

Alternative route towards the change of metric signature

Fan Zhang 

*Gravitational Wave and Cosmology Laboratory, Department of Astronomy,
Beijing Normal University, Beijing 100875, China
and Department of Physics, West Virginia University, PO Box 6315, Morgantown,
West Virginia 26506, USA*



(Received 31 July 2019; published 23 September 2019)

Beginning with Hartle and Hawking’s no-boundary proposal, it has long been known that the pathology of a big bang singularity can be suppressed if a transition into the Riemannian (Euclidean) metric signature (the usual singularity theorems become invalid in this region) occurs when we track back along cosmic time. A vital component of this type of models, which needs to be clarified, is the set of junction conditions at the boundary between the two signature regimes. In the traditional approach, the signature change occurs in the temporal sector through a switch of sign in the lapse-squared function. Motivated by more straightforward connections with the big bang cosmology, we explore here an alternative whereby the spatial metric eigenvalues change sign instead, so that the Riemannian side is purely timelike. We investigate the junction conditions required in this case.

DOI: [10.1103/PhysRevD.100.064043](https://doi.org/10.1103/PhysRevD.100.064043)

I. INTRODUCTION

A. Standard approaches

Discussions on signature changing spacetimes were arguably ignited by Hartle and Hawking’s no-boundary proposal for the initial conditions of the Universe [1,2]. The study of the semiclassical approximations [3] to the wave function of the Universe, especially the dominating real tunneling solutions (a real Riemannian spacetime joined onto a real Lorentzian one, with the Riemannian part determining the weighting in the path integral) [3,4], had garnered some interest.

While studying the tunneling solutions, it immediately became clear that it is impossible to transition a solution of the Einstein equations into the Riemannian signature in an uneventful manner because even a continuous metric will necessarily become either degenerate or divergent. In other words, General Relativity (GR), without any relaxations, is not intrinsically capable of dealing with signature changes (the transition surface is at best a mild singularity). However, if one is only interested in semiclassical approximations to quantum wave functions, the classical Einstein equations only need to be “almost” satisfied, in the sense that some pathologies on the transition surface are allowed so long as they do not spoil the steepest descent considerations by making a divergent contribution to the action [3]. Furthermore, even if we throw away such leniency afforded by quantum mechanics and consider, as in Refs. [5,6], purely classical setups, arguments can be made that suitably weaker versions of GR equations are not outrageous, since, after all, there are many situations that

the standard GR formalism cannot handle, such as when it comes to singularities inside black holes or impulsive gravitational waves, which do not appear to be prohibited by nature.

Broadly speaking, depending on the functional space from which one draws solutions to the Einstein’s equations that are formally ill defined (not just singular in a differential equation sense like when some higher-derivative terms vanish; some quantities appearing in the equations may become divergent and thus not defined) at a change-of-signature boundary, two types of junction conditions have been proposed in the literature (both for when a purely spatial Riemannian side is reached via the temporal eigenvalue of the metric switching sign):

- (i) 1: A more flexible one (e.g., Refs. [3,5–9]) allowing for discontinuous metrics with a continuous but not necessarily vanishing extrinsic curvature of the signature change surface Π , suitable for distributional solutions. The Einstein equations themselves remain ill defined at Π , so by “the distribution is a solution,” those authors mean that it satisfies the equation at any point away from that surface, while the equation is suspended on Π .
- (ii) 2: A more restrictive one (e.g., Refs. [10–14]) requiring the metric to be continuous and the extrinsic curvature to vanish when computed from both sides. This set of conditions is suitable for smoother solutions satisfying a regularized version of the Einstein equations that are not suspended on Π . Specifically, those offending ill-defined quantities are in fact well defined off of Π , so their limits

can possibly be obtained through a process asymptoting to Π , and the broken expressions are then replaced by such limits (and strong junction conditions are required for these limits to exist). One must note that only the covariant form of the equations is regularized, and the inverse metric still diverges, so not everything is made regular in this approach.

Since the extrinsic curvature is the time derivative of the spatial metric (its trace is essentially the rate at which spatial volume grows), its suppression is often said to imply stationarity. Indeed, similar analysis on other fields propagating on the signature-changing background also analogously possess vanishing velocities. This is easy to see from a naive limit-taking analysis of a toy massless Klein-Gordon equation,

$$\varphi^{;a}{}_{;a} = \frac{1}{\sqrt{|g|}} (\sqrt{|g|} g^{ab} \varphi_{;b})_{;a} = 0, \quad (1)$$

where the semicolon denotes a covariant derivative and the comma denotes a partial derivative. The early part of the latin alphabet will denote spacetime indices, and the middle part will denote the spatial ones. Let g^{ab} be diagonalized in our choice of two-dimensional (for illustration) coordinates (t, x) into

$$\begin{pmatrix} -\lambda_t(t, x) & 0 \\ 0 & \lambda_x(t, x) \end{pmatrix}; \quad (2)$$

then, the equation becomes

$$\begin{aligned} \frac{\lambda_x(\lambda_{t,x})}{\lambda_t} \varphi_{;x} + \lambda_{t,t} \varphi_{;t} + \lambda_t \left(2\varphi_{;tt} - \frac{\lambda_{x,t}}{\lambda_x} \varphi_{;t} \right) \\ - \lambda_{x,x} \varphi_{;x} - 2\lambda_x \varphi_{;xx} = 0. \end{aligned} \quad (3)$$

When approaching the temporal signature-change surface Π , we must have $\lambda_t \rightarrow \infty$ (since it is an entry in the inverse metric) and generically also $\lambda_{t,t} \rightarrow \infty$ at an even faster pace, resulting in the requirements of $\varphi_{;t} \rightarrow 0$ and $\varphi_{;tt} \rightarrow 0$ in order for the equation to admit a well-defined limit on Π .

There are essentially two steps involved in deriving these conditions. The first is to evoke more or less the generic Darmois junction condition (denoted \mathcal{C}_g below) that the surface metric implied (through pullbacks of the embedding maps) by either side must agree so there is a well-defined 3-geometry for the boundary surface and also that the extrinsic curvatures computed on either side must agree to avoid having to confine a stress-energy tensor onto the spacelike boundary (matter worldlines cannot be entirely confined to a spacelike surface) [15–17]. Although these conditions are derived in the constant signature case, they

essentially remain unchanged in the signature-changing situation (note that with condition 1 the jump is in the time-time component of the metric, while the spatial part remains continuous, so the implied intrinsic spatial geometries from the two sides still agree).

The second type of requirements (denoted \mathcal{C}_s below) is specific to the singular (with degenerate or discontinuous metric) signature-changing situation. With condition 2, \mathcal{C}_s is the vanishing of the matching extrinsic curvatures, which allows a version of the Einstein's equations to be imposed on the transition surface, but is unsurprisingly quite rigid [18]. The condition 1 approach, on the other hand, aims for more flexibility by not imposing any \mathcal{C}_s at all, arguing that the extra step of regularizing a singular equation is more a matter of choice than necessity [19]. The price it pays is a relaxation of the sense in which the resulting solutions are unique [5,18]. The differences between the approaches reflect alternative philosophies, perhaps of how universally valid the standard form of the Einstein equations should remain when their usual underlying assumptions are tempered with.

B. Alternative

In this paper, we investigate an alternative mechanism by which a signature change can be achieved, following more closely the approach of condition 2, since we wish to see if the restrictions imposed by the regularization procedure, onto the initial conditions (for our Lorentzian universe) laid down on our transition surface Σ , can help explain some cosmological fine-tuning issues. So, the equations of motion of metric and matter, for which the initial conditions are meant, must not be suspended on Σ .

We begin by noting that, while having the temporal metric eigenvalue [$1/\lambda_t$ in the notation of Eq. (3), since λ_t is an eigenvalue of the inverse metric] go through zero (we shall call this approach route A in this paper), either continuously or with a jump, is taken to be the default in previous literature, it is not the only way for the metric signature to change. Having it go through ∞ (equivalently λ_t through zero) is also valid, since ∞ is just the antipodal end of the stereographic projection circle of the real line. However, with this approach (route A') in its raw form, the integration measure $\sqrt{-g}$ diverges on Σ , which has adverse side effects with quantum path integrals (the logic of steepest descent that makes our classical investigation useful in a quantum context may be spoiled [3]). A related approach that removes this problem is to have $1/\lambda_x$ go through zero instead, so that the signature becomes Riemannian not because time changes sign but because the spatial signature reverses. This alternative (route B) is related to route A' since the $1/\tilde{\lambda}_t = \lambda_x/\lambda_t$ of the conformally rescaled metric $\tilde{g}_{ab} = g_{ab}\lambda_x$ (which shares the same causal structure as g_{ab} , such as those depicted in the figures below) goes through ∞ . In other words, when the physical metric

transitions via route B, the conformal metric changes via route A'.

This route B has not been examined in any of the previous literature that we found, and investigating it is the subject of this paper. Aside from filling in a gap in the literature to achieve pedagogical thoroughness, we note that route B possesses some features that might help make it physically relevant:

- (1) As compared to route A, it is more straightforward to make connections with our actual Universe when we adopt route B because the Friedman-Lemaître-Robertson-Walker (FLRW) metric

$$\begin{aligned} g_{ab}dx^a dx^b &= a(t)^2 \tilde{g}_{ab} dx^a dx^b \\ &= -dt^2 + a(t)^2 \hat{\gamma}_{ij} dx^i dx^j \\ &\equiv -dt^2 + a(t)^2 (d\chi^2 + \zeta_\kappa^2(\chi) d\Omega^2), \end{aligned} \quad (4)$$

whereby

$$d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2 \quad (5)$$

and

$$\zeta_\kappa = \begin{cases} \sin \chi, & \text{for } \kappa = 1 \\ \chi, & \text{for } \kappa = 0 \\ \sinh \chi, & \text{for } \kappa = -1, \end{cases} \quad (6)$$

is automatically a route B-compatible metric. This means that, within route B, a transition into a Riemannian region can occur at the beginning [where $a(t) = 0$ so the spatial metric becomes fully degenerate and ready to be continued further into negative definiteness] of the prevailing cosmological model, extending it beyond the big bang,¹ but without needing significant alterations to the currently prescribed post-big bang evolution, which would not have been economic since any such alterations must be reconciled with observations.

