Compactness bounds of spherically symmetric static objects

Lorena A. Contreras*

Department of Applied Mathematics, University of Waterloo,

Waterloo, Ontario, N2L 3G1, Canada

Abstract

We review the problem of how compact a static, spherically symmetric matter configuration can be without collapsing. After introducing the necessary background from differential geometry and general relativity, we obtain the most general metrics describing static, spherically symmetric spacetimes, both in the vacuum and in the presence of matter. We then prove Buchdahl's limit, which bounds the mass that a perfect fluid with non-negative energy density can hold in a stable configuration of hydrostatic equilibrium. Finally, we compile a collection of other compactness bounds that result from adopting different sets of assumptions than those used in Buchdahl's theorem. We conclude with an outlook on potential directions for future research on this topic.

CONTENTS

* lorena.contreraslemus@uwaterloo.ca

I. INTRODUCTION

Black holes are some of the most puzzling predictions of general relativity, and the existence of singularities is one of the many motivations to develop a quantum theory of gravity (where we would expect those singularities to become regular). However, even six years after the publication of the first picture of a black hole [1], deciding whether the dark, compact objects we observe are actual black holes in the sense of general relativity is still an open question (that keeps attracting interest [2–4]).

Currently, one of the main arguments in favour of general relativistic black holes lies precisely in that there seem to be upper bounds for how compact a star made of regular matter can be without collapsing, as well as for how redshifted is the light it emits. One example of this is Chandrasekhar's limit for the mass of white dwarf and neutron stars [5–7].

In this review, we will focus on Buchdahl's limit and its generalizations, which can be understood as mass limits, much like Chandrasekhar's, but applicable to static, spherically symmetric systems. The problem of studying the maximum compactness of static spherically symmetric objects remains open, as it has become increasingly clear that the assumptions underlying Buchdahl's limit are too restrictive. Even the assumption of non-negative energy density (which, as we will see, all generalizations still adopt) fails when quantum fields are considered [8]. Indeed, quantum fields can exhibit arbitrarily negative energy densities, although they are constrained by so-called quantum energy inequalities [9–11]. All in all, the situation calls for a systematic investigation of what happens when Buchdahl's assumptions are relaxed and replaced by alternative, possibly more realistic, conditions.

While ultimately this review aims to provide an overview of the question of how compact a configuration of static, spherically symmetric matter fields can be, we have taken a predominantly pedagogical approach. We have provided all the necessary tools (even if in compressed form) to study the problem and understand the results, with the goal of making the presentation as self-contained as possible. To this end, we start in Sec. II by reviewing essential notions of differential geometry and general relativity. In

Sec. III, we use those tools to study the general structure of static spherically symmetric spacetimes, both in the vacuum and in the presence of static matter. After this, in Sec. IV we introduce and prove the seminal contribution to this problem: Buchdahl's limit. In Sec. V, we offer a compilation of the compactness bounds that have been obtained since Buchdahl's work under different sets of assumptions, including some results regarding the formation of trapped surfaces (or rather its lack thereof) in these setups. We conclude in Sec. VI by highlighting several directions that, in our opinion, are worth pursuing to further advance this line of research.

Before proceeding, an observation regarding notation: at all points, we will be using natural units $G = c = 1$.

II. BACKGROUND

In this section, we review the fundamental tools of differential geometry and introduce the basic concepts of general relativity that will be necessary for studying bounds on the compactness of spherically symmetric, static objects. Our presentation loosely follows Chapters 2, 3, and 4 of Wald's book [12].

A. Differential geometry

We start by introducing manifolds and tangent spaces, which provide the underlying structure for all the remaining constructions. Vectors, dual vectors, and tensors, as well as their corresponding fields are subsequently defined, along with the notation that we will use throughout the text. Then, we move on to present the crucial concept of a metric, and, after introducing derivative operators and Christoffel symbols, we show how they can be made compatible with the metric. This allows us to define geodesics. We finish the review with the definition and properties of the Riemann tensor and its related quantities (namely, the Ricci tensor, and the curvature and Kretschmann scalars).

An *n-dimensional (smooth) manifold* $\mathcal{M}$ is a topological space endowed with an open cover $\{U_\alpha : \alpha \in I\} \subset \mathcal{M}$ (for some set of indices I) such that:

1. For each $\alpha \in I$, there exists a homeomorphism ψ_α between U_α and an open subset $O_\alpha \subset \mathbb{R}^n$ (where here $\mathbb{R}^n$ is understood to have the standard topology). Each homeomorphism ψ_α is called a *coordinate system*.

2. For every $\alpha, \beta \in I$ with $U_\alpha \cap U_\beta \neq \emptyset$, the homeomorphism $\psi_\beta \circ \psi_\alpha^{-1} : \psi_\alpha(U_\alpha \cap U_\beta) \to \psi_\beta(U_\alpha \cap U_\beta)$ is actually a smooth ($\mathcal{C}^\infty$) diffeomorphism, in the usual sense of real analysis. We then say that the coordinate systems are compatible with each other and that they form an *atlas*.

For an atlas to define a manifold, we require it to include all coordinate systems compatible with it.[1] The compatibility between different coordinate systems allows us to identify the infinitesimal structure of $\mathcal{M}$ with the infinitesimal structure of the subsets of $\mathbb{R}^n$ it is locally homeomorphic to, endowing it with a notion of differential calculus: for instance, we say that two manifolds $\mathcal{M}$ and $\mathcal{M}'$ are *diffeomorphic* if there exists a homeomorphism $f : \mathcal{M} \to \mathcal{M}'$ such that, for every pair of coordinate systems ψ_α and ψ'_β, the homeomorphism $\psi'_\beta \circ f \circ \psi_\alpha^{-1} : O_\alpha \to O'_\beta$ is in fact a smooth diffeomorphism.

To truly endow a manifold with a notion of differential calculus, it is necessary to define a notion of infinitesimal displacements, i.e., a notion of *tangent vector* at a point $p \in \mathcal{M}$. In order to do so, we first introduce the notion of *derivation*. Let $\mathcal{C}_p^\infty$ be the space of all smooth real functions defined on a neighbourhood of p. A derivation at p is a linear map $D : \mathcal{C}_p^\infty \to \mathbb{R}$ that satisfies the Leibniz's rule:

$$D(fg) = f(p)D(g) + g(p)D(f), \quad \forall\, f, g \in \mathcal{C}_p^\infty. \tag{1}$$

Now, denote with $(x^1, \ldots, x^n)$ the canonical coordinates of $\mathbb{R}^n$, and let $\psi : U \to \mathbb{R}^n$ be a coordinate system with $p \in U$, so that we denote $\psi = (\psi^1, \ldots, \psi^n)$. The vector space of derivations at a point p is then isomorphic to the space of directional derivatives in $\mathbb{R}^n$ through the following correspondence:

$$D(f) = \sum_{\mu=1}^{n} D(\psi^\mu)\, \partial_\mu f|_p, \quad \partial_\mu f|_p := \frac{\partial}{\partial x^\mu}(f \circ \psi^{-1})\Big|_{\psi(p)}. \tag{2}$$

[1] This requirement avoids the undesirable situation where two manifolds are considered different simply because they have different but compatible atlases.

Because the space of directional derivatives in $\mathbb{R}^n$ is isomorphic to its tangent space, we finally identify the derivations with the tangent vectors, and therefore the space of derivations with the tangent space at p, which we denote with $T_p\mathcal{M}$, of which $\{\partial_1, \dots, \partial_n\}$ is a coordinate basis. A map v that assigns a tangent vector $\mathsf{v}_p \in T_p\mathcal{M}$ to each point of the manifold $p \in \mathcal{M}$ defines a *vector field*, which we understand to be smooth if and only if, for each $p \in \mathcal{M}$, $\mathsf{v}(f)$ is smooth (on its domain) for all $f \in \mathcal{C}^\infty_p$. Because the coordinate fields $\{\partial_1, \dots, \partial_n\}$ can be defined by Eq. (2) in U, in such a neighbourhood of p we can write

$$\mathsf{v} = \sum_{\mu=1}^{n} \mathsf{v}(\psi^\mu)\partial_\mu \equiv v^\mu \partial_\mu, \tag{3}$$

where v^μ is the μ-th coordinate component of v with respect to the coordinates $(x^1, \dots, x^n)$, and we dropped the summation sign using Einstein's notation. Notice that, because the coordinate fields are trivially smooth, Eq. (3) implies that v is smooth if and only if each v^μ is a smooth function. Moreover, if we had been given a different coordinate system $\psi' : U \to \mathbb{R}^n$, with associated canonical coordinates $(x'^1, \dots, x'^n)$, then the components v'^ν of v with respect to the new coordinates are related to the old ones by

$$v'^\nu = \frac{\partial x'^\nu}{\partial x^\mu} v^\mu, \quad \text{where} \quad \frac{\partial x'^\nu}{\partial x^\mu} := \frac{\partial}{\partial x^\mu}(\psi'^\nu \circ \psi^{-1}). \tag{4}$$

