Star dynamics: collapse vs. expansion

Mahir Hadžić

Dedicated to Constantine M. Dafermos on the occasion of his 80-th birthday.

Abstract

We review a series of recent results on global dynamic properties of radially symmetric self-gravitating compressible Euler flows, which naturally arise in the mathematical description of stars. We focus on the role of scaling invariances and how they interact with nonlinearities to generate imploding finite-time singularities as well as expanding star solutions, arising from smooth initial data. This review paper is based on joint works with Y. Guo, J. Jang, and M. Schrecker.

Contents

1 Introduction .......................... 1

2 Self-gravitating Newtonian gases 4
   2.1 Lagrangian point of view .......... 4
   2.2 Scaling invariance ................. 6
   2.3 Existence via expansion .......... 6
      2.3.1 Nonlinear stability ............ 8
   2.4 Near-dust collapse ............... 10
      2.4.1 Self-gravitating dust .......... 10
      2.4.2 Near-dust collapse .......... 12
   2.5 Self-similar collapse .......... 15

3 Self-gravitating relativistic gases 21
   3.1 Formulation of the problem and the main result .......... 21
   3.2 Strategy of the proof and main ideas .......... 25

1 Introduction

First systematic mathematical studies of the global properties of self-gravitating fluids focused on steady states (i.e. time-independent solutions), which have played a central role in both astrophysics and mathematics literature, see for example the classical textbook of Chandrasekhar [10]. It is in particular natural to use stationarity to model objects such as the neutron stars and white dwarves, see the discussion in [56]. Different thermodynamic properties of the gas in the star are then encoded in a choice of the so-called equation of state, which can be rather complicated in realistic situations - we refer the interested reader to [10, 47, 88, 12, 94, 87] for more astrophysical background on the topic.

*Department of Mathematics, University College London, London UK, Email: m.hadzic@ucl.ac.uk.
Mathematically, the study of stationary stars and their stability properties is a beautiful topic with many unsolved open questions.

Our goal here is to move away from stationarity (and the associated stability analysis) and study time-dependent regimes wherein basic features of the physically important phenomena such as stellar collapse or supernova explosions can be captured in a mathematically rigorous way. We are in part motivated by the astrophysical expectation that the end-state of the gravitational collapse of a star should take on a self-similar form to leading order. This informal statement is sometimes referred to as the self-similarity hypothesis [7, 43, 44, 46]. It naturally leads to the basic question of the existence of physically meaningful self-similar solutions representing for example collapse - this topic is at the centre of this review.

Many of the mathematical difficulties associated with self-gravitating fluids originate in a very non-linear and nonlocal interaction between the pressure and the gravity. The former typically encourages dispersion, while the latter is clearly an attractive, focusing force. This antagonism, among other things, allows for the existence of steady states mentioned above – this is the case when these two forces are in perfect balance. However, it also leads to a number of different regimes, including most notably collapse and expansion.

The basic model of a self-gravitating dynamically evolving star is given by the compressible Euler-Poisson system, from now on the EP-system. The system takes the form

$$\frac{\partial_t \rho}{\rho} + \text{div}(\rho \mathbf{u}) = 0 \quad (1.1)$$

$$\rho D_t \mathbf{u} + \nabla p + \rho \nabla \phi = 0 \quad (1.2)$$

$$\Delta \phi = 4\pi \rho, \quad (1.3)$$

where the unknowns are the density \(\rho\), the pressure \(p\), the gravitational potential \(\phi\), and the velocity 3-vector \(\mathbf{u}\). To ensure that we model an isolated body, we demand asymptotic flatness encoded in the assumption \(\lim_{|x| \to \infty} \phi(x) = 0\). As written, (1.1)–(1.3) is underdetermined, with pressure \(p\) an unknown one too many. A standard way of closing the system is to impose an equation of state, which relates the pressure to the density. In this article, we are interested in the class of polytropic equations of state

$$p = \varepsilon \rho^\gamma, \quad 1 \leq \gamma < 2, \quad \varepsilon > 0. \quad (1.4)$$

The polytropic index \(\gamma\) parametrises the above family of problems and serves as a natural criticality parameter, see Section 2.2. The case \(\gamma = 1\) is sometimes referred to as the isothermal case and plays a particularly important role in the description of stellar collapse.

The EP-system has received a lot of attention in the astrophysics literature. A classical topic in this respect is the existence and (mostly linear) stability analysis of a well-known class of radial steady state solutions known as Lane-Emden stars [10, 94, 1], with many contributions in both mathematics and physics literature, see for example the introduction to [38] for the literature overview. At a nonlinear level, rigorous description of the phase space around the Lane-Emden stars remains an outstanding open problem. This includes the questions of the asymptotic-in-time behaviour, the existence of (quasi-) periodic solutions near the Lane-Emden stationary profiles, and the formation of shocks from smooth and small perturbations. There are only a few rigorous results available: conditional nonlinear stability by Rein [76] (see also the work of Luo and Smoller [57]) and rigorous nonlinear instability results in the range \(\frac{6}{5} \leq \gamma < \frac{4}{3}\) by Jang [50, 51].

In this review we focus on the nonlinear dynamics in a region of the phase space that is far from stationarity. Section 2 is devoted to the Newtonian theory, mostly the EP-system. We study two interesting dynamic scenarios which have a basis in astrophysical observations - stellar collapse and stellar expansion. The former is mathematically captured through finite-time singularity formation driven
by the density blow up, also referred to as implosion. Such singularities are therefore not shocks, as shocks are characterised by the formation of discontinuities in the field variables and give the archetype singular behaviour associated with compressible flows [21].

The common thread to all of our results is the decisive role of scaling invariances. In Sections 2.1 and 2.2 we present the Lagrangian formulation of the problem and explain the scaling symmetries respectively. We then show in Section 2.3 that one can find examples of global-in-time solutions to compressible Euler and Euler-Poisson flows which are driven by the expansion of the fluid (star) support - these are the main results of [38] and [39]. Working in the opposite direction, we also prove that there exist at least two types of collapsing stars in the Newtonian case. In Section 2.4.1 we explain the construction of collapsing stars from [34] that on approach to singularity behave like dust collapsing stars to the leading order. In contrast to such dust-like behaviour, in Section 2.5 we explain our recent construction [35, 37] of exactly self-similar collapsing flows, discovered numerically by Larson and Penston in 1969 [54, 75] when \( \gamma = 1 \), and by Yahil in 1983 [92] for certain ranges of \( \gamma > 1 \). The central mathematical challenge in our analysis is the presence of so-called sonic points. They encode natural singularities for the underlying dynamical system and correspond geometrically to the boundary of the backward acoustical cone emanating from the singularity. In a very recent work Merle, Raphaël, Rodnianski, and Szeftel [62] proved the existence of \( C^\infty \) imploding self-similar solutions for the Euler system (without gravity) featuring a sonic point.

In the context of general relativity, one must couple the Einstein field equations to the Euler equations. In this setting the gas density, pressure, and the 4-velocity \( u^\mu \) evolve simultaneously with the underlying Lorentzian manifold \((\mathcal{M}, g)\) through the Einstein equations:

\[
R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8 \pi T_{\mu\nu}, \quad \mu, \nu = 0, 1, 2, 3. \tag{1.5}
\]

Here, \( T_{\mu\nu} \) denotes the energy-momentum tensor which takes the form

\[
T_{\mu\nu} = (\rho + p) u_\mu u_\nu + pg_{\mu\nu}, \tag{1.6}
\]

and the 4-velocity is a future directed timelike vector satisfying the normalisation condition

\[
g_{\mu\nu}u^\mu u^\nu = -1. \tag{1.7}
\]

We refer to equations (1.5)–(1.7) as the Einstein-Euler system, from now on the EE-system. Here we shall assume the equation of state (1.4) with \( \gamma = 1 \). In this case the constant \( \varepsilon \) corresponds to the square of the speed of sound and to respect causality we assume \( 0 \leq \varepsilon \leq 1 \).

Under the right rescaling, the EP-system can be thought of as the Newtonian limit of the EE-system. The same questions as in the case of the EP-system apply to the EE-system, but due to the structure of the Einstein field equations and its interaction with the compressible Euler, much less is known. The analogues of the Lane-Emden stars have been known to exist since 1930’s, going back to the classical works of Tolman, Volkov, and Oppenheimer [91, 66]. Here the equation of state takes the form (1.4) with \( \gamma = 1 \) at high densities, and with \( \gamma \geq 1 \) at small densities close to the vacuum boundary. Formal linear stability analysis was initiated by Chandrasekhar in 1960’s [11] and the related turning point principle was formulated by Harrison, Thorne, Wakano, Wheeler, Zeldovitch et al. [47, 88, 93]. A rigorous linear (in)stability theory and the proof of the turning point principle were given recently by the author jointly with Z. Lin and G. Rein [41, 42].

In the absence of pressure, the relativistic fluid is often referred to as dust. In early 1930s Tolman [90] and Lemaître [55] solved the Einstein-dust system in radial symmetry (see also the later work of Bondi [2]), thus obtaining an infinite-dimensional family of solutions to the Einstein-dust system. Such solutions contain open families of initial data that lead to finite-time gravitational collapse, at first
instance measured by the blow-up of the dust mass-energy density. In their seminal work from 1939, Oppenheimer and Snyder [67] focused on a subclass of Tolman solutions characterised by the space-homogeneity of the dust density. By analysing the causal structure of the ensuing spacetime they laid the foundation for the modern notion of a black hole. However, in 1984 Christodoulou [16] examined the causal structure of the remaining, space-inhomogeneous (and therefore generic) Tolman spacetimes and came to a surprising conclusion. Instead of black holes, such spacetimes contained so-called naked singularities, which loosely speaking, emit signals that can be received by asymptotic observers. Such singularities are conjectured to be a non-generic feature of the theory of general relativity, and this is often mathematically codified via cosmic censorship hypotheses - we refer the reader to [19, 22, 79] for a rigorous discussion of this topic. In the context of radially symmetric matter models, Christodoulou proved the existence of a class of naked singularity spacetimes for the Einstein--scalar field model [17]. These solutions are self-similar in a suitable sense. In line with the weak cosmic censorship conjecture, he also proved that such naked singularities are unstable [19]. Much more recently, Rodnianski and Shlapentokh-Rothman [79, 83] proved the existence of naked singularity spacetimes solving the Einstein-vacuum system of equations.

In view of the above discussion, the conclusions of [16] can be taken as a motivation for study of the causal structure of implosion for more realistic fluid models which include pressure. In Section 3 we finally present our recent result [36] on the formation of self-similar imploding singularities for the EE-system with \( \gamma = 1 \) and \( 0 < \varepsilon \ll 1 \) in (1.4), and thereby make rigorous the pioneering predictions of Ori and Piran [68, 69, 70]. The resulting spacetimes are referred to as the relativistic Larson-Penston solutions. Again, here we are dynamically far from stationarity and the constructed spacetime, as we will show, contains a naked singularity. The latter statement requires an in-depth description of the causal structure of the spacetime and this interaction between the Lorentzian- and the acoustical geometry associated with the propagation of sound waves is one of the most interesting features of this problem.

Acknowledgments. The author acknowledges the support of the EPSRC Early Career Fellowship EP/S02218X/1. He thanks Y. Guo, J. Jang, and M. Schrecker for fruitful discussions.

2 Self-gravitating Newtonian gases

2.1 Lagrangian point of view

In the absence of vacuum regions, local well-posedness for the EP-system can be shown using for example the classical symmetrisation techniques [21] for hyperbolic conservation laws combined with an energy method. In the presence of vacuum the situation is more complicated, see for example [59, 60]. For more references on the topic and global-in-time weak solution theories in radial symmetry, we refer the reader to the recent work [14].

