Spike behavior in the approach to spacetime singularities

David Garfinkle[*]
Department of Physics, Oakland University, Rochester, Michigan 48309, USA
and Michigan Center for Theoretical Physics, Randall Laboratory of Physics, University of Michigan, Ann Arbor, Michigan 48109-1120, USA

Frans Pretorius[†]
Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544, USA

(Received 3 October 2020; accepted 8 December 2020; published 29 December 2020)

We perform numerical simulations of the approach to spacetime singularities. The simulations are done with sufficient resolution to resolve the small scale features (known as spikes) that form in this process. We find an analytical formula for the shape of the spikes and show that the spikes in the simulations are well described by this formula.

DOI: 10.1103/PhysRevD.102.124067

I. INTRODUCTION

Ever since the singularity theorem of Penrose [1], it has been known that spacetime singularities are a generic feature of gravitational collapse. However, Penrose's theorem gives very little information about the nature of these singularities, stating only that some light ray fails to be complete. In order to obtain a better understanding of the nature of singularities, Belinskii, Lifschitz, and Khalatnikov [2] (collectively known as BKL) conjectured an analytic approximation in which near the singularity, terms in the field equations containing spatial derivatives were negligible compared to those containing time derivatives. In order to test the correctness of the BKL conjecture, Berger and Moncrief [3] performed numerical simulations of the approach to the singularity in Gowdy spacetimes. The Gowdy spacetimes have two spatial Killing vectors and physically represent a closed universe with spatial topology T^3 containing plane gravitational waves and collapsing toward a singularity. The Gowdy spacetimes thus form a rather specialized class of spacetimes, which do not have direct astrophysical significance, but which can be thought of as a toy model for the general problem of gravitational collapse. Nonetheless, even in this special case, Berger and Moncrief found a new and unexpected feature of singularities: as the singularity was approached, the dynamics at almost all spatial points was in accord with the BKL conjecture; however, there were isolated points at which sharp features developed and became ever narrower the nearer one got to the singularity.

These sharp features later became known as spikes. The spikes represent a challenge for numerical simulations because an accurate numerical simulation requires that the spatial points that make up the numerical grid have sufficiently small separation to resolve all features. For a fixed spatial resolution, an ever narrowing spatial feature, such as the spikes found in [3] will eventually become too narrow to be resolved. However, because the Gowdy spacetimes have two spatial Killing fields, numerical simulations of these spacetimes require only a single spatial dimension, and thus can be done with a very fine spatial resolution. In [4], these fine scale numerical simulations were compared with an approximate analytical formula for the behavior of the spikes and were shown to match that formula. In [5], a class of exact analytic solutions was found for spikes in Gowdy spacetimes and shown in [6] to approach the late-time behavior of numerical simulations of spike formation in G_2 spacetimes (a generalization of the Gowdy model, but still with two spatial Killing vectors). Thus, despite the numerical challenges that they pose, spikes in Gowdy spacetimes are well understood.

The work of [3] was generalized to the case of only one Killing field [7,8] and later (using a different numerical method based on the analytical work of [9]) to the case of no symmetry [10]. However, the simulations of [7,10] did not have sufficient resolution to resolve the spikes. One method to obtain better resolution is adaptive mesh refinement (AMR) [11], which detects when resolution is about to become insufficient and then adds extra spatial points where they are needed. Indeed, AMR was used to resolve spikes in Gowdy spacetimes by Hern and Stewart [12]. However, though AMR is an effective method to use on Gowdy spacetimes, it is not so effective for the case of only one symmetry, or for the case of no symmetry. This is

[*]garfinkl@oakland.edu
[†]fpretori@princeton.edu

because AMR works well when the features that it needs to resolve occur at isolated spatial points, while (as we will see later) spikes are features of codimension one: that is, spikes occur at isolated points in the case of two symmetries, along curves in the case of one symmetry, and at surfaces in the case of no symmetry. Thus, in the latter two cases, the AMR would need to resolve too many regions and would quickly be overwhelmed. To obtain answers with adequate resolution in a reasonable amount of time thus requires that we parallelize the code; we use the PAMR/AMRD [13] libraries to do this (though again we only use its parallelization features and not AMR). Our highest resolution run used 112 cores of the *Perseus* cluster at Princeton, taking two days to complete.

In Sec. II, we present the field equations used in our simulations. These are the vacuum Einstein field equations expressed in terms of the scale invariant variables of [9]. Section III introduces a truncation of these equations obtained by applying the BKL approximation and derives an analytic formula for the spike from these truncated equations. Sections III A and III B explore the implications of the approximations made in Sec. III. Section IV presents one-dimensional (1D) (i.e., the case of two Killing fields) simulations of the equations of Sec. II and the comparison of those results to the analytic formula of Sec. III. Section V performs the same sort of simulations and comparison to analytic formula for the two-dimensional (i.e., one Killing field) case. Our conclusions are presented in Sec. VI.

II. EQUATIONS OF MOTION

The method we use to evolve the vacuum Einstein equations is the scale invariant tetrad method of Uggla *et al.* [9]. We use this method with constant mean curvature slicing as is done in the simulations of [14] (or equivalently as is done in the cosmological simulations of [15,16] but with no scalar field matter). More information on this type of method can be found in [9,14–16].

The spacetime is described in terms of a coordinate system (t, x^i) and a tetrad $(\mathbf{e}_0, \mathbf{e}_\alpha)$ where both the spatial coordinate index i and the spatial tetrad index α go from 1 to 3. Choose $\mathbf{e}_0$ to be hypersurface orthogonal with the relation between tetrad and coordinates of the form $\mathbf{e}_0 = N^{-1}\partial_t$, and $\mathbf{e}_\alpha = e_\alpha{}^i\partial_i$, where N is the lapse and the shift is chosen to be zero. Choose the spatial frame $\{\mathbf{e}_\alpha\}$ to be Fermi propagated along the integral curves of $\mathbf{e}_0$. The commutators of the tetrad components are decomposed as follows:

$$[\mathbf{e}_0, \mathbf{e}_\alpha] = \dot{u}_\alpha \mathbf{e}_0 - (H\delta_\alpha{}^\beta + \sigma_\alpha{}^\beta)\mathbf{e}_\beta, \tag{1}$$

$$[\mathbf{e}_\alpha, \mathbf{e}_\beta] = (2a_{[\alpha}\delta_{\beta]}{}^\gamma + \epsilon_{\alpha\beta\delta}n^{\delta\gamma})\mathbf{e}_\gamma, \tag{2}$$

where $n^{\alpha\beta}$ is symmetric and $\sigma^{\alpha\beta}$ is symmetric and trace free. The scale invariant tetrad variables are defined by $\partial_0 \equiv \mathbf{e}_0/H$ and $\partial_\alpha \equiv \mathbf{e}_\alpha/H$, while scale invariant versions of the other gravitational variables are given by

$$\{E_\alpha{}^i, \Sigma_{\alpha\beta}, A^\alpha, N_{\alpha\beta}\} \equiv \{e_\alpha{}^i, \sigma_{\alpha\beta}, a^\alpha, n_{\alpha\beta}\}/H. \tag{3}$$

Note that the relation between the scale invariant tetrad variables and the coordinate derivatives is

$$\partial_0 = \mathcal{N}^{-1}\partial_t, \tag{4}$$

$$\partial_\alpha = E_\alpha{}^i\partial_i, \tag{5}$$

where $\mathcal{N} = NH$ is the scale invariant lapse. The time coordinate t is chosen so that

$$e^{-t} = 3H. \tag{6}$$

Here we have used the scale invariance of the physical system to make both t and H dimensionless quantities. Note that Eq. (6) means that the surfaces of constant time are constant mean curvature surfaces. Note also that the singularity is approached as $t \to -\infty$.

Due to Eq. (6), the scale invariant lapse satisfies an elliptic equation

$$-\partial^\alpha\partial_\alpha\mathcal{N} + 2A^\alpha\partial_\alpha\mathcal{N} + \mathcal{N}(3 + \Sigma_{\alpha\beta}\Sigma^{\alpha\beta}) = 3. \tag{7}$$

The gravitational quantities $E_\alpha{}^i$, A_α, $N^{\alpha\beta}$, and $\Sigma_{\alpha\beta}$ satisfy the following evolution equations:

$$\partial_t E_\alpha{}^i = E_\alpha{}^i - \mathcal{N}(E_\alpha{}^i + \Sigma_\alpha{}^\beta E_\beta{}^i), \tag{8}$$

$$\partial_t A_\alpha = A_\alpha + \frac{1}{2}\Sigma_\alpha{}^\beta\partial_\beta\mathcal{N} - \partial_\alpha\mathcal{N} + \mathcal{N}\left(\frac{1}{2}\partial_\beta\Sigma_\alpha{}^\beta - A_\alpha - \Sigma_\alpha{}^\beta A_\beta\right), \tag{9}$$

$$\partial_t N^{\alpha\beta} = N^{\alpha\beta} - \epsilon^{\gamma\delta(\alpha}\Sigma_\delta{}^{\beta)}\partial_\gamma\mathcal{N} + \mathcal{N}(-N^{\alpha\beta} + 2N^{(\alpha}{}_\gamma\Sigma^{\beta)\gamma} - \epsilon^{\gamma\delta(\alpha}\partial_\gamma\Sigma_\delta{}^{\beta)}), \tag{10}$$

$$\begin{aligned}\partial_t\Sigma_{\alpha\beta} = {}& \Sigma_{\alpha\beta} + \partial_{\langle\alpha}\partial_{\beta\rangle}\mathcal{N} + A_{\langle\alpha}\partial_{\beta\rangle}\mathcal{N} \\ & + \epsilon_{\gamma\delta\langle\alpha}N_{\beta\rangle}{}^\delta\partial^\gamma\mathcal{N} + \mathcal{N}[-3\Sigma_{\alpha\beta} \\ & - \partial_{\langle\alpha}A_{\beta\rangle} - 2N_{\langle\alpha}{}^\gamma N_{\beta\rangle\gamma} + N^\gamma{}_\gamma N_{\langle\alpha\beta\rangle} \\ & + \epsilon_{\gamma\delta\langle\alpha}(\partial^\gamma N_{\beta\rangle}{}^\delta - 2A^\gamma N_{\beta\rangle}{}^\delta)].\end{aligned} \tag{11}$$

Here parentheses around a pair of indices denote the symmetric part, while angle brackets denote the symmetric trace-free part.

