Re-formulating the Einstein equations for stable numerical simulations
— Formulation Problem in Numerical Relativity —

Hisa-aki Shinkai

Computational Science Division,
The Institute of Physical and Chemical Research (RIKEN),
Wako, Saitama, 351-0198 Japan

Abstract

We review recent efforts to re-formulate the Einstein equations for fully relativistic numerical simulations. The so-called numerical relativity (computational simulations in general relativity) is a promising research field matching with ongoing astrophysical observations such as gravitational wave astronomy. Many trials for longterm stable and accurate simulations of binary compact objects have revealed that mathematically equivalent sets of evolution equations show different numerical stability in free evolution schemes. In this article, we first review the efforts of the community, categorizing them into the following three directions: (1) modifications of the standard Arnowitt-Deser-Misner equations initiated by the Kyoto group, (2) rewriting of the evolution equations in hyperbolic form, and (3) construction of an “asymptotically constrained” system. We next introduce our idea for explaining these evolution behaviors in a unified way using eigenvalue analysis of the constraint propagation equations. The modifications of (or adjustments to) the evolution equations change the character of constraint propagation, and several particular adjustments using constraints are expected to diminish the constraint-violating modes. We propose several new adjusted evolution equations, and include some numerical demonstrations. We conclude by discussing some directions for future research.

aMore detail review is available as [42].

1 Introduction

The theory of general relativity describes the nature of the strong gravitational field. The Einstein equation predicts quite unexpected phenomena such as gravitational collapse, gravitational waves, the expanding universe and so on, which are all attractive not only for researchers but also for the public. The Einstein equation consists of 10 partial differential equations (elliptic and hyperbolic) for 10 metric components, and it is not easy to solve them for any particular situation. Over the decades, people have tried to study the general-relativistic world by finding its exact solutions, by developing approximation methods, or by simplifying the situations. Among these approaches, direct numerical integration of the Einstein equations can be said to be the most robust way to study the strong gravitational field. This research field is often called “numerical relativity”.

Numerical relativity is now an essential field in gravity research. The current mainstream in numerical relativity is to analyze the final phase of compact binary objects (black holes and/or neutron stars) related to gravitational wave observations (see e.g. the conference proceedings [33]). Over the past decades, many groups have developed their numerical simulations by trial and error. Simulations require large-scale computational facilities, and long-time stable and accurate calculations. So far, we have achieved certain successes in simulating the coalescence of binary neutron stars (see e.g. [38]) and binary black holes (see e.g.[5]). However, people have still been faced with unreasonable numerical blow-ups at the end of simulations.

E-mail: hshinkai@postman.riken.go.jp
This article is for the proceedings of The 12th Japan General Relativity and Gravitation Workshop, held at Tokyo (Komaba) Univ., Nov 25-28, 2002.
Difficulties in accurate/stable long-term evolution were supposed to be overcome by choosing proper
gauge conditions and boundary conditions. However, recent several numerical experiments show that the
(standard) Arnowitt-Deser-Misner (ADM) approach [9, 44, 53] is not the best formulation for numerics,
and finding a better formulation has become one of the main research topics. A majority of workers in
the field now believe in the existence of constraint-violating modes in most of the formulations. Thus,
the stability problem is now shedding light on the mathematical structure of the Einstein equations.

The purpose of this article is to review the formulation problem in numerical relativity. Generally
speaking, there are many open issues in numerical relativity, both theoretical (mathematical or physical)
and numerical. We list several major topics in Box 1.2 in [42].

There are several different approaches to simulating the Einstein equations. Among them the most
robust way is to apply 3+1 (space + time) decomposition of space-time, as was first formulated by
Arnowitt, Deser and Misner (ADM) [9] (we call this the “original ADM” system). If we divide the space-
time into 3+1 dimensions, the Einstein equations form a constrained system: constraint equations and
evolution equations. The system is quite similar to that of the Maxwell equations, where people solve
constraint equations on the initial data, and use evolution equations to follow the dynamical behaviors.

In numerical relativity, this free-evolution approach is also the standard. This is because solving the
constraints (non-linear elliptic equations) is numerically expensive, and because free evolution allows us
to monitor the accuracy of numerical evolution. In black-hole treatments, recent “excision” techniques do
not require one to impose explicit boundary conditions on the horizon, which is also a reason to apply free
evolution scheme. As we will show in the next section, the standard ADM approach has two constraint
equations; the Hamiltonian (or energy) and momentum constraints.

Up to a couple of years ago, the “standard ADM” decomposition [44, 53] of the Einstein equation was
taken as the standard formulation for numerical relativists. However, numerical simulations were often
interrupted by unexplained blow-ups. This was thought due to the lack of resolution, or inappropriate
choice, or the particular numerical scheme which was applied. However, after the accumulation
of much experience, people have noticed the importance of the formulation of the evolution equations,
since there are apparent differences in numerical stability although the equations are mathematically

At this moment, there are three major ways to obtain longer time evolutions. Of course, the ideas,
procedures, and problems are mingled with each other. The purpose of this article is to review all three
approaches and to introduce our idea to view them in a unified way. Figure 1 is a chronological map of
the researches.

(1) The first direction is to use a modification of the ADM system developed by the Kyoto group [30, 31]
(see §2.1).

(2) The second direction is to re-formulate the Einstein equations in a first-order hyperbolic form. In
constructing hyperbolic systems, the essential procedures are to adjust equations using constraints
and to introduce new variables, normally the spatially derivatived metric (see §2.2).

