A Hyperbolic Tetrad Approach to Numerical Relativity

Luisa T. Buchman

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

2003

Program Authorized to Offer Degree: Astronomy

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Bell and Howell Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Luisa I. Buchman

Date 12/16/03

University of Washington

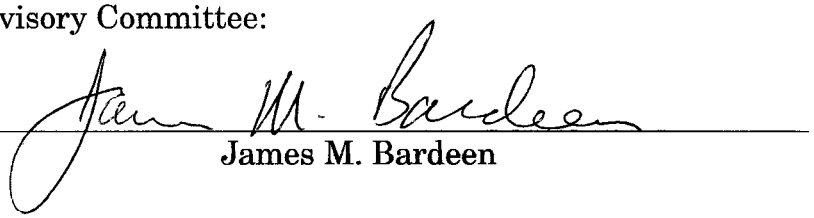
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Luisa T. Buchman

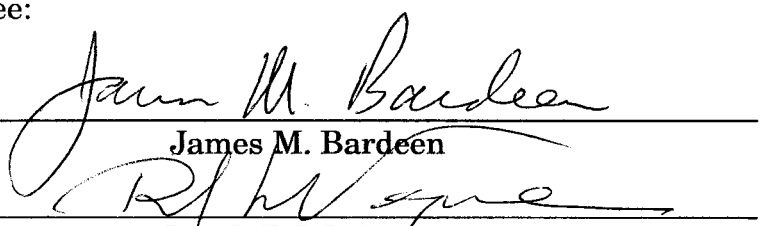
and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of Supervisory Committee:

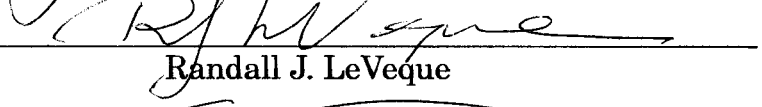


James M. Bardeen

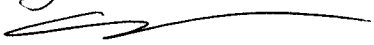
Reading Committee:



James M. Bardeen



Randall J. LeVeque



Craig Hogan

Date: 12/16/03

University of Washington

Abstract

A Hyperbolic Tetrad Approach to Numerical Relativity

Luisa T. Buchman

Chair of Supervisory Committee:
Professor James M. Bardeen
Physics

An in-depth numerical study using $3 + 1$ formulations of the Einstein equations for 1D colliding plane waves reveals factors which increase accuracy and stability: using "mixed" variables, hyperbolicity, satisfying the energy and momentum constraint equations at the boundaries, and changing the speeds of "constraint" eigenmodes so that they are not in step with dominant sources of error. In order to generalize these results, a novel tetrad approach to vacuum numerical relativity has been developed, based on the formalism of Estabrook, Robinson, and Wahlquist [24]. Clear advantages of this approach are that the variables are more naturally related to physical quantities, and the Minkowski metric is used to raise and lower indices on the tetrad variables. A potential disadvantage is that tetrads as well as coordinates and hypersurfaces must be evolved. A lapse function and a shift vector are introduced, which allow for a general choice of coordinates. The evolution equations expressed as coordinate-based partial differential equations are symmetrizable hyperbolic for Nester, Lorentz, and fixed gauge conditions. Implementation of this tetrad formulation for 1D colliding plane waves results in significantly greater accuracy than results using comparable $3 + 1$ approaches. Tests in spherically symmetric

Schwarzschild spacetime indicate that the Nester gauge gives a poor evolution of the congruence for non-stationary initial conditions. With the Lorentz gauge, the solution settles down to a stationary state, but second order convergence holds only to about $40M$. By changing the speeds of the "constraint" eigenmodes so that they propagate quickly through the region of instability around the event horizon, solid stability is achieved to at least $500M$. The black hole and plane wave results suggest that having the "constraint" eigenmodes travel away from dominant sources of error is key to stable and accurate numerical evolutions. Future work will involve understanding how this insight may be implemented in 3D.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my deep gratitude to all the members of my committee, and especially to my brilliant and patient advisor, J. M. Bardeen. I am also want to thank the Astronomy Department, for their extended support, and the NASA Graduate Students Researchers Program, which gave me the opportunity to work full-time on this research, and to collaborate with F. B. Estabrook and H. D. Wahlquist at the Jet Propulsion Laboratory.

Finally, this thesis would never have been completed without the loving devotion of my husband, the support of my family and all my sterling friends, and the prayers of the Sisters at St. Joseph's Carmelite monastery.

DEDICATION

–to Johnny–

In Wisdom is a Spirit
intelligent, holy, unique,
Manifold, subtle, agile, clear, unstained, certain,
Not baneful, loving the good, keen, unhampered, beneficent, kindly,
Firm, secure, tranquil, all-powerful, all-seeing,
And pervading all spirits, though they be intellingent, pure, and very subtle.
For Wisdom is mobile beyond all motion, and she penetrates and pervades
all things by reason of her purity.
For she is an aura of the might of God
and a pure effusion of the glory of the Almightly;
therefore nought that is sullied enters into her.
For she is the refulgence of eternal light,
the spotless mirror of the power of God, the image of His goodness.
And she, who is one, can do all things,
and renews everything while herself perduring;
And passing into holy souls from age to age,
she produces friends of God and prophets.
For there is nought God loves, be it not one who dwells with Wisdom.
For she is fairer than the sun
and surpasses every constellation of the stars.

Book of Wisdom 7 : 22b

Chapter 1

INTRODUCTION

Gravitational-wave detectors will soon bring us obervational maps of black holes colliding [...] Supercomputer simulations will attempt to replicate the symphonies and tell us what they mean, and black holes thereby will become the objects of detailed scrutiny. What will that scrutiny teach us? There will be surprises.–Kip Thorne[73]

The gravitational wave detector, LIGO (which stands for Laser Interferometer Gravitational Wave Observatory) has been online for about a year. It is planned to be replaced by Advanced LIGO in 2007. Once observations begin, probably around 2009, Advanced LIGO is expected to detect gravitational wave signals from coalescencing solar-mass binary black holes possibly as often as every day. LISA (Laser Interferometer Space Antenna) is a spaced-based gravitational wave detector, scheduled for launch around 2010. LISA is designed to detect low-frequency gravitational waves. A primary source for low-frequency gravity waves is the coalescence of supermassive black holes, a phenomenon which occurs during galaxy mergers. Coalescences of supermassive black holes, which range in mass from 30,000 to 10 million solar masses, are expected to be detectable by LISA with high signal to noise, on the order of thousands to one [29].

There are three phases in the coalescence process of two black holes: the inspiral, merger, and ringdown phases [30]. Each of these phases produces a

characteristic gravity wave signature, or waveform. The merger phase includes the transition from the inspiral to a freely falling plunge of the two black holes towards each other, their subsequent violent collision, and the beginning of the relaxation of the merged black hole to its final stationary state. This stage of the coalescence process is highly relativistic, dynamic, and nonlinear, and emits the strongest signal of the whole binary black hole coalescence process. Having theoretical waveforms, or templates, for the merger signals will contribute to both detecting and interpreting gravitational waves. These templates can only be calculated with numerical relativity, because the merger phase is so dynamic and nonlinear. Although considerable efforts have been devoted to developing three- dimensional (3D) codes to perform these calculations, the necessary long-term evolutions have not yet been achieved. (For recent reviews, see [67, 75, 15, 42, 41]).

The success of LIGO and LISA depends strongly on scientists ability to predict and interpret gravitational waveforms from the merger phase of the binary black hole collisions. If researchers can produce templates corresponding to different binary black hole merger scenarios, then LIGO and LISA scientists will be able to decode signals detected from the extreme strong-gravity spacetime of two coalescing black holes. Heretofore unknown phenomena could be uncovered. A match between templates and observed waveforms would confirm for the first time Einsteins theory of General Relativity in the strong field regime.

In addition, calculating merger templates will aid in the actual detection of signals, because accurate theoretical templates are used to design filters which greatly enhance the detection of weak signals. Further, since the signals emitted from the black hole merger phase are the strongest of the supermassive black hole coalescence process, it is predicted that LIGO and LISA will detect some binary black hole coalescence events primarily from their merger signal. Merger templates increase both the effective range and usable sensitivity of

the interferometers.

The difficulty in numerically calculating black hole merger templates arises because the code must not only be accurate, but long-term stable. Realistic initial data for the plunge must be obtained from simulating the slow inspiral of two unequal mass rotating black holes in 3D. This initial data then has to be evolved through the plunge of the two black holes toward each other, their collision, their merging into a single, highly distorted black hole, to the beginning of the relaxation of this distorted black hole towards its final stationary state. Only then can the desired gravitational waveforms be extracted.

In Chapter 2 of the dissertation, a systematic and careful numerical study of one-dimensional colliding gravitational plane waves is performed, in order to understand key issues involved in obtaining accurate and stable numerical simulations. This study indicates that three important factors which result in a significant increase in accuracy and stability are (1) using a hyperbolic formulation of the Einstein equations, (2) adjusting the speeds of the constraint eigenmodes of the hyperbolic system so they are different from the speeds of the constraint quantities which propagate in the full nonlinear solution, and (3) using a mixed form of the variables, which have a simple linear relation to the physical waves. It is difficult to generalize the third result to arbitrary 3D black hole spacetimes using the formulations traditionally employed in numerical relativity. This is because for general 3D black holes, there is no symmetry to the propagation of the gravitational waves. However, there is a natural mathematical construct for measuring observable physical quantities, namely, tetrad frames [see [26]]. In a tetrad formulation, the variables are resolved along orthonormal space and time axes of local Lorentz frames. The general concept is that these variables represent quantities that cameras at rest in the tetrad frames could observe. In light of the fact that the traditional formulations used have not yielded the necessary accuracy and sta-

bility after at least a decade of concentrated effort by the numerical relativity community, the viability for numerical relativity of a hyperbolic tetrad formalism published by Estabrook, Robinson, and Wahlquist [24] is investigated. In Chapter 3, the theoretical foundations for implementing this formalism in 3D black hole spacetime is developed. A completely arbitrary relationship between the tetrad orientations and the spacelike hypersurfaces on which the numerical calculations are performed is allowed. This is important especially when dealing with black hole event horizons. We have determined that when the tetrad equations are expressed in terms of partial derivatives, the system is symmetrizable hyperbolic. This is an important condition for stability, as established by Lindblom and Scheel [50]. Application of the tetrad formulation for 1D colliding gravitational plane waves is presented in Chapter 4, and for 1D Schwarzschild black holes in Chapter 5.

There has not been much published using the tetrad approach for Cauchy numerical evolutions of vacuum spacetimes. Most of the recent tetrad formalisms assume that the tetrads are tied to the physically defined flow of a fluid, primarily in the context of cosmology or of interior metrics of rotating stars [37, 74, 22]. To the best of our knowledge, the only published paper that uses a strictly tetrad approach for vacuum numerical relativity is that by van Putten [76], in 1D Gowdy wave spacetime. The tetrad formulation proposed in this dissertation has fewer variables than van Putten's, which may be advantageous in 3D codes.

Developing a code with greater accuracy and/ or longer-term stability than those already in use would bring numerical relativists closer to calculating templates for coalescing black hole mergers. This in turn would lead LIGO and LISA scientists to a deeper and more thorough exploration of the extreme gravity regions of the Universe.

Chapter 2

NUMERICAL TESTS OF EVOLUTION SYSTEMS, GAUGE CONDITIONS, AND BOUNDARY CONDITIONS FOR 1D COLLIDING GRAVITATIONAL PLANE WAVES

2.1 Abstract

We investigate how the accuracy and stability of numerical relativity simulations of 1D colliding plane waves depends on choices of equation formulations, gauge conditions, boundary conditions, and numerical methods, all in the context of a first-order 3+1 approach to the Einstein equations, with the basic variables being some combination of first derivatives of the spatial metric and components of the extrinsic curvature tensor. Hyperbolic schemes, specifically variations on schemes proposed by Bona and Massó and Anderson and York, are compared with variations of the Arnowitt-Deser-Misner formulation. Modifications of the three basic schemes include raising one index in the metric derivative and extrinsic curvature variables and adding a multiple of the energy constraint to the extrinsic curvature evolution equations. Redundant variables in the Bona-Massó formulation may be reset frequently or allowed to evolve freely. Gauge conditions which simplify the dynamical structure of the system are imposed during each time step, but the lapse and shift are reset periodically to control the evolution of the spacetime slicing and the longitudinal part of the metric. We show that physically correct boundary conditions, satisfying the energy and momentum constraint equations, generically require the presence of some ingoing eigenmodes of the characteristic matrix. Numer-

ical methods are developed for the hyperbolic systems based on decomposing flux differences into linear combinations of eigenvectors of the characteristic matrix. These methods are shown to be second-order accurate, and in practice second-order convergent, for smooth solutions, even when the eigenvectors and eigenvalues of the characteristic matrix are spatially varying. This chapter has been published (see reference [13]).

2.2 Introduction

The goal of projects such as the ground-based Laser Interferometric Gravitational Wave Observatory (LIGO) and the space-based Laser Interferometer Space Antenna (LISA) is to detect gravitational waves, and to use them as a new observational window for relativistic astrophysics. A primary source for these gravitational waves is the coalescence of binary black holes [30]. The highly nonlinear and dynamical merger phase of this coalescence process can only be calculated by numerical relativity, and obtaining merger gravitational waveforms, both for theoretical understanding and for detection, is dependent on long-term stable and accurate numerical evolutions. A worldwide collaboration of numerical relativists, physicists, mathematicians, and computer scientists has devoted considerable effort over the last 20 years to develop 3D codes to calculate black hole merger gravitational waveforms, and significant progress has been made, especially in the last few years. However, more groundwork is required before calculations of 3D binary black hole merger templates for a variety of scenarios can be completed. Greater understanding of equation formulations, boundary conditions, and dynamic gauge conditions, and the use of advanced numerical methods, is essential to achieve this goal. We believe that an important foundation for this understanding is extensive testing and analysis in 1D and 2D. Choptuik's discovery of black hole

critical phenomena in spherically symmetric gravitational collapse [19] is an example of the potential of careful numerical work in 1D.

This paper reports the methodology, results, and analysis of calculations of 1D nonlinear colliding gravitational planewave spacetimes. We have chosen to investigate hyperbolic formulations of the Einstein equations, as they are well-posed, they can be treated with advanced numerical methods, and they can help in the analysis of boundary conditions [32, 17]. We call a set of equations *hyperbolic* if the characteristic matrix can be diagonalized with a complete set of eigenvectors and real eigenvalues, following LeVeque [46]. This is called *strongly hyperbolic* [60] in much of the literature. The lapse and the shift are evolved during each time step in a manner which is consistent with a simple hyperbolic scheme. Between time steps, the lapse and the shift are reset according to conditions which are unconstrained by the need to preserve hyperbolicity. In this way, the evolution of the hypersurfaces and spatial coordinates can be controlled to prevent large gradients, coordinate pathologies, and instabilities. Some of the redundant variables of a hyperbolic formulation can also be reset between time steps. This resetting can have positive or negative effects on accuracy and stability, depending on the eigenmode structure of the reset system. Finally, we find that boundary conditions should not be based naively on the eigenmodes of the hyperbolic decomposition for two reasons: (a) satisfying the constraint equations at the boundaries generically requires the presence of incoming eigenmodes, and (b) even whether the "physical" eigenmodes are purely outgoing at the boundaries is gauge dependent.

Many ways of formulating evolution equations for the spatial metric in Einstein's theory of General Relativity are possible. The most thoroughly tested formulation in numerical relativity is the Arnowitt-Deser-Misner (ADM) set of

equations [8]. The standard ADM equations in vacuum are

$$(\partial_t - \mathcal{L}_\beta)h_{ij} = -2\alpha K_{ij}, \tag{2.1}$$

$$(\partial_t - \mathcal{L}_\beta)K_{ij} = -\alpha_{|ij} + \alpha[{}^{(3)}R_{ij} + KK_{ij} - 2K_i{}^l K_{lj}], \tag{2.2}$$

$$(\partial_t - \mathcal{L}_\beta)K^i{}_j = -\alpha^{|i}{}_{|j} + \alpha[{}^{(3)}R^i{}_j + KK^i{}_j]. \tag{2.3}$$

In these equations, K_{ij} is the extrinsic curvature, $K = K^l{}_l$, β^i is the shift, α is the lapse, h_{ij} is the 3-metric, and ${}^{(3)}R_{ij}$ is the 3D Ricci tensor. The vertical bar represents a covariant derivative taken with respect to the 3-geometry. Eq. (2.3) evolves what we call the "mixed" form of the extrinsic curvature tensor. The energy and momentum constraint equations are, respectively,

$$\mathcal{E} = 1/2[{}^{(3)}R - K_i{}^j K^i{}_j + K^2] = 0, \tag{2.4}$$

$$\mathcal{M}_i = K_i{}^j{}_{|j} - K_{|i} = 0. \tag{2.5}$$

While successful calculations using the ADM formulation have been done in 2D, 3D calculations generally crash after just a few dynamical times.

Alternative formalisms include many versions of hyperbolic systems, which add redundant variables and/or constraint terms to the equations to allow a complete set of eigenmodes describing evolution along the characteristics. As indicated by Reula [60], there are an infinite number of hyperbolic formulations. We focus on variations of relatively simple schemes proposed by Bona-Massó (BM) [16] and Anderson-York (AY) [5], in which the characteristics propagate either at local light speed or along the hypersurface normals, and in which the variables include first derivatives of the metric.

Initial attempts at using hyperbolic methods in 3D were based on the BM formulation [17], but did not use numerical methods which take advantage of the eigenfields of the system. These codes were not much more successful than ADM. Non-hyperbolic Baumgarte-Shapiro-Shibata-Nakamura (BSSN) schemes

[14, 66], based on conformal decomposition of the metric, have shown considerable success in improving the stability of 3D calculations for weak and strong gravitational fields and a variety of spacetime slicings [28]. Alcubierre *et al.* [3] report that a BSSN scheme, combined with excision and certain dynamic gauge conditions, allows accurate numerical evolutions of 3D distorted dynamic black holes up to hundreds of dynamical times.

In the context of considering only first derivative variables, a great variety of hyperbolic schemes have been proposed that involve adding constraint terms to the equations [34, 69, 68, 83, 84, 39]. Kidder, Scheel, and Teukolsky [39] examine a rather general class of such schemes, which include the AY [5] and Frittelli-Reula [34] formulations as special cases. Among these schemes are some which allow for long-term evolution of a Schwarschild black hole in 3D.

In this paper, we explore ways of using hyperbolic methods that combine superior accuracy with gauge conditions which maintain stability at least for the limited dynamical times we can explore with plane waves. Three basic first order systems are studied: BM, AY, and ADM. Hyperbolicity is obtained in our BM and AY formulations by adding momentum constraint terms to the ADM equations, as in the standard formulations. The BM, AY, and ADM formulations are modified by using "mixed" forms (with one index raised) of the extrinsic curvature and metric derivatives as variables. The BM formulations are further modified by resetting redundant variables, which gives an overall ADM-like evolution. Further, all our formulations are varied by adding a multiple of the energy constraint to the evolution equations for the extrinsic curvature. Specifically, $-n\alpha\mathcal{E}h_{ij}/2$ is added to Eq. (2.2), and $-n\alpha\mathcal{E}\delta^{i}{}_{j}/2$ is added to Eq. (2.3), where n, the energy constraint coefficient, is an arbitrary real number. The ADM formulation is actually hyperbolic as long as the longitudinal-transverse components of the metric and extrinsic curvature can be assumed to vanish identically, and $n < 0$ or $0 < n < 1$. Comparisons of results from our

various ADM, BM, and AY calculations allow us to identify and analyze aspects of equation formulation which significantly improve accuracy and/or stability. Some of these aspects are mixed variables, a separation of the constraint error speeds from the other characteristic speeds of the system, and maintaining long-term effective hyperbolicity (taking into account resetting of redundant variables, but ignoring deviation from strict hyperbolicity due to resetting the lapse and the shift).

While the same energy constraint terms as specified above are present in the standard BM formulation, numerical implementations have been carried out, as far as we are aware, only for $n = 0$ (the Ricci evolution system) and $n = 1$ (the Einstein evolution system). Adding these energy constraint terms to the AY formulation is a special case of the more general Kidder, Scheel, and Teukolsky [39] schemes. Shinkai and Yoneda [69, 68, 83] analyzed the stability and accuracy properties of first order hyperbolic systems using Ashtekar's connection variables in plane-symmetric spacetimes, and found that the addition of multiples of the constraints to the dynamical equations improved accuracy and stability. These results have been extended to ADM systems of equations [84].

