

Topical Review

The relativistic Euler equations: ESI notes on their geo-analytic structures and implications for shocks in 1D and multi-dimensions

Leonardo Abbrescia  and Jared Speck* 

Vanderbilt University, Nashville, TN, United States of America

E-mail: jared.speck@vanderbilt.edu

Received 5 May 2023; revised 25 August 2023

Accepted for publication 20 October 2023

Published 16 November 2023



CrossMark

Abstract

In this article, we provide notes that complement the lectures on the relativistic Euler equations and shocks that were given by the second author at the program *Mathematical Perspectives of Gravitation Beyond the Vacuum Regime*, which was hosted by the Erwin Schrödinger International Institute for Mathematics and Physics in Vienna in February 2022. We set the stage by introducing a standard first-order formulation of the relativistic Euler equations and providing a brief overview of local well-posedness in Sobolev spaces. Then, using Riemann invariants, we provide the first detailed construction of a localized subset of the maximal globally hyperbolic developments of an open set of initially smooth, shock-forming isentropic solutions in 1D, with a focus on describing the singular boundary and the Cauchy horizon that emerges from the singularity. Next, we provide an overview of the new second-order formulation of the 3D relativistic Euler equations derived in Disconzi and Speck (2019 *Ann. Henri Poincaré* **20** 2173–270), its rich geometric and analytic structures, their implications for the mathematical theory of shock waves, and their connection to the setup we use in our 1D analysis of shocks. We then highlight some key prior results on the study of shock formation and related

* Author to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

problems. Furthermore, we provide an overview of how the formulation of the flow derived in Disconzi and Speck (2019 *Ann. Henri Poincaré* **20** 2173–270) can be used to study shock formation in multiple spatial dimensions. Finally, we discuss various open problems tied to shocks.

Keywords: acoustical metric, Cauchy horizon, eikonal function, maximal globally hyperbolic development, Riemann invariant, shocks, singular boundary

Mathematics Subject Classification (2020) Primary: 35L67, Secondary, 35M11, 35L65, 35L81, 76L05, 76Y05

(Some figures may appear in colour only in the online journal)

Contents

1. Introduction	4
1.1. Motivation, context, and the structure of the article	5
1.1.1. Mathematical and physical motivations	5
2. Standard formulations of the relativistic Euler equations and local well-posedness	6
2.1. The basic setup	7
2.1.1. The basic fluid variables and relations	7
2.1.2. A first statement of the relativistic Euler equations	7
2.1.3. Equation of state and thermodynamic relations	8
2.2. Minkowski metric assumption and Minkowski-rectangular coordinates	8
2.2.1. Index conventions	9
2.3. A well-posed first-order formulation	9
2.3.1. Equations of variation	10
2.3.2. The acoustical metric	10
2.3.3. The arrays \mathbf{V} and $\dot{\mathbf{V}}$	12
2.3.4. Energy currents and energy identity in differential form	12
2.3.5. Local well-posedness and continuation principle in Sobolev spaces	15
3. A study of the maximal development in simple isentropic plane-symmetry	16
3.1. A description of the symmetry class	16
3.2. The basic setup and conventions in plane-symmetry	16
3.3. Null frame	17
3.4. Riemann invariants	18
3.5. Brief remarks on Riccati-type blowup and genuine nonlinearity	20
3.6. Simple isentropic plane-symmetry and nonlinear geometric optics	20
3.6.1. Explicit solution formula in geometric coordinates	24
3.6.2. The evolution equation for μ , the non-degeneracy condition, and explicit expressions for various solution variables	24
3.7. Initial data that lead to admissible shock-forming solutions	26
3.7.1. Assumptions on the ‘seed’ profile	26
3.7.2. The initial data for $\mathcal{R}_{(+)}^{\text{PS}}$ and amplitude smallness assumption	27
3.7.3. The time of first blowup	28
3.7.4. Sharp estimates for the inverse foliation density	29
3.8. The crease, singular curve, and Cauchy horizon in 1D simple isentropic plane-symmetry	30

3.8.1. Singular curve, crease, Cauchy horizon, and developments	31
3.8.2. The main theorem	33
4. The new formulation of the flow in three spatial dimensions	41
4.1. Connection to quasilinear wave equations	41
4.2. Additional fluid variables	41
4.2.1. Covariant wave operator	42
4.2.2. Null forms	42
4.2.3. Two vectorfields	43
4.2.4. The geometric wave-transport-div-curl formulation of the flow	43
4.2.5. Comments on the case of more general ambient spacetimes $(\mathcal{M}, \mathbf{g})$	44
5. Some prior works on shocks	44
5.1. Comments on the 1D theory	44
5.2. Blowup-results in multi-dimensions without symmetry: proofs by contradiction	45
5.3. Proofs of shock formation in multi-dimensions via nonlinear geometric optics	45
5.4. A different approach to multi-dimensional singularity formation	48
5.5. h -MGHDs	48
5.6. Shock development problem	51
5.7. Rarefaction waves	52
6. Some ideas behind the proof of shock formation in 3D without symmetry	52
6.1. Conjectures for the 3D relativistic Euler equations	52
6.2. Almost Riemann invariants	53
6.3. Nonlinear geometric optics	54
6.4. Geometric coordinates and vectorfield frames tied the eikonal function	55
6.4.1. t Geometric coordinates	56
6.4.2. Vectorfield frames	56
6.4.3. The first fundamental form of $\ell_{t,U}$ and related decompositions	57
6.5. Description of the compactly supported initial data of perturbations of simple isentropic plane-symmetric solutions	58
6.6. Bootstrap assumptions and the regions $\mathcal{M}_{[0,t'),[-U_1,U']}$	59
6.7. The behavior of μ and the formation of a shock	60
6.8. Vectorfield multiplier method and energies	61
6.8.1. Energies and null fluxes for wave equations	61
6.8.2. Energies and null fluxes for transport equations	63
6.8.3. Commutator method and the shock-driving terms	64
6.9. Energy estimates	67
6.9.1. Transport variable energy estimates	67
6.9.2. Elliptic-hyperbolic estimates for the vorticity and entropy	67
6.9.3. The wave energy estimate hierarchy	68
6.9.4. Discussion of the proof of the top-order inequality (198)	71
6.9.5. Discussion of the proof of the just-below-top-order inequality (201)	72
7. Open problems	73
Data availability statement	75
Acknowledgments	75
References	76

1. Introduction

In this article, we provide notes that complement the lectures on the relativistic Euler equations and shocks that were given by the second author at the program *Mathematical Perspectives of Gravitation Beyond the Vacuum Regime*, which was hosted by the Erwin Schrödinger International Institute for Mathematics and Physics in Vienna in February 2022. Broadly speaking, our main goals are the following:

- (1) To briefly introduce the equations.
- (2) To provide an overview of the new formulation of the equations derived in [42].
- (3) To describe the implications of the formulation from [42] for the study of multi-dimensional shocks.

Although our main interest is solutions in $3D$ *without* symmetry assumptions, we highlight section 3, in which we rigorously study a family of initially smooth, simple isentropic shock-forming *plane-symmetric* solutions and give a complete description of a portion of their maximal (classical) globally hyperbolic developments (MGHDs for short), up to the boundary. To the best of our knowledge, this is the first article on *any* quasilinear hyperbolic system in $1D$ that provides the detailed construction of the MGHD within the vicinity of a shock singularity. Section 3 in particular provides an introduction to some of the main ideas behind the study of shock formation in a gentle, semi-explicit $1D$ setting, where energy estimates are not needed. Roughly, the MGHD is the largest possible classical solution + globally hyperbolic region that is launched by the initial data, and when it exists and is unique (for some solutions to some hyperbolic PDEs, it is not unique—see section 5.5!), it is the holy grail object, at least from the point of view of classical solutions.

We stress outright that, although the relativistic Euler equations are posed on an ambient Lorentzian manifold $(\mathcal{M}, \mathbf{g})$, from the point of view of the causal structure of the fluid, the correct notion of ‘globally hyperbolic’ is not with respect to the spacetime metric \mathbf{g} , but rather with respect to the acoustical metric \mathbf{h} , introduced in definition 2.4. To avoid confusion, we will refer to this as \mathbf{h} -global hyperbolicity and the corresponding maximal (classical) \mathbf{h} -globally hyperbolic development (\mathbf{h} -MGHD).

Many of the techniques and geometric insights behind the modern approach to studying shocks have roots in mathematical general relativity, notably the celebrated proof by Christodoulou–Klainerman [33] of the nonlinear stability of Minkowski spacetime as a solution to Einstein’s equations. A key unifying theme between the global existence result [33] and the results on shocks that we discuss here is that *nonlinear geometric optics*, implemented via an *eikonal function* U , plays a central role in the study of the flow. We will discuss eikonal functions in detail in sections 3 and 6.

We highlight a key advantage of the formulation of the flow derived in [42].

For general solutions (in $3D$, i.e. three spatial dimensions), it allows one to implement a sharp version of nonlinear geometric optics. Nonlinear geometric optics is crucial for the study of multi-dimensional shocks, and it has other applications, such as low-regularity well-posedness (see section 2.3.5).

In line with points (1)–(3) above, this article is meant to give an introduction to the mathematical techniques and methods needed for the PDE analysis through the lens of nonlinear geometric optics. A comprehensive topical review of the rich history of shock singularities in the framework of hyperbolic conservation laws would encompass a work of substantially

larger order of magnitude. We refer the reader to [18, 40, 50] and the references therein for a good start.

1.1. Motivation, context, and the structure of the article

1.1.1. Mathematical and physical motivations. The relativistic Euler equations are an important matter model in mathematical general relativity. For example, relativistic fluids are often used in cosmology to model the average energy–matter content of spacetime, and they in particular play a central role in the standard model of cosmology. For a non-exhaustive yet extensive overview of these physical applications, we refer the reader to the texts [79, 98, 101], and the references therein. There has also been rigorous mathematical progress on the global structure of solutions when the spacetime is expanding [44, 45, 78, 85, 91]. We also highlight [30] and [31, chapter 1] for a modern introduction to the relativistic Euler equations and their connections to the laws of mechanics and thermodynamics.

Despite the above remarks, the rigorous mathematical theory of multi-dimensional solutions is far from complete. Even if one considers the relativistic Euler equations on a given spacetime background (i.e. without coupling to Einstein’s equations), there are many mathematically rich phenomena that have not been fully understood. Chief among these is the fundamental issue, going back to Riemann’s foundational work [84] on non-relativistic compressible fluids in $1D$, that initially smooth solutions can develop shock singularities in finite time. Shocks are singularities such that various fluid variables’ gradients blow up in finite time (in a precise, controlled fashion, as it turns out), though the fluid variables themselves (such as the velocity and density) remain bounded. This phenomenon is also known as *wave breaking*. The relatively mild nature of the singularity gives rise to the hope that one might—at least for short times—be able to uniquely continue the solution weakly past the first singularity¹, subject to suitable selection criteria, typically in the form of Rankine–Hugoniot-type jump conditions across a shock hypersurface (which is not known in advance) and an entropy-type condition. This weak continuation problem is known as the *shock development problem*, and in full generality, it remains open, for both the $3D$ relativistic Euler equations and their non-relativistic analog, i.e. the $3D$ compressible Euler equations; see section 5.6 for further discussion. The long-term goal is certainly the following:

Develop a rigorous *global-in-time-and-space* theory of existence and uniqueness for open sets of multi-dimensional solutions that are allowed to transition from classical to weak due to shock formation, and describe the interactions of different shock hypersurfaces, Cauchy horizons (see section 3.8), and other crucial qualitative features of the flow.

This goal is far out of reach as of present. If one could accomplish it even in a single perturbative regime, that would represent truly remarkable progress.

While there have been many works on shock formation for the multi-dimensional non-relativistic compressible Euler equations, such as [1–3, 21–24, 35, 66, 68, 92, 96], there have been relatively few works on multi-dimensional shock formation for the relativistic Euler equations. The most notable work in the relativistic case is Christodoulou’s breakthrough monograph [31], in which he proved shock formation and studied some aspects of the **h**-MGHD for open sets of irrotational and isentropic solutions in $3D$ that are compactly supported perturbations of non-vacuum constant fluid states. Unlike the non-relativistic case, there are

¹ In multi-dimensions, the correct notion of ‘first singularity’ is not a point in spacetime, but rather a co-dimension two submanifold of points; see section 5.5.

currently no rigorous results that prove multi-dimensional shock formation for the relativistic Euler equations in the presence of vorticity and entropy. However, in section 6, we provide a blueprint for how one can use the equations of [42] and the analytic framework of [1, 3, 31, 66, 68, 92, 96] to prove stable shock formation in the 3D relativistic case for open sets of solutions with vorticity and entropy.

The remainder of the article is organized as follows:

- In section 2, we set up the study of the relativistic Euler equations and provide an overview of basic ingredients that play a role in the proof of local well-posedness in Sobolev spaces. In particular, we introduce the *acoustical metric* \mathbf{h} , which is the solution-dependent Lorentzian metric that drives the propagation of sound waves; the acoustical metric is fundamental for implementing nonlinear geometric optics and studying multi-dimensional shocks. While we state the equations in the case of an arbitrary spacetime background $(\mathcal{M}, \mathbf{g})$, the vast majority of the article concerns the case in which $\mathcal{M} = \mathbb{R}^{1+3}$ and $\mathbf{g} = \mathbf{m}$ is the Minkowski metric on \mathcal{M} . We again point out that nearly all notions of hyperbolicity will refer to the acoustical metric \mathbf{h} and not \mathbf{g} or \mathbf{m} .
- In section 3, to help readers gain intuition for the subsequent discussion, we study the flow in one spatial dimension, which is equivalent to studying plane-symmetric solutions in 3D. In particular, in theorem 3.1, we provide a detailed analysis of the \mathbf{h} -MGHD for a large set of shock-forming relativistic Euler solutions in 1D. Some aspects of the theorem, notably the results concerning the formation of Cauchy horizons, have not been proved in detail elsewhere in the literature. Although theorem 3.1 specifically concerns the relativistic Euler equations, the techniques could be applied to a wide variety of strictly hyperbolic genuinely nonlinear PDEs in one spatial dimension.
- In section 4, we provide an overview of the new formulation of the 3D relativistic Euler equations derived in [42].
- In section 5, we describe some previous works on shock formation and related problems.
- In section 6, based in part on our prior experience [1, 3, 66, 68, 92, 96] in studying shock formation in multi-dimensional non-relativistic compressible fluids, we provide an overview of how the new formulation of the flow from [42] can be used to study shock formation for open sets of solutions to the 3D relativistic Euler equations. As of present, rigorous fluid shock formation results in multi-dimensions with vorticity and entropy have been proved only for the non-relativistic compressible Euler equations. Nonetheless, from the perspective of shock formation, the relativistic Euler equations and the non-relativistic compressible Euler equations enjoy many structural commonalities, and we expect that the overview we provide in section 6 could be turned (with substantial effort) into a complete proof.
- In section 7, we describe various open problems.

2. Standard formulations of the relativistic Euler equations and local well-posedness

In this section, we introduce some standard first-order formulations of the relativistic Euler equations. We also introduce the acoustical metric \mathbf{h} and set up a corresponding geometric version of the energy method that applies to a first-order formulation of the flow. Finally, we state proposition 2.3, which is a standard result on local well-posedness, and briefly discuss how its proof is connected to the energy method.

2.1. The basic setup

We start by discussing standard formulations of the relativistic Euler equations on an arbitrary four-dimensional Lorentzian manifold $(\mathcal{M}, \mathbf{g})$, where \mathbf{g} is the spacetime metric of signature $(-, +, +, +)$. Our discussion in this section is motivated by Christodoulou's presentation in [31, chapter 1]. We are somewhat terse in our presentation here, and thus we refer readers to [31, chapter 1] for additional details.

Convention 2.1 (moving indices with \mathbf{g} and \mathbf{m}). In this section, we will lower and raise indices with the spacetime metric \mathbf{g} and its inverse \mathbf{g}^{-1} respectively, i.e. $\mathbf{u}_\alpha := \mathbf{g}_{\alpha\kappa} \mathbf{u}^\kappa$ and $\xi^\alpha := (\mathbf{g}^{-1})^{\alpha\kappa} \xi_\kappa$. Similar comments apply to the bulk of the article, once we fix $\mathcal{M} = \mathbb{R}^{1+3}$ and $\mathbf{g} = \mathbf{m}$ as the Minkowski metric on \mathbb{R}^{1+3} ; see section 2.2.1. This will become especially important later on, when we introduce the acoustical metric \mathbf{h} , i.e. we do not raise or lower indices with \mathbf{h}^{-1} or \mathbf{h} .

2.1.1. The basic fluid variables and relations. We now introduce the basic fluid variables that will play a role in our study of the relativistic Euler equations: $\rho : \mathcal{M} \rightarrow [0, \infty)$ denotes the fluid's (proper) energy density, $p : \mathcal{M} \rightarrow [0, \infty)$ denotes the fluid pressure, $s : \mathcal{M} \rightarrow [0, \infty)$ denotes the entropy per particle ('entropy' for short form now on), $n : \mathcal{M} \rightarrow [0, \infty)$ denotes the proper number density, $\theta : \mathcal{M} \rightarrow [0, \infty)$ denotes the temperature,

$$H := \frac{\rho + p}{n} \quad (1)$$

denotes the enthalpy per particle, and \mathbf{u}^α denotes the fluid's four-velocity, which is a future-directed vectorfield on \mathcal{M} normalized by:

$$\mathbf{g}(\mathbf{u}, \mathbf{u}) = -1. \quad (2)$$

In the rest of the paper,

$$\bar{H} > 0 \quad (3)$$

denotes an arbitrary fixed, positive value of H ; we find it convenient to use \bar{H} to normalize various constructions.

We find it convenient to work with the natural log of the enthalpy.

Definition 2.1 (logarithmic enthalpy). We define the (dimensionless) logarithmic enthalpy h as follows:

$$h := \ln(H/\bar{H}). \quad (4)$$

Definition 2.2 (the quantity q). We define the quantity q as follows:

$$q := \frac{\theta}{H}. \quad (5)$$

2.1.2. A first statement of the relativistic Euler equations. Relative to arbitrary coordinates, the relativistic Euler equations on $(\mathcal{M}, \mathbf{g})$ can be expressed as the following system of *quasi-linear hyperbolic conservation laws*:

$$\nabla_\kappa \mathbf{T}^{\alpha\kappa} = 0, \quad (\alpha = 0, 1, 2, 3), \quad (6a)$$

$$\nabla_\kappa (n \mathbf{u}^\kappa) = 0, \quad (6b)$$

where ∇ is the Levi–Civita connection of the spacetime metric \mathbf{g} and:

$$\mathbf{T}^{\alpha\beta} := (\rho + p) \mathbf{u}^\alpha \mathbf{u}^\beta + p (\mathbf{g}^{-1})^{\alpha\beta}, \quad (\alpha, \beta = 0, 1, 2, 3), \quad (7)$$

is the energy–momentum tensor of the fluid. It turns out that under additional assumptions discussed in section 2.1.3, equation (2) can be viewed as a constraint that is preserved by the flow of the PDEs (6a); see equation (15) and the discussion surrounding it. We also emphasize that later on, we will discuss several other formulations of the flow.

2.1.3. Equation of state and thermodynamic relations. The system (2) + (6a)–(6b) + (7) is not closed because there are too many fluid variables and not enough equations. The standard path to closing the system begins with an *equation of state*, that is, an assumed functional relationship of the form $p = p(\rho, s)$. The basic hyperbolicity of the equations will require that *speed of sound* c , defined by:

$$c := \sqrt{\left. \frac{\partial p}{\partial \rho} \right|_s}, \quad (8)$$

should be real and non-negative, where $\left. \frac{\partial p}{\partial \rho} \right|_s$ denotes the partial derivative of p (i.e. of the equation of state) with respect to ρ at fixed s . The vanishing of c causes a severe degeneracy in the system, and we will therefore restrict our attention to solution regimes in which:

$$0 < c \leq 1, \quad (9)$$

where we have normalized our setup so that the speed of light is unity (i.e. the second inequality in (9) implies that the speed of sound is less than or equal to the speed of light). We also restrict our attention to solutions such that:

$$\rho > 0, p > 0, n > 0, \theta > 0, H > 0. \quad (10)$$

The laws of thermodynamics demand that the fluid variables satisfy the following functional relations:

$$H = \left. \frac{\partial \rho}{\partial n} \right|_s, \quad \theta = \left. \frac{1}{n} \frac{\partial \rho}{\partial s} \right|_n, \quad dH = \frac{dp}{n} + \theta ds, \quad (11)$$

where $\left. \frac{\partial}{\partial n} \right|_s$ denotes partial differentiation with respect to n at fixed s and $\left. \frac{\partial}{\partial s} \right|_n$ denotes partial differentiation with respect to s at fixed n .

2.2. Minkowski metric assumption and Minkowski-rectangular coordinates

In the rest of the paper, unless we explicitly state otherwise, it should be understood that the spacetime manifold \mathcal{M} is equal to \mathbb{R}^{1+3} and that \mathbf{g} is equal to the Minkowski metric, which for clarity we denote by \mathbf{m} . We fix a standard global Minkowski-rectangular coordinate system $\{x^\alpha\}_{\alpha=0,1,2,3}$ on \mathcal{M} , relative to which $\mathbf{m} = \text{diag}(-1, 1, 1, 1)$. We use the notation $\{\partial_\alpha\}_{\alpha=0,1,2,3}$ to denote the partial derivative vectorfields in this coordinate system. We also use the alternate notation $t := x^0$ and $\partial_t := \partial_0$.

2.2.1. Index conventions. From now until the end of the paper, we use the following conventions for indices.

- **(Lowercase Greek index conventions)** lowercase Greek spacetime indices α, β , etc correspond to the Minkowski-rectangular coordinates and vary over 0, 1, 2, 3. All lowercase Greek indices are lowered and raised with the Minkowski metric \mathbf{m} and its inverse \mathbf{m}^{-1} . Throughout the article, if ξ is a type $(m)_n$ spacetime tensorfield, then unless we indicate otherwise, in our identities and estimates, $\{\xi_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \alpha_m}\}_{\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n=0,1,2,3}$ **denotes its components with respect to the Minkowski-rectangular coordinates**.
- **(Lowercase Latin index conventions)** lowercase Latin spatial indices a, b , etc correspond to the Minkowski-rectangular spatial coordinates and vary over 1, 2, 3. Much like in the previous point, if ξ is a type $(m)_n$ Σ_t -tangent tensorfield, then $\{\xi_{b_1 \dots b_n}^{a_1 \dots a_m}\}_{a_1, \dots, a_m, b_1, \dots, b_n=1,2,3}$ **denotes its components with respect to the Minkowski-rectangular spatial coordinates**.
- **(Uppercase Latin index conventions)** uppercase Latin spatial indices A, B , etc correspond to the spatial coordinates (x^2, x^3) and vary over 2, 3. This will be important in section 6, where (x^2, x^3) will correspond to perturbations away from plane-symmetry.
- **(Einstein summation)** we use Einstein's summation convention in that repeated indices are summed, e.g. $L^A X^A := L^2 X^2 + L^3 X^3$.

2.3. A well-posed first-order formulation

In this section, we provide a first-order formulation of the flow in which the unknowns are taken to be $(h, \mathbf{u}^0, \mathbf{u}^1, \mathbf{u}^2, \mathbf{u}^3, s)$. Thanks to the assumptions state in section 2.1.2, the remaining fluid variables can be expressed² as functions of the unknowns.

Definition 2.3 (derivative of a scalar function with respect to a vectorfield). Throughout the paper, if \mathbf{Z} is a vectorfield and f is a scalar function, then $\mathbf{Z}f := \mathbf{Z}^\kappa \partial_\kappa f$ denotes the derivative of f in the direction of \mathbf{Z} .

With the help of (11), one can compute that for C^1 solutions, (6a)–(6b) is equivalent to the following first-order formulation of the flow:

$$\mathbf{u}^\kappa \partial_\kappa h + c^2 \partial_\kappa \mathbf{u}^\kappa = 0, \quad (12a)$$

$$\mathbf{u}^\kappa \partial_\kappa \mathbf{u}^\alpha + \Pi^{\alpha\kappa} \partial_\kappa h - q (\mathbf{m}^{-1})^{\alpha\kappa} \partial_\kappa s = 0, \quad (\alpha = 0, 1, 2, 3), \quad (12b)$$

$$\mathbf{u}^\kappa \partial_\kappa s = 0, \quad (12c)$$

$$\mathbf{m}(\mathbf{u}, \mathbf{u}) = -1, \quad (12d)$$

where:

$$\Pi^{\alpha\beta} := (\mathbf{m}^{-1})^{\alpha\beta} + \mathbf{u}^\alpha \mathbf{u}^\beta, \quad (\alpha, \beta = 0, 1, 2, 3), \quad (13)$$

denotes projection onto the \mathbf{m} -orthogonal complement of \mathbf{u} . In particular,

$$\Pi^{\alpha\kappa} \mathbf{u}_\kappa = 0, \quad (\alpha = 0, 1, 2, 3). \quad (14)$$

² By (2), one can also express $\mathbf{u}^0 = \sqrt{1 + \mathbf{u}_a \mathbf{u}^a}$.

Equations (12b)–(12d) + (13) are a first-order quasilinear hyperbolic system and are locally well-posed in suitable Sobolev spaces; see proposition 2.3. We also note by contracting (12b) against \mathbf{u}_β and using (12c), we deduce—without using (12d)—the following identity:

$$\begin{aligned} \mathbf{u}^\kappa \partial_\kappa \{ \mathbf{m}(\mathbf{u}, \mathbf{u}) + 1 \} &= \mathbf{u}^\kappa (\partial_\kappa \mathbf{u}_\lambda) \mathbf{u}^\lambda \\ &= -\{ \mathbf{m}(\mathbf{u}, \mathbf{u}) + 1 \} \mathbf{u}^\kappa \partial_\kappa h. \end{aligned} \quad (15)$$

Using (15), one can easily show via Grönwall's inequality that if $\mathbf{m}(\mathbf{u}, \mathbf{u}) + 1$ vanishes at $t = 0$, then $\mathbf{m}(\mathbf{u}, \mathbf{u}) + 1$ vanishes on any slab $[0, T] \times \mathbb{R}^3$ of classical existence on which the solution is C^1 . Hence, (12d) can be viewed as a constraint on the initial data that is preserved by the flow.

Remark 2.1 (additional terms in the context of curved spacetime backgrounds). When posed on a general smooth spacetime background $(\mathcal{M}, \mathbf{g})$, there are additional terms on RHS (12a) and (12b) of the schematic form $\boldsymbol{\Gamma} \cdot \mathbf{f}(\mathbf{V})$, where $\boldsymbol{\Gamma}$ schematically denotes Christoffel symbols of \mathbf{g} , \mathbf{V} denotes the array of fluid variable unknowns (see definition 23), and \mathbf{f} schematically denotes a smooth function. In the context of fluid shock formation, the dominant terms in the flow are Riccati-type nonlinearities in the fluid obtained after differentiating the equations one time, i.e. they are quadratic in the first derivatives of the fluid, which blow up (while \mathbf{V} remains bounded); see remark 3.5. Note that upon differentiating the term $\boldsymbol{\Gamma} \cdot \mathbf{f}(\mathbf{V})$, one obtains a term that is *linear* in the derivatives of the fluid variables. This means, in particular, that for large-gradient solutions, the differentiated terms $\partial_\alpha (\boldsymbol{\Gamma} \cdot \mathbf{f}(\mathbf{V}))$ will be small relative to the shock-driving Riccati-type fluid terms, i.e. the ‘new terms’ coming from the non-flat ambient geometry are *relatively* small. This observation is a key reason behind our expectation that one can use the methods described later in this article to prove shock formation for the relativistic Euler equations on a general spacetime $(\mathcal{M}, \mathbf{g})$, at least for initial data with suitable large gradients.

2.3.1. Equations of variation. As a preliminary step in studying the well-posedness of (12b)–(12d), we will discuss the equations satisfied by the derivatives of solutions. This is important in the sense that proofs of local well-posedness rely on differentiating the equations to obtain energy estimates for the solution's higher derivatives. Specifically, by differentiating the equations, it is straightforward to see that given a solution $(h, \mathbf{u}^\alpha, s)$ to (12a)–(12d), any derivative (of any order), which we denote by $(\dot{h}, \dot{\mathbf{u}}^\alpha, \dot{s})$, satisfies an inhomogeneous system, known as the *equations of variation*, of the following form:

$$\mathbf{u}^\kappa \partial_\kappa \dot{h} + c^2 \partial_\kappa \dot{\mathbf{u}}^\kappa = \dot{\mathfrak{F}}, \quad (16a)$$

$$\mathbf{u}^\kappa \partial_\kappa \dot{\mathbf{u}}^\alpha + \Pi^{\alpha\kappa} \partial_\kappa \dot{h} - q (\mathbf{m}^{-1})^{\alpha\kappa} \partial_\kappa \dot{s} = \dot{\mathfrak{G}}^\alpha, \quad (\alpha = 0, 1, 2, 3), \quad (16b)$$

$$\mathbf{u}^\kappa \partial_\kappa \dot{s} = \dot{\mathfrak{H}}, \quad (16c)$$

$$\mathbf{m}(\mathbf{u}, \dot{\mathbf{u}}) = \dot{\mathfrak{J}}. \quad (16d)$$

Although here we do not provide a detailed formula for the structure of the inhomogeneous terms $\dot{\mathfrak{F}}, \dots, \dot{\mathfrak{J}}$ on RHSs (16a)–(16d), it is easy to see that they feature terms involving \leq the number of derivatives represented by the variations $(\dot{h}, \dot{\mathbf{u}}^\alpha, \dot{s})$, i.e. fewer derivatives than the principal terms on LHSs (16a)–(16d).

2.3.2. The acoustical metric. A basic question about the equations of variation is: what kinds of energy estimates are available for solutions? It turns out that this question is connected to

the intrinsic geometry of the flow via the *acoustical metric*, which plays a fundamental role throughout the article.

Definition 2.4 (the acoustical metric and its inverse). We define:

$$\mathbf{h}_{\alpha\beta} := \mathbf{n} \{ c^{-2} \mathbf{m}_{\alpha\beta} + (c^{-2} - 1) \mathbf{u}_\alpha \mathbf{u}_\beta \}, \quad (17a)$$

$$(\mathbf{h}^{-1})^{\alpha\beta} := \mathbf{n}^{-1} \{ c^2 (\mathbf{m}^{-1})^{\alpha\beta} + (c^2 - 1) \mathbf{u}^\alpha \mathbf{u}^\beta \}, \quad (17b)$$

$$\mathbf{n} := c^2 + (1 - c^2) (\mathbf{u}^0)^2 \geq 1. \quad (17c)$$

We refer to \mathbf{h} as the *acoustical metric*. \mathbf{n} is just a convenient normalization factor.

With the help of (12d), one can compute that:

$$\mathbf{h}_{\alpha\gamma} (\mathbf{h}^{-1})^{\gamma\beta} = \delta_\alpha^\beta, \quad (18)$$

where δ_α^β is the Kronecker delta. That is, (18) shows that \mathbf{h}^{-1} is indeed the inverse of \mathbf{h} . Moreover, by construction, we have:

$$(\mathbf{h}^{-1})^{00} = -1. \quad (19)$$

Remark 2.2 (normalization condition for \mathbf{h}). Equation (19) is a convenient normalization condition. Other works (e.g. [31]) rely on a different normalization condition for the acoustical metric. That represents a minor difference that typically has no important implications for the study of solutions.

Remark 2.3 (subluminal speeds). Note that by (9), (17a), and (17c), $\mathbf{h}(\mathbf{Z}, \mathbf{Z}) < 0 \implies \mathbf{m}(\mathbf{Z}, \mathbf{Z}) < 0$, i.e. the \mathbf{h} -null cones are inside the \mathbf{m} -null cones. Note also that by (12d), (17a), and (17c), we have:

$$\mathbf{h}_{\kappa\lambda} \mathbf{u}^\kappa \mathbf{u}^\lambda = -\mathbf{n} < 0. \quad (20)$$

In particular, \mathbf{u} is \mathbf{h} -timelike (and thus \mathbf{m} -timelike, as one can explicitly see via (12d)).

Remark 2.4 (acoustical metric in a general ambient spacetime $(\mathcal{M}, \mathbf{g})$). When studying the relativistic Euler equations in a general ambient spacetime $(\mathcal{M}, \mathbf{g})$ equipped with coordinates (x^0, x^1, x^2, x^3) , where $t := x^0$ is a smooth time function (see section 4.2.5), the acoustical metric is (modulo the comments of remark 2.2):

$$\mathbf{h}_{\alpha\beta} := \mathbf{n} \{ c^{-2} \mathbf{g}_{\alpha\beta} + (c^{-2} - 1) \mathbf{u}_\alpha \mathbf{u}_\beta \}, \quad (21)$$

where

$$\mathbf{n} := -(\mathbf{g}^{-1})^{00} c^2 + (1 - c^2) (\mathbf{u}^0)^2 > 0, \quad (22)$$

where the positivity on RHS (22) stems from (9) and the assumption that t is a time function for \mathbf{g} , which in particular implies that relative to the coordinates (x^0, x^1, x^2, x^3) , we have $(\mathbf{g}^{-1})^{00} = (\mathbf{g}^{-1})^{\kappa\lambda} (\partial_\kappa t) \partial_\lambda t < 0$.

2.3.3. The arrays \mathbf{V} and $\dot{\mathbf{V}}$. For notational convenience, we define the *solution array* \mathbf{V} and the *variation array* $\dot{\mathbf{V}}$ as follows:

$$\mathbf{V} := (h, \mathbf{u}^0, \mathbf{u}^1, \mathbf{u}^2, \mathbf{u}^3, s), \quad (23)$$

$$\dot{\mathbf{V}} := (\dot{h}, \dot{\mathbf{u}}^0, \dot{\mathbf{u}}^1, \dot{\mathbf{u}}^2, \dot{\mathbf{u}}^3, \dot{s}). \quad (24)$$

2.3.4. Energy currents and energy identity in differential form. We now introduce energy currents, which form the basis for energy identities for the equations of variation.

Definition 2.5 (energy currents). Given a solution \mathbf{V} to (12a)–(12d) + (13), variations $\dot{\mathbf{V}} = (\dot{h}, \dot{\mathbf{u}}^\alpha, \dot{s})$ that satisfy (16a)–(16d), and scalar functions f_1 and f_2 , we define the corresponding *energy current* vectorfield $\dot{\mathbf{J}}_{(\mathbf{V}, f_1, f_2)}$ as follows:

$$\begin{aligned} \dot{\mathbf{J}}_{(\mathbf{V}, f_1, f_2)}^\alpha [\dot{\mathbf{V}}, \dot{\mathbf{V}}] := & f_1 \mathbf{u}^\alpha \dot{s}^2 - 2c^2 q \dot{\mathbf{u}}^\alpha \dot{s} - 2q \mathbf{u}^\alpha \dot{s} \dot{h} + \mathbf{u}^\alpha \dot{h}^2 + 2c^2 \dot{\mathbf{u}}^\alpha \dot{h} + c^2 \mathbf{u}^\alpha \dot{\mathbf{u}}_\kappa \dot{\mathbf{u}}^\kappa \\ & + 2c^2 \mathbf{u}^\alpha \mathbf{u}_\kappa \dot{\mathbf{u}}^\kappa \dot{h} + f_2 \mathbf{u}^\alpha (\mathbf{u}_\kappa \dot{\mathbf{u}}^\kappa)^2. \end{aligned} \quad (25)$$

Note that the components of $\dot{\mathbf{J}}$ can be viewed as quadratic forms in $\dot{\mathbf{V}}$ with coefficients that depend on the solution \mathbf{V} . Related energy currents were used by Christodoulou in [31]. A general framework for constructing energy currents in the context of *regularly hyperbolic* PDEs was developed by Christodoulou in [29].

The following lemma provides a differential version of the energy identities afforded by the energy current. The proof is a tedious but straightforward computation, omitted here (see [31, equation (1.41)]).

Lemma 2.1 (energy identity in differential form). *Under the assumptions of definition 2.5, with $\dot{\mathbf{J}}_{(\mathbf{V}, f_1, f_2)}^\alpha [\dot{\mathbf{V}}, \dot{\mathbf{V}}]$ denoting the energy current defined in (25), the following identity holds relative to the Minkowski-rectangular coordinates:*

$$\begin{aligned} \partial_\kappa \dot{\mathbf{J}}_{(\mathbf{V}, f_1, f_2)}^\kappa [\dot{\mathbf{V}}, \dot{\mathbf{V}}] = & 2f_1 \dot{s} \dot{h} - 2q \dot{h} \dot{s} - 2q \dot{s} \dot{h} + 2 \dot{h} \dot{h} + 2c^2 \dot{\mathbf{u}}_\kappa \dot{\mathbf{G}}^\kappa + 2c^2 \dot{h} \mathbf{u}_\kappa \dot{\mathbf{G}}^\kappa \\ & + 2f_2 q \mathbf{u}_\kappa \dot{\mathbf{u}}^\kappa \dot{h} + 2f_2 \mathbf{u}_\kappa \dot{\mathbf{u}}^\kappa \mathbf{u}_\lambda \dot{\mathbf{G}}^\lambda + \mathcal{Q}_{(\mathbf{V}, f_1, f_2, \partial \mathbf{V}, \partial f_1, \partial f_2)} [\dot{\mathbf{V}}, \dot{\mathbf{V}}], \end{aligned} \quad (26)$$

where $\mathcal{Q}_{(\mathbf{V}, f_1, f_2, \partial \mathbf{V}, \partial f_1, \partial f_2)} [\dot{\mathbf{V}}, \dot{\mathbf{V}}]$ is a quadratic form in $\dot{\mathbf{V}}$ with coefficients that depend on \mathbf{V} , f_1 , f_2 , and their first-order partial derivatives; it is straightforward to compute \mathcal{Q} in detail, though we do not do so here.