In addition, the variables are subject to the vanishing of the following constraint quantities:

$$(\mathcal{C}_{\text{com}})^{\lambda i} = \epsilon^{\alpha\beta\lambda}[\partial_\alpha E_\beta{}^i - A_\alpha E_\beta{}^i] - N^{\lambda\gamma}E_\gamma{}^i, \tag{12}$$

$$(\mathcal{C}_J)^\gamma = \partial_\alpha N^{\alpha\gamma} + \epsilon^{\alpha\beta\gamma}\partial_\alpha A_\beta - 2A_\alpha N^{\alpha\gamma}, \tag{13}$$

$$(\mathcal{C}_C)_\alpha = \partial_\beta \Sigma_\alpha{}^\beta - 3\Sigma_\alpha{}^\beta A_\beta - \epsilon_{\alpha\beta\gamma} N^{\beta\delta}\Sigma_\delta{}^\gamma, \tag{14}$$

$$\mathcal{C}_G = 1 + \frac{2}{3}\partial_\alpha A^\alpha - A^\alpha A_\alpha - \frac{1}{6}N^{\alpha\beta}N_{\alpha\beta} + \frac{1}{12}(N^\gamma{}_\gamma)^2 - \frac{1}{6}\Sigma^{\alpha\beta}\Sigma_{\alpha\beta}. \tag{15}$$

Initial data are chosen to satisfy the constraints of Eqs. (12)–(15), which are then preserved (to within numerical truncation error) under evolution. The data are evolved using the evolution equations, Eqs. (8)–(11), where to obtain a hyperbolic system a multiple of Eq. (14) is added to the right-hand side of Eq. (9) [17].

III. UNIVERSAL SPIKE BEHAVIOR

We now derive an analytic approximation for the shape of the spikes. The BKL conjecture for the system in this form says that sufficiently close to the singularity, A_α and $E_\alpha{}^i$ are small enough to be neglected. Note that all spatial derivatives occur in the equations of motion in the form $\partial_\alpha = E_\alpha{}^i\partial_i$, so it follows that all these terms can also be neglected. Subject to this approximation, we find the following: Eq. (7) becomes

$$\mathcal{N}^{-1} = 1 + \frac{1}{3}\Sigma_{\alpha\beta}\Sigma^{\alpha\beta}. \tag{16}$$

Equations (8) and (9) are automatically satisfied. Equations (10) and (11) become

$$\partial_t N^{\alpha\beta} = N^{\alpha\beta} + \mathcal{N}(-N^{\alpha\beta} + 2N^{(\alpha}{}_\gamma\Sigma^{\beta)\gamma}), \tag{17}$$

$$\partial_t \Sigma_{\alpha\beta} = \Sigma_{\alpha\beta} + \mathcal{N}[-3\Sigma_{\alpha\beta} - 2N_{\langle\alpha}{}^\gamma N_{\beta\rangle\gamma} + N^\gamma{}_\gamma N_{\langle\alpha\beta\rangle}]. \tag{18}$$

Equations (12) and (13) are automatically satisfied, while Eqs. (14) and (15) become

$$\epsilon_{\alpha\beta\gamma}N^{\beta\delta}\Sigma_\delta{}^\gamma = 0, \tag{19}$$

$$1 - \frac{1}{6}N^{\alpha\beta}N_{\alpha\beta} + \frac{1}{12}(N^\gamma{}_\gamma)^2 - \frac{1}{6}\Sigma^{\alpha\beta}\Sigma_{\alpha\beta} = 0. \tag{20}$$

We begin at an initial time close enough to the singularity that the conditions of the BKL conjecture are satisfied and follow the behavior through one bounce.[1] Equation (19) implies that the matrices $\Sigma^\alpha{}_\beta$ and $N^\alpha{}_\beta$ commute and therefore have a common basis of eigenvectors. It then follows from Eqs. (17) to (18) that the eigenvectors are constant in time (see Appendix B of [16] for more details), so all that we need to do is find the time dependence of the eigenvalues. Denote the eigenvalues of $\Sigma^\alpha{}_\beta$ by Σ_1, Σ_2, and Σ_3 with $\Sigma_1 \le \Sigma_2 \le \Sigma_3$ at the initial time. Let N_1 be the eigenvalue of $N^\alpha{}_\beta$ corresponding to the eigenvector of $\Sigma^\alpha{}_\beta$ that has eigenvalue Σ_1, and correspondingly for N_2 and N_3. We assume that at the initial time N_1, N_2, and N_3 are all negligibly small. Then it follows from Eq. (17) that N_1 grows in magnitude during the bounce process, but that N_2 and N_3 decrease in magnitude and therefore remain negligible. We then find from Eqs. (16) and (20) that

$$\mathcal{N}^{-1} = 3 - \frac{1}{6}(N_1)^2. \tag{21}$$

Using Eq. (21) in Eqs. (17) and (18), we then obtain

$$\partial_t \Sigma_1 = (1 - 3\mathcal{N})(\Sigma_1 + 4), \tag{22}$$

$$\partial_t \Sigma_2 = (1 - 3\mathcal{N})(\Sigma_2 - 2), \tag{23}$$

$$\partial_t \Sigma_3 = (1 - 3\mathcal{N})(\Sigma_3 - 2), \tag{24}$$

$$\partial_t N_1 = N_1(1 + \mathcal{N}(-1 + 2\Sigma_1)). \tag{25}$$

Now define the quantity Z by

$$Z \equiv \Sigma_1 + 4. \tag{26}$$

Then it follows from Eqs. (22) to (24) and the fact that $\Sigma_{\alpha\beta}$ is trace free that there is a constant b such that

$$\Sigma_2 = 2 + (b - 1)Z, \tag{27}$$

$$\Sigma_3 = 2 - bZ. \tag{28}$$

It then follows from Eqs. (16) and (26) and (28) that

$$\mathcal{N}^{-1} - 3 = \frac{2}{3}(b^2 - b + 1)Z^2 - 4Z + 6. \tag{29}$$

Thus, $\mathcal{N}^{-1} - 3$ is a quadratic in Z. Let Z_+ and Z_- be the roots of this quadratic. Then, we have

$$Z_\pm = \frac{3 \pm 3\sqrt{b(1-b)}}{b^2 - b + 1}. \tag{30}$$

Using Eq. (30) in Eq. (29), we obtain

$$\mathcal{N}^{-1} - 3 = \frac{6}{Z_+ Z_-}(Z - Z_+)(Z - Z_-). \tag{31}$$

It then follows from Eqs. (22), (26), and (31) that Z satisfies the equation of motion

[1] Here "bounce" refers to transition from one Kasner-like era to the next, not to be confused with the usage of the word in cosmology denoting transition from a contracting to expanding universe.

$$\partial_t Z = \frac{(Z - Z_+)(Z - Z_-)Z}{\frac{1}{2}Z_+Z_- + (Z - Z_+)(Z - Z_-)}. \tag{32}$$

From Eq. (32), we immediately obtain the following qualitative picture of spike formation: suppose that at the initial time there is a region where N_1 is positive and a region where N_1 is negative. Then by continuity there must be a surface where N_1 vanishes. It then follows from Eq. (25) that on this surface N_1 will always be zero, and it then follows from Eqs. (21) and (31) that $Z = Z_-$ on this surface. Now consider a point near this surface. Then N_1 is small but nonzero, and therefore Z is close to, but not equal to Z_-. It then follows from Eq. (32) that the evolution takes Z from near Z_- at the initial time to asymptotically close to Z_+ at large negative time (recall that the convention is that $t \to -\infty$ as the singularity is approached). Thus, the surface $N_1 = 0$ is stuck in the old phase, while all nearby points eventually bounce to the new phase. Thus, a feature of ever more narrow size forms in the vicinity of the surface.