(3) The third is to construct a system which is robust against the violation of constraints, such that
the constraint surface is an attractor. The idea was first proposed as a “λ-system” by Brodbeck et
al [16] in which they introduce artificial flow to the constraint surface using a new variable based
on the symmetric hyperbolic system (see §2.3).

The word stability is used quite different ways in the community.

- We mean by numerical stability a numerical simulation which continues without any blow-ups and in which data
remains on the constrained surface.

- Mathematical stability is defined in terms of the well-posedness in the theory of partial differential equations, such
that the norm of the variables is bounded by the initial data. See eq. (2.24) and around.

- For numerical treatments, there is also another notion of stability, the stability of finite differencing schemes. This
means that numerical errors (truncation, round-off, etc) are not growing by evolution, and the evaluation is obtained
by von Neumann’s analysis. Lax’s equivalence theorem says that if a numerical scheme is consistent (converging to
the original equations in its continuum limit) and stable (no error growing), then the simulation represents the right
(converging) solution. See [18] for the Einstein equations.
The third idea has been generalized by us as an asymptotically constrained system. The main procedure is to adjust the evolution equations using the constraint equations [49, 50, 41]. The method is also applied to explain why the above approach (1) works, and also to propose alternative systems based on the ADM [50, 41] and BSSN [51] equations. Section 3 is devoted to explain this idea with an analytical tool of the eigenvalue analysis of the constraint propagation.

We follow the notations of that of MTW[28]. We use $\mu, \nu = 0, \cdots, 3$ and $i, j = 1, \cdots, 3$ as space-time indices. The unit $c = 1$ is applied. The discussion is mostly to the vacuum space-time, but the inclusion of matter is straightforward.

2 The standard way and the three other roads

2.0 Strategy 0: The ADM formulation

2.0.1 The original ADM formulation

The idea of space-time evolution was first formulated by Arnowitt, Deser, and Misner (ADM) [9]. The formulation was first motivated by a desire to construct a canonical framework in general relativity, but it also gave the community to the fundamental idea of time evolution of space and time: such as foliations of 3-dimensional hypersurface (Figure 2). This original ADM formulation was translated to numerical relativists by Smarr [44] and York [53] in late 70s, with slightly different notations. We refer to the latter as the standard ADM formulation since this version is the starting point of the discussion.

The story begins by decomposing 4-dimensional space-time into 3 plus 1. The metric is expressed by

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -\alpha^2 dt^2 + \gamma_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt),$$

where $\alpha$ and $\beta_j$ are defined as $\alpha \equiv 1/\sqrt{-g^{00}}$ and $\beta_j \equiv g_{0j}$, and called the lapse function and shift vector, respectively. The projection operator or the intrinsic 3-metric $g_{ij}$ is defined as $\gamma_{ij} = g_{\mu\nu}n^\mu n^\nu$, where $n^\mu = (-\alpha, 0, 0, 0)$ and $n^\nu = g^{\mu\nu}n_\nu = (1/\alpha, -\beta^i/\alpha)$ is the unit normal vector of the spacelike hypersurface, $\Sigma$ (see Figure 2). By introducing the extrinsic curvature,

$$K_{ij} = -\frac{1}{2} \overline{\nabla} \gamma_{ij},$$

Figure 1: Chronological table of formulations and their numerical tests. Boxed ones are of proposals of formulation, circled ones are related numerical experiments. Please refer Table 1 in [42] for each references.
and using the Gauss-Codacci relation, the Hamiltonian density of the Einstein equations can be written as

$$H_{GR} = \pi^{ij} \gamma_{ij} - \mathcal{L}, \quad \text{where} \quad \mathcal{L} = \sqrt{-g} R = \alpha \sqrt{\gamma} [^{(3)}R - K^2 + K_{ij} K^{ij}], \quad (2.3)$$

where $\pi^{ij}$ is the canonically conjugate momentum to $\gamma_{ij}$,

$$\pi^{ij} = \frac{\partial \mathcal{L}}{\partial (\dot{\gamma}_{ij})} = -\sqrt{\gamma} (K^{ij} - K \gamma^{ij}), \quad (2.4)$$

omitting the boundary terms. The variation of $H_{GR}$ with respect to $\alpha$ and $\beta^i$ yields the constraints, and the dynamical equations are given by $\dot{\gamma}_{ij} = \frac{\delta H_{GR}}{\delta \pi^{ij}}$ and $\dot{\pi}_{ij} = -\frac{\delta H_{GR}}{\delta h_{ij}}$.

### 2.0.2 The standard ADM formulation

In the version of Smarr and York, $K_{ij}$ was used as a fundamental variable instead of the conjugate momentum $\pi^{ij}$. (We remark that there is one replacement in (2.6) using (2.7) in the process of conversion from the original ADM to the standard ADM equations.)

**The Standard ADM formulation [44, 53]:**

The fundamental dynamical variables are $(\gamma_{ij}, K_{ij})$, the three-metric and extrinsic curvature. The three-hypersurface $\Sigma$ is foliated with gauge functions, $(\alpha, \beta^i)$, the lapse and shift vector.

