Gravitational collapse and topology change in spherically symmetric dynamical systems

Péter Csizmadia and István Rácz

RMKI H-1121 Budapest, Konkoly Thege Miklós út 29-33, Hungary

E-mail: cspeter@rmki.kfki.hu and iracz@rmki.kfki.hu

Received 12 August 2009, in final form 6 November 2009
Published 15 December 2009
Online at stacks.iop.org/CQG/27/015001

Abstract

A new numerical framework, based on the use of a simple first-order strongly hyperbolic evolution equations, is introduced and tested in the case of four-dimensional spherically symmetric gravitating systems. The analytic setup is chosen such that our numerical method is capable of following the time evolution even after the appearance of trapped surfaces, more importantly, until the true physical singularities are reached. Using this framework, the gravitational collapse of various gravity-matter systems is investigated, with particular attention to the evolution in trapped regions. It is verified that, in advance of the formation of these curvature singularities, trapped regions develop in all cases, thereby supporting the validity of the weak cosmic censor hypothesis of Penrose. Various upper bounds on the rate of blow-up of the Ricci and Kretschmann scalars and the Misner–Sharp mass are provided. In spite of the unboundedness of the Ricci scalar, the Einstein–Hilbert action was found to remain finite in all the investigated cases. In addition, important conceptual issues related to the phenomenon of topology changes are discussed.

PACS numbers: 04.20.Dw, 04.20.Ex, 04.20.Gz, 04.25.dc, 04.25.dg

1. Introduction

The diffeomorphism invariance of Einstein’s theory of gravity is intimately related with the fact that there is a significant redundancy in the representation of the true physical degrees of freedom. It can therefore be a great challenge to carry out a faithful investigation of dynamical processes even in the case of spherically symmetric spacetimes, in spite of the simplifications offered by the symmetries. Correspondingly, the selection of the most appropriate variables, done by applying a suitable gauge fixing, i.e., the most effective framework to carry out the study of a given dynamical system, is considered to be something of an art.

In the numerical investigations of spherically symmetric dynamical systems, the method of Choptuik—which he first applied in [6] while exploring the critical phenomenon in the gravitational collapse of various gravity-matter systems—turned out to be the most successful
in the sense that it is still widely used. However, Choptuik’s choice has both advantages and disadvantages. Perhaps the most important advantage is that one has to solve only two first-order partial differential equations (PDEs) for the basic metric variables\footnote{In this particular case, the basic metric variables are the smooth functions $A$ and $B$ in terms of which the spacetime metric can be given as $\text{d}s^2 = A \text{d}t^2 - B \text{d}r^2 - r^2 (\text{d}\theta^2 + \sin^2 \theta \text{d}\phi^2)$.}, which, along with the matter field equations, determine the full evolution of the associated gravity-matter system. On the other hand, the following objections may also be raised. First of all, the radial coordinate $r$ is chosen so that the area $A$ of the $SO(3)$-invariant 2-spheres are given as $A = 4\pi r^2$. As is well known, coordinate systems of this type are not suitable to follow evolution in regions where ‘trapped surfaces’ are formed because a coordinate singularity also develops (for a short discussion see, e.g., the last paragraph of \cite{34}).

Another objection is that one of the first-order PDEs is hyperbolic while the other is elliptic. This means that one of the metric equations—which is a constraint equation—has to be repeatedly integrated on successive time level surfaces. This process slows down time integration and makes it hard to apply the powerful tool of adaptive mesh refinement (AMR) which ensures high numerical accuracy in strongly dynamical processes \cite{3, 12, 13}. There are two side-remarks in order. First of all, although it is possible to implement a variant of AMR for the numerical integration of mixed hyperbolic and elliptic equations \cite{33}, the precision is partly lost because the applicability of AMR necessitates the extrapolation of some variables in time, something which is better to avoid in a time integration process. Therefore, the use of a fully hyperbolic system is preferable in the case of numerical integration of the field equations based on a finite difference schema. Secondly, it was shown by one of the present authors in \cite{34} that by using the Kodama vector field as the time evolutions vector field, the mixed elliptic-hyperbolic system can be replaced by a fully hyperbolic one. However, we have found the inevitable formation of a coordinate singularity in numerical simulations, the appearance of which was anticipated in the discussion below equation (4.15) of \cite{34}. The above findings motivate the search for a better analytic set-up with a more appropriate choice of the basic variables.

Before proceeding and presenting our proposal for such a choice, we would like to mention that several attempts have also been made to use reduced versions of the ADM and BSSN \cite{1, 40} formalisms in spherical symmetry, see \cite{5, 22, 23}. These models have been yielded by the reduction of a more complicated evolutionary system; therefore they do not optimally fit the spherically symmetric setup. Either they apply coordinates that are not suitable to cover regions with trapped surfaces, as happens, e.g., in \cite{22, 23}, or they are simply too complicated—see, e.g., the basic set of field equations (9a)--(9f) and (10a)--(10c) in \cite{5}. They may also suffer from numerical instabilities at the origin because negative powers of $r$ appear in the evolutionary equations.

Therefore, it is of considerable interest to single out a simple and general enough framework within which time evolution can be investigated on the largest possible part of the physical spacetime up to the appearance of true geometrical singularities. To match this requirement, we shall start by choosing a fully hyperbolic evolutionary system that is automatically applicable to describe the evolution in trapped region(s). This choice should be such that the equations are free from the numerical instabilities that used to appear in the origin in spherically symmetric spacetimes.

An analytic framework fitting the above-outlined expectations may be chosen as follows. Based on the results of earlier investigations in spherically symmetric (see, e.g., \cite{9, 15, 29})
and also in generic (see, e.g., [31]) dynamical spacetimes, the metric of the four-dimensional spherically symmetric spacetime \((M, g_{ab})\) will be assumed to possess the form
\[
ds^2 = \alpha \beta^2 d\tau^2 - \alpha \, d\rho^2 - \beta^2 (d\delta^2 + \sin^2 \delta \, d\varphi^2),
\]
where the coordinates \(\tau\) and \(\rho\) label the points of the two-dimensional timelike surfaces transverse to the transitivity surfaces of the rotation group, and \(\alpha, \beta, r\) are smooth functions of \((\tau, \rho)\). Note that \(\rho\) generally differs from the area-radial coordinate. This condition, as we have already mentioned, is necessary to extend the domain of time evolution to include trapped surfaces, when they exist.

Using this geometrical framework, gravitational collapse has been investigated in some simple gravity-matter systems. In the simplest possible case of asymptotically flat configurations, the associated time evolution of the system is qualitatively expected \([8–10, 15, 16]\) to be as indicated in figure 1 where the event horizon\(^2\) and the apparent horizon are also shown. The latter is foliated by marginally trapped surfaces and it is represented by a curve connecting the two ends of the zigzag line depicting the singularity. A portion of matter either falls into the singularity or reaches future timelike or null infinity, \(i^+\) or \(\mathcal{I}^+\). The asymptotic structure of the spacetime is expected to approach that of the Schwarzschild solution as we get closer to \(i^+\) along the null generators of \(\mathcal{I}^+\) \([9]\). If the collapsing matter has no considerable radiative degrees of freedom, the mass of the developing black hole is expected to be close to the total mass on the initial data surface. Note that spherical symmetry makes it possible to define mass and energy inside an invariant metric sphere (see section 6). If the mass inside the marginally outer trapped surface is found to approach the total mass while moving outwards along the apparent horizon, then we have a strong indication that our spacetime grid covers the truly dynamical part of the collapse. Moreover, as it will be demonstrated in section 7, the lapse function \(\beta\) may always be chosen such that the Cauchy surfaces can get arbitrarily close to the singularity.

\(^2\) The event horizon is the boundary of the causal past of future null infinity, \(\mathcal{H} = \partial J^- [\mathcal{I}^+]\).
Figure 2. Topology changes of the time level surfaces are illustrated by the vertical figure sequences, with the suppression of one space dimension. Time progresses from bottom to top. Transitivity surfaces of the rotation group are represented by horizontal circles in the first two sequences (starting from $\mathbb{R}^3$ and $S^3$) and by vertical circles in the last two sequences ($\mathbb{R} \times S^2$ and $S^1 \times S^2$).

We were also interested in investigating the time evolution of more exotic initial data specifications. As in the standard Friedman–Robertson–Walker cosmological models—which are known to be spherically symmetrical around any of their spacetime events—there is a freedom in choosing the topology of the initial data surfaces.

The base manifold $M$ of the investigated spacetimes coincides with the future Cauchy development of some three-dimensional achronal hypersurface $\Sigma$. Thereby, $M = D^+[\Sigma]$ and it possesses the product space structure $\Sigma \times \mathbb{R}^+$ [25]. Since the spacetime is spherically symmetric, the $\tau = \text{constant}$ time level surfaces—which are diffeomorphic to $\Sigma$—can also be foliated by the transitivity surfaces of the rotation group. By virtue of the particular form of the applied line element (2), the metric induced on the transitivity surfaces of the rotation group is given as $ds^2|_{\mathbb{S}} = r^2(d\vartheta^2 + \sin^2 \vartheta d\phi^2)$. Thereby, the vanishing of $r$, which is, in fact, the area-radius function, is directly related to the existence of an origin. To see how many origins we might have, let us consider some simple choices for the topology of the initial data surface $\Sigma$. Whenever $\Sigma$ is a connected geodesically complete spacelike hypersurface possessing a trivial product bundle structure, its topology is either $\mathbb{R}^3 = [0, \infty) \times S^2$, $S^3 = [0, \pi] \times S^2$, $\mathbb{R} \times S^2$ or $S^1 \times S^2$ (see as an illustration the bottom line of figure 2). Accordingly, there may be one origin, two origins or no origin at all on our initial hypersurface. During time evolution, the geometrical properties of the time level surfaces may change such that new origins are produced, possibly indicating a change in the topology, as illustrated in figure 2.

At this point, it is important to clear up some of the related notions and potential misconceptions. While in the above, the terminology of topology change has been used, it should be kept in mind that it is the topology of the limit of time level surfaces—comprising the future boundary of the pertinent ‘domain of dependence’—that may differ from that of the time level surfaces. A careful approach is needed in applying notions such as domain of dependence, Cauchy development and Cauchy surface which were introduced in [25] by Geroch. We would like to emphasize first that there is a significant distinction between the
notions ‘initial data surface’ and ‘Cauchy surface’. We may choose an arbitrary achronal hypersurface \( \Sigma \) as an initial data surface. The initial data together with the field equations can be used to determine the solution everywhere in the domain of dependence \( D[\Sigma] \). As opposed to this, the use of the notion of ‘Cauchy surface’ tacitly requires additional knowledge about the global properties of the underlying spacetime. In particular, see theorem 11 of [25]: a spacetime is known to be globally hyperbolic if and only if it possesses a Cauchy surface.