But we can do even better than this qualitative picture and obtain a complete quantitative picture by integrating Eq. (32). Suppose that at some spatial point at time t_0 we have $Z = Z(t_0)$. Then some straightforward but tedious algebra leads to the following integral of Eq. (32):

$$\begin{aligned}&\exp\left[\frac{2}{Z_+}(Z_+ - Z_-)(t_0 - t)\right] \\ &\quad = \left(\frac{Z - Z_-}{Z(t_0) - Z_-}\right)\left(\frac{Z_+ - Z}{Z_+ - Z(t_0)}\right)^{-Z_-/Z_+} \\ &\qquad \times \left(\frac{Z}{Z(t_0)}\right)^{-3(Z_+ - Z_-)/Z_+}.\end{aligned} \tag{33}$$

Now consider Eq. (33) in the vicinity of a spike. Choose time t_0 sufficiently early in the process that no sharp features have formed, and choose a local coordinate x to vanish where N_1 vanishes. Then for sufficiently small x we have that N_1 is well approximated by $N_1 = \epsilon x$ where ϵ is a function of the coordinates transverse to x. It then follows from Eqs. (16) and (21) that near $x = 0$, we have

$$Z(t_0) = Z_-\left[1 + \frac{Z_+}{Z_+ - Z_-}\left(\frac{\epsilon x}{6}\right)^2\right]. \tag{34}$$

Then using Eq. (34) in Eq. (33), we obtain

$$\begin{aligned}&\left(\frac{\epsilon x}{6}\right)^2 \exp\left[\frac{2}{Z_+}(Z_+ - Z_-)(t_0 - t)\right] \\ &\quad = \left[\frac{(Z_+ - Z_-)^2}{Z_+Z_-}\right]\left[\frac{Z - Z_-}{Z_+ - Z}\right]\left[\frac{(Z_+ - Z)Z_-^3}{(Z_+ - Z_-)Z^3}\right]^{(Z_+ - Z_-)/Z_+}.\end{aligned} \tag{35}$$

Equation (35) shows that spikes are essentially a codimension one phenomenon, since everything can be expressed in terms of a single coordinate orthogonal to the spike surface. Thus, one should obtain essentially the same behavior in a two-dimensional simulation as in a one-dimensional simulation.

We now consider how to compare the results of the simulations to the prediction of Eq. (35). Though so far we have talked about the eigenvalues of $\Sigma^\alpha{}_\beta$ and $N^\alpha{}_\beta$, all the information about the eigenvalues of a matrix is contained in the invariants of that matrix and it is far simpler to compute invariants than to compute eigenvalues. In particular, since $N^\alpha{}_\beta$ has only one non-negligible eigenvalue, N_1, we find that the invariant $N^\alpha{}_\alpha$ is simply equal to N_1. It then follows from Eqs. (21) and (31) that

$$N^\alpha{}_\alpha = \frac{\pm 6}{\sqrt{Z_+Z_-}}(Z_+ - Z)^{1/2}(Z - Z_-)^{1/2}. \tag{36}$$

Note that Eqs. (36) and (35) together give a parametric equation for $N^\alpha{}_\alpha$ as a function of x (because the equations give both x and $N^\alpha{}_\alpha$ as functions of Z). Thus, to make a comparison with simulations, one should find from the simulation $N^\alpha{}_\alpha$ as a function of x and compare to this parametric curve.

We now consider the behavior of the invariants of $\Sigma^\alpha{}_\beta$. Define the quantity $\mathcal{S}$ by

$$\mathcal{S} = \Sigma^\alpha{}_\beta \Sigma^\beta{}_\gamma \Sigma^\gamma{}_\alpha. \tag{37}$$

Since $\Sigma^\alpha{}_\beta$ is trace free, it follows that the invariants of $\Sigma^\alpha{}_\beta$ are $\Sigma^{\alpha\beta}\Sigma_{\alpha\beta}$ and $\mathcal{S}$. However, from Eq. (20) and the fact that N_2 and N_3 are negligible, it follows that $\Sigma^{\alpha\beta}\Sigma_{\alpha\beta} = 6 - \frac{1}{2}(N^\alpha{}_\alpha)^2$ so there is no information in $\Sigma^{\alpha\beta}\Sigma_{\alpha\beta}$ that is not already contained in $N^\alpha{}_\alpha$. Therefore, in characterizing the invariants of $\Sigma^\alpha{}_\beta$, we can restrict our attention to $\mathcal{S}$. From Eqs. (26) and (27), we find

$$\mathcal{S} = 6[1 - (Z - 3)^2] + 3b(1 - b)Z^2(Z - 4). \tag{38}$$

Equations (35) and (38) give a parametric equation for $\mathcal{S}$ as a function of x. Thus, one should find from the simulations $\mathcal{S}$ as a function of x and compare to this parametric curve.

The formulas given in Eqs. (35), (36), and (38) contain two parameters: b and ϵ. Thus, to make comparisons with the simulations, we must specify how to determine these parameters from the simulations. To determine b, it is helpful to recall the definition of the BKL parameter u. Consider a time before the bounce when N_1 is negligible and the eigenvalues of $\Sigma^\alpha{}_\beta$ are approximately constant. This is a Kasner era, and the Kasner exponents p_1, p_2, and p_3 are expressed in terms of the corresponding eigenvalues of $\Sigma^\alpha{}_\beta$ by $p_i = (1 + \Sigma_i)/3$. The BKL parameter u is defined by

$$u = p_3/p_2. \tag{39}$$

Note that since $\Sigma_2 \leq \Sigma_3$ it follows that $p_2 \leq p_3$ and therefore that $u \geq 1$. Then it follows from Eqs. (27), (28), (30) using some straightforward algebra that

$$b = \frac{1}{u^2+1}, \tag{40}$$

$$Z_- = \frac{3(u^2+1)}{u^2+u+1}. \tag{41}$$

Before the bounce we have $Z \approx Z_-$. Let $\mathcal{S}_-$ denote the value of $\mathcal{S}$ before the bounce. Then using Eqs. (38), (40), and (41) straightforward but tedious algebra yields

$$\mathcal{S}_- = 6 - \frac{81u^2(u+1)^2}{(u^2+u+1)^3}. \tag{42}$$

As long as $-6 < \mathcal{S}_- < 6$, there is a unique $u > 1$ such that Eq. (42) is satisfied. Thus, to compute the parameter b in the simulations, we simply compute the invariant $\mathcal{S}$ before a bounce and then use Eq. (42) to determine u, and then use Eq. (40) to determine b.

There are two different ways to determine the parameter ϵ. This parameter is defined so that at time t_0 we have $N_1 = \epsilon x$, so we can simply use the definition to read off ϵ from the properties of N_1 at a time before the bounce. Alternatively, if we wait until a time t_1 at which a spike has formed, we can use the properties of the spike to determine ϵ as follows: from Eq. (36), it follows that the maximum value of $N^\alpha{}_\alpha$ occurs at $Z = (Z_+ + Z_-)/2$. (Note also that this maximum value is $3(Z_+ - Z_-)/\sqrt{Z_+ Z_-}$, a prediction that can easily be compared to the simulations). Let x_m be the value of x at which this maximum value of $N^\alpha{}_\alpha$ occurs. Then it follows from Eq. (35) that

$$\left(\frac{\epsilon x_m}{6}\right)^2 \exp\left[\frac{2}{Z_+}(Z_+ - Z_-)(t_0 - t_1)\right] = \left[\frac{(Z_+ - Z_-)^2}{Z_+ Z_-}\right]\left[\frac{4Z_-^3}{(Z_+ + Z_-)^3}\right]^{(Z_+ - Z_-)/Z_+}. \tag{43}$$

Combining Eqs. (35) and (43), we obtain

$$\left(\frac{x}{x_m}\right)^2 \exp\left[\frac{2}{Z_+}(Z_+ - Z_-)(t_1 - t)\right] = \left[\frac{Z - Z_-}{Z_+ - Z}\right]\left[\frac{(Z_+ + Z_-)^3(Z_+ - Z)}{4(Z_+ - Z_-)Z^3}\right]^{(Z_+ - Z_-)/Z_+}. \tag{44}$$

Thus, Eqs. (44) and (36) provide a parametric curve for $N^\alpha{}_\alpha$ versus x, while Eqs. (44) and (38) provide such a curve for $\mathcal{S}$ versus x.

A. Ephemeral nature of universal spike behavior

The universal spike formulas of the previous section were derived under the assumption that spatial derivatives are negligible. However, it follows from the spike formulas that spatial derivatives become arbitrarily large. Can a quantity be both arbitrarily large and negligible? In the equations of motion, all spatial derivatives appear multiplied by $E_\alpha{}^i$. Thus, spatial derivatives of a quantity F are negligible in the equations of motion provided that the quantity $E_\alpha{}^i \partial_i F$ is negligible. Specifically, we will use the spike formulas to calculate the quantity $E_\alpha{}^i \partial_i N^\beta{}_\beta$ at the center of the spike. Let the subscript I denote tetrad component in the direction of the Ith eigenvector of $\Sigma^\alpha{}_\beta$. Then using Eqs. (8) and (25) and the fact that $\mathcal{N} = 1/3$ in the center of the spike, we find that

$$\partial_t(E_I{}^i \partial_i N_1) = \frac{1}{3}[4 + 2\Sigma_1 - \Sigma_I](E_I{}^i \partial_i N_1). \tag{45}$$

It then follows that the magnitude of $E_I{}^i \partial_i N_1$ gets smaller as the singularity is approached if and only if the quantity in square brackets is positive. However, using Eqs. (26)–(28) and (40) and (41), we find

$$\begin{aligned} \frac{1}{3}[4 + 2\Sigma_1 - \Sigma_1] &= \frac{u^2+1}{u^2+u+1}, \\ \frac{1}{3}[4 + 2\Sigma_1 - \Sigma_2] &= \frac{u(u-2)}{u^2+u+1}, \\ \frac{1}{3}[4 + 2\Sigma_1 - \Sigma_3] &= \frac{2(1-u)}{u^2+u+1}. \end{aligned} \tag{46}$$

The first of these quantities is always positive, the last is always negative, and the one in the middle is positive when $u > 2$. What is going on is the following: during this particular epoch, as the singularity is approached, $E_1{}^i$ and $E_2{}^i$ are getting smaller, while $E_3{}^i$ is getting larger. Since the spatial derivative of N_1 is getting larger, it follows that the product of that spatial derivative and $E_3{}^i$ is always getting larger, though if at the beginning of the epoch $E_3{}^i$ starts out very small, it may take some time before this product is non-negligible. In contrast, $E_1{}^i$ is getting small faster than the spatial derivative of N_1 is getting large, so the product of these two quantities is always negligible. $E_2{}^i$ is getting small at a rate that is faster (resp. slower) than the spatial derivative of N_1 is getting large if $u > 2$ (resp. $u < 2$); thus, the product of the two quantities may be getting larger or smaller, depending on the value of u.