- **The evolution equations:**

  $$\partial_t \gamma_{ij} = -2\alpha K_{ij} + D_i \beta_j + D_j \beta_i, \quad (2.5)$$

  $$\partial_t K_{ij} = \alpha (^{(3)}R_{ij} + \alpha K K_{ij} - 2\alpha K_{ik} K^{kj} - D_i D_j \alpha + (D_i \beta^k) K_{kj} + (D_j \beta^k) K_{ki} + \beta^k D_k K_{ij}) \quad (2.6)$$

  where $K = K^i_i$, and $^{(3)}R_{ij}$ and $D_i$ denote three-dimensional Ricci curvature, and a covariant derivative on the three-surface, respectively.

- **Constraint equations:**

  $$H_{ADM} := (^{(3)}R + K^2 - K_{ij} K^{ij}) \approx 0, \quad (2.7)$$

  $$M_{ij}^{ADM} := D_j K^i_i - D_i K \approx 0, \quad (2.8)$$

  where $^{(3)}R = (^{(3)}R_{ii})$; these are called the Hamiltonian (or energy) and momentum constraint equations, respectively.
The formulation has 12 free first-order dynamical variables \( (\gamma_{ij}, K_{ij}) \), with 4 freedom of gauge choice \( (\alpha, \beta_i) \) and with 4 constraint equations, (2.7) and (2.8). The rest freedom expresses 2 modes of gravitational waves.

The constraint propagation equations [22], which are the time evolution equations of the Hamiltonian constraint (2.7) and the momentum constraints (2.8), can be written as

**The Constraint Propagations of the Standard ADM:**

<table>
<thead>
<tr>
<th>Box 2.2</th>
</tr>
</thead>
</table>
| \[ \begin{align*}
\partial_t \mathcal{H} &= \beta^i (\partial_i \mathcal{H}) + 2\alpha K \mathcal{H} - 2\alpha \gamma_{ij} (\partial_i M_j) \\
&\quad + \alpha (\partial_i \gamma_{mk}) (2\gamma_{ij} \gamma_{kj}) \mathcal{M}_j - 4\gamma_{ij} (\partial_i \alpha) M_j, \\
\partial_t M_i &= -(1/2) \alpha (\partial_t \mathcal{H}) - (\partial_i \alpha) \mathcal{H} + \beta^i (\partial_j M_i) \\
&\quad + \alpha K M_i - \beta^i \gamma_{ij} \gamma_{jk} M_j + (\partial_i \beta_k) \gamma_{kj} M_j. \tag{2.9} \end{align*} \] |

From these equations, we know that *if the constraints are satisfied on the initial slice \( \Sigma \), then the constraints are satisfied throughout evolution*. The normal numerical scheme is to solve the elliptic constraints for preparing the initial data, and to apply the free evolution (solving only the evolution equations). The constraints are used to monitor the accuracy of simulations.

The ADM formulation was the standard formulation for numerical relativity up to the middle 90s. Numerous successful simulations were obtained for the problems of gravitational collapse, critical behavior, cosmology, and so on. However, stability problems have arisen for the simulations such as the gravitational radiation from compact binary coalescence, because the models require quite a long-term time evolution.

The origin of the problem was that the above statement in *Italics* is true in principle, but is not always true in numerical applications. A long history of trial and error began in the early 90s. We showed that the standard ADM equations *has* a constraint violating mode in its constraint propagation equations even for a single black-hole (Schwarzschild) spacetime [41].

### 2.1 Strategy 1: Modified ADM formulation by Nakamura et al

Up to now, the most widely used formulation for large scale numerical simulations is a modified ADM system, which is now often cited as the Baumgarte-Shapiro-Shibata-Nakamura (BSSN) formulation. This reformulation was first introduced by Nakamura *et al.* [30, 31, 37]. The usefulness of this reformulation was re-introduced by Baumgarte and Shapiro [11], then was confirmed by other groups to show a long-term stable numerical evolution [4, 6].

#### 2.1.1 Basic variables and equations

The widely used notation [11] introduces the variables \( (\varphi, \bar{\gamma}_{ij}, K, \bar{A}_{ij}, \bar{\Gamma}^i) \) instead of \( (\gamma_{ij}, K_{ij}) \), where

\[
\begin{align*}
\varphi &= (1/2) \log(\det \gamma_{ij}), \\
\bar{\gamma}_{ij} &= e^{-4\varphi} \gamma_{ij}, \\
K &= \gamma^{ij} K_{ij}, \\
\bar{A}_{ij} &= e^{-4\varphi} (K_{ij} - (1/3) \gamma_{ij} K), \\
\bar{\Gamma}^i &= \bar{\Gamma}^i_{jk} \bar{\gamma}^{jk}.
\end{align*}
\]

The new variable \( \bar{\Gamma}^i \) was introduced in order to calculate Ricci curvature more accurately. In BSSN formulation, Ricci curvature is not calculated as \( R^\text{ADM}_{ij} = \partial_i \Gamma^k_{kj} - \partial_j \Gamma^k_{ki} + \Gamma^l_{ij} \Gamma^k_{lk} - \Gamma^l_{ik} \Gamma^k_{lj} \), but as \( R^\text{BSSN}_{ij} = \bar{R}^i_{ij} + \bar{R}_{ij} \), where the first term includes the conformal factor \( \varphi \) while the second term does not. These are approximately equivalent, but \( R^\text{BSSN}_{ij} \) does have wave operator apparently in the flat background limit, so that we can expect more natural wave propagation behavior.

