

Hyperbolic formulations and numerical relativity: experiments using Ashtekar's connection variables

Hisa-aki Shinkai[†] and Gen Yoneda[‡]

[†] Centre for Gravitational Physics and Geometry, 104 Davey Laboratory, Department of Physics,
The Pennsylvania State University, University Park, PA 16802-6300, USA

[‡] Department of Mathematical Sciences, Waseda University, Shinjuku, Tokyo, 169-8555, Japan

E-mail: shinkai@gravity.phys.psu.edu and yoneda@mn.waseda.ac.jp

Received 3 May 2000, in final form 13 September 2000

Abstract. In order to perform accurate and stable long-time numerical integration of the Einstein equation, several hyperbolic systems have been proposed. Here we present a numerical comparison between weakly hyperbolic, strongly hyperbolic and symmetric hyperbolic systems based on Ashtekar's connection variables. The primary advantage for using this connection formulation in this experiment is that we can keep using the same dynamical variables for all levels of hyperbolicity. Our numerical code demonstrates gravitational wave propagation in plane-symmetric spacetimes, and we compare the accuracy of the simulation by monitoring the violation of the constraints. By comparing with results obtained from the weakly hyperbolic system, we observe that the strongly and symmetric hyperbolic system show better numerical performance (yield less constraint violation), but not so much difference between the latter two. Rather, we find that the symmetric hyperbolic system is not always the best in terms of numerical performance.

This study is the first to present full numerical simulations using Ashtekar's variables. We also describe our procedures in detail.

(Some figures in this article are in colour only in the electronic version; see www.iop.org)

PACS numbers: 0420C, 0425, 0425D

1. Introduction

Numerical relativity—solving the Einstein equation numerically—is now an essential field in gravity research. As is well known, critical collapse in gravity systems was first discovered by numerical simulation [1]. The current mainstream of numerical relativity is to demonstrate the final phase of compact binary objects related to gravitational wave observations[†], and these efforts are now again shedding light on the mathematical structure of the Einstein equations.

Up to a couple of years ago, the standard Arnowitt–Deser–Misner (ADM) decomposition of the Einstein equation was taken as the standard formulation for numerical relativists. Difficulties in accurate/stable long-term evolutions were supposed to be overcome by choosing proper gauge conditions and boundary conditions. Recently, however, several numerical experiments show that the standard ADM is not the best formulation for numerics, and finding a better formulation has become one of the main research topics[‡].

[†] The latest reviews are available in [2].

[‡] Note that we are only concerned with the free evolution system of the initial data; that is, we only solve the constraint equations on the initial hypersurface. The accuracy and/or stability of the system is normally observed by monitoring the violation of constraints during the free evolution.

One direction in the community is to apply a conformally decoupled and trace-free reformulation of the ADM system which was first used by Nakamura *et al* [3]. The usefulness of this reformulation were confirmed by another groups to show long-term stable numerical evolution. [4, 5]. Although there is an effort to show why this reformulation is better than ADM [6], we do not yet know whether this method is robust for all situations.

Another alternative approach to ADM is to formulate the Einstein equations to reveal hyperbolicity[†]. A certain kind of hyperbolicity of the dynamical equations is essential to analyse their propagation features mathematically, and is known to be useful in numerical approximations (we explain these points in section 2). The propagation of the original ADM constraint equations obeys well posed behaviour [8], but the dynamical equations of the ADM system are not a hyperbolic system at all (and these facts can also be applied to the conformally decoupled version). Several hyperbolic formulations have been proposed to re-express the Einstein equation, with different levels: weakly, strongly and symmetric hyperbolic systems (we will discuss this in detail in section 2). Several numerical tests were also performed in this direction, and we can see advantages in numerical stability over the original ADM system (e.g. tests [9] of Bona–Massó’s symmetrizable form [10], tests [11] of Choquet-Bruhat and York (95)s symmetrizable form [12]), but the appearance of coordinate shocks has also been reported [13] in the system of [9]. A symmetric hyperbolic system of [14], on the other hand, has been studied numerically within the context of the ‘conformal Einstein’ approach [15].

The following questions, therefore, naturally present themselves (cf [16]).

- (a) Does hyperbolicity actually contribute to the numerical accuracy/stability?
- (b) If so, which level of hyperbolic formulation is practically useful for numerical applications? (or does the symmetric hyperbolicity solve all the difficulties?)
- (c) Are there any other approaches to improve the accuracy/stability of the system?

In this paper, we try to answer these questions with our simple numerical experiments. Such comparisons are appropriate when the fundamental equations are cast in the same interface, and that is possible at this moment only using Ashtekar’s connection variables [17, 18]. More precisely, the authors’ recent studies showed the following:

- (a) the original set of dynamical equations proposed by Ashtekar already forms a weakly hyperbolic system [19];
- (b) by requiring additional gauge conditions *or* adding constraints to the dynamical equations, we can obtain a strongly hyperbolic system [19];
- (c) by requiring additional gauge conditions *and* adding constraints to the dynamical equations, we can obtain a symmetric hyperbolic system [19, 20]; and finally
- (d) based on the above symmetric hyperbolic system, we can construct a set of dynamical systems which is robust against perturbative errors for constraints and reality conditions [21] (also known as a λ -system [22]).

Based on the above results (a)–(c), we developed a numerical code which handles gravitational wave propagation in the plane-symmetric spacetime. We performed the time evolutions using the above three levels of Ashtekar’s dynamical equations together with the standard ADM equation. We compare these for accuracy and stability by monitoring the violation of the constraints. We also show the demonstrations of our λ -system (above (d)) in a subsequent paper (paper II) [41], together with new proposal for controlling the stability.

[†] Recent reviews are given in, for example, [7].

It is worth remarking that this study is the first one which shows full numerical simulations of Lorentzian spacetime using Ashtekar's connection variables. This research direction was suggested [23] soon after Ashtekar completed his formulation, but has not yet been completed. Historically, an application to numerical relativity of the connection formulation was also suggested [18, 24] using the Capovilla–Dell–Jacobson version of the connection variables [25], which produce a direct relation to the Newman–Penrose Ψ s. However, here we apply Ashtekar's original formulation, because we know how to treat its reality conditions in detail [26, 27], and how they form hyperbolicities. We will also describe the basic numerical procedures in this paper.

The outline of this paper is as follows. In the next section, we review the mathematical background of the hyperbolic formulation briefly and present our fundamental dynamical equations. In section 3, we describe our numerical procedures. Our experiments are presented in sections 4 and 5, and we summarize them in section 6. Appendix A shows Ashtekar's basic equations in our notation, and we also present our experiments based on the Maxwell equation in appendix B. We also introduce briefly the discussion in our paper II in appendix C.

2. Hyperbolic formulations

2.1. Definitions, properties, mathematical backgrounds

We say that the system is a *first-order (quasi-linear) partial differential equation system*, if a certain set of (complex-valued) variables u_α ($\alpha = 1, \dots, n$) forms

$$\partial_t u_\alpha = \mathcal{M}^{l\beta}_\alpha(u) \partial_l u_\beta + \mathcal{N}_\alpha(u), \quad (1)$$

where \mathcal{M} (the characteristic matrix) and \mathcal{N} are functions of u but do not include any derivatives of u . If the characteristic matrix is a Hermitian matrix, then we say that (1) is a *symmetric hyperbolic system*.

Writing the system in a hyperbolic form is the essential step in proving the system is well posed. Here, *well posedness* of the system means (a) existence (of at least one solution u), (b) uniqueness (i.e. at most solutions) and (c) stability (or continuous dependence of solutions $\{u\}$ on the Cauchy data). The Cauchy problem under weak hyperbolicity is not, in general, C^∞ well posed. The symmetric hyperbolic system gives us the energy integral inequalities which are the primary tools for studying the stability of the system. 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 a limited proof for the well posedness. From the mathematical point of view, proving well posedness with less strict conditions is an old but active research problem.

We can define another hyperbolic system between the weakly and symmetric levels. For example, we say we have a *strongly hyperbolic* (or diagonalizable hyperbolic [19]) system, if the characteristic matrix is diagonalizable and has all real eigenvalues. The inclusion relation is, then,

$$\text{symmetric hyperbolic} \in \text{strongly hyperbolic} \in \text{weakly hyperbolic}, \quad (2)$$

(which means the symmetric hyperbolicity requires stronger conditions to be satisfied than the others). We do not repeat each level's features here (see section 2 of [19]). However, at the strongly hyperbolic level, we can prove the finiteness of the energy norm if the characteristic matrix is independent of u (cf [16]), which is one step further than a weakly hyperbolic form.

From the point of view of numerical applications, to write down the fundamental equation in an explicitly hyperbolic form is quite attractive, not only for its mathematically well posed features. It is well known that a certain flux-conservative hyperbolic system of equations is taken as an essential formulation in the computational Newtonian hydrodynamics [28]. There is also an effort to implement the boundary condition by using the characteristic speed (eigenvalues) of the system [29].

2.2. Hyperbolic formulations of the Einstein equation

As was discussed by Geroch [30], most physical systems can be expressed as symmetric hyperbolic systems. However, the standard ADM system does not form a first-order hyperbolic system. This can be seen immediately from the fact that the ADM dynamical equations,

$$\partial_t \gamma_{ij} = -2N K_{ij} + \nabla_j N_i + \nabla_i N_j, \quad (3)$$

$$\begin{aligned} \partial_t K_{ij} = & N({}^{(3)}R_{ij} + \text{tr } K K_{ij}) - 2N K_{im} K_j^m - \nabla_i \nabla_j N \\ & + (\nabla_j N^m) K_{mi} + (\nabla_i N^m) K_{mj} + N^m \nabla_m K_{ij}, \end{aligned} \quad (4)$$

have Ricci curvature ${}^{(3)}R_{ij}$ which involves second derivatives of the 3-metric γ_{ij} by definition. (The notation here is the standard one. K_{ij} is the extrinsic curvature, N and N^i are the lapse and shift vector, respectively. ∇ denotes a covariant derivative on the 3-surface.) For our later convenience, we also write down the ADM constraint equations,

$$\mathcal{C}_H^{\text{ADM}} := {}^{(3)}R + (\text{tr } K)^2 - K_{ij} K^{ij} \approx 0, \quad (5)$$

$$\mathcal{C}_M^{\text{ADM}} := \nabla_j (K^{ij} - \gamma^{ij} \text{tr } K) \approx 0, \quad (6)$$

which are called the Hamiltonian and momentum constraint equations, respectively.

So far, several first-order hyperbolic systems of the Einstein equation have been proposed; some of them are symmetrizable (strongly hyperbolic) [9, 10] or symmetric hyperbolic systems [12, 14, 31, 32]. There are many variations in the methods for constructing higher hyperbolic systems, but the number of fundamental dynamical variables is always larger than that of ADM (see a brief example by Anderson–York (1999) in [12]). Several numerical tests are reported (as we referred to in the introduction) using a particular hyperbolic formulation, but no numerical comparisons between these formulations are reported[†].

Using Ashtekar’s formulation, we can compare three levels of hyperbolicity in the same interface (same fundamental variables) as we describe next.