This line of reasoning suggests that spikes are ephemeral, or at least that there is only a limited time period under which each spike is accurately described by the spike formulas of the previous section. This is consistent with the transient spike solutions for Gowdy spacetimes found in [5].

B. Early spikes and late spikes

We now consider the extent to which we should expect the approximate formulas of Sec. III to match an actual

evolution of the Einstein field equations. The results of Sec. III are based on the assumptions of that section, namely, that $E_\alpha{}^i$ and A_α are negligibly small. We expect this assumption to be better and better satisfied the closer we are to the singularity, i.e., the longer the simulation is run. Or to put it another way: we expect late spikes (the ones that occur later in the simulation) to be better modeled by the analytic formulas of Sec. III. However, a simulation can only be run as long as it maintains enough resolution for accurate results. Since spikes are features that become very narrow, that means that eventually in every simulation some spike will become sufficiently narrow to make the simulation lose resolution. Or to put it another way: simulations can only see early spikes. Thus, there is something of a mismatch between the needs of the simulations and the needs of the spike formula: we expect the spike formula to be a crude, rather than exact, model for the early spikes produced in the simulations.

IV. 1D SIMULATIONS

Our methods for the one-dimensional (i.e., two Killing field) case are essentially those of [15,16] but without the scalar field matter. In particular, we must choose initial data that satisfy the constraint equations (12)–(15). We do this using the York method [18]. That is, we write the initial data in terms of a freely specifiable piece and an unknown conformal factor which we solve for numerically. The initial data are as follows:

$$E_\alpha{}^i = H^{-1}\psi^{-2}\delta_\alpha{}^i, \tag{47}$$

$$A_\alpha = -2\psi^{-1}E_\alpha{}^i\partial_i\psi, \tag{48}$$

$$N_{\alpha\beta} = 0, \tag{49}$$

$$\Sigma_{\alpha\beta} = \psi^{-6}Z_{\alpha\beta}. \tag{50}$$

Here ψ is the unknown conformal factor and H is a constant. The constraint equations require that Z_{ik} be divergence free: that is $\partial^i Z_{ik} = 0$. In addition, since $\Sigma_{\alpha\beta}$ is trace free, so is Z_{ik}. We choose the following simple Z_{ij} having both these properties:

$$Z_{ik} = \begin{bmatrix} b_2 & \kappa_1 & \kappa_2 \\ \kappa_1 & a_1\cos x + b_1 & a_2\cos x \\ \kappa_2 & a_2\cos x & -b_1 - b_2 - a_1\cos x \end{bmatrix}. \tag{51}$$

Here, a_1, a_2, b_1, b_2, κ_1, and κ_2 are constants. The constraint equations require that the conformal factor satisfy the equation

$$\partial^i\partial_i\psi = \frac{3}{4}H^2\psi^5 - \frac{1}{8}Z^{ik}Z_{ik}H^2\psi^{-7}, \tag{52}$$

which we solve numerically.

Figures 1–3 show the results of three simulations for three different choices of the parameters $(a_1, a_2, b_1, b_2, \kappa_1, \kappa_2)$. In each case, snapshots of $N^\alpha{}_\alpha$ versus x are plotted at different times in the simulation. The simulations are run with a spatial step size of $dx = 2\pi/2500 \approx 0.00251$ and each simulation is run only for as long as good resolution can be maintained.

Since the initial data have $N_{\alpha\beta} = 0$, it follows that initially $N^\alpha{}_\alpha$ vanishes, and its early structure reflects simple linear growth driven by the initial data for $\Sigma_{\alpha\beta}$ through the right-hand side of Eq. (10). However, eventually spikes form around points where $N^\alpha{}_\alpha$ remains zero. Thus, the later panels in these figures show structure that is mostly indicative of the presence of spikes.

It is clear from the figures that each simulation produces several spikes. However, as argued in the previous section, the early spikes of a simulation cannot be expected to be well described by the formulas of that section, and even the "late" spikes of a simulation are sufficiently "early" that the formulas of Sec. III can only be expected to be a fairly crude approximation. For this reason, we will examine one late spike per simulation.

Figure 4 shows the spike of the simulation of Fig. 1 that forms at $x = 5.6825$. In the figure, we have translated x so that the spike is centered at $x = 0$. The figure displays $N^\alpha{}_\alpha$ as a function of x for the times $t = -12, -13, -14, -15$, and -16.

Figure 4 clearly shows a narrowing feature. However, to compare with the formulas of Sec. III, we must perform a different type of comparison. Using Eq. (44), we define the rescaled spatial coordinate w by

$$w \equiv \left(\frac{x}{x_m}\right)\exp\left[\frac{1}{Z_+}(Z_+ - Z_-)(t_1 - t)\right]. \tag{53}$$

Then aside from the detailed shape, Eqs. (44) and (36) contained the prediction that $N^\alpha{}_\alpha$ plotted as a function of w will give the *same* shape regardless of time. Figure 5 contains such a plot. Here, six different curves are plotted: the five curves of Fig. 5, but now as a function of w, and a sixth curve given parametrically by Eqs. (44) and (36). To obtain the parameters in the analytic formula, we find u from the simulations and choose $t_1 = -12$ and find x_m for that time. Figure 6 contains the corresponding six curves for the quantity $\mathcal{S}$. It is clear from Figs. 5 and 6 that the formulas of Sec. III are a good, but by no means perfect, match to the results of the simulation.

Figures 7–9 do the same thing for the simulation of Fig. 2 that Figs. 4–6 do for the simulation of Fig. 1. That is, in Fig. 7, one of the late spikes of the simulation of Fig. 2 is plotted as a function of x for five different times. In Fig. 8,

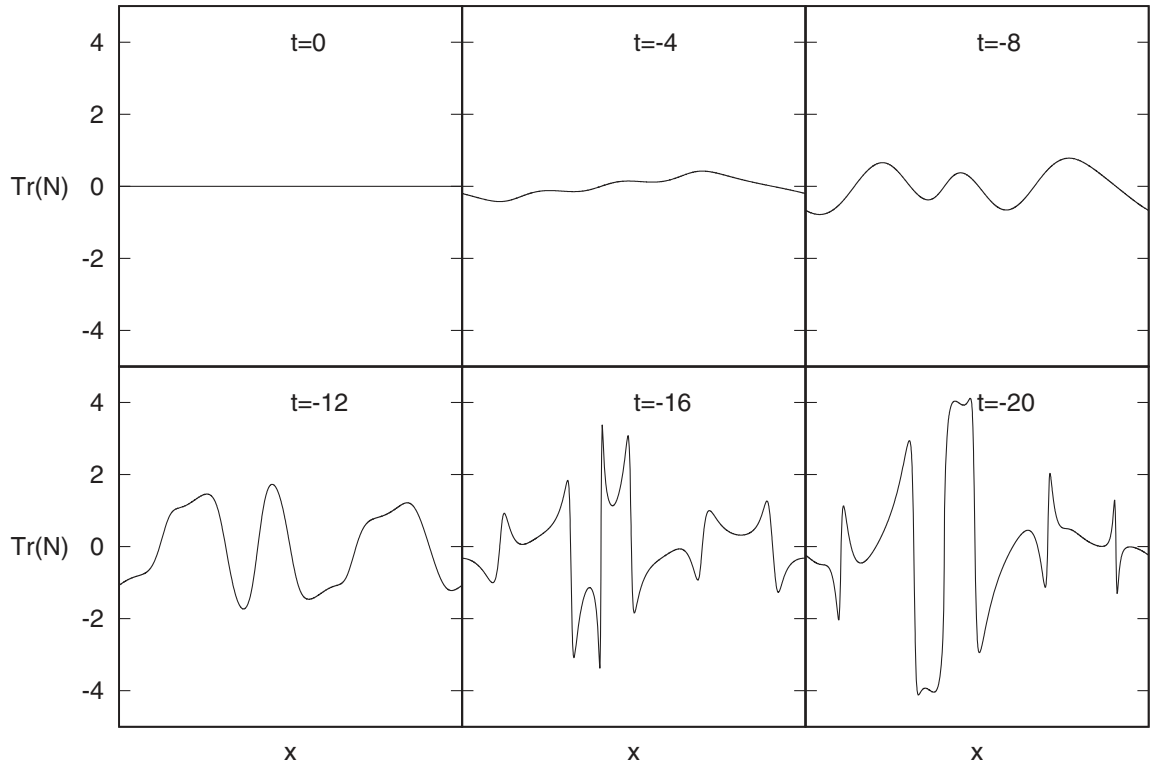

FIG. 1. $t =$ const snapshots of $N^{\alpha}{}_{\alpha}$ for $0 \leq x \leq 2\pi$ (with $x = 0$ and $x = 2\pi$ identified) for several different times. Here the parameters of the initial data (51) are $a_1 = 2.5$, $a_2 = 1.2$, $b_1 = 1.5$, $b_2 = 1.2$, $\kappa_1 = 0.4$, $\kappa_2 = 0.3$.