Gauge choices in most previous implementations of hyperbolic formulations have been limited in order to preserve the hyperbolicity of the system. Since no time derivatives of the lapse and the shift occur in the dynamical equations for the other variables, the lapse and the shift can be reset arbitrarily at any time during the numerical evolution, as pointed out by Balakrishna *et al.* [27]. Our gauge evolution maintains strict hyperbolicity during each time step, but the lapse and shift are reset periodically between time steps in order to control the long-term evolution of the coordinate system. The resetting may be accomplished by imposing algebraic conditions, by solving elliptic equations, or by evolving the lapse and/or shift through dynamical equations implemented

independently of the main hyperbolic system.

Poor boundary conditions can result in the introduction of instabilities or inaccuracies into the numerical grid. In numerical relativity, boundary conditions have usually been rather crudely implemented. Some sort of outgoing radiation conditions are imposed on all components of the metric, or boundary conditions are based on an analytic exterior solution [28, 61]. One attraction of hyperbolic methods has been the possibility of basing boundary conditions on the eigenmodes of the characteristic matrix. However, it is clear from our planewave calculations that, particularly for the "non-physical" eigenmodes involving the non-transverse-traceless parts of the metric, making the amplitudes of the incoming eigenmodes at the boundaries zero can lead to serious violations of the energy and momentum constraints. Furthermore, what constitutes an incoming eigenmode is dependent on the formulation of the equations as well as on gauge conditions. Even imposing purely outgoing boundary conditions on the "physical" eigenmodes of the hyperbolic system is not strictly correct, as nonlinear coupling between the "physical" and "non-physical" eigenmodes in the source terms can generate a gauge-dependent admixture of outgoing and incoming "physical" eigenmodes. Our boundary conditions are based on quadratic extrapolation of the variables from inside the grid to the first ghost cells on either side of the grid. The ghost cell values are then corrected to make sure the constraint equations are satisfied on the boundaries. For 1D plane waves, projection of the Weyl tensor onto a null tetrad gives a gauge-independent measure of the left and right-going components of the gravitational radiation. Our numerical solutions for colliding plane waves show that as the wave packets leave the grid, the incoming components of the Weyl tensor are in fact zero even though there are non-zero incoming "physical" eigenmodes of the characteristic matrix.

Our focus in applying hyperbolic methods to the Einstein equations is on

achieving second order accuracy for smooth solutions, when the eigenvectors and eigenvalues of the system are a function of position. Finite difference methods such as MacCormack, Lax-Wendroff, and staggered leapfrog [59], which are often used in numerical relativity, give good second order accuracy for smooth solutions, but standard wave propagation algorithms for hyperbolic systems as presented by LeVeque [47] are not second order accurate when the eigenvectors and eigenvalues are spatially varying. LeVeque suggested a new wave propagation method [44] for variable coefficient flux problems which we develop and apply to our 1D nonlinear gravitational planewave calculations. We show in Appendix A that the new methods are formally second order accurate even with varying eigenvectors and eigenvalues, and verify second order convergence in our numerical results.

2.3 Evolution Equations

The most general spatial metric for a nonlinear 1D plane wave traveling in the x-direction is

$$ds^2 = h_{xx}dx^2 + h_{yy}dy^2 + h_{zz}dz^2 + 2h_{yz}dydz, \tag{2.6}$$

in which h_{xx}, h_{yy}, h_{zz}, and h_{yz} are functions of x alone. We will restrict our discussion to a diagonal metric in this paper.

The standard ADM evolution equations are first order in time and second order in space. Most hyperbolic formalisms are first order in space and time, and incorporate first derivatives of the spatial metric as additional variables. The derivative variables as defined by BM are

$$D_{kij} = \frac{1}{2}\partial_k h_{ij}. \tag{2.7}$$

In our applications with a diagonal metric, we find that switching to the "mixed" variables

$$D_{ki}{}^{j} = \frac{1}{2}(\partial_k h_{il})h^{lj}, \quad K_i{}^{j} = K_{il}h^{lj} \tag{2.8}$$

from the "lowered" variables, D_{kij} and K_{ij}, improves accuracy significantly without other complications. However, with a non-diagonal metric, $D_{ki}{}^{j}$ and $K_i{}^{j}$ are not symmetric in i and j, and the evolution equations for $D_{ki}{}^{j}$ acquire complicated source terms.

Below, we present the first order evolution equations for 1D plane waves travelling along the x-direction and described by a diagonal spatial metric, using our mixed variables. The equations with D_{kij} and K_{ij} as variables are given in the BM papers [16]. A few points need to be made about notation. First, since our 1D problem involves derivatives only in the x-direction, we simplify our notation $D_{xi}{}^{j} \longrightarrow D_i{}^{j}$. Second, a prime indicates a spatial derivative with respect to x. Third, our symbol for the shift is simply β instead of β^i. We suppress the index on the shift because there is only one non-zero component in this 1D case.

The evolution equations for h_{ij} are obtained from the definition of the extrinsic curvature of the hypersurfaces, Eq. (2.1), and are

$$\partial_t h_{xx} = 2h_{xx}[\beta D_x{}^{x} + \beta' - \alpha K_x{}^{x}], \tag{2.9}$$

$$\partial_t h_{yy} = 2h_{yy}[\beta D_y{}^{y} - \alpha K_y{}^{y}], \quad \partial_t h_{zz} = 2h_{zz}[\beta D_z{}^{z} - \alpha K_z{}^{z}]. \tag{2.10}$$

The evolution equations for $D_i{}^{j}$ are obtained by taking the time derivative of $D_{ki}{}^{j}$ in Eq. (2.8), and interchanging space and time derivatives. The resulting equations are

$$\partial_t D_x{}^{x} + \partial_x[-\beta D_x{}^{x} - \beta' + \alpha K_x{}^{x}] = 0, \tag{2.11}$$

$$\partial_t D_y{}^{y} + \partial_x[-\beta D_y{}^{y} + \alpha K_y{}^{y}] = 0, \quad \partial_t D_z{}^{z} + \partial_x[-\beta D_z{}^{z} + \alpha K_z{}^{z}] = 0. \tag{2.12}$$

The $K_i{}^{j}$ variables are evolved from the Einstein equations, Eq. (2.3). We include the addition of an arbitrary multiple, n, of the energy constraint in these equations. After organization into a conservation law form, the $K_i{}^{j}$ evolution

equations are

$$\partial_t K_x{}^x + \partial_x \left[-\beta K_x{}^x + \frac{\alpha}{h_{xx}} \left(\frac{\alpha'}{\alpha} + D_y{}^y + D_z{}^z \right) \right] = -\beta' K_x{}^x + \alpha \left[K_x{}^x K_l{}^l + \frac{1}{h_{xx}} \left(\frac{\alpha'}{\alpha} (D_y{}^y + D_z{}^z) - D_y{}^y D_y{}^y - D_z{}^z D_z{}^z - D_x{}^x \left(\frac{\alpha'}{\alpha} + D_y{}^y + D_z{}^z \right) \right) \right] - \frac{n}{2} \alpha \mathcal{E}, \tag{2.13}$$

$$\partial_t K_y{}^y + \partial_x \left[-\beta K_y{}^y + \frac{\alpha}{h_{xx}} D_y{}^y \right] = -\beta' K_y{}^y + \alpha \left[K_y{}^y K_l{}^l - \frac{D_y{}^y D_l{}^l}{h_{xx}} \right] - \frac{n}{2} \alpha \mathcal{E},$$
$$\partial_t K_z{}^z + \partial_x \left[-\beta K_z{}^z + \frac{\alpha}{h_{xx}} D_z{}^z \right] = -\beta' K_z{}^z + \alpha \left[K_z{}^z K_l{}^l - \frac{D_z{}^z D_l{}^l}{h_{xx}} \right] - \frac{n}{2} \alpha \mathcal{E}, \tag{2.14}$$

where we write $\alpha\mathcal{E}$ so that the division between the flux terms and the source terms is apparent:

$$\alpha \mathcal{E} = -\partial_x \left[\frac{\alpha}{h_{xx}} (D_y{}^y + D_z{}^z) \right] + \alpha \left[K_x{}^x (K_y{}^y + K_z{}^z) + K_y{}^y K_z{}^z \right] + \frac{\alpha}{h_{xx}} \left[\left(\frac{\alpha'}{\alpha} - D_x{}^x \right) (D_y{}^y + D_z{}^z) - (D_y{}^y D_y{}^y + D_y{}^y D_z{}^z + D_z{}^z D_z{}^z) \right]. \tag{2.15}$$

2.4 Gauge Evolution

We let the lapse and the shift evolve during each time step according to a prescription which simplifies the hyperbolic system, and we periodically reset the lapse and the shift between time steps to control the longer term evolution of the coordinates and to keep gauge pathologies from developing. We defer discussion of resetting gauge conditions to Sec. 2.6. Here, we discuss how the gauge evolves between resettings.

For a hyperbolic formulation of the equations, the natural choice for the lapse between resettings is the Choquet-Bruhat algebraic gauge condition [21, 20], because it simplifies the fluxes and source terms in the hyperbolic system of equations considerably. This gauge condition is

$$\alpha = Q \sqrt{det(h_{ij})}, \tag{2.16}$$

where Q is a specified function of x, t.

We vary the Choquet-Bruhat algebraic gauge condition by making Q and Q' (which equals $\partial_x Q$) variables in the hyperbolic system, rather than specified functions of x, t. We choose Q and Q' as variables so that Q' can be included in the flux of ${K_x}^x$ as part of the hyperbolic system. Otherwise, Q'' would have to be considered part of the source of ${K_x}^x$, and evaluating Q'' from the lapse involves second derivatives of the metric. By advecting Q and Q' along the hypersurface normals, we incorporate them into the hyperbolic system in a consistent way. Our advection equations are

$$\partial_t Q - \beta Q' = 0, \quad \partial_t D_Q - \partial_x[\beta D_Q] = 0, \tag{2.17}$$

where $D_Q = Q'/Q$. Our advection of Q corresponds to harmonic slicing [21].

There is a danger with resetting the lapse and the shift, in that fluctuations in β' and D_Q can feed back on one another through the evolution equations for ${D_x}^x$ and ${K_x}^x$. The resetting gauge conditions of Sec. 2.6 imply that a fluctuation in ${K_x}^x$ is balanced by a fluctuation in β', and a fluctuation in ${D_x}^x$ is balanced by a fluctuation in D_Q. For certain time intervals between resetting, if these fluctuations propagate at different speeds, they may drift in such a way that they reinforce rather than cancel over much of the time interval. Although the standard procedure is to keep the shift constant in hyperbolic schemes, we find that if we advect D_Q with β' constant, such a positive feedback can occur, resulting in a runaway instability. However, if we advect both D_Q and β' along hypersurface normals, the evolution is stable. Our advection equations for β and β' are

$$\partial_t \beta - \beta\beta' = 0, \quad \partial_t \beta' - \partial_x[\beta\beta'] = 0. \tag{2.18}$$

2.5 Constraint Equations

The energy and momentum constraints must be satisfied by the initial conditions and throughout the evolution. We use these constraints to obtain the initial conditions. We do not impose the constraints during the evolution of the dynamical equations. However, we do insure that the boundary conditions are consistent with the constraint equations, and we use the constraints to check for accuracy and convergence as the numerical evolution proceeds. The energy and momentum constraint equations are, respectively,

$$\mathcal{E} = -\partial_x\left[\frac{1}{h_{xx}}(D_y{}^y + D_z{}^z)\right] - \frac{1}{h_{xx}}\left[D_y{}^y D_y{}^y + D_y{}^y D_z{}^z + D_z{}^z D_z{}^z + D_x{}^x(D_y{}^y + D_z{}^z)\right]$$
$$+ K_x{}^x(K_y{}^y + K_z{}^z) + K_y{}^y K_z{}^z = 0, \tag{2.19}$$

$$\mathcal{M}_x = -\partial_x(K_y{}^y + K_z{}^z) - D_y{}^y K_y{}^y - D_z{}^z K_z{}^z + (D_y{}^y + D_z{}^z)K_x{}^x = 0. \tag{2.20}$$

2.6 Resetting Gauge Conditions

The lapse and shift are periodically reset between time steps in order to implement a dynamic spacetime slicing which is unconstrained by the need to maintain a hyperbolic system. Our resetting gauge conditions are chosen to prevent pathologies and/or strong gradients from developing in the hypersurfaces and spatial coordinates, and to help stability properties at the boundaries of the grid. Resetting introduces discontinuities into the time evolution of the lapse and the shift at a given spatial location. However, in the 3+1 formalism, no time derivatives of the lapse and the shift appear in the equations. Resetting has no effect on the dynamical state of the system or on the gauge at a given time; it just affects how the coordinates evolve after the resetting.

Changes in spacetime slicing which maintain the explicit planar symmetry only directly impact $K_x{}^x$. Nonlinear source terms in the evolution equation for

${K_x}^x$ have the potential to generate runaway growth of ${K_x}^x$ when ${K_x}^x$ is positive. Our lapse resetting condition drives ${K_x}^x$ toward a small negative value to insure against this. In addition, a negative ${K_x}^x$ implies that the proper distance between hypersurface normals displaced in the x-direction increases with time. Together with our shift resetting condition, which keeps h_{xx} roughly constant, this results in hypersurface normals which point outward at the boundaries of the computational domain. Some features in ${K_x}^x$ and ${D_x}^x$ potentially associated with instability advect along hypersurface normals (see Sec. 2.10), and are then advected out of the grid before they can do much harm.

The equation for the lapse is derived by imposing the condition at the time of resetting

$$\partial_t {K_x}^x - \beta \partial_x {K_x}^x = -\Gamma[{K_x}^x - ({K_x}^x)_T], \tag{2.21}$$

where $({K_x}^x)_T$ is a specified "target" value, and Γ is a damping constant, which is chosen to be comparable to the characteristic frequency of the waves we are propagating. Substituting this condition into the evolution equation for ${K_x}^x$ (Eq. (2.13)) and simplifying using the energy constraint, we obtain our lapse resetting condition,

$$\partial_x \left(\frac{\alpha'}{h_{xx}}\right) = \alpha \left[\Gamma({K_x}^x - ({K_x}^x)_T) + {K_x}^x {K_x}^x - {K_y}^y {K_z}^z + \frac{{D_y}^y {D_z}^z}{h_{xx}}\right] - {D_x}^x \left(\frac{\alpha'}{h_{xx}}\right). \tag{2.22}$$

To limit initial transients in the lapse, given our initial condition ${K_x}^x = 0$, the target value is made proportional to $(1 - e^{-\Gamma t/4})$.

In our colliding wave calculations, Eq. (2.22) as it stands can cause the lapse to become negative at the edges of the grid, if the second derivative of the lapse becomes too negative. To prevent this, we replace $\mathcal{S}$, the expression in square brackets in Eq. (2.22), by

$$\mathcal{S} \longrightarrow \frac{\mathcal{S}}{\sqrt{1 + (\mathcal{S}/\mathcal{S}_{lim})^2}} \tag{2.23}$$

when $\mathcal{S}$ is negative, so $\mathcal{S} > -|\mathcal{S}_{lim}|$. A side effect of the limiter is to allow ${K_x}^x$ to become more negative than its target value.

We choose an equation for the shift so that at the time of resetting, h_{xx} is advected along hypersurface normals:

$$\partial_t h_{xx} - \beta \partial_x h_{xx} = 0. \tag{2.24}$$

Substituting this requirement into the evolution equation for h_{xx} (Eq. (2.9)), we obtain the shift resetting condition,

$$\partial_x \beta = \alpha {K_x}^x. \tag{2.25}$$

2.7 *Hyperbolic Systems*

The evolution equations presented in Sec. 2.3 have been cast in a first order, flux-conservative form, represented by the following set of l equations

$$\partial_t \mathbf{q} + \partial_x[\mathbf{F}(\mathbf{q}, x)] = \mathbf{S}(\mathbf{q}), \tag{2.26}$$

where $\mathbf{q}$ is a vector of l variables. The flux vector is given by

$$\mathbf{F}(\mathbf{q}, x) = \mathbf{A}(x)\mathbf{q}, \tag{2.27}$$

where the $l \times l$ characteristic matrix $\mathbf{A}(x)$ is the flux Jacobian, $\partial_{\mathbf{q}}[\mathbf{F}(\mathbf{q}, x)]$. The system is hyperbolic if $\mathbf{A}(x)$ has a complete set of eigenvectors and real eigenvalues.

2.7.1 *Modified Bona-Massó Formulation*

The standard BM formulation [16] creates a hyperbolic scheme by introducing the redundant variables V_i, which are defined as

$$V_i = {D_{ik}}^k - {D^k}_{ki} \Longrightarrow V_x = {D_y}^y + {D_z}^z. \tag{2.28}$$

The momentum constraint is used to evolve V_x:

$$\partial_t V_x + \partial_x[-\beta V_x] = \alpha[D_y{}^y K_y{}^y + D_z{}^z K_z{}^z - (D_y{}^y + D_z{}^z) K_x{}^x - (\alpha'/\alpha)(K_y{}^y + K_z{}^z)]. \quad (2.29)$$

We densitize the lapse according to the Choquet-Bruhat algebraic condition (see Sec. 2.4), which simplifies the standard BM system of equations considerably. With $\alpha \longrightarrow Q\sqrt{det(h_{ij})}$ and $(A_x = \partial_x \ln \alpha) \longrightarrow D_Q + D_x{}^x + D_y{}^y + D_z{}^z$, the fluxes for $K_i{}^j$ reduce to

$$F(K_x{}^x) = -\beta K_x{}^x + \frac{\alpha}{h_{xx}}\left(D_Q + D_x{}^x + \left(2 - \frac{n}{2}\right) V_x\right), \quad (2.30)$$

$$F(K_y{}^y) = -\beta K_y{}^y + \frac{\alpha}{h_{xx}}\left(D_y{}^y - \frac{n}{2} V_x\right), \quad F(K_z{}^z) = -\beta K_z{}^z + \frac{\alpha}{h_{xx}}\left(D_z{}^z - \frac{n}{2} V_x\right). \quad (2.31)$$

Our advection of Q and D_Q, as described in Sec. 2.4, corresponds to harmonic slicing, a special case of the standard BM lapse evolution equation. We also advect β and β', whereas the standard BM formulation specifies the shift as a known function of x and t.

2.7.2 Modified Anderson-York Formulation

The AY formulation differs from the BM scheme in how the momentum constraint is used to make the system hyperbolic. The AY scheme eliminates the need for the BM redundant variables V_i by incorporating the momentum constraint into the evolution equation for the f_{kij} variables, which are defined as

$$f_{kij} = D_{kij} + h_{ki} V_j + h_{kj} V_i. \quad (2.32)$$

The V_i variables in this equation are not separate variables, but rather denote the combinations of the D's given in Eq. (2.28).

The AY formulation replaces the BM D_{kij} with f_{kij}, which are simply the spatial metric derivative terms in the K_{ij} fluxes of the BM formulation. Using this as a guide, we generalize the AY scheme (whose original form is restricted

to Ricci evolution, $n = 0$) to allow for non-zero energy constraint contributions. This leads to

$$f_{kij} = D_{kij} + h_{ki}V_j + h_{kj}V_i - \frac{n}{2}V_k h_{ij}. \tag{2.33}$$

The generalization in Eq. (2.33) works as long as the inverse transformation from f_{kij} to D_{kij} exists, which is the case for $n \neq 1$. An evolution equation is obtained for f_{kij} from Eq. (2.33) by using the momentum constraint to eliminate the time derivative of the V_i variables. For our modified AY scheme, we then raise one index so that $f_{ki}{}^{j} = f_{kil}h^{lj}$ are our basic variables. A hyperbolic system results without the need for the BM redundant *variables* V_i.

For the diagonal metric planewave case under consideration, Eq. (2.33) reduces to

$$f_x{}^x = D_x{}^x + \left(2 - \frac{n}{2}\right) V_x, \tag{2.34}$$

$$f_y{}^y = D_y{}^y - \frac{n}{2}V_x, \quad f_z{}^z = D_z{}^z - \frac{n}{2}V_x. \tag{2.35}$$

We have simplified our notation in that $f_{ki}{}^{j} \longrightarrow f_i{}^j$ for this 1D problem. Notice that $f_{yx}{}^y = V_x$ and $f_{zx}{}^z = V_x$, which contribute to fluxes in the y and z directions, are not zero. However, with planar symmetry the divergence of these flux components vanishes identically.