Given the tangent space at a point p, $T_p\mathcal{M}$, we can define the *cotangent space*, $T^*_p\mathcal{M}$, as the space of linear maps defined from $T_p\mathcal{M}$ to $\mathbb{R}$, whose elements we call *dual vectors*. A map w^* that assigns a dual vector $\mathsf{w}^*_p \in T^*_p\mathcal{M}$ to each point of the manifold $p \in \mathcal{M}$ defines a *dual vector field*, which we understand to be smooth if an only if $\mathsf{w}^*(\mathsf{v})$ is smooth for every smooth vector field v. Now, from the coordinate vector fields $\{\partial_1, \dots, \partial_n\} \subset T_p\mathcal{M}$, one can define $\mathrm{d}x^\mu$ as the only dual vector field satisfying $\mathrm{d}x^\mu(\partial_\nu) = \delta^\mu_\nu$ (for each $\mu, \nu \in \{1, \dots, n\}$), so that

$$\mathsf{w}^* = \sum_{\mu=1}^{n} \mathsf{w}^*(\partial_\mu)\mathrm{d}x^\mu \equiv w^*_\mu \mathrm{d}x^\mu, \tag{5}$$

where, as before, w^*_μ denotes the μ-th coordinate component of w^* with respect to the coordinates $(x^1, \dots, x^n)$, and we used Einstein's convention to drop the summation sign, as will be standard practice in everything that follows. Again, we find w^* to be smooth if and only if w^*_μ are all smooth functions. Moreover, given another set of coordinates $(x'^1, \dots, x'^n)$, the components of w^* with respect to new basis relate to the old ones by

$$w'^*_\nu = \frac{\partial x^\mu}{\partial x'^\nu} w^*_\mu, \quad \text{where} \quad \frac{\partial x^\mu}{\partial x'^\nu} \coloneqq \frac{\partial}{\partial x'^\nu}(\psi^\mu \circ \psi'^{-1}). \tag{6}$$

From Eqs. (3) and (5), and by the definition of the dual coordinate fields, we have that

$$\mathsf{w}^*(\mathsf{v}) = w^*_\mu v^\mu. \tag{7}$$

Here we notice there is some symmetry between the role played by the vector and the dual vector: indeed, a tangent vector $\mathsf{v}_p \in T_p\mathcal{M}$ can in fact be understood as a *double dual vector* $\mathsf{v}^{**}_p \in T^{**}_p\mathcal{M}$, through

$$\mathsf{v}_p \equiv \mathsf{v}^{**}_p : T^*_p\mathcal{M} \to \mathbb{R},\ \mathsf{w}^*_p \mapsto \mathsf{w}^*_p(\mathsf{v}_p). \tag{8}$$

More generally, vector fields can be put in one-to-one correspondence with double dual vector fields by identifying $\mathsf{v} \equiv \mathsf{v}^{**}$, where

$$\mathsf{v}^{**}(\mathsf{w}^*) \coloneqq \mathsf{w}^*(\mathsf{v}), \tag{9}$$

for all dual vector fields w^*, and in particular $v^\mu = \mathsf{v}(\mathrm{d}x^\mu)$. The idea that dual vectors are linear maps on vectors, and vectors linear maps on dual vectors generalizes naturally to the concept of tensor.

A *tensor of type* (k, l) at a point $p \in \mathcal{M}$ is a multilinear map

$$\mathsf{S}_p : (T^*_p\mathcal{M})^k \times (T_p\mathcal{M})^l \to \mathbb{R}. \tag{10}$$

Notice that, with this definition, vectors and dual vectors are tensors of type $(1, 0)$ and $(0, 1)$, respectively. As before, a map S that assigns a tensor S_p of type (k, l) to each point $p \in \mathcal{M}$ defines a *tensor field*, which we understand to be smooth if and only if its output after applying it to smooth vector and dual vector fields is always a smooth function. To specify explicitly the type of a tensor, we imitate the notation followed for vectors and dual vectors, by which the former have upper indices and the latter have lower indices. Thus, a tensor S of type (k,l) is denoted as $S^{a_1\ldots a_k}{}_{b_1\ldots b_l}$, where we used the "abstract index notation", which uses Latin indices to distinguish the tensor as an object from its components, which are coordinate-dependent and are denoted with Greek indices. Specifically, from Eqs. (3) and (5), we denote

$$S^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} = \mathsf{S}(\mathrm{d}x^{\mu_1},\ldots,\mathrm{d}x^{\mu_k},\partial_{\nu_1},\ldots,\partial_{\nu_l}), \tag{11}$$

so that

$$\mathsf{S}(\mathsf{w}_1^*,\ldots,\mathsf{w}_k^*,\mathsf{v}_1,\ldots,\mathsf{v}_l) = S^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} w_{1,\mu_1}\cdots w_{k,\mu_k} v_1^{\nu_1}\cdots v_l^{\nu_l}, \tag{12}$$

and under a change of coordinates transforms as

$$S'^{\lambda_1\ldots\lambda_k}{}_{\rho_1\ldots\rho_l} = \frac{\partial x'^{\lambda_1}}{\partial x^{\mu_1}}\cdots\frac{\partial x'^{\lambda_k}}{\partial x^{\mu_k}}\frac{\partial x^{\nu_1}}{\partial x'^{\rho_1}}\cdots\frac{\partial x^{\nu_l}}{\partial x'^{\rho_l}} S^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}. \tag{13}$$

Two tensorial operations are worth mentioning: the product and the contraction. The *tensor product* of two tensors S and $\tilde{\mathsf{S}}$ of types (k,l) and (m,r), respectively, is another tensor of type $(k+m,l+r)$ denoted as $\mathsf{S}\otimes\tilde{\mathsf{S}}$, whose components are given by

$$(\mathsf{S}\otimes\tilde{\mathsf{S}})^{\mu_1\ldots\mu_k\lambda_1\ldots\lambda_m}{}_{\nu_1\ldots\nu_l\rho_1\ldots\rho_r} = S^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l}\tilde{S}^{\lambda_1\ldots\lambda_m}{}_{\rho_1\ldots\rho_r}, \tag{14}$$

which is easily shown to hold independently of the choice of coordinates. For this reason, in the abstract index notation we simply write the product of $S^{a_1\ldots a_k}{}_{b_1\ldots b_l}$ and $\tilde{S}^{c_1\ldots c_m}{}_{d_1\ldots d_r}$ as $S^{a_1\ldots a_k}{}_{b_1\ldots b_l}\tilde{S}^{c_1\ldots c_m}{}_{d_1\ldots d_r}$. Similarly, the *contraction* of two indices of a tensor $S^{a_1\ldots a_k}{}_{b_1\ldots b_l}$ of type (k,l) yields a tensor of type $(k-1,l-1)$ which we denote as $S^{a_1\ldots c\ldots a_k}{}_{b_1\ldots c\ldots b_l}$ and has components $S^{\mu_1\ldots\lambda\ldots\mu_k}{}_{\nu_1\ldots\lambda\ldots\nu_l}$ (notice here the use of Einstein's convention, and hence that the contraction essentially amounts to taking a trace). Finally, we also define the symmetrization

$$S_{(a_1\ldots a_m)} = \frac{1}{m!}\sum_\sigma S_{\sigma(a_1)\ldots\sigma(a_m)}, \tag{15}$$

and antisymmetrization

$$S_{[a_1\ldots a_m]} = \frac{1}{m!}\sum_\sigma (-1)^{\mathrm{sgn}(\sigma)} S_{\sigma(a_1)\ldots\sigma(a_m)}, \tag{16}$$

where σ runs over all the permutations of m elements, and $\mathrm{sgn}(\sigma)$ is the sign of the permutation. Eqs. (15) and (16) apply exactly to those cases where the indices that are being (anti)symmetrized are upper indices, as well as when they are only a subset of total set of indices of the tensor.

Once we have introduced all the machinery and notation surrounding manifolds and tensors (which are the language in which general relativity is formulated), we need to endow the manifold with a structure that allows us to measure (infinitesimal) distances. This is provided by a *metric*, which is a symmetric non-degenerate tensor field of type $(0, 2)$, which we denote

$$\mathsf{g} \equiv g_{ab} = g_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu. \tag{17}$$

The notation $\mathrm{d}s^2$ is often used instead of g, representing the fact that the metric is understood as encoding the product between two tangent vectors, and in particular applied on the same vector yields its "squared length". Since tangent vectors represent infinitesimal displacements, the metric provides a notion of infinitesimal lengths, as promised. The metric also provides a canonical way to define a correspondence between tangent and cotangent spaces,[2] and therefore between vector and dual vector fields, through the map

$$v^a \mapsto v_a \coloneqq g_{ab} v^b. \tag{18}$$

Because the metric is non-degenerate, it is invertible, and therefore the map given in Eq. (18) is in fact an isomorphism, and its inverse is given by

$$v_a \mapsto v^a = g^{ab} v_b, \tag{19}$$

where g^{ab} is the inverse of the metric, i.e, the tensor of type $(2, 0)$ that satisfies $g^{ab} g_{bc} = \delta^a_c$, where δ^a_c is the identity map, whose components are given by the Kronecker delta δ^μ_ν in all coordinate systems. The operations given by Eqs. (18) and (19) are called "lowering" and "rising" indices, and in general they can be applied to any upper or lower index of a tensor.