Additionally, the BSSN requires us to impose the conformal factor as \( \bar{\gamma}(= \det \bar{\gamma}_{ij}) = 1 \), during evolution. This is a kind of definition, but can also be treated as a constraint.
The BSSN formulation [30, 31, 37, 11]:

The fundamental dynamical variables are \((\phi, \gamma_{ij}, K, \tilde{A}_{ij}, \tilde{\Gamma}^i)\).

The three-hypersurface \(\Sigma\) is foliated with gauge functions, \((\alpha, \beta^i)\), the lapse and shift vector.

- The evolution equations:

\[
\begin{align*}
\partial_t^B \phi &= -(1/6)\alpha K + (1/6)\beta^i (\partial_i \phi) + (\partial_i \beta^i), \\
\partial_t^B \gamma_{ij} &= -2\alpha \tilde{A}_{ij} + \gamma_{ik}(\partial_j \beta^k) + \gamma_{jk}(\partial_i \beta^k) - (2/3)\tilde{\gamma}_{ij}(\partial_k \beta^k) + \beta^k(\partial_k \tilde{\gamma}_{ij}), \\
\partial_t^B K &= -D^i D_J \alpha + \alpha \tilde{A}_{ij} \tilde{\Gamma}^i + (1/3)\alpha K^2 + \beta^i (\partial_i K), \\
\partial_t^B \tilde{A}_{ij} &= -e^{-4\phi}(D_i D_J \alpha)^{TF} + e^{-4\phi} \alpha (R_{ij}^{BSSN})^{TF} + \alpha K \tilde{A}_{ij} - 2\alpha \tilde{A}_{ik} \tilde{A}^k_j + (\partial_i \beta^k) \tilde{A}_{kj} + (\partial_j \beta^k) \tilde{A}_{ki} - (2/3)(\partial_k \beta^k) \tilde{A}_{ij} + \beta^k (\partial_k \tilde{A}_{ij}), \\
\partial_t^B \tilde{\Gamma}^i &= -2(\partial_j \alpha) \tilde{\Gamma}^i_j + 2\alpha (\tilde{\Gamma}^i_{jk} \tilde{A}^k_j - (2/3)\tilde{\gamma}^i_{jk} (\partial_k K) + 6 \tilde{A}^i_j (\partial_j \phi)) - \partial_j (\beta^k (\partial_k \tilde{\gamma}^i_{jk}) - \tilde{\gamma}^l_{jk} (\partial_k \beta^l) - \tilde{\gamma}^i_{lj} (\partial_k \beta^j) + (2/3)\tilde{\gamma}^i_{lj} (\partial_k \beta^j)).
\end{align*}
\]

- Constraint equations:

\[
\begin{align*}
\mathcal{H}_{BSSN}^{BSSN} &= R^{BSSN} + K^2 - K_{ij} K^{ij}, \\
\mathcal{M}_{i}^{BSSN} &= \mathcal{M}_{i}^{ADM}, \\
\mathcal{G}^i &= \tilde{\Gamma}^i - \tilde{\gamma}^i_{jk} \tilde{\Gamma}^j_k, \\
\mathcal{A} &= \tilde{A}_{ij} \tilde{\gamma}^{ij}, \\
\mathcal{S} &= \tilde{\gamma} - 1.
\end{align*}
\]

(2.18) and (2.19) are the Hamiltonian and momentum constraints (the “kinematic” constraints), while the latter three are “algebraic” constraints due to the requirements of BSSN formulation.

2.1.2 Remarks

Why BSSN is better than the standard ADM? Together with numerical comparisons with the standard ADM case[6], this question has been studied by many groups using different approaches. Using numerical test evolution, Alcubierre et al [4] found that the essential improvement is in the process of replacing terms by the momentum constraints. They also pointed out that the eigenvalues of BSSN evolution equations have fewer “zero eigenvalues” than those of ADM, and they conjectured that the instability might be caused by these “zero eigenvalues”. An effort was made to understand the advantage of BSSN from the point of hyperbolization of the equations in its linearized limit [4, 34]. These studies provide some support regarding the advantage of BSSN, while it is also shown an example of an ill-posed solution in BSSN (as well in ADM) by Frittelli and Gomez [23].

As we discussed in [51], the stability of the BSSN formulation is due not only to the introductions of new variables, but also to the replacement of terms in the evolution equations using the constraints. Further, we will show several additional adjustments to the BSSN equations which are expected to give us more stable numerical simulations.

2.2 Strategy 2: Hyperbolic reformulations

2.2.1 Definitions, properties, mathematical backgrounds

The second effort to re-formulate the Einstein equations is to make the evolution equations reveal a first-order hyperbolic form explicitly. This is motivated by the expectation that the symmetric hyperbolic system has well-posed properties in its Cauchy treatment in many systems and also that the boundary treatment can be improved if we know the characteristic speed of the system.

We say that the system is a first-order (quasi-linear) partial differential equation system, if a certain set of (complex-valued) variables \( u_\alpha \) (\( \alpha = 1, \ldots, n \)) forms
\[
\partial_t u_\alpha = M^{\beta\alpha}(u) \partial_l u_\beta + N_\alpha(u),
\] (2.23)
where \( M \) (the characteristic matrix) and \( N \) are functions of \( u \) but do not include any derivatives of \( u \). Further we say the system is
- a weakly hyperbolic system, if all the eigenvalues of the characteristic matrix are real.
- a strongly hyperbolic system (or a diagonalizable / symmetrizable hyperbolic system), if the characteristic matrix is diagonalizable (has a complete set of eigenvectors) and has all real eigenvalues.
- a symmetric hyperbolic system, if the characteristic matrix is a Hermitian matrix.