The confusion might arise—and, in fact, apparently it does arise in many circumstances—as a consequence of the fact that the domain of dependence \( D[\Sigma] \) of any initial data surface \( \Sigma \) is a globally hyperbolic spacetime itself in its own right, i.e., \( \Sigma \) is a Cauchy surface for \( D[\Sigma] \). In many cases, it is obvious to recognize that there exists a globally hyperbolic global extension \( (M', g'_{ab}) \) of \( (D[\Sigma], g_{ab}) \) such that the boundary of \( D[\Sigma] \) is not empty in \( M' \), while the Cauchy horizon, \( H[\Sigma] = H^+[\Sigma] \cup H^-[\Sigma] \) of \( D[\Sigma] \), is obviously empty with respect to \( D[\Sigma] \). However, a generic method that could be applied in all the possible cases does not exist. Although the concept of ‘maximal Cauchy development’ was introduced in [7], it is based on Zorn’s lemma, which makes its use difficult in practice.

As a simple example, consider the maximal analytic extension of the Schwarzschild spacetime. In Kruskal coordinates \( (T, R, \vartheta, \varphi) \), the line element is given by

\[
\alpha = 32M^3 e^{-r/2GM},
\]

where the Schwarzschild coordinates \( r \) and \( t \) are determined by the following implicit relations (see, e.g., [44]):

\[
\left(1 - \frac{r}{2M}\right) e^{r/2GM} = T^2 - R^2,
\]

\[
\frac{t}{2M} = 2 \tanh^{-1} \left( \frac{T}{R} \right).
\]

By introducing suitable new coordinates \( \tau = \tau(T, R) \) and \( \rho = \rho(T, R) \), the line element of the Schwarzschild spacetime preserves the form (2). The time slicings—two of which are indicated in figure 3—and the corresponding lapse functions vary accordingly. In the first case (left panel), two points on the \( \Sigma_{\tau}^* \) hypersurface hit the \( r = 0 \) Schwarzschild singularity. The appearance of these points can be interpreted as the formation of two origins since \( r > 0 \) in all other points of \( \Sigma_{\tau}^* \). Correspondingly, the topology \( \mathbb{R} \times S^2 \) of \( \Sigma_{\tau} \) is replaced by the disjoint union of \( \mathbb{R}^3 = \mathbb{R}^+ \times S^2, \mathbb{S}^3 \) and \( \mathbb{R}^3 \) at \( \tau^* \) (see the third column of figure 2). It is not hard to modify \( \Sigma_{\tau} \) by inserting suitably located further ‘steps’ such that the number of \( \mathbb{S}^3 \) components of the pertinent \( \Sigma_{\tau} \) may take an arbitrary integer value.

Although the time level surfaces on the left panel do not cover the future domain of dependence, \( D^+[\Sigma_{\tau}] \), of \( \Sigma_{\tau} \), it is worthwhile to keep in mind that \( \Sigma_{\tau} \) is a Cauchy surface for the Schwarzschild spacetime. As indicated on the right panel of figure 3, the entire of \( D^+[\Sigma_{\tau}] \) can be covered by time level surfaces if the new coordinates \( \tau \) and \( \rho \), and in turn, the lapse \( \beta \) are chosen properly. There is no topology change and the section of the Cauchy surfaces within the black hole region uniformly converges to the \( r = 0 \) singularity. In section 7, the application of this type of dynamically determined \( \beta \) will be demonstrated via various examples.

3 For the precise notion of ‘local’ and ‘global’ extensions of spacetimes see, e.g., [36, 37].
We would like to mention that the type of topology changes indicated by the left panel of figure 3 showed up in many of our numerical simulations. We have found that, as a consequence of the developing inhomogeneities, certain parts of the time level surfaces may get closer to the singularity much faster than others. Nevertheless, as shown by the above example, the change of the topology refers to the limit of the applied time level surfaces rather than the underlying spacetimes. As it will be demonstrated in the succeeding sections, the $\beta$ function can be chosen such that the time evolution is slowed down in spacetime regions close to the singularity. This way we were able to enlarge the domain of dependences significantly and to get arbitrarily close to the developing spacetime singularities in our simulations. As a consequence, we were able to investigate the rate of blowing up of the curvature while approaching the singularity.

The structure of this paper is the following. In section 2, with the help of a simple-minded approach, we demonstrate the main technical issues in solving the Einstein-scalar field equations numerically. These difficulties are resolved in section 3, where the Misner–Sharp mass is introduced as an auxiliary variable to stabilize the applied numerical representation near the origin. Here we also describe the ‘trick’ which makes our framework capable of following the time evolution within the trapped region. Section 4 outlines the analytic procedure that guarantees the required stability of the time evolution in the neighbourhood of an origin. In section 5, a method is presented which provides the opportunity to monitor the energy transfer processes. Trapped surfaces, trapped regions and apparent horizons are defined in section 6. Investigations of concrete dynamical systems, in simple and in certain less obvious circumstances, are presented in section 7. The results of the convergence and accuracy tests are also presented in this section. The paper is closed by summarizing our results and by providing a short discussion concerning the interpretation and some of the immediate consequences of our findings.

2. A simple-minded approach

To have a truly dynamical spacetime, some matter fields also have to be specified in addition to the already introduced geometrical framework. For simplicity, we shall restrict our
considerations to a real self-interacting scalar field with Lagrangian
\[ \mathcal{L} = \frac{1}{2} g^{ij} \nabla_i \psi \nabla_j \psi - V, \]
where the potential possesses the form relevant to the Klein–Gordon field, \( V(\psi) = \frac{1}{2} \mu^2 \psi^2 \) with the mass parameter \( \mu \geq 0 \). Note, however, that \( V \) could be chosen to be any sufficiently regular function of the \( \psi \) field.

The field equations relevant for the gravity-matter system can be given as
\[ E_{ab} = R_{ab} - \frac{1}{2} g_{ab} R - 8 \pi T_{ab} = 0, \]
\[ \nabla_e \nabla_e \psi + \frac{\partial V}{\partial \psi} = 0, \]
where the energy-momentum tensor reads as
\[ T_{ab} = \nabla_a \psi \nabla_b \psi - g_{ab} \left[ \frac{1}{2} g^{ef} \nabla_e \psi \nabla_f \psi - V \right]. \]

It is straightforward to verify, by a direct calculation, that our basic variables \( \alpha, \beta, r \) and \( \psi \) satisfy the evolution equations
\[ \partial_\tau \alpha = \beta^2 \partial_\rho \alpha - 8 \pi \alpha \left( \psi^2 - \beta^2 \psi^2 \right) + \frac{\alpha^2 - \beta^2 \alpha^2}{\alpha} + 2 \alpha \beta^2 + r^2 - \beta^2 r^2 \rho \beta - 2 \alpha \beta^2 \rho \beta, \]
\[ \partial_\tau r = \beta^2 \partial_\rho r + 8 \pi r \alpha^2 V - \frac{\alpha^2 + r^2 - \beta^2 r^2 \rho}{\beta} + \beta^2 r^2, \]
\[ \partial_\tau \psi = \beta^2 \partial_\rho \psi - \frac{2}{r} \psi^2 r^2 - \beta^2 \psi^2 r^2 + \frac{\beta^2 \psi^2 r^2}{\beta} - \alpha \beta^2 \frac{\partial V}{\partial \psi}, \]
where the shorthand notation \( f_\sigma = \partial_\sigma f \) is applied (for \( f = \alpha, \beta, r \) and \( \psi \)). Similarly, the constraint equations are
\[ E^\tau = 0, \quad E^\rho = 0, \]
where
\[ E^\tau = \frac{\partial_\tau r^2}{\alpha^2 \beta^2 r} + \frac{\partial_\tau \psi}{\alpha^2 \beta^2 r} = \frac{2 \alpha \beta^2 (\partial_\rho r)^2 - r^2 \alpha^2 r^2 - \beta^2 r^2 \alpha^2}{\alpha^2 \beta^2 r} + 8 \pi \frac{\psi \psi}{\alpha \beta^2}, \]
\[ E^r = \frac{2 \alpha \beta^2 (\partial_\rho r)^2 - r^2 \alpha^2 r^2 - \beta^2 r^2 \alpha^2}{\alpha^2 \beta^2 r} + 4 \pi \frac{\psi \psi^2 + \beta^2 \psi^2 + 2 \alpha \beta^2 V}{\alpha \beta^2}. \]

The number of independent equations is in accordance with the fact that the evolution equation (13) for \( \psi \) is simply the actual form of (9), and that in a generic four-dimensional spherically symmetric spacetime there may only be four independent equations derived from Einstein’s equation (8). In particular, equations (11) and (12) can be seen to be equivalent to the combinations \( E^\rho + E^\psi - E^\tau = E^\rho = 0 \) and \( E^\tau + E^\rho = 0 \), whereas the constraints are \( E^\rho = 0 \) and \( E^\tau = 0 \).

Note that we have no evolution equation for \( \beta \). According to the terminology introduced by Friedrich in [21], this function can be seen to play a role analogous to the ‘gauge source
It is freely specifiable, provided that it is sufficiently regular on the domain of time evolution. By using this freedom, we are able to slow down evolution in our numerical simulations at the parts of the time level surfaces whenever they have got close to the singularity.

Defining the coordinate speed of light, \( c_{\tau,\rho} = \frac{d\rho}{d\tau} \) along the radial null geodesics in the \( \tau - \rho \) part—requiring the vanishing of the pertinent part of the line element—, it is straightforward to see that \( c_{\tau,\rho} = \beta \). Therefore, as long as \( \beta \) is guaranteed to be bounded so is \( c_{\tau,\rho} \). In our simulations, \( \beta \) was guaranteed to be less than or equal to 1.2 everywhere.

It is also remarkable that the system of evolution equations may be viewed as a system of coupled nonlinear wave equations for the variables \( \alpha, r \) and \( \psi \). These types of equations are known to possess a well-posed initial value problem (see, e.g., [11]). In addition, it can be checked by a direct calculation that whenever the evolution equations are satisfied—and, in turn, their derivatives also vanish—the following set of first-order strongly hyperbolic evolution equations can be derived:

\[
\begin{align*}
\partial_{\tau} E^\tau_{\rho} - \partial_{\rho} E^\tau_{\tau} + \frac{2}{r} \left[ r_{\tau} E^\tau_{\rho} - r_{\rho} E^\tau_{\tau} \right] + \frac{(\alpha_{\tau} \beta + \alpha \beta_{\tau}) E^\tau_{\rho} - (2\alpha_{\rho} \beta + \alpha \beta_{\rho}) E^\tau_{\tau}}{\alpha \beta} & = 0, \\
\partial_{\tau} E^\tau_{\tau} - \beta^2 \partial_{\rho} E^\tau_{\rho} + \frac{2}{r} \left[ r_{\tau} E^\tau_{\tau} - \beta^2 r_{\rho} E^\tau_{\rho} \right] + \frac{\alpha_{\tau} E^\tau_{\tau} - \beta (\alpha_{\rho} \beta + 3\alpha \beta_{\rho}) E^\tau_{\rho}}{\alpha} & = 0.
\end{align*}
\] (17) (18)

Since these equations are linear and homogeneous in the variables \( E^\tau_{\tau} \) and \( E^\tau_{\rho} \), they ensure that the constraints propagate with the evolution, and therefore it is enough to impose them only on the initial data hypersurface.