2.3. Hyperbolic formulations in the Ashtekar formulations

Here we present our fundamental dynamical equations. Our notation and a more detailed review are presented in appendix A, but we repeat them here if necessary.

The new basic variables are the densitized inverse triad, \tilde{E}_a^i , and the $SO(3, C)$ self-dual connection, \mathcal{A}_i^a , where the indices i, j, \dots indicate the 3-spacetime, and a, b, \dots are for $SO(3)$ space. The total four-dimensional spacetime is described together with the gauge variables

[†] Recently, we heard that Bardeen and Buchman are preparing a numerical comparison between the formulations of Bona–Massó and Anderson–York [33]. (After we submitted this paper, we noticed that Hern [42] compares numerically different levels of hyperbolicity based on the Frittelli–Reula system [32] applying it to Gowdy spacetime. There he obtains a similar conclusion to ours.)

\tilde{N} , N^i , \mathcal{A}_0^a , which we call the densitized lapse function, shift vector and the triad lapse function. The system has three constraint equations,

$$\mathcal{C}_H^{\text{ASH}} := (i/2)\epsilon^{ab}{}_{\tilde{c}} \tilde{E}_a^i \tilde{E}_b^j F_{ij}^c \approx 0, \quad (7)$$

$$\mathcal{C}_{Mi}^{\text{ASH}} := -F_{ij}^a \tilde{E}_a^j \approx 0, \quad (8)$$

$$\mathcal{C}_{Ga}^{\text{ASH}} := \mathcal{D}_i \tilde{E}_a^i \approx 0, \quad (9)$$

which are called the Hamiltonian, momentum and Gauss constraint equations, respectively. The dynamical equations for a set of $(\tilde{E}_a^i, \mathcal{A}_i^a)$ are

$$\partial_t \tilde{E}_a^i = -i\mathcal{D}_j (\epsilon^{cb}{}_{\tilde{a}} \tilde{N} \tilde{E}_c^j \tilde{E}_b^i) + 2\mathcal{D}_j (N^{[j} \tilde{E}_a^{i]}) + i\mathcal{A}_0^b \epsilon_{ab}{}^c \tilde{E}_c^i, \quad (10)$$

$$\partial_t \mathcal{A}_i^a = -i\epsilon^{ab}{}_{\tilde{c}} \tilde{N} \tilde{E}_b^j F_{ij}^c + N^j F_{ji}^a + \mathcal{D}_i \mathcal{A}_0^a, \quad (11)$$

where $F_{ij}^a := 2\partial_{[i} \mathcal{A}_{j]}^a - i\epsilon^a{}_{bc} \mathcal{A}_i^b \mathcal{A}_j^c$ is the curvature 2-form.

We have to consider the reality conditions when we use this formalism to describe the classical Lorentzian spacetime. As we review in appendix A.2, the metric will remain on its real-valued constraint surface during time evolution automatically if we prepare initial data which satisfy the reality condition. More practically, we also require that the triad be real-valued. However, again this reality condition appears as a gauge restriction on \mathcal{A}_0^a , (A11), which can be imposed at every time step. In our actual simulation, we prepare our initial data using the standard ADM approach, so that we have no difficulties in maintaining these reality conditions.

The set of dynamical equations (10) and (11) (hereafter we call these the *original* equations) does have a weakly hyperbolic form [19], so that we regard the mathematical structure of the original equations as one step advanced from the standard ADM. Furthermore, we can construct higher levels of hyperbolic systems by restricting the gauge condition and/or by adding constraint terms, $\mathcal{C}_H^{\text{ASH}}$, $\mathcal{C}_{Mi}^{\text{ASH}}$ and $\mathcal{C}_{Ga}^{\text{ASH}}$, to the original equations [19]. We extract only the final expressions here.

In order to obtain a symmetric hyperbolic system[†], we add constraint terms to the right-hand side of (10) and (11). The adjusted dynamical equations,

$$\partial_t \tilde{E}_a^i = -i\mathcal{D}_j (\epsilon^{cb}{}_{\tilde{a}} \tilde{N} \tilde{E}_c^j \tilde{E}_b^i) + 2\mathcal{D}_j (N^{[j} \tilde{E}_a^{i]}) + i\mathcal{A}_0^b \epsilon_{ab}{}^c \tilde{E}_c^i + P^i{}_{ab} \mathcal{C}_G^{\text{ASH}b}, \quad (12)$$

where

$$\begin{aligned} P^i{}_{ab} &\equiv N^i \delta_{ab} + i\tilde{N} \epsilon_{ab}{}^c \tilde{E}_c^i, \\ \partial_t \mathcal{A}_i^a &= -i\epsilon^{ab}{}_{\tilde{c}} \tilde{N} \tilde{E}_b^j F_{ij}^c + N^j F_{ji}^a + \mathcal{D}_i \mathcal{A}_0^a + Q_i^a \mathcal{C}_H^{\text{ASH}} + R_i{}^{ja} \mathcal{C}_{Mj}^{\text{ASH}}, \end{aligned} \quad (13)$$

where

$$Q_i^a \equiv e^{-2} \tilde{N} \tilde{E}_i^a, \quad R_i{}^{ja} \equiv ie^{-2} \tilde{N} \epsilon^{ac}{}_{\tilde{b}} \tilde{E}_i^b \tilde{E}_c^j$$

form a symmetric hyperbolicity if we further require the gauge conditions,

$$\mathcal{A}_0^a = \mathcal{A}_i^a N^i, \quad \partial_i N = 0. \quad (14)$$

We note that the adjusted coefficients, $P^i{}_{ab}$, Q_i^a , $R_i{}^{ja}$, for constructing the symmetric hyperbolic system are uniquely determined, and there are no other additional terms (say, no $\mathcal{C}_H^{\text{ASH}}$, $\mathcal{C}_M^{\text{ASH}}$ for $\partial_t \tilde{E}_a^i$, no $\mathcal{C}_G^{\text{ASH}}$ for $\partial_t \mathcal{A}_i^a$) [19]. The gauge conditions, (14), are consequences of the consistency with (triad) reality conditions.

[†] Iriando *et al* [34] presented a symmetric hyperbolic expression in a different form. The differences between ours and theirs are discussed in [19, 20]

Table 1. List of systems that we compare in this paper.

	System	Variables	Equations of motion	Remark
	ADM	(γ_{ij}, K_{ij})	(3), (4)	‘Standard ADM’
I	Ashtekar (weakly hyp.)	$(\tilde{E}_a^i, \mathcal{A}_i^a)$	(10), (11) (original)	‘Original Ashtekar’
II	Ashtekar (strongly hyp.)	$(\tilde{E}_a^i, \mathcal{A}_i^a)$	(12), (13) (adjusted)	(15) required
III	Ashtekar (symmetric hyp.)	$(\tilde{E}_a^i, \mathcal{A}_i^a)$	(12), (13) (adjusted)	(14) required
(In section 5)	Ashtekar (adjusted)	$(\tilde{E}_a^i, \mathcal{A}_i^a)$	(31), (32) (adjusted with κ)	

We can also construct a strongly (or diagonalizable) hyperbolic system by restricting to a gauge $N^l \neq 0, \pm N\sqrt{\gamma^{ll}}$ (where γ^{ll} is the 3-metric and we do not sum indices here) for the original equations (10) and (11). Or we can also construct from the adjusted equations, (12) and (13), together with the gauge condition

$$\mathcal{A}_0^a = \mathcal{A}_i^a N^i. \quad (15)$$

As for the strongly hyperbolic system, hereafter we take the latter expression.

In table 1, we have summarized the equations to be used throughout the remainder of this paper.

3. Numerical method

3.1. Overview

We coded up the program so as to compare the evolutions of spacetime with different sets of dynamical equations but with the common conditions: the same initial data, the same boundary conditions, the same slicing condition and the same evolution scheme.

We consider the plane-symmetric vacuum spacetime without a cosmological constant. This spacetime has the true freedom of gravitational waves of two polarized (+ and \times) modes. We apply the periodic boundary conditions to remove any difficulties caused by the numerical treatment of the boundary conditions. The initial data are given by solving constraint equations in ADM variables, using the standard conformal approach by O’Murchadha and York [35]. When we use Ashtekar’s variables for evolution, we transform the ADM initial data in terms of Ashtekar’s variables. The results are analysed by monitoring the violation of the constraint equations which are expressed using the same (or transformed if necessary) variables.

We describe our procedures in the following subsections in detail.

3.2. Metric and the initial data construction

We consider the plane-symmetric metric,

$$ds^2 = (-N^2 + N_x N^x) dt^2 + 2N_x dx dt + \gamma_{xx} dx^2 + \gamma_{yy} dy^2 + \gamma_{zz} dz^2 + 2\gamma_{yz} dy dz \quad (16)$$

where the components are the function of $N(x, t)$, $N_x(x, t)$, $\gamma_{xx}(x, t)$, $\gamma_{yy}(x, t)$, $\gamma_{zz}(x, t)$, $\gamma_{yz}(x, t)$. N and N^x are called the lapse function and the shift vector.

We prepare our initial data by solving the ADM constraint equations, (5) and (6), using the conformal approach [35]. Since we consider only the vacuum spacetime, the input quantities are the initial guess of the 3-metric $\hat{\gamma}_{ij}$, the trace part of the extrinsic curvature $\text{tr} K$, and the transverse traceless part of the extrinsic curvature \hat{A}_{TT} . For simplicity, we impose $\hat{A}_{TT} = 0$

and $\text{tr } K = K_0$ (constant). The Hamiltonian constraint then becomes an equation for the conformal factor, ψ :

$$8\hat{\Delta}\psi := 8\frac{1}{\sqrt{\hat{\gamma}}}\partial_i(\hat{\gamma}^{ij}\sqrt{\hat{\gamma}}\partial_j\psi) = \hat{R}\psi + \frac{2}{3}(K_0)^2\psi^5, \tag{17}$$

where $\hat{\gamma} = \det \hat{\gamma}_{ij}$. The momentum constraint is automatically satisfied by assumption. The initial dynamical quantities γ_{ij} , K_{ij} are given by the conformal transformation,

$$\gamma_{ij} = \psi^4\hat{\gamma}_{ij}, \quad K_{ij} = \frac{1}{3}\psi^4\hat{\gamma}_{ij}K_0. \tag{18}$$

We solve (17) under the periodic boundary conditions using the incomplete Cholesky conjugate gradient (ICCG) method.

We should remark here that we have to assume non-zero K_0 for a model of gravitational pulse waves under the periodic boundary conditions in this plane-symmetric spacetime. This can be seen as follows. Suppose we set $K_0 = 0$. From (17), we obtain

$$\partial_x\psi = \frac{1}{\sqrt{\gamma}g^{xx}}\int\sqrt{\gamma}R\psi dx. \tag{19}$$

If we set the boundary as $x = [A, B]$ and impose the periodic boundary conditions, then equation (19) becomes

$$\partial_x\psi|_{x=A} - \partial_x\psi|_{x=B} = \left(\frac{1}{\sqrt{\gamma}g^{xx}}\right)_{x=A=B}\int_B^A\sqrt{\gamma}R\psi dx. \tag{20}$$

However, when there exists a gravitational wave pulse which produces $R \neq 0$ in the region, this equation gives $\partial_x\psi|_{x=A} = \partial_x\psi|_{x=B}$, which is inconsistent with the periodic boundary conditions. Therefore, we need to assume non-zero K_0 in order to compensate the curvature which is produced by the pulse waves.