In the context of classical compressible fluid dynamics it is often advantageous to use the Lagrangian (or comoving) coordinates, which turn out to be particularly well-suited to the presence of vacuum regions. If we denote by \( \Omega(t) \subset \mathbb{R}^3 \) the support of the gas, the basic unknown is the flow map \( \eta(t, \cdot) \) defined on some reference domain \( \mathcal{M} \subset \mathbb{R}^3 \) so that \( \eta(t, \cdot) : \mathcal{M} \to \Omega(t) \) solves

\[
\partial_t \eta(t, x) = u(t, \eta(t, x)), \quad x \in \mathcal{M},
\]

\[
\eta(0, x) = \eta_0, \quad x \in \mathcal{M}.
\]

Here \( \eta_0 : \mathcal{M} \to \Omega(0) \) is the labelling map and its choice fixes the labelling gauge freedom in the problem; recall that \( u(t, \cdot) : \Omega(t) \to \mathbb{R}^3 \) is the velocity vectorfield associated with the gas flow. In the
case of a free-moving surface $\partial \Omega(t)$. Lagrangian coordinates can be used to pull the problem back to the fixed reference domain $M$. To do so, we introduce the new unknowns

\[ v := u \circ \eta \quad \text{(Lagrangian velocity)}, \]
\[ \mathcal{A} := [D\eta]^{-1} \quad \text{(Inverse of the Jacobian matrix)}, \]
\[ \mathcal{J} := \det[D\eta] \quad \text{(Jacobian determinant)}. \]

(2.10) \hspace{1cm} (2.11) \hspace{1cm} (2.12)

For notational brevity, it is useful to introduce

\[ \alpha := \frac{1}{\gamma - 1}. \]

(2.13)

A standard calculation [52] shows that the the compressible Euler equations take the form

\[ w^\alpha \partial_t \eta_i + \left( w^{1+\alpha} \mathcal{A}_i^k \mathcal{J}^{1-\gamma} \right)_{,k} = 0, \quad i = 1, 2, 3, \quad x \in M, \]

(2.14)

where

\[ w(x) := (\rho_0(\eta_0(x)) \mathcal{J}_0(x))^{\gamma - 1} \]

(2.15)

corresponds to the initial gas enthalpy. Here we have assumed $\varepsilon = 1$ in (1.4) as the exact value does not play a role for the considerations in this section. We note that $w(x) = \rho_0(x)^{\gamma - 1}$ if $M = \Omega(0)$ and $\eta_0 = \text{Id}$. In the presence of a self-induced gravitational field, the Lagrangian description of the Euler-Poisson system (1.1)–(1.3) takes the form

\[ w^\alpha \partial_t \eta_i + \left( w^{1+\alpha} \mathcal{A}_i^k \mathcal{J}^{1-\gamma} \right)_{,k} + \mathcal{A}_i^k \psi_{,k} = 0, \quad i = 1, 2, 3, \quad x \in M, \]

(2.16)
\[ \mathcal{A}_k^i \left( \mathcal{A}_j^i \psi_{,j} \right)_{,k} = 4\pi w^\alpha \mathcal{J}^{-1}, \]

(2.17)

where $\psi = \phi \circ \eta$ is the Lagrangian pull-back of the gravitational potential $\phi$. As we can see from (2.14) and (1.1)–(1.3), both systems of equations are quasilinear second order wave-like systems. The precise nature of the flow and the role of the enthalpy in the context of well-posedness for the free surface flows will become apparent in later sections.

If the flow is radially symmetric, we can write the flow map in the form

\[ \eta(t, x) = \chi(t, r)x, \]

where $r = |x|$ and $\chi(t, \cdot)$ is a non-negative function. It is straightforward to check [51] that (2.14) takes the form

\[ \chi_{tt} + \frac{\chi^2}{w^\alpha r} \partial_r \left( w^{1+\alpha} \mathcal{J}^{-\gamma} \right) = 0, \]

(2.18)

while the radial Euler-Poisson system (2.16)–(2.17) takes the form

\[ \chi_{tt} + \frac{\chi^2}{w^\alpha r} \partial_r \left( w^{1+\alpha} \mathcal{J}^{-\gamma} \right) + \frac{G(r)}{\chi^2} = 0, \]

(2.19)

where $G(r)$ is the average density expressed in Lagrangian coordinates and $\mathcal{J}[\chi]$ the Lagrangian determinant

\[ G(r) = \frac{1}{r^3} \int_0^r 4\pi \rho_0(\chi_0(s)) \mathcal{J}_0(s)s^2 ds, \]

(2.20)
\[ \mathcal{J}[\chi] = \chi^2 (\chi + r \partial_r \chi). \]

(2.21)
2.2 Scaling invariance

It is a simple algebraic exercise to verify that for any \( \lambda > 0 \) the transformation

\[
\begin{align*}
\rho &\mapsto \lambda^{-\frac{2}{3\gamma}}\rho(s, y), \\
\mathbf{u}(t, x) &\mapsto \lambda^{-\frac{2-\gamma}{3\gamma}}\mathbf{u}(s, y), \\
\phi(t, x) &\mapsto \lambda^{-\frac{-2\gamma+2}{3\gamma}}\phi(s, y)
\end{align*}
\]  

(2.22)

keeps the EP-system invariant, where the time and space scale according to

\[
s = \frac{t}{\lambda^{\frac{1}{3\gamma}}}, \quad y = \frac{x}{\lambda}
\]  

(2.23)

The polytropic index \( \gamma \) therefore acts as a natural criticality parameter, and it is readily verified that the total mass

\[
M[\rho] := \int_{\mathbb{R}^3} \rho \, dx,
\]  

(2.24)

and the total energy

\[
E[\rho, \mathbf{u}] := \int_{\mathbb{R}^3} \left( \frac{1}{2} \rho |\mathbf{u}|^2 + \frac{1}{\gamma} \rho^{\gamma} - \frac{1}{8\pi} |\nabla \phi|^2 \right) \, dx,
\]  

(2.25)

are formally preserved exactly when \( \gamma = \frac{4}{3} \) and \( \gamma = \frac{6}{5} \) respectively.

2.3 Existence via expansion

It is well-known that the tendency of fluid particles to expand and move away from each other can generate dispersion, which counteracts possible focusing of the characteristics. This mechanism therefore has the potential to generate solutions that exist globally-in-time. As it is expected for gravity to be a subdominant effect in such a situation, it is instructive to look for global-in-time solutions to the compressible Euler system

\[
\begin{align*}
\partial_t \rho + \text{div}(\rho \mathbf{u}) &= 0 \\
\rho (\partial_t + \mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p &= 0
\end{align*}
\]  

(2.26–2.27)

where we assume the equation of state (1.4) with \( \varepsilon = 1 \).

First solutions of this flavour were constructed in the important works of Ovsyannikov \[71\] and Dyson \[25\]. They belong to a class of so-called affine motions which will be discussed below. Expansion is also a mechanism for global existence in the works of Grassin \[31\], Serre \[81\], and Rozanova \[80\]. In a remarkable work \[85\], Sideris exhibited an explicit class of expanding solutions to the Euler system (2.26)–(2.27) that exist for all \( t > 0 \) and retain compactly supported density. The Sideris flows are most easily described using the comoving description of the flow. He restricts his attention to a special type of solutions, known as affine flows, which by definition take the form

\[
\eta(t, x) = A(t) x, \quad A \in \text{GL}^+(3, \mathbb{R}),
\]  

(2.28)

where we recall (2.8). The compressible Euler system then reduces to

\[
\ddot{A}(t) x + \frac{\gamma}{\gamma - 1} A(t)^{-\top} \det(A(t))^{1-\gamma} \nabla_x w = 0, \quad \gamma > 1,
\]  

(2.29)

where \( A^{-\top} \) stands for the transpose of the inverse of the matrix \( A \) and \( w = \rho_0^{\gamma-1} \) is the initial enthalpy. The idea is to separate variables. Given a \( \delta \in \mathbb{R} \) and \( (A_0, A_1) \in \text{GL}^+(3, \mathbb{R}) \times \mathbb{M}^3 \) we seek to solve the
Figure 1: Sideris expanding ellipsoids

The Goldreich-Weber solutions split into two groups: the expanding star solutions (not unlike the Sideris flows above) and the collapsing star solutions, which are very different from the Sideris flows. There also exists a purely Eulerian approach to the derivation of these affine flows developed by the author and Jang [39]. While less direct, it clarifies the precise role of symmetries of the Euler equation responsible for the existence of such special structures and contextualises them within the broader setting of dispersive PDE.

Both classes of affine motions described above feature a vacuum boundary separating regions of positive and vanishing pressure. More importantly, in both cases the fluid pressure is not a smooth
function at the vacuum boundary. This lack of smoothness is characterised by the behaviour of the fluid enthalpy (2.15) at the vacuum boundary, which both in the case of the Sideris- and the Goldreich-Weber flows reads

\[ w(x) \approx_{x \to \partial \Omega} \text{dist}(x, \partial \Omega). \]  

(2.37)

Condition (2.37) is known as the physical vacuum condition. In two independent works by Jang and Masmoudi [52] and Coutand-Shkoller [20] it was shown that the free surface Euler equations are well-posed locally-in-time precisely under the assumption of physical vacuum on the initial data. These works give the natural mathematical framework for the study of dynamic properties of solutions to Euler flows. The techniques in [52, 20] can be adapted to yield well-posedness for the Euler-Poisson system in the presence of physical vacuum [32, 40]. For an Eulerian approach to the well-posedness in the presence of physical vacuum see the recent work [49].

2.3.1 Nonlinear stability

The natural question is whether the Sideris and the expanding Goldreich-Weber solutions are dynamically stable in some reasonable topology. We answered this question to the affirmative, taking advantage of the expansion of the underlying solution.

Theorem 2.1 ([38], informal statement). The expanding Goldreich-Weber flows are nonlinearly stable in the class of radially symmetric perturbations.

Theorem 2.2 ([39], informal statement). Sideris flows are nonlinearly stable for \( 1 < \gamma \leq \frac{5}{3} \).

The main idea of the proofs of Theorems 2.2 and 2.1 is to capture the stabilising effect of the background expansion in the Lagrangian framework. If \((\rho, u)\) is an Eulerian description of a flow which starts off “close” to a fixed Sideris motion \((t, x) \mapsto A(t)x\), we consider the associated flow map \(\zeta\), i.e. \(\partial_t \zeta = u(\cdot, \cdot) \circ \zeta\). The idea is to renormalise the map \(\zeta\) by the background flow \(A(t)x\) in a suitable way - we introduce the vectorfield \(\theta\) through the relation

\[ A^{-1}(t)\zeta(t, x) = \text{Id} + \theta(t, x). \]  

(2.38)

To give a succinct explanation of the key stabilisation mechanism, let us choose a particularly convenient subclass of diagonal Sideris flows of the form

\[ A(t) = \text{diag}(\lambda(t), \lambda(t), \lambda(t)), \]  

(2.39)

where according to (2.30)–(2.31) \(\lambda(t)\) satisfies the second order ODE

\[ \dddot{\lambda} + \frac{3}{\gamma - 2} \lambda^{\gamma - 2} = \delta, \quad \lambda(0) = 1, \quad \dot{\lambda}(0) = \lambda_1 \in \mathbb{R}. \]  

(2.40)

A routine exercise in ODE shows that for any \( \delta > 0 \) there exists a global solution to (2.40) and \(\lambda(t) \approx_{t \to \infty} t\). In other words, the associated affine motion expands at a linear rate in \(t\). To highlight the stabilising effect of this expansion we switch to a logarithmic time-scale by introducing a new time variable \(\tau\) through \(\frac{d\tau}{dt} = \frac{1}{\lambda(t)}\). It is another routine exercise to derive the PDE satisfied by the perturbation \(\theta\):

\[ \lambda(\tau)^{3\gamma - 3} \left( \partial_{\tau \tau} \theta_i + \frac{\partial \lambda}{\lambda} \partial_{\tau} \theta_i \right) + \theta_i + \frac{\gamma}{\gamma - 1} w^{-\alpha} \partial_k \left( \omega_i^k \mathcal{J}^{-\frac{\gamma}{2}} - \delta_i^k \right) = 0, \quad i = 1, 2, 3. \]  