The evolution equations for $f_i{}^j$ are

$$\partial_t f_x{}^x + \partial_x[-\beta f_x{}^x - \beta' + \alpha K_x{}^x] = \left(2 - \frac{n}{2}\right)\alpha\mathcal{C}, \tag{2.36}$$

$$\partial_t f_y{}^y + \partial_x[-\beta f_y{}^y + \alpha K_y{}^y] = -\frac{n}{2}\alpha\mathcal{C}, \quad \partial_t f_z{}^z + \partial_x[-\beta f_z{}^z + \alpha K_z{}^z] = -\frac{n}{2}\alpha\mathcal{C}, \tag{2.37}$$

where

$$\mathcal{C} = [D_y{}^y K_y{}^y + D_z{}^z K_z{}^z - (D_y{}^y + D_z{}^z)K_x{}^x - (\alpha'/\alpha)(K_y{}^y + K_z{}^z)]. \tag{2.38}$$

The D's in Eq. (2.38) are not separate variables, but denote:

$$D_x{}^x = f_x{}^x - \left(\frac{2 - \frac{n}{2}}{1 - n}\right)[f_y{}^y + f_z{}^z], \tag{2.39}$$

$$D_y{}^y = \frac{1}{2(1-n)}[(2-n)f_y{}^y + nf_z{}^z], \quad D_z{}^z = \frac{1}{2(1-n)}[(2-n)f_z{}^z + nf_y{}^y]. \tag{2.40}$$

These relations are the inverse transformation of the system of Eqs. (2.34) to (2.35). One can see that $n = 1$ is not allowed.

The $K_i{}^j$ evolution equations are the same as in our modified BM scheme, with the understanding again that the D's in the source terms are not separate variables, but the above linear combinations of f's (Eqs. (2.39) to (2.40)). The fluxes are defined so the $f_i{}^j$ variables can replace the expressions involving $D_i{}^j$ in the fluxes of our modified BM scheme. The following fluxes result:

$$F(K_x{}^x) = -\beta K_x{}^x + \frac{\alpha}{h_{xx}}(D_Q + f_x{}^x), \tag{2.41}$$

$$F(K_y{}^y) = -\beta K_y{}^y + \frac{\alpha}{h_{xx}} f_y{}^y, \quad F(K_z{}^z) = -\beta K_z{}^z + \frac{\alpha}{h_{xx}} f_z{}^z. \tag{2.42}$$

The AY formalism imposes the Choquet-Bruhat algebraic condition on the lapse, as we did in our modified BM scheme. The evolution of the lapse and the shift between gauge resettings is treated in exactly the same way as in our modified BM formalism.

2.7.3 Modified Arnowitt-Deser-Misner Formulation

The simplest of the hyperbolic schemes we present is our modified ADM formulation, which consists of Eqs. (2.9) to (2.14), (2.17), and (2.18), with $\alpha = Q\sqrt{det(h_{ij})}$ and $\alpha'/\alpha = D_Q + D_x{}^x + D_y{}^y + D_z{}^z$. This system is hyperbolic when the metric is diagonal if $n < 0$ or $0 < n < 1$. The fluxes for $K_i{}^j$ are

$$F(K_x{}^x) = -\beta K_x{}^x + \frac{\alpha}{h_{xx}}\left[D_Q + D_x{}^x + \left(2 - \frac{n}{2}\right)(D_y{}^y + D_z{}^z)\right], \tag{2.43}$$

$$\begin{aligned} F(K_y{}^y) &= -\beta K_y{}^y + \frac{\alpha}{h_{xx}}\left[\left(1 - \frac{n}{2}\right) D_y{}^y - \frac{n}{2} D_z{}^z\right], \\ F(K_z{}^z) &= -\beta K_z{}^z + \frac{\alpha}{h_{xx}}\left[\left(1 - \frac{n}{2}\right) D_z{}^z - \frac{n}{2} D_y{}^y\right]. \end{aligned} \tag{2.44}$$

The hyperbolicity of our modified ADM system of equations breaks down for $n = 0$ and $n \geq 1$. Although our ADM formulation at $n = 0$ is non-hyperbolic, it is stable. At $n = 1$, however, the system is both non-hyperbolic and on the verge of being unstable. For $n > 1$, the equations have complex eigenvalues, giving unstable exponential growth of errors.

2.7.4 Wave Modes

BM

The hyperbolic system of equations obtained from the modified BM formulation described in Sec. 2.7.1 is

$$\partial_t \mathbf{q} + \partial_x [\mathbf{A}(x)\mathbf{q}] = \mathbf{S}(\mathbf{q}), \tag{2.45}$$

where

$$\mathbf{q} = \begin{pmatrix} D_x{}^x \\ D_y{}^y \\ D_z{}^z \\ K_x{}^x \\ K_y{}^y \\ K_z{}^z \\ V_x \\ D_Q \\ \beta' \end{pmatrix}, \tag{2.46}$$

and

$$\mathbf{A}(x) = \begin{pmatrix} -\beta & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & -1 \\ 0 & -\beta & 0 & 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & -\beta & 0 & 0 & \alpha & 0 & 0 & 0 \\ \frac{\alpha}{h_{xx}} & 0 & 0 & -\beta & 0 & 0 & \frac{\alpha}{h_{xx}}\left(2-\frac{n}{2}\right) & \frac{\alpha}{h_{xx}} & 0 \\ 0 & \frac{\alpha}{h_{xx}} & 0 & 0 & -\beta & 0 & -\frac{\alpha}{h_{xx}}\left(\frac{n}{2}\right) & 0 & 0 \\ 0 & 0 & \frac{\alpha}{h_{xx}} & 0 & 0 & -\beta & -\frac{\alpha}{h_{xx}}\left(\frac{n}{2}\right) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta \end{pmatrix}. \tag{2.47}$$

The nine eigenmodes of the homogeneous system are obtained from the characteristic matrix, $\mathbf{A}(x)$. Six of the eigenmodes travel along the light cones. They are:

$$\left.\begin{array}{c} \frac{1}{\sqrt{h_{xx}}}\left[{D_x}^x + D_Q + \left(2-\frac{n}{2}\right)V_x\right] \pm \left[{K_x}^x - \frac{\beta'}{\alpha}\right], \\ \frac{1}{\sqrt{h_{xx}}}\left[{D_y}^y - \frac{n}{2}V_x\right] \pm {K_y}^y,\ \frac{1}{\sqrt{h_{xx}}}\left[{D_z}^z - \frac{n}{2}V_x\right] \pm {K_z}^z \end{array}\right\} speeds = -\beta \pm \frac{\alpha}{\sqrt{h_{xx}}}. \tag{2.48}$$

The remaining three eigenmodes are simply the variables V_x, D_Q, and β', which travel along the hypersurface normals, with speeds $-\beta$.

The eigenmodes of the characteristic matrix, however, do not necessarily describe how solutions of the full nonlinear system of equations propagate. It is a special property of planewave systems that eigenmodes of the full nonlinear system of equations exist which consist of purely right-going waves with ${K_y}^y \pm {K_z}^z = ({D_y}^y \pm {D_z}^z)/\sqrt{h_{xx}}$, purely left-going waves with ${K_y}^y \pm {K_z}^z = -({D_y}^y \pm {D_z}^z)/\sqrt{h_{xx}}$, and ${D_x}^x = {K_x}^x = 0$. These are solutions of the Einstein equations in a gauge with $\alpha = 1$ and $\beta' = 0$. In our nonlinear colliding plane wave calculations, our initial conditions are such that the waves have this form. The right-going wave is in the left half of the grid, the left-going wave is in the right half of the grid, and they are just at the point of colliding. When

discussing solutions of the full nonlinear system of equations, we refer to the transverse-traceless quantities $(D_y{}^y - D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y - K_z{}^z)$, the constraint quantities $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y + K_z{}^z)$, and the longitudinal variables $D_x{}^x/\sqrt{h_{xx}}$ and $K_x{}^x$. After the waves pass through each other, it is only approximately true that the transverse-traceless quantities have the form of purely right-going and purely left-going waves as described above and it is not at all true that the constraint quantities have this form.

The characteristic speeds apply to small amplitude, short wavelength perturbations in the variables, so that the principal terms (which are first derivative terms) dominate over the source terms. The disturbances in the constraint quantities which propagate along the characteristics will generally be constraint-violating because the constraints explicitly tie the principal terms to the nonlinear source terms, and require that they cancel. The longitudinal variables, $D_x{}^x/\sqrt{h_{xx}}$ and $K_x{}^x$, are not eigenmodes of the homogeneous system. In the full nonlinear system, $D_x{}^x/\sqrt{h_{xx}}$ and $K_x{}^x$ have some features which propagate along the light cones, and some features which propagate along the hypersurface normals. The propagation of the longitudinal variables is strongly dependent on the choice of gauge.

AY

There is a complete set of eight eigenmodes of the modified AY homogeneous system of equations. The six eigenmodes which travel along the light cones are

$$\left.\begin{array}{c} \frac{1}{\sqrt{h_{xx}}}(f_x{}^x + D_Q) \pm \left(K_x{}^x - \frac{\beta'}{\alpha}\right), \\ \frac{f_y{}^y}{\sqrt{h_{xx}}} \pm K_y{}^y,\ \frac{f_z{}^z}{\sqrt{h_{xx}}} \pm K_z{}^z \end{array}\right\} speeds = -\beta \pm \frac{\alpha}{\sqrt{h_{xx}}}. \qquad (2.49)$$

The remaining two eigenmodes are the variables D_Q and β', which travel along the hypersurface normals, with speeds $-\beta$.

ADM

For the modified ADM homogeneous system, the eigenmodes, which form a complete hyperbolic system for $n < 0$ or $0 < n < 1$, consist of "longitudinal" and "physical" eigenmodes propagating along the light cone,

$$\frac{1}{\sqrt{h_{xx}}}\left[nD_x{}^x + nD_Q + \left(2-\frac{n}{2}\right)\left(D_y{}^y + D_z{}^z\right)\right] \pm \left[nK_x{}^x - \frac{n\beta'}{\alpha} + \left(2-\frac{n}{2}\right)\left(K_y{}^y + K_z{}^z\right)\right],$$

$$\frac{1}{\sqrt{h_{xx}}}[D_y{}^y - D_z{}^z] \pm [K_y{}^y - K_z{}^z], speeds = -\beta \pm \frac{\alpha}{\sqrt{h_{xx}}},$$

"constraint" eigenmodes propagating inside the light cone for $0 < n < 1$,

$$\sqrt{\tfrac{1-n}{h_{xx}}}\,[D_y{}^y + D_z{}^z] \pm [K_y{}^y + K_z{}^z] \Big\} speeds = -\beta \pm \frac{\alpha}{\sqrt{h_{xx}}}\sqrt{1-n}, \qquad (2.50)$$

and the D_Q and β' eigenmodes with speeds $-\beta$. Hyperbolicity fails for $n = 0$ because the "longitudinal" eigenmodes are not independent of the "constraint" eigenmodes, for $n = 1$ because the two "constraint" eigenmodes are not independent of each other, and for $n > 1$ because the "constraint" eigenvalues are complex.

2.8 Boundary Conditions

Since in numerical relativity, computations are usually performed on a limited grid within a much larger space, the boundary conditions should be designed to be consistent with how waves propagate while they are still inside the grid. Even more important, since the evolution equations admit constraint-violating solutions, constraint violations will propagate into the grid unless boundary conditions are carefully designed to suppress them.

Consider the "constraint" eigenmodes. They are $[D_y{}^y + D_z{}^z - nV_x]/\sqrt{h_{xx}} \pm [K_y{}^y + K_z{}^z]$ in BM and AY (though expressed in terms of different variables), and $\sqrt{1-n}\,[D_y{}^y + D_z{}^z]/\sqrt{h_{xx}} \pm [K_y{}^y + K_z{}^z]$ in ADM. Even for the same value of the energy constraint coefficient n, what is outgoing in BM and AY is different from

what is outgoing in ADM. Furthermore, for a given solution, the amplitudes of the BM, AY, and ADM modes depend on n. Whatever the correct boundary condition, its effect on the solution should be independent of the equation formulation. The relative amount of right and left-going "constraint" modes is also gauge dependent, in the sense that the choice of boundary conditions in solving the constraint equations in the initial conditions is a gauge choice, and this affects the relative values of $(D_y{}^y + D_z{}^z)$ and $(K_y{}^y + K_z{}^z)$ at all later times. The initial conditions symmetric about the midpoint of the grid at $x = 10$ give purely *incoming* "constraint" modes for $n = 0$ ($[D_y{}^y + D_z{}^z]/\sqrt{h_{xx}} = \pm[K_y{}^y + K_z{}^z]$ on the left/right edges of the grid) initially and at all times until the effects of the wave collision reach the boundaries.

The "longitudinal" eigenmodes involving $D_x{}^x$ and $K_x{}^x$ are also formulation dependent, since they are different in ADM from what they are in BM and AY, and they depend on n in all three formulations. There is gauge freedom to pose any boundary conditions one likes on these modes, but a poor choice might give rise to singularities in $D_x{}^x$ or $K_x{}^x$ inside the grid.

The "physical" eigenmodes $[D_y{}^y - D_z{}^z]/\sqrt{h_{xx}} \pm [K_y{}^y - K_z{}^z]$ are the same in all three formulations, and are independent of n. However, their time evolution is gauge-dependent because the nonlinear source terms in their evolution equations involve the gauge-dependent constraint quantities. With our choice of initial gauge, the amplitudes of $(D_y{}^y - D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y - K_z{}^z)$ differ by about 3 per cent after the wave collision, so there is typically a 3 per cent admixture of the incoming "physical" eigenmode as the outgoing waves approach the boundaries.

A gauge-independent measure of the amplitudes of left and right-going gravitational waves can be obtained by projecting the Weyl tensor onto a complex

null tetrad, as in the Newman-Penrose spin coefficient formalism [57],

$$\Psi^0_\pm = R_{(t)(y)(t)(y)} - R_{(t)(z)(t)(z)} + R_{(x)(y)(x)(y)} - R_{(x)(z)(x)(z)} \mp 2[R_{(t)(y)(x)(y)} - R_{(t)(z)(x)(z)}] \tag{2.51}$$

for right/left propagation. A purely right-going wave would have $\Psi^0_- = 0$. Our numerical results indicate that plane waves after a collision are indeed purely outgoing by this standard. As an outgoing wave boundary condition, using the evolution equations to evaluate the time derivatives of the extrinsic curvature in the Riemann tensor, this becomes

$$\begin{aligned} &\left(\partial_x \left[\frac{(D_y{}^y - D_z{}^z)}{\sqrt{h_{xx}}} \mp (K_y{}^y - K_z{}^z)\right]\right)\frac{1}{\sqrt{h_{xx}}} \\ &+\frac{1}{2}\frac{(D_y{}^y - D_z{}^z)}{\sqrt{h_{xx}}}\left[\frac{(D_y{}^y + D_z{}^z)}{\sqrt{h_{xx}}} \mp (K_y{}^y + K_z{}^z)\right] \\ &+\frac{1}{2}\frac{(D_y{}^y + D_z{}^z)}{\sqrt{h_{xx}}}\left[\frac{(D_y{}^y - D_z{}^z)}{\sqrt{h_{xx}}} \mp (K_y{}^y - K_z{}^z)\right] \;=\; 0 \end{aligned} \tag{2.52}$$

at the right/left boundaries. This expression is consistent with $[D_y{}^y - D_z{}^z]/\sqrt{h_{xx}} = \pm[K_y{}^y - K_z{}^z]$ if and only if $[D_y{}^y + D_z{}^z]/\sqrt{h_{xx}} = \pm[K_y{}^y + K_z{}^z]$.

Since conventional outgoing wave boundary conditions are not appropriate, our boundary conditions are based on a smooth second order extrapolation of the variables, which is corrected to make sure the energy and momentum constraint equations are satisfied on the boundaries. Eq. (2.52) could also be imposed at the boundaries to further improve the extrapolation, but we have not tried this. Our procedure is detailed further in Sec. 2.9.2.

In addition to the eigenmodes discussed above, there are eigenmodes propagating along the hypersurface normals, which can be incoming or outgoing, depending on the sign of the shift on the boundaries. It seems to be important for stability that the hypersurface normals do not point into the grid (see Secs. 2.10.3 and 2.10.4).

Our results show that quadratic extrapolation without correction for the energy and momentum constraints produces a significant but not dominant error

(see Sec. 2.10.4). However, errors from imposing outgoing boundary conditions on the "constraint" eigenmodes, or from using standard constant extrapolation, would swamp all other errors as they propagate into the grid. Standard constant extrapolation, which gives the same values for the variables, and therefore the fluxes, in the ghost cell and adjoining physical cell, also eliminates the incoming "constraint" eigenmodes.

2.9 Numerical Methods

2.9.1 Strang Splitting

As described in Sec. 2.7, all of the formulations we tested, both hyperbolic and non-hyperbolic, are in first order, flux conservative form. We solve all these sytems of equations using a Strang-split method [59]. In this method, the homogeneous transport part of Eq. (2.26) and the contributions from the source terms are treated separately. In particular, the following straightforward system of ordinary differential equations is first solved over half a time step

$$\partial_t \mathbf{q} = \mathbf{S}(\mathbf{q}). \tag{2.53}$$

Then, the transport part of Eq. (2.26), which contains the flux terms, is solved over a full time step

$$\partial_t \mathbf{q} + \partial_x[\mathbf{F}(\mathbf{q}, x)] = 0. \tag{2.54}$$

Our methods for solving the transport step are discussed in Sec. 2.9.2 below. The calculation is completed by again solving Eq. (2.53) over half a time step.

We choose to use the Strang-split method because it is simpler in the context of how we are handling boundary conditions. An iterative scheme such as the MacCormack method [52] requires repeated implementation of the boundary conditions each time step. However, in the Strang-split scheme, the boundary conditions are imposed only once during each time step. The fewer applications

of the boundary conditions in the Strang-split method is advantageous because we are using quadratic extrapolation to obtain ghost cell values. Quadratic extrapolation amplifies any jitter at the boundaries, and the frequent application of quadratic extrapolation in iterative schemes such as MacCormack could easily lead to an instability.

2.9.2 Transport Step

In the transport step, we solve Eq. (2.54) both with a finite difference method and with a wave propagation approach, which takes advantage of the eigenfields of a diagonalizable hyperbolic system. Advanced numerical methods for diagonalizable hyperbolic systems introduce limiter functions to resolve sharp discontinuities that typically arise in hydrodynamics problems. A smooth problem can be solved just as accurately and more efficiently with a finite difference method. In vacuum general relativity, discontinuities may or may not arise, depending on the gauge conditions. Commonly used gauge conditions lead to steep gradients near black hole horizons. One can deal with these gradients by using high resolution methods requiring diagonalizable hyperbolic formulations; or, one can dynamically adjust the gauge conditions so as to avoid the steep gradients altogether [6].

Whether one uses a finite difference method or a sophisticated hyperbolic technique, it is important to have a numerical scheme which is fully second order accurate for smooth solutions and generalizable to black hole spacetimes and higher dimensions. It is straightforward to devise a finite difference scheme based on a Taylor series expansion which is formally second order accurate. High resolution Riemann-based wave propagation algorithms introduced by LeVeque [47], which decompose $\triangle \mathbf{q}$ across a grid cell interface into a linear combination of eigenvectors of the $\mathbf{A}(x)$ matrix, are applicable to a wide variety of diagonalizable hyperbolic problems. Flux differences are calculated from

the $\triangle\mathbf{q}$ decomposition. We refer to these algorithms as "standard wave decomposition" methods. However, the standard wave decomposition methods are not second order accurate for smooth solutions when the characteristic matrix $\mathbf{A}(x)$ is a function of position, because the changes in $\mathbf{A}(x)$ across cell boundaries as well as $\triangle\mathbf{q}$'s must be accounted for in flux differences. In numerical relativity problems, $\mathbf{A}(x)$ depends on the lapse, the shift, and the spatial metric, and can have gradients comparable with the gradients of $\mathbf{q}$.

LeVeque has suggested a wave propagation approach for solving variable coefficient flux problems based on splitting up the jump in $\mathbf{F}(\mathbf{q}, x)$ rather than the jump in $\mathbf{q}$ [44]. We refer to this approach as "flux-based wave decomposition". We develop and apply this method to solve the Einstein equations for 1D nonlinear plane waves as described below in Sec. 2.9.2. We show in Appendix A that flux-based wave decomposition methods are formally second order accurate for sufficiently smooth solutions for arbitrary smooth variations of the eigenvalues and eigenvectors (see also Bale *et al.* [11]). For further discussion and analysis of flux-based wave decomposition methods, in the context of more general approximate Riemann solvers, see [49]. While it is difficult to formally prove second order convergence for numerical methods since this also requires proving stability, our numerical tests of these methods, and those of reference [11], typically exhibit second order convergence.

Flux-Based Wave Decomposition

Using Eq. (2.54) to update average grid cell values of the variables $\mathbf{q}$ requires knowing flux values at grid cell interfaces. The interface flux values are found by solving the following equation, obtained by multiplying Eq. (2.54) by $\mathbf{A}(x)$ on the left hand side:

$$\partial_t[\mathbf{F}(\mathbf{q}, x)] + \mathbf{A}(x)\partial_x[\mathbf{F}(\mathbf{q}, x)] = 0. \tag{2.55}$$

The time derivative of $\mathbf{A}(x)$ vanishes because it is time independent during the transport step. Using Eq. (2.55) to compute the interface fluxes was originally described by Bona *et al.* [16]. However, it is not clear from [16] how they handled problems in which $\mathbf{A}(x)$ varies from cell to cell.