Actually, the trace of the extrinsic curvature appears only in the quadratic form, so we can interpret that our (background) spacetime is either expanding, $K_0 < 0$, or contracting, $K_0 > 0$. However, this fact indicates that there is no known exact solution to compare with. If the background spacetime is allowed to be flat ($K_0 = 0$), then we know there is a series of exact solutions which describes a collision of plane gravitational waves which were originally found by Szekeres and Khan–Penrose [36]. The formation of a curvature singularity after such colliding waves is known to be generic, but that is not generalized to the expanding background (as discussed using numerical simulations [37, 38]).

We can set two different modes of gravitational waves. One is the +-mode waves, which is given by setting a conformal guess metric as (in a matrix form)

$$\hat{\gamma}_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ \text{sym.} & 1 + a \exp(-b(x - c)^2) & 0 \\ \text{sym.} & \text{sym.} & 1 - a \exp(-b(x - c)^2) \end{pmatrix} \tag{21}$$

where a, b, c are parameters. The other is the \times -mode waves, given by

$$\hat{\gamma}_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ \text{sym.} & 1 & a \exp(-b(x - c)^2) \\ \text{sym.} & \text{sym.} & 1 \end{pmatrix} \tag{22}$$

where a, b, c are parameters again. In both cases we expect nonlinear behaviour when the wave's curvature becomes quite large compared with the background. In the collision of a +-mode wave and a \times -mode wave, we also expect to see the mode-mixing phenomena which is known as the gravitational Faraday effect. These effects are confirmed in our numerical simulations.

3.3. Transformation of variables: from ADM to Ashtekar

We need to transform the dynamical variables on the initial data when we evolve them in the connection variables. We list the procedure to obtain $(\tilde{E}_a^i, \mathcal{A}_i^a)$ from (γ_{ij}, K_{ij}) . This procedure is also used when we evaluate the constraints, $C_H^{\text{ASH}}, C_{Mi}^{\text{ASH}}, C_{Ga}^{\text{ASH}}$ for the data evolved using ADM variables.

From the 3-metric γ_{ij} to \tilde{E}_a^i :

- (a) Define the triad E_i^a corresponding to the 3-metric γ_{ij} . We take

$$E_i^a = \begin{bmatrix} E_x^1 & E_y^1 & E_z^1 \\ E_x^2 & E_y^2 & E_z^2 \\ E_x^3 & E_y^3 & E_z^3 \end{bmatrix} = \begin{bmatrix} \sqrt{\gamma_{xx}} & 0 & 0 \\ 0 & e_{22} & e_{23} \\ 0 & e_{32} & e_{33} \end{bmatrix}, \quad (23)$$

and set simply $e_{23} = e_{32}$. The relation between the metric and the triad becomes

$$e_{22}^2 + e_{33}^2 = \gamma_{yy}, \quad e_{23}^2 + e_{33}^2 = \gamma_{zz}, \quad (e_{22} + e_{33})e_{23} = \gamma_{yz}. \quad (24)$$

For the case of +-mode waves, we define naturally, $e_{22} = \sqrt{\gamma_{yy}}$, $e_{33} = \sqrt{\gamma_{zz}}$, $e_{23} = 0$. For \times -mode waves, we also take a natural set of definitions, $e_{22} = e_{33} = [(\gamma_{yy} + (\gamma_{yy}^2 - \gamma_{yz}^2)^{1/2})/2]^{1/2}$ and $e_{23} = \gamma_{yz}/2e_{22}$ which are given by solving $e_{22}^2 + e_{33}^2 = \gamma_{yy}$ and $2e_{22}e_{23} = \gamma_{yz}$.

- (b) Obtain the inverse triad E_a^i from triad E_i^a .
(c) Calculate the density, e , as $e = \det E_i^a$.
(d) Obtain the densitized triad, $\tilde{E}_a^i = e E_i^a$.

From 3-metric (γ_{ij}, K_{ij}) to \mathcal{A}_i^a :

- (a) Prepare the triad E_i^a and its inverse E_a^i .
(b) Calculate the connection 1-form $\omega_i^{bc} = E^{b\mu} \nabla_i E_\mu^c$. This is expressed only using partial derivatives as[†]

$$\omega_i^{bc} = E^{jb} \partial_{[i} E_{j]}^c - E_{id} E^{kb} E^{jc} \partial_{[k} E_{j]}^d + E^{jc} \partial_{[j} E_{i]}^b. \quad (25)$$

- (c) $\mathcal{A}_i^a = -K_{ij} E^{ja} - \frac{1}{2} \epsilon^a{}_{bc} \omega_i^{bc}$.

3.4. Transformation of variables: from Ashtekar to ADM

In contrast to the previous transformation, we also need to obtain (γ_{ij}, K_{ij}) from $(\tilde{E}_a^i, \mathcal{A}_i^a)$ when we evaluate the metric output or ADM constraints when we evolve the spacetime using connection variables. This process is only required at evaluation times, i.e. not required at every time step (unless we use the gauge condition which is primarily defined using ADM quantities).

From the densitized inverse triad \tilde{E}_a^i to the 3-metric γ_{ij} :

[†] This is from the definitions, $\omega_i^{bc} := E^{jb} \nabla_i E_j^c$ and $\omega^{abc} := E^{ja} \omega_j^{bc}$, and a relation

$$3\omega^{[abc]} - 2\omega^{[bc]a} = \omega^{a[bc]} + \omega^{b[ca]} + \omega^{c[ab]} - \omega^{abc} + \omega^{cba} = \omega^{abc}.$$

Using the densitized triad, equation (25) can also be expressed as

$$\omega_i^{bc} = \frac{2}{e^2} \tilde{E}^{jb} (\partial_{[i} \tilde{E}_{j]}^c) + \frac{1}{e^4} \tilde{E}^{jb} \tilde{E}_i^c \tilde{E}_k^a (\partial_j \tilde{E}_a^k) + \frac{1}{4e^4} \tilde{E}_{ia} \tilde{E}^{kb} \tilde{E}_c^j (\partial_j \tilde{E}_k^a), \quad \text{taking } [bc].$$

- (a) Calculate the density e as $e = (\det \tilde{E}_a^i)^{1/2}$.
- (b) Get the inverse 3-metric as $\gamma^{ij} = \tilde{E}_a^i \tilde{E}_a^j / e^2$.
- (c) Obtain γ_{ij} .

From $(\tilde{E}_a^i, \mathcal{A}_i^a)$ to the extrinsic curvature K_{ij} :

- (a) Prepare the un-densitized inverse triad, $E_a^i = \tilde{E}_a^i / e$.
- (b) Prepare the triad E_i^a .
- (c) Calculate the connection 1-form $\epsilon^a{}_{bc} \omega_i^{bc}$.
- (d) Calculate Z_i^a , which is defined as[†] $Z_i^a := -\mathcal{A}_i^a + \frac{1}{2} i \epsilon^a{}_{bc} \omega_i^{bc}$ ($= K_{ij} E^{ja}$), and obtain $K_{ij} = Z_i^a E_{ja}$.

3.5. Gauge conditions

We evolve the initial data with different evolution equations and compare its accuracy/stability. As we summarized in table 1, we will compare time evolutions between ADM and Ashtekar (of the original system I) in section 4.1, and three of Ashtekar's systems (I, II and III: weakly, strongly and symmetric) in section 4.2. We, then, consider an alternative system (an adjusted κ system) in section 5.

Here we comment again on our choice of the slicing (gauge) condition. As for the primary tests of this subject, we apply the simplest slicing conditions we can take. That is,

- (1) the simplest geodesic slicing condition for the lapse function;
- (2) the simplest zero shift vector $N^x = 0$; and
- (3) the natural choice of triad lapse function $\mathcal{A}_0^a = \mathcal{A}_i^a N^i$ ($= 0$ if $N^x = 0$, which is suggested from (14) or (15)).

However, in the Ashtekar formalism, the densitized lapse function \underline{N} is the fundamental gauge quantity (rather than N). Therefore, we try two conditions for the lapse:

- (1a) the standard geodesic slicing condition $N = 1$, which will be transformed to $\underline{N} = 1/e$ when we apply this condition in Ashtekar's evolution system; and
- (1b) the densitized geodesic slicing condition $\underline{N} = 1$, which will be transformed to $N = e$ when we evolve the system using ADM equations.

In practice, such a transformation using the density e will not guarantee that the Courant condition holds if we fix the time evolution step Δt [‡]. Therefore, we need to rescale the transformed lapse (\underline{N} in (1a), N in (1b)) so that it has a maximum value of unity, in order to keep our evolution system stable.

If we apply the standard geodesic slice, then we can compare the weakly hyperbolic system with the symmetric hyperbolic one. Similarly, if we apply the densitized geodesic slice, then we can compare the (original) weakly hyperbolic system with the strongly hyperbolic one.

[†] This is from the original definition of $\mathcal{A}_i^a, \mathcal{A}_i^a := \omega_i^{0a} - (i/2)\epsilon^a{}_{bc} \omega_i^{bc}$.

[‡] Here we remind the reader of the stability condition, $N \Delta t \leq \Delta x$ for a standard forward-time centred-space (FTCS) scheme for a simple wave equation, in a $(\Delta t, \Delta x)$ -spaced numerical grid. Note that this condition will be changed due to the choice of the evolution scheme and the equations of the system.

3.6. Time integrating scheme

We applied two second-order evolution schemes, and confirmed that they give us nearly identical results.

One is the so-called iterative Crank–Nicholson scheme (cf [39]), which is now becoming the standard in the numerical relativity community. Suppose we have a dynamical equation in the form

$$\partial_t u(x, t) = f(u(x, t), \partial_x u(x, t)). \quad (26)$$

Then the scheme for updating u at a point x from t to $t + \Delta t$ consists of the following steps.

- (a) Use data on $t = t$ for the right-hand side and update $u(x, t)$ for the Δt step as $\tilde{u}(x, t + \Delta t)$,

$$\frac{\tilde{u}(x, t + \Delta t) - u(x, t)}{\Delta t} = f(u(x, t), \partial_x u(x, t)). \quad (27)$$

- (b) Take the average of $u(x, t)$ and $\tilde{u}(x, t + \Delta t)$, and let it represent a half-step value (say $v(x, t + \Delta t/2)$).
- (c) Update $u(x, t)$ for the Δt step again using $\hat{u}(x, t + \Delta t/2)$ in the argument of the right-hand side,

$$\frac{\tilde{u}(x, t + \Delta t) - u(x, t)}{\Delta t} = f(\hat{u}(x, t + \Delta t/2), \partial_x \hat{u}(x, t + \Delta t/2)). \quad (28)$$

- (d) Perform the above steps (b) and (c) once again (we assume a two-iteration Crank–Nicholson scheme), and take $\tilde{u}(x, t + \Delta t)$ to be the evolved quantity.