(2.41)

In self-similar variables the linear-in-\(t\) growth of \(\lambda(t)\) translates into the asymptotic property \(\lambda(\tau) \approx_{\tau \to \infty} e^\tau\). In particular \(\frac{\partial \Delta}{\partial \tau} \approx_{\tau \to \infty} 1\). It is therefore illustrative to consider the simpler model
normal and tangential derivatives inside $\partial M$ as in Masmoudi [52] where they implemented a particular weighting scheme sensitive to the number of operator on approach to the boundary $\partial w$ by-parts. The story is however not so simple since the presence of the weight $\partial^\theta$ highest-order spatial derivative that we commute $F_i[\theta]$ (2.43) is on the other hand well-known in the context of free-boundary vacuum problems. If $\partial$ denotes the highest-order spatial derivative that we commute $F_i[\theta]$ with, upon multiplying the commuted equation by $\partial^\theta$, our goal is to extract a (hopefully) positive-definite coercive energy-like term upon integrating-by-parts. The story is however not so simple since the presence of the weight $w$ degenerates the operator on approach to the boundary $\partial \Omega$. This issue is at the heart of the complications associated with the local-in-time well-posedness. The resolution came about through a deep idea of Jang and Masmoudi [52] where they implemented a particular weighting scheme sensitive to the number of normal and tangential derivatives inside $\partial$. If we decompose $\partial$ in a general form

$$\partial = \partial^k \nabla^\beta, \ | \beta | + k = N,$$

(2.45)

then whenever we commute the PDE with an operator of the form (2.45) we evaluate the inner product with $\partial^\theta$, in the weighted $L^2_{w^\alpha+k}$ energy space, where $\| f \|^2_{L^2_{w^\alpha+k}} = \int_{\Omega} f^2 w^{\alpha+k} \, dy$. Here $\nabla$ represents the tangential gradient. The role of the weight $w^{\alpha+k}$ is to offset the degeneracy caused by the vanishing of the enthalpy at the boundary. Using a high-order energy framework in such weighted spaces and Hardy-Sobolev embeddings one can adapt the well-posedness framework from [52] to our setting. The damping effect mentioned above then allows us to extend our local solution to a global one and thus prove Theorem 2.2. We refer the reader to [39] for more details. The proof of Theorem 2.1 follows similar ideas. It is different in that Goldreich-Weber stars can in fact expand at two different rates: the linear-in-$t$ rate (just like Sideris flows) and the non-generic self-similar rate $t^{\gamma/3}$. The latter solutions are only co-dimension 1 stable and to prove this one necessitates a refined spectral analysis around the solutions of the generalised Lane-Emden equation (2.36). We refer the reader to [38] for the details.

**Remark 2.3.** Stability of the collapsing Goldreich-Weber stars is an open problem.

**Remark 2.4.** The Sideris expanding flows in the regime $\gamma > \frac{5}{3}$ were shown to be stable by Shkoller and Sideris [82].

**Remark 2.5.** Ideas related to the above expansion mechanism have been further developed in [73, 72, 78].
2.4 Near-dust collapse

2.4.1 Self-gravitating dust

A historically integral part of the mathematical study of stellar collapse is the collapse of self-gravitating dust clouds. By definition, dust corresponds to a pressureless fluid, which amounts to the assumption \( p = 0 \) in (1.1)–(1.3) or \( p = 0 \) in (1.6) in the relativistic case. The absence of pressure simplifies the problem considerably, to the point where one can explicitly solve the resulting PDE under the further assumption of radial symmetry.

In the well-known works [90, 55] Tolman and Lemaître solved the dust-Einstein equations in radial symmetry, thus providing a large family of explicit solutions, parametrised by the choice of the initial mass-energy density \( \rho_0 \). In their works it is very natural to use the comoving coordinates and we next briefly explain how to derive the Newtonian analogue of Tolman-Lemaître solutions using Lagrangian coordinates. We begin with a physically natural assumption that the initial density \( \rho_0(r) \) is monotonically decreasing, i.e. \( \rho_0 < 0 \) on the interior of the support of \( \rho_0 \), which we take to be the interval \([0, 1]\).

If we assume that the dust-cloud is spherically symmetric and use the Lagrangian coordinates, then by (2.19), the dust-Poisson system takes the form

\[
\chi_{tt} + \frac{G(r)}{\chi^2} = 0,
\]

where we recall (2.20). This is remarkable, as in the Lagrangian variables, the problem reduces to a family of ODE parametrised by the choice of \( r \in [0, 1] \). Given the initial conditions

\[
\chi(0, r) = \chi_0(r) > 0, \quad \chi_t(0, r) = \chi_1(r),
\]

the solution can thus be determined uniquely. We consider the inward moving initial velocities with \( \chi_1 < 0 \). From the energy conservation

\[
E(t) := \frac{1}{2} \chi_t^2 - \frac{G(r)}{\chi} = E(0)
\]

we obtain the formula

\[
\chi_t = -\sqrt{\chi_1^2 + 2G \left( \frac{1}{\chi} - \frac{1}{\chi_0} \right)}.
\]

Integrating (2.49) one sees that for every \( r \) there exists a \( 0 < t^*(r) < \infty \) such that \( \chi(t^*(r), r) = 0 \). A simple calculation reveals that for any \( r \in [0, 1] \) we have the universal blow-up exponent 2/3

\[
\chi(t, r) \sim c(r)(t^*(r) - t)^{\frac{2}{3}}, \quad t \to t^*(r).
\]

We may further define the first blow-up time

\[
t^* := \min_{r \in [0, 1]} t^*(r).
\]

Observe that the Eulerian description of the solution seizes to make sense at and after time \( t \geq t^* \), as for example \( \mathcal{J}[\chi]|_{t = t^*} = 0 \), see (2.21). On the other hand, for different values of \( r \) the Lagrangian solution may make sense even after \( t^* \). In particular, when \( t^*(r) \) is a non-constant function, we can

\[\text{The monotonicity condition is for example satisfied by the Lane-Emden stars.}\]
speak of the continued stellar collapse, wherein particles with a different Lagrangian label $r$ collapse at different times. This behaviour is generic and is referred to as inhomogeneous collapse \[89\]. To further illustrate the nature of this type of singular behaviour we consider a particular subclass of solutions given by

$$\chi_{\text{dust}}(t,r) = (1 - g(r)t)^{\frac{2}{3}}, \quad g(r) := \left(\frac{9G(r)}{2}\right)^{\frac{1}{2}}. \quad (2.51)$$

Notice that this solution satisfies (2.46) with the initial values $\chi_0(r) \equiv 1$ and $\chi_1(r) = -\frac{2}{3}g(r) < 0$. In this case the average density $G(r)$ is given by

$$G(r) = \frac{1}{r^3} \int_0^r 4\pi \rho_0(s)s^2 \, ds, \quad (2.52)$$

which is a decreasing function since $\rho_0$ is decreasing. It follows that $\chi_{\text{dust}}$ becomes zero along the space-time curve

$$\Gamma := \{(t, r) \mid 1 - g(r)t = 0\}, \quad (2.53)$$

see Figure 2. The solution is only well-defined in the region

$$\mathcal{G} := \{(t, r) \mid 1 - g(r)t > 0\}.$$

After a simple calculation using (2.21) we have

$$\mathcal{J}[\chi_{\text{dust}}](t, r) = (1 - g(r)t)^2 \left(1 - \frac{2}{3} \frac{trg'(r)}{1 - g(r)t}\right), \quad (t, r) \in \mathcal{G}. \quad (2.54)$$

In particular, $\chi_{\text{dust}}$ and $\mathcal{J}[\chi_{\text{dust}}]$ vanish along $\Gamma$ and therefore, since the Eulerian density satisfies

$$\rho_{\text{dust}}(t, \chi_{\text{dust}}(t, r)x) = \rho_0(r) \mathcal{J}[\chi_{\text{dust}}](t, r)^{-1}, \quad r = |x|, \quad (t, r) \in \mathcal{G},$$

Figure 2: Dust collapse in Lagrangian coordinates
the value of \( \rho_{\text{dust}}(t, 0) \) diverges to infinity at the first blow-up time \( t^* := \frac{1}{g(0)} \). In the region \( \chi_{\text{dust}} > 0 \), the Eulerian density \( x \mapsto \rho_{\text{dust}}(t, x) \) is always well-defined away from the origin \( x = 0 \). Moreover for any \( r \in [0, 1] \)
\[
\lim_{t \to \frac{1}{g(r)}} \rho_{\text{dust}}(t, \chi_{\text{dust}}(t, r)x) = \infty.
\]
Since \( r \mapsto g(r) \) is monotonically decreasing, particles that start out closer to the boundary of the star take longer to vanish into the singularity.

Finally, the support of the collapsing dust star shrinks to zero as \( t \to \frac{1}{g(1)} \). This is clear, as the free boundary in the Eulerian description is at distance \( \chi_{\text{dust}}(t, 1) = (1 - g(1)t)^{\frac{2}{3}} \) from the origin. As \( t \to \frac{1}{g(1)} \) the star concentrates with its mass completely absorbed at the origin:
\[
\lim_{t \to \frac{1}{g(1)}} \chi_{\text{dust}}(t, 1) = 0.
\]
Therefore the time \( t = \frac{1}{g(1)} \) has a natural interpretation as the end-point of star collapse for the dust example considered here.

### 2.4.2 Near-dust collapse

It is natural to ask if there exist initial data that lead to finite time collapse that is qualitatively similar to the dust collapse described above, but in the presence of pressure. This is an a priori challenging problem, because pressure enters the equation at the highest order of regularity and counteracts density concentration. An affirmative answer to this question is formulated in the following theorem.

**Theorem 2.6 ([34]).** Let \( \gamma \in (1, \frac{4}{3}) \). Then there exist initial data \( \chi_0, \chi_1 \) to (2.19), a natural number \( n = n(\gamma) \) \( \in \mathbb{N} \) sufficiently large, and a compactly supported initial density profile \( \rho_0 : [0, 1] \to \mathbb{R}^+ \), satisfying
- \( \rho_0 > 0 \) on \( [0, 1) \) (density is positive);
- \( \rho_0(1) = 0 \) (vacuum at \( r = 1 \));
- There exists a constant \( c > 0 \) such that
  \[
  \rho_0(r) = \rho_0(0) - cr^n + O_{r \to 0+}(r^{n+1}), \quad \text{(the star is sufficiently flat at the centre),}
  \]
so that the resulting Euler-Poisson flow leads to finite time collapse along the spacetime surface \( \Gamma \) defined in (2.53), with \( g(r) \) as in (2.51) and (2.52).

To build some intuition, for any \( \gamma \in (1, \frac{4}{3}) \) consider a mass-invariant scaling of the EP-system given by the change of variables \( \rho \mapsto \lambda^{-3(\rho(s, y), \ u(t, x) \mapsto \lambda^{-\frac{2}{3}}u(s, y), \ \phi(t, x) \mapsto \lambda^{-1}\phi(s, y), \) where \( s = \frac{t}{\lambda^2}, \ y = \frac{x}{\lambda}. \) The rescaled system is similar to the original EP-system (1.1)–(1.3) and reads
\[
\partial_t \rho + \text{div}(\rho u) = 0
\]
\[
\rho D_t u + \epsilon \nabla p + \rho \nabla \phi = 0, \quad \epsilon := \lambda^{4-3\gamma},
\]
\[
\Delta \phi = 4\pi \rho.
\]
If we pretend for a moment that \( \lambda \) corresponds to some dynamically shrinking scale that leads to density blow-up, then we see from the rescaled momentum equation (2.57) that the coefficient \( \epsilon = \lambda^{4-3\gamma} \) is small when \( \lambda \ll 1 \) if and only if \( 1 < \gamma < \frac{4}{3} \). This reflects the mass-supercritical nature of the problem and suggests a possibility of taming the pressure term in (2.57). This intuition is also deeply
problematic, as the pressure term enters the equations at the top order of regularity and it is not clear whether the above scaling can conspire with the pressure to mimic a dust-like collapse described in Section 2.4.1.

