of k are electromagnetic null geodesics.

Any two functions $\varpi_{\mathrm{in}}, Q_{\mathrm{in}} \in C^\infty(\mathbb{R})$ determine a spherically symmetric solution to the system (7.1.1)–(7.1.4) by the formulas

$$g_{\mathrm{in}}[\varpi_{\mathrm{in}}, Q_{\mathrm{in}}] \doteq -D(V,r)\,dV^2 + 2\,dV dr + r^2\gamma, \tag{7.1.5}$$

$$F \doteq -\frac{Q_{\mathrm{in}}}{r^2}\,dV \wedge dr, \tag{7.1.6}$$

$$k \doteq \frac{\mathfrak{e}}{\dot{Q}_{\mathrm{in}}}\left(\dot{\varpi}_{\mathrm{in}} - \frac{Q_{\mathrm{in}}\dot{Q}_{\mathrm{in}}}{r}\right)(-\partial_r), \quad \rho \doteq \frac{(\dot{Q}_{\mathrm{in}})^2}{\mathfrak{e}^2 r^2}\left(\dot{\varpi}_{\mathrm{in}} - \frac{Q_{\mathrm{in}}\dot{Q}_{\mathrm{in}}}{r}\right)^{-1}, \tag{7.1.7}$$

where $\cdot$ denotes differentiation with respect to V and

$$D(V,r) \doteq 1 - \frac{2\varpi_{\mathrm{in}}(V)}{r} + \frac{Q_{\mathrm{in}}(V)^2}{r^2}.$$

The metric (7.1.5) is known as the *ingoing charged Vaidya metric* [PS68; BV70] and describes a "time dependent" Reissner–Nordström spacetime in ingoing Eddington–Finkelstein-type coordinates $(V, r, \vartheta, \varphi)$. The spacetime is time oriented by $-\partial_r$. The metric (7.1.5) and Maxwell field (7.1.6) are spherically symmetric and may therefore be considered as a spherically symmetric charged spacetime in the framework of Section 2.1. One easily sees that $D = 1 - \frac{2m}{r}$, $Q = Q_{\mathrm{in}}$, and $\varpi = \varpi_{\mathrm{in}}$.

We will always make the assumption $\dot{\varpi}_{\mathrm{in}} \geq 0$ so that $T^{\mu\nu} = \rho k^\mu k^\nu$ satisfies the weak energy condition for r sufficiently large. We also assume that $\mathfrak{e} > 0$ and impose the condition $\dot{Q}_{\mathrm{in}} \geq 0$ on the seed function Q_{in}, which just means that positively charged particles increase the charge of the spacetime. (If $\mathfrak{e} < 0$, we would instead assume $\dot{Q}_{\mathrm{in}} \leq 0$ and the discussion would otherwise remain unchanged.)

We define a function $r_\mathrm{b} = r_\mathrm{b}(V)$, called the *bounce radius*, by

$$r_\mathrm{b} \doteq \frac{Q_\mathrm{in}\dot{Q}_\mathrm{in}}{\dot{\varpi}_\mathrm{in}}$$

whenever $\dot{\varpi}_\mathrm{in}(V) > 0$. The reason for this terminology will become clear shortly. By inspection of (7.1.7), we observe the following: For $r > r_\mathrm{b}(V)$, $(g_\mathrm{in}, F, k, \rho)$ defines a solution of the Einstein–Maxwell-charged null dust system, k is future-directed null, and $\rho \geq 0$. If $r_\mathrm{b}(V) > 0$ and $r \searrow r_\mathrm{b}(V)$, then k and T^dust vanish. If also $\dot{Q}_\mathrm{in}(V) > 0$, then ρ blows up at $r = r_\mathrm{b}(V)$, but ρk is nonzero and bounded. Finally, for $r < r_\mathrm{b}(V)$, k is *past-directed* null and $\rho < 0$, so T^dust violates the weak energy condition.

Physically, the ingoing Vaidya metric and (7.1.7) describe an ingoing congruence of radial charged massless dust particles which interact with the electromagnetic field that they generate. One can interpret the vanishing of k as the dust being "stopped" by the resulting repulsive Lorentz force. Integral curves of k are ingoing radial electromagnetic null geodesics $\gamma(s)$ with limit points on the *bounce hypersurface* $\Sigma_\mathrm{b} \doteq \{r = r_\mathrm{b}\}$ as $s \to \infty$. The charged null dust system is actually *ill-posed* across Σ_b since the transport equation (7.1.3) breaks down there. Because of this, Ori argued in [Ori91] that the ingoing charged Vaidya metric (7.1.5) (and the associated formulas in (7.1.7)) should only be viewed as physical to the *past* of Σ_b and must be modified if we wish to continue the solution beyond Σ_b.

Remark 7.1.2. The divergence of ρ along Σ_b does not seem to have been explicitly mentioned by Ori, but it is one of the fundamentally singular features of charged null dust. One can also see that ρ can blow up if $\dot{Q}_\mathrm{in}/Q_\mathrm{in}$ blows up as a function of V, which occurs if the dust is injected into Minkowski space.

Remark 7.1.3. Before Ori's paper [Ori91], the "standard interpretation" [SI80; LZ91] of the ingoing Vaidya metric (7.1.5) did not actually involve Maxwell's equation and the fluid equation was simply taken to be the standard geodesic equation. The set $\{r < r_\mathrm{b}\}$ was included in the ingoing solution and the dust was thought to violate the weak energy condition in this region. We refer to [Ori91] for discussion.

In order to continue the dust solution across Σ_b, we must make some further (nontrivial!) assumptions on the seed functions ϖ_in and Q_in. In order to not trivially violate causality, we must demand that Σ_b is *spacelike*, so that the "other side" $\{r < r_\mathrm{b}\}$ of Σ_b does not intersect the past of Σ_b. This is equivalent to

$$D - 2\dot{r}_\mathrm{b} < 0 \quad \text{on } \Sigma_\mathrm{b}. \tag{7.1.8}$$

We further assume that $\Sigma_{\rm b}$ does not contain trapped symmetry spheres, which is equivalent to

$$D > 0 \quad \text{on } \Sigma_{\rm b}. \tag{7.1.9}$$

By examining the behavior of almost-radial electromagnetic null geodesics in Reissner–Nordström, Ori proposed the following *bouncing* continuation of the solution through $\Sigma_{\rm b}$: it should be as an *outgoing* charged Vaidya metric. This metric takes the form

$$g_{\rm out}[\varpi_{\rm out}, Q_{\rm out}] \doteq -\underline{D}(U,r)dU^2 - 2\,dUdr + r^2\gamma, \tag{7.1.10}$$

where

$$\underline{D}(U,r) \doteq 1 - \frac{2\varpi_{\rm out}(U)}{r} + \frac{Q_{\rm out}(U)^2}{r^2},$$

for free functions $\varpi_{\rm out}$ and $Q_{\rm out}$. The coordinates $(U, r, \vartheta, \varphi)$ are now *outgoing* Eddington–Finkelstein-like. Ori defined a procedure for gluing an outgoing Vaidya metric to the ingoing Vaidya metric along $\Sigma_{\rm b}$ by demanding continuity of the second fundamental form of $\Sigma_{\rm b}$ from both sides. One sets

$$(\varpi_{\rm out}, Q_{\rm out})(U) = (\varpi_{\rm in}, Q_{\rm in}) \circ \mathcal{G}^{-1}(U),$$

where the *gluing map* $\mathcal{G} = \mathcal{G}(V)$ is determined by

$$\frac{d\mathcal{G}}{dV} = \frac{D(V, r_{\rm b}(V)) - 2\dot{r}_{\rm b}(V)}{D(V, r_{\rm b}(V))}, \tag{7.1.11}$$

up to specification of the (unimportant) initial condition. Notice that $\mathcal{G}$ is strictly monotone decreasing on account of (7.1.8) and (7.1.9). It turns out that this continuation *preserves the weak energy condition* through $\Sigma_{\rm b}$. We formalize this choice of extension of the Vaidya metric with the following

Definition 7.1.4. Let $\varpi_{\rm in}$ and $Q_{\rm in}$ be nondecreasing charged Vaidya seed functions such that $\operatorname{spt}(\dot{\varpi}_{\rm in}) = \operatorname{spt}(\dot{Q}_{\rm in}) = [V_1, V_2]$ and $r_{\rm b}$ is well-defined and positive on $[V_1, V_2]$. Assume also the conditions (7.1.8) and (7.1.9). *Ori's bouncing charged null dust model* consists of the ingoing charged Vaidya metric $g_{\rm in}[\varpi_{\rm in}, Q_{\rm in}]$ on $\mathcal{M}_{\rm in} \doteq \{V \in \operatorname{spt}(\dot{\varpi}_{\rm in}), r \geq r_{\rm b}(V)\} \times S^2$ with spacelike, untrapped bounce hypersurface $\Sigma_{\rm b}^{\rm in} \doteq \{V \in \operatorname{spt}(\dot{\varpi}_{\rm in}), r = r_{\rm b}(V)\} \times S^2$ glued to the outgoing charged Vaidya metric $g_{\rm out}[\varpi_{\rm in} \circ \mathcal{G}^{-1}, Q_{\rm in} \circ \mathcal{G}^{-1}]$ on $\mathcal{M}_{\rm out} \doteq \{U \in \operatorname{spt}(\dot{\varpi}_{\rm out}), r \geq r_{\rm b} \circ \mathcal{G}^{-1}(U)\} \times S^2$ with spacelike, untrapped bounce hypersurface $\Sigma_{\rm b}^{\rm out} \doteq \{U \in \operatorname{spt}(\dot{\varpi}_{\rm out}), r = r_{\rm b}(\mathcal{G}^{-1}(U))\} \times S^2$ along the map

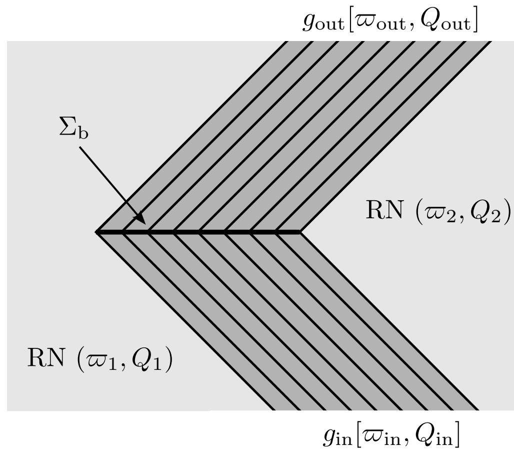

Figure 7.1: Penrose diagram of Ori's bouncing charged null dust model. The geometry of the beams is described by the ingoing and outgoing Vaidya metrics, g_{in} and g_{out}, which are related by the gluing map $\mathcal{G}$. The spacetime to the left and right of the bouncing beam is described by the Reissner–Nordström solution, with parameters (ϖ_1, Q_1) and (ϖ_2, Q_2) with $\varpi_1 < \varpi_2$ and $Q_1 < Q_2$. The endpoints of Σ_{b} correspond to symmetry spheres in these Reissner–Nordström spacetimes with radii $r_1 < r_2$. In this diagram, the V coordinate is normalized according to the ingoing solution. We have depicted here the case of a *totally geodesic* bounce hypersurface Σ_{b} and the outgoing beam is exactly the time-reflection of the ingoing beam.

$\mathcal{G} \times \mathrm{id}_r : \Sigma_{\text{b}}^{\text{in}} \to \Sigma_{\text{b}}^{\text{out}}$ defined by (7.1.11). Outside the support of the dust, Ori's bouncing charged null dust model extends by attaching two Reissner–Nordström solutions with parameters $(\varpi_1, Q_1) \doteq (\varpi_{\text{in}}, Q_{\text{in}})(V_1)$ and $(\varpi_2, Q_2) \doteq (\varpi_{\text{in}}, Q_{\text{in}})(V_2)$ as depicted in Fig. 7.1.

The model can be generalized to allow for multiple beams of dust by iterating the above definition in the obvious manner.

7.2 The radial parametrization of bouncing charged null dust spacetimes

It is not immediately clear that interesting seed functions ϖ_{in} and Q_{in} satisfying the requirements of Definition 7.1.4 exist. Therefore, it is helpful to *directly prescribe* the geometry of Σ_{b} and the dust along it in the following sense. Given a spacetime as in Fig. 7.1, we can parametrize Σ_{b} by the area-radius function r. Then the renormalized Hawking mass ϖ and charge Q, which are gauge-invariant quantities, can be viewed as functions of r on Σ_{b}, and we wish to prescribe these functions. We will also prescribe Σ_{b} to be *totally geodesic*. While not essential, this condition greatly simplifies Proposition 7.2.3 below and will later play a key role in our Vlasov construction in Chapter 8.

Definition 7.2.1. Let $\mathcal{P}$ denote the set of points $(r_1, r_2, \varpi_1, \varpi_2, Q_1, Q_2) \in \mathbb{R}^6_{\geq 0}$ subject to the conditions

$$0 < r_1 < r_2, \quad Q_1 < Q_2, \quad \varpi_1 < \varpi_2 \tag{7.2.1}$$

$$2r_1\varpi_1 \geq Q_1^2, \tag{7.2.2}$$

$$2r_1(\varpi_2 - \varpi_1) < Q_2^2 - Q_1^2 < 2r_2(\varpi_2 - \varpi_1), \tag{7.2.3}$$

$$\min_{r\in[r_1,r_2]}(r_1r^2 - 2\varpi_1 r_1 r + Q_1^2 r - Q_2^2 r + Q_2^2 r_1) > 0. \tag{7.2.4}$$

Elements of $\mathcal{P}$ will typically be denoted by the letter α and are called *admissible parameters.* Let $\mathfrak{V}$ denote the set of triples $(\alpha, \check{\varpi}, \check{Q}) \in \mathcal{P} \times C^\infty\big([0,\infty)\big) \times C^\infty\big([0,\infty)\big)$ such that the functions $\check{\varpi} = \check{\varpi}(r)$ and $\check{Q} = \check{Q}(r)$ are monotone increasing and satisfy

$$\operatorname{spt}(\check{\varpi}') = \operatorname{spt}(\check{Q}') = [r_1, r_2], \tag{7.2.5}$$

$$\frac{d}{dr}\check{Q}^2(r) = 2r\frac{d}{dr}\check{\varpi}(r), \tag{7.2.6}$$

$$\check{\varpi}(r_1) = \varpi_1, \quad \check{Q}(r_1) = Q_1, \quad \check{\varpi}(r_2) = \varpi_2, \quad \check{Q}(r_2) = Q_2, \tag{7.2.7}$$

where $'$ denotes differentiation with respect to r.

Remark 7.2.2. In the proof of Theorem 1.2.1 we will employ the *regular center* parameter space $\mathcal{P}_\Gamma$, consisting of those $\alpha \in \mathcal{P}$ with $\varpi_1 = Q_1 = 0$.

Proposition 7.2.3 (Radial parametrization of bouncing charged null dust)**.** *Let $(\alpha, \check{\varpi}, \check{Q}) \in \mathfrak{V}$, and define strictly monotone functions $\mathcal{V}, \mathcal{U} : [r_1, r_2] \to \mathbb{R}$ by*

$$\mathcal{V}(r) = -\mathcal{U}(r) = \int_{r_1}^{r} D(r')^{-1}\, dr',$$

where $D(r) \doteq 1 - \frac{2\check{\varpi}(r)}{r} + \frac{\check{Q}^2(r)}{r^2}$. Then:

1. *The seed functions $(\varpi_{\mathrm{in}}, Q_{\mathrm{in}}) \doteq (\check{\varpi}, \check{Q})\circ\mathcal{V}^{-1}$ and $(\varpi_{\mathrm{out}}, Q_{\mathrm{out}}) \doteq (\check{\varpi}, \check{Q})\circ\mathcal{U}^{-1}$ define a bouncing charged null dust spacetime as in Definition 7.1.4 with gluing map $\mathcal{G}(V) = -V$ and bounce radius $r_{\mathrm{b}}(V) = \mathcal{V}^{-1}(V)$.*

2. *The bounce hypersurface Σ_{b} is spacelike and untrapped. With the setup as in Fig. 7.1, the left edge of Σ_{b} has area-radius r_1 and Reissner–Nordström parameters (ϖ_1, Q_1) and the right edge has area-radius r_2 and Reissner–Nordström parameters (ϖ_2, Q_2). The Hawking mass m is nonnegative on Σ_{b}.*

3. The bounce hypersurface $\Sigma_{\rm b}$ is totally geodesic with respect to $g_{\rm in}$ and $g_{\rm out}$.

Proof. We must check that $\varpi_{\rm in} \doteq \check\varpi \circ \mathcal{V}^{-1}$ and $Q_{\rm in} \doteq \check Q \circ \mathcal{V}^{-1}$ satisfy the assumptions of Definition 7.1.4. Using the chain rule and (7.2.6), we compute

$$r_{\rm b}(V) = \frac{\check Q(\mathcal{V}^{-1}(V))\check Q'(\mathcal{V}^{-1}(V))}{\check\varpi'(\mathcal{V}^{-1}(V))} = \mathcal{V}^{-1}(V).$$

Differentiating, we obtain

$$\dot r_{\rm b}(V) = D(\mathcal{V}^{-1}(V)) = D(V, r_{\rm b}(V)), \tag{7.2.8}$$

which implies that $d\mathcal{G}/dV = -1$. To prove (7.1.9), we show that $D(r) > 0$ for $r \in [r_1, r_2]$. Integrating (7.2.6) in r and integrating by parts yields

$$\check\varpi(r) = \varpi_1 + \frac{1}{2}\int_{r_1}^{r} \frac{1}{r'}\frac{d}{dr'}\check Q^2(r')\,dr' = \varpi_1 + \frac{\check Q^2(r)}{2r} - \frac{Q_1^2}{2r_1} + \frac{1}{2}\int_{r_1}^{r}\frac{\check Q^2(r')}{r'^2}\,dr'. \tag{7.2.9}$$

Using condition (7.2.4) and $\check Q \le Q_2$, we then find

$$\begin{aligned} D(r) = 1 - \frac{2\check\varpi(r)}{r} + \frac{\check Q^2(r)}{r^2} &= 1 - \frac{2\varpi_1}{r} + \frac{Q_1^2}{r_1 r} - \frac{1}{r}\int_{r_1}^{r}\frac{\check Q^2(r')}{r'^2}\,dr' \\ &> \frac{1}{r_1 r^2}\left(r_1 r^2 - 2\varpi_1 r r_1 + Q_1^2 r - Q_2^2 r + Q_2^2 r_1\right) > 0 \end{aligned} \tag{7.2.10}$$

for $r \in [r_1, r_2]$. This proves (7.1.9) and since $D - 2\dot r_{\rm b} = -D$, also proves (7.1.8). Condition (7.2.2) implies that the Hawking mass is nonnegative at r_1. Finally, that $\Sigma_{\rm b}$ is a totally geodesic hypersurface is shown by directly computing its second fundamental form and using (7.2.8). □

The definition of $\mathfrak{V}$ involves many more conditions than just (7.1.8) and (7.1.9) alone, but it turns out that these are relatively easy to satisfy. In particular, we have:

Proposition 7.2.4. *The natural projection map $\mathfrak{V} \to \mathcal{P}$ admits a smooth section $\varsigma : \mathcal{P} \to \mathfrak{V}$. In other words, given any smooth family of parameters in $\mathcal{P}$ we may associate a smooth family of bouncing charged null dust spacetimes attaining those parameters, with totally geodesic bounce hypersurfaces.*

Remark 7.2.5. In the remainder of the dissertation (in particular, Chapter 8), we fix the choice of section to be the one constructed in the proof below.

Proof. Define a smooth, surjective function $\psi : \mathbb{R}^2 \to (0,1)$ by

$$\psi(x,\xi) = \frac{1}{1 + \exp\left[-(x-\xi)e^{(x-\xi)^2}\right]}.$$

Note that for each fixed $\xi \in \mathbb{R}$ the function $x \mapsto \psi(x,\xi)$ is strictly monotone increasing and surjective. Moreover, for $x \in \mathbb{R}$ we have $\psi(x,\xi) \to 0$ as $\xi \to \infty$ and $\psi(x,\xi) \to 1$ as $\xi \to -\infty$.

We now define the function $\check{Q} : (r_1, r_2) \times \mathbb{R} \times \mathcal{P} \to \mathbb{R}$ by

$$\check{Q}(r,\xi,\alpha) \doteq Q_1 + (Q_2 - Q_1)\,\psi\left(\log\left(\frac{r - r_1}{r_2 - r}\right), \xi\right). \tag{7.2.11}$$

By construction of ψ, the function $\check{Q}$ extends smoothly to $[0,\infty) \times \mathbb{R} \times \mathcal{P}$ by setting $\check{Q}(r,\xi,\alpha) = Q_1$ for $0 \le r \le r_1$ and $\check{Q}(r,\xi,\alpha) = Q_2$ for $r \ge r_2$.

With our family of candidate $\check{Q}$'s at hand, we aim to satisfy the constraint $\check{\varpi}(r_2) = \varpi_2$, where $\check{\varpi}(r)$ is defined by (7.2.9). Consider the smooth map $\Pi : (\xi,\alpha) \in \mathbb{R} \times \mathcal{P} \to \mathbb{R}$ defined by

$$\Pi(\xi,\alpha) \doteq \varpi_1 + \frac{Q_2^2}{2r_2} - \frac{Q_1^2}{2r_1} + \int_{r_1}^{r_2} \frac{\check{Q}^2(r',\xi,\alpha)}{2r'^2}\,dr'.$$

Since ψ satisfies $\frac{\partial\psi}{\partial\xi} < 0$ on $\mathbb{R}^2$, we have that $\frac{\partial\Pi}{\partial\xi} < 0$ on $\mathbb{R} \times \mathcal{P}$. Moreover, using the pointwise limits of ψ, a direct computation gives

$$\lim_{\xi\to\infty} \Pi(\alpha,\xi) = \varpi_1 + \frac{Q_2^2}{2r_2} - \frac{Q_1^2}{2r_2}, \qquad \lim_{\xi\to-\infty} \Pi(\alpha,\xi) = \varpi_1 + \frac{Q_2^2}{2r_1} - \frac{Q_1^2}{2r_1}.$$

By condition (7.2.3), this implies that

$$\lim_{\xi\to\infty} \Pi(\alpha,\xi) < \varpi_2 < \lim_{\xi\to-\infty} \Pi(\alpha,\xi).$$

Thus, the intermediate value theorem and the fact that $\frac{\partial\Pi}{\partial\xi} < 0$ show that there exists a unique $\xi(\alpha)$ such that $\Pi(\alpha,\xi(\alpha)) = \varpi_2$. Moreover, a direct consequence of the implicit function theorem is that the assignment $\mathcal{P} \ni \alpha \mapsto \xi(\alpha) \in \mathbb{R}$ is smooth. The above construction shows that the functions $\check{Q}(r,\xi(\alpha),\alpha)$ and $\check{\varpi}$ satisfy all required properties. □

The set $\mathcal{P}$ is defined by simple polynomial relations and includes many interesting examples as we will see in the next two sections.

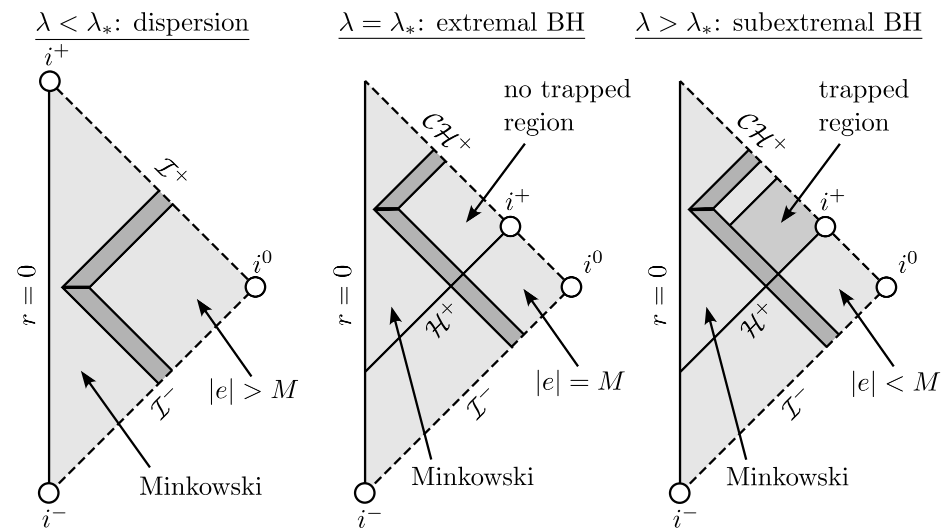

Figure 7.2: Penrose diagrams of extremal critical collapse in Ori's bouncing charged null dust model. Compare with Fig. 1.6. In Theorem 7.3.2, $\lambda_* = 1$.

7.3 Extremal critical collapse in Ori's model

The first application of Propositions 7.2.3 and 7.2.4 is the construction of one-parameter families of bouncing charged null dust spacetimes exhibiting *extremal critical collapse*. We first show that the regular center parameter space $\mathcal{P}_\Gamma$ contains elements with arbitrary final Reissner–Nordström parameters:

Lemma 7.3.1. *Let $\varpi_2, Q_2 > 0$. Then there exist $0 < r_1 < r_2$ such that $(r_1, r_2, 0, \varpi_2, 0, Q_2) \in \mathcal{P}_\Gamma$. If $\varpi_2 \geq Q_2$, then r_2 can moreover be chosen so that $r_2 < \varpi_2 - \sqrt{\varpi_2^2 - Q_2^2}$.*

Proof. Let

$$r_1 \doteq Q_2 \left(\frac{Q_2}{2\varpi_2} - \varepsilon \right), \quad r_2 \doteq Q_2 \left(\frac{Q_2}{2\varpi_2} + \varepsilon \right),$$

where $\varepsilon > 0$ is a small parameter to be determined. With this choice, (7.2.3) is clearly satisfied. Let $p(r) \doteq r_1 r^2 - Q_2^2 r + Q_2^2 r_1$ and observe that

$$\lim_{\varepsilon \to 0} p\left(\frac{Q_2^2}{2\varpi_2} \right) = \frac{Q_2^6}{8\varpi_2^3} > 0.$$

It follows that (7.2.4) is satisfied for ε sufficiently small. If $x \geq 1$, then $(2x)^{-1} < x - \sqrt{x^2 - 1}$, so taking ε perhaps smaller ensures that $r_2 < \varpi_2 - \sqrt{\varpi_2^2 - Q_2^2}$. □

Using this, we can show that Ori's model exhibits extremal critical collapse. Compare the following theorem with Theorem 1.2.1 and refer to Fig. 7.2 for Penrose diagrams.

Theorem 7.3.2. *For any $M > 0$ and fundamental charge $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$, there exist a small parameter $\delta > 0$ and a smooth two-parameter family of regular center parameters $\{\alpha_{\lambda,M'}\} \subset \mathcal{P}_\Gamma$ for $\lambda \in (0,2]$ and $M' \in [M-\delta, M+\delta]$ such that the two-parameter family of bouncing charged null dust spacetimes $\{\mathcal{D}_{\lambda,M'}\}$, obtained by applying Proposition 7.2.3 to $\varsigma(\alpha_{\lambda,M'})$, has the following properties:*

1. *For $0 < \lambda < 1$, $\mathcal{D}_{\lambda,M'}$ is isometric to Minkowski space for all sufficiently late retarded times u and hence future causally geodesically complete. In particular, it does not contain a black hole or naked singularity, and for $\lambda < 1$ sufficiently close to 1, sufficiently large advanced times $v \geq v_0$ and sufficiently small retarded times $u \leq u_0$, the spacetime is isometric to an appropriate causal diamond in a superextremal Reissner–Nordström solution. Moreover, $\mathcal{D}_{\lambda,M'}$ converges smoothly to Minkowski space as $\lambda \to 0$.*

2. *$\lambda = 1$ is critical: $\mathcal{D}_{1,M'}$ contains a nonempty black hole region $\mathcal{BH}$ and for sufficiently large advanced times $v \geq v_0$, the domain of outer communication, including the event horizon $\mathcal{H}^+$, is isometric to that of an extremal Reissner–Nordström solution of mass M'. The spacetime contains no trapped surfaces.*

3. *For $1 < \lambda \leq 2$, $\mathcal{D}_{\lambda,M'}$ contains a nonempty black hole region $\mathcal{BH}$ and for sufficiently large advanced times $v \geq v_0$, the domain of outer communication, including the event horizon $\mathcal{H}^+$, is isometric to that of a subextremal Reissner–Nordström solution. The spacetime contains an open set of trapped surfaces.*

In addition, for every $\lambda \in [0,2]$, $\mathcal{D}_{\lambda,M'}$ is isometric to Minkowski space for sufficiently early advanced time and near the center $\{r = 0\}$ for all time, and possesses complete null infinities $\mathcal{I}^+$ and $\mathcal{I}^-$.

Proof. Using Lemma 7.3.1, choose $0 < r_1 < r_2 < r_-(4M, 2M)$ such that $(r_1, r_2, 0, 4M, 0, 2M) \in \mathcal{P}_\Gamma$. We consider

$$\alpha_{\lambda,M'} \doteq (r_1, r_2, 0, \lambda^2 M', 0, \lambda M') \tag{7.3.1}$$

and note that $\alpha_{\lambda,M'}$ lies in $\mathcal{P}_\Gamma$ for $|\lambda - 2|$ sufficiently small and $|M - M'| \leq \delta$ sufficiently small by the openness of the conditions defining $\mathcal{P}_\Gamma$. Moreover, from the scaling properties of (7.2.3) and the monotonicity of (7.2.4), we observe that $\alpha_{\lambda,M'} \in \mathcal{P}_\Gamma$ for all $0 < \lambda \leq 2$ and $|M - M'| \leq \delta$.

After applying Proposition 7.2.3 to $\varsigma(\alpha_{\lambda,M'})$ for $\lambda > 0$, it remains only to show that $\mathcal{D}_{\lambda,M'}$ extends smoothly to Minkowski space as $\lambda \to 0$. Indeed, a direct inspection of the proof of Proposition 7.2.4 shows that $\xi(\alpha_{\lambda,M'})$ is independent of λ, so that the function $r \mapsto \check{Q}(r, \xi(\alpha_{\lambda,M'}), \alpha_{\lambda,M'})$ defined in (7.2.11) converges smoothly to the function $\check{Q} \equiv 0$ as $\lambda \to 0$. Therefore, $\check{\varpi}$ also converges smoothly to the zero function and hence $\mathcal{D}_{\lambda,M'}$ converges smoothly to Minkowski space as $\lambda \to 0$. $\square$

The construction in the proof of Theorem 1.2.1 can be thought of as a global-in-time desingularization of this family of dust solutions. In fact, we will make essential use of the one-parameter family $\{\varsigma(\alpha_{\lambda,M'})\}$ when constructing initial data for the Einstein–Maxwell–Vlasov system.

7.4 A counterexample to the third law of black hole thermodynamics in Ori's model

Using the radial parametrization, we can now give a very simple disproof of the third law:

Theorem 7.4.1. *There exist bouncing charged null dust spacetimes that violate the third law of black hole thermodynamics: a subextremal Reissner–Nordström apparent horizon can evolve into an extremal Reissner–Nordström event horizon in finite advanced time due to the incidence of charged null dust.*

Proof. Apply Propositions 7.2.3 and 7.2.4 to $(r_1, r_2, \varpi_1, \varpi_2, Q_1, Q_2) \in \mathcal{P}$ satisfying $r_2 < \varpi_2$, $Q_1 < \varpi_1$, and $Q_2 = \varpi_2$. For example, one may take $(0.85, 0.88, 0.56, 1, 0.5, 1) \in \mathcal{P}$. See Fig. 7.3. □

Remark 7.4.2. Since the energy-momentum tensor remains bounded in Ori's model and the weak energy condition is satisfied, this is indeed a counterexample to Israel's formulation of the third law [Isr86].

The counterexample in Theorem 7.4.1 explicitly displays the disconnectedness of the outermost apparent horizon which is also present in our charged scalar field counterexamples to the third law [KU22]. Note that the bouncing dust beam does not cross the subextremal apparent horizon, as is required by (7.1.9).

Remark 7.4.3. In the example depicted in Fig. 7.3, the parameters $\check{\varpi}$ and $\check{Q}$ satisfy $\check{\varpi}(r) < \check{Q}(r)$ for $r \in (r_2 - \varepsilon, r_2)$ and some $\varepsilon > 0$. Indeed, the ODE (7.2.6) implies

$$\check{Q}'(r) = \frac{r}{\check{Q}(r)}\check{\varpi}'(r) < \check{\varpi}'(r)$$

near r_2, where we have used $r_2 < Q_2$. The possibility (in fact, apparent inevitability) of the Vaidya parameters being superextremal right before extremality is reached seems to have been overlooked in the literature [SI80; FGS17].[1]

[1] The paper [FGS17] reexamines the third law in light of Ori's paper [Ori91], but always makes the assumption that the parameters satisfy $Q(V) < \varpi(V)$ right before extremality. Therefore, they seemingly reaffirm the third law!

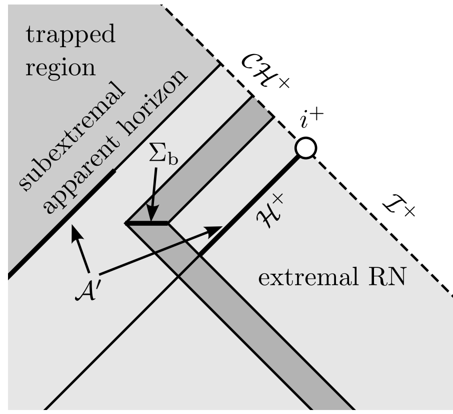

Figure 7.3: Penrose diagram of a counterexample to the third law of black hole thermodynamics in Ori's charged null dust model, Theorem 7.4.1. Note that the bounce Σ_b lies behind the extremal event horizon $\mathcal{H}^+$ since $r_2 < \varpi_2$ and ϖ_2 is the area-radius of the extremal horizon. The broken curve $\mathcal{A}'$ is the outermost apparent horizon of the spacetime. The disconnectedness of $\mathcal{A}'$ is necessary in third law violating spacetimes—refer back to the discussion in Section 1.1.3.3. A crucial feature of this counterexample is that Σ_b lies strictly between the (initially) subextremal apparent horizon and the (eventually) extremal event horizon. Compare with Fig. 1.8.

Remark 7.4.4. If one applies the old "standard interpretation" of the ingoing Vaidya metric from [SI80; LZ91] to the seed functions $\varpi_{\text{in}}(V)$ and $Q_{\text{in}}(V)$ constructed in the proof of Theorem 7.4.1, one sees that the beam will hit the subextremal apparent horizon with a negative energy density, which is consistent with [SI80].

Using methods from the proof of Theorem 1.2.1, the dust spacetimes in Theorem 7.4.1 can be "desingularized" to smooth Einstein–Maxwell–Vlasov solutions. The desingularized solutions can also be chosen to have the property that the matter remains strictly between the subextremal apparent horizon and the event horizon and we refer back to Section 1.2.7.

7.5 Issues with the bouncing charged null dust model

While Proposition 7.2.3 allows us to construct these interesting examples, the bouncing charged null dust model is unsatisfactory and we should seek to replace it for several reasons:

1. The model does not arise as a well-posed initial value problem for a system of PDEs. Pasting the ingoing and outgoing Vaidya solutions together is a deliberate surgery procedure that only works for seed functions ϖ_{in} and Q_{in} satisfying several nontrivial and nongeneric conditions.

2. The solutions are generally not smooth along Σ_b, nor along any cone where $Q = 0$ (recall Remark 7.1.2). The fluid density ρ is unbounded along Σ_b and the number current $N = \rho k$ is discontinuous across Σ_b.

3. Null dust is ill-posed once the dust reaches the center of symmetry [Mos17].

Nevertheless, we will show in Chapter 8 that the bouncing charged null dust model can be well-approximated (near Σ_{b}), in a precise manner, by smooth solutions of the Einstein–Maxwell–Vlasov system. See already Section 8.10.

7.6 The formal radial charged null dust system in double null gauge

In order to precisely phrase the manner in which Einstein–Maxwell–Vlasov approximates bouncing charged null dust, as well as to motivate the choice of Vlasov initial data, we now reformulate Ori's model in double null gauge. Following Moschidis [Mos17; Mos20], we reformulate the system by treating N and T as the fundamental variables. By eliminating the fluid variables k and ρ, we can view the ingoing and outgoing phases as two separate well-posed initial value problems, with data posed along the bounce hypersurface. This helpfully suppresses the issue of blowup of ρ on Σ_{b}.

Definition 7.6.1. The *spherically symmetric formal outgoing charged null dust* model for particles of fundamental charge $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$ consists of a smooth spherically symmetric charged spacetime $(\mathcal{Q}, r, \Omega^2, Q)$ and two nonnegative smooth functions N^v and T^{vv} on $\mathcal{Q}$.

The system satisfies the wave equations

$$\partial_u\partial_v r = -\frac{\Omega^2}{2r^2}\left(m - \frac{Q^2}{2r}\right), \tag{7.6.1}$$

$$\partial_u\partial_v \log\Omega^2 = \frac{\Omega^2 m}{r^3} - \frac{\Omega^2 Q^2}{r^4}, \tag{7.6.2}$$

the Raychaudhuri equations

$$\partial_u\left(\frac{\partial_u r}{\Omega^2}\right) = -\tfrac{1}{4} r\Omega^2 T^{vv}, \tag{7.6.3}$$

$$\partial_v\left(\frac{\partial_v r}{\Omega^2}\right) = 0, \tag{7.6.4}$$

and the Maxwell equations

$$\partial_u Q = -\tfrac{1}{2}\mathfrak{e} r^2\Omega^2 N^v, \tag{7.6.5}$$

$$\partial_v Q = 0. \tag{7.6.6}$$

The number current satisfies the conservation law

$$\partial_v(r^2\Omega^2 N^v) = 0 \tag{7.6.7}$$

and the energy-momentum tensor satisfies the Bianchi equation

$$\partial_v(r^2\Omega^4 T^{vv}) = +\mathfrak{e}\Omega^4 Q N^v. \tag{7.6.8}$$

In the outgoing model, we may think of N^u, T^{uu}, and T^{uv} to just be defined as identically zero. From (7.6.1) and (7.6.3)–(7.6.6) one easily derives

$$\partial_u m = -\tfrac{1}{2}r^2\Omega^2 T^{vv}\partial_v r + \frac{Q^2}{2r^2}\partial_u r, \quad \partial_v m = \frac{Q^2}{2r^2}\partial_v r, \tag{7.6.9}$$

$$\partial_u \varpi = -\tfrac{1}{2}r^2\Omega^2 T^{vv}\partial_v r - \tfrac{1}{2}\mathfrak{e}r\Omega^2 Q N^v, \quad \partial_v \varpi = 0. \tag{7.6.10}$$

Furthermore, if we set $k^v \doteq T^{vv}/N^v$, then

$$k^v\partial_v k^v + \partial_v \log\Omega^2 (k^v)^2 = +\mathfrak{e}\frac{Q}{r^2}k^v, \tag{7.6.11}$$

which is the spherically symmetric version of (7.1.3) for the vector field $k \doteq k^v\partial_v$. The energy density of the fluid is defined by $\rho \doteq (N^v)^2/T^{vv}$ whenever the denominator is nonvanishing.

Definition 7.6.2. The *spherically symmetric formal ingoing charged null dust* model for particles of fundamental charge $\mathfrak{e} \in \mathbb{R}\setminus\{0\}$ consists of a smooth spherically symmetric charged spacetime $(\mathcal{Q}, r, \Omega^2, Q)$ and two nonnegative smooth functions N^u and T^{uu} on $\mathcal{Q}$. The system satisfies the same equations as the ingoing system with $u \leftrightarrow v$ and the opposite sign in front of N^u.

In the ingoing case, $k^u \doteq T^{uu}/N^u$ and $\rho \doteq (N^u)^2/T^{uu}$.

Remark 7.6.3. By (7.6.11), these formal systems define solutions of the Einstein–Maxwell–charged null dust system (see Definition 7.1.1) whenever k and ρ are well-defined.

Remark 7.6.4. Inspection of (7.6.8) reveals that T^{vv} can reach zero in finite backwards time. If one were to continue the solution further, T^{vv} could become negative, which shows that the formal system actually reproduces the old "standard interpretation" of the charged Vaidya metric discussed in [Ori91]. As we will see, because the dominant energy condition holds in the Einstein–Maxwell–Vlasov model, only dust solutions with $T^{uu}, T^{vv} \geq 0$ will arise as limiting spacetimes, confirming Ori's heuristic picture discussed in [Ori91].

7.6.1 The Cauchy problem for outgoing formal charged null dust

Mirroring the treatment of time-symmetric[2] seed data for the Einstein–Maxwell–Vlasov system in Section 3.2.3, we make the following definition:

Definition 7.6.5. A *time-symmetric seed data set* $\mathcal{S}_\mathrm{d} \doteq (\mathring{\mathcal{N}}^v, \mathring{\mathcal{T}}^{vv}, r_2, \mathfrak{e})$ for the spherically symmetric formal outgoing charged null dust system consists of real numbers $r_2 \in \mathbb{R}_{>0}$ and $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$, together with nonnegative compactly supported smooth functions $\mathring{\mathcal{N}}^v$ and $\mathring{\mathcal{T}}^{vv}$ with support contained in $(0, r_2]$.

In the dust case, we define $\mathring{m}$ and $\mathring{Q}$ on $[0, r_2]$ with $\mathring{m}(0) = \mathring{Q}(0) = 0$ by solving

$$\frac{d}{dr}\mathring{m} = \frac{r^2}{4}\left(1 - \frac{2\mathring{m}}{r}\right)^{-2} \mathring{\mathcal{T}}^{vv} + \frac{\mathring{Q}^2}{2r^2}, \tag{7.6.12}$$

$$\frac{d}{dr}\mathring{Q} = \frac{1}{2}\mathfrak{e}r^2\left(1 - \frac{2\mathring{m}}{r}\right)^{-2} \mathring{\mathcal{N}}^v, \tag{7.6.13}$$

provided $2\mathring{m} < r$ on $[0, r_2]$. The remaining definitions from the Vlasov case, in particular Definition 3.2.10, can be carried over to dust with the obvious modification that $N^v = T^{vv} = 0$ along Γ.[3]

Proposition 7.6.6. *Let $\mathcal{S}_\mathrm{d}$ be an untrapped time-symmetric seed data set for outgoing dust. Then there exists a unique global smooth solution $(r, \Omega^2, Q, N^v, T^{vv})$ of the formal outgoing charged null dust system on $\mathcal{C}_{r_2}$ attaining the seed data.*

Remark 7.6.7. Let $r_1 \doteq \inf(\operatorname{spt} \mathring{\mathcal{N}}^v \cup \operatorname{spt} \mathring{\mathcal{T}}^{vv})$. Then (r, Ω^2, Q) is isometric to Minkowski space for $u \geq -r_1$.

Proof. This can be proved by applying a suitable coordinate transformation to a suitable outgoing charged Vaidya metric. However, it is instructive to give a direct proof using the evolution equations.

We pose initial data

$$\mathring{r}(r) = r, \quad \mathring{\Omega}^2(r) = \left(1 - \frac{2\mathring{m}}{r}\right)^{-1}, \quad \mathring{Q}(r) = \int_0^r \frac{1}{2}\mathfrak{e}r^2\left(1 - \frac{2\mathring{m}}{r}\right)^{-2} \mathring{\mathcal{N}}^v \, dr',$$

and for derivatives according to Definition 3.2.10, for the equations (7.6.1), (7.6.2), and (7.6.6). By a standard iteration argument, this determines the functions (r, Ω^2, Q) uniquely. The existence of a

[2] In the Vlasov case, time symmetry referred to both the geometry of the spacelike part of the initial data hypersurface and the matter configuration. Since purely outgoing dust is clearly not time symmetric, it refers here only to the geometry of the spacelike part of the initial data hypersurface.

[3] Since the dust here is purely outgoing, we do not have to be concerned about dust going into Γ.

global development is strictly easier than the corresponding proof in Proposition 8.2.3 once the rest of the system has been derived and is omitted. We now *define*

$$N^v \doteq -\frac{2}{\mathfrak{e}r^2\Omega^2}\partial_u Q, \quad T^{vv} \doteq -\frac{4}{r\Omega^2}\partial_u\left(\frac{\partial_u r}{\Omega^2}\right)$$

and aim to show that the rest of the equations in Definition 7.6.1 are satisfied.

To prove (7.6.7), simply rearrange the definition of N^v and use (7.6.6). Note that the definition of N^v is consistent with $\mathring{\mathcal{N}}^v = \mathring{\Omega}^{-2}\mathring{N}^v$ by (7.6.13).

Using (7.6.1), (7.6.2), and (7.6.6), a tedious calculation yields

$$\partial_u(r\partial_v^2 r - r\partial_v r\partial_v \log\Omega^2) = 0. \tag{7.6.14}$$

Arguing as in Proposition 3.2.12, we see that (7.6.4) holds on initial data and is therefore propagated by (7.6.14). This proves the evolution equation $\partial_v m = \partial_v r Q^2/(2r^2)$ and by using (7.6.1) once more, we see that

$$\partial_u m = -2r\partial_v r\partial_u\left(\frac{\partial_u r}{\Omega^2}\right) + \frac{Q^2}{2r^2}\partial_u r.$$

Comparing this with (7.6.12) and the definition of T^{vv} yields $\mathring{\mathcal{T}}^{vv} = \mathring{\Omega}^{-2}\mathring{T}^{vv}$, as desired. Finally, (7.6.8) is proved by directly differentiating the definition of T^{vv} and using (7.6.1), (7.6.2), and (7.6.5). □

7.6.2 Outgoing charged Vaidya as formal outgoing dust

We now want to represent the outgoing portion of a regular center bouncing charged null dust beam given by Proposition 7.2.3 in terms of the outgoing formal system. Let $\alpha \in \mathcal{P}_\Gamma$, $\varsigma(\alpha) = (\alpha, \check{\varpi}, \check{Q})$ be given by Proposition 7.2.4, and consider the time-symmetric dust seed data $\mathcal{S}_{\mathrm{d},\alpha} \doteq (\mathring{\mathcal{N}}^v_{\mathrm{d}}, 0, r_2, \mathfrak{e})$, where

$$\mathring{\mathcal{N}}^v_{\mathrm{d}} \doteq \frac{2}{\mathfrak{e}r^2}\left(1 - \frac{2\check{\varpi}}{r} + \frac{\check{Q}^2}{r^2}\right)^2 \check{Q}'. \tag{7.6.15}$$

For this choice of seed, the constraints (7.6.12)–(7.6.13) read

$$\frac{d}{dr}\mathring{m} = \frac{\mathring{Q}^2}{2r^2}, \tag{7.6.16}$$

$$\frac{d}{dr}\mathring{Q} = \left(1 - \frac{2\mathring{m}}{r}\right)^{-2}\left(1 - \frac{2\check{m}}{r}\right)^2 \check{Q}'. \tag{7.6.17}$$

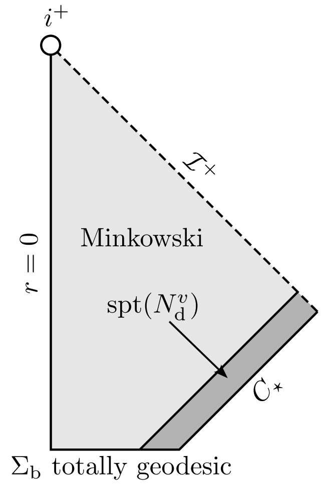

Figure 7.4: An outgoing charged null dust beam obtained by applying Proposition 7.6.6 to the seed $\mathcal{S}_{\mathrm{d},\alpha}$ for parameters $\alpha=(r_1,r_2,0,\varpi_2,0,Q_2)$. The electrovacuum boundary $C_\star$ can be attached to a Reissner–Nordström spacetime with parameters ϖ_2 and Q_2.

Therefore, by (7.2.6), $\mathring{m}=\check{m}$ and $\mathring{Q}=\check{Q}$, where $\check{m}\doteq\check{\varpi}-\check{Q}^2/(2r)$. It follows that the outgoing formal dust solution $(r_{\mathrm{d}},\Omega^2_{\mathrm{d}},Q_{\mathrm{d}},N^v_{\mathrm{d}},T^{vv}_{\mathrm{d}})$ provided by Proposition 7.6.6 on $\mathcal{C}_{r_2}$ is indeed the same as the outgoing charged Vaidya metric provided by the radial parametrization method, Proposition 7.2.3. See Fig. 7.4.

Chapter 8

Extremal black hole formation as a critical phenomenon

In this section, we prove Theorem 1.2.1 by constructing *bouncing charged Vlasov beams* as in Fig. 1.5 and Fig. 1.6 with prescribed final parameters. This is achieved by a very specific choice of time-symmetric Vlasov seed data and global estimates for the resulting developments. We give a detailed outline of the proof in Section 8.1 and the proof itself occupies Sections 8.2 to 8.9. In Section 8.10, we show as a consequence of the estimates in the previous sections that these bouncing charged Vlasov beams weak* converge to the bouncing charged null dust spacetimes of Proposition 7.2.3 in a hydrodynamic limit of the beam parameters. Finally, in Section 8.11 we disprove the third law in Einstein–Maxwell–Vlasov and construct examples of "event horizon jumping."

8.1 A guide to the proof of Theorem 1.2.1

8.1.1 The heuristic picture

The essential idea in the proof of Theorem 1.2.1 is to "approximate" the bouncing radial charged null dust solutions from Theorem 7.3.2 and Fig. 7.2 by smooth families of smooth Einstein–Maxwell–Vlasov solutions. Indeed, at least formally, dust can be viewed as Vlasov matter $f(x,p)$ concentrated on a single momentum $p = k(x)$ at each spacetime point x. One is faced with having to perform a *global-in-time* desingularization of families of dust solutions which are singular in both the space and momentum variables.

Assuming that this can be done, the heuristic picture is that of a focusing beam of Vlasov matter coming in from infinity with particles of mass $\mathfrak{m} = 0$ or $\mathfrak{m} \ll 1$ (so that the particles look almost massless for very large time scales) and very small angular momentum $0 < \ell \ll 1$, which are decelerated by the electromagnetic field that they generate. Then, along some "approximate bounce hypersurface," the congruence smoothly "turns around" and becomes outgoing, escaping to infinity if a black hole has not yet formed. Along the way, the particles do not hit the center of symmetry. By appropriately varying the beam parameters, we can construct families of spacetimes as depicted in Fig. 1.5 or Fig. 1.6.

As should be apparent from the treatment of the Cauchy problem for the Einstein–Maxwell–Vlasov system in Section 3.2.3 and for charged null dust in Section 7.6, we want to pose Cauchy data on (what will be) the approximate bounce hypersurface for the desingularized Vlasov solutions. We will choose the initial data for f to be supported on small angular momenta $\ell \sim \varepsilon$ and so that the charge $\mathring{Q}$ and Hawking mass $\mathring{m}$ profiles closely approximate the initial data for dust as described in Section 7.6.2. The Vlasov beam which is intended to approximate charged null dust is called the *main beam*.

As we will see, desingularizing bouncing charged null dust requires an ansatz for $\mathring{f}$ which necessarily degenerates in ε. Closing estimates in the region of spacetime where $Q \lesssim \varepsilon$ is then a fundamental issue because the repulsive effect of the electromagnetic field is relatively weak there. We overcome this issue by adding an *auxiliary beam* to the construction, which stabilizes the main beam by adding a small amount of charge on the order of $\eta \gg \varepsilon$. This beam is not dust-like, consists of particles with angular momentum ~ 1, and is repelled away from the center by the centrifugal force.

The goal will be to construct a smooth family of Vlasov seeds $\lambda \mapsto \mathcal{S}_\lambda$ for $\lambda \in [-1, 2]$ such that $\mathcal{S}_{-1}$ is trivial (i.e., evolves into Minkowski), $\mathcal{S}_2$ forms a subextremal Reissner–Nordström black hole with charge to mass ratio $\approx \frac{1}{2}$, and $\lambda_* \approx 1$ is the critical parameter for which an extremal Reissner–Nordström black hole with mass M forms. For $\lambda \in [0, 2]$, the Vlasov development $\mathcal{D}_\lambda$ closely approximates the dust developments from Theorem 7.3.2 (in a sense to be made precise in Section 8.10 below) and $\lambda \in [-1, 0]$ smoothly "turns on" the auxiliary beam. At the very end of the proof, λ is simply rescaled to have range $[0, 1]$.

In fact, our methods allow us to desingularize any bouncing charged null dust beam given by Proposition 7.2.4. Adding dependence on λ is then essentially only a notational hurdle. We now highlight specific aspects of the construction in more detail.

8.1.2 Time symmetry and reduction to the outgoing case

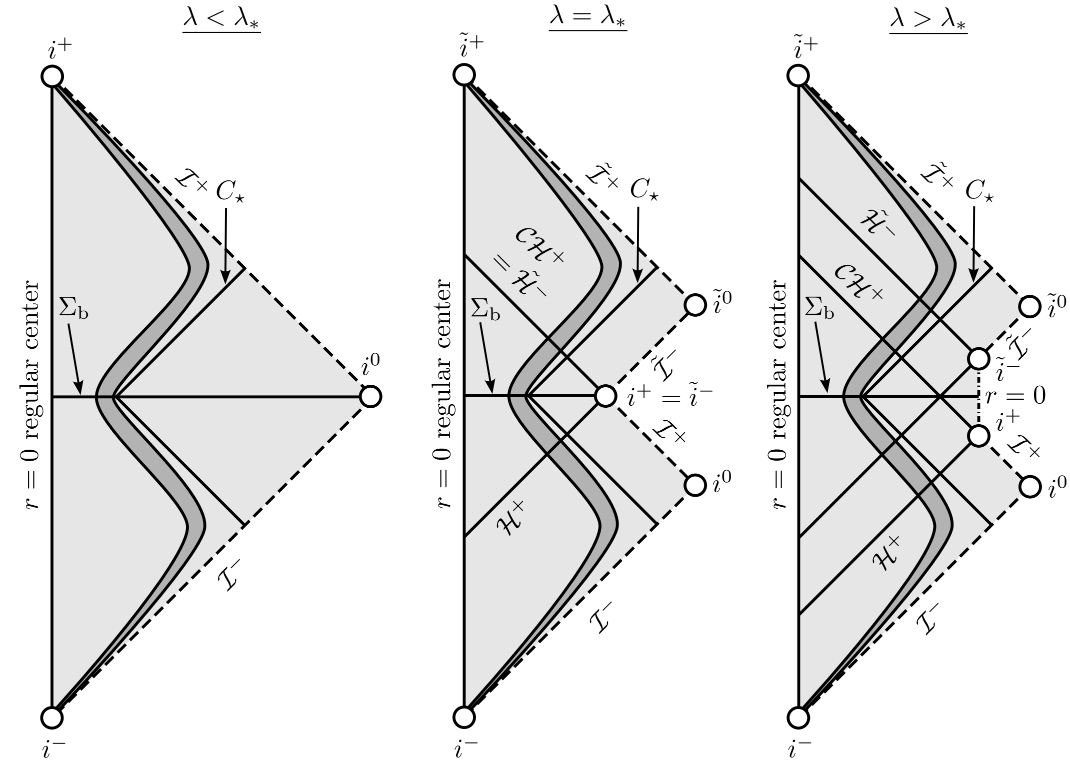

Figure 8.1: Penrose diagrams of the "maximal time-symmetric doubled spacetimes" used in the proof of Theorem 1.2.1 when $\mathfrak{m} > 0$. When $\lambda \geq \lambda_*$, these spacetimes are evidently not globally hyperbolic, but one can easily observe that the globally hyperbolic spacetimes depicted in Fig. 1.5 when $\lambda \geq \lambda_*$ are simply the above spacetimes restricted to the past of $\mathcal{CH}^+ \cup \{i^+\} \cup \mathcal{I}^+$. The exterior region is isometric to a subset of the maximally extended Reissner–Nordström solution with parameters depending on λ.

The starting point of the construction of bouncing charged Vlasov beams is the prescription of Cauchy data on an approximate bounce hypersurface Σ_b, using the radial parametrization of bouncing charged null dust as a guide. We can now see the utility of the time-symmetric ansatz in Section 3.2.3: it reduces the problem to constructing an *outgoing* beam, which is then reflected and glued to maximally extended Reissner–Nordström to construct a *time symmetric* spacetime as depicted in Fig. 8.1 below. These "maximal time-symmetric spacetimes" are constructed in Section 8.9.1. The globally hyperbolic developments in Theorem 1.2.1 are obtained by taking appropriate subsets and identifying suitable Cauchy hypersurfaces.

The problem now reduces to constructing the region bounded to the past by $C_\star$, Σ_b, and the center in Fig. 8.1. In this region, *the solution is always dispersive.* Therefore, we can actually treat the subextremal, extremal, and superextremal cases at once. Detection of whether a black hole forms in the doubled spacetime takes place *on the level of initial data* and we heavily exploit the global structure of the Reissner–Nordström family itself in this process. Note that while we prescribe data in the black hole interior when $\lambda \le 0$, there clearly exist Cauchy surfaces lying entirely in the domain of outer communication. In fact, the solutions are always past complete and disperse to the past. See already Proposition 8.9.9.

8.1.3 The choice of seed data

We now describe our desingularization procedure for bouncing charged null dust on the level of initial data. Consider the outgoing portion of a charged null dust beam $(r_\mathrm{d}, \Omega^2_\mathrm{d}, Q_\mathrm{d}, N^v_\mathrm{d}, T^{vv}_\mathrm{d})$ as in Section 7.6.2, with Cauchy data posed along the bounce hypersurface Σ_b. The geometry of the outgoing dust beam is entirely driven by the choice of renormalized number current $\mathring{\mathcal{N}}^v_\mathrm{d}$ in (7.6.15). Importantly, the energy-momentum tensor of dust vanishes identically along Σ_b.

Since radial charged null dust has $\ell = 0$, we wish to approximate dust with Vlasov matter consisting of particles with angular momentum $\ell \sim \varepsilon$, where $0 < \varepsilon \ll 1$ is a small parameter to be chosen. We want to choose the initial distribution function $\mathring{f}$ so that

$$\mathring{\mathcal{N}}^u + \mathring{\mathcal{N}}^v = \mathring{\mathcal{N}}^v_\mathrm{d} = \frac{1}{\mathfrak{e} r^2}\left(1 - \frac{2\check{\varpi}}{r} + \frac{\check{Q}^2}{r^2}\right)^2 \frac{d\check{Q}}{dr}, \tag{8.1.1}$$

$$\mathring{Q} \approx \check{Q}, \quad \mathring{\varpi} \approx \check{\varpi}, \quad \mathring{\mathcal{T}}^{uu}, \mathring{\mathcal{T}}^{uv}, \mathring{\mathcal{T}}^{vv} \approx 0 \tag{8.1.2}$$

on Σ_b, as $\varepsilon \to 0$. These conditions are satisfied if we choose

$$\mathring{f}^{\alpha,\varepsilon}_\mathrm{main}(r, p^u, p^v) \doteq \frac{c}{\mathfrak{e} r^2 \varepsilon}\left(1 - \frac{2\check{\varpi}}{r} + \frac{\check{Q}^2}{r^2}\right)^2 \frac{d\check{Q}}{dr}\delta_\varepsilon(p^u)\delta_\varepsilon(p^v) \tag{8.1.3}$$

for $r \in [r_1, r_2]$, where δ_ε are approximations of the identity with support $[\varepsilon, 2\varepsilon]$ and c is a normalization constant that depends on the precise choice of the family δ_ε. In order for the mass shell inequality $\Omega^2 p^u p^v \ge \mathfrak{m}^2$ to hold on the support of $\mathring{f}^{\alpha,\varepsilon}_\mathrm{main}$, (8.1.3) forces us to choose $\mathfrak{m} \in [0, \mathfrak{m}_0]$ with $0 < \mathfrak{m}_0 \ll \varepsilon$.

Remark 8.1.1. In the full bouncing null dust model, N is discontinuous across Σ_b. Indeed, to the past of Σ_b, N points in the u-direction and has a nonzero limit along Σ_b, but to the future points in the v-direction and also has a nonzero limit. In the Vlasov case, time symmetry demands N be smooth across, and orthogonal to, Σ_b. By comparing (3.2.30) with (7.6.13), we see that $\mathring{\mathcal{N}}^u + \mathring{\mathcal{N}}^v$ in Vlasov takes the role of $\mathring{\mathcal{N}}^v$ in dust.