We will keep the subsequent discussion in this paper general and not specialize to FLRW unless specifically noted. Nevertheless, it is helpful to always have this particularly well-studied and physically relevant special case in mind for intuition building.

- (2) Route B corresponds to the light cones opening up as one approaches the change-of-signature surface Σ from the Lorentzian side (see Fig. 1 for a visual depiction), since equal temporal increments would

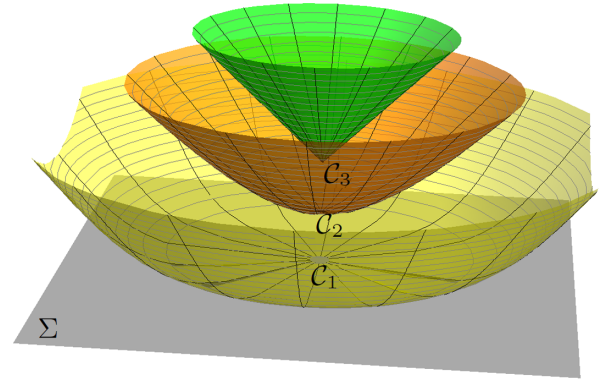


FIG. 1. As the base points C_3 , C_2 , and C_1 incrementally approach the signature change surface Σ , their null cones flatten out (and return to rising more steeply once they are sufficiently far away from Σ).

require increasingly greater spatial coordinate intervals to compensate in the $g_{ab}dx^a dx^b = 0$ equation for the null rays. As suggested by Fig. 1 and will be discussed in more details in Sec. II A below, the null cone structure is removed (it cannot exist in the Riemannian side) in route B via the future and past null cones opening up to collide with and annihilate each other, so it is the spacelike region that is squeezed out of existence, and the Riemannian side is purely temporal, as the metric signature obviously confirms. In contrast, the cones disappear in route A (condition 2) by separately closing up into a couple of half-lines, which then vanish beyond Π . In other words, the timelike regions are the ones taken out in that approach, and the Riemannian side is purely spatial.

A complication of that latter method is then that it takes constructive efforts (e.g., use different definitions for the geodesic Lagrangian when in alternative signature regimes [6]) to make timelike geodesics thread through Π , since if left alone they would have disappeared together with the timelike regions. In contrast, such intervention is unnecessary with route B, the Riemannian side of which is capable of hosting timelike curves. That such a continuation of timelike geodesics is required in the first place is due to the desire to show that the signature-change scenario no longer suffers geodesic incompleteness, so the big bang singularity is indeed removed in that particular sense, and one stays faithful to the original no-boundary proposal of Ref. [1]. This amelioration is possible because the usual singularity theorem [20,21] needs some causal properties that are no longer available in the Riemannian regime [5].

- (3) Following a procedure closely mimicking that of condition 2 but for route B, we obtain once again strong C_s conditions, but now including an additional one (C_s^2 of Sec. III A) enforcing the vanishing of spatial derivatives on Σ , in addition to the

¹Note that, contrary to common pictorial depictions, the big bang is not necessarily a single point, just a codimension-1 surface with a degenerate intrinsic metric—much like how distances along a null ray vanish, yet the null ray is not a single point. See Sec. II B below for more details.

temporal stationarity. Furthermore, the lapse function within route B can be set to a constant, so even more components of the 4-metric's derivatives vanish as compared to route A. Because these metric derivatives contribute to the curvature tensors, their suppression is beneficial for realizing the uniformity condition on the big bang, which is envisaged by the Weyl curvature conjecture [22]² to start the Universe off on low entropy (see also Ref. [23]). A signature-change universe via route B thus offers up an intriguing new way to supplement inflation in its quest to solve some cosmic puzzles.

In the rest of the paper, we turn to the details, beginning by establishing some basic properties of a route B transition in Sec. II, before finding the junction conditions in Sec. III. We finally conclude in Sec. IV with a discussion on the many studies required to more thoroughly explore the viability and properties of route B. For the Lorentzian side, we adopt signature $(-, +, +, +)$, and the Riemannian side subsequently has $(-, -, -, -)$.

II. LARGE-SCALE FEATURES

A. Signature morphology

A complication that route A did not suffer but route B must now face is that, since three dimensions now switch signature, there is the possibility that the three switches occur sequentially, instead of simultaneously as in the FLRW example. Specifically, consider the generic metric in the 3 + 1-form [24]

$$g_{ab}dx^a dx^b = -\alpha^2 dt^2 + \gamma_{ij}(\beta^i dt + dx^i)(\beta^j dt + dx^j), \quad (7)$$

where we fix the gauge freedoms by setting lapse $\alpha \equiv 1$ and shift $\beta^i \equiv 0$ so as to pick Gaussian normal (synchronous) coordinates, the temporal coordinate curves of which are timelike geodesics. Starting from an arbitrary coordinate system $x^{\bar{a}}$, we can find the Gaussian normal coordinates by solving the Hamilton-Jacobi equation

$$g^{\bar{a}\bar{b}}S_{,\bar{a}}S_{,\bar{b}} = -1, \quad (8)$$

for which the real solution S (it is to be the new time coordinate) exists even as the signature is allowed to vary, because we always have at least one timelike dimension within route B. A subtlety is that at places of signature change, some spatial³ components in the inverse metric $g^{\bar{a}\bar{b}}$ may diverge, but well-defined limits exist for these locations if the spatial derivatives of S simply vanish

²Incidentally, it was noted in this paper that something along the lines of the Hartle-Hawking no-boundary proposal may lead to the required condition.

³We will slightly abuse terminology in the interest of brevity and assign the label “spatial” to the other coordinates that are not t , even though their associated dimension can become null or timelike.

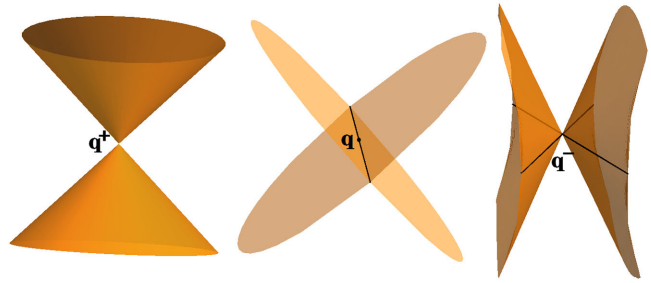


FIG. 2. If a single eigenvalue of γ_{ij} turns negative at q , then at a later time q^+ along a timelike path through q , the future and past null cones remain separated as per usual with a Lorentzian metric signature; at q , the two very elongated cones touch along a (black in the figure) line that is the eigenvector (eigenvectors are not directional, so the black line is a full and not a half-line) corresponding to the eigenvalue that now becomes zero, and thus vectors along this direction now have zero norms; at some earlier time q^- , the two cones partially “annihilate,” splitting the black line. The timelike interiors of the cones join up through the opened-up gap between the pair of black lines indicating the intersection lines between the cones. When all three eigenvalues change sign at the same point q (not plotted), the cones will open up towards each other as we approach q from q^+ , and the entireties of the two cones collide and annihilate at q , so there are no null cones at q^- , and all directions are timelike—the spacetime becomes Riemannian.

sufficiently quickly there. Once the solution is found, the Gaussian normal coordinate system can be constructed by following the standard textbook recipe. In this new coordinate system, which we adopt for expositional clarity, γ_{ij} is positive definite in a usual Lorentzian region, but its eigenvalues can transition, either one at a time or several together, into negative values. We can understand what this physically means by examining what happens to the null cones when one or more eigenvalues turns to zero and then negative through a transition point q . The situation is depicted in Fig. 2, and the local tangent space geometry can be intuited as future and past null cones colliding and annihilating, allowing their timelike interiors to merge.

The possibilities of partial sign switches and thus more diversified signature configurations are intriguing, but physically problematic. For example, a scalar field in a signature $(-, -, +, +)$ spacetime region would propagate via an ultrahyperbolic equation, which is generically (unless nonlocal constraints are imposed [25]) ill posed [26] when evolved off of any Cauchy surface (on the other hand, an elliptic equation in a Riemannian region admits well-posed boundary value problems). Beyond the scalar field, Ref. [27] also showed that fields with finite spins greater than zero cannot be defined in a signature $(-, -, +, +)$ spacetime region (they are, however, allowed in a Riemannian region). This implies a rather strange requirement in which a Dirac field describing, say, electrons can exist in the Lorentzian region but must somehow collude with spacetime in a fashion far beyond simply warping it and stop existing (not just becoming zero in amplitude) as soon as one spatial direction mutates.

We therefore need at least two spatial directions to switch simultaneously, leading to a time-space swapped Lorentzian spacetime of signature $(-, -, -, +)$. All massive particles must now become essentially long-lived “tachyons” [28], since they now move outside of light cones (centered on the remaining spatial direction) in order to follow timelike worldlines. Because the mathematics for a quantum field theory in this region is the same as in the regular Lorentzian signature (in fact, the sign convention adopted in particle physics is the opposite of that used by relativists, and this paper, so no sign changes are even needed when lifting formulas from books), one is then faced with all the vacuum instability issues and other pathologies that tachyons bring. We therefore assume, from here onward, that sequential sign switches are forbidden and that all three spatial dimensions switch simultaneously, giving us $(-, -, -, -)$ straight away.

B. FLRW junction surface

Continuity of the metric requires that any curve linking two points of different signatures must intersect the bounding wall Σ at least once, so Σ should at most have codimension 1 (the curve itself takes up one codimension, and if there exists another, the curve’s intersection point with Σ , as well as the surrounding sections to preserve continuity, can be shifted in that direction to avoid Σ) but does not need to be a constant t surface. The case of the highly symmetric FLRW is much simpler, though, and, due to its cosmological relevance, worthy of us taking a little detour to clarify. We emphasize, though, that most of our discussions on the junction conditions in Sec. III are not confined to this case and are valid for generic Σ s. They are local considerations relating to limit-taking procedures along an arbitrary single timelike curve threading through Σ at a single point and are as such independent of the larger-scale properties of Σ .

