NOTES ON "STABLE BIG BANG FORMATION FOR EINSTEIN'S EQUATIONS: THE COMPLETE SUBCRITICAL REGIME"

ONYX GAUTAM

These are notes on [FRS23] for the Cambridge–Princeton GR seminar talks given on November 21, 2025 and December 5, 2025.

1. INTRODUCTION

This paper is about the stability of spacelike singularities in general relativity. These singularities typically occur in two settings: in the interior of black holes, and in the cosmological setting, which is studied in this paper. By cosmological, we mean that the spatial topology is compact (in our case it is the torus).

The (generalized) Kasner solutions are explicit cosmological solutions to the Einstein vacuum equations (or the Einstein–scalar field system) with a spacelike singularity. The initial data at $\{t = 1\}$ is smooth, but a Big Bang singularity develops as $t \searrow 0$. That is, curvature blows up monotonically in time along an entire spacelike hypersurface.

The Kasner family is parameterized by a set of real-valued exponents that satisfy two algebraic constraints. Heuristics in the physics literature suggest that Kasner and its big band formation are dynamically stable if these exponents satisfy a certain "sub-criticality" condition (an open algebraic condition). This paper verifies this claim.

The setting is the Einstein–scalar field system in $\mathfrak{D} \geq 3$ spatial dimensions, or the vacuum equations in $\mathfrak{D} \geq 10$ spatial dimensions (the minimal number of dimensions for which there exist subcritical Kasner exponents in vacuum), or the vacuum equations in $\mathfrak{D} = 3$ spatial dimensions with the additional assumption of polarized-$U(1)$ symmetry.

1.1. **The Einstein–scalar field system.** The Einstein–scalar field system for a spacetime and scalar field $(\mathcal{M}^{1+\mathfrak{D}}, \overline{g}, \psi)$ takes the form

$$\begin{cases} \mathrm{Ric}_{\mu\nu}[\overline{g}] = \partial_\mu \psi \partial_\nu \psi \\ \Box_{\overline{g}} \psi = \overline{g}^{\alpha\beta} \overline{\nabla}_\alpha \overline{\nabla}_\beta \psi = 0. \end{cases} \tag{1}$$

(*We write $\overline{g}$ for a spacetime metric. Later we will use g for the induced metric on time slices.*) When $\psi \equiv 0$, this reduces to the Einstein vacuum equations. In our setting, $\mathcal{M}^{1+\mathfrak{D}} = I \times \mathbf{T}^{\mathfrak{D}}$, where the spatial topology is compact (a torus) and I plays the role of time.

1.1.1. *The initial value formulation.* A sufficiently regular initial data set $(\Sigma, g, k, \psi_0, \psi_1)$ of a Riemannian manifold (Σ, g) equipped with a symmetric $(0,2)$-tensor k and scalar fields ψ_0 and ψ_1 induces a maximal development $(\mathcal{M}, \overline{g}, \psi)$ (unique up to diffeomorphism) solving (1) (in which Σ is a Cauchy hypersurface with induced metric g and second fundamental form k and $(\psi|_\Sigma, \nu_\Sigma \psi|_\Sigma) = (\psi_0, \psi_1)$ for ν_Σ the unit normal to Σ), so long as they satisfy the constraint equations on Σ (where all operations are with respect to the metric g):

$$\begin{cases} R - |k|^2 + (\mathrm{tr}\,k)^2 = \psi_1^2 + |\nabla \psi_0|^2 & \text{(Hamiltonian constraint)}, \\ \mathrm{div}\,k - \nabla \mathrm{tr}\,k = -\psi_1 \nabla \psi_0 & \text{(momentum constraint)}. \end{cases} \tag{2}$$

1.1.2. *The ADM formalism.* A globally hyperbolic spacetime is foliated by Cauchy hypersurfaces. If we parametrize these Cauchy surfaces by a scalar function t (which we interpret as a time function) and equip each level set Σ_t of the time function with spatial coordinates x^i, then the metric takes the form

$$\overline{g}(t, x) = -n^2(t, x)\, \mathrm{d}t^2 + g_{ab}(t, x)\, \mathrm{d}x^a\, \mathrm{d}x^b, \tag{3}$$

where n is the *lapse function* (which measures the separation in proper time between the coordinate time slices) and g is the induced Riemannian metric on the time slice Σ_t.

Remark 1.1 (Presence of a shift vector field). In a general gauge, there will also be a shift vector field describing how the spatial coordinates change between hypersurfaces of constant time. We will end up choosing the coordinates so that this vector field vanishes, i.e. so that $(\nabla t)x^i = 0$.

The Einstein equations then become a first-order system of evolution equations for the dynamical variables g and k (the second fundamental form of Σ_t) coupled to constraint equations on Σ_t. We can compute

$$\begin{cases} \partial_t g_{ij} = -2n k_{ij}, \\ \partial_t k_{ij} = -\nabla_i \nabla_j n + n(-\overline{\mathrm{Ric}}_{ij} + \mathrm{Ric}_{ij} + \mathrm{tr}_g k\, k_{ij} - 2k_i{}^a k_{aj}), \end{cases} \tag{4}$$

and the following constraints must hold on Σ_t:

$$\begin{cases} R - |k|^2 + (\mathrm{tr}\, k)^2 = 2\overline{\mathrm{Ric}}_{00} + \overline{R} & \text{(Hamiltonian constraint)}, \\ \nabla^a k_{ai} - \nabla_i \mathrm{tr}\, k = \overline{\mathrm{Ric}}_{0i} & \text{(momentum constraint)}. \end{cases} \tag{5}$$

Remark 1.2 (Equation for the lapse). This is not a closed system; Einstein's equations possess a *gauge freedom*, and fixing the gauge will fix an equation for the lapse. In our setting we will fix a CMC gauge, in which the time slices have constant mean curvature, which will lead to an elliptic equation for the lapse.

Remark 1.3 (Derivation of the constraint equations). The Hamiltonian constraint is derived using the Gauss equation (which concerns the tangential component of Riemann curvature), while the momentum constraint is derived using the Codazzi equation (which concerns the normal component of Riemann curvature). (*The momentum constraint will be especially important for deriving energy estimates.*)

Remark 1.4 (Names of the constraint equations). Recall that the component $T^{\alpha\beta}$ of the energy-momentum tensor (in contravariant form) measures the flux of the α-component of the momentum through a hypersurface of constant x^β. Then we interpret T^{00} as the energy density and T^{0i} as (the i-component of) the momentum density. When Einstein's equations $\mathrm{Ric}_{\alpha\beta} - \frac{1}{2}R g_{\alpha\beta} = T_{\alpha\beta}$ are imposed, the left sides of the Hamiltonian and momentum constraints, respectively, are exactly T_{00} and T_{0i}.

1.2. Hawking's incompleteness theorem.

What happens to solutions of this equation in evolution? A general theorem predicts incompleteness for cosmological solutions.

Theorem 1.5 (Hawking incompleteness theorem, [Wal84, Thm. 9.5.1]). *Consider a C^2 globally hyperbolic spacetime with* compact *Cauchy surface Σ. Suppose that*

- *the strong energy condition holds ($\mathrm{Ric}(X, X) \geq 0$ for all causal vector fields X),*
- *and $\mathrm{tr}\, k|_\Sigma < 0$.*

Then every past-directed timelike geodesic is incomplete.

What is the nature of this incompleteness? Does it come from a singularity?

Remark 1.6 (Strong cosmic censorship conjecture). The proof of Hawking's theorem, like that of the Penrose incompleteness theorem, is by contradiction, so it does not provide any information on the nature of the incompleteness. Note that geodesic incompleteness does not imply inextendibility due to the presence of a singularity (as in Kerr, or Taub-NUT in the cosmological setting)! The strong cosmic censorship conjecture predicts that the incompleteness is due to singularity formation.

1.3. Kasner metrics.

Are there examples of spacetimes satisfying the hypotheses of Hawking's theorem?

Definition 1.7 (Kasner metrics). The *Kasner metric* on $(0, \infty) \times \mathbf{T}^{\mathfrak{D}}$ with exponents $\tilde{q}_i \in \mathbf{R}$ $(1 \leq i \leq \mathfrak{D})$ is

$$\overline{g} = -\mathrm{d}t^2 + \sum_{i=1}^{\mathfrak{D}} t^{2\tilde{q}_i}(\mathrm{d}x^i)^2. \tag{6}$$

It satisfies the Einstein vacuum equations when the following algebraic conditions hold:

$$\sum_{i=1}^{\mathfrak{D}} \tilde{q}_i = 1, \qquad \sum_{i=1}^{\mathfrak{D}} \tilde{q}_i^2 = 1. \tag{7}$$

Remark 1.8 (Kasner exponents live on a sphere). The Kasner exponents live on an $S^{\mathfrak{D}-2}$, being the intersection of a plane with an $S^{\mathfrak{D}-1}$. When $\mathfrak{D} = 3$, this is a circle.