Observe directly from (8.1.3) that $\mathring{f}^{\alpha,\varepsilon}_\mathrm{main}$ behaves pointwise like ε^{-3} and therefore pointwise

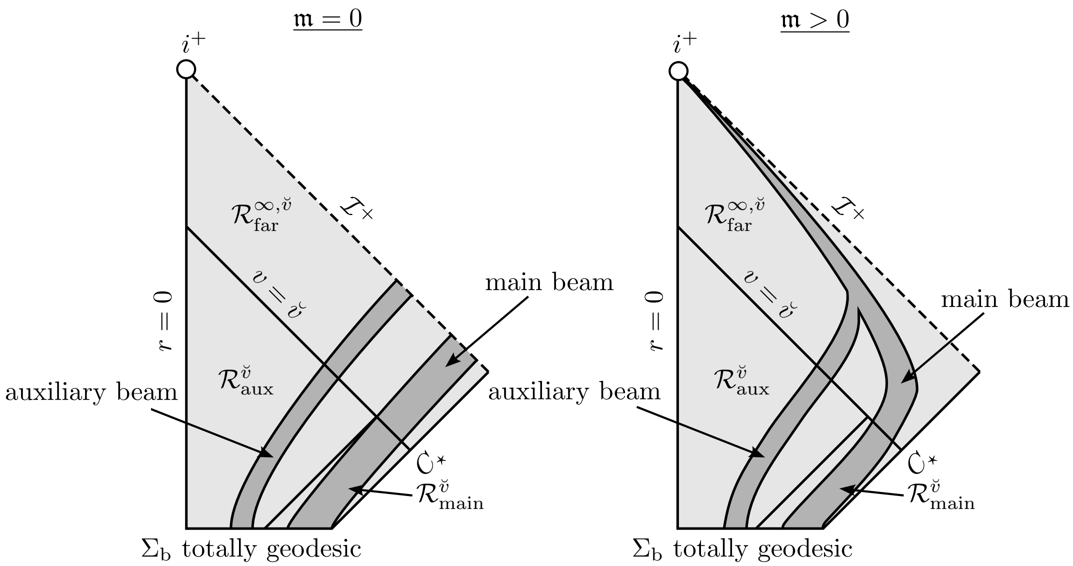

Figure 8.2: Penrose diagram of outgoing charged Vlasov beams (evolution of the seed data $\mathcal{S}_{\alpha,\eta,\varepsilon}$). Note that the beams do not intersect when $\mathfrak{m} = 0$. When $\mathfrak{m} > 0$, one can show that they do, but it is not necessary to do so for our purposes here.

estimates for N and T in evolution must utilize precise estimates of the electromagnetic flow to cancel factors of ε. Closing estimates independently of ε is the main challenge of this scheme and we directly exploit the null structure of the spherically symmetric Einstein–Maxwell–Vlasov system in the proof. The main mechanisms ensuring boundedness of N and T in the main beam are:

1. The angular momentum ℓ is conserved, so that $\ell \sim \varepsilon$ throughout the main beam.
2. If γ is an electromagnetic geodesic arising from the support of $\mathring{f}^{\alpha,\varepsilon}_{\text{main}}$, then p^v should rapidly increase due to electromagnetic repulsion. Dually, p^u should rapidly decrease, which ought to suppress the ingoing moments N^u, T^{uu}, and T^{uv}. This should be compared with the vanishing of N^u, T^{uu}, and T^{uv} in outgoing null dust. We say that the main beam *bounces due to electromagnetic repulsion.*

As is apparent from (2.1.22), the magnitude of the repulsive effect is proportional to Q. If we were to evolve the seed $\mathring{f}^{\alpha,\varepsilon}_{\text{main}}$ on its own, the inner edge of the beam would experience less electromagnetic repulsion since Q is potentially quite small in the inner region.

In order to reinforce the repulsive effect of the electric field in the main beam and get a consistent hierarchy of powers of ε, we introduce an *auxiliary beam* on the inside of the main beam which *bounces due to the centrifugal force* associated to electromagnetic geodesics with large angular momentum. The initial data for the auxiliary beam is chosen to be

$$\mathring{f}^{r_1,\eta}_{\text{aux}}(r, p^u, p^v) \doteq \eta\, \varphi(r, p^u, p^v), \tag{8.1.4}$$

where $\eta \gg \varepsilon$ is a constant determining the amplitude, φ is a cutoff function supported on the set $[\frac{1}{3}r_1, \frac{2}{3}r_1] \times [\Lambda - 1, \Lambda + 1] \times [\Lambda - 1, \Lambda + 1]$, and Λ is a fixed large constant that determines the strength of the centrifugal force felt by the auxiliary beam. The auxiliary beam ensures that the main beam always interacts with an electric field of amplitude $\gtrsim \eta$, which acts as a crucial stabilizing mechanism.

8.1.4 The near and far regions and the hierarchy of scales

The total seed for an outgoing Vlasov beam is taken to be $\mathcal{S}_{\alpha,\eta,\varepsilon} \doteq (\mathring{f}^{\alpha,\eta,\varepsilon}_{\mathrm{tot}}, r_2, \mathfrak{m}, \mathfrak{e})$, where

$$\mathring{f}^{\alpha,\eta,\varepsilon}_{\mathrm{tot}} \doteq \mathring{f}^{r_1,\eta}_{\mathrm{aux}} + \mathring{f}^{\alpha,\varepsilon}_{\mathrm{main}}, \tag{8.1.5}$$

the fundamental charge $\mathfrak{e} > 0$ is fixed, the mass $\mathfrak{m}$ lies in the interval $[0, \mathfrak{m}_0]$, and η, ε, and $\mathfrak{m}_0$ need to be chosen appropriately small.

To study the evolution of $\mathcal{S}_{\alpha,\eta,\varepsilon}$, depicted in Fig. 8.2, we distinguish between the *near region* $\{v \leq \check{v}\}$ and the *far region* $\mathcal{R}^{\check{v},\infty}_{\mathrm{far}} = \{v \geq \check{v}\}$, where $\check{v}$ is a large advanced time to be determined. Roughly, the ingoing cone $\{v = \check{v}\}$ is chosen so that the geometry is very close to Minkowskian and the Vlasov field is "strongly outgoing" and supported far away from the center, i.e.,

$$\frac{p^u}{p^v} \lesssim r^{-2} \ll 1 \tag{8.1.6}$$

for every p^u and p^v such that $f(\,\cdot\,, \check{v}, p^u, p^v) \neq 0$. The near region is further divided into the *main* and *auxiliary* regions, corresponding to the physical space support of the main and auxiliary beams and denoted by $\mathcal{R}^{\check{v}}_{\mathrm{main}}$ and $\mathcal{R}^{\check{v}}_{\mathrm{aux}}$, respectively.[1]

The beam parameters η, $\varepsilon, \mathfrak{m}_0$ and the auxiliary parameter $\check{v}$ satisfy the hierarchy

$$0 < \mathfrak{m}_0 \ll \varepsilon \ll \eta \ll \check{v}^{-1} \ll 1. \tag{8.1.7}$$

To prove the sharp rate of dispersion when $\mathfrak{m} > 0$, we augment this hierarchy with

$$0 < v_{\#}^{-1} \ll \mathfrak{m},$$

where $v_{\#}$ is a very large time after which the additional dispersion associated to massive particles

[1] For reasons of convenience, $\mathcal{R}_{\mathrm{main}}$ is defined slightly differently in the actual proof than the region depicted in Fig. 8.2, but this is inconsequential at this level of discussion.

kicks in.

The proof of Theorem 1.2.1 proceeds by showing that if (8.1.7) holds, then the solution exists, with certain properties, in the regions $\mathcal{R}^{\check{v}}_{\mathrm{main}}$, $\mathcal{R}^{\check{v}}_{\mathrm{aux}}$, and $\mathcal{R}^{\check{v},\infty}_{\mathrm{far}}$, in that order. The sharp decay rates of N and T are then shown a posteriori by re-analyzing the electromagnetic geodesic flow in the far region.

Remark 8.1.2. The final Reissner–Nordström parameters of the total Vlasov seed (8.1.5) depend on the approximation parameters η and ε, but are $O(\eta)$-close to (ϖ_2, Q_2). Therefore, in order to reach any fixed set of parameters, the background dust seed has to be appropriately modulated. See already Section 8.9.5.

8.1.5 Outline of the main estimates

The main beam in the near region: In this region, the main goal is proving smallness (in ε) of N^u, T^{uu}, and T^{uv}, which are identically zero in the background dust solution. We define the *phase space volume function* $\mathscr{V} : \mathcal{Q} \to \mathbb{R}_{\geq 0}$ by

$$\mathscr{V}(u,v) \doteq \Omega^2(u,v)|\{(p^u,p^v) : f(u,v,p^u,p^v) \neq 0\}|, \tag{8.1.8}$$

where $|\cdot|$ is the Lebesgue measure on $\mathbb{R}^2_{p^u,p^v}$. The function $\mathscr{V}$ is invariant under gauge transformations of u and v. Using the mass shell relation (2.3.12) and the change of variables formula, we find

$$\mathscr{V}(u,v) = \frac{2}{r^2}\int_0^\infty \int_{\{p^v : f(u,v,p^u,p^v)\neq 0\}} \frac{dp^v}{p^v}\,\ell\, d\ell, \tag{8.1.9}$$

where we view p^u as a function of p^v and ℓ. Because of the addition of $\mathring{f}^{r_1;\eta}_{\mathrm{aux}}$ to the seed data and the good monotonicity properties of (2.3.24) and (2.3.25), it holds that $Q \gtrsim \eta$ in $\mathcal{R}^{\check{v}}_{\mathrm{main}}$. Under relatively mild bootstrap assumptions, any electromagnetic geodesic γ in the main beam is accelerated outwards at a rate $\gtrsim \eta$, i.e.,

$$p^v \gtrsim \varepsilon + \eta \min\{\tau, 1\}, \quad p^u \lesssim \frac{\varepsilon^2}{r^2(\varepsilon + \eta\min\{\tau,1\})},$$

where $\tau \doteq \frac{1}{2}(u+v)$ is a "coordinate time." We also show that if γ_1 and γ_2 are two electromagnetic geodesics in the main beam which reach the same point $(u,v)\in\mathcal{R}^{\check{v}}_{\mathrm{main}}$, then

$$|p^v_1-p^v_2|\lesssim\frac{\varepsilon}{\eta^2}$$

at (u,v). Using these estimates, conservation of angular momentum, and the hierarchy (8.1.7), we show that

$$\mathscr{V}(u,v)\lesssim_\eta\frac{\varepsilon^3}{\varepsilon+\eta\min\{\tau,1\}},$$

where the notation $A\lesssim_\eta B$ means $A\le CB$, where C is a constant depending on η. Then, simply using the transport nature of the Vlasov equation, we obtain the estimates

$$T^{uu}(u,v)\lesssim\varepsilon^{1/2},\quad T^{uv}(u,v)\lesssim\varepsilon^{1/2},\quad \int_{-u}^{v}N^u(u,v')\,dv'\lesssim\varepsilon^{1/2},$$

which capture the fundamental characteristic of outgoing null dust. These estimates allow us to control the geometry at C^1 order, which is more than enough to use the generalized extension principle, Proposition 3.2.4, to extend the solution. For details, see Section 8.4. When $\lambda\in[-1,0]$ and the main beam has not yet been turned on, constructing the solution in this region is trivial since the solution is electrovacuum.

The auxiliary beam in the near region: Since the auxiliary beam is genuinely weak ($\mathring{f}^{r_1;\eta}_{\mathrm{aux}}\lesssim\eta$ pointwise), the bootstrap argument in $\mathcal{R}^{\check{v}}_{\mathrm{aux}}$ is a standard Cauchy stability argument, perturbing off of Minkowski space. We use explicit knowledge of the impact parameter and asymptotics of *null geodesics* with angular momentum $\sim\Lambda$ on Minkowski space and treat the charge as an error term in this region. For details, see Section 8.5.

Existence in the far region: The argument in this region is a refinement of Dafermos' proof of the stability of Minkowski space for the spherically symmetric Einstein–massless Vlasov system [Daf06] (see also [Tay15, Chapter 4]). Because of the singular nature of f_{main} in powers of ε, it seems difficult to obtain uniform in ε pointwise decay estimates for T^{uv} by the usual method of estimating decay of the phase space volume of the support of f at late times. Fortunately, we are able to exploit the a priori energy estimates

$$\int r^2\Omega^2T^{uv}\partial_ur\,du'\lesssim 1,\quad \int r^2\Omega^2T^{uv}\partial_vr\,dv'\lesssim 1 \tag{8.1.10}$$

coming from the monotonicity of the Hawking mass when $\partial_vr>0$ and $\partial_ur<0$ (see [Daf05b]). It is important to note that these energy estimates are independent of initial data and are a fundamental feature of the spherically symmetric Einstein equations. Under the bootstrap assumption that the electromagnetic geodesics making up the support of f are "outgoing" for $v \geq \check{v}$, the energy estimates (8.1.10) imply decay for the unweighted fluxes of T^{uv}. This shows that the geometry remains close to Minkowski in C^1 and recovering the bootstrap assumption on the support of f follows from good monotonicity properties of the electromagnetic geodesic flow when close to Minkowski. We also note that this approach using energy estimates allows us to treat the cases $\mathfrak{m} = 0$ and $\mathfrak{m} > 0$ simultaneously. For details, see Section 8.6.

Dispersion in the far region: Once the solution has been shown to exist globally, we prove sharp (in coordinate time τ) pointwise decay statements for N and T (see [RR92; Nou05; Tay15]). As the decay rates differ when $\mathfrak{m} = 0$ or $\mathfrak{m} > 0$, these two cases are treated separately.

The massless case. It follows immediately from the mass shell relation (2.3.12) that $p^u \lesssim r^{-2}$ in the far region. Since this is integrable, the beams are confined to null slabs and can even be shown to be disjoint as depicted in Fig. 8.2. Since each p^u contributes a factor of r^{-2} and our solutions have bounded angular momentum, we obtain the sharp dispersive hierarchy

$$N^v + T^{vv} \leq C(1+\tau)^{-2}, \quad N^u + T^{uv} \leq C(1+\tau)^{-4}, \quad T^{uu} \leq C(1+\tau)^{-6},$$

where the constant C depends on α, η, and ε. For details, see Section 8.7.

The massive case. When $\mathfrak{m} > 0$, p^u does not decay asymptotically. After a very late time $v_\# \gg \mathfrak{m}^{-1}$, $p^u \sim_\eta \mathfrak{m}^2$, which drives additional decay of the phase space volume. We prove this by a change of variables argument, turning volume in p^u at later times $v \geq v_\#$ into physical space volume of the support of f at time $v = v_\#$. This leads to the sharp isotropic decay rate

$$\mathcal{M} \leq C(1+\tau)^{-3}$$

for any moment $\mathcal{M}$ of f, where C depends on $\alpha, \eta, \varepsilon$, and a lower bound for $\mathfrak{m}$. For details, see Section 8.8.

8.2 Outgoing charged Vlasov beams

8.2.1 The beam parameters, fixed constants, and conventions

First, we fix once and for all the fundamental charge $\mathfrak{e} \in \mathbb{R} \setminus \{0\}$. Without loss of generality, we may take $\mathfrak{e} > 0$, as all of the arguments and definitions in the remainder of the present chapter require only minor cosmetic modifications to handle the case $\mathfrak{e} < 0$. Next, we fix an even function $\varphi \in C_c^\infty(\mathbb{R})$ satisfying $\operatorname{spt}\varphi = [-1,1]$, $\varphi \geq 0$, and

$$\int_{-1}^{1} \varphi\, dx = 1.$$

Let $\theta \in C^\infty(\mathbb{R})$ be a nondecreasing function such that $\theta(\lambda) = 0$ for $\lambda \leq -1$ and $\theta(\lambda) = 1$ for $\lambda \geq 0$. Let $\zeta \in C^\infty(\mathbb{R})$ be a nondecreasing function such that $\zeta(\lambda) = 0$ for $\lambda \leq 0$, $\zeta(\lambda) = \lambda$ for $\lambda \geq \frac{1}{2}$, and $\zeta'(\lambda) > 0$ for $\lambda \in (0, \frac{1}{2}]$. Finally, we fix a large[2] number $\Lambda \geq 1$, such that

$$\min_{p_1,p_2\in[\Lambda-1,\Lambda+1]} \frac{p_1 p_2}{(p_1+p_2)^2} \geq \tfrac{81}{400}. \tag{8.2.1}$$

We emphasize that:

The quintuple $(\mathfrak{e}, \varphi, \theta, \zeta, \Lambda)$ is fixed for the remainder of the chapter.

Recall the set $\mathcal{P}_\Gamma$ of regular center admissible parameters of the form $\alpha = (r_1, r_2, 0, \varpi_2, 0, Q_2)$ which was defined in Section 7.2. Let η, ε, and $\mathfrak{m}_0$ be positive real numbers. In the course of the proofs below, the particle mass $\mathfrak{m}$ will be restricted to satisfy $0 \leq \mathfrak{m} \leq \mathfrak{m}_0$.

Definition 8.2.1. Let $\alpha = (r_1, r_2, 0, \varpi_2, 0, Q_2) \in \mathcal{P}_\Gamma$, $\eta > 0$, and $\varepsilon > 0$. The *time-symmetric outgoing charged Vlasov beam seed* $\mathcal{S}_{\alpha,\eta,\varepsilon}$ is given by $(\mathring{f}^{r_1,\eta}_{\text{aux}} + \mathring{f}^{\alpha,\varepsilon}_{\text{main}}, r_2, \mathfrak{m}, \mathfrak{e})$, where

$$\mathring{f}^{r_1,\eta}_{\text{aux}}(r, p^u, p^v) \doteq \eta\, \varphi\left(\frac{6}{r_1} r - 3\right) \varphi(p^u - \Lambda)\varphi(p^v - \Lambda), \tag{8.2.2}$$

and

$$\mathring{f}^{\alpha,\varepsilon}_{\text{main}}(r, p^u, p^v) \doteq \frac{8}{3\pi\mathfrak{e}\varepsilon^3 r^2}\left(1 - \frac{2\check{\varpi}(r)}{r} + \frac{\check{Q}^2(r)}{r^2}\right)^2 \check{Q}'(r)\, \varphi\left(\frac{2p^u}{\varepsilon} - 3\right) \varphi\left(\frac{2p^v}{\varepsilon} - 3\right), \tag{8.2.3}$$

where $\check{\varpi}$ and $\check{Q}$ are taken from $\varsigma(\alpha)$, where the map ς was defined in Proposition 7.2.4 (cf. Remark 7.2.5).

Definition 8.2.2. Let $M > 0$, let $0 < r_1 < r_2$, and let $\{\alpha_{\lambda,M'}\}$ (with $|M - M'| \leq \delta$) be as in (7.3.1) in the proof of Theorem 7.3.2. For $\lambda \in [-1, 2]$, $\eta > 0$, $\varepsilon > 0$, $\mathfrak{m}_0$, we define

$$\mathcal{S}_{\lambda,M',\eta,\varepsilon} \doteq \mathcal{S}_{\alpha_{\zeta(\lambda),M'},\theta(\lambda)\eta,\varepsilon} \tag{8.2.4}$$

[2]Large relative to the other beam parameters. For instance, $\Lambda = 20$ suffices.

for particles of mass $0 \leq \mathfrak{m} \leq \mathfrak{m}_0$. For $\lambda \leq 0$, the $\check{\varpi}$ and $\check{Q}$ components of $\varsigma(\alpha_{\zeta(\lambda),M'})$ are interpreted as identically zero, in correspondence with the proof of Theorem 7.3.2.

Throughout the present chapter, the notation $A \lesssim B$ means that there exists a constant $C > 0$, which only depends on $\mathfrak{e}$, φ, θ, ζ, Λ, M, r_1, and r_2 such that $A \leq CB$. The notation $A \gtrsim B$ is defined similarly and $A \sim B$ means $A \lesssim B$ and $A \gtrsim B$. Moreover, we use the convention that all small (large) constants in "sufficiently small (large)" may depend on $\mathfrak{e}$, φ, θ, ζ, Λ, M, r_1, and r_2. In Section 8.7, we will also use the notation $A \lesssim_\eta B$ (resp., $A \lesssim_{\eta,\varepsilon} B$), in which we allow the constants to also depend on η (resp., η and ε). The relations $A \sim_\eta B$ and $A \sim_{\eta,\varepsilon} B$ are defined in the obvious way.

For the evolution problem, we will introduce a large parameter $\check{v}$ to separate $\mathcal{C}_{r_2}$ into the "near" and "far" regions. We will always assume that the parameter hierarchy

$$0 < \mathfrak{m}_0 \ll \varepsilon \ll \eta \ll \check{v}^{-1} \ll 1 \tag{8.2.5}$$

holds, by which we mean that any given statement holds for $\check{v}$ sufficiently large, η sufficiently small depending on $\check{v}$, ε sufficiently small depending on $\check{v}$ and η, and $\mathfrak{m}_0$ sufficiently small depending on $\check{v}$, η, and ε. To prove dispersion in the massive case, we introduce an even larger parameter $v_\#$ satisfying

$$0 < v_\#^{-1} \ll \mathfrak{m} \leq \mathfrak{m}_0,$$

so that $v_\#$ is chosen sufficiently large depending on $\check{v}, \eta, \varepsilon$, and $\mathfrak{m}$.

8.2.2 The global structure of outgoing charged Vlasov beams

Proposition 8.2.3. *Fix a fundamental charge $\mathfrak{e} > 0$, cutoff functions φ, θ, and ζ as in Section 8.2.1, a number $\Lambda \geq 1$ satisfying (8.2.1), $M > 0$, and $0 < r_1 < r_2 < r_-(4M, 2M)$. Let $\delta > 0$ be as in the statement of Theorem 7.3.2 and define $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$ as in Definition 8.2.2 for $\lambda \in [-1, 2]$, $|M' - M| \leq \delta$, $\eta > 0$, $\varepsilon > 0$, and for particles of mass $0 \leq \mathfrak{m} \leq \mathfrak{m}_0$, where $\mathfrak{m}_0 > 0$.*

If η is sufficiently small, ε is sufficiently small depending on η, and $\mathfrak{m}_0$ is sufficiently small depending on η and ε, then for any $\lambda \in [-1, 2]$ and $|M' - M| \leq \delta$, the seed $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$ is untrapped and consistent with particles of mass $\mathfrak{m}$. There exists a unique maximal normalized development $(\mathcal{U}, r, \Omega^2, Q, f)$ of $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$ for particles of charge $\mathfrak{e}$ and mass $\mathfrak{m}$ with the following properties.[3] *If*

[3]Here, uniqueness is in the class of normalized developments as in Definition 3.2.10. We have not shown an unconditional existence and uniqueness statement for maximal developments for the Einstein–Maxwell–Vlasov model in this dissertation (although this can be done) and will therefore infer uniqueness directly in the course of the construction.

$\mathfrak{m} > 0$:

1. *The development is global in the normalized double null gauge, i.e.,* $\mathcal{U} = \mathcal{C}_{r_2}$.

2. *The* $(3+1)$*-dimensional spacetime obtained by lifting* $(\mathcal{U}, r, \Omega^2)$ *is future causally geodesically complete and satisfies globally the estimates*

$$\partial_v r \sim 1, \quad |\partial_u r| \lesssim 1, \quad \Omega^2 \sim 1, \quad (1+u^2)|\partial_u \Omega^2| + (1+v^2)|\partial_v \Omega^2| \lesssim 1. \tag{8.2.6}$$

3. *Define the* final Reissner–Nordström parameters $\tilde{M}$ *and* $\tilde{e}$ *of* $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$ *to be the constant values of* ϖ *and* Q*, respectively, on the cone* C_{-r_2}*. Then* $\tilde{M}$ *and* $\tilde{e}$ *are smooth functions of* $(\lambda, M', \eta, \varepsilon, \mathfrak{m})$*, satisfy the estimate*

$$|\tilde{M} - \zeta(\lambda)^2 M'| + |\tilde{e} - \zeta(\lambda) M'| \lesssim \eta, \tag{8.2.7}$$

and extend smoothly to $\eta = \varepsilon = 0$*, where they equal* $\zeta(\lambda)^2 M'$ *and* $\zeta(\lambda) M'$*, respectively. The spacetime* $(\mathcal{U}, r, \Omega^2)$ *contains antitrapped surfaces (symmetry spheres where* $\partial_u r \geq 0$*) if and only if* $\tilde{e} \leq \tilde{M}$ *and* $r_2 < r_-$*, where* $r_\pm \doteq \tilde{M} \pm \sqrt{\tilde{M}^2 - \tilde{e}^2}$*. In this case, we nevertheless have* $\partial_u r \sim -1$ *for* v *sufficiently large and the antitrapped surfaces are restricted to lie in the slab* $\{2r_- - r_2 \leq v \leq 2r_+ - r_2\}$.

4. *The Vlasov distribution function* f *is quantitatively supported away from the center,*

$$\inf_{\pi(\operatorname{spt} f)} r \geq \tfrac{1}{6} r_1, \tag{8.2.8}$$

and the beam asymptotes to future timelike infinity i^+ *in the sense that*

$$\pi(\operatorname{spt} f) \subset \{C_1 v \leq u \leq C_2 v\}, \tag{8.2.9}$$

where C_1 *and* C_2 *are positive constants that may additionally depend on* $\eta, \varepsilon, \mathfrak{m}$*, and* λ*. The connected component of* $\mathcal{U} \setminus \pi(\operatorname{spt} f)$ *containing the center is isometric to Minkowski space. The connected component of* $\mathcal{U} \setminus \pi(\operatorname{spt} f)$ *containing future null infinity* $\mathcal{I}^+$ *is isometric to an appropriate neighborhood of future null infinity in the Reissner–Nordström solution with parameters* M *and* e.

5. *The Vlasov matter disperses in the sense that the macroscopic observables decay pointwise:*

$$\mathcal{M} \le C(1+\tau)^{-3}, \tag{8.2.10}$$

where $\mathcal{M} \in \{N^u, N^v, T^{uu}, T^{uv}, T^{vv}, S\}$ *and the constant* C *may additionally depend on* $\eta, \varepsilon, \mathfrak{m}$*, and* λ*.*

The same conclusions hold if $\mathfrak{m} = 0$*, but points 4. and 5. are improved to:*

4.′ *The estimate* (8.2.8) *still holds but* (8.2.9) *is improved to*

$$\pi(\operatorname{spt} f) \subset \{-r_2 \le u \le \tfrac{1}{3} r_1\}, \tag{8.2.11}$$

i.e., the beam is confined to a null slab. The spacetime is isometric to Minkowski space for $u \ge r_1$*.*

5.′ *The Vlasov matter disperses in the sense that the macroscopic observables decay pointwise:*

$$N^v + T^{vv} \le C(1+\tau)^{-2}, \tag{8.2.12}$$

$$N^u + T^{uv} \le C(1+\tau)^{-4}, \tag{8.2.13}$$

$$T^{uu} \le C(1+\tau)^{-6}, \tag{8.2.14}$$

where the constant C *may additionally depend on* η, ε*, and* λ*.*

Remark 8.2.4. An analogous version of Proposition 8.2.3 may be proved for any set of regular center parameters $\alpha = (r_1, r_2, 0, \varpi_2, 0, Q_2) \in \mathcal{P}_\Gamma$ or even $\alpha = (r_1, r_2, 0, 0, 0, 0)$ by evolving the seed data $\mathcal{S}_{\alpha,\eta,\varepsilon}$ given by Definition 8.2.1. In that case, (8.2.7) becomes

$$|\tilde{M} - \varpi_2| + |\tilde{e} - Q_2| \lesssim \eta. \tag{8.2.15}$$

Remark 8.2.5. The decay rate τ^{-3} in (8.2.10) is sharp for massive particles [RR92; Nou05]. The hierarchy of decay rates in (8.2.12)–(8.2.14) is sharp for massless particles [Tay15].

8.3 Estimates on the initial data

For the remainder of this chapter, we assume the notation and hypotheses of Proposition 8.2.3.

Lemma 8.3.1. *For η, ε, and $\mathfrak{m}_0$ sufficiently small and any $\lambda \in [-1,2]$ and $|M'-M| \le \delta$, $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$ is untrapped and consistent with particles of mass $0 \le \mathfrak{m} \le \mathfrak{m}_0$. Let $(\mathcal{U}, r, \Omega^2, Q, f)$ be a development of $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$, such as the one obtained from Proposition 3.2.12. Then the following holds:*

1. *The set $\mathcal{U}$ can be assumed to contain the corner $\mathcal{C}_{r_2} \cap \{v \le \frac{1}{3} r_1\}$. The solution is equal to Minkowski space in this region in the sense that*

$$r = \tfrac{1}{2}(v-u),$$

$$\Omega^2 = 1,$$

$$Q = f = 0$$

on $\mathcal{C}_{r_2} \cap \{v \le \frac{1}{3} r_1\}$.

2. *Estimates on the initial data for the auxiliary beam:*

$$\mathring{f}_{\mathrm{aux}}^{r_1,\theta(\lambda)\eta} \lesssim \theta(\lambda)\eta \mathbf{1}_{[\frac{1}{3}r_1,\frac{2}{3}r_1]\times[\Lambda-1,\Lambda+1]\times[\Lambda-1,\Lambda+1]}, \tag{8.3.1}$$

$$\sup_{v\in[\frac{1}{3}r_1,r_1]} |(\Omega^2-1, \partial_u \log\Omega^2, \partial_v \log\Omega^2, Q, m)(-v,v)| \lesssim \theta(\lambda)\eta, \tag{8.3.2}$$

$$Q(-\tfrac{2}{3}r_1, \tfrac{2}{3}r_1) \gtrsim \theta(\lambda)\eta, \tag{8.3.3}$$

$$\ell(-v,v,p^u,p^v) \approx 1 \quad \textit{for every } (v,p^u,p^v) \in \operatorname{spt}(\mathring{f}_{\mathrm{aux}}^{r_1,\theta(\lambda)\eta}). \tag{8.3.4}$$

3. *Estimates on the initial data for the main beam:*

$$\mathring{f}_{\mathrm{main}}^{\alpha_{\zeta(\lambda),M'},\theta(\lambda)\eta,\varepsilon} \lesssim \varepsilon^{-3}\mathbf{1}_{[r_1,r_2]\times[\varepsilon,2\varepsilon]\times[\varepsilon,2\varepsilon]} \cdot \begin{cases} 1 & \textit{if } \lambda > 0 \\ 0 & \textit{if } \lambda \le 0 \end{cases}, \tag{8.3.5}$$

$$\sup_{v\in[\frac{2}{3}r_1,r_2]} |(\Omega^2, \partial_u \log\Omega^2, \partial_v \log\Omega^2, Q, m)(-v,v) - (\check{\Omega}^2, \check{\underline{\omega}}, \check{\omega}, \check{Q}, \check{m})(v)| \lesssim \theta(\lambda)\eta, \tag{8.3.6}$$

$$\inf_{v\in[\frac{2}{3}r_1]} Q(-v,v) \gtrsim \theta(\lambda)\eta, \tag{8.3.7}$$

$$\ell(-v,v,p^u,p^v) \approx \varepsilon \quad \textit{for every } (v,p^u,p^v) \in \operatorname{spt}(\mathring{f}_{\mathrm{main}}^{\alpha_{\zeta(\lambda),M'},\theta(\lambda)\eta,\varepsilon}), \tag{8.3.8}$$

where

$$\check{m}(v) \doteq \check{\varpi}(v) - \frac{\check{Q}^2(v)}{2v}, \quad \check{\Omega}^2(v) \doteq \left(1 - \frac{2\check{m}(v)}{v}\right)^{-1}, \quad \check{\omega} \doteq -\check{\underline{\omega}} \doteq \frac{1}{2}\frac{d}{dv}\log\check{\Omega}^2(v). \tag{8.3.9}$$

4. Estimates on the initial outgoing cone C_{-r_2}:

$$\varpi(-r_2, v) = \varpi(-r_2, r_2), \tag{8.3.10}$$

$$Q(-r_2, v) = Q(-r_2, r_2), \tag{8.3.11}$$

$$0 \le m(-r_2, v) \le 10M, \tag{8.3.12}$$

$$r(-r_2, v) = \tfrac{1}{2}v + \tfrac{1}{2}r_2 \tag{8.3.13}$$

for $v \ge r_2$.

Proof. Consistency with particles of mass $\mathfrak{m} \le \mathfrak{m}_0$ follows immediately from Definition 8.2.1 and the estimates (8.3.2) and (8.3.6) by taking $\mathfrak{m}_0$ sufficiently small. We therefore focus on proving the estimates and as a byproduct infer the untrapped property of $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$.

Part 1. This is a restatement of Remark 3.2.13.

Part 2. The estimate (8.3.1) follows immediately from the definition (8.2.2). Inserting the ansatz (8.3.1) into (3.2.26)–(3.2.28), we find

$$\begin{aligned} \mathring{\mathcal{N}}^u(r) = \mathring{\mathcal{N}}^v(r) &= \pi\theta(\lambda)\eta\Lambda\varphi\left(\frac{6}{r_1}r - 3\right), \\ \mathring{\mathcal{T}}^{uu}(r) = \mathring{\mathcal{T}}^{vv}(r) &= \pi\theta(\lambda)\eta\left(\Lambda^2 + \int_{-1}^{1} x^2\varphi(x)\,dx\right)\varphi\left(\frac{6}{r_1}r - 3\right), \\ \mathring{\mathcal{T}}^{uv}(r) &= \pi\theta(\lambda)\eta\Lambda^2\varphi\left(\frac{6}{r_1}r - 3\right) \end{aligned}$$

for $r \in [\frac{1}{3}r_1, \frac{2}{3}r_1]$. For η sufficiently small, it then follows readily from the system (3.2.29) and (3.2.30) that $\mathring{Q}$ and $\mathring{m}$ are nonnegative, nondecreasing functions, and

$$0 \le \mathring{Q}(r) + \mathring{m}(r) \lesssim \theta(\lambda)\eta$$

for $r \in [\frac{1}{3}r_1, \frac{2}{3}r_1]$ and

$$\mathring{Q}(\tfrac{2}{3}r_1) \gtrsim \theta(\lambda)\eta.$$

Using the definition of $\mathring{\Omega}^2$, we infer $|\mathring{\Omega}^2 - 1| \lesssim \theta(\lambda)\eta$, and to estimate $|\partial_v \log\Omega^2(-v, v)|$, we observe that

$$
\begin{aligned}
|\partial_v \log \Omega^2(-v,v)| &= \left| \frac{d}{dv} \log \mathring{\Omega}^2(v) \right| \\
&= \left| \mathring{\Omega}^{-2} \left(\frac{2\mathring{m}(v)}{v^2} - \frac{2}{v} \frac{d}{dv} \mathring{m}(v) \right) \right| \\
&= \left| \frac{2\mathring{m}}{\mathring{\Omega}^2 v^2} - \frac{2}{\mathring{\Omega}^2 v} \left(\frac{\mathring{\Omega}^4 v^2}{4} \left(\mathring{T}^{uu} + 2\mathring{T}^{uv} + \mathring{T}^{vv} \right) + \frac{\mathring{Q}^2}{2v^2} \right) \right| \\
&\lesssim \theta(\lambda)\eta.
\end{aligned}
$$

The estimate for $\partial_u \log \Omega^2(-v,v)$ follows from (3.2.36). This establishes (8.3.2) and (8.3.3). Finally, (8.3.4) follows from the mass shell relation and (8.3.2), provided $\mathfrak{m}_0$ is chosen sufficiently small.

Part 3. The estimate (8.3.5) follows immediately from the definition (8.2.3). Inserting the ansatz (8.2.3) into (3.2.26)–(3.2.28), we find

$$
\mathring{\mathcal{N}}^u(r) = \mathring{\mathcal{N}}^v(r) = \frac{1}{\mathfrak{e} r^2} \left(1 - \frac{2\check{\varpi}(r)}{r} + \frac{\check{Q}^2(r)}{r^2} \right)^2 \check{Q}'(r), \tag{8.3.14}
$$

$$
\mathring{\mathcal{T}}^{uu}(r) = \mathring{\mathcal{T}}^{vv}(r) = \frac{\varepsilon}{6\mathfrak{e} r^2} \left(9 + \int_{-1}^{1} x^2 \varphi(x)\, dx \right) \left(1 - \frac{2\check{\varpi}(r)}{r} + \frac{\check{Q}^2(r)}{r^2} \right)^2 \check{Q}'(r), \tag{8.3.15}
$$

$$
\mathring{\mathcal{T}}^{uv}(r) = \frac{3\varepsilon}{4\mathfrak{e} r^2} \left(1 - \frac{2\check{\varpi}(r)}{r} + \frac{\check{Q}^2(r)}{r^2} \right)^2 \check{Q}'(r), \tag{8.3.16}
$$

where $\check{\varpi}$ and $\check{Q}$ are obtained from $\varsigma(\alpha_{\lambda,M'})$. Inserting (8.3.14)–(8.3.16) into (3.2.29) and (3.2.30) yields

$$
\begin{aligned}
\frac{d}{dr}\mathring{m} &= \frac{\mathring{Q}^2}{2r^2} + \mathrm{Err}, \\
\frac{d}{dr}\mathring{Q} &= \left(1 - \frac{2\mathring{m}}{r} \right)^{-2} \left(1 - \frac{2\check{m}}{r} \right)^2 \check{Q}',
\end{aligned}
$$

where

$$
|\mathrm{Err}| \lesssim \left(1 - \frac{2\mathring{m}}{r} \right)^{-2} \varepsilon
$$

and $\mathring{m}(r_1), \mathring{Q}(r_1) \lesssim \theta(\lambda)\eta$. Therefore, by (7.2.6), (7.2.10), and a simple Grönwall and bootstrap argument, $\mathring{m}$ and $\mathring{Q}$ exist on $[r_1, r_2]$ and satisfy

$$
\sup_{r \in [r_1, r_2]} |(\mathring{m} - \check{m}, \mathring{Q} - \check{Q})(r)| \lesssim \theta(\lambda)\eta.
$$

This implies the same estimate for $|\mathring{\Omega}^2 - \check{\Omega}^2|$ by definition. To estimate the other quantities, we may now argue as in the proof of Part 2.

Part 4. Equations (8.3.10), (8.3.11), and (8.3.13) follow immediately from the definitions. Inequality (8.3.12) follows from (8.3.6) provided η is chosen sufficiently small. □

8.4 The main beam in the near region

For $\tau_0 > 0$, let

$$\mathcal{R}^{\tau_0}_{\mathrm{main}} \doteq \{0 \le \tau \le \tau_0\} \cap \{-\tfrac{2}{3}r_1 \le u \le -r_2\} \subset \mathcal{C}_{r_2}. \tag{8.4.1}$$

Lemma 8.4.1. *For any $\breve{v}$, η, ε, and $\mathfrak{m}_0$ satisfying* (8.2.5)*, the following holds. Any normalized development $(\mathcal{U}, r, \Omega^2, Q, f)$ of $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$, such as the one obtained from Proposition 3.2.12, can be uniquely extended to $\mathcal{R}^{\frac{1}{2}\breve{v}-\frac{1}{3}r_1}_{\mathrm{main}}$. Moreover, the solution satisfies the estimates*

$$0 \le m \le 10M, \quad 0 \le Q \le 6M, \tag{8.4.2}$$

$$r \sim v, \quad \Omega^2 \sim 1, \tag{8.4.3}$$

$$\partial_v r \sim 1, \quad |\partial_u r| \lesssim 1, \tag{8.4.4}$$

$$|\partial_u \Omega^2| \lesssim 1, \quad |\partial_v \Omega^2| \lesssim v^{-3} \tag{8.4.5}$$

on $\mathcal{R}^{\frac{1}{2}\breve{v}-\frac{1}{3}r_1}_{\mathrm{main}}$ and

$$\tfrac{1}{3} \le \partial_v r \le \tfrac{2}{3}, \quad \partial_u r \sim -1 \tag{8.4.6}$$

on $\mathcal{R}^{\frac{1}{2}\breve{v}-\frac{1}{3}r_1}_{\mathrm{main}} \cap \underline{C}_{\breve{v}}$. Finally, the support of the distribution function satisfies

$$\pi(\mathrm{spt}\, f) \cap \mathcal{R}^{\frac{1}{2}\breve{v}-\frac{1}{3}r_1}_{\mathrm{main}} \subset \{-\tfrac{5}{6}r_1 \le u \le -r_2\} \tag{8.4.7}$$

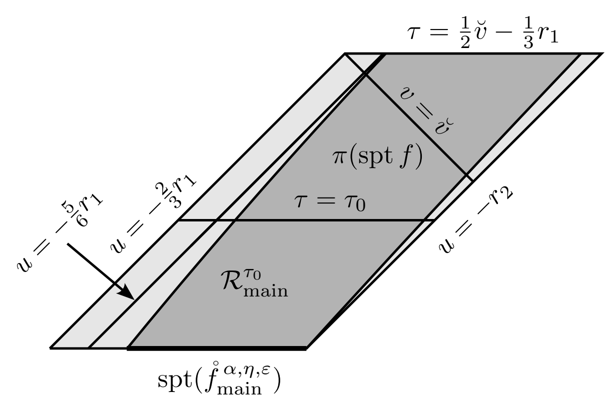

Figure 8.3: Penrose diagram of the bootstrap region $\mathcal{R}^{\tau_0}_{\mathrm{main}}$ used in the proof of Lemma 8.4.1.

and if $u \in [-r_2, -\frac{5}{6}r_1]$, p^u, *and* p^v *are such that* $f(u, \check{v}, p^u, p^v) \neq 0$, *then*

$$\frac{p^u}{p^v} \lesssim \check{v}^{-2}, \quad \frac{\ell^2}{p^v} \lesssim 1. \tag{8.4.8}$$

When $\lambda \leq 0$, (8.4.2) *reads instead*

$$0 \leq m \lesssim \theta(\lambda)\eta, \quad 0 \leq Q \lesssim \theta(\lambda)\eta.$$

The proof of Lemma 8.4.1 will be given on Page 227. We will make use of a bootstrap argument in the regions $\mathcal{R}^{\tau_0}_{\text{main}}$, where τ_0 ranges over $[0, \frac{1}{2}\check{v} - \frac{1}{3}r_1]$. For the basic geometric setup of the lemma and its proof, refer to Fig. 8.3. As the proof is much simpler when $\lambda \leq 0$ (the main beam is absent), we focus only on the case $\lambda > 0$, in which case $\theta(\lambda)\eta = \eta$.

We first make some definitions that will be used to define the bootstrap assumptions. Let $C_1 > 0$ be a constant such that

$$C_1^{-1} \leq \left(1 - \frac{2\check{m}(v)}{v}\right)^{-1} \leq C_1$$

for $v \in [\frac{2}{3}r_1, r_2]$, where $\check{m}$ is given by (8.3.9). (Recall that $\check{m}(v) = 0$ for $v \leq r_1$.) We then define

$$C_2 \doteq 8C_1 \left(\frac{3r_2}{2r_1} - 1\right)\left(5M + \frac{18M^2}{r_1}\right),$$

$$C_3 \doteq 2 \max_{v\in[\frac{2}{3}r_1, r_2]} |\underline{\check{\omega}}(v)| + 100C_1 e^{C_2}\left(5M + \frac{27M^2}{r_1}\right)\int_{\frac{2}{3}r_1}^{\infty} \frac{dv}{(\frac{2}{3}(1-\frac{1}{6}e^{-C_2})r_1 + \frac{1}{6}e^{-C_2}v)^3},$$

The constants C_1, C_2, and C_3 do not depend on η, ε, or $\mathfrak{m}_0$.

The quantitative bootstrap assumptions for the proof of Lemma 8.4.1 are

$$\tfrac{1}{6}e^{-C_2} \leq \partial_v r \leq \tfrac{3}{2}e^{C_2}, \tag{8.4.9}$$

$$\tfrac{1}{8}C_1^{-1} \leq \Omega^{-2}\partial_v r \leq C_1, \tag{8.4.10}$$

$$|\partial_u \log \Omega^2| \leq C_3, \tag{8.4.11}$$

$$\varpi \leq 5M, \tag{8.4.12}$$

$$N^v \leq Ae^{B\tau}, \tag{8.4.13}$$

on $\mathcal{R}^{\tau_0}_{\text{main}}$ where $A \geq 1$ and $B \geq 1$ are constants to be determined which may depend on $\check{v}$ and η, but not on ε. We now derive some consequences of the bootstrap assumptions for the geometry of the solution.

Lemma 8.4.2. *If* (8.2.5) *holds,* $\tau_0 \in [0, \frac{1}{2}\breve{v} - \frac{1}{3}r_1]$*,* $\mathcal{R}^{\tau_0}_{\text{main}} \subset \mathcal{U}$*, and the bootstrap assumptions* (8.4.9)–(8.4.12) *hold on* $\mathcal{R}^{\tau_0}_{\text{main}}$*, then*

$$\eta \lesssim Q \le 6M, \tag{8.4.14}$$

$$\tfrac{2}{3}(1 - \tfrac{1}{6}e^{-C_2})r_1 + \tfrac{1}{6}e^{-C_2}v \le r \le r_2 + \tfrac{3}{2}e^{C_2}v, \tag{8.4.15}$$

$$\Omega^2 \approx 1, \tag{8.4.16}$$

$$|\partial_u r| \lesssim 1 \tag{8.4.17}$$

on $\mathcal{R}^{\tau_0}_{\text{main}}$.

We will frequently use that (8.4.15) implies

$$r \sim v$$

on $\mathcal{R}^{\tau_0}_{\text{main}}$ without further comment.

Proof. For η and ε sufficiently small, $\eta \lesssim Q \le 6M$ on $\{\tau = 0\} \cap \{-r_2 \le u \le -\frac{2}{3}r_1\}$ and $\{\tau \ge 0\} \cap \{v = r_2\}$ by Lemma 8.3.1. Since $N^v \ge 0$ by definition, Maxwell's equation (2.3.24) implies the upper bound in (8.4.14). The lower bound also follows from Maxwell's equation (2.3.25) and $N^u \ge 0$. The inequality (8.4.15) follows from integrating the bootstrap assumption (8.4.9). The inequality (8.4.16) follows directly by multiplying the bootstrap assumptions (8.4.9) and (8.4.10). To estimate $\partial_u r$, we rewrite the definition of the Hawking mass (2.1.2) and the renormalized Hawking mass (2.1.19) as

$$\partial_u r = -\frac{1}{4}\left(1 - \frac{2\varpi}{r} + \frac{Q^2}{r^2}\right)\frac{\Omega^2}{\partial_v r}. \tag{8.4.18}$$

Now (8.4.17) follows immediately from (8.4.10), (8.4.12), and (8.4.15). □

We now use the basic geometric control obtained in Lemma 8.4.2 to obtain crucial control of the electromagnetic geodesic flow. It is convenient to first introduce some notation. Let Γ_f denote the set of maximally extended electromagnetic geodesics $\gamma : I \to \mathcal{U}$, where I is an interval, such that $(\gamma, p)(I) \subset \operatorname{spt} f$, where $p = d\gamma/ds$. If γ passes through the point (u, v), we denote by $s_{u,v}$ the parameter value such that $\gamma(s_{u,v}) = (u, v)$. Let $\Gamma_f(u, v)$ denote the subset of Γ_f consisting of curves passing through (u, v). Note that every curve in Γ_f intersects $\mathcal{C}_{r_2} \cap \{\tau = 0\}$.

Lemma 8.4.3. *If* (8.2.5) *holds, A is sufficiently large depending only on α,* $\tau_0 \in [0, \frac{1}{2}\breve{v} - \frac{1}{3}r_1]$,

$\mathcal{R}^{\tau_0}_{\text{main}} \subset \mathcal{U}$, the bootstrap assumptions (8.4.9)–(8.4.13) *hold on $\mathcal{R}^{\tau_0}_{\text{main}}$, and $(u,v) \in \mathcal{R}^{\tau_0}_{\text{main}}$, then*

$$\mathscr{V}(u,v) \lesssim \frac{\varepsilon^3}{\eta(\varepsilon + \eta \min\{\tau,1\})}\left(1 + \frac{A}{B}e^{B\tau}\right), \tag{8.4.19}$$

where $\mathscr{V}$ is the phase space volume function defined by (8.1.9)*. Furthermore, if $\gamma \in \Gamma_f(u,v)$, then*

$$0 < u - u_0 \lesssim \eta^{-1}\varepsilon, \tag{8.4.20}$$

$$\varepsilon + \eta \min\{\tau, 1\} \lesssim p^v(s_{u,v}) \lesssim \varepsilon + \min\{\tau, 1\}. \tag{8.4.21}$$

where u_0 is the retarded time coordinate of the intersection of γ with $\{\tau = 0\}$.

Proof. Let $\gamma \in \Gamma_f(u,v)$. We will use the Lorentz force written in the form of equation (2.1.25) to estimate p^v. Since $(u,v) \in \mathcal{R}^{\tau_0}_{\text{main}}$, γ intersects $\{\tau = 0\}$ in $\operatorname{spt}(\mathring{f}^{\alpha,\varepsilon}_{\text{main}})$ and therefore has angular momentum $\ell \sim \varepsilon$. By the bootstrap assumptions and Lemma 8.4.2, it holds that

$$\left|\left(\partial_u \log \Omega^2 - \frac{2\partial_u r}{r}\right)\frac{\ell^2}{r^2}\right| \lesssim \frac{\varepsilon^2}{v^2} \tag{8.4.22}$$

along the entire length of γ. Let (u_0, v_0) be the coordinate of the intersection of γ with $\{\tau = 0\}$. Using (8.4.14)–(8.4.16) and the fact that $p^v(s_{u_0,v_0}) \in [\varepsilon, 2\varepsilon]$, we have

$$\mathfrak{e}\frac{Q}{r^2}(\Omega^2 p^v)\bigg|_{s=0} \gtrsim \eta\varepsilon.$$

If ε is sufficiently small (independent of γ, but depending on η), these estimates show that

$$\frac{d}{ds}(\Omega^2 p^v) \gtrsim \eta\varepsilon > 0$$

along γ. Using (8.4.16), we see that

$$p^v \gtrsim \varepsilon \tag{8.4.23}$$

along γ.

It is now convenient to parametrize γ by the advanced time coordinate v of the spacetime. The

Lorentz force equation then becomes

$$\frac{d\gamma^u}{dv} = \frac{p^u}{p^v} = \frac{\ell^2 r^{-2} + \mathfrak{m}^2}{\Omega^2 (p^v)^2}, \tag{8.4.24}$$

$$\frac{d}{dv}(\Omega^2 p^v) = \frac{1}{p^v}\left(\partial_u \log\Omega^2 - \frac{2\partial_u r}{r}\right)\frac{\ell^2}{r^2} + \mathfrak{e}\frac{Q}{r^2}\Omega^2, \tag{8.4.25}$$

where however p^v is still given by $d\gamma^v/ds$ and we have used the mass shell relation in (8.4.24). By (8.4.22) and (8.4.23),

$$\left|\frac{1}{p^v}\left(\partial_u \log\Omega^2 - \frac{2\partial_u r}{r}\right)\frac{\ell^2}{r^2}\right| \lesssim \frac{\varepsilon}{v^2}$$

along γ. This implies, using $v_0 \geq r_1$ and hence $\int_{v_0}^{\infty} v'^{-2}\, dv' \lesssim 1$, that

$$\left|\Omega^2 p^v\big|_{(\gamma^u(v),v)} - \int_{v_0}^{v} \mathfrak{e}\frac{Q}{r^2}\Omega^2\Big|_{(\gamma^u(v'),v')} dv'\right| \lesssim \varepsilon. \tag{8.4.26}$$

Using Lemma 8.4.2, we readily deduce that

$$\int_{v_0}^{v} \mathfrak{e}\frac{Q}{r^2}\Omega^2\Big|_{(\gamma^u(v'),v')} dv' \lesssim 1 \tag{8.4.27}$$

and

$$\int_{v_0}^{v} \mathfrak{e}\frac{Q}{r^2}\Omega^2\Big|_{(\gamma^u(v'),v')} dv' \gtrsim \eta\left(\frac{1}{v_0} - \frac{1}{v}\right) \gtrsim \eta \min\{1, v - v_0\}.$$

Combining this with (8.4.23) and (8.4.26), we deduce

$$\varepsilon + \eta \min\{1, v - v_0\} \lesssim p^v(v) \lesssim 1 \tag{8.4.28}$$

along γ.

We are now able to prove (8.4.20). Since $r \lesssim v \lesssim \check{v}$, $r^2\mathfrak{m}^2 \lesssim \varepsilon^2 \lesssim \ell^2$ for $\mathfrak{m}_0$ sufficiently small while respecting the hierarchy (8.2.5). Therefore, using also Lemma 8.4.2 and (8.4.28), we find

$$\frac{d\gamma^u}{dv} \lesssim \frac{\varepsilon^2}{v^2(\varepsilon + \eta\min\{1, v - v_0\})^2}. \tag{8.4.29}$$

If $v \in [v_0, v_0 + 1]$, we compute

$$\int_{v_0}^{v} \frac{\varepsilon^2}{v'^2(\varepsilon + \eta(v' - v_0))^2}\, dv' \lesssim \frac{\varepsilon(v - v_0)}{\varepsilon + \eta(v - v_0)} \lesssim \eta^{-1}\varepsilon$$

and if $v \in [v_0+1, \infty)$, we compute

$$\int_{v_0+1}^{v} \frac{\varepsilon^2}{v'^2(\varepsilon+\eta)^2}\, dv' \lesssim \eta^{-2}\varepsilon^2 \lesssim \eta^{-1}\varepsilon,$$

for $\varepsilon \le \eta$. Therefore, integrating (8.4.29), we find

$$u - u_0 = \gamma^u(v) - \gamma^u(v_0) = \int_{v_0}^{v} \frac{d\gamma^u}{dv}\, dv' \lesssim \eta^{-1}\varepsilon, \tag{8.4.30}$$

which proves (8.4.20).

Now $u_0 = -v_0$, so (8.4.30) implies $u + v_0 \lesssim \eta^{-1}\varepsilon$. Therefore, we have

$$2\tau = v + u = (v - v_0) + (u + v_0) \lesssim \eta^{-1}\varepsilon + v - v_0$$

and

$$\varepsilon + \eta\tau \lesssim \varepsilon + \eta(v - v_0) \lesssim p^v(v)$$

for $\tau \lesssim 1$. This, together with (8.4.26) and (8.4.27), proves (8.4.21).

To prove (8.4.19), we use the approximate representation formula (8.4.26) for p^v and the change of variables formula (8.1.9). Using the bootstrap assumptions, Lemma 8.4.2, the mean value theorem, Maxwell's equation (2.3.24), and the estimate (8.4.20), we have

$$\left| \frac{Q}{r^2}\Omega^2\Big|_{(\gamma^u(v'),v')} - \frac{Q}{r^2}\Omega^2\Big|_{(u,v')} \right| \lesssim \left(1 + \sup_{[\gamma^u(v'),u]\times\{v'\}} |\partial_u Q|\right)(\gamma^u(v') - u) \lesssim A\eta^{-1}\varepsilon e^{B(u+v')/2}$$

for every $v' \in [v_0, v]$ and A sufficiently large depending only on α. Using this and (8.4.26), we find

$$\begin{aligned}
\left| \Omega^2 p^v\big|_{(\gamma^u(v),v)} - \int_{v_0}^{v} \mathfrak{e}\frac{Q}{r^2}\Omega^2\Big|_{(u,v')} dv' \right| &\lesssim \left| \Omega^2 p^v\big|_{(\gamma^u(v),v)} - \int_{v_0}^{v} \mathfrak{e}\frac{Q}{r^2}\Omega^2\Big|_{(\gamma^u(v'),v')} dv' \right| \\
&\quad + \int_{v_0}^{v} \left| \frac{Q}{r^2}\Omega^2\Big|_{(\gamma^u(v'),v')} - \frac{Q}{r^2}\Omega^2\Big|_{(u,v')} \right| dv' \\
&\lesssim \varepsilon + \int_{v_0}^{v} A\eta^{-1}\varepsilon e^{B(u+v')/2}\, dv' \\
&\le \varepsilon\left(1 + \frac{A}{B\eta} e^{B(u+v)/2}\right).
\end{aligned} \tag{8.4.31}$$

Next, we estimate

$$0 \le \int_{-u}^{v_0} \mathfrak{e}\frac{Q}{r^2}\Omega^2\Big|_{(u,v')} dv' \lesssim v_0 + u = u - u_0 \lesssim \eta^{-1}\varepsilon. \tag{8.4.32}$$

Combining (8.4.31) and (8.4.32) yields

$$\left|\Omega^2 p^v\big|_{(u,v)} - \int_{-u}^{v} \mathfrak{e}\frac{Q}{r^2}\Omega^2\bigg|_{(u,v')} dv'\right| \lesssim \frac{\varepsilon}{\eta}\left(1+\frac{A}{B}e^{B(u+v)/2}\right). \tag{8.4.33}$$

Therefore, if $\gamma_1, \gamma_2 \in \Gamma_f(u,v)$ and we parametrize both by advanced time, and denote the v-momentum of γ_i by p_i^v for $i = 1, 2$, we find

$$\begin{aligned} |p_1^v(v) - p_2^v(v)| &\le \left|p_1^v(v) - \frac{1}{\Omega^2(u,v)}\int_{-u}^{v} \mathfrak{e}\frac{Q}{r^2}\Omega^2\bigg|_{(u,v')} dv'\right| + \left|p_2^v(v) - \frac{1}{\Omega^2(u,v)}\int_{-u}^{v} \mathfrak{e}\frac{Q}{r^2}\Omega^2\bigg|_{(u,v')} dv'\right| \\ &\lesssim \frac{\varepsilon}{\eta}\left(1+\frac{A}{B}e^{B(u+v)/2}\right). \end{aligned}$$

Inserting this estimate, (8.4.21), and $\ell \sim \varepsilon$ in (8.1.9) yields (8.4.19), as desired. □

Proof of Lemma 8.4.1. The proof is a bootstrap argument based on the bootstrap assumptions (8.4.9)–(8.4.13) and continuation criterion given by the extension principle Proposition 3.2.4. Let

$$\mathcal{A} \doteq \{\tau_0 \in [0, \tfrac{1}{2}\breve{v} - \tfrac{1}{3}r_1] : \text{the solution extends uniquely to } \mathcal{R}^{\tau_0}_{\mathrm{main}} \text{ and (8.4.9)–(8.4.13) hold on } \mathcal{R}^{\tau_0}_{\mathrm{main}}\}.$$

The set $\mathcal{A}$ is nonempty by Proposition 3.2.12 if A is chosen sufficiently large and η, ε, and $\mathfrak{m}_0$ are sufficiently small. It is also manifestly connected and closed by continuity of the bootstrap assumptions. We now show that if A and B are sufficiently large depending on η and (8.2.5) holds, then $\mathcal{A}$ is also open.

Let $\tau_0 \in \mathcal{A}$. First, we use Lemmas 8.4.2 and 8.4.3 to estimate N^v and improve (8.4.13). Since f is transported along electromagnetic geodesics, we have $f(u,v,p^u,p^v) \lesssim \varepsilon^{-3}$ for $(u,v) \in \mathcal{R}^{\tau_0}_{\mathrm{main}}$. Using (8.4.19) and (8.4.21), we infer directly from the definition of N^v that

$$N^v(u,v) \lesssim (\varepsilon + \min\{\tau,1\})\varepsilon^{-3}\mathscr{V}(u,v) \lesssim \frac{\varepsilon + \min\{\tau,1\}}{\eta(\varepsilon + \eta\min\{\tau,1\})}\left(1+\frac{A}{B}\right)e^{B\tau} \lesssim \eta^{-2}\left(1+\frac{A}{B}\right)e^{B\tau}.$$

Letting $C_* = C_*(\mathfrak{e}, \varphi, \Lambda, \alpha)$ denote the implicit constant in this inequality, and choosing $A = 4C_*\eta^{-2}$ and $B \ge A$, we see that

$$N^v(u,v) \le \tfrac{1}{2}Ae^{B\tau},$$

which improves (8.4.13).

To continue, we now estimate N^u, T^{uu}, and T^{uv} in the same fashion, making use now of the strong decay of p^u. If $\gamma \in \Gamma_f(u,v)$, then

$$p^u(s_{u,v}) = \frac{\ell^2 r^{-2} + \mathfrak{m}^2}{\Omega^2 p^v(s_{u,v})} \lesssim \frac{\varepsilon^2}{\varepsilon + \eta \min\{\tau, 1\}}$$

by (8.4.21). Using (8.4.19), we therefore find

$$N^u \lesssim \frac{\varepsilon^2}{\eta} \frac{1}{(\varepsilon + \eta \min\{1,\tau\})^2} A e^{B\tau}, \tag{8.4.34}$$

$$T^{uv} \lesssim \frac{\varepsilon^2}{\eta} \frac{\varepsilon + \min\{\tau, 1\}}{\varepsilon + \eta \min\{1,\tau\}} A e^{B\tau}, \tag{8.4.35}$$

$$T^{uu} \lesssim \frac{\varepsilon^4}{\eta} \frac{1}{(\varepsilon + \eta \min\{1,\tau\})^3} A e^{B\tau}. \tag{8.4.36}$$

Using the hierarchy (8.2.5), these estimates imply

$$r^2 T^{uv} + r^2 T^{uu} + \int_{-u}^{v} r^2 N^u(u, v')\, dv' \lesssim \varepsilon^{1/2} \tag{8.4.37}$$

for any $(u,v) \in \mathcal{R}^{\tau_0}_{\text{main}}$. With these final estimates in hand, we may begin to improve the remaining bootstrap assumptions (8.4.9)–(8.4.12). We will then carry out the rest of the continuity argument and prove all of the stated conclusions of the lemma.

Improving (8.4.9): The wave equation (2.3.20) can be rewritten as

$$\partial_u \partial_v r = -\frac{1}{2r^2} \frac{\Omega^2}{\partial_v r} \left(\varpi - \frac{Q^2}{r} \right) \partial_v r + \frac{r}{4} \Omega^4 T^{uv}.$$

Using an integrating factor, we find

$$\partial_u \left[\exp\left(\int_{-r_2}^{u} \frac{1}{2r} \frac{\Omega^2}{\partial_v r} \left(\varpi - \frac{Q^2}{r} \right) du' \right) \partial_v r \right] = \frac{r}{4} \Omega^4 T^{uv} \exp\left(\int_{-r_2}^{u} \frac{1}{2r} \frac{\Omega^2}{\partial_v r} \left(\varpi - \frac{Q^2}{r} \right) du' \right), \tag{8.4.38}$$

where the integral is taken over fixed v. The bootstrap assumptions imply

$$\int_{-r_2}^{-\frac{2}{3} r_1} \frac{1}{2r} \frac{\Omega^2}{\partial_v r} \left| \varpi - \frac{Q^2}{r} \right| du' \le C_2.$$

For ε sufficiently small, the right-hand side of (8.4.38) is pointwise $\le \frac{1}{10}$ on $\mathcal{R}^{\tau_0}_{\text{main}}$, so integrating this equation yields

$$\tfrac{2}{5} e^{-C_2} \le \partial_v r(u,v) \le \tfrac{3}{5} e^{C_2}$$

for any $(u,v) \in \mathcal{R}^{\tau_0}_{\text{main}}$, which improves (8.4.9).