The first thing one notices is that the FLRW Σ is null, since $g_{ab}dx^a dx^b$ vanishes at $a = 0$ for any separations confined to Σ (with $dt = 0$ since a is a function of t only). It should be noted, though, that points on Σ can be macroscopically separated yet null related just like two points along a null ray, in which case Σ is not a single point as often depicted for the big bang, just like a null ray is not a single point. This situation arises because the Lorentzian metric is rather pathological for the purpose of defining open sets (metric balls are noncompact) and studying topology (thus the frequent adoption of a positive definite auxiliary metric in some topological studies; see, e.g., Refs. [29,30]). Baring any direct observational consequences of the Riemannian side from which the topology of Σ may become more obvious, the best hope we have to ascertain its nature may be to assume global hyperbolicity of the Lorentzian side, the topology of which should then be a direct product $\mathbb{R} \times \Sigma$ [3]. Since our Universe is not one dimensional, Σ cannot be a zero-dimensional point.

More explicitly, the observational evidence is that the spatial slices of our Universe appear to be flat [31] and thus could well be infinite in extent.⁴ It would then be quite strange for such an infinite noncompact plane to instantaneously collapse into a single point (a compact singleton) as soon as the scale factor reaches precisely zero, when it would still be noncompact for any infinitesimal value of $a > 0$. In that scenario, the early Universe would not resemble the collar neighborhood of Σ , which would obviously adversely affect our ability to evolve initial conditions off of Σ to uniquely determine the Lorentzian side of the Universe.

Incidentally, in the case of the flat slicing of de Sitter (see, e.g., Fig. 1 of Ref. [33]) serving as an isometry to an inflationary FLRW, the finite comoving observers (those labeled by finite spatial comoving coordinates) do in fact all get packed into a single asymptotic point of the de Sitter spacetime when traced back in time. The abrupt jump issue in this case is resolved by pulling in points from comoving spatial infinity to form an extended noncompact border surface. This is fine for de Sitter, since points on this surface are just regular points inside the actual de Sitter spacetime. Their carrying infinite spatial coordinates is simply due to the flat foliation coordinates being singular (a symptom is that this coordinate system cannot be extended beyond this border to cover the other half of de Sitter). However, for the actual FLRW universe, there is no reason to believe that the comoving coordinate system, as preferred by the observed motion of matter, is ill chosen and ill behaved, so similar infinity points would likely genuinely reside on the spatial compactification boundary. In other words, such infinity points are outside of the actual spacetime (similar to how the future null infinity \mathcal{I}^+ [34] is outside of an asymptotically flat spacetime itself), and are purely mathematical constructs. They are as such not in fact physically available to help smooth out the aforementioned “jump”, or as places to prescribe junction conditions on. In short, while the inflationary FLRW and de Sitter are isometric for the post-big bang segment, they likely differ when it comes to the topological structure of the big bang itself, which is not a radical prospect given that they already differ on what lies beyond.

Finally, as an aside, it is also worthwhile noting that the FLRW big bang is sometimes said to be spacelike, but this characterization is under the conformal metric rather than the physical metric, and the choice is not unique. Specifically, there is a well-established field of study on the “conformal gauge singularities” (regarding the big bang

⁴They could also be flat tori or other twisted alterations [32], but a point would still have the wrong dimension for a boundary of the four-dimensional Lorentzian universe and would instead be an interior point, resulting in the big bang cosmology already being “no boundary” even without introducing a Riemannian region, depriving us of a place on which to prescribe initial conditions (necessary for Cauchy evolution on the Lorentzian side).

singularity as being due to the special “conformal gauge choice” in a conformal class of mostly regular metrics) [35–39]. Even the well-posed-ness of the Cauchy problem for various matter content types has been proven for this construct [36,40–44]. It is also useful for us to think about the causal structure of the spacetime using the conformal metric \tilde{g}_{ab} , but we stop short of carrying out the additional temporal transformation $t \rightarrow \tau$ defined by

$$\frac{d\tau}{dt} = \frac{1}{a(t)}. \quad (9)$$

After this extra layer of coordinate transformation, the FLRW metric becomes conformally flat. However, Eq. (9) is singular at $a = 0$, and since t is the intrinsic clock carried by physical comoving observers, results obtained under τ must be fed through an additional singular transformation before it can be translated back into predictions on physical experimental outputs. The reward for this extra trouble is flexibility. Specifically, since $dt/d\tau = 0$ at the big bang, the condition of moving along the constant t surface, as expressed by $dt = 0$, can be satisfied by any finite $d\tau$ choice. Instinctively, one picks $d\tau = 0$, which gives a conformally spacelike (under the physical metric, it is still null) big bang, but one could actually equally well choose other $d\tau$'s that make it conformally null or even timelike. The arbitrariness is because that, essentially, via an infinite stretching, the zero-thickness three-dimensional Σ got stretched into a four-dimensional object. While people still customarily pick out a three-dimensional surface in there and call it the big bang, it perhaps should have been the whole four-dimensional totality. Regardless, if one holds the view that this newly inserted internal structure to the big bang is physical, then its flexibility would allow for establishing beautiful mathematical infrastructures. We will remain more parsimonious in this paper, though, and formulate the junction conditions under the physical metric.

III. JUNCTION CONDITIONS

A singular differential equation can be well defined at its singular set, e.g., $x\beta_{,xx} + \beta = 0$ at $x = 0$. However, with our toy Eq. (3) or Einstein's equations, the coefficient functions appearing in the equation or the curvature expressions become divergent or otherwise ill defined (e.g., $0/0$) on Σ . So, strictly speaking, the equations are not merely singular; they are not formally defined there. Nevertheless, we could follow the approach of Ref. [12] and regularize the offending divergences by imposing strict junction conditions so that the equations admit well-defined limits on Σ . Solutions satisfying such conditions can then be sought such that the Σ limits of the left- and right-hand sides of the equations match. This distills a set of equations of motion to be satisfied on Σ , so physics will not be left completely arbitrary there, but as already alluded to in Sec. I A, Σ cannot be rendered completely regular. At the

very least, the inverse metric still diverges, and some of the Carminati-McLenaghan curvature invariants [45] might do so as well. Our present endeavor is a modest attempt at partially resolving the big bang singularity in order to glean some information on the likely behavior of the important classical saddle-point solutions; it is not aimed at removing the singularity altogether, a task for which an understanding of quantum gravity is probably required (but the intriguing possibility of accomplishing it even at a classical level, perhaps through the adoption of more topology-friendly auxiliary metrics, should not be dismissed out of hand; for such an investigation, our study would serve as a first step to demonstrate how far one can go without bringing in additional infrastructures and to identify the remaining problems they must solve, thereby cluing us in on where new physics/mathematics might come in as well as what they might look like).

A. Einstein's equations

1. Method

We begin with the left-hand side of the Einstein equations. Following standard literature [17], under Gaussian normal coordinates, the Einstein tensor can be written in the $3 + 1$ -form as

$$G_{tt} = \frac{1}{2}({}^3R) + \frac{1}{2}[K^2 - \text{tr}(K \cdot K)], \quad (10)$$

$$G_{ti} = -K_i^m{}_{|m} + K_{|i}, \quad (11)$$

$$G_{ij} = {}^3G_{ij} - \left\{ (K_{ij} - K\gamma_{ij})_{,t} + 2K_{ik}K^k{}_j - 3KK_{ij} + \frac{1}{2}K^2\gamma_{ij} + \frac{1}{2}\text{tr}(K \cdot K)\gamma_{ij} \right\}, \quad (12)$$

where $K_{ij} = -\gamma_{ij,t}/2$ is the extrinsic curvature of the constant t slice (not necessarily coincident with Σ) and the vertical bar denotes a three-dimensional covariant derivative. The first two equations do not contain temporal derivatives and are the Hamiltonian and momentum constraints, respectively. The third equation tells us how to evolve the metric in time. Note that these expressions are valid on both sides of Σ , since, unlike with route A, there is no change to the norm of the normal vector ∂_t of the spatial slices within route B; thus, none of the explicit signs in Eqs. (10)–(12) needs to change; the signature changes are all hidden inside the symbolic spatial quantities, just as they are all hidden inside the γ_{ij} in the metric (see Eq. (7)). We therefore will not explicitly distinguish between the Lorentzian and Riemannian sides in the derivations below, since all expressions are identical.

Because γ^{ij} is the source of divergences at Σ , the terms

$$K \equiv \gamma^{ij}K_{ij}, \quad \text{tr}(K \cdot K) \equiv \gamma^{ij}\gamma^{kl}K_{ik}K_{jl},$$

as well as $K_i^m{}_{|m}$ and $K_{ik}K^k{}_j$, could all obviously diverge. Furthermore, since (also due to other contractions with the inverse metric in trace-taking computations)

$${}^{(3)}\Gamma_{jk}^i = \frac{1}{2}\gamma^{il}(\gamma_{kl,j} - \gamma_{jk,l} + \gamma_{lj,k}), \quad (13)$$

the terms ${}^{(3)}R$ and ${}^{(3)}G_{ij}$ could diverge there as well. The goal, following the arguments of Ref. [12], is to see what conditions arise from demanding that G_{ab} remains bounded in the Σ limit. It should be noted that this approach demands componentwise regularity for the Einstein tensor (because the Einstein equations are in component form), but being explicit tensor components, the expressions (10)–(12) depend on the underlying coordinate basis onto which the tensor is decomposed, and this basis could be ill behaved even when the underlying geometry is perfectly fine (e.g., if caustics develop for the congruence of timelike geodesics underlying the Gaussian normal coordinate system, due to a bad choice of initial velocities). This coordinate singularity issue is familiar and not specific to the problem at hand, but it is nevertheless worth emphasizing that it implies the conditions $\mathcal{C}_s^{1/2}$ we obtain in the next section are (unfortunately unavoidably) sufficient but not necessary.

They are also quite strong in another way, as they will demand that all the terms in G_{ab} that could possibly diverge would instead remain finite, in an individual termwise fashion. There is, of course, also the possibility that divergences cancel across terms. To find these cases, one needs to solve differential regularity equations derived from the condition that the divergent terms in Eqs. (10)–(12) are curbed, which is technically difficult without assuming symmetries to simplify expressions, but doing so would defeat the purpose of trying to find out what kind of constraints that regularity at Σ would place on our Universe. Instead, we deploy generic considerations to argue that such solutions would unlikely be numerous (or indeed exist at all), so at the very least, the solutions given by $\mathcal{C}_s^{1/2}$ would not be unlikely as physically relevant junction conditions from a statistical point of view.