By integrating (26) over a spacetime region and applying the divergence theorem, one obtains an energy identity, where the energies are defined along the boundary of the region. A fundamental question is: when are the energies obtained in this fashion positive definite³? The answer to this question, provided by lemma 2.2, is tied to the acoustic geometry of the system, captured by the acoustical metric. In short, the energies are positive definite whenever the boundary of the region has \mathbf{h} -timelike normals, i.e. whenever the boundary is spacelike with respect to \mathbf{h} .

Before proving the lemma, we first introduce a subset of state-space in which the relativistic Euler equations are hyperbolic.

³ In general, energies are not useful if they are indefinite.

Definition 2.6 (regime of hyperbolicity). We define the *regime of hyperbolicity* \mathcal{H} to be the following subset of solution-variable space:

$$\mathcal{H} := \{(h, \mathbf{u}^0, \mathbf{u}^1, \mathbf{u}^2, \mathbf{u}^3, s) \in \mathbb{R}^6 \mid 0 < c(h, s) \leq 1\}. \quad (27)$$

Lemma 2.2 (positive definiteness of $\mathbf{J}_{(\mathbf{V}, f_1, f_2)}[\dot{\mathbf{V}}, \dot{\mathbf{V}}]$). Let ξ be a past-directed \mathbf{h} -timelike one-form, i.e. assume that $(\mathbf{h}^{-1})^{\kappa\lambda}\xi_\kappa\xi_\lambda < 0$ and $\xi_0 := \xi(\partial_t) > 0$. Let \mathfrak{K} be a compact subset of the set \mathcal{H} defined in (27), and assume that the solution array \mathbf{V} defined in (23) satisfies $\mathbf{V} \in \mathfrak{K}$. Then if f_1 and f_2 are sufficiently large (in a manner that depends on ξ and \mathfrak{K}), there exists a $C = C(\xi, \mathfrak{K}, f_1, f_2) > 0$ such that the quadratic form $\xi_\alpha \mathbf{J}_{(\mathbf{V}, f_1, f_2)}^\alpha[\dot{\mathbf{V}}, \dot{\mathbf{V}}]$ is positive definite in the following sense:

$$\xi_\kappa \mathbf{J}_{(\mathbf{V}, f_1, f_2)}^\kappa[\dot{\mathbf{V}}, \dot{\mathbf{V}}] \geq C|\dot{\mathbf{V}}|^2. \quad (28)$$

Remark 2.5 (motivation for lemmas 2.1 and 2.2). Although lemmas 2.1 and 2.2 might seem at first glance like unmotivated technical results, they sit within the general framework of *multiplier methods* for hyperbolic equations. To give a rough idea of the multiplier method, consider a Schwartz-class solution ϕ to the linear wave equation $-\partial_t^2\phi + \Delta_x\phi = 0$ on Minkowski spacetime $(\mathbb{R}^{1+3}, \mathbf{m})$. We define the *energy current* vectorfield \mathbf{J} to be the vectorfield with the following components relative the standard Minkowski-rectangular coordinates:

$$\mathbf{J}^0 := \frac{1}{2}(\partial_t\phi)^2 + \frac{1}{2}|\nabla_x\phi|^2, \quad (29a)$$

$$\mathbf{J}^i := -(\partial_i\phi)\partial_t\phi, \quad (i = 1, 2, 3). \quad (29b)$$

In (29a), $\nabla_x\phi := (\partial_1\phi, \partial_2\phi, \partial_3\phi)$ denotes the spatial gradient of ϕ and $|\cdot|$ is the standard Euclidean norm on \mathbb{R}^3 . Then straightforward calculations yield that $\partial_\kappa \mathbf{J}^\kappa = 0$ and that $(\partial_t)_\kappa \mathbf{J}^\kappa = \frac{1}{2}(\partial_t\phi)^2 + \frac{1}{2}|\nabla_x\phi|^2$. These two identities are analogs of (26) and (28) respectively. The upshot is that by integrating $\partial_\alpha \mathbf{J}^\alpha = 0$ over $[0, t] \times \mathbb{R}^3$, one obtains the celebrated conservation of energy formula for solutions to the linear wave equation:

$$\frac{1}{2} \int_{\{t\} \times \mathbb{R}^3} \{(\partial_t\phi)^2 + |\nabla_x\phi|^2\} dx = \frac{1}{2} \int_{\{0\} \times \mathbb{R}^3} \{(\partial_t\phi)^2 + |\nabla_x\phi|^2\} dx. \quad (30)$$

The approach is called the ‘multiplier method’ because it is a geometric version of the more standard proof of (30), in which one multiplies the wave equation by $\partial_t\phi$ and integrates by parts over $[0, t] \times \mathbb{R}^3$.

Proof. Throughout, $C = C(\xi, \mathfrak{K}, f_1, f_2)$ denotes a positive constant (often smaller than 1) that can vary from line to line.

We start by defining the scalar function ξ^\parallel by:

$$\xi^\parallel := \xi_\kappa \mathbf{u}^\kappa. \quad (31)$$

Then, using (12d), we decompose:

$$\xi_\alpha = -\xi^\parallel \mathbf{u}_\alpha + \xi_\alpha^\perp, \quad (32)$$

where the one-form ξ_α^\perp satisfies $\xi_\kappa^\perp \mathbf{u}^\kappa = 0$. In particular, ξ_α^\perp is \mathbf{m} -orthogonal to the \mathbf{m} -timelike vectorfield \mathbf{u}^α and thus ξ_α^\perp is \mathbf{m} -spacelike whenever it is non-zero, i.e. $(\mathbf{m}^{-1})^{\kappa\lambda}\xi_\kappa^\perp \xi_\lambda^\perp \geq 0$.

Since $(\mathbf{h}^{-1})^{\kappa\lambda}\xi_\kappa\xi_\lambda < 0$ and $\mathbf{h}_{\kappa\lambda}\mathbf{u}^\kappa\mathbf{u}^\lambda < 0$ (see (20)), and since ξ_α is past-directed while \mathbf{u}^α is future-directed, it follows that there is a C with $0 < C < 1$ such that:

$$C \leq \xi^{\parallel} \leq \frac{1}{C}. \quad (33)$$

We can similarly decompose $\dot{\mathbf{u}}^\alpha$ into a variation that is parallel to \mathbf{u}^α and one that vanishes upon contraction with \mathbf{u}_α . We claim that prove (28), it suffices to show that when $f_1 = f_2 = 0$, $\dot{s} = 0$, and $\mathbf{u}_\kappa\dot{\mathbf{u}}^\kappa = 0$, we have:

$$\xi_\kappa \dot{\mathbf{J}}^\kappa [\dot{\mathbf{V}}, \dot{\mathbf{V}}] \geq C (\dot{h}^2 + \dot{\mathbf{u}}_\kappa \dot{\mathbf{u}}^\kappa). \quad (34)$$

The estimate (34) would yield control over \dot{h}^2 and variations such that $\dot{s} = \mathbf{u}_\kappa \dot{\mathbf{u}}^\kappa = 0$ (note that $\dot{\mathbf{u}}^\alpha$ is \mathbf{m} -spacelike for such variations and thus the term $\dot{\mathbf{u}}_\kappa \dot{\mathbf{u}}^\kappa$ on RHS (34) is positive-definite in such variations). Then to handle general variations, since the first and last terms on RHS (25) yield the two coercive (in view of (33)) terms $f_1 \xi^{\parallel} \dot{s}^2$ and $f_2 \xi^{\parallel} (\mathbf{u}_\kappa \dot{\mathbf{u}}^\kappa)^2$, we can choose f_1 and f_2 to be sufficiently large (for example, we can choose them to be large, positive constants) to obtain control over \dot{s}^2 and $(\mathbf{u}_\kappa \dot{\mathbf{u}}^\kappa)^2$, and large enough so that these two coercive terms and the coercivity (34) can collectively be used to absorb (via Young's inequality) all the remaining terms on RHS (25). In total, this would yield (28).

Hence, to complete the proof of (28), we can assume that $f_1 = f_2 = 0$, $\dot{s} = 0$, and $\mathbf{u}_\kappa \dot{\mathbf{u}}^\kappa = 0$. This implies that $\dot{\mathbf{u}}^\alpha$ is \mathbf{m} -spacelike. We use (12d), (25), (31), and (32) to compute that:

$$\xi_\kappa \dot{\mathbf{J}}^\kappa [\dot{\mathbf{V}}, \dot{\mathbf{V}}] = \xi^{\parallel} \dot{h}^2 + 2c^2 \xi_\kappa^\perp \dot{\mathbf{u}}^\kappa \dot{h} + c^2 \xi^{\parallel} \dot{\mathbf{u}}_\kappa \dot{\mathbf{u}}^\kappa. \quad (35)$$

Since $(\mathbf{h}^{-1})^{\kappa\lambda}\xi_\kappa\xi_\lambda < -C < 0$, $C < \mathbf{n} < \frac{1}{C}$, and $\mathbf{u}_\kappa \dot{\mathbf{u}}^\kappa = 0$, we compute using (12d), (17b), (31), and (32) that there is a $C' > 0$ such that:

$$\mathbf{n} (\mathbf{h}^{-1})^{\kappa\lambda} \xi_\kappa \xi_\lambda = -(\xi^{\parallel})^2 + c^2 (\mathbf{m}^{-1})^{\kappa\lambda} \xi_\kappa^\perp \xi_\lambda^\perp < -C'. \quad (36)$$

We are denoting the positive constant on RHS (36) by ' C' ' (rather than ' C ') for clarity. Since ξ_α^\perp is \mathbf{m} -spacelike, we have $c^2 (\mathbf{m}^{-1})^{\kappa\lambda} \xi_\kappa^\perp \xi_\lambda^\perp \geq 0$ and thus $C' < (\xi^{\parallel})^2$. Since ξ_α^\perp and $\dot{\mathbf{u}}^\alpha$ are \mathbf{m} -spacelike, for any $\sigma > 0$, we can use (36), the Cauchy–Schwarz inequality with respect to \mathbf{m} , and Young's inequality to bound the cross term on RHS (35) as follows:

$$\begin{aligned} |2c^2 \xi_\kappa^\perp \dot{\mathbf{u}}^\kappa \dot{h}| &\leq \frac{1}{\sigma} c^4 (\mathbf{m}^{-1})^{\alpha\beta} \xi_\alpha^\perp \xi_\beta^\perp \dot{\mathbf{u}}_\kappa \dot{\mathbf{u}}^\kappa + \sigma \dot{h}^2 \\ &\leq \frac{1}{\sigma} c^2 \left\{ (\xi^{\parallel})^2 - C' \right\} \dot{\mathbf{u}}_\kappa \dot{\mathbf{u}}^\kappa + \sigma \dot{h}^2 \\ &\leq \frac{1}{\sigma} c^2 (\xi^{\parallel})^2 \{1 - C' C^2\} \dot{\mathbf{u}}_\kappa \dot{\mathbf{u}}^\kappa + \sigma \dot{h}^2, \end{aligned} \quad (37)$$

where the last C on RHS (37) is from (33). We now set $\sigma := \xi^{\parallel} \{1 - C' C^2\}^{1/2}$. It follows that:

$$|2c^2 \xi_\kappa^\perp \dot{\mathbf{u}}^\kappa \dot{h}| \leq c^2 \xi^{\parallel} \{1 - C' C^2\}^{1/2} \dot{\mathbf{u}}_\kappa \dot{\mathbf{u}}^\kappa + \xi^{\parallel} \{1 - C' C^2\}^{1/2} \dot{h}^2. \quad (38)$$

Thus, since $1 - C' C^2 < 1$, we deduce from (38) that the cross term on RHS (35) can be absorbed into the two positive definite terms with room to spare. We have therefore proved (34) assuming $f_1 = f_2 = 0$, $\dot{s} = 0$, and $\mathbf{u}_\kappa \dot{\mathbf{u}}^\kappa = 0$, which completes the proof of the lemma. \square

2.3.5. Local well-posedness and continuation principle in Sobolev spaces. In proposition 2.3, we discuss local well-posedness for the relativistic Euler equations in Sobolev spaces.

Proposition 2.3 (local well-posedness and continuation principle in Sobolev spaces).

Recall the notation \mathbf{V} introduced in (23). Let $\mathring{\mathbf{V}} = (\mathring{h}, \mathring{\mathbf{u}}^0, \mathring{\mathbf{u}}^1, \mathring{\mathbf{u}}^2, \mathring{\mathbf{u}}^3, \mathring{s}) := \mathbf{V}|_{t=0}$ be initial data for the relativistic Euler equations (12a)–(12d) (in particular, assume that $\mathring{\mathbf{u}}$ satisfies (12d)). Assume that there is an integer $N \geq 3$ such that $\mathring{h}, \mathring{\mathbf{u}}^1, \mathring{\mathbf{u}}^2, \mathring{\mathbf{u}}^3, \mathring{s} \in H^N(\mathbb{R}^3)$ (and thus (12d)) and the standard Sobolev–Moser calculus imply that $\mathring{\mathbf{u}}^0 - 1 \in H^N(\mathbb{R}^3)$ and that $\mathring{\mathbf{V}}(\mathbb{R}^3)$ is contained in a compact subset of the interior of set \mathcal{H} defined in (27). Then there is a maximal $T_{\text{Lifespan}} > 0$ such that a unique solution to (12a)–(12d) exists and satisfies $h, \mathbf{u}^0 - 1, \mathbf{u}^1, \mathbf{u}^2, \mathbf{u}^3, s \in C^0([0, T_{\text{Lifespan}}], H^N(\mathbb{R}^3)) \cap C^1([0, T_{\text{Lifespan}}], H^{N-1}(\mathbb{R}^3))$. Moreover, either $T_{\text{Lifespan}} = \infty$, or one of the following two breakdown scenarios occurs:

(1) (**Blowup**). The solution's C^1 norm blows up:

$$\lim_{t \uparrow T_{\text{Lifespan}}} \|\mathbf{V}\|_{C^1([0, t] \times \mathbb{R}^3)} = \infty. \quad (39)$$

(2) (**Exiting the regime of hyperbolicity**). There exists a sequence of points $\{(t_m, x_m)\}_{m \in \mathbb{N}} \subset [0, T_{\text{Lifespan}}] \times \mathbb{R}^3$ such that for every compact subset \mathfrak{K} of the interior of \mathcal{H} , there exists an $m_{\mathfrak{K}} \in \mathbb{N}$ such that $\mathbf{V}(t_m, x_m) \notin \mathfrak{K}$ whenever $m \geq m_{\mathfrak{K}}$.

We make the following remarks.

- In sections 3 and 6, we will study the formation of shock singularities, in which the solution remains in the interior of the regime of hyperbolicity and the breakdown is precisely due to the scenario (39).
- The techniques of [51], which rely only on energy estimates, Sobolev embedding, and the fractional Sobolev–Moser calculus, can be used to save half a derivative compared to proposition 2.3, that is, to prove local well-posedness for the 3D relativistic Euler equations whenever $\mathring{h}, \mathring{\mathbf{u}}^1, \mathring{\mathbf{u}}^2, \mathring{\mathbf{u}}^3, \mathring{s} \in H^{5/2^+}(\mathbb{R}^3)$.
- In [103], a low-regularity local well-posedness result for the 3D relativistic Euler equations was proved. The low-regularity assumption was on the ‘wave-part’ of the initial data, which was assumed to belong to $H^{2^+}(\mathbb{R}^3)$, while the ‘transport-part’ (including the vorticity and entropy) was assumed to satisfy some additional smoothness assumptions, which were shown to be propagated by the flow; the terminology ‘wave-part’ and ‘transport’ is tied to the formulation of the flow provided by theorem 4.1 and is explained in [103]. Lindblad's work [65] implies that the assumptions on the wave-part of the initial data are *optimal* in the scale of Sobolev spaces due to the phenomenon of instantaneous shock formation for initial data in $H^2(\mathbb{R}^3)$. The analysis in [103] relies on the new formulation of the flow presented in theorem 4.1, which allows one to employ analytical techniques based on nonlinear geometric optics and Strichartz estimates. The techniques used in [103] have roots in the works [41, 100] (see also [104–106]) on low-regularity well-posedness for the non-relativistic compressible Euler equations, which in turn have roots in the works [56–61, 87, 99] on low-regularity well-posedness for quasilinear wave equations.

Discussion of proof of proposition 2.3. This is a standard result that can be proved by linearizing the equations, deriving *a priori* estimates for solutions to the linearized equations, and then invoking a standard iteration scheme. We refer to [88–90] for the main ideas, where an analog of proposition 2.3 was proved for the coupling of the relativistic Euler equations

to Nordström's theory of gravity. The main idea behind the *a priori* estimates (in H^N) is to construct energy currents for the linearized systems such that analogs of (26) and (28) hold, and to control RHS (26) in L^2 using the Sobolev–Moser calculus (which is standard in 3D when $N \geq 3$). \square

3. A study of the maximal development in simple isentropic plane-symmetry

Our goal in this section is to prove theorem 3.1, which yields a complete description of a localized portion of the boundary of the \mathbf{h} -MGHD for open sets of shock-forming solutions to the relativistic Euler equations (12a)–(12d) in simple isentropic plane-symmetry, where \mathbf{h} is the acoustical metric from definition 2.4; see figure 2. Roughly, the \mathbf{h} -MGHD is the largest possible classical solution + \mathbf{h} -globally hyperbolic region that is launched by the initial data. By a \mathbf{h} -globally hyperbolic region \mathfrak{R} of classical existence launched by the initial data, we mean that every inextendible \mathbf{h} -causal curve in \mathfrak{R} intersects $\Sigma_0 := \{t = 0\}$. In particular, the \mathbf{h} -MGHD cannot contain any points lying in the \mathbf{h} -causal future of any singularity, where in plane-symmetry, the \mathbf{h} -causal future of a point is effectively determined by integral curves of the two transversal null vectorfields L and \underline{L} from definition 3.1.

The portion of the boundary that we study includes a singular boundary, where the fluid's gradient blows up in a shock singularity, and a Cauchy horizon, along which the solution remains smooth. For the solutions under study, the past boundary of the Cauchy horizon intersects the past boundary of the singular boundary in a distinguished point that we refer to as ‘the crease’; see figure 3. The terminology ‘crease’ is motivated by its qualitative appearance in three spatial dimensions; see figure 6. The crease plays the role of the true initial singularity in a shock-forming relativistic Euler solution.

Section 3 is intended to provide a gentle introduction to the main ideas behind the study of the \mathbf{h} -MGHD in a simplified setting in which the geometry is easy to study and one can avoid the burden of energy estimates. It will also help set the stage for section 6, in which we outline some of the main ideas needed to study shock formation in 3D.

3.1. A description of the symmetry class

As we mentioned, we will study simple isentropic plane-symmetric solutions, which we now describe. First, by ‘isentropic solutions’, we mean that $s \equiv 0$; for any constant s_0 , we could treat the case $s \equiv s_0$ using the same arguments. From (12c), we infer that for C^1 isentropic solutions, the condition $s = s_0$ is preserved by the flow if it is satisfied at $t = 0$. By ‘plane-symmetry’, we mean that $\mathbf{u}^2 = \mathbf{u}^3 \equiv 0$ and that $h = h(t, x^1)$, $\mathbf{u}^1 = \mathbf{u}^1(t, x^1)$, and $\mathbf{u}^0 = \mathbf{u}^0(t, x^1)$. Such solutions can be viewed as solutions in 1D in which $s \equiv 0$, and throughout the rest of the paper, we make that identification. By ‘simple’, we mean that only one Riemann invariant (see section 3.4) is non-vanishing.

3.2. The basic setup and conventions in plane-symmetry

In section 3, since we are studying isentropic plane-symmetric solutions, we will slightly abuse notation and use \mathbf{h} to denote the restriction of the acoustical metric (17a) and (17b), to the (t, x^1) -plane. That is, in section 3, relative to the (t, x^1) -coordinates, we have:

$$\mathbf{h}_{\alpha\beta} := \mathbf{n} \{ c^{-2} \mathbf{m}_{\alpha\beta} + (c^{-2} - 1) \mathbf{u}_\alpha \mathbf{u}_\beta \}, \quad (\alpha, \beta = 0, 1), \quad (40a)$$

$$(\mathbf{h}^{-1})^{\alpha\beta} := \mathbf{n}^{-1} \{ c^2 (\mathbf{m}^{-1})^{\alpha\beta} + (c^2 - 1) \mathbf{u}^\alpha \mathbf{u}^\beta \}, \quad (\alpha, \beta = 0, 1), \quad (40b)$$

where n is defined by (17c), and we recall that (19) holds.

For future use, we note that for C^1 isentropic plane-symmetric solutions, the relativistic Euler equations (12a)–(12d) in Minkowski spacetime are equivalent to:

$$\mathbf{u}^\kappa \partial_\kappa h + c^2 \partial_\kappa \mathbf{u}^\kappa = 0, \quad (41a)$$

$$\mathbf{u}^\kappa \partial_\kappa \mathbf{u}^1 + \Pi^{1\kappa} \partial_\kappa h = 0, \quad (41b)$$

$$\mathbf{u}^0 = \sqrt{1 + (\mathbf{u}^1)^2}. \quad (41c)$$

For future use, we also note the following algebraic identity, which holds in plane-symmetry due to (17c) and (41c):

$$n = (\mathbf{u}^0 + \mathbf{u}^1 c) (\mathbf{u}^0 - \mathbf{u}^1 c). \quad (42)$$

To ensure that shocks can form in solutions near constant fluid states with constant enthalpy $\bar{H} > 0$, we assume the following non-degeneracy assumption:

$$1 - (\bar{c})^2 + \frac{\bar{c}'}{\bar{c}\bar{H}} \neq 0, \quad (43)$$

where in (43), we are viewing $c = c(H)$, $c' := \frac{dc}{dH}$, and LHS (43) is defined to be the constant obtained by evaluating $1 - c^2 + \frac{c'}{cH}$ at $H = \bar{H}$. The condition (43) implies in particular that the isentropic plane-symmetric relativistic Euler equations are genuinely nonlinear near the constant-state solution $(H, \mathbf{u}^1) \equiv (\bar{H}, 0)$; see also section 3.5. Christodoulou showed [31, chapter 1] that irrotational relativistic fluids (of which isentropic plane-symmetric fluids are a special case) admit a Lagrangian formulation for a potential function, and the only equation of state leading to $1 - c^2 + \frac{c'}{cH} \equiv 0$ is such that the potential function solves the timelike minimal surface equation⁴. For that equation, which satisfies a nonlinear version of the null condition that is stronger than the classic null condition identified by Klainerman [55], the first author proved that perturbations of plane-symmetric data (which are similar to the data we study here in section 3) lead to global, nonlinearly stable C^2 solutions in both \mathbb{R}^{1+1} and \mathbb{R}^{1+3} [4, 5].

3.3. Null frame

In 1D, it is possible to explicitly write down a pair of transversal null vectorfields that dictate the geometry of sound waves. As we explain in section 6.3, in more than one spatial dimension, it is not possible to explicitly write down null vectorfields in terms of the fluid variables. Rather, in that context, the null vectorfields depend on the gradient of an eikonal function U .

Definition 3.1 (null vectorfields in 1D). Relative to the Minkowski-rectangular coordinates (t, x^1) , we define the vectorfields L and \underline{L} as follows:

$$L := \partial_t + \frac{\frac{\mathbf{u}^1}{\mathbf{u}^0} + c}{1 + \frac{\mathbf{u}^1}{\mathbf{u}^0} c} \partial_{x^1}, \quad (44a)$$

$$\underline{L} := \partial_t + \frac{\frac{\mathbf{u}^1}{\mathbf{u}^0} - c}{1 - \frac{\mathbf{u}^1}{\mathbf{u}^0} c} \partial_{x^1}. \quad (44b)$$

⁴ For the minimal surface equation, $c \equiv 1$ corresponds to the trivial embedding $\mathbb{R}^{1+d} \hookrightarrow \mathbb{R}^{1+(d+1)}$ with respect to the Minkowski metric.

Straightforward calculations based on (40a) and (44a)–(44b) yield the following identities:

$$\mathbf{h}(L, L) = \mathbf{h}(\underline{L}, \underline{L}) = 0, \quad \mathbf{h}(L, \underline{L}) = -2. \quad (45)$$

In particular, (45) shows that L and \underline{L} are \mathbf{h} -null. The pair $\{L, \underline{L}\}$ is known as a **h-null frame**.

From (45) and straightforward calculations, we deduce the following lemma.

Lemma 3.1 (\mathbf{h}^{-1} in terms of L and \underline{L}). *In $1+1$ spacetime dimensions, the inverse acoustical metric \mathbf{h}^{-1} from (40b) satisfies:*

$$(\mathbf{h}^{-1})^{\alpha\beta} = -\frac{1}{2}L^\alpha \underline{L}^\beta - \frac{1}{2}\underline{L}^\alpha L^\beta, \quad (\alpha, \beta = 0, 1), \quad (46)$$

i.e. $\mathbf{h}^{-1} = -\frac{1}{2}L \otimes \underline{L} - \frac{1}{2}\underline{L} \otimes L$.

3.4. Riemann invariants

In this section, we introduce the *Riemann invariants*, which are a pair of fluid variables, one of which is constant along the integral curves of L and the other along the integral curves of \underline{L} . For isentropic plane-symmetric solutions to the relativistic Euler equations, the Riemann invariants can be used as state-space variables for the fluid, i.e. the other fluid variables are determined by them. Riemann invariants were discovered by Riemann in his proof of shock formation for the non-relativistic compressible Euler equations in $1D$. In [97], Taub discovered Riemann invariants for the isentropic relativistic Euler equations in $1D$. Studying the flow with respect to the Riemann invariants allows one to study the **h-MGHD** using techniques based on transport equations and ODE theory. In particular, energy estimates are not needed in $1D$. Although Riemann invariants are not available in multi-dimensions, approximate Riemann invariants *are* available, and they are convenient for studying perturbations of isentropic plane-symmetric solutions; see section 6.2.

To proceed, we let $F = F(H)$ be the solution to the initial value problem:

$$\frac{d}{dH}F(H) = \frac{1}{Hc}, \quad F(\bar{H}) = 0, \quad (47)$$

where $\bar{H} > 0$ is the constant (3), and in (47), we are viewing c as a function of H .

Note that since $c(\bar{H}) \neq 0$ by assumption, it follows from (47) that the function F has a local inverse F^{-1} defined near the origin such that $F^{-1}(0) = \bar{H}$.

Definition 3.2 (Riemann invariants). We define the *Riemann invariants* $\mathcal{R}_{(+)}^{\text{PS}}$ and $\mathcal{R}_{(-)}^{\text{PS}}$ as follows:

$$\mathcal{R}_{(+)}^{\text{PS}} := F(H) + \frac{1}{2} \ln \left(\frac{1 + \frac{u^1}{u^0}}{1 - \frac{u^1}{u^0}} \right), \quad (48a)$$

$$\mathcal{R}_{(-)}^{\text{PS}} := F(H) - \frac{1}{2} \ln \left(\frac{1 + \frac{u^1}{u^0}}{1 - \frac{u^1}{u^0}} \right). \quad (48b)$$

In the next proposition, for isentropic plane-symmetric solutions, we express the relativistic Euler equations in terms of $(\mathcal{R}_{(+)}^{\text{PS}}, \mathcal{R}_{(-)}^{\text{PS}})$, and we provide explicit formulae for various quantities in terms of $(\mathcal{R}_{(+)}^{\text{PS}}, \mathcal{R}_{(-)}^{\text{PS}})$.

Proposition 3.2 (isentropic plane-symmetric flow in terms of the Riemann invariants). Let $(h, \mathbf{u}^0, \mathbf{u}^1)$ be a C^1 solution to (41a)–(41c), and let L and \underline{L} be the vectorfields from definition 3.1. Then the Riemann invariants of definition 3.2 satisfy the following quasilinear transport system:

$$L\mathcal{R}_{(+)}^{\text{PS}} = 0, \quad (49a)$$

$$\underline{L}\mathcal{R}_{(-)}^{\text{PS}} = 0. \quad (49b)$$

Moreover, the components of the fluid velocity in the Minkowski-rectangular coordinate system (t, x^1) can be expressed in terms of the Riemann invariants as follows:

$$\mathbf{u}^0 = \cosh\left(\frac{\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}}{2}\right), \quad \mathbf{u}^1 = \sinh\left(\frac{\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}}{2}\right). \quad (50a)$$

In addition, with F^{-1} denoting the inverse function of the map F from definition 3.2, we have:

$$H = F^{-1}\left(\frac{1}{2}\left(\mathcal{R}_{(+)}^{\text{PS}} + \mathcal{R}_{(-)}^{\text{PS}}\right)\right). \quad (50b)$$

Finally, the following identities hold:

$$L = \partial_t + \left\{ \frac{\sinh\left(\frac{\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}}{2}\right) + c \cosh\left(\frac{\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}}{2}\right)}{\cosh\left(\frac{\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}}{2}\right) + c \sinh\left(\frac{\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}}{2}\right)} \right\} \partial_1, \quad (51a)$$

$$\underline{L} = \partial_t + \left\{ \frac{\sinh\left(\frac{\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}}{2}\right) - c \cosh\left(\frac{\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}}{2}\right)}{\cosh\left(\frac{\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}}{2}\right) - c \sinh\left(\frac{\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}}{2}\right)} \right\} \partial_1. \quad (51b)$$

Proof. Equations (49a) and (49b) follow from straightforward but tedious calculations involving definition 3.2, equations (41a)–(41c), the identity $\mathbf{u}^0 \partial_\alpha \mathbf{u}^0 = \mathbf{u}^1 \partial_\alpha \mathbf{u}^1$ (which follows from differentiating (41c) with ∂_α), and the identity $\partial_\alpha F(H) = \frac{1}{c} \partial_\alpha h$ (which follows from (4) and (47)).

To prove the identities in (50a), we first subtract (48b) from (48a) and use (41c) to compute that $\frac{1}{2}(\mathcal{R}_{(+)}^{\text{PS}} - \mathcal{R}_{(-)}^{\text{PS}}) = \ln(\mathbf{u}^0 + \mathbf{u}^1)$. Rewriting (41c) as $(\mathbf{u}^0 + \mathbf{u}^1)(\mathbf{u}^0 - \mathbf{u}^1) = 1$ and carrying out straightforward calculations based on the definitions of \cosh and \sinh , we conclude (50a). (50b) follows from adding (48a) to (48b), dividing by 2, and then applying F^{-1} to each side of the resulting identity.

Finally, (51a) and (51b) follow from definitions (44a) and (44b) and the already proven identity (50a). \square

Convention 3.1 (the fluid variables as functions of the Riemann invariants). In view of (50a) and (50b), throughout section 3, we often view all the fluid variables as smooth functions of the Riemann invariants (or of only $\mathcal{R}_{(+)}^{\text{PS}}$ when we are studying simple isentropic plane-symmetric solutions), e.g. $\mathbf{u}^0, \mathbf{u}^1, h, H = f(\mathcal{R}_{(+)}^{\text{PS}}, \mathcal{R}_{(-)}^{\text{PS}})$, where throughout the paper, f schematically denotes a smooth function that can vary from line to line. For example, we

use (47), (50b), the inverse function theorem, and the chain rule to compute the following identity, which holds for isentropic plane-symmetric solutions:

$$\partial_\alpha c = \frac{cH}{2} c' \left\{ \partial_\alpha \mathcal{R}_{(+)}^{\text{PS}} - \partial_\alpha \mathcal{R}_{(-)}^{\text{PS}} \right\}, \quad (52)$$

where on RHS (52), $c' := \frac{dc}{dH}$.

3.5. Brief remarks on Riccati-type blowup and genuine nonlinearity

Consider a simple isentropic plane-symmetric solution $\mathcal{R}_{(+)}^{\text{PS}}$ to (49a) (with $\mathcal{R}_{(-)}^{\text{PS}} = 0$). By differentiating the equation with respect to ∂_1 , one obtains a Riccati-type equation for $\partial_1 \mathcal{R}_{(+)}^{\text{PS}}$ of the form $L \partial_1 \mathcal{R}_{(+)}^{\text{PS}} = -\frac{d}{d\mathcal{R}_{(+)}^{\text{PS}}} L^1 \cdot (\partial_1 \mathcal{R}_{(+)}^{\text{PS}})^2$, where here we are viewing L^1 as a function of $\mathcal{R}_{(+)}^{\text{PS}}$, which is possible thanks to proposition 3.2 and convention 3.1. Hence, as long as the *genuine nonlinearity condition* $\frac{d}{d\mathcal{R}_{(+)}^{\text{PS}}} L^1 \neq 0$ holds, then for solutions $\mathcal{R}_{(+)}^{\text{PS}}$ launched by smooth compactly supported initial data, one can easily show that $\partial_1 \mathcal{R}_{(+)}^{\text{PS}}$ typically blows up along the integral curves of L , much like in the case of Burgers' equation $\partial_t \Psi + \Psi \partial_1 \Psi = 0$; we refer to [40, section 7.5] for a detailed discussion of the notion of genuine nonlinearity in the context of 1D hyperbolic conservation laws. One can check that the non-degeneracy assumption (43) indeed implies that $\frac{d}{d\mathcal{R}_{(+)}^{\text{PS}}} L^1|_{\mathcal{R}_{(+)}^{\text{PS}}=0} \neq 0$, i.e. that equation (49a) is genuinely nonlinear near $(\mathcal{R}_{(+)}^{\text{PS}}, \mathcal{R}_{(-)}^{\text{PS}}) = (0, 0)$.

Our study of the blowup in theorem 3.1 does not rely on differentiating the equations with ∂_1 to obtain a Riccati-type PDE. Instead, it relies on a much sharper approach tied to nonlinear geometric optics, which we set up in section 3.6. In 3D, such a sharpened approach seems essential, as there is no known simple approach for studying gradient-blowup based only on differentiating the equations with respect to a standard, solution-independent partial derivative operator. In particular, in our discussion of shock formation in 3D in section 6, nonlinear geometric optics will play a fundamental role.

3.6. Simple isentropic plane-symmetry and nonlinear geometric optics

Our study of the **h**-MGHD in theorem 3.1 intimately relies on nonlinear geometric optics, implemented via an acoustic eikonal function U , and a careful study of the corresponding acoustic geometry. The main idea of the proof of theorem 3.1 is to show that the solution remains smooth relative to the ‘geometric coordinates’ (t, U) all the way up to the shock and to recover the formation of the shock as a degeneracy between the geometric coordinates and the Minkowski-rectangular coordinates (t, x^1) .

In this section, we construct the eikonal function and corresponding acoustic geometry for the isentropic relativistic Euler equations in 1D, i.e. to the following initial value problem:

$$\begin{cases} L \mathcal{R}_{(+)}^{\text{PS}} = 0, \\ \mathcal{R}_{(+)}^{\text{PS}}|_{t=0} = \mathcal{R}_{(+)}^{\text{PS}}, \end{cases} \quad \begin{cases} \underline{L} \mathcal{R}_{(-)}^{\text{PS}} = 0, \\ \mathcal{R}_{(-)}^{\text{PS}}|_{t=0} = \mathcal{R}_{(-)}^{\text{PS}}. \end{cases} \quad (53)$$

More precisely, we restrict our attention to *simple* isentropic plane-symmetric solutions, defined here to be those solutions such that $\mathcal{R}_{(-)}^{\text{PS}} \equiv 0$. From (53) and (44b), it follows that such solutions arise from initial data with $\mathcal{R}_{(-)}^{\text{PS}} \equiv 0$. In this case, (50a) yields the identity

$\mathbf{u}^0 = \cosh(\frac{1}{2}\mathcal{R}_{(+)}^{\text{PS}})$. As we alluded to in convention 3.1, (50a) and (50b) also yield similar formulas for \mathbf{u}^1, H as functions of $\mathcal{R}_{(+)}^{\text{PS}}$.

We start by defining the acoustic eikonal function, which plays a central role in all of the forthcoming analysis.

Definition 3.3 (the acoustic eikonal function in 1D and the characteristics). Let \mathbf{h}^{-1} be the inverse acoustical metric defined in (46). Then we define the acoustic eikonal function (eikonal function for short) to be the solution U to the following fully nonlinear hyperbolic PDE with Cauchy data:

$$(\mathbf{h}^{-1})^{\kappa\lambda} \partial_\kappa U \partial_\lambda U = 0, \quad (54a)$$

$$\partial_t U > 0, \quad (54b)$$

$$U|_{t=0} = -x^1. \quad (54c)$$

We refer to the level sets of U as ‘the characteristics’.

Remark 3.1 (geometric optics is easier in 1D). Although understanding the fully nonlinear structure of (54a) is necessary in multiple space dimensions, in 1D, (54a) is equivalent to the *linear* (in U) transport equation for U derived in the following lemma. This makes the analysis of the acoustic geometry in this section considerably simpler than in multi- D . It is also unique to 1D that the \mathbf{h} -null vectorfield L can be explicitly expressed in terms of the fluid via the formula (44a). The correct analog of (44a) in higher dimensions is precisely $L = \frac{1}{L_{(\text{geo})}^0} L_{(\text{geo})}$, where $L_{(\text{geo})}^\alpha = -(\mathbf{h}^{-1})^{\alpha\kappa} \partial_\kappa U$ is a null geodesic vectorfield.