Nevertheless, if we formally set $\epsilon = 0$ in (2.57) we obtain the dust-Poisson system. This suggests the idea to try to build a collapsing solution to the EP-system as a “perturbation” of $\chi_{\text{dust}}$. To that end we first rectify the space-time support $G$ of the dust solution (see Figure 2) by foliating it with the level sets of the function $\chi_{\text{dust}}$. In other words, we introduce a new timelike coordinate $\tau$ and a new unknown $\phi$

$$\tau = 1 - g(r)t, \quad \phi(\tau, r) = \chi(t, r), \quad (2.59)$$

see Figure 3. The dust solution $\chi_{\text{dust}}$ (2.51) now corresponds to

$$\phi_0(\tau, r) \equiv \tau^2. \quad (2.60)$$

A simple change of variables implies that $\phi$ solves the PDE

$$\partial_{\tau} \phi + \frac{2}{9\phi^2} + \epsilon P[\phi] = 0, \quad (2.61)$$

where $P[\cdot]$ is the pull-back of the second order pressure operator, which in the new coordinates reads

$$P[\phi] = \frac{\phi^2}{g(r)w(r)^{\alpha r^2}} \Lambda \left(w(r)^{1+\alpha} J[\phi]^{-\gamma} \right), \quad (2.62)$$

$$J[\phi] = \phi^2 (\phi + \Lambda \phi). \quad (2.63)$$

Here $\Lambda$ is the pull-back of the scaling operator $r\partial_r$ and reads

$$\Lambda = M_g \partial_r + r \partial_r, \quad M_g(\tau, r) := (\tau - 1)r \partial_r (\log g).$$

13
The fundamental feature of the operator \( \Lambda \) is that in the vicinity of the origin it takes the approximate form
\[
\Lambda \approx r^n \frac{\partial}{\partial r} + \frac{\partial}{\partial r},
\]
where we recall (2.55). We see here that the first nontrivial term in the Taylor expansion of \( \rho_0 \) about the origin introduces a new effective scale in the problem in the vicinity of \( r = 0 \).

The main idea of the proof is to construct a sufficiently accurate approximate solution \( \phi_{\text{app}} \) to the problem so that
\[
\phi = \phi_{\text{app}} + \tau^m \frac{H}{r},
\]
where
\[
\phi_{\text{app}} = \phi_0 + \epsilon \phi_1 + \cdots + \epsilon^M \phi_M,
\]
where \( 1 \ll m, M \in \mathbb{N} \) are to be determined, and the remainder term \( \frac{H}{r} \) is to be of order 1. We emphasise here that all the difficulty in the problem is concentrated at the (new) spacetime origin \((\tau, r) = (0, 0)\). We therefore want to focus on the regime where both \( 0 < r \ll 1 \) and \( 0 < \tau \ll 1 \).

To construct the approximate solution, we plug in the expansion (2.66) in (2.61) and formally expand in the powers of \( \epsilon \). This leads to
\[
\partial_{\tau\tau} (\phi_0 + \epsilon \phi_1 + \cdots) + \frac{2}{9 (\phi_0 + \epsilon \phi_1 + \cdots)^2} + \epsilon P[\phi_0 + \epsilon \phi_1 + \cdots] = 0.
\]

Matching the powers of \( \epsilon \) we obtain a hierarchy of ODE of the form
\[
\partial_{\tau\tau} \phi_{j+1} - \frac{4}{9 r^2} \phi_{j+1} = f_{j+1} [\phi_0, \ldots, \phi_j], \quad j \geq 0,
\]
where \( f_{j+1} \) are complicated combinatorial expressions that depend only on \( \phi_1, \ldots, \phi_j \), but do not depend on \( \phi_{j+1} \) and therefore act as source terms in (2.68). For such an expansion to be meaningful, we want \( \phi_{j+1} \) to be effectively smaller than the previous iterate \( \phi_j \) in the near-collapse regime \( \tau \ll 1 \).

To illustrate this point, consider the equation satisfied by \( \phi_1 \):
\[
\partial_{\tau\tau} \phi_1 - \frac{4}{9 r^2} \phi_1 = -P[\phi_0].
\]

Recalling (2.64) it is not hard to see that to the leading order in the vicinity of \((0, 0)\) the quantity \( P[\phi_0] \) takes the form
\[
P[\phi_0] \approx \tau^{\frac{8}{3} - 2\gamma - \frac{2}{n}} \left( \frac{r^n}{1 + \frac{n}{r}} \right)^{1 - \frac{n}{r}}.
\]

Going back to (2.69), since it is a second order ODE it stands to reason that we can pick a solution \( \phi_1 \) so that it “gains” two powers of \( \tau \) with respect to the source term. In other words we solve (2.69), so that the resulting solution near the singular point \((0, 0)\) takes on the leading order form
\[
\phi_1 \approx \frac{\phi_0}{\tau^{\frac{8}{3} - 2\gamma - \frac{2}{n}} \left( 1 + \frac{n}{r} \right)^{1 - \frac{n}{r}}},
\]

We let
\[
\delta := \frac{8}{3} - 2\gamma - \frac{2}{n} = 2 \left( \frac{4}{3} - \gamma - \frac{1}{n} \right).
\]
It is now clear that for any $1 < \gamma < \frac{4}{3}$ there exists a choice of $n \in \mathbb{N}$ such that $\delta > 0$. Moreover, for any $r \in [0, 1]$ and $\tau > 0$ the fraction $\left(\frac{\tau}{r^n}\right)^{1 + \frac{\delta}{\tau^2}}$ is always bounded from above uniformly, and we conclude that

$$|\phi_1| \lesssim \tau^\delta |\phi_0| \ll \phi_0, \quad 0 < \tau \ll 1.$$  \hspace{1cm} (2.73)

This $\delta$-gain of smallness for the iterate $\phi_1$ is the fundamental feature of our strategy. It relies on both supercriticality of the problem, and the near-flatness assumption (2.55) with $n > \frac{1}{\sqrt{4 - \gamma}}$ chosen sufficiently large. The above calculation also illustrates the emergence and importance of the new characteristic spacetime scale $\tau^n$ in the region near the singularity $(r, \tau) = (0, 0)$. The gain (2.73) can be shown to be propagated with every next iterate and spacetime derivatives. In [34] we show inductively that

$$\left| \partial^m_t \left( r \partial_r \right)^\ell \phi_j \right| \lesssim m, \ell, j \tau^{\frac{3}{2} + j \delta - m}, \quad \ell, j, m \in \mathbb{N}.$$  \hspace{1cm} (2.74)

With (2.74) in hand, we can now probe the ansatz (2.65) by plugging it back into the original PDE (2.61). This will result in a new PDE for the remainder term $H$ with a source term which depends explicitly on $\phi_{\text{app}}$, see (2.66). This will result in a quasilinear wave equation of the schematic form

$$g^{00}[\phi] \partial_{\tau \tau} H + g^{01}[\phi] \partial_{\tau r} H + \left( \frac{M(M - 1)}{\tau^2} - \frac{2}{3\tau^2} \right) H - \epsilon c[\phi] \frac{1}{w^\alpha} \partial_r \left( w^{1 + \alpha} \frac{1}{\tau^2} \partial_r (r^2 H) \right) = S_M[\phi_{\text{app}}].$$  \hspace{1cm} (2.75)

The coefficients $g^{00}[\phi]$, $g^{01}[\phi]$, and $c[\phi]$ all exhibit singular/degenerate behaviour in the vicinity of $(0, 0)$ and therefore require care. To prove the existence of a solution $H$ to (2.75) we therefore develop a high-order energy scheme, which is suitably weighted with power-like weights in $\tau$ (to handle the degeneracy at $\tau = 0$) and powers of $w(r)$ (to handle the vacuum degeneracy at $r = 1$). Most importantly, we can choose the “depth” $M$ of the approximate solution $\phi_{\text{app}}$ (2.66) sufficiently large to show that the source term $S_M[\phi_{\text{app}}]$ is sufficiently small near $(\tau, r) = (0, 0)$. This scheme then leads to a global existence result for all $\tau \in (0, 1]$. By the change of variables (2.59) the data at $\tau = 1$ correspond to initial data at $t = 0$, which by construction lead to a dust-like collapsing solution, which to the leading order behaves like (2.51).

**Remark 2.7.** It is not clear whether there exist open classes of data in a reasonable topology that exhibit a dust-like collapse.

### 2.5 Self-similar collapse

We now turn our attention to the construction of self-similar solutions that lead to finite time gravitational collapse from smooth data. We shall see that this point of view is extremely fruitful, as it will additionally lead to new insights in the construction of collapsing relativistic flows in Section 3.

Based on the scaling invariance (2.22), we define the self-similar variable

$$y = \frac{r}{\sqrt{\varepsilon}(-t)^{2 - \gamma}}$$  \hspace{1cm} (2.76)

and look for solutions to (1.1)–(1.4) of the form

$$\rho(t, r) = (-t)^{-\gamma/2} \hat{\rho}(y),$$  \hspace{1cm} (2.77)

$$u(t, r) = \sqrt{\varepsilon}(-t)^{1 - \gamma} \hat{u}(y).$$  \hspace{1cm} (2.78)
We introduce the relative (sometimes called the wind-) velocity
\[ \omega := \frac{u + (2 - \gamma)y}{y}, \]
and plug in (2.77)–(2.78) into (1.1)–(1.3). This leads to the following system of ordinary differential equations
\[ \rho' = \frac{y h(\rho, \omega)}{G(y; \rho, \omega)}, \]
\[ \omega' = \frac{4 - 3\gamma - 3\omega}{y} - \frac{y\omega h(\rho, \omega)}{G(y; \rho, \omega)}, \]
where
\[ h(\rho, \omega) = 2\omega^2 + (\gamma - 1)\omega - \frac{4\pi\rho\omega}{4 - 3\gamma} + (\gamma - 1)(2 - \gamma), \]
\[ G(y; \rho, \omega) = \gamma\rho^{\gamma - 1} - y^2\omega^2. \]

In the isothermal case \( \gamma = 1 \), system (2.80)–(2.81) takes the form
\[ \rho' = \frac{2y\rho\omega(\omega - 2\pi\rho)}{1 - y^2\omega^2}, \]
\[ \omega' = \frac{1 - 3\omega}{y} - \frac{y\omega(\omega - 2\pi\rho)}{1 - y^2\omega^2}, \]

We shall be interested in smooth \( C^\infty \) solutions of (2.80)–(2.81) in the mass-supercritical range \( 1 \leq \gamma < \frac{4}{3} \). This amounts to solving the non-autonomous system of ODE (2.80)–(2.81) on the domain \([0, \infty)\). Any such solution can be interpreted as an imploding star solution whose density blows up at the rate \( \frac{1}{t^2} \) on approach to singularity along the cones of constant \( y \equiv c \), i.e.
\[ \rho(t, x) = \frac{1}{t^2}\rho(c), \quad \frac{|x|}{\sqrt{2(\gamma - 1)}} = c, \ t < 0. \]

In this formulation, the solution simply seizes to make sense as a classical solution at time \( t = 0^- \). Nevertheless, we shall see later that it makes sense to think of \( t = 0^- \) as the first singularity and consider a singular curve in the \( r - t \) plane along which the density blows up. A rigorous description of this phenomenon necessitates the use of Lagrangian coordinates, see Section 3.