In spite of the attractive features of the above-outlined evolutionary system—see equations (11)–(13) and (14) along with (17)–(18)—it has been found to be unstable in numerical simulation because truncation errors grow rapidly near the origin(s), i.e., where \( r \) tends to zero. The instability is caused by terms of the form

\[
\frac{\alpha \beta^2 + r^2 - \beta^2 r^2}{r^n},
\] (19)

appearing in (12) with \( n = 1 \) and in (11) with \( n = 2 \). The numerator of (19) consists of three independent variables, each of which contributes to its numerical error. Although it should vanish rapidly while approaching the origin, it does not exactly do that in practice and, in turn, its slightest error is amplified enormously via the division by \( r^n \).

Increasing the resolution does not help in such a circumstance; the simulations crash even earlier. We tried to overcome this instability in various ways such as extrapolation from outer points, decreasing the order of finite differences near the origin, increasing the numerical dissipation term, and also by using different parametrizations of the metric which were supposed to behave better close to the critical points. Neither of these attempts turned out to be successful.

3. Stabilizing with the Misner–Sharp mass

By the inspection of the term (19) responsible for the above-mentioned instabilities and of the Misner–Sharp mass \( m \) [30] defined by the relation

\[
m = \frac{r}{2} (1 + g^{ab} \partial_a r \partial_b r) = \frac{r}{2} \left( \frac{\alpha \beta^2 + r^2 - \beta^2 r^2}{\alpha \beta^2} \right),
\] (20)

it seems to be rewarding to introduce \( m \) as an auxiliary variable. By deriving it with respect to \( \tau \) and \( \rho \) and by using the pertinent form of Einstein’s equations (8) to eliminate the second derivatives of \( r \), an evolution and a constraint equation can be deduced:
\frac{\partial \tau}{m} = \frac{r^2}{2} \left\{ r \tau E^\rho \rho - r \rho E^\rho \tau + 8\pi \left( r \tau T^\rho \rho - r \rho T^\rho \tau \right) \right\}, \quad (21)

\frac{\partial \rho}{m} = \frac{r^2}{2} \left\{ r \rho E^\tau \tau - r \tau E^\rho \rho + 8\pi \left( r \rho T^\tau \tau - r \tau T^\rho \rho \right) \right\}. \quad (22)

What makes the introduction of \( m \) even more preferable is related to the following observations. First of all, Einstein’s equations \( E_{ab} = 0 \) can be seen to hold whenever

\begin{align*}
\partial_\tau \alpha_{\tau} &= \beta^2 \partial_\rho \alpha_{\rho} + \frac{4m\beta^2 \alpha^2}{r^3} + 8\pi \alpha^2 \beta^2 \left( T^\rho \rho + T^\tau \tau - T^\rho \rho \right) \\
&\quad + \frac{\alpha^2 - \beta^2 \alpha^2}{\alpha} \beta + \beta_\tau \alpha_{\tau} + \beta^2 \beta_\rho \alpha_{\rho} + 2\alpha \beta \partial_\rho \beta_{\rho}, \quad (23)
\end{align*}

\begin{align*}
\partial_\tau r_{\tau} &= \beta^2 \partial_\rho r_{\rho} + 4\pi r \alpha \beta^2 \left( T^\tau \tau + T^\rho \rho \right) - \frac{2m\beta^2 \alpha}{r^2} + \frac{\beta \tau r_{\tau} + \beta^2 \beta_\rho r_{\rho}}{\beta}, \quad (24)
\end{align*}

\begin{align*}
\partial_\tau m &= 4\pi r^2 \left( r_{\tau} T^\rho \rho - r_{\rho} T^\tau \tau \right), \quad (25)
\end{align*}

\begin{align*}
\partial_\rho m &= 4\pi r^2 \left( r_{\rho} T^\tau \tau - r_{\tau} T^\rho \rho \right). \quad (26)
\end{align*}

It is important to emphasize that instead of solving the constraint equations \( E_{\tau \tau} = 0 \) and \( E_{\tau \rho} = 0 \) on the initial data surfaces, it suffices to solve (26), provided that the evolution equations (23)–(25) hold. To see this, recall first that (24) is equivalent to

\begin{align*}
E_{\tau \tau} + E_{\tau \rho} = 0. \quad (27)
\end{align*}

By virtue of (21), (22), (25) and (26), we have that

\begin{align*}
r_{\tau} E^\rho \rho - r_{\rho} E^\rho \tau = 0, \quad r_{\rho} E^\tau \tau - r_{\tau} E^\tau \rho = 0. \quad (28)
\end{align*}

Equations (27)–(28) along with the relation \( E^\rho \rho = -\beta^2 E^\rho \rho \) imply then that for the variables \( E_{\tau \tau}, E_{\tau \rho}, \) the homogeneous linear equations

\begin{align*}
r_{\tau} E^\tau \tau - \beta^2 r_{\rho} E^\tau \rho = 0, \quad r_{\rho} E^\tau \tau - r_{\tau} E^\tau \rho = 0 \quad (29)
\end{align*}

hold. The algebraic sub-determinant of this system is \( r^2 - \beta^2 \tau^2 \), which will be shown in section 6 to vanish only at isolated marginally trapped surfaces. Therefore the system (29) possesses only the trivial solution which justifies our claim.

Restricting our considerations to the case of a self-interacting scalar field again, the components of the energy–momentum tensor can be given as

\begin{align*}
T^\tau \tau &= \frac{1}{2} \psi^2 + \frac{\beta^2 \psi^2}{\alpha \beta^2} + V, \quad (30)
\end{align*}

\begin{align*}
T^\tau \rho &= -\frac{1}{\beta^2} T^\rho \tau = \frac{\psi_{\rho} \psi_\tau}{\alpha \beta^2}, \quad (31)
\end{align*}

\begin{align*}
T^\rho \rho &= -\frac{1}{2} \frac{\psi^2 + \beta^2 \psi^2}{\alpha \beta^2} + V, \quad (32)
\end{align*}

\begin{align*}
T^\rho \theta &= T^\theta \rho = -\frac{1}{2} \frac{\psi^2 - \beta^2 \psi^2}{\alpha \beta^2} + V. \quad (33)
\end{align*}

In this case, the complete set of evolution equations consists of (23)–(25) and (13).
Note that the introduction of $m$ makes the use of (23) superfluous since $\alpha$ can be determined from (20) once the fields $m$ and $r$ are known. However, this way we get a $0/0$ type formula for $\alpha$ at 'marginally trapped surfaces' (for their precise definition see section 6) and an instability in the numerical simulation, thereby making it impossible to study the evolution inside the black hole region. In order to overcome these difficulties, we kept (23) as an evolution equation for $\alpha$. Then, as it will be demonstrated in section 7 (see figure 7), we could guarantee that our numerical framework does possess the desired capabilities without the loss of numerical accuracy anywhere in the domain of dependence.

By using standard procedures (described in details, e.g., in [11]), a first-order evolutionary system for our spherically symmetric dynamical system can be deduced as follows. In addition to $m$, $\alpha$, $r$ and $\psi$, the first derivatives $\alpha_{\tau}$, $\alpha_{\rho}$, $r_{\tau}$, $r_{\rho}$, $\psi_{\tau}$ and $\psi_{\rho}$ are also considered as being independent variables, and the evolution equations are complemented by the relations

\[
\partial_{\tau} \alpha = \alpha_{\tau}, \quad (34)
\]
\[
\partial_{\tau} r = r_{\tau}, \quad (35)
\]
\[
\partial_{\tau} \psi = \psi_{\tau}, \quad (36)
\]
\[
\partial_{\tau} \alpha_{\rho} = \partial_{\rho} \alpha_{\tau}, \quad (37)
\]
\[
\partial_{\tau} r_{\rho} = \partial_{\rho} r_{\tau}, \quad (38)
\]
\[
\partial_{\tau} \psi_{\rho} = \partial_{\rho} \psi_{\tau}. \quad (39)
\]

Note that the last three relations are, in fact, the integrability conditions for $\alpha$, $r$, and $\psi$. As a byproduct of this reduction process, in addition to the true dynamical constraint (26), we also have to take care of the trivial ones

\[
\partial_{\rho} \alpha = \alpha_{\rho}, \quad (40)
\]
\[
\partial_{\rho} r = r_{\rho}, \quad (41)
\]
\[
\partial_{\rho} \psi = \psi_{\rho}. \quad (42)
\]

The yielded first-order system can then be put into the form

\[
\partial_{\tau} u = A \partial_{\rho} u + B, \quad (43)
\]

where the associated ten-dimensional vector variable $u$, the $10 \times 10$ matrix $A$ and the ten-dimensional source vector $B$ are given as

\[
u = \begin{pmatrix}
m \\
\alpha \\
\alpha_{\tau} \\
\alpha_{\rho} \\
r \\
r_{\tau} \\
r_{\rho} \\
\psi \\
\psi_{\tau} \\
\psi_{\rho}
\end{pmatrix}, \quad A = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},
\]

(44)
Since the eigenvectors of the matrix \( A \) comprise a complete system and its eigenvalues are all real, this first-order system is strongly hyperbolic \([26]\). It is known that both the analytic and numerical well posedness of the initial value problem is guaranteed for this type of evolution equations \([26, 38]\). An additional preferable property of the investigated dynamical system is that for the coordinate speed of light, \( c_{\tau, \rho} \leq \beta_{\text{max}} \) holds, where \( \beta_{\text{max}} \) denotes the maximum value of \( \beta \).