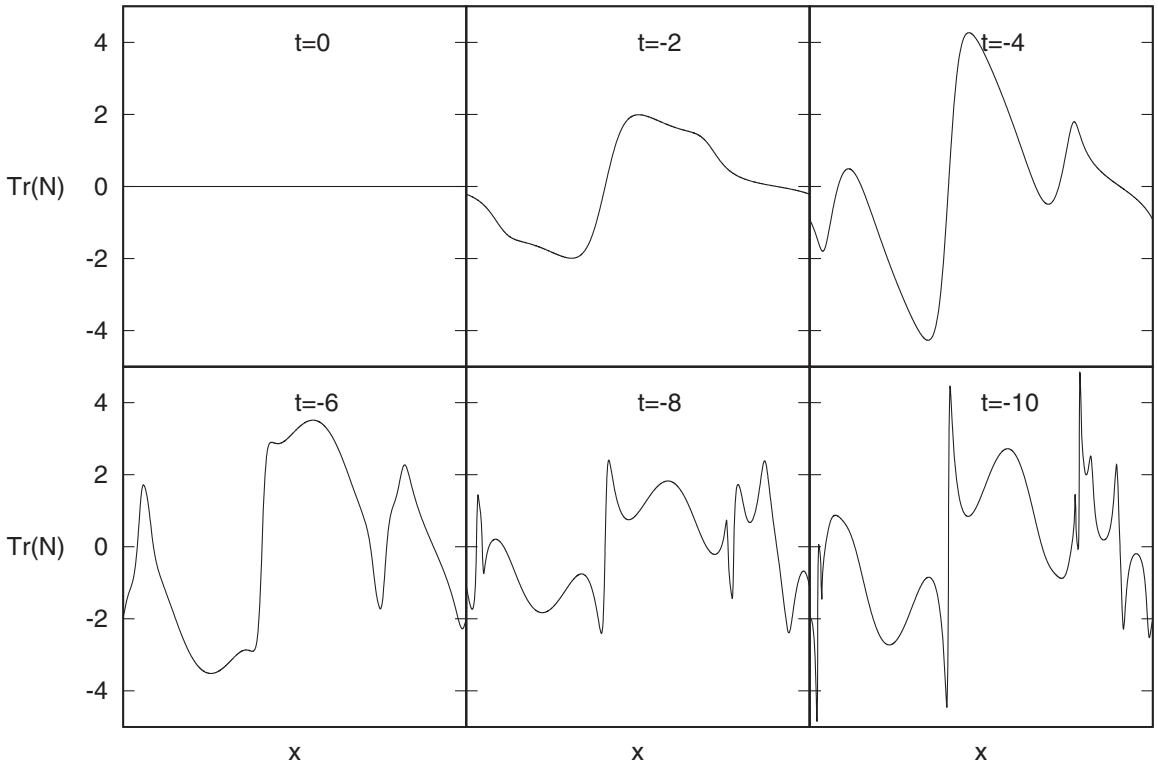

FIG. 2. $t =$ const snapshots of $N^{\alpha}{}_{\alpha}$ for $0 \leq x \leq 2\pi$ (with $x = 0$ and $x = 2\pi$ identified) for several different times. Here the parameters of the initial data (51) are $a_1 = 2.0$, $a_2 = 1.2$, $b_1 = 2.0$, $b_2 = -0.5$, $\kappa_1 = 0.2$, $\kappa_2 = 0.5$.

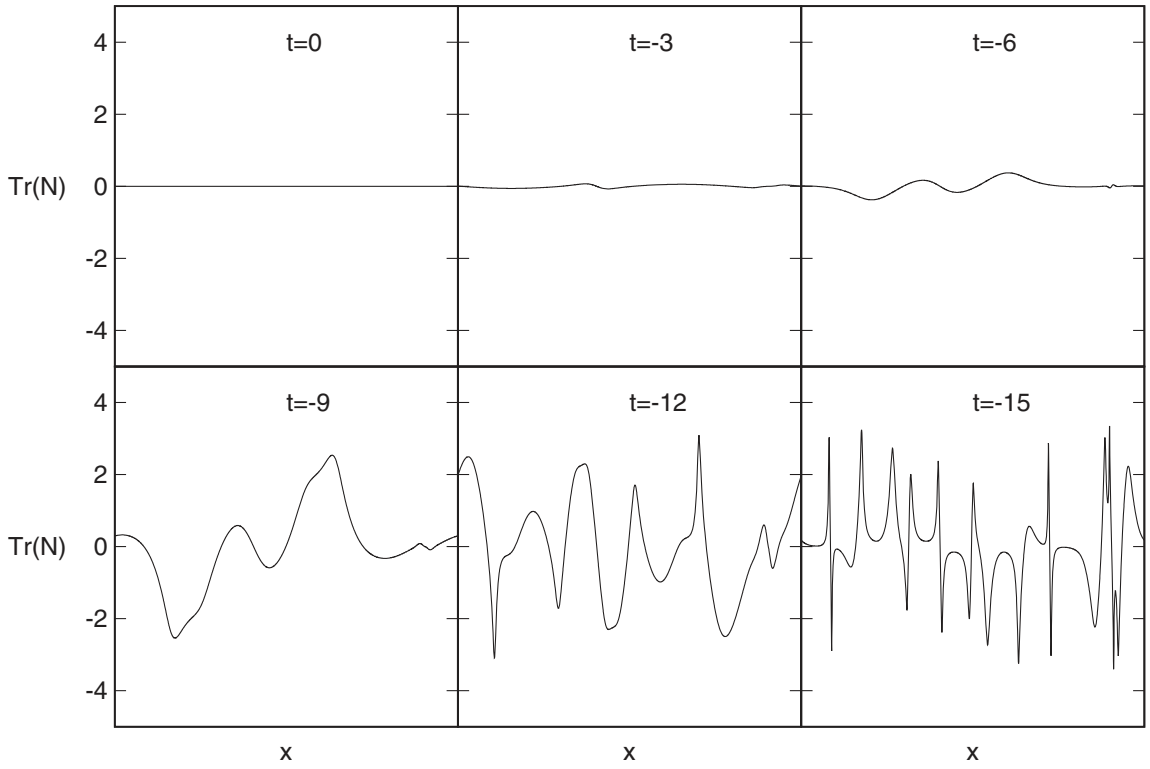

FIG. 3. $t =$ const snapshots of $N^{\alpha}{}_{\alpha}$ for $0 \leq x \leq 2\pi$ (with $x = 0$ and $x = 2\pi$ identified) for several different times. Here the parameters of the initial data (51) are $a_1 = 2.5$, $a_2 = 0.5$, $b_1 = 1.0$, $b_2 = -1.5$, $\kappa_1 = 0.6$, $\kappa_2 = 0.3$.

that same spike is plotted as a function of the rescaled coordinate w along with the corresponding formula from Sec. III. In Fig. 9, the quantity $\mathcal{S}$ for the 5 times is plotted as a function of w along with its formula. Correspondingly, Figs. 10–12 perform the same analysis of one of the late spikes of the simulation of Fig. 3.

In all cases, we find that the formulas of Sec. III are a good but not perfect fit for the results of the simulations. This is just what we expect from the analysis of that section, due to the fact that even the late spikes of our simulations are comparatively early in the sense of Sec. III.

V. 2D SIMULATIONS

The base of the code used for the 2D results is essentially identical to the 1D code, except now the fields can vary along two of the spatial dimensions x and y, and corresponding discretizations in the code are represented as 2D

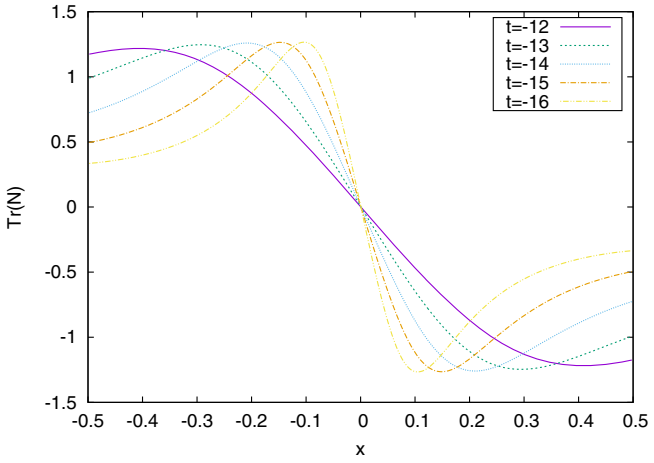

FIG. 4. $N^{\alpha}{}_{\alpha}$ versus x for $t = -12, -13, -14, -15$, and -16 for the spike located at $x = 5.6825$ from the evolution depicted in Fig. 1. Here we have translated x so that zero is the center of the spike.