Eq. (2.55) is a linear advection equation for the flux vector, $\mathbf{F}(\mathbf{q}, x)$. As such, flux values at cell interfaces can be updated by solving Riemann problems based on decomposing flux differences between adjacent grid cells into eigenvector expansions (see [46] for a discussion of solving Riemann problems for the advection equation), and including correction terms to give second order accuracy. We develop two wave propagation methods based on this idea which we call Methods **I** and **II**. A wave in this approach is defined as a discontinuity in the *flux* associated with a certain eigenmode across the characteristic corresponding to that eigenmode.

We explicitly deal with the fact that the eigenvalues and eigenvectors of the characteristic matrix are varying across the grid. The magnitudes of the eigenvalues give the wave speeds and the signs of the eigenvalues give the wave directions. In the flux decomposition for Method **I**, we need to decide if a wave is left-going or right-going at a given cell interface. This is determined by the sign of the average of the eigenvalues obtained from the characteristic matrices on either side of the interface. If the average eigenvalue for a particular eigenmode is negative, then the corresponding eigenvector is evaluated in the cell to the left of the interface. If the average eigenvalue is positive, then the eigenvector is evaluated in the cell to the right of the interface. In Method **II**, the eigenvalues and eigenvectors at a cell interface are obtained from the characteristic matrix at the interface, calculated as an average from the adjacent cells. For both methods, waves with zero interface speed still contribute to the flux difference. We can include these contributions in either the left-or right-going waves of Method **I**, as long as we do so consistently.

In Method **I**, the flux difference decomposition takes the following form at the interface between cells i and $i-1$:

$$\mathbf{F}(\mathbf{q}_i) - \mathbf{F}(\mathbf{q}_{i-1}) = \mathbf{A}_i\mathbf{q}_i - \mathbf{A}_{i-1}\mathbf{q}_{i-1} = \sum_{L=1}^{m} \gamma^L_{i-\frac{1}{2}}\mathbf{r}^L_{i-1} + \sum_{R=m+1}^{M} \gamma^R_{i-\frac{1}{2}}\mathbf{r}^R_i, \qquad (2.56)$$

where $\mathbf{r}$ are right eigenvectors of the characteristic matrix, and M is the total number of eigenmodes. We denote the left-going waves at this interface as $\mathbf{W}^L_{i-\frac{1}{2}} = \gamma^L_{i-\frac{1}{2}}\mathbf{r}^L_{i-1}$, where $1 \le L \le m$. The right-going waves are given by $\mathbf{W}^R_{i-\frac{1}{2}} = \gamma^R_{i-\frac{1}{2}}\mathbf{r}^R_i$, where $m+1 \le R \le M$. The number of left-going waves, m, can vary from interface to interface since the sign of the average eigenvalue can change from cell to cell. The eigenvectors $\mathbf{r}_{i-1}$ are evaluated in cell $i-1$. Likewise, $\mathbf{r}_i$ are evaluated in cell i. The coefficients $\gamma_{i-\frac{1}{2}}$ are obtained by solving Eq. (2.56); the subscripts $i-\frac{1}{2}$ indicate interface values. In Method **II**, the flux difference decomposition at a given interface between cells i and $i-1$ is the same as Eq. (2.56), except the eigenvectors $\mathbf{r}_{i-\frac{1}{2}}$ of the averaged characteristic matrix $\mathbf{A}_{i-\frac{1}{2}} = (\mathbf{A}_{i-1} + \mathbf{A}_i)/2$ replace both $\mathbf{r}_{i-1}$ and $\mathbf{r}_i$.

Method **I** is implemented in the context of the CLAWPACK software package [45]. The first order wave propagation and second order corrections in both Methods **I** and **II** are analogous to Eqs. (18) and (19) of LeVeque's paper on standard wave decomposition methods [47]. The updated value of $\mathbf{q}_i$ is given by

$$\overline{\mathbf{q}}_i = \mathbf{q}_i - \frac{\triangle t}{\triangle x}\left(\sum_R \mathbf{W}^R_{i-\frac{1}{2}} + \sum_L \mathbf{W}^L_{i+\frac{1}{2}}\right) - \frac{\triangle t}{\triangle x}\left(\tilde{\mathbf{F}}_{i+\frac{1}{2}} - \tilde{\mathbf{F}}_{i-\frac{1}{2}}\right). \qquad (2.57)$$

$\tilde{\mathbf{F}}_{i\pm\frac{1}{2}}$ are flux correction terms which can be reduced near discontinuities by introducing limiter functions. Limiters prevent the oscillatory behavior around discontinuities seen with finite difference methods. In the absence of limiting, the flux corrections are

$$\tilde{\mathbf{F}}_{i\pm\frac{1}{2}} = \frac{1}{2}\left(\sum_R \mathbf{W}^R_{i\pm\frac{1}{2}} - \sum_L \mathbf{W}^L_{i\pm\frac{1}{2}}\right) - \frac{1}{2}\frac{\triangle t}{\triangle x}\sum_{p=1}^{M} \lambda^p_{i\pm\frac{1}{2}}\mathbf{W}^p_{i\pm\frac{1}{2}}, \qquad (2.58)$$

where $\lambda^p_{i\pm\frac{1}{2}}$ denote cell-interface speeds.

Both flux-based wave decomposition methods **I** and **II** are successful in giving second order convergent results in our numerical calculations.

Finite Difference Method

To solve Eq. (2.54) using a Lax-Wendroff finite difference method, we first perform a second order Taylor expansion of $\mathbf{q}$ around t:

$$\mathbf{q}(x, t + \triangle t) = \mathbf{q}(x, t) + \triangle t\ \partial_t \mathbf{q}(x, t) + \frac{1}{2}\triangle t^2\ \partial_t^2 \mathbf{q}(x, t). \tag{2.59}$$

Observe that

$$\partial_t \mathbf{q} = -\partial_x[\mathbf{A}(x)\mathbf{q}], \tag{2.60}$$

and, taking another time derivative,

$$\partial_t^2 \mathbf{q} = -\partial_x[\mathbf{A}(x)(\partial_t \mathbf{q})] = \partial_x[\mathbf{A}(x)\partial_x(\mathbf{A}(x)\mathbf{q})]. \tag{2.61}$$

Note that the time derivative of $\mathbf{A}(x)$ vanishes as in Eq. (2.55). Plugging these expressions for $\partial_t \mathbf{q}$ and $\partial_t^2 \mathbf{q}$ into Eq. (2.59) gives

$$\mathbf{q}(x, t + \triangle t) = \mathbf{q}(x, t) - \triangle t\ \partial_x[\mathbf{A}(x)\mathbf{q}(x, t)] + \frac{1}{2}\triangle t^2\ \partial_x[\mathbf{A}(x)\partial_x(\mathbf{A}(x)\mathbf{q}(x, t))]. \tag{2.62}$$

Making the centered finite difference approximation to the derivatives in Eq. (2.62), the updated value of $\mathbf{q}_i$ is given by

$$\begin{aligned}\overline{\mathbf{q}}_i = \mathbf{q}_i &- \frac{\triangle t}{\triangle x}(\mathbf{A}_{i+1}\mathbf{q}_{i+1} - \mathbf{A}_i\mathbf{q}_i) \\ &+ \frac{1}{4}\left(\frac{\triangle t}{\triangle x}\right)^2 [(\mathbf{A}_i + \mathbf{A}_{i+1})(\mathbf{A}_{i+1}\mathbf{q}_{i+1} - \mathbf{A}_i\mathbf{q}_i) - (\mathbf{A}_{i-1} + \mathbf{A}_i)(\mathbf{A}_i\mathbf{q}_i - \mathbf{A}_{i-1}\mathbf{q}_{i-1})].\end{aligned} \tag{2.63}$$

Boundary Conditions

Our numerical methods only require values in one ghost cell at each boundary. We obtain values for all variables in the ghost cell by quadratic extrapolation

from the three adjacent physical cells. Numerical integration of the constraint equations from the last physical cell to the ghost cell by the trapezoidal rule is used to correct the constraint quantities $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y + K_z{}^z)$ in the ghost cell, with iteration to convergence.

2.10 Results

2.10.1 Initial Conditions

The initial conditions must satisfy the constraint equations, Eqs. (2.19) and (2.20). Since the constraint equations are differential equations, they require boundary conditions for their solutions. Different choices of boundary conditions correspond to different gauge conditions. We choose symmetric boundary conditions which give flat space between two waves. This means that we choose $(D_y{}^y + D_z{}^z)$ and $(K_y{}^y + K_z{}^z)$ to vanish initially between the waves.

The variables h_{xx}, $K_x{}^x$, and the combinations $(D_y{}^y - D_z{}^z)$ and $(K_y{}^y - K_z{}^z)$ are freely specifiable. We normally take $h_{xx} = 1$, $K_x{}^x = 0$, and

$$0.25 \ln\left(\frac{h_{yy}}{h_{zz}}\right) = \sum_{i=1}^{2} A_i \cos^2\left[\frac{\pi}{2}\frac{(x - x_{0i})}{w_i}\right] \sin\left[k_i(x - x_{0i}) + \delta_i\right], \tag{2.64}$$

for $-w_i < (x - x_{0i}) < w_i$, and zero outside that range. The x derivative of Eq. (2.64) gives $(D_y{}^y - D_z{}^z)/2$. In our standard initial conditions for colliding plane waves, one wave is initially on the left and moving to the right, with $[K_y{}^y - K_z{}^z]_1 = [(D_y{}^y - D_z{}^z)/\sqrt{h_{xx}}]_1$. The other wave is initially on the right and moving to the left, with $[K_y{}^y - K_z{}^z]_2 = -[(D_y{}^y - D_z{}^z)/\sqrt{h_{xx}}]_2$. The parameters for $0 \leq x \leq 20$ are $w_i = 4.0$, $k_i = 1.6$, $A_i = 0.08$, $x_{01} = 6.0$, $x_{02} = 14.0$, $\delta_1 = 0$, and $\delta_2 = \pi$. These initial conditions, depending on the variables, are symmetric (or antisymmetric) about $x = 10$, and symmetry (or antisymmetry) is preserved throughout the evolution. Hence, our figures only show the range $0 \leq x \leq 10$. Since the initial plane waves do not overlap, and $(D_y{}^y + D_z{}^z)$ and $(K_y{}^y + K_z{}^z)$

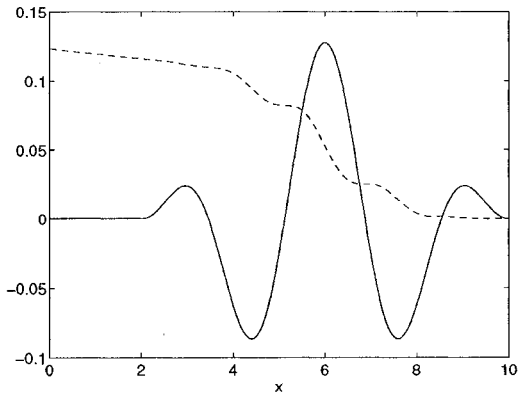


Figure 2.1: Initial conditions for derivatives of the transverse metric. The solid line is $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$ and the dashed line is $2(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$. Note that $x = 10$ is the center of the grid.

vanish at $x = 0$, the initial conditions are two analytic single plane waves of the type described by Misner, Thorne, and Wheeler [54].

Our initial conditions produce large amplitude, nonlinear colliding gravitational plane waves. Our measure of "large amplitude" is that h_{yy} and h_{zz} are substantially different from 1 by the time the waves have traversed the grid. It is known that nonlinear planewave spacetimes develop a singularity behind the wave [85, 86]. For a single plane wave, this is only a coordinate singularity, while for colliding plane waves, a physical singularity also develops. The values we take for our wave amplitudes are about as large as possible without allowing a singularity to develop during the crossing time of the waves. One can get a feel for this value by asking at what amplitude does a singularity develop at the left edge of the grid for a single plane wave exiting the right edge? For a single wave as given by Eq. (2.64) with the shape specified by our values for w_i and k_i, and a flat metric ahead of the wave, the answer is approximately 0.11. This is an upper limit, however, because the effects of colliding waves add together in a way which is hard to estimate. The initial conditions for $(D_y{}^y \pm D_z{}^z)$

are shown in Fig. 2.1.

2.10.2 Comparing Evolution Systems

Testing for the Optimal System

We experiment with several different formulations of the Einstein equations to determine the factors involved in improving the global accuracy of 1D colliding gravitational plane wave calculations. The basic formalisms we test are the modified BM, AY, and ADM schemes of Sec. 2.7. In all of these schemes, using mixed variables rather than lowered variables improves accuracy significantly. We also compare alternative ways of handling the redundant variable V_x in the BM schemes. V_x can be left to evolve independently (no-reset BM), or it can be reset periodically to enforce the constraint that $V_x = D_y{}^y + D_z{}^z$ (reset BM). Results have been calculated for a range of values of the coefficient n of the energy constraint term in the extrinsic curvature evolution equations, from about -0.4 to 1.0, and in some cases for values of $n > 1$. For the ADM and reset BM schemes, the results near 0 and 1 reflect the breakdown of hyperbolicity at these values of n. Results are primarily shown for $t = 12$ since this is the latest time at which the physical waves are largely within the grid.

Fig. 2.2 shows the evolution of linear combinations of metric derivatives appearing in the eigenmodes from $t = 8$ to $t = 12$, after the physical waves have finished colliding. In these high resolution (4000 cell) calculations, the numerical errors are negligible on the scale of the graph, and we have verified that all the different formulations seem to be converging to the same solution. The quantity $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$ is shown in Fig. 2.2a. The coordinate speed of propagation can be read off the graph: it is roughly two units in x for every two units of time until around $t = 12$, when the coordinate speed of light starts to differ significantly from one. Over the same range of times, the quan-

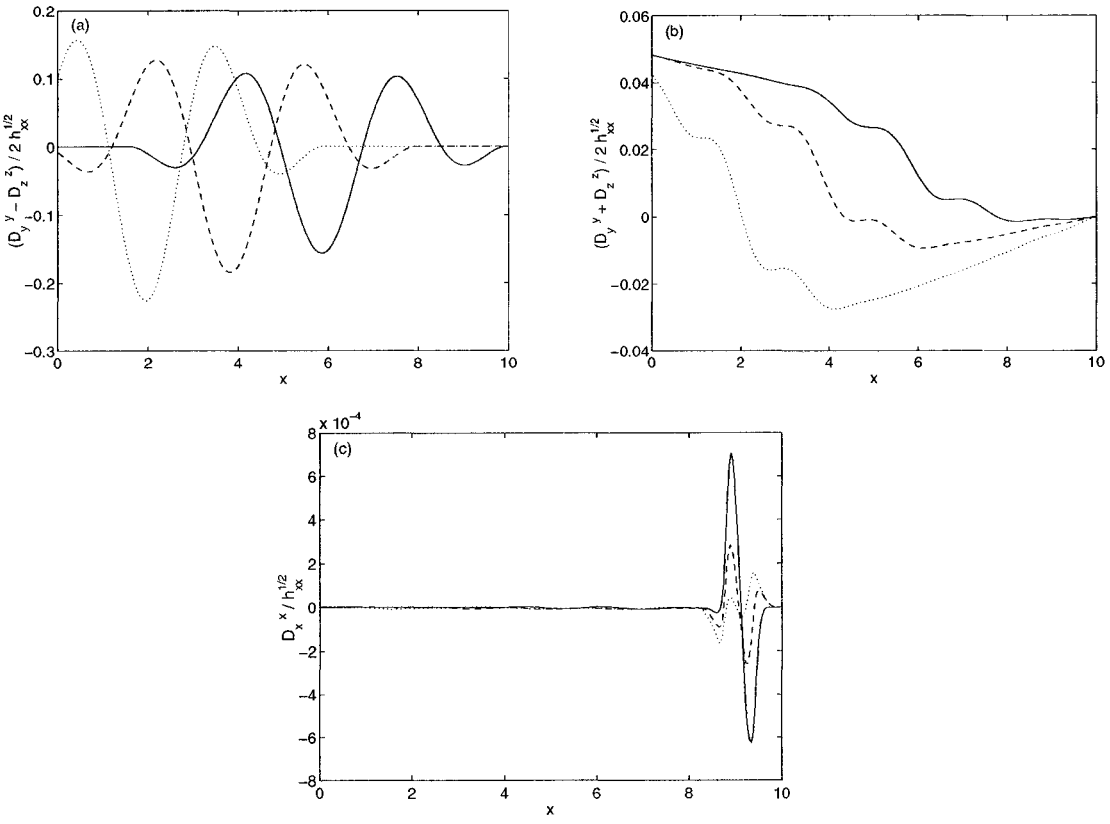


Figure 2.2: Evolution of the metric derivatives. The solid line is at $t = 8$, the dashed line at $t = 10$, and the dotted line at $t = 12$. (a) $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$, (b) $(D_y{}^y + D_z{}^z)/(2\sqrt{h_{xx}})$, and (c) $D_x{}^x/\sqrt{h_{xx}}$.

tity $(K_y{}^y - K_z{}^z)/2$ is within about 3 per cent of $-(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$, close to but not identical to what is expected from the left-propagating "physical" eigenmode. In Fig. 2.2b, we see that the steps in $(D_y{}^y + D_z{}^z)/(2\sqrt{h_{xx}})$ are associated with extrema of $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$. At these times, $(K_y{}^y + K_z{}^z)/2 = (D_y{}^y + D_z{}^z)/(2\sqrt{h_{xx}})$ to the left of the physical wave. In the vicinity of the physical wave, $(K_y{}^y + K_z{}^z)/2$ has step-like features associated with steps in $(D_y{}^y + D_z{}^z)/(2\sqrt{h_{xx}})$, but ascending to the right. In the region between the waves, $(K_y{}^y + K_z{}^z)/2$ is much larger than $(D_y{}^y + D_z{}^z)/(2\sqrt{h_{xx}})$ and increases with time. Fig. 2.2c shows the evolution of $D_x{}^x/\sqrt{h_{xx}}$. The prominent feature in this figure is a small residual effect (note that the scale of the graph is 10^{-4}) of the prominent feature in $K_x{}^x$ shown in Fig. 2.9 which survives the near cancellation of $K_x{}^x$ in the evolution equation of $D_x{}^x$ from our shift resetting condition (Eq. (2.25)). Since $\alpha K_x{}^x - \beta' = 0$ each time the shift is reset, this feature in $D_x{}^x/\sqrt{h_{xx}}$ tends to advect along the hypersurface normals. The generation and modification of features in $D_x{}^x/\sqrt{h_{xx}}$ is due to the different evolutions of $\alpha K_x{}^x$ and β' between gauge resettings. The feature in $K_x{}^x$ results from our lapse resetting condition, Eqs. (2.22) and (2.23), when a strong imbalance between the transverse D's and K's occurs near the center of the grid as the waves collide, creating negative values for $\mathcal{S}$. This in turn causes the limiter to take effect, which allows $K_x{}^x$ to dip in the negative direction.

To compare the overall accuracies of different formulations, we present 1-norms of the energy constraint errors in Fig. 2.3 and 1-norms of errors in $D_x{}^x/\sqrt{h_{xx}}$ in Fig. 2.4 at $t = 12$ for 500 cell grids. The constraint errors are predominantly errors in the derivatives of the constraint quantities, and are insensitive to errors in the longitudinal variables. For each scheme, the 1-norm errors are plotted for a number of values of the energy constraint coefficient, ranging from -0.25 to 0.95 at 0.05 increments. Since our ADM scheme is not hyperbolic for $n = 0$, the transport steps of the ADM calculations are solved

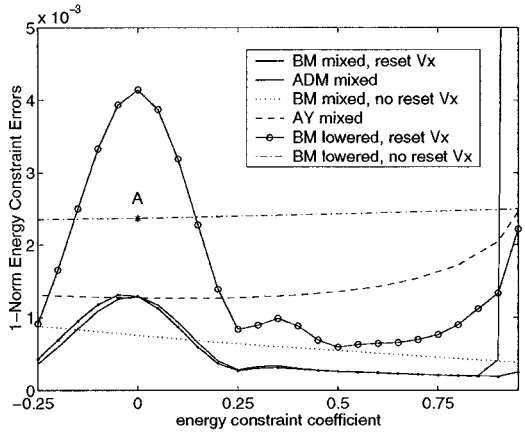


Figure 2.3: 1-norm errors of the energy constraint plotted against the energy constraint coefficient, n, for several different formulations of the Einstein equations. Evaluated at $t = 12$ with a grid resolution of 500 cells. Note that the largest value of n plotted is 0.95. Point A is the formulation closest to the standard BM scheme [17].

with the finite difference numerical method, whereas the transport steps of the BM and AY calculations use our flux-based wave decomposition methods. The choice of numerical method makes little difference to the results.