4. Stability in the origin(s)

The behaviour of the basic field variables \( m, \alpha, \beta, r, \psi \) close to an origin located, say, at \( \rho = \rho_0 \) can be explored by substituting the form of Taylor expansion

\[
 f(\tau, \rho) = \sum_{k=0}^{\ell} \frac{1}{k!} f_k(\tau)(\rho - \rho_0)^k + O_{\ell}( (\rho - \rho_0)^{\ell+1}),
\]

into the field equations for the functions \( m, \alpha, \beta, r, \psi \), and also by requiring the solutions to be at least of class \( C^\ell \) for some \( \ell \in \mathbb{N} \). Assuming that the solutions are at least of class \( C^4 \) in a neighbourhood of \( \rho_0 \), the resulting relations are

\[
 m = \frac{1}{6} m_3 (\rho - \rho_0)^3 + O_m((\rho - \rho_0)^6)
\]

\[
 \alpha = \alpha_0 + \frac{1}{2} \alpha_2 (\rho - \rho_0)^2 + \frac{1}{\pi} \alpha_4 (\rho - \rho_0)^4 + O_\alpha((\rho - \rho_0)^6)
\]

\[
 \beta = \beta_0 + \frac{1}{2} \beta_2 (\rho - \rho_0)^2 + \frac{1}{\pi} \beta_4 (\rho - \rho_0)^4 + O_\beta((\rho - \rho_0)^6)
\]

\[
 r = \text{sgn}(\rho - \rho_0) \sqrt{\alpha_0}(\rho - \rho_0) + \frac{1}{\pi} r_3 (\rho - \rho_0)^3 + O_r((\rho - \rho_0)^5)
\]

\[
 \psi = \psi_0 + \frac{1}{2} \psi_2 (\rho - \rho_0)^2 + \frac{1}{\pi} \psi_4 (\rho - \rho_0)^4 + O_\psi((\rho - \rho_0)^6).
\]

By using the indicated parity properties of the functions \( \alpha, \beta, r \) and \( \psi \), the grid boundary at the origin can be treated exactly the same way as it was done in \([19, 20]\). In particular, the method described in details in section D of \([20]\) makes it possible to determine all the spatial derivatives in a neighbourhood of an origin, by using a symmetric stencil.

Note that the factor \( \text{sgn}(\rho - \rho_0) \sqrt{\alpha_0} \) in the first-order term in \((50)\) has a simple geometrical explanation. In order to have a regular origin, i.e., to avoid the appearance of conical singularities, we have to guarantee that while approaching to the origin, \( \rho \to \rho_0 \), the ratio of

Note that defining \( \tilde{\alpha}_r = \beta \tilde{r}_r \), \( \tilde{\beta}_r = \beta \tilde{r}_r \), and \( \tilde{\psi}_r = \beta \tilde{\psi}_r \), the above set of field equations can be put into the form of a first-order symmetric hyperbolic system for the vector valued variable \((m, \alpha, \tilde{\alpha}_r, r, \tilde{\beta}_r, \psi, \tilde{\psi}_r, \tilde{\psi}_r)^T\).
the circumference of an infinitesimal, origin-centred circle and of the corresponding proper
radius tends to $2\pi$, i.e.,

\[
\frac{2\pi r}{\left| \int_{\rho_0}^{\rho} \sqrt{\alpha} \, d\rho \right|} \to \frac{2\pi \partial_\rho r}{\text{sgn}(\rho - \rho_0) \sqrt{\alpha}} \to 2\pi,
\]

(52)

where l'Hôpital’s rule is also applied.

Having the relations (47) and (50), it is also straightforward to see why the system
introduced in section 3 has been found stable in our numerical simulations. One of the reasons
for this unexpected stability is related to the fact that the term (19)—which was responsible
for the numerical instabilities in the case of the simple-minded system—reads as

\[
a\beta^2 + r^2 - \beta^2 r^2 \rho_n = \frac{2m \alpha^2}{3(\alpha+1)/2} |\rho - \rho_0|^{2-n}.
\]

(53)

Since $n$ takes the values 1 and 2, the term (19) is replaced by a completely regular one in our
new setup. It is also important to note that the Misner–Sharp mass $m$ is subject to a first-order
PDE (25) which does not involve a singular term at all.

If there is an origin, it is useful to have the limiting form of the source vector $B$ appearing
in (43). By using (47)–(51), it is straightforward to verify that the non-trivial components $B_\alpha$, $B_\tau$, and $B_\phi$, of $B$, as given in (45), tend to the following regular limits

\[
\lim_{\rho \to \rho_0} B_\alpha = \frac{3}{2} \left( \frac{\alpha_0}{\alpha} \right)^2 + \frac{2m \alpha^2}{\beta_0} - 2\text{sgn}(\rho - \rho_0) \sqrt{\alpha_0 \beta_0^2} \rho_3 + \alpha_2 \beta_0^2
\]

\[
+ 2\alpha_0 \beta_0 \beta_2 - 8\pi \alpha_0 (\partial_\tau \alpha_0)^2,
\]

(54)

\[
\lim_{\rho \to \rho_0} B_\tau = 0,
\]

(55)

\[
\lim_{\rho \to \rho_0} B_\phi = \frac{\alpha_0}{\alpha} \frac{\partial_\tau \phi_0}{\beta_0} + 2\beta_0^2 \psi_2 - \alpha_0 \beta_0 \frac{\partial V}{\partial \psi}.
\]

(56)

5. Energy balance

As it is well-known, energy transfer processes cannot be investigated in general relativity since
appropriate quasi-local quantities have not been found as yet. However, this can be done in
spherically symmetric dynamical spacetimes, using the Kodama vector field [29]

\[
K^a = \epsilon^{ae} \partial_e r,
\]

(57)

where $\epsilon^{ab} = q^{aj} \delta^b f_j$ and $\epsilon_{ab}$ denotes the volume form associated with the metric $q_{ab}$ induced
on the two-dimensional surface transverse to the $SO(3)$ group orbits. It was shown in [29]
that $K^a$ is divergence free, i.e., $\nabla^a K_a = 0$, and also that $G^{ef} \nabla_e K_f$ vanishes, where $G_{ab}$ denotes the Einstein tensor $R_{ab} - \frac{1}{2} R q_{ab}$. It follows then that the vector field

\[
J^a = T^a_b K^b
\]

(58)

is also divergence free. Accordingly, even though our spherically symmetric spacetime is
fully dynamical, $J^a$ behaves as a locally conserved energy flux-type vector field. Then, using
Stokes’ theorem and by choosing $Q$ to be a four-dimensional spacetime region with boundary
$\partial Q$, we have that

\[
\int_Q \nabla_a J^a \epsilon = \int_{\partial Q} n_a J^a \epsilon = 0,
\]

(59)
where \( \epsilon \) denotes the 4-volume element while \( \tilde{\epsilon} \) is the 3-volume element induced on the boundary \( \partial Q \) of \( Q \), which can be given as \( \tilde{\epsilon}_{abc} = \epsilon_{abc} n^c \), where \( n_a \) is the (outward pointing) unit normal 1-form field on \( \partial Q \).

With our particular choice of coordinates, the components of the Kodama vector field can be given as

\[
K^a = \begin{pmatrix} r_{\rho} & -r_{\tau} & 0 & 0 \end{pmatrix},
\]

which, in virtue of (58), implies that the non-zero components of the divergence-free vector field \( J^a \) read as

\[
J^\tau = \frac{1}{\alpha \beta} \left[ T^\pi r_{\rho} - T^\rho r_{\tau} \right] \quad \text{and} \quad J^\rho = \frac{1}{\alpha \beta} \left[ T^\rho r_{\rho} - T^\rho r_{\tau} \right].
\]

To characterize the energy transfer processes in a spacetime domain \( Q \) bounded by sections of \( \tau = \) constant and \( \rho = \) constant hypersurfaces, as it is depicted in figure 4, we introduce the following auxiliary notations.

Denote the origin-centred ball of radius \( \bar{\rho} \) on the \( \tau = \bar{\tau} \) hypersurface by \( B(\bar{\tau}, \bar{\rho}) \), and the part of the cylindrical hypersurface \( \rho = \bar{\rho} \) between \( \tau = \tau_1 \) and \( \tau_2 \) by \( C(\tau_2, \tau_1, \bar{\rho}) \). Then the boundaries \( \partial Q_{\tau i} \) \((i = 1, 2)\), as shown on figure 4, are represented by the differences \( \Delta B(\tau_i, \rho_2, \rho_1) = B(\tau_i, \rho_2) \setminus B(\tau_i, \rho_1) \) of the balls \( B(\tau_i, \rho_2) \) and \( B(\tau_i, \rho_1) \), while \( \partial Q_{\rho i} \) are the cylinders \( C(\tau_2, \tau_1, \rho_i) \) connecting the boundaries of the balls \( B(\tau_2, \rho_i) \) and \( B(\tau_1, \rho_i) \).

Then the energy contained in \( \Delta B(\tau, \rho_2, \rho_1) \) can be given as

\[
E(\bar{\tau}; \rho_2, \rho_1) = \int_{\Delta B(\tau, \rho_2, \rho_1)} n_a^{(\tau)} J^a \tilde{\epsilon}^{(\tau)},
\]

where \( \tilde{\epsilon}^{(\tau)} \) is the volume element on the \( \tau = \bar{\tau} \) hypersurface and \( n_a^{(\tau)} \) is its unit normal 1-form field, i.e.,

\[
n_a^{(\tau)} = \frac{\nabla_a \tau}{\sqrt{[\nabla_x \tau \cdot \nabla_x \tau]}} \bigg|_{\tau = \bar{\tau}}.
\]
Similarly, the energy transported through the cylindrical hypersurface \( C(\tau_2, \tau_1; \rho) \) is
\[
S(\tau_2, \tau_1; \rho) = \int_{C(\tau_2, \tau_1; \rho)} n^{(\rho)}_a J^a \tilde{\epsilon}^{(\rho)},
\]
where \( \tilde{\epsilon}^{(\rho)} \) is the volume element on the \( \rho = \bar{\rho} \) hypersurface and \( n^{(\rho)}_a \) is its unit normal 1-form field, i.e.,
\[
n^{(\rho)}_a = \frac{\nabla_a \rho}{\sqrt{\left(\nabla^\nu \rho \nabla_{\nu} \rho\right)}}|_{\rho = \bar{\rho}}.
\]
Using these notations, the energy balance equation (59) takes the form
\[
E(\tau_2; \rho_2, \rho_1) - E(\tau_1; \rho_2, \rho_1) + S(\tau_2, \tau_1; \rho_2) - S(\tau_2, \tau_1; \rho_1) = 0.
\]
The following relations can be verified by straightforward calculation:
\[
n^{(\tau)}_a = (\sqrt{\alpha} \beta, 0, 0, 0), \quad n^{(\rho)}_a = (0, \sqrt{\alpha}, 0, 0),
\]
\[
\tilde{\epsilon}^{(\tau)}_{\alpha \beta \gamma} = \sqrt{\alpha r^2} \sin \theta \cdot (d \rho)_\alpha \wedge (d \theta)_\beta \wedge (d \phi)_\gamma,
\]
\[
\tilde{\epsilon}^{(\rho)}_{\alpha \beta \gamma} = \sqrt{\alpha \beta r^2} \sin \theta \cdot (d \tau)_\alpha \wedge (d \theta)_\beta \wedge (d \phi)_\gamma.
\]
Thus \( E(\bar{\tau}; \rho_2, \rho_1) \) and \( S(\tau_2, \tau_1; \bar{\rho}) \) can be given as
\[
E(\bar{\tau}; \rho_2, \rho_1) = 4 \pi \int_{\rho_1}^{\rho_2} (\alpha \beta r^2 J^\tau)|_{\tau = \bar{\tau}} d\rho,
\]
\[
S(\tau_2, \tau_1; \bar{\rho}) = 4 \pi \int_{\tau_1}^{\tau_2} (\alpha \beta r^2 J^\rho)|_{\rho = \bar{\rho}} d\tau.
\]
It is important to mention here that, by virtue of the relations (26), (61) and (68)–(69), the Misner–Sharp mass function \( m \) can be given on a \( \tau = \bar{\tau} \) time level surface as
\[
m(\bar{\tau}, \rho) = E(\bar{\tau}; \rho, \rho_1 = 0) = m(\bar{\tau}, 0) + 4 \pi \int_0^\rho (\alpha \beta r^2 J^\tau)|_{\tau = \bar{\tau}} d\rho.
\]
Accordingly, by virtue of (68)–(69) and (72), the expression \( \epsilon_{GM} = \sqrt{\alpha \beta} J^\tau \) could also be interpreted as the combined energy density of our composed gravity-matter system.