Lemma 3.3 (linear transport equation for U). *Let L be the vectorfield defined in (44a). For simple isentropic plane-symmetric solutions, if $\mathcal{R}_{(+)}^{\text{PS}}$ is sufficiently small⁵, then the scalar function U solves (54a)–(54c) if and only if it solves the following transport equation initial value problem:*

$$LU = 0, \quad (55a)$$

$$U|_{t=0} = -x^1. \quad (55b)$$

Proof. Suppose U solves (54a) and (54b). Then by (46), we find that $0 = (LU) \cdot \underline{L}U$. If $LU = 0$, there is nothing to prove. Hence, we assume that $LU \neq 0$, and we will show that this leads to a contradiction. Then by (46), we have $\underline{L}U = 0$. By (9), (51b), (54c), (55b), and our smallness assumption on $\mathcal{R}_{(+)}^{\text{PS}}$, we have:

$$\begin{aligned} \partial_t U|_{t=0} &= - \left\{ \frac{\sinh\left(\frac{1}{2}\mathcal{R}_{(+)}^{\text{PS}}\right) - c \cosh\left(\frac{1}{2}\mathcal{R}_{(+)}^{\text{PS}}\right)}{\cosh\left(\frac{1}{2}\mathcal{R}_{(+)}^{\text{PS}}\right) - c \sinh\left(\frac{1}{2}\mathcal{R}_{(+)}^{\text{PS}}\right)} \right\} \partial_1 U|_{t=0} \\ &= \frac{\sinh\left(\frac{1}{2}\mathcal{R}_{(+)}^{\text{PS}}\right) - c \cosh\left(\frac{1}{2}\mathcal{R}_{(+)}^{\text{PS}}\right)}{\cosh\left(\frac{1}{2}\mathcal{R}_{(+)}^{\text{PS}}\right) - c \sinh\left(\frac{1}{2}\mathcal{R}_{(+)}^{\text{PS}}\right)} < 0. \end{aligned} \quad (56)$$

(56) contradicts (54b).

⁵ The solutions from theorem 3.1 satisfy the needed smallness. The key point is that if $\mathcal{R}_{(+)}^{\text{PS}}$ is large, then L and \underline{L} could both be right-pointing or both be left-pointing, in which case the initial condition (55b) would not be sufficient for identifying the ‘correct root of the eikonal equation’, namely the ‘root’ $LU = 0$.

Conversely, if (55a) and (55b) hold, then since (46) implies that $(\mathbf{h}^{-1})^{\alpha\beta}\partial_\alpha U\partial_\beta U = -(LU)\cdot\underline{L}U$, it immediately follows that (54a) and (54c) hold. To prove (54b), we use (9), (51a), (55a)–(55b), and the smallness of $\mathcal{R}_{(+)}^{\text{PS}}$, arguing as in the proof of (56) to deduce that $\partial_t U > 0$, as is desired. \square

Some remarks are in order.

- The minus sign in (55b) is just a convention imposed for consistency with related results that have appeared in the literature, such as [1, 66, 68, 96].
- The precise initial condition in (55b) is not important; small perturbations of this initial condition could also be used to study the \mathbf{h} -MGHD.
- We have adapted our eikonal function to L because in theorem 3.1, we will study initial data that lead to shock-forming solutions such that the singularity is tied to the infinite density of the L -characteristics and the blowup of $\partial_1 \mathcal{R}_{(+)}^{\text{PS}}$. There is nothing special about L compared to \underline{L} , and it is only for convenience that we have adapted our constructions to it. That is, using the methods of this section, one could also construct an eikonal function U adapted to \underline{L} and study different initial data that lead to a singularity tied to the infinite density of the \underline{L} -characteristics and the blowup of $\partial_1 \mathcal{R}_{(-)}^{\text{PS}}$. One could also try to study richer problems in which singularities occur along both families of characteristics.

We now define μ , the inverse foliation density of the characteristics. The name is justified by the fact that $\frac{1}{\mu}$ is a measure of the density of the stacking of the characteristics relative to the level-sets $\Sigma_{t'} := \{t = t'\}$. Roughly speaking, μ is initially positive, while shock formation corresponds to $\mu \downarrow 0$, which signifies the infinite density of the integral curves of L (or equivalently, the level sets of U in (t, x^1) -space); we refer to figure 3, a context in which μ vanishes along the curve portion denoted by ' $\Upsilon(\mathcal{B}_{[-U_\Phi, 0]})$ ' (we will explain these issues in much more detail later on).

Definition 3.4 (inverse foliation density). We define the *inverse foliation density* of the characteristics, denoted by μ , as follows:

$$\mu := -\frac{1}{\frac{c}{n}\partial_1 U}. \quad (57)$$

In section 6, which concerns solutions in 3D, we will define μ to be $-\frac{1}{(\mathbf{h}^{-1})^{\kappa\lambda}\partial_\kappa U\partial_\lambda U}$; see (156). Under the initial condition (155) that we will choose for U , that definition of μ can be shown to agree with the one for plane-symmetric solutions stated in (57).

We also note the following identities, which follow from (55b) and (57):

$$\mu|_{t=0} = \frac{n}{c}|_{t=0}, \quad \Rightarrow \quad \frac{c}{n}\mu|_{t=0} = 1. \quad (58)$$

We now define a collection of vectorfields associated to the eikonal function.

Definition 3.5 (the vectorfields X, \check{X}, \check{L}). Let L and \underline{L} be the vectorfields from definition 3.1 (they are \mathbf{h} -null by (45)). We define:

$$X := \frac{1}{2}(\underline{L} - L), \quad \check{X} := \mu X, \quad \check{L} := \mu \underline{L}. \quad (59)$$

Using (44a)–(44b), (45), and (59), it is straightforward to see that X is Σ_t -tangent (i.e. $Xt = 0$) and satisfies $\mathbf{h}(X, X) = 1$, that $\mathbf{h}(\check{X}, \check{X}) = \mu^2$, and that $\mathbf{h}(\check{L}, \check{L}) = 0$. We also note that

in plane-symmetry, using definition 3.1, (42), and definition 3.5, one can compute the following identity:

$$X = -\frac{c}{n} \partial_1. \quad (60)$$

Our main results concern solutions in which the factor $\frac{c}{n}$ on RHS (60) satisfies $\frac{c}{n} \approx 1$.

Definition 3.6 (geometric coordinates in 1D and the corresponding partial derivative vectorfields). We define (t, U) to be the *geometric coordinates*, and we denote the corresponding geometric coordinate partial derivative vectorfields by $\{\frac{\partial}{\partial t}, \frac{\partial}{\partial U}\}$ (which are not to be confused with the partial derivatives $\{\partial_t, \partial_1\}$ in the (t, x^1) -coordinate system).

Using (44a), (55a), (57), and (60), it is straightforward to compute that:

$$Lt = \check{X}U = 1, \quad (61a)$$

$$LU = \check{X}t = 0. \quad (61b)$$

From (61a) and (61b), it follows that in geometric coordinates, the following identities hold:

$$L = \frac{\partial}{\partial t}, \quad \check{X} = \frac{\partial}{\partial U}, \quad (62)$$

$$[L, \check{X}] = 0. \quad (63)$$

From (44b), (59), and the identities in (62), we deduce the following expression for \check{L} relative to the geometric coordinates:

$$\check{L} = \mu \frac{\partial}{\partial t} + 2 \frac{\partial}{\partial U}. \quad (64)$$

We will often silently use (62) and (63) in the rest of section 3.

The following sets play a fundamental role in our analysis of solutions.

Definition 3.7 (portions of submanifolds of geometric coordinate space $\mathbb{R}_t \times \mathbb{R}_U$). Given intervals $I, J \subset \mathbb{R}$, relative to the geometric coordinates, we define:

$$\mathcal{P}_U^I := \{(t, U) \in \mathbb{R}^2 \mid t \in I, U = U'\}, \quad (65a)$$

$$\Sigma_{t'}^J := \{(t, U) \in \mathbb{R}^2 \mid t = t', U \in J\}. \quad (65b)$$

We often write \mathcal{P}_U instead of $\mathcal{P}_U^{[0, \infty)}$ and Σ_t instead of $\Sigma_t^{(-\infty, \infty)}$. We often refer to the \mathcal{P}_U as *the characteristics*.

The following lemma shows that the geometric coordinates degenerate with respect to the Minkowski-rectangular coordinates precisely when $\mu = 0$. Despite the degeneracy, theorem 3.1 shows that for appropriate initial data, the change of variables map Υ will nonetheless be a *homeomorphism* on the subset of the boundary of the h-MGHD where μ vanishes.

Lemma 3.4 (the Jacobian determinant of $(t, U) \mapsto (t, x^1)$). *Let:*

$$\Upsilon(t, U) := (t, x^1) \quad (66)$$

denote the change of variables map from geometric coordinates to Minkowski-rectangular coordinates. Then the following identity holds:

$$\det d\Upsilon = -\frac{c}{n}\mu. \quad (67)$$

Proof. (67) follows from straightforward calculations based on the identity $\frac{\partial}{\partial U}x^1 = -\frac{c}{n}\mu$, which is implied by (59)–(60) and (62). \square

3.6.1. Explicit solution formula in geometric coordinates. From (62), we see that in simple isentropic plane-symmetry, the transport equation (49a) takes the following form in geometric coordinates:

$$\frac{\partial}{\partial t}\mathcal{R}_{(+)}^{\text{PS}}(t, U) = 0. \quad (68)$$

From (68) and (53), we see that in geometric coordinates, the solution to (68) is:

$$\mathcal{R}_{(+)}^{\text{PS}}(t, U) = \mathcal{R}_{(+)}^{\text{PS}}(U). \quad (69)$$

3.6.2. The evolution equation for μ , the non-degeneracy condition, and explicit expressions for various solution variables. In this section, we derive various identities involving μ . In particular, we will later use the identities to show that suitable assumptions on the initial data of $\mathcal{R}_{(+)}^{\text{PS}}$ lead to the vanishing of μ in finite time, i.e. the formation of a shock.

We begin with the following lemma, which provides the evolution equation for μ .

Lemma 3.5 (transport equation satisfied by μ). *For simple isentropic plane-symmetric solutions, the inverse foliation density, defined in (57), satisfies the following transport equation, where $c'(H) := \frac{d}{dH}c(H)$:*

$$L\left(\frac{c}{n}\mu\right) = G, \quad (70)$$

$$G := -\frac{1 - c^2 + \frac{c'}{cH}}{2(\mathbf{u}^0 + c\mathbf{u}^1)^2} \tilde{X}\mathcal{R}_{(+)}^{\text{PS}}. \quad (71)$$

Moreover,

$$LL\mu = LL\left(\frac{c}{n}\mu\right) = 0. \quad (72)$$

Proof. First, using (51a), (52), (60), the quotient rule, the identity $\cosh^2(z) - \sinh^2(z) = 1$, and straightforward calculations, we deduce the following commutator identity, valid for simple isentropic plane-symmetric solutions:

$$[L, \partial_1] = \frac{n}{c} \left\{ \frac{1 - c^2 + \frac{c'}{cH}}{2(\mathbf{u}^0 + c\mathbf{u}^1)^2} \right\} X\mathcal{R}_{(+)}^{\text{PS}} \partial_1.$$

Equation (70) then follows from multiplying (57) by c/n , differentiating both sides with respect to L , and using the following commutator identity, which is a consequence the eikonal equation $LU = 0$:

$$L\left(\frac{c}{n}\mu\right) = \frac{1}{(\partial_1 U)^2} L\partial_1 U = \left(\frac{c}{n}\mu\right)^2 L\partial_1 U = \left(\frac{c}{n}\mu\right)^2 [L, \partial_1] U.$$

Equation (72) then follows from differentiating (70) with respect to L and using (49a), (60), (63), and the assumption of simple isentropic plane-symmetry. \square

We find it convenient to work with an *anti-derivative* of the nonlinearity G defined in (71). To this end, viewing the quantity $-\frac{1-c^2+\frac{c'}{cH}}{2(\mathbf{u}^0+c\mathbf{u}^1)^2}$ as a function of $\mathcal{R}_{(+)}^{\text{PS}}$ (see convention 3.1), we define the function $\mathcal{R}_{(+)}^{\text{PS}} \rightarrow \mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}]$ as follows:

$$\mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}] := \int_0^{\mathcal{R}_{(+)}^{\text{PS}}} \left\{ -\frac{1-c^2+\frac{c'}{cH}}{2(\mathbf{u}^0+c\mathbf{u}^1)^2} \right\} (\mathcal{R}'_+) \, d\mathcal{R}'_+. \quad (73)$$

From (62) and (73), the fundamental theorem of calculus, and the chain rule, it follows that:

$$\frac{\partial}{\partial U} \mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}] = \check{X} \mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}] = G, \quad (74a)$$

$$\mathfrak{A}[0] = 0. \quad (74b)$$

We highlight the important fact that in simple isentropic plane-symmetry, the non-degeneracy assumption (43), the initial condition of (47), and (73) collectively imply that:

$$\frac{d}{d\mathcal{R}_{(+)}^{\text{PS}}} \mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}] |_{\mathcal{R}_{(+)}^{\text{PS}}=0} = -\frac{1}{2} \left\{ 1 - (\bar{c})^2 + \frac{\bar{c}'}{\bar{c}H} \right\} \neq 0. \quad (75)$$

From (73) and (75), it follows that the function $\mathcal{R}_{(+)}^{\text{PS}} \mapsto \mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}]$ is a diffeomorphism from a neighborhood of the origin (i.e. near $\mathcal{R}_{(+)}^{\text{PS}} = 0$) onto a neighborhood of the origin; we will silently use this basic fact in the rest of section 3. (75) is fundamental for ensuring the formation of shocks, for it implies that near $(\mathcal{R}_{(+)}^{\text{PS}}, \mathcal{R}_{(-)}^{\text{PS}}) = (0, 0)$, equation (49a) is genuinely nonlinear; see also the discussion in section 3.5.

Next, we note that in simple isentropic plane-symmetry, in view of (69), the quantities G , c , and \mathbf{n} can be viewed as functions of U alone. We sometimes emphasize this fact by using notation such as $G(U)$, $c(U)$, etc. The following corollary is a straightforward application of this fact and the above discussion; we omit the straightforward proof.

Corollary 3.6 (explicit expressions for μ , $L\mu$, $\check{X}\mu$, and $\partial_1 \mathcal{R}_{(+)}^{\text{PS}}$). *Let G be the source term from (71), viewed as a function of U . Let $\mathfrak{A} = \mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}]$ be the function of $\mathcal{R}_{(+)}^{\text{PS}}$ defined by (73). Then for simple isentropic plane-symmetric solutions, the following identities hold relative to the geometric coordinates, where $\mathfrak{A}'[\mathcal{R}_+] := \frac{d}{d\mathcal{R}_+} \mathfrak{A}[\mathcal{R}_+]$:*

$$L \left\{ \frac{c}{\mathbf{n}} \mu(t, U) \right\} = G(U) = \frac{d}{dU} \mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}(U)] = \mathfrak{A}'[\mathcal{R}_{(+)}^{\text{PS}}(U)] \frac{d}{dU} \mathcal{R}_{(+)}^{\text{PS}}(U), \quad (76)$$

$$\begin{aligned} \frac{c}{\mathbf{n}} \mu(t, U) &= 1 + tG(U) = 1 + t \frac{d}{dU} \mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}(U)] \\ &= 1 + t \mathfrak{A}'[\mathcal{R}_{(+)}^{\text{PS}}(U)] \frac{d}{dU} \mathcal{R}_{(+)}^{\text{PS}}(U), \end{aligned} \quad (77)$$

$$\begin{aligned}\check{X}\mu(t, U) &= t \left(\frac{n}{c} G' \right) (U) + \left\{ \frac{d}{dU} \left(\frac{n}{c} \right) (U) \right\} \{1 + tG(U)\} \\ &= \frac{n}{c} t \frac{d^2}{dU^2} \mathfrak{A} \left[\mathring{\mathcal{R}}_{(+)}^{\text{PS}} (U) \right] + \left\{ \frac{d}{dU} \left(\frac{n}{c} \right) (U) \right\} \left\{ 1 + t \frac{d}{dU} \mathfrak{A} \left[\mathring{\mathcal{R}}_{(+)}^{\text{PS}} (U) \right] \right\},\end{aligned}\quad (78)$$

$$\begin{aligned}\left[\partial_1 \mathcal{R}_{(+)}^{\text{PS}} \right] (t, U) &= -\frac{1}{\frac{c}{n} \mu(t, U)} \frac{\partial}{\partial U} \mathcal{R}_{(+)}^{\text{PS}} (t, U) = -\frac{\frac{d}{dU} \mathring{\mathcal{R}}_{(+)}^{\text{PS}} (U)}{1 + tG(U)} \\ &= -\frac{\frac{d}{dU} \mathring{\mathcal{R}}_{(+)}^{\text{PS}} (U)}{1 + t \mathfrak{A}' \left[\mathring{\mathcal{R}}_{(+)}^{\text{PS}} (U) \right] \frac{d}{dU} \mathring{\mathcal{R}}_{(+)}^{\text{PS}} (U)}.\end{aligned}\quad (79)$$

3.7. Initial data that lead to admissible shock-forming solutions

In this section, we construct examples of initial data that fall under the scope of theorem 3.1, i.e. initial data that launch shock-forming solutions such that we can derive the structure of a portion of the maximal classical development that includes a neighborhood of its boundary. Our construction could be substantially generalized to exhibit a much larger class of initial data for which the results of theorem 3.1 hold; here we have striven for simplicity.

3.7.1. Assumptions on the ‘seed’ profile. To start, we fix any scalar ‘seed profile’ $\mathring{\varphi}$ with the following properties (it is straightforward to show that such functions exist):

- $\mathring{\varphi} = \mathring{\varphi}(U)$ is a⁶ C^∞ function that is compactly supported in an interval $[-U_1, U_2]$ of U -values, where $U_1, U_2 > 1$.
- $\frac{d}{dU} \mathring{\varphi}(U)$ has a unique, non-degenerate minimum at $U = 0$ (our choice $U = 0$ is only for convenience). In particular, $\frac{d^3}{dU^3} \mathring{\varphi}(0) > 0$.
- Modifying $\mathring{\varphi}$ by multiplying it by a constant and composing it with a linear map of the form $U \rightarrow zU$ for some constant $z \in \mathbb{R}$ if necessary, we assume that (for the modified $\mathring{\varphi}$), there is a constant $\mathring{\delta}_*$ satisfying:

$$\mathring{\delta}_* > 0 \quad (80)$$

such that:

$$\left. \frac{d}{dU} \mathring{\varphi}(U) \right|_{U=0} = -\mathring{\delta}_*, \quad \left. \frac{d^2}{dU^2} \mathring{\varphi}(U) \right|_{U=0} = 0; \quad (81)$$

that there is a constant \mathfrak{b} satisfying:

$$\mathfrak{b} > 0 \quad (82)$$

such that:

$$\frac{d^3}{dU^3} \mathring{\varphi}(0) = \mathfrak{b}, \quad (83)$$

⁶ Our proof of theorem 3.1 does not actually require $\mathring{\varphi}$ to be C^∞ ; C^4 would suffice.

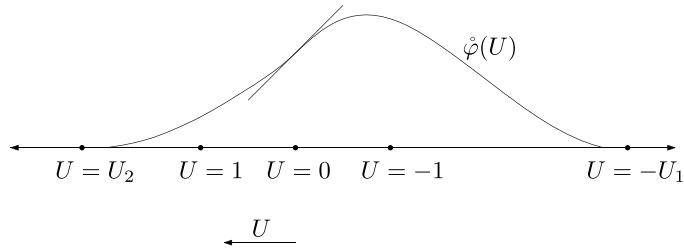


Figure 1. The graph of a representative ‘seed profile’, with U increasing from right to left.

and such that:

$$\frac{1}{2}\mathfrak{b} \leq \frac{d^3}{dU^3}\dot{\varphi}(U) \leq 2\mathfrak{b}, \quad \text{if } |U| \leq 1; \quad (84)$$

and that there is a constant \mathfrak{p} satisfying:

$$0 \leq \mathfrak{p} < 1 \quad (85)$$

such that:

$$\frac{d}{dU}\dot{\varphi}(U) > -\mathfrak{p}\delta_*, \quad \text{if } |U| \geq 1. \quad (86)$$

See figure 1 for the graph of a representative seed profile.

3.7.2. The initial data for $\mathcal{R}_{(+)}^{\text{PS}}$ and amplitude smallness assumption. We now construct the initial data for $\mathcal{R}_{(+)}^{\text{PS}}$. Let $\dot{\varphi}$ be as in section 3.7.1. We define the initial data function $\mathring{\mathcal{R}}_{(+)}^{\text{PS}} := \mathcal{R}_{(+)}^{\text{PS}}|_{t=0}$ to be the following function of U :

$$\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U) := \mathfrak{A}^{-1}[\dot{\varphi}(U)], \quad (87)$$

where \mathfrak{A}^{-1} is the inverse function of the map $\mathcal{R}_{(+)}^{\text{PS}} \rightarrow \mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}]$ defined by (73).

Amplitude smallness assumption on the initial data

By (75) and continuity, we can further modify $\dot{\varphi}$ by multiplying it by a constant to shrink the amplitude of $\mathring{\mathcal{R}}_{(+)}^{\text{PS}}$ and ensure that the following two conditions hold:

- (1) The isentropic plane-symmetric initial data array $\mathring{\mathbf{V}} := (h, \mathbf{u}^0, \mathbf{u}^1, 0, 0, 0)|_{t=0}$ determined by $\mathring{\mathcal{R}}_{(+)}^{\text{PS}}$ via (50a)–(50b) (with $\mathcal{R}_{(-)}^{\text{PS}} = 0$ in those formulas) satisfies $\mathring{\mathbf{V}} \in \text{interior}(\mathcal{H})$, where \mathcal{H} is the regime of hyperbolicity defined in (27).
- (2) The following non-degeneracy condition holds:

$$\frac{d}{d\mathcal{R}_{(+)}^{\text{PS}}}\mathfrak{A}[\mathcal{R}_{(+)}^{\text{PS}}] \Big|_{\mathcal{R}_{(+)}^{\text{PS}} = \mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)} \neq 0, \quad U \in [-U_1, U_2]. \quad (88)$$

In the rest of section 3, we will use the above two assumptions without always explicitly mentioning them.

Taking into account (43) and (74b), we deduce from Taylor expansions and standard calculus that the following conclusions hold, where $C > 0$ depends on $\dot{\varphi}$, and for functions $f = f(U)$ on $\Sigma_0 := \{t = 0\}$, $\|f\|_{C^0(\Sigma_0)} := \max_{U \in \mathbb{R}} |f(U)|$ and $\|f\|_{C^1(\Sigma_0)} := \|f\|_{C^0(\Sigma_0)} + \|\frac{d}{dU}f\|_{C^0(\Sigma_0)}$:

- $\mathring{\mathcal{R}}_{(+)}^{\text{PS}}$ is compactly supported in $[-U_1, U_2]$ and satisfies the following bounds:

$$\left\| \frac{d^M}{dU^M} \mathring{\mathcal{R}}_{(+)}^{\text{PS}} \right\|_{C^0(\Sigma_0)} \leq C, \quad M = 0, 1, 2, 3, 4. \quad (89)$$

- $\frac{d}{dU} \mathfrak{A}[\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)]$ has a unique, negative, non-degenerate minimum at $U = 0$. In particular, $\frac{d^3}{dU^3} \mathfrak{A}[\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)]|_{U=0} > 0$.
- There is a differentiable function $\mathcal{R} : \mathbb{R} \rightarrow \mathbb{R}$ such that $\|\mathcal{R}\|_{C^1(\Sigma_0)} \leq C$ and such that:

$$\frac{d}{dU} \mathfrak{A}[\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)] = -\dot{\delta}_* + \frac{1}{2} \mathfrak{b} U^2 + \mathcal{R}(U) U^3. \quad (90)$$

- For $j = 1, 2, 3$, there exists a continuous function $\Lambda_j : \mathbb{R} \rightarrow (0, \infty)$ such that:

$$\Lambda_j(U) \approx 1, \quad U \in [-U_1, U_2], \quad (91)$$

and such that for $M = 0, 1, 2$, we have:

$$\frac{d^M}{dU^M} \left(\frac{d}{dU} \mathfrak{A}[\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)] + \dot{\delta}_* \right) = \Lambda_{M+1}(U) U^{2-M}, \quad \text{if } |U| \leq 1, \quad (92a)$$

$$\frac{d}{dU} \mathfrak{A}[\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)] \geq -\mathfrak{p} \dot{\delta}_*, \quad \text{if } |U| \geq 1, \quad (92b)$$

where $\mathfrak{p} \in [0, 1)$ is the non-negative constant in (86), and in (91) and throughout, $A \approx B$ means that there is a $C \geq 1$ such that $\frac{1}{C}B \leq A \leq CB$.

- The constant $\dot{\delta}_*$ appearing in (81) satisfies:

$$\dot{\delta}_* = \max_{U \in [-U_1, U_2]} [G(U)]_- = \max_{U \in [-U_1, U_2]} \left[\frac{d}{dU} \mathfrak{A}[\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)] \right]_-, \quad (93)$$

where $[z]_- := \max\{-z, 0\}$ is the negative part of z .

3.7.3. The time of first blowup.

Definition 3.8 (time of first blowup). Given any initial data function $\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)$ constructed above, with $\dot{\delta}_* > 0$ as in (81), we define:

$$T_{\text{Shock}} := \frac{1}{\dot{\delta}_*}. \quad (94)$$

Note that (77) and (93) imply that T_{Shock} is the Minkowskian time at which μ first vanishes, which by (79) is the positive time of first blowup for $|\partial_1 \mathcal{R}_{(+)}^{\text{PS}}|$.

3.7.4. Sharp estimates for the inverse foliation density. In lemma 3.7, we provide various estimates for μ and its derivatives in a neighborhood of $[0, T_{\text{Shock}}] \times \{0\} \subset \mathbb{R}_t \times \mathbb{R}_U$. These estimates will be useful in the description of the boundary of the maximal development.

Lemma 3.7 (sharp estimates for μ and G). *Let $\mathcal{R}_{(+)}^{\text{PS}} = \mathcal{R}_{(+)}^{\text{PS}}(U)$ be compactly supported initial data constructed in and satisfying the assumptions stated in section 3.7.2. Recall that for the corresponding simple isentropic plane-symmetric solution, relative to the geometric coordinates (t, U) , $\mathcal{R}_{(+)}^{\text{PS}}$, c , and \mathbf{n} are functions of U . With this in mind, we define:*

$$\mathfrak{c} := \min_{U \in \mathbb{R}} \frac{c(U)}{\mathbf{n}(U)}, \quad \mathfrak{C} := \max_{U \in \mathbb{R}} \frac{c(U)}{\mathbf{n}(U)}. \quad (95)$$

Note that by (9) and (17c), we have:

$$0 < \mathfrak{c} \leq \mathfrak{C} < 1. \quad (96)$$

Then there exists⁷ a constant U_{shock} depending only on the seed profile φ and satisfying $0 < U_{\text{shock}} < 1$ such that the following hold.

Behavior of μ on $[-U_{\text{shock}}, U_{\text{shock}}]$. Let $T_{\text{Shock}} > 0$ be as defined in (94). Then the following estimates hold for $t \in [0, T_{\text{Shock}}]$:

$$-\mathfrak{c}^{-1} \dot{\delta}_* \leq \min_{\Sigma_t^{[-U_{\text{shock}}, U_{\text{shock}}]}} L\mu \leq \max_{\Sigma_t^{[-U_{\text{shock}}, U_{\text{shock}}]}} L\mu \leq -\mathfrak{C}^{-1} \dot{\delta}_*. \quad (97)$$

Moreover, for $(t, U) \in [0, T_{\text{Shock}}] \times [-U_{\text{shock}}, U_{\text{shock}}]$, the following estimates hold, where here and throughout, $A = \mathcal{O}_{\text{pos}}(B)$ means that there exist constants C_1, C_2 (depending on φ)⁸ such that $0 < C_1 \leq C_2$ and $C_1 B \leq A \leq C_2 B$, and $A = \mathcal{O}(B)$ means that there exists a $C > 0$ (depending on φ) such that $|A| \leq C|B|$:

$$\mu(t, U) = \mathcal{O}_{\text{pos}}(1) \frac{\mathfrak{b}}{2} U^2 + \mathcal{O}_{\text{pos}}(1) \dot{\delta}_* (T_{\text{Shock}} - t), \quad (98)$$

$$L\mu(t, U) = -\mathcal{O}_{\text{pos}}(1) \dot{\delta}_*, \quad (99)$$

$$\check{X}\mu(t, U) = \mathcal{O}_{\text{pos}}(1) \mathfrak{b} U + \mathcal{O}(1) (T_{\text{Shock}} - t), \quad (100)$$

$$L\check{X}\mu(t, U) = \mathcal{O}(1), \quad (101)$$

$$\check{X}\check{X}\mu(t, U) = \mathcal{O}_{\text{pos}}(1) \mathfrak{b} + \mathcal{O}(1) (T_{\text{Shock}} - t). \quad (102)$$

In addition, we have:

$$\{\check{X}\mu = 0\} \cap \Sigma_{T_{\text{Shock}}}^{[-U_{\text{shock}}, U_{\text{shock}}]} = \{(T_{\text{Shock}}, 0)\} \subset \Sigma_{T_{\text{Shock}}}^{[-\frac{1}{4}U_{\text{shock}}, \frac{1}{4}U_{\text{shock}}]}, \quad (103)$$

$$\min_{\Sigma_{T_{\text{Shock}}}^{[-U_{\text{shock}}, U_{\text{shock}}]} \setminus \Sigma_{T_{\text{Shock}}}^{[-\frac{1}{4}U_{\text{shock}}, \frac{1}{4}U_{\text{shock}}]}} |\check{X}\mu| \geq \frac{\mathfrak{b} U_{\text{shock}}}{8}, \quad (104)$$

⁷ With minor additional effort, one could explicitly compute such a constant U_{shock} in terms of the constants associated to φ such as $\dot{\delta}_*$, \mathfrak{b} , etc.

⁸ For example, the implicit constants present on RHS (99) can be taken to be $C_1 = \mathfrak{C}^{-1}$ and $C_2 = \mathfrak{c}^{-1}$; see (97).

$$\frac{1}{2} \mathfrak{C}^{-1} T_{\text{Shock}} \mathfrak{b} \leq \min_{\Sigma_{T_{\text{Shock}}}^{[-U_{\oplus}, U_{\oplus}]}} \check{X} \check{X} \mu, \leq \max_{\Sigma_{T_{\text{Shock}}}^{[-U_{\oplus}, U_{\oplus}]}} \check{X} \check{X} \mu, \leq 2 \mathfrak{c}^{-1} T_{\text{Shock}} \mathfrak{b}. \quad (105)$$

μ is uniformly positive away from the interesting region. With $\mathfrak{p} \in [0, 1)$ denoting the non-negative constant on RHS (92b), we have the following estimate:

$$\min_{\{(t, U) \mid t \in [0, T_{\text{Shock}}], |u| \geq U_{\oplus}\}} \mu(t, U) \geq \mathcal{O}_{\text{pos}}(1)(1 - \mathfrak{p}). \quad (106)$$

Properties of G in the interesting region. Let G be the source term from (71), viewed as a function of U , and let $G' := \frac{d}{dU} G$. Then for $U \in [-U_{\oplus}, U_{\oplus}]$, the following estimates hold:

$$G(U) = -\mathring{\delta}_* + \mathcal{O}_{\text{pos}}(1)U^2 = -\mathcal{O}_{\text{pos}}(1)\mathring{\delta}_*, \quad (107a)$$

$$\frac{1}{G(U)} = -\frac{1}{\mathring{\delta}_*} - \mathcal{O}_{\text{pos}}(1)U^2, \quad (107b)$$

$$G'(U) = \mathcal{O}_{\text{pos}}(1)U. \quad (107c)$$

In particular:

- $G(U) < 0$ for $U \in [-U_{\oplus}, U_{\oplus}]$
- $G'(0) = 0$
- $G'(U) > 0$ for $U \in [0, U_{\oplus}]$
- $G'(U) < 0$ for $U \in [-U_{\oplus}, 0]$.

Proof. All of the results except for those concerning G are straightforward consequences of corollary 3.6, the properties of the initial data $\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)$ from section 3.7.2, and Taylor expansions; we omit the details.

To prove the results for G , we first use (69) and (74a) to deduce that $G(U) = \frac{d}{dU} \mathfrak{A}[\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)]$. Using this identity and (90), we conclude (107c), provided $U_{\oplus} > 0$ is taken to be sufficiently small. \square

3.8. The crease, singular curve, and Cauchy horizon in 1D simple isentropic plane-symmetry

In this section, we provide a description of a localized portion of the boundary of the **h**-MGHD. We emphasize the following key point, which follows from (69), (77), and (79):

For the initial data from section 3.7.2, the corresponding fluid solution $\mathcal{R}_{(+)}^{\text{PS}}$ and the inverse foliation density μ are explicit smooth functions on all⁹ of \mathbb{R}^2 in the $((t, U))$ differential structure, even though the Minkowskian partial derivative $\partial_1 \mathcal{R}_{(+)}^{\text{PS}}$ blows up in finite time.

⁹ In multi-dimensions, the best known estimates allow for the possibility of high order energy-blowup; see section 6.9.3. However, even in multi-dimensions, the solution remains smooth at the low-to-mid order derivative levels.

Therefore, when we speak of the **h**-MGHD, we are referring to the standard differential structure corresponding to the (t, x^1) coordinates. Nonetheless, *we prefer to carry out most our analysis of the **h**-MGHD in (t, U) coordinates*, thanks to the explicit structure of the equations and solutions in geometric coordinates revealed in corollary 3.6. Hence, to reach conclusions about the behavior of solution in the (t, x^1) -differential structure, we have to control the change of variables map $\Upsilon(t, U) := (t, x^1)$; indeed, in theorem 3.1, we reveal various diffeomorphism and homeomorphism properties of Υ .

For the solutions and localized region under study, when described in geometric coordinates, the boundary of the **h**-MGHD consists of two pieces, depicted in figure 2:

- (1) A singular curve portion $\mathcal{B}_{[-U_\oplus, 0]}$, emanating from the first shock point—called the *crease* and denoted by $\partial_- \mathcal{B}$ —such that μ vanishes on $\mathcal{B}_{[-U_\oplus, 0]}$.
- (2) A Cauchy horizon portion, denoted by $\mathcal{C}_{[0, U_\oplus]}$, which is a portion of the integral curve of \underline{L} that emanates from $\partial_- \mathcal{B}$, and along which μ is positive, except at $\partial_- \mathcal{B}$.

In particular, the image $\Upsilon(\mathcal{B}_{[-U_\oplus, 0]})$, depicted in figure 3, is a boundary of the **h**-MGHD in the differential structure of the Minkowski-rectangular coordinates (t, x^1) in the sense that $|\partial_1 \mathcal{R}_{(+)}^{\text{PS}}|$ blows up when μ vanishes (as it does along $\Upsilon(\mathcal{B}_{[-U_\oplus, 0]})$). In contrast, even though $\mathcal{R}_{(+)}^{\text{PS}}$ remains smooth all the way up to $\Upsilon(\mathcal{C}_{[0, U_\oplus]} \setminus \partial_- \mathcal{B})$, $\Upsilon(\mathcal{C}_{[0, U_\oplus]})$ is a boundary of the **h**-MGHD because the crease $\Upsilon(\partial_- \mathcal{B})$ —a point at which $|\partial_1 \mathcal{R}_{(+)}^{\text{PS}}|$ blows up—lies in the **h**-causal past of any point in $\Upsilon(\mathcal{C}_{[0, U_\oplus]})$. We highlight again that the singular behavior is only visible relative to the differential structure of Minkowski space in (t, x^1) coordinates. The discrepancy between the behavior of the solution in the two coordinate systems is tied to the breakdown in the diffeomorphism property of Υ when μ vanishes; see (67).

3.8.1. Singular curve, crease, Cauchy horizon, and developments. We now define some subsets of geometric coordinate space $\mathbb{R}_t \times \mathbb{R}_U$ that, in theorem 3.1, will play an important role in describing the shape of the **h**-MGHD. Many of our definitions are localized to the subset $\{|U| \leq U_\oplus\}$, where $U_\oplus > 0$ is the constant from lemma 3.7. This is mainly for convenience, since lemma 3.7 provides sharp control over the solution when $|U| \leq U_\oplus$.

Definition 3.9 (truncated portions of the singular curve and the crease). For the solution launched by the compactly supported initial data $\mathcal{R}_{(+)}^{\text{PS}} = \mathcal{R}_{(+)}^{\text{PS}}(U)$ constructed in section 3.7.2, let G be the corresponding source term from (71), viewed as a function of U . We define:

$$t_{\text{Sing}}(U) := -\frac{1}{G(U)}. \quad (108)$$

Let $I \subset [-U_\oplus, U_\oplus]$ be a subset of U -values, where $U_\oplus > 0$ is the constant from lemma 3.7. We define the *truncated singular curve* to be the following subset of geometric coordinate space $\mathbb{R}_t \times \mathbb{R}_U$:

$$\begin{aligned} \mathcal{B}_I &:= \{(t, U) \in [0, \infty) \times I \mid \mu(t, U) = 0\} \cap \left\{ (t, U) \in [0, \infty) \times I \mid \check{X}\mu(t, U) \leq 0 \right\} \\ &= \{(t_{\text{Sing}}(U), U) \mid U \in I\} \cap \left\{ (t, U) \in [0, \infty) \times I \mid \check{X}\mu(t, U) \leq 0 \right\}, \end{aligned} \quad (109)$$

where \check{X} is the vectorfield defined in (59), and the second identity in (109) follows from (77), (96), and (108).

We define the *crease*, which we denote by $\partial_- \mathcal{B}$, to be the following subset of geometric coordinate space $\mathbb{R}_t \times \mathbb{R}_U$ (theorem 3.1 implies that $\partial_- \mathcal{B}$ is a single point for the solutions studied there):

$$\begin{aligned} \partial_- \mathcal{B} := & \left\{ (t, U) \in [0, \infty) \times [-U_{\oplus}, U_{\oplus}] \mid \mu(t, U) = 0 \right\} \\ & \cap \left\{ (t, U) \in [0, \infty) \times [-U_{\oplus}, U_{\oplus}] \mid \check{X}\mu(t, U) = 0 \right\}. \end{aligned} \quad (110)$$

From (77), (78), (96), (107b), (107c), and (108), it follows that if $(t', U') \in \left\{ (t, U) \in [0, \infty) \times [-U_{\oplus}, U_{\oplus}] \mid \mu(t, U) = 0 \right\}$, then $\check{X}\mu(t', U') = \mathcal{O}_{\text{pos}}(1)U'$ and thus $\check{X}\mu(t', U') \leq 0 \iff U' \in [-U_{\oplus}, 0]$. We therefore have the following alternate characterization of \mathcal{B}_I :

$$\mathcal{B}_I = \left\{ (\mathbf{t}_{\text{Sing}}(U), U) \mid U \in I \cap [-U_{\oplus}, 0] \right\} \cap [0, \infty) \times \mathbb{R}. \quad (111)$$

We make the following remarks.