Writing the system in a hyperbolic form is a quite useful step in proving that the system is well-posed. The mathematical well-posedness of the system means (1°) local existence (of at least one solution \( u \)), (2°) uniqueness (i.e., at most solutions), and (3°) stability (or continuous dependence of solutions \( \{u\} \) on the Cauchy data) of the solutions. The resultant statement expresses the existence of the energy inequality on its norm,
\[
||u(t)|| \leq e^{\alpha \tau} ||u(t = 0)||, \quad \text{where } 0 < \tau < t, \quad \alpha = \text{const.}
\] (2.24)
This indicates that the norm of \( u(t) \) is bounded by a certain function and the initial norm. Remark that this mathematical boundness does not mean that the norm \( u(t) \) decreases along the time evolution.

The inclusion relation of the hyperbolicities is,
\[
\text{symmetric hyperbolic} \subset \text{strongly hyperbolic} \subset \text{weakly hyperbolic}.
\] (2.25)

The Cauchy problem under weak hyperbolicity is not, in general, \( C^\infty \) well-posed. At the strongly hyperbolic level, we can prove the finiteness of the energy norm if the characteristic matrix is independent of \( u \) (cf [45]), that is one step definitely advanced over a weakly hyperbolic form. Similarly, the well-posedness of the symmetric hyperbolic is guaranteed if the characteristic matrix is independent of \( u \), while if it depends on \( u \) we have only limited proofs for the well-posedness.

From the point of numerical applications, to hyperbolize the evolution equations is quite attractive, not only for its mathematically well-posed features. The expected additional advantages are the following.

(a) It is well known that a certain flux conservative hyperbolic system is taken as an essential formulation in the computational Newtonian hydrodynamics when we control shock wave formations due to matter.

(b) The characteristic speed (eigenvalues of the principal matrix) is supposed to be the propagation speed of the information in that system. Therefore it is naturally imagined that these magnitudes are equivalent to the physical information speed of the model to be simulated.

(c) The existence of the characteristic speed of the system is expected to give us an improved treatment of the numerical boundary, and/or to give us a new well-defined Cauchy problem within a finite region (the so-called initial boundary value problem, IBVP).

These statements sound reasonable, but have not yet been generally confirmed in actual numerical simulations. But we are safe in saying that the formulations are not yet well developed to test these issues. For example, IBVP studies are preliminary yet, and most works are based on a particular symmetric hyperbolic and in a limited space-time symmetry. We will come back to this issue in §2.2.3, but meanwhile let us view the hyperbolic formulations from the comparisons of pure evolution equations.
2.2.2 Hyperbolic formulations of the Einstein equations

Most physical systems can be expressed as symmetric hyperbolic systems. In order to prove that the Einstein’s theory is a well-posed system, to hyperbolize the Einstein equations is a long-standing research area in mathematical relativity.

The standard ADM system does not form a first order hyperbolic system. This can be seen immediately from the fact that the ADM evolution equation (2.6) has Ricci curvature in RHS. So far, several first order hyperbolic systems of the Einstein equation have been proposed. In constructing hyperbolic systems, the essential procedures are (1°) to introduce new variables, normally the spatially derivatived metric, (2°) to adjust equations using constraints. Occasionally, (3°) to restrict the gauge conditions, and/or (4°) to rescale some variables. Due to process (1°), the number of fundamental dynamical variables is always larger than that of ADM.

Due to the limitation of space, we can only list several hyperbolic systems of the Einstein equations.

• The Bona-Massó formulation [13, 14]
• The Einstein-Ricci system [19, 1] / Einstein-Bianchi system [7]
• The Einstein-Christoffel system [8]
• The Ashtekar formulation [10]
• The Frittelli-Reula formulation [24, 45]
• The Conformal Field equations [21]
• The Kidder-Scheel-Teukolsky (KST) formulation [26]

Please refer [42] for each brief introductions.

2.2.3 Remarks

When we discuss hyperbolic systems in the context of numerical stability, the following questions should be considered:

Q From the point of the set of evolution equations, does hyperbolization actually contribute to numerical accuracy and stability? Under what conditions/situations will the advantages of hyperbolic formulation be observed?

Unfortunately, we do not have conclusive answers to these questions, but many experiences are being accumulated. Several earlier numerical comparisons reported the stability of hyperbolic formulations [14, 15, 35, 36]. But we have to remember that this statement went against the standard ADM formulation, which has a constraint-violating mode for Schwarzschild spacetime as has been shown recently [41].

These partial numerical successes encouraged the community to formulate various hyperbolic systems. Recently, Calabrese et al [17] reported there is a certain differences in the long-term convergence features between weakly and strongly hyperbolic systems on the Minkowskii background space-time. However, several numerical experiments also indicate that this direction is not a complete success.

Objections from numerical experiments

• Above earlier numerical successes were also terminated with blow-ups.
• If the gauge functions are evolved according to the hyperbolic equations, then their finite propagation speeds may cause pathological shock formations in simulations [2, 3].
• There are no drastic differences in the evolution properties between hyperbolic systems (weakly, strongly and symmetric hyperbolicity) by systematic numerical studies by Hern [25] based on Frittelli-Reula formulation [24], and by the authors [40] based on Ashtekar’s formulation [10, 48].
• Proposed symmetric hyperbolic systems were not always the best ones for numerical evolution. People are normally still required to reformulate them for suitable evolution. Such efforts are seen in the applications of the Einstein-Ricci system [36], the Einstein-Christoffel system [12], and so on.
Of course, these statements only cast on a particular formulation, and therefore we have to be careful not to over-emphasize the results. In order to figure out the reasons for the above objections, it is worth stating the following cautions:

Remarks on hyperbolic formulations

(a) Rigorous mathematical proofs of well-posedness of PDE are mostly for simple symmetric or strongly hyperbolic systems. If the matrix components or coefficients depend on dynamical variables (as in all any versions of hyperbolized Einstein equations), almost nothing was proved in more general situations.