6. Trapped surfaces

To define trapped surfaces, start with a smooth orientable two-dimensional compact manifold \( S \) with no boundary in a four-dimensional spacetime \((M, g_{ab})\). Let \( n^a \) be a smooth future-directed non-vanishing null vector field on \( S \) which is normal to \( S \), i.e., \( g_{ab} n^a X^b|_S = 0 \) for any vector field \( X^a \) tangent to \( S \). Consider then the null hypersurface \( N \) generated by geodesics starting on \( S \) with tangent \( n^a \). By parallel propagating \( n^a \) along these generators, it can be extended onto \( N \). Denote by \( u \) the synchronized affine parametrizations of these null geodesics. The hypersurface \( N \) is smooth in a neighbourhood of \( S \), and it is smoothly foliated by the \( u \)-level surfaces. Denote by \( \epsilon_q \) the volume element associated with the metric \( q_{ab} \) induced on these two-dimensional surfaces. Then the null expansion \( \theta^{(u)} \) with respect to \( n^a \) is defined as
\[
L_u \epsilon_q = \theta^{(u)} \epsilon_q,
\]
where \( L_u \) denotes the Lie derivative with respect to the null vector field \( n^a \).
It is straightforward to see that the sign of $\theta(n)$ remains intact under a positive rescaling of $na$, i.e., whenever $na$ is replaced by $n'a = fna$, where $f$ is a sufficiently smooth positive function on $\mathcal{I}$. Therefore, as far as the sign of $\theta(n)$ is concerned, $na$ and $n'a$ may be considered to be equivalent. It is well-known that there always exist two equivalence classes of future-directed non-vanishing null vector fields on $\mathcal{I}$ which are normal to $\mathcal{I}$. Let $na^+$ and $na^-$ be two such vector fields and denote the corresponding null expansions by $\theta^+$ and $\theta^-$.\footnote{Apart from the case of asymptotically flat spacetimes, the ‘+’ and ‘-’ signs have nothing to do with outward and inward pointing directions. Although these directions cannot be defined for generic spacetimes, it is worthwhile to mention that an adequate quasi-local notion of outward and inward direction can be introduced—without referring to global properties of the underlying spacetimes—provided that attention is restricted to untrapped or marginally trapped surfaces (see, e.g. \cite{35}).} According to the original definition of Penrose \cite{32}, a two-dimensional surface $\mathcal{I}$ in a four-dimensional spacetime is future trapped or untrapped if both of the future-directed null geodesic congruences orthogonal to $\mathcal{I}$ are converging, or one of them is diverging while the other is converging at $\mathcal{I}$, i.e., $\theta^\pm \leq 0$ or $\theta^+$ and $\theta^-$ have opposite signs. If one of the expansions vanishes identically while the other is everywhere non-positive, then $\mathcal{I}$ is called future marginally trapped surface. Past trapped and past marginally trapped surfaces can be defined analogously by reversing the time orientation applied tacitly in the above definitions.

Choose now $\mathcal{I}$ to be a 2-sphere of radius $r$, invariant under the $SO(3)$ symmetry of our spherically symmetric spacetime. In virtue of the $\rho - \tau$ part of the metric (2), an obvious choice for the non-vanishing future-directed null vector fields $na^+$ and $na^-$ on $\mathcal{I}$ is
\begin{equation}
na^\pm = \left(\frac{\partial}{\partial \tau}\right)^a \pm \beta \left(\frac{\partial}{\partial \rho}\right)^a,
\end{equation}
while the 2-volume element $\epsilon_q$ on the 2-spheres $\mathcal{I}_{\tau, \rho}$ foliating the $\mathcal{N}_\pm$ null hypersurfaces is
\begin{equation}
\epsilon_q = r^2 \sin \vartheta \cdot \epsilon,
\end{equation}
where $\epsilon$ denotes the Levi-Civita symbol. It is straightforward to check that the pertinent null expansions are
\begin{equation}
\theta^\pm = \frac{2}{r} \cdot (r \tau \pm \beta r \rho).
\end{equation}
Equations (20) and (76) imply that the following relation holds on $\mathcal{I}$:
\begin{equation}
1 - \frac{2m}{r} = \frac{\beta^2 r^2 - r_+^2}{4 \alpha \beta^2} = - \frac{r^2}{4 \alpha \beta^2} \cdot \theta^+ \cdot \theta^-.
\end{equation}
Since all the functions $\alpha$, $\beta$ and $r^2$ are positive, apart from origin(s) or the spacetime singularity, an $SO(3)$ invariant 2-sphere $\mathcal{I}$ of radius $r$ is untrapped or trapped (future or past) if and only if $1 - \frac{2m}{r}$ is positive or negative, respectively.

We shall call a spacetime region $\mathcal{T}$, future trapped if each point of it belongs to an $SO(3)$ invariant trapped 2-sphere. The past boundary, $\partial^+ \mathcal{T} = \mathcal{T} \cap \mathcal{I}^-[\mathcal{T}]$, of such a future trapped region will be referred as future apparent horizon $\mathcal{A}_+$.\footnote{It is important to keep in mind that whenever non-invariant topological 2-spheres are involved, the extent of the trapped region might be larger than the one determined by $SO(3)$-invariant 2-spheres \cite{2}.} The notion of past trapped region, $\mathcal{T}$, and past apparent horizon, $\mathcal{A}_-$, may be introduced analogously.
The time integration process uses the method of lines based on a fourth-order Runge-Kutta scheme for a detailed description of this method, for the case of fixed spatial resolutions, see, e.g., and as indicated by its name, adaptive mesh refinement techniques are also incorporated.

7.1. Initial data

Before determining the time evolution of our system using equation (43), we have to specify suitable initial data on a spacelike hypersurface \(\Sigma_0\)—which hereafter will always be chosen as the \(\tau = 0\) hypersurface—so that all the constraint equations are satisfied. Recall that the components of the vector variable \(u\) in (43) are \(m, \alpha, \alpha_\rho, \alpha_\tau, r, r_\rho, r_\tau, \psi, \psi_\rho, \text{ and } \psi_\tau\), and, as follows from the discussion in section 3, we apparently have only the constraint

\[
\partial_\rho m = \frac{2\pi r^2}{\alpha} \left[ r_\rho (\psi_\tau^2 + \beta^2 \psi_\rho^2 + 2\alpha \beta^2 V) - 2 r_\tau \psi_\tau \psi_\rho \right]
\]

beside the three trivial ones (40)–(42). However, the price to be paid for using the ‘superfluous’ evolution equation (23) is the appearance of an additional constraint (20), which implies that only two of the four variables, \(\alpha, \alpha_\tau, r\) and \(r_\tau\), may be chosen freely. To satisfy all constraints on the initial hypersurface, we have applied either of the following two selection procedures.

(A) Fixing of \(r\) and \(r_\tau\) from among the functions \(\alpha, \alpha_\tau, r\) and \(r_\tau\) on \(\Sigma_0\). Start by specifying the sufficiently regular but otherwise arbitrary functions \(r, r_\tau, \psi\) and \(\psi_\tau\), then choose \(r_\rho\) and \(\psi_\rho\) so as to satisfy the trivial constraints (41) and (42). By substituting the relation

\[
\alpha = \left(1 - \frac{2m}{r}\right)^{-1} \frac{\beta^2 r_\rho^2 - r_\tau^2}{\beta^2}
\]

into (78), the initial data for \(m\) can be determined numerically. Then the initial \(\alpha\) and \(\alpha_\rho\) are determined by (79) and the trivial constraint (40), respectively. Finally, \(\alpha_\tau\) is the \(\tau\)-derivative of (79),

\[
\alpha_\tau = 2 \left( \frac{r m_\tau - r_\tau}{r} \right) \left( \beta^2 r_\rho^2 - r_\tau^2 \right) + r (r - 2m) \left( \beta^2 r_\rho \cdot \partial_\rho r_\tau - r_\tau \cdot \partial_\tau r_\tau + r_\tau^2 \beta_\tau / \beta \right) / \beta^2 (r - 2m)^2
\]

(B) Fixing of \(\alpha, r, \psi\) and \(\psi_\tau\) on \(\Sigma_0\). Choose \(\alpha_\rho, r_\rho\) and \(\psi_\rho\) so as to satisfy the trivial constraints (40)–(42). From (20), we get

\[
r_\tau = \pm \beta \sqrt{r_\rho^2 - \alpha \left(1 - \frac{2m}{r}\right)},
\]

where the sign depends on the specific physical problem we intend to investigate. By substituting the right-hand side of (81) into (78), the yielded equation can be solved (numerically) for \(m\) on the initial data surface \(\Sigma_0\). Then the initial data for \(r_\tau\) are determined by the relation (81). Finally, \(\alpha_\tau\) can be determined similarly as in the case A. Procedure A is simpler but B provides a convenient control of both the temporal geometrical setup and the energy distribution of the matter field. The latter is also more appropriate to investigations of expanding cosmological models. Other choices may also be possible; however, the use of these two cases was found to be completely satisfactory. It is worth emphasizing that either of these initial data specifications is suitable to host all the possible Einstein-scalar field systems in the selected setup.