[2] Given a coordinate system, the linear extension of the map that associates ∂_μ and $\mathrm{d}x^\mu$ defines an isomorphism between vectors and dual vectors. However, this map is coordinate-dependent. It is worth remarking that, on the contrary, the correspondence we established before between vectors and double dual vectors is indeed canonical, in the sense that it can be established without reference to a particular coordinate system (cf. Eq. (8)).

Endowing a manifold with a metric also induces a preferred notion of how tensor fields *change* from one point to another. In order to understand this we first need to introduce the concept of *covariant derivative* or *connection*: a (covariant) derivative ∇ is a linear operator that maps tensor fields of type (k,l) to tensor fields of type $(k,l+1)$,

$$\nabla : S^{a_1\ldots a_k}{}_{b_1\ldots b_l} \mapsto \nabla_c S^{a_1\ldots a_k}{}_{b_1\ldots b_l}, \tag{20}$$

that 1) generalizes the notion of derivation, in the sense that it acts on smooth functions just like directional derivatives, i.e.,

$$\nabla_\mu f = \partial_\mu f \quad \rightsquigarrow \quad \mathsf{v}(f) = v^a \nabla_a f, \tag{21}$$

for all vector fields v and smooth functions f; and 2) satisfies the Leibniz's rule,

$$\nabla_e(S^{a_1\ldots a_k}{}_{b_1\ldots b_l}\tilde{S}^{c_1\ldots c_m}{}_{d_1\ldots d_r}) = \nabla_e(S^{a_1\ldots a_k}{}_{b_1\ldots b_l})\tilde{S}^{c_1\ldots c_m}{}_{d_1\ldots d_r} + S^{a_1\ldots a_k}{}_{b_1\ldots b_l}\nabla_e(\tilde{S}^{c_1\ldots c_m}{}_{d_1\ldots d_r}). \tag{22}$$

Here, we also demand that our derivative be *torsion-free*, i.e., that

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f, \quad \forall f \in \mathcal{C}^\infty(\mathcal{M}). \tag{23}$$

Given a coordinate system, one can easily check that the associated ordinary derivative ∂ is a covariant derivative. More generally, a covariant derivative can always be (perhaps, locally) related to the ordinary derivative by

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu{}_{\mu\lambda} v^\lambda, \tag{24}$$

where $\Gamma^\nu{}_{\mu\lambda}$ are the so-called *Christoffel symbols*, which, because we have demanded covariant derivatives to be torsion-free, must be symmetric in their lower indices. By Leibniz's rule, Eq. (24) generalizes to arbitrary tensors as

$$\nabla_\lambda S^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} = \partial_\lambda S^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\nu_l} + \sum_{i=1}^{k} \Gamma^{\mu_i}{}_{\lambda\rho} S^{\mu_1\ldots\rho\ldots\mu_k}{}_{\nu_1\ldots\nu_l} - \sum_{j=1}^{l} \Gamma^{\rho}{}_{\lambda\nu_j} S^{\mu_1\ldots\mu_k}{}_{\nu_1\ldots\rho\ldots\nu_l}. \tag{25}$$

Given a covariant derivative, we can define the notion of *parallel transport* along a curve, as the operation by which a vector can be "transported along a curve without changing". Concretely, given a curve $C : I \subset \mathbb{R} \to \mathcal{M}$ parametrized by t, we define its tangent vector as the derivation c that satisfies

$$c(f) = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ C) = c^\mu \partial_\mu f = \frac{\mathrm{d}f}{\mathrm{d}t}, \quad \text{where} \quad c^\mu = \frac{\mathrm{d}}{\mathrm{d}t}(\psi^\mu \circ C) \equiv \frac{\mathrm{d}x^\mu}{\mathrm{d}t}. \tag{26}$$

Because the covariant derivative is understood as a generalization of derivations, we say that a vector v^a is parallel-transported along C if

$$c^b \nabla_b v^a = 0 \ \Leftrightarrow \ c^\nu \partial_\nu v^\mu + \Gamma^\mu_{\ \nu\lambda} c^\nu v^\lambda = \frac{\mathrm{d}v^\mu}{\mathrm{d}t} + \Gamma^\mu_{\ \nu\lambda} c^\nu v^\lambda = 0, \ \forall \mu. \tag{27}$$

Notice that parallel transport allows us to establish a (curve-dependent) correspondence between different tangent spaces. This is the subtle reason why covariant derivatives are also called connections.

In this context, the metric allows us to choose an otherwise non-unique covariant derivative by demanding that parallel-transported vectors preserve their product, which requires that $\nabla_c g_{ab} = 0$. This is the condition that characterizes the *Levi-Civita connection*, for which the Christoffel symbols are given by

$$\Gamma^\mu_{\ \nu\lambda} = \frac{g^{\mu\rho}}{2}\Big(\partial_\nu g_{\lambda\rho} + \partial_\lambda g_{\nu\rho} - \partial_\rho g_{\nu\lambda}\Big). \tag{28}$$

It is worth noting that Eq. (28) includes derivatives of the metric, and therefore depends on how the metric *changes* across the manifold. For an Euclidean space (i.e., $\mathbb{R}^n$ with the standard inner product), these derivatives vanish, and therefore the Levi-Civita covariant derivative coincides with the ordinary derivative.

The Levi-Civita connection allows us to define *geodesics*, which are a generalization of the Euclidean notion of "straight line". By analogy with a straight line, a curve C is said to be a geodesic (with a *affine parametrization*) if its tangent vector "does not change along it" or, in the language we have been introducing here, if it is parallel-transported along it. Namely, C is a geodesic if an only if it satisfies the *geodesic equation*,

$$c^a \nabla_a c^b = 0 \ \Leftrightarrow \ c^\mu \partial_\mu c^\nu + \Gamma^\nu_{\ \mu\lambda} c^\mu c^\lambda = \frac{\mathrm{d}^2 x^\nu}{\mathrm{d}t^2} + \Gamma^\nu_{\ \mu\lambda} \frac{\mathrm{d}x^\mu}{\mathrm{d}t}\frac{\mathrm{d}x^\lambda}{\mathrm{d}t} = 0, \ \forall \nu, \tag{29}$$

where c^a is the tangent vector of C.

In Euclidean spaces, because the Levi-Civita derivative coincides with the ordinary derivative, it is straightforward to see that if a vector is parallel-transported along a closed curve, the final result of the parallel transport coincides with the initial one. When this property fails to be satisfied in general, we say that the manifold $\mathcal{M}$ with the metric g is *curved*. Otherwise, we say that it is *flat*. To measure curvature in a manifold, we use the *Riemann (curvature) tensor*, which is defined as the tensor field ${R_{abc}}^d$ that satisfies

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)w_c = {R_{abc}}^d\, w_d, \tag{30}$$

for all dual vector fields w_d, which, because the connection is torsion-free, also satisfies

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)v^c = -{R_{abd}}^c\, v^d, \tag{31}$$

and more generally

$$(\nabla_c\nabla_d - \nabla_c\nabla_d)S^{a_1\ldots a_k}{}_{b_1\ldots b_l} = -\sum_{i=1}^{k} {R_{cde}}^{a_i}\, S^{a_1\ldots e\ldots a_k}{}_{b_1\ldots b_l} + \sum_{j=1}^{l} {R_{cdb_j}}^{e}\, S^{a_1\ldots a_k}{}_{b_1\ldots e\ldots b_l}. \tag{32}$$

From Eq. (30), one can see that the properties of the covariant derivative impose certain symmetries on the Riemann tensor, reducing its number of independent components. Namely, the tensor that results from lowering its only upper index, $R_{abcd} = g_{de}{R_{abc}}^e$, is antisymmetric in the first two and last two indices, i.e.,

$$R_{abcd} = -R_{bacd} = -R_{abdc} = R_{badc}, \tag{33}$$

it is symmetric under swap of the first two and last two indices, i.e.,

$$R_{abcd} = R_{cdab}, \tag{34}$$

and satisfies the *Bianchi identity*

$$\nabla_{[a}R_{bc]de} = 0. \tag{35}$$

There a few relevant quantities that can be defined from the Riemann tensor: contracting its second and fourth indices, we obtain the *Ricci tensor*,

$$R_{ac} \coloneqq {R_{abc}}^b = g^{bd}R_{abcd}, \tag{36}$$

which, by Eq. (34), is a symmetric tensor, $R_{ac} = R_{ca}$. The trace of the Ricci tensor is called the *Ricci scalar* or *curvature scalar*,

$$R \coloneqq R_a{}^a = g^{ac} R_{ac}. \tag{37}$$

Finally, the *Kretschmann scalar* or *Kretschmann invariant*, which is often useful to detect the existence of (curvature) singularities, is given by

$$K \coloneqq R_{abcd} R^{abcd}. \tag{38}$$

B. General relativity

In general relativity, space and time are described jointly in a *spacetime manifold* $\mathcal{M}$, which is endowed with a metric g. The metric describing a spacetime manifold is always *Lorentzian*, which means that at each point $p \in \mathcal{M}$ it is always possible to find a basis $\{e_0, e_1 \ldots, e_n\}$ such that $\mathsf{g}_p(e_\mu, e_\nu) = 0$ when $\mu \neq \nu$, $\mathsf{g}_p(e_0, e_0) = -1$, and $\mathsf{g}_p(e_j, e_j) = 1$ for all $j \in \{1, \ldots, n\}$. We say that the pair formed by the spacetime manifold and the metric, $(\mathcal{M}, \mathsf{g})$, define an $(1+n)$*-dimensional Lorentzian manifold*. Here, we will always consider $n = 3$.