We begin by noting that Eqs. (11) and (12) contain terms involving both one and two factors of the inverse spatial metric γ^{ij} , which diverge at different rates and have to be treated separately. Schematically, write $\gamma^{ij} \sim 1/\zeta$ with $\zeta \rightarrow 0$ when approaching Σ ; then, Eq. (11) or (12) could be stylized as

$$\frac{A(\zeta)}{\zeta} + \frac{B(\zeta)}{\zeta^2} = \frac{1}{\zeta} \left(A + \frac{B}{\zeta} \right), \quad (14)$$

where A and B are nondivergent at $\zeta = 0$ since we have collected all the problematic terms into powers of $1/\zeta$. Regularity then requires that

$$A = -\frac{B}{\zeta} + \mathcal{O}(\zeta), \quad (15)$$

and since A cannot diverge (but can be nonvanishing) when $\zeta \rightarrow 0$, we also need

$$B = \mathcal{O}(\zeta). \quad (16)$$

The cross-cancellations thus allow for more relaxed A and B than what termwise regularity would demand, which is $A = \mathcal{O}(\zeta)$ and $B = \mathcal{O}(\zeta^2)$ [these select a subset of solutions to Eqs. (15) and (16) and are not alternatives to them].

The catch is that for each original Eq. (14) we end up with twice as many regularity conditions (15) and (16). This means that Eqs. (11) and (12) would demand 2×3 and 2×6 regularity equations, respectively, while Eq. (10) adds another few. Furthermore, the three eigenvalues γ_i^e could all vanish at different rates, so instead of just two powers as in our stylized example, Eq. (14), there are in fact more distinct divergence rates, spawning a great many regularity equations. On the other hand, there are only 12 independent components in the variables γ_{ij} and K_{ij} ,⁵ so the coupled set of regularity equations is heavily overdetermined, thus generically not admitting solutions beyond the trivial ones identified by $\mathcal{C}_s^{1/2}$. By triviality,⁶ we mean that individual terms in each equation are all pushed below the “error

⁵Here, we are taking the Hamiltonian approach of Arnowitt-Deser-Misner [24], in which γ_{ij} and K_{ij} are regarded as independent variables, each marching forward according to a first-derivative-in-time evolution equation. One can, of course, also take the Lagrangian view and see γ_{ij} as the only fundamental variable, governed by a second-derivative-in-time evolution equation. By definition, the regularity equations are there to limit what initial conditions one can place on Σ (they are allowed to be underdetermining), and for these initial conditions, one can either lay down 6 initial values for γ_{ij} and K_{ij} each or 12 for γ_{ij} alone (since it is then governed by a second-order equation, one should give both Dirichlet and Neumann conditions). The number of required initial values is always 12, and they must satisfy the Σ limit of the regularity equations, which is generically impossible if there are more than 12 such equations.

⁶The regularity equations are stronger than just implicit initial conditions confined to Σ , since they also constrain the variables at small but nonvanishing ζ values (for well-defined limits to exist for G_{ab} , just having the coefficients to the various powers of $1/\zeta$ vanishing on Σ is not enough; they also need to vanish sufficiently quickly as Σ is approached). Yet they are more relaxed than the usual exact equations in the interior, since small errors are allowed [e.g., stray ζ^2 terms are allowed for Eq. (15) due to the $\mathcal{O}(\zeta)$ provision]. For interior equations, the relevant number of free variables drops to 6 (the number of physical freedoms in a metric; the six freedoms in K_{ij} are removed by its definition as the time derivative of γ_{ij} , which gives a set of six constraint equations) from the 12 as for the initial conditions, exacerbating the overdeterminacy. Alternatively, one may stay with the boundary view and note that each of the regularity equations is stronger than just one initial condition, since one could expand it into powers of ζ (as a surrogate for t) around $\zeta = 0$ (i.e., the expansion coefficients are evaluated on Σ), and observe that coefficients to all powers lower than that inside \mathcal{O} must vanish (there is always at least one such coefficient, that of ζ^0 , but there could be more), translating into multiple initial conditions. Thus, the severity of overdeterminacy is underestimated in the main text, although already sufficient for our purpose. The triviality discussed here is in regard to these relaxed interior equations, and a more familiar notion of trivial solutions to overdetermined exact equation systems can be recovered by confining the discussion to Σ itself, as we have also done in this footnote.

budget” of that equation [e.g., $\mathcal{O}(\zeta)$ for Eq. (15)] by $\mathcal{C}_s^{1/2}$, so no strict equalities need to be actually enforced (to precisely balance/cancel out between quantities above the error tolerance threshold), resulting in the overabundance of equations all being rendered inert, left with no chance to conflict with one another.

For an illustrative example of how the error budget bestows flexibility, take the case $A \propto B^p$ with some fixed p prescribed by physics [i.e., there is only one free variable B , and Eqs. (15) and (16) are overdetermined], and let $B \propto \zeta^q$ be an ansatz solution of which the q is up to us to pick. Then, if we take up the more relaxed $q = 1$ as allowed by Eq. (16), we would have a chance of balancing Eq. (15) only in the fine-tuned case of $p = 0$. However, if $q = n \geq 2$ as required by $\mathcal{C}_s^{1/2}$, then any $p \geq 1/n$ is comfortably accommodated. This is because, in the latter case, the B/ζ term does not rise above the error tolerance $\mathcal{O}(\zeta)$ of the overall equation and thus does not require careful cancellation from A (as a result, p does not have to be of any particular value), which is not freely variable and thus is defective for fulfilling this role. Note that, although we have used cancellation across different powers of ζ for our example, analogous considerations as well as the triviality discussion of the last paragraph in general also apply to cancellations between terms contributing to the same power (i.e., A , and/or B , alone could further subdivide into a small number of contributors), which $\mathcal{C}_s^{1/2}$ also excludes.

2. Conditions

Near a temporal coordinate geodesic ξ of our Gaussian normal coordinate system x^a that threads through Σ , we can construct a principal coordinate system $x^{\hat{a}}$ under which $\gamma_{\hat{i}\hat{j}}$ is diagonalized on ξ , by first applying an $O(3)$ transformation within the spatial tangent space to diagonalize γ_{ij} there (this is always possible according to the spectral theorem since γ_{ij} is nonsingular real symmetric; we do not need $x^{\hat{a}}$ to be unique), and then lay down the spatial coordinates in an open tube surrounding ξ via the exponential map. Note that we do not normalize $\gamma_{\hat{i}\hat{j}}$ to unity (the coordinate basis $\{\partial_{x^{\hat{a}}}\}$ is not orthonormal), so the Jacobian transforming between the two coordinate systems remains well behaved along ξ , even as we approach Σ (always just a block-diagonal matrix with a regular orthogonal matrix for the spatial sector and unity for the temporal sector). Along ξ (where the spatial tangent spaces according to x^a and $x^{\hat{a}}$ coincide), the spatial tensors on the right-hand sides of Eqs. (10)–(12) can be computed as their counterparts in the $x^{\hat{i}}$ coordinate system multiplied for an appropriate number of times by the $O(3)$ spatial Jacobian, which is never divergent nor degenerate (always full ranked); thus, it suffices to examine the divergences under $x^{\hat{i}}$ in which the algebraic matrix operations reduce to those between the three eigenvalues $\gamma_i^e, i \in \{1, 2, 3\}$ shared by γ_{ij} and $\gamma_{\hat{i}\hat{j}}$,

which are positive on the Lorentzian side and negative on the Riemannian side.

Essentially, we have here a Fermi normal coordinate construction [46,47] with the addition of a rescaling step (on the parallelly transported spatial basis vectors) to recover the eigenvalues, thus ensuring that the divergences are not appropriated by the coordinates and are captured by $\gamma^{\hat{i}\hat{j}}$. Just like the Fermi coordinates, our principal coordinates cover the entire open tube, but the nice properties such as the metric $\gamma_{\hat{i}\hat{j}}$ being diagonal are only true on the geodesic ξ itself (a $[\xi]$ prefix below signifies expressions valid only on ξ). This is fine for us, though, since we are studying the limiting behaviors of quantities as we approach Σ along ξ , so we only ever need to evaluate such quantities on ξ . Therefore, in our computations, $\gamma_{\hat{i}\hat{j}}, \gamma^{\hat{i}\hat{j}}$, and their temporal partial derivatives (measuring changes along ξ) to any order (including $K_{\hat{i}\hat{j}}$ in particular) are diagonal,

$$\begin{aligned} [\xi]: \gamma_{\hat{i}\hat{j}} &= [\text{diag}_i \gamma_i^e]_{\hat{i}\hat{j}}, & [\xi]: \gamma^{\hat{i}\hat{j}} &= \left[\text{diag}_i \frac{1}{\gamma_i^e} \right]^{\hat{i}\hat{j}}, \\ [\xi]: K_{\hat{i}\hat{j}} &= \left[-\frac{1}{2} \text{diag}_i \gamma_{i,i}^e \right]_{\hat{i}\hat{j}}, \\ [\xi]: K &= -\frac{1}{2} \sum_i \frac{\gamma_{i,i}^e}{\gamma_i^e}, \\ [\xi]: K_{,\hat{i}} &= -\frac{1}{2} \sum_i \left(\frac{\gamma_{i,i\hat{i}}^e}{\gamma_i^e} - \left(\frac{\gamma_{i,i}^e}{\gamma_i^e} \right)^2 \right), \\ [\xi]: \text{tr}(K \cdot K) &= \frac{1}{4} \sum_i \left(\frac{\gamma_{i,i}^e}{\gamma_i^e} \right)^2, \\ [\xi]: K_{\hat{i}}^{\hat{m}} &= \left[-\frac{1}{2} \text{diag}_i \frac{\gamma_{i,i}^e}{\gamma_i^e} \right]_{\hat{i}}^{\hat{m}}, \\ [\xi]: K_{\hat{i}\hat{k}} K^{\hat{k}\hat{j}} &= \frac{1}{4} \left[\text{diag}_i \frac{(\gamma_{i,i}^e)^2}{\gamma_i^e} \right]_{\hat{i}\hat{j}}. \end{aligned} \quad (17)$$

We then immediately see that the necessary and sufficient requirement for the following terms

$$K^2, \quad \text{tr}(K \cdot K), \quad (K_{\hat{i}\hat{j}} - K \gamma_{\hat{i}\hat{j}})_{,\hat{i}}, \quad K_{\hat{i}\hat{k}} K^{\hat{k}\hat{j}}, \quad K K_{\hat{i}\hat{j}} \quad (18)$$

inside $G_{\hat{a}\hat{b}}$ to individually remain bounded is the following (c.f., Ref. [12] for route A): for \mathcal{C}_s^1 , temporal derivatives of the spatial metric vanish at least as quickly as the spatial metric itself as Σ is approached, in the sense that $\gamma_{i,i}^e = \mathcal{O}(\gamma_i^e)$, $\gamma_{i,i\hat{i}}^e = \mathcal{O}(\gamma_i^e)$, $\forall i$.