To summarize, we have free control only of the four dynamical variables \(r, r_\tau, \psi\) and \(\psi_\tau\) or \(\alpha, r, \psi\) and \(\psi_\tau\) in the two cases, respectively. In addition, the specification of \(\beta\) is also required. Once these functions are specified, they determine all other functions by either of the outlined procedures on \(\Sigma_0\). In our numerical simulations, we always used \(\beta = 1\) on \(\Sigma_0\).
7.2. Dynamical lapse function

As was emphasized earlier, the lapse function $\beta$ is freely specifiable. Also, as was indicated by the simple example in the introduction, the extent of the spacetime domain covered by the time level surfaces is influenced significantly by the chosen lapse function. To achieve the largest possible domain, we use a dynamically determined $\beta$ satisfying the evolution equation

$$\beta_{\tau} = -p \left( \frac{m r - 3 m r_{\tau}}{r^4} \right) \left( \frac{r_{\tau}}{r} \right)^3 \beta,$$

(82)

where $p$ is a positive real parameter, starting with the initial value $\beta|_{\Sigma_0} = 1$. The value of the parameter was chosen to be $p = 10^{-4}$ in all of our simulations.

Equation (82) was motivated as follows. To slow down the evolution close to the developing singularities, we need a function which tends to zero there rapidly enough. For this reason, first we applied $\beta = \exp(-pm/r^3)$. Besides decaying fast enough, this function is also regular at the origin since, in virtue of (47) and (50), $m/r^3$ remains finite while $\rho \to \rho_0$. The $\tau$-derivative of this $\beta$ is ‘almost’ (82), but without the factor $(r_{\tau}/r)^3$ on the right-hand side. The inclusion of this factor was found to be necessary in stabilizing the time evolution close to the singularity. Its use is supported by the observation that in a neighbourhood of a point where the function $r$ takes its minimum the factor $(r_{\tau}/r)^3$ gets to be sufficiently small providing thereby some additional slowing down to the evolution of $\beta$ there.

7.3. Gravitational collapse with spatial topology $\mathbb{R}^3 = [0, \infty) \times S^2$

Let us start by investigating the gravitational collapse of a massive real scalar field in the ‘conventional’ case, i.e., with the assumption that the initial data are specified on a hypersurface $\Sigma_0$ possessing the topology of $\mathbb{R}^3$. According to procedure A, we specify two metric functions on $\Sigma_0$ as

$$r(\rho) = \rho, \quad r_{\tau}(\rho) = 0.$$  

(83)

The scalar field $\psi$ is chosen to be a smooth hunch with compact support

$$\psi(\rho) = \begin{cases} \exp \left( d + \frac{\rho - a}{\rho - a - b} \right) & \text{if } |\rho - a| < b, \\ 0 & \text{if } |\rho - a| \geq b, \end{cases}$$

(84)

with vanishing $\tau$-derivative on $\Sigma_0$

$$\psi_{\tau}(\rho) = 0.$$  

(85)

The self-interaction potential is assumed to possess the form $V(\psi) = \frac{1}{2} \mu^2 \psi^2$. The parameter values are

$$a = 0, \quad b = 70, \quad c = 0.0795, \quad d = 100, \quad \mu = 0.8.$$  

(86)

In the numerical simulation, AMR was used with 7000 spatial points on the base grid and five refinement levels (refinement ratio 2).

To demonstrate the significance of the freedom in specifying $\beta$, we shall compare the result yielded in two cases, $\beta \equiv 1$ and the dynamical lapse determined by (82). The initial state is the same in these two cases, as shown in figure 5 where the initial matter and gravity-matter energy density distributions associated with a shell of radius $\rho$, $\rho_M = 4\pi r^2 \sqrt{\alpha} T_{\tau\tau}$ and $\rho_{GM} = 4\pi r^2 \alpha \beta J_{\tau}$, are plotted. The $\rho$ integrals of these quantities are, respectively, the matter and gravity-matter energy contained in an origin-centred ball of radius $\rho$ on the $\tau = \text{constant}$ time slices, $E_M(\tau, 0, \rho) = \int_0^\rho \rho_M \, d\rho$ and $E_{GM}(\tau, 0, \rho) = \int_0^\rho \rho_{GM} \, d\rho$. 

17
Figure 5. The matter and gravity-matter energy density distributions associated with a shell of radius $\rho$ on the initial data surface $\Sigma_0$.

Figure 6. The area radial coordinate function $r(\rho)$ is plotted on various time slices linearly (left) and logarithmically (right). A new origin can be seen to develop at $\rho_\ast \approx 2.1$ on both plots.

7.3.1. Time evolution with unit lapse. The time evolution of this initial state, with $\beta \equiv 1$, yields the gravitational collapse of the scalar field until a spacetime singularity is formed at $\tau_\ast \approx 17.583$. The change of the metric function $r$ during the pertinent time evolution is depicted in figure 6, where the appearance of the second origin is clearly visible at $\rho_\ast \approx 2.1$.

Figure 7 shows $\alpha, \alpha' = (r^2_\rho - r^2_\tau)/(1 - 2m/r), 1 - 2m/r$ and $r^2_\rho - r^2_\tau$ on an intermediate time level surface where trapped surfaces are already present. The value of $\alpha$ is calculated in two ways: by solving the evolution equation (23) and by using the algebraic relation (79). At the first sight, the two calculations yield approximately the same result. However, there is a tiny difference close to the marginally trapped surfaces which, in virtue of (76) and (77), are associated with the zeros of $1 - 2m/r$ and $r^2_\rho - r^2_\tau$ in the present case. This difference is enlarged by $\rho$-derivation, thereby $\alpha_\rho$ and $\alpha'_\rho$ differ significantly here. Whenever we started to
evolve the system using (79), the error became enormous in a couple of time steps, making it impossible to carry on with the time evolution in the trapped region. To reach long-term stability even in regions with trapped surfaces, it turned out to be essential to apply (23) as an evolution equation.

Figure 8 is to demonstrate that the combined gravity-matter energy is preserved with high precision during the entire history. More precisely, the energy balance relation (66), defined by using the Kodama vector field in the previous section, is found to hold during the entire evolution. The wiggly ends of the curves at $\tau \approx \tau_*$ are associated with the very appearance of the scalar curvature singularity at $(\rho_*, \tau_*)$. In the evaluation of the energy balance relation (66), the particular choices $\rho_1 = 0$ and $\rho_2 = 69$ have been made. Moreover, $E$ and $E_0$ are defined as

$$E = E(\tau; \rho_2, \rho_1) + S(\tau, 0; \rho_2), \quad E_0 = E(0; \rho_2, \rho_1).$$

(87)

Figure 7. Functions $r_2^2 - r_1^2$, $1 - 2m/r$, $\alpha$ and $\alpha'$ are plotted on a $\Sigma_1$ hypersurface. On the right, the corresponding difference $\Delta\alpha = \alpha - \alpha'$ is shown. Thin vertical broken lines indicate the location of marginally trapped surfaces. Although the error of $\alpha'$ appears to be small, the spikes on the right plot indicate that its $\rho$-derivative has large errors at the marginally trapped surfaces.

Figure 8. The validity of the energy balance relation (66) is shown for various spatial resolution during the gravitational collapse of a massive scalar field.
Note that doubling the number of grid points shifts the curve downward by about a factor $1/16$. Therefore figure 8 provides a verification that our numerical method is fourth-order accurate even with respect to $E$, which is a complicated function of the elementary field variables. Note that the experienced convergence order is in accordance with the applied fourth-order finite difference schemes, both for the spatial derivatives and for the time integration.

Figure 9 shows that the constraints (26) and (79) are preserved with high numerical accuracy even on the last time slice before reaching the singularity. The Misner–Sharp mass $m$, its derivative, $m_{\rho}$, and the $\alpha$ metric component are also plotted for comparison. Note that $m$ appears to blow-up at the point where the singularity is about to form, its maximum value reaches $m_{\text{max}} \sim 5 \times 10^2$ on the time slice shown. The error of the constraints (thick solid curves) also has peaks in this point.

7.3.2. Time evolution with dynamical lapse. In the case of dynamical lapse, the value of the Ricci scalar is of the order $10^{10}$ within the interval $3 \lesssim \rho \lesssim 17$ at $\tau = 29$. Constraint preservation on this time level surface is shown in figure 10.

In figure 11, the time evolution of the gravity-matter energy density distribution associated with a shell of radius $\rho$, $\mathcal{E}_{\text{GM}}$, is shown. This spacetime diagram depicts the variation of $\mathcal{E}_{\text{GM}}$ above the $\rho - \tau_c$ coordinate plane, where $\tau_c$ denotes the conformal time defined, along the constant $\rho$-coordinate lines, as $\tau_c = \int_0^{\tau} \beta \, d\tau$. The use of $\tau_c$ is advantageous since, as is indicated by figure 12, the radial null geodesics (see the thin dotted lines) are well approximated by $\tau_c \pm \rho = \text{constant}$ lines. On top and bottom, the orthogonal projections of the world sheet of $\mathcal{E}_{\text{GM}}$ can be seen. The curvy boundary represents the singularity. It is remarkable that $\mathcal{E}_{\text{GM}}$ is positive everywhere outside the apparent horizon. However,

---

7 Note that instead of the $\rho$-dependent values of the constraint violations, the $\rho$-dependence of their averaged values is shown, in order to smooth out the noise, in Figures 9–10. The averaging for each $\rho$ is done by using the interval $[\rho - 0.05, \rho + 0.05]$ around $\rho$.

8 To see this, recall that along radial null geodesics $\beta \, d\tau = \pm d\rho = 0$. This, along with $d\tau_c = \beta \, d\tau + \left[ \int_0^\tau \beta_c \, d\tau \right] \, d\rho$ and that the $\tau$-integral of $\beta_c$ is negligible, implies that in the $\rho - \tau_c$ coordinate plane they satisfy the differential equation $d\tau_c/\, d\rho = \mp 1 + \left[ \int_0^\tau \beta_c \, d\tau \right] \approx \mp 1$. 

---
Figure 10. Constraint preservation in the case of dynamical lapse.

Figure 11. Time evolution of the gravity-matter energy density distribution, $\delta_{GM}$. The black curve indicates the apparent horizon. (Colour and interactive versions of this and some other 3D figures can be found at http://www.kfki.hu/~cspeter/numrel/2009-ekg/index.html.)

Beyond this boundary, it starts to oscillate such that the amplitude of this oscillation is increasing towards the intersection of the singularity and the $\rho = 0$ line. It is important to note that in spite of this oscillatory behaviour of $\delta_{GM}$, the Misner–Sharp mass is always nonnegative.

In figure 12, the location of the apparent horizon is plotted on the $\rho - \tau_c$ coordinate plane. In accordance with the generic expectations which are also justified by figures 1 and 7, there are two marginally trapped surfaces on each $\tau = $ constant time level surface. The
Figure 12. On the left, the $r = 0$ singularity (dashed line) and the apparent horizon (continuous line) are shown. On the right, the $\rho$-dependence of the Misner–Sharp mass along this apparent horizon is plotted.

Figure 13. Left: blow-up rates of the square root of the Kretschmann scalar, $R_{abcd} R^{abcd}$, the Ricci scalar curvature $R_{sc} = g^{ab} R_{ab}$, and the Misner–Sharp mass $m$ along a timelike curve, the $\rho = 2.5$ line. Right: $\rho$-dependence of the critical exponents of the blow-up.