The Lorentzian character of the spacetime manifold implies that we can always classify a tangent vector v into three categories, depending on the value of its "squared length" $\mathsf{g}(\mathsf{v}, \mathsf{v}) = g_{\mu\nu} v^\mu v^\nu = v_\mu v^\mu$: (i) *timelike*, if $v_\mu v^\mu < 0$, (ii) *spacelike*, if $v_\mu v^\mu > 0$, and (iii) *null*, if $v_\mu v^\mu = 0$. By analogy, we say that a curve is timelike, spacelike, or null, if its tangent vector is so at all points. Similarly, we say that a hypersurface (i.e., a submanifold with codimension 1) is timelike, spacelike, or null, if its normal vector is so at all points. Notice that while a vector can always be classified as timelike, spacelike, or null, a given curve or hypersurface need not fall in any of these categories.

Given a spacelike curve C parametrized by s with tangent vector c^a, we can compute its *proper length* between $s = s_1$ and $s = s_2$ as

$$\ell = \int_{s_1}^{s_2} \sqrt{g_{ab} c^a c^b}\, \mathrm{d}s = \int_{s_1}^{s_2} \sqrt{g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}s} \frac{\mathrm{d}x^\nu}{\mathrm{d}s}}\, \mathrm{d}s, \tag{39}$$

where, as in Sec. II A, in the second equality we have particularized the expression for a given coordinate system, with respect to which $\mathrm{d}x^\mu/\mathrm{d}s$ is the tangent vector of the curve. Analogously, given a timelike curve γ parametrized by t with tangent vector v^a, we can compute its *proper time* between $t = t_1$ and $t = t_2$ as

$$\tau = \int_{t_1}^{t_2} \sqrt{-g_{ab} v^a v^b}\, \mathrm{d}t = \int_{t_1}^{t_2} \sqrt{-g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}t} \frac{\mathrm{d}x^\nu}{\mathrm{d}t}}\, \mathrm{d}t, \tag{40}$$

where the minus sign inside the square root is due to the timelike character of v^a. The proper time between two points of a timelike curve is understood as the time interval that an observer following the curve would measure, i.e., the time interval measured in the "proper reference frame" of the trajectory. One can always reparametrize a timelike curve in terms of τ, and the tangent vector in that case is called the *4-velocity* of the curve, u^a, which from Eq. (40) must satisfy $u_a u^a = -1$. Now, (massive) particles always follow timelike curves. Specifically, if a particle is free, i.e., it is not subjected to any forces (besides gravity), then its 4-velocity "does not change", that is, it is parallel-transported along the trajectory, therefore satisfying

$$u^a \nabla_a u^b = 0. \tag{41}$$

Hence, freely falling particles follow the geodesics of the spacetime metric g, with their equations of motion given by the geodesic equation, Eq. (41), along with the constraint $u_a u^a = -1$. It is worth remarking that the latter constraint only needs to be imposed as an initial condition, since the quantity $u_a u^a$ is preserved along a geodesic, which is why we can talk about timelike, spacelike, and null geodesics.

Now, in general relativity, the spacetime metric is not a given fixed background. On the contrary, it depends on the matter content of the universe, which is described in general through the *stress-energy-momentum tensor* (SET), T_{ab}, defined as [13]

$$T^{ab} = \frac{2}{\sqrt{-g}} \frac{\delta S_{\text{matter}}}{\delta g_{ab}}, \tag{42}$$

where g is the determinant of the metric g, and

$$S_{\text{matter}} = \int_{\mathcal{M}} \mathcal{L}_{\text{matter}} \sqrt{-g}\, \mathrm{d}\mathsf{x} \tag{43}$$

is the *action* of the matter fields, computed from its Lagrangian function $\mathcal{L}_{\text{matter}}$. Given an observer with 4-velocity v^a and a spacelike vector field along its trajectory s^b,

- $T_{ab}v^a v^b$ is interpreted as the *energy density* of the matter field as measured by the observer, and

- $T_{ab}v^a s^b$ is interpreted as the *momentum density* of the matter field in the direction of s^b, as measured by the observer

If additionally, we consider a second spacelike vector field r^b, then $T_{ab}s^a r^b$ is interpreted as the *stress* of the matter field along the hyperplane of normal s^a in the direction of r^b. It is for this reason that we refer to T_{ab} as the "stress-energy-momentum" tensor. Of particular interest to us will be the case of *perfect fluids*, which are continuously distributed matter fields that in their proper reference frames are characterized solely by their energy density ρ, and their (isotropic) pressure p as the only internal stress. Hence, given the fluid's 4-velocity field u^a, its SET is given by

$$T_{ab} = \rho u_a u_b + p(g_{ab} + u_a u_b). \tag{44}$$

Once the stress-energy-momentum tensor has been introduced, we are in conditions to formulate the central pillar of general relativity, *Einstein's equation*, which relates the curvature of spacetime and the matter's SET through

$$R_{ab} - \frac{1}{2}Rg_{ab} = 8\pi T_{ab}, \tag{45}$$

where recall that R_{ab} and R are the Ricci tensor and the curvature scalar, respectively. The LHS of Einstein's equation, which only depends on the geometry, is usually denoted G_{ab} and we call it the *Einstein tensor*. From the contraction of the Bianchi identity (Eq. (35)) we find that

$$\nabla^a G_{ab} = g^{ac}\nabla_c G_{ab} = g^{ac}\nabla_c\Big(R_{ab} + \frac{1}{2}Rg_{ab}\Big) = 0, \tag{46}$$

which by Einstein's equation implies that

$$\nabla^a T_{ab} = 0. \tag{47}$$

In flat spacetimes, Eq. (47) implies the conservation of energy. In general curved spacetimes, however, this is no longer the case, although the equation can still be interpreted as a "local conservation of energy", in the sense that energy is indeed approximately conserved in sufficiently small regions.

III. STATIC SPHERICALLY SYMMETRIC SPACETIMES

In this review we are concerned with static spherically symmetric objects, and therefore we will always be working with *static spherically symmetric spacetimes*. These are spacetimes for which there exists a set of coordinates $\{t, r, \theta, \varphi\}$ such that the metric can be written as[3]

$$\mathsf{g} = -f(r)\mathrm{d}t^2 + h(r)\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2, \quad \text{with} \quad \mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2, \tag{48}$$

where r is the areal radius, defined by the property that the area A of each 2-sphere of constant r satisfies $A = 4\pi r^2$, and f and h are $\mathcal{C}^2$ functions of r. From the metric given in Eq. (48), we find that the Einstein tensor reads

$$G_t{}^t = -\frac{h'}{rh^2} - \frac{1}{r^2}\left(1 - \frac{1}{h}\right), \tag{49}$$

$$G_r{}^r = \frac{f'}{rfh} - \frac{1}{r^2}\left(1 - \frac{1}{h}\right), \tag{50}$$

$$G_\theta{}^\theta = G_\varphi{}^\varphi = \frac{1}{2\sqrt{fh}}\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{f'}{\sqrt{fh}}\right) + \frac{f'}{2rfh} - \frac{h'}{2rh^2}, \tag{51}$$

where the primes $'$ denote differentiation with respect to r.

A. Vacuum solutions: the Schwarzschild metric

The first thing we can use Eqs. (49)–(51) for is to analyze the static spherically symmetric *vacuum solutions* of Einstein's equation, i.e., the ones that result from imposing $T_a{}^b = 0$. In that case, we have that $G_t{}^t = G_r{}^r = G_\theta{}^\theta = G_\varphi{}^\varphi = 0$. Substracting

[3] See [12] for a first-principles definition of staticity and spherical symmetry, as well as for the full derivation of how these properties allow one to write the metric as in Eq. (48).