Fig. 2.3 identifies factors which affect accuracy as measured by 1-norm energy constraint errors. It is apparent that using mixed variables improves accuracy significantly in the BM formulation. Similar improvements occur in the AY and ADM formulations. The 1-norm energy constraint errors for ADM and the ADM-like reset BM schemes are almost identical, differing by only 1 to 2 per cent over the range $-0.25 \leq n \leq 0.85$, and are minimized for $0.25 \leq n \leq 0.80$. Point "A" on Fig. 2.3 is the formulation closest to the BM scheme as implemented in reference [17]. The identical formulation using mixed instead of lowered variables decreases the 1-norm energy constraint error 3.1 times. If the mixed BM system of equations is transformed into an ADM-like scheme by frequently resetting V_x, and an energy constraint coefficient of 0.5 is used,

a 9.3-fold decrease in the 1-norm energy constraint error compared to point "A" is obtained. Both the ADM and the reset BM error curves peak at $n = 0$, and increase rapidly as $n \to 1$, though the increase as $n \to 1$ for mixed reset BM occurs too close to $n = 1$ to be apparent in Fig. 2.3. The rise in energy constraint errors at $n = 0$ and $n = 1$ reflects in part the breakdown of hyperbolicity in ADM at these values of n. The effects of this breakdown are more severe at $n = 1$ than at $n = 0$ because ADM is unstable for $n > 1$. Momentum constraint errors are similar to or smaller than the energy constraint errors.

The 1-norm energy constraint errors in the no-reset BM schemes vary slowly for all n. These schemes are well-behaved for $n \geq 1$, and the errors for the mixed version continue to decrease. Despite the breakdown in the AY scheme at $n = 1$, the constraint errors do not increase strongly until n gets close to 1. The AY formulation is well-behaved for $n > 1$.

Since the true value of ${D_x}^x/\sqrt{h_{xx}}$ is not known exactly, we must extrapolate to estimate the true value and calculate the 1-norm errors shown in Fig. 2.4. Assuming quadratic convergence, the error estimate at each grid cell of a 500 cell calculation is $\frac{4}{3}$ times the difference between the 500 and 1000 cell results. For ${D_x}^x/\sqrt{h_{xx}}$ at time $t = 12$, the 500 cell errors deviate from quadratic scaling in the region $8 < x < 10$, where the feature in ${D_x}^x/\sqrt{h_{xx}}$ associated with the spike in ${K_x}^x$ is located. Here, our standard extrapolation underestimates the errors for the ADM and no-reset BM formulations, and overestimates the errors for the reset BM formulations, but the effects on Fig. 2.4 are not very significant.

Fig. 2.4 shows that the same factors which decrease the 1-norm energy constraint errors also decrease the 1-norm errors in ${D_x}^x/\sqrt{h_{xx}}$. Specifically, using mixed variables instead of lowered variables in the no-reset BM formulation at $n = 0$ decreases the 1-norm error in ${D_x}^x/\sqrt{h_{xx}}$ 1.5 times. The shapes of the curves fall into the same two classes as described above for Fig. 2.3. If one

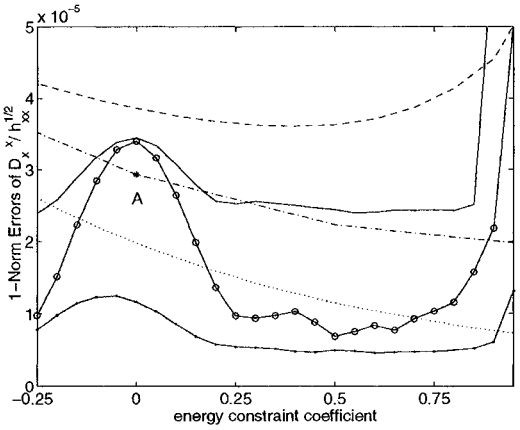


Figure 2.4: 1-norm errors of $D_x{}^x/\sqrt{h_{xx}}$ plotted against the energy constraint coefficient, n, for several different formulations of the Einstein equations. Evaluated at $t = 12$ with a grid resolution of 500 cells. Errors are estimated from comparisons with 1000 cell calculations, assuming quadratic convergence. Note that the largest value of n plotted is 0.95. The legend is the same as in Fig. 2.3 *except* that the AY mixed values are multiplied by 0.045.

frequently resets V_x and sets $n = 0.5$ in mixed BM, the 1-norm error decreases 6-fold from point "A". The errors in $D_x{}^x/\sqrt{h_{xx}}$ in the ADM and reset BM schemes peak at $n = 0$ and increase rapidly as $n \to 1$, again reflecting, in part, the failure of hyperbolicity at these values of n.

The fact that Fig. 2.4 is at all similar to Fig. 2.3 is because the $D_x{}^x/\sqrt{h_{xx}}$ errors and the constraint errors behave similarly in the region of the physical wave (this is described in detail in Sec. 2.10.2). However, there are several differences between these two figures. First, the ADM curve in Fig. 2.4 is well above the reset BM curve. Second, the reset BM curves have smaller drop-offs from $n = 0$ in Fig. 2.4. Third, the decreasing slopes of the no-reset BM curves are bigger in Fig. 2.4 than in Fig. 2.3. These differences are due to relatively large formulation-dependent numerical errors in $D_x{}^x/\sqrt{h_{xx}}$ in the region $8 < x < 10$, which contribute to the 1-norms. Recall that the evolution

of the feature in ${D_x}^x/\sqrt{h_{xx}}$ in this region depends mainly on the evolutions of ${K_x}^x$ and β' between gauge resettings. We expect the numerical errors for ${K_x}^x$ to differ from those for β', since these two variables evolve by quite different equations. Further, we expect these numerical errors to be larger where ${K_x}^x$ (and β') vary rapidly (see Fig. 2.9). Small differences in these numerical errors among the different formulations result in large differences in numerical errors for ${D_x}^x/\sqrt{h_{xx}}$ in this region. For example, the errors in ${D_x}^x/\sqrt{h_{xx}}$ for $8 < x < 10$ are larger at $n = 0$ than at $n = 0.5$ for the mixed no-reset BM scheme, explaining the decreasing slopes in Fig. 2.4, and are larger for ADM than for reset BM, explaining the displacement between the mixed ADM and mixed reset BM curves.

Another difference between Figs. 2.3 and 2.4 is that the 1-norm errors of ${D_x}^x/\sqrt{h_{xx}}$ are one to two orders of magnitude higher for AY than for the other formulations, whereas the 1-norm energy constraint errors for AY and the other formulations are comparable. For example, the 1-norm error in ${D_x}^x/\sqrt{h_{xx}}$ at $n = 0.5$ is about 70 times higher for the mixed AY formulation than for the mixed no-reset BM formulation, whereas the 1-norm AY energy constraint error is about 2.5 times higher. The subtraction of two numbers with large errors in Eq. (2.39) gives a large error for ${D_x}^x/\sqrt{h_{xx}}$ in the AY formulation. When we introduce substantial variations in h_{xx} in the initial conditions, so that ${D_x}^x$ is much larger than $({D_y}^y + {D_z}^z)$, and h_{xx} is substantially different from 1, then equally large errors are introduced into all the formulations. In our particular calculations, the errors already present in the AY scheme fortuitously cancel the introduced errors, resulting in ${D_x}^x/\sqrt{h_{xx}}$ errors for mixed AY and mixed no-reset BM which are within a factor of 2.

These results demonstrate our ability to significantly increase the accuracy of 1D highly nonlinear colliding gravitational plane wave calculations through equation formulation. In particular, the choice of a mixed set of in-

dices improves the accuracy of the formulations for all values of the energy constraint coefficient tested. Resetting V_x in the BM formulations creates ADM-like schemes. This is an advantage for $0.25 \leq n \leq 0.80$, where ADM is hyperbolic, and a disadvantage for $n = 0$ or $n = 1$, where ADM is not hyperbolic.

Error Propagation

In order to understand the variations in accuracy among the formulations, it is instructive to look at how errors vary with position, and how they propagate over time. We focus on the energy constraint errors and the errors in $D_x{}^x/\sqrt{h_{xx}}$, since they are representative of errors in the transverse and longitudinal parts of the metric, respectively. Momentum constraint errors are comparable to or less than the energy constraint errors.

Since the principal parts of the energy and momentum constraints are the derivatives of the constraint quantities, the constraint errors propagate with the same speeds as the "constraint" eigenmodes (which are given in Sec. 2.7.4 for BM, AY, and ADM). The constraint error propagation can also be obtained from the evolution equations for the energy and momentum constraints, which form their own closed hyperbolic system. The constraint errors propagate along the light cones for the no-reset BM and AY formulations. For ADM, the constraint errors propagate at

$$v_{ADM} = -\beta \pm \frac{\alpha}{\sqrt{h_{xx}}}\sqrt{1-n}. \tag{2.65}$$

We expect the constraint errors to also propagate at v_{ADM} for the reset BM scheme. Note that for $n < 0$ or $0 < n < 1$, v_{ADM} is different from any of the other characteristic speeds of the system. We find that a separation of the constraint error speeds from the other characteristic speeds improves accuracy.

Fig. 2.5a shows the energy constraint error propagation for reset BM and $n = 0$. First, notice that the constraint errors are large where the physical

wave is present, and that at a given location, the error decreases almost to zero when the physical wave has passed. However, there is a rapid increase in the energy constraint errors propagating with the physical wave. Second, from the graph we see that the waveform of the energy constraint errors propagates at roughly unit coordinate speed, which, over the times we are considering, is approximately coordinate light speed. The energy constraint errors are predominantly errors in the derivatives of $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$. Errors in the propagation of $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$ are largest where its second derivative is largest, at the corners of the steps visible in Fig. 2.2b. From the energy constraint equation, the steps in $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$ are associated with large values of the physical quantities $(D_y{}^y - D_z{}^z)/\sqrt{h_{xx}}$ and $(K_y{}^y - K_z{}^z)$. These physical quantities propagate at light speed, and constraint errors, once generated, propagate with the velocity of the "constraint" eigenmodes, which is also light speed for reset BM with $n = 0$. Since new errors remain in phase with the propagating old errors, the constraint errors are continuously reinforced. Careful comparison of Fig. 2.2b with Fig. 2.5a shows that the constraint error peaks are coincident with the corners of the steps in $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$ at all three times shown.

The same argument applies to AY and no-reset BM, since these formulations also have "constraint" mode errors propagating at light speed, but Fig. 2.3 shows larger errors for ADM and reset BM at $n = 0$ than for AY and no-reset BM. We attribute the larger ADM and reset BM errors to the breakdown of hyperbolicity in ADM at $n = 0$, so that the "constraint" eigenmodes, which propagate at light speed with constant amplitude, interact with the "longitudinal" eigenmodes through the constraint quantities. As $n \to 0$, any errors in the constraint quantities tend to produce amplified errors in the longitudinal variables, because $n(D_x{}^x + D_Q)$ and $n(K_x{}^x - \beta'/\alpha)$ are the same order of magnitude as the constraint quantities in the ADM "longitudinal" eigenmodes. The errors in $D_x{}^x$ and $K_x{}^x$ then feed back to the constraint quantities through the source

terms of the evolution equation for $(K_y{}^y + K_z{}^z)$.

Fig. 2.5b shows a dramatic decrease in both the energy constraint errors and the growth rate of the errors in the mixed reset BM formulation when $n = 0.5$. Furthermore, the full waveform of the energy constraint errors does not maintain its shape as it propagates, as does the waveform in Fig. 2.5a. By tracking the rightmost bump in this figure, one can determine the speed of the errors to be approximately 0.7 times light speed, which agrees with the value predicted by Eq. (2.65). Because the energy constraint errors lag behind the source of the errors, namely, the steps in $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$, there is not constant reinforcement and rapid growth of the errors in the region of the physical wave. This results in greater overall accuracy for reset BM than for no-reset BM at $n = 0.5$, as seen in Fig. 2.3.

Fig. 2.5c shows the energy constraint error propagation for mixed no-reset BM at $n = 0.5$. In contrast to Fig. 2.5b, the waveform is maintained reasonably well, and the errors grow more rapidly in time. This is because the constraint errors travel at light speed; so, as in Fig. 2.5a, there is a continuous reinforcement of the errors. For no-reset BM at $n = 0$, the curve is practically the same as what we show here at $n = 0.5$. The error waveform has a smaller growth rate than that for reset BM at $n = 0$, presumably because of the hyperbolicity of the no-reset formulation, as discussed earlier. We have also looked at the AY constraint error propagation and confirm that errors propagate at light speed for all values of n. Because the constraint errors travel with the physical waves for all n in these formulations, their 1-norm errors vary slowly with n in Fig. 2.3.

The errors in $D_x{}^x/\sqrt{h_{xx}}$ versus x at $t = 12$ for the reset BM formulations are shown in Fig. 2.6, for energy constraint coefficients of (a) 0 and (b) 0.5. Fig. 2.6a shows a localization of the errors in the region of the physical wave $(0 \leq x \leq 6)$. Fig. 2.6b shows the dramatic decrease in the errors for $0 \leq x \leq 6$

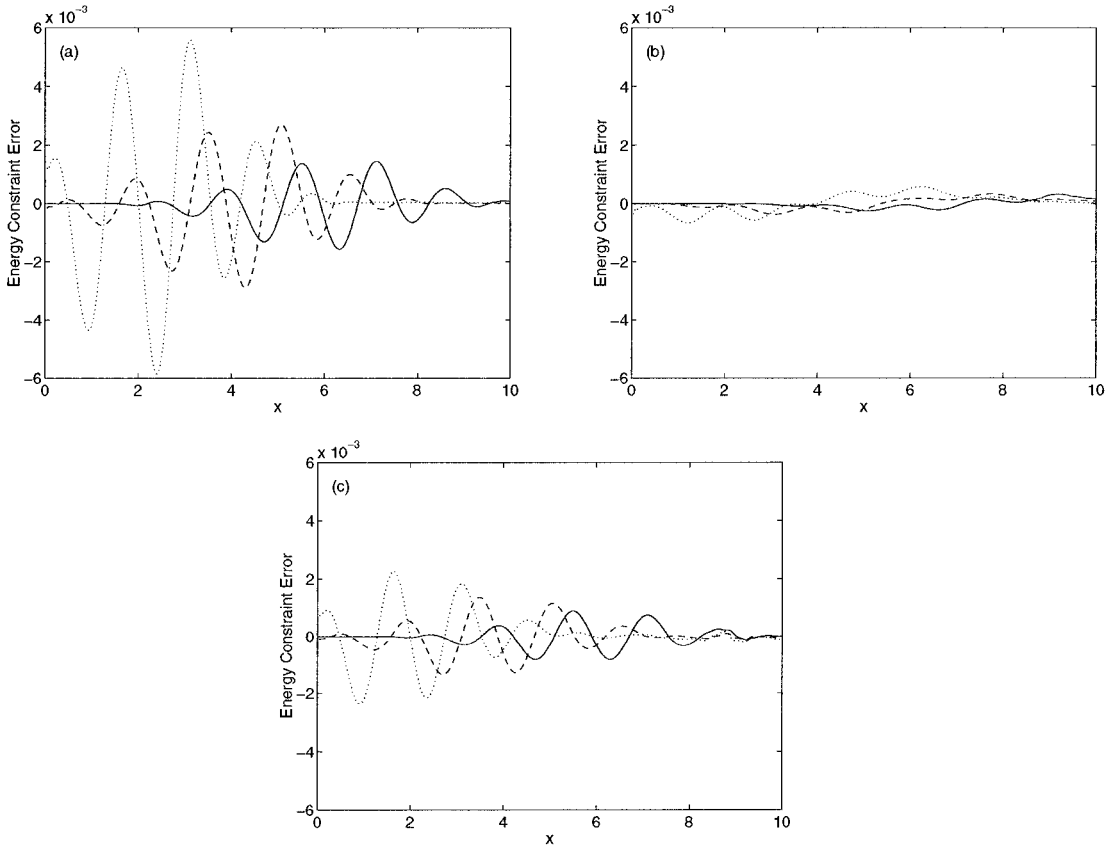


Figure 2.5: Energy constraint error propagation using our mixed BM formulation. The solid line is at $t = 8$, the dashed line at $t = 10$, and the dotted line at $t = 12$. Evaluated with a grid resolution of 500 cells, for (a) $n = 0$, with V_x resetting, (b) $n = 0.5$, with V_x resetting, (c) $n = 0.5$, with *no* V_x resetting.

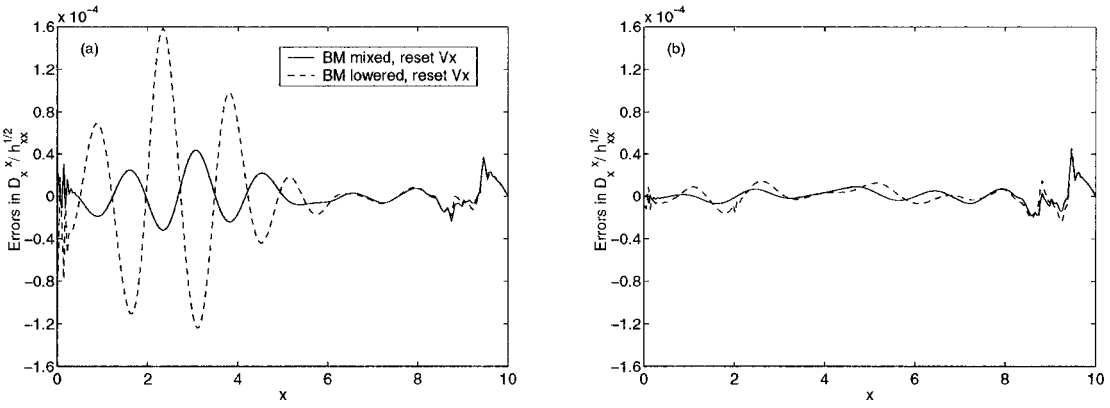


Figure 2.6: Errors in ${D_x}^x/\sqrt{h_{xx}}$ versus x for our BM mixed and lowered formulations, with V_x resetting. Evaluated at $t = 12$ with a grid resolution of 500 cells. Errors are estimated from comparisons with 1000 cell calculations, assuming quadratic convergence, for (a) $n = 0$, and (b) $n = 0.5$.

when $n = 0.5$. The mixed no-reset errors in this region at $n = 0.5$ are larger by a factor of about 2 than the mixed reset errors at $n = 0.5$, and only slightly smaller than the mixed no-reset errors at $n = 0$. The fact that there is a similar reduction in constraint errors and ${D_x}^x/\sqrt{h_{xx}}$ errors when going from no-reset BM to reset BM at $n = 0.5$ suggests that the constraint errors are a major source of errors for ${D_x}^x/\sqrt{h_{xx}}$. From the failure of hyperbolicity in ADM at $n = 0$, one expects the errors in ${D_x}^x/\sqrt{h_{xx}}$ to increase more rapidly for reset BM than the constraint errors, but we do not see clear evidence for this. The increase in errors is about the same going from $t = 8$ to $t = 12$, perhaps because the ${D_x}^x/\sqrt{h_{xx}}$ errors have not yet reached their asymptotic limit. The spikey errors at approximately $8.7 \leq x \leq 10$ in both Figs. 2.6a and b are discussed in Sec. 2.10.2.

The ADM and ADM-like reset BM systems show rapid increases in errors as $n \to 1$ in Figs. 2.3 and 2.4. The breakdown of hyperbolicity at $n = 1$ results in errors in $({D_y}^y + {D_z}^z)/\sqrt{h_{xx}}$ becoming large compared to errors in

$(K_y{}^y + K_z{}^z)$. The contribution to the energy constraint errors from errors in the derivatives of $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$ increases correspondingly. Since the "constraint" eigenmodes propagate along the hypersurface normals at $n = 1$, the errors in $(D_y{}^y + D_z{}^z)/\sqrt{h_{xx}}$ also reinforce errors in variables which propagate along the hypersurface normals, namely, $D_x{}^x/\sqrt{h_{xx}}$, β', and D_Q.

Fig. 2.7 shows the errors in $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$ as a function of x for three formulations at an n of 0.5. The amplitudes and shapes of the errors for the different formulations are roughly the same, because the evolution equations for $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$ are similar for the different formalisms. The curve shapes do not change significantly when one uses $n = 0$ instead of $n = 0.5$ because n does not enter into the evolution equations for $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$. The entire graph shown in Fig. 2.7 converges quadratically except for the bump around $x = 6$, which converges linearly. This bump is near the trailing edge of the physical wave, where the initial conditions are not smooth enough to give second order accuracy. The errors for the formulations which are not shown are similar.