Remark 1.9 (Source of the Kasner constraints). The condition on the squares stems from the Hamiltonian constraint.

Remark 1.10 (Properties of Kasner spacetimes). The Kasner solutions are spatially homogeneous (flat in space). When one of the exponents is 1, the solution is flat (isometric to Minkowski in Rindler coordinates), and we call it trivial. The condition on the squares then implies that $|\tilde{q}_i| < 1$, and so the first condition implies that some $\tilde{q}_i$ is negative. In particular, not all directions are collapsing; there is at least one direction that is expanding. In particular, the expansion/contraction of spacetime is *anisotropic*, and observers experience *spaghettification* as they approach the singularity.

Remark 1.11 (Kasner and Schwarzschild). When Schwarzschild is written in the above gauge, the metric in the black hole interior converges to a Kasner spacetime with exponents $(-1/3, 2/3, 2/3)$.

Remark 1.12 (Kasner satisfies the hypotheses of Hawking's singularity theorem). Since the Kasner exponents add up to 1, we have

$$\mathrm{Vol}(\Sigma_t) = t. \tag{8}$$

By the first variation of area formula, we conclude that the mean curvature of Σ_t is negative. That is, $\mathrm{tr}\, k < 0$, and so Hawking's theorem holds on Kasner. In fact, one can compute

$$k_{ij} = -\delta_{ij} \cdot \tilde{q}_i \cdot \frac{1}{t} \implies \mathrm{tr}\, k = -\frac{1}{t} \tag{9}$$

Since the negativity of mean curvature is an open condition, the theorem also applies to perturbations of Kasner.

Remark 1.13 (Kasner exhibits Big Bang formation). The Kretschmann scalar $\mathrm{Riem}^{\alpha\beta\gamma\delta}\mathrm{Riem}_{\alpha\beta\gamma\delta}$ of a nontrivial Kasner solution is a multiple of t^{-4} (where the constant depends on the Kasner exponents). In particular, there is a spacelike singularity occurring at each point on $\mathbf{T}^{\mathfrak{D}}$.

Remark 1.14 (Kasner satisfies the strong cosmic censorship conjecture). The monotonic curvature blowup shows that Kasner is C^2-inextendible. In fact, Kasner is C^0-inextendible [Mie24] by an argument adapting Sbierski's proof of the C^0-inextendibility of Schwarzschild [Sbi18].

Remark 1.15 (Spacelike singularity formation in Kasner as compared to in black hole interiors). The spacelike singularity formation in Kasner is unlike the case of the black hole interior, where, even outside of symmetry, the boundary of the maximal globally hyperbolic development contains a null piece (which is expected to be singular) [DL25]. Work on the interior of spherically symmetric black holes by Dafermos [Daf05; Daf14] shows that there is a null piece of the singularity, work by Li [Li25] (see also [AZ20; AG23]) shows that spacelike singularities in the black hole exhibit Kasner asymptotics, and work by Van de Moortel [Moo22; Moo25b; Moo25a] shows that spacelike and null singularities coexist (and Van de Moortel analyzes the Kasner asymptotics towards the spacelike singularity and the junction between spacelike and null singularity).

What about the case of scalar field matter?

Definition 1.16 (Generalized Kasner metrics). Let $\bar{g}$ be a Kasner metric. If

$$\sum_{i=1}^{\mathfrak{D}} \tilde{q}_i = 1, \qquad \sum_{i=1}^{\mathfrak{D}} \tilde{q}_i^2 = 1 - \tilde{B}^2 \tag{10}$$

for some $\tilde{B} \in \mathbf{R}$, then $\bar{g}$ together with $\tilde{\psi} := \tilde{B} \log t$ solve the Einstein–scalar field system on $(0, \infty) \times \mathbf{T}^{\mathfrak{D}}$.

Remark 1.17. The extra flexibility afforded by $\tilde{B}$ allows one to choose all $\tilde{q}_i > 0$.

1.4. **Rough statement of the main theorem.** What happens to perturbations of Kasner? Is the Big Bang formation in Kasner special, or is it generic?

The heuristic due to [DHS85] is that Kasner should be stable if the exponents verify the following (open) algebraic condition:

$$\tilde{q}_I + \tilde{q}_J - \tilde{q}_K < 1 \text{ for all } I \neq J. \tag{11}$$

Remark 1.18 (Three distinct indices). Note that the condition is trivially satisfied if $B = I$ or $B = J$, so the relevant case is when the three indices are distinct.

Remark 1.19 (When can (11) be satisfied?). In vacuum, the condition (11) can only be satisfied when $\mathfrak{D} \geq 10$. With scalar matter, (11) can hold when $\mathfrak{D} \geq 3$.

Theorem 1.20 (Main theorem of [FRS23], rough version)**.** *All Kasner spacetimes for which (11) holds are stable. Namely, the solution arising from perturbed data at $\{t = 1\}$ exhibits Big Bang formation, in the sense that it*

- *exists on $t \in (0, 1]$, where t is a CMC-normalized time function (which satisfies $\operatorname{tr} k = -1/t$),*
- *and satisfies $\operatorname{Riem} \cdot \operatorname{Riem} \sim t^{-4}$ (so there is C^2-inextendibility).*

Moreover, the singularity is of Kasner-type and is close to the original Kasner solution, in the sense that

- *$t\partial_t \psi \to B^\infty(x)$ as $t \searrow 0$ for some $B^\infty(x)$ close to B,*
- *there exists a frame in which $tk_{IJ} \to k_{IJ}^\infty(x)$ as $t \searrow 0$,*
- *the eigenvalues $q_I^\infty(x)$ of $-k^\infty$ are close to $\tilde{q}_I$,*
- *and the following Kasner constraints hold*

$$\sum_{I=1}^{\mathfrak{D}} q_I^\infty(x) = 1, \qquad \sum_{I=1}^{\mathfrak{D}} q_I^\infty(x)^2 = 1 - B^\infty(x)^2. \tag{12}$$

Remark 1.21 (Gauge choice and synchronization of the singularity). One does not expect the singularity to occur at the same proper time from the data at different points in space. The CMC gauge used in theorem 1.20 *synchronizes* the singularity in coordinate time. Morally, this is possible because the CMC condition leads to an elliptic equation for the lapse, which has infinite speed of propagation.

Remark 1.22 (Polarized $U(1)$-symmetry in $(1 + 3)$ dimensions). When $\mathfrak{D} = 3$ (where (11) never holds), the conclusions of theorem 1.20 hold for *all* Kasner solutions when the perturbations have polarized $U(1)$-symmetry. Polarized $U(1)$-symmetric data[1] can be equipped with coordinates x^i in which g and k are independent of x^3 (so ∂_{x^3} is Killing), and g and k are block-diagonal, with:

$$g_{13} = g_{23} = k_{13} = k_{23} = 0. \tag{13}$$

1.5. **Related works.**

Remark 1.23 (Previous works on Big Bang formation outside of symmetry). There are previous works outside of symmetry that concern various subregimes of the complete subcritical regime considered in theorem 1.20.

- Early work of Andersson–Rendall [AR01] considers the Einstein–scalar field system in the analytic setting where all $\tilde{q}_i > 0$ (with general topology when $\mathfrak{D} = 3$).

There are works of Rodnianski–Speck [RS18b; RS18a; RS21] in:

- the *nearly isotropic* setting with scalar field (or stiff fluid)[2] matter, namely the stability of FLRW (with $\mathbf{T}^3$ topology) when $\mathfrak{D} = 3$ (where $\tilde{q}_i = 1/3$ and $\tilde{B} = 2/3$), later generalized to the case of FLRW with $\mathbf{S}^3$ topology by Speck [Spe18],[3]
- and the *moderately anisotropic* setting in vacuum in high space dimensions (where $|\tilde{q}_i| < 1/6$ and $\mathfrak{D} \geq 38$).

Remark 1.24 (Later works on stable Big Bang formation outside of symmetry). There are a number of relevant works after [FRS23]:

- (Oude Groeniger–Petersen–Ringström [GPR25]) Big Bang formation under a largeness assumption on the initial mean curvature and quantitative distinctness of the initial eigenvalues of $(\operatorname{tr}_g k)^{-1} \cdot k^\sharp$ (the expansion-normalized Weingarten map). In particular, this is not a result about stability of specific solution, and the data can be far from spatially homogeneous.
- (Beyer–Oliynyk–Zheng [BOZ25]) localized version of theorem 1.20 for Einstein–scalar field with $\mathfrak{D} = 3$, using a wave gauge (as opposed to the non-local CMC gauge), following [BO23] near FLRW; interestingly, they use the scalar field as a time function (as opposed to the inverse of the mean curvature) to synchronize the singularity

---

[1]A more geometric formulation of polarized $U(1)$-symmetry is that there exists a non-vanishing hypersurface-orthogonal spacelike Killing vector field X with $\mathcal{L}_X g = \mathcal{L}_X k = 0$. Polarization means that $k(X, Y) = 0$ whenever $g(X, Y) = 0$.