**Remark 2.8.** There is a parallel here to the dust and near-dust collapse from Sections 2.4.1 and 2.4.2 in that they also allow for an extension that allows us to attribute different collapse times to different particles inside the gas. A decisive difference however, is that in the case of self-similar collapse, the rates of blow-up at the scaling origin \( \mathcal{O} = (0, 0) \) (first time collapse point) and along the boundary of the maximal extension in the Lagrangian coordinates are different.

The basic challenge in the analysis of (2.80)–(2.81) is the possible presence of singularities in the denominator \( G(y; \rho, \omega) \) on the right-hand side. Any such point \( y_* \in [0, \infty) \) by definition satisfies
\[ \gamma\rho(y_*)^{\gamma - 1} - y_*^2\omega(y_*)^2 = 0, \]
and is referred to as a sonic point. Geometrically it corresponds to the boundary of the backward acoustical cone emanating from the singularity at \( (0, 0) \). Assuming smoothness at \( y = 0 \) and that the
solution is smooth and the density vanishes on approach to \(y = \infty\), a simple Taylor expansion analysis shows that we must have

\[
\rho \sim y^{-\frac{2}{2-\gamma}}, \quad \omega \sim y^{-1}, \quad \rho(0) > 0, \quad \omega(0) = \frac{4-3\gamma}{3}. \tag{2.87}
\]

**Remark 2.9.** There are in fact two explicit solutions to (2.80)–(2.81), each consistent with exactly one of the two boundary conditions (2.87)–(2.88). They are the far-field solution

\[
\omega_f = 2 - \gamma, \quad \rho_f = ky^{-\frac{2}{2-\gamma}}, \text{ where } k = \left(\frac{\gamma(4-3\gamma)}{2\pi(2-\gamma)^2}\right)^{\frac{1}{4-3\gamma}}, \tag{2.89}
\]

and the Friedman solution

\[
\omega_F = \frac{4-3\gamma}{3}, \quad \rho_F = \frac{1}{6\pi}. \tag{2.90}
\]

Note that \(\rho_f\) blows up as \(y \to 0^+\) and \(\rho_F\) does not decay to 0 as \(y \to \infty\). We may therefore, informally speaking, think of the smooth solution to (2.80)–(2.81) satisfying (2.87)–(2.88) as a heteroclinic connection between the far-field and the Friedman solution.

It is straightforward to see that any continuous solution \((\rho, \omega)\) satisfying (2.87)–(2.88) leads to the boundary behaviour \(G(0; \rho, \omega) > 0\) and \(\lim_{y \to \infty} G(y; \rho, \omega) = -\infty\). Therefore by the intermediate value theorem any \(C^1\)-solution \((\rho, \omega)\) has at least one sonic point \(\bar{y}_s \in (0, \infty)\). So by necessity, we are interested in smooth solutions that feature a sonic point. We shall see that this is challenging, as the requirement of smoothness places severe restrictions on the behaviour of the solution at the sonic point.

In a recent pioneering work [62] Merle, Raphaël, Rodnianski, and Szefetl systematically constructed smooth self-similar solutions to the compressible Euler equations with the \(\gamma\)-law (1.4) (for a.e. \(\gamma > 1\)). Their work gives the first rigorous construction of physically relevant \(C^\infty\) self-similar solutions featuring a sonic point in the setting of compressible fluids. The smoothness plays a crucial role in the dynamic stability analysis [63]; we emphasise that the underlying self-similar reduction gives an autonomous \(2 \times 2\) system of ODE in contrast to (2.80)–(2.81). For further existence results on smooth self-similar solutions of the Euler equation, see recent preprint [4].

When \(\gamma = 1\) the \(C^\infty\) solutions to the system (2.84)–(2.85) were first analysed numerically in the seminal works of Larson [54] and Penston [75] in 1969. These works recognised the importance such a solution may play for the description of the late stage of the gravitational collapse of a Newtonian star. It is commonly referred to as the Larson-Penston (LP)-solution. Further milestone in the subject were the works of Hunter [48] and Shu [84], which among other things, suggested numerically the existence of a discrete family of \(C^\infty\)-self-similar solutions. In this picture, the LP-solution can be interpreted as the “ground state” self-similar flow, which is expected to be the most stable member of the family. The next “excited” state, the so-called Hunter A solution, plays an important role in the Newtonian criticality theory [58, 33]. Numerically, the LP-solutions are distinguished as the only member of the discrete family where the radial velocity \(u\) is strictly positive. This implies that the flow is strictly inflowing in contrast to the excited solutions found by Hunter that exhibit regions of both positive and negative fluid velocity [48, 58, 5].

Our main result in this setting is the existence of an LP-solution

**Theorem 2.10 ([35]).** Let \(\gamma = 1\). Then there exists a real-analytic LP-solution \((\rho, \omega)\) to (2.84)–(2.85) satisfying the boundary conditions (2.87)–(2.88) with \(\gamma = 1\).
Naturally, the assumption \( \gamma = 1 \) may seem restrictive.\(^2\) In fact, the physical relevance of the LP and Hunter solutions was criticised by Yahil [92] on the basis of having both infinite mass and energy. It is of course a separate question, whether one can truncate or suitably modify the LP solution in the asymptotic region so that we obtain a collapsing star of finite mass and energy (this is indeed possible, the proof is implicit in the results of [36]). But, if one insists on exact self-similarity then there can be no finite mass collapsing flows when \( \gamma \in [1, \frac{4}{3}] \). However, when \( \gamma \in \left[ \frac{6}{5}, \frac{4}{3} \right] \) self-similar solutions to \((2.80)-(2.81)\) do have finite energy, which can be read off from the way scaling affects the total energy \((2.25)\). This is the energy-critical (\( \gamma = \frac{6}{5} \)) and the energy-subcritical (\( \gamma \in \left( \frac{6}{5}, \frac{4}{3} \right) \)) range in which Yahil [92] numerically established the existence of ground state analogues of the LP-solutions. We refer to them as Yahil solutions. Our next main result is the proof of existence of such solutions.

**Theorem 2.11** ([37]). For any \( \gamma \in (1, \frac{4}{3}) \) there exists a real-analytic Yahil-type solution \((\rho, \omega)\) to \((2.80)-(2.81)\) satisfying the boundary conditions \((2.87)-(2.88)\).

We devote the rest of this section to the main ideas behind the proofs of Theorems 2.10-2.11. We will for expository brevity focus on the proof of Theorem 2.10 and indicate some conceptual differences to the proof of Theorem 2.11. Theorem 2.11 is considerably more complex and its proof required some fundamental new ideas with respect to the proof of Theorem 2.10.

By a slight abuse of notation we can rewrite \((2.84)-(2.85)\) in the form

\[
\rho' = \frac{2y\rho\omega(\omega - \rho)}{1 - y^2\omega^2}, \\
\omega' = \frac{1 - 3\omega}{y} - \frac{2y\omega^2(\omega - \rho)}{1 - y^2\omega^2},
\]

where we changed the unknowns \( \rho \mapsto 2\pi \rho \). The proof of Theorem 2.10 contains three main steps.

**Step 1. Sonic point and the sonic window.** As explained above, any smooth solution satisfying the asymptotic behaviour \((2.87)-(2.88)\) must feature a sonic point, i.e. a point \( y_\ast \) where \( 1 - y_\ast^2\omega(y_\ast)^2 = 0 \). Such a singular point is a priori unknown and clearly depends on the solution itself. Our first task is to make an educated guess as to where one might expect such a point to be. It is not hard to see that the sonic point associated with the Friedman solution \((2.90)\) in the case \( \gamma = 1 \) is precisely \( y_\ast = 3 \). We choose 3 to be the right end-point of our sonic window, while the left end-point is chosen to be \( y_\ast = 2 \) for reasons that will become apparent below. Numerical investigations \([5, 58]\) suggest that the sonic point value associated with the actual LP-solution is roughly \( y_\ast \approx 2.41 \).

Assuming the existence of a putative sonic point \( y_\ast \in [2, 3] \), our first goal is to show the existence of a smooth solution of the dynamical system \((2.91)-(2.92)\) in a small neighbourhood \((y_\ast - \delta, y_\ast + \delta)\) for some \( \delta > 0 \). Assuming that there indeed exists a solution, it must have a Taylor expansion of the general form

\[
\rho(y) = \sum_{N=0}^{\infty} \rho_N(y - y_\ast)^N, \quad \omega(y) = \sum_{N=0}^{\infty} \omega_N(y - y_\ast)^N.
\]

Our key idea is to prove convergence of the above series to obtain a local-around-\( y_\ast \) real-analytic solution. Since the denominator \( 1 - y_\ast^2\omega(y_\ast)^2 \) vanishes we infer that i) \( \omega_0 = \frac{1}{y_\ast} \) and ii) as a consistency condition the numerator must vanish at \( y_\ast \), i.e. \( \omega_0 = \rho_0 \). In conclusion

\[
\rho_0 = \omega_0 = \frac{1}{y_\ast}.
\]

\(^2\)Understanding the isothermal case \( \gamma = 1 \) is morally important for the fully relativistic setting, where the linear relation \( p = \varepsilon \rho \) is indeed the correct equation of state at high densities \([43, 46, 41]\).
We then plug this information back into (2.91)–(2.92) and expand to the next order in order to obtain the values of $\omega_1$ and $\rho_1$. In doing so, an interesting dichotomy arises: it turns out that $\omega_1$ satisfies a quadratic equation and two choices/types of solutions emerge: $(\rho_1, \omega_1)$ is

- either of Type 1:
  \[
  \rho_1 = -\frac{1}{y_*}, \quad \omega_1 = -\frac{1}{y_*}(1 - \frac{2}{y_*}),
  \]
  \hspace{1cm} (2.94)

- or of Type 2:
  \[
  \rho_1 = \frac{1}{y_*} \left( 1 - \frac{3}{y_*} \right), \quad \omega_1 = 0.
  \]
  \hspace{1cm} (2.95)

It was recognised early [48] that the numerical solutions discovered by Larson and Penston are in fact of Type 1. Upon fixing the choice (2.94) it is then shown that for all $N \geq 2$ the Taylor coefficients $(\rho_N, \omega_N)$ are uniquely determined from the recursive relationship

\[
A_N \left( \rho_N, \omega_N \right) = \left( F_N G_N \right),
\]

where $F_N, G_N$ are source terms that depend only on $(\rho_k, \omega_k)$, $k \leq N - 1$ and

\[
A_N(\omega_0, \omega_1, \rho_1) = \left( \begin{array}{c}
-2N + 2 - 2N \frac{\omega_1}{\omega_0} \\
-2 - 2N - 4 + \frac{2\rho_1}{\omega_0} - (2N + 2) \frac{\omega_1}{\omega_0}
\end{array} \right).
\]

\hspace{1cm} (2.96)

**Remark 2.12.** A similar recursion can be derived if we make the choice (2.95) instead. The choice (2.95) subsequently led Hunter to numerically discover a whole discrete family of further smooth self-similar solutions, sometimes labelled as Hunter A, B, C, D, . . . solutions, mentioned above. They are particularly interesting as they provide a link to so-called criticality theory in gravitational collapse [33, 65]. It is an important open problem to construct Hunter solutions rigorously. They should be thought of as excited states in a discrete family of self-similar blow-up solutions, the ground state being the LP-solution. It is likely that the methods developed in [35, 37] will play an important role in such a construction.