(b) The statement of “stability” in the discussion of well-posedness refers to the bounded growth of the norm, and does not indicate a decay of the norm in time evolution.

(c) The discussion of hyperbolicity only uses the characteristic part of the evolution equations, and ignores the rest.

We think the origin of confusion in the community results from over-expectation on the above issues. Mostly, point (c) is the biggest problem. The above numerical claims from Ashtekar and Frittelli-Reula formulations were mostly due to the contribution (or interposition) of non-principal parts in evolution. Regarding this issue, the recent KST formulation finally opens the door. KST’s “kinematic” parameters enable us to reduce the non-principal part, so that numerical experiments are hopefully expected to represent predicted evolution features from PDE theories. At this moment, the agreement between numerical behavior and theoretical prediction is not yet perfect but close [27].

If further studies reveal the direct correspondences between theories and numerical results, then the direction of hyperbolization will remain as the essential approach in numerical relativity, and the related IBVP researches will become a main research subject in the future. Meanwhile, it will be useful if we have an alternative procedure to predict stability including the effects of the non-principal parts of the equations. Our proposal of adjusted system in the next subsection may be one of them.

2.3 Strategy 3: Asymptotically constrained systems

The third strategy is to construct a robust system against the violation of constraints, such that the constraint surface is an attractor. The idea was first proposed as “λ-system” by Brodbeck et al [16], and then developed in more general situations as “adjusted system” by the authors [49].

2.3.1 The “λ-system”

Brodbeck et al [16] proposed a system which has additional variables λ that obey artificial dissipative equations. The variable As are supposed to indicate the violation of constraints and the target of the system is to get λ = 0 as its attractor.

The “λ-system” (Brodbeck-Frittelli-Hübner-Reula) [16]:

For a symmetric hyperbolic system, add additional variables λ and artificial force to reduce the violation of constraints.

The procedure:

1. Prepare a symmetric hyperbolic evolution system
   \[ \partial_t u = M \partial_i u + N \]

2. Introduce λ as an indicator of violation of constraint which obeys dissipative eqs. of motion
   \[ \partial_t \lambda = \alpha C - \beta \lambda \]  \( \alpha \neq 0, \beta > 0 \)

3. Take a set of \((u, \lambda)\) as dynamical variables
   \[ \partial_t \begin{pmatrix} u \\ \lambda \end{pmatrix} \simeq \begin{pmatrix} A & 0 \\ F & 0 \end{pmatrix} \partial_i \begin{pmatrix} u \\ \lambda \end{pmatrix} \]

4. Modify evolution eqs so as to form a symmetric hyperbolic system
   \[ \begin{pmatrix} u \\ \lambda \end{pmatrix} = \begin{pmatrix} A & F \\ F & 0 \end{pmatrix} \partial_i \begin{pmatrix} u \\ \lambda \end{pmatrix} \]
Since the total system is designed to have symmetric hyperbolicity, the evolution is supposed to be unique. Brodbeck et al showed analytically that such a decay of $\lambda$s can be seen for sufficiently small $\lambda(>0)$ with a choice of appropriate combinations of $\alpha$s and $\beta$s.

Brodbeck et al presented a set of equations based on Frittelli-Reula’s symmetric hyperbolic formulation [24]. The version of Ashtekar’s variables was presented by the authors [39] for controlling the constraints or reality conditions or both. The numerical tests of both the Maxwell-$\lambda$-system and the Ashtekar-$\lambda$-system were performed [49], and confirmed to work as expected. Although it is questionable whether the recovered solution is true evolution or not [43], we think the idea is quite attractive. To enforce the decay of errors in its initial perturbative stage seems the key to the next improvements, which are also developed in the next section on “adjusted systems”.

However, there is a high price to pay for constructing a $\lambda$-system. The $\lambda$-system can not be introduced generally, because (i) the construction of $\lambda$-system requires the original evolution equations to have a symmetric hyperbolic form, which is quite restrictive for the Einstein equations, (ii) the final system requires many additional variables and we also need to evaluate all the constraint equations at every time step, which is a hard task in computation. Moreover, (iii) it is not clear that the $\lambda$-system is robust enough for non-linear violation of constraints, or that $\lambda$-system can control constraints which do not have any spatial differential terms.

2.3.2 The “adjusted system”

Next, we propose an alternative system which also tries to control the violation of constraint equations actively, which we named “adjusted system”. We think that this system is more practical and robust than the previous $\lambda$-system.