- One might wonder why in the singular curve definition (109), there is a restriction to regions in which $\check{X}\mu \leq 0$. This is because the portion of $\{\mu = 0\}$ on which we formally have $\check{X}\mu > 0$ never actually dynamically develops in the **h**-MGHD; that ‘fictitious portion’ is cut off by the Cauchy horizon before it has a chance to dynamically form; see figure 2 and remark 3.2.
- Our terminology ‘crease’ is motivated by picture of the **h**-MGHD in multi-dimensions, where the analog of $\partial_- \mathcal{B}$ is a co-dimension two hypersurface given by the intersection of the singular boundary and the Cauchy horizon; see figure 6.
- From (79), (109), and (111), we deduce the following key result: *in regions where we can justify the use of geometric coordinates for reaching conclusions about the behavior of the solution in (t, x^1) space, \mathcal{B}_I is the subset of points $(t, U) \in [0, \infty) \times (I \cap [-U_{\oplus}, 0])$ such that $|\partial_1 \mathcal{R}_{(+)}^{\text{PS}}(t, U)|$ blows up at $\Upsilon(t, U)$* . In theorem 3.1, we will indeed justify the use of the geometric coordinates when $I = [-U_{\oplus}, 0]$.
- (77), (93), and (96) imply that T_{Shock} is the smallest value of t along $\mathcal{B}_{[-U_{\oplus}, 0]}$.
- To further flesh out the previous point, we note that for the solutions we study in theorem 3.1, the smallest value of t on $\mathcal{B}_{[-U_{\oplus}, 0]}$ will occur precisely at $U = 0$, and we will have $\mathbf{t}_{\text{Sing}}(0) = T_{\text{Shock}}$.

Definition 3.10 (truncated Cauchy horizon). Assume that $\partial_- \mathcal{B} = (T_{\text{Shock}}, 0)$, where T_{Shock} is defined by (94) (theorem 3.1 shows that this assumption is satisfied for the solutions under study here). Let \check{L} be the vectorfield defined in (59). Let $\check{\gamma}$ be the integral curve of $\frac{1}{2}\check{L}$ in geometric coordinate space $\mathbb{R}_t \times \mathbb{R}_U$ parametrized by U (note that $\frac{1}{2}\check{L}U = 1$ by (64)), with the initial condition $\check{\gamma}(0) = \partial_- \mathcal{B}$. That is, $\check{\gamma}$ solves the following initial value problem:

$$\frac{d}{dU} \check{\gamma}(U) = \frac{1}{2} \check{L} \circ \check{\gamma}(U), \quad (112a)$$

$$\check{\gamma}(0) = (T_{\text{Shock}}, 0). \quad (112b)$$

We use the notation \mathbf{t}_{CH} to denote the t -component of the curve $\check{\gamma}$. That is, by (64), we have:

$$\check{\gamma}(U) = (\mathbf{t}_{\text{CH}}(U), U), \quad (113)$$

where t_{CH} solves the following initial value problem:

$$\frac{d}{dU} t_{\text{CH}}(U) = \frac{1}{2} \mu(t_{\text{CH}}(U), U), \quad t_{\text{CH}}(0) = T_{\text{Shock}}. \quad (114)$$

Then for an interval $I \subset [0, \infty)$ of non-negative¹⁰ U -values, we define the truncated Cauchy horizon to be:

$$\underline{\mathcal{C}}_I := \underline{\gamma}(I). \quad (115)$$

We now define \mathcal{M}_* , which is a subset of $\mathbb{R}_t \times \mathbb{R}_U$ that is tailored to the shape of the **h**-MGHD.

Definition 3.11 (classical development regions). Let $\mathcal{B}_{[-U_{\oplus}, 0]}$ and $\underline{\mathcal{C}}_{[0, U_{\oplus}]}$ denote the truncated singular curve and Cauchy horizon from definitions 3.9 and 3.10, where $0 < U_{\oplus} < 1$ is the constant from lemma 3.7. Let t_{Sing} and t_{CH} be the functions of U from (108) and (113). We define the following subsets of geometric coordinate space $\mathbb{R}_t \times \mathbb{R}_U$:

$$\mathcal{M}_{\text{Sing}} := \{(t, U) \in \mathbb{R}^2 \mid 0 \leq t < t_{\text{Sing}}(U), U \in [-U_{\oplus}, 0]\}, \quad (116)$$

$$\mathcal{M}_{\text{Reg}} := \{(t, U) \in \mathbb{R}^2 \mid 0 \leq t \leq t_{\text{CH}}(U), U \in (0, U_{\oplus}]\}, \quad (117)$$

$$\mathcal{M}_* := \mathcal{M}_{\text{Sing}} \cup \mathcal{M}_{\text{Reg}}. \quad (118)$$

The set \mathcal{M}_* defined by (118) is the region of geometric coordinate space lying between $\Sigma_0^{[-U_{\oplus}, U_{\oplus}]}$ and $\underline{\mathcal{C}}_{[0, U_{\oplus}]} \cup \mathcal{B}_{[-U_{\oplus}, 0]}$, where the boundary portions $\underline{\mathcal{C}}_{[0, U_{\oplus}]}$, $\mathcal{P}_{-U_{\oplus}}^{[0, t_{\text{Sing}}(-U_{\oplus})]}$, and $\mathcal{P}_{U_{\oplus}}^{[0, t_{\text{CH}}(U_{\oplus})]}$ are included, but $\mathcal{B}_{[-U_{\oplus}, 0]}$ is not; see figure 2. Theorem 3.1 shows that $\Upsilon(\mathcal{M}_*) \subset \mathbb{R}_t \times \mathbb{R}_{x^1}$ is a subset of Minkowski-rectangular coordinate space where we can *justify* the use of (t, U) coordinates to reach conclusions about the behavior of the solution in (t, x^1) coordinates¹¹.

3.8.2. The main theorem. We now prove theorem 3.1, which is the main result of section 3. On the region \mathcal{M}_* , the theorem provides a sharp description of the solution in the differential structure of the geometric coordinates (t, U) as well as the behavior of the solution on $\Upsilon(\mathcal{M}_*)$ in differential structure of the Minkowski-rectangular coordinates (t, x^1) , where we recall that $\Upsilon(t, U) = (t, x^1)$ is the change of variables map. In particular, the theorem reveals various diffeomorphism and homeomorphism properties of Υ and exhibits the blowup of $|\partial_1 \mathcal{R}_{(+)}^{\text{PS}}|$ as $\Upsilon(\mathcal{B}_{[-U_{\oplus}, 0]})$ is approached by points in $\Upsilon(\mathcal{M}_*)$. In figure 2, we have depicted the region in geometric coordinate space $\mathbb{R} \times \mathbb{R}_U$. In figure 3, we have depicted the region in Minkowski-rectangular coordinate space $\mathbb{R} \times \mathbb{R}_{x^1}$, i.e. the image under Υ of the region from geometric coordinate space.

¹⁰ In theorem 3.1, the Cauchy horizon will correspond to non-negative values of U .

¹¹ With some additional effort, given the results of theorem 3.1, we could extend the region of spacetime on which $\Upsilon(t, U) = (t, x^1)$ is a homeomorphism to include the infinite region $\mathcal{M}_*^{\text{Large}} := \mathcal{M}_* \cup ([0, T_{\text{Shock}}] \times (-\infty, -U_1) \cup (U_2, \infty))$. We refrain from giving the proof because it would require some technical, but uninteresting, adjustments of the arguments presented in the proof of the theorem. The details are uninteresting in the sense that μ is uniformly positive on $\mathcal{M}_*^{\text{Large}} \setminus \mathcal{M}_*$ (see (106)) and thus no singularity occurs in that region.

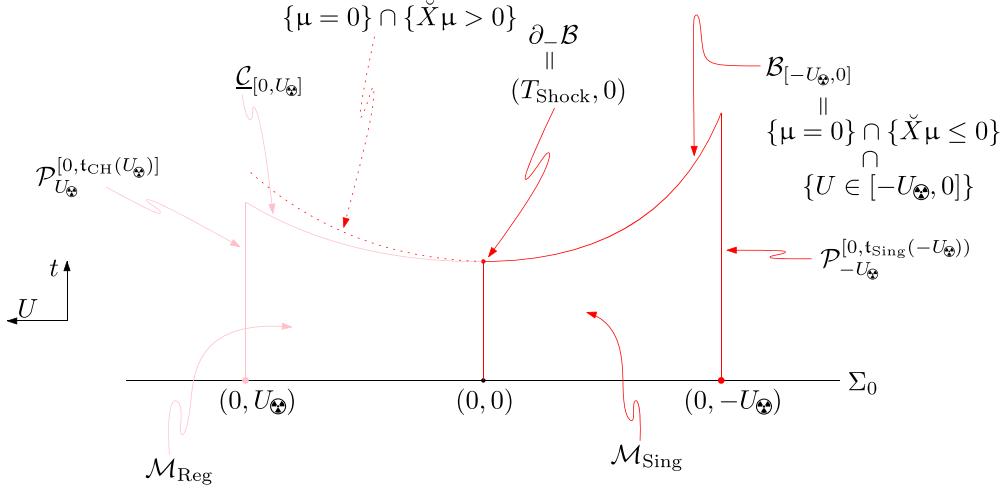


Figure 2. \mathcal{M}_* in geometric coordinate space, with U increasing from right to left.

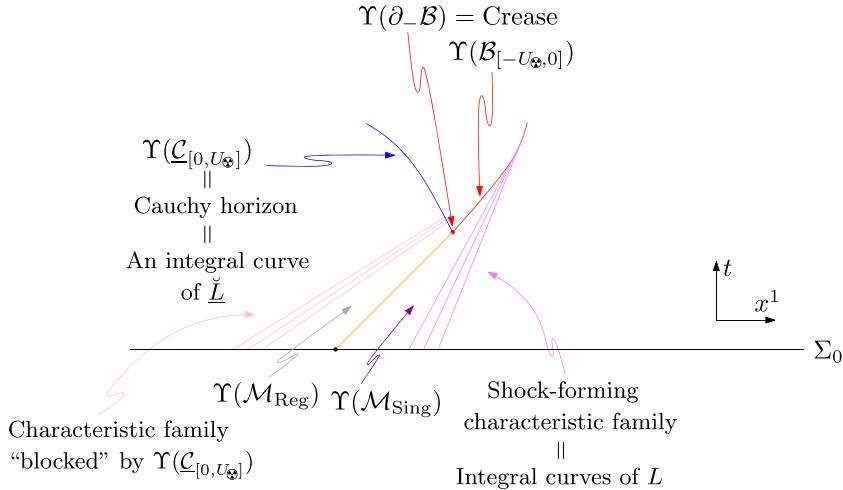


Figure 3. $\Upsilon(\mathcal{M}_*)$ in Minkowski-rectangular coordinate space.

Remark 3.2 (only a subset of $\{\mu = 0\}$ is relevant for the h-MGHD). The dotted curve in figure 2 labeled ‘ $\{\mu = 0\} \cap \{\check{X}\mu > 0\}$ ’ is not part of the h-MGHD or its boundary. Formally, $\{\mu = 0\} \cap \{\check{X}\mu > 0\}$ is the portion of the level set $\{\mu = 0\}$ that never has a chance to dynamically emerge because it is cut off by the Cauchy horizon portion $\underline{C}_{(0, U_*)}$, which lies below it (as the proof of theorem 3.1 shows). For this reason, we do not display $\Upsilon(\{\mu = 0\} \cap \{\check{X}\mu > 0\})$ in figure 3. One can use Taylor expansions to show that formally, $\Upsilon(\{\mu = 0\} \cap \{\check{X}\mu > 0\})$ would be a curve in $\Upsilon(\mathcal{M}_{Sing})$ that lies below $\Upsilon(\mathcal{B}_{[-U_*, 0]})$ and thus formally, Υ would not even be injective if its domain extended up to $\{\mu = 0\} \cap \{\check{X}\mu > 0\}$.

Theorem 3.1 (a large portion of the h-MGHD, including a localized portion of its boundary). Let $\tilde{\mathcal{R}}_{(+)}^{\text{PS}}$ be initial data satisfying the assumptions of section 3.7.2, and let $\mathcal{R}_{(+)}^{\text{PS}}$

be the corresponding solution of the quasilinear initial value problem (53) (with $\mathcal{R}_{(-)}^{\text{PS}} \equiv 0$). Under the assumptions and conclusions of lemma 3.7, perhaps shrinking the constant $U_{\boxtimes} > 0$ if necessary, the following results hold, where the notation $\mathcal{O}_{\text{pos}}(1)$ is defined in the statement of lemma 3.7.

Classical existence with respect to the geometric coordinates. With respect to the geometric coordinates (t, U) , the following results hold.

- $\mathcal{R}_{(+)}^{\text{PS}}$ is a smooth function of U alone, i.e. $\mathcal{R}_{(+)}^{\text{PS}}(t, U) = \mathcal{R}_{(+)}^{\text{PS}}(U)$.
- $\mathcal{R}_{(+)}^{\text{PS}}$, the null vectorfields L and \underline{L} , the vectorfield \dot{X} , and the inverse foliation density μ exist classically on $\mathbb{R}_t \times \mathbb{R}_U$.
- On $\{(t, U) \in \mathbb{R}^2 \mid U \notin [-U_1, U_2]\}$, $\mathcal{R}_{(+)}^{\text{PS}}$ vanishes, $L^1 := Lx^1 = \bar{c}$, $\underline{L}^1 := \underline{L}x^1 = -\bar{c}$, and $\mu = 1/\bar{c}$, where the constant $\bar{c} > 0$ denotes the speed of sound when the enthalpy is constantly \bar{H} , where $\bar{H} > 0$ is the constant fixed in (3).
- The following estimate holds:

$$L\mu(t, U) = -\mathcal{O}_{\text{pos}}(1) \dot{\delta}_*, \quad \text{for } (t, U) \in \mathbb{R}^2 \times [-U_{\boxtimes}, U_{\boxtimes}]. \quad (119)$$

In particular, with $\overline{\mathcal{M}_*}$ denoting the closure of the region \mathcal{M}_* defined in (118), since definition 3.11 implies that $\overline{\mathcal{M}_*} \subset \mathbb{R}^2 \times [-U_{\boxtimes}, U_{\boxtimes}]$, it follows that the estimate (119) holds on $\overline{\mathcal{M}_*}$.

Description of the h-MGHD in geometric coordinates.

- Let $\mathcal{B}_{[-U_{\boxtimes}, 0]}$ be the singular curve portion from (111), and let \mathcal{M}_* be the subset of geometric coordinate space defined in (118). Then the following disjoint union holds, where $\overline{\mathcal{M}_*}$ is the closure of \mathcal{M}_* :

$$\overline{\mathcal{M}_*} = \mathcal{M}_* \sqcup \mathcal{B}_{[-U_{\boxtimes}, 0]}. \quad (120)$$

- The *crease*, which is defined in (110), is a single point:

$$\partial_- \mathcal{B} = \{(T_{\text{Shock}}, 0)\}. \quad (121)$$

- The singular curve function $t_{\text{Sing}} = t_{\text{Sing}}(U)$ from (108) and (111) satisfies the following estimates:

$$\frac{d}{dU} t_{\text{Sing}}(U) = \mathcal{O}_{\text{pos}}(1) U, \quad U \in [-U_{\boxtimes}, 0], \quad (122a)$$

$$t_{\text{Sing}}(U) = T_{\text{Shock}} + \mathcal{O}_{\text{pos}}(1) U^2, \quad U \in [-U_{\boxtimes}, 0]. \quad (122b)$$

- The rectangular spatial coordinate x^1 satisfies the following estimate on $\mathcal{B}_{[-U_{\boxtimes}, 0]}$:

$$\frac{d}{dU} x^1(t_{\text{CH}}(U), U) = \mathcal{O}_{\text{pos}}(1) U, \quad U \in [-U_{\boxtimes}, 0]. \quad (123)$$

- The integral curve initial value problem for $\dot{\gamma}(U)$ given by (112a)–(112b) has a unique solution on the interval $U \in [0, U_{\boxtimes}]$. Hence, the truncated Cauchy horizon $\mathcal{C}_{[0, U_{\boxtimes}]}$, defined in (115), can be parameterized as follows:

$$\mathcal{C}_{[0, U_{\boxtimes}]} = \{(t_{\text{CH}}(U), U) \mid U \in [0, U_{\boxtimes}]\}, \quad (124)$$

as in (115). Moreover, $t_{\text{CH}}(U)$ satisfies the following estimates:

$$\frac{d}{dU} t_{\text{CH}}(U) = \mathcal{O}_{\text{pos}}(1) U^2, \quad U \in [0, U_{\oplus}], \quad (125a)$$

$$t_{\text{CH}}(U) = T_{\text{Shock}} + \mathcal{O}_{\text{pos}}(1) U^3, \quad U \in [0, U_{\oplus}]. \quad (125b)$$

- The inverse foliation density μ , defined in (57), satisfies the following estimate along the Cauchy horizon portion $\underline{\mathcal{C}}_{[0, U_{\oplus}]}$:

$$\mu(t_{\text{CH}}(U), U) = \mathcal{O}_{\text{pos}}(1) U^2, \quad U \in [0, U_{\oplus}]. \quad (126)$$

- The rectangular spatial coordinate x^1 satisfies the following estimate along the Cauchy horizon portion $\underline{\mathcal{C}}_{[0, U_{\oplus}]}$:

$$\frac{d}{dU} x^1(t_{\text{CH}}(U), U) = -\mathcal{O}_{\text{pos}}(1) U^2, \quad U \in [0, U_{\oplus}]. \quad (127)$$

μ is positive on \mathcal{M}_* and on $\overline{\mathcal{M}_*} \setminus \mathcal{B}_{[-U_{\oplus}, 0]}$, but vanishes on $\mathcal{B}_{[-U_{\oplus}, 0]}$. In particular, (126) implies that μ is positive on $\underline{\mathcal{C}}_{(0, U_{\oplus})}$ but vanishes on the crease $\partial_- \mathcal{B} \subset \underline{\mathcal{C}}_{[0, U_{\oplus}]}$, which corresponds to $U = 0$ in (126).

Description of the singularity formation and the h-MGHD in Minkowski-rectangular coordinates.

- Let $\Upsilon(t, U) := (t, x^1)$ be the change of variables map from geometric coordinates to Minkowski-rectangular coordinates. Then Υ is a diffeomorphism from \mathcal{M}_* onto its image. Moreover, Υ is a **homeomorphism** from $\overline{\mathcal{M}_*}$ onto its image.
- The solution $\mathcal{R}_{(+)}^{\text{PS}}$ exists classically relative to the Minkowski-rectangular coordinates (t, x^1) on the subset $\Upsilon(\mathcal{M}_*)$ of $\mathbb{R}_t \times \mathbb{R}_{x^1}$. In particular, since the Cauchy horizon portion $\underline{\mathcal{C}}_{(0, U_{\oplus})}$ is contained in \mathcal{M}_* , $\mathcal{R}_{(+)}^{\text{PS}}$ exists classically relative to the Minkowski-rectangular coordinates (t, x^1) on $\Upsilon(\underline{\mathcal{C}}_{(0, U_{\oplus})})$.
- The following estimate holds on $\Upsilon(\mathcal{M}_*)$:

$$\left| \left[\partial_1 \mathcal{R}_{(+)}^{\text{PS}} \right] \circ \Upsilon^{-1}(t, x^1) \right| = \frac{1}{\mu} \left| \frac{n}{c} \check{X} \mathcal{R}_{(+)}^{\text{PS}} \right| \circ \Upsilon^{-1}(t, x^1) = \frac{1}{\mu} \mathcal{O}_{\text{pos}}(1), \quad (t, x^1) \in \Upsilon(\mathcal{M}_*). \quad (128)$$

In particular, since μ vanishes along $\Upsilon(\mathcal{B}_{[0, U_{\oplus}]})$, given any point $(t_*, x_*^1) \in \Upsilon(\mathcal{B}_{[0, U_{\oplus}]})$ and any sequence of points $\{(t_m, x_m^1)\}_{m \in \mathbb{N}} \subset \Upsilon(\mathcal{M}_*)$ with $\lim_{m \rightarrow \infty} (t_m, x_m^1) = (t_*, x_*^1)$, we have that $|\partial_1 \mathcal{R}_{(+)}^{\text{PS}}| \circ \Upsilon^{-1}((t_m, x_m^1)) \rightarrow \infty$ as $m \rightarrow \infty$.

- Recall that the vectorfield L is defined relative to the (t, x^1) -coordinates in (44a) (see also (51a)). With respect to the (t, x^1) -coordinates, the solution $\mathcal{R}_{(+)}^{\text{PS}}$ and all of its L -derivatives of any order, i.e. $L^M \mathcal{R}_{(+)}^{\text{PS}}$ for any M , are continuous on $\Upsilon(\overline{\mathcal{M}_*})$, including the singular curve portion $\Upsilon(\mathcal{B}_{[0, U_{\oplus}]})$. Moreover, on $\Upsilon(\overline{\mathcal{M}_*})$, relative to the (t, x^1) -coordinates, the 1D acoustical metric $\mathbf{h}_{\alpha\beta}$ from (40a) is a continuous, non-degenerate, 2×2 matrix of signature $(-, +)$.

The causal structure of the Cauchy horizon.

- $\underline{\mathcal{C}}_{[0, U_{\Theta}]}$ is a \mathbf{h} -null curve portion, i.e. its tangent vector is null with respect to \mathbf{h} .

The causal structure of the singular curve.

- Relative to the geometric coordinates (t, U) , the vectorfield Q defined by:

$$Q := L - \frac{L\left(\frac{c}{n}\mu\right)}{\dot{X}\left(\frac{c}{n}\mu\right)} \dot{X} \quad (129)$$

is well-defined on the singular curve portion $\mathcal{B}_{[-U_{\Theta}, 0]}$, and Q is tangent to $\mathcal{B}_{[-U_{\Theta}, 0]}$.

- On $\mathcal{B}_{[-U_{\Theta}, 0]}$, we have $Qt > 0$ and $QU < 0$. In particular, Q is future-directed and transversal to the characteristics \mathcal{P}_U .
- For every $q \in \mathcal{B}_{[-U_{\Theta}, 0]}$, the pushforward vectorfield $[d\Upsilon(q)] \cdot Q(q)$, which is tangent to the curve $\Upsilon(\mathcal{B}_{[-U_{\Theta}, 0]}) \subset \mathbb{R}_t \times \mathbb{R}_{x^1}$, is equal to $L|_{\Upsilon(q)} = [L^\kappa \partial_\kappa]|_{\Upsilon(q)}$.
- For every $q \in \mathcal{B}_{[-U_{\Theta}, 0]}$, $L|_{\Upsilon(q)}$ is \mathbf{h} -orthogonal to the tangent space of $\Upsilon(\mathcal{B}_{[-U_{\Theta}, 0]})$.

Since there exist integral curves of the \mathbf{h} -null vectorfield $L^\kappa \partial_\kappa$ that foliate and are \mathbf{h} -orthogonal to $\Upsilon(\mathcal{B}_{[-U_{\Theta}, 0]})$, it follows that $\Upsilon(\mathcal{B}_{[-U_{\Theta}, 0]})$ is a \mathbf{h} -null curve in the Minkowskian coordinate differential structure.

Before proving the theorem, we make some remarks.

- All the important results of the theorem are stable under perturbations of the initial data of $\mathcal{R}_{(+)}^{\text{PS}}$. That is, even though the theorem was stated only for initial data that satisfy the assumptions of section 3.7.2, the proof shows that if one perturbs any of those data by an arbitrary (small) C^4 function, then the corresponding perturbed simple isentropic plane-wave solution has an \mathbf{h} -MGHD/exhibits singularity formation that is quantitatively and qualitatively close to the un-perturbed solution.
- In a similar vein as the previous point, with only modest additional effort, the results of theorem 3.1 could be extended to a large, open set of (non-simple) isentropic plane-symmetric solutions in which $\mathcal{R}_{(-)}^{\text{PS}}$ is sufficiently small compared to $\mathcal{R}_{(+)}^{\text{PS}}$ (or, alternatively, in which $\mathcal{R}_{(+)}^{\text{PS}}$ is sufficiently small compared to $\mathcal{R}_{(-)}^{\text{PS}}$).
- The emergence of the Cauchy horizon from the crease is fundamentally tied to the fact that in 1D, the isentropic compressible Euler equations have two characteristic directions, L and \underline{L} , as is evident from equations (49a) and (49b). In particular, even within the class of simple isentropic plane-symmetric solutions in which $\mathcal{R}_{(-)}^{\text{PS}}$ (which is transported by \underline{L}) vanishes, when investigating the structure of the \mathbf{h} -MGHD, one cannot ‘ignore’ the \underline{L} direction. This is in contrast to Burgers’ equation, in which there is only one characteristic direction. Relatedly, for Burgers’ equation, there does not exist a notion of ‘globally hyperbolic’ that enjoys the same fundamental significance as the notion of ‘globally hyperbolic’ in the context of Lorentzian geometry, relativistic Euler flow, or wave equations.
- From the point of view of determinism, the Cauchy horizon is an ally in that it ‘blocks’ the non-uniqueness of classical solutions (in \mathbf{h} -globally hyperbolic regions). In particular, in figure 3, one can see that the Cauchy horizon prevents the characteristics in the left of the figure from entering into the region on the right, where the fluid’s gradient blows up. That is, the Cauchy horizon prevents crossing of the characteristics $\Upsilon(\mathcal{M}_*)$, which in turn prevents multi-valuedness/non-uniqueness of classical solutions in this region. This should be contrasted against Burgers’ equation, which exhibits non-uniqueness of classical solutions due

to multi-valuedness stemming from the crossing of characteristics emanating from widely separated regions in the initial data hypersurface.

- In figure 3, along $\Upsilon(\mathcal{B}_{[-U_\Theta, 0]})$, one sees the interesting phenomenon of *non-uniqueness* of integral curves of L . In particular, L is tangent to $\Upsilon(\mathcal{B}_{[-U_\Theta, 0]})$ and the characteristics, which are the straight lines in the figure that intersect $\Upsilon(\mathcal{B}_{[-U_\Theta, 0]})$ tangentially. The non-uniqueness is tied to the blowup of $|\partial_1 L^1|$ along $\Upsilon(\mathcal{B}_{[-U_\Theta, 0]})$, which in particular shows that L^α does not enjoy the Lipschitz regularity that lies behind standard ODE uniqueness theorems for integral curves.
- One can show that in 1D, the acoustical metric takes the following form in geometric coordinates: $\mathbf{h} = -\mu dt \otimes dU - \mu dU \otimes dt + \mu^2 dU \otimes dU$. In particular, in geometric coordinates, \mathbf{h} *vanishes* at points such that $\mu = 0$. Hence, \mathbf{h} is not capable of ‘measuring’ causality along the singular curve portion $\mathcal{B}_{[-U_\Theta, 0]}$, along which μ vanishes. We note, however, that in more than one space dimension, even when $\mu = 0$, there are some remaining non-degenerate components of \mathbf{h} in geometric coordinates. Relatedly, in our 3D paper [1] on the non-relativistic compressible Euler equations, we prove that the $\partial_- \mathcal{B}$ is a co-dimension two \mathbf{h} -spacelike submanifold of geometric coordinate space.

Proof of theorem 3.1. Throughout the proof, we will silently shrink the size of U_Θ as necessary; this will in particular guarantee that some important terms arising in various Taylor expansions have a sign.

Classical existence and properties of the solution relative to the geometric coordinates:

The fact that $\mathcal{R}_{(+)}^{\text{PS}}(t, U) = \mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U)$ is smooth on all of $\mathbb{R}_t \times \mathbb{R}_U$ was proved in (69). In particular, in geometric coordinates, $\mathcal{R}_{(+)}^{\text{PS}}$ smoothly extends to the closure $\overline{\mathcal{M}_*}$. The fact that $\mathcal{R}_{(+)}^{\text{PS}}$ vanishes on the complement of $\{(t, U) \in \mathbb{R}^2 \mid U \in [-U_1, U_2]\}$ then follows from our assumptions on the initial data, which in particular imply that $\mathring{\mathcal{R}}_{(+)}^{\text{PS}}$ is supported in $[-U_1, U_2]$.

Next, using the identity (77), the estimate (96), and the fact that $\mathcal{R}_{(+)}^{\text{PS}}$ is supported in $\{(t, U) \in \mathbb{R}^2 \mid U \in [-U_1, U_2]\}$, we conclude the asserted properties of μ in geometric coordinates, except for the positivity of μ on \mathcal{M}_* , which we prove below. The asserted properties of L , \underline{L} , Lx^1 , and $\underline{L}x^1$ in geometric coordinates then follow from (51a)–(51b), (62), (63), and the fact that $\mathcal{R}_{(+)}^{\text{PS}}$ is supported in $\{(t, U) \in \mathbb{R}^2 \mid U \in [-U_1, U_2]\}$.

Proof of the properties of and relative to the geometric coordinates: The fact that the crease is the single point $(T_{\text{Shock}}, 0)$ is a consequence of definition (110), (98), and (103). We have therefore proved (121).

From (77), (96), (108), (109), (111), and definition (116), we see that $\mu > 0$ on $\mathcal{M}_{\text{Sing}}$ while μ vanishes on $\mathcal{B}_{[-U_\Theta, 0]}$.

The estimate (119) follows from the first equality in (76) and the estimate (107a).

The estimates (122a)–(122b) follow from definitions (94) and (108) and the estimates (107b)–(107c).

Proof of the properties of relative to the geometric coordinates: We first prove that the integral curve initial value problem for $\dot{\gamma}(U)$ given by (112a)–(112b) (or equivalently, (114)) has a unique solution on the interval $\bar{U} \in [0, U_\Theta]$. Since μ is a smooth function on all of $(t, U) \in \mathbb{R}^2$, the standard existence and uniqueness theorem for ODEs yields a unique solution to (114) for $U \in [0, U_\Theta]$, provided we shrink the size of U_Θ if necessary.

We now prove (125a) and (125b). Differentiating (114) and using (62) and the chain rule, we deduce that:

$$\frac{d^2}{dU^2} t_{\text{CH}}(U) = \frac{1}{4} [\mu(L\mu)]|_{(t_{\text{CH}}(U), U)} + \frac{1}{2} \dot{X}\mu|_{(t_{\text{CH}}(U), U)}, \quad (130)$$

$$\begin{aligned} \frac{d^3}{dU^3} t_{\text{CH}}(U) &= \frac{1}{8} [\mu^2 (LL\mu)]|_{(t_{\text{CH}}(U), U)} + \frac{1}{8} [\mu (L\mu)^2]|_{(t_{\text{CH}}(U), U)} + \frac{1}{4} [\check{X}\mu (L\mu)]|_{(t_{\text{CH}}(U), U)} \\ &\quad + \frac{1}{2} [\mu (\check{X}L\mu)]|_{(t_{\text{CH}}(U), U)} + \frac{1}{2} \check{X}\check{X}\mu|_{(t_{\text{CH}}(U), U)}. \end{aligned} \quad (131)$$

Using (62), (114), (130)–(131), and the expansions (98)–(102), we can Taylor expand the solution $t_{\text{CH}}(U)$ to (114) around $U = 0$, thereby concluding (125a)–(125b).

Since (114) implies that $\mu(t_{\text{CH}}(U), U) = 2 \frac{d}{dU} t_{\text{CH}}(U)$ along $\mathcal{C}_{(0, U_{\oplus})}$, we can use the expansions described in the previous paragraph to conclude the estimate (126) for μ .

To prove (127), we first use (51b), (59), (112a), our assumed smallness on the amplitude of the initial data (which in particular implies that the ratio $\frac{\sinh(\frac{\mathcal{R}_{\text{PS}}^{(+)}}{2}) - c \cosh(\frac{\mathcal{R}_{\text{PS}}^{(+)}}{2})}{\cosh(\frac{\mathcal{R}_{\text{PS}}^{(+)}}{2}) - c \sinh(\frac{\mathcal{R}_{\text{PS}}^{(+)}}{2})}$ from RHS (44b) with $\mathcal{R}_{(-)} \equiv 0$ is $-\mathcal{O}_{\text{pos}}(1)$), and the chain rule to deduce that $\frac{d}{dU} x^1(t_{\text{CH}}(U), U) = \frac{1}{2} \check{L}^1(t_{\text{CH}}(U), U) = \frac{1}{2} [\mu \underline{L}^1](t_{\text{CH}}(U), U) = -\mu \mathcal{O}_{\text{pos}}(1)$. From this relation and (126), we conclude (127).

Finally, $\mathcal{C}_{[0, U_{\oplus}]}$ is a **h**-null curve portion because its tangent vector is everywhere proportional to the vectorfield \check{L} , which is **h**-null by (45) and (59).

Proof of (120): we have already shown that $t_{\text{Sing}}(U)$ is smooth on $[-U_{\oplus}, 0]$, that $t_{\text{CH}}(U)$ is smooth on $[0, U_{\oplus}]$, and that $t_{\text{Sing}}(0) = t_{\text{CH}}(0) = T_{\text{Shock}}$. From these facts, (111), and (116)–(118), we conclude (120).

Proof that $\mu > 0$ on \mathcal{M}_* : We have already shown that $\mu > 0$ on $\mathcal{M}_{\text{Sing}}$. Hence, in view of definition 118, we see that to prove $\mu > 0$ on \mathcal{M}_* , it suffices to prove $\mu > 0$ on \mathcal{M}_{Reg} . To this end, we first note that the same arguments given above imply that the curve $U \rightarrow (t_{\text{Sing}}(U), U)$ is smooth for $U \in [-U_{\oplus}, U_{\oplus}]$ and that $t_{\text{Sing}}(U)$ satisfies the estimate stated in (122b) for $U \in [-U_{\oplus}, U_{\oplus}]$ (not just for $U \in [-U_{\oplus}, 0]$). The portion of this curve corresponding to $U \in (0, U_{\oplus})$ is shown as a dotted curve in figure 2; for purposes of this proof, we denote this dotted curve, which by definition does not contain the crease, by ' $\mathcal{B}_{\text{Fictitious}}$ ' ('fictitious' because the image of $\mathcal{B}_{\text{Fictitious}}$ under Υ in (t, x^1) -space is not part of the **h**-MGHD). The same arguments that we used to prove that $\mu > 0$ on $\mathcal{M}_{\text{Sing}}$ (which relied on (77), (96), and (108)) imply that $\mu > 0$ on $\mathcal{D} := \{(t, U) \in \mathbb{R}^2 \mid 0 \leq t < t_{\text{Sing}}(U), U \in (0, U_{\oplus})\}$, where we note that the top boundary of this region is $\mathcal{B}_{\text{Fictitious}}$. The key point is that (113), (115), and the estimates (122b) (valid also for $U \in (0, U_{\oplus})$) and (125b) imply that $t_{\text{CH}}(U) < t_{\text{Sing}}(U)$ for $U \in (0, U_{\oplus})$ and thus $\mathcal{C}_{(0, U_{\oplus})}$ lies *below* $\mathcal{B}_{\text{Fictitious}}$. In view of definition (117), we see that $\mathcal{M}_* \subset \mathcal{D}$ and hence $\mu > 0$ on \mathcal{M}_{Reg} , as desired.

Proof of the diffeomorphism and homeomorphism properties of Υ : From definition 3.11, we see that the top boundary of \mathcal{M}_* is the following U -parameterized curve (note that the curve is well-defined since $t_{\text{Sing}}(0) = t_{\text{CH}}(0) = T_{\text{Shock}}$):

$$\gamma_{\text{top}}(U) = (t_{\text{top}}(U), U), \quad U \in [-U_{\oplus}, U_{\oplus}], \quad (132)$$

$$t_{\text{top}}(U) := \begin{cases} (t_{\text{Sing}}(U), U), & U \in [-U_{\oplus}, 0], \\ (t_{\text{CH}}(U), U), & U \in [0, U_{\oplus}]. \end{cases} \quad (133)$$

Using (123), (127), and the mean value theorem, we see that the function $U \rightarrow x^1(t_{\text{top}}(U), U)$ is strictly decreasing for $U \in [-U_{\oplus}, U_{\oplus}]$, i.e. the x^1 coordinate function strictly decreases as we move from right to left along the top boundary of \mathcal{M}_* in figure 2. Next, we note that (59) and (60) imply that $\frac{\partial}{\partial U} x^1 = -\frac{c}{n} \mu$. Using this identity, the fact that $\frac{c}{n} > 0$, and the fact that

we have already shown that $\mu > 0$ on \mathcal{M}_* and μ vanishes precisely along the curve portion $\mathcal{B}_{[-U_\Theta, 0]}$ in $\overline{\mathcal{M}_*}$, we see that for any $(t, U) \in \overline{\mathcal{M}_*}$, the map $U \rightarrow x^1(t, U)$ is strictly decreasing. From these two monotonicity properties of x^1 , we conclude, given the shape of $\overline{\mathcal{M}_*}$ (see figure 2), that the map $\Upsilon(t, U) = (t, x^1(t, U))$ is injective on the compact set $\overline{\mathcal{M}_*}$ and thus is a homeomorphism from $\overline{\mathcal{M}_*}$ onto $\Upsilon(\overline{\mathcal{M}_*})$. Also considering the identity $\det d\Upsilon = -\frac{c}{n}\mu$ proved in (67), we further conclude that Υ is a diffeomorphism on \mathcal{M}_* , where $\mu > 0$.