**Remark 2.13.** A surprising outcome of the sonic-point analysis in the polytropic setting $\gamma > 1$ is that the two seemingly disjoint branches (2.94)–(2.95) lend themselves to a different interpretation. They can be naturally viewed as members of the same branch which exhibits a discontinuity in the limit as $\gamma \to 1^+$. In [35] it is shown that the formal power series (2.93) indeed converges on a small open interval around $y = y_*$. The challenge here is that there are no known general ODE results that can be cited to show the convergence in presence of sonic point singularities. Instead we resort to a combinatorial-type argument where we set up a careful inductive strategy and show that there exists a constant $C > 0$ such that

\[
|\rho_N|, |\omega_N| \leq \frac{CN}{N^2}, \quad N \geq 2.
\]

\hspace{1cm} (2.97)

This in turn implies the power series convergence. In conclusion, for any $y_* \in [2, 3]$ we obtain a solution

\[
(\rho(\cdot; y_*), \omega(\cdot; y_*))
\]

\hspace{1cm} (2.98)

defined on an interval $(y_* - \delta, y_* + \delta)$, where $\delta > 0$ can be chosen independently of $y_* \in [2, 3]$. By standard continuity arguments we extend such a solution to its maximal interval of existence to the left of the form $(s(y_*), y_*]$, where $s(y_*) \in [0, y_* - \delta]$ for all $y_* \in [2, 3]$. 
Step 2. The Friedman connection. Our goal is to identify a $y_* \in [2, 3]$ such that the local solution $(\rho(y_*), \omega(y_*))$ exists all the way to $y = 0$ and satisfies the boundary condition

$$\omega(0; y_*) = \frac{1}{3}. \tag{2.99}$$

To this end we develop an effective shooting method, which turns out to be a very robust idea applicable also in the relativistic case, see Section 3. Using the precise knowledge of the Taylor coefficients in the expansion for $(\rho, \omega)$ near the sonic point, one can show that for all $y_* < 3$, but sufficiently close to $y_* = 3$, the associated local LP-type solution “dips” below $\omega = \frac{1}{3}$. Assuming that $y'_3(y_*)$ is the supremum of all such points - i.e. the “first” point going to the left from the sonic point so that $\omega(y'_3(y_*) ; y_*) = \frac{1}{3}$, we see from (2.92) that

$$\omega'(y'_3) = -\frac{2y'_3(\frac{1}{3} - \rho(y'_3))}{9(1 - \frac{y'_3}{3})}. \tag{2.100}$$

On the other hand, by subtracting (2.91) from (2.92) we also get

$$(\omega - \rho)' = \frac{1}{y} - \frac{2y\omega}{1 - y^2\omega^2} (\rho + \omega). \tag{2.101}$$

Based on this relation and the strict inequality $\rho - \omega > 0$ which holds “initially” on a small interval $(y_* - \delta, y_*)$ to the left of the sonic point, we can show that $\rho > \omega$ on the maximal interval of existence (to the left) of the solution. It follows in particular that $\omega$ dips strictly below $\frac{1}{3}$ locally to the left of $y'_3$. If $\omega$ were to ever cross $\frac{1}{3}$ again on its maximal interval of existence, this would contradict the relation (2.100) and the strict inequality $\rho > \omega$. Therefore the region $\{y \mid \omega < \frac{1}{3}\}$ is invariant under the flow, and as a consequence any solution that does go below $\frac{1}{3}$ cannot be a regular solution of the problem as it cannot satisfy the boundary condition (2.99). This then suggests a strategy to construct a regular connection. We define $\tilde{y}_*$ via

$$\tilde{y}_* \equiv \inf_{y_* \in Y} y_* \tag{2.102}$$

where

$$Y \equiv \{ y_* \in [2, 3] \mid \exists y \in (s(\tilde{y}_*), y_*) \text{ such that } \omega(y ; \tilde{y}_*) = \frac{1}{3} \text{ for all } \tilde{y}_* \in [y_*, 3] \}. \tag{2.103}$$

In other words, $\tilde{y}_*$ is the “first” $y_*$ smaller than 3 such that the associated solution $\omega(y ; \tilde{y}_*)$ never takes on a value strictly smaller than $\frac{1}{3}$ on its maximal interval of existence to the left $(s(\tilde{y}_*), \tilde{y}_*) \subset [0, y_*]$. This “shooting” method is illustrated in Figure 4. With a little bit of extra work it is possible to show that this minimality property embedded in the definition (2.102) guarantees that $s(\tilde{y}_*) = 0$, i.e. the solution exists on the whole interval $[0, \tilde{y}_*]$. The final step is to show that the thus constructed solution $\omega(y ; \tilde{y}_*)$ satisfies the boundary condition (2.99) and does not, for example, blow up on approach to $y = 0$. In the work [35] this is accomplished through a series of rather technical lemmas, but the procedure can be considerably simplified by following the strategy introduced in the polytropic context [37]. One can in fact show that $\omega(y ; \tilde{y}_*)$ is also a monotonically increasing function (i.e. decreasing in value as we move to the left). The proof is delicate and uses both the minimality property (2.102)–(2.103) and nonlinear invariances in a nontrivial way, see [37] for details.

Step 3. The far-field connection. The final step in the proof of Theorem 2.10 is to ensure that we can solve (2.91)–(2.92) “to the right of the sonic point”, i.e. on the interval $[\tilde{y}_*, \infty)$ so that the asymptotic
behaviour (2.87) (with $\gamma = 1$) is honoured. This far-field connection turns out to be easier to construct than the Friedman connection from Step 2 above, and the main heuristic reason is that the flow is more stable to the right of the sonic point. We use basic invariances of the system to ensure that for any $y_\ast \in [2, 3]$ the solution $\omega(\cdot, y_\ast)$ stays trapped in the interval $\left(\frac{1}{3}, 1\right)$ for all $y > y_\ast$, as well as to ensure the inequality $\omega > \rho$ for all $y > y_\ast$ on the maximal interval of existence. With these two properties in place, it is not hard to see that the solution $(\rho(\cdot; y_\ast), \omega(\cdot; y_\ast))$ then exists globally on the interval $[y_\ast, \infty)$ for all $y_\ast \in [2, 3]$ and satisfies the boundary condition (2.87). Choosing $y_\ast = \bar{y}_\ast$ constructed in Step 2 above, we finally obtain a globally defined real analytic solution on $[0, \infty)$ satisfying the correct boundary conditions. The resulting flow is the Larson-Penston solution claimed in Theorem 2.10.

### 3 Self-gravitating relativistic gases

Our goal here is to explain the construction of a family of self-similar imploding relativistic solutions from [36]. These solutions are analogous to the Larson-Penston flows constructed above, but with a new twist. The analysis of the causal structure of such relativistic Larson-Penston spacetimes implies that the spacetime contains a so-called naked singularity. This notion on the other hand is intimately tied to the weak cosmic censorship hypothesis of Penrose [74, 19, 79]. We work here with a rigorous mathematical definition of naked singularities given by Rodnianski and Shlapentokh-Rothman - Definition 1.1 in [79]. In a nutshell, we say that a spacetime contains a naked singularity if

- it is a maximal global hyperbolic development with a complete Cauchy hypersurface,
- and it possesses an incomplete future null-infinity.

The latter by definition means that exists a uniform upper bound on the affine length of a sequence of (suitably normalised) ingoing geodesics approaching null-infinity, starting from a regular outgoing cone - see Figure 5.

#### 3.1 Formulation of the problem and the main result

Self-similar spacetimes $(M, g)$ are by definition characterised by the existence of a homothetic Killing vectorfield $\xi$ with the property

$$\mathcal{L}_\xi g = 2g.$$ 

We assume radial symmetry and wish to choose coordinates $(\tau, R)$ in a suitable conformal compactification so that $\xi = \tau \partial_\tau + R \partial_R$ indeed corresponds to the usual scaling vectorfield. Following the
pioneering work of Ori and Piran [70] we work with the so-called comoving coordinates and the metric takes the form

\[ g = -e^{2\mu(\tau,R)}d\tau^2 + e^{2\lambda(\tau,R)}dR^2 + r^2(\tau,R)\gamma. \] (3.104)

Here \( R \) is the comoving label, \( \gamma \) the standard metric on \( S^2 \), and the vectorfield \( \partial_\tau \) is chosen to be parallel to the 4-velocity \( u^\mu \), which in light of the normalisation (1.7) implies \( u = e^{-\mu}\partial_\tau \). In this choice of coordinates, the Einstein-Euler system takes the form

\[ \partial_\tau \rho + (1 + \varepsilon)\rho \left( \frac{\partial RV}{\partial R} + 2\frac{V}{\tau} \right) e^\mu = 0, \] (3.105)

\[ \partial_\tau \lambda = e^\mu \frac{\partial RV}{\partial R}, \] (3.106)

\[ e^{-\mu}\partial_\tau V + \frac{\varepsilon}{1 + \varepsilon} \frac{\partial R e^{-2\lambda}}{\rho} \partial R \rho + 4\pi \left( \frac{1}{3}G + \varepsilon \rho \right) = 0, \] (3.107)

\[ (\partial R)^2 e^{-2\lambda} = 1 + V^2 - \frac{8\pi}{3}Gr^2, \] (3.108)

where \( G(\tau,R) := \frac{m(\tau,R)}{4\pi r^3(\tau,R)} \) is the average mass, \( V := e^{-\mu}\partial_\tau r \) the radial velocity, and \( \varepsilon \) the square of the speed of sound (1.4). In this formulation it is very easy to see that (3.105)–(3.108) is invariant under the scaling transformation

\[ \rho \mapsto a^{-2}\rho(s,y), \enspace r \mapsto ar(s,y), \enspace V \mapsto V(s,y), \enspace \lambda \mapsto \lambda(s,y), \enspace \mu \mapsto \mu(s,y), \]

where the comoving “time” \( \tau \) and the label \( R \) scale according to

\[ s = \frac{\tau}{a}, \enspace y = \frac{R}{a}, \enspace a > 0. \]

By analogy to the Newtonian case this motivates us to look for self-similar solutions, with the fundamental new variable

\[ y = \frac{R}{-\sqrt{\varepsilon \tau}}. \]

\[ ^{1} \text{Up to our knowledge, self-similar formulation of the EE-system was first given in [86], see also [6]. For the derivation of the EE-system in comoving coordinates see for example [64, 26].} \]
We look for a solution to (3.105)–(3.108) of the form

\[ \rho(\tau, R) = \frac{1}{2\pi \tau^2} \Sigma(y), \quad r(\tau, R) = -\sqrt{\varepsilon \tau \tilde{r}(y)}, \quad V(\tau, R) = \sqrt{\varepsilon} V(y), \]

(3.109)

\[ \lambda(\tau, R) = \lambda(y), \quad \mu(\tau, R) = \mu(y). \]

(3.110)

**Remark 3.1.** If we can indeed find a solution of the form (3.109)–(3.110) such that \( \Sigma \) is bounded, then we conclude that \( \rho(\tau, R) \propto \frac{1}{\tau^2} \) as we approach the scaling origin \( \mathcal{O} \) and we thus have implosion.

Instead of working with the above variables directly, for the purpose of rigorous analysis, it is advantageous to work with the unknowns:

\[ d := \Sigma^{1+\varepsilon}, \]

\[ w := (1 + \varepsilon) \frac{e^{\mu V} + \tilde{r}}{\tilde{r}} - \varepsilon. \]

It can be then checked that \((d, w)\) solve

\[ d' = \frac{2(1 - \varepsilon)d(d - w)}{(1 + \varepsilon)y(e^{2\mu - 2\lambda y^2} - 1)}, \]

(3.111)

\[ w' = \frac{(w + \varepsilon)(1 - 3w)}{(1 + \varepsilon)y} + \frac{2w(d - w)}{y(e^{2\mu - 2\lambda y^2} - 1)}. \]

(3.112)

**Remark 3.2.** The system (3.111)–(3.112) is a non-autonomous \(2 \times 2\) system and the sonic point corresponds to a point \(y\) such that

\[ e^{2\mu - 2\lambda y^2} - 1 = 0. \]

(3.113)

The sonic point in fact corresponds to a hypersurface in the original spacetime \((\mathcal{M}, g)\), whose projection on the \((\tau, R)\)-plane is schematically portrayed in Figure 7. Just like in the Newtonian case, it corresponds to the boundary of the backward acoustical cone emanating from the scaling origin \(\mathcal{O}\).