<table>
<thead>
<tr>
<th>The Adjusted system (procedures):</th>
<th>Box 2.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Prepare a set of evolution eqs.</td>
<td>$\partial_t u = J \partial_i u + K$</td>
</tr>
<tr>
<td>2. Add constraints in RHS</td>
<td>$\partial_t u = J \partial_i u + K + \kappa C$</td>
</tr>
<tr>
<td>3. Choose the coeff. $\kappa$ so as to make the eigenvalues of the homogenized adjusted $\partial_t C$ eqs negative reals or pure imaginary.</td>
<td>$\partial_t C = D \partial_i C + EC$</td>
</tr>
</tbody>
</table>

The process of adjusting equations is a common technique in other re-formulating efforts as we reviewed. However, we try to employ the evaluation process of constraint amplification factors as an alternative guideline to hyperbolization of the system. We will explain these issues in the next section.

3 A unified treatment: Adjusted System

This section is devoted to present our idea of “asymptotically constrained system”. Original references can be found in [49, 50, 41, 51].

3.1 Procedures : Constraint propagation equations and Proposals

Suppose we have a set of dynamical variables $u^\alpha(x^i, t)$, and their evolution equations,

$$\partial_t u^\alpha = f(u^\alpha, \partial_i u^\alpha, \cdots), \quad (3.1)$$

and the (first class) constraints,

$$C^\alpha(u^\alpha, \partial_i u^\alpha, \cdots) \approx 0. \quad (3.2)$$

Note that we do not require (3.1) forms a first order hyperbolic form. We propose to investigate the evolution equation of $C^\alpha$ (constraint propagation),

$$\partial_t C^\alpha = g(C^\alpha, \partial_t C^\alpha, \cdots), \quad (3.3)$$
for predicting the violation behavior of constraints in time evolution. We do not mean to integrate (3.3) numerically together with the original evolution equations (3.1), but mean to evaluate them analytically in advance in order to reformulate the equations (3.1).

There may be two major analyses of (3.3); (a) the hyperbolicity of (3.3) when (3.3) is a first order system, and (b) the eigenvalue analysis of the whole RHS in (3.3) after a suitable homogenization. However, as we critically viewed the hyperbolization road in the previous section, we prefer to proceed the road (b).

Amplification Factors of Constraint Propagation equations: 

\[
\partial_t \hat{C}^\alpha = \hat{g}(\hat{C}^\alpha) = M^\alpha_\beta \hat{C}^\beta,
\]

where \( C(x,t)^\rho = \int \hat{C}(k,t)^\rho \exp(ik \cdot x) d^3k \),

then to analyze the set of eigenvalues, say \( \Lambda_s \), of the coefficient matrix, \( M^\alpha_\beta \), in (3.4). We call \( \Lambda_s \) the constraint amplification factors (CAFs) of (3.3).

The CAFs predict the evolution of constraint violations. We therefore can discuss the “distance” to the constraint surface using the “norm” or “compactness” of the constraint violations (although we do not have exact definitions of these “…” words).

The next conjecture seems to be quite useful to predict the evolution feature of constraints:

Conjecture on Constraint Amplification Factors (CAFs):

(A) If CAF has a negative real-part (the constraints are forced to be diminished), then we see more stable evolution than a system which has positive CAF.

(B) If CAF has a non-zero imaginary-part (the constraints are propagating away), then we see more stable evolution than a system which has zero CAF.

We found that the system becomes more stable when more \( \Lambda_s \) satisfy the above criteria. A general feature of the constraint propagation is reported in [52].

The above features of the constraint propagation, (3.3), will differ when we modify the original evolution equations. Suppose we add (adjust) the evolution equations using constraints

\[
\partial_t u^a = f(u^a, \partial_i u^a, \cdots) + F(C^\alpha, \partial_i C^\alpha, \cdots),
\]

then (3.3) will also be modified as

\[
\partial_t C^\alpha = g(C^\alpha, \partial_i C^\alpha, \cdots) + G(C^\alpha, \partial_i C^\alpha, \cdots).
\]

Therefore, the problem is how to adjust the evolution equations so that their constraint propagations satisfy the above criteria as much as possible.

3.2 Applications

For the Maxwell equation and the Ashtekar version of the Einstein equations, we numerically found that this idea works to reduce the violation of constraints, and that the effects are much better than by constructing its symmetric hyperbolic versions [40, 49].
Applications to ADM  The idea was applied to the standard ADM formulation which is not hyperbolic and several attractive adjustments were proposed [50, 41]. We made various predictions how additional adjusted terms will change the constraint propagation. Systematic numerical comparisons are also progressing, and we show two sample plots here.

Figure 3 (a) is a test numerical evolution of Detweiler-type adjustment [20] on the Minkowskii background. We see the adjusted version gives convergence on to the constraint surface by arranging the magnitude of the adjusting parameter, $\kappa$. Figure 3 (b) is obtained by a 3-dimensional numerical evolution of weak gravitational wave, the so-called Teukolsky wave [46]. The lines are of the original/standard ADM evolution equations, Detweiler-type adjustment, and a part of Detweiler-type adjustment. For a particular choice of $\kappa$, we observe again the L2 norm of constraint (violation of constraints) is reduced than the standard ADM case, and can evolve longer than that.

Notion of Time Reversal Symmetry  During the comparisons of adjustments, we found that it is necessary to create time asymmetric structure of evolution equations in order to force the evolution on to the constraint surface. There are infinite ways of adjusting equations, but we found that if we follow the guideline Box 3.5, then such an adjustment will give us time asymmetric evolution.

---

**Box 3.5**

---

**Trick to obtain asymptotically constrained system:**

Break the time reversal symmetry (TRS) of the evolution equation.