Eqs. (50) and (49), multiplying them by rh, and equating the result to zero we find that

$$\frac{h'}{h}+\frac{f'}{f}=0 \;\Leftrightarrow\; \frac{\mathrm{d}}{\mathrm{d}r}\big(\log h+\log f\big)=\frac{\mathrm{d}}{\mathrm{d}r}\log(hf)=0 \;\Leftrightarrow\; fh=\tilde{C}, \tag{52}$$

for some fixed constant $\tilde{C}\in\mathbb{R}$. Hence, $h=\tilde{C}/f$, and since this constant can be reabsorbed into the definition of t (by taking $\sqrt{\tilde{C}}t$ as the new time coordinate), we can simply take $h=1/f$. Inserting this in Eq. (49), multiplying it by r^2 and equating it to zero we find

$$rf'+f-1=\frac{\mathrm{d}}{\mathrm{d}r}(rf)-1=0 \;\Rightarrow\; rf=r+C, \tag{53}$$

for some integration constants $C\in\mathbb{R}$. This implies that

$$f(r)=1+\frac{C}{r} \quad\rightsquigarrow\; h=\left(1+\frac{C}{r}\right)^{-1}. \tag{54}$$

Notice that equating Eqs. (50) and (51) to zero does not add any additional constraints, since after imposing $h=1/f$, Eq. (50) becomes Eq. (49), and equating Eq. (51) to zero yields

$$\frac{f''}{2}+\frac{f'}{r}=0, \tag{55}$$

which is a consequence of Eq. (53). Hence, the most general metric that describes a static spherically symmetric spacetime without matter content is given by the so-called *Schwarzschild metric*,

$$\mathsf{g}=-\left(1+\frac{C}{r}\right)\mathrm{d}t^2+\left(1+\frac{C}{r}\right)^{-1}\mathrm{d}r^2+r^2(\mathrm{d}\theta^2+\sin^2\mathrm{d}\varphi^2), \tag{56}$$

for some $C\in\mathbb{R}$, and some appropriately chosen coordinates $\{t,r,\theta,\varphi\}$. Notice that for $C=0$ we recover the metric of a flat spacetime, i.e., the *Minkowski metric*

$$\eta=-\mathrm{d}t^2+\mathrm{d}r^2+r^2(\mathrm{d}\theta^2+\sin^2\mathrm{d}\varphi^2)=\mathrm{d}t^2+\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2. \tag{57}$$

In order to conclude that the metric has the form of Eq. (56) we only assumed that the metric could be written as in Eq. (48), and that Eqs. (49)–(51) were equal to zero. In particular, we did not use any properties that involved the whole manifold. Therefore, we just proved the following theorem:

Theorem (Schwarzschild)**.** *Let* $(\mathcal{M}, \mathsf{g})$ *be a* $(1+3)$*-dimensional Lorentzian manifold, and let* $\mathcal{V} \subset \mathcal{M}$ *be a subset where there exist coordinates* $\{t, r, \theta, \varphi\}$ *such that, restricted to* $\mathcal{V}$*, the metric can be written as*

$$\mathsf{g} = -f(r)\,\mathrm{d}t^2 + h(r)\,\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2, \tag{58}$$

for some $f, h \in \mathcal{C}^2$*. Then, if* g *satisfies Einstein's equation in the vacuum, there exists* $C \in \mathbb{R}$ *such that*

$$f(r) = 1 + \frac{C}{r} \quad \textit{and} \quad h = \left(1 + \frac{C}{r}\right)^{-1}. \tag{59}$$

It is worth remarking that here we are assuming that the spacetime is static, but this assumption can be dropped: indeed, Birkhoff's theorem [13] shows that any spherically symmetric spacetime must be, in fact, static, and therefore Schwarzschild's theorem holds for general spherically symmetric spacetimes.

Now, in the previous derivation we were not particularly careful regarding the extent to which the coordinates $\{t, r, \theta, \varphi\}$ cover the totality of the spacetime manifold, or whether the functions h and f vanish at some point, which would make the metric degenerate at that point. While Schwarzschild's theorem still stands as long as the ability to write the metric in the form of Eq. (48) is part of the assumptions, a few remarks are in order.

First, if the $\{t, r, \theta, \varphi\}$ coordinates are constructed placing the center of symmetry of the spacetime at $r = 0$, then obviously $r = 0$ is not covered by the coordinates. However, if the coordinates cover a neighbourhood of this point, and in this neighbourhood the stress-energy-momentum tensor vanishes, then the theorem still applies, and all the relevant geometric quantities at $r = 0$ can be computed by taking the limit $r \to 0^+$. The Kretschmann scalar can then be computed to be (see, e.g., [14])

$$K = \frac{12C^2}{r^6}, \tag{60}$$

which diverges as $r \to 0^+$ as long as $C \neq 0$. Hence, either $C = 0$ (and the spacetime is flat around the center of symmetry) or there is a (curvature) singularity at $r = 0$. This singularity, if it exists, is a physical one, and cannot be tamed by a change of coordinates—since this would not change the fact that K diverges as we approach the center of symmetry.

Second, if $C < 0$, then it is possible that the $\{t, r, \theta, \varphi\}$ cover the region where $r = -C$. At this 2-sphere, f would vanish and h would diverge, signalling that the coordinate system breaks. Looking at Eq. (60), however, we see that the Kretschmann invariant does not diverge, which can be a sign that the lack of regularity of the metric at this surface is not due to a true physical singularity, but to the choice of coordinates. Indeed, what happens here is that the vector fields $(\partial_t)^a$ and $\nabla^a r$ become collinear at this surface, and the construction of the spherical coordinates breaks. The metric, however, can be made regular by changing the coordinate system (typically to Eddington-Finkelstein or Kruskal-Szekeres coordinates [12]). Notice that for the region $r < -C$ the metric can still be described by the coordinates used for the $r > -C$ region, but there the t coordinate becomes spacelike and the r coordinate becomes timelike. At the boundary, $r = -C$, something interesting happens: all null geodesics point inwards. This means, in particular, that light rays cannot escape the $r < -C$ region, in which case we say that it is a *black hole region*, and $r = -C$ is what we call a *marginally trapped surface*.[4] On the other hand, the region $r > -C$ is asymptotically flat (i.e., $g_{ab} \to \eta_{ab}$ as $r \to \infty$), and the equation of motion of a free test particle in that region is given by (see, e.g., [12])

$$\frac{\dot{r}^2}{2} + \frac{C}{2r}\left(\frac{L}{r^2} + 1\right) = \frac{E^2 - 1}{2}, \tag{61}$$

where the dot $\dot{}$ denotes differentiation with respect to the proper time, and $E = (1 + C/r)\dot{t}$ and $L = r^2\dot{\varphi}$ are conserved quantities of the motion. For $r \to \infty$, Eq. (61) can be approximated as

$$\frac{\dot{r}^2}{2} + \frac{C}{2r} \approx \text{constant}, \tag{62}$$

which, by comparison with the equation for conservation of energy in Newtonian gravity,

$$\frac{\dot{r}^2}{2} - \frac{M}{r} = \text{constant}, \tag{63}$$

[4] Trapped surfaces and physical singularities are intimately connected, inasmuch as the existence of the former is a fundamental ingredient for the formation of the latter [15].

motivates the identification $C \equiv -2M$, where M is the effective mass of the inner regions as perceived by an asymptotic observer. With this identification, Eq. (56) can be rewritten in the familiar form

$$\mathsf{g} = -\left(1 - \frac{2M}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2M}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2. \tag{64}$$

B. Solutions in the presence of matter

We can now study what happens in the presence of a static fluid. In this case, $T_a{}^b$ does not necessarily vanish, but the restrictions that staticity and spherical symmetry impose on the Einstein tensor must also be satisfied by the SET: namely, the only possibly non-zero components are $T_t{}^t$, $T_r{}^r$, and $T_\theta{}^\theta = T_\varphi{}^\varphi$, and since the metric (and therefore the Einstein tensor) only depends on r, the corresponding components of the SET may only depend on r, as well. We denote

$$T_t{}^t \equiv -\rho(r), \quad T_r{}^r \equiv p(r), \quad \text{and} \quad T_\theta{}^\theta = T_\varphi{}^\varphi \equiv q(r), \tag{65}$$

which we identify with the energy density, the radial pressure, and the tangential pressure of the fluid, respectively. With this notation, and from Eqs. (49)–(51), Einstein's equations read

$$8\pi\rho = \frac{h'}{rh^2} + \frac{1}{r^2}\left(1 - \frac{1}{h}\right), \tag{66}$$

$$8\pi p = \frac{f'}{rfh} - \frac{1}{r^2}\left(1 - \frac{1}{h}\right), \tag{67}$$

$$8\pi q = \frac{1}{2\sqrt{fh}}\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{f'}{\sqrt{fh}}\right) + \frac{f'}{2rfh} - \frac{h'}{2rh^2}. \tag{68}$$