[2]Scalar field matter is a special case of Einstein–Euler with a stiff fluid with vanishing vorticity. A fluid is called *stiff* if the equation of state is $p = \rho$, in which case the speed of sound equals the speed of light.

[3]In this setting spacetime also collapses in the future; this scenario is termed the Big Crunch.

- (Fajman–Urban [FU25]) stability of FLRW for Einstein–scalar field–Vlasov

Remark 1.25 (Previous works on stable Big Bang formation within symmetry). The most definitive result is due to Ringström [Rin09], who established strong cosmic censorship in vacuum with $\mathfrak{D} = 3$ under Gowdy symmetry[4] by showing that the solution exhibits Kasner asymptotics towards the past, except possibly at a finite number of *spikes*, which are roughly regions where terms with spatial derivatives are not negligible when compared to time derivative terms (failure of AVTD behaviour). Solutions with Gowdy symmetry exhibiting spikes were previously constructed by Rendall and Weaver [RW01].

1.6. **Heuristics associated to Kasner-type metrics.** Why is it reasonable to expect Kasner-type behaviour at the singularity when the subcriticality condition (11) holds? Consider the following class of "Kasner-type" metrics:

$$\overline{g}(t,x) = -\mathrm{d}t^2 + g(t,x), \qquad g(t,x) \approx_{t\searrow 0} \sum_{I=1}^{\mathfrak{D}} t^{2q_I(x)} \theta^I(x) \otimes \theta^I(x) \tag{14}$$

$$\theta^I(x) = \theta^I_a(x)\,\mathrm{d}x^a, \qquad \psi = B(x)\log t, \tag{15}$$

where the following constraints are satisfied:

$$\sum_{I=1}^{\mathfrak{D}} q_I(x) = 1, \qquad \sum_{I=1}^{\mathfrak{D}} q_I(x)^2 = 1 - B(x)^2. \tag{16}$$

Remark 1.26 (Kasner metrics are of Kasner-type). The exact Kasner spacetimes have $\theta^I(x) = \mathrm{d}x^I$ and $q_I^\infty(x) = \tilde{q}_I$ and $B^\infty(x) = \tilde{B}$.

Remark 1.27 (Solutions to Einstein's equations). Metrics of the form (14) generally do not solve Einstein's equations, though they might approximate solutions. Indeed, (14) is a valid leading-order term in a formal power series solution to Einstein's equations (given a version of the momentum constraint and the integrability condition $\theta^I \wedge \mathrm{d}\theta^I = 0$ whenever $p_I < 0$; see remark 1.33).

Remark 1.28 (Character of the singularity). The ansatz (14) *assumes* that the singularity is spacelike and synchronized at $\{t = 0\}$. In a dynamical problem starting from data at $\{t = 1\}$, one must *locate* the singularity!

Remark 1.29 (The statements of this paper). This paper does *not* prove that perturbations of Kasner have the asymptotics of (14); the argument is closed using much weaker information.

Consider the type-$(1,1)$ second fundamental form with respect to the spatial coframe θ^I of Kasner-like directions and the corresponding basis-dual spatial frame e_I:

$$k^I{}_J = \theta^I_a e^b_J k^a{}_b \tag{17}$$

The form of g from (14) gives

$$k^I{}_J = -\delta^I_J \frac{q_I(x)}{t} \underset{(16)}{\Longrightarrow} \operatorname{tr} k = -t^{-1}. \tag{18}$$

On the other hand, we have an evolution equation

$$\partial_t k^I{}_J - \operatorname{tr} k k^I{}_J = -\overline{\operatorname{Ric}}^I{}_J + \operatorname{Ric}^I{}_J. \tag{19}$$

Remark 1.30 (Origin of (19)). This equation is obtained by setting the lapse to 1 in (4) and multiplying by g^{ia} (i.e. raising the index) and using the equation $\partial_t g^{ij} = \partial_t(g^{-1})_{ij} = -(g^{-1} \cdot \partial_t g \cdot g^{-1})_{ij}$ (which holds for any matrix g). This gives an equation in local coordinates; to obtain the equation with frame indices, simply contract with the frame e_I and coframe θ^I (which are time-independent).

We could *deduce* the behaviour $k_{IJ} \sim t^{-1}$ directly from the evolution equation (19) if we knew that for some $\sigma > 0$ we had

(i) $\operatorname{tr} k = -t^{-1} + O(t^{-1+\sigma})$,

---

[4]Gowdy spacetimes have compact spatial topology and admit a $U(1) \times U(1)$ isometry group acting effectively with spacelike orbits, so that there are two commuting linearly independent spacelike Killing vector fields. If X and Y are the corresponding one-forms, then the Gowdy condition is that $X \wedge Y \wedge \mathrm{d}X = X \wedge Y \wedge \mathrm{d}Y = 0$. When in addition the Killing vector fields be taken to be orthogonal, the spacetime is called *polarized*.

(ii) (RHS of (19)) $= O(t^{-2+\sigma})$.

This is because we could integrate the equation

$$\partial_t k^I{}_J + \frac{1}{t} k^I{}_J = O(t^{-2+\sigma}) \implies \partial_t(tk^I{}_J) = O(t^{-1+\sigma}) \tag{20}$$

to the data at $\{t = 1\}$, the point being that $t^{-1+\sigma}$ is integrable as $t \searrow 0$.

How do we obtain items (i) and (ii)? In the proof, (i) is obtained by using a CMC gauge that *imposes* $\operatorname{tr} k = -t^{-1}$.

Remark 1.31 (A technical remark on (i) and the CMC gauge). The CMC gauge is technically not compatible with the ansatz (14), since the lapse in that gauge is not identically 1. However, the proof will show that the lapse decays to 1 as $t \searrow 0$, so this difference is negligible.

The first part of (ii) is $|\operatorname{Ric}[\overline{g}]| \lesssim t^{-2+\sigma}$, which is trivially satisfied when we impose the Einstein vacuum equations; in the presence of matter, we would have to show this, but this only involves the scalar field (which is easier to estimate). Thus the key estimate is

$$|\operatorname{Ric}[g]| \lesssim t^{-2+\sigma}. \tag{21}$$

(*Note that* $\operatorname{Ric}[g] = 0$ *on exact Kasner, since the spatial geometry is flat.*) From (14), one computes that (in the absence of special algebraic structure in the one-forms θ^I)

$$|\operatorname{Ric}[g]|(x) \sim \max_{\substack{1 \leq I,J,B \leq \mathfrak{D} \\ I < J}} t^{-2(q_I(x)+q_J(x)-q_B(x))} \tag{22}$$

Thus (21) holds for Kasner-type metrics when the subcriticality condition (11) holds. The terminology "subcritical" refers to the fact that certain terms are subcritical/lower-order with respect to growth in t as $t \searrow 0$.

Remark 1.32 (AVTD behaviour in (21)). The bound (21) is a manifestation of *AVTD behaviour*, that spatial derivative terms are negligible when compared to time derivative terms.

Remark 1.33 (Special algebraic structure can recover (21)). In $(1+3)$ dimensions in vacuum (where the subcriticality condition (11) never holds), there is a unique negative Kasner exponent q_-. Let θ^- be the corresponding one-form. One can show that (21) holds if

$$\theta^- \wedge \mathrm{d}\theta^- = 0 \iff \theta^- = u\,\mathrm{d}v \text{ locally for some } u,v : \mathbf{T}^3 \to \mathbf{R} \iff \ker\theta^- \text{ is integrable}, \tag{23}$$

where the equivalences hold by the Frobenius theorem. In particular, this vanishing condition kills the coefficient of the worst power of t in the expression for $\operatorname{Ric}^I{}_J$.

Note that polarized $U(1)$-symmetric Kasner-type metrics satisfy (23).