Improving (8.4.10): Raychaudhuri's equation (2.3.23) can be rewritten as

$$\partial_v \log\left(\frac{\partial_v r}{\Omega^2}\right) = -\frac{r}{4}\partial_v r T^{uu}. \tag{8.4.39}$$

To improve the upper bound in (8.4.10), note that the right-hand side of (8.4.39) is nonpositive by (8.4.9) and hence

$$\frac{\partial_v r}{\Omega^2}(u,v) \le \frac{\partial_v r}{\Omega^2}(-v,v) = \frac{1}{2}\left(1 - \frac{2\check{m}(v)}{v}\right) \le \frac{C_1}{2}.$$

To improve the lower bound, note that the right-hand side of (8.4.39) is bounded by $\log 2$ in absolute value for ε sufficiently small and hence

$$\frac{\partial_v r}{\Omega^2}(u,v) \ge \frac{1}{2}\frac{\partial_v r}{\Omega^2}(-v,v) \ge \frac{1}{4C_1},$$

which improves (8.4.10).

Improving (8.4.11): The wave equation (2.3.21) can be rewritten as

$$\partial_v\partial_u \log\Omega^2 = \frac{\Omega^2}{r^3}\left(\varpi - \frac{3Q^2}{2r}\right) - \frac{1}{2}\Omega^4 T^{uv} - \Omega^2 S. \tag{8.4.40}$$

Integrating this equation in v and using the bootstrap assumptions, (8.3.6), and (8.4.37) yields

$$|\partial_u \log\Omega^2| \le \frac{3}{4}C_3$$

for η and ε sufficiently small, which improves (8.4.11).

Improving (8.4.12): Integrating the evolution equation for the renormalized Hawking mass (2.3.31) and using (8.4.37), we have

$$|\varpi(u,v) - \varpi(-v,v)| \lesssim \varepsilon^{1/2},$$

which improves (8.4.12).

We have thus improved the constants in all of the bootstrap assumptions (8.4.9)–(8.4.13). Using the local existence theory Proposition 3.2.3 and generalized extension principle Proposition 3.2.4, there exists a $\tau_0' > \tau_0$ such that $\mathcal{U} \subset \mathcal{R}_{\mathrm{main}}^{\tau_0'}$. Choosing $\tau_0' > \tau_0$ perhaps smaller, the bootstrap assumptions (8.4.9)–(8.4.13) extend to $\mathcal{R}_{\mathrm{main}}^{\tau_0'}$ by continuity. Therefore, $\mathcal{A}$ is open and the bootstrap argument is complete.

We now prove the remaining conclusions of the lemma. First, $m(u,v) \ge 0$ for every $(u,v) \in$

$\mathcal{R}_{\text{main}}^{\frac{1}{2}\check{v}-\frac{1}{3}r_1}$ because either $\partial_u r(u,v) < 0$ (and by Raychaudhuri (2.3.22) also for any $u' > u$) or $\partial_u r(u,v) \geq 0$ and then $m(u,v) \geq 0$ directly from the definition (2.1.2). In the first case, the evolution equation (2.3.29) implies m is nondecreasing along the outgoing cone terminating at (u,v). Since this cone either intersects $\{\tau = 0\}$, where $m \geq 0$, or a sphere where $\partial_u r \leq 0$ (and hence $m \geq 0$), we conclude $m(u,v) \geq 0$. Now integrating the wave equation (2.3.21) in u and using (8.4.37), we see that $|\partial_v \log \Omega^2| \lesssim v^{-3}$. Together with the bootstrap assumptions and Lemma 8.4.2, this proves (8.4.2)–(8.4.5). Next, (8.4.6) follows from integrating the wave equation (2.3.20) in u along $\underline{C}_{\check{v}}$ and taking $\check{v} \sim r$ sufficiently large and similarly in (8.4.18). The inclusion (8.4.7) follows immediately from the u-deflection estimate (8.4.20) for electromagnetic geodesics in Γ_f and the hierarchy (8.2.5). Finally, to prove (8.4.8) we use the mass shell relation, (8.4.21), and the parameter hierarchy to estimate

$$r^2 \frac{p^u}{p^v} = \frac{\ell^2 + r^2 \mathfrak{m}^2}{\Omega^2 (p^v)^2} \lesssim \frac{\varepsilon^2}{\eta^2} \lesssim 1,$$
$$\frac{\ell^2}{p^v} \lesssim \frac{\varepsilon^2}{\eta} \lesssim 1,$$

which completes the proof. □

8.5 The auxiliary beam in the near region

For $v_0 > 0$, let

$$\mathcal{R}_{\text{aux}}^{v_0} \doteq \{v \geq u\} \cap \{\tau \geq 0\} \cap \{-\tfrac{2}{3} r_1 \leq u \leq \tfrac{1}{3} r_1\} \cap \{\tfrac{1}{3} r_1 \leq v \leq v_0\}, \tag{8.5.1}$$

$$\tilde{\mathcal{R}}_{\text{aux}}^{v_0} \doteq \{v \geq u\} \cap \{\tau \geq 0\} \cap \{u \geq -\tfrac{2}{3} r_1\} \cap \{v \leq v_0\}. \tag{8.5.2}$$

Lemma 8.5.1. *For any $\check{v}$, η, ε, and $\mathfrak{m}_0$ satisfying* (8.2.5)*, the following holds. The development of $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$ obtained in Lemma 8.4.1 can be uniquely extended to $\tilde{\mathcal{R}}_{\text{aux}}^{\check{v}}$. The spacetime is vacuum for $u \geq \frac{1}{3} r_1$ and $v \leq \check{v}$. Moreover, the solution satisfies the estimates*

$$0 \leq m \lesssim \theta(\lambda)\eta, \quad 0 \leq Q \lesssim \theta(\lambda)\eta,$$

$$\Omega^2 \sim 1, \quad \partial_v r \sim -\partial_u r \sim 1,$$

$$(1+u^2)|\partial_u \Omega^2| + (1+v^2)|\partial_v \Omega|^2 \lesssim 1$$

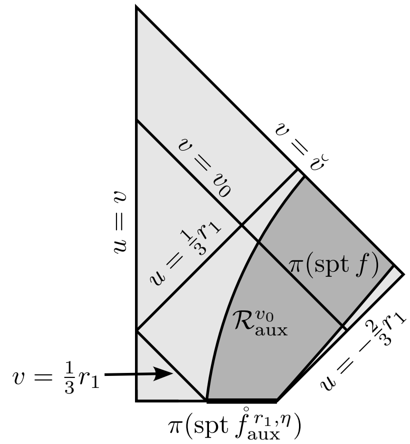

Figure 8.4: Penrose diagram of the bootstrap region $\mathcal{R}_{\text{aux}}^{v_0}$ used in the proof of Lemma 8.5.1.

on $\tilde{\mathcal{R}}_{\text{aux}}^{\check{v}}$ *and*

$$\tfrac{1}{4} \le \partial_v r \le \tfrac{3}{4}$$

on $\tilde{\mathcal{R}}_{\text{aux}}^{\check{v}} \cap \underline{C}_{\check{v}}$. *Finally, the support of the distribution function satisfies*

$$\pi(\operatorname{spt} f) \cap \tilde{\mathcal{R}}_{\text{aux}}^{\check{v}} \subset \{-\tfrac{2}{3}r_1 \le u \le \tfrac{1}{6}r_1\},$$

$$\inf_{\operatorname{spt}(f)\cap\tilde{\mathcal{R}}_{\text{aux}}^{\check{v}}} r \ge \tfrac{1}{6}r_1,$$

and if $u \in [-\frac{2}{3}r_1, \frac{1}{3}r_1]$, p^u, *and* p^v *are such that* $f(u, \check{v}, p^u, p^v) \neq 0$, *then*

$$\frac{p^u}{p^v} \lesssim \check{v}^{-2}, \quad \frac{\ell^2}{p^v} \lesssim 1. \tag{8.5.3}$$

The proof of Lemma 8.5.1 will be given on Page 234. We will make use of a bootstrap argument in the regions $\mathcal{R}_{\text{aux}}^{v_0}$, where v_0 ranges over $[\frac{1}{3}r_1, \check{v}]$. The triangle $\{v \ge u\} \cap \{u \ge \frac{1}{3}r_1\} \cap \{v \le \check{v}\}$ is Minkowskian and can simply be attached at the very end of the argument, cf. Lemma 3.2.16. For the basic geometric setup of the lemma and its proof, refer to Fig. 8.4.

For $(u, v) \in \mathcal{R}_{\text{aux}}^{\check{v}}$, let

$$\hat{r}(u, v) = \frac{r_1}{6} - \frac{u}{2} + \frac{1}{2}\int_{\frac{1}{3}r_1}^{v} \beta(v')\, dv',$$

$$\hat{\Omega}^2(u, v) = \beta(v),$$

where

$$\beta(v) \doteq \begin{cases} 1 & \text{if } v < \frac{2}{3}r_1 \\ \Omega^2(-\frac{2}{3}r_1, v) & \text{if } v \geq \frac{2}{3}r_1 \end{cases}.$$

It is easily verified that $(\hat{r}, \hat{\Omega}^2)$ is a solution of the spherically symmetric Einstein vacuum equations and matches smoothly in v with (r, Ω^2) from $\mathcal{R}^{\frac{1}{2}\check{v}-\frac{1}{3}r_1}_{\text{main}}$ along $C_{-\frac{2}{3}r_1}$.

The first bootstrap assumption is

$$|Q| + |\varpi| + |\partial_v r - \partial_v \hat{r}| + |\partial_u r - \partial_u \hat{r}| + |\Omega^2 - \hat{\Omega}^2| + |\partial_v \Omega^2 - \partial_v \hat{\Omega}^2| + |\partial_u \Omega^2| \leq A\theta(\lambda)\eta e^{B\tau}, \quad (8.5.4)$$

where $A \geq 1$ and $B \geq 1$ are constants to be determined that may depend on $\check{v}$ but not η. We also make the following assumption on the electromagnetic geodesic flow. For $(v_0', p_0^u, p_0^v) \in \operatorname{spt}(\mathring{f}^{r_1,\theta(\lambda)\eta}_{\text{aux}})$, let γ be an electromagnetic geodesic of mass $\mathfrak{m}$ for (r, Ω^2, Q) starting at $(-v_0', v_0', p_0^u, p_0^v)$, and let $\hat{\gamma}$ be a *null geodesic* for $(\hat{r}, \hat{\Omega}^2)$ starting at $(-v_0', v_0', p_0^u, p_0^v)$. Then, assuming both γ and $\hat{\gamma}$ remain within $\mathcal{R}^{v_0}_{\text{aux}}$, we assume that

$$|\Omega^2 p^u - \hat{\Omega}^2 \hat{p}^u| + |\Omega^2 p^v - \hat{\Omega}^2 \hat{p}^v| \leq A\theta(\lambda)\eta e^{B(\gamma^v - v_0')}. \quad (8.5.5)$$

First, we note the following immediate consequences of the first bootstrap assumption:

Lemma 8.5.2. *If* (8.2.5) *holds,* $v_0 \in [\frac{1}{3}r_1, \check{v}]$*,* $\mathcal{R}^{v_0}_{\text{aux}} \subset \mathcal{U}$*, the bootstrap assumption* (8.5.4) *holds on* $\mathcal{R}^{v_0}_{\text{aux}}$*, and* η *is sufficiently small depending on* A *and* B*, then on* $\mathcal{R}^{v_f}_{\text{aux}}$ *it holds that*

$$0 \leq \varpi \lesssim \theta(\lambda)\eta, \quad 0 \leq Q \lesssim \theta(\lambda)\eta,$$
$$|\log \partial_v r| + |\log(-\partial_u r)| + |\log \Omega^2| \lesssim 1,$$
$$|\partial_v \Omega^2| + |\partial_u \Omega^2| \lesssim 1.$$

Next, we use the second bootstrap assumption to obtain

Lemma 8.5.3. *If* (8.2.5) *holds,* $v_0 \in [\frac{1}{3}r_1, \check{v}]$*,* $\mathcal{R}^{v_0}_{\text{aux}} \subset \mathcal{U}$*, the bootstrap assumptions* (8.5.4) *and* (8.5.5) *hold on* $\mathcal{R}^{v_0}_{\text{aux}}$*,* B *is sufficiently large, and* η *is sufficiently small depending on* A *and* B*, then the following holds. Let* $\gamma : [0, S] \to \mathcal{R}^{v_0}_{\text{aux}}$ *be a future-directed electromagnetic geodesic starting in*

$\mathrm{spt}(\mathring{f}_{\mathrm{aux}}^{r_1,\theta(\lambda)\eta})$, *then*

$$r \geq \tfrac{1}{6} r_1,$$
$$u \leq \tfrac{1}{3} r_1,$$
$$r^2 \frac{p^u}{p^v} \lesssim 1, \quad \frac{\ell^2}{p^v} \lesssim 1$$

along γ.

Proof. Upon making the coordinate transformation

$$\tilde{u} = u, \quad \tilde{v} = \tfrac{1}{3} r_1 + \int_{\frac{1}{3} r_1}^{v} \beta(v')\, dv', \tag{8.5.6}$$

the metric $(\hat{r}, \hat{\Omega}^2)$ is brought into the standard Minkowski form $(\frac{1}{2}(\tilde{v} - \tilde{u}), 1)$. If $\tilde{t} \doteq \frac{1}{2}(\tilde{v} + \tilde{u})$, $\hat{\gamma}$ is a null geodesic in $\mathcal{R}_{\mathrm{aux}}^{\check{v}}$ with respect to $(\hat{r}, \hat{\Omega}^2)$ intersecting $\{\tau = 0\}$ with momentum $(p_0^u, p_0^v) = (p_0^{\tilde{u}}, p_0^{\tilde{v}})$ at an area-radius of $\hat{r}_0$, then it is easy to check that

$$\hat{r}^2 = \left(\tilde{t} + \mathrm{sign}(p_0^v - p_0^u) \sqrt{\hat{r}_0^2 - \frac{\hat{\ell}^2}{\hat{E}^2}} \right)^2 + \frac{\hat{\ell}^2}{\hat{E}^2}, \tag{8.5.7}$$

$$p^{\tilde{u}} = \begin{cases} \hat{E} + \sqrt{\hat{E}^2 - \hat{\ell}^2/\hat{r}^2} & \text{if } \tilde{t} < -\,\mathrm{sign}(p_0^v - p_0^u)\sqrt{\hat{r}_0^2 - \hat{\ell}^2/\hat{E}^2} \\ \hat{E} - \sqrt{\hat{E}^2 - \hat{\ell}^2/\hat{r}^2} & \text{if } \tilde{t} \geq -\,\mathrm{sign}(p_0^v - p_0^u)\sqrt{\hat{r}_0^2 - \hat{\ell}^2/\hat{E}^2} \end{cases}, \tag{8.5.8}$$

$$p^{\tilde{v}} = \begin{cases} \hat{E} - \sqrt{\hat{E}^2 - \hat{\ell}^2/\hat{r}^2} & \text{if } \tilde{t} < -\,\mathrm{sign}(p_0^v - p_0^u)\sqrt{\hat{r}_0^2 - \hat{\ell}^2/\hat{E}^2} \\ \hat{E} + \sqrt{\hat{E}^2 - \hat{\ell}^2/\hat{r}^2} & \text{if } \tilde{t} \geq -\,\mathrm{sign}(p_0^v - p_0^u)\sqrt{\hat{r}_0^2 - \hat{\ell}^2/\hat{E}^2} \end{cases} \tag{8.5.9}$$

along $\hat{\gamma}$, where $\hat{\ell}^2 \doteq \hat{r}^2 p^{\tilde{u}} p^{\tilde{v}}$ and $\hat{E} \doteq \frac{1}{2}(p^{\tilde{v}} + p^{\tilde{u}})$ are conserved quantities. If $(p_0^u, p_0^v) \in [\Lambda - 1, \Lambda + 1]^2$ and Λ satisfies (8.2.1), then

$$\tfrac{81}{100} \hat{r}_0^2 \leq \frac{\hat{\ell}^2}{\hat{E}^2} \leq \hat{r}_0^2.$$

From (8.5.7) it is apparent that

$$\min_{\hat{\gamma}} r \geq \frac{\hat{\ell}}{\hat{E}} \geq \tfrac{9}{10} \hat{r}_0, \tag{8.5.10}$$

$$\sup_{\hat{\gamma}} \tilde{u} = -\,\mathrm{sign}(p_0^v - p_0^u) \sqrt{\hat{r}_0^2 - \frac{\hat{\ell}^2}{\hat{E}^2}} \leq \tfrac{45}{100} \hat{r}_0, \tag{8.5.11}$$

and by inspection of (8.5.8) and (8.5.9) that

$$\hat{r}^2\frac{p^{\tilde{u}}}{p^{\tilde{v}}}\lesssim 1,\quad \frac{\hat{\ell}^2}{p^{\tilde{v}}}\lesssim 1 \tag{8.5.12}$$

along $\hat{\gamma}$.

Let γ and $\hat{\gamma}$ be as defined before (8.5.5). Parametrize γ and $\hat{\gamma}$ by v as in the proof of Lemma 8.4.3. Then

$$\left|\frac{d}{dv}(\gamma^u-\hat{\gamma}^u)\right|=\left|\frac{p^u}{p^v}-\frac{\hat{p}^u}{\hat{p}^v}\right|\lesssim A\theta(\lambda)\eta e^{B(v-v_0')}$$

by (8.5.5) and the observation that $p^v\gtrsim 1$ and $p^{\tilde{v}}\gtrsim 1$, so that

$$|\gamma^u-\hat{\gamma}^u|\le\theta(\lambda)\eta e^{B(v-v_0')} \tag{8.5.13}$$

for B chosen sufficiently large. Therefore, the conclusions of the lemma follow from the estimates (8.5.10)–(8.5.12) and the fact that $\hat{r}_0\in[\frac{1}{3}r_1,\frac{2}{3}r_1]$ after undoing the coordinate transformation (8.5.6) and applying the bootstrap assumptions. □

Proof of Lemma 8.5.1. The proof is a bootstrap argument based on the bootstrap assumptions (8.5.4) and (8.5.5), the continuation criterion Proposition 3.2.4, and Lemma 3.2.16. Define the bootstrap set

$$\mathcal{A}\doteq\{v_0\in[\tfrac{1}{3}r_1,\check{v}]:\text{the solution extends uniquely to }\mathcal{R}^{v_0}_{\text{aux}}\text{ and (8.5.4), (8.5.5) hold on }\mathcal{R}^{v_0}_{\text{aux}}\}.$$

The set $\mathcal{A}$ is nonempty by Propositions 3.2.6 and 3.2.12 if A is chosen sufficiently large and is manifestly closed and connected. We now show that if the parameters satisfy (8.2.5), then $\mathcal{A}$ is also open.

Let $v_0\in\mathcal{A}$. Taking $\mathfrak{m}_0$ sufficiently small and using the formula (8.1.9), the initial data estimate Lemma 8.3.1, and Lemmas 8.5.2 and 8.5.3, we immediately find

$$N^u+N^v+T^{uu}+T^{uv}+T^{vv}+S\lesssim r^{-2}\theta(\lambda)\eta\mathbf{1}_{\{r\ge\frac{1}{6}r_1\}} \tag{8.5.14}$$

on $\mathcal{R}^{v_0}_{\text{aux}}$. By (8.3.2) and the observation that $f=0$ along $C_{-\frac{2}{3}r_1}\cap\mathcal{R}^{v_0}_{\text{aux}}$, we have

$$|Q|+|\varpi|+|\partial_v r-\partial_v\hat{r}|+|\Omega^{-2}\partial_v r-\hat{\Omega}^{-2}\partial_v\hat{r}|+|\Omega^2-\hat{\Omega}^2|+|\partial_v\Omega^2-\partial_v\hat{\Omega}^2|+|\partial_u\Omega^2|\lesssim\theta(\lambda)\eta \tag{8.5.15}$$

along $\{\tau = 0\}$ and $C_{-\frac{2}{3}r_1}$ in $\mathcal{R}^{v_0}_{\mathrm{aux}}$. Then, using (8.5.14) and the Einstein–Maxwell–Vlasov system, we see that (8.5.15) holds on $\mathcal{R}^{v_0}_{\mathrm{aux}}$. This improves (8.5.4) for an appropriate choice of A (independent of η).

We now improve (8.5.5). Using the Lorentz force equations (2.1.24) and (2.1.25), we have

$$\frac{d}{dv}(\Omega^2 p^u - \hat{\Omega}^2 \hat{p}^u) = \left(\partial_v \log \Omega^2 - \frac{2\partial_v r}{r}\right) \frac{\ell^2}{r^2 p^v}\bigg|_{\gamma} - \left(\partial_v \log \hat{\Omega}^2 - \frac{2\partial_v \hat{r}}{\hat{r}}\right) \frac{\hat{\ell}^2}{\hat{r}^2 \hat{p}^v}\bigg|_{\hat{\gamma}} - \mathfrak{e}\frac{\Omega^2 Q}{r^2}\frac{p^u}{p^v},$$

$$\frac{d}{dv}(\Omega^2 p^v - \hat{\Omega}^2 \hat{p}^v) = \left(\partial_u \log \Omega^2 - \frac{2\partial_u r}{r}\right) \frac{\ell^2}{r^2 p^v}\bigg|_{\gamma} + \frac{2\partial_u \hat{r}}{\hat{r}} \frac{\hat{\ell}^2}{\hat{r}^2 \hat{p}^v}\bigg|_{\hat{\gamma}} + \mathfrak{e}\frac{\Omega^2 Q}{r^2}.$$

Using the parameter hierarchy (8.2.5), the bootstrap assumptions (8.5.4) and (8.5.5), Lemmas 8.5.2 and 8.5.3, and the bound (8.5.13), we can estimate

$$\left|\frac{d}{dv}(\Omega^2 p^u - \hat{\Omega}^2 \hat{p}^u)\right| + \left|\frac{d}{dv}(\Omega^2 p^u - \hat{\Omega}^2 \hat{p}^u)\right| \lesssim \theta(\lambda)\eta + A\theta(\lambda)\eta e^{B(v-v_0)}.$$

Integrating and choosing the constants A and B sufficiently large in terms of the implied constants and $\check{v}$ improves (8.5.5) and shows that the solution extends to $\mathcal{R}^{\check{v}}_{\mathrm{aux}}$.

The solution is at once extended to $\tilde{\mathcal{R}}^{\check{v}}_{\mathrm{aux}}$ by Lemma 3.2.16. The rest of the conclusions of the lemma follow immediately from Lemmas 3.2.16, 8.4.1, 8.5.2 and 8.5.3 and (8.5.14). □

8.6 The far region

Lemma 8.6.1. *For any $\check{v}$, η, ε, and $\mathfrak{m}_0$ satisfying* (8.2.5)*, there exists a constant $C_\nu > 0$ such that the following holds. The development of $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$ obtained in Lemma 8.5.1 can be uniquely extended*

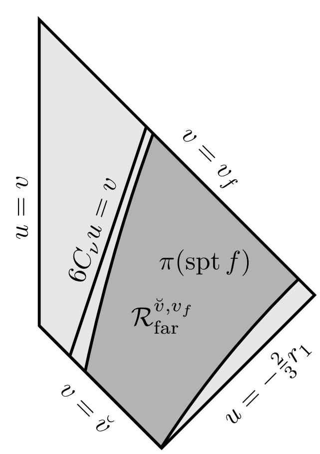

Figure 8.5: Penrose diagram of the bootstrap region $\mathcal{R}^{\check{v},v_f}_{\mathrm{far}}$ used in the proof of Lemma 8.6.1.

globally to $\mathcal{C}_{r_2}$. Moreover, the solution satisfies the estimates

$$0 \le m \le 10M, \quad 0 \le Q \le 6M, \tag{8.6.1}$$

$$\Omega^2 \sim 1, \quad \partial_v r \sim -\partial_u r \sim 1, \tag{8.6.2}$$

$$(1+u^2)|\partial_u \Omega^2| \lesssim 1, \quad v^2|\partial_v \Omega^2| \lesssim 1 \tag{8.6.3}$$

on $\mathcal{C}_{r_2} \cap \{v \ge \check{v}\}$, and the distribution function satisfies

$$\pi(\operatorname{spt} f) \cap \{v \ge \check{v}\} \subset \{6C_\nu u \le v\}.$$

We will make use of a bootstrap argument in the regions

$$\mathcal{R}_{\text{far}}^{\check{v},v_f} \doteq \{v \ge u\} \cap \{\tau \ge 0\} \cap \{u \ge -r_2\} \cap \{\check{v} \le v \le v_f\},$$

where $v_f \ge \check{v}$. Refer to Fig. 8.5. The bootstrap assumptions are

$$-C_\nu \le \partial_u r \le -C_\nu^{-1}, \tag{8.6.4}$$

$$\tfrac{1}{5} \le \partial_v r \le 1, \tag{8.6.5}$$

$$\pi(\operatorname{spt} f) \cap \mathcal{R}_{\text{far}}^{\check{v},v_f} \subset \mathcal{W}, \tag{8.6.6}$$

where

$$\mathcal{W} \doteq \{6C_\nu u \le v\} \cap \{v \ge \check{v}\}$$

and the constant $10 \le C_\nu \lesssim 1$ is chosen so that $-\frac{1}{2}C_\nu \le \partial_u r \le -2C_\nu^{-1}$ on $\underline{C}_{\check{v}}$. Such a constant exists by Lemmas 8.4.1 and 8.5.1.

Lemma 8.6.2. *If* (8.2.5) *holds, $v_f \ge \check{v}$, $\mathcal{R}_{\text{far}}^{\check{v},v_f} \subset \mathcal{U}$, and the bootstrap assumptions* (8.6.4)–(8.6.6) *hold on $\mathcal{R}_{\text{far}}^{\check{v},v_f}$, then*

$$0 \le m \le 10M, \quad 0 \le Q \le 6M, \tag{8.6.7}$$

$$\Omega^2 \sim 1, \tag{8.6.8}$$

$$\partial_v \log \Omega^2 \lesssim v^{-2}, \tag{8.6.9}$$

$$\tfrac{1}{2} \le 1 - \frac{2m}{r} \le 1 \tag{8.6.10}$$

on $\mathcal{R}_{\mathrm{far}}^{\check{v},v_f}$ *and*

$$r \sim v, \tag{8.6.11}$$

$$\partial_v \log \Omega^2 < \frac{2\partial_v r}{r} \tag{8.6.12}$$

on $\mathcal{W}$*. Furthermore,*

$$m = Q = 0 \tag{8.6.13}$$

on $\mathcal{R}_{\mathrm{far}}^{\check{v},v_f} \setminus \mathcal{W}$*.*

Proof. Proof of (8.6.7) and (8.6.13): The bootstrap assumptions (8.6.4) and (8.6.5) imply $\partial_u r < 0$ and $\partial_v r > 0$. Therefore, (8.6.7) follows from the monotonicity properties of the Einstein–Maxwell–Vlasov system, Lemmas 8.3.1, 8.4.1 and 8.5.1, and the boundary condition (3.2.42).

Proof of (8.6.11): By the bootstrap assumption (8.6.4),

$$r(u,v) - r(-r_2, v) \geq -C_\nu(u + r_2) \geq -\tfrac{1}{6}v - C_\nu r_2$$

for $(u,v) \in \mathcal{W}$. By (8.3.13), the lower bound in (8.6.11) easily follows if $\check{v}$ is taken sufficiently large. Since $\partial_u r < 0$, $r(u,v) \leq r(-r_2, v) \lesssim v$ for $v \geq \check{v}$ and $\check{v}$ sufficiently large, which proves the upper bound in (8.6.11).

Proof of (8.6.10): This is immediate for $\check{v}$ chosen sufficiently large in light of (8.6.7) and the fact that

$$\inf_{\mathcal{W}} r \gtrsim \check{v}, \tag{8.6.14}$$

which follows from (8.6.11).

Proof of (8.6.8): This follows from (2.1.3) by combining the bootstrap assumptions (8.6.4) and (8.6.5) with (8.6.10).

Proof of (8.6.9): Let $(u,v) \in \mathcal{W}$. We will show that

$$\int_{-r_2}^{u} T^{uv}(u', v)\, du' \lesssim v^{-2}, \tag{8.6.15}$$

which together with (2.3.21), (8.6.7), (8.6.8), and (8.6.11), readily implies (8.6.9). To prove (8.6.15), we observe that by the bootstrap assumptions (8.6.4) and (8.6.5) and the evolution equation (2.3.28),

$\partial_u m \leq 0$. Using that $m \geq 0$ in $\mathcal{R}_{\text{aux}}^{\check{v},v_f}$, we therefore infer

$$\int_{-r_2}^{u} \left[\tfrac{1}{2}r^2\Omega^2(T^{uv}(-\partial_u r) + T^{vv}\partial_v r) + \frac{Q^2}{2r^2}(-\partial_u r)\right]\Big|_{(u',v)} du' = m(-r_2,v) - m(u,v) \leq 2\varpi_2.$$

Since all three integrands are nonnegative, this energy estimate, taken together with (8.6.4), (8.6.8) and (8.6.11) imply

$$v^2 \int_{-r_2}^{u} T^{uv}(u',v)\, du' \lesssim \int_{-r_2}^{u} r^2\Omega^2 T^{uv}\partial_v r\big|_{(u',v)}\, du' \lesssim 1$$

for $(u,v) \in \mathcal{W}$, which proves (8.6.15).

Proof of (8.6.12): This follows from the fact that $\partial_v r/r \gtrsim v^{-1}$ in $\mathcal{W}$ by (8.6.5) and (8.6.11). □

Proof of Lemma 8.6.1. The proof is a bootstrap argument based on the bootstrap assumptions (8.6.4)–(8.6.6). Let

$$\mathcal{A} \doteq \{v_f \in [\check{v},\infty) : \mathcal{R}_{\text{aux}}^{\check{v},v_f} \subset \mathcal{U} \text{ and (8.6.4)–(8.6.6) hold on } \mathcal{R}_{\text{aux}}^{\check{v},v_0}\}.$$

The set $\mathcal{A}$ is nonempty by Proposition 3.2.12, Lemma 8.4.1, and Lemma 8.5.1. It is also manifestly closed by continuity of the bootstrap assumptions. We now show that if (8.2.5) holds, then $\mathcal{A}$ is also open. Let $v_f \in \mathcal{A}$.

Improving (8.6.5): Let $(u,v) \in \mathcal{R}_{\text{aux}}^{\check{v},v_f}$. Integrating the wave equation (2.3.20) in u starting at $u' = -r_2$ and using the estimates of Lemma 8.6.2 yields

$$|\partial_v r(u,v) - \tfrac{1}{2}| \leq \int_{-r_2}^{u} \left(\frac{\Omega^2}{2r^2}\left(m + \frac{Q^2}{2r}\right) + \tfrac{1}{4}r\Omega^4 T^{uv}\right)\Big|_{(u',v)} du' \lesssim v^{-1} \leq \check{v}^{-1},$$

which improves (8.6.5) for $\check{v}$ sufficiently large.

Improving (8.6.4): Using (2.3.29), $\partial_v m \geq 0$, and (8.6.14), we obtain (similarly to (8.6.15))

$$\int_{v_1}^{v_2} rT^{uv}|_{(u,v')}\, dv' \lesssim \check{v}^{-1} \tag{8.6.16}$$

for any $(u,v_1),(u,v_2) \in \mathcal{R}_{\text{aux}}^{\check{v},v_f}$. Let $(u,v) \in \mathcal{R}_{\text{aux}}^{\check{v},v_f}$. We integrate the wave equation (2.3.20) in v starting at $\underline{C}_{\check{v}}$ if $u \leq \check{v}$ and at $(u,u) \in \Gamma$ if $u > \check{v}$. In the former case,

$$|\partial_u r(u,v) - \partial_u r(u,\check{v})| \leq \int_{\check{v}}^{v} \left(\frac{\Omega^2}{2r^2}\left(m + \frac{Q^2}{2r}\right) + \tfrac{1}{4}r\Omega^2 T^{uv}\right)\Big|_{(u,v')} dv' \lesssim \check{v}^{-1},$$

which improves (8.6.4) for $\breve{v}$ sufficiently large by the definition of C_ν. In the latter case, the boundary condition (3.2.42) implies $\partial_u r(u,u) = -\partial_v r(u,u)$, so

$$|\partial_u r(u,v) + \partial_v r(u,u)| \le \int_u^v \left(\frac{\Omega^2}{2r^2}\left(m + \frac{Q^2}{2r}\right) + \tfrac{1}{4} r\Omega^2 T^{uv}\right)\bigg|_{(u,v')} dv' \lesssim \breve{v}^{-1},$$

which by (8.6.5) improves (8.6.4) for $\breve{v}$ sufficiently large by the definition of C_ν.

Improving (8.6.6): Let $\gamma : [0,S) \to \mathcal{R}_{\mathrm{aux}}^{\breve{v},v_f}$ be an electromagnetic geodesic in the support of f starting at $\underline{C}_{\breve{v}}$ at $s=0$. By (2.1.24) and (8.6.12),

$$\frac{d}{ds}(\Omega^2 p^u) \le 0,$$

so by (8.6.8),

$$p^u(s) \lesssim p^u(0). \tag{8.6.17}$$

Using (2.1.22), the signs of $\partial_u r$ and Q, and parametrizing γ by v yields

$$\frac{d}{dv}\log p^v \ge -\partial_v \log\Omega^2. \tag{8.6.18}$$

By (8.6.9), it is easy to see that

$$\int_{\breve{v}}^v |\partial_v \log\Omega^2(\gamma^u(v'),v')|\, dv' \lesssim \breve{v}^{-1}$$

and therefore

$$\exp\left(-\int_{\breve{v}}^v \partial_v \log\Omega^2(\gamma^u(v'),v')\, dv'\right) \ge \frac{1}{2}$$

for $\breve{v}$ sufficiently large. It follows from (8.6.18) that

$$p^v(s) \gtrsim p^v(0). \tag{8.6.19}$$

Combining (8.6.17) and (8.6.19) yields

$$\frac{p^u}{p^v}(s) \le \frac{1}{7C_\nu},$$

for $\breve{v}$ sufficiently large by (8.4.8) and (8.5.3). It is then easy to show that $\gamma(s)$ stays in $\{7C_\nu u \le v\}$ for every $s \in [0,S)$, which quantitatively improves (8.6.6). The rest of the existence and uniqueness proof now follows a standard continuity argument using Proposition 3.2.6 and Lemma 3.2.16.

To estimate $|\partial_u\Omega^2|$, note that $|\partial_u\Omega^2| \lesssim (1+u^2)^{-1}$ along $\underline{C}_{\check{v}} \cup \Gamma$ by Lemma 8.5.1 and (3.2.44). Observe that

$$\inf_{C_u\cap\mathcal{W}\cap\mathcal{R}^{\check{v},\infty}_{\mathrm{far}}} r \sim \inf_{C_u\cap\mathcal{W}\cap\mathcal{R}^{\check{v},\infty}_{\mathrm{far}}} v = \max\{\check{v}, 6C_\nu u\} \gtrsim 1+u,$$

so the energy estimate (8.6.16) can be improved to

$$\int_{v_1}^{v_2} T^{uv}(u,v')\,dv' \lesssim (1+u)^{-2}$$

for any $u \le v_1 \le v_2$. Therefore the desired estimate can be propagated to the interior by integrating the wave equation (2.3.21) in v. Together with the bootstrap assumptions and Lemma 8.6.2, this completes the proof of the estimates (8.6.1)–(8.6.3). □

8.7 The dispersive estimate in the massive case

Let $\mathfrak{m} > 0$ and consider the solution (r, Ω^2, Q, f) given by Lemma 8.6.1, defined globally on $\mathcal{C}_{r_2}$. We augment the hierarchy (8.2.5) with a large parameter $v_\#$ satisfying

$$0 < v_\#^{-1} \ll \mathfrak{m} \le \mathfrak{m}_0 \tag{8.7.1}$$

and aim to prove the following

Lemma 8.7.1. *For any $\check{v}, \eta, \varepsilon, \mathfrak{m}_0, \mathfrak{m}$, and $v_\#$ satisfying* (8.2.5) *and* (8.7.1)*, we have the decay*

$$\mathcal{M} \le Cv^{-3}$$

for $v \ge v_\#$ and any $\mathcal{M} \in \{N^u, N^v, T^{uu}, T^{uv}, T^{vv}, S\}$, where C may depend on η, ε, $\mathfrak{m}$, and $v_\#$.

The proof is based on a bootstrap argument for the dispersion of ingoing momentum p^u along $\operatorname{spt} f$ as $v \to \infty$, which leads to cubic decay of the phase space volume $\mathscr{V}$, which was defined in (8.1.8). Using the mass shell relation (2.3.12) and the change of variables formula, we have

$$\mathscr{V}(u,v) = \frac{2}{r^2}\int_0^\infty \int_{\{p^u : f(u,v,p^u,p^v)\neq 0\}} \frac{dp^u}{p^u}\,\ell\,d\ell, \tag{8.7.2}$$

where we view p^v as a function of p^u and ℓ. Compare with (8.1.9).

Lemma 8.7.2. *If* (8.2.5) *and* (8.7.1) *hold and* $(u, v, p^u, p^v) \in \operatorname{spt} f$*, then* $p^v \lesssim 1$ *if* $v \ge \check{v}$ *and* $p^u \sim_\eta \mathfrak{m}^2$ *if* $v \ge v_\#$*.*

Remark 8.7.3. The estimate $p^v \lesssim 1$ also holds in the massless case. The non-decay of p^u for massive particles drives the decay rate v^{-3}, but only at very late times.

Proof. Let $\gamma \in \Gamma_f$ and $v \geq \check{v}$. Parametrizing γ by v and using (8.4.8), (8.4.25), (8.5.3), and the estimates in Lemma 8.6.1, we infer

$$\left| \frac{d}{dv}(\Omega^2 p^v) \right| \lesssim v^{-2},$$

which is integrable and hence shows that $p^v \lesssim 1$. By (8.4.8) and (8.5.3), we have $p^v \gtrsim_\eta 1$. We then obtain

$$p^u = \frac{\ell^2 r^{-2} + \mathfrak{m}^2}{\Omega^2 p^v} \sim_\eta \frac{\ell^2}{r^2} + \mathfrak{m}^2 \sim \mathfrak{m}^2$$

for $r \gtrsim v_\#$ sufficiently large. □

Using this lemma and (8.4.24), we immediately infer:

Lemma 8.7.4. *If* (8.2.5) *and* (8.7.1) *hold,* $(u, v, p^u, p^v) \in \operatorname{spt} f$*, and* $v \geq v_\#$*, then* $u \sim_\eta \mathfrak{m}^2 v$*. If* $\gamma_1, \gamma_2 \in \Gamma_f(u, v)$*, then we have* $|\gamma_1^u(s^1_{v_\#}) - \gamma_2^u(s^2_{v_\#})|(s_{v_\#}) \lesssim_\eta \mathfrak{m}^2(v - v_\#)$*, where* $s^i_{v_\#}$ *is the parameter time for which* $\gamma_i^v(s^i_\#) = v_\#$*.*

Let $(u_0, v_0) \in \mathcal{C}_{r_2}$ and let $\gamma \in \Gamma_f(u_0, v_0)$ have ingoing momentum p_0^u and angular momentum ℓ at (u_0, v_0). We parametrize γ by v going backwards in time and denote this by

$$\gamma^u(v) \doteq \gamma^u(v; u_0, v_0, p_0^u, \ell),$$
$$p^u(v) \doteq p^u(v; u_0, v_0, p_0^u, \ell).$$

We readily derive the equations

$$\frac{d}{dv}\gamma^u = \frac{(\Omega^2 p^u)^2}{\Omega^2(\ell^2 r^{-2} + \mathfrak{m}^2)}, \tag{8.7.3}$$
$$\frac{d}{dv}(\Omega^2 p^u) = \left(\partial_v \log \Omega^2 - \frac{2\partial_v r}{r}\right)\frac{\ell^2 \Omega^2 p^u}{r^2\mathfrak{m}^2 + \ell^2} - \mathfrak{e}\frac{Q}{r^2\mathfrak{m}^2 + \ell^2}(\Omega^2 p^u)^2. \tag{8.7.4}$$

Next, we define the variational quantities

$$\mathfrak{u}(v; u_0, v_0, p_0^u, \ell) \doteq \frac{\partial}{\partial p_0^u}\gamma^u(v; u_0, v_0, p_0^u, \ell), \quad \mathfrak{p}(v; u_0, v_0, p_0^u, \ell) \doteq \frac{\partial}{\partial p_0^u}(\Omega^2 p^u)(v; u_0, v_0, p_0^u, \ell),$$

where we emphasize that the derivative in p_0^u is taken with ℓ fixed. From (8.7.3) and (8.7.4) we obtain

$$\frac{d}{dv}\mathfrak{u} = \frac{2p^u}{\ell^2 r^{-2} + \mathfrak{m}^2}\mathfrak{p} - \frac{(p^u)^2}{(\ell^2 r^{-2} + \mathfrak{m}^2)^2}\left[\partial_u\Omega^2\left(\frac{\ell^2}{r^2} + \mathfrak{m}^2\right) - \frac{2\Omega^2\ell^2\partial_u r}{r^3}\right]\mathfrak{u}, \tag{8.7.5}$$

$$\begin{aligned}\frac{d}{dv}\mathfrak{p} = &\left[\left(\partial_v \log\Omega^2 - \frac{2\partial_v r}{r}\right)\frac{\ell^2}{r^2\mathfrak{m}^2 + \ell^2} - \frac{2\mathfrak{e}Q\Omega^2 p^u}{r^2\mathfrak{m}^2 + \ell^2}\right]\mathfrak{p}\\ &+ \left[\left(\partial_u\partial_v \log\Omega^2 - \frac{2\partial_u\partial_v r}{r} + 2\frac{\partial_v r\partial_u r}{r^2}\right)\frac{\ell^2\Omega^2 p^u}{r^2\mathfrak{m}^2 + \ell^2} - \left(\partial_v \log\Omega^2 - \frac{2\partial_v r}{r}\right)\frac{2\mathfrak{m}^2 r\partial_u r\ell^2\Omega^2 p^u}{(r^2\mathfrak{m}^2 + \ell^2)^2}\right.\\ &\left.-\frac{\mathfrak{e}(\Omega^2 p^u)^2}{r^2\mathfrak{m}^2 + \ell^2}\partial_u Q + \frac{2\mathfrak{e}Q\mathfrak{m}^2 r\partial_u r(\Omega^2 p^u)^2}{(r^2\mathfrak{m}^2 + \ell^2)^2}\right]\mathfrak{u}. \end{aligned} \tag{8.7.6}$$

Note that $\mathfrak{u}(v_0) = 0$ and $\mathfrak{p}(v_0) = \Omega^2(u_0, v_0) \sim 1$.

Lemma 8.7.5. *If* (8.2.5) *and* (8.7.1) *hold, and* $v \geq v_\#$*, then*

$$\mathscr{V}(u, v) \lesssim_{\eta,\varepsilon} \frac{v_\#}{v^3}. \tag{8.7.7}$$

Proof. We claim that there exists a constant C_*, depending on η and ε, such that

$$C_*^{-1} \leq \frac{\mathfrak{u}(v)}{v - v_0} \leq C_* \tag{8.7.8}$$

for any $v_\# \leq v \leq v_0$. To see how this proves (8.7.7), let $\Phi_{u_0,v_0,\ell}(p_0^u) \doteq \gamma^u(v_\#; u_0, v_0, p_0^u, \ell)$ and observe that $\Phi'_{u_0,v_0,\ell}(p_0^u) = \mathfrak{u}(v_\#) < 0$. Changing variables in the p^u integral in (8.7.2) to $\gamma^u(v_\#)$ and using Lemma 8.7.4 to estimate the u-dispersion along $\underline{C}_{v_\#}$, we find

$$\mathscr{V}(u_0, v_0) \lesssim_{\eta,\varepsilon} \frac{1}{r^2\mathfrak{m}^2}\frac{C_*}{|v_\# - v_0|}\min\{\mathfrak{m}^2 v_\#, \mathfrak{m}^2(v_0 - v_\#)\} \lesssim C_*\frac{v_\#}{v_0^3}. \tag{8.7.9}$$

We prove (8.7.8) by a bootstrap argument as follows. Let $v_f \geq v_\#$ and assume (8.7.8) holds for all $(v, u_0, v_0, p_0^u, \ell)$ with $v_\# \leq v \leq v_0 \leq v_f$. The assumption is clearly satisfied for some choice of C_* for v_f sufficiently close to $v_\#$ on account of $\mathfrak{u}(v_0) = 0$.

We now show that for $\mathfrak{m}$ sufficiently small and $v_\#$ sufficiently large, we can improve the constant in (8.7.8). Using (8.7.7), we estimate

$$N^v \lesssim_{\eta,\varepsilon} C_*\frac{v_\#}{v^3}, \quad T^{uv} \lesssim_{\eta,\varepsilon} C_*\frac{\mathfrak{m}^2 v_\#}{v^3}$$

for $v \in [v_\#, v_f]$. Using this, as well as (2.3.21), (2.3.24), Lemma 8.6.1, and $p^u \sim_\eta \mathfrak{m}^2$ in (8.7.6) yields

$$\left|\frac{d}{dv}\mathfrak{p}\right| \lesssim_{\eta,\varepsilon} \left(\frac{1}{\mathfrak{m}^2 v^3} + \frac{1}{v^2}\right)|\mathfrak{p}| + \left(\frac{1}{v^4} + \frac{\mathfrak{m}^2}{v^3} + C_*\frac{\mathfrak{m}^2 v_\#}{v^3}\right)|\mathfrak{u}|.$$

We use the bootstrap assumption (8.7.8) and (8.7.1) to infer

$$\left|\frac{d}{dv}(\mathfrak{p}-\mathfrak{p}(v_0))\right| \lesssim_{\eta,\varepsilon} \frac{1}{v^2}|\mathfrak{p}-\mathfrak{p}(v_0)| + \frac{1}{v^2} + C_*^2\frac{\mathfrak{m}^2 v_\#}{v^3}|v-v_0|.$$

and then use Grönwall's inequality to obtain

$$|\mathfrak{p}(v)-\mathfrak{p}(v_0)| \lesssim_{\eta,\varepsilon} v_\#^{-1} + C_*^2\frac{\mathfrak{m}^2 v_\#}{v^2}|v-v_0|.$$

Using $\mathfrak{p}(v_0)\sim 1$, we therefore have

$$\int_v^{v_0}\mathfrak{p}(v')\,dv' \sim_{\eta,\varepsilon} v_0 - v + O_{\eta,\varepsilon}(C_*^2\mathfrak{m}^2|v-v_0|) \sim v_0 - v \tag{8.7.10}$$

for $\mathfrak{m}$ sufficiently small. Finally, we use (8.6.3) and Lemma 8.7.4 to estimate

$$\int_v^{v_0}|\partial_u\Omega^2|(\gamma^u(v'),v')\left(\frac{\ell^2}{r^2}+\mathfrak{m}^2\right)dv' \lesssim_\eta \int_{\gamma^u(v)}^{\gamma^u(v_0)}\frac{1}{1+u'^2}\,du' \lesssim \frac{1}{\gamma^u(v_\#)} \lesssim \frac{1}{\mathfrak{m}^2 v_\#}. \tag{8.7.11}$$

Integrating (8.7.5) and using (8.7.10) and (8.7.11) improves the constant in (8.7.8) for $\mathfrak{m}$ sufficiently small and $v_\#$ sufficiently large, which completes the proof. □

Proof of Lemma 8.7.1. Immediate from (8.7.7). □

8.8 Proof of Proposition 8.2.3

Proof. Part 1. This follows immediately from combining Lemmas 8.4.1, 8.5.1 and 8.6.1.

Part 2. The estimates in (8.2.6) follow directly from the estimates in Lemmas 8.4.1, 8.5.1 and 8.6.1. We will now prove causal geodesic completeness of the spacetime. Let γ be a future-directed causal geodesic in the $(3+1)$-dimensional spacetime. Then the projection of γ to the reduced spacetime, still denoted γ, satisfies

$$\frac{d}{ds}(\Omega^2 p^u) = \left(\partial_v \log\Omega^2 - \frac{2\partial_v r}{r}\right)\frac{\ell^2}{r^2}, \tag{8.8.1}$$

$$\frac{d}{ds}(\Omega^2 p^v) = \left(\partial_u \log\Omega^2 - \frac{2\partial_u r}{r}\right)\frac{\ell^2}{r^2}, \tag{8.8.2}$$

$$\Omega^2 p^u p^v = \frac{\ell^2}{r^2} + \mathfrak{m}^2, \tag{8.8.3}$$

where s is an affine parameter and γ is continued through the center according to Part 4 of Definition 3.2.10. We will show for any future-directed causal geodesic $\gamma : [0,S) \to \mathcal{C}_{r_2}$, p^τ is uniformly bounded in any compact interval of coordinate time $0 \le \tau \le \tau_0$ along γ. This implies that γ can be extended to $[0,\infty)$ by the normal neighborhood property of the geodesic flow in Lorentzian manifolds [ONe83].

By (8.8.1)–(8.8.3) and (8.2.6),

$$\left|\frac{d}{ds}(\Omega^2 p^\tau)\right| \lesssim \left(1 + \left|\frac{\partial_v r + \partial_u r}{r}\right|\right)(\Omega^2 p^\tau)^2. \tag{8.8.4}$$

When $r \ge \frac{1}{6} r_1$, the term in the absolute value on the right-hand side of this estimate is clearly bounded by (8.2.6). When $r \le \frac{1}{6} r_1$, the spacetime is Minkowski and the formulas in Lemma 3.2.16 can be used to show that

$$\partial_v r + \partial_u r \lesssim v - u \lesssim r. \tag{8.8.5}$$

Therefore, passing to a τ parametrization of γ, (8.8.4) implies

$$\left|\frac{d}{d\tau}(\Omega^2 p^\tau)\right| \lesssim \Omega^2 p^\tau,$$

and the proof is completed by an application of Grönwall's lemma.[4]

<u>Part 3.</u> The estimate (8.2.7) for the final parameters follows from (8.3.6). The remaining claims in this part follow from the proof of Lemma 8.3.1, (8.4.18) on $C_{-\frac{2}{3}r_1}$, and Lemma 8.6.1.

<u>Part 4.</u> The estimate (8.2.8), the upper bound in (8.2.9), and the claim about the neighborhood of the center follow from Lemma 8.6.1. The lower bound in (8.2.9) follows from Lemma 8.7.2, which implies that γ^u grows linearly in v at very late times for any electromagnetic geodesic γ in the support of f. Since a neighborhood of $\mathcal{I}^+$ is then electrovacuum, it is isometric to an appropriate Reissner–Nordström solution by Birkhoff's theorem.

<u>Part 5.</u> This follows immediately from Lemma 8.7.1.

<u>Part 4′.</u> We now take $\mathfrak{m} = 0$ and seek to prove the upper bound in (8.2.11), the rest following immediately from Birkhoff's theorem. Let $\gamma(s)$ be an electromagnetic geodesic lying in the support of f with $\gamma^v(0) = \check{v}$. When $\mathfrak{m} = 0$, the mass shell relation, together with (8.4.8), (8.5.3), and (8.6.19) gives the estimate

$$\frac{p^u}{p^v}(s) = \frac{r^2(0)}{r^2(s)}\frac{\Omega^2(0)}{\Omega^2(s)}\frac{p^u(0)}{p^v(s)}\frac{p^v(0)}{p^v(s)} \lesssim v^{-2},$$

[4] We have based this argument off of a non-gauge-invariant energy $\Omega^2 p^\tau$, which is why the estimate (8.8.5) has to be performed even in the Minkowski region of the spacetime. One could alternatively consider the gauge-invariant energy $(\partial_v r)p^u - (\partial_u r)p^v$, which is constant in the Minkowski region, but satisfies a more complicated evolution equation where $f \neq 0$.

which has integral $\lesssim \breve{v}^{-1}$. Therefore, the upper bound in (8.2.11) follows from (8.4.24) and taking $\breve{v}$ sufficiently large.

Part 5′. When $\mathfrak{m} = 0$, the mass shell relation implies $p^u \lesssim v^{-2}$ for any electromagnetic geodesic lying in the support of f. The estimates (8.2.12)–(8.2.14) follow from this and the formula (8.1.9). □

8.9 Patching together the ingoing and outgoing beams

8.9.1 The maximal time-symmetric doubled spacetime

Let $\alpha \in \mathcal{P}_\Gamma$ or be of the form $(r_1, r_2, 0, 0, 0, 0)$. Let η and ε be beam parameters for which the conclusion of Proposition 8.2.3 holds, recalling also Remark 8.2.4. Let $\tilde{M}$ and $\tilde{e}$ denote the final Reissner–Nordström parameters of $\mathcal{S} = \mathcal{S}_{\alpha,\eta,\varepsilon}$. We say that $\mathcal{S}$ is *subextremal* if $\tilde{e} < \tilde{M}$, *extremal* if $\tilde{e} = \tilde{M}$, and *superextremal* if $\tilde{e} > \tilde{M}$.[5] If $\mathcal{S}$ is not superextremal, we may define

$$r_\pm = \tilde{M} \pm \sqrt{\tilde{M}^2 - \tilde{e}^2},$$

which is the formula for the area radii of the outer and inner horizons in Reissner–Nordström.

The following lemma is an easy consequence of the well-known structure of the maximally extended Reissner–Nordström solution:

Lemma 8.9.1. *For any $\tilde{M}, \tilde{e} > 0$ and $0 < r_2 < r_-$ (if $\tilde{e} \le \tilde{M}$), there exists a relatively open set*

$$\mathcal{E}_{\tilde{M},\tilde{e},r_2} \subset \{u \le -r_2\} \cap \{v \ge r_2\} \subset \mathbb{R}^2_{u,v}$$

and an analytic spherically symmetric solution (r, Ω^2, Q) of the Einstein–Maxwell equations on $\mathcal{E}_{\tilde{M},\tilde{e},r_2}$ with the following properties: The renormalized Hawking mass $\varpi = \tilde{M}$ globally, the charge $Q = \tilde{e}$ globally, $r(-r_2, r_2) = r_2$, $\partial_v r(-r_2, r_2) = -\partial_u r(-r_2, r_2) = \frac{1}{2}$, and Ω^2 is constant along the cones $\{u = -r_2\} \cap \{v \ge r_2\}$ and $\{v = r_2\} \cap \{u \le -r_2\}$. Moreover, we may choose $(\mathcal{E}_{\tilde{M},\tilde{e},r_2}, r, \Omega^2, Q)$ to be maximal with these properties, and it will then be unique.

Remark 8.9.2. The (3 + 1)-dimensional lift of $(\mathcal{E}_{\tilde{M},\tilde{e},r_2}, r, \Omega^2, Q)$ is isometric to a subset of the maximally extended Reissner–Nordström solution with parameters $\tilde{M}$ and $\tilde{e}$. The hypersurface $(\{\tau = 0\} \cap \mathcal{E}_{\tilde{M},\tilde{e},r_2}) \times S^2$ is then totally geodesic.

Remark 8.9.3. If $\tilde{M} < \tilde{e}$, then $\mathcal{E}_{\tilde{M},\tilde{e},r_2} = \{u \le -r_2\} \cap \{v \ge r_2\}$. In the case $\tilde{e} \le \tilde{M}$, $\mathcal{E}_{\tilde{M},\tilde{e},r_2}$ is a strict subset of $\{u \le -r_2\} \cap \{v \ge r_2\}$ with this choice of gauge.

[5] Recall that we are always taking $\mathfrak{e} > 0$ and hence $\tilde{e} > 0$.

Definition 8.9.4. Let $(\mathcal{C}_{r_2}, r, \Omega^2, Q, f)$ be the maximal normalized development of $\mathcal{S} = \mathcal{S}_{\alpha,\eta,\varepsilon}$ for particles of mass $\mathfrak{m}$ with final Reissner–Nordström parameters $(\tilde{M}, \tilde{e})$ for which the conclusion of Proposition 8.2.3 holds, recalling also Remark 8.2.4. Let

$$\tilde{\mathcal{C}}_{r_2} \doteq \{\tau \leq 0\} \cap \{v \geq u\} \cap \{v \leq r_2\}$$

be the time-reflection of $\mathcal{C}_{r_2}$. For $(u,v) \in \tilde{\mathcal{C}}_{r_2}$, define

$$\begin{aligned}
\tilde{r}(u,v) &= r(-v,-u),\\
\tilde{\Omega}^2(u,v) &= \Omega^2(-v,-u),\\
\tilde{Q}(u,v) &= Q(-v,-u),\\
\tilde{f}(u,v,p^u,p^v) &= f(-v,-u,p^v,p^u).
\end{aligned}$$

Let $\mathscr{M} \doteq \mathcal{C}_{r_2} \cup \tilde{\mathcal{C}}_{r_2} \cup \mathcal{E}_{\tilde{M},\tilde{e},r_2}$ and define (r, Ω^2, Q, f) on $\mathscr{M}$ by simply gluing the corresponding functions across the boundaries of the sets $\mathcal{C}_{r_2}$, $\tilde{\mathcal{C}}_{r_2}$, and $\mathcal{E}_{\tilde{M},\tilde{e},r_2}$. (We therefore now drop the tilde notation on the solution in $\mathcal{C}_{r_2}$, except for in the proof of Lemma 8.9.5 below.) The tuple $(\mathscr{M}, r, \Omega^2, Q, f)$ is called the *maximal time-symmetric doubled spacetime* associated to $\mathcal{S}$.

Lemma 8.9.5. *$(\mathscr{M}, r, \Omega^2, Q, f)$ is a smooth solution of the spherically symmetric Einstein–Maxwell–Vlasov system. The hypersurface $\{\tau = 0\} \cap \mathscr{M}$ is a totally geodesic hypersurface once lifted to the $(3+1)$-dimensional spacetime by Proposition 2.3.10.*

Proof. This is immediate except perhaps across $\{\tau = 0\} \cap \{v \leq r_2\}$. We only have to show that the solution is $C^2 \times C^2 \times C^1 \times C^1$ regular across this interface by the regularity theory of Proposition 3.2.3. By Definition 3.2.7 and Definition 3.2.10, the solution is clearly $C^1 \times C^1 \times C^0 \times C^0$ regular. Using Raychaudhuri's equations (2.3.22), (2.3.23) and Maxwell's equations (2.3.24), (2.3.25), it is easy to see that r is C^2 and Q is C^1 across $\{\tau = 0\}$. To check second derivatives of Ω^2, differentiate $(\partial_u + \partial_v)\Omega^2(-v,v) = 0$ in v to obtain $\partial_u^2\Omega^2(-v,v) = \partial_v^2\Omega^2(-v,v)$, which together with the wave equation (2.3.21) implies C^2 matching. The p^u and p^v derivatives of f are also continuous by inspection and continuity of spatial derivatives can be proved as follows: On $\{\tau = 0\}$, $\partial_v f$ can be eliminated in terms of $\partial_u f$ and $\frac{d}{dv}\mathring{f}$ by (3.2.37). Then the Maxwell–Vlasov equation (2.3.26) can be solved for $\partial_u f$. Performing the same calculation for $\tilde{f}$ shows that $\partial_u f$ is continuous across $\{\tau = 0\}$. The same argument applies for $\partial_v f$ and the proof is complete. $\square$

8.9.2 The anchored Cauchy hypersurface

In order to define the family of Cauchy data in Theorem 1.2.1, we need to identify an appropriate Cauchy hypersurface in each $\mathscr{M}$. Let

$$v_\infty \doteq \begin{cases} 2r_- - r_2 & \text{if } \tilde{e} \le \tilde{M} \\ \infty & \text{if } \tilde{e} > \tilde{M} \end{cases} \tag{8.9.1}$$

and let $\Sigma^u : [-50\tilde{M}, v_\infty) \to \mathbb{R}$ be the smooth function defined by

$$\Sigma^u(v) = -100\tilde{M} - v$$

for $v \in [-50\tilde{M}, r_1]$, by solving the ODE

$$\frac{d}{dv}\Sigma^u(v) = \frac{\partial_v r}{\partial_u r}\bigg|_{(\Sigma^u(v),v)} \tag{8.9.2}$$

with initial condition $\Sigma^u(r_2) = -100\tilde{M} - r_2$ for $v \in [r_2, v_\infty)$, and by applying the following easy consequence of Borel's lemma for $v \in [r_1, r_2]$:

Lemma 8.9.6. *Given $0 < r_1 < r_2$, $\tilde{M} > 0$, and a sequence of real numbers $a_1, a_2, \dots$ with $a_1 < 0$, there exists a smooth function $\Sigma^u : [r_1, r_2] \to [-100\tilde{M} - r_2, -100\tilde{M} - r_1]$ such that $\frac{d}{dv}\Sigma^u(v) < 0$ for $v \in [r_1, r_2]$, Σ^u has Taylor coefficients $(-100\tilde{M} - r_1, -1, 0, 0, \dots)$ at r_1, and Taylor coefficients $(-100\tilde{M} - r_2, a_1, a_2, \dots)$ at r_2. Moreover, if $\tilde{M}$ and each a_j are smooth functions of the parameters $\varpi_2, Q_2, \varepsilon, \eta$, then f can be chosen to depend smoothly on $\varpi_2, Q_2, \varepsilon, \eta$.*

Remark 8.9.7. For $v \ge r_2$, the curve $\Sigma : v \mapsto (\Sigma^u(v), v) \in \mathcal{E}_{\tilde{M},\tilde{e},r_2}$ lies in the domain of outer communication if $\tilde{e} \le \tilde{M}$ and is contained in a constant time hypersurface in Schwarzschild coordinates. Indeed, the time-translation Killing vector field in $\mathcal{E}_{\tilde{M},\tilde{e},r_2}$ is given by the Kodama vector field

$$K \doteq 2\Omega^{-2}\partial_v r\, \partial_u - 2\Omega^{-2}\partial_u r\, \partial_v,$$

which is clearly orthogonal to Σ.