Now, if we multiply Eq. (66) by r^2 we get

$$8\pi r^2\rho = r\frac{h'}{h^2} + \left(1 - \frac{1}{h}\right) = \frac{\mathrm{d}}{\mathrm{d}r}\left[r\left(1 - \frac{1}{h}\right)\right], \tag{69}$$

which can be integrated right away yielding

$$h(r) = \left(1 - \frac{2m(r)}{r}\right)^{-1}, \quad \text{with} \quad m(r) = m_0 + 4\pi\int_0^r \rho(s)s^2\,\mathrm{d}s, \tag{70}$$

where $m(r)$ is the *Misner-Sharp mass function* [16, 17]. Even though in the analysis of the vacuum solutions we allowed the presence of singularities at $r = 0$, here we assume that the fluid is regular at the center of symmetry, and therefore so should be the metric. This means that the Einstein tensor should not diverge as $r \to 0^+$, which from Eq. (49) implies

$$1 - \frac{1}{h} \sim \mathcal{O}(r^2) \;\Rightarrow\; m(r) \sim \mathcal{O}(r^3) \quad \text{as} \quad r \to 0^+, \tag{71}$$

enforcing $m_0 = 0$. This, however, does not mean that we are assuming $2m(r)/r < 1$ everywhere. Rather, we are assuming that the center of symmetry is regular, and that the coordinates $\{t, r, \theta, \varphi\}$ describe the spacetime metric at least up to the smallest r such that $2m(r)/r = 1$, i.e., at least up to the first marginally trapped surface, if it exists. If such trapped surface does not exist, then the coordinates will cover the whole spacetime manifold. Similarly, as long as $2m(r)/r < 1$ and the radial pressure of the fluid p is well defined, by Eq. (67) it must be that $f > 0$. Therefore, in this domain of r we can write

$$f(r) = e^{2\Phi(r)}, \tag{72}$$

where $\Phi(r)$ is the *redshift function*. Introducing this and the expression for h in terms of m into Eq. (67) yields

$$\Phi' = \frac{m + 4\pi r^3 p}{r(r-2m)}. \tag{73}$$

To obtain the last equation of the system, which specifies p', we could use Eq. (68), but it is much more efficient to use the local conservation of the SET, $\nabla_\mu T_\nu{}^\mu = 0$, which, for the particular case $\nu = r$ yields

$$p' + \left(\frac{f'}{2f} + \frac{h'}{2h} + \frac{2}{r}\right)p + \frac{f'}{2f}\rho - \frac{h'}{2h}p - \frac{2}{r}q = p' + \frac{f'}{2f}(\rho + p) + \frac{2}{r}(p - q) = 0. \tag{74}$$

Using that $f'/2f = \Phi'$, and by Eq. (73), we obtain

$$p' = -(\rho + p)\frac{m + 4\pi r^3 p}{r(r-2m)} - \frac{2}{r}(p - q). \tag{75}$$

This is the equation of *hydrostatic equilibrium*, which can be interpreted as the equation that needs to be satisfied for the fluid to "sustain itself". Notice that from this equation we observe that for the metric to be regular at the center of symmetry, the radial and tangential pressures must coincide in the origin $r = 0$. For the particular case in which the fluid is perfect, and therefore isotropic (cf. Eq. (44)), $p = q$, and therefore Eq. (75) reduces to

$$p' = -(\rho + p)\frac{m + 4\pi r^3 p}{r(r - 2m)}, \tag{76}$$

which is known as the *Tolman-Oppenheimer-Volkoff equation* (TOV). All in all, we have shown that the most general metric that describes a static spherically symmetric spacetime with a fluid can always be written as

$$\mathsf{g} = -e^{2\Phi(r)}\mathrm{d}t^2 + \left(1 - \frac{2m(r)}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2, \tag{77}$$

with m given by Eq. (71) (with $m_0 = 0$), Φ obtained from Eq. (73), and p and q so that Eq. (75) is satisfied.

Of particular interest to us will be the case in which the fluid is contained on a region $r \leq R$ for which the metric is regular, and for $r > R$ the vacuum Einstein's equations are satisfied. In that case, by Schwarzschild's theorem, the metric for $r > R$ must be given by Eq. (56), for some appropriately chosen C, while for $r < R$ it will be given by Eq. (77). Now, it is apparent from examining Einstein's equations, Eqs. (66)–(68), that for them to be well-defined, f must be at least piecewise $\mathcal{C}^2$, and h must be at least piecewise $\mathcal{C}^1$. As a consequence, f' and h are at least continuous at the interface $r = R$. By Eq. (67), this implies that p is continuous at $r = R$; however, while m must also be continuous, ρ need not be, and neither does q, as can be seen from Eq. (68). Hence, the "radius of the fluid", R, is fixed by the condition $p(R) = 0$, while ρ and q do not necessarily vanish as we take their limits $r \to R^-$. Moreover, because h is continuous at $r = R$, this also fixes the value of C for the Schwarzschild metric, which must be $C = -2m(R)$, or, taking the more standard form Eq. (64) as a reference, $M = m(R)$. Hence, under these assumptions, the metric takes the following piecewise form:

$$\mathsf{g} = \begin{cases} -e^{2\Phi(r)}\mathrm{d}t^2 + \left(1 - \frac{2m(r)}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2, & r < R \\ -\left(1 - \frac{2m(R)}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2m(R)}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2, & r > R \end{cases} \tag{78}$$

where one should also guarantee that, at the interface, Israel's junction conditions are satisfied [18]. Then, we define the *compactness* of the fluid as

$$\alpha \coloneqq \frac{m(R)}{R} \equiv \frac{M}{R}, \tag{79}$$

which here we implicitly assumed to satisfy $\alpha < 1/2$.

IV. BUCHDAHL'S LIMIT

Despite the generality of the metric given by Eq. (78), it turns out that for a perfect fluid, and under fairly reasonable assumptions, the compactness α is upper-bounded by a number strictly below $1/2$. This tells us that in general relativity there is a maximum mass that a perfect fluid star of a given radius R can sustain without collapsing, and that maximum mass actually lies strictly below its Schwarzschild mass, i.e., the mass required to form a black hole region in $r \leq R$.

Theorem (Buchdahl [19])**.** *Under the conditions with which we constructed the metric given in Eq. (78), a perfect fluid whose energy density ρ is non-negative and a monotonically decreasing function of the areal radius r satisfies*

$$\alpha < \frac{4}{9}. \tag{80}$$

Proof. Because we are considering a perfect fluid, $p = q$, and therefore we can equate the RHS of Eqs. (67) and (68), which after introducing the expression for h in terms of m yields

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{1}{r\sqrt{h}}\frac{\mathrm{d}\sqrt{f}}{\mathrm{d}r}\right) = \sqrt{fh}\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{m}{r^3}\right). \tag{81}$$

Now, because ρ is monotonically decreasing, from the definition of m (Eq. (71)) we get that

$$m(r) \geq 4\pi\rho(r)\int_0^r s^2\,\mathrm{d}s = \frac{4\pi r^3}{3}\rho(r) = \frac{r}{3}m'(r), \tag{82}$$

which can be rewritten as

$$r^3\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{m}{r^3}\right) \leq 0 \;\Rightarrow\; \frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{m}{r^3}\right) \leq 0. \tag{83}$$

Hence, from Eq. (81),

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{1}{r\sqrt{h}}\frac{\mathrm{d}\sqrt{f}}{\mathrm{d}r}\right) \leq 0. \tag{84}$$

Integrating this inequality between an arbitrary $r < R$ and the surface of the fluid at $r = R$, and using the continuity of h and f' at the interface,

$$\frac{1}{r\sqrt{h}}\frac{\mathrm{d}\sqrt{f}}{\mathrm{d}r} \geq \frac{1}{R}\sqrt{1-\frac{2M}{R}}\frac{\mathrm{d}}{\mathrm{d}r}\sqrt{1-\frac{2M}{r}}\bigg|_{r=R} = \frac{M}{R^3} \;\Rightarrow\; \frac{\mathrm{d}\sqrt{f}}{\mathrm{d}r} \geq \frac{M}{R^3}\frac{r}{\sqrt{1-\frac{2m(r)}{r}}}, \tag{85}$$

where, as at the end of Sec. III B, we have denoted $M \equiv m(R)$. Finally, we integrate this inequality from $r = 0$ to $r = R$, finding

$$\sqrt{f(0)} \leq \sqrt{1-\frac{2M}{R}} - \frac{M}{R^3}\int_0^R \frac{r}{\sqrt{1-\frac{2m(r)}{r}}}\,\mathrm{d}r. \tag{86}$$