The other scheme we applied is the Brailovskaya integration scheme, which is a second-order predictor–corrector method [40] and is rather easy to code. The first step (predictor step) is the same as (27), and the second step (corrector step) simply switches the right-hand side using the updated $\tilde{u}(x, t + \Delta t)$ to be

$$\frac{u(x, t + \Delta t) - u(x, t)}{\Delta t} = f(\tilde{u}(x, t + \Delta t), \partial_x \tilde{u}(x, t + \Delta t)). \quad (29)$$

Note that all derivatives here on the right-hand side are assumed to use a central difference.

The latter scheme is quite simple, but gives us reasonably accurate and stable evolutions for our problems. We confirmed that both give us nearly identical evolutions (which will be shown in figure 2(b)), but the Brailovskaya method requires less computational time.

3.7. Checking the constraints

We compare the violation of the constraint equations during the time evolution. We have ADM constraint equations, $\mathcal{C}_H^{\text{ADM}}$ and $\mathcal{C}_{Mi}^{\text{ADM}}$ (equations (5) and (6)), and also Ashtekar’s constraint equations, $\mathcal{C}_H^{\text{ASH}}$, $\mathcal{C}_{Mi}^{\text{ASH}}$ and $\mathcal{C}_{Ga}^{\text{ASH}}$ (equations (7)–(9), respectively). By means of the transformation between (γ_{ij}, K_{ij}) and $(\tilde{E}_a^i, \mathcal{A}_i^a)$, we can evaluate ADM constraints even if we evolved the system using Ashtekar’s variables and vice versa.

We measure the violation of a constraint by its

- (a) maximum, $\max_x |\mathcal{C}(x)|$,
- (b) L1 norm, $(\sum_{x=1}^{n_x} \mathcal{C}(x))/n_x$, and
- (c) L2 norm, $(\sum_{x=1}^{n_x} |\mathcal{C}(x)|^2/n_x)^{1/2}$,

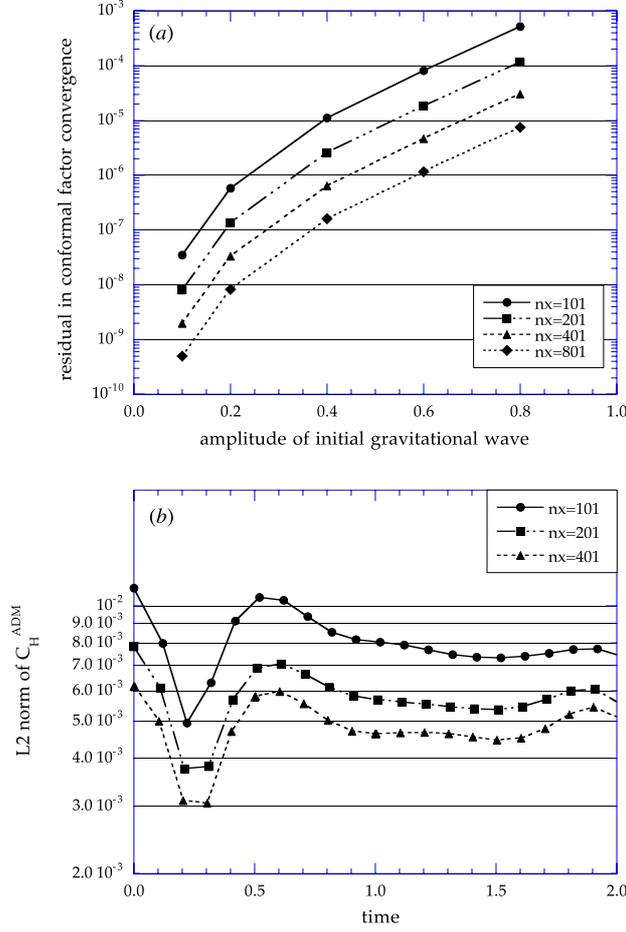


Figure 1. Examples of the convergence tests. (a) Convergence of the initial data solver (Hamiltonian constraint (17) solver). We plot the residual of the conformal factor, L2 norm of $|(\psi_n - \psi_{n-1})/\psi_n|^2$, when it converges, where n is the iteration number in the ICCG routine. The horizontal value is the amplitude of the gravitational wave (a in equation (21)) where we assume $+$ -mode single-pulse wave and fix $b = 2.0$, $c = 0.0$ in equation (21), and $K_0 = -0.025$. According to the resolutions (grid points = 101, . . . , 801 for the range of $x = [-5, +5]$), we see second-order convergence. (b) Convergence behaviour of ADM evolution code. L2 norms of C_H^{ADM} is plotted for the above initial data for the amplitude $a = 0.2$. We applied geodesic slicing condition ($N = 1$). (c) Convergence behaviour of Ashtekar evolution code. L2 norms of C_H^{ASH} is plotted for the above initial data for the amplitude $a = 0.2$. We applied geodesic slicing condition ($N = 1$). We can see clearly that the error norms in evolutions will decrease in high-resolution cases.

where n_x is the number of grid points. When we compare them during the evolution, we measure them at the same proper time, τ , for the two different evolution systems. The proper time is defined locally as $d\tau = N dt$, which is also $d\tau = e\tilde{N} dt$, but here we apply its averaged value on the whole $t = \text{constant}$ surface, (say $\langle N \rangle = (1/n_x) \sum_{x=1}^{n_x} N$),

$$\tau = \int_0^t \langle N \rangle dt, \quad \text{or} \quad \tau = \int_0^t \langle e\tilde{N} \rangle dt, \quad (30)$$

to characterize the ‘time’ of evolution.

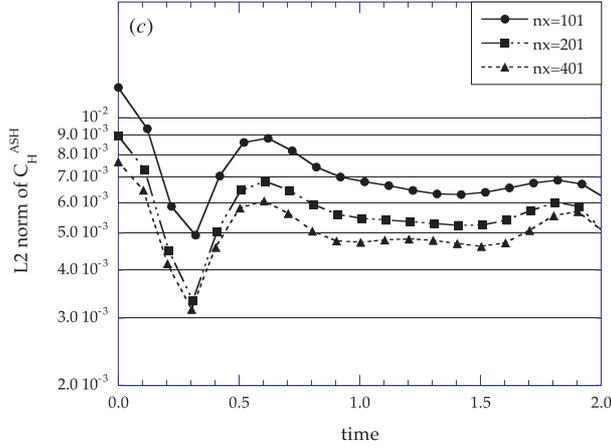


Figure 1. Continued.

The numerical code passed convergence tests, and the results shown in this paper are all obtained with acceptable accuracy. In figure 1, we show a result of convergence tests. We show the convergence behaviour of our initial data solver in figure 1(a), together with the convergence behaviour of the evolution codes both for ADM and Ashtekar variables in figures 1(b) and (c). We plotted the residual of the Hamiltonian constraint solver when it was minimized, and the L2 norm of C_H^{ADM} and C_H^{ASH} for the evolution of a +-mode single-pulse wave (the model is described in the next section). We can see that all errors are diminished by finer resolutions. The order of the convergence[†] is 1.98 to 2.01 for the initial data solver, and at best 1.96 (e.g. at $\tau = 0.5$) for the evolution code.

All the results we present in this paper are obtained using 401 grid points for the range $x = [-5, +5]$; that is gravitational waves traverse the entire numerical region in proper time 10 if the background expansion K_0 is close to zero. We use the Courant number $\nu = \Delta t / \Delta x = 0.2$.

We coded all our fundamental quantities (metric, gauge variables, etc) as complex, but we observed that the evolution from our initial data never violate its metric reality conditions. Due to the gauge condition for \mathcal{A}_0^q (A11), we also confirmed that our evolutions preserve the triad reality condition.

4. Experiments 1: differences between hyperbolicities

In this section, we examine the accuracy/stability of the numerical evolutions by comparing the different hyperbolic systems. We begin by showing how the evolution of Ashtekar's equations look, comparing with those of the ADM equations.

4.1. ADM versus Ashtekar

We start by describing our model—plane-wave propagation in an expanding/collapsing spacetime. We prepare the initial data with one or two gravitational pulse waves in our numerical region. The pulses then start propagating in both $\pm x$ directions at the speed of light, and appear on the other side of the numerical region due to the periodic boundary condition.

[†] Here we used the definition of the order of convergence following Bona *et al* [29].

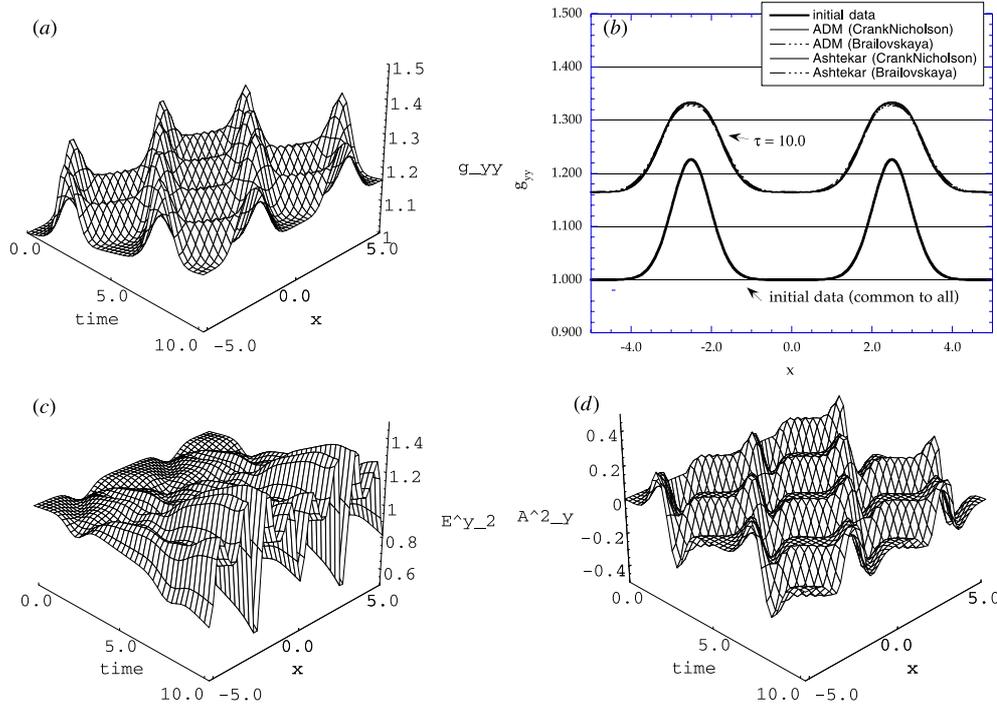


Figure 2. Images of gravitational wave propagation and comparisons of dynamical behaviour of Ashtekar’s variables and ADM variables. We applied the same initial data of two +-mode pulse waves ($a = 0.2, b = 2.0, c = \pm 2.5$ in equation (21) and $K_0 = -0.025$), and the same slicing condition, the standard geodesic slicing condition ($N = 1$). (a) Image of the 3-metric component g_{yy} of a function of proper time τ and coordinate x . This behaviour can be seen identically both in ADM and Ashtekar evolutions, and both with the Brailovskaya and Crank–Nicholson time-integration scheme. Part (b) explains this fact by comparing the snapshot of g_{yy} at the same proper time slice ($\tau = 10$), where four lines at $\tau = 10$ are looked at identically. Parts (c) and (d) are of the real part of the densitized triad \tilde{E}_2^y , and the real part of the connection A_y^2 , respectively, obtained from the evolution of the Ashtekar variables.