8.9.3 Cauchy data for the Einstein–Maxwell–Vlasov system

Let $(\mathcal{M}, g, F, f)$ be a solution of the $(3+1)$-dimensional Einstein–Maxwell–Vlasov system as defined in Section 2.3.1. Let $i : \mathbb{R}^3 \to \tilde{\Sigma} \subset \mathcal{M}$ be a spacelike embedding with future-directed unit timelike normal n to $\tilde{\Sigma}$. As usual, we may consider the induced metric $\bar{g} \doteq i^*g$ and second fundamental form k (pulled back to $\mathbb{R}^3$ along i) of $\tilde{\Sigma}$. The electric field is defined by $\bar{E} \doteq i^*F(\cdot, n)$ and the magnetic field is defined by $\bar{B} \doteq (\star_{\bar{g}} i^*F)^\sharp$, where $\sharp$ is taken relative to $\bar{g}$. Since the domain of f is the spacetime mass shell $P^{\mathfrak{m}}$ which is not intrinsic to $\tilde{\Sigma}$, one first has to define a projection procedure to $T\tilde{\Sigma} \cong T\mathbb{R}^3$, after which the restriction of f to $\tilde{\Sigma}$ can be thought of as a function $\bar{f} : T\mathbb{R}^3 \to [0, \infty)$. Similarly, the volume forms in the Vlasov energy momentum tensor T and number current N have to be written in terms of $\bar{g}$, after which $\bar{\rho}_T \doteq i^*T(n, n)$, $\bar{j}_T \doteq i^*T(n, \cdot)$ and $\bar{\rho}_N \doteq i^*N(n)$ can be evaluated on $\tilde{\Sigma}$ only in terms of $\bar{g}$ and $\bar{f}$. For details of this procedure, we refer to [Rin13, Section 13.4].

Definition 8.9.8. A *Cauchy data set* for the Einstein–Maxwell–Vlasov system for particles of mass $\mathfrak{m}$ and fundamental charge $\mathfrak{e}$ consists of the tuple $\Psi = (\bar{g}, \bar{k}, \bar{E}, \bar{B}, \bar{f})$ on $\mathbb{R}^3$ satisfying the constraint equations[6]

$$
\begin{aligned}
R_{\bar{g}} - |\bar{k}|^2_{\bar{g}} + (\operatorname{tr}_{\bar{g}} \bar{k})^2 &= |\bar{E}|^2_{\bar{g}} + |\bar{B}|^2_{\bar{g}} + 2\bar{\rho}_T[\bar{f}], \\
\operatorname{div}_{\bar{g}} \bar{k} - d \operatorname{tr}_{\bar{g}} \bar{k} &= -(\star_{\bar{g}}(\bar{E}^\flat \wedge \bar{B}^\flat))^\sharp + \bar{j}_T[\bar{f}] \\
\operatorname{div}_{\bar{g}} \bar{E} &= \mathfrak{e}\bar{\rho}_N[\bar{f}], \\
\operatorname{div}_{\bar{g}} \bar{B} &= 0.
\end{aligned}
$$

We denote by $\mathfrak{M}_\infty(\mathbb{R}^3, \mathfrak{m}, \mathfrak{e})$ the set of solutions Ψ of the Einstein–Maxwell–Vlasov constraint system on $\mathbb{R}^3$ with the C^∞_{loc} subspace topology.

8.9.4 The globally hyperbolic region

By Proposition 8.2.3, Remark 8.2.4, their time-reversed ($u \mapsto -v$, $v \mapsto -u$) versions, Remark 8.9.7, and the structure of the Reissner–Nordström family, we have:

Proposition 8.9.9. *Let $\mathcal{S}$ and $(\mathscr{M}, r, \Omega^2, Q, f)$ be as in Definition 8.9.4 and let $\Sigma^u(v)$ be the function defined in Section 8.9.2. Let*

$$\mathcal{X} \doteq \{v \geq u\} \cap \{v < v_\infty\} \subset \mathscr{M}$$

and let $\Sigma = \{(\Sigma^u(v), v) : v \in [-50\tilde{M}, v_\infty)\}$. Then the following holds:

1. *The manifold $\mathcal{M} \doteq ((\mathcal{X}\setminus\Gamma)\times S^2)\cup\Gamma$ with metric $g = -\Omega^2\, dudv + r^2\gamma$, electromagnetic field, and Vlasov field lifted according to Proposition 2.3.10 is a globally hyperbolic, asymptotically flat*

[6] The particle mass $\mathfrak{m}$ is implicitly contained in the formulas for $\bar{T}$ and $\bar{N}$.

spacetime, free of antitrapped surfaces, with Cauchy hypersurface $\tilde{\Sigma} \doteq ((\Sigma \setminus \Gamma) \times S^2) \cup (\Sigma \cap \Gamma)$.

2. $(\mathcal{M}, g)$ *possesses complete null infinities* $\mathcal{I}^\pm$ *and is past causally geodesically complete.*

3. *If* $\tilde{e} > \tilde{M}$*, then* $(\mathcal{M}, g)$ *is future causally geodesically complete.*

4. *If* $\tilde{e} \le \tilde{M}$*, there are two options. (If* $\mathcal{S}_{\alpha,\eta,\varepsilon}$ *is to be untrapped, then necessarily* $r_2 \notin [r_-, r_+]$*.)*

   (a) *If* $r_2 < r_-$*, then* $(\mathcal{M}, g)$ *is future causally geodesically incomplete. The spacetime contains a nonempty black hole region, i.e.,* $\mathcal{BH} \doteq \mathcal{M} \setminus J^-(\mathcal{I}^+) \neq \emptyset$*. The Cauchy hypersurface* $\tilde{\Sigma}$ *is disjoint from* $\mathcal{BH}$*. The event horizon* $\mathcal{H}^+ = \partial(\mathcal{BH})$ *is located at* $u = r_2 - 2r_+$*.*

   (b) *If* $r_2 > r_+$*, then* $(\mathcal{M}, g)$ *is future causally geodesically complete.*

8.9.5 Proof of the main theorem

Proof of Theorem 1.2.1. Fix a fundamental charge $\mathfrak{e} > 0$, cutoff functions φ, θ, and ζ as in Section 8.2.1, and a number $\Lambda \ge 1$ satisfying (8.2.1). Fix the extremal black hole target mass $M > 0$ and let $0 < r_1 < r_2$ and $\delta > 0$ be as in Theorem 7.3.2. Let $\eta_0 > 0$ be such that Proposition 8.2.3 applies to the multiparameter family of seed data $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$ (which was defined in Definition 8.2.2) for $\lambda \in [-1, 2]$, $|M' - M| \le \delta$, $0 < \eta \le \eta_0$, $\varepsilon > 0$ sufficiently small depending on η, and particle mass $0 \le \mathfrak{m} \le \mathfrak{m}_0$, where $\mathfrak{m}_0$ is sufficiently small depending on η and ε.

Let $\mathcal{F} : [0, \infty)^2 \to [0, \infty)^2$ be defined by $\mathcal{F}(\lambda, M') = (\lambda^2 M', \lambda M')$, which is easily verified to be smoothly invertible on $(0, \infty)^2$. Define the function $F_{\eta,\varepsilon}(\lambda, M') = (\tilde{M}, \tilde{e})$, the final Reissner–Nordström parameters of $\mathcal{S}_{\lambda,M',\eta,\varepsilon}$. By (8.2.7), we find

$$|F_{\eta,\varepsilon}(\lambda, M') - \mathcal{F}(\zeta(\lambda), M')| \lesssim \eta \tag{8.9.3}$$

for $\lambda \in [-1, 2]$, $|M' - M| \le \delta$, and $0 < \eta \le \eta_0$. There exists a constant $0 < \lambda_0 \ll 1$ depending on η_0 such that if $\lambda \in [-1, \lambda_0]$, then $|F_{\eta,\varepsilon}(\lambda, M')| < \frac{1}{2} r_1$. Also, (8.9.3) implies

$$\sup_{\lambda_0 \le \lambda \le 2, |M' - M| \le \delta} \left| \frac{\tilde{e}}{\tilde{M}} - \frac{1}{\zeta(\lambda)} \right| \lesssim_{\lambda_0} \eta. \tag{8.9.4}$$

From smooth convergence of $F_{\eta,\varepsilon}$ to $\mathcal{F}$ as $\eta \to 0$, we obtain

$$\sup_{\lambda_0 \le \lambda \le 2, |M' - M| \le \delta} \left| \frac{d}{d\lambda} \left(\frac{\tilde{e}}{\tilde{M}} \right) + \frac{\zeta'(\lambda)}{\zeta(\lambda)^2} \right| \lesssim_{\lambda_0} \eta. \tag{8.9.5}$$

It follows that the charge to mass ratio $\tilde{e}/\tilde{M}$ is *strictly decreasing* as a function of λ, for $\lambda \ge \lambda_0$.

On any fixed neighborhood of $(1, M) \in \mathbb{R}^2$, $\mathcal{F}^{-1} \circ F_{\eta,\varepsilon}$ converges uniformly to the identity map by (8.9.3). Therefore, by a simple degree argument[7] we find an assignment $(\eta, \varepsilon) \to (\lambda(\eta, \varepsilon), M'(\eta, \varepsilon))$ such that $F_{\eta,\varepsilon}(\lambda(\eta, \varepsilon), M'(\eta, \varepsilon)) = (M, M)$. We claim that for η sufficiently small, the family of seed data $\lambda \mapsto \mathcal{S}_\lambda \doteq \mathcal{S}_{\lambda, M'(\eta,\varepsilon),\eta,\varepsilon}$ gives rise to the desired family of spacetimes and Cauchy data:

1. For $\lambda \in [-1, \lambda_0]$, the final Reissner–Nordström parameters of $\mathcal{S}_\lambda$ are $< \frac{1}{2} r_1$ by definition of λ_0 and hence the globally hyperbolic spacetime $\mathcal{D}_\lambda$ associated to $\mathcal{S}_\lambda$ by Proposition 8.9.9 is future causally geodesically complete and dispersive. At $\lambda = -1$, the seed data is trivial and hence the development is isometric to Minkowski space.

2. For $\lambda \in [\lambda_0, \lambda(\eta, \varepsilon))$, the final charge to mass ratio $\tilde{e}/\tilde{M}$ is strictly decreasing towards 1 by (8.9.5). Therefore, $\mathcal{D}_\lambda$ is future causally geodesically complete and dispersive. A neighborhood of spatial infinity i^0 is isometric to a superextremal Reissner–Nordström solution.

3. $\lambda = \lambda(\eta, \varepsilon)$ is, by construction, critical, with parameter ratio $\tilde{e}/\tilde{M} = 1$. $\mathcal{D}_\lambda$ contains a nonempty black hole region and for sufficiently large advanced time, the domain of outer communication and event horizon are isometric to an appropriate portion of an extremal Reissner–Nordström black hole.

4. For $\lambda \in (\lambda(\eta, \varepsilon), 2]$, $\tilde{e}/\tilde{M}$ decreases away from 1 by (8.9.5). By definition of r_2, $r_- > r_2$ for η sufficiently small and hence $\mathcal{D}_\lambda$ contains a nonempty black hole region and for sufficiently large advanced time, the domain of outer communication and event horizon are isometric to an appropriate portion of a subextremal Reissner–Nordström black hole. By (8.9.4), the charge to mass ratio at $\lambda = 2$ can be made arbitrarily close to $\frac{1}{2}$.

To complete the proof, we assign a smooth family of Cauchy data to $\mathcal{D}_\lambda$. Let $i_\lambda : [0, \infty) \to \Sigma_\lambda$ be the arc length parametrization (with respect to the metric g_λ) of the Cauchy surface associated to $\mathcal{D}_\lambda$ by Proposition 8.9.9. Then we may define the embedding $\tilde{i}_\lambda = i_\lambda \times \mathrm{id}_{S^2}/\!\sim : \mathbb{R}^3 \to \tilde{\Sigma}_\lambda \subset \mathcal{M}_\lambda$ (where the central sphere is collapsed to a point). The natural map $\lambda \mapsto \Psi_\lambda$, where Ψ_λ is the Cauchy data induced on $\tilde{\Sigma}_\lambda$ by pullback along $\tilde{i}_\lambda$, is smooth. This completes the proof of the theorem. □

8.10 Weak* convergence to dust

In this section, we show that the spacetimes constructed in Proposition 8.9.9 weak* converge in an appropriate sense to the bouncing charged null dust spacetimes given by Proposition 7.2.3. First,

[7] Specifically, we use the following statement: Let $f : \overline{B}_1 \to \mathbb{R}^d$ be a continuous map, where $\overline{B}_1$ is the closed unit ball in $\mathbb{R}^d$. If $\sup_{\overline{B}_1} |f - \mathrm{id}| < \frac{1}{2}$, then the image of f contains $B_{1/2}$.

we show convergence of an outgoing Vlasov beam to the underlying outgoing formal dust beam:

Proposition 8.10.1. *Let $\alpha \in \mathcal{P}_\Gamma$, let $\{\eta_i\}$ and $\{\varepsilon_j\}$ be decreasing sequences of positive numbers tending to zero, let (r, Ω^2, Q, f) be the solution of the Einstein–Maxwell–Vlasov system associated to $(\alpha, \eta_i, \varepsilon_j)$ by Proposition 8.2.3 for $j \gg i$ and arbitrary allowed mass, and let $(r_\mathrm{d}, \Omega^2_\mathrm{d}, Q_\mathrm{d}, N^v_\mathrm{d}, T^{vv}_\mathrm{d})$ be the outgoing formal dust solution from Section 7.6.2 on $\mathcal{C}_{r_2}$. Then the following holds for any relatively compact and relatively open set $U \subset \mathcal{C}_{r_2}$:*

1. *We have the following strong convergence:*

$$\lim_{\substack{i\to\infty\\ j\gg i}} \left(\|r - r_\mathrm{d}\|_{C^1(U)} + \|\Omega^2 - \Omega^2_\mathrm{d}\|_{C^1(U)} + \|Q - Q_\mathrm{d}\|_{C^0(U)} + \|T^{uu}\|_{C^0(U)} + \|T^{uv}\|_{C^0(U)}\right) = 0, \tag{8.10.1}$$

   where the limit is to be understood as taking $i \to \infty$ while keeping j sufficiently large for each i such that the conclusion of Proposition 8.2.3 applies for $\mathcal{S}_{\alpha,\eta_i,\varepsilon_j}$.

2. *If $U' \subset U$ is disjoint from a neighborhood of $\{\tau = 0\}$, then*

$$\lim_{\substack{i\to\infty\\ j\gg i}} \|N^u\|_{C^0(U')} = 0. \tag{8.10.2}$$

3. *We have the following weak* convergence: For any $\varphi \in C^1_c(\mathbb{R}^2)$ with $\operatorname{spt}\varphi \cap \mathcal{C}_{r_2} \subset U$,*

$$\lim_{\substack{i\to\infty\\ j\gg i}} \int_U (N^u, N^v, T^{vv})\varphi\, du dv = \int_U (0, N^v_\mathrm{d}, T^{vv}_\mathrm{d})\varphi\, du dv. \tag{8.10.3}$$

4. *We have the following weak* convergence: For any $\varphi \in L^1(U)$,*

$$\lim_{i\to\infty} \lim_{j\to\infty} \int_U (N^u, N^v, T^{vv})\varphi\, du dv = \int_U (0, N^v_\mathrm{d}, T^{vv}_\mathrm{d})\varphi\, du dv. \tag{8.10.4}$$

Proof. It is clear from the estimates used in the proof of Lemma 8.5.1 that the Vlasov solutions converge strongly to Minkowski space in the regions $\mathcal{R}^{\check v}_{\mathrm{aux}}$. It therefore suffices to prove the proposition only for the case $U = \mathcal{R}^{\frac12 \check v - \frac13 r_1}_{\mathrm{main}}$, where we use the estimates of Lemma 8.4.1.

Part 1. Using Lemma 8.3.1, (2.3.25), and (8.4.37), we already obtain

$$|Q(u,v) - \check Q(u)| \lesssim \eta$$

for any $(u,v) \in U \doteq \mathcal{R}^{\frac12 \check v - \frac13 r_1}_{\mathrm{main}}$ and note that $\check Q(u) = Q_\mathrm{d}(u,v)$ by (7.6.6). Using this estimate, Lemma 8.3.1, (8.4.37), and Grönwall's inequality on the differences of (2.3.20), (2.3.21) and (7.6.1), (7.6.2), we readily infer

$$|r - r_{\rm d}| + |\Omega^2 - \Omega^2_{\rm d}| \lesssim \eta$$

on U, which completes the proof of (8.10.1).

Part 2. This follows immediately from (8.4.34) since τ is bounded below on U'.

Part 3. Since φ is bounded,

$$\left|\int_U N^u \varphi\, dudv\right| \lesssim \int_{-r_2}^{-\frac{2}{3}r_1} \int_{-u}^{\check{v}} N^u\, dvdu \lesssim \varepsilon^{1/2}$$

by (8.4.37). Next, let $\tilde{\varphi} \doteq \varphi/(-\mathfrak{e}r^2\Omega^2)$, use Maxwell's equation (2.3.24), and integrate by parts:

$$\int_U N^v \varphi\, dudv = \int_U \partial_u Q \tilde{\varphi}\, dudv = -\int_U Q \partial_u \tilde{\varphi}\, dudv + \int_{\partial U} Q\tilde{\varphi} \tag{8.10.5}$$

By Part 1, $Q\partial_u\tilde{\varphi} \to Q_{\rm d}\partial_u\tilde{\varphi}_{\rm d}$ and $Q\tilde{\varphi} \to Q_{\rm d}\tilde{\varphi}_{\rm d}$ uniformly as $i \to \infty$ and $j \ll i$, where $\tilde{\varphi}_{\rm d} \doteq \varphi/(-\mathfrak{e}r_{\rm d}^2\Omega_{\rm d}^2)$. Therefore, passing to the limit, we have

$$\lim_{\substack{i\to\infty \\ j\gg i}} \int_U N^v\varphi\, dudv = -\int_U Q_{\rm d}\partial_u\tilde{\varphi}_{\rm d}\, dudv + \int_U Q_{\rm d}\tilde{\varphi}_{\rm d} = \int_U \partial_u Q_{\rm d}\tilde{\varphi}_{\rm d} = \int_U N^v_{\rm d}\, dudv,$$

where we have again integrated by parts and used (7.6.5). A similar argument applies for the convergence of T^{vv} using the Raychaudhuri equations (2.3.22) and (7.6.3). This completes the proof of (8.10.3).

Part 4. We first fix i and take $j \to \infty$. By (8.4.34), $N^u \lesssim_i 1$, for every $j \ll i$, where we use the notation $A \lesssim_i B$ to denote $A \le CB$, where C may depend on i. By the Banach–Alaoglu theorem, there exists a subsequence j_n and an $L^\infty(U)$ function h such that $N^u \overset{*}{\rightharpoonup} h$ in $L^\infty(U)$. However, by (8.10.2), it is clear that $h = 0$ almost everywhere. Since the subsequential limit is unique, $N^u \overset{*}{\rightharpoonup} 0$ as $j \to \infty$. This completes the proof of (8.10.4) for N^u.

Let $\mathring{m}_{{\rm V},i}$ and $\mathring{Q}_{{\rm V},i}$ be the values of $\mathring{m}$ and $\mathring{Q}$ at $r = \frac{2}{3}r_1$ for the Vlasov seed $\mathcal{S}_{\alpha,\eta_i,\varepsilon_j}$. Note that these numbers do not depend on j. Let $\mathring{m}_i$ and $\mathring{Q}_i$ be the solutions of the system (7.6.12) and (7.6.13) on $[\frac{2}{3}r_1, r_2]$ with initial conditions $\mathring{m}_{{\rm V},i}$ and $\mathring{Q}_{{\rm V},i}$ at $r = \frac{2}{3}r_1$, current $\mathring{\mathcal{N}}^v$ given by (7.6.15), and identically vanishing energy-momentum tensor $\mathring{\mathcal{T}}^{vv}$. Following the proof of Proposition 7.6.6, we obtain a unique smooth solution $(r_{{\rm d},i}, \Omega^2_{{\rm d},i}, Q_{{\rm d},i}, N^v_{{\rm d},i}, T^{vv}_{{\rm d},i})$ of the formal outgoing charged null dust system on U which attains the initial data just described.

By repeating the proof of Part 1 of the present proposition, we see that

$$\lim_{j\to\infty}\left(\|r-r_{\mathrm{d},i}\|_{C^1(U)}+\|\Omega^2-\Omega^2_{\mathrm{d},i}\|_{C^1(U)}+\|Q-Q_{\mathrm{d},i}\|_{C^0(U)}\right)=0$$

for fixed i. Arguing as in Part 2, we then see that

$$\lim_{j\to\infty}\int_U (N^v,T^{vv})\varphi\,dudv=\int_U (N^v_{\mathrm{d},i},T^{vv}_{\mathrm{d},i})\varphi\,dudv \tag{8.10.6}$$

for any $\varphi\in C^1_c(\mathbb{R}^2)$. Since $N^v,T^{vv}\lesssim_i 1$, a standard triangle inequality argument shows that φ can be replaced by any L^1 function in (8.10.6). Now it follows by construction that

$$|N^v_{\mathrm{d},i}-N^v_{\mathrm{d}}|+|T^{vv}_{\mathrm{d},i}-T^{vv}_{\mathrm{d}}|\lesssim\eta_i$$

on U, so we can safely take $i\to\infty$ in (8.10.6), which completes the proof of (8.10.4). □

In order to globalize this, we must first define bouncing charged null dust in the formal system in double null gauge. So consider again the outgoing solution $(r_{\mathrm{d}},\Omega^2_{\mathrm{d}},Q_{\mathrm{d}},N^v_{\mathrm{d}},T^{vv}_{\mathrm{d}})$ on $\mathcal{C}_{r_2}$ with seed data $(\mathring{\mathcal{N}}^v_{\mathrm{d}},0,r_2,\mathfrak{e})$ given by (7.6.15). As in Definition 8.9.4, we extend r_{d}, Ω^2_{d}, and Q_{d} to $\tilde{\mathcal{C}}_{r_2}$ by reflection, and set

$$N^u_{\mathrm{d}}(u,v)=N^v_{\mathrm{d}}(-v,-u),$$
$$T^{uu}_{\mathrm{d}}(u,v)=T^{vv}_{\mathrm{d}}(-v,-u)$$

for $(u,v)\in\tilde{\mathcal{C}}_{r_2}\setminus\{\tau=0\}$. We extend N^v_{d} and T^{vv}_{d} to zero in $\tilde{\mathcal{C}}_{r_2}\setminus\{\tau=0\}$ and similarly extend N^u_{d} and T^{uu}_{d} to zero in $\mathcal{C}_{r_2}$. Using Lemma 8.9.1, we attach a maximal piece of Reissner–Nordström with parameters ϖ_2 and Q_2 to $\mathcal{C}_{r_2}\cup\tilde{\mathcal{C}}_{r_2}$. Let v_∞ be as defined in (8.9.1).

Definition 8.10.2. The *globally hyperbolic bouncing formal charged null dust spacetime* associated to a set of parameters $\alpha\in\mathcal{P}_\Gamma$ is the tuple $(\mathcal{X}_{\mathrm{d}},r_{\mathrm{d}},\Omega^2_{\mathrm{d}},Q_{\mathrm{d}},N^u_{\mathrm{d}},N^v_{\mathrm{d}},T^{uu}_{\mathrm{d}},T^{vv}_{\mathrm{d}})$, where $\mathcal{X}_{\mathrm{d}}\doteq\{v\geq u\}\cap\{v<v_\infty\}$.

For $\tau\geq 0$, $(r_d,\Omega^2_{\mathrm{d}},N^v_{\mathrm{d}},T^{vv}_{\mathrm{d}})$ solves the outgoing formal dust system and for $\tau<0$, $(r_d,\Omega^2_{\mathrm{d}},N^u_{\mathrm{d}},T^{uu}_{\mathrm{d}})$ solves the ingoing formal dust system. The functions r_{d} and Ω^2_{d} are C^1 across $\{\tau=0\}$ and Q_{d}, T^{uu}_{d}, and T^{vv}_{d} are C^0 across $\{\tau=0\}$. The currents N^u_{d} and N^v_{d} are *discontinuous* across $\{\tau=0\}$ (since we extended by zero), but of course $N^u_{\mathrm{d}}=N^v_{\mathrm{d}}$ at $\{\tau=0\}$.

Using this definition and Proposition 8.10.1, we immediately obtain the following

Theorem 8.10.3. *Let $\alpha \in \mathcal{P}_\Gamma$, let $\{\eta_i\}$ and $\{\varepsilon_j\}$ be decreasing sequences of positive numbers tending to zero, let $(\mathcal{X}, r, \Omega^2, Q, f)$ be the globally hyperbolic solution of the Einstein–Maxwell–Vlasov system associated to $(\alpha, \eta_i, \varepsilon_j)$ by Proposition 8.9.9 for $j \ll i$ and arbitrary allowed mass, and let $(\mathcal{X}_\mathrm{d}, r_\mathrm{d}, \Omega^2_\mathrm{d}, Q_\mathrm{d}, N^u_\mathrm{d}, N^v_\mathrm{d}, T^{uu}_\mathrm{d}, T^{vv}_\mathrm{d})$ be the globally hyperbolic bouncing formal charged null dust spacetime associated to α by Definition 8.10.2. Then the following holds for any relatively compact open set $U \subset \mathcal{X}_\mathrm{d}$:*

1. *We have the following strong convergence:*

$$\lim_{\substack{i\to\infty\\ j\gg i}} \left(\|r - r_\mathrm{d}\|_{C^1(U)} + \|\Omega^2 - \Omega^2_\mathrm{d}\|_{C^1(U)} + \|Q - Q_\mathrm{d}\|_{C^0(U)} + \|T^{uv}\|_{C^0(U)}\right) = 0.$$

2. *We have the following weak* convergence: For any $\varphi \in C^1_c(U)$,*

$$\lim_{\substack{i\to\infty\\ j\gg i}} \int_U (N^u, N^v, T^{uu}, T^{vv})\varphi \, dudv = \lim_{\substack{i\to\infty\\ j\ll i}} \int_U (N^u_\mathrm{d}, N^v_\mathrm{d}, T^{uu}_\mathrm{d}, T^{vv}_\mathrm{d})\varphi \, dudv.$$

3. *We have the following weak* convergence: For any $\varphi \in L^1(\mathbb{R}^2)$ with $\operatorname{spt}\varphi \subset U$,*

$$\lim_{i\to\infty} \lim_{j\to\infty} \int_U (N^u, N^v, T^{uu}, T^{vv})\varphi \, dudv = \lim_{\substack{i\to\infty\\ j\ll i}} \int_U (N^u_\mathrm{d}, N^v_\mathrm{d}, T^{uu}_\mathrm{d}, T^{vv}_\mathrm{d})\varphi \, dudv.$$

Remark 8.10.4. The globally hyperbolic region $\mathcal{X}$ depends on η_i and ε_j, but it always holds that $U \subset \mathcal{X}$ for i sufficiently large and $j \gg i$.

8.11 The third law in Einstein–Maxwell–Vlasov and event horizon jumping at extremality

Using the technology of bouncing Vlasov beams developed in the proof of Theorem 1.2.1, we are now able to quickly prove Theorems 1.2.22 and 1.2.23. As complete proofs would require more lengthy setup, we only sketch the proofs of these results. We refer the reader back to Section 1.2.7 for the theorem statements and discussion.

8.11.1 Counterexamples to the third law

Refer to Fig. 8.6 for global Penrose diagrams.

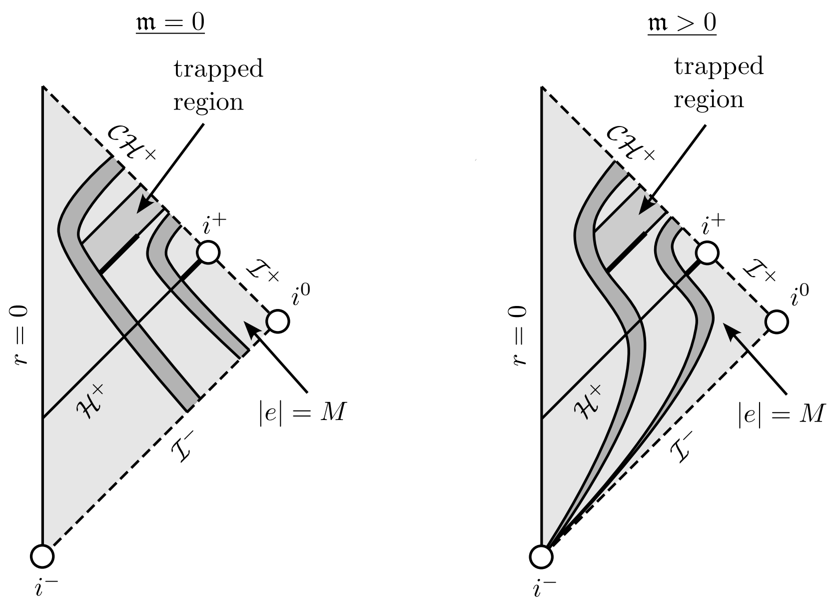

Figure 8.6: Penrose diagrams of counterexamples to the third law of black hole thermodynamics in the Einstein–Maxwell–Vlasov model. The disconnected thick black curve denotes the outermost apparent horizon $\mathcal{A}'$, which jumps outward as the black hole becomes extremal. This behavior is necessary in third law violating spacetimes which obey the weak energy condition, see Proposition 1.1.10 and [Isr86].

Proof of Theorem 1.2.22. Let $\alpha \in \mathcal{P}$ be third law violating dust parameters as in Theorem 7.4.1. We desingularize this dust beam as in the proof of Theorem 1.2.1, noting that the charge on the inner edge of the beam is bounded below and hence no auxiliary beam is required. In order to achieve extremality we must modulate α slightly as in Theorem 1.2.1, but all required inequalities for this construction are strict, so this can be done. By this procedure we obtain the time-symmetric solution $\mathcal{D}_{\text{ext}}$ depicted in Fig. 8.7 below.

Let (ϖ_1, Q_1) be the initial Reissner–Nordström parameters of α, which will also be the initial parameters of $\mathcal{D}_{\text{ext}}$. Using the same methods as Theorem 1.2.1 and Lemma 7.3.1, we can construct a solution $\mathcal{D}_{\text{sub}}$ of Einstein–Maxwell–Vlasov collapsing to a subextremal Reissner–Nordström black hole with parameters ϖ_1 and Q_1. In the case of massless particles, the desired third law violating spacetime is then obtained by deleting an appropriate double null rectangle from $\mathcal{D}_{\text{sub}}$ and gluing in an appropriate piece of $\mathcal{D}_{\text{ext}}$. In the case of massive particles, the beams from $\mathcal{D}_{\text{ext}}$ and $\mathcal{D}_{\text{sub}}$ will possibly interact in the early past, but as is clear from the proof of Proposition 8.2.3, this happens in the dispersive region and the proof of Lemma 8.6.1 can be repeated to show global existence and causal geodesic completeness in the past. $\square$

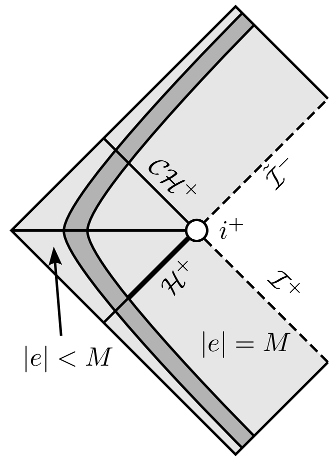

Figure 8.7: Penrose diagram of the time symmetric Einstein–Maxwell–Vlasov solution $\mathcal{D}_{\text{ext}}$ interpolating between subextremality and extremality. This diagram is valid for both massive and massless particles.

8.11.2 (Semi)continuity of the location of the event horizon

8.11.2.1 General results for spherically symmetric event horizons

We now consider general *weakly tame* spherically symmetric Einstein-matter systems, i.e., those satisfying the dominant energy condition and the weak extension principle. We refer to [Daf05b; Kom13] for the precise definition of the weak extension principle, but note that it is a strictly weaker condition than the generalized extension principle as formulated in Proposition 3.2.4 and therefore holds for the Einstein–Maxwell–Klein–Gordon, Einstein–Higgs, and Einstein–Maxwell–Vlasov systems [Chr93; Daf05a; Kom13; DR16].

Let $\Psi = (\Sigma, \bar{g}, \bar{k}, \dots)$ be a spherically symmetric, asymptotically flat Cauchy data set on $\mathbb{R}^3$ with a regular center in the given matter model. We will assume that Ψ contains no spherically symmetric antitrapped surfaces, i.e., that $\partial_u r < 0$ on Σ. This assumption is physically motivated by the observation that if the maximal Cauchy development $\mathcal{D} = (\mathcal{Q}, r, \Omega^2, \dots)$ does not contain a white hole, then there are no antitrapped surfaces in the spacetime. Furthermore, by Raychaudhuri's equation (2.1.7) and the dominant energy condition, $\partial_u r < 0$ is propagated to the future of the Cauchy hypersurface Σ.

Under these assumptions, a very general a priori characterization of the boundary of $(\mathcal{Q}, r, \Omega^2)$ is available due to the work of Dafermos [Daf05b] and we refer to Kommemi [Kom13] for a detailed account. We will utilize the following two facts:

Fact 1. If $(\mathcal{Q}, r, \Omega^2)$ contains a trapped or marginally trapped surface, i.e., $\partial_v r(u_0, v_0) \le 0$ for some $(u_0, v_0) \in \mathcal{Q}$, then the black hole region is nonempty ($\mathcal{BH} \doteq \mathcal{Q} \setminus J^-(\mathcal{I}^+) \neq \emptyset$), future null infinity is complete in the sense of Christodoulou [Chr99a], and $(u_0, v_0) \in \mathcal{BH}$.

Fact 2. The Hawking mass m extends to a (not necessarily continuous) nonincreasing and nonnegative function on future null infinity $\mathcal{I}^+$, called the *Bondi mass*. If the Hawking mass of Σ is bounded,[8] then the Bondi mass is also bounded and the *final Bondi mass* $M_f \doteq \inf_{\mathcal{I}^+} m$ is finite. Then the "event horizon Penrose inequality" $\sup_{\mathcal{H}^+} r \leq 2M_f$ holds and by the no antitrapped surfaces condition we also obtain

$$\sup_{\mathcal{BH}} r \leq 2M_f. \tag{8.11.1}$$

We wish to consider sequences of initial data and their developments. In order to compare them, we have to ensure that the double null gauges are synchronized in an appropriate sense. We consider only developments $\mathcal{D} = (\mathcal{Q}, r, \Omega^2, \dots)$ for which the center Γ is a subset of $\{u = v\} \subset \mathbb{R}^2_{u,v}$ and if $i : [0, \infty) \to \Sigma$ denotes the embedding map of the Cauchy hypersurface into $\mathcal{Q}$, we demand that $i(0) \in \Gamma$. Clearly, these conditions can always be enforced by an appropriate transformation of the double null gauge.

Assumption 8.11.1. Let $\{\Psi_j\}$ be a sequence of spherically symmetric asymptotically flat Cauchy data for a weakly tame Einstein-matter system. Let $\mathcal{D}_j = (\mathcal{Q}_j, r_j, \Omega_j^2, \dots)$ denote the maximal development of Ψ_j with Cauchy hypersurface $\Sigma_j \subset \mathcal{Q}_j$ and embedding map $i_j : [0, \infty) \to \Sigma_j$ normalized as above. We assume that Ψ_j converges to another data set Ψ_∞ and the developments converge in the following sense:

1. Gauge condition: Let $\mathcal{D}_\infty$ denote the maximal development of Ψ_∞ with Cauchy hypersurface and embedding map $i_\infty : [0, \infty) \to \Sigma_\infty$. We assume that $\mathcal{D}_j$ and $\mathcal{D}_\infty$ have *continuously synchronized gauges* in the sense that $(u, v) \circ i_j : [0, \infty) \to \mathbb{R}^2$ converges uniformly on compact sets as $j \to \infty$.

2. Cauchy stability of the area-radius: If $U \subset \mathcal{Q}_\infty$ is a relatively compact open set, then $U \subset \mathcal{Q}_j$ for j sufficiently large and $r_j \to r_\infty$ in $C^1(U)$.

Remark 8.11.2. The notion of continuous synchronization also makes sense for continuous one-parameter families of Cauchy data $\lambda \mapsto \Psi_\lambda$. In this case we require the maps $(u, v) \circ i_\lambda(x)$ to be jointly continuous in λ and $x \in [0, \infty)$.

Remark 8.11.3. The initial data and developments given by Proposition 8.9.9 are continuously synchronized as functions of the beam parameters $(\alpha, \eta, \varepsilon, \mathfrak{m})$.

[8] This can be regarded as a part of the definition of asymptotic flatness.

Proposition 8.11.4. *Let $\Psi_j \to \Psi_\infty$ be a convergent sequence of one-ended asymptotically flat Cauchy data for a weakly tame spherically symmetric Einstein-matter system, containing no spherically symmetric antitrapped surfaces, and satisfying Assumption 8.11.1. Assume that the sequence has uniformly bounded Bondi mass and that the development $\mathcal{D}_j$ of Ψ_j contains a black hole for each $j \in \mathbb{N}$. Let $r_{j,\mathcal{H}^+}$ denote the limiting area-radius of the event horizon $\mathcal{H}_j^+$ of $\mathcal{D}_j$ and $u_{j,\mathcal{H}^+}$ its retarded time coordinate (also for $j = \infty$ if $\mathcal{D}_\infty$ contains a black hole). Then the following holds:*

1. *If future timelike infinity i_∞^+ is a limit point of the center Γ in $\mathcal{D}_\infty$ (in particular, $\mathcal{D}_\infty$ does not contain a black hole), then*
$$\lim_{j\to\infty} u_{j,\mathcal{H}^+} = \sup_{\mathcal{Q}_\infty} u. \tag{8.11.2}$$

2. *If $\mathcal{D}_\infty$ contains a black hole, then*

   (a) *The retarded time of the event horizon is lower semicontinuous:*
$$\liminf_{j\to\infty} u_{j,\mathcal{H}^+} \geq u_{\infty,\mathcal{H}^+}. \tag{8.11.3}$$

   (b) *Assume further that there are trapped surfaces asymptoting to i_∞^+ in the following sense: Let $(u_{\infty,i^+}, v_{\infty,i^+})$ denote the coordinates[9] of i_∞^+ and suppose there exist sequences $u_1 > u_2 > \cdots \to u_{\infty,i^+}$ and $v_1 < v_2 < \cdots \to v_{\infty,i^+}$ such that $\partial_v r_\infty(u_i, v_i) < 0$ for every $i \geq 1$. Then* (8.11.3) *is upgraded to*
$$\lim_{j\to\infty} u_{j,\mathcal{H}^+} = u_{\infty,\mathcal{H}^+} \tag{8.11.4}$$
   and it additionally holds that
$$\liminf_{j\to\infty} r_{j,\mathcal{H}^+} \geq r_{\infty,\mathcal{H}^+}. \tag{8.11.5}$$

Proof. Part 1: The inequality $\leq$ in (8.11.2) follows directly from Cauchy stability, which implies the stronger statement
$$\limsup_{j\to\infty} \sup_{\mathcal{Q}_j} u \leq \sup_{\mathcal{Q}_\infty} u.$$
We now prove the inequality $\geq$. Let $u_0 < \sup_{\mathcal{Q}_\infty} u$ and let M be an upper bound for the Bondi mass of the sequence $\{\mathcal{D}_j\}$. Since the cone C_{u_0} in $\mathcal{Q}_\infty$ reaches future null infinity, we can choose v_0 such that $r_\infty(u_0, v_0) > 2M$. By Cauchy stability, $r_j(u_0, v_0) > 2M$ for j sufficiently large, so $(u_0, v_0) \notin \mathcal{BH}_j$ by (8.11.1). This implies $u_{j,\mathcal{H}^+} \geq u_0$ which completes the proof.

[9] u_{∞,i^+} is necessarily finite and equals $u_{\infty,\mathcal{H}^+}$, but v_{∞,i^+} could be $+\infty$.

Part 2 (a): The argument to prove (8.11.3) is the same as the proof of Part 1, since C_{u_0} reaches future null infinity in $\mathcal{Q}_\infty$ for any $u_0 < u_{\infty,\mathcal{H}^+}$.

Part 2 (b): By (8.11.3) we must now show that for $u_0 > u_{\infty,\mathcal{H}^+}$, $u_{j,\mathcal{H}^+} \le u_0$ for j sufficiently large. Let (u_i, v_i) be a trapped sphere for $\mathcal{D}_\infty$ as in the statement, with $u_0 > u_i > u_{\infty,\mathcal{H}^+}$. By Cauchy stability, (u_i, v_i) is then trapped in $\mathcal{D}_j$ for j sufficiently large and therefore $u_i > u_{j,\mathcal{H}^+}$, which completes the proof of (8.11.4). Using this, we now have by Cauchy stability and monotonicity of r_j along $\mathcal{H}^+_j$ that

$$\liminf_{j\to\infty} r_{j,\mathcal{H}^+} \ge \lim_{j\to\infty} r_j(u_{j,\mathcal{H}^+}, v_0) = r_\infty(u_{\infty,\mathcal{H}^+}, v_0)$$

for any $v_0 < v_{\infty,i^+}$. Letting $v_0 \to v_{\infty,i^+}$ completes the proof. □

Remark 8.11.5. It is natural to ask if the "reverse" of (8.11.5), i.e.,

$$\limsup_{j\to\infty} r_{j,\mathcal{H}^+} \le r_{\infty,\mathcal{H}^+}, \tag{8.11.6}$$

holds at this level of generality (even assuming trapped surfaces asymptoting towards i^+_∞). It turns out that (8.11.6) is *false* without additional assumptions. On the one hand, assuming additionally a strict inequality in (8.11.3), a minor modification of the arguments used to show (8.11.3) can be used to show (8.11.6).[10] On the other hand, without the assumption of a strict inequality in (8.11.3), using the ingoing (uncharged) Vaidya metric, one can construct a counterexample to (8.11.6) which moreover satisfies the asymptotic trapped surface assumption of Part 2 (b). One is to imagine inflating the event horizon of a Schwarzschild black hole by injecting a fixed dust packet at later and later advanced times $v \sim j$. In the limit $j \to \infty$, the dust disappears, and $\limsup r_{j,\mathcal{H}^+} > r_{\infty,\mathcal{H}^+}$. Curiously, since black holes in the Vaidya model always have trapped surfaces behind the horizon, Part 2 (b) of the proposition implies that $u_{j,\mathcal{H}^+}$ is actually continuous in this process. This is because injecting a fixed amount of matter at later and later times causes the horizon to move outwards less and less (in u), causing it to converge back to the original Schwarzschild horizon as $j \to \infty$. Therefore, in order for (8.11.6) to hold, one must assume that Ψ_j converges to Ψ_∞ in a norm that sufficiently respects the asymptotically flat structure. We emphasize at this point that the conclusions of Proposition 8.11.4 hold only under an assumption of *local Cauchy stability*—no asymptotic stability or weighted convergence is required.

[10] Note that in this case, there cannot be trapped surfaces asymptoting towards i^+_∞!

8.11.2.2 Proof of event horizon jumping in the Einstein–Maxwell–Vlasov model

Proof of Theorem 1.2.23. This is proved by following the proof of Theorem 1.2.22 and varying the final parameters of $\mathcal{D}_{\mathrm{ext}}$ as in the proof of Theorem 1.2.1. □

This theorem shows that it is not always possible to have equality in (8.11.3) when $\mathcal{D}_\infty$ is extremal: the event horizon can and does jump as a function of the initial data.

Part II

Positive mass theorems with black holes and arbitrary ends

Chapter 9

Overview of Part II

9.1 Introduction

Let M^n be a smooth manifold, g a complete Riemannian metric on M, and k a symmetric $(0,2)$-tensor on M^n. Such a triple naturally arises as an *initial data set*[1] for the Cauchy problem in general relativity. We say that (M^n, g, k) is *asymptotically flat with compact core* if there exists a compact set $K \subset M$ such that $M \setminus K$ is diffeomorphic to $\mathbb{R}^n \setminus \overline{B}$, and in the coordinates x^i induced by this diffeomorphism, $g_{ij} - \delta_{ij}$ and k_{ij} decay suitably. This definition is intended to capture the configuration of an isolated system in a Cauchy hypersurface of an asymptotically flat spacetime.

Given an initial data set, the *mass density* μ (a scalar) and the *momentum density* J (a vector field) are defined by

$$\mu = \tfrac{1}{2}\left(R_g - |k|_g^2 + (\operatorname{tr}_g k)^2\right),$$
$$J^i = (\operatorname{div}_g k)^i - \nabla^i(\operatorname{tr}_g k).$$

If (M, g, k) arises as a spacelike hypersurface in a Lorentzian manifold satisfying the Einstein equations (1.0.1), then $\mu = T^{00}$ and $J^i = T^{0i}$, where 0 denotes the timelike direction normal to the hypersurface. We say that (M, g, k) satisfies the *dominant energy condition* (*DEC*) if $\mu \geq |J|_g$. This condition is satisfied by reasonable matter systems, such as those considered in Part I of this dissertation. As a special case, (M, g, k) is *vacuum* if μ and J vanish identically.[2] In this case, (M, g, k)

[1] This terminology, though standard, is misleading. The triple (M, g, k) need not solve constraints for any specific Einstein-matter system to be called an "initial data set" in this part of the dissertation.

[2] The vacuum constraints have the character of an underdetermined elliptic system. The DEC $\mu - |J|_g \geq 0$ may be thought of as an underdetermined elliptic *inequality*.

can be developed into a globally hyperbolic spacetime satisfying the Einstein vacuum equations (1.1.6) [Fou52; CG69].

Associated to an asymptotically flat end is the *ADM energy-momentum* vector $(E, P^1, \ldots, P^n) \in \mathbb{R}^{n+1}$, which captures a notion of total energy and momentum of the initial data set at infinity. The celebrated *spacetime positive mass theorem* of Schoen and Yau states that the total energy-momentum of an asymptotically flat data set is nonnegative if the local energy-momentum density is:

Theorem 9.1.1 (Schoen–Yau [SY81a])**.** *Let (M^3, g, k) be a 3-dimensional asymptotically flat initial data set with compact core and satisfying the dominant energy condition. Then (E, P) is a future-directed causal vector, that is, $E \geq |P|$. In particular, the total mass $E \geq 0$. If $E = |P|$, then (M^3, g, k) embeds into Minkowski space.*

For the history of this theorem, we refer the reader to the excellent textbook [Lee19]. Schoen and Yau's proof of Theorem 9.1.1 involves a reduction to the *Riemannian positive mass theorem*, which is itself a fundamental result in scalar curvature geometry:[3]

Theorem 9.1.2 (Schoen–Yau [SY79a])**.** *Let (M^3, g) be a 3-dimensional asymptotically flat Riemannian manifold with compact core and nonnegative scalar curvature. Then $m_{\mathrm{ADM}} \geq 0$ and $m_{\mathrm{ADM}} = 0$ if and only if g is flat.*

In this part of the dissertation, we consider two natural generalizations of Theorem 9.1.1:

1. In Chapter 10, we show that Theorem 9.1.1 holds for asymptotically flat data sets with compact core and *nonempty compact boundary* in dimensions 3 through 7.[4] The natural boundary condition is for ∂M to be *weakly outer trapped*, which means M contains an apparent horizon. We therefore call this result the *spacetime positive mass theorem for black holes.* This result removes the restrictive topological assumption of spin from the theorem of Gibbons, Hawking, Horowitz, and Perry [GHHP83].

2. In Chapter 11, we show that Theorem 9.1.2 holds for complete asymptotically flat manifolds—without the compact core assumption. The manifolds considered may have other non-asymptotically flat ends. We therefore call this result the *positive mass theorem with arbitrary ends.* This theorem was conjectured by Schoen–Yau in [SY88] in connection with conformal geometry, which

[3] In the context of the Riemannian positive mass theorem, it is customary to denote E by m_{ADM}.

[4] In this part of the dissertation, we always take $3 \leq n \leq 7$, where n is the dimension of the manifold. The dimension restriction comes from the regularity theory of minimal hypersurfaces. Removing this bound is a very difficult open problem, though there have been some attempts and partial results in recent years. The results in this dissertation can be applied as a black box to the higher dimensional case once the relevant regularity theory has been developed.

will be discussed in Section 9.4. Furthermore, our methods imply corollaries for incomplete manifolds containing a *positive scalar curvature shield.*

Both of these results also have applications to extremal black hole spacetimes, as these often do not contain trapped surfaces or asymptotically flat Cauchy hypersurfaces with compact cores.

Density theorems

These main common thread between these two lines of results is the use of *density theorems* to perturb the original asymptotically data into something easier to work with. The first result of this kind was proved by Schoen and Yau in [SY81b], who showed how to perturb an asymptotically flat, scalar-flat metric

$$g_{ij}(x) = \delta_{ij} + O_2(|x|^{-1}), \tag{9.1.1}$$

to an asymptotically flat, scalar-flat metric satisfying

$$\tilde{g}_{ij}(x) = u^4(x)\delta_{ij} \tag{9.1.2}$$

outside a compact set, where u is a harmonic function with respect to the underlying Euclidean structure and $u(x) \to 1$ as $|x| \to \infty$. Such a metric has so-called *Schwarzschild asymptotics*, which can be seen by expanding u in spherical harmonics, and so may be treated using the methods of [SY79a]. This result can be generalized to show that an asymptotically flat, nonnegative scalar curvature metric can also be perturbed so that (9.1.2) holds outside a compact set, while maintaining nonnegative scalar curvature [LP87; Kuw90]. One may characterize these results as saying that metrics with asymptotic expansion (9.1.2) are *dense* among metrics with asymptotic expansion (9.1.1), either subject to the constraint $R_g = 0$, or the constraint $R_g \geq 0$, where R_g denotes the scalar curvature of g.

Density theorems of this kind are used in every minimal hypersurface, marginally outer trapped surface (MOTS), and Jang reduction proof of the positive mass theorem [SY79a; SY81b; SY81a; Sch89; Loh99; Eic13; EHLS16; Loh16; Loh17; SY19]. The improved asymptotics are crucially used to construct barriers and perform asymptotic analysis on area-minimizing hypersurfaces (or stable MOTS) spanning large spheres in the asymptotically flat region. Therefore, most of the analytic work in proving the positive mass theorem lies in establishing an appropriate density theorem—the geometric arguments take place on the perturbed, simplified manifold. Most of Chapter 10 and Chapter 11 is indeed concerned with proving density theorems in the relevant settings.

9.2 The spacetime positive mass theorem for black holes

In Chapter 10, we settle the full spacetime positive mass theorem with boundary and without a spin assumption, at least in dimensions where we have regularity of C-minimizing integral currents. This theorem was proved in joint work with Dan A. Lee and Martin Lesourd [LLU22].

Theorem 9.2.1 (Spacetime positive mass theorem with boundary)**.** *Let $3 \leq n \leq 7$, and let (M^n, g, k) be a complete asymptotically flat initial data set with compact boundary ∂M such that the dominant energy condition holds on M, and each component of ∂M is either weakly outer trapped ($\theta^+ \leq 0$) or weakly inner untrapped ($\theta^- \geq 0$), with respect to the normal pointing into M. Then $E \geq |P|$ in each end, where (E, P) denotes the ADM energy-momentum vector of (g, k).*

The quantities $\theta^\pm$ are the null expansions of the boundary. The definitions and precise assumptions for this theorem will be given in Section 10.2.1.

The $\partial M = \emptyset$ case of Theorem 9.2.1 was proved by Eichmair, Huang, Lee, and Schoen in [EHLS16]. The $\partial M \neq \emptyset$ case is desirable from a physical perspective, since it verifies the intuition that the geometry lying behind an "apparent horizon" cannot influence the asymptotic geometry. In general relativity, a surface with $\theta^+ \leq 0$ cannot be seen from null infinity, and therefore must lie inside a black hole [Haw72b]. Previously, the $\partial M \neq \emptyset$ case was only known to be true for spin manifolds, by work of Gibbons, Hawking, Horowitz, and Perry [GHHP83] (see also [Her98]), who implemented Witten's spinor method [Wit81] with a boundary condition. In 3 dimensions, Theorem 9.2.1 also follows from recent work of Hirsch, Kazaras, and Khuri [HKK21], using an unrelated method. The short note of Galloway and Lee [GL21] proves Theorem 9.2.1 under the stronger assumption that each component of ∂M either has $\theta^+ < 0$ or $\theta^- > 0$.

Despite these advances and the general belief that Theorem 9.2.1 was true, the problem remained open for a long time. It is natural to adapt the proof of the $\partial M = \emptyset$ case in [EHLS16], and in fact, the proof is essentially unchanged for $\partial M \neq \emptyset$ *if* one already has so-called *harmonic asymptotics*, which is the "initial data set version" of condition (9.1.2). However, it is not clear how to prove a *density theorem* to achieve harmonic asymptotics as in [EHLS16, Theorem 18] when a boundary is present. This is what we accomplish with Theorem 10.1.1, and we explain how Theorem 9.2.1 follows from Theorem 10.1.1 and [EHLS16] in Section 10.4.1. The reason why Theorem 10.1.1 is a nontrivial generalization of Theorem 18 of [EHLS16] is that the latter is proved by solving an elliptic system, and the weakly outer trapped condition on the boundary is not an elliptic boundary condition for this system. We solve this problem by supplementing the weakly outer trapped condition with other conditions to create an elliptic boundary condition.

Since one expects that $E = |P|$ is only possible if the initial data set sits inside Minkowski space, which does not contain weakly outer trapped surfaces, one should be able to strengthen Theorem 9.2.1 to conclude that $E > |P|$. Indeed, we are able to do this if one is willing to make stronger assumptions about the asymptotics.

Theorem 9.2.2. *Assume the hypotheses of Theorem 9.2.1 with $\partial M \neq \emptyset$, and furthermore, assume that (M, g, k) satisfies the stronger asymptotic assumption appearing in Theorem 10.4.1. Then $E > |P|$.*

By work of Beig and Chruściel [BC96] (see also [CM06]), this result should also hold for all spin manifolds, and this argument is sketched[5] in [BC03, Remark 11.5] in dimension 3. The 3-dimensional case was also obtained by [HKK21], and more recently, Hirsch and Yiyue Zhang used this approach to remove the "stronger asymptotic assumption" in 3 dimensions [HZ23]. In the spin case, Hirsch–Zhang have characterized $E = |P|$ spacetimes with "weak" decay as "pp wave" spacetimes in the very recent paper [HZ24].

Our proof of Theorem 9.2.2 follows fairly easily from Theorem 9.2.1 combined with known techniques in the $\partial M = \emptyset$ case. Specifically, we break the proof into two parts. In the first part, presented in Section Section 10.4.2, we suppose that $E = |P|$ and conclude that $E = |P| = 0$ by adapting the $\partial M = \emptyset$ proof by Huang and the first author [HL20]. This is where the stronger asymptotic assumption in Theorem 10.4.1 is needed. In the second part, presented in Section Section 10.4.3, we show that $E = 0$ is impossible by examining Eichmair's Jang equation proof (in the $\partial M = \emptyset$ case) that $E \geq 0$ in [Eic13] (which itself generalized Schoen–Yau's pioneering result in dimension 3 [SY81b]). Technically, in dimension 3 our argument requires the assumption that $\operatorname{tr}_g k = O(|x|^{-\gamma})$ for some $\gamma > 2$, but we choose to leave this assumption out of the statement of Theorem 9.2.2 by explicitly relying on either [HKK21] or [BC03, Remark 11.5] (both of which require very different methods than the ones discussed in this dissertation).

Remark 9.2.3. Huang and Lee have given a very beautiful and completely different proof of $E = 0$ rigidity in [HL23], which makes use of some results from [LLU22].

Remark 9.2.4.

Besides the positive mass theorem, another application of Theorem 10.1.1 concerns the gluing problem for initial data sets. Indeed, since the gluing-across-annulus theorem of Corvino–Schoen [CS06] is appropriately local and done in a region where the data is vacuum and has good asymptotics, we can combine it with Theorem 10.1.1 to obtain the following.

[5] Note that this discussion only appears in the arXiv version of the paper.

Corollary 9.2.5. *Let (M, g, k) be an asymptotically flat initial data set satisfying the assumptions of Theorem 10.1.1, such that $E > |P|$. Then, for any $\varepsilon > 0$, there is an initial data set $(\tilde{g}, \tilde{k})$ with the following properties:*

- *$(\tilde{g}, \tilde{k})$ satisfies the dominant energy condition,*
- *the outer null expansion of ∂M with respect to $(\tilde{g}, \tilde{k})$ is unchanged, that is, $\tilde{\theta}^+ = \theta^+$,*
- *$(\tilde{g}, \tilde{k})$ is ε-close to (g, k) in $W^{2,p}_{-q} \times W^{1,p}_{-q-1}$,*
- *$(\tilde{\mu}, \tilde{J})$ is ε-close to (μ, J) in L^1,*
- *outside a compact set containing ∂M, $(\tilde{g}, \tilde{k})$ is isometric to an initial data set for a Kerr spacetime[6] with ADM energy-momentum $(\tilde{E}, \tilde{P})$, where $|\tilde{E} - E| + |\tilde{P} - P| < \varepsilon$.*

9.3 The positive mass theorem with arbitrary ends

An interesting geometric feature of extremal black holes is that Cauchy surfaces for their domains of outer communication can have *arbitrary ends.* For example, the $t = 0$ slice of extremal Reissner–Nordström has one asymptotically flat end as $r \to \infty$ but is actually asymptotically cylindrical as $r \to 0$. This Cauchy surface does not have a compact core, and therefore Theorem 9.1.2 does not apply. As it turns out, the arbitrary ends cause difficulties in every step of the proof.

In joint work with Lesourd and Yau [LUY21] and Lee and Lesourd [LLU23], we were able to remove the compact core assumption on Theorem 9.1.2, while keeping the most general possible asymotptics on the asymptotically flat end:

Theorem 9.3.1 (The positive mass theorem with arbitrary ends)**.** *Let (M^n, g), $3 \le n \le 7$, be a complete manifold with nonnegative scalar curvature and at least one asymptotically flat end $\mathcal{E}$ of Sobolev type (p, q), where $p > n$ and $q > \frac{n-2}{2}$. Then the ADM mass of $\mathcal{E}$ is nonnegative. Furthermore, if the mass is zero, then (M, g) is isometric to Euclidean space.*

Remark 9.3.2. In [LUY21], we proved this theorem under a stronger decay assumption on $\mathcal{E}$ and were unable to conclude a rigidty statement. We revisted the problem in [LLU23] with more robust methods, which resulted in the stronger version of the theorem stated here.

Our proof makes use of Gromov's *μ-bubble* technique [Gro96; Gro18] to localize the Schoen–Yau descent scheme of [SY79b] to the asymptotically flat end (see already Section 9.3.2). Our methods

[6] More specifically, this Kerr initial data comes from an element of the "reference family" for Kerr, as described in [CD03].

actually prove the following stronger result, which is a positive mass analogue of Gromov's *band width inequality.*

Theorem 9.3.3 (Quantitative shielding theorem)**.** *Let (M^n, g), $3 \leq n \leq 7$, be an asymptotically flat manifold of Sobolev type (p, q), with $p > n$ and $q > \frac{n-2}{2}$, not assumed to be complete or to have nonnegative scalar curvature everywhere. Let U_0, U_1, and U_2 be neighborhoods of an asymptotically flat end $\mathcal{E}$ such that $\overline{U_2} \subset U_1$, $\overline{U_1} \subset U_0$, and $\overline{U_0 \setminus \mathcal{E}}$ is compact, and let*

$$D_0 = \mathrm{dist}_g(\partial U_0, U_1) \quad \textit{and} \quad D_1 = \mathrm{dist}_g(U_2, \partial U_1).$$

If the following hold:

1. *g has no points of incompleteness in U_0,*
2. *$R_g \geq 0$ on U_0, and*
3. *the scalar curvature satisfies the largeness assumption*

$$R_g > \frac{4}{D_0 D_1} \quad \textit{on} \quad \overline{U_1} \setminus U_2, \tag{9.3.1}$$

then the ADM mass of the asymptotically flat end $\mathcal{E}$ is strictly positive.

We call this a *shielded* version of the positive mass theorem because the positive scalar curvature band shields $\mathcal{E}$ from $M \setminus U_0$, on which we make no assumption.

Our methods also imply an inextendibility result: Given an asymptotically flat end $\mathcal{E}$ with nonnegative scalar curvature and negative mass, Theorem 9.3.1 tells us that it is impossible to extend $\mathcal{E}$ to be a complete manifold with nonnegative scalar curvature. The following corollary of Theorem 9.3.3 states that, in fact, there is a *fixed distance* D that puts a limit on how far we can extend the metric away from $\mathcal{E}$ before hitting either a point of incompleteness or a point of negative scalar curvature.

Corollary 9.3.4. *Let (M^n, g), $3 \leq n \leq 7$, be a Riemannian manifold, not assumed to be complete, with an asymptotically flat end $\mathcal{E}$ of Sobolev type (p, q), where $p > n$ and $q > \frac{n-2}{2}$. If $m_{\mathrm{ADM}}(\mathcal{E}, g) < 0$, then there exists a constant D, depending only on $m_{\mathrm{ADM}}(\mathcal{E}, g)$ and $\|g - \delta\|_{W^{2,p}_{-q}(\mathcal{E})}$, with the following property. In the D-neighborhood $N_D(\mathcal{E})$ of $\mathcal{E}$, one or both of the following must be true:*

1. *$R_g < 0$ somewhere in $N_D(\mathcal{E})$, or*

2. $N_D(\mathcal{E})$ contains an incomplete point.

We have given further applications of these techniques to questions about quasilocal mass in [LLU24].

Remark 9.3.5 (Spinor methods). Using Dirac operators on a space with a strong weight on the other ends, Bartnik and Chruściel are able to prove a remarkable *spacetime* positive mass theorem with arbitary ends under a spin assumption [BC03]. After [LUY21] appeared, Cecchini and Zeidler revisted the Riemannian positive mass theorem in the spin setting [CZ21]. They use Callias operators (i.e., Dirac operators with a potential) as a localization tool. They obtain analogues of all our results stated above, which is not needed for spin arguments. Also, their asymptotics are slightly stronger than ours, but we do not believe that this is essential to their method. We would also like to point the reader to the very interesting paper [CZ20].