Finally, we note that the assumptions for the construction of the metric in Eq. (78) include that the fluid is regular at the center of symmetry, and therefore $f(0) > 0$. Moreover, Eq. (83) can also be integrated from r to the surface of the fluid, yielding

$$\frac{M}{R^3} - \frac{m}{r^3} \leq 0 \;\Rightarrow\; m(r) \geq \frac{Mr^3}{R^3}. \tag{87}$$

Hence, Eq. (86) implies that

$$0 < \sqrt{1-\frac{2M}{R}} - \frac{M}{R^3}\int_0^R \frac{r}{\sqrt{1-\frac{2Mr^2}{R^3}}}\,\mathrm{d}r = \frac{3}{2}\sqrt{1-\frac{2M}{R}} - \frac{1}{2} \;\Rightarrow\; \alpha = \frac{M}{R} < \frac{4}{9}, \tag{88}$$

which is precisely the aforementioned *Buchdahl's limit*. □

After completing the proof of the theorem, some remarks are in order:

1. As a byproduct of the proof, we have obtained that the density profile $\rho(r)$ that optimizes the compactness is the constant one, where the energy density $\rho(r) = \rho_c$ is equal to the central density at all radii $r < R$. In this very simple case, the TOV equation, Eq. (76), can be integrated right away to obtain the pressure profile,

$$p(r) = \rho_c\left(\frac{\sqrt{1-2Mr^2/R^3} - \sqrt{1-2M/R}}{3\sqrt{1-2M/R} - \sqrt{1-2Mr^2/R^3}}\right), \tag{89}$$

from which the central pressure (i.e., the pressure at $r = 0$) is given by

$$p_c = \rho_c \left(\frac{1 - \sqrt{1 - 2M/R}}{3\sqrt{1 - 2M/R} - 1} \right). \tag{90}$$

We readily see that for $\alpha = 4/9$, p_c diverges. The upper-bound on the compactness appears because to sustain a higher mass in the same radius would require an infinite central pressure.

2. The bound on the compactness given by Buchdahl's theorem is *sharp* under the conditions of its statement (i.e., $\rho \geq 0$ and monotonically decreasing), meaning that for any $\epsilon > 0$ there exists a (constant) density profile whose compactness satisfies $\frac{4}{9} - \epsilon < \alpha < \frac{4}{9}$, and therefore this is the best upper-bound one can possibly give under these assumptions.

3. Remarkably, Buchdahl's limit can be obtained without making any assumptions about the pressure. However, it is worth mentioning that, for barotropic fluids, i.e., those for which the pressure is a function of ρ, the condition that ρ is monotonically decreasing is equivalent to demanding $p \geq 0$ and that p be a monotonically increasing function of ρ. Indeed, these two conditions, together with $\rho \geq 0$, imply that p be a monotonically decreasing function of r, by the TOV equation. And if p increases with ρ and decreases with r, it must be the case that ρ decreases with r. The significance of this equivalence lies in the fact that the conditions $p \geq 0$ and $\mathrm{d}p/\mathrm{d}\rho \geq 0$ can be argued to be necessary for the stability of the fluid [20, 21].

V. OTHER COMPACTNESS BOUNDS

While Buchdahl's result is powerful, its assumptions can be considered too restrictive. Isotropy, for instance, is hardly justified, as there are numerous astrophysical models of stars that include anisotropies (see, e.g., [22–25]). Furthermore, as pointed out by[26], once isotropy is lifted, there is no compelling reason to demand $\mathrm{d}\rho/\mathrm{d}r \leq 0$, since now the tangential pressure may be able to sustain an outward increase in the density without rendering the system unstable. The simplest example of this is a soap bubble. More complex but still reasonable models, such as the Einstein–Vlasov system (which describes a self-gravitating collisionless gas), also fall outside the assumptions of Buchdahl's theorem, and yet they do obey Buchdahl's limit [27].

Even within the realm of perfect fluids, however, one may challenge Buchdahl's constraints by comparing them with the standard *energy conditions* (see, e.g., [11]), which, for perfect fluids, can be written in terms of the density and the pressure as follows:

- *Null energy condition* (NEC): $\rho + p \geq 0$.
- *Weak energy condition* (WEC): $\rho \geq 0$ and $\rho + p \geq 0$.
- *Dominant energy condition* (DEC): $\rho \geq |p|$.
- *Strong energy condition* (SEC): $\rho + p \geq 0$ and $\rho + 3p \geq 0$.

While the weak energy condition guarantees $\rho \geq 0$, the monotonicity condition needs to be justified from stability arguments, as it is not implied by any energy condition. As a matter of fact, the constant density solutions that saturate Buchdahl's limit violate the dominant energy condition, as remarked in [28] and [29], among several others, which may indicate that for perfect fluids Buchdahl's assumptions are actually too lax.

For all these reasons, in the last decades there has been an effort to study upper bounds on the compactness under different sets of assumptions, both generalizing and constraining Buchdahl's result. A wide variety of methods have been employed by different authors, and it is beyond the scope of this review to revisit them. However, it is still useful to list here all the generalizations that (to the best of our knowledge) have been proved since Buchdahl's pioneering paper [19]. In what follows, as for Buchdahl's theorem, we work under the conditions with which we constructed the metric given in Eq. (78), namely, that f and h in Eq. (48) are at least piecewise $\mathcal{C}^2$ and $\mathcal{C}^1$, respectively, and that $f > 0$ and $2m(r)/r < 1$. Moreover, ρ_c and p_c denote the energy density and pressure at the center $r = 0$, as before, and $\bar{\rho} = 3m(r)/4\pi r^3$ denotes the average density.

Theorem (Guven and Ó Murchadha [26])**.** *Let* $\rho \geq 0$. *Then,*

1. *If* $\rho' \leq 0$ *and* $p \geq q$, *then*
$$\alpha \leq \frac{4}{9}, \tag{91}$$
that is, Buchdahl's limit still holds.

2. *If* $p \geq 0$, *and*
$$(\rho - \bar{\rho}) + 2(q - p) \leq \beta(p + \bar{\rho}/3), \tag{92}$$
for some $\beta \geq 0$. *Then,*
$$\alpha \leq \frac{(\beta+2)^2}{4}\left[\sqrt{1 + \frac{4}{(\beta+2)^2}} - 1\right]. \tag{93}$$

Theorem (Ivanov [28])**.** *Let* $\rho > 0$, $\rho' \leq 0$, $q \geq p \geq 0$. *If* $2(q-p) > \bar{\rho} - \rho$ *and* $q \leq \epsilon\rho$ *for some* $\epsilon \geq 0$, *then*
$$4\sqrt{1-2\alpha}(2\alpha)^{3\epsilon} \geq \int_0^{2\alpha} \frac{x^{3\epsilon}}{\sqrt{1-x}}\,\mathrm{d}x. \tag{94}$$

Theorem (Barraco and Hamity [30])**.** *Let* $\rho \geq 0$ *be a monotonically decreasing function of* r, *and let* $q = p \leq \beta\rho/3$, *for some* $\beta \geq 0$. *Then,*
$$\alpha \leq \frac{1}{2}\left[1 - \left(\frac{1+\xi}{1+3\xi}\right)^2\right], \tag{95}$$
where $\xi = \beta\rho_c/3\bar{\rho}$. *In particular, for a constant density profile with* $\beta = 3$ (DEC), $\alpha \leq 3/8$. *For* $\beta \to \infty$, *we recover Buchdahl's limit.*

The last two theorems we mention in this section have the remarkable property that the bounds they establish are sharp (i.e., optimal), a fact that they both prove constructively by providing sequences of solutions that asymptotically approach the bound.

Theorem (Andréasson [31])**.** *Let* $\rho \geq 0$ *and* $p \geq 0$. *If* $p + q \leq \Omega\rho$ *for some* $\Omega > 0$, *then*
$$\sup_{r>0} \frac{2m(r)}{r} \leq \frac{(1+2\Omega)^2 - 1}{(1+2\Omega)^2}. \tag{96}$$
In particular,
$$\alpha \leq \frac{(1+2\Omega)^2 - 1}{2(1+2\Omega)^2}. \tag{97}$$

Theorem (Bondi, Karageorgis, and Stalker [32, 33]). *Let $\rho, p, q \geq 0$ and such that the mass function m is everywhere finite.*

(i) If $p + 2q \leq \rho$ (Vlasov-Einstein)*, then $\alpha \leq 4/9$.*

(ii) If $p = q$ (perfect fluid)*, then $\alpha \leq 6\sqrt{2} - 8$.*

(iii) If $q = p \leq \rho$ (perfect fluid with DEC)*, then $\alpha \lesssim 0.433$.*

(iv) If $q \leq \rho$ (DEC in tangential direction)*, then $\alpha \leq \frac{1+\sqrt{2}}{5}$.*

(v) If $q, p \leq \rho$ (DEC)*, then $\alpha \lesssim 0.482$.*

The numerical values provided in (iii) and (v) can be computed numerically up to arbitrary precision.