Calculations of the distribution and propagation of energy constraint errors and errors in $D_x{}^x/\sqrt{h_{xx}}$ have illuminated how resetting V_x in the BM formulations to create overall ADM-like evolutions increases or decreases accuracy and stability, depending on the value of n. Intermediate values of n separate constraint error speeds from the other characteristic speeds of the system, resulting in significant improvements in accuracy. Values of n for which ADM hyperbolicity fails result in rapid increases in errors.

The higher accuracy of reset BM compared to the other formulations when $0.25 \leq n \leq 0.80$ is seen when $h_{xx} = 1$ and $D_x{}^x = 0$ in our initial conditions. When we introduce substantial variations in h_{xx} in the initial conditions, so that $D_x{}^x$ is at least as large as $(D_y{}^y + D_z{}^z)$, and $\sqrt{h_{xx}}$ varies by a factor of 2 to 3, the accuracy results are dominated by the errors directly associated with the

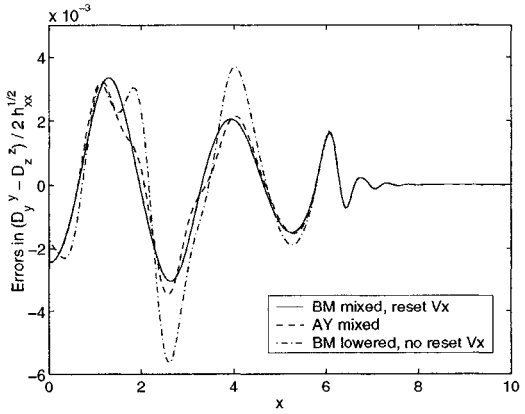


Figure 2.7: Errors in $(D_y{}^y - D_z{}^z)/(2\sqrt{h_{xx}})$ versus x for different formulations. Evaluated at $t = 12$ with a grid resolution of 500 cells. Errors are estimated from comparisons with 1000 cell calculations, assuming quadratic convergence, for $n = 0.5$.

variations in h_{xx}. These errors depend primarily on the amplitude of $D_x{}^x/\sqrt{h_{xx}}$, and are largest where the derivative of $D_x{}^x/\sqrt{h_{xx}}$ is largest. They are similar in all the formulations and are independent of n; thus, they tend to equalize the results from the different formulations.

2.10.3 Gauge Conditions

Our gauge conditions involve both a periodic resetting of the lapse and the shift, and an evolution of the lapse and the shift between resettings. The lapse and shift are reset according to Eqs. (2.22) and (2.25), with $\Gamma = 1$, and $(K_x{}^x)_T = -0.02$. They are reset at constant time intervals rather than after a specific number of time steps, to give resolution-independent results. Between resettings, β, β', Q, and D_Q are advected along hypersurface normals.

Fig. 2.8 shows the behavior of the lapse and the shift at the left boundary of the grid with time. The values of $\ln(\alpha)$ and β are fixed at zero at $x = 10$, the grid center. $\ln(\alpha)$ and β vary approximately monotonically from the grid center

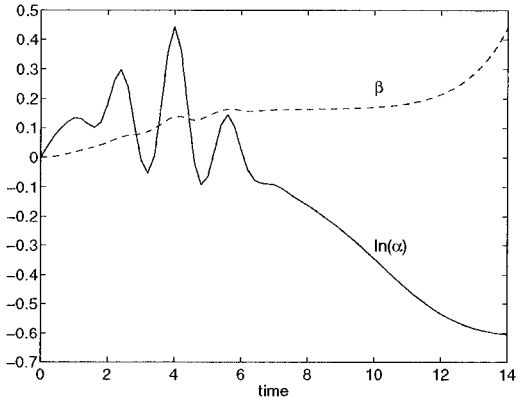


Figure 2.8: The shift, β, and the logarithm of the lapse, $\ln(\alpha)$, at the left grid edge plotted against time at intervals $\triangle t = 0.2$. The values of β and $\ln(\alpha)$ are fixed at zero at the grid center.

to the edges; therefore, we illustrate how they behave as a function of time by graphing their values at the edge, where typically $\ln(\alpha)$ and β are largest in magnitude.

In Fig. 2.8, the lapse oscillates at the grid edge until about $t = 6$. From $t = 6$ to $t = 12$, $\ln(\alpha)$ becomes increasingly negative until it starts leveling off around $t = 12$. The lapse is determined by Eq. (2.22). Since we set the lapse to one and its first derivative to zero at the center of the grid, whether the lapse is greater than or less than one at the edge is determined by the sign of the second derivative of the lapse with respect to proper distance, the quantity $\mathcal{S}$ in Eq. (2.22). The large terms in $\mathcal{S}$ are typically $-K_y{}^y K_z{}^z$ and $D_y{}^y D_z{}^z / h_{xx}$. These nearly cancel when $D_y{}^y$, $D_z{}^z$, $K_y{}^y$, and $K_z{}^z$ are dominated by a single propagating "physical" wave pulse. As two such pulses collide, constructive interference in $D_y{}^y$ and $D_z{}^z$ implies destructive interference in $K_y{}^y$ and $K_z{}^z$, and vice versa, which causes rather large oscillations in the value of the lapse at the edge of the grid. Note that at $t = 4$, there is maximum constructive interference of $D_y{}^y$ and $D_z{}^z$. Once the waves have largely passed through each

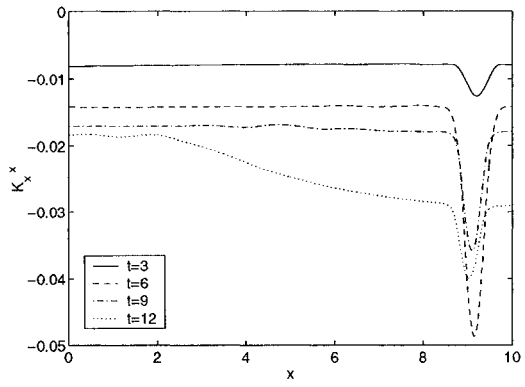


Figure 2.9: ${K_x}^x$ versus x at the times indicated in the legend. ${K_x}^x$ is driven to a target value of -0.02 by our lapse resetting condition. The spike at $x \approx 9$ develops as the waves collide, due to the limiter in the equation for the lapse. The effects of the limiter are felt over a wider range of x at $t = 12$.

other ($t > 6$), ${K_y}^y$ and ${K_z}^z$ are large, and ${D_y}^y$ and ${D_z}^z$ are small in the region between the separating waves, so the second derivative of the lapse is strongly negative, and becomes more strongly negative as ${K_y}^y$ and ${K_z}^z$ become steadily larger. By $t = 12$, the limiter in Eq. (2.22), which is necessary to prevent the lapse from becoming negative at the edge of the grid, largely stops further decrease of the lapse at the edge.

The behavior of the shift in Fig. 2.8 reflects the behavior of ${K_x}^x$ in Fig. 2.9, since β' is proportional to ${K_x}^x$ by Eq. (2.25). The resetting of the lapse keeps ${K_x}^x$ near its target value of -0.02 until $t > 10$, when the limiter in Eq. (2.22) takes hold, allowing ${K_x}^x$ to become steadily more negative. Since the derivative of the shift is proportional to ${K_x}^x$, and the shift is clamped to zero at the center of the grid, the shift at the left edge of the grid is positive and starts getting steadily larger for $t > 10$.

Recall that a negative ${K_x}^x$, in combination with our shift resetting condition, causes the hypersurface normals to point away from the computational

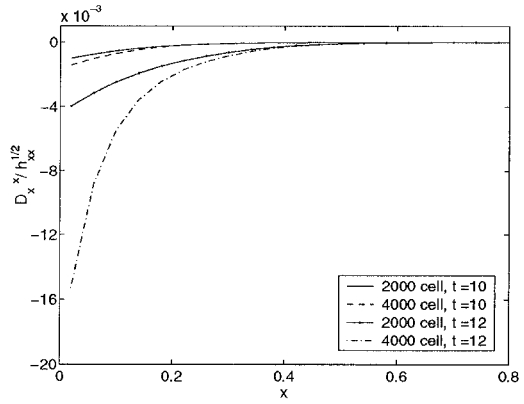


Figure 2.10: Development of an instability at the left boundary in the evolution of ${D_x}^x/\sqrt{h_{xx}}$ when the ${K_x}^x$ target value is $+0.02$ in our lapse resetting condition. The instability is substantially stronger with the higher resolution grid. Note that the physical wave reaches the left edge at $t = 10$.

grid at the boundaries (see Sec. 2.6). These resetting gauge conditions are designed so that Q and β are advected out of the grid, in order to suppress the development of instabilities associated with extrapolation at the grid edge. As long as $({K_x}^x)_T \leq 0$, the shift is positive at the left edge of the grid and the advection velocity is negative. However, Fig. 2.10 shows that when $({K_x}^x)_T = +0.02$, so that the shift is negative at the left edge, the solution becomes unstable at the grid edge once the physical wave reaches the edge.

2.10.4 Boundary Conditions

Fig. 2.11 compares two methods for implementing boundary conditions: quadratic extrapolation of all the variables with and without correcting the extrapolated values of $({D_y}^y + {D_z}^z)/\sqrt{h_{xx}}$ and $({K_y}^y + {K_z}^z)$ to insure that the energy and momentum constraint equations are satisfied at the boundaries. Failure to strictly enforce the constraints at the edges of the grid results in significant constraint errors propagating over much of the grid by $t = 8$.

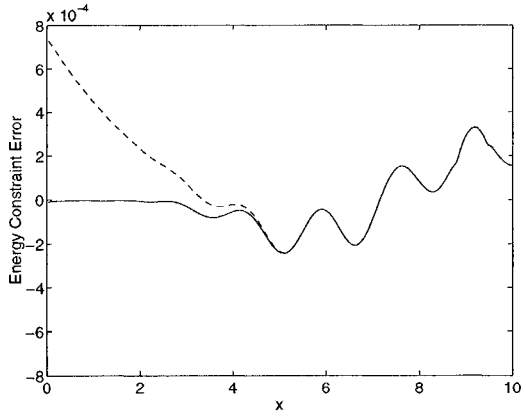


Figure 2.11: Energy constraint error for our standard boundary condition, which corrects quadratic extrapolation of the variables at the boundaries using the constraint equations (solid line), versus quadratic extrapolation only (dashed line). Evaluated at $t = 8$ with a grid resolution of 500 cells using our mixed BM formulation with V_x resetting and $n = 0.5$.

The jitter at the left edge of Fig. 2.6, which graphs the errors in $D_x{}^x/\sqrt{h_{xx}}$ versus x, is due to numerical errors and the use of quadratic extrapolation in our boundary conditions. It is most prominent when the physical wave is crossing the boundary, but it is not unstable as long as the hypersurface normal at the boundary does not point into the grid (see Sec. 2.6). We will explore ways of reducing this jitter either through the use of limiters near the boundary and/or improved extrapolation methods.

2.10.5 Numerical Methods

We compare convergence results from our two flux-based wave decomposition schemes with those from the traditional finite difference approach. Tables 2.1 and 2.2 show quadratic convergence for errors in the metric derivative variables and in the momentum and energy constraints at $t = 12$ using our optimal evolution scheme and flux-based wave decomposition method **I**. Quadratic con-

Table 2.1: Convergence results for errors in the metric derivative variables. Evaluated at $t = 12$ using our mixed BM formalism, resetting V_x, with $n = 0.5$, and flux-based wave decomposition method **I**. We take the 1-norms of the differences of the results obtained using the indicated grid resolutions.

Grid resolutions	$\frac{1}{2}\left[\frac{(D_y{}^y - D_z{}^z)}{\sqrt{h_{xx}}}\right]$	$\frac{1}{2}\left[\frac{(D_y{}^y + D_z{}^z)}{\sqrt{h_{xx}}}\right]$	$\frac{D_x{}^x}{\sqrt{h_{xx}}}$
$(500-1000)/(1000-2000)$	3.98	3.99	5.09
$(1000-2000)/(2000-4000)$	3.97	3.99	4.38

Table 2.2: Convergence results for the energy and momentum constraint errors. Evaluated at $t = 12$ using our mixed BM formalism, resetting V_x, with $n = 0.5$, and flux-based wave decomposition method **I**. We take the 1-norms of the results obtained using the indicated grid resolutions.

Grid resolutions	Energy constraint	Momemtum constraint
500/1000	4.02	3.97
1000/2000	4.00	3.98
2000/4000	3.99	3.99

vergence for the metric derivative variables is tested by comparing 1-norms of differences between results evaluated at 500 and 1000 cell resolutions, 1000 and 2000 cell resolutions, and 2000 and 4000 cell resolutions. Convergence is calculated in this way because true values are unavailable for the metric derivative variables. Constraint errors can be calculated directly so quadratic convergence is determined simply by taking the ratios of results from grid resolutions that differ by a factor of two. There is very little difference in the results of the three numerical methods tested. Further, the 1-norm results from the three numerical methods all converge to the same answer.

Tables 2.1 and 2.2 show that our flux-based wave decomposition methods are second order convergent when the characteristic matrix $\mathbf{A}(\mathbf{x})$ has significant spatial variation, because of the spatial dependence of the lapse and the shift (see Eq. (2.47) and Fig. 2.8). However, in these calculations the eigenvectors of $\mathbf{A}(\mathbf{x})$, which depend only on $\sqrt{h_{xx}}$, are roughly constant. We have tested convergence of both our flux-based wave decomposition methods when the eigenvectors of $\mathbf{A}(\mathbf{x})$ vary rapidly, by introducing a 3-fold variation in $\sqrt{h_{xx}}$. Our results are still second order convergent.

2.11 Discussion

2.11.1 Summary

We have identified ways of improving the accuracy and stability of 1D nonlinear colliding gravitational plane wave calculations through an in-depth study of equation formulations, dynamic gauge conditions, boundary conditions, and numerical methods. Three issues stand out in our study of equation formulations. The first issue which improves the accuracy of all the formulations tested is raising an index in the metric derivative and extrinsic curvature variables. The separation of constraint and physical behavior in our calculations

is much simpler with mixed variables than with lowered variables. Since the dominant feature of plane wave solutions is the physical wave, and since the mixed variables have a simple linear relation to the "physical" eigenmodes, the mixed form gives an advantage in accuracy. In addition, the variables in the source terms of the evolution equations are in mixed form, so a cleaner system of equations results.

Second, it is advantageous if the "constraint" eigenmodes, which are constraint violating, propagate at different speeds from other features in the solution, so that nothing in the actual solution is constantly in step with the constraint errors. Our gauge conditions result in features which propagate on the hypersurface normals as well as on the light cones. With a multiple between 0.25 and 0.80 of the energy constraint equation added to the evolution equations for the extrinsic curvature in the ADM and reset BM schemes, the "constraint" eigenmodes propagate neither on the light cones, nor on the hypersurface normals, and the growth rates of constraint errors and errors in ${D_x}^x/\sqrt{h_{xx}}$ significantly decrease. However, when the amplitude of ${D_x}^x$ is large compared to $({D_y}^y + {D_z}^z)$, formulation-independent errors associated with the derivatives of ${D_x}^x/\sqrt{h_{xx}}$ dominate the overall errors for $0.25 \leq n \leq 0.80$.

Third, we find that when hyperbolicity fails in ADM, for values of the energy constraint coefficient equal to 0 or 1, the constraint errors and errors in ${D_x}^x/\sqrt{h_{xx}}$ increase much more quickly in both ADM and reset BM than they do in no-reset BM. At $n = 0$, the "constraint" and "longitudinal" eigenmodes are no longer independent, causing a rapid growth of errors which travel with the physical waves. These errors decrease and stabilize, however, when the waves exit the grid. At $n = 1$, the two "constraint" eigenmodes are not independent, causing errors in $({D_y}^y + {D_z}^z)/\sqrt{h_{xx}}$ to increase rapidly. This results in large errors in the energy constraint and in variables which propagate along the hypersurface normals.

The key to our approach to dynamic gauge conditions for hyperbolic calculations is to maintain a simple diagonalizable hyperbolic evolution during each time step, while allowing periodic, flexible resetting of the lapse and the shift between time steps. By resetting the lapse and the shift, we can control the coordinate system in a dynamic and (in principle) arbitrary way, unconstrained by the need to maintain hyperbolicity. Our gauge resetting conditions control the longitudinal components of the extrinsic curvature and the spatial metric so as to prevent pathologies and strong gradients from developing in the hypersurfaces and spatial coordinates. Further, our gauge resetting conditions cause the hypersurface normals to point away from the grid at the edges. This helps to suppress the development of instabilities at the boundaries which are associated with features advecting along the hypersurface normals.

A careful study of boundary conditions has led us to the conclusion that it is incorrect to impose outgoing wave boundary conditions based on the eigenmodes of the hyperbolic decomposition. A substantial contribution from the incoming "constraint" eigenmodes is necessary to satisfy the constraint equations at the boundaries. In the presence of nonlinearities, the interaction of the "physical" eigenmodes with the incoming "constraint" eigenmodes generates an admixture of incoming and outgoing "physical" eigenmodes. The right and left-going gravitational wave amplitudes can be determined by projecting the Weyl tensor onto a null tetrad, and our numerical results indicate that the gravitational waves after the collision are purely outgoing according to this definition. Our boundary condition procedure consists of an accurate calculation of the incoming eigenmodes by correcting the ghost cell values using the energy and momentum constraint equations, after quadratically extrapolating all the variables to the ghost cells. This procedure, in combination with our lapse and shift resetting conditions which insure that the hypersurface normals at the boundaries do not point into the grid, has given stable, accurate, and second

order convergent results.

Finally, we have developed flux splitting numerical methods for solving hyperbolic formulations of the Einstein equations which are second order accurate for smooth solutions, even when the eigenvalues and eigenvectors of the characteristic matrix are spatially varying. These methods are based on decomposing flux differences between adjacent grid cells into linear combinations of the eigenvectors of the characteristic matrix. We show that these methods are formally second order accurate and, in practice, second order convergent.

2.11.2 Relevance of Results and Future Directions

Our results suggest that hyperbolicity can be a useful guide to picking equation formulations for numerical integration of the Einstein equations. When eigenvectors are incomplete, or some eigenvectors are nearly linearly dependent, some solutions to the equations will tend to grow without bound or by large factors. In 2D and 3D, the ADM and ADM-like equation formulations for non-diagonal metrics fail to have a complete set of characteristic eigenvectors for *any* value of the energy constraint coefficient. This may be at the root of some of the instabilities seen in 3D ADM codes.

However, hyperbolicity, which involves only the principal terms in the equations, is not the whole story. The constraint equations relate derivatives of some quantities (the constraint quantities) to nonlinear terms involving additional quantities, so the actual evolution of the constraint quantities for constraint-satisfying solutions may be very different from the evolution implied by the "constraint" eigenmodes, which satisfy equations without source terms. It seems to be advantageous if the speeds of the short-wavelength errors in the constraint quantities, which are the "constraint" mode eigenvalues, differ substantially from the propagation speeds of the major features in the constraint quantities. In 1D, both the ADM and the no-reset BM formulations are safely hyper-

bolic for $n \approx 0.5$, but the former is substantially more accurate because no-reset BM has the same speeds for the "constraint" eigenmodes and the features in the constraint quantities, whereas ADM has different speeds. How much of an advantage it is to have different speeds depends on the dominant source of errors. In a gauge where h_{xx} is close to 1, the dominant numerical errors in the constraint quantities are associated with features in the same quantities, but in a gauge where the dominant numerical errors are associated with features in h_{xx}, it does not make much difference whether the "constraint" eigenmodes and features in the constraint quantities have the same velocities. In generic black hole spacetimes, it is probably not possible to find a gauge where the physical waves dominate the metric; therefore, the separation of speeds may not make much difference in accuracy for 2D and 3D calculations of greatest physical interest.

The eigenmode decomposition in higher dimensions depends on a chosen direction of propagation, as do the combinations of the actual variables which can be identified as "longitudinal", "constraint", and "physical" quantities. In addition, putting the equations into first-order form requires a choice of derivative ordering, which has important effects on the hyperbolicity of the system. A "directional splitting" approach to solving the transport steps, in which $\partial_t \mathbf{q} + \partial_k \mathbf{F}^k = 0$ is solved separately for each coordinate direction k, does, at least for some choices of derivative ordering, allow the identification of "longitudinal", "constraint", and "physical" eigenmodes for one coordinate direction at a time. The decomposition involves projecting the D_{kij}'s and K_{ij}'s perpendicular to and into the constant-x^k surfaces. The eigenvectors can be constructed explicitly.

Using the mixed coordinate components $D_{ki}{}^j$ and $K_i{}^j$ as variables when the metric is not diagonal gives complicated source terms in the $D_{ki}{}^j$ evolution equations, because the $D_{ki}{}^j$ are no longer pure derivatives. Another possibility is to take as the variables the components of an orthonormal triad and the

projection of the extrinsic curvature onto the triad (or the Ashtekar variables [69, 68, 83]). The triad formalism, because the extrinsic curvature tensor projected on the triad is symmetric, has fewer variables than the mixed coordinate component formalism. In neither case would the variables be simply related to the eigenvectors of the characteristic matrix when the metric is non-diagonal or the triad vectors are not along the coordinate directions; thus, there is no obvious advantage to trying to generalize our mixed variables in 1D to generic 2D and 3D calculations. However, we plan to explore these questions further.