9.3.1 Outline of the proof of Theorem 9.3.1

An essential ingredient to obtaining the positive mass theorem with the weakest decay assumptions is a *density theorem* that reduces the problem to studying an asymptotically flat manifold with harmonic asymptotics (i.e., of the form (9.1.2)). See already Theorem 11.1.1 for the precise statement. With such a density theorem in hand, Theorem 9.3.1 can be proved in one of two ways:

Proof 1. Assume for the sake of contradiction that $m_{\mathrm{ADM}} < 0$. One now makes a slight conformal change of the metric that makes the scalar curvature strictly positive in an annular region surrounding the asymptotically flat end $\mathcal{E}$. (This procedure is called "bumping up," see already Proposition 11.4.2 below.) The mass is still negative after this perturbation and one has a contradiction to Theorem 9.3.3 after choosing U_0, U_1, and U_2 appropriately.

We give the details of this proof in Section 11.5.2 below.

Proof 2. Assume for the sake of contradiction that $m_{\mathrm{ADM}} < 0$. Applying the density theorem, we perturb the metric to be conformally flat far out on $\mathcal{E}$ while keeping the mass negative. Lohkamp [Loh99] observed that one can then deform u to be constant outside of a (possibly very large) compact set, while staying *superharmonic* (see also [CP11]).[7] This means that the deformed metric has nonnegative scalar curvature everywhere, and is flat outside of this compact set. By identifying sides of a cube surrounding the non-flat region, and taking X to be the one-point compactification

[7] This deformation step uses $m_{\mathrm{ADM}} < 0$ crucially and gives the positive mass theorem the flavor of a Liouville theorem.

of M along $\mathcal{E}$, one obtains a complete metric on $X^n \# T^n$ with positive scalar curvature.[8] When X^n is compact, a classical result of Schoen and Yau [SY79b] states that $X^n \# T^n$ does not admit a metric of positive scalar curvature, which would give the required contradiction. However, if M has arbitrary ends, then $X^n \# T^n$ is noncompact and the proof of [SY79b] breaks down in a serious manner. Nevertheless, using Gromov's μ-bubble technique, Chodosh and Li were able to extend Schoen and Yau's result to the complete noncompact case:

Theorem 9.3.6 (Chodosh–Li [CL24])**.** *Let $3 \leq n \leq 7$ and let X^n be a smooth manifold of dimension n. Then $X^n \# T^n$ does not admit complete a metric of positive scalar curvature.*

This gives the required contradiction.

9.3.2 Gromov's μ-bubble technique

A natural question in scalar curvature geometry is whether minimal hypersurface and Dirac operator arguments can be effectively localized around a particular geometric feature. This is particularly important when working in ambient spaces that are noncompact, incomplete, or contain boundaries. In the case of minimal hypersurfaces, minimizing sequences are susceptible to various problems: they can escape every compact set and fail to converge, they can degenerate to something noncompact and unwieldy, or they can hit points of incompleteness and become singular.

To get around these issues, Gromov introduced the technique of μ-bubbles as a way of localizing the Schoen–Yau minimal hypersurface descent scheme [SY79b] to stay within a well understood and usually compact region of the ambient space [Gro96; Gro18]. The basic idea is to replace stable minimal hypersurfaces by stable prescribed mean curvature hypersurfaces (called μ-bubbles), where the prescription function (called the *potential*) blows up in the "forbidden" region of the manifold. There is now a wealth of examples where μ-bubbles have been used to study problems in scalar curvature curvature which previously seemed out of reach [CL24; Gro23; Gro20; Zhu23; Zhu21; LUY21; CLSZ21].

To make the idea clear, we give the following definition.

Definition 9.3.7. Let (M, g) be a Riemannian manifold and let $h : M \to \overline{\mathbb{R}}$ (called the *potential*) be a continuous function with the following properties:

1. h is smooth on the open set $\{-\infty < h < \infty\}$ in M.

[8] Lohkamp's work [Loh99] was concerned with asymptotically flat manifolds with compact core. However, his arguments go through unchanged in the case of arbitrary ends once the density theorem is proved.

2. Let M_0 denote the closure of $\{-\infty < h < \infty\}$ in M. Then $\partial M_0 = \partial M_+ \cup \partial M_-$, where $\partial M_\pm$ are *nonempty* smooth closed embedded hypersurfaces and $h|_{\partial M_\pm} = \pm\infty$.

Let Ω_0 be a Caccioppoli set in M_0 which contains ∂M_+. For Ω another such Caccioppoli set, we define

$$\mathcal{F}(\Omega) \doteq \mathcal{H}^{n-1}_g(\partial^*\Omega) - \int_{M_0} (\chi_\Omega - \chi_{\Omega_0})\, h\, d\mu_g.$$

We say that Ω is a *μ-bubble* if it is a critical point of this functional under variations satisfying $\Omega \,\Delta\, \Omega_0 \subset\subset M_0$, or equivalently, if $\partial\Omega$ has prescribed mean curvature h (with the normal oriented pointing towards ∂M_-). *Stable* and *minimizing* μ-bubbles are defined in the obvious way.

It is relatively straightforward to show that in the above setting, stable μ-bubbles always exist.

Lemma 9.3.8. *If $2 \le n \le 7$, nonempty stable μ-bubbles exist and are smooth.*

Sketch of Proof. Let $\{\Omega_i\}$ be a minimizing sequence for $\mathcal{F}$. This sequence can be modified to avoid a neighborhood of ∂M_0 and still be minimizing. Since $\partial M_\pm$ are closed, smooth hypersurfaces, any foliation of a tubular neighborhood has uniformly bounded mean curvature. Therefore, $H < h$ near ∂M_+ (with the normal pointing away from ∂M_+) and $H > h$ near ∂M_- (with the normal pointing away from ∂M_-). These surfaces act as barriers and by performing a suitable replacement, for each i we can find a Caccioppoli set $\Omega_i' \subset \{|h| \le C\}$, where C is a sufficiently large constant independent of i, satisfying also $\Omega_i' \,\Delta\, \Omega_i \subset\subset M_0$ and $\mathcal{F}(\Omega_i') \le \mathcal{F}(\Omega_i)$. For details of this argument, we refer to [CL24].

Since M_0 is compact and h is bounded on $\Omega_i' \Delta \Omega_0$, $\mathcal{F}(\Omega_i') \gtrsim -1$. We obviously have $\mathbf{M}(\partial^*\Omega_i') \lesssim 1$ and $\mathbf{M}(\Omega_i') \lesssim 1$ by previous observations. So by the BV compactness theorem, there exists a subsequence of $\{\Omega_i'\}$ converging in the sense of Caccioppoli sets. By standard regularity theory, the limiting set will have smooth boundary with mean curvature h. As it is $\mathcal{F}$-minmizing, it is also automatically stable. □

The utility of stable μ-bubbles is explained by the following computation:

Lemma 9.3.9. *If Ω is a smooth stable μ-bubble with boundary Σ, then $\Sigma = \partial\Omega$ satisfies the stability inequality*

$$\int_\Sigma |\nabla\varphi|^2 + \tfrac{1}{2} R_\Sigma \varphi^2 - \tfrac{1}{2}\left(R_g + \tfrac{n}{n-1} h^2 + 2\nu(h)\right)\varphi^2\, d\mu_\Sigma \ge 0, \tag{9.3.2}$$

for every $\varphi \in C^1(\Sigma)$, where ν is the unit normal pointing away from $\partial_+ M$.

Proof. The second variation of the μ-bubble functional is given by (see [LUY21, Proposition 2.3])

$$\int_\Sigma |\nabla\varphi|^2 - \tfrac{1}{2}\left(R_g - R_\Sigma + |A|^2\right)\varphi^2 - \tfrac{1}{2}\left(h^2 + 2\nu(h)\right)\varphi^2\, d\mu_\Sigma \geq 0. \tag{9.3.3}$$

We then write $|A|^2 = \frac{1}{n}H^2 + |\mathring{A}|^2 = \frac{1}{n}h^2 + |\mathring{A}|^2$, insert this into the second variation, and rearrange to obtain (9.3.2). □

Recall that a closed Riemannian manifold (M^n, g) is *Yamabe positive (nonnegative)* if it contains a positive (nonnegative) scalar curvature metric in its conformal class. Alternatively, (M^n, g) is Yamabe positive (nonnegative) if

$$\int_M \varphi L\varphi > 0 \quad (\geq 0) \tag{9.3.4}$$

for every $\varphi \in C^1(M)$, where L is the *conformal Laplacian*

$$L \doteq -4\frac{n-1}{n-2}\Delta_g + R_g. \tag{9.3.5}$$

From the stability inequality (9.3.2), we see that if

$$R_g + \tfrac{n}{n-1}h^2 - 2|\nabla h| > 0 \quad (\geq 0), \tag{$\star$}$$

then Σ is Yamabe positive (nonnegative). We will refer to this inequality as condition ($\star$).[9] Therefore, μ-bubbles may be used in place of stable minimal hypersurfaces to study scalar curvature as long as one can construct suitable potential functions. This is exactly how Chodosh and Li proved Theorem 9.3.6. One of the innovations of [LLU23] is an interpretation of the μ-bubble technique in terms of *marginally outer trapped surfaces* (MOTS). See already Section 11.5 and "Proof 2" in Section 9.3.3 below.

9.3.3 Outline of the proof of Theorem 9.3.3

We now explain two different proofs of Theorem 9.3.3 based on the μ-bubble idea.

Proof 1. Assume for sake of contradiction that $m_{\mathrm{ADM}} < 0$. By applying a density theorem, one can assume harmonic asymptotics on $\mathcal{E}$. Then $m_{\mathrm{ADM}} < 0$ implies large cylinders in the asymptotic region are mean-convex, and one can apply a "Plateau problem" version of the μ-bubble technique to construct large stable μ-bubbles spanning cross-sectional spheres of these cylinders. Arguing

[9] In [LUY21], the condition is written as $R_M + h^2 - 2|\nabla h| > 0$. This implies the current condition since $\frac{n}{n-1} > 1$. The difference comes from keeping the trace part of $|A|^2$ versus just throwing it away.

similarly to [Sch89], one constructs a complete "strongly stable" μ-bubble which can be conformally deformed to produce a counterexample to the positive mass theorem in dimension $n-1$. One then argues by induction on dimension, with the $n = 3$ case following from the Gauss–Bonnet theorem and the logarithmic cutoff trick.

Remark 9.3.10. In [LUY21], we did not have a density theorem available for asymptotically flat manifolds with arbitrary ends. Therefore, we were not able to obtain Theorem 9.3.3 with the most general assumptions on $\mathcal{E}$. It turns out that the argument just described works directly for *asymptotically Schwarzschild* manifolds with a fair bit of work, without a density theorem.

Proof 2. Using the same μ-bubble potential function h as in Proof 1, define an "auxiliary second fundamental form" k_h by

$$k_h \doteq -\frac{h}{n-1}g. \tag{9.3.6}$$

It follows from a very short computation that (M', g, k_h) satisfies the hypotheses of Theorem 9.2.2 where $M' \doteq \{|h| \leq C\}$ for C chosen sufficiently large. Therefore $m_{\mathrm{ADM}} > 0$.

This very short and surprising proof was found in [LLU23]. We give the details in Section 11.5.1 below.

9.4 The Liouville theorem for locally conformally flat manifolds

Let (M, g) and (N, h) be two Riemannian manifolds of the same dimension. A smooth map $\varphi : M \to N$ is said to be *conformal* if there exists a smooth positive function u on M such that $\varphi^* h = ug$. It is easy to see that any conformal map is an immersion. If φ is in addition a diffeomorphism, we say that it is a conformal diffeomorphism. An n-dimensional Riemannian manifold (M, g) is said to be *locally conformally flat* (LCF) if is it locally conformally diffeomorphic to S^n with the round metric.

A fundamental theorem of Kuiper states that any LCF manifold M^n ($n \geq 3$) can be conformally immersed in S^n [Kui49]. A conformal immersion $\Phi : M^n \to S^n$ is called a *developing map*. In [SY88], Schoen and Yau studied the question of when Φ is injective, as this has implications for the higher homotopy groups of LCF manifolds. By combining the work of Schoen–Yau in [SY88], the scalar curvature rigidity theorem of Chodosh–Li (Theorem 9.3.6), and a Lohkamp-type compactification argument, we were able to prove the following theorem in joint work with Lesourd and Yau [LUY20].

Theorem 9.4.1 (Liouville theorem for LCF manifolds)**.** *Let (M^n, g), $n \geq 3$, be a complete locally conformally flat Riemannian manifold with with nonnegative scalar curvature. If $\Phi : M^n \to S^n$ is a conformal map, then Φ is injective and $\partial\Phi(M)$ has zero capacity.*

This theorem was proved in [SY88] with various additional assumptions (either assuming $n \geq 7$ or some global geometric restrictions). The original idea in [SY88] for proving this most general form of the Liouville theorem was to reduce it to the positive mass theorem. However, when M^n is noncompact, this requires presisely the positive mass theorem with arbitrary ends. Therefore, given our Theorem 9.3.1, we are able to carry out the original idea of [SY88] in one stroke. The details of this argument will be given in Section 11.6 below.

9.A Positive scalar curvature on noncompact surfaces

In this appendix, we prove the $n = 2$ version of Theorem 9.3.6, which has a cute and elementary proof. Since the 2-torus has vanishing Euler characteristic, the Gauss–Bonnet theorem immediately implies:

Proposition 9.A.1. *The torus T^2 does not admit a metric of positive scalar curvature.*

One can alternatively argue using closed geodesics (which incidentally motivates the stable minimal hypersurface technique for studying positive scalar curvature). Indeed, the torus contains a nontrivial free homotopy class of curves $\mathcal{L}$. Using the Arzela–Ascoli theorem, we may find a closed geodesic $\gamma \in \mathcal{L}$ of minimal length. The second variation formula implies this geodesic is unstable because of positive curvature, which contradicts its minimality.

The compactness of the torus is used in two places. Firstly, to apply the Arzela–Ascoli theorem to maps $S^1 \to T^2$ and secondly, to ensure the positivity of the convexity radius of (T^2, g) which is used to conclude that the minimizer is closed. However, the argument actually goes through if we know that the "relevant" curves all lie in some large ball.

Theorem 9.A.2. *Let X^2 be a smooth surface. Then $X^2 \# T^2$ does not admit a complete metric of positive scalar curvature.*

Proof. Suppose g is a complete metric on $X^2 \# T^2$ with positive scalar curvature. Consider a free homotopy class $\mathcal{L}$ corresponding to going once around one of the circular factors in T^2. Let $C \subset T^2 \# X^2$ be a "surgery circle" associated to the connected sum. Then $C \notin \mathcal{L}$. In T^2, C bounds an open disk. Call the complement of this disk $K \subset X^2 \# T^2$. Let $\{\gamma_i\} \subset \mathcal{L}$ be a length-minimizing sequence of smooth curves. For each i, $\gamma_i \cap K \neq \emptyset$ for otherwise γ_i would be nullhomotopic in T^2, contradicting the choice of $\mathcal{L}$. Let $\ell = \sup_i L(\gamma_i) < \infty$. Then by the triangle inequality, γ_i is contained in the 2ℓ-neighborhood of K for i sufficiently large. Since g is complete, this neighborhood is relatively compact. Now a subsequence of $\{\gamma_i\}$ tends to a closed length-minimizing geodesic which can be shown to be unstable by the second variation formula, and we have a contradiction. □

The topology of $X^2 \# T^2$ prevents the minimizing sequence from "escaping off to infinity"—the compact set K *anchors* $\mathcal{L}$. In fact, the completeness of g prevented the curves from entering the noncompact end at all. This is much stronger than what can be expected for minimal hypersurfaces and in general we will have degeneration of topology at infinity.

Chapter 10

The spacetime positive mass theorem for black holes

10.1 The density theorem for initial data sets

When discussing initial data sets (M^n, g, k), it is often convenient to replace the second fundamental form k_{ij} by the *conjugate momentum tensor*

$$\pi^{ij} = k^{ij} - (\operatorname{tr}_g k)g^{ij},$$

and we will also refer to (M, g, π) as an initial data set. Then the formulas for μ and J become

$$\mu = \frac{1}{2}\left(R_g + \frac{1}{n-1}(\operatorname{tr}_g \pi)^2 - |\pi|_g^2\right)$$
$$J = \operatorname{div}_g \pi.$$

The first density theorem for initial data sets was proved by Corvino and Schoen [CS06], who showed that asymptotically flat vacuum initial data sets (g, π) can be approximated by vacuum initial data satisfying the *harmonic asymptotics* condition

$$\tilde{g}_{ij} = u^4\delta_{ij}, \qquad \tilde{\pi}^{ij} = u^{-6}\mathfrak{L}_\delta Y^{ij} \tag{10.1.1}$$

outside a compact set, for some function u and vector field Y which have good asymptotic expansions. The notation $\mathfrak{L}$ is defined by the formula $\mathfrak{L}_g Y^{ij} \doteq (\mathcal{L}_Y g)^{ij} - (\operatorname{div}_g Y)g^{ij}$ for an arbitrary metric g, where $\mathcal{L}_Y$ denotes the Lie derivative. In the extension of the positive mass theorem to initial data sets in higher dimensions [EHLS16], Eichmair, Huang, Lee, and Schoen proved a harmonic asymptotics density theorem for initial data sets satisfying the dominant energy condition. This theorem also plays a crucial role in Eichmair's Jang reduction and rigidity theorems [Eic13], as well as Lohkamp's compactification approach to the spacetime positive mass theorem [Loh17].

In this chapter, we generalize the density theorems of Corvino–Schoen [CS06] and Eichmair–Huang–Lee–Schoen [EHLS16] to allow for initial data sets with compact boundary.

Theorem 10.1.1 (Density theorem for initial data sets with boundary)**.** *Let (M^n, g, k) be a complete asymptotically flat initial data set with compact boundary ∂M, such that the dominant energy condition, $\mu \geq |J|_g$, holds on M. Let $p > n$ and $\frac{n-2}{2} < q < n-2$ such that q is less than the decay rate of (g, k). Let θ^+ denote the outer null expansion of ∂M.*

Then for any $\varepsilon > 0$, there exists an asymptotically flat initial data set $(\tilde{g}, \tilde{k})$ on M also satsifying the dominant energy condition such that $(\tilde{g}, \tilde{k})$ has harmonic asymptotics in each end of M, $(\tilde{g}, \tilde{k})$ is ε-close to (g, k) in $W^{2,p}_{-q} \times W^{1,p}_{-q-1}$, the new constraints $(\tilde{\mu}, \tilde{J})$ are ε-close to (μ, J) in L^1, and the new outer null expansion $\tilde{\theta}^+$ is equal to *θ^+ on ∂M.*

Furthermore, we can choose $(\tilde{g}, \tilde{k})$ such that the strict *dominant energy condition holds, $\tilde{\mu} > |\tilde{J}|_g$. Simultaneously, $(\tilde{\mu}, \tilde{J})$ may be chosen to decay as fast as we like.*

Alternatively, we can choose $(\tilde{g}, \tilde{k})$ to be vacuum (that is, $\tilde{\mu} = |\tilde{J}|_g = 0$) outside a compact set. Moreover, if (g, k) is vacuum everywhere, then $(\tilde{g}, \tilde{k})$ can be chosen to be vacuum everywhere.

Remark 10.1.2. More generally, we may prescribe $\tilde{\theta}^+$ to be any function sufficiently close to θ^+ in the fractional Sobolev space $W^{1-\frac{1}{p},p}(\partial M)$. This theorem is more precisely stated as Theorem 10.3.7 below. In particular, $\tilde{\theta}^+$ may be chosen to be strictly less than θ^+ at every point. Moreover, we may alternatively choose to prescribe the inner expansion θ^- (instead of θ^+) on any given components of ∂M.

The proof of the density theorem is given in Section 10.3 and the proof of the spacetime positive mass theorem with boundary is given in Section 10.4.

10.2 Preliminaries

10.2.1 Notation and definitions

Let M^n be a smooth n-dimensional manifold ($n \geq 3$) with compact boundary ∂M, and fix a smooth background metric $\overline{g}$ which is identically Euclidean on the finitely many noncompact ends of M, which are all diffeomorphic to $\mathbb{R}^n$ minus a ball.

Let $\mathcal{E}$ denote such an end with coordinates x^i. For $p \geq 1$ and $s \in \mathbb{R}$, we define a weighted L^p norm by

$$\|u\|_{L^p_s(\mathcal{E})} \doteq \left(\int_{\mathcal{E}} ||x|^{-s} u|^p \frac{dx}{|x|^n} \right)^{1/p}.$$

For $k \in \mathbb{N}$, weighted Sobolev norms on $\mathcal{E}$ are defined by

$$\|u\|_{W^{k,p}_s(\mathcal{E})} \doteq \sum_{i=0}^{k} \|\partial^i u\|_{L^p_{s-i}(\mathcal{E})}.$$

Weighted C^k norms are defined by

$$\|u\|_{C^k_s(\mathcal{E})} \doteq \sum_{i=0}^{k} \sup_{\mathcal{E}} ||x|^{i-s} \partial^i u|.$$

For $\alpha \in (0,1)$, weighted C^α seminorms are defined by

$$[\partial^k u]_{C^\alpha_s(\mathcal{E})} \doteq \sup_{x \in \mathcal{E}} \left(|x|^{k+\alpha-s} \sup_{y \in \mathcal{E}} \frac{|\partial^k u(x) - \partial^k u(y)|}{|x-y|^\alpha} \right).$$

Let $\mathcal{E}_1, \ldots, \mathcal{E}_N$ denote the collection of Euclidean ends of M and let K denote a compact set such that $M \setminus K = \bigcup_{j=1}^N \mathcal{E}_j$. The weighted norms on $(M, \overline{g})$ are defined by:

$$\|u\|_{W^{k,p}_s(M)} \doteq \|u\|_{W^{k,p}_s(\mathcal{E})} + \|u\|_{W^{k,p}(K)},$$
$$\|u\|_{C^k_s(M)} \doteq \|u\|_{C^k_s(\mathcal{E})} + \|u\|_{C^k(K)},$$
$$\|u\|_{C^{k,\alpha}_s(M)} \doteq \|u\|_{C^k_s(M)} + [\partial^k u]_{C^\alpha_s(\mathcal{E})}$$

and the corresponding spaces are defined to the corresponding collections of functions (or tensor fields) for which the norms are well-defined and finite. See [Lee19] for more details.

Definition 10.2.1. We say that an initial data set (M^n, g, k) is *asymptotically flat* if (g, k) is locally $C^{2,\alpha} \times C^{1,\alpha}$ for some $0 < \alpha < 1$, and there exists a compact set $K \subset M$ such that $M \setminus K$ is a disjoint union of Euclidean ends such that in the associated coordinate charts,

$$g_{ij}(x) = \delta_{ij} + O_2(|x|^{-q}) \tag{10.2.1}$$

$$k_{ij} = O_1(|x|^{-q-1}) \tag{10.2.2}$$

for some $q > \frac{n-2}{2}$, and also $(\mu, J) \in L^1(M)$. We refer to this q as the asymptotic decay rate of (g, k).

In this case, the ADM energy-momentum (E, P) is well-defined. We refer the reader to [Lee19] for details and references.

Throughout most of this chapter, we will use π in place of k. Given a fixed manifold M, we define the *constraint map* Φ by

$$\Phi(g, \pi) \doteq (2\mu, J),$$

for any initial data (g, π) on M.

Given a hypersurface Σ with unit normal ν in an initial data set (M, g, k), we define the *outer and inner null expansions* θ^+_Σ and θ^-_Σ, respectively, with respect to (g, π) by

$$\theta^\pm_\Sigma \doteq \pm H_\Sigma + \operatorname{tr}_\Sigma k,$$

where H_Σ is the mean curvature of ∂M with respect to g and ν, and

$$\operatorname{tr}_\Sigma k \doteq (g^{ij} - \nu^i \nu^j) k_{ij} = -\pi^{ij} \nu_i \nu_j$$

is the trace of k over $T\Sigma$. In the case when (M, g, k) sits inside a spacetime, $\theta^\pm_\Sigma$ can be interpreted in terms of Lorentzian geometry, but we shall not need this viewpoint here. In this chapter we will always choose Σ to be ∂M, and we choose ν to be the unit normal pointing into M. We will want to prescribe either $\theta^+_{\partial M}$ or $\theta^-_{\partial M}$ on each boundary component, so we make the following definition.

Definition 10.2.2. Let M be a fixed manifold with boundary, and let $\partial^+ M$ and $\partial^- M$ designate unions of components of ∂M such that $\partial M = \partial^+ M \cup \partial^- M$. Given initial data (g, π) on M, define $\Theta(g, \pi)$ to be the function ∂M that is equal to $\theta^\pm_{\partial M}$ on $\partial^\pm M$ with respect to the data (g, π) and the normal pointing into M.

For PDE purposes, it is convenient to slightly enlarge the space of data sets under consideration. We will consider initial data (g, π), where $g - \overline{g} \in W^{2,p}_{-q}(T^*M \odot T^*M)$ and $\pi \in W^{1,p}_{-q-1}(TM \odot TM)$, where $p > n$, $\frac{n-2}{2} < q < n-2$, and q is less than the decay rate in Definition 10.2.1. Note that such a pair (g, π) need not satisfy our definition of asymptotic flatness, and in particular, need not have well-defined ADM energy-momentum. We define

$$\mathcal{D} \doteq \left(\overline{g} + W^{2,p}_{-q}(T^*M \odot T^*M)\right) \times W^{1,p}_{-q-1}(TM \odot TM), \tag{10.2.3}$$

so that $\mathcal{D}$ is a (affine) Banach space of initial data sets. Note that the tangent space of $\mathcal{D}$ at (g, π),

$T_{(g,\pi)}\mathcal{D}$, can be identified with $W^{2,p}_{-q}(T^*M \odot T^*M) \times W^{1,p}_{-q-1}(TM \odot TM)$.

Lemma 10.2.3. *Let $p > n$ and $\frac{n-2}{2} < q < n-2$. On a fixed asymptotically flat manifold M^n with compact boundary decomposed as in Definition 10.2.2, the descriptions of Φ and Θ given above define a smooth map of Banach spaces*

$$(\Phi, \Theta) : \mathcal{D} \to \mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M),$$

where

$$\mathcal{L} \doteq L^p_{-q-2}(M) \times L^p_{-q-2}(TM), \tag{10.2.4}$$

and $W^{1-\frac{1}{p},p}(\partial M)$ is a fractional Sobolev space on ∂M. (See, for example, [Gri85, Section 1.4].)

Proof. The claim about Φ is standard and easy to verify, so we focus on the map Θ. We can realize ∂M as a level set of a smooth function f, which has no critical points in a small neighborhood U of ∂M. Then the formula

$$\nu^i \doteq g^{ij} \frac{\partial_i f}{|\nabla f|_g}$$

defines a vector field on U which is the unit normal to the level sets of f in U, and it is has $W^{2,p}$ regularity. Next, the formula

$$H_g \doteq (g^{ij} - \nu^j \nu^j) \frac{1}{|\nabla f|_g} (\partial_{ij} f - \Gamma^k_{ij} \partial_k f),$$

defines a function on U which is equal to the mean curvature of the level sets of f in U. We can also see that H_g has $W^{1,p}$ regularity since $\Gamma^k_{ij} \in W^{1,p}$ and $W^{1,p}$ is a Banach algebra. Similary, the quantity $-\pi_{ij}\nu^i\nu^j$ is a $W^{1,p}$ function on U. More precisely, we observe that we have a bounded map from from $(g, \pi) \in \mathcal{D}$ to $\pm H_g - \pi_{ij}\nu^i\nu^j \in W^{1,p}(U)$. The result follows from viewing Θ as the composition of this map with the usual bounded trace operator from $W^{1,p}(U)$ to $W^{1-\frac{1}{p},p}(\partial M)$. □

The proof above made use of a trace theorem. Later on, we will need the following *sharp trace theorem.*

Lemma 10.2.4. *Let $(M^n, \overline{g})$ be as above and let $g \in \overline{g} + W^{2,p}_{-q}(T^*M \odot T^*M)$. Then, for any $s \in \mathbb{R}$, the weighted Sobolev space $W^{2,p}_s(M)$ enjoys a bounded trace operator*

$$T^g_2 : W^{2,p}_s(M) \to W^{2-\frac{1}{p},p}(\partial M) \times W^{1-\frac{1}{p},p}(\partial M)$$

which is the unique extension of

$$u \longmapsto \left(u|_{\partial M}, \left.\frac{\partial u}{\partial \nu_g}\right|_{\partial M} \right)$$

for $u \in C^2(M) \cap W^{2,p}_s(M)$*. The mapping* T_2^g *is surjective and admits a bounded right inverse. The operator norms of* T_2^g *and its right inverse depend only on* $\|g - \overline{g}\|_{W^{2,p}_{-q}}$*.*

We emphasize that the normal vector field ν_g is the one corresponding to the metric g.

Proof. The existence and properties of T_2^g are easily reduced to the case of bounded domains [Gri85, Theorem 1.5.1.2] by means of cutoff functions. In particular, we may take elements in the image of the right inverse to be supported in a neighborhood of ∂M. □

10.2.2 "Conformal" initial data sets

Conformal transformations play a special role in the study of mass and the Riemannian positive mass theorem. The following notion of conformal transformations of initial data sets plays a crucial role in the density theorem and the positive mass theorem [CS06; EHLS16].

Let

$$\mathcal{C} \doteq \left(1 + W^{2,p}_{-q}(M)\right) \times W^{2,p}_{-q}(TM) \tag{10.2.5}$$

denote the (affine) Banach space of conformal deformations. Note that the tangent space of $\mathcal{C}$ at $(1,0)$, $T_{(1,0)}\mathcal{C}$, can be identified with $W^{2,p}_{-q}(M) \times W^{2,p}_{-q}(TM)$. For $(g,\pi) \in \mathcal{D}$ fixed and $(u,Y) \in \mathcal{C}$, we define

$$\begin{aligned} \tilde{g} &\doteq u^s g, \\ \tilde{\pi} &\doteq u^{-\frac{3}{2}s}(\pi + \mathfrak{L}_g Y), \end{aligned}$$

where $s = \frac{4}{n-2}$ is the conformal exponent and $\mathfrak{L}_g Y$ was defined in the introduction. We denote

$$\begin{aligned} \Psi_{(g,\pi)} &: \mathcal{C} \to \mathcal{D} \\ &(u,Y) \mapsto (\tilde{g}, \tilde{\pi}). \end{aligned}$$

Definition 10.2.5. Let (M^n, g, π) be an asymptotically flat initial data set. We say that (M, g, π) has *harmonic asymptotics* in a particular end if in the asymptotically flat coordinates, the initial data takes the form

$$(g,\pi) = \Psi_{(\overline{g},0)}(u,Y)$$

outside a compact set, where u and Y are a function and vector field pair satisfying

$$u(x) = 1 + a|x|^{2-n} + O_{2,\alpha}(|x|^{1-n}) \tag{10.2.6}$$

$$Y^i(x) = b_i|x|^{2-n} + O_{2,\alpha}(|x|^{1-n}), \tag{10.2.7}$$

for some $\alpha \in (0,1)$. When the initial data set is of this form, the ADM energy-momentum has a particularly simple expression: $E = 2a$ and $P_i = -\frac{n-2}{n-1}b_i$.

Initial data sets in the image of $\Psi_{(\overline{g},0)}$ have harmonic asymptotics if the constraints decay quickly enough:

Lemma 10.2.6 ([EHLS16, Proposition 24])**.** *Suppose there exist $(u,Y) \in \mathcal{C}$ such that $(g,\pi) = \Psi_{(\overline{g},0)}(u,Y)$ outside a compact set. If $(\mu, J) \in C^{0,\alpha}_{-n-1-\delta}$ for some $\delta > 0$, then u and Y admit the expansions* (10.2.6) *and* (10.2.7)*. Hence (g,π) has harmonic asymptotics.*

Next we define

$$\mathcal{P} \doteq (T, \Upsilon) \doteq (\Phi, \Theta) \circ \Psi_{(g,\pi)} : \mathcal{C} \to \mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M). \tag{10.2.8}$$

In the following, we let (μ, J, θ) be the value of (Φ, Θ) on the fixed data set (g,π).

Proposition 10.2.7. *The map $\mathcal{P}$ is a smooth map of Banach spaces and is explicitly given by*

$$T(u,Y) = \left(u^{-s}\left[\frac{L_g u}{u} + \tfrac{1}{n-1}(\operatorname{tr}_g \pi + \operatorname{tr}_g \mathfrak{L}_g Y)^2 - (|\pi|_g^2 + 2\langle \mathfrak{L}_g Y, \pi\rangle + |\mathfrak{L}_g Y|_g^2)\right],\right. \tag{10.2.9}$$
$$\left. u^{-\frac{3}{2}s}\left[(\operatorname{div}_g \mathfrak{L}_g Y + \operatorname{div}_g \pi)^i + \tfrac{s(n-1)}{2}(\pi + \mathfrak{L}_g Y)^{ij}\frac{\nabla_j u}{u} - \tfrac{s}{2}\operatorname{tr}_g(\pi + \mathfrak{L}_g Y) g^{ij}\frac{\nabla_j u}{u}\right]\right)$$
$$\Upsilon(u,Y) = u^{-\frac{s}{2}}\left(\theta \pm \tfrac{s(n-1)}{2}\frac{\partial}{\partial \nu}(\log u) + \operatorname{div}_{\partial M} Y^\top + H\langle Y, \nu\rangle - \langle \nabla_\nu Y^\perp, \nu\rangle + \langle Y^\top, \nabla_\nu \nu\rangle\right), \tag{10.2.10}$$

where the $\pm$ depends on whether the point lies in $\partial^\pm M$. Here ν is any extension of the g-unit normal vector field of ∂M (pointing into M), $Y^\perp = \langle Y, \nu\rangle \nu$, and $Y^\top = Y - Y^\perp$. Note that the quantities $\nabla_\nu Y^\perp$ and $\nabla_\nu \nu$ depend on the particular extension chosen, but $-\langle \nabla_\nu Y^\perp, \nu\rangle + \langle Y^\top, \nabla_\nu \nu\rangle$ is an invariant quantity. The linearization of $\mathcal{P}$ at $(1,0)$ is given by

$$D\mathcal{P}_{(1,0)}(v,Z) = (DT|_{(1,0)}, D\Upsilon|_{(1,0)})(v,Z),$$

where

$$DT|_{(1,0)}(v,Z) = \Big(-s(n-1)\Delta_g v + \tfrac{2}{n-1}(\operatorname{tr}_g \pi)(\operatorname{div}_g Z) - 4\,\pi\cdot\nabla_g Z - 2s\mu v, \tag{10.2.11}$$
$$(\operatorname{div}_g \mathfrak{L}_g Z)^i + \tfrac{s(n-1)}{2}\pi^{ij}\nabla_j v - \tfrac{s}{2}(\operatorname{tr}_g \pi) g^{ij}\nabla_j v - \tfrac{3}{2} s J^i v\Big),$$
$$D\Upsilon|_{(1,0)}(v,Z) = -\tfrac{s}{2}\theta v \pm \tfrac{s(n-1)}{2}\frac{\partial v}{\partial \nu} + \operatorname{div}_{\partial M} Z^\top + H\langle Z,\nu\rangle - \langle \nabla_\nu Z^\perp, \nu\rangle + \langle Z^\top, \nabla_\nu \nu\rangle. \tag{10.2.12}$$

Proof. The formula (10.2.9) is given in the erratum for Exercise 9.7 in [Lee19]. To prove (10.2.10), we first use the standard formula

$$\tilde{H} = u^{-\frac{s}{2}}\left(H + \tfrac{s(n-1)}{2}\frac{\partial}{\partial \nu}(\log u)\right). \tag{10.2.13}$$

Using $\tilde{\nu} = u^{-\frac{s}{2}}\nu$, we compute

$$\tilde{\pi}(\tilde{\nu},\tilde{\nu}) = u^{-\frac{s}{2}}\left(\pi(\nu,\nu) + \mathfrak{L}_g Y(\nu,\nu)\right). \tag{10.2.14}$$

Finally, extend ν off of ∂M, let $Y^\perp = \langle \nu, Y\rangle \nu$ and $Y^\top = Y - Y^\perp$. Then, on ∂M,

$$\begin{aligned}
\mathfrak{L}_g Y(\nu,\nu) &= 2\langle \nabla_\nu Y, \nu\rangle - \operatorname{div}_g Y \\
&= \langle \nabla_\nu Y, \nu\rangle - \operatorname{div}_{\partial M} Y \\
&= -\langle Y^\top, \nabla_\nu \nu\rangle + \langle \nabla_\nu Y^\perp, \nu\rangle - \operatorname{div}_{\partial M} Y^\top - H\langle Y, \nu\rangle.
\end{aligned}$$

Combining these computations gives (10.2.10), and the linearizations are then computed in the obvious way. ☐

Using the formulas for the linearization and the sharp trace theorem, we can prove the following crucial result.

Lemma 10.2.8. *The maps*

$$D\Upsilon|_{(1,0)} : T_{(1,0)}\mathcal{C} \to W^{1-\frac{1}{p},p}(\partial M)$$

and

$$D\Theta|_{(g,\pi)} : T_{(g,\pi)}\mathcal{D} \to W^{1-\frac{1}{p},p}(\partial M)$$

are surjective and their kernels split.

Proof. From the formula for $D\Upsilon|_{(1,0)}$, we see that we want to solve the equation

$$-\tfrac{s}{2}\theta v \pm \tfrac{s(n-1)}{2}\frac{\partial v}{\partial \nu} + \operatorname{div}_{\partial M} Z^{\top} + H\langle Z, \nu\rangle - \langle \nabla_\nu Z^{\perp}, \nu\rangle + \langle Z^{\top}, \nabla_\nu \nu\rangle = f,$$

for any given $f \in W^{1-\frac{1}{p},p}(\partial M)$. We set $Z \equiv 0$, reducing this to

$$-\tfrac{s}{2}\theta v \pm \tfrac{s(n-1)}{2}\frac{\partial v}{\partial \nu} = f \quad \text{on } \partial M.$$

By the sharp trace theorem (Lemma 10.2.4), we can find a $v \in W^{2,p}_{-q}(M)$ such that

$$\left(v|_{\partial M}, \left.\frac{\partial v}{\partial \nu}\right|_{\partial M} \right) = \left(0, \frac{\pm 2f}{s(n-1)} \right)$$

and

$$\|v\|_{W^{2,p}_{-q}(M)} \leq C\|f\|_{W^{1-\frac{1}{p},p}(\partial M)}.$$

Therefore, $D\Upsilon|_{(1,0)}(v,0) = f$ as desired, with an estimate, which proves $D\Upsilon|_{(1,0)}$ has a bounded right inverse. By standard functional analysis [Bre11, Theorem 2.12], this implies that the kernel splits.

The corresponding statement for $D\Theta|_{(g,\pi)}$ follows by choosing first-order deformations

$$(h, w) = D\Psi_{(g,\pi)}|_{(1,0)}(v, 0) = \left(svg, -\tfrac{3}{2}sv\pi\right).$$

By the definition of Υ and the chain rule,

$$D\Theta|_{(g,\pi)}(h, w) = f$$

and

$$\|(h,w)\|_{T_{(g,\pi)}\mathcal{D}} \leq C\|v\|_{W^{2,p}_{-q}(M)} \leq C\|f\|_{W^{1-\frac{1}{p},p}(\partial M)}.$$

The same functional analysis argument as before completes the proof. □

Unfortunately, $D\mathcal{P}|_{(1,0)}$ does not define an elliptic boundary value problem. This is evident from the fact that the boundary operator $D\Upsilon|_{(1,0)}$ is a scalar operator while $DT|_{(1,0)}$ describes an elliptic system of $n+1$ equations. Therefore, we introduce boundary operators describing $n+1$ equations on the boundary:

$$
\begin{aligned}
B_1(u,Y) &\doteq u^{-\frac{s}{2}}\left[\theta \pm \tfrac{s(n-1)}{2}\frac{\partial}{\partial\nu}(\log u)\right],\\
B_2(u,Y) &\doteq Y^\top,\\
B_3(u,Y) &\doteq H\langle Y,\nu\rangle - \langle \nabla_\nu Y^\perp,\nu\rangle + \langle Y^\top,\nabla_\nu \nu\rangle,
\end{aligned}
$$

where H is the mean curvature of ∂M with respect to g. Altogether, these define a map

$$
B=(B_1,B_2,B_3):\mathcal{C}\to W^{1-\frac{1}{p},p}(\partial M)\times W^{2-\frac{1}{p},p}(T(\partial M))\times W^{1-\frac{1}{p},p}(\partial M). \tag{10.2.15}
$$

Clearly, we have

$$
\Upsilon = B_1 + u^{-\frac{s}{2}}\operatorname{div}_{\partial M} B_2 + u^{-\frac{s}{2}} B_3
$$

on $\mathcal{C}$. It follows that

$$
D\Upsilon|_{(1,0)} = DB_1|_{(1,0)} + \operatorname{div}_{\partial M} DB_2|_{(1,0)} + DB_3|_{(1,0)} \tag{10.2.16}
$$

on $T_{(1,0)}\mathcal{C}$, where

$$
\begin{aligned}
DB_1|_{(1,0)}(v,Z) &= -\tfrac{s}{2}\theta v \pm \tfrac{s(n-1)}{2}\frac{\partial v}{\partial\nu}\\
DB_2|_{(1,0)}(v,Z) &= Z^\top\\
DB_3|_{(1,0)}(v,Z) &= H\langle Z,\nu\rangle - \langle \nabla_\nu Z^\perp,\nu\rangle + \langle Z^\top,\nabla_\nu\nu\rangle.
\end{aligned}
$$

For ease of reading, we will remove $|_{(1,0)}$ when there is no risk for confusion.

Proposition 10.2.9. $(DT|_{(1,0)}, DB|_{(1,0)})$ *defines an elliptic system in the following sense:*

1. *There exists a relatively compact set $U \subset M$ so that the elliptic estimate*

$$
\|(v,Z)\|_{W^{2,p}_{-q}} \lesssim \|DT(v,Z)\|_{\mathcal{L}} + \|DB(v,Z)\|_{W^{1-\frac{1}{p},p}\times W^{2-\frac{1}{p},p}\times W^{1-\frac{1}{p},p}} + \|(v,Z)\|_{L^p(U)}
$$

holds for every $(v,Z)\in T_{(1,0)}\mathcal{C}$, where $\mathcal{L}$ and $\mathcal{C}$ were defined in (10.2.4) *and* (10.2.5)*; and*

2. *the mapping*

$$
(DT,DB): T_{(1,0)}\mathcal{C}\to \mathcal{L}\times W^{1-\frac{1}{p},p}(\partial M)\times W^{2-\frac{1}{p},p}(T(\partial M))\times W^{1-\frac{1}{p},p}(\partial M)
$$

is Fredholm.

The proof is unfortunately complicated by some technicalities. In order to apply the theory of elliptic systems, we have to check that the boundary operator DB is elliptic. However, we defined B relative to an orthogonal splitting of Y near the boundary. Hence, we must choose coordinates which respect this splitting. The most natural choice would be Fermi coordinates. However, since their construction involves solving the geodesic equation, it is well known that the resulting metric coefficients would only be $C^{0,\alpha}$ [DK81]. This causes problems for the regularity theory, but thankfully exact Fermi coordinates are not needed. Instead, we use the following:

Definition 10.2.10 (Andersson–Chruściel)**.** Let g be a $C^{k,\alpha}$ Riemannian metric on a manifold M with compact smooth boundary ∂M. Let $p \in \partial M$ and U be a neighborhood of p in M. We say that coordinates $(x^1, \dots, x^n) : U \to \mathbb{R}^n$ form an *almost-Fermi* coordinate system at p if

1. Each x^i is a $C^{k+1,\alpha}$ function on U relative to the smooth structure of M;

2. $x^1, \dots, x^{n-1}$ form a coordinate system for a neighborhood of p in ∂M when restricted to $x^n = 0$, consequently the coordinate partial derivatives $\partial_1, \dots, \partial_{n-1}$ are a frame for $T(\partial M)$ along ∂M; and

3. the "bottom row" of metric components satisfy $g_{nn}(x) = 1 + O((x^n)^{k+\alpha})$ and $g_{ni}(x) = O((x^n)^{k+\alpha})$ for $i = 1, \dots, n-1$.

The existence of almost-Fermi coordinate systems is proved in [AC96, Appendix B] (where they are called "almost Gaussian"). Note that the conclusions of (3) are proved directly there for the inverse metric, but can easily be seen to hold for the metric components themselves by Taylor expansion of the matrix inverse function. We note three more facts:

4. The metric components are $C^{k,\alpha}$ up to the boundary;

5. the coordinate vector field ∂_n agrees with the g-unit normal vector ν along ∂M; and

6. the coordinate vector field ∂_n satisfies $\nabla_{\partial_n} \partial_n = 0$ along ∂M.

Property (4) follows directly from (1) in Definition Definition 10.2.10. Property (5) follows from (2) and (3), because ∂_n has unit length on ∂M and is orthogonal to ∂M. Property (6) follows from the definition $\nabla_{\partial_n} \partial_n = \Gamma^k_{nn} \partial_k$ and the Christoffel symbols $\Gamma^1_{nn}, \dots, \Gamma^n_{nn}$ all vanish along ∂M by (3).

With this out of the way, we can prove the proposition.

Proof of Proposition 10.2.9. We compute the boundary operator DB in almost-Fermi coordinates $x^1, \ldots, x^n$. Let the extension of ν be ∂_n, so that $\nabla_\nu \nu = 0$ along ∂M. Then

$$\langle \nabla_\nu Z^\perp, \nu \rangle = \nabla_\nu \langle Z, \nu \rangle$$

along ∂M. We conclude that

$$\begin{aligned} DB_1(v, Z) &= -\tfrac{s}{2}\theta v \pm \tfrac{s(n-1)}{2} \frac{\partial v}{\partial x^n}, \\ DB_2(v, Z) &= \sum_{i=1}^{n-1} Z^i \partial_i, \\ DB_3(v, Z) &= HZ^n - \frac{\partial Z^n}{\partial x^n}. \end{aligned}$$

We furthermore observe that

$$\nabla_j \mathfrak{L}_g Z^{ij} = \Delta Z^i + \nabla_j \nabla^i Z^j - \nabla^i \nabla_k Z^k = \Delta Z^i + R^i{}_j Z^j, \tag{10.2.17}$$

so to leading order DT is diagonal and equal to the Laplacian in each component. Moreover, in these coordinates, up to leading order, DB is diagonal and gives a Dirichlet boundary condition in the DB_2 components, while giving a Neumann boundary condition in the DB_1 and DB_3 components. Therefore it is clear that (DT, DB) is properly elliptic in M and satisfies the complementary condition of Agmon–Douglis–Nirenberg on ∂M [ADN64]. Therefore we have elliptic boundary estimates in addition to interior estimates, which can now be combined with the asymptotic flatness assumption to obtain the global estimate (1) of Proposition 10.2.9 in routine way. Specifically, we use a partition of unity and a scaling argument to obtain a global weighted estimate (for example, see [Lee19, Theorem A.33]), and then a cutoff argument to replace $L^p_{-q}(M)$ by $L^p(U)$ on the right-hand side of the global weighted estimate (as in [Bar86, Theorem 1.10], or see [Lee19, Lemma A.41]).

If the coefficients of (DT, DB) were smooth, then the Fredholm property (2) of Proposition 10.2.9 would also follow, as in [LM85], from the fact that DB is an elliptic boundary condition for DT, which is asymptotic to the Laplacian in each component. To account for the lack of smoothness (the coefficients are $C^{1,\alpha}$ at worst), we adapt an argument of D. Maxwell [Max05]. Although the Fredholm property does not follow directly from (1), the elliptic estimate (1) combined with compactness of the map $W^{2,p}_{-q}(M) \to L^p(U)$ does show that the map (DT, DB) is *semi-Fredholm*[1] via standard

[1] A bounded linear operator $T : X \to Y$ is *semi-Fredholm* if $\dim \ker T < \infty$ and $T(X)$ is closed in Y.

arguments [Sch02, Theorem 5.21].

Standard smoothing arguments allow us to construct a continuous one-parameter family of initial data sets in $(g_\mu, \pi_\mu) \in \mathcal{D}$ for $\mu \in [0,1]$, such that $(g_\mu, \pi_\mu) \in C^\infty \times C^\infty$ for $\mu > 0$, and $(g_0, \pi_0) = (g, \pi)$. Since the associated operators (DT_μ, DB_μ) have smooth coefficients for $\mu > 0$, we know that they are Fredholm on the relevant Sobolev spaces [LM85]. Since the index of semi-Fredholm operators is a homotopy invariant [Sch02, Theorem 5.22], we have $\operatorname{ind}(DT_0, DB_0) = \operatorname{ind}(DT_\mu, DB_\mu)$ for any $\mu > 0$. Since the index for $\mu > 0$ is finite, this implies that $(DT, DB) = (DT_0, DB_0)$ is itself Fredholm. $\square$

10.3 Density theorems

10.3.1 Prescribed constraint density theorem

Our first density theorem generalizes the vacuum density theorem of Corvino–Schoen [CS06].

Theorem 10.3.1 (Density theorem for prescribed constraints)**.** *Let (M^n, g, π) be a complete asymptotically flat initial data set with constraints (μ, J) and compact boundary ∂M having outer null expansion θ on $\partial^+ M$ and inner null expansion θ on $\partial^- M$. Let $p > n$ and $\frac{n-2}{2} < q < n-2$ be strictly less than the decay rate of (g, π). Recall the definitions of $\mathcal{L}$ and $\mathcal{D}$ from* (10.2.4) *and* (10.2.3)*. There exist constants $\delta > 0$ and C so that the following is true:*

If $(\tilde\mu, \tilde J) \in \mathcal{L}$, $\tilde\theta \in W^{1-\frac{1}{p},p}(\partial M)$, and $\|(\tilde\mu, \tilde J, \tilde\theta) - (\mu, J, \theta)\|_{\mathcal{L}\times W^{1-\frac{1}{p},p}} < \delta$, then there exists an asymptotically flat initial data set $(\tilde g, \tilde\pi) \in \mathcal{D}$, whose constraints are $(\tilde\mu, \tilde J)$ and with outer/inner null expansion $\tilde\theta$ on $\partial^\pm M$, which satisfies

$$\|(\tilde g, \tilde\pi) - (g, \pi)\|_{\mathcal{D}} \leq C \|(\tilde\mu, \tilde J, \tilde\theta) - (\mu, J, \theta)\|_{\mathcal{L}\times W^{1-\frac{1}{p},p}}.$$

Furthermore, there exists $(u, Y) \in \mathcal{C}$ such that

$$(\tilde g, \tilde\pi) = \Psi_{(\overline{g},0)}(u, Y)$$

outside a compact set, where $(\overline{g}, 0)$ is the flat data set on the Euclidean end.

In particular, if $(\tilde\mu, \tilde J)$ is $C^{0,\alpha}_{-n-1-\varepsilon}$ up to the boundary, and $\tilde\theta \in C^{1,\alpha}(\partial M)$ for some $\alpha \in (0,1)$ and $\varepsilon > 0$, then $(\tilde g, \tilde\pi)$ is $C^{2,\alpha}_{2-n} \times C^{1,\alpha}_{1-n}$ up to the boundary and has harmonic asymptotics.[2]

[2]In this theorem and throughout this chapter, whenever we refer to H́older spaces on M, we mean that they are regular up to the boundary.

Remark 10.3.2. If (g, π) is $C^{k+2,\alpha} \times C^{k+1,\alpha}$ up to the boundary, $(\tilde{\mu}, \tilde{J})$ is $C^{k,\alpha}$ up to the boundary, and $\tilde{\theta}$ is $C^{k+1,\alpha}$ on the boundary, then $(\tilde{g}, \tilde{\pi})$ will be $C^{k+2,\alpha} \times C^{k+1,\alpha}$ up to the boundary.

We fix (g, π) be asymptotically flat according to Definition Definition 10.2.1, with $(\Phi, \Theta)(g, \pi) = (2\mu, J, \theta)$. We now define the operator used in the proof of Theorem Theorem 10.3.1 and study its linearization.

Let χ be a smooth nonnegative cutoff function on $\mathbb{R}^n$ equal to 1 on B_1 and vanishing outside B_2. Define $\chi_\lambda(x) = \chi(\frac{x}{\lambda})$. For sufficiently large λ, χ_λ is defined by extending it to be 1 on the connected compact subset of M that strictly contains the boundary ∂M. Now define

$$g_\lambda = \chi_\lambda g + (1 - \chi_\lambda)\overline{g}$$

$$\pi_\lambda = \chi_\lambda \pi$$

so that $g_\lambda = \overline{g}$ and $\pi_\lambda = 0$ for $|x| \geq 2\lambda$ and $\Theta(g_\lambda, \pi_\lambda) = \Theta(g, \pi) = \theta$. It is convenient to set $(g_\infty, \pi_\infty) = (g, \pi)$.

The basic idea of the density theorem (going back to [SY81b; CS06; EHLS16]) is to make a conformal change to (g_λ, π_λ) in order to reimpose the "constraint" (either prescribed Φ or modified Φ) lost in the cutoff process by taking λ large and using the inverse function theorem. However, $(DT|_{(1,0)}, DB_{(1,0)})$ is not necessarily an isomorphism. This issue is also present in [CS06; EHLS16]. The solution is to change the domain of (T, Υ) to create an operator whose differential at (g_∞, π_∞) is an isomorphism. The alteration will only introduce "compactly supported" deformations, so that the final data set will still have harmonic asymptotics.

Lemma 10.3.3. *Fix initial data (g, π) as in Theorem 10.3.1. There exists a closed subspace $K_1 \subset T_{(1,0)}\mathcal{C}$, a finite dimensional subspace $K_2 \subset T_{(g,\pi)}\mathcal{D}$ spanned by compactly supported smooth functions, and constants $r_0 > 0$ and C such that the following holds for all λ sufficiently large:*

The differentials of the operators

$$\hat{\mathcal{P}}_\lambda : [(1,0) + K_1] \times K_2 \to \mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M)$$
$$((u, Y), (h, w)) \mapsto (\Phi, \Theta)[\Psi_{(g_\lambda,\pi_\lambda)}(u, Y) + (h, w)]$$

are isomorphisms at the point $((1,0),(0,0))$. In fact, we have

$$\|D\hat{\mathcal{P}}_\lambda|^{-1}_{((1,0),(0,0))}\|_{\text{op}} \leq C \tag{10.3.1}$$

and the Lipschitz constant of $D\hat{\mathcal{P}}_\lambda$ is bounded by C on $B_{r_0}((1,0),(0,0))$.

For $\lambda = \infty$, the relevant differential is

$$D\hat{\mathcal{P}}_\infty|_{((1,0),(0,0))}((v,Z),(h,w)) = (DT, D\Upsilon)|_{(1,0)}(u,Y) + (D\Phi, D\Theta)|_{(g,\pi)}(h,w). \tag{10.3.2}$$

To construct K_1, we will require the following lemma.

Lemma 10.3.4. *$(DT_{(1,0)}, DB_1|_{(1,0)})$ restricted to $\ker DB_2|_{(1,0)} \cap \ker DB_3|_{(1,0)}$ is Fredholm.*

Proof. By Proposition 10.2.9 (1), we have the estimate

$$\|(v,Z)\|_{W^{2,p}_{-q}} \lesssim \|DT(v,Z)\|_{\mathcal{L}} + \|DB_1(v,Z)\|_{W^{1-\frac{1}{p},p}} + \|(v,Z)\|_{L^p(U)}$$

for every $(v,Z) \in \ker DB_2 \cap \ker DB_3$. As mentioned in the proof of Proposition 10.2.9, it follows from [Sch02, Theorem 5.21] that (DT, DB_1) is semi-Fredholm. It remains to show that $(DT, DB_1)[\ker DB_2 \cap DB_3]$ has finite codimension. If not, then there exists an infinite-dimensional subspace $X \subset \mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M)$ such that $(DT, DB_1)[\ker DB_2 \cap DB_3] \cap X = 0$. But that would imply that

$$(DT, DB)[T_{(1,0)}\mathcal{C}] \cap [X \times \{0\} \times \{0\}] = 0,$$

which is impossible since the full elliptic operator is Fredholm. □

To construct K_2, we need to know that the lineariation of (Φ, Θ) is surjective, which generalizes Proposition 3.1 in [CS06]. See also the related work by Zhongshan An [An22].

Proposition 10.3.5. *$(D\Phi, D\Theta)|_{(g,\pi)} : T_{(g,\pi)}\mathcal{D} \to \mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M)$ is surjective.*

Proof. Since $D\Theta|_{(g,\pi)} : T_{(g,\pi)}\mathcal{D} \to W^{1-\frac{1}{p},p}(\partial M)$ is surjective (Lemma 10.2.8), it suffices to show that $D\Phi|_{(g,\pi)} : \ker(D\Theta_{(g,\pi)}) \to \mathcal{L}$ is surjective.

First we claim that

$$D\Phi|_{(g,\pi)}[\ker D\Theta|_{(g,\pi)}] \subset \mathcal{L}$$

is closed and has finite codimension. It suffices to observe that

$$DT|_{(1,0)}[\ker DB|_{(1,0)}] \subset D\Phi|_{(g,\pi)}[\ker D\Theta|_{(g,\pi)}], \tag{10.3.3}$$

as the former is closed with finite codimension by repeating the argument of Lemma 10.3.4.

As a consequence of the Hahn–Banach theorem [Bre11, Corollary 1.8], if $D\Phi|_{(g,\pi)}[\ker D\Theta|_{(g,\pi)}] \neq \mathcal{L}$, there is a nontrivial bounded linear functional $(\xi, V) \in \mathcal{L}^* = L^{p'}_{2+q-n}$ which annihilates it. By the inclusion (10.3.3), (ξ, V) annihilates $DT|_{(1,0)}(\ker DB|_{(1,0)})$. By considering arbitrary test data $(h, w) \in T_{(g,\pi)}\mathcal{D}$ compactly supported away from ∂M (whence $(h, w) \in \ker DB|_{(1,0)}$), we see that (ξ, V) solves the equation

$$D\Phi|^*_{(g,\pi)}(\xi, V) = 0 \tag{10.3.4}$$

in the sense of distributions. Arguing as in [HL20, Appendix B], we conclude that ξ and V are C^2 in the interior and hence solve the equation classically. Note that the boundary behavior of (ξ, V) is not needed for our argument.

The result now follows from arguments in [CS06], as explained in detail in [Lee19, Theorem 9.9]. (The presence of a boundary is irrelevant to this part of the argument.) For the reader's convenience, we summarize the basic argument: The equations (10.3.4) imply homogeneous Hessian-type equations for (ξ, V), with coefficients decaying according to the asymptotic flatness assumption. Using this Hessian-type system, the $L^{p'}$ decay can be bootstrapped to become pointwise C^1 decay. Next, initial decay of (ξ, V) then implies Hessian decay that is more than 2 orders faster, which gives improved decay on (ξ, V) simply by (twice) integrating along coordinate rays to infinity. Bootstrapping in this way, (ξ, V) must have infinite-order pointwise decay. From here, one can use a unique continuation argument (as in [CS06]) to see that (ξ, V) vanishes identically, or alternatively, as explained in the proof of [Lee19, Theorem 9.9], for each $p \in M$ one can directly use the second-order system of ODEs satisfied by (ξ, V) along a curve from p to infinity to see that infinite-order decay implies vanishing at p. This ODE argument originated in the work of Huang–Martin–Miao [HMM18, Lemma B.3]. □

With our preparatory results in place, we can construct K_1 and K_2.

Proof of Lemma 10.3.3. Let $K_1 \subset \ker DB_2 \cap \ker DB_3$ (note that this intersection is closed in $T_{(1,0)}\mathcal{C}$ and hence is a Banach space) be a complementing subspace for the kernel of (DT, DB_1) inside $\ker DB_2 \cap \ker DB_3$. This subspace exists because the kernel is finite dimensional (Lemma 10.3.4) and finite dimensional subspaces are always complemented. We then note that by the formula for $D\Upsilon$, for $(v, Z) \in K_1$,

$$D\Upsilon(v, Z) = DB_1(v, Z).$$

It follows that

$$(DT, D\Upsilon) : K_1 \to \mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M)$$

is injective. Furthermore, its range $R = (DT, DB_1)[\ker DB_2 \cap \ker DB_3]$ is a closed subspace of $\mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M)$ with finite codimension by Lemma 10.3.4.

Let A be a finite-dimensional subspace of $\mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M)$ which complements R in $\mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M)$. Using Proposition 10.3.5, we can find a finite-dimensional subspace $K_2' \subset T_{(g,\pi)}\mathcal{D}$ so that

$$(D\Phi, D\Theta)|_{(g,\pi)}(K_2) = A.$$

Note that smooth, compactly supported sections of $(T^*M \odot T^*M) \times (TM \odot TM)$ are dense in $T_{(g,\pi)}\mathcal{D} = W^{2,p}_{-q}(T^*M \odot T^*M) \times W^{1,p}_{-q-1}(TM \odot TM)$, so we can find a finite-dimensional space K_2 made up of smooth, compactly supported sections that closely approximates K_2'. By choosing a good enough approximation, the image of K_2 will be close enough to A to still be complementary to R. With this choice of K_1 and K_2, $D\hat{\mathcal{P}}_\infty|_{((1,0),(0,0))}$ is an isomorphism.

Since $(g_\lambda, \pi_\lambda) \to (g_\infty, \pi_\infty)$ in $\mathcal{D}$, it follows from (10.A.1) of Lemma 10.A.2 that

$$D\hat{\mathcal{P}}_\lambda|_{((1,0),(0,0))} \to D\hat{\mathcal{P}}_\infty|_{((1,0),(0,0))}$$

in operator norm as $\lambda \to \infty$. Therefore $D\hat{\mathcal{P}}_\lambda|_{((1,0),(0,0))}$ is also an isomorphism for sufficiently large λ, and its inverse satisfies (10.3.1). Finally, the Lipschitz constant bound follows from the Hessian bound (10.A.2) of Lemma 10.A.2. □

We state the relevant standard elliptic regularity fact needed to establish the boundary regularity in Theorem Theorem 10.3.1. The only subtlety is that we are not assuming u to be C^2. The version here (for Sobolev u) can be read off from [Mor66, Theorem 6.4.8].