Note that, although $K_{\hat{i}\hat{j}}$ vanishes on Σ , its trace K does not need to, since \mathcal{C}_s^1 allows for $\gamma_{i,i}^e = \Theta(\gamma_i^e)$; i.e., it allows the numerator and denominator in the K expression in Eq. (17) to vanish equally quickly when approaching Σ , so the limit of the ratio can be finite but nonvanishing.

The other terms in $G_{\hat{a}\hat{b}}$ not appearing in Eq. (19) involve spatial derivatives. In general, spatial derivatives of even the

off-diagonal entries in the tensorial quantities appearing in Eq. (17) do not necessarily vanish, since these quantities can be nondiagonal off ξ . Nonetheless, because the principal coordinate system is constructed via the exponential map, we must have a vanishing connection,

$$[\xi]: {}^{(3)}\Gamma_{\check{j}\check{k}}^{\check{i}} = 0, \quad (19)$$

and subsequently

$$[\xi]: \gamma_{\check{i}\check{j},\check{k}} = \gamma_{\check{i}\check{l}}{}^{(3)}\Gamma_{\check{j}\check{k}}^{\check{l}} + \gamma_{\check{j}\check{l}}{}^{(3)}\Gamma_{\check{i}\check{k}}^{\check{l}} = 0. \quad (20)$$

Since

$$0 = \delta_{\check{i}}^{\check{j}} = (\gamma_{\check{i}\check{l}}\gamma^{\check{l}\check{j}})_{,\check{k}} = \gamma_{\check{i}\check{l},\check{k}}\gamma^{\check{l}\check{j}} + \gamma_{\check{i}\check{l}}\gamma^{\check{l}\check{j}}_{,\check{k}}, \quad (21)$$

Eq. (20) further yields

$$[\xi]: \gamma^{\check{m}\check{i}}\gamma_{\check{i}\check{l}}\gamma^{\check{l}\check{j}}_{,\check{k}} = \delta^{\check{m}\check{i}}\gamma^{\check{l}\check{j}}_{,\check{k}} = \gamma^{\check{m}\check{j}}_{,\check{k}} = 0. \quad (22)$$

Furthermore, since Eqs. (20) and (22) are true everywhere along ξ , temporal derivatives can be added to yield

$$[\xi]: \gamma_{\check{i}\check{j},\check{k}\check{i}} = 0 = \gamma^{\check{i}\check{j}}_{,\check{k}\check{i}}. \quad (23)$$

Equipped with these tools, we are now ready to tackle the first derivative terms in $G_{\check{i}\check{i}}$. Equation (19) reduces the covariant derivative to a partial derivative, and then by Eqs. (22) and (23), we have

$$\begin{aligned} [\xi]: K_{\check{i}}^{\check{m}}|_{\check{m}} &= -\frac{1}{2}(\gamma^{\check{m}\check{j}}\gamma_{\check{i}\check{j},\check{m}\check{i}} + \gamma^{\check{m}\check{j}}_{,\check{m}}\gamma_{\check{i}\check{j},\check{i}}) = 0, \\ [\xi]: K_{\check{i}} &= -\frac{1}{2}(\gamma^{\check{j}\check{k}}\gamma_{\check{j}\check{k},\check{i}\check{i}} + \gamma^{\check{j}\check{k}}_{,\check{i}}\gamma_{\check{j}\check{k},\check{i}}) = 0, \end{aligned} \quad (24)$$

which are automatically regular without requiring any additional conditions.

The spatial curvatures are then the only ones left, with the Ricci tensor given by

$$[\xi]: {}^{(3)}R_{\check{j}\check{l}} = {}^{(3)}R_{\check{j}\check{i}\check{l}}^{\check{i}} = \sum_{\check{i}} \frac{{}^{(3)}R_{\check{i}\check{j}\check{l}}^{\check{i}}}{\gamma_{\check{i}\check{i}}^e}, \quad (25)$$

where the correspondence relation like $\iota \triangleq \check{i}$ means that the principal coordinate base $\partial_{\check{i}}$ should be along the eigenvector direction corresponding to $\gamma_{\check{i}}^e$. There is also a further contraction with the problematic $\gamma^{\check{i}\check{j}}$ to get to

$$[\xi]: {}^{(3)}R = \sum_{\check{i}\check{j}} \frac{{}^{(3)}R_{\check{i}\check{j}\check{i}\check{j}}}{\gamma_{\check{i}\check{i}}^e\gamma_{\check{j}\check{j}}^e}, \quad (26)$$

and subsequently ${}^{(3)}G_{\check{i}\check{j}}$. Because the three γ^e s generically decline at different rates, we need the Riemann tensor

components in each term of the summations in Eqs. (25) and (26) to separately decline sufficiently quickly. In fact, even if all the eigenvalues share the same rate of decline, there will still be seven regularity equations between Eqs. (25) and (26) but only six independent components in the three-dimensional Riemann tensor; thus, the equation set is overdetermining and generically only admits trivial solutions in which each variable individually “vanishes” (sinks below the error budget). In either case, we have explicitly

$$\begin{aligned} [\xi]: {}^{(3)}R_{\check{i}\check{j}\check{i}\check{i}} &= \mathcal{O}(\gamma_{\check{i}\check{i}}^e), \\ [\xi]: {}^{(3)}R_{\check{i}\check{j}\check{i}\check{j}} &= \mathcal{O}(\gamma_{\check{i}\check{i}}^e\gamma_{\check{j}\check{j}}^e), \end{aligned} \quad (27)$$

where \check{i} , \check{j} , and \check{l} all take different values. Through index symmetries, Eq. (27) accounts for all six freedoms in the spatial Riemann tensor (explicitly, the $\check{i} = 1, 2, \text{ or } 3$ possibilities for the first line and the three inequivalent pairs $(\check{i}, \check{j}) = (1, 2), (1, 3), \text{ or } (2, 3)$ for the second line).

These conditions can be further transcribed onto the second spatial derivatives of $\gamma_{\check{i}\check{j}}$. To this end, note that our principal coordinates are just rescalings of the Fermi coordinates, so the coordinate transformations between them are achieved via the Jacobian, and some simple Jacobian gymnastics allow us to import the standard Fermi result [48] to produce (note the sign difference with Ref. [47], stemming from the different conventions in the definition of the Riemann tensor)

$$\gamma_{\check{i}\check{j}} = [\text{diag}\gamma_{\check{i}}^e]_{\check{i}\check{j}}|_{\xi} - \frac{1}{3}R_{\check{i}\check{j}\check{m}}^{\check{m}}|_{\xi}x^{\check{j}}x^{\check{m}} + \mathcal{O}((x)^3), \quad (28)$$

which extends Eq. (17) off ξ (the expansion coefficients labeled with $|_{\xi}$ are to be evaluated on ξ). Applying the Gauss-Codazzi equation, Eq. (28) then implies

$$\begin{aligned} [\xi]: \gamma_{\check{i}\check{j},\check{p}\check{q}} &= \frac{2}{3}R_{\check{i}(\check{p}\check{q})\check{j}} \\ &= \frac{2}{3}({}^{(3)}R_{\check{i}(\check{p}\check{q})\check{j}} + K_{\check{i}(\check{p}}K_{\check{q})\check{j}} - K_{\check{i}\check{j}}K_{\check{p}\check{q}}). \end{aligned} \quad (29)$$

We note that there are four indices in the second derivatives of the metric, yet only three spatial dimensions to choose from, so at least one of the four indices repeats. On the other hand, if any index repeats three times or more, the four-dimensional Riemann tensor in the first line of the right-hand side of Eq. (29) vanishes due to its index antisymmetry properties. Applying these properties to the rest and using the fact that the extrinsic curvature is diagonal on ξ , we obtain that all of the components in these second derivatives that are not automatically precisely zero are

$$\begin{aligned}
[\xi]: \gamma_{\check{i}\check{j}\check{i}} &= \frac{1}{3}({}^{(3)}R_{\check{i}\check{j}\check{i}} + K_{\check{i}\check{i}}K_{\check{j}\check{j}}) = \mathcal{O}(\gamma_{i\check{i}}^e \gamma_{i\check{j}}^e), \\
[\xi]: \gamma_{\check{i}\check{i}\check{j}} &= -\frac{2}{3}({}^{(3)}R_{\check{i}\check{j}\check{i}} + K_{\check{i}\check{i}}K_{\check{j}\check{j}}) = \mathcal{O}(\gamma_{i\check{i}}^e \gamma_{i\check{j}}^e), \\
[\xi]: \gamma_{\check{i}\check{j}\check{i}\check{i}} &= \frac{1}{3}({}^{(3)}R_{\check{i}\check{j}\check{i}\check{i}}) = \mathcal{O}(\gamma_{i\check{i}}^e), \\
[\xi]: \gamma_{\check{i}\check{i}\check{j}\check{i}} &= -\frac{2}{3}({}^{(3)}R_{\check{i}\check{j}\check{i}\check{i}}) = \mathcal{O}(\gamma_{i\check{i}}^e), \tag{30}
\end{aligned}$$