Proof of (128) and the structure of the h-MGHD in Minkowski-rectangular coordinates (t, x^1) : We deduce from (59), (60), (69), (76), the chain rule, the non-degeneracy assumption (88), and (119) that for $(t, U) \in \mathcal{M}_*$, the following estimate holds:

$$\begin{aligned} |\mu \partial_1 \mathcal{R}_{(+)}^{\text{PS}}|(t, U) &= \frac{n(U)}{c(U)} |\check{X} \mathcal{R}_{(+)}^{\text{PS}}(U)| = \frac{n(U)}{c(U)} \left| \frac{d}{dU} \mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U) \right| = \frac{n(U)}{c(U)} \left| \frac{\frac{d}{dU} \mathfrak{A} \left[\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U) \right]}{\frac{d}{d\mathcal{R}_{(+)}^{\text{PS}}} \mathfrak{A} \left[\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U) \right]} \right| \\ &= \left| \frac{L\mu(U)}{\frac{d}{d\mathcal{R}_{(+)}^{\text{PS}}} \mathfrak{A} \left[\mathring{\mathcal{R}}_{(+)}^{\text{PS}}(U) \right]} \right| = \mathcal{O}_{\text{pos}}(1). \end{aligned} \quad (134)$$

From (134), we conclude (128).

The continuity of $\mathcal{R}_{(+)}^{\text{PS}}, L^M \mathcal{R}_{(+)}^{\text{PS}}$ and $\mathbf{h}_{\alpha\beta} = \mathbf{h}_{\alpha\beta}(\mathcal{R}_{(+)}^{\text{PS}})$ all the way up to $\Upsilon(\overline{\mathcal{M}_*})$ in Minkowski-rectangular coordinates (t, x^1) follows from the fact that $\mathcal{R}_{(+)}^{\text{PS}}(t, U)$ is smooth on $\mathbb{R}_t \times \mathbb{R}_U$ and the already proven fact that Υ is a homeomorphism from $\overline{\mathcal{M}_*}$ onto $\Upsilon(\overline{\mathcal{M}_*})$. The non-degeneracy of the 2×2 matrix $\mathbf{h}_{\alpha\beta}$ relative to the (t, x^1) -coordinates follows from the expression (40a) for $\mathbf{h}_{\alpha\beta} = \mathbf{h}_{\alpha\beta}(\mathcal{R}_{(+)}^{\text{PS}})$, our assumed smallness on the amplitude of the initial data (which in particular implies that $0 < c \leq 1$), and the identities stated in proposition 3.2.

Proof of the causal structure of the singular curve in Minkowski-rectangular coordinates:

Using (62), (76)–(77), and (108)–(109), we deduce that along $\mathcal{B}_{[-U_\Theta, 0]}$, the vectorfield Q defined by (129) can be expressed as follows, where G is the source term from (71), viewed as a function of U , and G' is its derivative:

$$Q = \frac{\partial}{\partial t} - \frac{G(U)}{tG'(U)} \frac{\partial}{\partial U} = \frac{\partial}{\partial t} + \frac{G^2(U)}{G'(U)} \frac{\partial}{\partial U}, \quad \text{along } \mathcal{B}_{[-U_\Theta, 0]}. \quad (135)$$

Hence, using (107a) and (107c), we see that Q is well-defined on $\mathcal{B}_{[-U_\Theta, 0]}$, i.e. the denominator term $G'(U)$ on RHS (135) is non-zero for $U \in [-U_\Theta, 0)$. Next, we deduce from (129) that $Q(\frac{c}{n}\mu) = 0$, i.e. that Q is tangent to the level sets of $\frac{c}{n}\mu$. Using the fact that $\frac{c}{n} > 0$ (see (96)), we see that the zero level set of $\frac{c}{n}\mu$ coincides with the zero level set of μ . Hence, in view of definition (109), we see that Q is tangent to $\mathcal{B}_{[-U_\Theta, 0]}$. The inequalities $Qt > 0$ and $QU < 0$ follow easily from (135), (107a), and (107c).

Next, using (59), (60), and (62), we see that the pushforward of $\frac{\partial}{\partial U}$ by Υ , i.e. $d\Upsilon \cdot \frac{\partial}{\partial U}$ is equal to $-\frac{c}{n}\mu \partial_1$. Hence, $d\Upsilon \cdot \frac{\partial}{\partial U}$ vanishes along all of $\Upsilon(\mathcal{B}_{[-U_\Theta, 0]})$, where μ is 0. Moreover, from (62), we deduce that $d\Upsilon \cdot \frac{\partial}{\partial t} = L^\kappa \partial_\kappa$. From these facts and (135), we deduce that for every $q \in \mathcal{B}_{[-U_\Theta, 0]}$, we have $[d\Upsilon(q)] \cdot Q(q) = L|_{\Upsilon(q)}$. In particular, since $L|_{\Upsilon(q)}$ spans the tangent space of $\Upsilon(\mathcal{B}_{[-U_\Theta, 0]})$ and is \mathbf{h} -null (see (45)), it follows that $\Upsilon(\mathcal{B}_{[-U_\Theta, 0]})$ is a \mathbf{h} -null curve. \square

4. The new formulation of the flow in three spatial dimensions

In section 6, we provide an overview of how to extend some aspects of the simple isentropic plane-symmetric shock formation results yielded theorem 3.1 to the much harder case of three spatial dimensions in the presence of vorticity and entropy. The strategy we present in section 6 is based on the approach we used in studying shock formation for the non-relativistic multi-dimensional compressible Euler equations [1, 66, 68]. Fundamental to that strategy is the availability of a new formulation of the flow as a system of geometric wave-transport-div-curl equations that exhibit remarkable regularity properties and geometric null structures. As we explain in section 6, such a formulation allows one to implement multi-dimensional non-linear geometric optics, which is important for the study of shocks (as is already evident from theorem 3.1). In the non-relativistic case, such a formulation was derived in [67, 95], while in the relativistic case, such a formulation was derived in [42]. In section 4, in the relativistic case, we set up the machinery needed to state the formulation from [42] and present it in a simplified, schematic form as theorem 4.1. In section 4.2.5, we make some brief comments on how the new formulation can be extended to more general spacetimes $(\mathcal{M}, \mathbf{g})$.

4.1. Connection to quasilinear wave equations

Although our study of isentropic plane-symmetric solutions in section 3 was based on analyzing the first-order Riemann invariant system (49a) and (49b), as we will see in section 6, in multi-dimensions, it is advantageous to work with the formulation of the flow provided by theorem 4.1, which features geometric wave equations. The main advantage is that for geometric wave equations, there is an advanced geo-analytic machinery available, tied to the vectorfield method, for studying the global behavior of solutions. To motivate the relevance of wave equations for the study of relativistic fluids, we now establish a simple connection between quasilinear wave equations and the Riemann invariant formulation of isentropic plane-symmetric solutions that we provided in section 4. More precisely, upon differentiating (49a) and (49b), we deduce that for isentropic plane-symmetric solutions, the Riemann invariants $(\mathcal{R}_{(+)}^{\text{PS}}, \mathcal{R}_{(-)}^{\text{PS}})$ satisfy the following coupled system of quasilinear wave equations:

$$\underline{L} \mathcal{R}_{(+)}^{\text{PS}} = 0, \quad (136a)$$

$$L \underline{L} \mathcal{R}_{(-)}^{\text{PS}} = 0. \quad (136b)$$

Equations (136a) and (136b) are quasilinear wave equations because by lemma 3.1, the principal operator on the LHSs is proportional to $(\mathbf{h}^{-1})^{\alpha\beta} \partial_\alpha \partial_\beta$, where \mathbf{h} is the acoustical metric.

4.2. Additional fluid variables

The remainder of section 4 concerns solutions $(h, \mathbf{u}^0, \mathbf{u}^1, \mathbf{u}^2, \mathbf{u}^3, s)$ to the relativistic Euler equations (12a)–(12d) in three spatial dimensions without any symmetry, isentropicity, or irrotationality assumptions.

In section 4.2, we introduce some additional fluid variables that complement the ones we introduced in section 2. The new variables play a fundamental role in the new formulation of relativistic Euler flow provided by theorem 4.1.

Definition 4.1 (the \mathbf{u} -orthogonal vorticity of a one-form). Let ξ be a one-form. We define its \mathbf{u} -orthogonal vorticity, denoted by $\text{vort}^\alpha(\xi)$, to be the following vectorfield, where ϵ is the

fully antisymmetric symbol normalized by $\epsilon_{0123} = 1$:

$$\text{vort}^\alpha(\xi) := -\epsilon^{\alpha\beta\gamma\delta} \mathbf{u}_\beta \partial_\gamma \xi_\delta, \quad (137)$$

and we recall that throughout the paper, we use the Minkowski metric \mathbf{m} —as opposed to \mathbf{h} —to lower and raise indices. In particular, $\epsilon^{0123} = -1$.

Note that $\mathbf{u}_\alpha \text{vort}^\alpha(\xi) = 0$, which explains our terminology ‘ \mathbf{u} -orthogonal vorticity’ of ξ .

Definition 4.2 (vorticity vectorfield). We define the vorticity ϖ^α to be the following vectorfield:

$$\varpi^\alpha := \text{vort}^\alpha(H\mathbf{u}) = -\epsilon^{\alpha\beta\gamma\delta} \mathbf{u}_\beta \partial_\gamma (H\mathbf{u}_\delta). \quad (138)$$

Definition 4.3 (entropy gradient one-form). We define S_α to be the following one-form:

$$S_\alpha := \partial_\alpha s. \quad (139)$$

The fluid variables \mathcal{C} and \mathcal{D} from the next definition play a fundamental role in the setup of the new formulation of the flow. In particular, theorem 4.1 shows that they satisfy transport equations with source terms that enjoy remarkable regularity and null structures.

Definition 4.4 (modified fluid variables). We define \mathcal{C}^α and \mathcal{D} to respectively be the following vectorfield and scalar function:

$$\begin{aligned} \mathcal{C}^\alpha &:= \text{vort}^\alpha(\varpi) + c^{-2} \epsilon^{\alpha\beta\gamma\delta} \mathbf{u}_\beta (\partial_\gamma h) \varpi_\delta + (\theta - \theta_{;h}) S^\alpha (\partial_\kappa u^\kappa) + (\theta - \theta_{;h}) u^\alpha (S^\kappa \partial_\kappa h) \\ &\quad + (\theta_{;h} - \theta) S^\kappa ((\mathbf{m})^{-1})^{\alpha\lambda} \partial_\lambda \mathbf{u}_\kappa, \end{aligned} \quad (140a)$$

$$\mathcal{D} := \frac{1}{n} (\partial_\kappa S^\kappa) + \frac{1}{n} (S^\kappa \partial_\kappa h) - \frac{1}{n} c^{-2} (S^\kappa \partial_\kappa h). \quad (140b)$$

We clarify that in (140a), $\theta_{;h}$ denotes the partial derivative of the temperature θ with respect to h at fixed s .

4.2.1. Covariant wave operator. Covariant wave equations play a central role in theorem 4.1. In definition 4.5, we recall the standard definition of a covariant wave operator acting on scalar functions; in theorem 4.1, we apply the covariant wave operator *only* to quantities that we view to be scalar functions.

Definition 4.5 (covariant wave operator acting on a scalar function). The covariant wave operator $\square_{\mathbf{h}}$ acts on scalar functions φ as follows:

$$\square_{\mathbf{h}} \varphi := \frac{1}{\sqrt{|\det \mathbf{h}|}} \partial_\kappa \left\{ \sqrt{|\det \mathbf{h}|} (\mathbf{h}^{-1})^{\kappa\lambda} \partial_\lambda \varphi \right\}. \quad (141)$$

4.2.2. Null forms. In definition 4.6, we recall the definition of standard null forms relative to \mathbf{h} . The key point is that in theorem 4.1, all derivative-quadratic terms on the RHS of the equations are null forms relative to the acoustical metric. In section 6.8.3, we will provide a detailed explanation as to why, in the context of shock formation, such null forms are harmless error terms. However, at this point in the paper, we merely state that \mathbf{h} -null forms are derivative-quadratic nonlinearities that, when a shock forms, are *strictly weaker than the Riccati-type*

terms that drive the blowup (see section 3.5). As we will further explain in remark 6.4, the dangerous Riccati-type terms are ‘hiding’ in the definition of the covariant wave operator on LHS (146a).

Definition 4.6 (standard null forms relative to \mathbf{h}). We define the standard null forms relative to \mathbf{h} (which we refer to as ‘standard \mathbf{h} -null forms’ for short) to be the following derivative-quadratic terms, where ϕ and ψ are scalar functions and $0 \leq \alpha < \beta \leq 3$:

$$\mathfrak{Q}^{(\mathbf{h})}(\partial\phi, \partial\psi) := (\mathbf{h}^{-1})^{\kappa\lambda} \partial_\kappa \phi \partial_\lambda \psi, \quad (142a)$$

$$\mathfrak{Q}_{\alpha\beta}(\partial\phi, \partial\psi) := \partial_\alpha \phi \partial_\beta \psi - \partial_\beta \phi \partial_\alpha \psi. \quad (142b)$$

4.2.3. Two vectorfields. The two vectorfields featured in the following definition will play a role in our forthcoming discussion.

Definition 4.7 (B and N). We define \mathbf{B} and \mathbf{N} to respectively be the vectorfields with the following components relative to the Minkowski-rectangular coordinates (t, x^1, x^2, x^3) :

$$\mathbf{B}^\alpha := \frac{1}{\mathbf{u}^0} \mathbf{u}^\alpha, \quad (143a)$$

$$\mathbf{N}^\alpha := -(\mathbf{h}^{-1})^{\alpha\beta} \partial_\beta t = -(\mathbf{h}^{-1})^{\alpha 0}. \quad (143b)$$

Note that $\mathbf{B}t = 1$, while by (19), \mathbf{N} is the future-directed \mathbf{h} -unit normal to Σ_t . In particular,

$$\mathbf{h}(\mathbf{N}, \mathbf{N}) = -1. \quad (144)$$

We also compute that:

$$\mathbf{h}(\mathbf{B}, \mathbf{N}) = -\mathbf{B}^0 = -1. \quad (145)$$

4.2.4. The geometric wave-transport-div-curl formulation of the flow. We now state—in a somewhat abbreviated form—the formulation of relativistic Euler flow derived in [42]. The proof of theorem is based on differentiating the equations (12a)–(12d) in well-chosen directions and finding a myriad of cancellations and special structures.

Theorem 4.1 ([42, the geometric wave-transport-div-curl formulation of the flow]). Let $\vec{\Psi} = (h, \mathbf{u}^0, \mathbf{u}^1, \mathbf{u}^2, \mathbf{u}^3, s)$ be a C^3 solution to the relativistic Euler equations (12a)–(12d) on the Minkowski spacetime background. Then for $\Psi \in \vec{\Psi}$, the following equations hold, where \mathfrak{Q} denotes a linear combination of standard \mathbf{h} -null forms (see definition 4.6) and f denotes a smooth function, both of which are free to vary from line to line:

$$\square_{\mathbf{h}(\vec{\Psi})} \Psi = f(\vec{\Psi}) \cdot (\mathcal{C}, \mathcal{D}) + f(\vec{\Psi}) \cdot \mathfrak{Q}(\partial\vec{\Psi}, \partial\vec{\Psi}) + f(\vec{\Psi}, S) \cdot (\varpi, S) \cdot \partial\vec{\Psi}, \quad (146a)$$

$$\mathbf{B}(\varpi^i, S^i) = f(\vec{\Psi}, S) \cdot (\varpi, S) \cdot \partial\vec{\Psi}, \quad (146b)$$

$$\begin{aligned} \mathbf{B}(\mathcal{C}^i, \mathcal{D}) &= f(\vec{\Psi}) \cdot \mathfrak{Q}(\partial\varpi, \partial\vec{\Psi}) + f(\vec{\Psi}) \cdot \mathfrak{Q}(\partial S, \partial\vec{\Psi}) + f(\vec{\Psi}) \cdot (\varpi, S) \cdot \mathfrak{Q}(\partial\vec{\Psi}, \partial\vec{\Psi}) \\ &+ f(\vec{\Psi}) \cdot \partial\vec{\Psi} \cdot (\mathcal{C}, \mathcal{D}) + f(\vec{\Psi}) \cdot S \cdot \partial S + f(\vec{\Psi}, \varpi, S) \cdot (\varpi, S) \cdot \partial\vec{\Psi}, \end{aligned} \quad (146c)$$

$$\partial_\kappa \varpi^\kappa = f(\vec{\Psi}) \cdot \varpi \cdot \partial\vec{\Psi}, \quad \text{vort}^\alpha(S) = 0. \quad (146d)$$

4.2.5. Comments on the case of more general ambient spacetimes $(\mathcal{M}, \mathbf{g})$. Here we make some brief comments on some of the adjustments that would be needed to extend the results of section 4 to general \mathbf{g} -globally hyperbolic spacetimes $(\mathcal{M}, \mathbf{g})$. By \mathbf{g} -globally hyperbolic, we mean that there exists a \mathbf{g} -Cauchy hypersurface $\Sigma_0 \subset \mathcal{M}$. By a ‘ \mathbf{g} -Cauchy hypersurface’, we mean a co-dimension one submanifold Σ_0 such that every inextendible \mathbf{g} -causal curve in \mathcal{M} intersects Σ_0 exactly once. It is well known [17, 46] that \mathbf{g} -globally hyperbolic spacetimes admit a smooth time function $t: \mathcal{M} \rightarrow \mathbb{R}$. By a ‘time function’, we mean that $t^{-1}(0) = \Sigma_0$, that the level sets $\Sigma_{t'} := \{t \equiv t'\}$ are \mathbf{g} -spacelike Cauchy hypersurfaces, that $\mathcal{M} = \cup_t \Sigma_t$, and that t has a past-directed timelike gradient $\nabla^\# t := \mathbf{g}^{-1} \cdot dt$ (which is everywhere normal to Σ_t). Here, $\nabla^\# t$ denotes the vectorfield equal to the dual of ∇t with respect to \mathbf{g} .

First, we note that the extra terms described in remark 2.1 would lead to the presence of extra terms in the system (146a)–(146d). However, as we described in remark 2.1, in the context of the study of shock formation, in which fluid gradients are large, the extra terms would be small relative to the shock-driving Riccati-type fluid terms (which are ‘hidden’ in $\square_{\mathbf{h}(\vec{\Psi})} \Psi$; see remark 6.4).

Next, we note that under the above assumptions, we can define a lapse function Φ by $\Phi := (-\mathbf{g}(\nabla t, \nabla t))^{-1/2}$. We can then define the vectorfield $T := -\Phi^2 \nabla^\# t$, which is \mathbf{g} -normal to Σ_t and satisfies $Tt = 1$. Hence, given a coordinate system (x^1, x^2, x^3) on Σ_0 , it can be propagated to any Σ_t by the flow of T , thereby yielding a coordinate system (t, x^1, x^2, x^3) on spacetime. In this setup, the ambient spacetime metric can be decomposed as follows:

$$\mathbf{g} = -\Phi^2 dt \otimes dt + g, \quad (147)$$

where $g = g(t)$ is the Riemannian metric on Σ_t induced by \mathbf{g} , i.e. the first fundamental form of Σ_t with respect to \mathbf{g} . Given \mathbf{g} and the relativistic fluid, the corresponding acoustical metric \mathbf{h} is defined by (21).

The upshot of introducing a time function t is that it allows for many of the constructions from this section to easily be generalized. For example, using (2) and (147), one sees that $\mathbf{u}^0 := \mathbf{u} \cdot \nabla t := \mathbf{u}^\kappa \partial_\kappa t > 0$ and hence one could naturally define the analog of (143a) by $\mathbf{B} := \frac{1}{\mathbf{u}^0} \mathbf{u}$ and the analog of (143b) by $\mathbf{N} := -\mathbf{h}^{-1} \cdot dt$.

5. Some prior works on shocks

In this section, we discuss some prior works on shocks. There is a vast literature, and we cannot hope to mention all of it here. We have aimed to discuss works that provide further context and motivation for sections 2 and 4 and to help prepare the reader for sections 6 and 7.

5.1. Comments on the 1D theory

In Riemann’s foundational paper [84], he developed the method of Riemann invariants and combined it with the method of characteristics to prove that for the 1D non-relativistic compressible Euler equations, there exist large sets of initial data such that a shock forms in the sense that the fluid’s gradient blows up in finite time, though the fluid variables themselves remain bounded. Once one has formulated the flow in terms of Riemann invariants, the proof of the blowup is based on differentiating the equations to obtain a coupled system of Riccati-type equations along the characteristics. This can be viewed as a multi-directional (i.e. there are two characteristic directions in 1D isentropic compressible Euler flow) version of the blowup of $\partial_x \Psi$ that often occurs in solutions to Burgers’ equation $\partial_t \Psi(t, x) + \Psi \partial_x \Psi(t, x) = 0$;

in section 3.5, we discussed this approach in more detail in the case of simple isentropic plane-symmetric solutions to the relativistic Euler equations. Lax later generalized [63] Riemann's results to 2×2 genuinely nonlinear hyperbolic systems in $1D$. John later extended [54] Lax's blowup-results to apply to some systems in $1D$ with more than two unknowns. Christodoulou and Raoul-Perez used a sharpened, more geometric version (in the style of our proof of theorem 3.1) of Lax's approach to give a detailed proof of shock formation in $1D$ for electromagnetic waves in nonlinear crystals [36]. Related results have recently been proved for plane-symmetric solutions to the equations of elasticity [12] and the equations of non-relativistic magnetohydrodynamics [13], and in these works, the plane-symmetric blowup was also used to prove ill-posedness results in $3D$; see also the survey article [14]. For the relativistic Euler equations in $1D$, stable blowup-results have been proved for large sets of initial data (which are allowed to be large) for isentropic flows [16] as well as flows with dynamic entropy [15]. Theorem 3.1 provides a sharpened version of these kind of results for simple isentropic plane-symmetric solutions to the relativistic Euler equations; by 'sharpened', we mean that, different from the works mentioned above, theorem 3.1 follows the solution all the way up to the singular boundary and Cauchy horizon, rather than just up to the time of first blowup.

Even though shocks can form for large classes of hyperbolic PDEs, signifying the end of the classical evolution, ideally, one would like to develop a theory of weak solutions that can accommodate the formation of shocks and their subsequent interactions. In $1D$, a robust such theory exists, at least for strictly hyperbolic systems and initial data that are small in a suitable bounded variation (BV) function space. We refer to [40] for an extensive discussion of the $1D$ theory. In multi-dimensions, much less is known. A key obstacle is that it is known, thanks to the important paper [81] by Rauch, that quasilinear hyperbolic PDEs are typically ill-posed in BV spaces. In fact, for most multi-dimensional quasilinear hyperbolic systems of interest, even local well-posedness is known *only* in L^2 -type Sobolev spaces. In practice, this means that to prove shock formation in multi-dimensions, one must commute the equations and derive energy estimates up the singularity; this turns out to be a difficult task, for reasons we discuss in detail in section 6.

5.2. Blowup-results in multi-dimensions without symmetry: proofs by contradiction

In light of the previous paragraph, it is perhaps not surprising that the first blowup-results without symmetry assumptions in fluid mechanics were non-constructive. Specifically, in [86], Sideris studied the $3D$ compressible Euler equations under equations of state that satisfy a convexity assumption. For open sets of initial data (including near-constant-state-data) without symmetry or irrotationality assumptions, he used arguments based on integrated quantities to give a proof by contradiction that a smooth solution cannot exist for all time. For the $3D$ relativistic Euler equations and relativistic Euler–Maxwell equations, Guo and Tahvildar-Zadeh proved similar results [49] for open sets of initial data that are *large* perturbations of constant-states with positive density. In contrast to the present work, [49, 86] relied on convexity assumptions on the equation of state.

5.3. Proofs of shock formation in multi-dimensions via nonlinear geometric optics

Alinhac [8–11] was the first to prove stable shock formation without symmetry assumptions for scalar quasilinear wave equations in $2D$ and $3D$ of the form:

$$(\mathbf{h}^{-1})^{\kappa\lambda}(\partial\Phi)\partial_\kappa\partial_\lambda\Phi = 0 \quad (148)$$

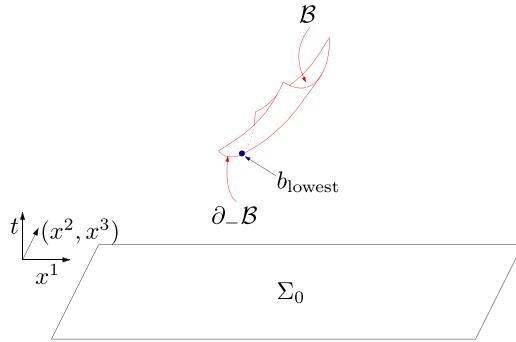


Figure 4. Strictly convex crease and singular boundary.

whenever the nonlinearities in (148) fail to satisfy the null condition. In (148), \mathbf{h} is a Lorentzian metric whose components $\mathbf{h}_{\alpha\beta}$ in the standard coordinate system are prescribed functions of $\partial\Phi$, i.e. functions of the spacetime-gradient of Φ in the standard coordinate system. Alinhac's proof applied to open sets of initial data that satisfy a non-degeneracy assumption that guaranteed that within the constant-time hypersurface of first blowup, the singularity is an isolated point. His non-degeneracy assumption guaranteed that the singular boundary \mathcal{B} and its past boundary $\partial_{-}\mathcal{B}$ are strictly convex, as in figure 4, and in the context of the figure, his approach allowed him to follow the solution up to the flat constant-time hypersurface that contains the point b_{lowest} ; in section 5.5, we discuss \mathcal{B} and $\partial_{-}\mathcal{B}$ in more detail. Alinhac's main results showed that at b_{lowest} , a shock forms in the sense that $\partial^2\Phi$ blows up while $\partial^{\leq 1}\Phi$ remains bounded. His proof relied on nonlinear geometric optics (i.e. eikonal functions and the method of characteristics, as in sections 3.6 and 6.3), and his proof showed that relative to a *geometric coordinate system* constructed out of the eikonal function, Φ and its partial derivatives remain bounded, except possibly at the high derivative levels. To close the energy estimates in the geometric coordinate system without derivative loss, he relied on a Nash–Moser iteration scheme that necessarily terminated at the time of first blowup.

In Christodoulou's breakthrough monograph [31], he proved a substantially sharper result for the 3D relativistic Euler equations in regions where the solution is irrotational and isentropic. The dynamics in such regions is described by a scalar quasilinear wave equation for a potential function Φ (i.e. $\partial\Phi$ determines the four-velocity and proper energy density) that is of the form (148) and that is invariant under the Poincaré group. Like Alinhac, Christodoulou used an eikonal function in his proof, that is, a solution U to the eikonal equation:

$$(\mathbf{h}^{-1})^{\kappa\lambda}(\partial\Phi)\partial_{\kappa}U\partial_{\lambda}U=0. \quad (149)$$

The level sets of U are characteristic for equation (148), i.e. hypersurfaces that are null with respect to \mathbf{h} . Also like Alinhac, Christodoulou constructed a geometric coordinate system tied to U and proved that relative to the geometric coordinates, the solution remains smooth, except possibly at the high derivative levels. He showed that for open sets of initial data, the shock singularity—i.e. the blowing-up of the second-order Minkowski-rectangular derivatives $\partial^2\Phi$ while $\partial^{\leq 1}\Phi$ remains bounded—is tied to the vanishing of the *inverse foliation density* μ , defined by:

$$\mu:=-\frac{1}{(\mathbf{h}^{-1})^{\kappa\lambda}(\partial\Phi)\partial_{\kappa}U\partial_{\lambda}t}=-\frac{1}{(\mathbf{h}^{-1})^{\kappa 0}(\partial\Phi)\partial_{\kappa}U}, \quad (150)$$

where t is the standard time coordinate on Minkowski spacetime. The vanishing of μ signifies the infinite density of the level-sets of U in Minkowski-spacetime, i.e. the infinite density of the characteristics and the blowup of ∂U , as in theorem 3.1. Motivated by techniques used in his joint proof of the stability of Minkowski spacetime with Klainerman [33], Christodoulou was able to avoid Nash–Moser estimates. Rather, to avoid derivative loss in the eikonal function and to close the proof, he used a *geometric energy method* for wave equations and U ; see section 6 for further discussion of a related geometric energy method in the context of the 3D relativistic Euler equations with vorticity and entropy.

The results of [31] go beyond Alinhac’s in several crucial ways. First, for smooth compactly supported perturbations of non-vacuum constant-state initial data, Christodoulou proved that the solution remains global unless μ vanishes, i.e. he proved a conditional global existence result. Second, his geometric energy method allowed for commutator and multiplier vector-fields with time and radial weights, which allowed him to simultaneously study the interaction between the dispersive tendency of waves and the fact that nonlinearities can focus waves and cause singularities. Third, for open sets of initial data that do not have to satisfy Alinhac’s non-degeneracy condition, Christodoulou proved that μ vanishes in finite time, leading to the blowup of $\partial^2 \Phi$. In particular, even *without* strict convexity of the type shown in figure 4, Christodoulou was able to show that there is at least one singular point. Finally, for open sets of initial data that satisfy a non-degeneracy assumption that is weaker than Alinhac’s, Christodoulou was able to follow the solution beyond the time of first blowup and to reveal a portion of the **h**-MGHD, up to the boundary. The portion of the boundary of the **h**-MGHD revealed by [31] was not described explicitly and thus we cannot make straightforward comparisons with theorem 3.1. The weaker notion of non-degeneracy used by Christodoulou coincides with our notion of transversal convexity, which we describe in section 5.5.

Christodoulou’s results [31] have been extended and applied to other equations, including:

- Under the assumption of an irrotational and isentropic flow, the (non-relativistic) 3D compressible Euler equations were handled in [35].
- A larger class of wave equations (e.g. without the assumption of Poincaré invariance) was treated in [76, 92]. See also the survey article [50], which is centered around the works [31, 92].
- Solution regimes that are different than the small, compactly supported data regime were treated in [76, 96].
- Solutions that exist classically precisely on a past-infinite half-slab were studied in [75].
- Some systems in 1D involving multiple speeds of propagation were treated in [12, 13, 36].
- Other multiple speed systems in multi-dimensions were treated in [93, 94].
- 2D compressible Euler solutions with vorticity were handled in [66] with the help of the new formulation of the flow derived in [67], which yields a non-relativistic analog of theorem 4.1 for barotropic equations of state, which by definition are such that the pressure is a function of the density.
- 3D compressible Euler solutions with vorticity and entropy were treated in [68] with the help of the new formulation of the flow derived in [95], which extends the results of [67] to general equations of state in which the pressure is a function of the density and entropy.
- For the 3D compressible Euler equations with vorticity and entropy, a complete description of open sets of solutions up to a neighborhood of the boundary of the **h**-MGHD is provided by our series [1–3], which collectively yield a multi-dimensional version of theorem 3.1 for the non-relativistic equations. See also section 5.5.

5.4. A different approach to multi-dimensional singularity formation

In the works [23, 24], the authors developed a new approach for proving gradient-blowup for solutions the 3D compressible Euler equations under adiabatic equations of state with vorticity (and entropy in [24]). Instead of using nonlinear geometric optics, the authors used modulation parameters to show that for open sets of initial data with large gradients, the solution forms a gradient singularity in finite time that is a perturbation of a self-similar Burgers-type shock. The approach allows one to follow the solution to the time of first blowup, but not further. It applies to non-degenerate initial data such that within the constant-time hypersurface of first blowup, the singularity occurs at an isolated point. That is, the approach applies when the singular boundary \mathcal{B} and its past boundary $\partial_{-}\mathcal{B}$ are strictly convex, as in figure 4, and in the context of the figure, it allows one to follow the solution up to the flat constant-time hypersurface that contains the point b_{lowest} , where the singularity occurs. Such singularities can be viewed as fluid analogs of the ones studied by Alinhac in his aforementioned works [8–11] on quasilinear wave equations. See also the precursor work [22], in which the authors studied the same problem for the 2D compressible Euler equations in azimuthal symmetry, and the work [21], which, in the same symmetry class, constructs unstable self-similar solutions whose cusp-like behavior at the singularity is non-generic.

Self-similar blowup is a phenomenon that occurs in many other PDEs besides those of fluid mechanics. In particular, singularity formation modeled on a self-similar Burgers'-type shock has been proved for various *non-hyperbolic PDEs* [37–39, 62, 77, 80, 102].

5.5. \mathbf{h} -MGHDs

It is of mathematical and physical interest to study shock-forming solutions in a region of classical existence that goes beyond the constant-time hypersurface of first blowup, e.g. in a region larger than the one studied by Alinhac in [8–11], as we described in section 5.3. The holy grail object in this vein is the \mathbf{h} -MGHD, which is what we studied in theorem 3.1 within the class of simple isentropic plane-symmetric solutions to the relativistic Euler equations. Recall that, roughly speaking, the \mathbf{h} -MGHD the largest possible classical solution + \mathbf{h} -globally hyperbolic region that is uniquely determined by the initial data. In particular, the \mathbf{h} -MGHD can have a complicated boundary that includes points lying to the future of the constant-time hypersurface of first blowup, as in figures 3 and 6. Here, ‘ \mathbf{h} -globally hyperbolic’ means that the development contains a Cauchy hypersurface, i.e. a surface such that every inextendible \mathbf{h} -causal curve (where \mathbf{h} is the acoustical metric) in the development intersects it. In the important work [43], the authors showed that for general quasilinear hyperbolic PDEs, *one cannot ensure uniqueness of the MGHD* until one constructs it and shows that it enjoys some crucial structural properties, which are global in nature.

In our series [1–3], for open sets of initial data for the 3D compressible Euler equations leading to shock-forming solutions, we constructed a large (though bounded) portion of the \mathbf{h} -MGHD, up to the boundary, i.e. in the non-relativistic case without symmetry assumptions, we proved an analog of theorem 3.1. In particular, these are the first results to fully justify figure 6 for open sets of solutions to the 3D compressible Euler equations with vorticity and dynamic entropy. Although these works are of mathematical interest in themselves, they are also fundamental for setting up the shock development problem, as we describe in section 5.6.

Our results [1–3] exhibit a localized version of the crucial property that the \mathbf{h} -MGHD ‘lies on one side of its boundary’, as is shown in figure 6. In [43], the authors showed, roughly, that if this ‘one-sidedness’ property holds globally, then the \mathbf{h} -MGHD is unique. They also showed, for a specific hyperbolic PDE and specific initial data, that *uniqueness of the MGHD*

fails and the property fails! As in theorem 3.1, the boundary portion we construct in [1–3] has two kinds of components:

- A singular boundary \mathcal{B} , constructed in [3], along which the fluid's gradient blows up, the past boundary of which is a two-dimensional acoustically spacelike¹² submanifold known as *the crease* and denoted by ' $\partial_{-}\mathcal{B}$ ' in figure 6.
- A Cauchy horizon, constructed in [2] and denoted by ' \mathcal{C} ' in figure 6, which is an acoustically null hypersurface that emanates from the crease and along which the solution extends smoothly (except at the crease). The crease plays the role of the *true initial singularity* in shock-forming solutions.

Our analysis in [1, 2] relies on an assumption of *transversal convexity*, which is a weak, stable form of convexity that can be read off the initial data and propagated by the flow. Roughly, transversal convexity means that when the inverse foliation density μ (which in the context of [1, 2] is an analog of (150) or (156) for the 3D compressible Euler equations) is small, μ has a positive second derivative in a direction transversal to the level sets of the eikonal function U . For the simple isentropic plane-symmetric solutions we studied in section 3, transversal convexity is ensured by the estimate (102) and is manifested in figure 3 by the upward-bending nature of $\Upsilon(\mathcal{B}_{[-U_{\Phi},0]})$. We highlight the importance of transversal convexity with the following remarks.

Even for simple isentropic plane-symmetric solutions, in the absence of transversal convexity, the qualitative structure of the singularity can be radically altered. For example, if instead of the estimate (102) there held $\tilde{X}\tilde{X}\mu \equiv 0$ at the first singularity, then in the (t, x^1) -coordinates picture, an entire continuum of characteristics would fold into a single point, and the change of variables map Υ from theorem 3.1 would dramatically fail to be an injection. This would be a serious obstacle to even the local well-posedness of the shock development problem, which we describe in section 5.6. This is in contrast to the ‘favorable’ situation shown in figure 3, in which, thanks to transversal convexity, the characteristics *graze* the singular boundary $\Upsilon(\mathcal{B}_{[-U_{\Phi},0]})$, but distinct characteristics do not actually intersect along $\Upsilon(\mathcal{B}_{[-U_{\Phi},0]})$.

In the context of the 3D solutions depicted in figure 6, the transversal convexity is manifested by \mathcal{B} being ‘convex in the x^1 -direction’ (i.e. upward bending in the x^1 direction) but *not necessarily convex in the (x^2, x^3) -directions*. Note that if we view plane-symmetric solutions as solutions in 3D with symmetry, then the corresponding singular boundary \mathcal{B} is not strictly convex because of the ‘symmetric directions’ (x^2, x^3) .

Of the works on shock formation cited in section 5.3, Christodoulou's monograph [31] on 3D irrotational and isentropic relativistic fluids is the only one that follows the solution beyond the time of first blowup. He used eikonal functions and foliations of spacetime by \mathbf{h} -spacelike planes Σ (more precisely, the Σ were coordinate planes with respect to the standard Minkowski-rectangular coordinates) to reveal an implicit portion of the boundary of the \mathbf{h} -MGHD, where, for example, the portion of the crease $\partial_{-}\mathcal{B}$ that was revealed is such that the crease had to lie to the future of any Σ that is tangent to it. For initial data such that the crease is a

¹² Acoustically spacelike means spacelike with respect to the acoustical metric of the 3D compressible Euler equations. Acoustically null means null with respect to the acoustical metric.