Lemma 10.3.6. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with $C^{2,\alpha}$ boundary, $\alpha \in (0,1)$. Suppose that $a^{ij} \in C^{0,\alpha}(\overline{\Omega})$ is positive definite, and suppose that $f \in C^{0,\alpha}(\overline{\Omega})$ and $g \in C^{2,\alpha}(\partial\Omega)$. If $u \in W^{2,p}(\Omega)$ $(p > n)$ is a strong solution of*

$$a^{ij}\partial_i\partial_j u = f \quad \text{in } \Omega,$$
$$u = g \quad \text{on } \partial\Omega,$$

then $u \in C^{2,\alpha}(\overline{\Omega})$.

If instead $\beta^i \in C^{1,\alpha}(\partial\Omega)$ is an oblique vector field and $g \in C^{1,\alpha}(\partial\Omega)$ and $u \in W^{2,p}(\Omega)$ $(p > n)$

satisfies

$$a^{ij}\partial_i\partial_j u = f \quad \textit{in } \Omega,$$
$$\beta^i\partial_i u = g \quad \textit{on } \partial\Omega,$$

then $u \in C^{2,\alpha}(\overline{\Omega})$.

We now use the inverse function theorem to prove Theorem 10.3.1. The proof is a boundary version of the argument given for [Lee19, Theorem 9.10 and Proposition 9.11], which itself is based on [CS06].

Proof of Theorem Theorem 10.3.1. By Lemma 10.3.3 and the inverse function theorem for Banach spaces (see in particular the "quantitative" version [Lee19, Theorem A.43]), there exists a constant C such that for $r > 0$ sufficiently small and λ sufficiently large, $\hat{\mathcal{P}}_\lambda^{-1}$ exists and maps $B_r(2\mu_\lambda, J_\lambda, \theta) \subset \mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M)$ into $B_{Cr}((1,0),(0,0)) \subset [(1,0)+K_1] \times K_2$, where (μ_λ, J_λ) are the constraints of (g_λ, π_λ). So if $(\tilde{\mu}, \tilde{J}, \tilde{\theta})$ satisfy the hypotheses of the theorem with

$$\|(2\tilde{\mu}, \tilde{J}, \tilde{\theta}) - (2\mu, J, \theta)\|_{\mathcal{L}\times W^{1-\frac{1}{p},p}} < \frac{r}{2}$$

and λ is sufficiently large that $\|(2\mu_\lambda, J_\lambda) - (2\mu, J)\|_{\mathcal{L}} < \frac{r}{2}$, then there exist $\alpha_\lambda \in [(1,0)+K_1] \times K_2$ such that $\hat{\mathcal{P}}_\lambda(\alpha_\lambda) = (2\tilde{\mu}, \tilde{J}, \tilde{\theta})$, and

$$\|\alpha_\lambda - ((1,0),(0,0))\|_{\mathcal{C}\times T_{(g,\pi)}\mathcal{D}} \leq C\|(2\tilde{\mu}, \tilde{J}, \tilde{\theta}) - (2\mu, J, \theta)\|_{\mathcal{L}\times W^{1-\frac{1}{p},p}} \leq Cr.$$

By choosing r smaller (and hence also λ larger) depending on the constant in Morrey's inequality, we can ensure $|u_\lambda - 1| < 1$ everywhere, so that it is a valid conformal factor. Set $\delta = \frac{r}{2}$.

Having made all of these choices, write $\alpha_\lambda = ((u,Y),(h,w))$. We claim that the initial data $(\tilde{g}, \tilde{\pi}) = \Psi_{(g_\lambda,\pi_\lambda)}(u,Y) + (h,w)$ is the desired solution in the conclusion of Theorem 10.3.1. By construction, it has the desired constraints $(\tilde{\mu}, \tilde{J})$ and outer/inner null expansion $\tilde{\theta}$ on $\partial^\pm M$, and it satisfies the desired estimate

$$\|(\tilde{g}, \tilde{\pi}) - (g, \pi)\|_{\mathcal{D}} \leq C\|(\tilde{\mu}, \tilde{J}, \tilde{\theta}) - (\mu, J, \theta)\|_{\mathcal{L}\times W^{1-\frac{1}{p},p}}.$$

And since (h_λ, w_λ) is compactly supported,

$$(\tilde{g}, \tilde{\pi}) = \Psi_{(\overline{g},0)}(u_\lambda, Y_\lambda)$$

for $|x| \geq 2\lambda$, as desired.

It only remains to show that show that if $(\tilde{\mu}, \tilde{J})$ is $C^{0,\alpha}_{-n-1-\varepsilon}$ up to the boundary and $\tilde{\theta}$ is $C^{1,\alpha}$, then $(\tilde{g}, \tilde{\pi})$ is $C^{2,\alpha}_{2-n} \times C^{1,\alpha}_{1-n}$ up to the boundary and has harmonic asymptotics. Since (h, w) is compactly supported and smooth, it suffices to show that (u, Y) satisfies the asymptotic expansion in Definition 10.2.5 and is $C^{2,\alpha}$ up to the boundary. The asymptotic expansion is well-known from earlier references such as [Lee19, Lemma 9.8], but here we carefully account for the presence of the (h, w) term and the boundary in order to prove regularity up to the boundary.

To prove the desired result, we re-write

$$\hat{\mathcal{P}}_\lambda((u, Y), (h, w)) = (2\tilde{\mu}, \tilde{J}, \tilde{\theta})$$

as a set of linear elliptic equations in (u, Y), viewing the nonlinearities as either coefficients or nonhomogenous terms. More precisely, after we choose local coordinates (which is fine since we are proving a local regularity result now), we will have equations for $u, Y_1, \ldots, Y_n$ of the form described in Lemma 10.3.6 above. We will now explain this in detail.

By the Morrey embedding theorem, we already know that $(u, Y) \in C^{1,\alpha}_{\text{loc}}$ on M. We first examine the $\tilde{\mu}$ equation as our equation for u. First, the non-scalar curvature terms of $\tilde{\mu}$ are $C^{0,\alpha}_{\text{loc}}$ and therefore can be viewed as part of the $C^{0,\alpha}_{\text{loc}}$ nonhomogenous term. Moreover, using [Lee19, Exercise 1.2] or otherwise, we can re-write

$$R(u^s g_\lambda + h) = s u^{s-1}(\tilde{g}^{ik}\tilde{g}^{jl} - \tilde{g}^{ij}\tilde{g}^{kl})(g_\lambda)_{kl}\partial_i\partial_j u + (\text{terms in } C^{0,\alpha}_{\text{loc}}) \qquad (10.3.5)$$

Let U be a bounded set large enough so that every element of K_2 vanishes outside U. So in the complement of U, where $\tilde{g}$ is just a conformal change, this becomes

$$R(u^s g_\lambda + h) = -s(n-1)u^{-s-1}(g_\lambda)^{ij}\partial_i\partial_j u + (\text{terms in } C^{0,\alpha}_{\text{loc}}),$$

whose coefficient matrix is obviously negative definite. Meanwhile, $g_\lambda = g$ on U for sufficiently large λ, so for δ sufficiently small, (10.3.5) will be strictly elliptic because $|u-1|$ and h can be made uniformly small enough so that the second-order coefficients in (10.3.5) can be made uniformly close to

$$s(g^{ik}g^{jl} - g^{ij}g^{kl})g_{kl} = -s(n-1)g^{ij},$$

which we know is negative definite. In either case, u solves an elliptic equation of the type described in Lemma 10.3.6.

For the $\tilde{J}$ equations, we compute

$$\operatorname{div}_{\tilde{g}} \tilde{\pi}^i = u^{-\frac{3}{2}s} \left[-\tfrac{3}{2} s \frac{\nabla_j u}{u} (\pi_\lambda^{ij} + \mathfrak{L}_{g_\lambda} Y^{ij}) + \tilde{\nabla}_j (\pi_\lambda^{ij} + \mathfrak{L}_{g_\lambda} Y^{ij} + w^{ij}) \right].$$

Again, the lower order terms are clearly $C^{0,\alpha}_{\text{loc}}$, so we have

$$\tilde{\nabla}_j \mathfrak{L}_{g_\lambda} Y^{ij} \in C^{0,\alpha}_{\text{loc}}. \tag{10.3.6}$$

Meanwhile, by looking at the first order terms of $\mathfrak{L}_{g_\lambda} Y$, we have

$$\mathfrak{L}_{g_\lambda} Y^{ij} = g_\lambda^{ik} \partial_k Y^j + g_\lambda^{jk} \partial_k Y^i - \partial_k Y^k g_\lambda^{ij} + (\text{terms in } C^{0,\alpha}_{\text{loc}}).$$

By taking ∂_j, we can see that

$$\begin{aligned} \tilde{\nabla}_j \mathfrak{L}_{g_\lambda} Y^{ij} &= g_\lambda^{ik} \partial_j \partial_k Y^j + g_\lambda^{jk} \partial_j \partial_k Y^i - \partial_j \partial_k Y^k g_\lambda^{ij} + (\text{terms in } C^{0,\alpha}_{\text{loc}}) \\ &= g_\lambda^{jk} \partial_k \partial_j Y^i + (\text{terms in } C^{0,\alpha}_{\text{loc}}). \end{aligned}$$

Combining this with (10.3.6), we see that

$$g_\lambda^{jk} \partial_k \partial_j Y^i \in C^{0,\alpha}_{\text{loc}},$$

where g_λ is positive definite, so each Y^i also satisfies an elliptic equation of the of the type described in Lemma 10.3.6.

The only thing left to check is that $u, Y^1, \ldots, Y^n$ satisfy boundary conditions of the type described in Lemma 10.3.6. We have stipulated that $\Theta(\tilde{g}, \tilde{\pi}) = \tilde{\theta}$, which is only one boundary condition. However, the space K_1 in the definition of $\hat{\mathcal{P}}$ contains the other n boundary conditions we need.

Indeed, since $(u, Y) \in K_1$, then by definition

$$Y^\top = 0$$
$$HY^n - \frac{\partial Y^n}{\partial x^n} = 0,$$

where we have expressed the second condition in an almost-Fermi coordinate system. (We may do this since regularity is a local property.) So we see that $Y^1, \ldots, Y^{n-1}$ satisfy the Dirichlet boundary condition, and we claim that Y^n and u satisfy Neumann-type conditions of the type described in Lemma 10.3.6. For Y^n, this is immediate from observing that HY^{n-1} is $C^{1,\alpha}$ up to the boundary. Hence Lemma 10.3.6 implies that Y is $C^{2,\alpha}$ up to the boundary.

Using this upgraded regularity for Y, we can interpret $\Theta(\tilde{g}, \tilde{\pi}) = \tilde{\theta}$ as a Neumann-type boundary condition for u. We clearly have $-\tilde{\pi}^{ij}\tilde{\nu}_i\tilde{\nu}_j \in C^{1,\alpha}(\partial M)$, so we just have to investigate the mean curvature of ∂M with respect to $\tilde{g}$, which we denote by $H_{\tilde{g}}$. As in the proof of Lemma 10.2.3, we write ∂M as a regular level set of a smooth function f, so that

$$H_{\tilde{g}} = \gamma^{ij} \frac{1}{|\nabla f|_{\tilde{g}}} (\partial_{ij} f - \tilde{\Gamma}^k_{ij} \partial_k f),$$

where $\gamma^{ij} = \tilde{g}^{ij} - \tilde{\nu}^i \tilde{\nu}^j$. By focusing only on the terms with derivatives of u (in the Christoffel symbols), we have

$$(-2\gamma^{ij}\tilde{g}^{kl} + \gamma^{lj}\tilde{g}^{ik})\tilde{\nu}_k s u^{s-1} g_{jl} \partial_i u \in C^{1,\alpha}(\partial M),$$

which is an equation of the form $\beta^i \partial_i u \in C^{1,\alpha}(\partial M)$, where each $\beta^i \in C^{1,\alpha}(\partial M)$ as well. The only thing left to check is that β^i is not tangent to ∂M. To see this, observe that when $h = 0$, we have $\beta^i = (n-1) s u^{-1} \nu^i$, and therefore β^i is oblique for sufficiently small h. □

10.3.2 Dominant energy condition density theorem

We begin with a precise re-statement of Theorem 10.1.1:

Theorem 10.3.7 (Density theorem for DEC)**.** *Let (M^n, g, π) be a complete asymptotically flat initial data set satisfying the dominant energy condition $\mu \geq |J|$ and with compact boundary ∂M having outer null expansion θ on $\partial^+ M$ and inner null expansion θ on $\partial^- M$. Let $p > n$ and $\frac{n-2}{2} < q < n-2$ be strictly less than the decay rate of (g, π). For any $\varepsilon > 0$ there exists a constant $\delta > 0$ so that the following is true:*

For any $\tilde{\theta} \in C^{1,\alpha}(\partial M)$ satisfying $\|\tilde{\theta} - \theta\|_{W^{1-\frac{1}{p},p}} < \delta$ there exists an asymptotically flat initial data *set* $(\tilde{g}, \tilde{\pi})$, $C^{2,\alpha} \times C^{1,\alpha}$*-regular up to the boundary, also satisfying the dominant energy condition, such that* $(\tilde{g}, \tilde{\pi})$ *has harmonic asymptotics in each end of* M*, the new outer/inner null expansion on* $\partial^{\pm} M$ *is* $\tilde{\theta}$*, and the new data set satisfies*

$$\|(\tilde{g}, \tilde{\pi}) - (g, \pi)\|_{\mathcal{D}} < \varepsilon \quad \text{and} \quad \|(\tilde{\mu}, \tilde{J}) - (\mu, J)\|_{L^1} < \varepsilon.$$

Furthermore, we can choose $(\tilde{g}, \tilde{\pi})$ *such that the* strict *dominant energy condition holds,* $\tilde{\mu} > |\tilde{J}|$*. Simultaneously,* $(\tilde{\mu}, \tilde{J})$ *may be chosen to decay as fast as we like in the sense that if* f *is any positive smooth function, then we can demand* $\tilde{\mu} + |\tilde{J}| \leq f(|x|)$ *on the end.*

Alternatively, we can choose $(\tilde{g}, \tilde{\pi})$ *to be vacuum outside a compact set, that is,* $\tilde{\mu} = |\tilde{J}| = 0$ *outside a compact set.*

Given a fixed initial data set (M^n, g, π), the *modified constraint operator* $\overline{\Phi}_{(g,\pi)}$ is defined by

$$\overline{\Phi}_{(g,\pi)}(\gamma, \tau) = \Phi(\gamma, \tau) + (0, \tfrac{1}{2} g^{ij} \gamma_{jk} J^k)$$

for $(\gamma, \tau) \in \mathcal{D}$.

Lemma 10.3.8 (Corvino–Huang [CH20])**.** *Let* (g, π) *and* $(\tilde{g}, \tilde{\pi})$ *be initial data, and assume that*

$$\overline{\Phi}_{(g,\pi)}(\tilde{g}, \tilde{\pi}) - \overline{\Phi}_{(g,\pi)}(g, \pi) = (2\psi, 0)$$

for some function ψ*. If additionally* $|\tilde{g} - g|_g \leq 3$*, then*

$$|\tilde{J}|_{\tilde{g}} \leq |J|_g.$$

The linearization of the modified constraint operator at (g, π) is given by

$$D\overline{\Phi}_{(g,\pi)}(h, w) = D\Phi|_{(g,\pi)}(h, w) + (0, \tfrac{1}{2} g^{ij} h_{jk} J^k)$$

The addtional zeroth order term $(0, \frac{1}{2} g^{ij} h_{jk} J^k)$ does not affect the proof of Proposition 10.3.5, and so we obtain the surjectivity result.

Proposition 10.3.9. $(D\overline{\Phi}_{(g,\pi)}, D\Theta)|_{(g,\pi)} : T_{(g,\pi)}\mathcal{D} \to \mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M)$ *is surjective.*

By Proposition 10.3.9, we can use the construction of Lemma 10.3.3 to construct the analogous subspaces K_1, K_2. Using this and the inverse function theorem argument in Theorem 10.3.1, we show that it is possible to perturb the initial data to strict DEC. Proposition 10.3.10 does not attempt to produce harmonic asymptotics, which we get to in Theorem 10.3.7.

Proposition 10.3.10 (Perturbing to strict DEC)**.** *Let (M^n, g, π) be a complete asymptotically flat initial data set satisfying the dominant energy condition $\mu \geq |J|$ and with $\Theta(g,\pi) = \theta$. Let $p > n$ and $\frac{n-2}{2} < q < n-2$ be strictly less than the decay rate of (g,π). For any $\varepsilon > 0$ there exist constants $\delta > 0$ and $\gamma > 0$ so that the following is true:*

For any $\tilde{\theta} \in C^{1,\alpha}(\partial M)$ satisfying $\|\tilde{\theta} - \theta\|_{W^{1-\frac{1}{p},p}} < \delta$ there exists an asymptotically flat initial data set $(\tilde{g}, \tilde{\pi})$, $C^{2,\alpha} \times C^{1,\alpha}$-regular up to the boundary, with $\Theta(\tilde{g},\tilde{\pi}) = \tilde{\theta}$, that satisfies the following "uniform" strict dominant energy condition

$$\tilde{\mu} > (1+\gamma)|\tilde{J}|_{\tilde{g}},$$

as well as the estimates

$$\|(\tilde{g},\tilde{\pi}) - (g,\pi)\|_{\mathcal{D}} < \varepsilon \quad \text{and} \quad \|(\tilde{\mu},\tilde{J}) - (\mu, J)\|_{L^1} < \varepsilon.$$

Proof. Let f be a smooth positive function on M decaying exponentially at infinity. By essentially repeating the proof of Lemma 10.3.3, we can construct subspaces $K_1 \subset \ker DB_2 \cap \ker DB_3$ and $K_2 \subset T_{(g,\pi)}\mathcal{D}$ (consisting of compactly supported smooth functions) to define an operator

$$\begin{aligned} \hat{\mathcal{P}} : [(1,0) + K_1] \times K_2 &\to \mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M) \\ ((u,Y),(h,w)) &\mapsto (\overline{\Phi}_{(g,\pi)}, \Theta)[\Psi_{(g,\pi)}(u,Y) + (h,w)] \end{aligned}$$

whose differential at $((1,0),(0,0))$ is an isomorphism. We now proceed as in the proof of Theorem 10.3.1. In particular, we use the inverse function theorem [Lee19, Theorem A.43] to solve

$$\hat{\mathcal{P}}((u_t, Y_t),(h_t,w_t)) = \left(\overline{\Phi}_{(g,\pi)}(g,\pi) + (2t(f+|J|_g),0),\, \tilde{\theta}\right) \tag{10.3.7}$$

for $((u_t,Y_t),(h_t,w_t))$, which is possible for sufficiently small $\delta > 0$ and $t > 0$.

Arguing as in the proof of Theorem 10.3.1, we can show that (u_t, Y_t) decays enough so that $(\tilde{g}_t, \tilde{\pi}_t) = \Psi_{(g,\pi)}(u_t,Y_t) + (h_t,w_t)$ is asymptotically flat, and (u_t,Y_t) is $C^{2,\alpha}$ up to the boundary.

We now see about the other claims in the proposition. First, we claim that $(\tilde{g}_t, \tilde{\pi}_t)$ satisfies the strict DEC for t small enough. By the Sobolev inequality, we may assume $\tilde{g}_t - g$ is uniformly small pointwise. Therefore, by Lemma 10.3.8,

$$\tilde{\mu}_t = \mu + t(f + |J|_g) > \mu + t|J|_g \geq (1+t)|J|_g \geq (1+t)|\tilde{J}_t|_{\tilde{g}_t}.$$

Finally, we need to show that $(\tilde{\mu}_t, \tilde{J}_t) \to (\mu, J)$ in L^1. Since $f + |J|_g \in L^1$, it is clear from the equation $\tilde{\mu}_t = \mu + t(f + |J|_g)$ that $\tilde{\mu}_t \to \mu$. For the momentum we have $\tilde{J}^i_t - J^i = \frac{1}{2} g^{ij}(\tilde{g}_t - g)_{jk} J^k$, but $\tilde{g}_t \to g$ uniformly, so $\tilde{J}_t \to J$. We take γ to be the t satisfying these conditions, which completes the proof. □

Finally, we prove Theorem 10.3.7.

Proof of Theorem 10.3.7. First, perturb the data set according to Proposition 10.3.10 to ensure the strict DEC holds in the form $\mu > (1+\gamma)|J|_g$ everywhere on M. Call the perturbed data set (g, π) for simplicity. Let χ_λ be the family of cutoff functions used in Theorem 10.3.1 and f a rapidly decaying positive function on M. By Theorem 10.3.1, we may construct a data set $(\tilde{g}_\lambda, \tilde{\pi}_\lambda)$ with harmonic asymptotics satisfying

$$(\Phi, \Theta)(\tilde{g}_\lambda, \tilde{\pi}_\lambda) = \left(\chi_\lambda \Phi(g, \pi) + \tfrac{2}{\lambda}(f, 0),\, \tilde{\theta}\right)$$

if λ is large enough and δ is small enough. Note that since we can choose f to decay as rapidly as we like, our prescribed $(\tilde{\mu}, \tilde{J})$ will certainly lie in $C^{0,\alpha}_{-n-1-\varepsilon}$.

The only thing left to check is that the strict DEC is satisfied:

$$\begin{aligned}
\tilde{\mu}_\lambda - \frac{1}{\lambda} f &= \chi_\lambda \mu \\
&\geq \chi_\lambda (1+\gamma)|J|_g \\
&\geq \chi_\lambda \left(|J|_{\tilde{g}_\lambda} - |\tilde{g}_\lambda - g|_g^{\frac{1}{2}} |J|_g + \gamma |J|_g\right) \\
&= |\tilde{J}_\lambda|_{\tilde{g}_\lambda} + \chi_\lambda |J|_g (\gamma - |\tilde{g}_\lambda - g|_g^{\frac{1}{2}}).
\end{aligned}$$

For λ large, $|\tilde{g}_\lambda - g|_g^{\frac{1}{2}} < \gamma$ by the estimates in Theorem 10.3.1, so we then have

$$\tilde{\mu}_\lambda - |\tilde{J}_\lambda|_{\tilde{g}_\lambda} \geq \frac{1}{\lambda} f,$$

which implies the strict DEC. Note that f also controls the decay of the DEC scalar $\tilde{\mu}_\lambda - |\tilde{J}_\lambda|_{\tilde{g}_\lambda}$, as it is identically equal to $\frac{1}{\lambda} f$ for $|x| \geq 2\lambda$.

If we instead wish to prescribe vacuum outside a compact set, we perform the same argument as above but with $f \equiv 0$. □

10.4 Positive mass theorem with boundary

10.4.1 Proof of the inequality $E \geq |P|$

In this subsection we explain how Theorem 9.2.1 follows from combining Theorem 10.3.7 with the proof of the boundaryless case from [EHLS16]. Suppose that there exists a complete asymptotically flat initial data set (M, g, π) satisfying the DEC, whose compact boundary is made up of components which are either weakly outer trapped ($\theta^+ \leq 0$) or weakly inner untrapped ($\theta^- \geq 0$) with respect to the normal pointing into M, such that $E < |P|$. By Theorem 10.3.7, we can perturb (g, π) to new initial data $(\tilde{g}, \tilde{\pi})$ that has harmonic asymptotics, satisfies the strict DEC, and has compact boundary made up of components that have either $\theta^+ < 0$ or $\theta^- > 0$, while maintaining the inequality $\tilde{E} < |\tilde{P}|$. From here the exact same argument as in [EHLS16] (after the application of the density theorem there) results in a contradiction. The only thing to note is that the boundary acts as a barrier for the MOTS ($\theta^+ = 0$ hypersurfaces) that are constructed in the proof. This part of the proof is also identical to the reasoning used in [GL21]. (Note that the H́older decay assumption on (μ, J) in [EHLS16] is unnecessary, as can be seen from our proof and was observed in [Lee19; GL21].)

To be more precise, in [EHLS16], one seeks to construct a stable MOTS hypersurface in M with prescribed boundary equal to a large sphere Γ^{n-2} of constant height on a large cylinder C (with smoothed corners). Theorem 1.1 of [Eic09] states that this is possible if one can find a compact Ω such that $\Gamma \subset \partial\Omega$ divides $\partial\Omega$ into $\partial_1\Omega$ and $\partial_2\Omega$ such that $\theta^+_{\partial_1\Omega} > 0$ with respect to the normal pointing *out of* Ω, and $\theta^+_{\partial_2\Omega} < 0$ with respect to the normal pointing *into* Ω. If M has one end and no boundary, then we choose Ω to be the region enclosed by C, $\partial_1\Omega$ to be the part of C lying above Γ, and $\partial_2\Omega$ to be the part of C lying below Γ. Harmonic asymptotics guarantee that if C is big enough, these choices satisfy the hypotheses on θ^+ needed to apply [Eic09, Theorem 1.1]. If there are multiple ends, then we choose Ω to be enclosed by C in the end of interest and large celestial spheres in all other ends. Those celestial spheres have $\theta^+ < 0$ with respect to the normal pointing into Ω, and hence those spheres can be included as part of $\partial_2\Omega$. Finally, we come to the case of interest where M has a boundary. We define Ω the same way, except now we can treat any $\theta^+ < 0$ components of ∂M as part of $\partial_2\Omega$ while treating any $\theta^- > 0$ components of ∂M as part of $\partial_1\Omega$, because the condition $\theta^- > 0$ with respect to the normal pointing into M is equivalent to the condition $\theta^+ > 0$ with respect to the normal pointing out of Ω. □

10.4.2 The equality case $E = |P|$

Here we explain how the arguments in [HL20] can be adapted to handle a boundary, using the results of this chapter. Explicitly, we prove the following theorem, which constitutes the first part of the proof of Theorem 9.2.2.

Theorem 10.4.1. *Let $n \geq 3$, $p > n$, $q > \frac{n-2}{2}$, and $0 < \alpha < 1$, and assume that*

$$q + \alpha > n - 2. \tag{10.4.1}$$

Suppose that (M^n, g, π) is a complete asymptotically flat initial data set with boundary (in the sense defined in Definition 10.2.1) with the stronger decay assumption that

$$g - \overline{g} \in C^{2,\alpha}_{-q}(T^*M \odot T^*M) \tag{10.4.2}$$

$$\pi \in C^{1,\alpha}_{-1-q}(TM \odot TM). \tag{10.4.3}$$

Then if (M, g, π) satisfies the DEC, each component of ∂M is either weakly outer trapped or weakly inner untrapped, and its ADM energy-momentum satisfies $E = |P|$, then $E = |P| = 0$.

As explained in [HL24], in the case without boundary, the theorem is actually false without the stronger decay assumption.

Proof. Assume that (M, g, π) satisfies the hypotheses of Theorem 10.4.1. The basic strategy is the following: Using the first part of Theorem 9.2.1, we can see that (g, π) minimizes a "modified Regge–Teitelboim Hamiltonian" among all nearby initial data sets that have the same values of $(\overline{\Phi}_{(g,\pi)}, \Theta)$. By Proposition 10.3.9, $(D\overline{\Phi}_{(g,\pi)}, D\Theta)|_{(g,\pi)}$ is surjective, and hence we can apply Lagrange multipliers. These Lagrange multipliers give rise to a solution (f, X) of the adjoint equations $D\overline{\Phi}_{(g,\pi)}|^*_{(g,\pi)}(f, X) = 0$ such that (f, X) is asymptotic to the constant $(E, -2P)$. Once we have that, a result of Beig and Chruściel [BC96] (see also [HL20, Theorem A.2]) implies that $E = |P| = 0$.

For the analysis that follows, select $q \in (\frac{n-2}{2}, n-2)$ that is smaller than the q in statement of Theorem Theorem 10.4.1. Let (f_0, X_0) be a function and a vector field on M such that (f_0, X_0) is smooth, supported in the asymptotically flat coordinate chart, and exactly equal to the constant $(E, -2P)$ outside some compact set, where (E, P) denotes the fixed ADM energy-momentum of (g, π). We define the *modified Regge–Teitelboim Hamiltonian* $\mathcal{H} : \mathcal{D} \to \mathbb{R}$ corresponding to (g, π) by

$$\mathcal{H}(\gamma, \tau) = 2(n-1)\omega_{n-1}\left[E \cdot E(\gamma, \tau) - P \cdot P(\gamma, \tau)\right] - \int_M \overline{\Phi}_{(g,\pi)}(\gamma, \tau) \cdot (f_0, X_0)\, d\mu_g, \tag{10.4.4}$$

for all $(\gamma,\tau)\in\mathcal{D}$, where the volume measure $d\mu_g$ and the inner product in the integral are both with respect to g. Although $E(\gamma,\tau)$, $P(\gamma,\tau)$, and the integral need not exist for elements $(\gamma,\tau)\in\mathcal{D}$ whose constraints are not integrable, the expression $\mathcal{H}(\gamma,\tau)$ can be given meaning by using the alternative formula:

$$\begin{aligned}\mathcal{H}(\gamma,\tau)=\int_M &\left[(\operatorname{div}_g[\operatorname{div}_{\overline{g}}\gamma-d(\operatorname{tr}_{\overline{g}}\gamma)],\operatorname{div}_g\tau)-\Phi(\gamma,\tau)-\left(0,\tfrac{1}{2}\gamma\cdot J\right)\right]\cdot(f_0,X_0)\,d\mu_g\\&+\int_M([\operatorname{div}_{\overline{g}}\gamma-d(\operatorname{tr}_{\overline{g}}\gamma)],\tau)\cdot(\nabla f_0,\nabla X_0)\,d\mu_g,\end{aligned}\tag{10.4.5}$$

where $\overline{g}$ is a globally defined background metric that is Euclidean in the asymptotically flat end. As in [HL20], we can compute the linearization $D\mathcal{H}|_{(g,\pi)}:T_{(g,\pi)}\mathcal{D}\to\mathbb{R}$ to be

$$D\mathcal{H}|_{(g,\pi)}(h,w)=-\int_M(h,w)\cdot(D\overline{\Phi}_{(g,\pi)}|_{(g,\pi)})^*(f_0,X_0)\,d\mu_g,\tag{10.4.6}$$

for all $(h,w)\in T_{(g,\pi)}\mathcal{D}$. Note that ∂M can be ignored in all of these formulae because (f_0,X_0) vanishes near ∂M.

Next we define a constraint space

$$\mathfrak{C}_{(g,\pi)}=\left\{(\gamma,\tau)\in\mathcal{D}:\overline{\Phi}_{(g,\pi)}(\gamma,\tau)=\overline{\Phi}_{(g,\pi)}(g,\pi)\text{ and }\Theta(\gamma,\tau)=\Theta(g,\pi)\right\}.$$

By our assumptions on (g,π), each data set (γ,τ) in $\mathfrak{C}_{(g,\pi)}$ has weakly outer trapped or inner untrapped boundary components, and thanks to Lemma 10.3.8, it also satisfies the DEC. Moreover, since (γ,τ) has the same modified constraints as (g,π), it also follows that its constraints are integrable. Then Theorem 9.2.1 implies that if (γ,τ) is near enough to (g,π), then $E(\gamma,\tau)\geq|P(\gamma,\tau)|$. (We will discuss this more below. See Lemma 10.4.2.) From this, we see that (g,π) locally minimizes $\mathcal{H}$ on $\mathfrak{C}_{(g,\pi)}$ since

$$\begin{aligned}E\cdot E(\gamma,\tau)-P\cdot P(\gamma,\tau)&\geq E\cdot E(\gamma,\tau)-|P||P(\gamma,\tau)|\\&=E(E(\gamma,\tau)-|P(\gamma,\tau)|)\geq 0=E\cdot E(g,\pi)-P\cdot P(g,\pi),\end{aligned}$$

and the integral term in (10.4.4) is constant over $\mathfrak{C}_{(g,\pi)}$.

In other words, (g,π) locally minimizes $\mathcal{H}$ over a level set of $(\overline{\Phi}_{(g,\pi)},\Theta):\mathcal{D}\to\mathcal{L}\times W^{1-\frac{1}{p},p}(\partial M)$, so by surjectivity of its linearization (Proposition 10.3.9), there exist Lagrange multipliers (see [HL20,

Appendix D]) $(f_1, X_1) \in \mathcal{L}^*$, and $\lambda \in W^{1-\frac{1}{p},p}(\partial M)^*$ such that for all $(h, w) \in T_{(g,\pi)}\mathcal{D}$,

$$D\mathcal{H}|_{(g,\pi)}(h,w) = \int_M (f_1, X_1) \cdot D\overline{\Phi}_{(g,\pi)}|_{(g,\pi)}(h,w)\, d\mu_g + \int_{\partial M} \lambda \cdot D\Theta|_{(g,\pi)}(h,w).$$

Combining this with (10.4.6) and by choosing (h, w) to be arbitrary smooth test data that is compactly supported away from ∂M, we see that (f_1, X_1) must be a solution (in the distributional sense) of

$$D\overline{\Phi}_{(g,\pi)}|^*_{(g,\pi)}(f_1, X_1) = -D\overline{\Phi}_{(g,\pi)}|^*_{(g,\pi)}(f_0, X_0)$$

in the interior of M. As argued in the proof of Proposition 10.3.5, ellipticity of $DT|_{(1,0)}$ implies that (f_1, X_1) is actually smooth in the interior of M. Moreover, using the initial decay from being in $\mathcal{L}^*$ together with elliptic estimates, it follows that (f_1, X_1) has $C^{2,\alpha}_{-q}$ decay. (See [HL20, Proposition B.4] for details.) Thus $(f, X) \doteq (f_0, X_0) + (f_1, X_1)$ solves

$$D\overline{\Phi}_{(g,\pi)}|^*_{(g,\pi)}(f, X) = 0$$

in the interior of M, and $(f, X) - (E, -2P)$ has $C^{2,\alpha}_{-q}$ decay. The result now follows from Theorem A.2 of [HL20]. Note that it does not matter what what (f, X) does near the boundary ∂M since Theorem A.2 of [HL20] is only a statement about asymptotics and makes no global assumptions. □

There is one step in the proof above that requires further justification. We claimed that elements (γ, τ) of $\mathfrak{C}_{(g,\pi)}$ must satisfy $E(\gamma, \tau) \leq |P(\gamma, \tau)|$, but the problem is that (γ, τ) may only have Sobolev regularity and decay, but our positive mass theorem (Theorem 9.2.1) requires at least $C^{2,\alpha} \times C^{1,\alpha}$ local regularity as well as pointwise decay. Although we do not have a positive mass theorem for initial data in $\mathcal{D}$ with integrable constraints, we can at least prove it for data that is *near* the smooth data (g, π). This is the same idea that was used in [HL20, Theorem 4.1].

Lemma 10.4.2 (Sobolev version of positive mass inequality, with boundary)**.** *Let $3 \leq n \leq 7$, and let (M^n, g, π) be a complete asymptotically flat manifold, as in Definition 10.2.1, satsifying the DEC and with compact boundary such that each component is either weakly outer trapped or weakly inner untrapped. Let $p > n$ and let $q \in (\frac{n-2}{2}, n-2)$ be smaller than the assumed asymptotic decay rate of (g, π). Then there is an open ball $U \subset \mathcal{D}$ containing (g, π) such that if $(\gamma, \tau) \in U$, $\overline{\Phi}_{(g,\pi)}(\gamma, \tau) = \overline{\Phi}_{(g,\pi)}(g, \pi)$, and $\Theta(\gamma, \tau) = \Theta(g, \pi)$, then*

$$E(\gamma, \tau) \geq |P(\gamma, \tau)|.$$

Proof. The proof is essentially the same as in the proof of [HL20, Theorem 4.1], except that we use our new results to deal with the boundary. Define K_1, K_2, and $\hat{\mathcal{P}}$ as in the proof of Proposition 10.3.10. More generally, for $(\gamma,\tau)\in\mathcal{D}$, we define

$$\hat{\mathcal{P}}_{(\gamma,\tau)} : [(1,0)+K_1]\times K_2 \to \mathcal{L}\times W^{1-\frac{1}{p},p}(\partial M)$$
$$((u,Y),(h,w)) \mapsto (\overline{\Phi}_{(g,\pi)},\Theta)[\Psi_{(\gamma,\tau)}(u,Y)+(h,w)],$$

so that in particular, $\hat{\mathcal{P}}_{(g,\pi)}=\hat{\mathcal{P}}$. Using similar reasoning as in the proof of Lemma 10.3.3, we can use the inverse function theorem (together with estimates from Lemma 10.A.2) to see that there exists open ball $U\subset\mathcal{D}$ containing (g,π) and constants $\delta>0$ and $C_1>0$ with the property that for all (γ,τ) in U, $\hat{\mathcal{P}}_{(\gamma,\tau)}$ is a diffeomorphism between a neighborhood of $((1,0),(0,0))$ in $[(1,0)+K_1]\times K_2$ and the ball of radius δ around $\hat{\mathcal{P}}_{(\gamma,\tau)}((1,0),(0,0))$ in $\mathcal{L}\times W^{1-\frac{1}{p},p}(\partial M)$, and

$$\|((u-1,Y),(h,w))\|_{K_1\times K_2}\le C_1\|\hat{\mathcal{P}}_{(\gamma,\tau)}((u,Y),(h,w))-\hat{\mathcal{P}}_{(\gamma,\tau)}((1,0),(0,0))\|_{\mathcal{L}\times W^{1-\frac{1}{p},p}}.$$

We claim that the conclusion of Lemma 10.4.2 holds with this choice of U. We now assume (γ,τ) satisfies the hypotheses described in Lemma 10.4.2, that is, $(\gamma,\tau)\in U$ such that $\overline{\Phi}_{(g,\pi)}(\gamma,\tau)=\overline{\Phi}_{(g,\pi)}(g,\pi)$, and $\Theta(\gamma,\tau)=\Theta(g,\pi)$. We want to show that $E(\gamma,\tau)\ge|P(\gamma,\tau)|$, and we will do this by constructing a sequence $(\bar{\gamma}_k,\bar{\tau}_k)$ that converges to (γ,τ), to which we can apply Theorem 9.2.1).

Select a sequence of smooth asymptotically flat initial data (γ_k,τ_k) converging to (γ,τ) in $\mathcal{D}$. This implies that

$$\hat{\mathcal{P}}_{(\gamma_k,\tau_k)}((1,0),(0,0))=\left(\overline{\Phi}_{(g,\pi)}(\gamma_k,\tau_k),\Theta(\gamma_k,\tau_k)\right)\to\hat{\mathcal{P}}_{(\gamma,\tau)}((1,0),(0,0))=\left(\overline{\Phi}_{(g,\pi)}(\gamma,\tau),\Theta(\gamma,\tau)\right)$$

in $\mathcal{L}\times W^{1-\frac{1}{p},p}$.

In particular, for large enough k, $(\gamma_k,\tau_k)\in U$ and $\hat{\mathcal{P}}_{(\gamma,\tau)}((1,0),(0,0))$ lies in the δ-ball centered around $\hat{\mathcal{P}}_{(\gamma_k,\tau_k)}((1,0),(0,0))$, and hence, by our construction of U, there exists $((u_k,Y_k),(h_k,w_k))\in[(1,0)+K_1]\times K_2$ such that

$$\hat{\mathcal{P}}_{(\gamma_k,\tau_k)}((u_k,Y_k),(h_k,w_k))=\hat{\mathcal{P}}_{(\gamma,\tau)}((1,0),(0,0))$$

and

$$\|((u_k-1,Y_k),(h_k,w_k))\|_{K_1\times K_2} \le C_1\|\hat{\mathcal{P}}_{(\gamma_k,\tau_k)}((u_k,Y_k),(h_k,w_k)) - \hat{\mathcal{P}}_{(\gamma_k,\tau_k)}((1,0),(0,0))\|_{\mathcal{L}\times W^{1-\frac{1}{p},p}}$$
$$= C_1\|\hat{\mathcal{P}}_{(\gamma,\tau)}((1,0),(0,0)) - \hat{\mathcal{P}}_{(\gamma_k,\tau_k)}((1,0),(0,0))\|_{\mathcal{L}\times W^{1-\frac{1}{p},p}}.$$

Setting $(\bar{\gamma}_k,\bar{\tau}_k) \doteq \Psi_{(\gamma_k,\tau_k)}(u_k,Y_k)+(h_k,w_k)$, the inequality above shows that $(\bar{\gamma}_k,\bar{\tau}_k)$ converges to (γ,τ) in $\mathcal{D}$. Note that

$$(\overline{\Phi}_{(g,\pi)},\Theta)(\bar{\gamma}_k,\bar{\tau}_k) = \hat{\mathcal{P}}_{(\gamma,\tau)}((1,0),(0,0)) = (\overline{\Phi}_{(g,\pi)},\Theta)(\gamma,\tau) = (\overline{\Phi}_{(g,\pi)},\Theta)(g,\pi),$$

and thus, unlike the arbitrary smoothing (γ_k,τ_k), $(\bar{\gamma}_k,\bar{\tau}_k)$ satisfies the DEC (by Lemma 10.3.8) and has weakly outer trapped or inner untrapped boundary components. Moreover, by the same regularity argument used in the proof of Theorem 10.3.1, $(\bar{\gamma}_k,\bar{\tau}_k)$ is smooth enough and decays enough so that the positive mass inequality (Theorem 9.2.1) applies to $(\bar{\gamma}_k,\bar{\tau}_k)$, and hence $E(\bar{\gamma}_k,\bar{\tau}_k) \ge |P(\bar{\gamma}_k,\bar{\tau}_k)|$. Finally, we take the limit as $k\to\infty$ and use continuity of ADM energy-momentum [Lee19, Lemma 8.4] to conclude that $E(\gamma,\tau) \ge |P(\gamma,\tau)|$. (Note that $(\bar{\gamma}_k,\bar{\tau}_k)$ has the same modified constraints as (γ,τ), and thus the constraints of $(\bar{\gamma}_k,\bar{\tau}_k)$ converge to the constraints of (γ,τ) in L^1.) □

10.4.3 Embedding in Minkowski space when $E=0$

In this section, we use the Jang reduction method to show that if $\partial M \neq \emptyset$, then $E>0$. Combined with Theorem 9.2.1 and Theorem 10.4.1, this implies Theorem 9.2.2. We recall that a function f defined on an open set U in an initial data set (M,g,k) solves *Jang's equation* [Jan78; SY81a] if

$$H_g(f) - \mathrm{tr}_g(k)(f) = 0, \tag{10.4.7}$$

where

$$H_g(f) = \mathrm{div}_g\left(\frac{\nabla f}{\sqrt{1+|\nabla f|^2}}\right)$$

is the mean curvature of the graph of f in the cylinder over (M,g) and

$$\mathrm{tr}_g(k)(f) = \mathrm{tr}_g\, k - \frac{k(\nabla f,\nabla f)}{1+|\nabla f|^2}$$

is the trace of k (extended trivially in the vertical direction) over the tangent spaces of the graph of f.

Jang's equation (10.4.7) is a quasilinear elliptic equation for f, but the presence of the lower order term $\operatorname{tr}_g k$ precludes the use of the maximum principle to obtain a supremum estimate for f.[3] The lack of such an a priori estimate is an obstacle for proving solutions of Jang's equation exist. Schoen and Yau [SY81a] overcame this by instead considering the *capillary regularized Jang's equation*

$$H_g(f_\tau) - \operatorname{tr}_g(k)(f_\tau) = \tau f_\tau, \tag{10.4.8}$$

where τ is a positive real parameter which we want to send to zero.

The maximum principle now yields $\|f_\tau\|_{L^\infty} \lesssim \tau^{-1}$, which is singular but suffices to show that the solutions f_τ exist globally. It follows that any global *nonparametric* estimates (in the sense of minimal graphs) will grow like τ^{-1} as $\tau \to 0$. However, crucially, in the asymptotically flat setting $f_\tau(x)$ is bounded and even decays, uniformly in τ, for $|x|$ sufficiently large in the asymptotically flat region [Eic13, Proposition 5]. To study the convergence of the f_τ's in the "core," Schoen and Yau considered *parametric* estimates, i.e. geometric estimates for the graphs of f_τ. In fact, these graphs are C-*minimizing* for some constant C independent of τ (for this definition we refer to [DS93; Eic09]). For such hypersurfaces the compactness and regularity theory is essentially the same as for area minimizing hypersurfaces. It follows that the *graphs* of f_τ converge smoothly as hypersurfaces in $M \times \mathbb{R}$ as $\tau \to 0$. The components of the limit are either graphs of solutions to Jang's equation (10.4.7) or cylinders over MOTS or MITS in the data set (M, g, k).[4] Since $|f_\tau(x)| \lesssim 1$ for $|x| \gtrsim 1$, the limiting hypersurface contains a graphical component, defined over some set $\mathcal{U}$ which must contain a neighborhood of infinity. It is precisely this exterior graphical component that is studied in the works [SY81a; Eic13]. These basic properties of Jang's equation are summarized neatly in [Eic13, Proposition 7].

We can now describe our modification of Eichmair's argument in the boundary case.

Theorem 10.4.3. *Let $3 \le n \le 7$, and let (M^n, g, k) be a complete asymptotically flat initial data set with* nonempty *compact boundary ∂M such that the dominant energy condition holds on M and each component of ∂M is either weakly outer trapped or weakly inner untrapped. In the case $n = 3$, we also assume that* $\operatorname{tr}_g k = O(|x|^{-\gamma})$ *for some $\gamma > 2$. Then $E > 0$.*

Proof. Let (M, g, k) be as described in the hypotheses. By our DEC density theorem, Theorem Theorem 10.3.7, there exists a sequence of initial data $(g_j, k_j) \to (g, k)$ on M with harmonic asymptotics, satisfying the strict dominant energy condition, $\theta_j^+ < 0$ on $\partial^+ M$, $\theta_j^- > 0$ on $\partial^- M$, and

[3] In the case when $\operatorname{tr}_g k$ has a good sign, see [Met10, Theorem 3.4].

[4] The graphical components tend to $\pm\infty$ on approach to these cylinders. We say that the Jang graph "blows up" over the MOTS or MITS.

$|\operatorname{tr}_{g_j} k_j| \le C|x|^{-\gamma}$ uniformly in j when $n = 3$.[5]

The strict sign for $\theta^{\pm}$ on ∂M relative to (g_j, k_j) allows ∂M to act as a barrier for the capillary regularized Jang equation

$$H_{g_j}(f_\tau) - \operatorname{tr}_{g_j}(k_j)(f_\tau) = \tau f_\tau$$

as in [AM09, Proposition 3.5]. On the asymptotically flat end of M, we prescribe $f_\tau \to 0$ and proceed to solve as in [SY81a; AM09; Eic13]. We also obtain the usual parametric and nonparametric estimates associated to Jang's equation. Letting $\tau \to 0$, we obtain open sets $\mathcal{U}_j \subset M$, as described above, containing $\{|x| \ge R_0\}$ for some large R_0, equipped with a function $f_j \in C^{3,\alpha}_{-n+2+\eta}(\mathcal{U}_j)$ for any fixed $\eta > 0$ which solves Jang's equation and satisfies the properties proved in [Eic13, Proposition 7], with the same proof. The barrier property of ∂M implies that f_j is unbounded. In particular, f_j must blow up over some nonempty union of closed MOTS and MITS enclosing ∂M. More specifically, the graph of f_j must have at least one end that is asymptotic to a cylinder [Eic13, Proposition 7 (c)].

Claim 10.4.4. *After passing to a subsequence, the graphs of f_j converge in $C^{3,\alpha}_{\mathrm{loc}}$ to the graph of a Jang solution $f : \mathcal{U} \to \mathbb{R}$, which blows up over some nonempty union of closed MOTS and MITS in (M, g, k). With its induced metric, the graph of f is asymptotically flat with a nonzero number of ends that are asymptotically cylindrical.*

The main nontrivial claim here is that the property of having an asymptotically cylindrical end persists in the limit. By the Harnack inequality for Jang's equation, it suffices to show that the limiting function f is unbounded. The only thing we must rule out is cylindrical ends collapsing into the boundary ∂M. To do this, we extend the manifold M to a slightly larger manifold $\tilde{M}$ so that ∂M lies in the interior of M. We extend each metric g_j (including g) to $\tilde{M}$ so that on every compact set, $\|g_j - g\|_{C^{2,\alpha}} \to 0$. Then we apply the compactness and regularity theory for C-minimizing graphs on the extended manifold. The Jang graphs no longer approach the boundary, so the cylindrical ends cannot disappear in the limit. The claims about the blow-up locus being a collection of MOTS/MITS in (M, g, k) and the graph being asymptotically flat follow easily from [Eic13, Proposition 7]. This proves Claim 10.4.4.

We assume now that $E = 0$ and work to obtain a contradiction. It follows that $E_j \to 0$, and arguing as in [Eic13, Proposition 16], we see that the graph of f, which we denote by Σ, is scalar-flat and has zero mass. (This part of the argument is highly nontrivial but is agnostic to the presence

[5] The $n = 3$ claim follows from observing that the only term appearing in $\operatorname{tr}_{g_j} k_j$ (when written in terms of g, k, λ, and the deformations u, Yh, w) that is not directly controlled is the $\operatorname{tr}_g k$ term, which is controlled by assumption. See [Eic13, Proposition 15].

of the boundary ∂M; it only relies on the conclusion of Claim 10.4.4.) Now (Σ, g_Σ), viewed as a time-symmetric initial data set, must be diffeomorphic to $\mathbb{R}^n$. This can be viewed as an extension of the rigidity of the Riemannian positive mass theorem to manifolds with cylindrical ends. The argument is outlined in [Eic13], but is a special case of the more recent positive mass theorem with *arbitrary ends* [LUY21; LLU23]. One first shows that Σ is Ricci-flat and then has only one end by the Cheeger–Gromoll splitting theorem. However, Σ having only one end is in contradiction to Claim 10.4.4. □

10.A Second differential of the constraint-null expansion system

In this chapter, we utilize the inverse function theorem to perturb *families* of initial data sets. To this end, we need to control the constants appearing in the following "quantitative" version of the inverse function theorem (see [Lee19, Theorem A.43]).

Theorem 10.A.1. *Let X and Y be Banach spaces, $x_0 \in X$, and $r_0 > 0$. Suppose that $F : B_{r_0}(x_0) \to Y$ is C^1 and that DF is Lipschitz in this ball with constant C_L. Assume that $DF|_{x_0}$ is invertible with inverse bounded in operator norm by C_I. Then F^{-1} is defined on $B_{r_*}(y_0)$, where $y_0 = F(x_0)$ and r_* is explicitly determined by r_0, C_L, and C_I. Finally, $F^{-1}(B_{r_*}(y_0)) \subset B_{r_1}(x_0)$, where again r_1 is explicitly calculable in terms of r_0, C_L, and C_I.*

The estimates we require are as follows:

Lemma 10.A.2. *Let (M^n, g, π) be an asymptotically flat data set as in Section 10.2.1. Let $K_1 \subset T_{(1,0)}\mathcal{C}$ be a closed subspace and $K_2 \subset T_{(g,\pi)}\mathcal{D}$ be a finite-dimensional subspace. There exists a constant C_0 such that for any $r_0 > 0$ sufficiently small, the following is true.*

Let $(\gamma, \tau) \in \mathcal{D}$ with $\|(\gamma,\tau) - (g,\pi)\|_{\mathcal{D}} \leq r_0$ and define

$$\begin{aligned}\hat{\mathcal{P}}_{(\gamma,\tau)} : [(1,0) + K_1] \times K_2 &\to \mathcal{L} \times W^{1-\frac{1}{p},p}(\partial M)\\ ((u,Y),(h,w)) &\mapsto (\Phi,\Theta)[\Psi_{(\gamma,\tau)}(u,Y) + (h,w)].\end{aligned}$$

Then

$$\|D\hat{\mathcal{P}}_{(\gamma,\tau)}|_{((1,0),(0,0))} - D\hat{\mathcal{P}}_{(g,\pi)}|_{((1,0),(0,0))}\|_{L(K_1\times K_2, \mathcal{L}\times W^{1-\frac{1}{p},p})} \leq C_0 r_0 \tag{10.A.1}$$

and

$$\|D^2\hat{\mathcal{P}}_{(\gamma,\tau)}|_{((u,Y),(h,w))}\|_{L_2(K_1\times K_2,\mathcal{L}\times W^{1-\frac{1}{p},p})} \leq C_0 \tag{10.A.2}$$

for any $((u,Y),(h,w)) \in B_{r_0}((1,0),(0,0))$.

This lemma also holds if in the definition of the map $\hat{\mathcal{P}}$*, we use the modified constraint operator* $\overline{\Phi}_{(g,\pi)}$ *instead of* Φ.

Here $L_2(X,Y)$ refers to the space of bounded multilinear maps $X\times X \to Y$. Note that a Lipschitz bound for $D\hat{\mathcal{P}}_{(\gamma,\tau)}$ follows from the Hessian bound by the mean value theorem in Banach spaces. The proof proceeds with a computation of $D\Phi, D^2\Phi, D\Theta$, and $D^2\Theta$.

Lemma 10.A.3. *The first derivative (linearization) of the constraint operator is given by*

$$\begin{aligned} D\Phi|_{(g,\pi)}(h,w) = \Big(& -\Delta_g(\operatorname{tr}_g h) + \operatorname{div}_g(\operatorname{div}_g h) - \langle \operatorname{Ric}_g, h\rangle_g + \tfrac{2}{n-1}(\operatorname{tr}_g \pi)\left(\pi^{ij}h_{ij} + \operatorname{tr}_g w\right) \\ & - 2g_{kl}\pi^{ik}\pi^{jl}h_{ij} - 2\langle \pi, w\rangle_g, \\ & (\operatorname{div}_g w)^i - \tfrac{1}{2}g^{ij}\pi^{kl}\nabla_j h_{kl} + g^{ij}\pi^{kl}\nabla_k h_{jl} + \tfrac{1}{2}\pi^{ij}\nabla_j(\operatorname{tr}_g h)\Big) \end{aligned} \tag{10.A.3}$$

Schematically, the second derivative is given by

$$\begin{aligned} D^2\Phi|_{(g,\pi)}((h_1,w_1),(h_2,w_2)) = \Big(& \sum_{0\leq i_1+i_2\leq 2} \nabla^{i_1}h_1 * \nabla^{i_2}h_2 + \operatorname{Riem} * h_1 * h_2 \\ & + \pi*\pi*h_1*h_2 + w_1*w_2 + \pi*h_1*w_2 + \pi*h_2*w_1, \\ & w_1*\nabla h_2 + w_2*\nabla h_1 + \pi*h_1*\nabla h_2 + \pi*\nabla h_1 * h_2\Big). \end{aligned} \tag{10.A.4}$$

Here we use the usual schematic notation where $A*B$ *denotes linear combinations and contractions of the components of* A *and* B *with respect to the metric* g.

The schematic notation misses factors of g and g^{-1} but these are pointwise bounded by Morrey's inequality. In the following calculation, we use the shorthand $\delta_g F = DF|_g(h)$.

Proof. The formula for $D\Phi$ is well known in the literature [FM73]. It depends on the linearization of the scalar curvature, which can be found in [Lee19], for instance. To obtain the formula for $D^2\Phi$, we simply differentiate (10.A.3), making note of the following rules:

- $\delta_g \nabla T = \nabla h_2 * T + \nabla \delta_g T$ for any tensor T, and
- contractions produce terms of $h_2 *$ what was being contracted.

Finally, we also note that the variation of the Ricci tensor is given by

$$-2\delta_g R_{ij} = \Delta_L h_{2\,ij} + \nabla_i \nabla_j \operatorname{tr}_g h_2 - \nabla_i (\operatorname{div}_g h_2)_j - \nabla_j (\operatorname{div}_g h_2)_i,$$

where Δ_L is the Lichnerowicz Laplacian. In our schematic notation, this becomes

$$\delta_g \mathrm{Ric} = \nabla^2 h_2 + \mathrm{Riem} * h_2.$$

The variation in π is much more straightforward and (10.A.4) is easily obtained along these lines. □

Lemma 10.A.4. *The first derivative of the boundary null expansion is given by*

$$D\Theta|_{(g,\pi)}(h,w) = \tfrac{1}{2}\operatorname{tr}_{\partial M}(\nabla_\nu h) - \operatorname{div}_{\partial M}\omega - \tfrac{1}{2}h(\nu,\nu)H - h(\nu,\nu)\pi(\nu^\flat,\nu^\flat) - w(\nu^\flat,\nu^\flat), \tag{10.A.5}$$

where $\omega_i = h_{ij}\nu^j - h(\nu,\nu)\nu_i$ *and* $\nu^\flat$ *denotes the* 1*-form dual to* ν*. Schematically, the second derivative is given by*

$$D^2\Theta|_{(g,\pi)}((h_1,w_1),(h_2,w_2)) = \sum_{0\le i_1+i_2\le 1} \nabla^{i_1}h_1 * \nabla^{i_2}h_2 + h_1 * w_2 + h_2 * w_1 + w_1 * w_2, \tag{10.A.6}$$

where schematic notation here is omitting terms like ν *and* H*.*

Proof. We first compute the linearization of the normal. Varying $g(\nu,\nu) = 1$ gives

$$h(\nu,\nu) + 2g(\nu,\delta_g\nu) = 0,$$

while varying $g(X,\nu) = 0$ for $X \in T\Sigma$ gives

$$h(X,\nu) + g(X,\delta_g\nu) = 0.$$

It follows that

$$\delta_g\nu^i = -h^{ij}\nu_j + \tfrac{1}{2}h(\nu,\nu)\nu^i. \tag{10.A.7}$$

Secondly, we compute the linearization of the second fundamental form. For X and Y tangent to Σ, we have

$$A(X,Y) = -g(\nu,\nabla_X Y) = -g_{ij}\nu^i X^k \nabla_k Y^j.$$

Taking the variation, we have

$$\begin{aligned}\delta_g A(X,Y) &= -h_{ij}\nu^i X^k \nabla_k Y^j - g_{ij}\left(-h^{il}\nu_l + \tfrac{1}{2}h(\nu,\nu)\nu^i\right) X^k \nabla_k Y^j - g_{ij}\nu^i X^k \delta_g(\nabla_k Y^j)\\ &= \tfrac{1}{2}h(\nu,\nu)\left(-g_{ij}\nu^i X^k \nabla_k Y^j\right) - g_{ij}\nu^i X^k \delta_g(\nabla_k Y^j)\\ &= \tfrac{1}{2}h(\nu,\nu)A(X,Y) - \tfrac{1}{2}g_{ij}\nu^i g^{jm}\left(\nabla_k h_{lm} + \nabla_l h_{km} - \nabla_m h_{kl}\right) X^k Y^l\\ &= \tfrac{1}{2}\left(h(\nu,\nu)A_{kl} - \nu^m \nabla_k h_{lm} - \nu^m \nabla_l h_{km} + \nu^m \nabla_m h_{kl}\right) X^k Y^l.\end{aligned}$$

Now

$$\nu^m \nabla_k h_{lm} = \nabla_k \omega_l + h(\nu,\nu)A_{kl} - h_{ln}A_k{}^n,$$

so that finally

$$\delta_g A(X,Y) = \tfrac{1}{2}(\nabla_\nu h_{ij} - \nabla_i^{\partial M}\omega_j - \nabla_j^{\partial M}\omega_i + h_{ik}A_j{}^k + h_{jk}A_i{}^k - h(\nu,\nu)A_{ij})X^i Y^j. \tag{10.A.8}$$

The mean curvature of the boundary is given by

$$H = \operatorname{tr}_{\partial M} A = (g^{ij} - \nu^i\nu^j)A_{ij},$$

so taking the variation and using (10.A.8) yields

$$\delta_g H = \tfrac{1}{2}\operatorname{tr}_{\partial M}(\nabla_\nu h) - \operatorname{div}_{\partial M}\omega - \tfrac{1}{2}h(\nu,\nu)H. \tag{10.A.9}$$

The formula for $D\Theta$ follows easily, where also note that

$$\delta_g \nu^\flat = \tfrac{1}{2}h(\nu,\nu)\nu^\flat.$$

The schematic computation for $D^2\Theta$ also follows easily using the rules establised in the proof of Lemma 10.A.3. □

From these formulas, we deduce:

Lemma 10.A.5. *There exists a constant C_0 such that for any sufficiently small $r_0 > 0$ the following is true. If $(\gamma,\tau) \in \mathcal{D}$ satisfies $\|(\gamma,\tau) - (g,\pi)\|_{\mathcal{D}} \le r_0$, then*

$$\|D(\Phi,\Theta)|_{(\gamma,\tau)} - D(\Phi,\Theta)|_{(g,\pi)}\| \le C_0 r_0 \tag{10.A.10}$$

and

$$\|D^2(\Phi,\Theta)|_{(g,\pi)}\| \le C_0. \tag{10.A.11}$$

Proof. We first remark that the constants appearing in the Sobolev, Morrey, and trace inequalities associated to the metric γ can be bounded in terms of r_0. The first estimate (10.A.10) can be read off from the explicit formulas (10.A.3) and (10.A.5). For example, consider

$$g^{ij}\partial_i\partial_j(g^{kl}h_{kl}) - \gamma^{ij}\partial_i\partial_j(\gamma^{kl}h_{kl}).$$

We rewrite this as

$$(g^{ij}-\gamma^{ij})\partial_i\partial_j(\gamma^{kl}h_{kl}) + g^{ij}\partial_i\partial_j((g^{kl}-\gamma^{kl})h_{kl})$$

and from this it is not hard to see that the L^p_{-q} norm can be estimated by $\lesssim r_0\|h\|_{W^{2,p}_{-q}}$.

To prove the estimate (10.A.11), we examine the bilinear structure of the schematic formulas (10.A.4) and (10.A.6). For $D^2\Phi$, we put the highest number of derivatives in L^p_{-q} and the lowest number of derivatives in L^∞ using Morrey's inequality. Special care must be taken with the Riem $*$ $h_1 * h_2$ term, as the curvature is not assumed to be pointwise bounded. However, it is in L^p_{-q}, so we just put h_1 and h_2 in L^∞. Altogether, we obtain the estimate

$$\|D^2\Phi|_{(\gamma,\tau)}((h_1,w_1),(h_2,w_2))\|_{\mathcal{L}} \lesssim \|(h_1,w_1)\|_{W^{2,p}_{-q}\times W^{1,p}_{-q-1}}\|(h_2,w_2)\|_{W^{2,p}_{-q}\times W^{1,p}_{-q-1}}.$$

For $D^2\Theta$, we estimate each of the terms appearing in (10.A.6) in $W^{1-\frac{1}{p},p}(\partial\Omega)$. Terms with derivatives are handled using Lemma 10.2.4 instead of Morrey's inequality. Note that our schematic notation omits the normal ν_g and mean curvature H_g, however both of these are pointwise bounded in terms of γ. Therefore, we obtain the estimate

$$\|D^2\Theta|_{(\gamma,\tau)}\|_{W^{1-\frac{1}{p},p}} \lesssim \|(h_1,w_1)\|_{W^{2,p}_{-q}\times W^{1,p}_{-q-1}}\|(h_2,w_2)\|_{W^{2,p}_{-q}\times W^{1,p}_{-q-1}},$$

as desired. □

We can now prove the main result of this appendix, Lemma 10.A.2.