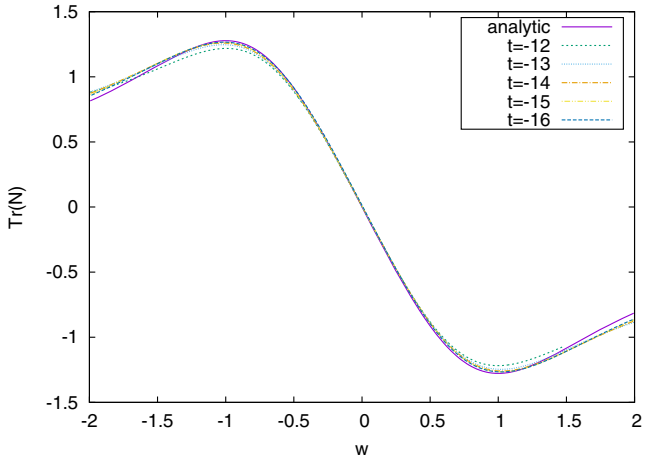

FIG. 5. $N^{\alpha}{}_{\alpha}$ versus the rescaled coordinate w for $t = -12, -13, -14, -15$, and -16 for the same data depicted in Fig. 4, along with the spike formula.

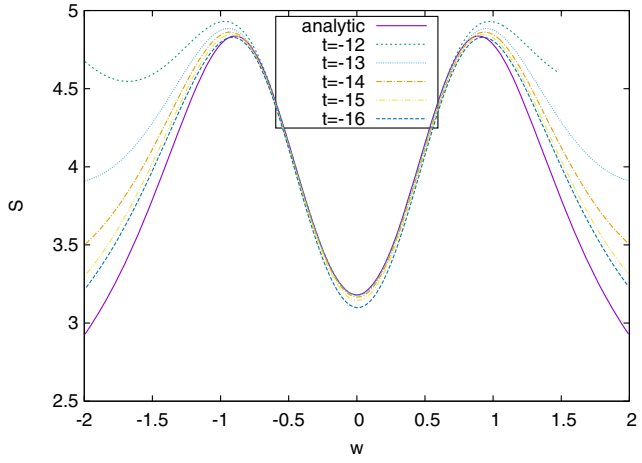

FIG. 6. $\mathcal{S}$ versus the rescaled coordinate w for $t = -12, -13, -14, -15$, and -16, along with the spike formula, for the spike at $x = 5.6825$ shown in Fig. 1.

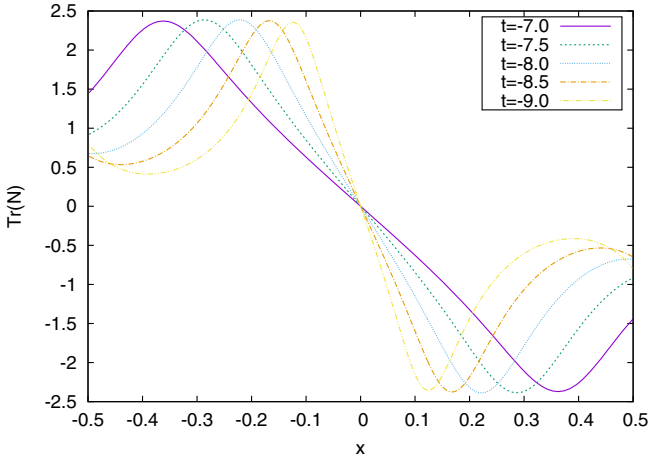

FIG. 7. $N^{\alpha}{}_{\alpha}$ versus x for $t = -7, -7.5, -8, -8.5$, and -9 for the spike located at $x = 5.683$ from the evolution depicted in Fig. 2. Here we have translated x so that zero is the center of the spike.

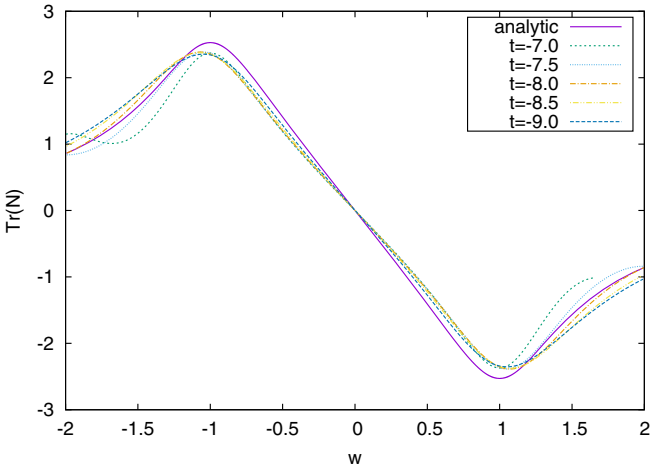

FIG. 8. $N^{\alpha}{}_{\alpha}$ versus the rescaled coordinate w for $t = -7, -7.5, -8, -8.5$, and -9 for the same data depicted in Fig. 7, along with the spike formula.

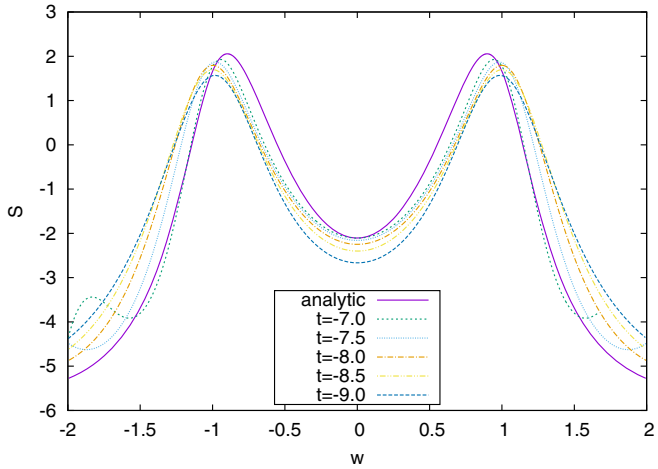

FIG. 9. $\mathcal{S}$ versus the rescaled coordinate w for $t = -7, -7.5, -8, -8.5$, and -9, along with the spike formula, for the spike at $x = 5.683$ depicted in Fig. 2.

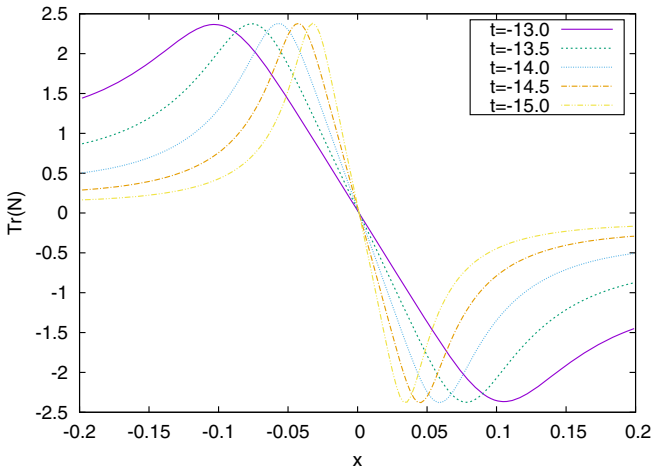

FIG. 10. $N^{\alpha}{}_{\alpha}$ versus x for $t = -13, -13.5, -14, -14.5$, and -15 for the spike located at $x = 2.5405$ from the evolution depicted in Fig. 3. Here we have translated x so that zero is the center of the spike.

arrays. We compactify on a torus, identifying $x = 0$ ($y = 0$) with $x = 2\pi$ ($y = 2\pi$). As mentioned, the PAMR/AMRD framework allows for adaptive mesh refinement; however, here the spikes are essentially volume filling (see Fig. 14), and little benefit is achieved compared to unigrid evolution; hence, all our runs are unigrid.

PAMR achieves parallelization via the standard domain decomposition approach: the full 2D spatial domain $x \in 0..2\pi, y \in 0..2\pi$ is broken up into $m \times n$ rectangles, where $N = m \times n$ is the number of processors in the run. Each processor solves the equations on one of these rectangles. The actual domain stored on a given processor is the corresponding $2\pi/m \times 2\pi/n$ piece of the full domain, plus (in this case) two additional *ghost cells* around each local boundary. For each iteration step of the numerical evolution, the ghost cells of all fields are updated from the

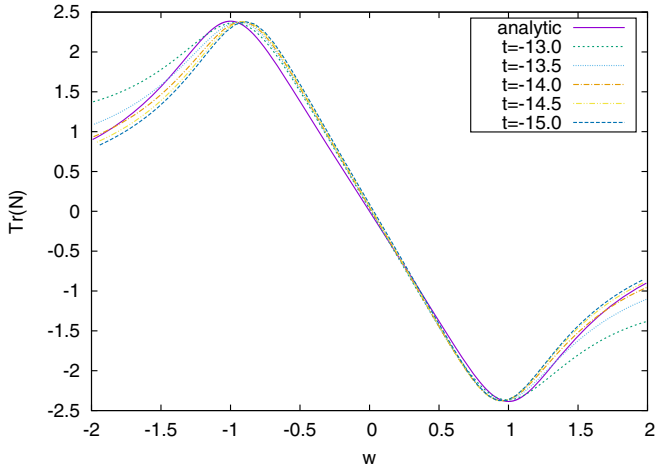

FIG. 11. ${N^{\alpha}}_{\alpha}$ versus the rescaled coordinate w for $t = -13$, -13.5, -14, -14.5, and -15, for the same data depicted in Fig. 10, along with the spike formula.