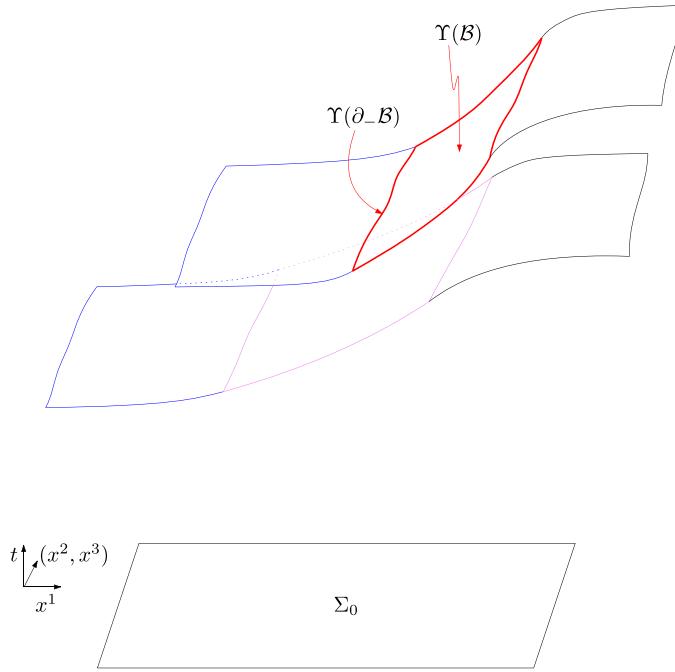


Figure 5. Rough foliations and the singular boundary from [1], depicted in Cartesian coordinate space.

strictly convex subset of Minkowski-rectangular coordinate space (as in figure 4), his approach can be used to study the entire crease and a neighborhood of the singular boundary that emerges from it. However, strict convexity does not hold for all shock-forming solutions. In particular, as we noted in the previous paragraph, strict convexity fails for symmetric solutions (roughly, strict convexity will fail in the directions of symmetry) and general small perturbations of them. In [1], under the weaker assumption of transversal convexity, we construct the entire crease and a neighborhood of the singular boundary that contains it for solutions with vorticity and dynamic entropy. To handle the lack of strict convexity, we dynamically construct a new foliation of spacetime by ‘rough acoustically spacelike hypersurfaces’, depicted in figure 5, that are precisely adapted to the shape of the crease and singular boundary. We refer to the foliations as ‘rough’ because they are less differentiable than the fluid solution, which is a key difficulty that has to be overcome in the PDE analysis. To handle the presence of vorticity and dynamic entropy, we combine the special structure of the geometric wave-transport-div-curl formulation of the flow derived in [95] with upgraded versions of the integral identities derived in [3], which in total yield a regularity theory for the vorticity and entropy that is adapted to the rough foliations (and hence the shape of the singular boundary) and that allows one to treat the fluid equations as a perturbation of a wave equation system. In particular, as in [66, 68], the framework yields (and requires) one extra degree of differentiability for the vorticity and entropy compared to standard estimates. See also section 6.9.2, in which we discuss the extra differentiability of the vorticity and entropy in the context of the 3D relativistic Euler equations.

5.6. Shock development problem

In Majda's celebrated works [71, 72], he solved the *shock front problem* for open sets of solutions to the non-relativistic 3D compressible Euler equations. Roughly, this means that he considered piecewise smooth initial data that have a jump discontinuity across a smooth hypersurface, and he then constructed a local weak solution to the equations and a shock hypersurface such that the solution is piecewise smooth on either side of shock hypersurface but has a jump discontinuity across it and satisfies the Rankine–Hugoniot condition and an entropy-type condition.

The *shock development problem* is closely related to the shock front problem but is laden with additional technical difficulties. It is the problem of describing how initially smooth solutions form a ‘first singularity’ (which is the crease, as described in section 5.5), from which emerges a shock hypersurface (which is not known in advance, as in the shock front problem), and describing the transition of the solution from classical to weak such that in the ‘weak solution region’, the solution jumps across the shock hypersurface and satisfies the Rankine–Hugoniot jump condition and an entropy-type condition. In particular, in the shock development problem, there is no jump discontinuity in the initial data; rather, the jump emerges dynamically, starting from the crease. In problems that are effectively one-dimensional (and thus energy estimates can be avoided), the shock development problem has been solved for various hyperbolic PDEs [20, 28, 34, 52, 53, 64].

The only solution to a multi-dimensional shock development problem without symmetry assumptions was provided by Christodoulou's recent breakthrough monograph [32], in which he used nonlinear geometric optics (i.e. a pair of acoustic eikonal functions, one ingoing and one outgoing) to solve the *restricted shock development problem* in an arbitrary number of spatial dimensions for the compressible Euler equations and the relativistic Euler equations. Roughly, the word ‘restricted’ means that he solved an idealized problem in which he ignored the jump in entropy and vorticity across the shock hypersurface, thereby producing a weak solution to a hyperbolic PDE system that approximates the true one; in the real problem, which remains unsolved, the vorticity and entropy must jump across the shock hypersurface. Compared to the 1D case, the main new difficulty in multiple spatial dimensions is that one must derive energy estimates, which are degenerate near the first singularity for reasons related to the energy degeneracies that occur in the shock formation problem, as we discuss in section 6.9.3.

We highlight the following key issue, which is a primary motivating factor for our construction of large portions of the **h**-MGHD in [1–3]:

The setup of the shock development problem (see [32, 34]) is such that the data for it are an **h**-MGHD¹³ of a shock-forming solution launched by smooth initial data, with a precise description of the gradient-blowup along a singular boundary, a precise description of the classical solution's regular behavior along a Cauchy horizon \mathcal{C} , and a sharp description of the structure of the ‘first singularity’, which in section 5.5 (also in theorem 3.1), we referred to as ‘the crease’ and which we denote by ‘ $\partial_{-}\mathcal{B}$ ’ in figure 6. Crucially, under the ‘transversal convexity’ assumption described in section 5.5, $\partial_{-}\mathcal{B}$ is a co-dimension two, **h**-spacelike submanifold equal to the intersection of the Cauchy horizon

¹³ The shock development problem is local, so to locally study the solution, one only needs as initial data a portion of the **h**-MGHD, specifically a portion that contains a crease and a portion of the singular boundary and Cauchy horizon that emerge from it, which we constructed in [1–3].

and the singular boundary; without this structure, it is not even clear whether the shock development problem is well-posed. These structures, as well as the full state of the fluid up to $\partial_- \mathcal{B}, \mathcal{L}$, were *posed* as data in [32, 34]. In 1D, theorem 3.1 provides the first rigorous *construction* of these qualitative and quantitative aspects of the data for the shock development problem, starting from smooth initial conditions on a spacelike Cauchy hypersurface. For the non-relativistic Euler equations in 3D and without symmetry assumptions, our works [1–3] collectively provide an analog of theorem 3.1.

5.7. Rarefaction waves

In his foundational papers [6, 7], for a large class of multi-dimensional hyperbolic systems that includes scalar conservation laws and the compressible Euler equations as special cases, Alinhac proved local existence and uniqueness for rarefaction wave solutions. Multi-dimensional rarefaction waves are analogs of a class of solutions to the well-known Riemann problem in 1D, in which the initial data are piecewise smooth and discontinuous, and the initial discontinuity is immediately smoothed out by the flow. His approach relied on Nash–Moser estimates to overcome derivative loss in linearized versions of the equations. Alinhac proved [6, 7] that, as in the 1D case, in the corresponding multi-dimensional region + solution, the initial discontinuity is immediately smoothed out by the flow. In the important recent works [69, 70], the authors study the isentropic 2D compressible Euler equations with a family of irrotational discontinuous initial data that are (asymmetric) perturbations of plane-symmetric data for a corresponding 1D Riemann problem. For irrotational data, their main result provides a sharpened version of the local existence and uniqueness results of Alinhac [6, 7]. Their work shows in particular that the 2D irrotational rarefaction solution is a perturbation of the standard 1D rarefaction wave solution to the Riemann problem. Compared to [6, 7], the works [69, 70] yield two important advances. First, the techniques of [69, 70] avoid a loss of derivatives in a corresponding linearized problem and consequently, the authors were able to close the energy estimates without Nash–Moser estimates. Second, [69, 70] provide a complete description of the characteristic geometry in the problem.

6. Some ideas behind the proof of shock formation in 3D without symmetry

In this section, we provide an outline (without complete proofs) of how to extend some aspects of theorem 3.1 to apply to open sets of 3D relativistic Euler solutions without symmetry, isentropicity, or irrotationality assumptions. We anticipate that *all* aspects of theorem 3.1 can be extended. A key reason behind our expectation is the availability of the new formulation of the flow provided by theorem 4.1, which is qualitatively similar to the formulation of the non-relativistic flow from [95], which was used in prior related works [1–3, 66, 68] on shocks for the (non-relativistic) 3D compressible Euler equations. To keep the discussion short, we only sketch some key ideas behind the proof of conjecture 1, which essentially concerns studying 3D solutions until their time of first blowup. Our discussion here mirrors the strategy we used in our work [68], in which we proved an analog of conjecture 1 for the non-relativistic 3D compressible Euler equations.

6.1. Conjectures for the 3D relativistic Euler equations

To set the stage, we state three conjectures for the 3D relativistic Euler equations that are tied to the following figure:

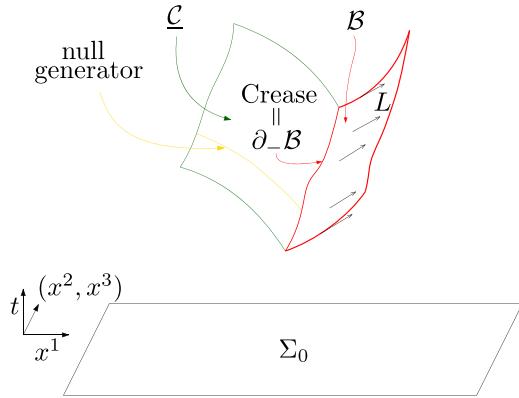


Figure 6. Conjectured local picture of the **h**-MGHD for general small perturbations of simple isentropic plane-symmetric solutions to the 3D relativistic Euler equations in Minkowski-rectangular coordinate space, with the map $\Upsilon(t, U, x^2, x^3) = (t, x^1, x^2, x^3)$ suppressed.

For the 3D compressible Euler equations, figure 6 has been justified [1–3] for open sets of initial data.

Conjecture 1 (time of first blowup for the 3D relativistic Euler equations with vorticity and entropy). Consider the simple isentropic plane-symmetric solutions of theorem 3.1 as ‘background solutions’ to the 3D relativistic Euler equations with symmetry. Consider smooth initial data in 3D—without symmetry, irrotationality, or isentropicity assumptions—that are close (in a sufficiently high-order Sobolev space) to the data of one of the background solutions. Then the perturbed solution forms a shock at a time that is a perturbation of T_{Shock} , i.e. a perturbation of the shock-formation-time of the background solution; the conjectured perturbed first blowup-time is the smallest value of t along the crease $\partial_{-}\mathcal{B}$ depicted in figure 6.

Conjecture 2 (the structure of the crease and singular boundary for the 3D relativistic Euler equations with vorticity and entropy). Under the assumptions of conjecture 1, the perturbed solution has a crease and singular boundary that are perturbations of the crease singular boundary of the background solution; the background crease and singular boundary are depicted in figure 3, while the conjectured perturbed singular boundary and crease are depicted in figure 6.

Conjecture 3 (the structure of the Cauchy horizon with vorticity and entropy). Under the assumptions of conjecture 1, the perturbed solution has a Cauchy horizon that is a perturbation of the Cauchy horizon of the background solution; the background Cauchy horizon is depicted in figure 3, while the conjectured perturbed Cauchy horizon is depicted in figure 6.

Conjecture 2 is strictly harder than conjecture 1, while conjecture 3 is independent (though closely related).

6.2. Almost Riemann invariants

To study general (asymmetric) perturbations of simple isentropic plane-symmetric solutions, it is convenient to use analogs of the Riemann invariants from definition 3.2.

Definition 6.1 (almost Riemann invariants). Let $F = F(H, s)$ be the solution to the following transport equation initial value problem:

$$\frac{\partial}{\partial H} F(H, s) = \frac{1}{Hc(H, s)}, \quad F(\bar{H}, s) = 0, \quad (151)$$

where $\bar{H} > 0$ is the constant (3), and in (151), we are viewing the speed of sound c as a function of H and s .

We define the *almost Riemann invariants* $\mathcal{R}_{(+)}$ and $\mathcal{R}_{(-)}$ as follows:

$$\mathcal{R}_{(+)} := F(H, s) + \frac{1}{2} \ln \left(\frac{1 + \frac{\mathbf{u}^1}{\sqrt{1+(\mathbf{u}^1)^2}}}{1 - \frac{\mathbf{u}^1}{\sqrt{1+(\mathbf{u}^1)^2}}} \right), \quad (152a)$$

$$\mathcal{R}_{(-)} := F(H, s) - \frac{1}{2} \ln \left(\frac{1 + \frac{\mathbf{u}^1}{\sqrt{1+(\mathbf{u}^1)^2}}}{1 - \frac{\mathbf{u}^1}{\sqrt{1+(\mathbf{u}^1)^2}}} \right). \quad (152b)$$

Note that for plane-symmetric solutions with $s \equiv 0$, the quantities defined by (152a) and (152b) coincide with the ones defined in (48a) and (48b).

We view $\mathcal{R}_{(+)}$ and $\mathcal{R}_{(-)}$ as replacements for H and \mathbf{u}^1 that are convenient for studying perturbations of simple isentropic plane-symmetric solutions. In particular, for values of (H, s) near $(\bar{H}, 0)$, the factor $\frac{1}{Hc(H, s)}$ in (151) is positive. The implicit function theorem then allows one to solve for H as a smooth function of F and s . We can then use the relations (152a) and (152b) to express (H, \mathbf{u}^1) as smooth functions of $(\mathcal{R}_{(+)}, \mathcal{R}_{(-)}, s)$. Then, under the algebraic relation (12d), a complete set of state-space variables for the 3D relativistic Euler equations is given by:

$$\vec{\Psi} := (\mathcal{R}_{(+)}, \mathcal{R}_{(-)}, \mathbf{u}^2, \mathbf{u}^3, s). \quad (153)$$

Although we have previously used the symbol $\vec{\Psi}$ to denote the array $(h, \mathbf{u}^0, \mathbf{u}^1, \mathbf{u}^2, \mathbf{u}^3, s)$, in the rest of section 6, we use $\vec{\Psi}$ to denote the array in (153).

Away from isentropic plane-symmetry, $\mathcal{R}_{(+)}$ and $\mathcal{R}_{(-)}$ no longer solve the homogeneous transport equations (49a) and (49b). However, from (146a), the chain rule, and definition 6.1, it follows that for general solutions, $\mathcal{R}_{(+)}$ and $\mathcal{R}_{(-)}$ satisfy geometric wave equations that have the same schematic form as (146a). In the rest of section 6, we use this fact without always explicitly mentioning it.

6.3. Nonlinear geometric optics

As in the 1D case treated in section 3, the study of shock formation in 3D relies on nonlinear geometric optics, i.e. multi-dimensional analogs of the constructions from section 3.6. The main object behind the construction is an eikonal function. The use of eikonal functions to study the *global* properties of multi-dimensional quasilinear hyperbolic PDE solutions originated Christodoulou–Klainerman’s proof [33] of the stability of Minkowski spacetime. See also section 5.3 for a discussion of nonlinear geometric optics in the context of proofs of multi-dimensional shock formation.

As we discussed at the beginning of section 3.6, for simple isentropic plane-symmetric solutions, the role of nonlinear geometric optics/geometric coordinates is to yield a framework/differential structure in which the solution remains smooth all the way up to the shock, thereby allowing one to reduce the problem of shock formation to a more traditional problem in which one studies long-time existence. The same remarks hold in the present context of three spatial dimensions, although many additional geometric constructions are required. In particular, all of the geometric quantities we introduce in sections 6.3 and 6.8, including the acoustic eikonal function, the geometric coordinates, various geometric vectorfields, and various geometric energies are introduced in order regularize the problem. However, unlike in the simple isentropic plane-symmetric case, in the present multi-dimensional context, one needs to derive energy estimates, and it turns out that a remnant of the shock singularity survives, even in the ‘good’ geometric coordinates. The remnant manifests as singular estimates for the high-order geometric energies, though it is crucial for our approach that the mid-order-and-below geometric energies remain bounded; see section 6.9.3 for further discussion.

Definition 6.2 (acoustic eikonal function). An acoustic eikonal function (eikonal function for short) is a solution to the eikonal equation, which is the following fully nonlinear hyperbolic PDE, where \mathbf{h}^{-1} is the acoustical inverse metric from definition 2.4:

$$(\mathbf{h}^{-1})^{\kappa\lambda} (\vec{\Psi}) \partial_\kappa U \partial_\lambda U = 0, \quad \partial_t U > 0. \quad (154)$$

When studying shocks close to plane-symmetry, it is convenient to assume the following initial condition for U , which is the exact same initial condition (55b) that we used in our study of isentropic plane-symmetric solutions:

$$U|_{t=0} = -x^1. \quad (155)$$

We refer to the level sets of U as *the characteristics*. One can show that for isentropic plane-symmetric solutions, the solution ‘ U ’ to (55a) and (55b) coincides with the solution to (154) and (155).

Definition 6.3 (inverse foliation density). We define the inverse foliation density of the characteristics to be the following scalar function:

$$\mu := -\frac{1}{(\mathbf{h}^{-1})^{\kappa\lambda} (\vec{\Psi}) \partial_\kappa t \partial_\lambda U} = \frac{1}{NU}, \quad (156)$$

where the second equality in (156) follows from (143b).

One can show that for isentropic plane-symmetric solutions, the quantity ‘ μ ’ defined by (57) coincides with the one defined in (156). As in section 3, the formation of a shock occurs when the density of the characteristics becomes infinite, that is, when $\mu \downarrow 0$.

6.4. Geometric coordinates and vectorfield frames tied the eikonal function

In this section, we introduce geometric coordinates and related vectorfields that are analogs of quantities we used in our study isentropic plane-symmetric solutions.

6.4.1. t Geometric coordinates.

Definition 6.4 (geometric coordinates and the corresponding partial derivative vector-fields). We refer to $\{t, U, x^2, x^3\}$ as the *geometric coordinates* on spacetime, where U is the eikonal function and t, x^2 , and x^3 are the Minkowski-rectangular coordinates, with $t := x^0$. We denote the corresponding partial derivatives by $\{\frac{\partial}{\partial t}, \frac{\partial}{\partial U}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\}$ (which are not to be confused with the partial derivatives ∂_α in the (t, x^1, x^2, x^3) -coordinate system).

Definition 6.5 (some important submanifolds of geometric coordinate space).

- We denote the level set $\{U = U'\}$ by $\mathcal{P}_{U'}$, and we refer to these level sets as *the characteristics*; see figure 7.
- We define $\Sigma_{t'}$ to be the hypersurface $\{t = t'\}$, where t is the Minkowski time function.
- We define the two-dimensional surfaces $\ell_{t', U'} := \Sigma_{t'} \cap \mathcal{P}_{U'}$; see figure 7.

The surfaces \mathcal{P}_U are characteristic for the wave operator $\square_{\mathbf{h}}$ on LHS (146a), or equivalently, null with respect to the acoustical metric \mathbf{h} (\mathbf{h} -null for short).

As we mentioned in section 3 and just above definition 6.2, a key strategy behind the proof of shock formation is to show that the solution remains smooth with respect to the geometric coordinates. The blowup of the solution's gradient in (t, x^1, x^2, x^3) -coordinates, i.e. the blowup of $\partial \vec{\Psi}$, occurs when μ vanishes. One can prove (see e.g. [68, lemmas 2.22 and 2.23]) the following relationship between the partial derivatives in the two coordinate systems, where the a_α^β are smooth (solution-dependent) functions of the geometric coordinates:

$$\partial_\alpha = a_\alpha^0 \frac{1}{\mu} \frac{\partial}{\partial U} + a_\alpha^1 \frac{\partial}{\partial t} + a_\alpha^2 \frac{\partial}{\partial x^2} + a_\alpha^3 \frac{\partial}{\partial x^3}. \quad (157)$$

The first product on RHS (157) shows in particular why $\partial \vec{\Psi}$ can blow up as $\mu \downarrow 0$, even when the derivatives of $\vec{\Psi}$ with respect to elements of $\{\frac{\partial}{\partial t}, \frac{\partial}{\partial U}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\}$ remain bounded. The use of geometric coordinates allows one to study the problem of shock formation by using ideas that have traditionally been used in long-time existence problems. However, different from the 1D case, in 3D, some difficult degeneracies can occur in the energy high-order energy estimates; we discuss this difficulty in more detail in section 6.9.3.

6.4.2. Vectorfield frames. Experience has shown (e.g. [31] and the related works cited in section 5.3) that in the study of shocks, to avoid loss of derivatives in commutator estimates and other difficulties, it is advantageous to use vectorfield frames adapted to the eikonal function. In particular, once one has constructed U , one can construct the following vectorfield frames, depicted in figure 7:

$$\mathcal{X} := \{L, \check{X}, Y_{(2)}, Y_{(3)}\}, \quad \mathcal{P} := \{L, Y_{(2)}, Y_{(3)}\}, \quad \mathcal{Y} := \{Y_{(2)}, Y_{(3)}\}, \quad (158)$$

which we describe below. The frame \mathcal{X} spans the tangent space of geometric coordinate space, the subset \mathcal{P} spans the tangent space of the characteristics \mathcal{P}_U , and the subset \mathcal{Y} spans the tangent space of the two-dimensional surfaces $\ell_{t, U}$, i.e. the span of \mathcal{Y} is equal to that of $\{\frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\}$.

We now briefly describe how to construct the vectorfields. First, L is defined by:

$$L^\alpha := -\mu (\mathbf{h}^{-1})^{\alpha\kappa} (\vec{\Psi}) \partial_\kappa U. \quad (159)$$

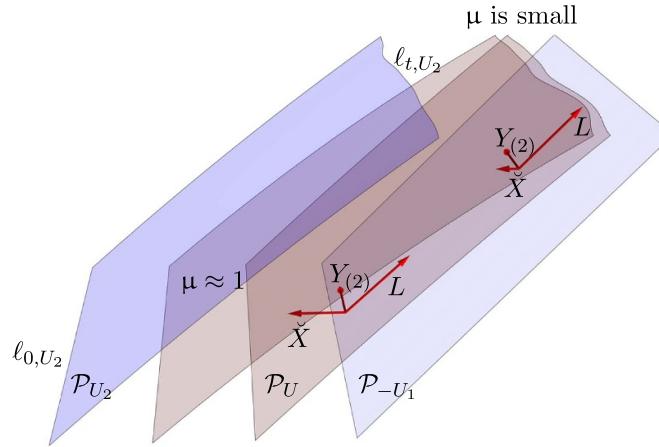


Figure 7. The frame \mathcal{L} at two distinct points on \mathcal{P}_U with the x^3 -direction suppressed.

L is clearly \mathbf{h} -orthogonal to the characteristics \mathcal{P}_U , and by (154) and (156), we see that $\mathbf{h}(L, L) = 0$ and $Lt = 1$. In particular, since L is \mathbf{h} -orthogonal to \mathcal{P}_U and \mathbf{h} -null, \mathcal{P}_U is a \mathbf{h} -null hypersurface. One can show that for isentropic plane-symmetric solutions, the quantity ‘ L ’ defined in this fashion coincides with the one defined in (44a). However, for general solutions in 3D, a big difference with the plane-symmetric case is that there does not exist any explicit formula for L in the spirit of (44a). That is, in general, L depends on the gradient of U , which in turns depends implicitly on the fluid solution via the coefficients in the eikonal equation (154).

Next, we define X to be the Σ_t -tangent vectorfield that is \mathbf{h} -orthogonal to the $\ell_{t,U}$, normalized by $\mathbf{h}(X, X) = 1$, and is such that U increases along the integral curves of X , i.e. $XU > 0$. We then define $\check{X} := \mu X$, where μ is defined in (156). It follows that $\mathbf{h}(\check{X}, \check{X}) = \mu^2$. With the help of (19), one can show that $\check{X}U = 1$. That is, \check{X} is a geometric replacement¹⁴ for $\frac{\partial}{\partial U}$, and among the elements of \mathcal{L} , it is the one vectorfield that is transversal to the characteristics. One can show that for isentropic plane-symmetric solutions, the quantities ‘ X ’ and ‘ \check{X} ’ defined in this fashion coincide with the ones defined in (60).

Next, for $A = 2, 3$, we define $Y_{(A)}$ to be the \mathbf{h} -orthogonal projection of ∂_A onto $\ell_{t,U}$, where ∂_A is the standard Minkowski-rectangular partial derivative vectorfield. In our study of isentropic plane-symmetric solutions, we did not rely on any analogs of the $Y_{(A)}$.

For future use, we use (12d), (17a), and (143a) to compute that:

$$\mathbf{h}(\mathbf{B}, L) = \mathbf{n} \mathbf{m}(\mathbf{B}, L). \quad (160)$$

Note that RHS (160) < 0 because $\mathbf{n} > 0$, \mathbf{B} is future-directed and \mathbf{h} -timelike, L is future-directed and \mathbf{h} -null, and the \mathbf{h} -null cones are inside the \mathbf{m} -null cones (see remark 2.3).

6.4.3. The first fundamental form of $\ell_{t,U}$ and related decompositions. Standard calculations (see, e.g. [96, equation (2.40b)]) imply that the inverse acoustical metric from (17b) can be

¹⁴ (63) shows that in plane-symmetry, $\check{X} = \frac{\partial}{\partial U}$. However, in the general 3D case, \check{X} is equal to $\frac{\partial}{\partial U}$ plus corrections belonging to $\text{span} \left\{ \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3} \right\}$.

decomposed as follows:

$$(\mathbf{h}^{-1})^{\alpha\beta} = -L^\alpha L^\beta - L^\alpha X^\beta - X^\alpha L^\beta + (\mathbb{h}^{-1})^{\alpha\beta}, \quad (161)$$

where \mathbb{h} is the Riemannian metric on $\ell_{t,U}$ induced by \mathbf{h} , i.e. the first fundamental form of $\ell_{t,U}$. We view \mathbb{h} to be a spacetime tensor that vanishes on contraction with L or X and that agrees with \mathbf{h} on $\ell_{t,U}$ -tangent vectors. Moreover, in (161), $(\mathbb{h}^{-1})^{\alpha\beta} := (\mathbf{h}^{-1})^{\alpha\gamma}(\mathbf{h}^{-1})^{\beta\delta}\mathbb{h}_{\gamma\delta}$ is the corresponding inverse first fundamental form.

In what follows, ∇ denotes the Levi–Civita connection of \mathbb{h} and $|\cdot|_{\mathbb{h}}$ denotes the pointwise norms of tensors with respect to \mathbb{h} , e.g. for scalar functions f , we have $|\nabla f|_{\mathbb{h}}^2 = (\mathbb{h}^{-1})^{\kappa\lambda}\nabla_\kappa f \nabla_\lambda f$.

For future use, we note that for scalar functions f , we have the following decomposition of $\mu\square_{\mathbf{h}(\bar{\Psi})}f$ (see, e.g. [96, proposition 2.16]), where Δ denotes the covariant Laplacian associated to \mathbb{h} :

$$\mu\square_{\mathbf{h}(\bar{\Psi})}f = -L(\mu Lf + 2\check{X}f) + \mu\Delta f + \dots \quad (162)$$

On RHS (162), \dots denotes terms that depend on the first derivatives of f and that we will not discuss in detail here.

Remark 6.1 (comments on the case of more general ambient spacetimes $(\mathcal{M}, \mathbf{g})$). Under the setup described in section 4.2.5, almost all the constructions from this section have natural analogs for the study of the relativistic Euler equations on a \mathbf{g} -globally hyperbolic spacetime $(\mathcal{M}, \mathbf{g})$. That is, consider the time function $t: \mathcal{M} \rightarrow \mathbb{R}$ and the foliations $\mathcal{M} = \cup_t \Sigma_t$. Here are three concrete examples:

- (I) The definition of the eikonal function (17a) would remain unchanged.
- (II) With $\mathbf{N} := -\mathbf{h}^{-1} \cdot dt$, (156) would again yield a suitable definition of the inverse foliation density (of the characteristics, with respect to Σ_t).
- (III) X could be defined to be a Σ_t -tangent vectorfield that is \mathbf{h} -orthogonal to $\ell_{t,U}$ and normalized by $\mathbf{h}(X, X) = 1$ (this would define X up to an overall minus sign, where in section 3, we defined X such that is ‘left-pointing’).

We do, however, highlight the following point: general spacetime manifolds have no symmetries and thus do not generally allow for the existence of the simple plane-symmetric solutions constructed in section 3. Nonetheless, we expect the robust geometric framework described above and many of the key ideas presented in the subsequent sections to be useful for proving shock formation results for initial data with large gradients. We also note that the initial data (155) for the eikonal function generally needs to be adjusted to fit the regime one is studying.

6.5. Description of the compactly supported initial data of perturbations of simple isentropic plane-symmetric solutions

We now describe the initial data in conjecture 1 in more detail. One simply takes any of the ‘background’ simple isentropic plane-symmetric shock-forming solutions from theorem 3.1 and then considers a general small (asymmetric) perturbation of its initial data on Σ_0 . For the proof to close, the perturbed initial data need to belong to a sufficiently high order Sobolev space; we will discuss the issue of regularity in more detail later on. For convenience, it is easiest to consider perturbed initial data that, like the background data, are compactly supported in $\Sigma_0^{[-U_1, U_2]} := \Sigma_0 \cap \{(t, U, x^2, x^3) \mid -U_1 \leq U \leq U_2\}$. Then by finite speed of propagation,

the corresponding perturbed solution remains spatially compactly supported for all time and in particular vanishes along the characteristic \mathcal{P}_{-U_1} (which therefore appears flat in figure 6). The compact support allows one to avoid difficult boundary terms in the elliptic estimates for the vorticity and entropy; see section 6.9.2.

For such initial data, all of the wave variables in the array $\vec{\Psi}$ defined in (153)—except for $\mathcal{R}_{(+)}$ —will initially have a negligible effect on the dynamics. Similarly, the fluid variables ϖ^α , S_α , \mathcal{C}^α , and \mathcal{D} from section 4.2 will also be small initially. The challenge is to propagate suitable versions of this smallness all the way up to the shock.

6.6. Bootstrap assumptions and the regions $\mathcal{M}_{[0,t'),[-U_1,U']}$

To prove that a shock forms, one commutes the equations up to N_{top} times and derives energy estimates up to top-order and L^∞ estimates at the lower orders. For reasons described later on, the proof requires N_{top} to be rather large (i.e. substantially larger than a proof of local well-posedness requires) and thus the initial data need to belong to a high order Sobolev space. To close the proof, it is convenient to make L^∞ bootstrap assumptions that capture the expectation that at the lower derivative levels, the solution behaves like a perturbation of a simple isentropic plane-symmetric solution (in which only $\mathcal{R}_{(+)}$ is non-vanishing). That is, one aims to propagate various aspects of the smallness enjoyed by the initial data described in section 6.5. The bootstrap assumptions are convenient because explicit formulae in the spirit of corollary 3.6 are not available in 3D. They capture the expectation that nonetheless, the 3D solution should obey estimates that are similar to the ones proved in lemma 3.7 and theorem 3.1 in the simple isentropic plane-symmetric case.

Before introducing the bootstrap assumptions, we first introduce some notation. In what follows, \mathcal{P}^N denotes an arbitrary order N string of elements of the \mathcal{P}_U -tangent vectorfields \mathcal{P} defined in (158). We refer to such vectorfields as ‘tangential’ (to the characteristics \mathcal{P}_U). In contrast, the vectorfield \tilde{X} from (158) is transversal (because it satisfies $\tilde{X}U = 1$).

The bootstrap assumptions are made on a region $\mathcal{M}_{[0,T_{\text{Boot}}),[-U_1,U_3]}$ for some $T_{\text{Boot}} > 0$ and $U_3 > U_2$ (discussed further below, where U_1 and U_2 are as in section 6.5), where in the rest of section 6, given any $t' > 0$ and $U' \in [-U_1, U_2]$, $\mathcal{M}_{[0,t'),[-U_1,U']}$ denotes the following open-at-the-top slab in geometric coordinate space:

$$\mathcal{M}_{[0,t'),[-U_1,U']} := \{(t, U, x^2, x^3) \mid t \in [0, t')\} \cap \{(t, U, x^2, x^3) \mid -U_1 \leq U \leq U'\}. \quad (163)$$

The proof will show that μ first vanishes at a time that is a perturbation of the plane-symmetric blowup-time T_{Shock} from theorem 3.1. That is, in all estimates and arguments, one can safely assume that $0 < T_{\text{Boot}} \leq 2T_{\text{Shock}}$. From this assumption, the assumption that the data are supported in $\Sigma_0^{[-U_1,U_2]}$, and finite speed of propagation, one can compute/choose $U_3 > U_2$ such that for $t \in [0, T_{\text{Boot}})$, the solution is trivial¹⁵ at points in Σ_t belonging to the complement of $\Sigma_t \cap \{U \in [-U_1, U_3]\}$.

Specifically, for N up to mid-order, i.e. roughly, for $N \leq \frac{N_{\text{top}}}{2}$, one assumes L^∞ smallness on $\mathcal{M}_{[0,T_{\text{Boot}}),[-U_1,U_3]}$ for the following quantities, which, aside from $L\mu$, vanish in the case of simple isentropic plane-symmetric solutions:

- The pure \mathcal{P}_u -tangential derivatives $\mathcal{P}^N \mathcal{R}_{(+)}$, $\mathcal{P}^N \varpi^\alpha$, $\mathcal{P}^N S_\alpha$, $\mathcal{P}^N \mathcal{C}^\alpha$, $\mathcal{P}^N \mathcal{D}$, $\mathcal{P}^N L^i$, and $\mathcal{P}^N \mu$, except $L\mu$ does not have to be small (see (119)).

¹⁵ For general perturbations of the background solutions, the perturbed solution will not be spatially supported in $\Sigma_t \cap \{U \in [-U_1, U_2]\}$; this is why we have introduced the parameter U_3 .

- Mixed tangential-transversal derivatives of all these quantities except μ , e.g. $L\check{X}\mathcal{R}_{(+)}$.
- Pure transversal derivatives of all fluid variables excluding $\mathcal{R}_{(+)}$, e.g. for $\check{X}\mathcal{R}_{(-)}$, $\check{X}v^2$, $\check{X}\varpi^\alpha$, $\check{X}\mathcal{C}^\alpha$.

One also assumes L^∞ *boundedness* (not smallness) on $\mathcal{M}_{[0, T_{\text{Boot}}), [-U_1, U_3]}$ for the following quantities, which are non-zero for the simple isentropic plane-symmetric solutions we studied in theorem 3.1:

- $\check{X}\mathcal{R}_{(+)}$
- $\check{X}L^i, L\mu$.

One also assumes that:

$$\mu > 0, \quad \text{on } \mathcal{M}_{[0, T_{\text{Boot}}), [-U_1, U_3]}. \quad (164)$$

The assumption (164) captures that no shock has occurred on $\mathcal{M}_{[0, T_{\text{Boot}}), [-U_1, U_3]}$, though it leaves open the possibility that the shock happens exactly at time T_{Boot} .

The difficult part of the proof is proving energy estimates that allow one to derive, through a combination of Sobolev embedding and transport estimates, improvements of the bootstrap assumptions. Then through a standard continuation argument, one can extend the solution all the way up until the first time that μ vanishes.

6.7. The behavior of μ and the formation of a shock

Given the framework established above and the bootstrap assumptions, it is not difficult to prove that shocks can form for open sets of initial data that are perturbations of simple isentropic plane-symmetric data. We now sketch the argument. Using the eikonal equation (154) and the bootstrap assumptions of section 6.6 one can show that for perturbations of simple isentropic plane-symmetric solutions, the inverse foliation density defined in (156) satisfies an evolution equation of the following schematic form (see lemma 3.5) on $\mathcal{M}_{[0, T_{\text{Boot}}), [-U_1, U_3]}$:

$$L\mu(t, U, x^2, x^3) = \check{X}\mathcal{R}_{(+)}(t, U, x^2, x^3) + \dots, \quad (165)$$

where \dots denotes small error terms. In addition, using that $L\check{X}\mathcal{R}_{(+)}$ is small by the bootstrap assumptions, one can show that on $\mathcal{M}_{[0, T_{\text{Boot}}), [-U_1, U_3]}$, we have:

$$\check{X}\mathcal{R}_{(+)}(t, U, x^2, x^3) = \check{X}\mathcal{R}_{(+)}(0, U, x^2, x^3) + \dots, \quad (166)$$

that is, $\check{X}\mathcal{R}_{(+)}$ stays close to its initial condition (see (69), which shows that in simple isentropic plane-symmetry $\mathcal{R}_{(+)}$ depends only on U). Combining (165) and (166), we see that on $\mathcal{M}_{[0, T_{\text{Boot}}), [-U_1, U_3]}$, we have:

$$L\mu(t, U, x^2, x^3) = \check{X}\mathcal{R}_{(+)}(0, U, x^2, x^3) + \dots. \quad (167)$$

Since $Lt = 1$ (i.e. $L = \frac{d}{dt}$ along the integral curves of L), and since the $\ell_{t,U}$ -tangential derivative $\frac{\partial}{\partial x^A} \mu$ is small by the bootstrap assumptions, we deduce from (167) and the fundamental theorem of calculus that on $\mathcal{M}_{[0, T_{\text{Boot}}), [-U_1, U_3]}$, we have:

$$\mu(t, U, x^2, x^3) = \mu(0, U, x^2, x^3) + t \left\{ \check{X}\mathcal{R}_{(+)}(0, U, x^2, x^3) + \dots \right\}. \quad (168)$$

Thus, for initial data such that there are points where $\check{X}\mathcal{R}_{(+)}(0, U, x^2, x^3)$ is negative and large enough to dominate the terms \dots on RHS (168), we infer from (168) that $\min_{\Sigma_t} \mu$ will vanish in finite time, and that the time of first vanishing can be controlled by the initial data. Moreover, since this argument implies that $|\check{X}\mathcal{R}_{(+)}| \neq 0$ at the points where μ vanishes, and since:

$$|X\mathcal{R}_{(+)}| = \frac{1}{\mu} |\check{X}\mathcal{R}_{(+)}|, \quad (169)$$

it follows that $|X\mathcal{R}_{(+)}|$ blows up like $\frac{C}{\mu}$ at the points where μ vanishes (see (128) and (157)). In total, these arguments show why μ can vanish in finite time and reveal how the vanishing is connected to the blowup of $|X\mathcal{R}_{(+)}|$.