Proof of Lemma 10.A.2. We first define a function

$$\hat{\Psi}_{(\gamma,\tau)} : [(1,0)+K_1]\times K_2 \to \mathcal{D}$$
$$((u,Y),(h,w)) \mapsto \Psi_{(\gamma,\tau)}(u,Y) + (h,w),$$

so that

$$\hat{\mathcal{P}}_{(\gamma,\tau)} = (\Phi, \Theta) \circ \hat{\Psi}_{(\gamma,\tau)}.$$

By the chain rule for functions on Banach spaces,

$$D\hat{\mathcal{P}}_{(\gamma,\tau)}((v_1, Z_1), (h_1, w_1)) = D(\Phi, \Theta) \circ D\hat{\Psi}_{(\gamma,\tau)}((v_1, Z_1), (h_1, w_1)). \tag{10.A.12}$$

The second derivative is given by

$$\begin{aligned} D^2\hat{\mathcal{P}}_{(\gamma,\tau)}\Big(((v_1, Z_1), (h_1, w_1)), ((v_2, Z_2), (h_2, w_2))\Big) = \\ D^2(\Theta, \Phi)\Big(D\hat{\Psi}_{(\gamma,\tau)}((v_1, Z_1), (h_1, w_1)), D\hat{\Psi}_{(\gamma,\tau)}((v_2, Z_2), (h_2, w_2))\Big) \\ + D(\Theta, \Phi) \circ D^2\hat{\Psi}_{(\gamma,\tau)}\Big(((v_1, Z_1), (h_1, w_1)), ((v_2, Z_2), (h_2, w_2))\Big). \end{aligned} \tag{10.A.13}$$

The derivatives of $\hat{\Psi}_{(\gamma,\tau)}$ are given by

$$D\hat{\Psi}_{(\gamma,\tau)}((v_1, Z_1), (h_1, w_1)) = (su^{s-1}v_1\gamma + h_1, -\tfrac{3}{2}su^{-\frac{3}{2}s-1}v(\tau + \mathfrak{L}_\gamma Y) + u^{-\frac{3}{2}s}\mathfrak{L}_\gamma Z_1 + w_1) \tag{10.A.14}$$

and

$$\begin{aligned} D^2\hat{\Psi}_{(\gamma,\tau)}\Big(((v_1, Z_1), (h_1, w_1)), ((v_2, Z_2), (h_2, w_2))\Big) = (s(s-1)u^{s-2}v_1v_2, \\ \tfrac{3}{2}s(\tfrac{3}{2}s + 1)u^{-\frac{3}{2}s-2}v_1v_2(\tau + \mathfrak{L}_\gamma Y) - \tfrac{3}{2}su^{-\frac{3}{2}s-1}v_2\mathfrak{L}_\gamma Z_1 - \tfrac{3}{2}su^{-\frac{3}{2}s-1}v_1\mathfrak{L}_\gamma Z_2). \end{aligned} \tag{10.A.15}$$

In these formulas, the differentials are being evaluated at $((u, Y), (h, w))$ or $\hat{\Psi}_{(\gamma,\tau)}((u, Y), (h, w))$, wherever appropriate.

To prove (10.A.1), we use (10.A.12) for (γ, τ) and (g, π) at $((1, 0), (0, 0))$, which yields

$$D\hat{\mathcal{P}}_{(\gamma,\tau)} - D\hat{\mathcal{P}}_{(g,\pi)} = D(\Phi, \Theta)|_{(\gamma,\tau)}(D\hat{\Psi}_{(\gamma,\tau)} - D\hat{\Psi}_{(g,\pi)}) + (D(\Phi, \Theta)|_{(\gamma,\tau)} - D(\Phi, \Theta)|_{(g,\pi)})D\hat{\Psi}_{(g,\pi)}.$$

For (γ, τ) sufficiently close to (g, π), we may evidently estimate both of these terms (in operator norm) using (10.A.14) and the estimate (10.A.10).

To prove (10.A.2), we note that (10.A.13) implies

$$\|D^2\hat{\mathcal{P}}_{(\gamma,\tau)}|_{((u,Y),(h,w))}\| \leq \left\|D^2(\Theta,\Phi)|_{\hat{\Psi}_{(\gamma,\tau)}((u,Y),(h,w))}\right\| \cdot \left\|D\hat{\Psi}_{(\gamma,\tau)}|_{((u,Y),(h,w))}\right\|^2 + \left\|D(\Theta,\Phi)|_{\hat{\Psi}_{(\gamma,\tau)}((u,Y),(h,w))}\right\| \cdot \left\|D^2\hat{\Psi}_{(\gamma,\tau)}|_{((u,Y),(h,w))}\right\|.$$

For $((u,Y),(h,w))$ small, $\hat{\Psi}_{(\gamma,\tau)}((u,Y),(h,w))$ is close to (g,π) in $\mathcal{D}$, so we may apply (10.A.10) and (10.A.11). Furthermore, the same estimates may be derived for $D\hat{\Psi}$ and $D^2\hat{\Psi}$ from (10.A.14) and (10.A.15). This completes the proof of (10.A.2). □

Chapter 11

The positive mass theorem with arbitrary ends

11.1 The density theorem

The proofs of the theorems described in Section 9.3 rest on the following *density theorem*, which sharpens and generalizes previous results in this direction [SY81b; LP87; Kuw90].

Theorem 11.1.1. *Let (M^n, g), $n \geq 3$, be a Riemannian manifold, not assumed to be complete, with an asymptotically flat end $\mathcal{E}$ of Sobolev type (p, q), where $p > n$ and $q > \frac{n-2}{2}$. For any $\varepsilon > 0$, $\frac{n-2}{2} < q' < q$, and any compact set $K \subset M$, there exists another asymptotically flat metric $\tilde{g}$ of Sobolev type (p, q') on M with the following properties:*

1. *$\tilde{g}$ is harmonically flat outside a bounded set in $\mathcal{E}$, that is, $\tilde{g} = u^{\frac{4}{n-2}}\overline{g}$, where $\overline{g}$ is the Euclidean metric on $\mathcal{E}$, and u is a $\overline{g}$-harmonic function with expansion*

$$u(x) = 1 + \frac{A}{|x|^{n-2}} + O_\infty(|x|^{-n-1}),$$

2. *The ADM mass of $\tilde{g}$ is $2A$ and we have*

$$\|\tilde{g} - g\|_{W^{2,p}_{-q'}(K\cup\mathcal{E})} + \|R_{\tilde{g}} - R_g\|_{L^1(K\cup\mathcal{E})} + |2A - m_{\mathrm{ADM}}(\mathcal{E}, g)| < \varepsilon,$$

3. *$\sup_K |R_{\tilde{g}} - R_g| < \varepsilon$,*

4. *If $R_g(x) \geq 0$ at a point x, then $R_{\tilde{g}}(x) \geq 0$, and*

5. *The metrics g and $\tilde{g}$ are ε-close as bilinear forms everywhere on M:*

$$(1-\varepsilon)g \le \tilde{g} \le (1+\varepsilon)g.$$

Remark 11.1.2. It will be clear from the proof of the density theorem that we can also accommodate the case when M additionally has compact boundary components.

11.2 Asymptotic analysis in the presence of arbitrary ends

We begin this section with a precise statement of our definitions, which are slightly different than the usual ones because we allow for incompleteness and arbitrary ends.

Definition 11.2.1. Let (X, d) be a metric space and $(\overline{X}, d)$ be its completion. For example, $\overline{X}$ can be constructed by taking appropriate equivalence classes of Cauchy sequences. A point in $\overline{X} \setminus X$ is called a called a *point of incompleteness* for X. A set $S \subset X$ is said to be *complete* if its closure in X remains closed under the inclusion $X \to \overline{X}$.

Definition 11.2.2. Let M^n be a noncompact manifold with a distinguished end $\mathcal{E}$. We say that (M, g) possesses a *structure of infinity* along $\mathcal{E}$ if $\mathcal{E}$ possesses no points of incompleteness and there exists a diffeomorphism

$$\Phi : \mathcal{E} \to \mathbb{R}^n \setminus B_{r_0} \tag{11.2.1}$$

for some positive number r_0 and the coordinate norm $|x|$ diverges as we go out along the end. The set $\mathring{M} = M \setminus \mathcal{E}$ is called the *core*. Note that in our definition, the core is not assumed to be compact and (M, g) is not assumed to be complete. We will often identify $\mathcal{E}$ with the set $\{|x| \ge r_0\}$. The coordinates x^i induce a natural flat metric on $\mathcal{E}$, which we extend arbitrarily to a complete metric on all of M and denote by $\overline{g}$.

We also allow for M to have a boundary, but of course require ∂M to not intersect $\mathcal{E}$.

Definition 11.2.3. Let (M^n, g), $\mathcal{E}$, and Φ be as in the previous definition. Let N be a closed subset of M which contains $\mathcal{E}$ and such that $N \setminus \mathcal{E}$ is compact. (For example, we might take $N = \mathcal{E}$.) Given $k \in \mathbb{N}$, $p \ge 1$, and $s \in \mathbb{R}$, we define the *weighted Sobolev space* $W^{k,p}_s(N)$ to be the space of functions $u \in W^{k,p}_{\text{loc}}(N)$ with finite norm

$$\|u\|_{W^{k,p}_s(N)} = \|u\|_{W^{k,p}(N\setminus\mathcal{E})} + \sum_{i=0}^{k} \|\partial^i u\|_{L^p_{s-i}(\mathcal{E})},$$

where the weighted L^p norm is defined by

$$\|u\|_{L^p_s(\mathcal{E})} = \left(\int_{\mathcal{E}} ||x|^{-s} u|^p \, \frac{dx}{|x|^n} \right)^{\frac{1}{p}}.$$

Note that $r^s \notin L^p_s$ but $r^{s-\delta} \in L^p_s$ for any $\delta > 0$. Note also that $L^p_{s'} \subset L^p_s$ if $s' \geq s$. We also remark that in our definition, the weighted spaces are to be constructed relative to the reference metric $\overline{g}$ and do not reference the (later) geometric metric g at all. This is because g will be changing at some points in the proof and we do not wish to have the norm changing as well. (This is only a minor point.)

Definition 11.2.4. Let (M^n, g) be a noncompact smooth Riemannian manifold possessing a structure of infinity Φ along $\mathcal{E}$. Let $p > n$ and $q > \frac{n-2}{2}$. We say that $\mathcal{E}$ is *asymptotically flat (AF) of Sobolev type* (p, q) if in the coordinates x^i defined by Φ,

$$g_{ij} - \delta_{ij} \in W^{2,p}_{-q}(\mathcal{E}).$$

Furthermore, we assume that the scalar curvature of g, R_g, lies in $L^1(\mathcal{E})$. A (p, q) Sobolev asymptotically flat metric is also (p, q') Sobolev asymptotically flat for any $q' < q$.

The *ADM mass of* $\mathcal{E}$ is defined by

$$m_{\mathrm{ADM}}(\mathcal{E}, g) = \lim_{r \to \infty} \frac{1}{2(n-1)\omega_{n-1}} \int_{|x|=r} (g_{ij,i} - g_{ii,j}) \frac{x^j}{|x|} \, d\mu_{S_r, \overline{g}}. \tag{11.2.2}$$

The condition $R_g \in L^1(\mathcal{E})$ guarantees that this is well-defined.

Remark 11.2.5. Our definition of asymptotically flat is the weakest "standard" definition used in the positive mass literature. A stronger condition would be *asymptotically flat with pointwise decay* q, that is,

$$g_{ij}(x) = \delta_{ij} + O_2(|x|^{-q}).$$

This will be of Sobolev type (p, q) for any $p > n$ and $q' < q$. Yet another definition would be with weighted Hölder spaces.

One reason to consider Sobolev decay is that Jang graphs in asymptotically flat initial data sets with pointwise decay are only Sobolev asymptotically flat [Eic13, Proposition 7].

Since $p > n$, the weighed Morrey inequality implies $g_{ij} - \delta_{ij} \in C^{1,\alpha}_{-q}$ for some $\alpha \in (0, 1)$. We define the *pointwise decay constant* $C_g < \infty$ of g by

$$C_g \doteq \sum_{ijk} \sup_{x \in \mathcal{E}} \left(|x|^q |g_{ij} - \delta_{ij}| + |x|^{q+1} |\partial_k g_{ij}| \right) . \tag{11.2.3}$$

The following lemma is standard, cf. [Lee19, Lemma 3.35].

Lemma 11.2.6. *Suppose that g_i is a sequence of (p,q) Sobolev asymptotically flat metrics converging in $W^{2,p}_{-q}(N)$ for N as in Definition Definition 11.2.3. Assume that $R_{g_i} \to R_g$ in $L^1(\mathcal{E})$. Then*

$$m_{\mathrm{ADM}}(\mathcal{E}, g_i) \to m_{\mathrm{ADM}}(\mathcal{E}, g).$$

This motivates the following definition of closeness of asymptotically flat metrics.

Definition 11.2.7. Given $\varepsilon > 0$, we say that two (p,q) Sobolev asymptotically flat metrics g_1 and g_2 on M are *ε-close in the asymptotic topology* if the following inequalities are satisfied:

(i) $\|g_1 - g_2\|_{W^{2,p}_{-q}(\mathcal{E})} < \varepsilon$,

(ii) $\|R_{g_1} - R_{g_2}\|_{L^1(\mathcal{E})} < \varepsilon$,

(iii) $|m_{\mathrm{ADM}}(\mathcal{E}, g_1) - m_{\mathrm{ADM}}(\mathcal{E}, g_2)| < \varepsilon$,

(iv) $(1-\varepsilon) g_1 \le g_2 \le (1+\varepsilon) g_1$ as bilinear forms, globally.

To begin our construction, we define a distinguished function ρ on M.

Lemma 11.2.8. *Let $(M, g, \mathcal{E})$ be asymptotically flat, not assumed to be complete. There exists a C^∞ proper function $\rho : M \to (0, \infty)$ which equals $|x|$ on $\mathcal{E}$ and for every sequence $\{p_i\} \subset M \setminus \mathcal{E}$ which eventually leaves every compact set, $\rho(p_i) \to 0$. We may additionally suppose that $\rho < r_0$ on $M \setminus \mathcal{E}$.*

Proof. By a partition of unity argument, there exists a function $\sigma : M \to [1, \infty)$ such that $\sigma \to \infty$ outside of every compact set. We then interpolate between σ^{-1} and $|x|$ near $|x| = r_0$ to obtain the desired function ρ. □

We use ρ to construct a "compact" exhaustion of the core which avoids incomplete points: For $\sigma > 0$, let $M_\sigma = \{\rho \ge \sigma\}$. For σ a regular value of ρ, ∂M_σ is a smooth hypersurface. Then M_σ is an asymptotically flat manifold with boundary and containing no incomplete points.

The analytical content of this section is to construct solutions to certain Schrödinger equations on M_σ and let $\sigma \to 0$. This is only possible because the potentials will vanish identically away from

M_{σ_0} for some fixed $\sigma_0 > 0$. We will need precise a priori estimates and so work very carefully and keep track of the constants. First, we note a standard Euclidean-type Sobolev inequality on M_σ.

Lemma 11.2.9 (Sobolev inequality for M_σ)**.** *Let $(M^n, \overline{g}, \mathcal{E})$ be a fixed asymptotically flat manifold, and define M_σ as above. Let $\Lambda > 0$ be a fixed constant, and assume that g is another metric on M such that $\Lambda^{-1}\overline{g} \leq g \leq \Lambda\overline{g}$. Then for all $\sigma > 0$, there exists a constant C_σ depending only on σ and Λ, such that if $u : M_\sigma \to \mathbb{R}$ is a C^1 function for which $du \in L^2(M_\sigma)$, then $u \in L^{\frac{2n}{n-2}}(M_\sigma)$ and*

$$\left(\int_{M_\sigma} |u|^{\frac{2n}{n-2}}\, d\mu_g\right)^{\frac{n-2}{n}} \leq C_\sigma \int_{M_\sigma} |du|_g^2\, d\mu_g. \tag{11.2.4}$$

Proof. For a fixed metric $\overline{g}$, this was proved in [SY79a], and then it is clear that one can switch from $\overline{g}$ to g simply by multiplying C_σ by a power of the uniform bound Λ. □

Remark 11.2.10. It is crucial for the method here that this Sobolev inequality holds without a "boundary condition" for u on ∂M_σ. Indeed, imposing $u \to 0$ along $\mathcal{E}$ acts like a boundary condition.

We also require a scale-broken weighted elliptic estimate on M_σ.

Lemma 11.2.11. *Let $(M^n, g, \mathcal{E})$ be an asymptotically flat manifold of Sobolev type (p, q) and $0 < \delta < \sigma_0$. Then there exists a constant C, depending only on n, p, q, δ, and the $C^{1,\alpha}_{-q}(M_{\sigma_0/2})$ norm of $g - \overline{g}$, such that for any $u \in W^{2,p}_{-q}(M_{\sigma_0})$,*

$$\|u\|_{W^{2,p}_{-q}(M_{\sigma_0})} \leq C\left(\|\Delta_g u\|_{L^p_{-q-2}(M_{\sigma_0-\delta})} + \|u\|_{L^{\frac{2n}{n-2}}(M_{\sigma_0-\delta})}\right). \tag{11.2.5}$$

We emphasize that this is an "interior-type" estimate since $\overline{M}_{\sigma_0} \subset M_{\sigma_0-\delta}$. We also take this opportunity to define a cutoff function that will be used later as well. Let $\chi(|x|)$ be a radial cutoff in $\mathbb{R}^n$ which is one on the ball B_1 and zero outside the ball B_2. For $\lambda \geq r_0$ we define χ_λ on M by setting it equal to $\chi(|x|/\lambda)$ on $\mathcal{E}$ and extending to the rest of the manifold by one.

Proof of Lemma 11.2.11. Let $\lambda \geq r_0$ and set $u_0 = \chi_\lambda u$, $u_1 = (1-\chi_\lambda)u$, so that $u = u_0 + u_1$. Then $u_1 \in W^{2,p}_{-q}(\mathbb{R}^n)$, so we have the sharp estimate [McO79, Theorem 0]

$$\|u_1\|_{W^{2,p}_{-q}(\mathbb{R}^n)} \leq C\|\Delta_{\overline{g}} u_1\|_{L^p_{-q-2}(\mathbb{R}^n)}.$$

Since $\Delta_g - \Delta_{\overline{g}}$ is the zero operator for $|x| \leq \lambda$, we then have

$$\begin{aligned}\|\Delta_{\overline{g}} u_1\|_{L^p_{-q-2}(\mathbb{R}^n)} &\leq \|(\Delta_g - \Delta_{\overline{g}}) u_1\|_{L^p_{-q-2}(\mathbb{R}^n)} + \|\Delta_g u_1\|_{L^p_{-q-2}(\mathbb{R}^n)} \\ &\leq C\lambda^{-q} \|u_1\|_{W^{2,p}_{-q}(\mathbb{R}^n)} + \|\Delta_g u_1\|_{L^p_{-q-2}(\mathbb{R}^n)}.\end{aligned}$$

Choosing λ large enough (depending only on the quantities listed in the statement of the lemma), we have

$$\|u_1\|_{W^{2,p}_{-q}(\mathbb{R}^n)} \leq C\|\Delta_g u_1\|_{L^p_{-q-2}(\mathbb{R}^n)}.$$

Inserting the definition of u_1 into the right-hand side and carrying out the differentiations, we obtain error terms over the fixed compact set $K = \operatorname{spt} \nabla\chi_\lambda$. Lower order terms are moved to the left-hand side by interpolation, leaving us with

$$\|u_1\|_{W^{2,p}_{-q}(\mathbb{R}^n)} \leq C \left(\|\Delta_g u\|_{L^p_{-q-2}(M)} + \|u\|_{L^p(K)} \right).$$

If $p \leq \frac{2n}{n-2}$ (only possible when $n = 3$ since $p > n$), we use Hölder's inequality to estimate $\|u\|_{L^p(K)} \leq C\|u\|_{L^{\frac{2n}{n-2}}(K)}$. If $p > \frac{2n}{n-2}$, we trivially estimate $\|u\|_{L^p(K)} \leq C\|u\|_{L^\infty}$ and use the De Giorgi–Nash–Moser theorem to estimate

$$\|u\|_{L^\infty(K)} \leq C \left(\|\Delta_g u\|_{L^p(K')} + \|u\|_{L^{\frac{2n}{n-2}}(K')} \right),$$

where K' is a slightly larger compact set containing K. The unweighted L^p norm on the right-hand side can be absorbed into a weighted L^p norm. Altogether, we have proved

$$\|u_1\|_{W^{2,p}_{-q}(\mathbb{R}^n)} \leq C \left(\|\Delta_g u\|_{L^p_{-q-2}(M)} + \|u\|_{L^{\frac{2n}{n-2}}(K)} \right). \tag{11.2.6}$$

The estimate

$$\|u_0\|_{W^{2,p}_{-q}(M_{\sigma_0})} \leq C \left(\|\Delta_g u\|_{L^p_{-q-2}(M)} + \|u\|_{L^{\frac{2n}{n-2}}(M_{\sigma_0-\delta})} \right). \tag{11.2.7}$$

for u_0 is much simpler since it vanishes on $\mathcal{E}$. It may be achieved by applying the L^p theory for elliptic equations on the compact set $M_{\sigma_0-\delta} \setminus \{|x| \geq \lambda\}$. Combining (11.2.6) and (11.2.7) gives (11.2.5), as desired. □

The following is the main result of this section and can be compared to [SY79a, Lemma 3.2].

Proposition 11.2.12. *Let $(M^n, g, \mathcal{E})$ be a (p, q) asymptotically flat manifold with $p > n$, $q > \frac{n-2}{2}$, and $\sigma_0 > 0$. Let V be a smooth function on M, and let V^- denote its negative part. Assume that* $\mathrm{spt}(V) \subset M_{\sigma_0}$ *and* $\|V^-\|_{L^{\frac{n}{2}}(M_{\sigma_0})} < C_{\sigma_0}^{-1}$, *where C_{σ_0} is a constant as in Lemma 11.2.9 that works for this specific g). Then the Neumann problem*

$$(-\Delta_g + V, \nu_g|_{\partial M_\sigma}) : W^{2,p}_{-q}(M_\sigma) \to L^p_{-q-2}(M_\sigma) \times W^{2-\frac{1}{p},p}(\partial M_\sigma)$$

is an isomorphism for any $\sigma \in (0, \sigma_0)$ a regular value of ρ.

There exist constants $\varepsilon_0 > 0$ and C, depending only on n, p, q, and the $C^{1,\alpha}_{-q}(M_{\sigma_0/2})$ norm of $g - \overline{g}$, such that if $\sigma \in (0, \frac{\sigma_0}{2})$, $f \in L^p_{-q-2}(M_{\sigma_0}) \cap L^{\frac{2n}{n+2}}(M_{\sigma_0})$ is also supported in M_{σ_0}, $\|V^-\|_{L^{\frac{n}{2}}(M_{\sigma_0})} + \|V\|_{L^p_{-q-2}(M_{\sigma_0})} < \varepsilon_0$, and $u \in W^{2,p}_{-q}(M_\sigma)$ solves

$$-\Delta_g u + Vu = f \quad \text{in } M_\sigma \tag{11.2.8}$$

$$\nu_g(u) = 0 \quad \text{on } \partial M_\sigma, \tag{11.2.9}$$

then

$$\sup_{M_\sigma} |u| + \|u\|_{W^{2,p}_{-q}(M_{\sigma_0})} \le C \left(\|f\|_{L^p_{-q-2}(M_{\sigma_0})} + \|f\|_{L^{\frac{2n}{n+2}}(M_{\sigma_0})} \right). \tag{11.2.10}$$

In practice, we will only apply this lemma in the situation where $f = -V$.

Remark 11.2.13. To get a sense for how we will use this proposition to prove the density theorem, see the beginning of Section 11.3. Our basic observation is that the equations to be solved have "small" potentials which are identically zero away from $\mathcal{E}$. (See (11.3.1) below.) The control over $\sup_{M_\sigma} |u|$, rather than simply $\sup_{M_{\sigma_0}} |u|$ in Proposition 11.2.12 comes from the maximum principle. This addresses the essential issue of completeness because we can make this as small as we like. Eichmair used a similar observation to prove a version of the positive mass theorem for manifolds with cylindrical ends [Eic13].

Proof of Proposition 11.2.12. Since the problem is self-adjoint, we only need to show that the operator has no kernel. Suppose $w \in W^{2,p}_{-q}(M_\sigma)$ satisfies

$$-\Delta_g w + Vw = 0 \quad \text{in } M_\sigma$$

$$\nu_g(w) = 0 \quad \text{on } \partial M_\sigma.$$

Invoking Lemma 11.2.9 and then integrating by parts (which we justify in Lemma 11.2.14 below), we have

$$\begin{aligned}
\left(\int_{M_{\sigma_0}} |w|^{\frac{2n}{n-2}}\,d\mu_g\right)^{\frac{n-2}{n}} &\leq C_{\sigma_0}\int_{M_\sigma} |dw|_g^2\,d\mu_g \\
&= C_{\sigma_0}\int_{M_\sigma} -(\Delta_g w)w\,d\mu_g \\
&= C_{\sigma_0}\int_{M_\sigma} -Vw^2\,d\mu_g \\
&\leq C_{\sigma_0}\int_{M_{\sigma_0}} V^- w^2\,d\mu_g \\
&\leq C_{\sigma_0}\left(\int_{M_{\sigma_0}} |V^-|^{\frac{n}{2}}\,d\mu_g\right)^{\frac{2}{n}}\left(\int_{M_{\sigma_0}} w^{\frac{2n}{n-2}}\,d\mu_g\right)^{\frac{n-2}{n}}.
\end{aligned}$$

Our hypothesis that $\|V^-\|_{L^{\frac{n}{2}}(M_{\sigma_0})} < C_{\sigma_0}^{-1}$ implies that w vanishes on M_{σ_0}. But w is harmonic on a neighborhood of $M_\sigma \setminus M_{\sigma_0}$ in M_σ, so it must vanish on all of M_σ.

Now suppose u satisfies (11.2.8) and (11.2.9). To prove (11.2.10), we first note that since u is harmonic on $M_\sigma \setminus M_{\sigma_0}$ and satisfies a Neumann condition on ∂M_σ, the Hopf lemma implies

$$\sup_{M_\sigma} |u| = \sup_{M_{\sigma_0}} |u|,$$

and by the weighted Morrey inequality,

$$\sup_{M_{\sigma_0}} |u| \leq C\|u\|_{W^{1,p}_{-q}(M_{\sigma_0})} \leq C\|u\|_{W^{2,p}_{-q}(M_{\sigma_0})},$$

where C depends only on σ_0. We now apply the interior estimate (11.2.5) with $\delta = \sigma_0/2$ and note that f and V are supported in M_{σ_0} to obtain

$$\begin{aligned}
\sup_{M_\sigma}|u| + \|u\|_{W^{2,p}_{-q}(M_{\sigma_0})} &\leq C\left(\|Vu\|_{L^p_{-q-2}(M_{\sigma_0/2})} + \|f\|_{L^p_{-q-2}(M_{\sigma_0/2})} + \|u\|_{L^{\frac{2n}{n-2}}(M_{\sigma_0/2})}\right) \\
&\leq C\left(\|V\|_{L^p_{-q-2}(M_{\sigma_0})}\sup_{M_\sigma}|u| + \|f\|_{L^p_{-q-2}(M_{\sigma_0})} + \|u\|_{L^{\frac{2n}{n-2}}(M_{\sigma_0/2})}\right).
\end{aligned}$$

By choosing ε_0 small enough, $\|V\|_{L^p_{-q-2}}(M_{\sigma_0})$ will be small enough so that we have

$$\sup_{M_\sigma}|u| + \|u\|_{W^{2,p}_{-q}(M_{\sigma_0})} \leq 2C\left(\|f\|_{L^p_{-q-2}(M_{\sigma_0})} + \|u\|_{L^{\frac{2n}{n-2}}(M_{\sigma_0/2})}\right). \tag{11.2.11}$$

It only remains to estimate $\|u\|_{L^{\frac{2n}{n-2}}(M_{\sigma_0/2})}$. Using the Sobolev inequality and integrating by parts as we did for w above,

$$\begin{aligned}
C_{\sigma_0/2}^{-1}\left(\int_{M_{\sigma_0/2}}|u|^{\frac{2n}{n-2}}\,d\mu_g\right)^{\frac{n-2}{n}} &\leq \int_{M_{\sigma_0/2}}(f-Vu)u\,d\mu_g \\
&\leq \left(\int_{M_{\sigma_0/2}}|f|^{\frac{2n}{n+2}}\,d\mu_g\right)^{\frac{n+2}{2n}}\left(\int_{M_{\sigma_0/2}}|u|^{\frac{2n}{n-2}}\,d\mu_g\right)^{\frac{n-2}{2n}} \\
&\quad + \left(\int_{M_{\sigma_0/2}}|V^-|^{\frac{n}{2}}\,d\mu_g\right)^{\frac{2}{n}}\left(\int_{M_{\sigma_0/2}}|u|^{\frac{2n}{n-2}}\,d\mu_g\right)^{\frac{n-2}{n}}.
\end{aligned}$$

$$C_{\sigma_0/2}^{-1}\|u\|^2_{L^{\frac{2n}{n-2}}(M_{\sigma_0/2})} \leq \|f\|_{L^{\frac{2n}{n+2}}(M_{\sigma_0})}\|u\|_{L^{\frac{2n}{n-2}}(M_{\sigma_0/2})} + \|V^-\|_{L^{\frac{n}{2}}(M_{\sigma_0})}\|u\|^2_{L^{\frac{2n}{n-2}}(M_{\sigma_0/2})}.$$

So long as $\varepsilon_0 < \frac{1}{2}C_{\sigma_0/2}^{-1}$, we can absorb the V^- term to obtain

$$\|u\|_{L^{\frac{2n}{n-2}}(M_{\sigma_0/2})} \leq 2C_{\sigma_0/2}\|f\|_{L^{\frac{2n}{n+2}}(M_{\sigma_0})} \tag{11.2.12}$$

and the result follows. □

Lemma 11.2.14. *Let $w \in W^{2,p}_{-q}(M_\sigma)$ for $p > n$. Then $dw \in L^2(M_\sigma)$ and if $\nu_g(w) = 0$ on ∂M_σ, then*

$$\int_{M_\sigma}(-\Delta_g w)w\,d\mu_g = \int_{M_\sigma}|dw|^2_g\,d\mu_g.$$

Proof. By Morrey's inequality, $w \in C^1_{-q}$ so both sides of the equality are defined. Furthermore, integrating by parts on the compact domain $\{\sigma \leq \rho \leq r\}$, we pick up a boundary term

$$\int_{|x|=r}\nu_g(w)w\,d\mu_{S_r,g}.$$

By inspection the integrand is $O(r^{-2q-1})$, so it must disappear in the limit since $q > \frac{n-2}{2}$. □

Because we will use them many times, we record some basic facts about conformal metrics constructed using the previous proposition.

Proposition 11.2.15. *Let $(M^n, g, \mathcal{E})$ be a (p, q) asymptotically flat manifold with $p > n$, $q > \frac{n-2}{2}$, and $\sigma_0 > 0$. Let V be a smooth integrable function on M that is compactly supported in M_{σ_0}. There exists a constant $\varepsilon_0 > 0$, depending only on n, p, q, and the $C^{1,\alpha}_{-q}(M_{\sigma_0/2})$ norm of $g - \overline{g}$, such that if*

$$\|V^-\|_{L^{\frac{n}{2}}(M_{\sigma_0})} + \|V\|_{L^p_{-q-2}(M_{\sigma_0})} + \|V\|_{L^{\frac{2n}{n+2}}(M_{\sigma_0})} < \varepsilon_0,$$

then there exists a globally defined function u on M such that

$$-a\Delta_g u + Vu = 0,$$

everywhere, where $a = 4\frac{n-1}{n-2}$, such that $u - 1 \in W^{2,p}_{-q}(M_{\sigma_0})$. Moreover, u has positive upper and lower bounds, and we can define the metric $\tilde{g} = u^{\frac{4}{n-2}}g$. This metric $\tilde{g}$ is asymptotically flat of Sobolev type (p,q), with scalar curvature

$$R_{\tilde{g}} = (R_g - V)u^{-\frac{4}{n-2}} \tag{11.2.13}$$

and ADM mass

$$m_{\text{ADM}}(\mathcal{E}, \tilde{g}) = m_{\text{ADM}}(\mathcal{E}, g) - \frac{1}{2(n-1)\omega_{n-1}} \int_M Vu\, d\mu_g. \tag{11.2.14}$$

Proof. We first invoke Proposition 11.2.12 with $f = -V$ to see that for any $\sigma \in (0, \sigma_0)$ that is a regular value of ρ, there exists solution u_σ to the problem

$$-a\Delta_g u_\sigma + Vu_\sigma = 0 \quad \text{in } M_\sigma \tag{11.2.15}$$

$$\nu_g(u_\sigma) = 0 \quad \text{on } \partial M_\sigma, \tag{11.2.16}$$

$$u_\sigma - 1 \in W^{2,p}_{-q}(M_\sigma) \tag{11.2.17}$$

Using the global estimates (11.2.10) together with local elliptic theory, it follows that for some sequence of σ's converging zero, the u_σ's converge locally in $W^{2,p}$ to some globally define function u. By (11.2.10) and smallness of ε_0, we can ensure that u has a positive upper and lower bound. (In fact, we can choose $\frac{1}{2} < u < \frac{3}{2}$.) The formula (11.2.13) follows from the standard formula for scalar curvature of a conformal metric.

It is a standard fact that $(1+v)^{\frac{4}{n-2}} - 1 \in W^{2,p}_{-q}$ if $v \in W^{2,p}_{-q}$ [Kuw90, Lemma 2.2(i)]. We claim that $(u^{\frac{4}{n-2}} - 1)g_{ij} \in W^{2,p}_{-q}$, for then

$$\tilde{g}_{ij} - \delta_{ij} = (u^{\frac{4}{n-2}} g_{ij} - 1)g_{ij} + (g_{ij} - \delta_{ij}) \in W^{2,p}_{-q}.$$

It is an easy matter to check that the claim is true, and thus $\tilde{g}$ is asymptotically flat of Sobolev type (p,q).

Since V is integrable, (11.2.13) implies that $R_{\tilde{g}}$ is as well. To compute the mass of $\tilde{g}$, we compute the masses of the metrics $\tilde{g}_\sigma = u_\sigma^{\frac{4}{n-2}} g$ on M_σ. A standard computation shows that

$$\begin{aligned}
m_{\mathrm{ADM}}(\mathcal{E}, \tilde{g}_\sigma) - m_{\mathrm{ADM}}(\mathcal{E}, g) &= \lim_{r\to\infty} \frac{-2}{(n-2)\omega_{n-1}} \int_{|x|=r} \nu_g(u_\sigma)\, d\mu_{S_r,g} \\
&= \frac{-2}{(n-2)\omega_{n-1}} \int_{M_\sigma} \Delta_g u_\sigma \, d\mu_g \\
&= \frac{-1}{2(n-1)\omega_{n-1}} \int_M V u_\sigma \, d\mu_g,
\end{aligned}$$

where the inner boundary term of the integration by parts vanishes due to the Neumann condition for u_σ. The formula (11.2.14) now follows because $m_{\mathrm{ADM}}(\mathcal{E}, \tilde{g}_\sigma) \to m_{\mathrm{ADM}}(\mathcal{E}, \tilde{g})$ by Lemma 11.2.6, and the corresponding integrals obviously as well since $u_\sigma \to u$ uniformly on compact sets. $\square$

11.3 Proof of the density theorem, Theorem 11.1.1

Let $(M^n, g, \mathcal{E})$, (p, q), ρ, ε, and K be as in the statement of Theorem 11.1.1 and Section 11.2. Let $\chi_\lambda(x) = \chi(x/\lambda)$ be the family of cutoff functions defined below Lemma 11.2.11. Define $g_\lambda = \chi_\lambda g + (1 - \chi_\lambda)\overline{g}$, where $\overline{g}$ is the background flat metric on $\mathcal{E}$. We may take σ_0 to be any positive regular value of ρ such that $K \subset M_{\sigma_0}$.

For $0 < \sigma < \sigma_0$ a regular value of ρ we consider the conformal Laplace-type equation

$$\begin{aligned}
-a\Delta_\lambda u_{\lambda,\sigma} + (R_\lambda - \chi_\lambda R_g) u_{\lambda,\sigma} &= 0 \quad \text{in } M_\sigma, \\
\nu_g(u_{\lambda,\sigma}) &= 0 \quad \text{on } \partial M_\sigma, \\
u_{\lambda,\sigma} &\to 1 \quad \text{on } \mathcal{E},
\end{aligned} \tag{11.3.1}$$

where $u_{\lambda,\sigma} : M_\sigma \to \mathbb{R}$ and Δ_λ and R_λ refer to the Laplacian and scalar curvature of the metric g_λ, respectively. Setting $v_{\lambda,\sigma} = u_{\lambda,\sigma} - 1$, we therefore solve

$$\begin{aligned}
-a\Delta_\lambda v_{\lambda,\sigma} + (R_\lambda - \chi_\lambda R_g) v_{\lambda,\sigma} &= -(R_\lambda - \chi_\lambda R_g) \quad \text{in } M_\sigma, \\
\nu_g(v_{\lambda,\sigma}) &= 0 \quad \text{on } \partial M_\sigma, \\
v_{\lambda,\sigma} &\in W^{2,p}_{-q'}(M_\sigma)
\end{aligned} \tag{11.3.2}$$

as in Proposition 11.2.12 and Proposition 11.2.15, where $q' \in (\frac{2n}{n-2}, q)$. (The reason for the change from q to q' will become apparent in the proof.) To verify the hypotheses of these propositions, we first require a basic integration lemma.

Lemma 11.3.1. *Let $f \in L^p_{-q-2}$, where $p > n$ and $q > \frac{n-2}{2}$. Let A_i denote the dyadic annulus $2^i \leq |x| \leq 2^{i+1}$. For any $s \in [\frac{2n}{n+2}, p]$ and i sufficiently large there exists a constant C independent of f and i such that*

$$\|f\|_{L^s(A_i)} \leq C2^{-\eta i}\|f\|_{L^p_{-q-2}},$$

where $\eta = q - \frac{n-2}{2} > 0$.

Proof. We first use Hölder's inequality to estimate

$$\begin{aligned}\left(\int_{A_i} |f|^s \, dx\right)^{\frac{p}{s}} &\leq \mathrm{vol}(A_i)^{\frac{p}{s}-1} \int_{A_i} |f|^p \, dx \\ &\leq C(2^i)^{n(\frac{p}{s}-1)} \int_{A^i} |f|^p \, dx.\end{aligned}$$

Now

$$(2^i)^{n(\frac{p}{s}-1)} = \left((2^i)^{\frac{n}{s}-(q+2)}\right)^p \cdot \left((2^i)^{q+2}\right)^p (2^i)^{-n},$$

so that

$$\begin{aligned}(2^i)^{n(\frac{p}{s}-1)} \int_{A^i} |f|^p \, dx &= \left((2^i)^{\frac{n}{s}-(q+2)}\right)^p \int_{A_i} |(2^i)^{q+2} f|^p \, \frac{dx}{(2^i)^n} \\ &\leq C \left((2^i)^{\frac{n}{s}-(q+2)}\right)^p \int_{A_i} ||x|^{q+2} f|^p \, \frac{dx}{|x|^n} \\ &\leq C(2^{-i\eta})^p \|f\|^p_{L^p_{-q-2}}.\end{aligned}$$

□

Lemma 11.3.2. *Assume the hypotheses and notation of Theorem 11.1.1. There exist constants C and I with the following property. Let $\lambda_i = 2^i$. Then for any $i \geq I$ and $\sigma \in (0, \frac{\sigma_0}{2})$, there exists a unique solution $v_{\lambda_i,\sigma} \in W^{2,p}_{-q'}(M_\sigma)$ of (11.3.2) satisfying*

$$\lim_{i\to\infty} \left(\sup_{M_\sigma} |v_{\lambda_i,\sigma}| + \|v_{\lambda_i,\sigma}\|_{W^{2,p}_{-q'}(M_{\sigma_0})}\right) = 0.$$

Proof. We apply Proposition 11.2.12 with $g = g_{\lambda_i}$ and $V = -f = R_{\lambda_i} - \chi_{\lambda_i} R_g$. We only need to check that the relevant estimates are satisfied. To obtain the required weighted L^p smallness for the potential V, we will be required to lower the decay rate from q to q'.

First, we claim that the constants in the pointwise decay for g_λ, (11.2.3), are uniformly bounded. Indeed, in coordinates,

$$g_{\lambda\, ij} - \delta_{ij} = \chi_\lambda (g_{ij} - \delta_{ij})$$

and

$$\partial_k g_{\lambda\, ij} = \partial_k \chi_\lambda (g_{ij} - \delta_{ij}) + \chi_\lambda \partial_k g_{ij}.$$

We clearly have

$$|g_{\lambda\, ij} - \delta_{ij}| \leq C_g |x|^{-q} \tag{11.3.3}$$

independently of λ. For the derivative, note that $|\partial \chi_\lambda| \approx \lambda^{-1} \approx |x|^{-1}$ on $\mathrm{spt}(\partial\chi_\lambda)$, so

$$|\partial_k g_{\lambda\, ij}| \leq C|x|^{-1} \cdot C_g |x|^{-q} + C_g |x|^{-q-1}, \tag{11.3.4}$$

which is also uniform in λ.

We now claim that

$$\lim_{\lambda\to\infty} \|R_\lambda - \chi_\lambda R_g\|_{L^p_{-q'-2}} = 0. \tag{11.3.5}$$

The unweighted estimates for $R_{\lambda_i} - \chi_{\lambda_i} R_g$ then follow from Lemma 11.3.1 with $s = \frac{n}{2}$ and $\frac{2n}{n+2}$:

$$\|R_{\lambda_i} - \chi_{\lambda_i} R_g\|_{L^{\frac{2n}{n+2}} \cap L^{\frac{n}{2}}} \leq C 2^{-\left(q' - \frac{n-2}{2}\right) i} \|R_{\lambda_i} - \chi_{\lambda_i} R_g\|_{L^p_{-q'-2}} \underset{i\to 0}{\to} 0.$$

A proof of (11.3.5) can be found in [Kuw90], for example, but we give the argument here for completeness. It is easiest to compute $R_\lambda - \chi_\lambda R_g$ and note that the worst decaying second derivatives of g cancel out. Indeed, we have

$$R_\lambda = \partial_j(\partial_i g_{\lambda\, ij} - \partial_j g_{\lambda\, ii}) + O((g_\lambda - \delta)\partial^2 g_\lambda) + O((\partial g_\lambda)^2),$$

with summation over i and j implied in the first term on the right. This term can be expanded as

$$\begin{aligned} \partial_j(\partial_i g_{\lambda\, ij} - \partial_j g_{\lambda\, ii}) &= \chi_\lambda \partial_j(\partial_i g_{ij} - \partial_j g_{ii}) + \partial\chi_\lambda \partial g + \partial^2 \chi_\lambda (g - \delta) \\ &= \chi_\lambda \partial_j(\partial_i g_{ij} - \partial_j g_{ii}) + O(|x|^{-q-2}). \end{aligned}$$

For first derivatives we again have $|\partial g_\lambda| \leq C|x|^{-q}$ and for second derivatives

$$\begin{aligned} \partial^2 g_\lambda = \partial^2\left(\chi_\lambda(g - \delta)\right) &= \partial^2\chi_\lambda(g-\delta) + \partial\chi_\lambda \partial(g - \delta) + \chi_\lambda \partial^2 g \\ &= \chi_\lambda \partial^2 g + O(|x|^{-q-2}). \end{aligned}$$

Here all instances of Landau notation occur with implied constants independent of λ. Putting everything together, we find that

$$|R_\lambda - \chi_\lambda R_g| \le C|x|^{-q-2} + C|x|^{-q}|\partial^2 g|$$

when $\lambda \le |x| \le 2\lambda$ and this difference vanishes everywhere else. For the second term, we have

$$\||x|^{-q}\partial^2 g\|_{L^p_{-q-2}(\{\lambda\le|x|\le 2\lambda\})} \le C\lambda^{-q} \to 0,$$

so the same is true with decay $q' < q$. Now q' becomes crucial for the first term, since we have

$$\||x|^{-q-2}\|^p_{L^p_{-q'-2}(\{\lambda\le|x|\le 2\lambda\})} \le C\int_\lambda^{2\lambda} r^{-p(q-q')-1}\,dr \le C\lambda^{-p(q-q')} \to 0.$$

This completes the proof of the claim and hence the lemma follows. □

Proof of Theorem 11.1.1. For $\lambda_i = 2^i$, let u_{λ_i} be the functions whose existence is guaranteed by applying Proposition 11.2.15 to the metric g_{λ_i} with $V = R_\lambda - \chi_\lambda R_g$, and let $\tilde{g}_i = u_{\lambda_i}^{\frac{4}{n-2}} g_{\lambda_i}$. By construction, these are harmonically flat and ε-close to g in the asymptotic topology and on K. We first check part (i) of Definition 11.2.7. This follows from smallness of $g_{\lambda_i} - g$ and $u_{\lambda_i} - 1$ in $W^{2,p}_{-q}$. As the second claim is a part of the package in Proposition 11.2.15, we only need to prove the first. First, we observe that

$$\|g_\lambda - g\|_{W^{2,p}_{-q}} \le C\|\chi_\lambda - 1\|_{C^2}\|g - \delta\|_{W^{2,p}_{-q}} \le C.$$

Then we note that $\operatorname{spt}(g_\lambda - g) \subset \{|x| \ge \lambda\}$, which implies

$$\|g_\lambda - g\|^p_{L^p_{-q'-2}} = \int_{|x|\ge\lambda} |x|^{-(q-q')}\||x|^q|g_\lambda - g|^p \frac{dx}{|x|^n} \le C\lambda^{-(q-q')} \to 0.$$

Similar considerations apply to the derivatives and hence part (i) is proved. By (11.2.13), the scalar curvature is given by

$$R_{\tilde{g}_i} = \chi_{\lambda_i} R_g u_{\lambda_i}^{\frac{4}{n-2}}.$$

From this we see that $R_{\tilde{g}_i}(x) \ge 0$ whenever $R_g(x) \ge 0$. We now prove part (ii). We have

$$\begin{aligned} R_{\tilde{g}_i} - R_g &= \chi_{\lambda_i} R_g u_{\lambda_i}^{\frac{4}{n-2}} - R_g \\ &= \chi_{\lambda_i}(u_{\lambda_i}^{\frac{4}{n-2}} - 1)R_g + (\chi_\lambda - 1)R_g. \end{aligned}$$

For the first term, we estimate

$$\int_{\mathcal{E}} \chi_{\lambda_i} |u_{\lambda_i}^{\frac{4}{n-2}} - 1||R_g| \leq \sup |u_{\lambda_i}^{\frac{4}{n-2}} - 1| \int_{\mathcal{E}} |R_g| \to 0.$$

For the second term, we have

$$\int_{\mathcal{E}} |\chi_\lambda - 1||R_g| \leq \int_{|x|\geq\lambda} |R_g| \to 0$$

by elementary measure theory. Part (iii) now follows from parts (i) and (ii) together with Lemma 11.2.6.

Since K is compact, the cutoff region $|x| \geq 2^i$ misses K for i sufficiently large and there exists constant C such that $\sup_K |R_g| \leq C$. It follows that

$$\sup_K |R_{\tilde{g}_i} - R_g| \leq C \sup_K |1 - u_{\lambda_i}^{\frac{4}{n-2}}| = o(1)$$

as $i \to \infty$, where $o(1)$ follows from the estimate (11.2.10) for $v_{\lambda,\sigma}$. □

11.4 Pushing the scalar curvature up and down

In this section we explicitly describe a well-known mechanism for increasing or decreasing mass by making appropriate conformal changes. The precise statements, which we prove are valid in the context of incomplete manifolds, will be required in our proof of rigidity in the positive mass theorem, and in Corollary 9.3.4.

Proposition 11.4.1 (Pushing down)**.** *Let (M^n, g) be an asymptotically flat manifold of Sobolev type (p, q). Suppose $R_g > 0$ somewhere on M. For any $\varepsilon > 0$, there exists a (p, q) Sobolev asymptotically flat metric $\tilde{g}$ on M which is ε-close to g in the asymptotic topology, with* $\mathrm{spt}(R_{\tilde{g}}^-) = \mathrm{spt}(R_g^-)$ *and*

$$m_{\mathrm{ADM}}(\mathcal{E}, \tilde{g}) < m_{\mathrm{ADM}}(\mathcal{E}, g).$$

Proof. Let B be a ball on which $R_g > 0$. Let η be a smooth cutoff function for B such that $0 < \eta < 1$ on B and $\eta = 0$ on $M \setminus B$. For $\delta > 0$ we now solve the equation

$$-a\Delta_g u_\delta + \delta\eta R_g u_\delta = 0$$

for $u_\delta - 1 \in W^{2,p}_{-q}(M)$ in the sense of Proposition 11.2.12 and Proposition 11.2.15. The relevant norms are $O(\delta)$, so for δ sufficiently small we obtain a unique solution of this equation with the desired asymptotic behavior. Our previous computation shows that

$$R(u_\delta^{\frac{4}{n-2}} g) = (1 - \delta\eta) R_g u_\delta^{-\frac{4}{n-2}}.$$

Since $1-\delta\eta > 0$, the sign of the scalar curvature remains pointwise unchanged. Finally, by (11.2.14), we have

$$m_{\mathrm{ADM}}(\mathcal{E}, \tilde{g}_\delta) - m_{\mathrm{ADM}}(\mathcal{E}, g) = -\frac{1}{2(n-1)\omega_{n-1}} \int_M \delta\eta R_g u_\delta \, d\mu_g < 0,$$

so the mass strictly decreases. □

Proposition 11.4.2 (Bumping up)**.** *Let (M^n, g) be an asymptotically flat manifold of Sobolev type (p, q) with nonnegative scalar curvature on $\mathcal{E}$. Let $f : \mathbb{R} \to [0,1]$ be an exponentially decreasing smooth function with $f(x) > 0$ for $x > 2r_0$ and f vanishing on $M \setminus \mathcal{E}$. For sufficiently small $\varepsilon > 0$, depending only on f and $\|g - \delta\|_{W^{2,p}_{-q}(\mathcal{E})}$, there exists a (p, q) Sobolev asymptotically flat metric $\tilde{g}$ which is ε-close to g in the asymptotic topology and satisfies* $\mathrm{spt}(R^-_{\tilde{g}}) \subset \mathrm{spt}(R^-_g)$ *and $R_{\tilde{g}}(x) \geq cf(|x|)$ for $|x| \geq 2r_0$ and a constant $c > 0$ depending only on ε and the other stated parameters.*

Proof. For $\delta > 0$ we solve the equation

$$-a\Delta_g u_\delta - \delta f u_g = 0$$

in the sense of Proposition 11.2.12 and Proposition 11.2.15. The norms are again $O(\delta)$, so we can solve the equation with the desired asymptotics, for δ small depending on f, the geometry, and ε. The scalar curvature of the conformal metric is

$$R(\tilde{g}_\delta) = (R_g + \delta f) u_\delta^{-\frac{4}{n-2}}.$$

For δ small depending only on the allowable parameters, $u_\delta \leq 2$. It follows that $R(\tilde{g}_\delta) \geq \delta 2^{-\frac{4}{n-2}} f$. Finally, we note that the mass strictly increases in this process. □

11.5 Analogy between μ-bubbles and MOTS

Recall from Section 10.2.1 that a hypersurface Σ in an initial data set (M^n, g, k) with distinguished choice of normal ν is a *marginally outer trapped surface (MOTS)* if

$$\theta^+ = H + P = 0,$$

where H is the mean curvature and $P = \operatorname{tr}_\Sigma k = (g^{ij} - \nu^i \nu^j)k_{ij}$. MOTS do not satisfy a variational criterion,[1] but there is a naturally associated stability operator [AMS05]

$$L = -\Delta\varphi + 2\langle W_\Sigma, \nabla u\rangle + (\operatorname{div}_\Sigma W_\Sigma - |W_\Sigma|^2 + Q_\Sigma)\varphi,$$

where $W_\Sigma = k(\nu, \cdot)$ restricted to $T\Sigma$,

$$Q_\Sigma = \tfrac{1}{2}R_\Sigma - \mu - \langle J, \nu\rangle - \tfrac{1}{2}|k_\Sigma + A|^2,$$

where k_Σ is k restricted to $T\Sigma$. A MOTS is *stable* if $\lambda_1(L) \geq 0$.[2] Closely related is the *symmetrized MOTS stability operator* of Galloway–Schoen [GS06]

$$L_{\text{sym}} = -\Delta_\Sigma + Q_\Sigma.$$

Notably, Galloway and Schoen showed that $\lambda_1(L) \leq \lambda_1(L_{\text{sym}})$ for any MOTS. We now explain a relationship between stable MOTS and stable μ-bubbles.

Definition 11.5.1. Given (M, g, h) as in Definition 9.3.7, we define a data set (M', g, k_h) by taking $M' \doteq \{|h| < \infty\}$ and setting

$$k_h \doteq -\frac{h}{n-1} g.$$

The first observation is that if Σ is any hypersurface in M', then $P = -h$, so that

$$\theta^+ = H - h,$$

and thus Ω is a μ-bubble with respect to (M, g, h) if and only if $\partial\Omega$ is a MOTS with with respect to (M', g, k_h). In fact, more is true.

[1] That is, there is no operator defined on the space of hypersurfaces or boundaries whose critical points are MOTS.

[2] The operator L is not self-adjoint, but it still has a real principal eigenvalue and an associated real positive eigenfunction. See [AMS05].

Proposition 11.5.2. *Let (M, g, h) and $\Sigma = \partial\Omega$ smooth be as in Definition 9.3.7. Then:*

1. *Σ is a stable μ-bubble if and only if Σ is a stable MOTS with respect to (g, k_h).*

2. *(g, h) satisfies condition $(\star)$,*

$$R_g + \tfrac{n}{n-1}h^2 - 2|\nabla h| \geq 0,$$

if and only if (g, k_h) satisfies the dominant energy condition.

Proof. The first thing to note is that $W_\Sigma = 0$ for our choice of k. Therefore, $L = L_{\text{sym}}$ in this setting. To compute Q_Σ, we note the following, which the reader may easily verify:

$$\begin{aligned}
\operatorname{tr} k &= -\tfrac{n}{n-1}h, \\
|k|^2 &= \tfrac{n}{(n-1)^2}h^2, \\
\nabla^j k_{ij} &= -\tfrac{1}{n-1}\nabla_i h, \\
\nabla_i \operatorname{tr} k &= -\tfrac{n}{n-1}\nabla_i h.
\end{aligned}$$

It follows that

$$\begin{aligned}
2\mu &= R_g + \tfrac{n}{n-1}h^2 \\
J^i &= \nabla^i h.
\end{aligned}$$

Putting these together yields

$$\begin{aligned}
2(\mu + \langle J, \nu\rangle) &= R_g + \tfrac{n}{n-1}h^2 + 2\langle \nabla h, \nu\rangle, \\
2(\mu - |J|) &= R_g + \tfrac{n}{n-1}h^2 - 2|\nabla h|,
\end{aligned}$$

The second equation verifies part (2). To complete the proof of part (1), if Ω is a μ-bubble, then $H = h$, and we can see that

$$\begin{aligned}
|k_\Sigma + A|^2 &= \tfrac{1}{n-1}h^2 - \tfrac{2}{n-1}hH + |A|^2 \\
&= -\tfrac{1}{n-1}h^2 + |A|^2.
\end{aligned}$$

Thus

$$Q_\Sigma = \tfrac{1}{2}R_\Sigma - \tfrac{1}{2}R_g - \tfrac{1}{2}|A|^2 - \tfrac{1}{2}h^2 - \langle \nabla h, \nu\rangle,$$

and we can explicitly see that the MOTS stability inequality $\lambda(L_{\text{sym}}) \geq 0$ is the same as the μ-bubble stability inequality (9.3.3). $\square$

Using this, we observe that Lemma 9.3.8 can be seen as a special case of the existence theorem for stable MOTS.

Proof of Lemma 9.3.8 using MOTS. Since $h \to \pm\infty$ as we approach $\partial_\pm M$, hypersurfaces foliating a small neighborhood of ∂M_+ and ∂M_- will be strictly trapped and untrapped, respectively. Hence we may apply the MOTS existence theory of L. Andersson, M. Eichmair, and J. Metzger [Eic09; AM09; AEM11] to find a nontrivial MOTS in M_0. By part (1) of Proposition 11.5.2, this gives us the desired nontrivial stable μ-bubble. $\square$

This is a more complicated proof of the lemma, but it illustrates the general principle of the analogy.

11.5.1 Proof of the quantitative shielding theorem

We now use this viewpoint to give a very short proof of the quantitative shielding theorem, Theorem 9.3.3, which we restate here for the convenience of the reader.

Theorem 11.5.3 (Quantitative shielding theorem). *Let (M^n, g), $3 \leq n \leq 7$, be an asymptotically flat manifold of Sobolev type (p, q), with $p > n$ and $q > \frac{n-2}{2}$, not assumed to be complete or to have nonnegative scalar curvature everywhere. Let U_0, U_1, and U_2 be neighborhoods of an asymptotically flat end $\mathcal{E}$ such that $\overline{U_2} \subset U_1$, $\overline{U_1} \subset U_0$, and $\overline{U_0 \setminus \mathcal{E}}$ is compact, and let*

$$D_0 = \operatorname{dist}_g(\partial U_0, U_1) \quad \textit{and} \quad D_1 = \operatorname{dist}_g(U_2, \partial U_1).$$

If the following hold:

1. *g has no points of incompleteness in U_0,*
2. *$R_g \geq 0$ on U_0, and*
3. *the scalar curvature satisfies the largeness assumption*

$$R_g > \frac{4}{D_0 D_1} \quad \textit{on} \quad \overline{U_1} \setminus U_2, \tag{11.5.1}$$

then the ADM mass of the asymptotically flat end $\mathcal{E}$ is strictly positive.

Proof. We construct a potential function h. Let $\varepsilon > 0$ be small enough so that $\varepsilon < \frac{1}{n-1}$ and

$$R_g > \frac{4(1+\varepsilon)}{(D_0 - 2\varepsilon)D_1} \tag{11.5.2}$$

on the compact set $\overline{U}_1 \setminus U_2$. Let $\rho = \mathrm{dist}_g(x, U_1)$ and let $\tilde{\rho}$ be a smoothing of ρ on U_0 such that ρ also vanishes on U_1 and the following inequalities hold:

- $\sup_{U_0} |\tilde{\rho} - \rho| < \varepsilon$,
- $|\nabla \tilde{\rho}| < 1 + \varepsilon$.

One can construct $\tilde{\rho}$ by mollifying the Lipschitz distance function from the $\varepsilon/2$-neighborhood of U_1. Also, let ϑ be a smooth cutoff function such that $\vartheta = 1$ on $M \setminus U_1$, $\vartheta = 0$ on U_2, and $|\nabla \vartheta| \leq (1+\varepsilon)D_1^{-1}$ on $\overline{U}_1 \setminus U_2$. To see that such a ϑ exists, first construct a Lipschitz function that is 1 and 0 on *neighborhoods* of $M \setminus U_1$ and $\overline{U}_2$, respectively, and then mollify.

Now select any α between $D_0 - 2\varepsilon$ and $D_0 - \varepsilon$ such that the level set $\tilde{\rho}^{-1}(\alpha)$ is smooth. Then we define

$$h_1(x) = \begin{cases} \frac{2}{\alpha - \tilde{\rho}} & \text{if } \tilde{\rho}(x) < \alpha \\ +\infty & \text{if } \tilde{\rho}(x) \geq \alpha \end{cases}$$

and

$$h = \vartheta h_1.$$

We claim that this choice of h satisfies condition $(\star)$. Note that $M_0 = \overline{\{h < \infty\}}$ is contained in U_0, so there are three regions to analyze: U_2, $\overline{U}_1 \setminus U_2$, and $U_0 \setminus U_1$. The potential h vanishes identically on U_2, so $(\star)$ is trivially satisfied since $R_g \geq 0$ on U_2. On the region $\overline{U}_1 \setminus U_2$, h_1 is a constant equal to $\frac{2}{\alpha}$, so we have

$$\tfrac{n}{n-1} h^2 - 2|\nabla h| \geq -\frac{4(1+\varepsilon)}{\alpha D_1}. \tag{11.5.3}$$

Combining this with (11.5.2) and the definition of α, we see that condition $(\star)$ holds in this region as well. Finally, on $U_0 \setminus U_1$, we have

$$2|\nabla h| = \frac{4}{(\alpha - \tilde{\rho})^2} |\nabla \tilde{\rho}| \leq \frac{4}{(\alpha - \tilde{\rho})^2}(1+\varepsilon) = (1+\varepsilon)h_1^2 \leq \tfrac{n}{n-1} h_1^2$$

Since $R_g \geq 0$ on U_0, we see that $(\star)$ holds on $U_0 \setminus U_1$, and hence we have shown that it holds everywhere.

By Proposition 11.5.2, we observe that (g, k_h) defines asymptotically flat initial data on M_0 that satisfies the dominant energy condition, and moreover, since ∂M_0 is smooth, the level sets of h near ∂M_0, where h is large, must be strictly outer trapped surfaces. The strict positivity of $m_{\mathrm{ADM}}(\mathcal{E}, g)$ now follows immediately from Theorem 9.2.2.