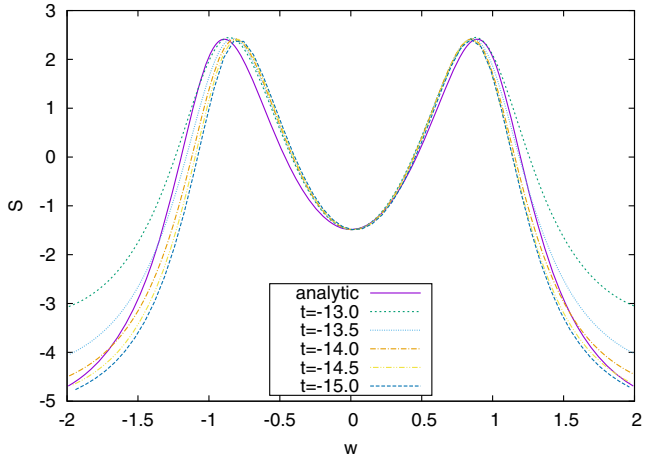

FIG. 12. $\mathcal{S}$ versus the rescaled coordinate w for $t = -13$, -13.5, -14, -14.5, and -15, along with the spike formula, for the spike at $x = 2.5405$ depicted in Fig. 3.

interior (nonghost) cells of adjacent processors, then the equations are solved at the interior points. The ghost cells thus allow standard, centered finite difference operators to be applied at all interior cells, and are also the vehicles whereby local boundary information is communicated from one processor to its neighbors, and vice versa. The hyperbolic equations are solved with a second order iterative Crank-Nicolson scheme, and the elliptic slicing condition is solved with a full approximation storage multigrid algorithm [19] using Gauss-Seidel relaxation as the smoother; both these methods are straightforwardly compatible with the domain decomposition algorithm just described.

The same initial data procedure is used for the 2D versus 1D code, though modifying the ansatz for Z_{ik} to

$$Z_{ik} = \begin{bmatrix} b_2 + a_y \cos(y + \phi_y) & \kappa_1 & \kappa_2 \\ \kappa_1 & a_1 \cos(x + \phi_x) + b_1 & 0 \\ \kappa_2 & 0 & -b_1 - b_2 - a_1 \cos(x + \phi_x) - a_y \cos(y + \phi_y) \end{bmatrix}. \tag{54}$$

Here, a_1, a_y, b_1, b_2, ϕ_x, ϕ_y, κ_1, and κ_2 are constants. The 2D simulations are computationally quite expensive compared to the 1D case, so here we only show results for a single set of initial data: $a_1 = 0.2$, $a_2 = 0.7$, $b_1 = 1.80$, $b_2 = -0.15$, $\phi_x = 0.15$, $\phi_y = 0.25$, $\kappa_1 = 0.5$, and $\kappa_2 = 0.3$. To check convergence, the above initial data were evolved with resolutions 192^2, 384^2, 768^2, 1536^2; see Fig. 13 for a plot of the norm of the constraints with time. The comparison figures shown below were obtained from the highest resolution data.

As discussed, the hypothesis is that spikes form along codimension one volumes of the spacetime where ${N^{\alpha}}_{\alpha} = 0$. For the 2D case then, this would correspond to lines within the (x, y) subspace, and the analytic approximation for the spike profiles should approximate the full (numerical) results on any slice orthogonal to a given point along the spike line. The parameters ϵ and b (see Sec. III) governing the spike profile can vary along the spike line. For a given point that we want to compare, we measure these parameters at one time within the simulation. We find that the extracted value for b, the quantity characterizing the

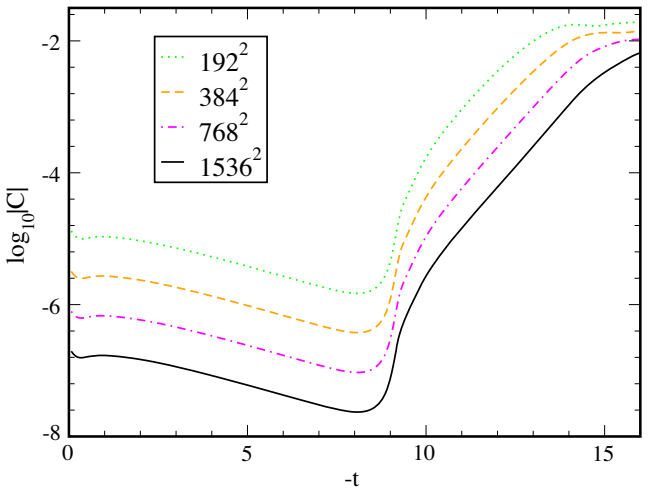

FIG. 13. An ℓ_2 norm of all terms in the constraint equations (12)–(15) over the computational domain, versus time, from four different resolution runs of the 2D case discussed in Sec. V. This shows close to second order convergence to zero for most of the run time; the drop in the rate toward the end is due to the spike regions becoming under-resolved.

geometry of the spike point (40), varies by a few percent depending on what time we choose to measure it; this is not unexpected, in particular because we only have the resolution to uncover the early time evolution of the spike, whereas the analytical formula should govern its late time behavior. The parameter ϵ sets the scale of the spike at a given time, so is more a function of the initial data than intrinsic to the spike geometry; thus, we set it to give a best fit to ${N^\alpha}_\alpha$ at the time b is measured.

In the 2D case, there is also more gauge ambiguity in performing the comparison than the 1D case; in particular, how to define "orthogonal" far from the spike line, as well as defining the coordinate measure w (53) along the spike. Here we simply define tangent/orthogonal to a spike line as measured in coordinate space (x, y), setting the overall scale (ϵ) for the orthogonal direction w at the time the spike parameters are measured, and then assuming the scale narrows with time as predicted by the analytic formula (i.e., we cannot distinguish between the differences in scale that arise with time from gauge effects vs limitations of the approximation).

Here we show a comparison of the numerical results versus analytic formulas along two slices of the simulation, as depicted in Fig. 14. Figure 15 shows ${N^\alpha}_\alpha$ and $\mathcal{S}$ orthogonal to a point on the spike line at $(x, y) = (\pi, \pi)$

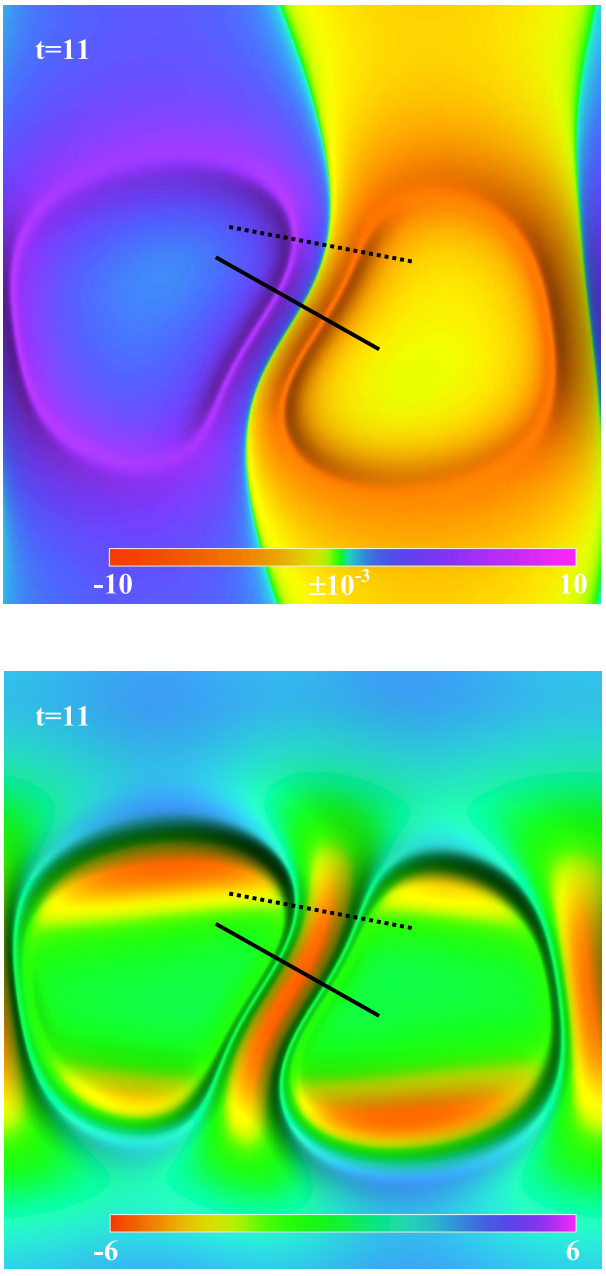

FIG. 14. Snapshots of ${N^\alpha}_\alpha$ (top) and $\mathcal{S}$ (bottom) at $|t| = 11$ of the 2D simulation. The solid (dashed) line illustrates the slice of the domain where the spike profiles centered at $(x, y) = (\pi, \pi)$ $[(x, y) = (3.37, 3.75)]$, depicted in Fig. 15 (Fig. 16) below, was measured. The width and height of each picture cover $x = [0, 2\pi]$ and $y = [0, 2\pi]$, respectively.