1. Evaluate the parity of the evolution equation.
   
   By reversing the time ($\partial_t \rightarrow -\partial_t$), there are variables which change their signatures (parity ($-$)) [e.g. $K_{ij}, \partial_t \gamma_{ij}, M_i, \cdots$], while not (parity ($+$)) [e.g. $g_{ij}, \partial_t K_{ij}, \mathcal{H}, \cdots$].

2. Add adjustments which have different parity of that equation.
   
   For example, for the parity ($-$) equation $\partial_t \gamma_{ij}$, add a parity ($+$) adjustment $\kappa \mathcal{H}$.

---

One of our criteria, the negative real CAFs, requires breaking the time-symmetric features of the original evolution equations. Such CAFs are obtained by adjusting the terms which break the TRS of the evolution equations, and this is available even at the standard ADM system.

Applications to BSSN  This analysis was also applied to explain the advantages of the BSSN formulation, and again several alternative adjustments to BSSN equations were proposed [51]. Recently Yo et al [47] reported their simulations of stationary rotating black hole, and mentioned that one of our proposal was contributed to maintain their evolution of Kerr black hole ($J/M$ up to 0.9$M$) for long time ($t \sim 6000M$). Their results also indicates that the evolved solution is closed to the exact one, that is, the constrained surface.

4  Outlook

4.1  What we have achieved

We reviewed recent efforts on this problem by categorizing them into (0) The standard ADM formulation (§2.0, Box 2.1), (1) The modified ADM (so-called BSSN) formulation (§2.1, Box 2.3), (2) Hyperbolic formulations (§2.2, Box 2.5), and (3) Asymptotically constrained formulations (§2.3). Among them, the approach (2) is perhaps best justified on mathematical grounds. However, as we critically reviewed in §2.2.3, the practical advantages may not be available unless we kill the lower-order terms in its hyperbolized equation, as in KST’s formulation.

We therefore proceeded in the direction (3). Our approach, which we term **adjusted system**, is to construct a system that has its constraint surface as an attractor. Our unified view is to understand the evolution system by evaluating its constraint propagation. Especially we proposed to analyze the
constraint amplification factors (Box 3.1) which are the eigenvalues of the homogenized constraint propagation equations. We analyzed the system based on our conjecture (Box 3.2) whether the constraint amplification factors suggest the constraint to decay/propagate or not. We concluded that

- The constraint propagation features become different by simply adding constraint terms to the original evolution equations (we call this the adjustment of the evolution equations).

- There is a constraint-violating mode in the standard ADM evolution system when we apply it to a single non-rotating black hole space-time, and its growth rate is larger near the black-hole horizon.

- Such a constraint-violating mode can be killed if we adjust the evolution equations with a particular modification using constraint terms (Box 2.7). An effective guideline is to adjust terms as they break the time-reversal symmetry of the equations (Box 3.5).

- Our expectations are borne out in simple numerical experiments using the Maxwell, Ashtekar, and ADM systems. However the modifications are not yet perfect to prevent non-linear growth of the constraint violation.

- We understand why the BSSN formulation works better than the ADM one in the limited case (perturbative analysis in the flat background), and further we proposed modified evolution equations along the lines of our previous procedure.

The common key to the problem is how to adjust the evolution equations with constraints. Any adjusted systems are mathematically equivalent if the constraints are completely satisfied, but this is not the case for numerical simulations. Replacing terms with constraints is one of the normal steps when people hyperbolize equations. Our approach is to employ the evaluation process of constraint amplification factors for an alternative guideline to hyperbolization of the system.

4.2 Final remarks

If we say the final goal of this project is to find a robust algorithm to obtain long-term accurate and stable time-evolution method, then the recipe should be a combination of (a) formulations of the evolution equations, (b) choice of gauge conditions, (c) treatment of boundary conditions, and (d) numerical integration methods. We are in the stages of solving this mixed puzzle.
may not exit, but we believe our accumulating experience will make the ones we do have more robust and automatic.

We have written this review from the viewpoint that the general relativity is a constrained dynamical system. This is not only a proper problem in general relativity, but also in many physical systems such as electrodynamics, magnetohydrodynamics, molecular dynamics, mechanical dynamics, and so on. Therefore sharing the thoughts between different field will definitely accelerate the progress.

When we discuss asymptotically constrained manifolds, we implicitly assume that the dynamics could be expressed on a wider manifold. Recently we found that such a proposal is quite similar with techniques in e.g. molecular dynamics. For example, people assume an extended environment when simulating molecular dynamics under constant pressure (with a potential piston) or a constant temperature (with a potential thermostat) (see e.g. [32]). We have also noticed that a dynamical adjusting method of Lagrange multipliers has been developed in multi-body mechanical dynamics (see e.g. [29]). We are now trying to apply these ideas back into numerical relativity.

In such a way, communication and interaction between different fields is encouraged. Cooperation between numerical and mathematical scientists is necessary. By interchanging ideas, we hope we will reach our goal in next few years, and obtain interesting physical results and predictions. It is our personal view that exciting revolutions in numerical relativity are coming soon.

Acknowledgment

This review is based on the series of works with Gen Yoneda. HS is supported by the special postdoctoral researchers program at RIKEN, and this work was supported partially by the Grant-in-Aid for Scientific Research Fund of Japan Society of the Promotion of Science, No. 14740179.

References

[27] L. Lindblom and M. Scheel, gr-qc/0206035
[42] H. Shinkai and G. Yoneda, in *Progress in Astronomy and Astrophysics* (Nova Science Publ) to be published. The manuscript is available as gr-qc/0209111.