$\rho$-coordinate of the inner one is decreasing while that of the outer one is increasing. The plotted radial null geodesics indicate that the apparent horizon on the left panel is everywhere achronal, as expected. It is also remarkable that the outer part of the apparent horizon tends to become almost null very rapidly. On the right panel of figure 12, the Misner–Sharp mass is plotted as a function of $\rho$ along the future apparent horizon, $A^+$. As is clearly visible, $m$ is increasing monotonously along $A^+$ and it tends to the value of the ADM mass $m_{\text{ADM}} = m(\tau = 0, \rho \to \infty)$ which is equal to the Misner–Sharp mass $m(\tau = 0, \rho_{\text{max}})$ in the present case since the matter field is of compact support. As the area of the apparent horizon is $A = 4\pi r^2$ and $r/2 = m_{\text{adm}} \leq m_{\text{ADM}}$, this plot is also in accordance with the Penrose inequality $\sqrt{A/16\pi} \leq m_{\text{ADM}}$.

As mentioned above, a true scalar curvature singularity develops by the end of the time evolution inside the black hole region when both the Kretschmann scalar, $R_{abcd} R^{abcd}$, and the Ricci scalar curvature, $R_{sc} = g^{ab} R_{ab}$, tend to infinity. As it is justified by the log–log plots on the left panel of figure 13, the blow-up rate of $R_{abcd} R^{abcd}$, $R_{sc} = g^{ab} R_{ab}$ and $m$, at fixed $\rho = 2.5$ location, are $r^{-8}$, $r^{-4}$ and $r^{-0.95}$, respectively. On the right panel of figure 13, the $\rho$-dependence of the critical exponents $\gamma_m$, $\gamma_K$ and $\gamma_K - \gamma_m$ of the blow-up rates $m \approx A_m r^{\gamma_m}$ and $(R_{abcd} R^{abcd})^{1/2} \approx A_K r^{\gamma_K}$ are shown. It can be seen that these curvature scalars, along with the Misner–Sharp mass, diverge faster close to the origin. We would like to point to the fact that the blow-up of the Kretschmann scalar is completely consistent with the
Whenever the topology of the initial data surface, relation \( (R_{abcd} R^{abcd})^{1/2} \geq 4\sqrt{2} m/r^3 \) derived by Christodoulou for the case of the collapse of a massless scalar field in [9].

Interestingly, in spite of the divergence of the Ricci scalar, the quantity \( a \beta r^2 R_{sc} \) remains finite. Therefore the Einstein–Hilbert action

\[
S_{EH} = \int_{M_0=\Sigma \times [\rho_0, \rho_1]} R_{sc} \epsilon = 4\pi \int_{\rho_1}^{\rho_2} \int_{\tau_1}^{\tau_2} a \beta r^2 R_{sc} \, d\rho \, d\tau, \tag{88}
\]

remains finite as well, for any particular choice of \( \rho_1, \rho_2 \) with \( 0 \leq \rho_1 < \rho_2 < \rho_{\text{max}} \) in \( \Sigma' = [\rho_1, \rho_2] \times \mathbb{S}^2 \).

### 7.4. Gravitational collapse with spatial topology \( \Sigma_0 = \mathbb{S}^3 \)

Whenever the topology of the initial data surface, \( \Sigma_0 \), is chosen to be \( \mathbb{S}^3 = [0, \pi] \times \mathbb{S}^2 \), there have to be two origins on \( \Sigma_0 \) in the considered spherically symmetric setting. Motivated by and mimicking the simplicity of the geometric setup of the FRW cosmological model, we used procedure B (see section 7.1) and chose the functional form of the initial \( \alpha \) and \( r \) as

\[
\alpha = R_0^2 \quad \text{and} \quad r = R_0 \sin \rho, \tag{89}
\]

To have a nearly central symmetric time evolution, the initial data for the scalar field were chosen to be slightly asymmetric by requiring \( \psi \) and \( \psi_\tau \) to take the form

\[
\psi = \psi_0 \cos n\rho + \psi_1 \cos n'\rho \quad \text{and} \quad \psi_\tau = \psi_0 \cos n\rho + \psi_1 \cos n'\rho, \tag{90}
\]

The parameters \( R_0, \psi_0, \psi_1, \psi_0, \psi_1, n \) and \( n' \) in the above functional expressions were fixed as

\[
R_0 = 10, \quad \psi_0 = 0.4, \quad \psi_1 = 0.005, \quad \psi_0 = 1.2, \quad \psi_1 = 0.01, \quad n = 3 \quad \text{and} \quad n' = 2, \tag{91}
\]

while the number of the spatial grid points \( N \) was chosen to be 1000.

In figure 14, the time evolution of the gravity-matter energy density distribution associated with a shell of radius \( \rho \), \( \epsilon_{\text{GM}} \), is shown. The left panel of figure 15 shows a somewhat exotic distribution of the marginally trapped surfaces produced by the time evolution. This plot verifies that the formation of the curvature singularity is censored, i.e., a connected future distribution of the marginally trapped 2-surfaces produced by the time evolution. This plot resembles the present-time evolution of the parameters relevant for the detection of the marginally trapped surfaces. Thereby the associated past apparent horizon may also be referred to as the boundary of the dynamical white hole region.

On the right panel of figure 15, the time dependence of the spatial 3-volume, \( V = \int_0^\tau \sqrt{\alpha \beta} \, d\rho \), is shown. It increases exponentially up to \( \tau \approx 0.48 \) after which the tendency is reversed and an even faster decay of \( V \) is experienced.

An interesting new feature of the evolution is that the metric function \( \alpha \) appears to blow-up while approaching the singularity. However, its blow-up is compensated by the decay of \( \beta \); the proper time \( t(\rho) = \int_0^\tau \sqrt{\alpha \beta} \, d\tau \) was found to be finite along any constant \( \rho \) world-line—meaning that the singularity is in finite proper time distance from the initial data surface.

The \( r \)-dependence of the square root of the Kretschmann scalar and the Misner–Sharp mass, along the world-line \( \rho = 2.5 \), is indicated on the left panel of figure 16. On the right, the \( \rho \)-dependence of the critical exponents \( \gamma_K, \gamma_m \) and \( \gamma_K - \gamma_m \) is given. Interestingly, Christodoulou’s relation \( (R_{abcd} R^{abcd})^{1/2} \geq 4\sqrt{2} m/r^3 \) derived only for the case of the
collapse of a massless scalar field in a spacetime with spatial topology \( \mathbb{R}^3 \) in [9]—does apply here as well.

The value of the Einstein–Hilbert action,

\[
S_{\text{EH}} = \int_{\mathcal{M} = \mathbb{S}^1 \times (0, \tau_*)} \mathcal{R} \mathcal{c} \epsilon = \int_0^{\tau_*} L \, d\tau, \quad (92)
\]

remains finite as it is indicated by the \( \tau \)-dependence of the Lagrangian, \( L = 4\pi \int_0^\pi \alpha \beta r^2 R_{sc} \, d\rho \), shown on the left panel of figure 15.

7.5. **Gravitational collapse with spatial topology** \( \Sigma_0 = \mathbb{S}^1 \times \mathbb{S}^2 \)

Whenever the topology of the initial data surface \( \Sigma_0 \) is \( \mathbb{S}^1 \times \mathbb{S}^2 \), the simplest initial configuration to start with is a homogeneous ‘torus’ with constant initial radius and with constant initial
expansion. Accordingly, we used scheme A and chose the initial \( r \) and \( r_\tau \) as
\[
  r(\rho) = r_0 \quad \text{and} \quad r_\tau (\rho) = r_\tau _0,
\]
while the number of the spatial grid points \( N \) was chosen to be 4000.

For the initial \( \psi \) and \( \psi_\tau \), we used the relations in (90) with the parameter values in (91), while the values of the metric parameters were chosen as
\[
  r_0 = 1 \quad \text{and} \quad \dot{r}_0 = 1.
\]

It is important to keep in mind that now \( \rho \) is a periodic coordinate, along the \( S^1 \) factor. This periodicity length was chosen to be \( 2\pi \). Note also that according to the above choice of initial data, there is no origin at all on \( \Sigma_0 \).

In figure 17, the time evolution of the gravity-matter energy density distribution associated with a shell of radius \( \rho \), \( \varepsilon_{\text{GM}} \), is shown. As opposed to the other cases, \( \varepsilon_{\text{GM}} \) tends to zero at
the singularity. The Kretschmann scalar blows up in the singularity such that Christodoulou’s relation \((R_{abcd} R^{abcd})^{1/2} \geq 4\sqrt{2} m/r^3\) holds. The location of the future and past apparent horizons is also indicated on the left panel of figure 18.

The appearance of curvature singularities is preceeded by the formation of trapped regions. As in the previous case, both future and past apparent horizons are formed. They intersect at the 2-surfaces with \(\rho = 0, \pi/2, \pi, 3\pi/2\). Interestingly enough, while the 2-surfaces with \(\rho = 0\) and \(\rho = \pi\) are maximal the other two 2-surfaces have no definite character. More precisely, they are locally maximal in the spatial \(\rho\)-direction whereas they are locally minimal in the timelike \(\tau\)-direction. The spacetime region with untrapped surfaces is very limited in the present case. It is represented by the points between the future (thick solid line) and the past (thin solid line) apparent horizons on the left panel of figure 18.

The \(\tau\)-dependence of the Lagrangian, \(L = 4\pi \int_0^{2\pi} \sqrt{\alpha \beta} \, r^2 R_{\rho \rho} \, d\rho\), indicates (see the right panel of figure 18) that the Einstein–Hilbert action,

\[
S_{\text{EH}} = \int_{M = S^2 \times S^1 \times [0, \tau_\ast]} R_{\infty} \, \epsilon = \int_0^{\tau_\ast} L \, d\tau,
\]

remains finite as in the previous cases.

Finally, the \(\tau\)-dependence of the spatial 3-volume, \(V = \int_0^{2\pi} \sqrt{\alpha \beta} \, d\rho\), is also shown on the right. The volume starts to increase rapidly, and then a short oscillation followed by an extremely fast contraction can be seen.

8. Summary

In most previous numerical simulations, one of the aims was to minimize the extent of the trapped region. This was achieved by using techniques like singularity (and trapped region) avoiding slicings and black hole excision—the latter was originally suggested by Unruh in 1984 (for its first adaptation in numerical simulations see, e.g., [39, 43]). The unsatisfactory aspect of this trouble-avoiding attitude, and the reason for choosing here the simplest possible framework of spherically symmetric dynamical configurations, are justified by the following comments of David Hilbert (1902):

\[
\text{Class. Quantum Grav. 27 (2010) 015001 P Csizmadia and I Rácz}
\]
'In dealing with mathematical problems, specialization plays, as I believe, a still more important part than generalization. Perhaps in most cases where we unsuccessfully seek the answer to a question, the cause of the failure lies in the fact that problems simpler and easier than the one in hand have been either incompletely solved, or not solved at all. Everything depends then, on finding those easier problems and on solving them by means of devices as perfect as possible and of concepts capable of generalizations.'