**Remark 3.3.** An approach inspired by dynamical systems ideas and phase space analysis to describe and classify smooth self-similar solutions of the EE-system was presented in [30, 8].

The self-similar reduction (3.111)–(3.112) originates from a manifestly Lagrangian description of the flow. On the other hand, the Newtonian version of the problem (2.84)–(2.85) is derived from the Eulerian description. To link the two flows (for the moment just formally), we wish to formulate a suitable “Eulerian” self-similar reduction of the problem. There is however no canonical choice of such coordinates in general relativity, but a choice that is well-suited to our needs are the so-called Schwarzschild coordinates, also used by Ori and Piran [70]. At the level of self-similar reduction (3.111)–(3.112) this amounts to introducing new variables

\[ x := \tilde{r}(y), \]
\[ D(x) := d(y), \quad W(x) := w(y). \]

A calculation then reveals that
\[ D'(x) = -\frac{2x(1-\varepsilon)D(W+\varepsilon)(D-W)}{B}, \tag{3.114} \]
\[ W'(x) = \frac{(1-3W)}{x} + \frac{2x(1+\varepsilon)W(W+\varepsilon)(D-W)}{B}, \tag{3.115} \]
where
\[ B = B[x; D, W] := D - 2\varepsilon - [(W+\varepsilon)^2 - \varepsilon(W-1)^2 + 4\varepsilon DW] x^2, \tag{3.116} \]
Notice that in this description sonic points correspond to the \( x \in \mathbb{R} \) such that \( B(x) = 0 \). We make here an important observation - the formal Newtonian limiting system is obtained by setting \( \varepsilon = 0 \), which corresponds precisely to (2.84)–(2.85). This suggests that in the \( 0 < \varepsilon \ll 1 \) regime we might hope to solve (3.114)–(3.115) by emulating the strategy developed in Section 2.5. In that spirit first that by analogy to Remark 2.9 observe that there are two explicit solutions. The far-field solution is given by
\[ D_f(x) = (1 - \varepsilon)^{-\frac{2(1+\varepsilon)}{1+\varepsilon}} x^{-\frac{2(1+\varepsilon)}{1+\varepsilon}}, \quad W_f(x) = 1; \tag{3.117} \]
we note that the density \( D_f \) blows up at \( x = 0 \) and decays to 0 as \( x \to \infty \). By contrast, the Friedmann solution
\[ D_F(x) = \frac{1}{3}, \quad W_F(x) = \frac{1}{3}, \]
has bounded density at 0, but it does not decay as \( x \to \infty \). Our goal is to find a smooth solution connecting Friedmann to far-field, so that in particular
\[ \lim_{x \to \infty} W(x) = 1, \quad \lim_{x \to 0^+} W(x) = \frac{1}{3}. \]
It is easily checked that this requirement implies the existence of at least one sonic point.

**Theorem 3.4** ([36], Informal statement). *There exists an \( 0 < \varepsilon_0 < 1 \) such that for all \( 0 < \varepsilon \leq \varepsilon_0 \) there exists a relativistic Larson-Penston solution (from now on RLP solution) with the following properties:

1. The maximal analytic extension of the self-similar spacetime has a singular boundary consisting of two components: the scaling origin \( O = (0,0) \) and the Massive Singularity \( MS_\varepsilon \) depicted in Figure 7. On approach to the scaling origin the fluid density / Ricci scalar blow up like \( \tau^{-2} \);
2. There exists exactly one sonic line - the boundary of the backward acoustical cone emanating from \( O \), and exactly one null-curve \( N \) corresponding to the boundary of the backward null cone emanating from \( O \);
3. There exist at least two future oriented null-geodesics \( B_1 \) and \( B_2 \) emanating from the scaling origin \( O \) with the property that \( r \) increases to \( \infty \) as the affine parameter tends to \( \infty \). Moreover, they are similarity curves of the spacetime (i.e. straight lines in the \( R-\tau \) diagram) and \( B_1 \) is the “first” such outgoing null geodesic in the sense that any outgoing null-geodesic initiated in the causal past of \( B_1 \) does not cross \( B_1 \) nor it goes to/out of \( O \), see Figure 7.

Moreover, the solution can be flattened in the asymptotic region to give examples of naked singularity spacetimes. More precisely, for any \( 0 < \varepsilon \leq \varepsilon_0 \) there exists a family of solutions to the Einstein-Euler system with the Penrose diagram like in Figure 8. The solution coincides with the exact self-similar RLP spacetime in the causal past of the curve \( N \), is asymptotically flat, and possesses an incomplete future null infinity. The singularity at \( O \) is visible to distant observers and the Cauchy Horizon \( \mathcal{CH}^+ \) corresponds to the future null-geodesic emanating from \( O \) as in Figure 8, which effectively terminates the future null-infinity \( I^+ \).
Remark 3.5. Our proof is nonperturbative - it is not the case that we use for example an implicit function theorem to perturb away from the Newtonian solution and construct the relativistic one. The nonlinear structure of the relativistic problem is rather delicate and does not lend itself to an easy perturbation argument. Instead, while our results do use the smallness of \( \epsilon \) in a crucial way, we do so to enforce the validity of various nonlinear invariances of the ODE flow, catered to the problem.

3.2 Strategy of the proof and main ideas

The proof of Theorem 3.4 proceeds in six steps. The first four steps are devoted to the construction of the maximally extended RLP spacetime, in step 5 we study the causal structure of the self-similar spacetime, and finally in step 6 we explain how to flatten the self-similar solution asymptotically to obtain a spacetime with naked singularity in line with Figure 8. Structurally, steps 1-3 echo the construction of the Larson-Penston spacetime explained in Section 2.5 - step 2 corresponds to the relativistic version of the Friedman shooting argument, while step 3 establishes the existence of the far-field connection, see Figure 9. From the technical point of view, both steps are considerably more challenging than in the Newtonian case due to the highly nonlinear structure of the right-hand side of (3.114)--(3.115). Step 4 is new by comparison to the Larson-Penston analysis, as it provides an extension of the solution into the upper-half plane (i.e. positive times), which by definition is only possible in a suitable choice of comoving coordinates (i.e. not in the “Eulerian” description). This leads to a nonstandard ODE problem and the question of how to extend the solution uniquely. In step 5 we demonstrate the existence of outgoing null-geodesics \( B_1 \) and \( B_2 \) and in Step 6 we show that \( B_1 \) can be suitably deformed in the asymptotic region to correspond to the Cauchy horizon of an asymptotically flat spacetime with an incomplete future null-infinity in the sense of Christodoulou and Rodnianski–Shlapentokh-Rothman [18, 79].

Step 1. Sonic point \( x_* \) is an a priori unknown singular point for the flow. By analogy to Step 1 of the proof of Theorem 2.10, we make an educated guess \( x_* \in [2, 3] \). This guess turns out to be slightly too
optimistic and one instead works on the interval $[x_{\min}(\varepsilon), x_{\max}(\varepsilon)]$ where $\lim_{\varepsilon \to 0}[x_{\min}(\varepsilon), x_{\max}(\varepsilon)] = [2, 3]$, which is consistent with the Newtonian limit. The exact formulas for $x_{\min}(\varepsilon), x_{\max}(\varepsilon)$ are not so important - they play the same role as 2 and 3 do in the Newtonian problem in ensuring later on that our shooting argument works. We prove the local existence on an interval of the form $(x^\ast - \delta, x^\ast + \delta)$, where $\delta > 0$ is independent $x^\ast$. We do this by considering the Taylor series expansion

$$D = \sum D_N (x - x^\ast)^N, \quad W = \sum W_N (x - x^\ast)^N$$

at the sonic point $x^\ast$, anticipating that the requirement of smoothness will introduce a hierarchy of constraints for the sequence $(D_N, W_N), N \geq 0$. It is immediate from (3.114)–(3.115) that we must have $W_0 = D_0$. By analogy to the Newtonian case, the next order coefficient $W_1$ is shown to satisfy a quadratic equation with $W_0$-dependent coefficients. Much like in the Newtonian case, the two choices are the relativistic analogues of choosing a Larson-Penston type or a Hunter-type expansion. In either case, the resulting sequence is uniquely specified once this choice is made, and we obtain a recursive relation of the form

$$A_N \begin{pmatrix} D_N \\ W_N \end{pmatrix} = \begin{pmatrix} S_{1N}^1 \\ S_{1N}^2 \end{pmatrix}, \quad N \geq 2,$$

where $S_{1N}^j$ are source terms that depend only on $(D_k, W_k), k \leq N - 1$ and $A_N$ is an invertible matrix whose entries depend explicitly on $x^\ast$ and $N \in \mathbb{N}$.

Remark 3.6. Note that the existence of Hunter-type solutions is not pursued in our work [36]. By analogy to the Newtonian case, see Remark 2.12, the Hunter family of solutions provides a link to the criticality theory for relativistic gases [27, 8, 43, 44, 33], which has its roots in the well-known numerical work of Choptuik [15] on the collapse of massive scalar field.

Steps 2 and 3: Friedman and far-field connections. Inspired by our Newtonian approach we wish to develop an effective shooting argument which will allow us to select a value $\bar{x}^\ast$, so that the local solution around $\bar{x}^\ast$ of RLP-type extends all the way to the left to $x = 0$. We refer to it as the Friedman connection as any such solution, if smooth, must attain the value $W(0) = \frac{1}{3}$. Similarly, in order to extend the solution to the right of the sonic point all the way to $x = \infty$ it must qualitatively behave like the far-field solution (3.117) - hence far-field connection. The proof to the left relies on a shooting procedure analogous to the one depicted in Figure 4 which allows us to single out a specific value of $\bar{x}^\ast$ with the right properties. The analysis to the right is less sensitive to the choice of $x^\ast$ and indeed we show that for any value of $x^\ast \in [x_{\min}(\varepsilon), x_{\max}(\varepsilon)]$ there exists a global solution on $[x^\ast, \infty)$. However, in the relativistic case it is considerably harder to show that the sign of the sonic denominator $B$ remains strictly positive (resp. negative) for all $x \in (0, \bar{x}^\ast)$ (resp. $x \in (x^\ast, \infty)$), by comparison to
the Newtonian case. A key new technical tool in our analysis is a Monotonicity Lemma designed to capture a novel nonlinear invariance in the problem, in the regime where $0 < \varepsilon \ll 1$.

We first recall the sonic denominator (3.116)

$$B = D^{-\frac{2\varepsilon}{1 - \varepsilon}} - [(W + \varepsilon)^2 - \varepsilon(W - 1)^2 + 4\varepsilon DW] x^2.$$  

To control its sign we introduce the following simple algebraic decomposition:

$$B[x; D, W] = [J - xW(x)] [H + xW(x)],$$

where

$$J := -\frac{2\varepsilon}{1 - \varepsilon}(1 + D)x + \sqrt{\frac{4\varepsilon^2}{(1 - \varepsilon)^2}(1 + D)^2 x^2 + \varepsilon x^2 + \frac{D - \frac{2\varepsilon}{1 - \varepsilon}}},$$

$$H := J + \frac{4\varepsilon}{1 - \varepsilon}(1 + D)x.$$

It is now clear that the sign of $B$ corresponds to the sign of $J - xW$. The key quantity to monitor turns out to be

$$f(x) := J - xD,$$  

and a direct calculation [36] shows that for any smooth solution of (3.114)–(3.115) we have

Lemma 3.7 (Monotonicity Lemma).

$$f' + a[x; D, W]f = b[x; D, W],$$

where $b$ has “good sign properties in the regime we care about”.

To show that the sign of $B$ remains positive to the left of the sonic point, we then show with the help of Lemma 3.7 that the following “sandwich” relations

$$xW < xD < J, \quad x < x_*$$  

$$xW > xD > J, \quad x > x_*$$

are dynamically propagated to the left (resp. right) of $x = \bar{x}_*$. It then remains to show that it is also satisfied in a small left (resp. right) neighbourhood of $x = \bar{x}_*$ which is accomplished using the precise information on the Taylor coefficients obtained in Step 1.