Finally, we note that Theorem 9.2.2 requires C^2_{-q} decay of the metric rather than the Sobolev decay as in Definition 11.2.4, but Theorem 11.1.1 can be used to assume we have this decay without loss of generality when proving the *nonnegativity* of mass. It is easy to see that strict positivity follows from nonnegativity using the pushing down technique Proposition 11.4.1. □

We can now also prove Corollary 9.3.4, which we restate here for the convenience of the reader.

Corollary 11.5.4. *Let (M^n, g), $3 \leq n \leq 7$, be a Riemannian manifold, not assumed to be complete, with an asymptotically flat end $\mathcal{E}$ of Sobolev type (p, q), where $p > n$ and $q > \frac{n-2}{2}$. If $m_{\mathrm{ADM}}(\mathcal{E}, g) < 0$, then there exists a constant D, depending only on $m_{\mathrm{ADM}}(\mathcal{E}, g)$ and $\|g - \delta\|_{W^{2,p}_{-q}(\mathcal{E})}$, with the following property. In the D-neighborhood $N_D(\mathcal{E})$ of $\mathcal{E}$, one or both of the following must be true:*

1. *$R_g < 0$ somewhere in $N_D(\mathcal{E})$, or*

2. *$N_D(\mathcal{E})$ contains an incomplete point.*

Proof. Apply Proposition 11.4.2 with an arbitrarily chosen f and with ε sufficiently small that $m_{\mathrm{ADM}}(\mathcal{E}, \tilde{g}) < 0$. Then we know that the hypotheses of Theorem 9.3.3 are violated in $N^{\tilde{g}}_D(\mathcal{E})$ for $\tilde{g}$, where D is such that (9.3.1) is satisfied on some annular region where $f > 0$. However, $\tilde{g}$ does not have any new points of negative scalar curvature and no new incomplete points, so one of the hypotheses must be violated for g in $N^{\tilde{g}}_D(\mathcal{E})$ as well. Since g and $\tilde{g}$ are uniformly equivalent, $N^{\tilde{g}}_D(\mathcal{E}) \subset N^g_{D'}(\mathcal{E})$ for some D' close to D. □

11.5.2 Proof of the positive mass theorem with arbitrary ends

We now prove the main theorem of this chapter, Theorem 9.3.1, which we restate here for the convenience of the reader.

Theorem 11.5.5 (The positive mass theorem with arbitrary ends)**.** *Let (M^n, g), $3 \leq n \leq 7$, be a complete manifold with nonnegative scalar curvature and at least one asymptotically flat end $\mathcal{E}$ of Sobolev type (p, q), where $p > n$ and $q > \frac{n-2}{2}$. Then the ADM mass of $\mathcal{E}$ is nonnegative. Furthermore, if the mass is zero, then (M, g) is isometric to Euclidean space.*

Proof. We first prove the inequality $m_{\mathrm{ADM}}(\mathcal{E}, g) \geq 0$. Suppose otherwise. By bumping the scalar curvature up with Proposition 11.4.2, we obtain a new asymptotically flat metric $\tilde{g}$ with $R_{\tilde{g}} > \eta$ for $|x| \in [3r_0, 4r_0]$ and negative mass. Therefore, taking $U_0 = \{|x| > 4r_0\}$, $U_1 = \{|x| > 3r_0\}$, and U_0 a sufficiently large open set containing U_1, we obtain a contradiction to Theorem 9.3.3.

We now prove rigidity, which roughly follows the standard conformal approach, except that we use our new results from Section 11.2 and Section 11.4. Assume $m_{\mathrm{ADM}}(\mathcal{E}, g) = 0$. First we claim that g is scalar-flat. Otherwise we can use Proposition 11.4.1 to obtain a new metric which still has nonnegative scalar curvature but has negative mass, contradicting the positive mass inequality that we already proved. Next, we show that g is Ricci-flat as well.

Let η be a compactly supported cutoff function and $t \in \mathbb{R}$. We consider the deformed metrics $g_t = g + t\eta \mathrm{Ric}_g$. For t sufficiently small, these will indeed be Riemannian metrics and will satsify the analytic hypotheses of Section Section 11.2 uniformly. We solve the equations

$$-a\Delta_{g_t} u_t + R_{g_t} u_t = 0$$

to obtain a scalar-flat metric $\tilde{g}_t := u_t^{\frac{4}{n-2}} g_t$. One can see that this is possible by Proposition 11.2.12 and Proposition 11.2.15, for $|t|$ sufficiently small. By the positive mass inequality for $\tilde{g}$, we know that $\frac{m_{\mathrm{ADM}}(\mathcal{E}, \tilde{g}_t)}{t} \geq 0$ for $t > 0$ and $\frac{m_{\mathrm{ADM}}(\mathcal{E}, \tilde{g}_t)}{t} \leq 0$ for $t < 0$. Then by (11.2.14), we can see that

$$-\frac{1}{2(n-1)\omega_{n-1}} \lim_{t \to 0} \frac{1}{t} \int_M R_{g_t} u_t \, d\mu_{g_t} = \lim_{t \to 0} \frac{m_{\mathrm{ADM}}(\mathcal{E}, \tilde{g}_t)}{t} = 0.$$

Using the dominated convergence theorem and the calculation as in [Lee19, page 96], we can also see that

$$\lim_{t \to 0} \int_M \frac{R_{g_t}}{t} u_t \, d\mu_{g_t} = \int_M \eta |\mathrm{Ric}_g|^2,$$

Combining the two equalities above, we see that $\mathrm{Ric}_g = 0$ on the support of η. Since η was arbitrary, this implies g is Ricci-flat.

Now we show that (M, g) has only one end. If M has a second end, then it contains a geodesic line γ which goes out to infinity along the asymptotically flat end [Pet16]. By the Cheeger–Gromoll theorem, (M, g) splits isometrically along this line as $(\mathbb{R} \times N, dt^2 + h)$. Concretely, there is a smooth function $f : M \to \mathbb{R}$ whose level sets foliate M and are all isometric to (N, h), the isometry being the gradient flow of f. If N is not flat, then there exists a ball $B \subset M$ such that

$$\int_B |\mathrm{Rm}_g|^p \, d\mu_g > 0. \tag{11.5.4}$$

We flow B along the gradient flow of f in the direction of the asymptotically flat end. Since the gradient flow is an isometry, the distance between the ball and γ is unchanged, as is the integral (11.5.4). But $\mathrm{Rm}_g \in L^p(\mathcal{E})$, so the integral (11.5.4) must limit to zero along a sequence as the ball goes further and further out, which is a contradiction.

Since N is flat, it is either isometric to $\mathbb{R}^{n-1}$ or a nontrivial quotient, in which case $\pi_1(N) \neq 0$. We must rule out the latter case. Let ℓ be a homotopically nontrivial loop in N. Then γ is contained in some compact ball and we can push it along the line γ. It enters $\mathcal{E} \approx \mathbb{R} \times S^{n-1}$ and can be contracted, a contradiction.

We conclude that (M, g) cannot have two or more ends. We now show that (M^n, g) is flat. Take a point $p \in M$ and consider the volume ratio

$$V(r) = \frac{\mathrm{vol}_g(B_r(p))}{\alpha_n r^n}.$$

We have

$$\lim_{r\to 0} V(r) = \lim_{r\to\infty} V(r) = 1.$$

The limit as $r \to 0$ is true for any manifold, the limit as $r \to \infty$ is a consequence of one-endedness and asymptotic flatness. A careful proof can be found in [Li18, Lemma 2.6]. Now the Bishop–Gromov volume comparison theorem implies $V(r) = 1$ for any r and hence (M, g) is flat. □

11.5.3 A positive mass theorem with a non-mean convex boundary

The second proof has the following immediate application to manifolds with boundary. We assume that M now has a *non-mean convex*[3] boundary with respect to the normal pointing out of the manifold.[4]

Theorem 11.5.6. *Let (M^n, g), $3 \leq n \leq 7$, be a complete asymptotically flat manifold with nonempty compact boundary ∂M. Set $U_0 = M$, and let U_1, U_2, D_0, D_1 be as in Theorem Theorem 9.3.3, with the exception of item* (3). *If we instead assume that*

1. *$R_g > \kappa$ on $\overline{U}_1 \setminus U_2$ for a positive constant κ,*

2. *the parameters satisfy $\kappa < \frac{4}{D_0 D_1}$ (see Remark Remark 11.5.7 below), and*

[3]Traditionally, the boundary mean curvature in the Riemannian PMT is measured with respect to the outward normal. However, in view of conventions used in the spacetime positive mass theorem with boundary (see [GL21; LLU22]), we opt to state Corollary Theorem 11.5.6 with respect to the inner pointing normal.

[4]The case of mean convex boundaries is already implicitly contained in the original work of Schoen–Yau [SY79a; SY81b].

3. the mean curvature of ∂M with respect to the normal pointing into M satisfies

$$H < \frac{2\kappa D_1}{4 - \kappa D_0 D_1}, \tag{11.5.5}$$

then the ADM mass is strictly positive.

Remark 11.5.7. Fixing D_0 and D_1 in (11.5.5) and letting $\kappa \to 0$ recovers the classical condition $H \leq 0$. If $\kappa \geq \frac{4}{D_0 D_1}$, then Theorem Theorem 9.3.3 applies without any consideration of the boundary mean curvature.

Proof. The goal of this proof is to choose α so that ∂M is outer trapped with respect to the data set (g, k_h), while maintaining the DEC. The result then follows from the spacetime positive mass theorem with boundary, as above.

According to (11.5.3), we can maintain DEC, which is equivalent to $(\star)$, by setting

$$\alpha = \frac{4(1+\varepsilon)}{\kappa D_1}.$$

Meanwhile, ∂M being outer trapped means that $H < \min_{\partial M} h$, so we compute

$$\min_{\partial M} h \geq \frac{2}{\alpha - D_0 + \varepsilon} = \frac{2}{\frac{4(1+\varepsilon)}{\kappa D_1} - D_0 + \varepsilon}.$$

Taking $\varepsilon \to 0$ yields the desired bound appearing in Theorem Theorem 11.5.6.

From here we can now invoke again Theorem 9.2.2. Although that result assumes C^2_{-q} decay, we can again reduce to this case using Theorem 11.1.1. □

11.6 Proof of the Liouville theorem for locally conformally flat manifolds

In this section, we give a proof of Theorem 9.4.1 using Theorem 9.3.1. This is not excactly how the original proof in [LUY20] proceeded, but rather follows the original intention of Schoen and Yau in [SY88].

11.6.1 Existence of the conformal Green's Function

Recall the conformal laplacian

$$Lu \doteq -a\Delta u + R_g u,$$

where $a = 4(n-1)/(n-2)$. In this section we prove the existence of a unique minimal Green's function for L on certain open manifolds.

Proposition 11.6.1. *Let (M, g) be a complete Riemannian manifold of dimension $n \geq 3$ which admits a developing map $\Phi : M \to S^n$. For any $p \in M$, there is unique minimal and positive Green's function for the conformal Laplacian with pole p. It is C^∞ away from p and satisfies*

$$LG = a\delta_p$$

in the sense of distributions.

Here *minimal* means that if G' is any other positive Green's function for the conformal Laplacian with pole p, $G \leq G'$ on $M \setminus \{p\}$. The proof of Proposition 11.6.1 is outlined in [SY88, Corollary 1.3] and [SY94, Proposition VI.2.4].

Sketch of Proof. Let $\Omega_i \nearrow M$ be a compact exhaustion with $p \in \Omega_i$ a common point. On each domain Ω_i there is a unique Dirichlet Green's function G_i for the conformal Laplacian with pole p. Near p, we have

$$G_i(x) = \frac{c_n}{r^{n-2}}(1 + o(1)),$$

where $c_n^{-1} = n(n-1)|B_1^n|$. This is classical, but a careful proof can be found in [DHR04, Appendix A]. Since each G_i has the same growth rate at p, the maximum principle shows $G_i \leq G_{i+1}$ pointwise.[5] We wish to define G to be the limit of this monotone sequence, but we first need a barrier to prove that G is finite away from p.

Since $\Phi : M \to S^n$ is conformal, $\Phi^* g_0 = |d\Phi|^2 g$, where g_0 is the round metric on S^n. Let $y = \Phi(p)$ and denote by H the conformal Green's function of (S^n, g_0) with pole y. One can check that

$$L\overline{H} = \sum_{q \in \Phi^{-1}(y)} |d\Phi(q)|^{-\frac{n-2}{2}} a\delta_q$$

[5] This argument is a bit subtle. Since we are not assuming any sign condition on R_g, the operator L does not actually satisfy the maximum principle. It is a nontrivial theorem that L is a positive operator on Ω_i for i sufficiently large [SY94, Theorem VI.2.2]. Therefore, one can make a conformal change to positive scalar curvature and then apply the maximium principle, making crucial use of the conformal invariance of the conformal Green's function.

in the sense of distributions, where $\overline{H} := |d\Phi|^{\frac{n-2}{2}} H \circ \Phi$. Rescaling by an overall constant, we obtain a function $\overline{G}$ with poles located at every point in $\Phi^{-1}(y)$ and such that

$$L\overline{G} = \sum_{q \in \Phi^{-1}(y)} b_q a \delta_q, \quad b_q > 0, \quad b_p = 1.$$

Using $\overline{G}$ as a barrier, it can be shown that $G < \infty$ away from $\Phi^{-1}(y)$. That G is bounded near each point in $\Phi^{-1}(y) \setminus \{p\}$ follows from the Harnack inequality applied to the sequence $\{G_i\}$. Since the sequence $\{G_i\}$ converges uniformly away from p, it is not hard to now show that $LG = a\delta_p$. □

Remark 11.6.2. It follows from this construction that

$$G = \frac{c_n}{r^{n-2}}(1 + o(1)).$$

Indeed, G is dominated by $\overline{G}$, which has this growth.

11.6.2 Blowing up and finishing the proof

We will use the Green's function G as a conformal factor to obtain an asymptotically flat manifold. In the following lemma we establish a precise decay rate.

Lemma 11.6.3. *The function $v = G/\overline{G}$ is a positive harmonic function with respect to the metric $\overline{g} = \overline{G}^{\frac{4}{n-2}} g$, smooth across p, and for any normal coordinate system $\{x^i\}$ centered at p,*

$$v = 1 + cr^{n-2} + o(r^{n-2})$$

for some constant c, where $r = |x|$.

Proof. We outline the steps already contained in [SY88; SY94]. Let $\pi : S^n \setminus \{y\} \to \mathbb{R}^n$ be the stereographic projection. Since the Green's function H on (S^n, g_0) is actually just the conformal factor associated to stereographic projection, $\pi^*\delta = H^{\frac{4}{n-2}} g_0$, where δ is the flat metric on $\mathbb{R}^n$. It follows that $\Phi^*\pi^*\delta = \overline{g}$.

It is easy to see that v is harmonic with respect to $\overline{g}$ and also that

$$v(p) = \lim_{x \to p} \frac{G(x)}{\overline{G}(x)} = 1,$$

see Remark 11.6.2. Setting $h = v - 1$, $L(hG) = 0$ since $hG = \overline{G} - G$. It follows from the removable singularities theorem for elliptic equations that hG extends smoothly over p. For convenience we set $f = hG$. Then $h = fG^{-1}$. Since f is C^∞ and we have the expansion $G = c_n r^{2-n} + o(r^{2-n})$, it follows that $h = cr^{n-2} + o(r^{n-2})$. $\square$

Proof of Theorem 9.4.1. As shown in [SY88; SY94], Φ is injective and $\partial\Phi(M)$ has zero Newtonian capacity if and only if $v \equiv 1$. The conformal blowup

$$\tilde{g}_\varepsilon \doteq (G + \varepsilon)^{\frac{4}{n-2}} g,$$

is a complete metric on $\tilde{M} = M \setminus \{p\}$. The formula for scalar curvature under a conformal deformation implies

$$R(\tilde{g}_\varepsilon) = (G + \varepsilon)^{-\frac{n+2}{n-2}} L_g(G + \varepsilon) = (G + \varepsilon)^{-\frac{n+2}{n-2}} \varepsilon R(g) \geq 0.$$

In a neighborhood of p,

$$\tilde{g}_\varepsilon = \left(v + \frac{\varepsilon}{\overline{G}}\right)^{\frac{4}{n-2}} \overline{g} = \Phi^* \pi^* \left(v_\varepsilon^{\frac{4}{n-2}} \delta\right),$$

where $v_\varepsilon : \mathbb{R}^n \setminus B \to (0, \infty)$ satisfies

$$v_\varepsilon \circ \pi \circ \Phi = v + \frac{\varepsilon}{\overline{G}}.$$

By the expansion of v shown in Lemma 11.6.3 and the expansion $\overline{G} = c_n r^{2-n} + o(r^{2-n})$ it follows (after performing a coordinate inversion) that

$$v_\varepsilon = 1 + c_\varepsilon r^{2-n} + o(r^{2-n})$$

for a constant c_ε which is proportional to the ADM mass. By Theorem 9.3.1, $c_\varepsilon \geq 0$. Letting $\varepsilon \to 0$ shows that the ADM mass of $\tilde{g}_0$ is nonnegative. It then follows immediately from a simple maximum principle argument (see [SY88, Proposition 4.2]) that $v \equiv 1$, which completes the proof. $\square$

Bibliography

[ADN64] S. Agmon, A. Douglis, and L. Nirenberg. "Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. II". *Comm. Pure Appl. Math.* 17 (1964), pp. 35–92.

[AAR21] E. Ames, H. Andréasson, and O. Rinne. "Dynamics of gravitational collapse in the axisymmetric Einstein-Vlasov system". *Classical Quantum Gravity* 38.10 (2021), Paper No. 105003, 34.

[AH23] X. An and T. He. "Dynamics of apparent horizon and a null comparison principle" (2023). arXiv: 2312.15666.

[AL22a] X. An and Z. F. Lim. "Trapped surface formation for spherically symmetric Einstein-Maxwell-charged scalar field system with double null foliation". *Ann. Henri Poincaré* 23.9 (2022), pp. 3159–3190.

[An22] Z. An. "On mass-minimizing extensions of Bartnik boundary data". To appear in *Comm. Anal. Geom.* (2022).

[AC96] L. Andersson and P. T. Chruściel. "Solutions of the constraint equations in general relativity satisfying "hyperboloidal boundary conditions"". *Dissertationes Math. (Rozprawy Mat.)* 355 (1996), p. 100.

[AEM11] L. Andersson, M. Eichmair, and J. Metzger. "Jang's equation and its applications to marginally trapped surfaces". *Complex analysis and dynamical systems IV. Part 2.* Vol. 554. Contemp. Math. Amer. Math. Soc., Providence, RI, 2011, pp. 13–45.

[AMS05] L. Andersson, M. Mars, and W. Simon. "Local Existence of Dynamical and Trapping Horizons". *Phys. Rev. Lett.* 95 (11 2005), p. 111102.

[AM09] L. Andersson and J. Metzger. "The Area of Horizons and the Trapped Region". *Communications in Mathematical Physics* 290 (2009), pp. 941–972.

[And14] H. Andréasson. “Black hole formation from a complete regular past for Vlasov matter”. *XVIIth ICMP*. World Sci. Publ., Hackensack, NJ, 2014, pp. 365–372.

[And09] H. Andréasson. “Sharp bounds on the critical stability radius for relativistic charged spheres”. *Communications in Mathematical Physics* 288.2 (2009), pp. 715–730.

[And21] H. Andréasson. “Existence of steady states of the massless Einstein-Vlasov system surrounding a Schwarzschild black hole”. *Ann. Henri Poincaré* 22.12 (2021), pp. 4271–4297.

[AER09] H. Andréasson, M. Eklund, and G. Rein. “A numerical investigation of the steady states of the spherically symmetric Einstein–Vlasov–Maxwell system”. *Classical and Quantum Gravity* 26.14 (2009), p. 145003.

[AR06] H. Andréasson and G. Rein. “A numerical investigation of the stability of steady states and critical phenomena for the spherically symmetric Einstein-Vlasov system”. *Classical Quantum Gravity* 23.11 (2006), pp. 3659–3677.

[Ang16] Y. Angelopoulos. “Global spherically symmetric solutions of non-linear wave equations with null condition on extremal Reissner-Nordström spacetimes”. *Int. Math. Res. Not. IMRN* 11 (2016), pp. 3279–3355.

[AAG20] Y. Angelopoulos, S. Aretakis, and D. Gajic. “Nonlinear scalar perturbations of extremal Reissner-Nordström spacetimes”. *Ann. PDE* 6.2 (2020), Paper No. 12, 124.

[Ape22] M. A. Apetroaie. “Instability of gravitational and electromagnetic perturbations of extremal Reissner-Nordström spacetime” (2022). arXiv: 2211.09182.

[Are11a] S. Aretakis. “Stability and instability of extreme Reissner-Nordström black hole spacetimes for linear scalar perturbations I”. *Comm. Math. Phys.* 307.1 (2011), pp. 17–63.

[Are11b] S. Aretakis. “Stability and instability of extreme Reissner-Nordström black hole spacetimes for linear scalar perturbations II”. *Ann. Henri Poincaré* 12.8 (2011), pp. 1491–1538.

[Are12] S. Aretakis. “Decay of axisymmetric solutions of the wave equation on extreme Kerr backgrounds”. *J. Funct. Anal.* 263.9 (2012), pp. 2770–2831.

[Are15] S. Aretakis. “Horizon instability of extremal black holes”. *Adv. Theor. Math. Phys.* 19.3 (2015), pp. 507–530.

[Are17] S. Aretakis. “The characteristic gluing problem and conservation laws for the wave equation on null hypersurfaces”. *Ann. PDE* 3.1 (2017), Paper No. 3, 56.

[ACR21] S. Aretakis, S. Czimek, and I. Rodnianski. "The characteristic gluing problem for the Einstein equations and applications" (2021). arXiv: 2107.02441.

[ACR23a] S. Aretakis, S. Czimek, and I. Rodnianski. "Characteristic gluing to the Kerr family and application to spacelike gluing". *Communications in Mathematical Physics* 403.1 (2023), pp. 275–327.

[ACR23b] S. Aretakis, S. Czimek, and I. Rodnianski. "The characteristic gluing problem for the Einstein vacuum equations: linear and nonlinear analysis". *Annales Henri Poincaré*. Springer. 2023, pp. 1–125.

[AL22b] N. Athanasiou and M. Lesourd. "Construction of Cauchy data for the dynamical formation of apparent horizons and the Penrose Inequality". *Advances in Theoretical and Mathematical Physics* 26.8 (2022), pp. 2379–2496.

[BC73] D. Bancel and Y. Choquet-Bruhat. "Existence, uniqueness, and local stability for the Einstein-Maxwell-Boltzman system". *Comm. Math. Phys.* 33 (1973), pp. 83–96.

[Bar70] J. M. Bardeen. "Kerr metric black holes". *Nature* 226.5240 (1970), pp. 64–65.

[BW71] J. M. Bardeen and R. V. Wagoner. "Relativistic disks. I. Uniform rotation". *Astrophysical Journal, vol. 167, p. 359* 167 (1971), p. 359.

[BCH73] J. M. Bardeen, B. Carter, and S. W. Hawking. "The four laws of black hole mechanics". *Comm. Math. Phys.* 31 (1973), pp. 161–170.

[Bar86] R. Bartnik. "The mass of an asymptotically flat manifold". *Comm. Pure Appl. Math.* 39.5 (1986), pp. 661–693.

[BC03] R. Bartnik and P. T. Chruściel. "Boundary value problems for Dirac–type equations, with applications". arXiv:math/0307278 (2003).

[Bau+23] T. W. Baumgarte, B. Brügmann, D. Cors, C. Gundlach, D. Hilditch, A. Khirnov, T. Ledvinka, S. Renkhoff, and I. S. Fernández. "Critical phenomena in the collapse of gravitational waves". *Physical Review Letters* 131.18 (2023), p. 181401.

[BC96] R. Beig and P. T. Chruściel. "Killing vectors in asymptotically flat space-times. I. Asymptotically translational Killing vectors and the rigid positive energy theorem". *J. Math. Phys.* 37.4 (1996), pp. 1939–1961.

[Bek72] J. Bekenstein. "Black holes and the second law". *Lettere Al Nuovo Cimento Series 2* 4.15 (1972), pp. 737–740.

[Big22] L. Bigorgne. "Asymptotic properties of small data solutions of the Vlasov-Maxwell system in high dimensions". *Mém. Soc. Math. Fr. (N.S.)* 172 (2022), pp. vi+123.

[Big23] L. Bigorgne. "Decay estimates for the massless Vlasov equation on Schwarzschild spacetimes". *Ann. Henri Poincaré* 24.11 (2023), pp. 3763–3836.

[BV70] W. B. Bonnor and P. C. Vaidya. "Spherically symmetric radiation of charge in Einstein-Maxwell theory". *Gen. Relativity Gravitation* 1.2 (1970), pp. 127–130.

[Boo16] I. Booth. "Evolutions from extremality". *Phys. Rev. D* 93.8 (2016), pp. 084005, 15.

[BF08] I. Booth and S. Fairhurst. "Extremality conditions for isolated and dynamical horizons". *Phys. Rev. D* 77.8 (2008), p. 084005.

[Bor33] K. Borsuk. "Drei Sätze über die n-dimensionale euklidische Sphäre". ger. *Fundamenta Mathematicae* 20.1 (1933), pp. 177–190.

[Bou73] D. G. Boulware. "Naked singularities, thin shells, and the Reissner-Nordström metric". *Phys. Rev. D* 8.8 (1973), p. 2363.

[BCG97] P. R. Brady, C. M. Chambers, and S. M. Goncalves. "Phases of massive scalar field collapse". *Physical Review D* 56.10 (1997), R6057.

[Bre11] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. Springer, New York, 2011, pp. xiv+599.

[Buc59] H. A. Buchdahl. "General relativistic fluid spheres". *Physical Review* 116.4 (1959), p. 1027.

[CS16] A. Carlotto and R. Schoen. "Localizing solutions of the Einstein constraint equations". *Invent. Math.* 205.3 (2016), pp. 559–615.

[CZ20] S. Cecchini and R. Zeidler. "Scalar and mean curvature comparison via the Dirac operator". To appear in *Geometry & Topology* (2020).

[CZ21] S. Cecchini and R. Zeidler. "The positive mass theorem and distance estimates in the spin setting". To appear in *Trans. Amer. Math. Soc.* (2021).

[CLSZ21] J. Chen, P. Liu, Y. Shi, and J. Zhu. "Incompressible hypersurface, positive scalar curvature and positive mass theorem" (2021). arXiv: 2112.14442.

[CL24] O. Chodosh and C. Li. "Generalized soap bubbles and the topology of manifolds with positive scalar curvature". *Ann. of Math. (2)* 199.2 (2024), pp. 707–740.

[Cho93] M. W. Choptuik. “Universality and scaling in gravitational collapse of a massless scalar field”. *Phys. Rev. Lett.* 70.1 (1993), p. 9.

[Cho09] Y. Choquet-Bruhat. *General relativity and the Einstein equations.* Oxford Mathematical Monographs. Oxford University Press, Oxford, 2009, pp. xxvi+785.

[CG69] Y. Choquet-Bruhat and R. Geroch. “Global aspects of the Cauchy problem in general relativity”. *Comm. Math. Phys.* 14 (1969), pp. 329–335.

[Chr70] D. Christodoulou. “Reversible and irreversible transformations in black-hole physics”. *Phys. Rev. Lett.* 25.22 (1970), p. 1596.

[Chr71] D. Christodoulou. “Investigations in gravitational collapse and the physics of black holes”. PhD thesis. Princeton University, 1971.

[Chr86] D. Christodoulou. “The problem of a self-gravitating scalar field”. *Comm. Math. Phys.* 105.3 (1986), pp. 337–361.

[Chr87] D. Christodoulou. “A mathematical theory of gravitational collapse”. *Communications in mathematical physics* 109.4 (1987), pp. 613–647.

[Chr91a] D. Christodoulou. *Notes on the geometry of null hypersurfaces.* U.S. Copyright Office, Library of Congress, Registration Number TXu 832-728, 1991.

[Chr91b] D. Christodoulou. “The formation of black holes and singularities in spherically symmetric gravitational collapse”. *Comm. Pure Appl. Math.* 44.3 (1991), pp. 339–373.

[Chr91c] D. Christodoulou. “The formation of black holes and singularities in spherically symmetric gravitational collapse”. *Comm. Pure Appl. Math.* 44.3 (1991), pp. 339–373.

[Chr93] D. Christodoulou. “Bounded variation solutions of the spherically symmetric Einstein-scalar field equations”. *Comm. Pure Appl. Math.* 46.8 (1993), pp. 1131–1220.

[Chr94] D. Christodoulou. “Examples of naked singularity formation in the gravitational collapse of a scalar field”. *Ann. of Math. (2)* 140.3 (1994), pp. 607–653.

[Chr99a] D. Christodoulou. “On the global initial value problem and the issue of singularities”. *Classical and Quantum Gravity* 16.12A (1999), A23.

[Chr99b] D. Christodoulou. “The instability of naked singularities in the gravitational collapse of a scalar field”. *Ann. of Math. (2)* 149.1 (1999), pp. 183–217.

[Chr02] D. Christodoulou. "The global initial value problem in general relativity". *The Ninth Marcel Grossmann Meeting: On Recent Developments in Theoretical and Experimental General Relativity, Gravitation and Relativistic Field Theories (In 3 Volumes)*. World Scientific. 2002.

[Chr09] D. Christodoulou. *The formation of black holes in general relativity*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2009, pp. x+589.

[CCG24] P. T. Chruściel, W. Cong, and F. Gray. "Characteristic Gluing with Λ: II. Linearised equations in higher dimensions" (2024). arXiv: 2401.04442.

[Chr13] P. T. Chruściel. "On maximal globally hyperbolic vacuum space-times". *J. Fixed Point Theory Appl.* 14.1 (2013), pp. 325–353.

[CC22] P. T. Chruściel and W. Cong. "Characteristic Gluing with Λ 1. Linearised Einstein equations on four-dimensional spacetimes" (2022). arXiv: 2212.10052.

[CC23] P. T. Chruściel and W. Cong. "Gluing variations" (2023). arXiv: 2302.06928.

[CD03] P. T. Chruściel and E. Delay. "On mapping properties of the general relativistic constraints operator in weighted function spaces, with applications". *Mém. Soc. Math. Fr. (N.S.)* 94 (2003), pp. vi+103.

[CM06] P. T. Chruściel and D. Maerten. "Killing vectors in asymptotically flat space-times. II. Asymptotically translational Killing vectors and the rigid positive energy theorem in higher dimensions". *J. Math. Phys.* 47.2 (2006), pp. 022502, 10.

[CRT06] P. T. Chruściel, H. S. Reall, and P. Tod. "On Israel-Wilson-Perjés black holes". *Classical Quantum Gravity* 23.7 (2006), pp. 2519–2540.

[CIP21] F. Corelli, T. Ikeda, and P. Pani. "Challenging cosmic censorship in Einstein-Maxwell-scalar theory with numerically simulated gedanken experiments". *Phys. Rev. D* 104.8 (2021), p. 084069.

[Cor00] J. Corvino. "Scalar curvature deformation and a gluing construction for the Einstein constraint equations". *Comm. Math. Phys.* 214.1 (2000), pp. 137–189.

[CH20] J. Corvino and L.-H. Huang. "Localized deformation for initial data sets with the dominant energy condition". *Calc. Var. Partial Differential Equations* 59.1 (2020), Paper No. 42, 43.

[CP11] J. Corvino and D. Pollack. "Scalar curvature and the Einstein constraint equations". *Surveys in geometric analysis and relativity.* Vol. 20. Adv. Lect. Math. (ALM). Int. Press, Somerville, MA, 2011, pp. 145–188.

[CS06] J. Corvino and R. Schoen. "On the asymptotics for the vacuum Einstein constraint equations". *J. Differential Geom.* 73.2 (2006), pp. 185–217.

[CR22] S. Czimek and I. Rodnianski. "Obstruction-free gluing for the Einstein equations" (2022). arXiv: 2210.09663.

[Daf03] M. Dafermos. "Stability and instability of the Cauchy horizon for the spherically symmetric Einstein-Maxwell-scalar field equations". *Ann. of Math. (2)* 158.3 (2003), pp. 875–928.

[Daf05a] M. Dafermos. "On naked singularities and the collapse of self-gravitating Higgs fields". *Advances in Theoretical and Mathematical Physics* 9.4 (2005), pp. 575–591.

[Daf05b] M. Dafermos. "Spherically symmetric spacetimes with a trapped surface". *Classical and Quantum Gravity* 22.11 (2005), p. 2221.

[Daf05c] M. Dafermos. "The interior of charged black holes and the problem of uniqueness in general relativity". *Comm. Pure Appl. Math.* 58.4 (2005), pp. 445–504.

[Daf06] M. Dafermos. "A note on the collapse of small data self-gravitating massless collisionless matter". *J. Hyperbolic Differ. Equ.* 3.4 (2006), pp. 589–598.

[Daf09] M. Dafermos. "Black hole formation from a complete regular past". *Comm. Math. Phys.* 289.2 (2009), pp. 579–596.

[DHR] M. Dafermos, G. Holzegel, and I. Rodnianski. "The linear stability of the Schwarzschild solution to gravitational perturbations". *Acta Math.* 222.1 (2019), pp. 1–214.

[DHRT] M. Dafermos, G. Holzegel, I. Rodnianski, and M. Taylor. "The non-linear stability of the Schwarzschild family of black holes" (2021). arXiv: 2104.08222.

[DL17] M. Dafermos and J. Luk. "The interior of dynamical vacuum black holes I: The C^0-stability of the Kerr Cauchy horizon" (2017). arXiv: 1710.01722.

[DR05a] M. Dafermos and A. D. Rendall. "An extension principle for the Einstein-Vlasov system in spherical symmetry". *Ann. Henri Poincaré* 6.6 (2005), pp. 1137–1155.

[DR16] M. Dafermos and A. D. Rendall. "Strong cosmic censorship for surface-symmetric cosmological spacetimes with collisionless matter". *Comm. Pure Appl. Math.* 69.5 (2016), pp. 815–908.

[DR05b] M. Dafermos and I. Rodnianski. "A proof of Price's law for the collapse of a self-gravitating scalar field". *Invent. Math.* 162.2 (2005), pp. 381–457.

[DS18] M. Dafermos and Y. Shlapentokh-Rothman. "Rough initial data and the strength of the blue-shift instability on cosmological black holes with $\Lambda > 0$". *Classical and Quantum Gravity* 35.19 (2018), p. 195010.

[DI67] V. De la Cruz and W. Israel. "Gravitational bounce". *Il Nuovo Cimento A (1965-1970)* 51.3 (1967), pp. 744–760.

[DK81] D. M. DeTurck and J. L. Kazdan. "Some regularity theorems in Riemannian geometry". *Ann. Sci. École Norm. Sup. (4)* 14.3 (1981), pp. 249–260.

[DPV12] E. Di Nezza, G. Palatucci, and E. Valdinoci. "Hitchhiker's guide to the fractional Sobolev spaces". *Bull. Sci. Math.* 136.5 (2012), pp. 521–573.

[DHR04] O. Druet, E. Hebey, and F. Robert. *Blow-up theory for elliptic PDEs in Riemannian geometry.* Vol. 45. Mathematical Notes. Princeton University Press, Princeton, NJ, 2004, pp. viii+218.

[DS93] F. Duzaar and K. Steffen. "λ minimizing currents". *Manuscripta Math.* 80.4 (1993), pp. 403–447.

[Eic09] M. Eichmair. "The Plateau problem for marginally outer trapped surfaces". *J. Differential Geom.* 83.3 (2009), pp. 551–583.

[Eic13] M. Eichmair. "The Jang Equation Reduction of the Spacetime Positive Energy Theorem in Dimensions Less Than Eight". *Comm. Math. Phys.* 319 (2013), pp. 575–593.

[EHLS16] M. Eichmair, L.-H. Huang, D. A. Lee, and R. Schoen. "The spacetime positive mass theorem in dimensions less than eight". *J. Eur. Math. Soc. (JEMS)* 18.1 (2016), pp. 83–121.

[Ein15] A. Einstein. "Die feldgleichungen der gravitation". *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften* (1915), pp. 844–847.

[FGS17] C. Fairoos, A. Ghosh, and S. Sarkar. "Massless charged particles: Cosmic censorship, and the third law of black hole mechanics". *Physical Review D* 96.8 (2017), p. 084013.

[FH79] C. J. Farrugia and P. Hajicek. "The third law of black hole mechanics: a counterexample". *Comm. Math. Phys.* 68.3 (1979), pp. 291–299.

[FM73] A. E. Fischer and J. E. Marsden. "Linearization stability of the Einstein equations". *Bull. Amer. Math. Soc.* 79 (1973), pp. 997–1003.

[Fou52] Y. Fourès-Bruhat. "Théorème d'existence pour certains systèmes d'équations aux dérivées partielles non linéaires". *Acta Math.* 88 (1952), pp. 141–225.

[Gaj17a] D. Gajic. "Linear waves in the interior of extremal black holes I". *Comm. Math. Phys.* 353.2 (2017), pp. 717–770.

[Gaj17b] D. Gajic. "Linear waves in the interior of extremal black holes II". *Ann. Henri Poincaré* 18.12 (2017), pp. 4005–4081.

[Gaj23] D. Gajic. "Azimuthal instabilities on extremal Kerr" (2023). arXiv: 2302.06636.

[GL19] D. Gajic and J. Luk. "The interior of dynamical extremal black holes in spherical symmetry". *Pure Appl. Anal.* 1.2 (2019), pp. 263–326.

[Gal00] G. J. Galloway. "Maximum principles for null hypersurfaces and null splitting theorems". *Ann. Henri Poincaré* 1.3 (2000), pp. 543–567.

[GL21] G. J. Galloway and D. A. Lee. "A note on the positive mass theorem with boundary". *Lett. Math. Phys.* 111.4 (2021), Paper No. 111, 10.

[GS06] G. J. Galloway and R. Schoen. "A generalization of Hawking's black hole topology theorem to higher dimensions". *Comm. Math. Phys.* 266.2 (2006), pp. 571–576.

[GHHP83] G. W. Gibbons, S. W. Hawking, G. T. Horowitz, and M. J. Perry. "Positive mass theorems for black holes". *Communications in Mathematical Physics* 88 (1983), pp. 295–308.

[Gla96] R. T. Glassey. *The Cauchy problem in kinetic theory.* SIAM, 1996.

[Gri85] P. Grisvard. *Elliptic problems in nonsmooth domains.* Vol. 69. Classics in Applied Mathematics. 1985, pp. xx+410.

[Gro96] M. Gromov. *Positive curvature, macroscopic dimension, spectral gaps and higher signatures.* Functional analysis on the eve of the 21st century, Vol. II (New Brunswick), NJ, 1993, Progr. Math., vol. 132. 1996, pp. 1–123.

[Gro18] M. Gromov. "Metric inequalities with scalar curvature". *GAFA* 28.6 (2018), pp. 645–726.

[Gro20] M. Gromov. "No metrics with Positive Scalar Curvature on Aspherical 5-Manifolds" (2020). arXiv: 2009.05332.

[Gro23] M. Gromov. "Four Lectures on Scalar Curvature". *Perspectives in scalar curvature.* World Scientific, 2023, pp. 1–514.

[GM07] C. Gundlach and J. M. Martín-García. "Critical Phenomena in Gravitational Collapse". *Living Rev. Relativ.* 10.1 (2007).

[Haw71] S. W. Hawking. "Gravitational radiation from colliding black holes". *Phys. Rev. Lett.* 26.21 (1971), p. 1344.

[Haw72a] S. W. Hawking. "Black holes in general relativity". *Communications in Mathematical Physics* 25.2 (1972), pp. 152–166.

[Haw72b] S. W. Hawking. "Black holes in general relativity". *Comm. Math. Phys.* 25 (1972), pp. 152–166.

[Haw75] S. W. Hawking. "Particle creation by black holes". *Comm. Math. Phys.* 43.3 (1975), pp. 199–220.

[HE73] 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, pp. xi+391.

[Her98] M. Herzlich. "The positive mass theorem for black holes revisited". *J. Geom. Phys.* 26.2 (1998).

[HKK21] S. Hirsch, D. Kazaras, and M. Khuri. "Spacetime harmonic functions and the mass of 3-dimensional asymptotically flat initial data for the Einstein equations". To appear in *J. Differential Geom.* (2021).

[HZ23] S. Hirsch and Y. Zhang. "The case of equality for the spacetime positive mass theorem". *The Journal of Geometric Analysis* 33.1 (2023), p. 30.

[HZ24] S. Hirsch and Y. Zhang. "Initial data sets with vanishing mass are contained in pp-wave spacetimes" (2024). arXiv: 2403.15984.

[HL24] L.-H. Huang and D. A. Lee. "Bartnik mass minimizing initial data sets and improvability of the dominant energy scalar". *Journal of Differential Geometry* 126.2 (2024), pp. 741–800.

[HL20] L.-H. Huang and D. A. Lee. "Equality in the spacetime positive mass theorem". *Comm. Math. Phys.* 376.3 (2020), pp. 2379–2407.

[HL23] L.-H. Huang and D. A. Lee. "Equality in the spacetime positive mass theorem II" (2023). arXiv: 2302.06040.

[HMM18] L.-H. Huang, D. Martin, and P. Miao. "Static potentials and area minimizing hypersurfaces". *Proc. Amer. Math. Soc.* 146.6 (2018), pp. 2647–2661.

[Hub99] V. E. Hubeny. "Overcharging a black hole and cosmic censorship". *Physical Review D* 59.6 (1999), p. 064013.

[HKM76] T. J. R. Hughes, T. Kato, and J. E. Marsden. "Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity". *Arch. Rational Mech. Anal.* 63.3 (1976), 273–294 (1977).

[Isr86] W. Israel. "Third law of black-hole dynamics: a formulation and proof". *Phys. Rev. Lett.* 57.4 (1986), pp. 397–399.

[Isr66] W. Israel. "Singular hypersurfaces and thin shells in general relativity". *Il Nuovo Cimento B (1965-1970)* 44.1 (1966), pp. 1–14.

[Isr92] W. Israel. "Thermodynamics and internal dynamics of black holes: Some recent developments". *Black hole physics.* Springer, 1992, pp. 147–183.

[Jab21] F. E. Jabiri. "Static spherically symmetric Einstein-Vlasov bifurcations of the Schwarzschild spacetime". *Ann. Henri Poincaré* 22.7 (2021), pp. 2355–2406.

[JS09] T. Jacobson and T. P. Sotiriou. "Overspinning a black hole with a test body". *Physical review letters* 103.14 (2009), p. 141101.

[Jan78] P. S. Jang. "On the positivity of energy in general relativity". *J. Math. Phys.* 19.5 (1978), pp. 1152–1155.

[KM24] L. Kehberger and H. Masaood. "The Case Against Smooth Null Infinity V: Early-Time Asymptotics of Linearised Gravity Around Schwarzschild for Fixed Spherical Harmonic Modes" (2024). arXiv: 2401.04179.

[Keh22a] C. Kehle. "Diophantine approximation as cosmic censor for Kerr-AdS black holes". *Invent. Math.* 227.3 (2022), pp. 1169–1321.

[KU22] C. Kehle and R. Unger. "Gravitational collapse to extremal black holes and the third law of black hole thermodynamics". To appear in *J. Eur. Math. Soc. (JEMS)* (2022).

[KU23] C. Kehle and R. Unger. "Event horizon gluing and black hole formation in vacuum: the very slowly rotating case" (2023). arXiv: 2304.08455.

[KV21] C. Kehle and M. Van de Moortel. "Strong Cosmic Censorship in the presence of matter: the decisive effect of horizon oscillations on the black hole interior geometry". To appear in *Analysis & PDE* (2021).

[Keh23] L. Kehrberger. "Mathematical Studies on the Asymptotic Behaviour of Gravitational Radiation in General Relativity". PhD thesis. University of Cambridge, 2023.

[Keh22b] L. M. A. Kehrberger. “The case against smooth null infinity I: Heuristics and counter-examples”. *Ann. Henri Poincaré* 23.3 (2022), pp. 829–921.

[KLM11] A. Kleinwächter, H. Labranche, and R. Meinel. “On the black hole limit of rotating discs and rings”. *General Relativity and Gravitation* 43 (2011), pp. 1469–1486.

[Kom13] J. Kommemi. “The global structure of spherically symmetric charged scalar field spacetimes”. *Comm. Math. Phys.* 323.1 (2013), pp. 35–106.

[Kuc68] K. Kuchař. “Charged shells in general relativity and their gravitational collapse”. *Czechoslovak Journal of Physics B* 18.4 (1968), pp. 435–463.

[Kui49] N. H. Kuiper. “On conformally-flat spaces in the large”. *Ann. of Math. (2)* 50 (1949), pp. 916–924.

[Kuw90] E. C. Kuwert. Der Minimalflächenbeweis des Positive Energy Theorem. Diplomarbeit, Universität Bonn, 1990, pp. xiii+491.

[LZ91] K. Lake and T. Zannias. “Structure of singularities in the spherical gravitational collapse of a charged null fluid”. *Physical Review D* 43.6 (1991), p. 1798.

[Lee19] D. A. Lee. *Geometric Relativity. Graduate Studies in Mathematics, Vol. 201, AMS*, 2019.

[LLU22] D. A. Lee, M. Lesourd, and R. Unger. “Density and positive mass theorems for initial data sets with boundary”. *Comm. Math. Phys.* 395.2 (2022), pp. 643–677.

[LLU23] D. A. Lee, M. Lesourd, and R. Unger. “Density and positive mass theorems for incomplete manifolds”. *Calc. Var. Partial Differential Equations* 62.7 (2023), Paper No. 194.

[LLU24] D. A. Lee, M. Lesourd, and R. Unger. “Noncompact fill-ins of Bartnik data”. *The Journal of Geometric Analysis* 34.4 (2024), p. 102.

[Lee13] J. M. Lee. *Introduction to smooth manifolds.* Second Ed. Vol. 218. Graduate Texts in Mathematics. Springer, New York, 2013, pp. xvi+708.

[LP87] J. M. Lee and T. H. Parker. “The Yamabe problem”. *Bull. Amer. Math. Soc. (N.S.)* 17.1 (1987), pp. 37–91.

[Lem33] G. Lemaître. “L’univers en expansion”. *Annales de la Société scientifique de Bruxelles.* Vol. 53. 1933, p. 51.

[LUY20] M. Lesourd, R. Unger, and S.-T. Yau. "Positive scalar curvature on noncompact manifolds and the Liouville theorem". To appear in *Comm. Anal. Geom.* (2020).

[LUY21] M. Lesourd, R. Unger, and S.-T. Yau. "The positive mass theorem with arbitrary ends". To appear in *J. Differential Geom.* (2021).

[LM20] J. Li and H. Mei. "A construction of collapsing spacetimes in vacuum". *Comm. Math. Phys.* 378.2 (2020), pp. 1343–1389.

[LY15] J. Li and P. Yu. "Construction of Cauchy data of vacuum Einstein field equations evolving to black holes". *Ann. of Math. (2)* 181.2 (2015), pp. 699–768.

[Li18] Y. Li. "Ricci flow on asymptotically Euclidean manifolds". *Geom. Topol.* 22.3 (2018), pp. 1837–1891.

[Lic55] A. Lichnerowicz. *Théories relativistes de la gravitation et de l'électromagnetisme Relativité Gén- érale et Théories Unitaires.* Paris, 1955.

[LM85] R. B. Lockhart and R. C. McOwen. "Elliptic differential operators on noncompact manifolds". *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 12.3 (1985), pp. 409–447.

[Loh99] J. Lohkamp. "Scalar curvature and hammocks". *Math. Ann.* 313.3 (1999), pp. 385–407.

[Loh16] J. Lohkamp. "The higher dimensional positive mass theorem I" (2016). arXiv: `math/0608795`.

[Loh17] J. Lohkamp. "The higher dimensional positive mass theorem II" (2017). arXiv: `1612.07505`.

[LMRT13] J. Lucietti, K. Murata, H. S. Reall, and N. Tanahashi. "On the horizon instability of an extreme Reissner-Nordström black hole". *J. High Energy Phys.* 3 (2013), 035, front matter+43.

[LW99] A. Lue and E. J. Weinberg. "Magnetic monopoles near the black hole threshold". *Physical Review D* 60.8 (1999), p. 084025.

[LW00] A. Lue and E. J. Weinberg. "Gravitational properties of monopole spacetimes near the black hole threshold". *Physical Review D* 61.12 (2000), p. 124003.

[Luk12] J. Luk. "On the local existence for the characteristic initial value problem in general relativity". *Int. Math. Res. Not. IMRN* 20 (2012), pp. 4625–4678.

[LO15] J. Luk and S.-J. Oh. "Quantitative decay rates for dispersive solutions to the Einstein-scalar field system in spherical symmetry". *Analysis & PDE* 8.7 (2015), pp. 1603–1674.

[LO19] J. Luk and S.-J. Oh. "Strong cosmic censorship in spherical symmetry for two-ended asymptotically flat initial data I. The interior of the black hole region". *Ann. of Math. (2)* 190.1 (2019), pp. 1–111.

[LOY18] J. Luk, S.-J. Oh, and S. Yang. "Solutions to the Einstein-scalar-field system in spherical symmetry with large bounded variation norms". *Ann. PDE* 4.1 (2018), Paper No. 3, 59.

[MOT23] Y. Mao, S.-J. Oh, and Z. Tao. "Initial data gluing in the asymptotically flat regime via solution operators with prescribed support properties" (2023). arXiv: `2308.13031`.

[Max05] D. Maxwell. "Solutions of the Einstein constraint equations with apparent horizon boundaries". *Comm. Math. Phys.* 253.3 (2005), pp. 561–583.

[McO79] R. C. McOwen. "The behavior of the Laplacian on weighted Sobolev spaces". *Comm. Pure Appl. Math.* 32.6 (1979), pp. 783–795.

[Mei06] R. Meinel. "On the black hole limit of rotating fluid bodies in equilibrium". *Classical and Quantum Gravity* 23.4 (2006), p. 1359.

[Mei+08] R. Meinel, M. Ansorg, A. Kleinwächter, G. Neugebauer, and D. Petroff. *Relativistic figures of equilibrium*. Cambridge University Press, 2008.

[MH11] R. Meinel and M. Hütten. "On the black hole limit of electrically counterpoised dust configurations". *Classical and Quantum Gravity* 28.22 (2011), p. 225010.

[Met10] J. Metzger. "Blowup of Jang's equation at outermost marginally trapped surfaces". *Comm. Math. Phys.* 294.1 (2010), pp. 61–72.

[Mor66] C. B. Morrey Jr. *Multiple integrals in the calculus of variations*. Die Grundlehren der mathematischen Wissenschaften, Band 130. Springer-Verlag New York, Inc., New York, 1966, pp. ix+506.

[Mos17] G. Moschidis. "The Einstein–null dust system in spherical symmetry with an inner mirror: structure of the maximal development and Cauchy stability" (2017). arXiv: `1704.08685`.

[Mos18] G. Moschidis. "The characteristic initial-boundary value problem for the Einstein–massless Vlasov system in spherical symmetry" (2018). arXiv: `1812.04274`.

[Mos20] G. Moschidis. "A proof of the instability of AdS for the Einstein-null dust system with an inner mirror". *Analysis & PDE* 13.6 (2020), pp. 1671–1754.

[Mos23] G. Moschidis. "A proof of the instability of AdS for the Einstein-massless Vlasov system". *Invent. Math.* 231.2 (2023), pp. 467–672.

[MRT13] K. Murata, H. S. Reall, and N. Tanahashi. "What happens at the horizon (s) of an extreme black hole?" *Classical Quantum Gravity* 30.23 (2013), p. 235007.

[Ner26] W. Nernst. *The New Heat Theorem: Its foundations in theory and experiment.* EP Dutton, 1926.

[NM95] G. Neugebauer and R. Meinel. "General relativistic gravitational field of a rigidly rotating disk of dust: Solution in terms of ultraelliptic functions". *Physical Review Letters* 75.17 (1995), p. 3046.

[Nir01] L. Nirenberg. *Topics in nonlinear functional analysis.* Vol. 6. Courant Lecture Notes in Mathematics. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2001, pp. xii+145.

[Nou05] P. Noundjeu. "The Einstein-Vlasov-Maxwell (EVM) system with spherical symmetry". *Classical Quantum Gravity* 22.24 (2005), pp. 5365–5384.

[NN04] P. Noundjeu and N. Noutchegueme. "Local existence and continuation criterion for solutions of the spherically symmetric Einstein-Vlasov-Maxwell system". *Gen. Relativity Gravitation* 36.6 (2004), pp. 1373–1398.

[NNR04] P. Noundjeu, N. Noutchegueme, and A. D. Rendall. "Existence of initial data satisfying the constraints for the spherically symmetric Einstein–Vlasov–Maxwell system". *Journal of Mathematical Physics* 45.2 (2004), pp. 668–676.

[ONe83] B. O'Neill. *Semi-Riemannian geometry with applications to relativity.* Academic press, 1983.

[ONe95] B. O'Neill. *The geometry of Kerr black holes.* A K Peters, Ltd., Wellesley, MA, 1995, pp. xviii+381.

[OC02] I. Olabarrieta and M. W. Choptuik. "Critical phenomena at the threshold of black hole formation for collisionless matter in spherical symmetry". *Phys. Rev. D (3)* 65.2 (2002), pp. 024007, 10.

[OS39] J. R. Oppenheimer and H. Snyder. "On continued gravitational contraction". *Phys. Rev. (2)* 56.5 (1939), pp. 455–459.

[Ori91] A. Ori. "Charged null fluid and the weak energy condition". *Classical Quantum Gravity* 8.8 (1991), pp. 1559–1575.

[PT74] D. N. Page and K. S. Thorne. "Disk-accretion onto a black hole. Time-averaged structure of accretion disk". *The Astrophysical Journal* 191 (1974), pp. 499–506.

[Pen65] R. Penrose. "Gravitational collapse and space-time singularities". *Phys. Rev. Lett.* 14 (1965), pp. 57–59.

[Pet16] P. Petersen. *Riemannian geometry.* Third. Vol. 171. Graduate Texts in Mathematics. Springer, Cham, 2016, pp. xviii+499.

[PR07] F. Planchon and I. Rodnianski. "On uniqueness for the Cauchy problem in general relativity" (2007).

[PS68] J. Plebanski and J. Stachel. "Einstein tensor and spherical symmetry". *Journal of Mathematical Physics* 9.2 (1968), pp. 269–283.

[Pró83] M. Prószyński. "Thin charged shells and the violation of the third law of black hole mechanics". *Gen. Relativity Gravitation* 15.5 (1983), pp. 403–415.

[RR92] G. Rein and A. D. Rendall. "Global existence of solutions of the spherically symmetric Vlasov-Einstein system with small initial data". *Comm. Math. Phys.* 150.3 (1992), pp. 561–583. Corrected in "Erratum: Global existence of solutions of the spherically symmetric Vlasov-Einstein system with small initial data". *Comm. Math. Phys.* 176.2 (1996), pp. 475–478.

[Rei94] G. Rein. "Static solutions of the spherically symmetric Vlasov-Einstein system". *Math. Proc. Cambridge Philos. Soc.* 115.3 (1994), pp. 559–570.

[RR93] G. Rein and A. D. Rendall. "Smooth static solutions of the spherically symmetric Vlasov-Einstein system". *Annales de l'IHP Physique théorique.* Vol. 59. 4. 1993, pp. 383–397.

[RRS98] G. Rein, A. D. Rendall, and J. Schaeffer. "Critical collapse of collisionless matter: A numerical investigation". *Phys. Rev. D* 58 (4 1998), p. 044007.

[Rin13] H. Ringström. *On the topology and future stability of the universe.* Oxford Mathematical Monographs. Oxford University Press, Oxford, 2013, pp. xiv+718.

[RSR23] I. Rodnianski and Y. Shlapentokh-Rothman. "Naked singularities for the Einstein vacuum equations: the exterior solution". *Ann. of Math. (2)* 198.1 (2023), pp. 231–391.

[SW77] R. K. Sachs and H. H. Wu. *General relativity for mathematicians.* Vol. Vol. 48. Graduate Texts in Mathematics. Springer-Verlag, New York-Heidelberg, 1977, pp. xii+291.

[SZ14] O. Sarbach and T. Zannias. "Tangent bundle formulation of a charged gas". *AIP Conference Proceedings*. Vol. 1577. 1. American Institute of Physics. 2014, pp. 192–207.

[Sbi16] J. Sbierski. "On the existence of a maximal Cauchy development for the Einstein equations: a dezornification". *Ann. Henri Poincaré* 17.2 (2016), pp. 301–329.

[Sch02] M. Schechter. *Principles of functional analysis*. Second. Vol. 36. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2002, pp. xxii+425.

[Sch89] R. Schoen. "Variational theory for the total scalar curvature functional for Riemannian metrics and related topics". *Topics in calculus of variations (Montecatini Terme, 1987)*. Vol. 1365. Lecture Notes in Math. Springer, Berlin, 1989, pp. 120–154.

[SY79a] R. Schoen and S.-T. Yau. "On the proof of the positive mass conjecture in general relativity". *Comm. Math. Phys.* 65.1 (1979), pp. 45–76.

[SY79b] R. Schoen and S.-T. Yau. "On the structure of manifolds with positive scalar curvature". *Manuscripta Math.* 28.1-3 (1979), pp. 159–183.

[SY81a] R. Schoen and S.-T. Yau. "Proof of the positive mass theorem. II". *Comm. Math. Phys.* 79.2 (1981), pp. 231–260.

[SY81b] R. Schoen and S.-T. Yau. "The energy and the linear momentum of space-times in general relativity". *Comm. Math. Phys.* 79.1 (1981), pp. 47–51.

[SY88] R. Schoen and S.-T. Yau. "Conformally flat manifolds, Kleinian groups and scalar curvature". *Invent. Math.* 92.1 (1988), pp. 47–71.

[SY94] R. Schoen and S.-T. Yau. *Lectures on Differential Geometry.* International Press, 1994.

[SY19] R. Schoen and S.-T. Yau. "Positive scalar curvature and minimal hypersurface singularities". *Surveys in Differential Geometry* 24.1 (2019), pp. 441–480.

[Sch16] K. Schwarzschild. "Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie". *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften* (1916), pp. 189–196.

[Sin24] J. Singh. "High regularity waves on self-similar naked singularity interiors: decay and the role of blue-shift" (2024). arXiv: 2402.00062.

[SW17] J. Sorce and R. M. Wald. "Gedanken experiments to destroy a black hole. II. Kerr-Newman black holes cannot be overcharged or overspun". *Phys. Rev. D* 96.10 (2017), pp. 104014, 19.

[SI80] B. T. Sullivan and W. Israel. "The third law of black hole mechanics: What is it?" *Phys. Lett. A* 79.5-6 (1980), pp. 371–372.

[Tay15] M. Taylor. "The Global Nonlinear Stability of Minkowski Space for the Massless Einstein–Vlasov System". PhD thesis. University of Cambridge, 2015.

[TdC20] R. Teixeira da Costa. "Mode stability for the Teukolsky equation on extremal and subextremal Kerr spacetimes". *Comm. Math. Phys.* 378.1 (2020), pp. 705–781.

[Tha19] M. Thaller. "Existence of static solutions of the Einstein-Vlasov-Maxwell system and the thin shell limit". *SIAM J. Math. Anal.* 51.3 (2019), pp. 2231–2260.

[Tha20] M. Thaller. "Rotating clouds of charged Vlasov matter in general relativity". *Classical Quantum Gravity* 37.3 (2020), pp. 035008, 36.

[Tho74] K. S. Thorne. "Disk-accretion onto a black hole. II. Evolution of the hole". *Astrophysical Journal, Vol. 191, pp. 507-520 (1974)* 191 (1974), pp. 507–520.

[TA14] J. M. Torres and M. Alcubierre. "Gravitational collapse of charged scalar fields". *Gen. Relativity Gravitation* 46.9 (2014), pp. 1–36.

[Van18a] M. Van de Moortel. "Stability and instability of the sub-extremal Reissner–Nordström black hole interior for the Einstein–Maxwell–Klein–Gordon equations in spherical symmetry". *Communications in Mathematical Physics* 360 (2018), pp. 103–168.

[Van18b] M. Van de Moortel. "Stability and instability of the sub-extremal Reissner-Nordström black hole interior for the Einstein-Maxwell-Klein-Gordon equations in spherical symmetry". *Comm. Math. Phys.* 360.1 (2018), pp. 103–168.

[Van23] M. Van de Moortel. "The breakdown of weak null singularities inside black holes". *Duke Math. J.* 172.15 (2023), pp. 2957–3012.

[Vel23] R. A. Velozo Ruiz. "Linear and non-linear collisionless many-particle systems". PhD thesis. University of Cambridge, 2023.

[Wal74] R. Wald. "Gedanken experiments to destroy a black hole". *Annals of Physics* 82.2 (1974), pp. 548–556.

[Wal84] R. M. Wald. *General relativity.* University of Chicago Press, Chicago, IL, 1984, pp. xiii+491.

[Wal97] R. M. Wald. "The Nernst theorem and black hole thermodynamics". *Phys. Rev. D* 56.10 (1997), p. 6467.

[Wal01] R. M. Wald. "The Thermodynamics of Black Holes". *Living Rev. in Rel.* 4.6 (2001).

[Wei23] M. Weissenbacher. "Decay and non-decay for the massless Vlasov equation on subextremal and extremal Reissner–Nordström black holes" (2023). arXiv: 2303.15338.

[Wit81] E. Witten. "A new proof of the positive energy theorem". *Comm. Math. Phys.* 80.3 (1981), pp. 381–402.

[Zhu21] J. Zhu. "Width estimate and doubly warped product". *Trans. Amer. Math. Soc.* 374.2 (2021), pp. 1497–1511.

[Zhu23] J. Zhu. "Rigidity results for complete manifolds with nonnegative scalar curvature". *Journal of Differential Geometry* 125.3 (2023), pp. 623–644.