1.7. **The picture outside of the subcritical regime: the BKL conjecture.** Recall that (11) is never satisfied in vacuum in $(3+1)$ dimensions. Moreover, by a function counting argument, [KL63] suggest that (23) generically fails. What happens then? The generic picture is governed by the BKL conjecture, loosely formulated in [KL63; BKL71; BKL82] (and supported numerically in [Ber+98]). Generic cosmological singularities are spacelike, local, and oscillatory. That is:

- There exists a large class of solutions outside of symmetry with Big Bang formation along a spacelike hypersurface.
- Distinct causal curves lose their ability to communicate in the direction of the singularity.
- Unlike the Kasner-type solutions, these spacetimes are highly oscillatory in time as $t \searrow 0$ (that is, q_i and θ^I can depend also on t). This prediction comes from an analysis of spatially homogeneous (Bianchi VIII and IX) spacetimes.[5] This is sometimes called the "Mixmaster scenario," after Misner [Mis94]. The chaotic dynamics has been modelled in the physics literature by billiard motion on a region in the hyperbolic space [DHN03].

The "locality" is related to the following idea:

---

[5]A manifold is called *homogeneous* if it admits a group G of isometries that act transitively. Bianchi IX symmetry is the case $G = \mathrm{SU}(2)$ (with Lie algebra $\mathfrak{so}(3)$). Bianchi VIII symmetry is the case when G has Lie algebra $\mathfrak{sl}(2)$.

- Solutions become *asymptotically velocity term dominated (AVTD)* as $t \searrow 0$. That is, spatial derivative terms are less singular than time derivative terms, so they can be discarded. For each spatial point, the evolution of the geometry therefore becomes an ODE in time.
- As a result, the dynamics in future light cones emanating from distinct points on the singular hypersurface decouple.
- In such a cone, space is approximately homogeneous; in this setting, Einstein's equations become an autonomous finite-dimensional ODE system.
- Fixed points of this system are exact Kasner solutions, which are unstable. The unstable manifolds correspond to heteroclinic orbits (connecting two fixed points), called *BKL bounces*.
- BKL suggest that the Kasner-type ansatz is valid for a certain time period (a so-called Kasner epoch), but then the metric rapidly transitions to a different set of Kasner exponents, and they provide a formula relating the two sets of exponents.
- In later work, BKL [BK73] suggest that "matter doesn't matter," i.e. most matter models (perfect fluids, electromagnetism, Vlasov, etc.) do not change the chaotic dynamics. However, certain "stiff" models, in particular scalar field in $1 + \mathfrak{D}$ dimensions with $\mathfrak{D} \geq 3$ should silence these oscillations, as should vacuum when $\mathfrak{D} \geq 10$ [DHS85].

Remark 1.34 (Rigorous result in the oscillatory regime in the spatially homogeneous setting). The BKL heuristics are based on analysis in the spatially homogeneous setting, where Einstein's equations become a finite-dimensional ODE system. A rigorous result due to Ringström [Rin01] shows that in vacuum under Bianchi IX symmetry, solutions generically exhibit oscillatory behaviour towards a spacelike singularity (in particular, the expansion-normalized eigenvalues of the second fundamental form do not converge).

Remark 1.35 (Rigorous result in the oscillatory regime outside of spatial homogeneity). In view of the existence of spikes, the relevance of the BKL heuristics to the spatially inhomogeneous setting remains unclear. However, there is the work of Li [Li24], which constructs a solution in Gowdy symmetry with one BKL bounce in the presence of spikes.

Remark 1.36 (Existence of non-generic Kasner-like solutions in $(3 + 1)$ vacuum). Despite the conjectured instability of Kasner solutions in vacuum in $(3+1)$ dimensions outside of symmetry, Fournodavlos–Luk [FL23] construct Sobolev-class solutions with a Kasner-type singularity. These solutions satisfy the polarization condition (23), so they have three functional degrees of freedom, as opposed to the full four. There is also previous work of Klinger [Kli15], who constructed analytic Kasner-type solutions. These metrics are constructed by solving a singular initial value problem (of Fuschian-type) with "data prescribed at $\{t = 0\}$" to generate an approximate solution, and then use energy estimates to construct a nearby exact solution.

2. Outline of the proof

2.1. The constant-mean-curvature (CMC) gauge.
We use a time function $t \in (0, 1]$ with level sets Σ_t. The spatial coordinates x^i on the data are transported by $(\nabla t) x^i = 0$. In this gauge, the metric is

$$\overline{g} = -n^2 \, \mathrm{d}t^2 + g_{ab} \, \mathrm{d}x^a \, \mathrm{d}x^b, \tag{24}$$

where n is the lapse and g is the induced metric on Σ_t. We impose that the time function is CMC (constant mean curvature)-normalized, in the sense that $\mathrm{tr}\, k = -1/t$. Then the lapse solves an elliptic equation.

2.2. The orthonormal frame.
Write $e_0 = n^{-1} \partial_t$. We get a $\overline{g}$-orthonormal spacetime frame on Σ_1 by supplementing e_0 with a g-orthonormal spatial frame on Σ_1 constructed by applying the Gram–Schmidt algorithm to the coordinate vector fields $\partial_{x^1}, \ldots, \partial_{x^\mathfrak{D}}$. We then propagate this frame via Fermi–Walker transport from Σ_1 along integral curves of e_0:

$$\overline{\nabla}_{e_0} e_I = (e_I \log n) e_0. \tag{25}$$

Lemma 2.1. *The frame* $\{e_\alpha\}_{\alpha=0}^{\mathfrak{D}}$ *is orthonormal. In particular, the metric takes the form*

$$\overline{g}^{\alpha\beta} = m^{\alpha\beta} e_\alpha \otimes e_\beta. \tag{26}$$

Proof. Use the Fermi–Walker transport equation together with the calculation

$$\overline{\nabla}_{e_0} e_0 = n^{-1}(e_C n) e_C \tag{27}$$

obtained from the form of the metric $\overline{g}$. $\square$

Remark 2.2 (Parallel transport is a special case of Fermi–Walker transport). Fermi–Walker transport along a geodesic reduces to parallel transport. In general, we are trying to define a non-rotating spatial frame for an observer that may have acceleration (i.e. the curve may not be a geodesic) such that supplementing this spatial frame with the unit velocity vector e_0 gives a spacetime orthonormal frame.

Remark 2.3 (Physically realizing Fermi–Walker transport). A physical way to realize Fermi–Walker transport of a spatial frame is to "attach" the vectors to a gyroscope, which precesses due to the acceleration of the worldline of the observer. By the previous remark, a free-falling observer transports their basis vectors.

Remark 2.4 (The frame is not adapted to Kasner directions). The frame is *not* adapted to any particular one-forms, since we do not yet know that there is a Kasner-type singularity.

2.3. **The unknowns.** We formulate the equations in terms of:

- the lapse n,
- the coefficients e_I^i of the spatial orthonormal frame with respect to the spatial coordinates,
- the frame components $k_{IJ} := g(\overline{\nabla}_{e_I} e_J, e_0) = k_{JI}$ of the second fundamental form,
- the connection coefficients $\gamma_{IJK} := g(\nabla_{e_I} e_J, e_K) = -\gamma_{IKJ}$ (where the symmetry comes from differentiating $g(e_J, e_K) = \delta_{JB}$),
- and the frame derivatives $e_0 \psi$ and $e_I \psi$ of the scalar field.

Remark 2.5 (Structure coefficients). It will also be important to consider the structure coefficients

$$S_{IJK} = g([e_I, e_J], e_K) = g(\nabla_{e_I} e_J, e_K) - g(\nabla_{e_J} e_I, e_K) = \gamma_{IJK} - \gamma_{JIK} = \gamma_{IJK} + \gamma_{JKI}. \tag{28}$$

By the Koszul formula (used in the proof of unique existence of the Levi–Civita connection), one can express the connection coefficients in terms of the structure coefficients:

$$\gamma_{IJK} = \frac{1}{2}(S_{IJK} + S_{KJI} + S_{KIJ}) \tag{29}$$

We call all the non-lapse variables *dynamical*, since they solve equations with time derivatives. In the proof, one shows that the lapse can be controlled by the dynamical quantities.

Remark 2.6 (Kasner background values). We have

$$\tilde{n} = 1, \qquad \tilde{e}_I^i = t^{-\tilde{q}_I} \delta_I^i, \qquad \tilde{k}_{IJ} = -\frac{\tilde{q}_I}{t} \delta_{IJ}, \qquad \tilde{\gamma}_{IJK} = 0, \qquad \tilde{\psi} = \tilde{B} \log t, \tag{30}$$

where we do *not* sum over repeated indices. Also, the frame indices are with respect to the background Kasner frame, and not the $\overline{g}$-orthonormal frame we just discussed.

2.4. **The reduced equations.**

Remark 2.7 (Relation between the reduced equations and the original equations). Solutions $(n, e_I^i, k_{IJ}, \gamma_{IJK}, \psi)$ to the following system of reduced equations are equivalent to solutions $\overline{g}$ of the Einstein equations. Moreover, smallness of the distance of *geometric initial data* $(g_\circ, k_\circ, \psi_\circ, \varphi_\circ)$ to Kasner data (in Sobolev spaces) implies smallness of the differences for the gauge-dependent quantities (including the lapse).