When the pulses collide, then the amplitude seems simply to double, as they are superposed, and the pulses keep travelling in their original propagation direction. That is, we observe something like solitonic wave pulse propagation.

As we mentioned in section 3.2, we have to assume our background not to be flat, therefore there are no exact solutions. The reader might think that if we set $|\text{tr } K|$ to be small and pulse wave shapes to be quite sharp then our simulations will be close to the analytic colliding plane-wave solutions which produce the curvature singularity. However, from the numerical side, these two requirements are contradictory (e.g. sharp wave input produces large curvature which should be compensated by $|\text{tr } K|$ in order to construct our initial data). Thus it is not so surprising that our waves propagate like solitons, not forming a singularity.

In figure 2(a), we plot an image of wave propagation (a metric component g_{yy}) up to $\tau = 10$, of +-mode pulse waves initially located at $x = \pm 2.5$. We took a small negative K_0 , so that the background spacetime is slowly expanding.

Figure 2(b), then, tells us that our ADM evolution code and Ashtekar’s variable code give us identical evolutions. We plotted a snapshot of g_{yy} on the initial data (which is common to all models here), and its snapshot at $\tau = 10.0$. The fact that all four lines (ADM/Ashtekar, of

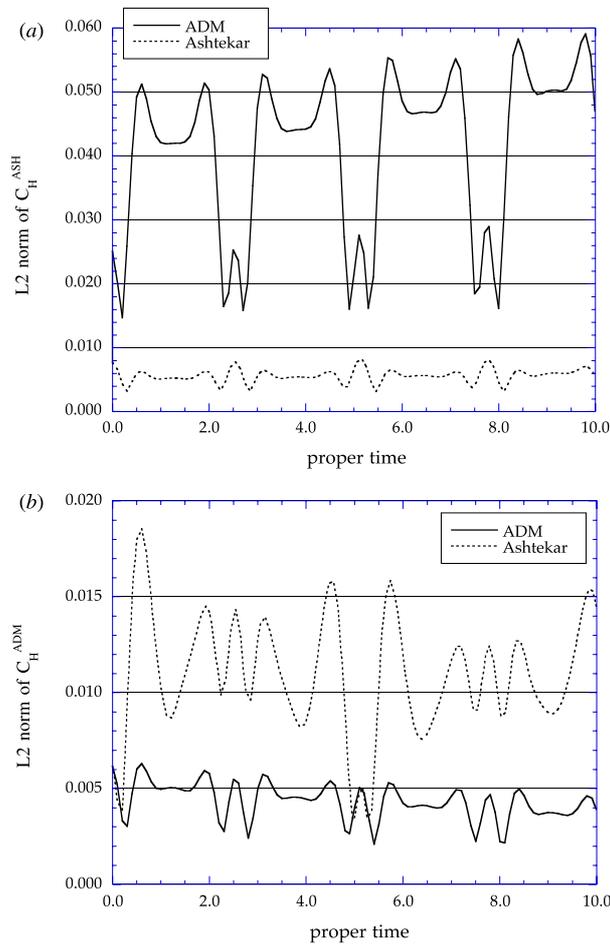


Figure 3. Comparisons of the constraint violation of Ashtekar's equation with that of ADM. (a) L2 norm of Ashtekar's Hamiltonian constraint equation, C_H^{ASH} as a function of averaged proper time. (b) L2 norm of the ADM Hamiltonian constraint equation, C_H^{ADM} as a function of averaged proper time. The plots are of the same parameters as those of figure 2.

their Brailovskaya/Crank–Nicholson evolution schemes) overlapped clearly indicates that we are showing exact evolutions.

We also plotted a typical evolution of the fundamental dynamical quantities \tilde{E}_2^y and \mathcal{A}_y^2 in figures 2(c) and (d).

Next we compare constraint violations by Ashtekar's equation with that of ADM. In figure 3, we plot the L2 norm of C_H^{ASH} and C_H^{ADM} . We see that ADM evolution shows less violation in measuring C_H^{ADM} , and the Ashtekar evolution shows less violation in measuring C_H^{ASH} . The magnitudes of these violations are similar. Thus, we believe that these violations are within the numerical truncation errors in the process of numerical transformation of variables (ADM to Ashtekar/ Ashtekar to ADM), and therefore it is not appropriate to conclude here which formulation is better.

As the reader may guess, the violations of constraints reduce if the background spacetime is expanding ($K_0 < 0$). Therefore, we will use the collapsing background spacetime ($K_0 > 0$)

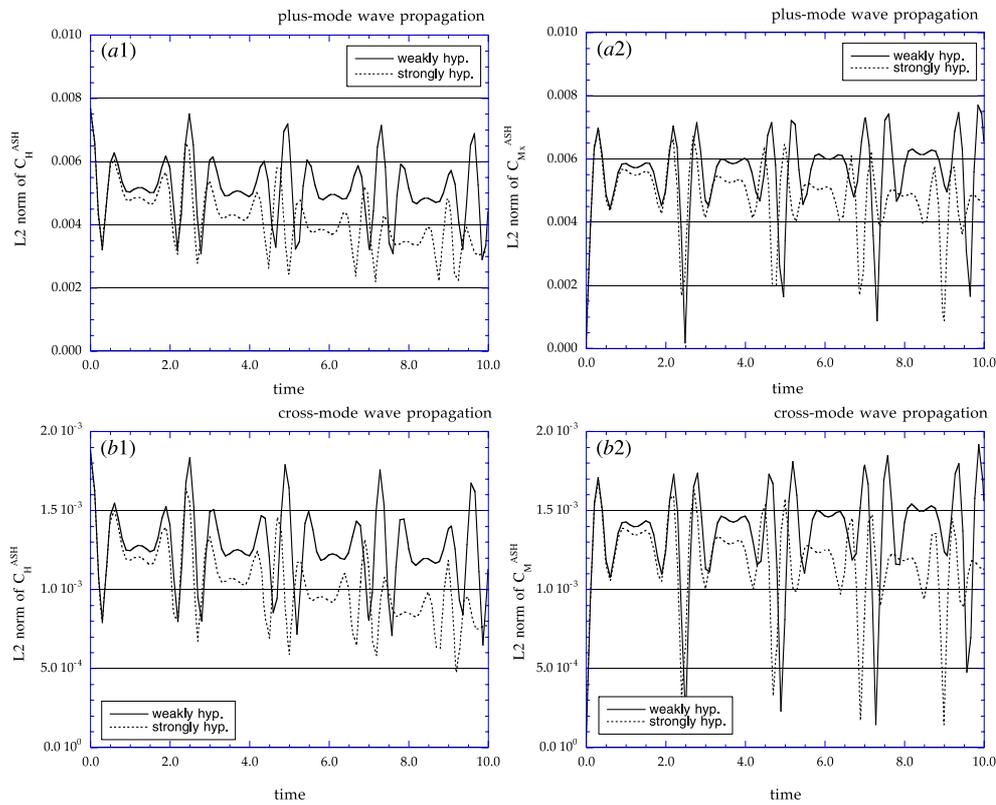


Figure 4. Comparisons of the strongly hyperbolic system (Ashtekar II) with the weakly hyperbolic system (Ashtekar original) ($N = 1$ slice). Panels (a) are of +-mode waves ($a = 0.2, b = 2.0, c = \pm 2.5$ in equation (21)), while (b)s are of \times -mode waves ($a = 0.1, b = 2.0, c = \pm 2.5$ in equation (22)), in a background spacetime with $K_0 = 0.025$). We plot the L2 norm of the Hamiltonian and momentum constraints, C_H^{ASH} and C_M^{ASH} , for each case. We see from all of them that a strongly hyperbolic system improves the violation of the constraints.

hereafter for presentations, with the expectation of having more nonlinear effects; however, this direction also stops the evolution after the finite time (e.g. for flat initial data with $K_0 = +0.025$, the spacetime will collapse to zero volume at around $t = 60$).

4.2. Comparison between hyperbolicities

Here we present our comparisons of the accuracy and/or stability between the different hyperbolicities. Since all examples we show in this section are not for the case of unstable evolution (no exponential growth of the constraint violation), our experiments can be said to be comparisons of the accuracy of the evolution, conservatively.

We first compare the (original) weakly hyperbolic system (system I in table 1) with the strongly hyperbolic system (system II in table 1). This comparison can be made under the densitized geodesic slicing condition, $N = 1$. We prepare two initial gravitational pulses (both +-+ or $\times-\times$ modes) and take our background spacetime to be collapsing ($K_0 > 0$). In figure 4, we show the constraint errors, C_H^{ASH} and C_M^{ASH} . In both two situations, we observe that the strongly hyperbolic system has slightly improved the violation of the constraints, but

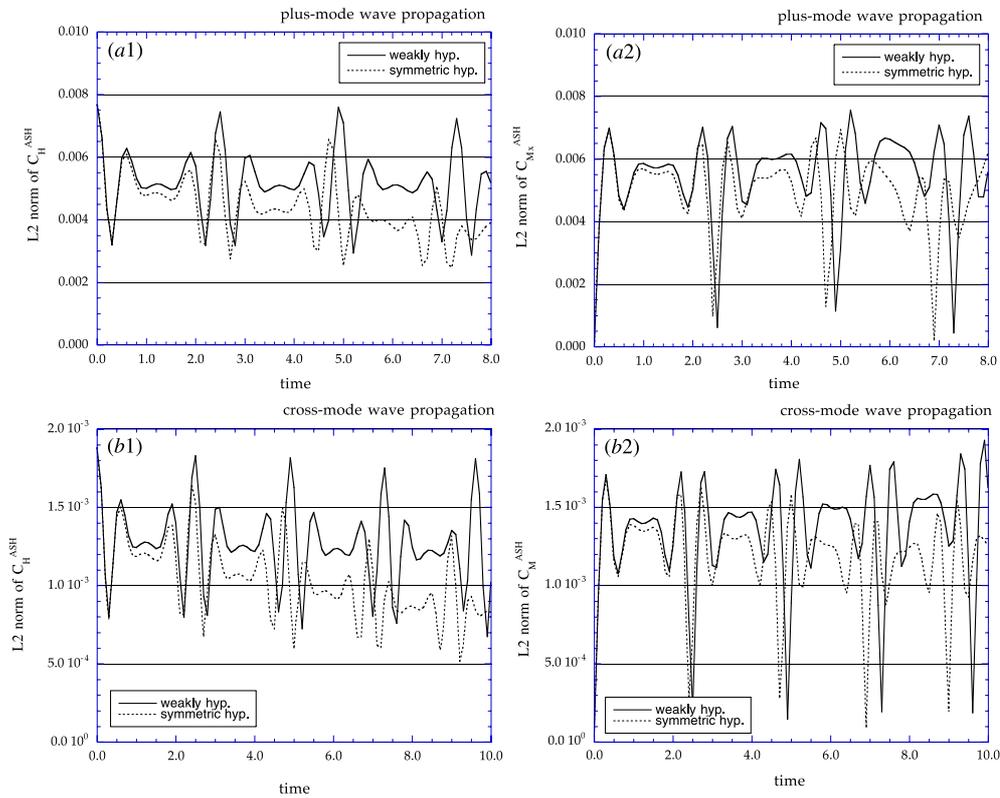


Figure 5. Comparisons of the symmetric hyperbolic system (Ashtekar III) with the weakly hyperbolic system (Ashtekar original) ($N = 1$ slice), in the same way as in figure 4. We applied the same parameters as those of figure 4. Figures (a1) and (a2) are of $+$ -mode wave propagation and the L2 norm of C_H^{ASH} and C_M^{ASH} , respectively. Figures (b1) and (b2) are of \times -mode wave propagation and the L2 norm of C_H^{ASH} and C_M^{ASH} , respectively. We see from all of them that a symmetric hyperbolic system improves the violation of the constraints.

we cannot see the orders of magnitude differences.