where we have used Eq. (27) and \mathcal{C}_s^1 , and once again no two of \check{i} , \check{j} , and \check{l} can equal each other. The conditions in Eq. (30) can be summarized as follows: for \mathcal{C}_s^2 , spatial derivatives of the spatial metric vanish at least as quickly as the spatial metric itself as Σ is approached, in the sense of letting \check{p} be the doubly repeated index appearing in the second spatial derivative of the spatial metric and then that second derivative must belong to $\mathcal{O}(\gamma_{i\check{i}}^e)$, and when there are two doubly repeated indices (say, \check{p} and \check{q}), the derivative belongs to $\mathcal{O}(\gamma_{i\check{i}}^e \gamma_{i\check{j}}^e)$. The first spatial derivatives of $\gamma_{\check{i}\check{j}}$ already vanish according to Eq. (20), and these conditions on the second derivatives enforce a constraint on inhomogeneity in the early Universe. Importantly, \mathcal{C}_s^2 is to be satisfied along every, and not just one, temporal coordinate curve of the Gaussian normal system. Just like $\beta_{,x} = 0$ for some function $\beta(x)$ at one particular $x = x_0$ value would be a mere boundary condition that is not very constraining, having it satisfied everywhere will force β to be a constant. In our case, there is a complication that the first derivatives are made to vanish due to the choice of the principal coordinate system, which is schematically akin to going into local coordinate patches $(x', y')_q$ individually rotated to adapt to the slope of β (the x' axis is chosen to be parallel to this slope) at each point $q \in \beta$, so $\beta_{,x'}|_q = 0$ is guaranteed whatever the shape of β is (besides being sufficiently smooth to allow derivatives). Now, the vanishing of the second derivative $\beta_{,x'x'}|_q = 0$ carries the weight instead. It is a nontrivial condition that ensures the infinitesimally close neighboring local patches $(x', y')_{q \pm \delta q}$ do not need to be rotated against $(x', y')_q$ (the Jacobian is simply the identity matrix). The same argument continues on and propagates out further away from q if the vanishing of the second derivative is to be satisfied everywhere, so β is forced to be a straight line again, which can be made into a constant if a boundary condition,

$$\beta_{,\hat{x}}|_{\hat{x}_0} = 0, \tag{31}$$

is supplied at any single point \hat{x}_0 in some global coordinate system (\hat{x}, \hat{y}) .

This last step amounts to judiciously choosing the global/finite-regional coordinate system, which is necessary in our case also, since the metric is not spatially

constant under arbitrary coordinate systems even for the FLRW spacetime. In particular, the metrics as they are written under polar coordinates in Eq. (4) are spatially variable (the basis vectors for this coordinate system are not parallelly transported; thus, there are many nonvanishing spin coefficients even in a flat spacetime) and cannot be directly plugged into \mathcal{C}_s^2 that is instead stated under the more physical principal coordinates (geodetically constructed, somewhat like Cartesian coordinates in flat spacetime). With the toy example, Eq. (31) can be achieved by simply extending the local $(x', y')_q$ for an arbitrary q into a global coordinate system (\hat{x}, \hat{y}) , which in our context is mimicked by using the principal coordinates associated with an arbitrary ξ within entire finite regions surrounding that geodesic.

When homogeneity is coupled with an initial $K > 0$ (growing spatial volumes) allowed by \mathcal{C}_s^1 , we have the basic ingredients underlying Wald's theorem [49], as a concrete realization of the more general cosmological “no-hair” conjecture [50,51], which shows isotropy and spatial flatness (local resemblance to de Sitter) can possibly be achieved later through accelerated expansion, due to inflation (with any vestiges plausibly manifesting as the low multipole temperature anomalies of the cosmic microwave background [52], provided those are not simply statistical fluctuations accentuated by cosmic variance). Within the proof of Wald's theorem, homogeneity is required to maintain ${}^{(3)}R \leq 0$ but can be slightly relaxed to allow small perturbations on top of a homogeneous background [53–56]. Furthermore, if one is only interested in isotropy, then it has long been known [57,58] that anisotropy drops off rapidly with the effective spatial scale factor in a homogeneous universe, even without inflation. In this sense, isotropy may be seen as a secondary consequence of \mathcal{C}_s^2 , provided that the universe subsequently expands.

Finally, we note that the conditions $\mathcal{C}_s^{1/2}$ are to be applied in conjunction with the generic condition \mathcal{C}_g . Namely, the spatial metrics induced on Σ from the two sides match up, and the extrinsic curvature of the two sides should also suitably agree. For the latter condition, it is worth noting that, in principle, \mathcal{C}_g allows a surface layer of radiation or gravitational impulsive wave [17,59–61] to reside on the null surface Σ , permitting the extrinsic curvature to jump and the curvature tensors to become distributional at Σ [62]. However, the specific condition \mathcal{C}_s^1 removes such scenarios. In other words, the discontinuities and mild “zero-width blowup” [63] of a Dirac delta-type distribution become collateral casualties of our attempt to avoid more severe divergences.

B. Klein-Gordon equation

1. Conditions

We have regularized the left-hand side of the Einstein equations in the last section and now turn to the right-hand

side, the matter stress energy. We also need to make sure that the equation of motion for the matter itself is well behaved. As a tractable representative case (particularly relevant for those single-field inflation scenarios without other fields before reheating), we concentrate on the scalar field, which satisfies the Klein-Gordon equation

$$g^{\dot{a}\dot{b}}\varphi_{,\dot{a}\dot{b}} - g^{\dot{a}\dot{b}}\Gamma_{\dot{a}\dot{b}}^{\dot{c}}\varphi_{,\dot{c}} = \mathcal{V}'(\varphi), \quad (32)$$

where the prime denotes a derivative against φ . On ξ , the four-dimensional metric is block diagonal, so the equation becomes

$$[\xi]: -\varphi_{,\dot{i}\dot{i}} + \gamma^{\dot{i}\dot{j}}\varphi_{,\dot{i}\dot{j}} + \Gamma_{\dot{i}\dot{i}}^{\dot{c}}\varphi_{,\dot{c}} - \gamma^{\dot{i}\dot{j}}\Gamma_{\dot{i}\dot{j}}^{\dot{c}}\varphi_{,\dot{c}} = \mathcal{V}'(\varphi). \quad (33)$$

Since ξ is a geodesic always at the origin of the $x^{\dot{a}}$ coordinate system, and \dot{i} in this coordinate system is its affine parameter, we have by the geodesic equation that $\Gamma_{\dot{i}\dot{i}}^{\dot{c}} = 0$. Furthermore, from the same procedure that yielded Eq. (28), we see that, just as within the Fermi coordinates, the first spatial derivatives of $g_{\dot{a}\dot{b}}$ vanishes on ξ (but different from the Fermi case, the temporal derivatives do not vanish, since our $\gamma_{\dot{i}\dot{j}}$ is not constant along ξ), so

$$[\xi]: \Gamma_{\dot{i}\dot{j}}^{\dot{i}} = \frac{1}{2}\gamma_{\dot{i}\dot{j},\dot{i}} = -K_{\dot{i}\dot{j}}, \quad \Gamma_{\dot{i}\dot{j}}^{\dot{k}} = 0. \quad (34)$$

Therefore, the Klein-Gordon equation reduces to

$$[\xi]: \mathcal{V}'(\varphi) = -\varphi_{,\dot{i}\dot{i}} + \gamma^{\dot{i}\dot{j}}\varphi_{,\dot{i}\dot{j}} - \frac{1}{2}\gamma^{\dot{i}\dot{j}}\gamma_{\dot{i}\dot{j},\dot{i}}\varphi_{,\dot{i}} \\ = -\varphi_{,\dot{i}\dot{i}} + \sum_{\dot{i}} \frac{\varphi_{,\dot{i}\dot{i}}}{\gamma_{\dot{i}\dot{i}}^e} + K\varphi_{,\dot{i}}, \quad (35)$$

where the condition \mathcal{C}_s^1 ensures that the coefficient to the $\varphi_{,\dot{i}}$ term is regular, so the temporal derivatives of φ do not need to vanish. Only the spatial derivative needs to decline sufficiently quickly to ensure that the equation of motion admits a well-defined limit on Σ .

The other condition for φ is that the stress-energy tensor $T_{\dot{a}\dot{b}}$ that equates to $G_{\dot{a}\dot{b}}$ in Eqs. (10)–(12) should not diverge. Explicitly (including the contribution from a cosmological constant Λ),

$$[\xi]: T_{\dot{i}\dot{i}} = \frac{1}{2}\varphi_{,\dot{i}}^2 + \frac{1}{2}\gamma^{\dot{i}\dot{j}}\varphi_{,\dot{i}}\varphi_{,\dot{j}} + \mathcal{V}(\varphi) + \frac{\Lambda}{8\pi}, \\ [\xi]: T_{\dot{i}\dot{j}} = \gamma_{\dot{i}\dot{j}}\left(\frac{1}{2}\varphi_{,\dot{i}}^2 - \frac{1}{2}\gamma^{\dot{k}\dot{l}}\varphi_{,\dot{k}}\varphi_{,\dot{l}} - \mathcal{V}(\varphi) - \frac{\Lambda}{8\pi}\right) + \varphi_{,\dot{i}}\varphi_{,\dot{j}}, \\ [\xi]: T_{\dot{i}\dot{i}} = \varphi_{,\dot{i}}\varphi_{,\dot{i}}, \quad (36)$$

and the only dangerous term is

$$[\xi]: \gamma^{\dot{k}\dot{l}}\varphi_{,\dot{k}}\varphi_{,\dot{l}} = \sum_{\dot{k}} \frac{(\varphi_{,\dot{k}})^2}{\gamma_{\dot{i}\dot{i}}^e}. \quad (37)$$

Combining with our earlier discussion on the Klein-Gordon equation and noting that the $\gamma_{\dot{i}\dot{i}}^e$ can decline at different rates [or that there are still two regularity equations arising from Eqs. (35) and (37) even if they do share the same rate]—but there is only one variable φ —we obtain the following conditions: for \mathcal{C}_s^3 , the first and second spatial derivatives of φ must vanish sufficiently quickly as compared to the spatial metric, in the sense that $\varphi_{,\dot{i}} = \mathcal{O}((\gamma_{\dot{i}\dot{i}}^e)^{1/2})$ and $\varphi_{,\dot{i}\dot{i}} = \mathcal{O}(\gamma_{\dot{i}\dot{i}}^e)$. These spatial homogeneity conditions ensure that $T_{\dot{i}\dot{j}}$ and $T_{\dot{i}\dot{j}}$ vanish on Σ , but none of the terms in $T_{\dot{i}\dot{i}}$ needs to. Unfortunately then, \mathcal{C}_s^3 alone is not sufficient to ensure potential energy dominance to launch inflation if φ is the inflaton.