Our approach to dynamic gauge conditions, namely, implementing a simple hyperbolic evolution during each time step, then resetting the lapse and the shift between time steps, is general. A hyperbolic evolution during each time step allows the use of flux-based wave decomposition numerical methods. A long-term strictly hyperbolic evolution, however, is destroyed in all our hyperbolic formulations when we reset the lapse and shift. Despite this, resetting the gauge can be used to increase the accuracy and stability of our 1D calculations. Achieving the same goal in 2D or 3D calculations will be more difficult. There is no one longitudinal direction, and there is not enough gauge freedom to control the longitudinal components of the metric and extrinsic curvature for all directions. What gauge resetting conditions are most effective in 2D and 3D remains an open question.

Our boundary condition results indicate that simple outgoing wave boundary conditions based on the eigenmodes of the characteristic matrix are not valid in general and are inconsistent with the constraint equations. On the other hand, extrapolation (in particular, quadratic extrapolation) is a numerically dangerous procedure. Although one can use constraint extrapolation to constrain certain linear combinations of the variables, this technique is likely to be less effective in higher dimensions than in 1D because the number of variables increases more rapidly than the number of constraints. Using an outgo-

ing wave boundary condition based on the Weyl tensor to further constrain the variables is problematic in the general case. While the peeling theorem [58] in an asymptotically flat spacetime does show that the incoming wave projection of the Weyl tensor falls off much more rapidly than the outgoing wave projection in the wave zone, one cannot assume that the outgoing wave dominates in 3D numerical relativity calculations, which typically have to be truncated at best in the inner part of the wave zone. The ingoing part of the Weyl tensor contains non-radiative quasi-static quadrupole moment contributions which cannot be assumed to vanish.

The flux-based wave decomposition numerical methods we have presented for hyperbolic formulations of the Einstein equations are completely generalizable, and should prove useful for calculations in black hole and Brill wave spacetimes where the eigenvectors and eigenvalues have significant spatial variation. We hope to extend our exploration of hyperbolic methods to include the use of limiters and upwind differencing. Limiters will be used to suppress short wavelength numerical instabilities. With upwind differencing, one does not need ghost cells at the apparent horizon boundary of an excised black hole, where the eigenmodes are purely outgoing (relative to the computational domain).

In conclusion, we have developed basic methodologies for hyperbolic formulations of the Einstein equations, which improve accuracy and stability in 1D, and which we think merit further exploration in the context of numerical relativity calculations of more substantial physical interest.

2.12 Acknowledgements

This paper presents the results of one phase of research supported by the NASA Graduate Student Researchers Program under Grant No. NGT5-50298 in col-

laboration with the Jet Propulsion Laboratory. LTB also gratefully acknowledges partial support from NSF grant DMS-9803442. We thank R. LeVeque and F. Estabrook for discussions and for careful reading of the manuscript. Gratitude is also extended to D. Bale, J. Rossmanith, M. Alcubierre, E. Seidel, and K. Thorne for insightful discussions, and to the Numerical Relativity group in Potsdam, Germany for their warm generosity during an extended visit. Finally, JMB would like to thank the Institute for Theoretical Physics at UCSB for their support during the initial stages of this research.

Chapter 3

A HYPERBOLIC TETRAD FORMULATION OF THE EINSTEIN EQUATIONS FOR NUMERICAL RELATIVITY

3.1 Abstract

The tetrad-based equations for vacuum gravity published by Estabrook, Robinson, and Wahlquist [24] are simplified and adapted for numerical relativity. We show that the evolution equations as partial differential equations for the Ricci rotation coefficients constitute a rather simple first-order symmetrizable hyperbolic system, not only for the Nester gauge condition on the acceleration and angular velocity of the tetrad frames considered by Estabrook *et al.*, but also for the Lorentz gauge condition of van Putten and Eardley [77] and for a fixed gauge condition. We introduce a lapse function and a shift vector to allow general coordinate evolution relative to the timelike congruence defined by the tetrad vector field. This chapter has been published (see reference [18]).

3.2 Introduction

Gravitational wave detection is soon to become a reality. However, the scientific community has not yet been able to calculate gravitational waveforms from the most likely source, namely the violent and dynamic merger phase of binary black hole collisions [42]. We hope to contribute to this effort with detailed small-scale studies. Our first such study of 1D colliding gravitational plane waves [13] indicated that two important factors which improve the accu-

racy and stability of the numerical calculations are hyperbolicity of the equations and evolving variables which are related to physical quantities. As a way to generalize our results to black hole spacetimes and higher dimensions, we are investigating a tetrad approach based on the formalism published by Estabrook, Robinson, and Wahlquist (ERW) [24]. In this paper, we present a modified version of the formalism, adapted for numerical relativity. Subsequent papers will present our numerical results.

The standard approach to vacuum numerical relativity is a $3 + 1$ decomposition, of the type introduced by Arnowitt-Deser-Misner [8]. The $3 + 1$ formulations slice four-dimensional spacetime into three-dimensional spacelike hypersurfaces (see [67] for a recent review). They evolve the spatial metric and extrinsic curvature of the hypersurfaces, which are expressed in a coordinate basis. The evolution of the coordinates is described by a lapse function and shift vector, which may or may not be dynamic.

A tetrad formulation uses orthonormal basis vectors, $\boldsymbol{e}_\alpha$ ($\alpha = 0, 1, 2, 3$), which describe local Lorentz frames. The spacetime metric, the dot product of the basis vectors, is everywhere the Minkowski metric, $g_{\alpha\beta} = \boldsymbol{e}_\alpha \cdot \boldsymbol{e}_\beta = \eta_{\alpha\beta}$. The timelike vector field of the orthonormal frames, $\boldsymbol{e}_0$, defines a preferred timelike congruence, to which it is tangent. The spatial triad vectors in a particular rotational orientation with respect to $\boldsymbol{e}_0$ are $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3$. The dual basis of orthonormal one-forms is $\boldsymbol{e}^\alpha$ such that $\langle \boldsymbol{e}^\alpha, \boldsymbol{e}_\beta \rangle = \delta^\alpha_\beta$ and $\boldsymbol{e}^\alpha \cdot \boldsymbol{e}^\beta = \eta^{\alpha\beta}$. As Estabrook and Wahlquist point out [26], these tetrad frames are natural for measuring observable physical quantities.

Most of the recent tetrad formalisms assume that the tetrads are tied to the physically defined flow of a fluid, primarily in the context of cosmology or of interior metrics of rotating stars. Here, we advocate that a tetrad formalism may be useful even in vacuum blackhole spacetimes. The variables in a tetrad formalism are the connection coefficients (often called the Ricci rotation coef-

ficients), the tetrad vector components, and, typically, the tetrad components of the Weyl tensor or the Riemann tensor. The spatial coordinates are often assumed from the start to be comoving with the timelike congruence generated by the tetrad field. In dealing with black hole event horizons, however, it is important to allow general choices of coordinates. We do this by introducing a lapse function and a shift vector defined relative to the congruence world lines (see also van Putten and Eardley [77]).

The ERW formulation is a quasi-FOSH (first order symmetric hyperbolic) system. We use the term "quasi" because the directional derivatives along the spatial tetrad legs contain partial time derivatives (also see [22]). The basic quasi-FOSH structure of the equations involves the Weyl tensor components as variables. The quasi-FOSH system also includes equations derived from the Nester gauge conditions [56]. These equations evolve the acceleration of the congruence worldlines, $\boldsymbol{a}$, and the angular velocity of the spatial tetrad vectors relative to Fermi-Walker transport, $\boldsymbol{\omega}$. The tetrad components a_a and ω_a are gauge quantities since the spacetime orientation of the tetrad is not fixed by the spacetime geometry. Both $\boldsymbol{a}$ and $\boldsymbol{\omega}$ are orthogonal to $\boldsymbol{e}_0$, and therefore have only the three spatial tetrad components. In addition to evolution equations, the Nester gauge conditions provide equations which constrain the spatial dependence of the a_a and the ω_a at any given time.

ERW's formalism can be modified to give a particularly simple quasi-FOSH system, as suggested by Estabrook and Wahlquist (EW) [25], by eliminating the Weyl tensor components as separate variables, and adding the Nester constraint equations to the evolution equations.

In this paper, we discuss the derivation of the EW formalism in a way which is perhaps more accessible to those familiar with standard tensor analysis, as opposed to Cartan differential form analysis. What distinguishes the EW formalism is that the basic quasi-FOSH structure involves only the connection

coefficients. Most of the tetrad formulations in the literature include the Weyl tensor or the Riemann tensor to get a simple quasi-hyperbolic structure. Unless one adds the constraints to the Einstein equations in the particular way we present in this paper, the quasi-hyperbolic system involving only the connenction coefficients is quite complicated. We do not claim that the EW system is simpler or more elegant that those involving the Weyl or Riemann tensor; however, it has fewer variables and fewer constraint modes, which may be advantageous in 3D codes.

After deriving the EW formalism, we proceed to extend it into a form useful for vacuum numerical relativity. We consider two gauge conditions other than the Nester gauge which also give quasi-FOSH systems of equations. These are the Lorentz gauge used by van Putten and Eardley [77], and a fixed gauge, where a_a and ω_a are fixed functions of time and space. Additionally, we allow for a completely arbitrary relationship between the congruence and the hypersurface. Finally, we analyze the true hyperbolicity of the equations when expressed in terms of partial derivatives. The partial differential system contains evolution equations for eighteen non-gauge connection coefficients, six gauge quantities (if a dynamic gauge condition is chosen), three components of a vector describing the velocity of the congruence relative to the hypersurface, and nine components of the spatial tetrad vectors projected into a $t =$ constant hypersurface. These last twelve variables, together with our lapse function and shift vector (which we do not evolve), completely determine the sixteen tetrad vector components. The flux Jacobian of the partial differential system of equations has a complete set of eigenvectors and real eigenvalues, thus satisfying the requirement for hyperbolicity as per LeVeque [48]. Furthermore, these equations are symmetrizable hyperbolic as defined by Lindblom and Scheel [50].

Tetrad formulations for general relativity other than ERW and EW include

those by Friedrich [31], Jantzen, Carini, and Bini [37], van Elst and Uggla [74], Choquet-Bruhat and York [22], and van Putten and Eardley [77]. Except for EW, all of these systems include the Weyl or Riemann tensor components as fundamental variables. Friedrich's paper is a definitive discussion of hyperbolicity for both tetrad and $3+1$ representations. Janzten, Carini, and Bini give an extensive historical review of the tetrad and $3+1$ approaches. They provide a unified framework for the two approaches in what they call "gravitoelectromagnetism", in which, however, they only consider co-moving spatial coordinates. van Elst and Uggla's formalism applies only to non-rotating congruences and orthogonal hypersurface slicings. Also, they do not bring the equations into a quasi-FOSH form. The main emphasis of Choquet-Bruhat and York's paper is a tetrad formulation for fluids, where the congruence is aligned with the fluid flow lines, and the spatial coordinates are co-moving with the congruence. The acceleration of the congruence worldlines is given by the acceleration of the fluid, and the angular velocity variables (our ω_a's) are fixed functions. Choquet-Bruhat and York do briefly mention the vacuum case, in which their system is quasi-FOSH for arbitrary given acceleration and angular velocity. van Putten and Eardley use the Lorentz gauge condition to obtain second-order wave equations for the connection coefficients. van Putten gives a first-order form of these equations, involving the Riemann tensor as well as the connection coefficients, for his numerical implementation in 1D Gowdy wave vacuum spacetime [76]. The van Putten-Eardley formulation allows for complete freedom in the choice of spacetime foliation.

An example of using orthonormal frames in spatial hypersurfaces is the Ashtekar formulation (with subsequent modifications) [9, 10, 36, 81, 82]. Ashtekar follows a $3+1$ split in the sense that his variables are defined relative to a hypersurface. However, his description of the geometry of the hypersurface is in terms of orthonormal triads instead of the metric. He uses complex variables

to give a compact formalism, requiring the use of reality constraints to recover real spacetime. As subsequently modified by Yoneda and Shinkai, Ashtekar's formulation becomes a FOSH system of PDE's. Shinkai and Yoneda [68, 83] have published numerical studies using this formulation in 1D planewave vacuum spacetime.

3.3 Variables

Throughout this paper, lower case Greek letters denote spacetime indices (0-3), and lower case Latin letters denote only spatial indices (1-3). The letters in the beginning of the alphabets, (α, β, γ, δ, ϵ) and (a, b, c, d, e, f), denote tetrad indices. Mid-alphabet letters (λ, μ, ν) and (i, j, k, l) denote coordinate indices. Repeated indices are summed in all cases.

For an orthonormal tetrad of basis vector fields, there are twenty-four distinct connection coefficients. These coefficients are scalar fields under coordinate transformations, and are defined as

$$\Gamma_{\alpha\beta\gamma} = \boldsymbol{e}_\alpha \cdot \nabla_\gamma \boldsymbol{e}_\beta = -\Gamma_{\beta\alpha\gamma}, \tag{3.1}$$

with ∇ the covariant derivative operator. The $\Gamma_{\alpha\beta\gamma}$ are the same as Ricci rotation coefficients (see Wald [79]). They can be written in terms of commutators of the basis vectors:

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2}\{\langle \boldsymbol{e}^\delta, [\boldsymbol{e}_\alpha, \boldsymbol{e}_\beta]\rangle \eta_{\delta\gamma} + \langle \boldsymbol{e}^\delta, [\boldsymbol{e}_\alpha, \boldsymbol{e}_\gamma]\rangle \eta_{\delta\beta} - \langle \boldsymbol{e}^\delta, [\boldsymbol{e}_\beta, \boldsymbol{e}_\gamma]\rangle \eta_{\delta\alpha}\}. \tag{3.2}$$

Just as in the $3+1$ formalisms, it is convenient to make a space-time split in the tetrad formulation. Here, however, the split is relative to the timelike congruence defined by the tetrad rather than the constant-t spacelike hypersurface. The connection coefficients can then be re-labeled as 3D quantities with spatial triad indices (see Wahlquist [78] and ERW). To begin,

$$K_{ba} \equiv \Gamma_{a0b}, \tag{3.3}$$

where the symmetric part of K_{ba} is the rate of strain of the congruence, and the antisymmetric part,

$$\Omega_a \equiv \frac{1}{2}\,\varepsilon_{abc}\,K_{bc}, \tag{3.4}$$

is the vorticity of the congruence. If $\Omega_a = 0$, the congruence is hypersurface orthogonal, and K_{ba} is the traditional extrinsic curvature of the orthogonal hypersurface. Note that the sign of K_{ba} here is the same as in ERW and Wald, but opposite to that of Misner, Thorne, and Wheeler [54].

The spatial tetrad connection coefficients can be expressed more compactly as a two-index quantity defined by

$$N_{ba} \equiv \frac{1}{2}\,\varepsilon_{acd}\,\Gamma_{cdb}. \tag{3.5}$$

The diagonal components, N_{11}, N_{22}, and N_{33}, describe the twists of the spatial triads along the 1, 2, or 3 directions, respectively. The combinations $N_{ab} + N_{ba}$ of the non-diagonal components represent gravitational wave degrees of freedom. It is sometimes convenient to represent the antisymmetric part of N_{ba} by its own symbol

$$n_a \equiv \frac{1}{2}\,\varepsilon_{abc}\,N_{bc}. \tag{3.6}$$

The acceleration of the congruence is

$$a_a \equiv \Gamma_{a00}, \tag{3.7}$$

and the angular velocity of the spacelike triads relative to Fermi propagated axes is

$$\omega_a \equiv \frac{1}{2}\,\varepsilon_{abc}\,\Gamma_{cb0}. \tag{3.8}$$

There are nine K_{ba} and nine N_{ba}, giving eighteen primary variables to be evolved. The three a_a and three ω_a are gauge quantities, which, in this paper, are evolved by either the Nester or Lorentz gauge, or kept fixed.

Numerical calculations are performed with a particular choice of coordinates x^μ. Hyperbolic evolution consists in calculating variables on the spacelike

hypersurfaces characterized by $x^0 = t_2$ from an initial state specified by values of the variables on an earlier spacelike hypersurface $x^0 = t_1$. Let λ^μ_α denote the transformation matrix between coordinate basis vectors and the tetrad basis vectors:

$$\boldsymbol{e}_\alpha = \lambda^\mu_\alpha \, \boldsymbol{e}_\mu. \tag{3.9}$$

From Eq. (3.9), we obtain directional derivatives along the tetrad directions in terms of the partial derivatives along the coordinate directions,

$$D_\alpha = \lambda^\mu_\alpha \frac{\partial}{\partial x^\mu}. \tag{3.10}$$

The λ^μ_α are the coordinate components of the tetrad basis vectors. The coordinate metric is constructed from these tetrad vector components as $g^{\mu\nu} = \eta^{\alpha\beta} \, \lambda^\mu_\alpha \, \lambda^\nu_\beta$.

We find it simpler to use as variables not the sixteen λ^μ_α, but the following sixteen quantities which completely determine the tetrad vector components: nine B^k_a, the coordinate components of projections of the spatial tetrad vectors into the hypersurface; three A_a, which measure the tilts of the spatial tetrad vectors relative to the hypersurface and are (minus) the tetrad components of the 3-velocity of the hypersurface frame relative to the tetrad frame; the tetrad lapse function α and the three coordinate components of the tetrad shift vector β^k, which describe the evolution of the coordinates relative to the tetrad congruence. In this paper we evolve the B^k_a and the A_a as dynamic variables, but take α and the β^k to be fixed functions of the coordinates. Eventually we may want to expand the hyperbolic system to include dynamic equations for the lapse and the shift.

In a $3+1$ formalism, the lapse function N is the rate of change of proper time with respect to coordinate time along the hypersurface normal, and the shift vector N^k is the rate of displacement of the spatial coordinates with respect to coordinate time relative to the hypersurface normal, such that the coordinate

velocity of the normal worldline $dx^k/dt = -N^k$. Our tetrad lapse function α is the rate of change of proper time with respect to coordinate time along the tetrad congruence, and our tetrad shift vector β^k is the rate of displacement of the spatial coordinates relative to the tetrad congruence worldlines per unit coordinate time. For simple comoving coordinates, the spatial coordinates are constant along the congruence, and $\beta^k = 0$. However, comoving coordinates are not desirable in black hole calculations, since with a finite acceleration tetrad worldlines will be continuously advected inward across the event horizon. The tetrad lapse must be chosen so that the constant-time hypersurfaces remain spacelike, which is equivalent to the condition $A_a\, A_a < 1$.

The projection of $\boldsymbol{e}_a$ into the hypersurface is done along the congruence worldlines, such that

$$\boldsymbol{e}_a = A_a\, \boldsymbol{e}_0 + \boldsymbol{B}_a. \tag{3.11}$$

Fig. 3.1 shows how these various vectors are related from the point of view of a frame at rest with respect to the hypersurface.

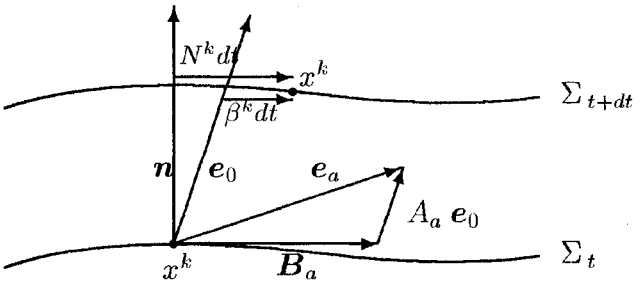


Figure 3.1: Decomposition of $\boldsymbol{e}_a$ into a vector tangent to the hypersurface, $\boldsymbol{B}_a$, and a vector parallel to the congruence, $A_a\, \boldsymbol{e}_0$. Displacements of the spatial coordinates, x^k, relative to the congruence worldline equal $\beta^k dt$, where β^k is our tetrad shift vector. Displacements of x^k relative to the hypersurface normal $\boldsymbol{n}$ are $N^k dt$, where N^k is the $3+1$ shift vector.