Similarly, we next compare the (original) weakly hyperbolic system (system I in table 1) with the symmetric hyperbolic system (system III in table 1). This comparison can be made under the standard geodesic slicing condition, $N = 1$. We repeat the same experiments as above and show plots in figure 5. We again see that the symmetric hyperbolic system slightly improves the situation, but not so drastically.

From both figures 4 and 5, we see that the strongly and symmetric hyperbolic systems produce less violation of constraints than the original weakly hyperbolic system. Therefore, one conclusion is that adjusting the equation of motion with constraint terms *does* definitely make the system accurate. However, the constraint violation remains the same order of magnitude.

From each figure, we may conclude that a higher-level hyperbolic system gives us slightly more accurate evolutions. However, if we evaluate the magnitude of the L2 norms, then we also conclude that there is no measurable differences between strongly and symmetric hyperbolicities. This last fact will be supported more affirmatively in the next experiment.

5. Experiments 2: another way to control the accuracy/stability

The results we have presented in the previous section indicate that both strongly and symmetric hyperbolic systems show better performance than the original weakly hyperbolic system. These systems are obtained by adding constraint terms (or ‘adjusted’ terms) to the right-hand side of the original equations, (10) and (11). In this section, we report on simple experiments in changing the magnitude of the multipliers of such adjusted terms.

We consider the following system, where the equations of motion are adjusted in the same way as before, but with a real-valued constant multiplier κ :

$$\partial_t \tilde{E}_a^i = -i\mathcal{D}_j(\epsilon^{cb} N \tilde{E}_c^j \tilde{E}_b^i) + 2\mathcal{D}_j(N^{[j} \tilde{E}_a^{i]}) + i\mathcal{A}_0^b \epsilon_{ab}{}^c \tilde{E}_c^i + \kappa P^i{}_{ab} C_G^{\text{ASH}b}, \quad (31)$$

$$\text{where} \quad P^i{}_{ab} \equiv N^i \delta_{ab} + iN \epsilon_{ab}{}^c \tilde{E}_c^i,$$

$$\partial_t \mathcal{A}_i^a = -i\epsilon^{ab}{}^c N \tilde{E}_b^j F_{ij}^c + N^j F_{ji}^a + \mathcal{D}_i \mathcal{A}_0^a + \kappa Q_i^a C_H^{\text{ASH}} + \kappa R_i{}^{ja} C_{Mj}^{\text{ASH}}, \quad (32)$$

$$\text{where} \quad Q_i^a \equiv e^{-2} N \tilde{E}_i^a, \quad R_i{}^{ja} \equiv ie^{-2} N \epsilon^{ac}{}_{b} \tilde{E}_i^b \tilde{E}_c^j.$$

The set of equations (31) and (32) becomes the original weakly hyperbolic system if $\kappa = 0$, becomes the symmetric hyperbolic system if $\kappa = 1$ and $N = \text{constant}$, and remains a strongly hyperbolic system for other choices of κ except $\kappa = \frac{1}{2}$ which only forms a weakly hyperbolic system. We again remark that the coefficients for constructing the symmetric hyperbolic system are uniquely determined.

We tried the same evolutions as in the previous section for different value of κ . In figure 6, we plot the L2 norm of the Hamiltonian and momentum constraint equations, C_H^{ASH} and C_M^{ASH} . We checked first that $\kappa = 0$ and 1 produce the same results as those of weakly and symmetric hyperbolic systems. What is interesting is the case of $\kappa = 2$ and 3. These κ s produce better performance than the symmetric hyperbolic system, although these cases are of strongly hyperbolic levels. Therefore, as far as monitoring the violation of the constraints is concerned, we may say that the symmetric hyperbolic form is *not always the best*. We note that the negative κ will produce unstable evolution as we plotted, while too a large positive κ will also result in unstable evolution in the end (see the $\kappa = 3$ lines).

We also tried similar experiments with the vacuum Maxwell equation. The original Maxwell equation has a symmetric hyperbolicity, and additional constraint terms (with multiplier κ) reduce the hyperbolicity to the strong or weak level. We show the details and a figure in appendix B, but in short there may be no measurable differences between strongly and symmetric hyperbolicities.

These experiments in changing κ are now reported in our paper II [41] more extensively. There, we propose a plausible explanation as to why such adjusted terms work for stabilizing the system. We introduce the idea in appendix C. Briefly, we will conjecture a criterion using the eigenvalues of the ‘adjusted version’ of the constraint propagation equations. This analysis may explain the appearance of phase differences between two systems, which is observed in figures 4–6.

6. Discussion

Motivated by many recent proposals for hyperbolic formulations of the Einstein equation, we studied numerically these accuracy/stability properties with the purpose of comparing three mathematical levels of hyperbolicity: weakly hyperbolic, strongly hyperbolic and symmetric hyperbolic systems. We apply Ashtekar’s connection formulation, because this is the only known system in which we can compare three hyperbolic levels with the same interface.

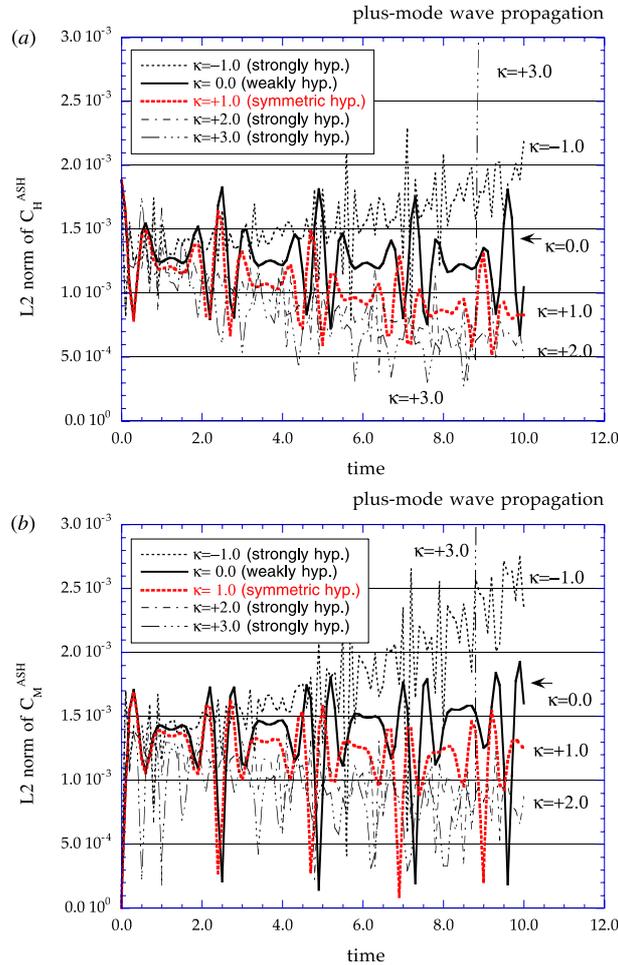


Figure 6. Comparisons of the ‘adjusted’ system with the different multiplier, κ , in equations (31) and (32). The model uses +-mode pulse waves ($a = 0.1$, $b = 2.0$, $c = \pm 2.5$) in equation (21) in a background $K_0 = -0.025$. Plots are of the L2 norm of the Hamiltonian and momentum constraint equations, C_H^{ASH} and C_M^{ASH} ((a) and (b), respectively). We see some κ produce a better performance than the symmetric hyperbolic system.

Our numerical code demonstrates gravitational wave propagation in plane-symmetric spacetime, and we compare the ‘accuracy’ and/or ‘stability’ by monitoring the violation of the constraints. Actually, our experiments in section 4 were the comparisons of accuracy in evolutions, while in section 5 we observed cases of unstable evolution. By comparing with the results obtained from the weakly hyperbolic system, we observe that the strongly and symmetric hyperbolic system show better properties with little differences between them. Therefore, we may conclude that higher levels of hyperbolic formulations help the numerics more, though the differences are small.

However, we also found that the symmetric hyperbolic system is not always the best one for controlling accuracy or stability, by introducing a multiplier for adjusted terms in the equations of motion. This result suggests that a certain kind of hyperbolicity is enough to control the violation of the constraint equation. In our case it is the strongly hyperbolic

level. This statement is supported by an experiment in the Maxwell system as we describe in appendix B.

The remaining question is: why can we obtain better performance by adding constraint terms in the dynamical equations? The added terms are basically *error* terms during the evolution for the original dynamical equations. Nevertheless, these terms improve the accuracy of the evolution. We now have a plausible way to explain the reason which is discussed in our paper II [41] (a brief introduction is given in appendix C in this paper). There we evaluate the eigenvalues of the adjusted version of the constraint propagation equations, and propose a criteria for obtaining the stability of the system. In some cases, for example, the decay/growth of the constraints can predict the signature of the eigenvalues of the adjusted version of the constraint propagation equations. In [41], we will discuss this point in detail together with a numerical demonstration of λ -systems [21, 22]. There we also show that some choices of adjusted terms may produce an unstable evolution.

To conclude, we are glad to announce that Ashtekar's connection variables have finally been applied in numerical simulations. This new approach, we hope, will contribute to further understanding of gravitational physics, and will open a new window for peeling off interesting nonlinear natures together with a step towards a numerical treatment of quantum gravity.

Acknowledgments

HS appreciates helpful comments by Abhay Ashtekar, Jorge Pullin, Douglas Arnold, L Samuel Finn and Mijan Huq, and the hospitality of the CGPG group. He thanks Simonetta Frittelli for pointing out the reference [42]. We thank Bernard Kelly for a careful reading of the manuscript. Numerical computations were performed using machines at CGPG. This work was supported in part by the NSF grants PHYS95-14240, and the Everly research funds of Penn State. HS was supported by the Japan Society for the Promotion of Science as a research fellow abroad.