Finally, it is worth noting that configurations allowing for negative energy densities or unrestricted anisotropies can attain compactness arbitrarily close to 1/2 (the compactness of a black hole), as investigated in [29] through specific models.

A. Formation of trapped surfaces

We preambled the exhibition of the various results presented in Sec. V by remarking that imposing $2m/r < 1$ is a common assumption for all of them. However, it is not always the case that this needs to be assumed: indeed, there are some results regarding conditions under which one can guarantee that the radius R of the fluid (where $p(R) = 0$) is reached without the formation of any trapped surface. The first such result, to the best of our knowledge, was proven by Baumgarte and Rendall in [34]:

Theorem (Baumgarte and Rendall [34]). *Let $\rho \geq 0$ be a continuous function of r.*

(i) (Isotropic case) *If $p = q$ and $0 < p_c < \infty$, then $2m(r)/r < 1$ for all $r \leq R$.*

(ii) (Anisotropic case) *If q and p are $\mathcal{C}^1$ functions (with $q(0) = p(0) = p_c > 0$), and p is bounded, then $2m(r)/r < 1$ for all $r \leq R$.*

This result was later generalized by Mars, Martín-Prats, and Senovilla in [35]:

Theorem (Mars, Martín-Prats, and Senovilla [35]). *Assume that the spacetime metric is piecewise $\mathcal{C}^2$ and regular at the center of symmetry, and let it satisfy Israel's junction conditions on (at most) a finite number of hypersurfaces, where appropriate. If $\rho + p + 2q \geq 0$, then $2m(r)/r \leq 1$ everywhere. If, furthermore, $p \geq 0$, then $2m(r)/r < 1$ everywhere.*

It is worth remarking that the condition $\rho + p + 2q \geq 0$ can be obtained as a consequence of the SEC. We close this section by pointing out that the aforementioned theorem by Bondi, Karageorgis, and Stalker, was actually proven in [33] without assuming $2m(r)/r < 1$. By arriving at the corresponding bounds, therefore, they also end up showing that no trapped surfaces can form in each one of those scenarios.

VI. CONCLUSION AND OUTLOOK

We have introduced from scratch all the tools needed to study the basic properties of static spherically symmetric spacetimes, to obtain Buchdahl's limit on the compactness of stable perfect fluids, and to understand the subsequent generalizations of Buchdahl's result obtained by other authors.

Besides its value as a pedagogical introduction to this fundamental problem, this review also provides a useful perspective of its current landscape. Several strong results have been established since Buchdahl's seminal work, but there is still much to be explored. Here, we highlight four clear directions in which there is room for further progress:

- Sharp bounds should be found (or proven) in those cases where optimality has not yet been established.

- So far, each result concerns a set of "reasonable" conditions. However, an exhaustive analysis of all the potentially "reasonable" conditions and their interactions (i.e., when more than one is imposed simultaneously) is still missing.

- The potential formation of trapped surfaces under different sets of assumptions should be analyzed systematically.
- In cases where the constructions proving sharpness involve interfaces, the fulfillment of Israel's junction conditions should be explicitly verified.

We hope to address at least some of these questions in future work.

---

[1] K. e. a. Akiyama, First m87 event horizon telescope results. i. the shadow of the supermassive black hole, Astrophys. J. Lett. **875**, L1 (2019).

[2] W. Collmar, N. Straumann, S. K. Chakrabarti, G. 't Hooft, E. Seidel, and W. Israel, Panel Discussion: The Definitive Proofs of the Existence of Black Holes, in *Black Holes: Theory and Observation*, edited by F. W. Hehl, C. Kiefer, and R. J. K. Metzler (Springer Berlin Heidelberg, Berlin, Heidelberg, 1998) pp. 481–489.

[3] M. A. Abramowicz, W. Kluźniak, and J.-P. Lasota, No Observational Proof of the Black-Hole Event Horizon, Astron. Astrophys. **396**, L31 (2002).

[4] R. Carballo-Rubio, V. Cardoso, and Z. Younsi, Toward very large baseline interferometry observations of black hole structure, Phys. Rev. D **106**, 084038 (2022), arXiv:2208.00704 [gr-qc].

[5] S. Chandrasekhar, Xlviii. the density of white dwarf stars, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **11**, 592 (1931), https://doi.org/10.1080/14786443109461710.

[6] S. Chandrasekhar, The Maximum Mass of Ideal White Dwarfs, Astrophys. J. **74**, 81 (1931).

[7] J. R. Oppenheimer and G. M. Volkoff, On massive neutron cores, Phys. Rev. **55**, 374 (1939).

[8] H. Epstein, V. Glaser, and A. Jaffe, Nonpositivity of the energy density in quantized field theories, Nuovo Cimento **36**, 1016 (1965).

[9] L. H. Ford, Constraints on negative-energy fluxes, Phys. Rev. D **43**, 3972 (1991).

[10] E. E. Flanagan, Quantum inequalities in two-dimensional minkowski spacetime, Phys. Rev. D **56**, 4922 (1997).

[11] C. J. Fewster, Lectures on quantum energy inequalities (2012), arXiv:1208.5399 [gr-qc].

[12] R. M. Wald, *General Relativity* (Chicago University Press., 1984).

[13] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, 1973).

[14] R. C. Henry, Kretschmann Scalar for a Kerr-Newman Black Hole, Astrophys. J. **535**, 350 (2000).

[15] J. M. M. Senovilla, Trapped Surfaces, Int. J. Mod. Phys. D **20**, 2139 (2011).

[16] C. W. Misner and D. H. Sharp, Relativistic equations for adiabatic, spherically symmetric gravitational collapse, Phys. Rev. **136**, B571 (1964).

[17] J. Hernandez, Walter C. and C. W. Misner, Observer Time as a Coordinate in Relativistic Spherical Hydrodynamics, Astrophys. J. **143**, 452 (1966).

[18] W. Israel, Singular Hypersurfaces and Thin Shells in General Relativity, Nuovo Cimento B **44**, 1 (1966).

[19] H. A. Buchdahl, General Relativistic Fluid Spheres, Phys. Rev. **116**, 1027 (1959).

[20] Z. Roupas, Thermodynamical Instabilities of Perfect Fluid Spheres in General Relativity, Class. Quantum Grav. **30**, 115018 (2013).

[21] Z. Roupas, Corrigendum: Thermodynamical Instabilities of Perfect Fluid Spheres in General Relativity (2013 Class. Quantum Grav. 30 115018), Class. Quantum Grav. **32**, 119501 (2015).

[22] P. S. Florides, A New Interior Schwarzschild Solution, Proc. R. Soc. Lond. A **337**, 529 (1974).

[23] R. L. Bowers and E. P. T. Liang, Anisotropic Spheres in General Relativity, Astrophys. J. **188**, 657 (1974).

[24] H. Heintzmann and W. Hillebrandt, Neutron Stars with an Anisotropic Equation of State: Mass, Redshift and Stability, Astron. Astrophys. **38**, 51 (1975).

[25] A. Füzfa, J.-M. Gerard, and D. Lambert, The Lemaître–Schwarzschild Problem Revisited, Gen. Relativ. Gravit. **34**, 1411 (2002).

[26] J. Guven and N. Ó Murchadha, Bounds on $2m/R$ for static spherical objects, Phys. Rev. D **60**, 084020 (1999).

[27] H. Andréasson, On Static Shells and the Buchdahl Inequality for the Spherically Symmetric Einstein-Vlasov System, Commun. Math. Phys. **274**, 409 (2007).

[28] B. V. Ivanov, Maximum bounds on the surface redshift of anisotropic stars, Phys. Rev. D **65**, 104011 (2002).

[29] J. Arrechea, C. Barceló, G. García-Moreno, and J. Polo-Gómez, Dissecting Buchdahl's limit: A surgeon's guide to compact objects (2024), arXiv:2411.14018 [gr-qc].

[30] D. Barraco and V. H. Hamity, Maximum mass of a spherically symmetric isotropic star, Phys. Rev. D **65**, 124028 (2002).

[31] H. Andréasson, Sharp Bounds on $2m/r$ of General Spherically Symmetric Static Objects, J. Differ. Equations **245**, 2243 (2008).

[32] H. Bondi, Massive Spheres in General Relativity, Proc. R. Soc. Lond. A **282**, 303 (1964).

[33] P. Karageorgis and J. G. Stalker, Sharp Bounds on $2m/r$ for Static Spherical Objects, Class. Quantum Grav. **25**, 195021 (2008).

[34] T. W. Baumgarte and A. D. Rendall, Regularity of Spherically Symmetric Static Solutions of the Einstein Equations, Class. Quantum Grav. **10**, 327 (1993).

[35] M. Mars, M. M. Martín-Prats, and J. M. M. Senovilla, The $2m \leq r$ Property of Spherically Symmetric Static Spacetimes, Phys. Lett. A **218**, 147 (1996).