We also note that the above arguments yield the following:

$$L\mu(t, U, x^2, x^3) < 0, \quad \text{at points where } \mu(t, U, x^2, x^3) \text{ is close to 0,} \quad (170)$$

which is an analog of the estimate (119) in simple isentropic plane-symmetry. In particular, at fixed (U, x^2, x^3) , μ vanishes linearly in t . These facts have important implications for the energy estimates, described below; see in particular remark 6.3.

6.8. Vectorfield multiplier method and energies

In this section, we describe some basic ingredients that are needed to set up the energy estimates. To control the solution up to the shock, one needs geometric machinery that is much more advanced than that of section 2.3.4 and that takes advantage of the structure of the equations provided by theorem 4.1.

6.8.1. Energies and null fluxes for wave equations. In this section, we set up the vectorfield multiplier method for the wave equations (146a).

Given a scalar function f , we define the energy–momentum tensor associated to it to be the following symmetric type $(0, 2)$ tensor:

$$\mathbf{Q}_{\alpha\beta} = \mathbf{Q}_{\alpha\beta}[f] := \mathbf{D}_{\alpha}f \mathbf{D}_{\beta}f - \frac{1}{2}\mathbf{h}_{\alpha\beta}(\mathbf{h}^{-1})^{\kappa\lambda}\mathbf{D}_{\kappa}f \mathbf{D}_{\lambda}f. \quad (171)$$

In (171) and throughout, \mathbf{D} denotes the Levi–Civita connection of \mathbf{h} .

Straightforward calculations, based on the symmetry property $\mathbf{D}_\alpha \mathbf{D}_\beta f = \mathbf{D}_\beta \mathbf{D}_\alpha f$ and the Leibniz rule, yield that given any *multiplier vectorfield* Z , we have the following identity:

$$(\mathbf{h}^{-1})^{\alpha\beta} (\mathbf{D}_\alpha \mathbf{Q}_{\beta\kappa} [f] Z^\kappa) = \left(\square_{\mathbf{h}(\tilde{\Psi})} f \right) Z f + \frac{1}{2} (\mathbf{h}^{-1})^{\alpha\beta} (\mathbf{h}^{-1})^{\kappa\lambda} \mathbf{Q}_{\alpha\kappa}^{(Z)} \boldsymbol{\pi}_{\beta\lambda}, \quad (172)$$

where:

$${}^{(Z)}\boldsymbol{\pi}_{\alpha\beta} := \mathbf{h}_{\beta\kappa} \mathbf{D}_\alpha Z^\kappa + \mathbf{h}_{\alpha\kappa} \mathbf{D}_\beta Z^\kappa \quad (173)$$

is the deformation tensor of Z (with respect to \mathbf{h}). The identity (172) is convenient for book-keeping in the divergence theorem when one integrates by parts.

To derive energy-null flux identities for wave equations, we will use (172) with the multiplier vectorfield Z equal to \check{T} , which is defined as follows, where L and \check{X} are as in section 6.4.2:

$$\check{T} := (1 + 2\mu) L + 2\check{X}. \quad (174)$$

One can compute that $\mathbf{h}(\check{T}, \check{T}) = -4\mu(1 + \mu)$ and thus \check{T} is \mathbf{h} -timelike whenever $\mu > 0$. The \mathbf{h} -timelike nature of \check{T} is crucial for generating coercive energies. The μ -weights in (174) have been carefully placed. Recall that $\check{X} = \mu X$, where $\mathbf{h}(X, X) = 1$. Thus, \check{X} effectively contains a μ -weight.

To close the L^2 estimates, one needs energies on:

$$\Sigma_{t'}^{[-U_1, U']} := \{(t, U, x^2, x^3) \mid t = t'\} \cap \{(t, U, x^2, x^3) \mid -U_1 \leq U \leq U'\} \quad (175)$$

and null fluxes on:

$$\mathcal{P}_U^{[0, t']} := \{(t, U, x^2, x^3) \mid 0 \leq t < t'\} \cap \{(t, U, x^2, x^3) \mid U = U'\}. \quad (176)$$

Recall that \mathbf{N} is the future-directed \mathbf{h} -unit normal to Σ_t defined in (143b). The strength of our energies on the $\Sigma_t^{[-U_1, U]}$ is determined by the following identity (see, e.g. [96, lemma 3.4]):

$$\mathbf{Q}[f](\check{T}, \mathbf{N}) = \frac{1}{2} (1 + 2\mu) \mu (Lf)^2 + 2\mu(\check{X}f)Lf + 2(\check{X}f)^2 + \frac{1}{2} (1 + 2\mu) \mu |\nabla f|_{\mathbf{h}}^2. \quad (177)$$

Similarly, the strength of our null fluxes on the $\mathcal{P}_U^{[0, t']}$ is determined by the following identity (again, see, e.g. [96, lemma 3.4]):

$$\mathbf{Q}[f](\check{T}, L) := (1 + \mu) (Lf)^2 + \mu |\nabla f|_{\mathbf{h}}^2. \quad (178)$$

These identities (177)–(178) motivate the following definition.

Definition 6.6 (wave equation energies and null-fluxes). Given a scalar function f , we define the associated energy $\mathbb{E}_{(\text{Wave})}[f]$ and null flux $\mathbb{F}_{(\text{Wave})}[f]$ as follows:

$$\begin{aligned} \mathbb{E}_{(\text{Wave})}[f](t, U) \\ := \int_{\Sigma_t^{[-U_1, U]}} \left\{ \frac{1}{2} (1 + 2\mu) \mu (Lf)^2 + 2\mu(\check{X}f)Lf + 2(\check{X}f)^2 + \frac{1}{2} (1 + 2\mu) \mu |\nabla f|_{\mathbf{h}}^2 \right\} d\varpi_{\mathbf{h}} dU', \end{aligned} \quad (179a)$$

$$\mathbb{F}_{(\text{Wave})}[f](t, U) := \int_{\mathcal{P}_U^{[0, t]}} \left\{ (1 + \mu) (Lf)^2 + \mu |\nabla f|_{\mathbf{h}}^2 \right\} d\varpi_{\mathbf{h}} dt', \quad (179b)$$

where \hbar and $|\nabla f|_\hbar^2$ are as in section 6.4.3, and on each $\ell_{t',U'}$,

$$d\omega_\hbar := \sqrt{\det \hbar(t', U', x^2, x^3)} dx^2 dx^3, \quad (180)$$

is the canonical volume form induced by \hbar . In particular, the determinant in (180) is taken relative to the coordinates (x^2, x^3) on $\ell_{t',U'}$.

Remark 6.2 (μ weights and error terms). Note that one L -derivative involving term on RHS (179b) lacks a μ -weight. This means that $\mathbb{F}_{(\text{Wave})}[f](t, U)$ is strongly coercive in the L -derivatives even in regions where μ is small. This is crucial for controlling L -derivative-involving energy estimate error terms that lack μ -weights. In contrast, all terms on RHS (179a) and (179b) involving the ∇ -derivatives of f have a μ -weight. Hence, to control dangerous ∇ -derivative-involving energy estimate error terms that lack μ -weights, one must rely on the special structure afforded by the coercive spacetime integral described in remark 6.3.

Using (172) with $Z := \check{T}$, one can prove the following lemma (see, e.g. [96, proposition 3.5]), which forms the starting point for the L^2 -analysis of solutions to the wave equations (146a).

Lemma 6.1 (energy-null flux identity for wave equations). *For scalar functions f on $\mathcal{M}_{[0, T_{\text{Boot}}), [-U_1, U_3]}$ that vanish along¹⁶ \mathcal{P}_{-U_1} , the following energy-null flux identity holds for $(t, U) \in [0, T_{\text{Boot}}) \times [-U_1, U_3]$:*

$$\begin{aligned} \mathbb{E}_{(\text{Wave})}[f](t, U) + \mathbb{F}_{(\text{Wave})}[f](t, U) &= \mathbb{E}_{(\text{Wave})}[f](0, U) \\ &\quad - \frac{1}{2} \int_{\mathcal{M}_{[0,t), [-U_1, U]}} \mu(\mathbf{h}^{-1})^{\alpha\beta} (\mathbf{h}^{-1})^{\kappa\lambda} \mathbf{Q}_{\alpha\kappa}(\check{T}) \boldsymbol{\pi}_{\beta\lambda} d\omega_\hbar dU' dt' \\ &\quad - \int_{\mathcal{M}_{[0,t), [-U_1, U]}} (\check{T}f) (\mu \square_{\mathbf{h}(\check{Y})} f) d\omega_\hbar dU' dt'. \end{aligned} \quad (181)$$

Remark 6.3 (the coercive spacetime integral $\mathbb{K}[f](t, U)$). Upon decomposing the \mathbf{Q} -involving integral on RHS (181) relative to the vectorfields in \mathcal{Z} , one finds that it contains the following term, which, after moving it to LHS (181), takes the following form:

$$\mathbb{K}[f](t, U) = \frac{1}{2} \int_{\mathcal{M}_{[0,t), [-U_1, U]}} [L\mu]_- |\nabla f|_\hbar^2 d\omega_\hbar dU' dt'. \quad (182)$$

In (182), $[z]_- := \max\{0, -z\}$ denotes the negative part of z . Importantly, the discussion in section 6.7 shows that $L\mu$ is negative in regions close to the shock (i.e. in regions where μ is small) and thus in such regions, $\mathbb{K}[f](t, U)$, which we again emphasize has the favorable sign (182) when moved to LHS (181), yields *non- μ -weighted spacetime L^2 control* of $|\nabla f|_\hbar^2$. As we mentioned above, the control afforded by this term is crucial for controlling error terms in the energy estimates that depend on ∇f but lack μ -weights.

6.8.2. Energies and null fluxes for transport equations. We now define energies and null fluxes for transport equations of the form $\mathbf{B}f = \dots$, where the vectorfield \mathbf{B} is defined in (143a). The energies and null fluxes are relevant for controlling solutions to the transport equations from theorem 4.1.

¹⁶ This vanishing assumption holds under the compact support assumptions on the initial data stated in section 6.5.

Specifically, motivated by the identities (145) and (160), we define the following energies and null fluxes.

Definition 6.7 (transport equation energies and null-fluxes).

$$\mathbb{E}_{(\text{Transport})}[f](t, U) := \int_{\Sigma_t^U} \mu f^2 d\varpi_{\#} dU', \quad (183a)$$

$$\mathbb{F}_{(\text{Transport})}[f](t, U) := \int_{\mathcal{P}_U^{[0,t)}} [-\mathbf{n}\mathbf{m}(\mathbf{B}, L)] f^2 d\varpi_{\#} dt'. \quad (183b)$$

Note that the discussion below (160) implies that the factor $-\mathbf{n}\mathbf{m}(\mathbf{B}, L)$ on RHS (183b) is positive, which of course is important for the coercivity of $\mathbb{F}_{(\text{Transport})}[f](t, U)$. For the same reasons given in section 6.8.1, it is crucially important that RHS (183b) lacks a μ -weight.

By applying the divergence theorem to the vectorfield $f^2 \mathbf{B}$ on $\mathcal{M}_{[0,t), [-U_1, U]}$, using (145) and (160), and using that in geometric coordinates we have (see, e.g. [92, corollary 3.47]):

$$\sqrt{|\det \mathbf{h}|} dx^2 dx^3 dU' dt' = \mu d\varpi_{\#} dU' dt', \quad (184)$$

one can prove the following lemma, which forms the starting point for the L^2 -analysis of solutions to the transport equations (146b) and (146c).

Lemma 6.2 (energy-null flux identity for transport equations). *For scalar functions f on $\mathcal{M}_{[0, T_{\text{Boot}}), [-U_1, U_3]}$ that vanish along \mathcal{P}_{-U_1} (footnote 16 applies here as well), the following energy-null flux identity holds for $(t, U) \in [0, T_{\text{Boot}}) \times [-U_1, U_3]$:*

$$\begin{aligned} \mathbb{E}_{(\text{Transport})}[f](t, U) + \mathbb{F}_{(\text{Transport})}[f](t, U) &= \mathbb{E}_{(\text{Transport})}[f](0, U) \\ &+ \int_{\mathcal{M}_{[0,t), [-U_1, U]}} \mu(\mathbf{D}_\alpha \mathbf{B}^\alpha) f^2 d\varpi_{\#} dU' dt' \\ &+ \int_{\mathcal{M}_{[0,t), [-U_1, U]}} 2f(\mu \mathbf{B} f) d\varpi_{\#} dU' dt'. \end{aligned} \quad (185)$$

6.8.3. Commutator method and the shock-driving terms. To derive suitable energy estimates, one commutes μ -weighted¹⁷ versions of the equations of theorem 4.1 with the elements of the \mathcal{P}_U -tangent subset \mathcal{P} defined in (158) and then applies the energy-null flux identities of lemmas 6.1 and 6.2 and the elliptic estimates described in section 6.9.2. In particular, it has been understood since [68] that to close the energy estimates in a shock formation problem, one does not need to commute the equations with the vectorfield \check{X} , which is transversal to the characteristics. We stress that it does not seem possible to close the estimates by commuting the equations with the Minkowski-rectangular partial derivatives ∂_α ; such an approach would generate error terms that we do not know how to control.

Although one must commute with all the elements of \mathcal{P} to close the energy estimates, the most difficult terms arise from commuting the wave equation (146a) with a string of elements of the $\ell_{t,U}$ subset \mathcal{Y} defined in (158). Schematically, for $\Psi \in \vec{\Psi}$, with \mathcal{Y}^N denoting an order N string of elements of \mathcal{Y} , we have:

$$\mu \square_{\mathbf{h}(\vec{\Psi})} \mathcal{Y}^N \Psi = (\check{X} \Psi) \cdot \mathcal{Y}^N \text{tr}_{\#} \chi + \mu f(\vec{\Psi}) \mathcal{Y}^N \mathcal{C} + \mu f(\vec{\Psi}) \mathcal{Y}^N \mathcal{D} + \dots, \quad (186)$$

¹⁷ The μ -weights allow one to avoid problematic error terms, and they are compatible with the μ -weights in lemmas 6.1 and 6.2.

where χ is the null second fundamental form of \mathcal{P}_U with respect to \mathbf{h} , $\text{tr}_{\mathbf{h}}\chi$ denotes its trace with respect to \mathbf{h} , and \dots denotes easier error terms.

The $\mathcal{Y}^N\mathcal{C}$ - and $\mathcal{Y}^N\mathcal{D}$ -involving terms on RHS (186) clearly arise from the first term on RHS (146a), and we will discuss how to control them in section 6.9.2. We refer to the proofs of [96, lemmas 2.18 and 4.2] for details on the origin of the commutator term $(\check{X}\Psi) \cdot \mathcal{Y}^N \text{tr}_{\mathbf{h}}\chi$, which we describe how to control just below; this term would lead to the loss of a derivative at the top order if handled in a naive fashion.

Before proceeding, we first highlight that we have relegated all null-form-involving terms on RHS (146a) to the terms \dots on RHS (186). The reason is that one can show (see [66, lemma 2.56]) that \mathbf{h} -null forms (see definition 4.6) enjoy the following good decomposition, expressed schematically, relative to the frame \mathcal{Z} (158), where P schematically denotes elements of \mathcal{P} :

$$\mu \mathfrak{Q}(\partial\phi, \partial\psi) = \mu P\phi \cdot P\psi + \check{X}\phi \cdot P\psi + \check{X}\psi \cdot P\phi. \quad (187)$$

The key point is that there are not any terms on RHS (187) that are proportional to $\check{X}\psi \cdot \check{X}\phi$; signature considerations imply that such terms, if present, would be multiplied by a dangerous factor of $\frac{1}{\mu}$. Such a factor would become singular as $\mu \downarrow 0$, which would spoil the philosophy of proving that the solution remains quite regular in geometric coordinates, even as μ vanishes. A decomposition similar to (187) also holds for the up-to-top-order derivatives of the μ -weighted \mathbf{h} -null forms. We stress that the absence of terms proportional to $\check{X}\psi \cdot \check{X}\phi$ on RHS (187) is a statement about the *full nonlinear structure* of the \mathbf{h} -null forms from definition 4.6. In particular, this notion of null form is tied to the acoustical metric \mathbf{h} and is stronger than the ‘classic null condition’ established by Klainerman in [55], which is tied to the Minkowski metric and is indifferent to the structure of most cubic nonlinearities.

Remark 6.4 (the shock-driving terms are hidden in $\square_{\mathbf{h}(\vec{\Psi})}\Psi$). The upshot is that all terms in (146a) that drive the formation of the shock are hidden in the covariant wave operator term $\square_{\mathbf{h}(\vec{\Psi})}\Psi$ on the LHS. More precisely, the shock-driving terms are semilinear Riccati-type terms of type $X\Psi \cdot X\Psi$ (see section 3.5) that become visible if one first expands LHS (146a) relative to the Minkowski-rectangular coordinates and then decomposes the semilinear terms with respect to the frame \mathcal{Z} from (158).

In view of the above discussion, it follows that the goal is to show that all terms on RHS (186) are error terms that do not interfere with the shock formation processes. As we have alluded to earlier, this is indeed possible except at the high derivative levels, where the difficult regularity theory of the eikonal function (in particular, the difficult regularity theory of χ) leads to singular high-order estimates; see section 6.9.3.

We now explain how to control the first product on RHS (186). One starts with the Raychaudhuri equation [82], which plays an important role in mathematical general relativity, and which takes the following form in the present context:

$$L\text{tr}_{\mathbf{h}}\chi = \frac{L\mu}{\mu} \text{tr}_{\mathbf{h}}\chi - \mathbf{Ric}_{LL} - |\chi|_{\mathbf{h}}^2. \quad (188)$$

In (188), \mathbf{Ric} is the Ricci curvature of the acoustical metric \mathbf{h} and $\mathbf{Ric}_{LL} := \mathbf{Ric}_{\alpha\beta}L^\alpha L^\beta$. The difficulty is that since $\mathbf{h} = \mathbf{h}(\vec{\Psi})$, \mathbf{Ric} depends on the second derivatives of $\vec{\Psi}$. This suggests that control of the term $\mathcal{Y}^N \text{tr}_{\mathbf{h}}\chi$ on RHS (186) requires control over $N+2$ derivatives of $\vec{\Psi}$, which is one too many derivatives to be compatible with the regularity for $\vec{\Psi}$ afforded by

energy estimates for the wave equation (186), i.e. one has to worry about losing a derivative at the top-order.

Fortunately, based on ideas originating in [33, 56], there is a way to avoid the loss of derivatives by combining the special structure of \mathbf{Ric}_{LL} with the special structures of the equations of theorem 4.1. We will briefly explain the main ideas of the argument. First, one decomposes \mathbf{Ric}_{LL} to deduce an algebraic identity that can schematically be expressed as:

$$\mu \mathbf{Ric}_{LL} = L \left\{ \mu f(\vec{\Psi}) \cdot L \vec{\Psi} + \mu f(\vec{\Psi}) \cdot \nabla \vec{\Psi} \right\} + \mu f(\vec{\Psi}) \cdot \Delta \vec{\Psi} + \dots, \quad (189)$$

where f schematically denotes a smooth function that is allowed to vary from line to line and \dots denotes lower order terms that depend on the first derivatives of $\vec{\Psi}$. Next, one uses (162) and (165) to replace the term $\mu f(\vec{\Psi}) \cdot \Delta \vec{\Psi}$ on RHS (189) with a perfect L -derivative term plus $\mu \square_{h(\vec{\Psi})} \vec{\Psi}$ plus lower-order terms. One then uses the relativistic Euler wave equation (146a) to replace $\mu \square_{h(\vec{\Psi})} \vec{\Psi}$ with the inhomogeneous terms on RHS (146a). Combining with (188) and bringing all perfect L -derivative terms to the left, we deduce a transport equation that can schematically be expressed as:

$$L \left\{ \mu \text{tr}_{\tilde{h}} \chi + f(\vec{\Psi}) \check{X} \vec{\Psi} + \mu f(\vec{\Psi}) L \Psi + \mu f(\vec{\Psi}) \nabla \Psi \right\} = -\mu |\chi|_{\tilde{h}}^2 + \mu f(\vec{\Psi}) \mathcal{C} + \mu f(\vec{\Psi}) \mathcal{D} + \dots, \quad (190)$$

where \dots denotes easier error terms that involve ϖ , S , and the first derivatives of $\vec{\Psi}$. In particular, RHS (190) does not involve the second derivatives of $\vec{\Psi}$, which is its main advantage compared to equation (188).

Similar results hold for the commuted equations, i.e. given any order N string of $\ell_{t,U}$ -tangent vectorfields \mathcal{Y}^N , with $(\mathcal{Y}^N)\mathcal{X}$ denoting the ‘fully modified’ order N version of $\text{tr}_{\tilde{h}} \chi$ defined by:

$$(\mathcal{Y}^N)\mathcal{X} := \mu \mathcal{Y}^N \text{tr}_{\tilde{h}} \chi + \mathcal{Y}^N \left\{ f(\vec{\Psi}) \check{X} \vec{\Psi} + \mu f(\vec{\Psi}) L \Psi + \mu f(\vec{\Psi}) \nabla \Psi \right\}, \quad (191)$$

one can derive a transport equation of the following schematic form (see, e.g. the proof of [92, proposition 11.10] for the main ideas behind the proof):

$$L^{(\mathcal{Y}^N)}\mathcal{X} = -\mu \mathcal{Y}^N (|\chi|_{\tilde{h}}^2) + \mu f(\vec{\Psi}) \mathcal{Y}^N \mathcal{C} + \mu f(\vec{\Psi}) \mathcal{Y}^N \mathcal{D} + \dots. \quad (192)$$

Equation (192) is an order N version of equation (190) that allows one to control $(\mathcal{Y}^N)\mathcal{X}$ without derivative loss, assuming that one can adequately control RHS (192). In section 6.9.2, we will explain how to control the $\mathcal{Y}^N \mathcal{C}$ - and $\mathcal{Y}^N \mathcal{D}$ -involving terms on RHS (192). One remaining difficulty is that the term $\mu \mathcal{Y}^N (|\chi|_{\tilde{h}}^2)$ on RHS (192) involves the full tensor χ , rather than just the trace part featured on LHS (192). However, one can control this term using a strategy that goes back to [33, 56]. Specifically, one can use elliptic estimates on the two-dimensional surfaces $\ell_{t,U}$, based on the Codazzi equations from geometry, to control the $\ell_{t,U}$ -tangent derivatives of $|\chi|_{\tilde{h}}^2$ in terms of the $\ell_{t,U}$ -tangent derivatives of $\text{tr}_{\tilde{h}} \chi$ plus error terms with an admissible amount of regularity; similar results hold up to top-order, and we refer, for example, to the proof [92, lemma 20.20] for the main ideas behind the analysis. In total, the approach described above allows one to control the product $(\check{X} \Psi) \cdot \mathcal{Y}^N \text{tr}_{\tilde{h}} \chi$ on RHS (186) in appropriate L^2 -based Sobolev spaces. However, the argument introduces a difficult factor of $\frac{1}{\mu}$ into the top-order estimates, as we explain in section 6.9.

6.9. Energy estimates

In this section, we will describe some features of the proof of the energy estimates, which are the most technical and difficult aspect of studying shock formation.

6.9.1. Transport variable energy estimates. We now explain how to derive energy estimates for (ϖ, S) , which solve the transport equation (146b). These estimates are relatively easy to derive. Let \mathcal{P}^N denote an arbitrary order N string of elements of the \mathcal{P}_U -tangent vectorfields \mathcal{P} defined in (158). One multiplies equation (146b) by μ , commutes with \mathcal{P}^N for $N \leq N_{\text{top}}$, and uses the bootstrap assumptions and the transport equation energy identity (185) to deduce that for $(t, U) \in [0, T_{\text{Boot}}) \times [-U_1, U_3]$, we have:

$$\begin{aligned} & \mathbb{E}_{(\text{Transport})}[\mathcal{P}^{\leq N}(\varpi, S)](t, U) + \mathbb{F}_{(\text{Transport})}[\mathcal{P}^{\leq N}(\varpi, S)](t, U) \\ & \lesssim \text{data} + \int_{\mathcal{M}_{[0, t), [-U_1, U]}} \left\{ \mu\text{-regular } \vec{\Psi} \text{ terms} \right\} d\varpi \# dU' dt' \\ & \quad + \int_{-U_1}^U \mathbb{F}_{(\text{Transport})}[\mathcal{P}^{\leq N}(\varpi, S)](t, U') dU'. \end{aligned} \quad (193)$$

From (193) and Grönwall's inequality in U , we deduce that:

$$\begin{aligned} & \mathbb{E}_{(\text{Transport})}[\mathcal{P}^{\leq N_{\text{top}}}(\varpi, S)](t, U) + \mathbb{F}_{(\text{Transport})}[\mathcal{P}^{\leq N_{\text{top}}}(\varpi, S)](t, U) \\ & \lesssim \text{data} \cdot \int_{\mathcal{M}_{[0, t), [-U_1, U]}} \left\{ \mu\text{-regular } \vec{\Psi} \text{ terms} \right\} d\varpi \# dU' dt', \end{aligned} \quad (194)$$

where here and throughout, ‘data’ denotes a small term that depends on the initial data and that measures the perturbation of the data away from simple isentropic plane-symmetric data. The terms ‘ μ -regular $\vec{\Psi}$ terms’ in (194) are easily controllable by the wave energies and null fluxes, and in particular, no singular factor of $\frac{1}{\mu}$ is present in these terms. Hence, the estimate (194) shows that the behavior of $\mathcal{P}^{\leq N_{\text{top}}}(\varpi, S)$ is effectively determined by the behavior of the wave variables $\vec{\Psi}$ (defined in (153)). We clarify that even though there is no singular factor of $\frac{1}{\mu}$ present in the μ -regular $\vec{\Psi}$ terms, these terms can still blow up at the high derivative levels as $\mu \downarrow 0$, for reasons we describe in section 6.9.3. Hence, in view of the coupling between the transport and wave energies shown by (194), we see that the blowup of the $\vec{\Psi}$ -energies can cause the blowup of the (ϖ, S) -energies.

6.9.2. Elliptic-hyperbolic estimates for the vorticity and entropy. Recall that the terms $\mathcal{Y}^N \mathcal{C}$ and $\mathcal{Y}^N \mathcal{D}$ appear on RHSs (186) and (192). To close the energy estimates, we need to control $\mathcal{P}^{\leq N_{\text{top}}}(\mathcal{C}, \mathcal{D})$ in L^2 . These terms would lead to a loss of derivatives at the top-order if handled in a naive fashion. There are several reasons why one cannot control these terms by using only the same transport equation energy methods we used to derive (194). One reason is that definition 4.4 shows that from the point of view of regularity, we have $(\mathcal{C}, \mathcal{D}) \sim \partial(\varpi, S) + \dots$; since the inhomogeneous terms in the transport equation (146b) satisfied by (ϖ, S) suggest (incorrectly, as it fortunately turns out) that $\partial(\varpi, S)$ have at best the same regularity as $\partial^2 \vec{\Psi}$, this is formally inconsistent (from the point of view of regularity) with having $\partial(\varpi, S)$ as a source term in the wave equations (146a) for $\vec{\Psi}$. A second reason is that the transport equations (146c) satisfied by $(\mathcal{C}, \mathcal{D})$ feature the inhomogeneous terms $\mathcal{Q}(\partial \varpi, \partial \vec{\Psi})$ and $\mathcal{Q}(\partial S, \partial \vec{\Psi})$, which depend on the general first derivatives of (ϖ, S) , rather than the special combinations of first

derivatives of (ϖ, S) present in the definitions (140a) and (140b) of $(\mathcal{C}, \mathcal{D})$. That is, due to these inhomogeneous terms, (146c) cannot be treated as a pure transport equation in $(\mathcal{C}, \mathcal{D})$.

To overcome the difficulties highlighted in the previous paragraph, one can treat (146b)–(146d) as a coupled transport-div-curl system that yields L^2 -control of $(\mathcal{C}, \mathcal{D})$ and $(\partial\varpi, \partial S)$, where $(\mathcal{C}, \mathcal{D})$ are controlled with hyperbolic transport energy estimates in the spirit of (194), and the elliptic estimates are used only to handle the inhomogeneous terms $\mathfrak{Q}(\partial\varpi, \partial\vec{\Psi})$ and $\mathfrak{Q}(\partial S, \partial\vec{\Psi})$ on RHS (146c). Similar results hold up to top-order and yield L^2 control over $\mathfrak{D}^{\leq N_{\text{top}}}(\varpi, S)$ and $\mathfrak{P}^{\leq N_{\text{top}}}(\mathcal{C}, \mathcal{D})$. One key difficulty in this argument is that to obtain elliptic estimates on the spacelike hypersurfaces Σ_t , one needs to extract a *spatial* div-curl subsystem from (146b)–(146d); the difficulty is that equations (146b)–(146d) appear to involve *spacetime* div-curl equations. Nevertheless, by splitting various derivative operators into a \mathbf{B} -parallel part and a Σ_t -tangent part, one can extract the desired spatial div-curl subsystem; we refer to the proof of [42, lemma 9.21] for details on this extraction in the context of a local well-posedness argument. We also refer to [68, sections 11.2 and 11.3] for similar but simpler elliptic estimates in the context of shock formation for the non-relativistic 3D compressible Euler equations.

The argument described in the previous paragraph applies when the initial data of the vorticity and entropy are compactly supported, which is the case for the initial data described in section 6.5. The key point is that under compact support, one can avoid *boundary terms* in the elliptic estimates (the boundary terms vanish thanks to the compact support). To treat solutions that are not (spatially) compactly supported, one would need to handle the boundary terms that arise in the elliptic estimates. In the case of non-relativistic 3D compressible Euler equations, based on the special structures found in the formulation of the flow derived in [95], localized spacetime ‘elliptic-hyperbolic’ integral identities for the vorticity and entropy were derived in [3]. The identities of [3] allow one, in the non-relativistic case, to handle the boundary terms. In particular, in [1, 2], we used specialized versions of those identities to study the structure of \mathbf{h} -MGHD for the non-relativistic 3D compressible Euler equations. To extend these results to 3D relativistic Euler solutions without compact support, one would need to derive relativistic analogs of the integral identities from [3].

6.9.3. The wave energy estimate hierarchy. We now discuss the energy estimates for the wave variables $\vec{\Psi}$, defined in (153). As before, let \mathcal{P}^N denote an arbitrary order N string of elements of the \mathcal{P}_U -tangent vectorfields \mathcal{P} defined in (158). The main difficulty in closing the energy estimates in multi-dimensions is that the best estimates we know how to derive allow for the possibility that the top-order energies blow up as $\mu \downarrow 0$. Before proceeding, for notational convenience, we define:

$$\begin{aligned} \mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^N \vec{\Psi} \right] (t, U) &:= \max_{\Psi \in \vec{\Psi}} \mathbb{E}_{(\text{Wave})} [\mathcal{P}^N \Psi] (t, U), \\ \mathbb{F}_{(\text{Wave})} \left[\mathcal{P}^N \vec{\Psi} \right] (t, U) &:= \max_{\Psi \in \vec{\Psi}} \mathbb{F}_{(\text{Wave})} [\mathcal{P}^N \Psi] (t, U), \\ \mathbb{K} \left[\mathcal{P}^N \vec{\Psi} \right] (t, U) &:= \max_{\Psi \in \vec{\Psi}} \mathbb{K} [\mathcal{P}^N \Psi] (t, U), \end{aligned} \quad (195)$$

where $\mathbb{E}_{(\text{Wave})}$ and $\mathbb{F}_{(\text{Wave})}$ are as in definition 6.6 and \mathbb{K} is as in remark 6.3.

The estimates for the wave energies and null fluxes take the following hierarchical form on $(t, U) \in [0, T_{\text{Boot}}) \times [-U_1, U_3]$ (see, e.g. [66, proposition 14.1] for complete proofs in the case of the 2D non-relativistic compressible Euler equations with vorticity):

$$\mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}} \vec{\Psi} \right] (t, U) + \mathbb{F}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}} \vec{\Psi} \right] (t, U) + \mathbb{K} \left[\mathcal{P}^{N_{\text{top}}} \vec{\Psi} \right] (t, U)$$

$$\leq \text{data} \cdot \mu_{\star}^{-A}(t, U), \quad (196a)$$

$$\begin{aligned} & \mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) + \mathbb{F}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) + \mathbb{K} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) \\ & \leq \text{data} \cdot \mu_{\star}^{-(A-2)}(t, U), \end{aligned} \quad (196b)$$

$$\begin{aligned} & \mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}-2} \vec{\Psi} \right] (t, U) + \mathbb{F}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}-2} \vec{\Psi} \right] (t, U) + \mathbb{K} \left[\mathcal{P}^{N_{\text{top}}-2} \vec{\Psi} \right] (t, U) \\ & \leq \text{data} \cdot \mu_{\star}^{-(A-4)}(t, U), \end{aligned} \quad (196c)$$

$$\vdots$$

$$\begin{aligned} & \mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^{[1, \frac{A}{2}]} \vec{\Psi} \right] (t, U) + \mathbb{F}_{(\text{Wave})} \left[\mathcal{P}^{[1, \frac{A}{2}]} \vec{\Psi} \right] (t, U) + \mathbb{K} \left[\mathcal{P}^{[1, \frac{A}{2}]} \vec{\Psi} \right] (t, U) \\ & \leq \text{data}, \end{aligned} \quad (196d)$$

where as before, ‘data’ denotes a small term that depends on the initial data and that measures the perturbation of the data away from simple isentropic plane-symmetric data,

$$\mu_{\star}(t, U) := \min \left\{ 1, \min_{\Sigma_t^{[-U_1, U]}} \mu \right\}, \quad (197)$$

$\mathcal{P}^{[1, \frac{A}{2}]}$ denotes an arbitrary string of elements of the \mathcal{P}_U -tangent vectorfields \mathcal{P} of order in between¹⁸ 1 and $\frac{A}{2}$, and for reasons described below, $A \gg 1$ is a universal constant (independent of N_{top} , the initial data, and the equation of state).

Energy estimates in the spirit of (196a)–(196d), which are singular at the high derivative levels, are the only kinds of energy estimates that are known in the context of multi-dimensional shock formation. One might be concerned that the high-order energies are allowed to blow up when μ vanishes, as this seems to be in conflict with the philosophy described in section 6.3, namely that the solution should look regular in geometric coordinates, all the way up to the shock. However, the full hierarchy (196a)–(196d) shows that the wave energies become less singular by two powers of μ_{\star}^{-1} with each level of descent below the top, until one reaches the level (196d) at which the energies remain bounded. In particular, the boundedness of the mid-order and below geometric energies, as shown by the estimate (196d), capture the sense in which the solution remains regular relative to the geometric coordinates. Due to coupling (see, e.g. (194)), these estimates imply that the fluid variables $\mathcal{C}, \mathcal{D}, \varpi, S$ obey a similar energy hierarchy (omitted here for brevity, but see [68] for a detailed statement and proof of the hierarchy in the case of the non-relativistic 3D compressible Euler equations) featuring related but distinct blowup-rates, which are nonetheless compatible with proving (196a)–(196d). We highlight the following key point:

¹⁸ One can close the proof without deriving energy estimates in the case $N = 0$. This is convenient because the order 0 wave energies can initially be large, stemming from the initial largeness of the data for $\mathcal{X}\mathcal{R}_{(+)}$ (largeness that is present even for the simple isentropic plane-symmetric solutions treated in theorem 3.1). In contrast, the energies in (196a)–(196d) are initially small (in fact, they vanish for the solutions in theorem 3.1).

The non-singular estimates (196d) are what allow one to improve, through Sobolev embedding, a smallness assumption on ‘data’, and derivative-losing transport-equation-type estimates, the L^∞ bootstrap assumptions described in section 6.6.

The singular top-order wave equation energy estimate (196a) stems from the following integral inequality, whose proof we describe in section 6.9.4:

$$\begin{aligned} & \mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}} \vec{\Psi} \right] (t, U) + \mathbb{F}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}} \vec{\Psi} \right] (t, U) + \mathbb{K} \left[\mathcal{P}^{N_{\text{top}}} \vec{\Psi} \right] (t, U) \\ & \leq \text{data} + A \int_{t'=0}^t \left\| \frac{L\mu}{\mu} \right\|_{L^\infty(\Sigma_{t'}^{[-U_1, U]})} \mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}} \vec{\Psi} \right] (t', U) dt' + \dots \end{aligned} \quad (198)$$

In (198), $A \gg 1$ is the universal positive constant mentioned above and \dots denotes similar or easier¹⁹ error terms. From (198) and Grönwall’s inequality, we deduce that:

$$\begin{aligned} & \mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}} \Psi \right] (t, U) + \mathbb{F}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}} \Psi \right] (t, U) + \mathbb{K} \left[\mathcal{P}^{N_{\text{top}}} \Psi \right] (t, U) \\ & \leq \text{data} \cdot \exp \left(A \int_{t'=0}^t \left\| \frac{L\mu}{\mu} \right\|_{L^\infty(\Sigma_{t'}^{[-U_1, U]})} dt' \right) + \dots \end{aligned} \quad (199)$$

An argument based on precise, refined versions of (167) and (168) yields the following crucial bound, which can be proved using the same arguments given in [96, section 10] (the proof is trivial if one ignores the terms \dots in (167) and (168)):

$$\exp \left(\int_{t'=0}^t \left\| \frac{L\mu}{\mu} \right\|_{L^\infty(\Sigma_{t'}^{[-U_1, U]})} dt' \right) \leq C\mu_*^{-1}(t, U) + C, \quad (200)$$

where μ_* is defined in (197). Combining (199) and (200), we conclude (196a).