Step 4. Maximal analytic extension. The metric $g$ becomes singular on approach to the horizontal axis $\tau = 0$, i.e. as $x \to \infty$ or equivalently $y \to \infty$. We prove this by using the precise asymptotic behaviour as $x \to \infty$ and show that

$$e^{2\mu} \asymp y \to \infty y^{\frac{1}{1 + \varepsilon}},$$

where we have conveniently reverted back to the Lagrangian variable $y$. The goal is to show that the singular behaviour (3.121) is merely a coordinate singularity and it is possible to extend the metric smoothly (in fact analytically) across the surface $\tau = 0$ into the upper half plane. We achieve this by introducing adapted comoving coordinates:

$$Y = y^{-\frac{1}{1 + \varepsilon}},$$
and a new set of unknowns

\[
\chi(Y) := \frac{\tilde{r}(y)}{y}, \quad d(Y) := d(y), \quad w(Y) := w(y).
\] (3.123)

The choice (3.122) of $Y$ is clearly motivated by (3.121). A simple calculation reveals that the system (3.111)–(3.112) in the new variables takes the form

\[
d' = \frac{2\chi^2}{Y} d(w + \varepsilon)^2(d - w),
\] (3.124)

\[
w' = -\frac{(w + \varepsilon)(1 - 3w)}{(1 - \varepsilon)Y} - \frac{2(1 + \varepsilon)\chi^2 w(w + \varepsilon)^2(d - w)}{(1 - \varepsilon)Y C},
\] (3.125)

\[
\chi' = \frac{1 - w}{(1 - \varepsilon)Y} \chi,
\] (3.126)

where

\[
C := (dY^{-2})^{-n} Y^2 - \chi^2 \left[(w + \varepsilon)^2 - \varepsilon(w - 1)^2 + 4\varepsilon wd\right]
\]

is the sonic denominator in the new variables. The sharp asymptotic understanding of the solution as $y \to +\infty$ allows us to feed “initial data” for the new system of ODE at $Y = 0$ ($y = \infty$) and show a local-in-$Y$ well-posedness result around $Y = 0$. However, to extend the solution to the maximal interval of existence to the left of $Y$ we then prove the following key lemma.

**Lemma 3.8** (The sandwich lemma). *On the maximal interval of existence to the left of $Y = 0$, we have the dynamic “sandwich” bound:*

\[
1 \lesssim \frac{d}{w} < 1, \quad \text{for } Y < 0.
\]

The proof of Lemma 3.8 relies on careful exploitation of nonlinear invariances of (3.124)–(3.126). This in turn allows us to show that the maximal interval of existence is characterised by the simultaneous blow up of $d, w$, and $\frac{1}{\chi}$, see Figure 10, which is proved by showing that $w$ solves a Riccati type equation.

**Step 5. Causal structure and the existence of an outgoing null geodesic from $O$.**

To complete the proof of Theorem 3.4 we must show that there exists an outgoing radial null geodesic (RNG), denoted by $B_1$ in Figure 7, emanating from $O$ and reaching infinity. To that end we recall the metric expression (3.104) - an RNG by definition satisfies $\left(\frac{dR}{d\tau}\right)^2 = e^{2\mu - 2\lambda}$. Here we switched back to the comoving coordinates $(\mu, \lambda, y)$ which continue to make sense in the upper half-plane strictly away from the horizontal axis $\tau = 0$. We first restrict our attention to a particular
subclass of such geodesics that are simultaneously similarity lines in the problem so that there exists a \( \sigma \in \mathbb{R} \) such that \( R(\tau) = \sigma \tau \) (we refer to them as simple RNG-s following the terminology of Ori and Piran [70]). This trivially implies the relation

\[
y(\tau) = \frac{R(\tau)}{\sqrt{\varepsilon \tau}} = -\frac{\sigma}{\sqrt{\varepsilon}},
\]

so we conclude that \( \frac{\sigma}{\sqrt{\varepsilon}} \) is a root of

\[
y \mapsto F_{\varepsilon}(y)
\]

\[
F_{\varepsilon}(y) := \varepsilon y^2 e^{2\lambda - 2\mu} - 1
\]

A simple, but effective idea to prove the existence of a root of \( F_{\varepsilon} \) is to use intermediate value theorem and goes back to Ori and Piran [70]. The sharp asymptotic description of the self-similar flow from Steps 3 and 4 allows us to show that

- \[ \lim_{y \to y_{\text{ms}}} F_{\varepsilon}(y) = \lim_{y \to -\infty} F_{\varepsilon}(y) = \infty \] (3.127)
- and there exists a \( y_0 \in (-\infty, y_{\text{ms}}) \) such that
  \[ F_{\varepsilon}(y_0) < 0 \] (3.128)

if \( \varepsilon \) is sufficiently small.

By the intermediate value theorem there must exist at least two slopes \( y_1 \) and \( y_2 \) corresponding to outgoing null-geodesics, which correspond precisely to curves \( B_1 \) and \( B_2 \) on Figure 7. We emphasise here that the proof of (3.127) relies on the leading order asymptotics of the dynamical system (3.124)–(3.126) on approach to the terminal point \( Y = Y_{\text{ms}} \), or equivalently \( y = y_{\text{ms}} \), while the second limit claimed in (3.127) mirrors the coordinate singularity at the horizontal axis \( \{ \tau = 0 \} \) explained in Step 4. The key is therefore to establish (3.128), which schematically speaking, follows from very precise \( \varepsilon \)-independent bounds on the solution of (3.124)–(3.126) in a small (but \( \varepsilon \)-independent) left neighbourhood of \( Y = 0 \) (or equivalently \( y = -\infty \)), see Figure 11. The full description of (not necessarily simple) RNG-s as well as nonradial null-geodesics can be found in Section 8 and Appendix B of [36]. We note that after the publication [70], nonradial null-geodesics were analysed by Joshi and Dwivedi [53], see also the work of Carr and Gundlach [9].

The above analysis shows that informally \( B_1 \) is the first outgoing null-geodesic that “carries information” from the singularity \( \mathcal{O} \) to infinity. In particular, the extension that we have constructed all the way to the massive singularity \( \mathcal{M}_{\varepsilon} \) is in general not unique and other extensions pass \( B_1 \) in principle possible. Our extension however is unique in the class of exact self-similar extensions - this situation parallels the one in [17].

**Step 6. Asymptotic flattening and naked singularities.**

The exactly self-similar RLP spacetime constructed above is, due to simple scaling considerations, not asymptotically flat. In fact, one can easily show that

\[
m(\tau, r) \asymp r \rightarrow \infty \ r,
\]

where \( m(\tau, r) \) is the local mass and therefore the total ADM-mass of the RLP-spacetime is infinite. One way to produce an asymptotically flat spacetime with a naked singularity is to truncate it in the asymptotic region (like in [17]), but this is tricky as the Euler-part of the evolution becomes highly degenerate in the presence of vacuum. Instead, we resort to a tail dampening procedure so to produce a density profile that decays sufficiently fast at null-infinity to give us an asymptotically flat spacetime.
To do this, we employ the so-called double null gauge in which the metric takes the form

\[ g = -\Omega^2 \, dp \, dq + r^2 \gamma, \quad (3.129) \]

where surfaces of constant \( p \) are outgoing null-surfaces and surfaces of constant \( q \) are the ingoing null-surfaces. Our goal is to recast Einstein-Euler system in this gauge and formulate the corresponding characteristic value problem in a semi-infinite characteristic rectangle depicted by the grey shaded region \( D \) in Figure 12. The Einstein field equations in double-null gauge and radial symmetry take the form \([23, 36]\]

\[ \partial_p \partial_q \Omega = -\frac{\Omega^2}{4r} - \frac{1}{r} \partial_p r \partial_q r + \pi r \Omega^4 T^{pq}, \quad (3.130) \]

\[ \partial_q (\Omega^{-2} \partial_q r) = -\pi r \Omega^2 T^{pp}, \quad (3.132) \]

\[ \partial_p (\Omega^{-2} \partial_p r) = -\pi r \Omega^2 T^{qq}. \quad (3.133) \]

The fluid 4-velocity in the new frame is written as

\[ u = u^p \partial_p + u^q \partial_q \]

and one can readily compute the components of the energy-momentum tensor:

\[ T^{pp} = (1 + \varepsilon) \rho (u^p)^2, \quad T^{vv} = (1 + \varepsilon) \rho (u^q)^2, \quad T^{pq} = (1 - \varepsilon) \rho \Omega^{-2}, \quad (3.134) \]

\[ T^{AB} = \varepsilon \rho r^{-2} \gamma^{AB}, \quad T^{pA} = T^{qA} = 0, \quad A, B = 2, 3. \quad (3.135) \]

The normalisation condition \((1.7)\) now reads

\[ T^{pp} T^{qq} = \left( \frac{1 + \varepsilon}{1 - \varepsilon} \right)^2 (T^{pq})^2. \quad (3.136) \]

To complete the formulation of the problem, we must explain how the fluid unknowns \((\rho, u^p, u^q)\) evolve in time. The corresponding Euler evolution is a simple consequence of the Bianchi identities and reads:

\[ \partial_p (\Omega^4 r^2 T^{pp}) + \frac{\Omega^2}{r \Omega^4} \partial_q (\Omega^2 r^{\frac{2+2\varepsilon}{4r}} T^{pq}) = 0, \quad (3.137) \]

\[ \partial_p (\Omega^2 r^{\frac{2+2\varepsilon}{4r}} T^{pq}) + r^{\frac{4\varepsilon}{\Omega^2}} \partial_q (\Omega^4 r^2 T^{qq}) = 0. \quad (3.138) \]
Equations (3.137)–(3.138) are nonlinear transport equations and the full coupling (3.130)–(3.138) is a nonlinear system of coupled wave- and transport-like equations. By the finite speed of propagation, we may hope to modify the RLP-solution in an asymptotic region in such a way that the solution is exactly self-similar and agrees with the constructed RLP-solution in a region in the causal past of some ingoing curve $\mathcal{C}$ (the white region in Figure 12), and otherwise solve a characteristic initial value problem in the gray shaded region $\mathcal{D}$. To make this work, we prescribe the data along the two characteristic characteristic curves $\mathcal{C}$ and $\mathcal{C}'$, so that the constraint equations (3.132)–(3.133) are satisfied respectively.

The final and key technical step is the proof of local well-posedness for this characteristic value problem. Although various local well-posedness results for the Einstein-Euler system are available (see for example [13, 77, 26, 3] and references therein), we have not been able to find such a reference that treats isothermal EE-system in the double-null coordinate gauge. To do this, we make a change of fluid unknowns that effectively “diagonalises” (3.137)–(3.138) in the spirit of Riemann variables in the theory of hyperbolic conservation laws. We do not present the details (an interested reader can see Section 9.1 of [36]), but the key geometric feature of this reformulation is to highlight the relationship of the two characteristic geometries in the problem - the light and the acoustical null-cones, see Figure 9 in [36].

Finally, we impose data on the outgoing piece of the boundary $\mathcal{C}$ that are asymptotically flat and belong to a suitable weighted Sobolev space, while the data on the ingoing piece coincide precisely with the self-similar RLP-solution. Using the method of characteristics we prove local-in-$\delta$ well-posedness of the characteristic value problem, where $\delta$ is the width of the semi-infinite characteristic rectangle $\mathcal{D}$. The resulting spacetime, obtained by gluing the solution inside $\mathcal{D}$ with the exact self-similar solution is a classical solution by the finite speed of propagation with a Penrose diagram given by Figure 8. It is then not hard to show that the Cauchy horizon of the resulting spacetime is but a suitable “deformation” of $B_1$ from Figure 7 and that it contains a naked singularity in the sense of Definition 1.1 in [79], see also Figure 5.

References