Remark 2.8 (Constructing a CMC data set). One can solve the Einstein equations locally in harmonic gauge. If the geometric data set producing this solution is close to Kasner data, then there is a spatial slice in this local spacetime solution that

(i) has constant mean curvature,
(ii) is close to the initial slice (in the sense that it is a graph over the initial slice with Sobolev norm small in terms of the smallness of data),
(iii) and the induced data on this slice are close to the data on the original slice, and hence close to Kasner initial data.

The lapse satisfies an elliptic equation

$$e_C e_C (n-1) - \frac{1}{t^2}(n-1) = \gamma \cdot en + n \cdot e\gamma + n \cdot \gamma \cdot \gamma + n \cdot e\psi \cdot e\psi. \tag{31}$$

The very singular t^{-2}-term on the left has a good sign. This is why the lapse will decay to 1.

Remark 2.9 (Schematic notation). In our schematic notation, a dot $(\cdot)$ represents an arbitrary product, and e will only denote derivatives with respect to the spatial frame (which always appear on the right side of the equations).

The frame coefficients satisfy a system of transport equations.

$$e_0 e_I^i = k_{IC} e_C^i. \tag{32}$$

Remark 2.10 (Gain in regularity for the frame). From this equation, one might expect that the frame components are at the same level of regularity as k_{IJ}. However, γ_{IJK} has the same level of differentiability as k_{IJ}. Since the connection coefficients γ_{IJK} are at the level of one derivative of the frame, this represents a *gain in regularity* for the frame. This comes from the fact that γ and k solve a coupled first-order hyperbolic system, so we can propagate their regularity from initial data.

The scalar field satisfies a wave equation

$$e_0 e_0 \psi + \frac{1}{t} e_0 \psi = e_C e_C \psi = n^{-1} en \cdot e\psi + \gamma \cdot e\psi \tag{33}$$

The first-order time-derivative term on the left has a good sign.

Once the lapse, frame, and scalar field are estimated, the equations for k and γ become a coupled first-order system. A caricature of the equations for k and γ is

$$\begin{cases} e_0(k_{IJ} - \tilde{k}_{IJ}) + \dfrac{1}{t}(k_{IJ} - \tilde{k}_{IJ}) = e\gamma + \gamma \cdot \gamma + \cdots \\ e_0 \gamma_{IJK} = ek + k \cdot \gamma + \cdots \end{cases} \tag{34}$$

Remark 2.11 (Using the constraints to generate energy identities). To derive energy estimates which do not lose derivatives, one must use the momentum constraint equation.

2.5. Approximately diagonal structure in the evolution equations for the structure coefficients.
Why do we need to introduce the structure coefficients?

Once γ is estimated, the equation for k in (34) leads to an estimate $k \lesssim t^{-1}$. However, the equation for γ in (34) is not diagonal: the term $k \cdot \gamma$ mixes together the components of γ, so an energy estimate seems elusive.

It turns out that $k \cdot \gamma$ has good structure: if we think of it as $k \cdot \gamma = \tilde{k} \cdot \gamma + (k - \tilde{k}) \cdot \gamma$, then we can estimate $k - \tilde{k}$ using the k-equation. Recall that $\tilde{k}_{IJ} = t^{-1} \tilde{q}_I \delta_{IJ}$. The exact structure of the $\tilde{k} \cdot \gamma$ terms are such that the system for $S_{IJK} = \gamma_{IJK} + \gamma_{JKI}$ diagonalizes up to error terms:

$$e_0 S_{IJK} + \frac{\tilde{q}_I + \tilde{q}_J - \tilde{q}_K}{t} S_{IJK} = \cdots \tag{35}$$

In particular, by the subcriticality condition (11), we can take

$$\max(|\tilde{q}_I|, \tilde{q}_I + \tilde{q}_J - \tilde{q}_K) < q < 1, \tag{36}$$

This lets us obtain

$$t^q |S_{IJK}| \lesssim \text{data.} \tag{37}$$

Remark 2.12 (Role of the subcriticality condition). This is the *only* place the subcriticality condition is used!

Remark 2.13 (Why the proof works in polarized-$U(1)$ symmetry). Under polarized-$U(1)$ symmetry in $3+1$ dimensions in vacuum (taking $e_3 = (g_{33})^{-1}\partial_3$, where ∂_3 is the Killing direction), all structure coefficients with three distinct indices (namely S_{123}, S_{231}, and S_{312})—in view of remark 1.18, these are the dangerous ones—vanish.

2.6. The high-level overview.
The proof has three ingredients:
- *highly degenerate* top-order energy estimates (in L^2),
- *sharp* low-order uncommuted pointwise estimates (in L^∞),
- an interpolation argument for intermediate derivatives.

2.6.1. *Top order*. At top order, we commute the equation with $t^A \partial^N$ for $N \gg A \gg 1$.

Remark 2.14 (Meaning of ∂). Here ∂ is a coordinate derivative in the CMC-transported spatial coordinates.

Remark 2.15 (Order of choosing constants). First we choose the degeneracy in the top order estimates to be large ($A \gg 1$). Then we choose the regularity parameter to be large ($N \gg A$), and finally we choose the data to be small ($\epsilon \ll_{A,N} 1$).

Definition 2.16 (Top order energy). At top order, we control the energy

$$\mathbf{H}(t) := t^A (\|k\|_{\dot{H}^N(\Sigma_t)} + \|\gamma\|_{\dot{H}^N(\Sigma_t)} + \|e_0\psi\|_{\dot{H}^N(\Sigma_t)} + \|e\psi\|_{\dot{H}^N(\Sigma_t)})$$
$$+ t^{A-1+q}\|e^i\|_{\dot{H}^N(\Sigma_t)} + t^{A-1}\|n\|_{\dot{H}^N(\Sigma_t)} + t^{A-1+q}\|en\|_{\dot{H}^N(\Sigma_t)}. \tag{38}$$

Remark 2.17 (Purpose of the t-weights). Observe that

$$\partial_t u = F \implies \partial_t(t^A \partial^N u) = \frac{A}{t} t^A \partial^N u + t^A \partial^N F. \tag{39}$$

Integrating on $(t, 1]$, we get a term with coefficient A of a *good sign*:

$$t^A \partial^N u = \partial^N u|_{t=1} - \boxed{\int_t^1 \frac{A}{s} s^A \partial^N u \, \mathrm{d}s} + \int_t^1 s^A \partial^N F \, \mathrm{d}s. \tag{40}$$

We call this term *borderline* because it has a borderline-non-integrable t^{-1}-weight. Such terms would be problematic for a Grönwall argument if they had a bad sign. This borderline term of a good sign will help us to cancel those terms.

2.6.2. *Low order*.

Definition 2.18 (Low order norm). At low order, we control the norm

$$\mathbf{L}(t) := t(\|k - \tilde{k}\|_{W^{1,\infty}(\Sigma_t)} + \|e_0\psi - \partial_t\tilde{\psi}\|_{W^{1,\infty}(\Sigma_t)}) + t^q(\|e - \tilde{e}\|_{L^\infty(\Sigma_t)} + \|\gamma\|_{L^\infty(\Sigma_t)})$$
$$+ t^{-\sigma}(\|n - 1\|_{L^\infty(\Sigma_t)} + t^q\|en\|_{L^\infty(\Sigma_t)}), \tag{41}$$

where $0 < \sigma < 1 - q$ is small.

Remark 2.19 (Comparing the estimate for k with those for S and e). Although our top-order estimates are very singular, our low-order estimates must be sharp in the sense that k_{IJ} and $e_0\psi$ blow up no faster than t^{-1}, which is the behaviour on Kasner.

However, there is room in the estimates for S and $e - \tilde{e}$, since q is a bit larger than $\tilde{q}$ and $\tilde{q}_I + \tilde{q}_J - \tilde{q}_K$ for $I \neq J$.

Remark 2.20 (Strategy to estimate the low order norm). We treat the evolution equations for the dynamical variables as ODEs in time with source terms coming from spatial derivatives. Some of these terms will be borderline due to the presence of $k - \tilde{k}$ terms, but those will have the smallness of ϵ. Some of these terms are not borderline, but they lose derivatives; we estimate these using an interpolation argument.

Remark 2.21 (Another manifestation of AVTD behaviour). Recall that the philosophy is that spatial derivatives behave better than time derivatives! This is why we can treat these terms as source terms.

Remark 2.22 (Relation to the heuristics for Kasner-type metrics). We have

$$\mathrm{Ric}_{IJ} = e\gamma + \gamma \cdot \gamma, \tag{42}$$

and so once we have (46) (and similar estimates for eS), we obtain the key bound

$$|\mathrm{Ric}_{IJ}| \lesssim t^{-2+\sigma} \tag{43}$$

that we needed in section 1.6.