The less singular below-top-order estimate (196b) stems from the following integral inequality, whose proof we describe in section 6.9.5:

$$\begin{aligned} & \mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) + \mathbb{F}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) + \mathbb{K} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) \\ & \leq \text{data} + C \int_{t'=0}^t \frac{1}{\mu_*^{1/2}(t', U)} \mathbb{E}_{(\text{Wave})}^{1/2} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t', U) \\ & \quad \times \int_{t''=0}^{t'} \frac{1}{\mu_*^{1/2}(t'', U)} \mathbb{E}_{(\text{Wave})}^{1/2} \left[\mathcal{P}^{N_{\text{top}}} \vec{\Psi} \right] (t'', U) dt'' dt' + \dots \end{aligned} \quad (201)$$

Note that RHS (201) involves a coupling between the top-order energies and the just-below-top-order energies. The actual proof of (196b) involves a delicate Grönwall argument that is coupled to the proof of the top-order estimate (196a) as well as the following estimates, which

¹⁹ In reality, some of the ‘easier’ error terms also require substantial effort to treat. For example, to handle some of the error terms coming from the $(\mathcal{I})\pi_{\beta\lambda}$ -involving integral on RHS (181), one needs sharp information about the behavior of μ and its derivatives in regions where μ is small; see, for example, [66, lemma 14.10].

hold for numbers $B > 1$ and which can be proved by using arguments similar to the ones needed to prove (200):

$$\int_{t'=0}^t \mu_*^{-B} (t', U) dt' \leq C \mu_*^{1-B} (t, U), \quad (202)$$

$$\int_{t'=0}^t \mu_*^{-\frac{9}{10}} (t', U) dt' \leq C. \quad (203)$$

We stress that (202) is a ‘quasilinear version’ of the following simple estimate, which is valid for $t \in [0, 1]$: $\int_{t'=0}^t (1-t')^{-B} dt' \leq C(1-t)^{1-B}$. Similarly, (203) is a quasilinear version of the estimate $\int_{t'=0}^t (1-t')^{-\frac{9}{10}} dt' \leq C$. The proof of (202) and (203) fundamentally relies on the fact that μ vanishes linearly in t , as we described at the end of section 6.7. Here, we will just use (201) and (202) to explain why the desired estimates (196a) and (196b) are *consistent* with respect to powers of $\mu_*^{-1}(t, U)$. To confirm the consistency, we plug the estimates (196a) and (196b) into the integrals on RHS (201) to deduce:

$$\begin{aligned} & \mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) + \mathbb{F}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) + \mathbb{K} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) \\ & \leq \text{data} + \text{data} \cdot \int_{t'=0}^t \mu_*^{-\frac{A}{2} + \frac{1}{2}} (t', U) \int_{t''=0}^{t'} \mu_*^{-\frac{A}{2} - \frac{1}{2}} (t'', U) dt'' dt' + \dots \end{aligned} \quad (204)$$

Using (204) and twice using (202), we deduce that:

$$\begin{aligned} & \mathbb{E}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) + \mathbb{F}_{(\text{Wave})} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) + \mathbb{K} \left[\mathcal{P}^{N_{\text{top}}-1} \vec{\Psi} \right] (t, U) \\ & \leq \text{data} \cdot \mu_*^{-(A-2)} (t, U), \end{aligned} \quad (205)$$

which is indeed consistent with the desired estimate (196b). This concludes our proof sketch of (196b).

Using the same strategy, one can continue the descent in the energy estimate hierarchy, gaining two powers of μ_* with each level of descent until one finally can use (203) to reach the level (196d), where the energies remain uniformly bounded.

6.9.4. Discussion of the proof of the top-order inequality (198). We now sketch some key ideas behind the proof of (198). We focus on the most difficult case in which the operator $\mathcal{P}^{N_{\text{top}}}$ in (198) is of the form²⁰ $\mathcal{P}^{N_{\text{top}}} = \mathcal{Y}^{N_{\text{top}}}$, where $\mathcal{Y}^{N_{\text{top}}}$ is a string of elements of the $\ell_{t,U}$ -tangent subset \mathcal{Y} defined in (158). Using the wave equation energy-null flux identity (181), the commuted wave equation (186), the transport/elliptic estimates for the vorticity and entropy described in sections 6.9.1 and 6.9.2, and the coerciveness of the spacetime integral from (182), one can show that for $\Psi \in \vec{\Psi}$ (which is defined in (153)), the following identity holds for $(t, U) \in [0, T_{\text{Boot}}) \times [-U_1, U_3]$:

$$\begin{aligned} & \mathbb{E}_{(\text{Wave})} [\mathcal{Y}^{N_{\text{top}}} \Psi] (t, U) + \mathbb{F}_{(\text{Wave})} [\mathcal{Y}^{N_{\text{top}}} \Psi] (t, U) + \mathbb{K} [\mathcal{Y}^{N_{\text{top}}} \Psi] (t, U) \\ & = \text{data} + \int_{\mathcal{M}_{[0,t], [-U_1, U]}} (\check{T} \mathcal{Y}^{N_{\text{top}}} \Psi) \cdot (\check{X} \Psi) \cdot \mathcal{Y}^{N_{\text{top}}} \text{tr}_{\mathcal{Y}} \chi d\varpi_{\mathcal{Y}} dU' dt' \end{aligned}$$

²⁰ The case in which $\mathcal{P}^{N_{\text{top}}}$ contains an L -differentiation is much easier because in that case, one can directly use equation (188) to estimate the L -derivatives of $\text{tr}_{\mathcal{Y}} \chi$.

$$\begin{aligned}
& + \int_{\mathcal{M}_{[0,t),[-U_1,U]}} \mu(\check{T}\mathcal{Y}^{N_{\text{top}}}\Psi) \cdot \mathcal{Y}^{N_{\text{top}}}\mathcal{C} d\varpi_{\hbar} dU' dt' \\
& + \int_{\mathcal{M}_{[0,t),[-U_1,U]}} \mu(\check{T}\mathcal{Y}^{N_{\text{top}}}\Psi) \cdot \mathcal{Y}^{N_{\text{top}}}\mathcal{D} d\varpi_{\hbar} dU' dt' + \dots, \tag{206}
\end{aligned}$$

where \dots denotes easier error terms.

We first consider the difficult case $\Psi = \mathcal{R}_{(+)}$. We recall the definition (174) of \check{T} . We consider the part of the integral $\int_{\mathcal{M}_{[0,t),[-U_1,U]}} (\check{T}\mathcal{Y}^{N_{\text{top}}}\Psi) \cdot (\check{X}\Psi) \cdot \mathcal{Y}^{N_{\text{top}}}\text{tr}_{\hbar}\chi d\varpi_{\hbar} dU' dt'$ on RHS (206) that is generated by the $2\check{X}$ term in (174). Using (191), and using (165) to replace $\check{X}\mathcal{R}_{(+)}$ with $L\mu$ up to error terms, we can express the $2\check{X}$ -involving portion of the integral under consideration as follows, where we highlight the crucial factors of $\frac{1}{\mu}$ on RHS (207), and for simplicity, we have ignored the factors of $f(\vec{\Psi})$ on RHS (191):

$$\begin{aligned}
& 2 \int_{\mathcal{M}_{[0,t),[-U_1,U]}} (\check{X}\mathcal{Y}^{N_{\text{top}}}\mathcal{R}_{(+)}) \cdot (\check{X}\mathcal{R}_{(+)}) \cdot \mathcal{Y}^{N_{\text{top}}}\text{tr}_{\hbar}\chi d\varpi_{\hbar} dU' dt' \\
& = -2 \int_{\mathcal{M}_{[0,t),[-U_1,U]}} \frac{L\mu}{\mu} \cdot (\check{X}\mathcal{Y}^{N_{\text{top}}}\mathcal{R}_{(+)}) \cdot (\check{X}\mathcal{Y}^{N_{\text{top}}}\mathcal{R}_{(+)}) d\varpi_{\hbar} dU' dt' \\
& + 2 \int_{\mathcal{M}_{[0,t),[-U_1,U]}} \frac{L\mu}{\mu} \cdot (\check{X}\mathcal{Y}^{N_{\text{top}}}\mathcal{R}_{(+)}) \cdot (\mathcal{Y}^{N_{\text{top}}}\mathcal{X}) d\varpi_{\hbar} dU' dt' + \dots \tag{207}
\end{aligned}$$

From (179a), it follows that the first integral on RHS (207) is bounded by the A -multiplied integral on RHS (198) (this integral contributes a ‘portion of’ the ‘ A ’). The $(\mathcal{Y}^{N_{\text{top}}})\mathcal{X}$ -involving integral on RHS (207) can be bounded by a similar error term involving a double time-integral, where one of the time integrations comes from integrating (192); we refer to the proof of [96, proposition 14.2] for the details. The part of the first integral on RHS (206) that is generated by the $(1+2\mu)L$ term in (174) can be handled through an argument involving integration by parts in L , which leads to difficult critical-strength boundary terms that make a similar contribution to the blowup of the top-order energies as the A -multiplied integral on RHS (198); the same arguments given in the proof of [96, proposition 14.2] can be used to handle these terms. The $\mathcal{Y}^{N_{\text{top}}}\mathcal{C}$ - and $\mathcal{Y}^{N_{\text{top}}}\mathcal{D}$ -involving integrals on RHS (206) are easy to handle, thanks to the helpful factor of μ in the integrand and the strategy for bounding $\mathcal{Y}^{N_{\text{top}}}\mathcal{C}$ and $\mathcal{Y}^{N_{\text{top}}}\mathcal{D}$ via transport-div-curl estimates that we described in section 6.9.2.

We now consider the remaining cases, namely $\Psi \in \{\mathcal{R}_{(-)}, s, \mathbf{u}^2, \mathbf{u}^3\}$. The identity (206) holds in these cases too. However, the factor of $\check{X}\Psi$ in the identity is now very small in L^∞ in these cases, since we are considering perturbations of simple isentropic plane-symmetric solutions (for which these quantities all vanish). Hence, due to the smallness, all error terms on RHS (206) are much easier to treat and can in fact be relegated to the error terms ‘ \dots ’ on RHS (198); see [66, page 154] for further details.

6.9.5. Discussion of the proof of the just-below-top-order inequality (201). We now sketch some key ideas behind the proof of the just-below-top-order inequality (201). First, one uses the identity (206), which, for any $\Psi \in \vec{\Psi}$, holds with $\mathcal{P}^{N_{\text{top}}-1}$ in the role of $\mathcal{Y}^{N_{\text{top}}}$. The key step, which is the one that is different compared to the proof of (198), is that one bounds the error integral on LHS (207), namely

$$2 \int_{\mathcal{M}_{[0,t),[-U_1,U]}} (\check{T}\mathcal{P}^{N_{\text{top}}-1}\Psi) \cdot (\check{X}\Psi) \cdot \mathcal{P}^{N_{\text{top}}-1}\text{tr}_{\hbar}\chi d\varpi_{\hbar} dU' dt', \tag{208}$$

in a different way. First, one uses the bootstrap assumptions to bound the factor $|\check{X}\Psi|$ in (208) by $\lesssim 1$. Next, one uses (174), the Cauchy–Schwarz inequality, and (179a) to bound the integral in (208) by:

$$\lesssim \int_0^t \frac{1}{\mu_*^{1/2}(t', U)} \mathbb{E}_{(\text{Wave})}^{1/2} [\mathcal{P}^{N_{\text{top}}-1} \Psi](t', U) \left\| \mathcal{P}^{N_{\text{top}}-1} \text{tr}_{\vec{\mu}} \chi \right\|_{L^2(\Sigma_{t'}^{[-U_1, U]})} dt' + \dots, \quad (209)$$

where we have used that the \mathcal{P}_U -tangent derivative terms in the energy (179a) contain a μ -weight, while there is an L factor in the definition of \check{T} that does *not* contain a μ weight. Next, to bound the factor $\left\| \mathcal{P}^{N_{\text{top}}-1} \text{tr}_{\vec{\mu}} \chi \right\|_{L^2(\Sigma_{t'}^{[-U_1, U]})}$, one commutes the Raychaudhuri equation (188) with $\mathcal{P}^{N_{\text{top}}-1}$ and uses the schematic relation $\mathbf{Ric}_{LL} = \mathcal{P}^2 \vec{\Psi} + \dots$ to deduce the following evolution equation, schematically depicted: $L \mathcal{P}^{N_{\text{top}}-1} \text{tr}_{\vec{\mu}} \chi = \mathcal{P}^{N_{\text{top}}+1} \vec{\Psi} + \dots$, where \dots denotes lower order error terms. This equation represents a loss of one derivative in the estimates and leads to the coupling between the different order energies described below (201). However, such a loss is permissible below the top-order. Since $Lt = 1$ and $LU = 0$, we can integrate this evolution equation and use (179a) to deduce that:

$$\begin{aligned} \left\| \mathcal{P}^{N_{\text{top}}-1} \text{tr}_{\vec{\mu}} \chi \right\|_{L^2(\Sigma_{t'}^{[-U_1, U]})} &\lesssim \text{data} + \int_0^{t'} \left\| \mathcal{P}^{N_{\text{top}}+1} \vec{\Psi} \right\|_{L^2(\Sigma_{t''}^{[-U_1, U]})} dt'' + \dots \\ &\lesssim \text{data} + \int_0^{t'} \frac{1}{\mu_*^{1/2}(t'', U)} \mathbb{E}_{(\text{Wave})}^{1/2} [\mathcal{P}^{N_{\text{top}}-1} \Psi](t'', U) dt'' + \dots, \end{aligned} \quad (210)$$

where to obtain the last line of (210), we have again used that the \mathcal{P}_U -tangent derivative terms in the energy (179a) contain a μ -weight. Inserting (210) into (209), we conclude (201).

7. Open problems

In this section, we describe various open problems tied to shocks.

- (1) **Prove conjecture 1.** In section 6, we outlined how to achieve this. We are confident that the conjecture can be proved, especially since the formulation of relativistic Euler flow provided by theorem 4.1 is qualitatively similar to the formulation of 3D non-relativistic compressible Euler flow derived in [95], a system for which shock formation results have been derived (see section 5). Nonetheless, there are many non-trivial details that have to be checked, and this would be a good project for someone who wants to learn the field.
- (2) **Prove conjecture 2.** This problem is much more technically demanding than proving conjecture 1. However, for the same reasons mentioned above, we expect that this can be achieved by adapting the methods used in the non-relativistic work [1] to the equations of theorem 4.1.
- (3) **Prove conjecture 3.** This problem is also much more technically demanding than proving conjecture 1. For the same reasons mentioned above, we expect that this can be achieved by adapting the methods that we are using in our forthcoming non-relativistic work [2] to the equations of theorem 4.1.
- (4) **The shock development problem.** Recall that in [32], Christodoulou solved the restricted shock development problem for the relativistic Euler equations and the non-relativistic compressible Euler equations in an arbitrary number of spatial dimensions;

see section 5.6. The general problem, i.e. the shock development problem with vorticity and entropy, remains an outstanding open problem for both systems in two or more spatial dimensions.

- (5) **The global behavior of solutions with shocks.** After the shock development problem (which is local-in-time) is solved, a particularly compelling problem will be to understand the *global*-in-time-and-space behavior of the corresponding weak solutions, at least in a perturbative regime. Among the many outstanding challenges in this problem is that of understanding the long-time behavior of the vorticity and entropy; even in the context of smooth incompressible non-relativistic flows, the long-time behavior of the vorticity is not well understood. We do, however, mention that in the breakthrough works [26, 27], for axisymmetric *incompressible* non-relativistic flows in 3D with a boundary, smooth initial conditions were identified such that the solution (including the vorticity) blows up in finite time in an approximately self-similar fashion; such vorticity blowup, if present in multi-dimensional relativistic Euler flow, would be a serious obstacle to finding a meaningful way to continue the solution weakly past the singularity.
- (6) **Extending the results to the coupled Einstein–Euler system.** It is of great interest to extend the above results to the Einstein–Euler system in multiple spatial dimensions. The expectation is that while the fluid will exhibit shock wave phenomena (as in the uncoupled problem), the gravitational metric will exhibit less singular behavior. As of present, the only known singularity formation result for the coupled system is in the plane-symmetric case [83], in which the dynamics are described by 1 + 1-dimensional hyperbolic PDEs. In [83], it was shown that for many equations of state and a large class of plane-symmetric initial data, the Einstein–Euler solution breaks down in finite time. Although [83] provided heuristic arguments suggesting that the singularity is of shock-type, a precise description of the singularity was not given. The multi-dimensional problem is a difficult PDE problem because surfaces that are null or ‘barely spacelike’ with respect to the acoustical metric are in fact timelike with respect to the gravitational metric; the reason is that speed of sound is slower than the speed of propagation of gravitational waves (at least when the speed of sound is less than unity). A corresponding key difficulty that arises in the context of the PDE estimates is that one must control the spacetime metric on various gravitationally timelike surfaces on which the fluid is singular; this is difficult at the top derivative level because generally, energy estimates for the spacetime metric are not available on surfaces that are timelike with respect to the spacetime metric. In contrast, in multiple wave speed systems such that the shock forms in the *fastest* wave, it is possible to close the top-order energy estimates and thus prove stable shock formation for the coupled problem [93].
- (7) **Extending the results to more complicated multiple speed systems.** It is also of great physical and mathematical interest to extend the above results to other multiple speed systems, such as the equations of compressible magnetohydrodynamics, the GRMHD equations (i.e. the coupling of Einstein’s equations to the equations of relativistic magnetohydrodynamics), the equations of elasticity, the equations of nonlinear electromagnetism, and the equations of crystal optics. A key difficulty of these systems is that their principal symbols are more complicated than wave operators and transport operators. Relatedly, their corresponding geometry is more complicated than Lorentzian geometry, and the characteristics comprise multiple sheets, which can be singular even at the tangent space level. These difficulties are a serious obstacle to implementing the kind of sharp version of nonlinear geometric optics (i.e. eikonal functions) that has been so successfully employed to study shocks for wave equations and fluid equations.

(8) **Implosion singularities.** In the breakthrough works [73, 74], the authors proved implosion singularity formation for some initially C^∞ spherically symmetric solutions to the compressible Euler equations and Navier–Stokes equations under adiabatic equations of state $p = \rho^\gamma$ for $\gamma > 1$, outside of a countable set of γ -values. In [19], the results were extended to allow for all $\gamma > 1$. Implosion singularities are much more severe than shocks in the sense that the density itself blows up in finite time at the center of symmetry (in shock singularities, the density’s gradient blows up but the density itself remains bounded). The proof depends on the detailed structure of the equations and in particular relies on a careful analysis of a phase portrait for self-similar solutions. Hence, it is of interest to decide whether a similar result holds for the relativistic Euler equations. We highlight the important related work [48], in which the authors studied spherically symmetric solutions the Einstein–Euler equations under adiabatic equations of state $p = \varepsilon \rho$ for $0 < \varepsilon \ll 1$ sufficiently small and proved the existence of self-similar solutions that form a naked singularity. See also [47], in which the authors studied solutions to the non-relativistic Euler–Poisson system under adiabatic equations of state $p = \rho^\gamma$ for $\gamma \in (1, \frac{4}{3})$ and showed the existence of spherically symmetric self-similar imploding solutions modeling gravitational collapse, i.e. initially smooth solutions such that the density blows up in finite time.

(9) **Inviscid limits.** In [25], the authors studied the inviscid limit of solutions to the 1D viscous Burgers’ equation all the way up to the time of first shock formation in the inviscid solution. They decomposed the viscous solution into a singular piece and a smoother piece and proved that the viscous solution converges to the singular piece in L^∞ as the viscosity vanishes, where the L^∞ norm is taken over the entire slab of classical existence of the inviscid solution. This is the first result of its type that extends all the way to the time of first singularity formation. Important open problems include extending this result to the 1D compressible Euler equations (where the corresponding viscous equations are the Navier–Stokes equations) and, after that, to multi-dimensions. It would also be of interest to extend the result to the relativistic Euler equations, but as of present, it is not clear if there exist any well-posed relativistic viscous fluid models that suppress the singularity formation while retaining physically desirable features such as causality.

(10) **Rarefaction waves.** In section 5.7, we described the recent works [69, 70] on irrotational rarefaction wave solutions to the 2D compressible Euler equations. It is of interest to eliminate the irrotationality assumption and to extend these results to the relativistic Euler equations, and perhaps even the Einstein–Euler equations.

Data availability statement

No new data were created or analyzed in this study.

Acknowledgments

L A gratefully acknowledges support from an NSF Postdoctoral Fellowship. J S gratefully acknowledges support from NSF Grant # DMS-2054184 and NSF CAREER Grant # DMS-1914537.

ORCID iDs

Leonardo Abbrescia  <https://orcid.org/0000-0001-5188-5265>
 Jared Speck  <https://orcid.org/0009-0007-8698-5515>

References

- [1] Abbrescia L and Speck J 2022 The emergence of the singular boundary from the crease in 3D compressible Euler flow (arXiv:[2207.07107](https://arxiv.org/abs/2207.07107))
- [2] Abbrescia L and Speck J The emergence of the Cauchy horizon from the crease in 3D compressible Euler flow (in preparation)
- [3] Abbrescia L and Speck J 2020 Remarkable localized integral identities for 3D compressible Euler flow and the double-null framework (arXiv:[2003.02815](https://arxiv.org/abs/2003.02815))
- [4] Abbrescia L and Wong W W Y 2020 Global nearly-plane-symmetric solutions to the membrane equation *Forum Math. Pi* **8** e13
- [5] Abbrescia L E and Wong W W Y 2019 Geometric analysis of 1 + 1 dimensional quasilinear wave equations (arXiv:[1912.04692](https://arxiv.org/abs/1912.04692))
- [6] Alinhac S 1989 Existence d'ondes de rarefaction pour des systèmes quasi-linéaires hyperboliques multidimensionnels *Commun. PDE* **14** 173–230
- [7] Alinhac S 1989 Unicité d'ondes de rarefaction pour des systèmes quasi-linéaires hyperboliques multidimensionnels *Indiana Univ. Math. J.* **38** 345–63
- [8] Alinhac S 1999 Blowup of small data solutions for a class of quasilinear wave equations in two space dimensions. II *Acta Math.* **182** 1–23
- [9] Alinhac S 1999 Blowup of small data solutions for a quasilinear wave equation in two space dimensions *Ann. Math.* **149** 97–127
- [10] Alinhac S 2001 The null condition for quasilinear wave equations in two space dimensions. II *Am. J. Math.* **123** 1071–101
- [11] Alinhac S 2002 A minicourse on global existence and blowup of classical solutions to multidimensional quasilinear wave equations *Journées “Équations aux Dérivées Partielles”* (Forges-les-Eaux, 2002) (Univ. Nantes) p 33
- [12] An X, Chen H and Yin S 2020 Low regularity ill-posedness for elastic waves driven by shock formation (arXiv:[2003.03195](https://arxiv.org/abs/2003.03195))
- [13] An X, Chen H and Yin S 2021 Low regularity ill-posedness and shock formation for 3D ideal compressible MHD (arXiv:[2110.10647](https://arxiv.org/abs/2110.10647))
- [14] An X, Chen H and Yin S 2022 Low regularity ill-posedness for non-strictly hyperbolic systems in three dimensions *J. Math. Phys.* **63** 051503
- [15] Athanasiou N, Bayles-Rea T and Zhu S 2023 Development of singularities in the relativistic Euler equations *Trans. Am. Math. Soc.* **376** 2325–72
- [16] Athanasiou N and Zhu S 2021 Formation of singularities for the relativistic Euler equations *J. Differ. Equ.* **284** 284–317
- [17] Bernal A N and Sánchez M 2006 Further results on the smoothability of Cauchy hypersurfaces and Cauchy time functions *Lett. Math. Phys.* **77** 183–97
- [18] Bressan A, Chen G-Q G, Lewicka M and Wang D 2011 *Nonlinear Conservation Laws and Applications* (Springer)
- [19] Buckmaster T, Cao-Labora G and Gómez-Serrano J 2022 Smooth imploding solutions for 3D compressible fluids (arXiv:[2208.09445](https://arxiv.org/abs/2208.09445))
- [20] Buckmaster T, Drivas T D, Shkoller S and Vicol V 2022 Simultaneous development of shocks and cusps for 2D Euler with azimuthal symmetry from smooth data *Ann. PDE* **8** 26
- [21] Buckmaster T and Iyer S 2022 Formation of unstable shocks for 2D isentropic compressible Euler *Commun. Math. Phys.* **389** 197–271
- [22] Buckmaster T, Shkoller S and Vicol V 2022 Formation of shocks for 2D isentropic compressible Euler *Commun. Pure Appl. Math.* **75** 2069–120
- [23] Buckmaster T, Shkoller S and Vicol V 2023 Formation of point shocks for 3D compressible Euler *Commun. Pure Appl. Math.* **76** 2073–191
- [24] Buckmaster T, Shkoller S and Vicol V 2023 Shock formation and vorticity creation for 3D Euler *Commun. Pure Appl. Math.* **76** 1965–2072

[25] Chaturvedi S and Graham C 2023 The inviscid limit of viscous Burgers at nondegenerate shock formation *Ann. PDE* **9** 1

[26] Chen J and Hou T Y 2022 Stable nearly self-similar blowup of the 2D Boussinesq and 3D Euler equations with smooth data I: analysis (arXiv:2210.07191)

[27] Chen J and Hou T Y 2023 Stable nearly self-similar blowup of the 2D Boussinesq and 3D Euler equations with smooth data II: rigorous numerics (arXiv:2305.05660)

[28] Chen S and Dong L 2001 Formation and construction of shock for p -system *Sci. China A* **44** 1139–47

[29] Christodoulou D 2000 *The Action Principle and Partial Differential Equations (Annals of Mathematics Studies, vol 146)* (Princeton University Press)

[30] Christodoulou D 2007 The Euler equations of compressible fluid flow *Bull. Am. Math. Soc.* **44** 581–602

[31] Christodoulou D 2007 *The Formation of Shocks in 3-Dimensional Fluids (EMS Monographs in Mathematics)* (European Mathematical Society (EMS))

[32] Christodoulou D 2019 *The Shock Development Problem (EMS Monographs in Mathematics)* (European Mathematical Society)

[33] Christodoulou D and Klainerman S 1993 *The Global Nonlinear Stability of the Minkowski Space (Princeton Mathematical Series, vol 41)* (Princeton University Press)

[34] Christodoulou D and Lisibach A 2016 Shock development in spherical symmetry *Ann. PDE* **2** 1–246

[35] Christodoulou D and Miao S 2014 *Compressible Flow and Euler's Equations (Surveys of Modern Mathematics vol 9)* (International Press, Higher Education Press)

[36] Christodoulou D and Perez D R 2016 On the formation of shocks of electromagnetic plane waves in non-linear crystals *J. Math. Phys.* **57** 081506

[37] Collot C, Ghoul T-E, Ibrahim S and Masmoudi N 2022 On singularity formation for the two-dimensional unsteady Prandtl system around the axis *J. Eur. Math. Soc.* **24** 3703–800

[38] Collot C, Ghoul T-E and Masmoudi N 2021 Singularities and unsteady separation for the inviscid two-dimensional Prandtl system *Arch. Ration. Mech. Anal.* **240** 1349–430

[39] Collot C, Ghoul T-E and Masmoudi N 2022 Singularity formation for Burgers' equation with transverse viscosity *Ann. Sci. Éc. Norm. Supér.* **55** 1047–133

[40] Dafermos C M 2010 *Hyperbolic Conservation Laws in Continuum Physics (Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol 325)* 3rd edn (Springer)

[41] Disconzi M M, Luo C, Mazzone G and Speck J 2022 Rough sound waves in 3D compressible Euler flow with vorticity *Sel. Math.* **28** 41

[42] Disconzi M M and Speck J 2019 The relativistic Euler equations: remarkable null structures and regularity properties *Ann. Henri Poincaré* **20** 2173–270

[43] Eperon F C, Reall H S and Sbierski J J 2019 Predictability of subluminal and superluminal wave equations *Commun. Math. Phys.* **368** 585–626

[44] Fajman D, Ofner M, Oliynyk T A and Wyatt Z 2023 The stability of relativistic fluids in linearly expanding cosmologies (arXiv:2301.11191)

[45] Fajman D, Oliynyk T A and Wyatt Z 2021 Stabilizing relativistic fluids on spacetimes with non-accelerated expansion *Commun. Math. Phys.* **383** 401–26

[46] Geroch R 1970 Domain of dependence *J. Math. Phys.* **11** 437–49

[47] Guo Y, Hadzic M, Jang J and Schrecker M 2022 Gravitational collapse for polytropic gaseous stars: self-similar solutions *Arch. Ration. Mech. Anal.* **246** 957–1066

[48] Guo Y, Hadzic M and Jang J 2023 Naked singularities in the Einstein-Euler system *Ann. PDE* **9** 4

[49] Guo Y and Tahvildar-Zadeh A S 1999 Formation of singularities in relativistic fluid dynamics and in spherically symmetric plasma dynamics *Nonlinear Partial Differential Equations (Contemporary Mathematics vol 238)* (American Mathematical Society) pp 151–61

[50] Holzegel G, Klainerman S, Speck J and Wong W W-Y 2016 Small-data shock formation in solutions to 3D quasilinear wave equations: an overview *J. Hyperbolic Differ. Equ.* **13** 1–105

[51] Hughes T J R, Kato T and Marsden J E 1976 Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity *Arch. Ration. Mech. Anal.* **63** 273–94

[52] Huicheng Y and Lu Z 2022 Formation and construction of a multidimensional shock wave for the first-order hyperbolic conservation law with smooth initial data *SIAM J. Math. Anal.* **54** 2587–610

[53] Huicheng Y and Lu Z 2022 The shock formation and optimal regularities of the resulting shock curves for 1D scalar conservation laws *Nonlinearity* **35** 954–97

[54] John F 1974 Formation of singularities in one-dimensional nonlinear wave propagation *Commun. Pure Appl. Math.* **27** 377–405

[55] Klainerman S 1984 Long time behaviour of solutions to nonlinear wave equations *Proc. Int. Congress of Mathematicians (Warsaw, 1983)* vol 1, 2 (PWN) pp 1209–15

[56] Klainerman S and Rodnianski I 2003 Improved local well-posedness for quasilinear wave equations in dimension three *Duke Math. J.* **117** 1–124

[57] Klainerman S and Rodnianski I 2005 Causal geometry of Einstein-vacuum spacetimes with finite curvature flux *Invent. Math.* **159** 437–529

[58] Klainerman S and Rodnianski I 2005 The causal structure of microlocalized rough Einstein metrics *Ann. Math.* **161** 1195–243

[59] Klainerman S and Rodnianski I 2005 Rough solutions of the Einstein-vacuum equations *Ann. Math.* **161** 1143–93

[60] Klainerman S and Rodnianski I 2006 A geometric approach to the Littlewood–Paley theory *Geom. Funct. Anal.* **16** 126–63

[61] Klainerman S and Rodnianski I 2008 On the radius of injectivity of null hypersurfaces *J. Am. Math. Soc.* **21** 775–95

[62] Pandya G, Chickering K R and Moreno-Vasquez R C 2021 Asymptotically self-similar shock formation for 1d fractal Burgers equation (arXiv:2105.15128)

[63] Lax P D 1964 Development of singularities of solutions of nonlinear hyperbolic partial differential equations *J. Math. Phys.* **5** 611–3

[64] Lebaud M-P 1994 Description de la formation d'un choc dans le p -système *J. Math. Pures Appl.* **73** 523–65

[65] Lindblad H 1998 Counterexamples to local existence for quasilinear wave equations *Math. Res. Lett.* **5** 605–22

[66] Luk J and Speck J 2018 Shock formation in solutions to the 2D compressible Euler equations in the presence of non-zero vorticity *Invent. Math.* **214** 1–169

[67] Luk J and Speck J 2020 The hidden null structure of the compressible Euler equations and a prelude to applications *J. Hyperbolic Differ. Equ.* **17** 1–60

[68] Luk J and Speck J 2021 The stability of simple plane-symmetric shock formation for 3D compressible Euler flow with vorticity and entropy (arXiv:2107.03426)

[69] Luo T-W and Yu P 2023 On the stability of multi-dimensional rarefaction waves I: the energy estimates (arXiv:2302.09714)

[70] Luo T-W and Yu P 2023 On the stability of multi-dimensional rarefaction waves II: existence of solutions and applications to Riemann problem (arXiv:2305.06308)

[71] Majda A 1981 The existence and stability of multidimensional shock fronts *Bull. Am. Math. Soc.* **4** 342–4

[72] Majda A 1983 The existence of multidimensional shock fronts *Mem. Am. Math. Soc.* **43** v+93

[73] Merle F, Raphaël P, Rodnianski I and Szeftel J 2022 On the implosion of a compressible fluid I: smooth self-similar inviscid profiles *Ann. Math.* **196** 567–778

[74] Merle F, Raphaël P, Rodnianski I and Szeftel J 2022 On the implosion of a compressible fluid II: singularity formation *Ann. Math.* **196** 779–889

[75] Miao S 2018 On the formation of shock for quasilinear wave equations with weak intensity pulse *Ann. PDE* **4** 1–140

[76] Miao S and Yu P 2017 On the formation of shocks for quasilinear wave equations *Invent. Math.* **207** 697–831

[77] Oh S-J and Pasqualotto F 2021 Gradient blow-up for dispersive and dissipative perturbations of the Burgers equation (arXiv:2107.07172)

[78] Oliynyk T A 2016 Future stability of the FLRW fluid solutions in the presence of a positive cosmological constant *Commun. Math. Phys.* **346** 293–312

[79] O’Neill B 1983 *Semi-Riemannian Geometry (Pure and Applied Mathematics* vol 103) (Academic Inc. [Harcourt Brace Jovanovich Publishers])

[80] Qiu Y and Zhao L 2021 Shock formation of 3D Euler–Poisson system for electron fluid with steady ion background (arXiv:2108.09972)

[81] Rauch J 1986 BV estimates fail for most quasilinear hyperbolic systems in dimensions greater than one *Commun. Math. Phys.* **106** 481–4

[82] Raychaudhuri A 1955 Relativistic cosmology. I *Phys. Rev.* **98** 1123–6

[83] Rendall A D and Ståhl F 2008 Shock waves in plane symmetric spacetimes *Commun. PDE* **33** 2020–39

[84] Riemann B 1860 Über die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite *Abh. Königlichen Ges. Wiss. Goett.* **8** 43–66

[85] Rodnianski I and Speck J 2013 The nonlinear future stability of the FLRW family of solutions to the irrotational Euler–Einstein system with a positive cosmological constant *J. Eur. Math. Soc.* **15** 2369–462

[86] Sideris T C 1985 Formation of singularities in three-dimensional compressible fluids *Commun. Math. Phys.* **101** 475–85

[87] Smith H F and Tataru D 2005 Sharp local well-posedness results for the nonlinear wave equation *Ann. Math.* **162** 291–366

[88] Speck J 2008 On the questions of local and global well-posedness for the hyperbolic PDEs occurring in some relativistic theories of gravity and electromagnetism *PhD Dissertation* Rutgers University

[89] Speck J 2009 The non-relativistic limit of the Euler–Nordström system with cosmological constant *Rev. Math. Phys.* **21** 821–76

[90] Speck J 2009 Well-posedness for the Euler–Nordström system with cosmological constant *J. Hyperbolic Differ. Equ.* **6** 313–58

[91] Speck J 2012 The nonlinear future stability of the FLRW family of solutions to the Euler–Einstein system with a positive cosmological constant *Sel. Math.* **18** 633–715

[92] Speck J 2016 *Shock Formation in Small-Data Solutions to 3D Quasilinear Wave Equations (Mathematical Surveys and Monographs)* (American Mathematical Society)

[93] Speck J 2018 Shock formation for 2D quasilinear wave systems featuring multiple speeds: blowup for the fastest wave, with non-trivial interactions up to the singularity *Ann. PDE* **4** 6

[94] Speck J 2019 Multidimensional nonlinear geometric optics for transport operators with applications to stable shock formation *Pure Appl. Anal.* **1** 447–514

[95] Speck J 2019 A new formulation of the 3D compressible euler equations with dynamic entropy: remarkable null structures and regularity properties *Arch. Ration. Mech. Anal.* **234** 1223–79

[96] Speck J, Holzegel G, Luk J and Wong W 2016 Stable shock formation for nearly simple outgoing plane symmetric waves *Ann. PDE* **2** 1–198

[97] Taub A H 1948 Relativistic Rankine–Hugoniot equations *Phys. Rev.* **74** 328–34

[98] Wald R M 1984 *General Relativity* (University of Chicago Press)

[99] Wang Q 2017 A geometric approach for sharp local well-posedness of quasilinear wave equations *Ann. PDE* **3** 12

[100] Wang Q 2022 Rough solutions of the 3-D compressible Euler equations *Ann. Math.* **195** 509–654

[101] Weinberg S 2008 *Cosmology* (Oxford University Press)

[102] Yang R 2021 Shock formation of the Burgers–Hilbert equation *SIAM J. Math. Anal.* **53** 5756–802

[103] Yu S 2022 Rough solutions of the relativistic Euler equations (arXiv:2203.11746)

[104] Zhang H 2021 Local existence with low regularity for the 2D compressible Euler equations *J. Hyperbolic Differ. Equ.* **18** 701–28

[105] Zhang H and Andersson L 2021 On the rough solutions of 3D compressible Euler equations: an alternative proof (arXiv:2104.12299)

[106] Zhang H and Andersson L 2022 Improvement on the rough solutions of 3D compressible Euler equations (arXiv:2208.10132)