2. Inflationary universe

Nevertheless, additional supplementary junction conditions \mathcal{C}_{sup} can be obtained through physical considerations. Such conditions are not needed by the mathematical regularity of the various equations of motion and so are not strictly the subject of the present paper. Nevertheless, they owe their appearance to \mathcal{C}_s and are thus interesting to investigate.

It is to be noted that the energy density $T_{\dot{i}\dot{i}}$ as given by Eq. (36) is a special (scalar field) case of the matter density ρ that appears in FLRW derivations (the FLRW comoving coordinates are Gaussian normal, so ξ is automatically the worldline of a comoving observer) and will scale as $\rho \propto a^{-3(1+w)}$, when the equation of state is $P = w\rho$. So, if ρ is finite at $a = 0$, it will vanish at later times unless $w \leq -1$. Looking at the same issue in reverse, if instead $\rho \neq 0$ when $a > 0$ with a $w > -1$, the FLRW will have a diverging Hubble's parameter $\mathcal{H} = a_{,\dot{i}}/a$ on Σ , defying \mathcal{C}_s^1 , which requires it to be regular. We can see this quite readily from the Friedmann equation

$$3\mathcal{H}^2 = 8\pi\rho - \frac{{}^{(3)}R}{2} + \Lambda. \quad (38)$$

Since we have required ${}^{(3)}R$ to remain regular on Σ (for FLRW, this requires flatness $\kappa = 0$; thus, ${}^{(3)}R|_{\Sigma} = 0$) and Λ is just a constant, there is nothing to cancel with the divergence from ρ . Furthermore, adding anisotropy would unlikely be helpful, because, while it adds a shear scalar term [49]

$$\sigma^2 \equiv \frac{1}{2}\left(K_{\dot{i}\dot{j}} - \frac{1}{3}K\gamma_{\dot{i}\dot{j}}\right)\left(K_{\dot{k}\dot{l}} - \frac{1}{3}K\gamma_{\dot{k}\dot{l}}\right)\gamma^{\dot{i}\dot{k}}\gamma^{\dot{j}\dot{l}} \\ = \frac{1}{2}\text{tr}(K \cdot K) - \frac{1}{6}K^2 \quad (39)$$

into the right-hand side of Eq. (38), this term is regulated by \mathcal{C}_s^1 to be nondivergent.

In summary, \mathcal{C}_s^1 and the fact that our Universe has a nonvanishing matter energy density today, together, force a scalar inflaton field φ to be the only matter near Σ (cf., Ref. [33]), which must also behave like a perfect cosmological constant with $w = -1$ (we ignore the $w < -1$ case since there are no accepted matter models with that kind of equation of state). This makes physical sense, since traditional particles of constant finite spatial metric sizes (the standard assumption is that the sizes of particles like electrons are determined by local physics and will not scale with the cosmic size a) should not already exist all the way back at Σ , or else they will each engulf the entire spatial slice and overlap with one another thus at the very least significantly deviate from our normal intuition of how they behave. A potential energy-dominated inflaton field or cosmological constant does not need to possess any finite spatial-size features, on the other hand, and can be accommodated quite easily. They will also not dilute or concentrate and so will not produce diverging stress-energy tensors when $a = 0$. We therefore impose the following condition: for \mathcal{C}_{sup} , $\varphi_{,\dot{a}} \rightarrow 0$ sufficiently quickly, so $T_{\dot{\tau}\dot{\tau}}|_{\Sigma}$ as given by Eq. (36) is contributed only by the potential \mathcal{V} and the cosmological constant. This condition translates directly into the Cartesian coordinates for FLRW (recall $\kappa = 0$), which coincide with the principal coordinates associated with the timelike geodesic at the arbitrarily chosen spatial origin. As a consequence, the Lorentzian universe will be born directly into an inflationary period with [solving Eq. (38) for the FLRW case]

$$a(t) \approx \mathcal{B}e^{\lambda t}, \quad \lambda = \mathcal{H} = \sqrt{\frac{8\pi\mathcal{V}|_{\Sigma} + \Lambda}{3}}, \quad (40)$$

where \mathcal{B} is a constant, and the approximation is valid regardless of the shape of the potential \mathcal{V} because \mathcal{C}_{sup} sets inflation off with an instantaneous no-rolling configuration $\varphi_{,t}|_{\Sigma} = 0$. However, since [cf. Eq. (35)]

$$\varphi_{,tt} = -3\mathcal{H}\varphi_{,t} - \mathcal{V}' \quad (41)$$

does not need to vanish initially, the φ field will eventually begin to roll. A flattish \mathcal{V} could significantly prolong inflation, though, depending on the location of the flat region in relation to $\varphi|_{\Sigma}$, either by reducing initial $\varphi_{,tt}|_{\Sigma} = -\mathcal{V}'|_{\Sigma}$ to delay rolling [if $\mathcal{V}'|_{\Sigma} = 0$ precisely, the Klein-Gordon equation (41) is satisfied at all times without φ ever changing] or/and to allow the φ field to settle into a standard slow-roll regime of $\varphi_{,t} \approx -\mathcal{V}'/(3\mathcal{H})$ at a later time. Regardless, the constraints on regularity within route B, through \mathcal{C}_{sup} (as a consequence of \mathcal{C}_s^1) specifically, compel inflation to start without delay (in reverse, such an inflationary homogeneous early Universe is in compliance with all the conditions in this paper). In other words, there is not

a preinflationary radiation- or kinetic-dominated deceleration phase, the signatures of which have been searched for but indeed not found in observational data [64]. Furthermore, the conditions \mathcal{C}_s^2 , \mathcal{C}_s^3 , and \mathcal{C}_{sup} are beneficial to the inflation paradigm in another sense: that they could conceivably take us to the required initial homogeneity [58,65,66] (a more precise quantitative and nonperturbative statement of this requirement would facilitate further analysis).

IV. CONCLUSION

In this paper, we have examined what classical junction conditions would be required for a transition of our Universe into a purely timelike Riemannian regime through the big bang. So far, the restrictions they impose do not appear to raise immediate contradictions that would spoil the viability of the signature change scenario in terms of describing our physical Universe. Instead, useful constraints seem to arise. For example, the conditions \mathcal{C}_s^1 and \mathcal{C}_s^2 [particularly in the form of Eq. (27)] enforce that as we approach Σ along a timelike geodesic ξ the geodetically developed spatial slices in the principal coordinate system associated with ξ become intrinsically and extrinsically flat. Because the direction of ξ can be chosen freely (ξ is any temporal coordinate curve of any Gaussian normal system, which can be built out of arbitrary timelike congruences), this means that, via the Gauss-Codazzi equation, the projection of the four-dimensional covariant Riemann curvature tensor onto any spatial tangent plane at a point near Σ must be small (when written in sensible coordinates of which the Jacobian against the principal coordinates associated with the timelike geodesic orthogonal to that plane does not diverge). Although the spatial projection operator is rank deficient, its kernel is only one dimensional (specifically, the tangential direction to ξ ; the projection will yield zero for nonvanishing vectors only if the vector is precisely along this direction), so if the projection is vanishingly small for any arbitrary ξ , the full four-dimensional covariant Riemann tensor should be nearly zero (because any large component hidden inside the kernel of one projection operator would have been exposed by a different operator). In this sense, a strong version of the low gravitational entropy condition for the early Universe, mentioned in item 3 of Sec. IB, is realized. In particular, the inflationary FLRW discussed in Sec. III B 2, which is compatible with $\mathcal{C}_s^{1/2}$, not only has a vanishing four-dimensional Weyl curvature as FLRW metrics always do due to their symmetries, but the entire four-dimensional Riemann curvature vanishes when $a \rightarrow 0$. In contrast, this is not the case with dust- or radiation-dominated FLRWs that do not satisfy \mathcal{C}_s^1 .

Our conditions $\mathcal{C}_s^{1/2/3}$ and \mathcal{C}_{sup} are, however, not yet as strong as they can be. While they ensure the existence of one-sided limits such as ${}^{\mp}G_{ab}|_{\Sigma}$, so that the equations of

motion for metric and matter can be extended onto the big bang Σ from either side, they do not require that the \mp limits match up, which would force the two signature regimes to connect up in a smoother manner. This omission is intentional (besides trying to be conservative, given our ignorance of whether the matching is absolutely necessary) because then the one-sided conditions enumerated in this paper would admit physical interpretations independent of signature change. Namely, they need to be satisfied if the equations of motion are to be extended onto the big bang itself. Without including the big bang in the domain of validity for these equations, the Lorentzian universe will become an open set, without a suitable boundary on which to impose boundary (initial) conditions. In other words, regardless of one's view on what happens beyond the big bang, the main result of this paper can be read as necessary conditions for our Lorentzian universe to admit a Cauchy description. The utility of this paper thus does not fully diminish even if the signature change scenario is not physically realized in nature.

There are many important issues that we have not been able to tackle. In particular, unlike in Refs. [5,6], in which the genuinely classical transition into a spacelike Riemannian region occurs prior to the Planck time, it seems more difficult for us to circumvent the issue of quantum gravity because the scale factor does need to vanish in our case. The theory of quantum gravity is as yet unavailable; thus, our discussion merely aims to shed some light on the possible behavior of the classical saddle-point solutions that hopefully would dominate the full quantum path integral. Having said that, it must be noted, though, that whether such a semiclassical approach even makes sense in the gravitational context is presently subject to debate [67,68]. Furthermore, one could also note that the criteria for the onset of quantum gravity, based on dimensional analysis, are not Lorentz invariant unless one demands macroscopically separated events connected by null rays also to be treated quantum gravitationally [69], a prospect that has not been shown to be necessary. Taken to the extreme, this appears to indicate that the distances computed with the Lorentzian metric may not be the sole determining factor regarding the onset of quantum gravity, and one should perhaps be more circumspect when stating that quantum gravity must be evoked near the transition

surface, which in our case could just be another macroscopic null surface. In other words, the trans-Planckian problem [70] of inflation might not necessarily arise.