Remark 2.23 (Kinetic terms are estimated at higher order). We control all quantities in L^∞ except for the *kinetic terms* $k - \tilde{k}$ and $e_0\psi - \partial_t\tilde{\psi}$, for which we also control one derivative. This will be because the right side of these equations will only see spatial derivative terms, and so there is room to use the interpolation argument.

2.6.3. *Interpolation.*

Lemma 2.24 (Interpolation). *We have*

$$\|\partial f\|_{L^\infty(\Sigma_t)} \lesssim_N \|f\|_{L^\infty(\Sigma_t)} + \|f\|_{L^\infty(\Sigma_t)}^{1-\delta_N}\|f\|_{\dot H^N(\Sigma_t)}^{\delta_N}, \qquad \delta_N \to 0 \text{ as } N \to \infty. \tag{44}$$

Proof. Sobolev embedding and interpolation (i.e. Hölder's inequality on the Fourier side). Note that $L^\infty \subset L^2$ since we are on the torus, which is compact. $\qquad\square$

Remark 2.25. The point is that, when N is large, $\partial\gamma$ is not much more singular than γ itself, even though $\|\gamma\|_{\dot H^N(\Sigma_t)}$ is very singular.

2.7. **Pointwise estimates for the low-order L^∞ norms.** We make the bootstrap assumption that

$$\mathbf{L}(t) \leq \epsilon, \tag{45}$$

and we want to improve this to

$$\mathbf{L}(t) \lesssim \text{data.} \tag{46}$$

As a caricature (with slightly more detail than before), we have

$$\begin{cases} \partial_t(k_{IJ} - \tilde k_{IJ}) + \dfrac{1}{t}(k_{IJ} - \tilde k_{IJ}) = e\gamma + \gamma \cdot \gamma + \cdots \\[2mm] \partial_t(e_I^i - \tilde e_i^I) + \dfrac{\tilde q_I}{t}(e_I^i - e_i^I) = (k - \tilde k) \cdot (e - \tilde e) + \cdots \\[2mm] \partial_t S_{IJK} + \dfrac{\tilde q_I + \tilde q_J - \tilde q_K}{t} S_{IJK} = (k - \tilde k) \cdot \gamma + \boxed{(e - \tilde e) \cdot \partial k} + \cdots \end{cases} \tag{47}$$

Remark 2.26 (The basic principle and the omitted terms). The basic principle is that

$$\partial_t u + \frac{\alpha}{t}u = 0 \implies t^\alpha u \leq \text{data.} \tag{48}$$

Of course, the right side of our equation is not zero, but includes borderline terms (as well as integrable terms)! There is lots of work to be done to control the "$\cdots$" terms (in particular, they couple the equations), but we ignore this for now.

Remark 2.27 (Schematic notation). By $k - \tilde k$ and $e - \tilde e$, we mean terms of the form $k_{AB} - \tilde k_{AB}$ and $e_C^c - \tilde e_C^c$. That is, the indices are the same in both the dynamical and the background quantities.

Remark 2.28 (The boxed term). The presence of the ∂k in the boxed term is why we must estimate $k - \tilde k$ in $W^{1,\infty}$ as opposed to just in L^∞. The reason we can improve this assumption is that the equation for $k - \tilde k$ does not see any borderline terms.

There are two issues here:

(i) the loss of derivatives on the right side of the equations for $k - \tilde k$ and for S

(ii) the presence of *borderline terms* on the right side, namely those with t^{-1}-weights (since $k - \tilde k \sim t^{-1}$).

We deal with (i) via interpolation and with (ii) using the smallness of ϵ and the room between q and the background quantities.

We first explain (ii). The equation for S gives

$$\partial_t(t^q S_{IJK}) = \frac{q - (\tilde q_I + \tilde q_J - \tilde q_K)}{t} t^q S_{IJK} + O(\epsilon t^{-1}) t^q \gamma. \tag{49}$$

Multiplying by $2t^q S_{IJK}$ and summing, we get (for some $\sigma > 0$ such that $\max_{I \neq J}(\tilde q_I + \tilde q_J - \tilde q_K) + \sigma < q$)

$$\begin{aligned} \sum_{I,J,K} \partial_t((t^q S_{IJK})^2) &= 2\sum_{I,J,K} \frac{q - (\tilde q_I + \tilde q_J - \tilde q_K)}{t}(t^q S_{IJK})^2 + O(\epsilon t^{-1}) t^q \gamma \cdot \sum_{I,J,K} t^q S_{IJK} + \cdots \\ &\geq \frac{1}{t}(2\sigma - C\epsilon) \sum_{I,J,K}(t^q S_{IJK})^2 \\ &\geq 0. \end{aligned} \tag{50}$$

Integrating on $(t, 1]$, we get the desired estimate

$$\text{data}^2 - \sum_{I,J,K} (t^q S_{IJK})^2 \geq 0. \tag{51}$$

The estimate for $e - \tilde{e}$ is obtained similarly. We now know that

$$t^q |S| + t^q |e - \tilde{e}| \lesssim \text{data} \tag{52}$$

Now we need to estimate $k - \tilde{k}$ from this equation:

$$\partial_t (k_{IJ} - \tilde{k}_{IJ}) + \frac{1}{t}(k_{IJ} - \tilde{k}_{IJ}) = e\gamma + \gamma \cdot \gamma + \cdots \tag{53}$$

Since γ is a linear combination of S, we have

$$\|\gamma \cdot \gamma\|_{L^\infty(\Sigma_t)} \lesssim t^{-2q}, \tag{54}$$

which is less singular than t^{-2}, hence integrable even after multiplying by t.

How do we control the term $e\gamma$? We know $|e_I^i| \lesssim t^{-q}$, so we just need to control $\partial\gamma$. Interpolation (lemma 2.24) and the estimates we already obtained imply

$$|\partial\gamma| \lesssim t^{-q-A\delta_N} \cdot \text{data} \implies |e\gamma| \lesssim t^{-2q-A\delta_N} \cdot \text{data}. \tag{55}$$

Since $N \gg A$, we can make this $A\delta_N$-term as small as we need to keep the $|e\gamma|$ term less singular than t^{-2} (which is possible since $q < 1$). Now the fundamental theorem of calculus gives

$$\partial_t(t(k - \tilde{k})) = O(t^{-1+\sigma}) \cdot \text{data} \implies t|k - \tilde{k}| \leq |k - \tilde{k}|_{t=1} + \text{data} \cdot \int_t^1 O(s^{-1+\sigma}) \, ds \lesssim \text{data}. \tag{56}$$

Remark 2.29 (The two lies). In reality, the estimates are more coupled than we have presented (due to omitted terms), and so one cannot improve the bootstrap assumptions for $e - \tilde{e}$ and S independently of the one for k. Even if there were no omitted terms, we cannot improve the assumptions for the low-order norm independently of the assumptions for the high-order norm (due to the interpolation argument used in the estimates for $k - \tilde{k}$).

2.8. Energy estimates for the high-order L^2 norms.

Two questions remain:

- How do we obtain the high-order L^2 estimates?
- Why do we need to take A large?

There are two difficulties associated to the energy estimates: the presence of borderline terms and the potential loss of derivatives. Recall the evolution equation for γ has the form

$$\partial_t \gamma = ek + k \cdot \gamma + \cdots \tag{57}$$

The first term on the right looks like it leads to a loss of derivatives, and the second term leads to borderline non-integrable terms.

2.8.1. *Borderline terms and the largeness of A.* There are a large number of borderline top-order error terms in the energy estimates. An example of a borderline term (coming from (57), the equation for γ) is

$$\tilde{k} \cdot \partial^N \gamma \cdot \partial^N \gamma \tag{58}$$

which is like $t^{-1}\tilde{q}\mathbf{H}(t)^2 \leq t^{-1}\mathbf{H}(t)^2$ in L^2. The point is that t^{-1} is not integrable, so this causes trouble for a Grönwall argument. On the other hand, commuting with the degenerate t^{A+1} weight produces another borderline term with a *good sign*. We want A to be larger than the size of remaining borderline terms, which can be done as long as this size is independent of N and A. In the end, we will see an estimate like

$$\mathbf{H}(t)^2 \leq \text{data}^2 + (C_* - A) \int_t^1 s^{-1}\mathbf{H}(s)^2 \, ds + C_N \int_t^1 s^{-1+\sigma}\mathbf{H}(s)^2 \, ds. \tag{59}$$

If $A > C_*$, we can drop this term, and then use Grönwall's inequality to conclude the desired estimate

$$\mathbf{H}(t)^2 \leq C_N \cdot \text{data}^2. \tag{60}$$

12

Remark 2.30 (The role of the t-weights in the energy and the sharp L^∞ estimate for k). The integral with a good sign comes exactly from the t^{A+1} weights in our energy norm.