The key technical achievements that, as we believe, have not been applied before in numerical simulations, are the following.

1. A strongly hyperbolic (symmetrizable) first-order system of evolution equations was singled out for four-dimensional spherically symmetric gravitating systems.
2. The analytic setup ensures that time evolution can be studied on equal footing in trapped and untrapped regions.
3. The numerical framework applies this analytic setup and incorporates the techniques of AMR.

By using these technical developments, we achieved the following results.

1. By introducing a suitable evolution equation for the lapse function $\beta$, the extent of the investigated spacetime domain was enlarged significantly and the physical singularities were approached arbitrarily closely everywhere.
2. The location of the future and past apparent horizons has been determined.
3. Detailed investigation of the rate of curvature blow-up while approaching the singularity.
4. The Einstein–Hilbert action remained finite in all the investigated cases, in spite of the blow-up of the Ricci scalar.

As one of our aims was to clear up some conceptual issues of topology change, let us close this paper with some related comments.

In dealing with the problem of topology changes, one should start by recalling the result of Geroch [24] which asserts that if topology change develops then there must exist either closed causal curves beyond the Cauchy horizon\footnote{As was discussed in the introduction, in the present context one can always think of the limit of the time level surfaces as the Cauchy horizon.} or a spacetime singularity has to appear. Due to Tipler’s theorem [42], the latter case manifests itself if Einstein’s equations are imposed. Based on these results, it has been widely held that no indication of topology changes will ever show up in classical general relativity; therefore, quantum theory should be investigated to see whether they may occur.

The idea that the topology may change in quantum gravity was originally proposed by Wheeler [45, 46]. Since then, one of the most important questions in any quantum theory of gravity is whether there is a non-zero probability for the topology of space to change. There have been a number of classical investigations aiming to demonstrate the feasibility of topology change by using the techniques of differential geometry and topology (see, e.g., [4, 17, 28, 41]). Nevertheless, no quantitative investigations have yet been carried out. Therefore, it is important to emphasize that some of our findings provide the first definite quantitative support for all the former speculations concerning the existence of topology changes.

In summarizing our pertinent observations, we can say, in accordance with the above recalled results of Geroch and Tipler, that instead of having a regular Cauchy horizon, along with a causality violating region beyond, a spacetime singularity develops at the ‘new origins’.
It was also found that the Kretschmann scalar always blows up there. It is important to emphasize that the Einstein–Hilbert action $S_{\text{EH}} = \int_{M=\Sigma \times (t_0, t_1)} R \, \epsilon$ was found to be bounded in all of our investigations.

The existence of time developments with apparent topology changes may be significant in quantum theoretical considerations. For instance, in the sum over histories approach to quantum gravity, the transition amplitude for the topology change between the two Riemannian manifolds $(\Sigma_1, h_1)$ and $(\Sigma_2, h_2)$ is supposed to be given by the formula

$$
\langle (\Sigma_1, h_1) | (\Sigma_2, h_2) \rangle = \sum_{M} \int_{\partial M} \exp \{i S_{\text{EH}}\} \mathcal{D}g_{ab},
$$

(96)

where the boundary of $M$ is the disjoint union of $\Sigma_1$ and $\Sigma_2$, $S_{\text{EH}}$ denote the Einstein–Hilbert action and $\mathcal{D}g_{ab}$ is the space of four-dimensional Lorentzian geometries on $M$, while the sum is over all 4-manifolds whose boundary is the disjoint union of $\Sigma_1$ and $\Sigma_2$ with Riemannian metrics $h_1$ and $h_2$ (see, e.g., [18, 27, 28, 45, 46]).

Admittedly, the right-hand side of (96) is completely formal and is far from being defined as yet. Nevertheless, regardless of the specific form of the measure on the space of metrics, $\mathcal{D}g_{ab}$, it is widely held that not only the smooth Lorentzian metrics but all the metrics with finite Einstein–Hilbert action should be taken into account in evaluating the functional integral. Therefore the spacetimes, with topology changes investigated in this paper, might be of interest in quantum theoretical considerations.

It is of obvious importance to know whether the methods introduced in this paper could be adopted in more generic geometrical setup allowing the presence of gravitational waves and with the inclusion of binary black holes. We would like to mention that by using the analytic setup suggested in [31] such a generalization seems to be possible. The results of our corresponding investigations will be published elsewhere.

Acknowledgement

This research was supported in part by OTKA grant K67942.

References

[1] Baumgarte T W and Shapiro S L 1999 On the numerical integration of Einstein’s field equations Phys. Rev. D59 024007
[2] Bengtsson I and Senovilla J M M 2009 A note on trapped surfaces in the vaidya solution Phys. Rev. D79 024027
[3] Berger M J and Oliger J 1984 Adaptive mesh refinement for hyperbolic partial differential equations J. Comput. Phys. 53 484–512
[4] Borde A 1997 How impossible is topology change? Bull. Astron. Soc. India 25 571–7
[5] Brown D 2008 BSSN in spherical symmetry Class. Quantum Grav. 25 205004
[6] Choptuik M W 1993 Universality and scaling in gravitational collapse of a massless scalar field Phys. Rev. Lett. 70 9–12
[7] Choquet-Bruhat Y and Geroch R 1969 Global aspects of the Cauchy problem in general relativity Commun. Math. Phys. 14 329–35
[8] Christodoulou D 1986 The problem of a self-gravitating scalar field Commun. Math. Phys. 105 337–61
[9] Christodoulou D 1991 The formation of black holes and singularities in spherically symmetric gravitational collapse Commun. Pure Appl. Math. XLIV 339–73
[10] Christodoulou D 1994 Examples of naked singularity formation in the gravitational collapse of a scalar field Ann. Math. 140 607–53
[11] Courant R and Hilbert D 1962 Methods of Mathematical Physics vol II (New York: Interscience)
[12] Csizmadia P 2006 Testing a new mesh refinement code in the evolution of a spherically symmetric Klein–Gordon field Int. J. Mod. Phys. D 15 107–19
[13] Csizmadia P 2007 Fourth order AMR and nonlinear dynamical systems in compactified space Class. Quantum Grav. 24 S369
[14] http://www.kfki.hu/~cspeter/gridripper/index.html
[15] Dafermos M 2005 Spherically symmetric space-times with a trapped surface Class. Quantum Grav. 22 2221–32
[16] Dafermos M 2005 A note on naked singularities and the collapse of self-gravitating Higgs fields Adv. Theor. Math. Phys. 9 575–91
[17] Dowker H F and Garcia R S 1998 A handlebody calculus for topology change Class. Quantum Grav. 15 1859–79
[18] Dowker H F 2003 Topology change in quantum gravity The Future of Theoretical Physics and Cosmology ed G W Gibbons, S J Rankin and E P S Shellard (Cambridge: Cambridge University Press) pp 879
[19] Fodor G and Rácz I 2004 What does a strongly excited ’t Hooft–Polyakov magnetic monopole do? Phys. Rev. Lett. 92 151801
[20] Fodor G and Rácz I 2008 Numerical investigation of highly excited magnetic monopoles in SU(2) Yang–Mills–Higgs theory Phys. Rev. D 77 025019
[21] Friedrich H 1996 Hyperbolic reductions for Einstein’s equations Class. Quantum Grav. 13 1451–69
[22] Frittelli S 2006 Well-posed ADM equivalent of the Bondi–Sachs problem Phys. Rev. D 73 124001
[23] Frittelli S and Gomez R 2007 Initial-boundary-value problem of the self-gravitating scalar field in the Bondi–Sachs gauge Phys. Rev. D 75 044021
[24] Geroch R 1967 Topology in general relativity J. Math. Phys. 8 782–6
[25] Geroch R 1970 Domain of dependence J. Math. Phys. 11 437–49
[26] Gustafsson B, Kreiss H-O and Oliger J 1995 Time Dependent Problems and Difference Methods (Pure and Applied Mathematics) (New York: Wiley)
[27] Hawking S W 1978 Quantum gravity and path integrals Phys. Rev. D 18 1747–53
[28] Horowitz G T 1991 Topology change in classical and quantum gravity Class. Quantum Grav. 8 587–601
[29] Kodama H 1980 Conserved energy flux from the spherically symmetric system and the back reaction problem in the black hole evaporation Prog. Theor. Phys. 63 1217
[30] Misner C W and Sharp D H 1964 Relativistic equations for adiabatic, spherically symmetric gravitational collapse Phys. Rev. 136 B571–6
[31] Müller O and Sánchez M 2008 Lorentzian manifolds isometrically embeddable in $L^n$ arXiv:0812.4439
[32] Penrose R 1965 Gravitational collapse and space-time singularities Phys. Rev. Lett. 14 54–9
[33] Pretorius F and Choptuik M W 2006 Adaptive mesh refinement for coupled elliptic-hyperbolic systems J. Comput. Phys. 218 246–74
[34] Rácz I 2006 On the use of the Kodama vector field in spherically symmetric dynamical problems Class. Quantum Grav. 23 115–24
[35] Rácz I 2009 On the topology of untrapped surfaces Class. Quantum Grav. 26 055017
[36] Rácz I 1993 Space-time extensions: I. J. Math. Phys. 34 2448–64
[37] Rácz I 2008 Space-time extensions: II arXiv:0803.0648
[38] Reula O A 1998 Hyperbolic methods for Einstein’s equations Living Rev. Relativity 1 3 URL: http://www.livingreviews.org/lrr-1998-3
[39] Seidel E and Suen W-M 1992 Towards a singularity-proof scheme in numerical relativity Phys. Rev. Lett. 69 1845–9
[40] Shibata M and Nakamura T 1995 Evolution of three-dimensional gravitational waves: harmonic slicing case Phys. Rev. D 52 5428–44
[41] Sorkin R D and Surya S 1998 An analysis of the representations of the mapping class group of a multi-geon three-manifold Int. J. Mod. Phys. A 13 3749–90
[42] Tipler F 1977 Singularities and causality violation Ann. Phys., NY. 108 1–36
[43] Thornburg J 1987 Coordinates and boundary conditions for the general relativistic initial data problem Class. Quantum Grav. 4 1119–31
[44] Wald R M 1984 General relativity (Chicago, IL: University of Chicago Press)
[45] Wheeler J A 1957 On the nature of quantum geometrodynamics Ann. Phys., NY. 2 604–14
[46] Wheeler J A 1962 Geometrodynamics (New York: Academic)