Finally, even staying at the purely classical level, the junction conditions examined in this paper are minimal, in that, while they ensure initial conditions can be imposed on the big bang, they do not tell us whether the evolution off of such compliant (with the junction conditions) initial data sets can be a well-posed initial value problem. In other words, they do not guarantee that physically interesting solutions (not plagued by wild exponentially growing perturbations, which inevitably lead to an extreme prevalence of singularities that appear to arise spontaneously) exist (the inflationary FLRW does satisfy the junction conditions, but its stability may need further scrutiny within our context). For different purposes, the required level of well-posed-ness is different. When trying to simulate the Universe on a computer, initial conditions even off of the constraint surface (i.e., do not strictly satisfy the Hamiltonian and momentum constraints) are relevant, since numerical errors are inevitable, not least because computers cannot store numbers to infinite digits (i.e., we always have truncation error). There is, of course, the possibility that our physical Universe is not amenable to being studied this way, and the well-posed-ness condition can presumably be relaxed to considerations on only a neighborhood of the constraint-satisfying initial conditions space, surrounding that of our actual Universe. Regardless, answering this well-posed-ness question demands substantial technical dexterity (as attested by the already strenuous work that went into proving the well-posed-ness of specific formulations of Einstein's equations off more familiar spacelike Cauchy surfaces) and will have to be addressed in future works.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China Grants No. 11503003 and No. 11633001, the Interdiscipline Research Funds of Beijing Normal University, and the Strategic Priority Research Program of the Chinese Academy of Sciences Grant No. XDB23000000.

-
- [1] J. B. Hartle and S. W. Hawking, *Phys. Rev. D* **28**, 2960 (1983).
 - [2] S. W. Hawking, *Nucl. Phys.* **B239**, 257 (1984).
 - [3] G. W. Gibbons and J. B. Hartle, *Phys. Rev. D* **42**, 2458 (1990).
 - [4] A. Vilenkin, *Phys. Rev. D* **27**, 2848 (1983).

- [5] G. Ellis, A. Sumeruk, D. Coule, and C. Hellaby, *Classical Quantum Gravity* **9**, 1535 (1992).
- [6] G. F. R. Ellis, *Gen. Relativ. Gravit.* **24**, 1047 (1992).
- [7] T. Dray, C. A. Manogue, and R. W. Tucker, *Gen. Relativ. Gravit.* **23**, 967 (1991).

- [8] T. Dray, C. A. Manogue, and R. W. Tucker, *Phys. Rev. D* **48**, 2587 (1993).
- [9] T. Dray, *J. Math. Phys. (N.Y.)* **37**, 5627 (1996).
- [10] S. A. Hayward, *Classical Quantum Gravity* **9**, 1851 (1992).
- [11] S. A. Hayward, [arXiv:gr-qc/9303034](https://arxiv.org/abs/gr-qc/9303034).
- [12] M. Kossowski and M. Kriele, *Classical Quantum Gravity* **10**, 2363 (1993).
- [13] M. Kossowski and M. Kriele, *Classical Quantum Gravity* **10**, 1157 (1993).
- [14] M. Kossowski and M. Kriele, *Proc. R. Soc. A* **444**, 297 (1994).
- [15] G. Darmon, *Memorial des Sciences Mathematiques* (Gauthier-Villars, Paris, 1927).
- [16] W. Israel, *Nuovo Cimento B* **44**, 1 (1966).
- [17] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, New York, 1973).
- [18] S. A. Hayward, *Classical Quantum Gravity* **11**, L87 (1994).
- [19] T. Dray, G. Ellis, and C. Hellaby, *Gen. Relativ. Gravit.* **33**, 1041 (2001).
- [20] R. Penrose, *Phys. Rev. Lett.* **14**, 57 (1965).
- [21] S. W. Hawking, Ph. D. thesis, University of Cambridge, 1966.
- [22] R. Penrose, *Ann. N.Y. Acad. Sci.* **571**, 249 (1989).
- [23] S. M. Carroll, [arXiv:1406.3057](https://arxiv.org/abs/1406.3057).
- [24] R. Arnowitt, S. Deser, and C. W. Misner, *Gen. Relativ. Gravit.* **40**, 1997 (2008).
- [25] W. Craig and S. Weinstein, *Proc. R. Soc. A* **465**, 3023 (2009).
- [26] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publication, New York, 1953), Vols. 1 and 2.
- [27] H. van Dam and Y. J. Ng, *Phys. Lett. B* **520**, 159 (2001).
- [28] M. Tegmark, *Classical Quantum Gravity* **14**, L69 (1997).
- [29] V. Perlick, *Gen. Relativ. Gravit.* **38**, 991 (2006).
- [30] R. Penrose, in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979), pp. 581–638.
- [31] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo *et al.* (Planck Collaboration), [arXiv:1807.06209](https://arxiv.org/abs/1807.06209).
- [32] A. Riazuelo, J. Weeks, J.-P. Uzan, R. Lehoucq, and J.-P. Luminet, *Phys. Rev. D* **69**, 103518 (2004).
- [33] A. Aguirre and S. Gratton, *Phys. Rev. D* **65**, 083507 (2002).
- [34] R. Penrose, *Proc. R. Soc. A* **284**, 159 (1965).
- [35] S. W. Goode and J. Wainwright, *Classical Quantum Gravity* **2**, 99 (1985).
- [36] K. Anguige and K. P. Tod, *Ann. Phys. (N.Y.)* **276**, 257 (1999).
- [37] C. Lübbe and P. Tod, *Ann. Phys. (Amsterdam)* **323**, 2905 (2008).
- [38] C. Lübbe and P. Tod, *J. Math. Phys. (N.Y.)* **50**, 112501 (2009).
- [39] P. Tod, *J. Phys. Conf. Ser.* **229**, 012013 (2010).
- [40] R. P. A. C. Newman, *Proc. R. Soc. A* **443**, 473 (1993).
- [41] K. Anguige and K. P. Tod, *Ann. Phys. (N.Y.)* **276**, 294 (1999).
- [42] K. Anguige, *Ann. Phys. (N.Y.)* **282**, 395 (2000).
- [43] K. P. Tod, *Classical Quantum Gravity* **20**, 521 (2003).
- [44] P. Tod, *Classical Quantum Gravity* **24**, 2415 (2007).
- [45] J. Carminati and R. McLenaghan, *J. Math. Phys. (N.Y.)* **32**, 3135 (1991).
- [46] E. Fermi, *Rend. Fis. Acc. Lincei* **31**, 21 (1922) (in Italian).
- [47] F. K. Manasse and C. W. Misner, *J. Math. Phys. (N.Y.)* **4**, 735 (1963).
- [48] W.-T. Ni and M. Zimmermann, *Phys. Rev. D* **17**, 1473 (1978).
- [49] R. M. Wald, *Phys. Rev. D* **28**, 2118 (1983).
- [50] G. W. Gibbons and S. W. Hawking, *Phys. Rev. D* **15**, 2738 (1977).
- [51] S. W. Hawking and I. L. Moss, *Phys. Lett. B* **110**, 35 (1982).
- [52] Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak *et al.* (Planck Collaboration), [arXiv:1906.02552](https://arxiv.org/abs/1906.02552).
- [53] A. A. Starobinskiĭ, *Sov. J. Exp. Theor. Phys. Lett.* **37**, 66 (1983).
- [54] W. Boucher and G. W. Gibbons, in *Very Early Universe*, edited by G. W. Gibbons, S. W. Hawking, and S. T. C. Siklos (Cambridge University Press, Cambridge, UK, 1983), pp. 273–278.
- [55] J. D. Barrow, in *Very Early Universe*, edited by G. W. Gibbons, S. W. Hawking, and S. T. C. Siklos (Cambridge University Press, Cambridge, UK, 1983), pp. 267–271.
- [56] T. Piran, in *NATO Advanced Science Institutes (ASI) Series C*, edited by W. G. Unruh and G. W. Semenoff, NATO Advanced Science Institutes (ASI) Series C (D. Reidel, Dordrecht, Netherlands, 1983), Vol. 219, pp. 261–282.
- [57] C. W. Misner, *Astrophys. J.* **151**, 431 (1968).
- [58] D. S. Goldwirth and T. Piran, *Phys. Rep.* **214**, 223 (1992).
- [59] Y. Choquet-Bruhat, *Ann. Inst. Henri Poincaré, A* **8**, 327 (1968).
- [60] R. Penrose and M. A. H. MacCallum, *Phys. Rep.* **6**, 241 (1973).
- [61] R. Penrose, in *General Relativity: Papers in Honour of J.L. Synge*, edited by L. O’Raifeartaigh (Clarendon Press, Oxford, UK, 1972), pp. 101–115.
- [62] A. H. Taub, *J. Math. Phys. (N.Y.)* **21**, 1423 (1980).
- [63] J. Mikusiński, *Fundam. Math.* **35**, 235 (1948).
- [64] Y. Akrami *et al.* (Planck Collaboration), [arXiv:1807.06211](https://arxiv.org/abs/1807.06211).
- [65] T. Vachaspati and M. Trodden, *Phys. Rev. D* **61**, 023502 (1999).
- [66] A. Aguirre and S. Gratton, *Phys. Rev. D* **67**, 083515 (2003).
- [67] J. Feldbrugge, J.-L. Lehners, and N. Turok, *Phys. Rev. Lett.* **119**, 171301 (2017).
- [68] J. D. Dorronsoro, J. J. Halliwell, J. B. Hartle, T. Hertog, and O. Janssen, *Phys. Rev. D* **96**, 043505 (2017).
- [69] S. Hollands and R. M. Wald, *Gen. Relativ. Gravit.* **34**, 2043 (2002).
- [70] J. Martin and R. H. Brandenberger, *Phys. Rev. D* **63**, 123501 (2001).