In order to estimate the error term arising from (58), we used the *sharp* estimate $\|k_{IJ}\|_{L^\infty(\Sigma_t)} \lesssim C_*/t$ derived in the previous section. If we instead had $\|k_{IJ}\|_{L^\infty(\Sigma_t)} \lesssim C_*/t^{1+\eta}$, then we would see an error term $C_* \int_t^1 s^{-1-\eta} \mathbf{H}(s)^2 \, \mathrm{d}s$, and so Grönwall's inequality would only provide an estimate in which the energy *grows*.

Remark 2.31 (How large is A?). In [RS18a; RS18b], Rodnianski–Speck prove stable Big Bang formation in the nearly isotropic setting (i.e. near FLRW). In this setting, striking cancellations occur, and so C_* is small, which means A can be small (i.e. the top-order energy estimates do not have to be very degenerate). In our setting, C_* is like the total number of borderline terms, plus $C_N \epsilon$.

Where does the borderline term (58) come from, and how many such terms are there? The term $k \cdot \gamma$ on the right side of (57) can be split as

$$k \cdot \gamma = \tilde{k} \cdot \gamma + (k - \tilde{k}) \cdot \gamma \overset{\partial^N}{\rightsquigarrow} \tilde{k} \cdot \partial^N \gamma + \sum_{n_1 + n_2 = N} \partial^{n_1}(k - \tilde{k}) \cdot \partial^{n_2}\gamma. \tag{61}$$

Remark 2.32 (The two types of borderline terms). There is one term of the first type ($\tilde{k} \cdot \partial^N \gamma$), and there are an N-dependent number of terms of the second type ($\partial^{n_1}(k - \tilde{k}) \cdot \partial^{n_2}\gamma$), but terms of the second type come with smallness from the data (since $k - \tilde{k}$ is small).

Thus in the energy identity one sees

$$\partial_t((\partial^N \gamma)^2) + \cdots = \tilde{k} \cdot \partial^N \gamma \cdot \partial^N \gamma + \sum_{n_1 + n_2 = N} \partial^{n_1}(k - \tilde{k}) \cdot \partial^{n_2}\gamma \cdot \partial^N \gamma + \cdots \tag{62}$$

Remark 2.33 (Omitted terms). The omitted term on the left is $\partial_t((\partial^N k)^2)$, since the estimates for k and γ are coupled. There are also terms involving the frame and the lapse on the right, but we omit those for the purposes of this discussion.

After commuting with t^{2A}, we get

$$\partial_t(t^{2A}(\partial^N \gamma)^2) - \frac{2A}{t}t^{2A}(\partial^N \gamma)^2 + \cdots = \tilde{k} \cdot t^A \partial^N \gamma \cdot t^A \partial^N \gamma + \sum_{n_1 + n_2 = N} \partial^{n_1}(k - \tilde{k}) \cdot t^A \partial^{n_2}\gamma \cdot t^A \partial^N \gamma. \tag{63}$$

Upon integrating on $[t, 1] \times \mathbf{T}^{\mathfrak{D}}$, we get

$$t^{2A}\|\partial^N \gamma\|_{L^2}(t)^2 + \cdots \lesssim \text{data} - \int_t^1 \frac{2A}{s} s^{2A}\|\partial^N \gamma\|_{L^2}(s)^2 \, \mathrm{d}s + \text{(I)} + \text{(II)}, \tag{64}$$

where

$$
\begin{aligned}
\text{(I)} &:= \int_t^1 \int_{\Sigma_s} \tilde{k} \cdot t^A \partial^N \gamma \cdot t^A \partial^N \gamma \, \mathrm{d}x \, \mathrm{d}s \\
\text{(II)} &:= \sum_{n_1 + n_2 = N} \int_t^1 \int_{\Sigma_s} \partial^{n_1}(k - \tilde{k}) \cdot t^A \partial^{n_2}\gamma \cdot t^A \partial^N \gamma \, \mathrm{d}x \, \mathrm{d}s.
\end{aligned} \tag{65}
$$

Term (I) is like

$$\text{(I)} \leq \int_t^1 \frac{\tilde{q}}{s}\|\partial^N \gamma\|_{L^2}(s)^2 \, \mathrm{d}s \leq \int_t^1 \frac{1}{s}\|\partial^N \gamma\|_{L^2}(s)^2 \, \mathrm{d}s. \tag{66}$$

The number of such terms is independent of N and A.

For term (II), we use Cauchy–Schwarz and the interpolation inequality (for $n_1 + n_2 = N$)

$$\|\partial^{n_1}u \partial^{n_2}v\|_{L^2} \lesssim_N \|\partial^N u\|_{L^2}\|v\|_{L^\infty} + \|u\|_{L^\infty}\|\partial^N v\|_{L^2} \tag{67}$$

13

to get

$$\text{summand of (II)} \lesssim \int_t^1 \|\partial^{n_1}(k - \tilde{k}) \cdot t^A \partial^{n_2}\gamma\|_{L^2}(s)\|\partial^N \gamma\|_{L^2}(s)\,\mathrm{d}s$$

$$\lesssim_N \int_t^1 \left(\|k - \tilde{k}\|_{L^\infty}(s)\|\partial^N \gamma\|_{L^2}(s) + \|\gamma\|_{L^\infty}(s)\|\partial^N k\|_{L^2}(s)\right)\|\partial^N \gamma\|_{L^2}(s)\,\mathrm{d}s \tag{68}$$

$$\lesssim_N \int_t^1 \frac{\epsilon}{s}\|\partial^N \gamma\|_{L^2}(s)^2\,\mathrm{d}s + \int_t^1 \epsilon s^{-1+\sigma}\|\partial^N k\|_{L^2}\|\partial^N \gamma\|_{L^2}(s)\,\mathrm{d}s$$

$$\lesssim_N \int_t^1 \frac{\epsilon}{s}\|\partial^N \gamma\|_{L^2}(s)^2\,\mathrm{d}s + \int_t^1 \epsilon s^{-1+\sigma}\mathbf{H}(s)^2\,\mathrm{d}s.$$

There are an N-dependent number of such terms, but they come with a factor of ϵ, so the total size is still independent of N. The second term has an integrable-in-time weight, so it is admissible for a Grönwall argument (and would be even if it didn't come with smallness).

2.8.2. *Avoiding a loss of derivatives.* In the above discussion, we did not explain how the energy identity itself is derived. This is non-trivial, because the equations look like

$$\begin{cases} e_0 k_{IJ} + \dfrac{1}{t}k_{IJ} = e\gamma + n^{-1}een\cdots \\ e_0 \gamma_{IJK} = ek + k \cdot \gamma + \cdots \end{cases} \tag{69}$$

How do we deal with the extra spatial derivatives on the right side? The ek and $e\gamma$ terms have a special structure such that, although we cannot derive energy estimates for any given component, when we add all the components, a term from the k-equation cancels up to error terms with terms from the γ-equation, and the remaining terms can be differentiated by parts to create error terms plus terms involving the divergence of k, which is exactly what the momentum constraint expresses:

$$e_C k_{CI} = k \cdot \gamma + e_0 \psi \cdot e\psi. \tag{70}$$

More precisely, the equations are

$$\begin{cases} e_0 k_{IJ} + \dfrac{1}{t}k_{IJ} = e_C \gamma_{IJC} - e_I \gamma_{CJC} - n^{-1}e_I e_J n \cdots \\ e_0 \gamma_{IJK} = e_K k_{IJ} - e_J k_{IK} + k \cdot \gamma + \cdots \end{cases} \tag{71}$$

Up to forgetting about the lapse, we get

$$\begin{cases} \partial_t(k_{IJ}k_{IJ}) + \dfrac{2}{t}k_{IJ}k_{IJ} = \underline{2e_C \gamma_{IJC}k_{IJ}} - \boxed{2e_I \gamma_{CJC}k_{IJ}} - \boxed{2e_I e_J n k_{IJ}} \\ \dfrac{1}{2}\partial_t(\gamma_{IJK}\gamma_{IJK}) = e_K k_{IJ}\gamma_{IJK} - e_J k_{IK}\gamma_{IJK} = e_K k_{IJ}\gamma_{IJK} + e_J k_{IK}\gamma_{IKJ} = \underline{2e_K k_{IJ}\gamma_{IJK}} \end{cases} \tag{72}$$

Add these (invoking the summation convention). The underlined terms cancel up to errors involving ∂e after differentiation by parts (and relabeling of indices). Indeed, we have

$$2e_C \gamma_{IJC}k_{IJ} = 2e_C^c \partial_c \gamma_{IJC}k_{IJ} = -2e_C k_{IJ}\gamma_{IJC} + \partial e \cdot \gamma \cdot k = -2e_K k_{IJ}\gamma_{IJK} + \partial e \cdot \gamma \cdot k. \tag{73}$$

For the first boxed term, differentiate by parts and use the momentum constraint to get

$$e_I \gamma_{CJC}k_{IJ} = \gamma_{CJC} \cdot \underline{e_I k_{IJ}} + \partial e \cdot \gamma \cdot k + \partial(\cdots) \tag{74}$$

Similarly for the second boxed term, we have

$$e_I e_J n k_{IJ} = en \cdot \underline{e_I k_{IJ}} + \partial e \cdot en \cdot k + \partial(\cdots) \tag{75}$$

This leads to an identity of the form

$$\partial_t(k_{IJ}k_{IJ}) + \frac{1}{2}\partial_t(\gamma_{IJK}\gamma_{IJK}) = -\frac{2}{t}k_{IJ}k_{IJ} + \partial e \cdot en \cdot k + \partial(ne) \cdot k \cdot \gamma + \partial(\cdots) + \cdots \tag{76}$$

In practice, one needs to commute this identity with operators of the form $t^A \partial^N$.