Appendix A. Ashtekar's formulation of general relativity

We give a brief review of the Ashtekar formulation and the way to handle reality conditions. This appendix describes our notation.

A.1. Variables and equations

The key feature of Ashtekar's formulation of general relativity [17] is the introduction of a self-dual connection as one of the basic dynamical variables. Let us write the metric $g_{\mu\nu}$ using the tetrad E_μ^I as $g_{\mu\nu} = E_\mu^I E_\nu^J \eta_{IJ}$ †. Define its inverse, E_I^μ , by $E_I^\mu := E_\nu^J g^{\mu\nu} \eta_{IJ}$ and we impose $E_a^0 = 0$ as the gauge condition. We define $SO(3, C)$ as self-dual and anti-self-dual connections ${}^\pm \mathcal{A}_\mu^a := \omega_\mu^{0a} \mp \frac{1}{2} i \epsilon^{abc} \omega_\mu^{bc}$, where ω_μ^{IJ} is a spin connection 1-form (Ricci connection), $\omega_\mu^{IJ} := E^{I\nu} \nabla_\nu E_\mu^J$. Ashtekar's plan is to use only the self-dual part of the connection ${}^+ \mathcal{A}_\mu^a$ and to use its spatial part ${}^+ \mathcal{A}_i^a$ as a dynamical variable. Hereafter, we simply denote ${}^+ \mathcal{A}_\mu^a$ as \mathcal{A}_μ^a .

The lapse function, N , and shift vector, N^i , both of which we treat as real-valued functions, are expressed as $E_0^\mu = (1/N, -N^i/N)$. This allows us to think of E_0^μ as a normal vector

† We use $\mu, \nu = 0, \dots, 3$ and $i, j = 1, \dots, 3$ as spacetime indices, while $I, J = (0), \dots, (3)$ and $a, b = (1), \dots, (3)$ are $SO(1, 3)$, $SO(3)$ indices, respectively. We raise and lower μ, ν, \dots by $g^{\mu\nu}$ and $g_{\mu\nu}$ (the Lorentzian metric); I, J, \dots by $\eta^{IJ} = \text{diag}(-1, 1, 1, 1)$ and η_{IJ} ; i, j, \dots by γ^{ij} and γ_{ij} (the 3-metric); a, b, \dots by δ^{ab} and δ_{ab} . We also use volume forms ϵ_{abc} : $\epsilon_{abc} \epsilon^{abc} = 3!$.

field to Σ spanned by the condition $t = x^0 = \text{constant}$, which plays the same role as that of the Arnowitt–Deser–Misner formulation. Ashtekar treated the set $(\tilde{E}_a^i, \mathcal{A}_i^a)$ as basic dynamical variables, where \tilde{E}_a^i is an inverse of the densitized triad defined by $\tilde{E}_a^i := e E_a^i$, where $e := \det E_a^i$ is a density. This pair forms the canonical set.

In the case of pure gravitational spacetime, the Hilbert action takes the form

$$S = \int d^4x \left[(\partial_t \mathcal{A}_i^a) \tilde{E}_a^i + \frac{1}{2} i \tilde{N} \tilde{E}_a^i \tilde{E}_b^j F_{ij}^c \epsilon^{abc} - e^2 \Lambda \tilde{N} - N^i F_{ij}^a \tilde{E}_a^j + \mathcal{A}_0^a \mathcal{D}_i \tilde{E}_a^i \right], \quad (\text{A1})$$

where $\tilde{N} := e^{-1} N$, $F_{\mu\nu}^a := 2\partial_{[\mu} \mathcal{A}_{\nu]}^a - i\epsilon^{abc} \mathcal{A}_\mu^b \mathcal{A}_\nu^c$ is the curvature 2-form, Λ is the cosmological constant, $\mathcal{D}_i \tilde{E}_a^j := \partial_i \tilde{E}_a^j - i\epsilon_{ab}^c \mathcal{A}_i^b \tilde{E}_c^j$, and $e^2 = \det \tilde{E}_a^i = (\det E_a^i)^2$ is defined to be $\det \tilde{E}_a^i = \frac{1}{6} \epsilon^{abc} \epsilon_{ijk} \tilde{E}_a^i \tilde{E}_b^j \tilde{E}_c^k$, where $\epsilon_{ijk} := \epsilon_{abc} E_i^a E_j^b E_k^c$ and $\epsilon_{ijk} := e^{-1} \epsilon_{ijk}^\dagger$.

Varying the action with respect to the non-dynamical variables \tilde{N} , N^i and \mathcal{A}_0^a yields the constraint equations,

$$\mathcal{C}_H^{\text{ASH}} := \frac{1}{2} i \epsilon^{abc} \tilde{E}_a^i \tilde{E}_b^j F_{ij}^c - \Lambda \det \tilde{E} \approx 0, \quad (\text{A2})$$

$$\mathcal{C}_{Mi}^{\text{ASH}} := -F_{ij}^a \tilde{E}_a^j \approx 0, \quad (\text{A3})$$

$$\mathcal{C}_{Ga}^{\text{ASH}} := \mathcal{D}_i \tilde{E}_a^i \approx 0. \quad (\text{A4})$$

The equations of motion for the dynamical variables $(\tilde{E}_a^i$ and $\mathcal{A}_i^a)$ are

$$\partial_t \tilde{E}_a^i = -i \mathcal{D}_j (\epsilon^{cb} \tilde{N} \tilde{E}_c^j \tilde{E}_b^i) + 2 \mathcal{D}_j (N^{[j} \tilde{E}_a^{i]}) + i \mathcal{A}_0^b \epsilon_{ab}^c \tilde{E}_c^i, \quad (\text{A5})$$

$$\partial_t \mathcal{A}_i^a = -i \epsilon^{ab} \tilde{N} \tilde{E}_b^j F_{ji}^c + N^j F_{ji}^a + \mathcal{D}_i \mathcal{A}_0^a + \Lambda \tilde{N} \tilde{E}_i^a, \quad (\text{A6})$$

where $\mathcal{D}_j X_a^{ji} := \partial_j X_a^{ji} - i \epsilon_{ab}^c \mathcal{A}_j^b X_c^{ji}$, for $X_a^{ij} + X_a^{ji} = 0$.

A.2. Reality conditions

In order to construct the metric from the variables $(\tilde{E}_a^i, \mathcal{A}_i^a, \tilde{N}, N^i)$, we first prepare the tetrad E_I^μ as $E_0^\mu = (1/e\tilde{N}, -N^i/e\tilde{N})$ and $E_a^\mu = (0, \tilde{E}_a^i/e)$. Using them, we obtain the metric $g^{\mu\nu}$ such that $g^{\mu\nu} := E_I^\mu E_J^\nu \eta^{IJ}$.

This metric, in general, is not real-valued in the Ashtekar formulation. To ensure that the metric is real-valued, we need to impose real lapse and shift vectors together with two *metric reality* conditions;

$$\text{Im}(\tilde{E}_a^i \tilde{E}^{ja}) = 0, \quad (\text{A7})$$

$$W^{ij} := \text{Re}(\epsilon^{abc} \tilde{E}_a^k \tilde{E}_b^{(i} \mathcal{D}_k \tilde{E}^{j)}) = 0, \quad (\text{A8})$$

where the latter comes from the secondary reality condition of the metric $\text{Im}\{\partial_t(\tilde{E}_a^i \tilde{E}^{ja})\} = 0$ [26], and we assume $\det \tilde{E} > 0$ (see [27]).

For later convenience, we also prepare stronger reality conditions, *triad reality* conditions. The primary and secondary conditions are written, respectively, as

$$U_a^i := \text{Im}(\tilde{E}_a^i) = 0, \quad (\text{A9})$$

$$\text{Im}(\partial_t \tilde{E}_a^i) = 0. \quad (\text{A10})$$

Using the equations of motion of \tilde{E}_a^i , the gauge constraint (A4), the metric reality conditions (A7), (A8) and the primary condition (A9), we see that (A10) is equivalent to [27]

$$\text{Re}(\mathcal{A}_0^a) = \partial_i (N) \tilde{E}^{ia} + (1/2e) E_i^b \tilde{N} \tilde{E}^{ja} \partial_j \tilde{E}_b^i + N^i \text{Re}(\mathcal{A}_i^a), \quad (\text{A11})$$

† When $(i, j, k) = (1, 2, 3)$, we have $\epsilon_{ijk} = e$, $\epsilon_{ijk} = 1$, $\epsilon^{ijk} = e^{-1}$ and $\tilde{\epsilon}^{ijk} = 1$.

or with undensitized variables,

$$\text{Re}(\mathcal{A}_0^a) = \partial_i(N)E^{ia} + N^i \text{Re}(\mathcal{A}_i^a). \quad (\text{A12})$$

From this expression we see that the secondary triad reality condition restricts the three components of the ‘triad lapse’ vector \mathcal{A}_0^a . Therefore, equation (A11) is not a restriction on the dynamical variables (\tilde{E}_a^i and \mathcal{A}_i^a) but on the slicing, which we should impose on each hypersurface.

Throughout the discussion in this paper, we assume that the initial data of ($\tilde{E}_a^i, \mathcal{A}_i^a$) for evolution are solved so as to satisfy all three constraint equations and the metric reality condition (A7) and (A8). Practically, this is obtained, for example, by solving ADM constraints and by transforming a set of initial data to Ashtekar’s notation.

Appendix B. Experiments using the Maxwell equation

In this appendix, using the Maxwell equation of the vacuum field, we show that the symmetric hyperbolic system does not change the stability feature drastically. The result here supports the discussion in section 5. More detail analysis can be found in our paper II [41].

The Maxwell equation has two constraint equations,

$$C_E := \partial_i E^i \approx 0, \quad C_B := \partial_i B^i \approx 0, \quad (\text{B1})$$

and two dynamical equations

$$\partial_t E_i = c\epsilon_i^{jk} \partial_j B_k, \quad \partial_t B_i = -c\epsilon_i^{jk} \partial_j E_k \quad (\text{B2})$$

for the field (E_i, B_i).