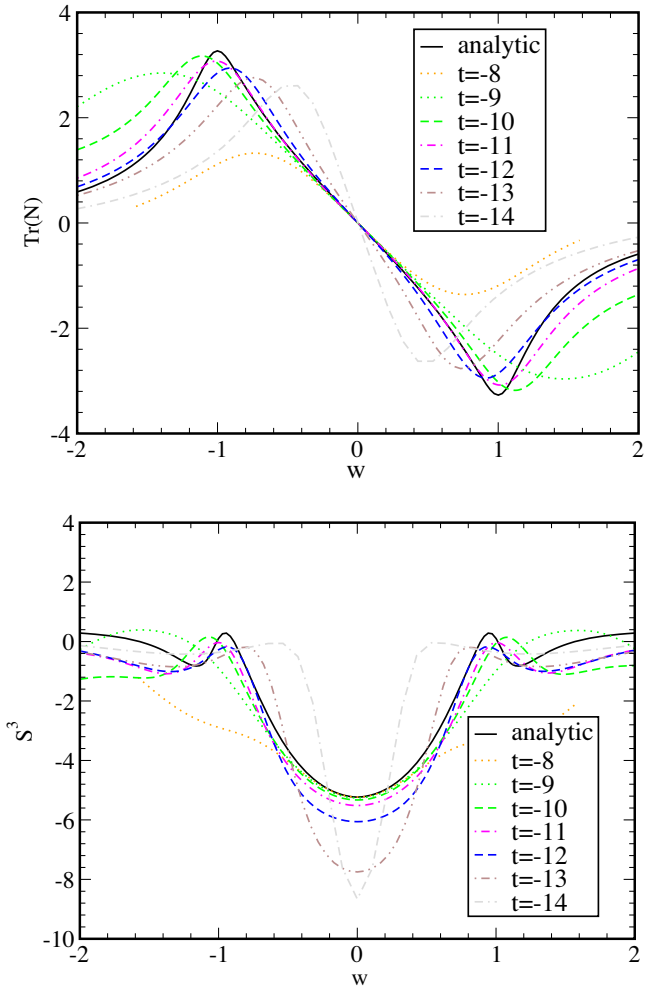

FIG. 15. ${N^\alpha}_\alpha$ (top) and $\mathcal{S}$ (bottom) measured along the slice of the 2D simulation orthogonal to the spike centered at $(x, y) = (\pi, \pi)$ (solid line in Fig. 14), at several times, together with the analytic approximations (for the latter, the spike parameter b was measured at $t = -8$ to be $b \sim 0.35$).

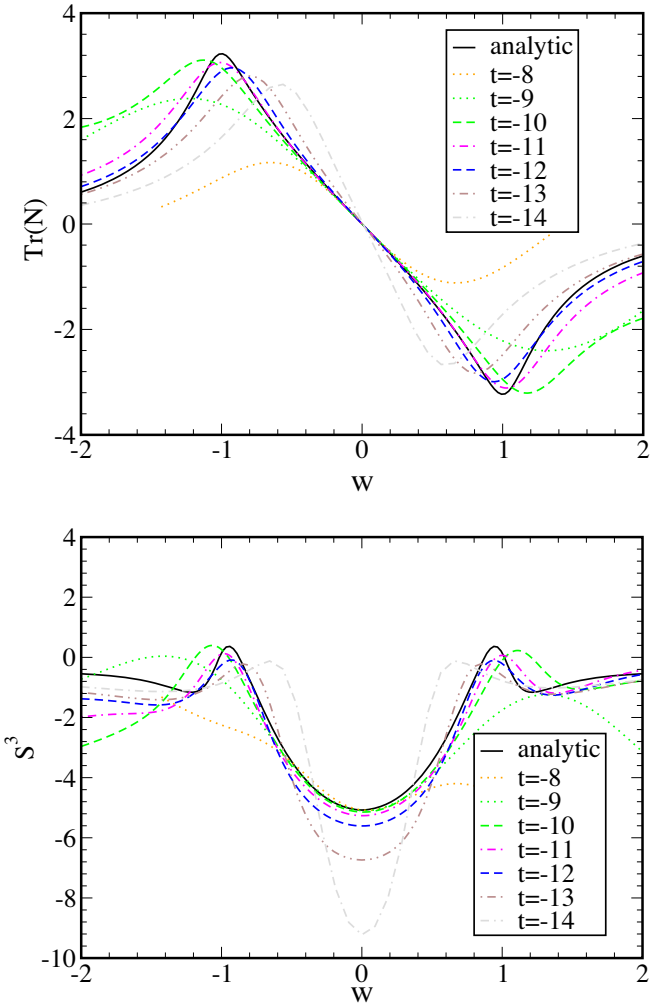

FIG. 16. $N^{\alpha}{}_{\alpha}$ (top) and $\mathcal{S}$ (bottom) measured along the slice of the 2D simulation orthogonal to the spike centered at $(x, y) = (3.37, 3.75)$ (dashed line in Fig. 14), at several times, together with the analytic approximations (for the latter, the spike parameter b was measured at $t = -8$ to be $b \sim 0.32$).

and Fig. 16 for that at $(x, y) = (3.37, 3.75)$. The results for the 2D runs are thus qualitatively consistent with that demonstrated for the 1D case: the formulas show decent agreement at intermediate times of the runs (late enough that a spike has clearly formed, but not so late that the spike has become under-resolved).

VI. CONCLUSIONS

BKL dynamics consists of a sequence of bounces in the approach to the singularity. When spikes were first found in the simulations of [3], they seemed like a mysterious exception to the behavior of the rest of the spacetime. Instead, we see that spikes are a straightforward consequence of BKL behavior. Each bounce is driven by growth in $N^{\alpha}{}_{\alpha}$. But in general $N^{\alpha}{}_{\alpha}$ vanishes on surfaces of codimension one. Points on that surface do not bounce, while nearby points do, leading to an ever narrower feature: the spike. This qualitative picture gives rise to a quantitative description encapsulated in the formulas of Sec. III for the behavior of the invariants of $N_{\alpha\beta}$ and $\Sigma_{\alpha\beta}$ as functions of transverse distance from the spike.

Spikes are a significant challenge for numerics, due to the need to resolve small scale features at so many points as to make adaptive mesh refinement impractical. This places severe limitations on the amount of time for which such a simulation can be run. However, the BKL approximation itself (and its consequences like the spike formulas) gets better the closer the singularity is approached, and thus the longer the simulation is run. The simulations of this paper are a compromise between these two stringent requirements: long enough to come within the regime of validity of the BKL approximation, but short enough that resolution is not overwhelmed.

Within this uneasy compromise, we find compelling evidence for the picture of Sec. III. That is, the simulations match the formulas of that section as well as can be expected. This characterization of spikes completes the numerical evidence that BKL behavior describes the approach to the singularity in spacetimes with compact Cauchy surfaces.

ACKNOWLEDGMENTS

This work was supported in part by NSF Grants No. PHY-1505565 and No. PHY-1806219 (D. G.), and NSF Grants No. PHY-1607449 and No. PHY-1912171, the Simons Foundation, and the Canadian Institute for Advanced Research (F. P.). Computational resources were provided by the Perseus cluster at Princeton University.

[1] R. Penrose, Phys. Rev. Lett. **14**, 57 (1965).
[2] V. Belinskii, E. Lifschitz, and E. Khalatnikov, Adv. Phys. **19**, 525 (1970).
[3] B. Berger and V. Moncrief, Phys. Rev. D **48**, 4676 (1993).
[4] B. Berger and D. Garfinkle, Phys. Rev. D **57**, 4767 (1998).
[5] W. C. Lim, Classical Quantum Gravity **25**, 045014 (2008).
[6] W. C. Lim, L. Andersson, D. Garfinkle, and F. Pretorius, Phys. Rev. D **79**, 123526 (2009).
[7] B. Berger and V. Moncrief, Phys. Rev. D **58**, 064023 (1998).

[8] B. K. Berger, D. Garfinkle, J. Isenberg, V. Moncrief, and M. Weaver, Mod. Phys. Lett. **13A**, 1565 (1998).
[9] C. Uggla, H. van Elst, J. Wainwright, and G. F. R. Ellis, Phys. Rev. D **68**, 103502 (2003).
[10] D. Garfinkle, Phys. Rev. Lett. **93**, 161101 (2004).
[11] M. J. Berger and J. Oliger, J. Comput. Phys. **53**, 484 (1984).
[12] S. Hern and J. Stewart, Classical Quantum Gravity **15**, 1581 (1998).
[13] PAMR (Parallel Adaptive Mesh Refinement) and AMRD (Adaptive Mesh Refinement Driver) libraries (http://laplace.physics.ubc.ca/Group/Software.html).
[14] D. Garfinkle, Classical Quantum Gravity **24**, S295 (2007).
[15] D. Garfinkle, W. C. Lim, F. Pretorius, and P. J. Steinhardt, Phys. Rev. D **78**, 083537 (2008).
[16] A. Ijjas, W. G. Cook, F. Pretorius, P. J. Steinhardt, and E. Y. Davies, J. Cosmol. Astropart. Phys. 08 (2020) 030.
[17] D. Garfinkle and C. Gundlach, Classical Quantum Gravity **22**, 2679 (2005).
[18] J. W. York, Phys. Rev. Lett. **26**, 1656 (1971).
[19] A. Brandt, Math. Comput. **31**, 333 (1977).