References

[AG23]   Xinliang An and Dejan Gajic. "Curvature blow-up and mass inflation in spherically symmetric collapse to a Schwarzschild black hole". en. *Arch. Ration. Mech. Anal.* 247.3 (June 2023).

[AR01]   Lars Andersson and Alan D. Rendall. "Quiescent Cosmological Singularities". *Communications in Mathematical Physics* 218.3 (May 2001), pp. 479–511.

[AZ20]   Xinliang An and Ruixiang Zhang. "Polynomial Blow-Up Upper Bounds for the Einstein-Scalar Field System Under Spherical Symmetry". *Communications in Mathematical Physics* 376.2 (Jan. 2020), pp. 1671–1704.

[Ber+98]   B. K. Berger et al. "The singularity in generic gravitational collapse is spacelike, local and oscillatory". *Modern Physics Letters A* 13.19 (June 1998), pp. 1565–1573.

[BK73]   V. A. Belinskii and I. M. Khalatnikov. "Effect of Scalar and Vector Fields on the Nature of the Cosmological Singularity". *Sov. Phys. JETP* 36 (1973), pp. 591–597.

[BKL71]   V. A. Belinskiĭ, I. M. Khalatnikov, and E. M. Lifshitz. "Oscillatory approach to the singular point in relativistic cosmology". *Soviet Physics Uspekhi* 13.6 (June 1971), p. 745.

[BKL82]   V.A. Belinskii, I.M. Khalatnikov, and E.M. Lifshitz. "A general solution of the Einstein equations with a time singularity". *Advances in Physics* 31.6 (1982), pp. 639–667.

[BO23]   Florian Beyer and Todd A. Oliynyk. "Localized Big Bang Stability for the Einstein-Scalar Field Equations". *Archive for Rational Mechanics and Analysis* 248.1 (Dec. 2023).

[BOZ25]   F. Beyer, T. A. Oliynyk, and W. Zheng. *Localized past stability of the subcritical Kasner-scalar field spacetimes*. 2025. arXiv: 2502.09210 [gr-qc].

[Daf05]   Mihalis Dafermos. "The interior of charged black holes and the problem of uniqueness in general relativity". *Comm. Pure and Appl. Math* 58.4 (2005), pp. 445–504.

[Daf14]   Mihalis Dafermos. "Black Holes Without Spacelike Singularities". *Comm. Math. Phys.* 332.2 (May 2014), pp. 729–757.

[DHN03]   T. Damour, M. Henneaux, and H. Nicolai. "Cosmological billiards". *Classical and Quantum Gravity* 20.9 (Apr. 2003), R145.

[DHS85]   J. Demaret, M. Henneaux, and P. Spindel. "Non-oscillatory behavior in vacuum Kaluza-Klein cosmologies". *Phys. Lett. B* 164 (1985), pp. 27–30.

[DL25]   Mihalis Dafermos and Jonathan Luk. "The interior of dynamical vacuum black holes I: The C^0-stability of the Kerr Cauchy horizon". *Annals of Mathematics* 202.2 (2025), pp. 309–630.

[FL23]   Grigorios Fournodavlos and Jonathan Luk. "Asymptotically Kasner-like singularities". en. *Amer. J. Math.* 145.4 (Aug. 2023), pp. 1183–1272.

[FRS23]   Grigorios Fournodavlos, Igor Rodnianski, and Jared Speck. "Stable Big Bang formation for Einstein's equations: The complete sub-critical regime". *Journal of the American Mathematical Society* 36.3 (Apr. 2023), pp. 827–916.

[FU25]   David Fajman and Liam Urban. *On the past maximal development of near-FLRW data for the Einstein scalar-field Vlasov system*. 2025. arXiv: 2402.08544 [gr-qc].

[GPR25]   Hans Oude Groeniger, Oliver Petersen, and Hans Ringström. *Formation of quiescent big bang singularities*. 2025. arXiv: 2309.11370 [gr-qc].

[KL63]   I.M. Khalatnikov and E.M. Lifshitz. "Investigations in relativistic cosmology". *Advances in Physics* 12.46 (1963), pp. 185–249.

[Kli15]   Paul Klinger. "A new class of asymptotically non-chaotic vacuum singularities". *Annals of Physics* 363 (Dec. 2015), pp. 1–35.

[Li24]   Warren Li. *BKL bounces outside homogeneity: Gowdy symmetric spacetimes*. 2024. arXiv: 2408.12427 [gr-qc].

[Li25]   Warren Li. "Kasner-like description of spacelike singularities in spherically symmetric spacetimes with scalar matter". en. *Ann. PDE* 11.1 (June 2025).

[Mie24]   Benedikt Daniel Miethke. "C^0-inextendibility of the Kasner spacetime" (2024).

[Mis94]   Charles W. Misner. "The Mixmaster Cosmological Metrics". In: *Deterministic Chaos in General Relativity*. Ed. by David Hobill, Adrian Burd, and Alan Coley. Boston, MA: Springer US, 1994, pp. 317–328. ISBN: 978-1-4757-9993-4.

[Moo22]   Maxime Van de Moortel. *The breakdown of weak null singularities inside black holes*. 2022. arXiv: 1912.10890 [gr-qc].

[Moo25a]   Maxime Van de Moortel. *Asymptotically flat black holes with a singular Cauchy horizon and a spacelike singularity*. 2025. arXiv: 2510.07431 [gr-qc].

[Moo25b]   Maxime Van de Moortel. *The coexistence of null and spacelike singularities inside spherically symmetric black holes*. 2025. arXiv: 2504.12370 [gr-qc].

[Rin01]   Hans Ringström. "The Bianchi IX Attractor". *Annales Henri Poincaré* 2.3 (June 2001), pp. 405–500.

[Rin09]   Hans Ringström. "Strong cosmic censorship in T^3-Gowdy spacetimes". en. *Ann. Math.* 170.3 (Nov. 2009), pp. 1181–1240.

[RS18a]   Igor Rodnianski and Jared Speck. "A regime of linear stability for the Einstein-scalar field system with applications to nonlinear Big Bang formation". *Annals of Mathematics* 187.1 (Jan. 2018).

[RS18b]   Igor Rodnianski and Jared Speck. "Stable Big Bang formation in near-FLRW solutions to the Einstein-scalar field and Einstein-stiff fluid systems". en. *Sel. Math. New Ser.* 24.5 (Nov. 2018), pp. 4293–4459.

[RS21]   Igor Rodnianski and Jared Speck. "On the nature of Hawking's incompleteness for the Einstein-vacuum equations: The regime of moderately spatially anisotropic initial data". en. *J. Eur. Math. Soc.* 24.1 (Aug. 2021), pp. 167–263.

[RW01]   Alan D Rendall and Marsha Weaver. "Manufacture of Gowdy spacetimes with spikes". *Classical and Quantum Gravity* 18.15 (July 2001), pp. 2959–2975.

[Sbi18]   Jan Sbierski. "The C^0-inextendibility of the Schwarzschild spacetime and the spacelike diameter in Lorentzian geometry". *Journal of Differential Geometry* 108.2 (2018), pp. 319–378.

[Spe18]   Jared Speck. "The Maximal Development of Near-FLRW Data for the Einstein-Scalar Field System with Spatial Topology $\mathbb{S}^3$". *Communications in Mathematical Physics* 364.3 (Oct. 2018), pp. 879–979.

[Wal84]   Robert M. Wald. *General relativity*. University of Chicago Press, Chicago, IL, 1984, pp. xiii+491. ISBN: 0-226-87032-4.