Suppose we have adjusted (B2) using the constraint terms, (B1), with a multiplier, κ ,

$$\partial_t E_i = c\epsilon_i^{jk} \partial_j B_k + \kappa_i C_E, \quad \partial_t B_i = -c\epsilon_i^{jk} \partial_j E_k + \kappa_i C_B \quad (\text{B3})$$

where $\kappa_i = (\kappa, \kappa, \kappa)$ for simplicity. This matrix expression

$$\partial_t \begin{pmatrix} E_i \\ B_i \end{pmatrix} \approx \begin{pmatrix} \delta^{jl} \kappa_i & -c\epsilon_i^{jl} \\ c\epsilon_i^{jl} & \delta^{jl} \kappa_i \end{pmatrix} \partial_l \begin{pmatrix} E_j \\ B_j \end{pmatrix} \quad (\text{B4})$$

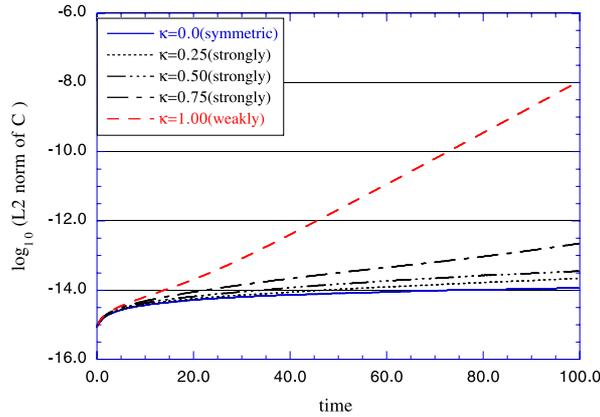


Figure B1. Comparisons of the ‘adjusted’ system with the different multiplier, κ , in equation (B3). Plots are of the L2 norm of the constraint equations, (B1).

immediately tells us its hyperbolicity depending on κ as follows. The system, (B4), becomes of symmetric hyperbolic form when $\kappa = 0$ (that is the original Maxwell equation), becomes of weakly hyperbolic form when $\kappa = \pm c$, and becomes strongly hyperbolic otherwise. The eigenvalues of the dynamical equation can be written as $(c, c, -c, -c, \kappa, \kappa)$.

We made a numerical code to demonstrate a propagation of plane electromagnetic wave,

$$E^i(x, t) = \left(0, 0, -\frac{1}{\sqrt{2}} \sin\left(\frac{x+y}{\sqrt{2}} - ct\right) \right), \quad (\text{B5})$$

$$B^i(x, t) = \left(-\frac{1}{2} \sin\left(\frac{x+y}{\sqrt{2}} - ct\right), \frac{1}{2} \sin\left(\frac{x+y}{\sqrt{2}} - ct\right), 0 \right) \quad (\text{B6})$$

in two-dimensional spacetime with periodic boundary condition. We use (B6) as our initial data, and monitor its numerical error during its evolution by evaluating constraint equations and by checking the error from the exact solution. The error itself is quite small, but as we show in figure B1 we found the difference due to the multiplier of the adjusted terms κ . We see that the symmetric hyperbolic equation shows the best performance for the stability, but does not show very different performance from the strongly hyperbolic system.

Appendix C. Why do adjusted equations have better performance?

Here, we try to explain briefly why the adjusted equations ((31) and (32) for Ashtekar's system, (B3) for Maxwell equations) reduce the violation of the constraints in the evolution. The detailed explanations and numerical experiments are given in our paper II [41], and this appendix describes the essential idea of the mechanism.

Suppose we have constraint equations, $\mathcal{C}_1 \approx 0, \mathcal{C}_2 \approx 0, \dots$, in a system. Normally, we monitor the error of the evolution by evaluating these constraint equations on the each constant-time hypersurface. Such monitoring, on the other hand, can also be performed by checking the evolution equations of the constraint, which we denote as constraint propagation equations (cf [8]). We, therefore, consider constraint propagation equation transformed in Fourier components, $\hat{\mathcal{C}}$,

$$\partial_t \begin{pmatrix} \hat{\mathcal{C}}_1 \\ \hat{\mathcal{C}}_2 \\ \vdots \end{pmatrix} = M \begin{pmatrix} \hat{\mathcal{C}}_1 \\ \hat{\mathcal{C}}_2 \\ \vdots \end{pmatrix}. \quad (\text{C1})$$

The idea here is to estimate the eigenvalues of the matrix, M , after we took its leading-order quantity in linearization against a particular background. Clearly, if all of the real part of the eigenvalues are negative, then all constraints decays to zero along the system's evolution. In our paper II [41], we show that such a case can be obtained by adding 'adjusted terms' both for Ashtekar's and Maxwell's systems. There we also show examples of unstable evolution by choosing adjusted terms which produce positive eigenvalues of M . The imaginary part of the eigenvalues are also supposed to contribute to the appearance of the phase differences of the system.

At this point, we can say that adjusted terms are responsible for obtaining the stable and/or accurate evolution system, and this is a way to control the stability of the simulation, which affects more than the system's hyperbolicity.

References

- [1] Choptuik M W 1993 *Phys. Rev. Lett.* **70** 9
- [2] Nakamura T and Kodama H (eds) 1999 *Proc. Black Holes and Gravitational Waves (Prog. Theor. Phys. Suppl.* **136**)
- [3] Nakamura T, Oohara K and Kojima Y 1987 *Prog. Theor. Phys. Suppl.* **90** 1
Shibata M and Nakamura T 1995 *Phys. Rev. D* **52** 5428
- [4] Baumgarte T W and Shapiro S L 1999 *Phys. Rev. D* **59** 024007
Baumgarte T W, Hughes S A and Shapiro S L 1999 *Phys. Rev. D* **60** 087501
- [5] Alcubierre M *et al* 2000 *Preprint* gr-qc/0003071
- [6] Alcubierre M *et al* 1999 *Preprint* gr-qc/9908079
- [7] Reula O A 1998 *Living Rev. Relativ.* **1998-3** <http://www.livingreviews.org/>
Friedrich H and Rendall A 2000 *Einstein's Field Equations and Their Physical Interpretation* ed B G Schmidt (Berlin: Springer)
(Friedrich H and Rendall A 2000 *Preprint* gr-qc/0002074)
- [8] Frittelli S 1997 *Phys. Rev. D* **55** 5992
- [9] Bona C, Massó J, Seidel E and Stela J 1995 *Phys. Rev. Lett.* **75** 600
Bona C, Massó J, Seidel E and Stela J 1997 *Phys. Rev. D* **56** 3405
- [10] Bona C and Massó J 1992 *Phys. Rev. Lett.* **68** 1097
- [11] Scheel M A, Baumgarte T W, Cook G B, Shapiro S L and Teukolsky S A 1997 *D* **56** 6320
Scheel M A, Baumgarte T W, Cook G B, Shapiro S L and Teukolsky S A 1998 *Phys. Rev. D* **58** 044020
- [12] Choquet-Bruhat Y and York J W Jr 1995 *C. R. Acad. Sci. Paris I* **321** Série I, 1089
(Choquet-Bruhat Y and York J W Jr 1995 *Preprint* gr-qc/9506071)
Abrahams A, Anderson A, Choquet-Bruhat Y and York J W Jr 1995 *Phys. Rev. Lett.* **75** 3377
Abrahams A, Anderson A, Choquet-Bruhat Y and York J W Jr 1996 *C. R. Acad. Sci. Paris II* **323** 835
Abrahams A, Anderson A, Choquet-Bruhat Y and York J W Jr 1997 *Class. Quantum Grav.* **14** A9
Anderson A and York J W Jr 1999 *Phys. Rev. Lett.* **82** 4384
- [13] Alcubierre M 1997 *Phys. Rev. D* **55** 5981
Alcubierre M and Massó J 1998 *Phys. Rev. D* **57** 4511
- [14] Friedrich H 1981 *Proc. R. Soc. A* **375** 169
Friedrich H 1981 *Proc. R. Soc. A* **378** 401
Friedrich H 1985 *Commun. Math. Phys.* **100** 525
Friedrich H 1991 *J. Differ. Geom.* **34** 275
Friedrich H 1996 *Class. Quantum Grav.* **13** 1451
- [15] Frauendiener J 1998 *Phys. Rev. D* **58** 064002
Frauendiener J 1998 *Phys. Rev. D* **58** 064003
Frauendiener J 2000 *Class. Quantum Grav.* **17** 373
Hübner P 1999 *Class. Quantum Grav.* **16** 2145
Hübner P 1999 *Class. Quantum Grav.* **16** 2823
- [16] Stewart J M 1998 *Class. Quantum Grav.* **15** 2865
- [17] Ashtekar A 1986 *Phys. Rev. Lett.* **57** 2244
Ashtekar A 1987 *Phys. Rev. D* **36** 1587
- [18] Ashtekar A 1991 *Lectures on Non-Perturbative Canonical Gravity* (Singapore: World Scientific)
- [19] Yoneda G and Shinkai H 2000 *Int. J. Mod. Phys. D* **9** 13
- [20] Yoneda G and Shinkai H 1999 *Phys. Rev. Lett.* **82** 263
- [21] Shinkai H and Yoneda G 1999 *Phys. Rev. D* **60** 101502
- [22] Brodbeck O, Frittelli S, Hübner P and Reula O A 1999 *J. Math. Phys.* **40** 909
- [23] Ashtekar A and Romano J D 1989 Key (3+1)-equations in terms of new variables for numerical relativity *Report* Syracuse University
- [24] Salisbury D C, Shepley L C, Adams A, Mann D, Turvan L and Turner B 1994 *Class. Quantum Grav.* **11** 2789
- [25] Capovilla R, Jacobson T and Dell J 1989 *Phys. Rev. Lett.* **63** 2325
Capovilla R, Jacobson T and Dell J 1991 *Class. Quantum Grav.* **8** 59
- [26] Ashtekar A, Romano J D and Tate R S 1989 *Phys. Rev. D* **40** 2572
- [27] Yoneda G and Shinkai H 1996 *Class. Quantum Grav.* **13** 783
- [28] Hirsch C 1988 *Numerical Computation of Internal and External Flows* vols I, II (New York: Wiley)
- [29] Bona C, Massó J, Seidel E and Walker P 1998 *Preprint* gr-qc/9804052
- [30] Geroch R 1996 Partial differential equations in physics *Preprint* gr-qc/9602055
- [31] Fischer A E and Marsden J E 1972 *Commun. Math. Phys.* **28** 1

- [32] Frittelli S and Reula O A 1996 *Phys. Rev. Lett.* **76** 4667
- [33] Bardeen J and Buchman L 2000 in preparation
- [34] Iriondo M S, Leguizamón E O and Reula O A 1997 *Phys. Rev. Lett.* **79** 4732
Iriondo M S, Leguizamón E O and Reula O A 1998 *Adv. Theor. Math. Phys.* **2** 1075
- [35] O'Murchadha N and York J W Jr 1974 *Phys. Rev. D* **10** 428
- [36] Szekeres P 1970 *Nature* **228** 1183
Szekeres P 1972 *J. Math. Phys.* **13** 286
Khan V A and Penrose R 1971 *Nature* **229** 185
- [37] Centrella J and Matzner R A 1979 *Astrophys. J.* **230** 311
Centrella J 1980 *Astrophys. J.* **241** 875
Centrella J and Matzner R A 1982 *Phys. Rev. D* **25** 930
- [38] Shinkai H and Maeda K 1993 *Phys. Rev. D* **48** 3910
- [39] Teukolsky S A 2000 *Phys. Rev. D* **61** 087501
- [40] Bernstein D, Hobill D W and Smarr L L 1989 *Frontiers in Numerical Relativity* ed C R Evans, L S Finn and D W Hobill (Cambridge: Cambridge University Press)
- [41] Yoneda G and Shinkai H 2000 *Preprint* gr-qc/0007034 (paper II)
- [42] Hern S D 2000 *PhD Thesis Preprint* gr-qc/0004036