The directional derivative along a spatial tetrad direction is

$$D_a = A_a\, D_0 + B_a^k\, \frac{\partial}{\partial x^k}, \tag{3.12}$$

and the directional derivative along the congruence is

$$D_0 = \frac{1}{\alpha}\left(\frac{\partial}{\partial t} - \beta^k\, \frac{\partial}{\partial x^k}\right). \tag{3.13}$$

Plugging Eq. (3.13) into Eq. (3.12), we get

$$D_a = \left(B_a^k - \frac{A_a}{\alpha}\,\beta^k\right)\frac{\partial}{\partial x^k} + \frac{A_a}{\alpha}\,\frac{\partial}{\partial t}. \tag{3.14}$$

Comparing Eqs. (3.13) and (3.14) with Eq. (3.10), we can read off the tetrad vector components. They are

$$\lambda^\mu_\alpha = \begin{bmatrix} \lambda^0_0 & \lambda^k_0 \\ \\ \lambda^0_a & \lambda^k_a \end{bmatrix} = \begin{bmatrix} \frac{1}{\alpha} & -\frac{\beta^k}{\alpha} \\ \\ \frac{A_a}{\alpha} & \left(B_a^k - \frac{A_a}{\alpha}\,\beta^k\right) \end{bmatrix}. \tag{3.15}$$

The $3+1$ lapse function, N, and shift vector components, N^k, can be expressed in terms of α, β^k, A_a, and B_a^k by constructing g^{tt} and g^{tk} from the tetrad vector components using $g^{\mu\nu} = \eta^{\alpha\beta}\,\lambda^\mu_\alpha\,\lambda^\nu_\beta$. With the relations

$$g^{tt} = -\frac{1}{N^2} \quad \text{and} \quad g^{tk} = -\frac{N^k}{N^2}, \tag{3.16}$$

we get

$$N = \frac{\alpha}{\sqrt{1 - A_a\, A_a}}, \tag{3.17}$$

$$N^k = \beta^k + \frac{\alpha\, A_a\, B_a^k}{1 - A_a\, A_a}. \tag{3.18}$$

The tetrad lapse is smaller than the 3+1 lapse due to the time dilation of the tetrad observer in the rest frame of the hypersurface.

3.4 The tetrad equations

In this section, we present the evolution and constraint equations for the connection coefficients defined in Sec. 3.3 in terms of directional derivative operators along the tetrad directions. The structure of the equations in this form is deceptively simple. One has to keep in mind that partial differential equations are solved in numerical relativity, so the directional derivatives have to be expanded into partial derivatives using Eqs. (3.13) and (3.14). In terms of directional derivatives, the equations are quasi-FOSH. We will present the necessary steps to obtain the true hyperbolic form in Sec. 3.5. Furthermore, we call the "constraint" equations presented in this section quasi-constraint equations, as is done in [22], because although they contain only spatial directional derivatives, these spatial directionals contain time partials (see Eq. (3.14)). The details of converting the quasi-constraint equations into true constraint equations are given in Appendix B.

In both the tetrad and $3+1$ approaches, there is an ordering ambiguity inherent in the derivation of a first order hyperbolic evolution system from the Einstein equations. In first order $3+1$ formalisms, the spatial derivatives of the metric are independent variables, D_{jkl} (which equal $\frac{1}{2}\partial_j h_{kl}$, where h_{kl} is the spatial metric). Index re-ordering of the Riemann tensor is an interchange of spatial partial derivatives such that

$$\partial_i D_{jkl} = \partial_j D_{ikl}. \tag{3.19}$$

Exploiting this freedom leads to a wide variety of hyperbolic formulations [39]. The standard energy and momentum constraint equations can also be added to the evolution system to get additional hyperbolic or non-hyperbolic structures. In the context of a tetrad approach, the connection coefficients are given in terms of the commutation relations, Eq. (3.2). To derive integrability conditions from Eq. (3.2) is messy. We find it is much easier to approach the question

of ordering ambiguity from the symmetries of the Riemann tensor, especially since these symmetries are explicit when the indices are all up or down, and it is trivial to raise and lower indices with the Minkowski metric. We do not consider a whole range of schemes as do Kidder, Scheel, and Teukolsky [39], but rather a particular scheme which leads to a simpler hyperbolic structure than the others. To focus on the quasi-FOSH structure of the equations, we present only the principal terms here. The second order source terms are given in Appendix B.

The Riemann tensor projected onto a tetrad is

$$\begin{aligned} R_{\alpha\beta\gamma\delta} &= \boldsymbol{e}_\alpha \cdot \left(\nabla_\gamma \nabla_\delta - \nabla_\delta \nabla_\gamma - \nabla_{[\boldsymbol{e}_\gamma, \boldsymbol{e}_\delta]}\right) \boldsymbol{e}_\beta \\ &= D_\gamma\, \Gamma_{\alpha\beta\delta} - D_\delta\, \Gamma_{\alpha\beta\gamma} + \Gamma_{\alpha\epsilon\gamma}\, \Gamma^{\epsilon}{}_{\beta\delta} - \Gamma_{\alpha\epsilon\delta}\, \Gamma^{\epsilon}{}_{\beta\gamma} + \Gamma_{\alpha\beta\epsilon}\, (\Gamma^{\epsilon}{}_{\gamma\delta} - \Gamma^{\epsilon}{}_{\delta\gamma}), \end{aligned}$$

where D_α represents directional derivatives along the tetrad directions. The antisymmetry of $R_{\alpha\beta\gamma\delta}$ on the second pair of indices is explicit in the above equation. The antisymmetry on the first pair is also trivial because of the antisymmetry of $\Gamma_{\alpha\beta\gamma}$ on the first two indices. However, the Riemann identities

$$R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}, \tag{3.20}$$

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\delta\beta\gamma} + R_{\alpha\gamma\delta\beta} = 0 \tag{3.21}$$

lead to new Riemann constraints, which we exploit in the following.

First we derive all the possible quasi-constraint equations, noting the use of Eq. 14.7 from Misner, Thorne, and Wheeler. The energy quasi-constraint equation is

$$G_{00} = R_{1212} + R_{2323} + R_{3131} \sim 2\, D_a\, n_a. \tag{3.22}$$

An analogous quasi-constraint equation for Ω_a is derived using the cyclic identity, Eq. (3.21), on $0abc$:

$$0 = R_{0213} + R_{0321} + R_{0132} \sim 2\, D_a\, \Omega_a. \tag{3.23}$$

The momentum quasi-constraint equations are obtained from $G_{0a} = 0$:

$$\begin{aligned} G_{01} &= R_{0212} + R_{0313} \\ &\sim -D_1 \left(K_{22} + K_{33}\right) + D_2 \, K_{12} + D_3 \, K_{13}, \end{aligned} \tag{3.24}$$

$$\begin{aligned} G_{02} &= R_{0121} + R_{0323} \\ &\sim -D_2 \left(K_{11} + K_{33}\right) + D_1 \, K_{21} + D_3 \, K_{23}, \end{aligned} \tag{3.25}$$

$$\begin{aligned} G_{03} &= R_{0131} + R_{0232} \\ &\sim -D_3 \left(K_{11} + K_{22}\right) + D_1 \, K_{31} + D_2 \, K_{32}. \end{aligned} \tag{3.26}$$

Similar quasi-constraint equations involving N_{ab} are derived solely from Eq. (3.20) applied to spatial indices of the Riemann tensor:

$$R_{1213} - R_{1312} \sim D_1 \left(N_{22} + N_{33}\right) - D_2 \, N_{12} - D_3 \, N_{13}, \tag{3.27}$$

$$R_{2321} - R_{2123} \sim D_2 \left(N_{11} + N_{33}\right) - D_1 \, N_{21} - D_3 \, N_{23}, \tag{3.28}$$

$$R_{3132} - R_{3231} \sim D_3 \left(N_{11} + N_{22}\right) - D_1 \, N_{31} - D_2 \, N_{32}. \tag{3.29}$$

We now to turn to evolution equations. By using Eq. (3.20) in the momentum quasi-constraint equations, we calculate *evolution* equations for the non-diagonal components of N_{ab}. For example, interchanging the first and second pairs of indices in Eq. (3.26) as shown below gives evolution equations for N_{21} and N_{12}:

$$\begin{aligned} G_{03} &= R_{0131} + R_{3202} \\ &\sim -D_0 \, N_{21} - D_2 \, \omega_1 - D_3 \, K_{11} + D_1 \, K_{31}, \end{aligned} \tag{3.30}$$

$$\begin{aligned} G_{03} &= R_{3101} + R_{0232} \\ &\sim D_0 \, N_{12} + D_1 \, \omega_2 + D_2 \, K_{32} - D_3 \, K_{22}. \end{aligned} \tag{3.31}$$

Evolution equations for the diagonal components of N_{ab} are obtained solely from applications of Eq. (3.20):

$$R_{2301} - R_{0123} \sim D_0\, N_{11} + D_1\, \omega_1 + D_2\, K_{31} - D_3\, K_{21}, \tag{3.32}$$

$$R_{3102} - R_{0231} \sim D_0\, N_{22} + D_2\, \omega_2 + D_3\, K_{12} - D_1\, K_{32}, \tag{3.33}$$

$$R_{1203} - R_{0312} \sim D_0\, N_{33} + D_3\, \omega_3 + D_1\, K_{23} - D_2\, K_{13}. \tag{3.34}$$

Evolution equations for K_{ab} are obtained from the Einstein equations. As is done elsewhere (see for example [16]), a multiple of the energy quasi-constraint equation times the spatial metric can be subtracted from the K_{ab} evolution equations. In the formulation we are presenting, this amounts to subtracting a multiple of Eq. (3.22) from the diagonal spatial components of the Ricci tensor, since our metric is just δ^b_a. The same type of procedure in $3+1$ formulations affects the evolution equations for the non-diagonal as well as the diagonal components of K_{ab}, since the spatial metric there is, in general, non-diagonal. The number we choose to multiply the energy quasi-constraint equation by is 1, to bring the evolution equations for the diagonal K_{ab} into the same form as those for the non-diagonal components. With the aid of Eq. (3.20), we derive sample evolution equations for K_{ab} below:

$$\begin{aligned} R_{12} &= -R_{1020} + R_{1323} = -R_{1020} + R_{2313} \\ &\sim D_0\, K_{21} - D_2\, a_1 - D_3\, N_{11} + D_1\, N_{31}, \end{aligned} \tag{3.35}$$

$$\begin{aligned} R_{12} &= -R_{2010} + R_{1323} \\ &\sim D_0\, K_{12} - D_1\, a_2 - D_2\, N_{32} + D_3\, N_{22}, \end{aligned} \tag{3.36}$$

$$\begin{aligned} R_{11} - G_{00} &= -R_{0101} - R_{2323} \\ &\sim D_0\, K_{11} - D_1\, a_1 - D_2\, N_{31} + D_3\, N_{21}. \end{aligned} \tag{3.37}$$

If the gauge quantities a_a and ω_a are fixed functions of time, then the quasi-FOSH structure is now complete and is represented by the following system of

eighteen equations for the nine K_{ab} and nine N_{ab}:

$$D_0\, K_{ab} - D_a\, a_b - \varepsilon_{acd}\, D_c\, N_{db} = S\_K_{ab}, \tag{3.38}$$

$$D_0\, N_{ab} + D_a\, \omega_b + \varepsilon_{acd}\, D_c\, K_{db} = S\_N_{ab}, \tag{3.39}$$

where S_K_{ab} and S_N_{ab} are source terms quadratic in the connection coefficient variables. (See Appendix C.)

Alternatively, one can implement a dynamic gauge. Both the Nester gauge used by ERW and EW and the Lorentz gauge used by van Putten and Eardley result in evolution equations for a_a and ω_a which , when added to Eqs. (3.38) and (3.39), form a quasi-FOSH system.

The Nester gauge conditions in [56] are defined for an arbitrary number of dimensions. In 4D spacetime, the Nester conditions state that two 1-forms, $\tilde{\boldsymbol{q}}$ and $\hat{\boldsymbol{q}}$, whose tetrad components are

$$\tilde{q}_\alpha = \Gamma_\alpha{}^\gamma{}_\gamma, \quad \hat{q}_\alpha = \varepsilon_{\alpha\beta\gamma\delta}\, \Gamma^{\beta\gamma\delta}, \tag{3.40}$$

are closed. This implies vanishing exterior derivatives so that, in tetrad component form, the Nester conditions are

$$\nabla_\alpha\, \tilde{q}_\beta - \nabla_\beta\, \tilde{q}_\alpha = 0, \tag{3.41}$$

$$\nabla_\alpha\, \hat{q}_\beta - \nabla_\beta\, \hat{q}_\alpha = 0. \tag{3.42}$$

The Nester conditions result in six evolution equations,

$$\begin{aligned} D_0\, a_b - D_c\, K_{cb} &= -(\omega_c - 2\,\Omega_c)\, N_{cb} + (TrN)\, \omega_b, \\ D_0\, \omega_b + D_c\, N_{cb} &= -(\omega_c - 2\,\Omega_c)\, K_{cb} - (TrN)\, a_b, \end{aligned} \tag{3.43}$$

and six quasi-constraint equations,

$$\begin{aligned} \varepsilon_{abc}\, D_b\, (a_c - 2\, n_c) &= 2\, (TrK)\, \Omega_a - (TrN)\, (a_a - 2\, n_a) + (a_d - 2\, n_d)\, N_{da}, \\ \varepsilon_{abc}\, D_b\, (\omega_c - 2\, \Omega_c) &= -(TrN)\, \omega_a + (\omega_d - 2\, \Omega_d)\, N_{da}. \end{aligned} \tag{3.44}$$

The Lorentz gauge condition in [77] is $\nabla^{\gamma}\omega_{\gamma\alpha\beta} = 0$, where $\omega_{\gamma\alpha\beta} = \Gamma_{\alpha\beta\gamma}$. For fixed α and β, the $\omega_{\gamma\alpha\beta}$ are the components of connection 1-forms (see Wald [79]). Expressed in terms of $\Gamma_{\alpha\beta\gamma}$, the Lorentz gauge condition is

$$D_{\delta}\,\Gamma_{\alpha\beta}{}^{\delta} + \Gamma_{\alpha\beta}{}^{\gamma}\,\Gamma^{\delta}{}_{\gamma\delta} = 0. \tag{3.45}$$

The Lorentz gauge results in six evolution equations:

$$\begin{aligned} D_0\,a_b - D_c\,K_{cb} &= (a_c - 2\,n_c)\,K_{cb} - (TrK)\,a_b, \\ D_0\,\omega_b + D_c\,N_{cb} &= -(a_c - 2\,n_c)\,N_{cb} - (TrK)\,\omega_b, \end{aligned} \tag{3.46}$$

and no additional constraint equations. Note that the principal terms in Eq. (3.46) are identical to those for the Nester gauge evolution equations, Eq. (3.43).

The evolution equations expressed in terms of directional derivatives, Eqs. (3.38), (3.39), and (with a dynamic gauge) (3.43) or (3.46), can be written in a condensed notation,

$$D_0\,\mathbf{q} + \boldsymbol{M}^a\,D_a\,\mathbf{q} = \mathbf{S}, \tag{3.47}$$

where $\mathbf{q}$ is a vector of the twenty-four variables N_{ab}, K_{ab}, a_a, and ω_a. Furthermore, $\boldsymbol{M}^a$ are three sparse 24×24 matrices whose only nonzero elements are $\pm\,1$, and $\mathbf{S}$ is a vector of source terms. If one orders the variables so that

$$\begin{aligned} \mathbf{q} = (&N_{11}, N_{21}, N_{31}, a_1, K_{11}, K_{21}, K_{31}, \omega_1, \\ &N_{12}, N_{22}, N_{32}, a_2, K_{12}, K_{22}, K_{32}, \omega_2, N_{13}, N_{23}, N_{33}, a_3, K_{13}, K_{23}, K_{33}, \omega_3), \end{aligned} \tag{3.48}$$

then the $\boldsymbol{M}^a$ matrices have a simple block diagonal structure. Each is com-

posed of three identical 8×8 blocks, respectively:

$$\boldsymbol{M}^1_{block} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{3.49}$$

$$\boldsymbol{M}^2_{block} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{3.50}$$

$$\boldsymbol{M}^3_{block} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{3.51}$$

3.5 Hyperbolic structure of the coordinate equations

The beautiful quasi-FOSH tetrad formulation with constant coefficients given in Eqs. (3.47) to (3.51) is not so simple, or so beautiful, when expressed in terms of coordinates. However, we proceed to show that the coordinate equations are still quite manageable and are, indeed, symmetrizable hyperbolic.

Substituting Eqs. (3.13) and (3.14) into Eq. (3.47), we express the tetrad evolution equations in terms of partial derivatives along coordinate directions:

$$\boldsymbol{T}\, D_0\, \mathbf{q} + \boldsymbol{M}^a\, B_a^k\, \frac{\partial\, \mathbf{q}}{\partial x^k} = \mathbf{S}, \tag{3.52}$$

where

$$\boldsymbol{T} \equiv \left[\boldsymbol{I} + \boldsymbol{M}^a\, A_a\right], \tag{3.53}$$

and $\boldsymbol{I}$ is the identity matrix. For clarity of notation, we let D_0 represent the partials in Eq. (3.13). The $\boldsymbol{T}$ matrix has a block diagonal structure composed of three identical 8×8 entries which are

$$\boldsymbol{T}_{block} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -A_3 & A_2 & A_1 \\ 0 & 1 & 0 & 0 & A_3 & 0 & -A_1 & A_2 \\ 0 & 0 & 1 & 0 & -A_2 & A_1 & 0 & A_3 \\ 0 & 0 & 0 & 1 & -A_1 & -A_2 & -A_3 & 0 \\ 0 & A_3 & -A_2 & -A_1 & 1 & 0 & 0 & 0 \\ -A_3 & 0 & A_1 & -A_2 & 0 & 1 & 0 & 0 \\ A_2 & -A_1 & 0 & -A_3 & 0 & 0 & 1 & 0 \\ A_1 & A_2 & A_3 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.54}$$

We now multiply Eq. (3.52) by $\boldsymbol{T}^{-1}$ to give

$$D_0\, \mathbf{q} + \boldsymbol{C}^a\, B_a^k\, \frac{\partial\, \mathbf{q}}{\partial x^k} = \boldsymbol{T}^{-1}\mathbf{S}, \tag{3.55}$$

where

$$\boldsymbol{C}^a \equiv \boldsymbol{T}^{-1}\boldsymbol{M}^a. \tag{3.56}$$

$\boldsymbol{T}^{-1}$ is straightforward to calculate, since it is a block diagonal matrix with each block equal to the inverse of $\boldsymbol{T}_{block}$ in Eq. (3.54), where

$$\boldsymbol{T}^{-1}_{block} = \frac{1}{1 - A_a\, A_a} \times \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & A_3 & -A_2 & -A_1 \\ 0 & 1 & 0 & 0 & -A_3 & 0 & A_1 & -A_2 \\ 0 & 0 & 1 & 0 & A_2 & -A_1 & 0 & -A_3 \\ 0 & 0 & 0 & 1 & A_1 & A_2 & A_3 & 0 \\ 0 & -A_3 & A_2 & A_1 & 1 & 0 & 0 & 0 \\ A_3 & 0 & -A_1 & A_2 & 0 & 1 & 0 & 0 \\ -A_2 & A_1 & 0 & A_3 & 0 & 0 & 1 & 0 \\ -A_1 & -A_2 & -A_3 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.57}$$

The system given in Eq. (3.55) is hyperbolic according to the definition given in [48] if any linear combination $b_a\, \boldsymbol{C}^a$ of $\boldsymbol{C}^1$, $\boldsymbol{C}^2$, and $\boldsymbol{C}^3$ can be diagonalized with a complete set of eigenvectors and real eigenvalues. The linear combination for propagation in the x^k coordinate direction has $b_a = B^k_a$. Solving the eigensystem of the combined 8×8 matrix in Mathematica gives for the eigenvalues

$$\frac{\boldsymbol{b}\cdot\boldsymbol{A} \mp \sqrt{\boldsymbol{b}\cdot\boldsymbol{b} - (\boldsymbol{b}\times\boldsymbol{A})\cdot(\boldsymbol{b}\times\boldsymbol{A})}}{1 - \boldsymbol{A}\cdot\boldsymbol{A}}, \tag{3.58}$$

where $\boldsymbol{b}$ and $\boldsymbol{A}$ are the 3-vectors in the spatial orthonormal frame with components b_a and A_a, respectively. The dot and cross products are the standard vector operations, and the upper/ lower signs on the square root are for left/ right- propagating (relative to $\boldsymbol{b}$) modes. The eigenvalues are real as long as $\boldsymbol{A}\cdot\boldsymbol{A} = A_a\, A_a < 1$. The eigenvectors (given in Appendix C) form a complete set. The lapse and shift hidden in the D_0 operator modify the eigenvalues in

a trivial way (multiply by the lapse α, then subtract the shift β^k), but have no effect on the eigenvectors.

For Eq. (3.55) to also be symmetrizable by the definition of Lindblom and Scheel [50], a positive definite symmetric matrix must be found which multiplies the $\boldsymbol{C}^a$ matrices to give symmetric matrices. The obvious candidate for such a symmetrizer is $\boldsymbol{T}$, since $\boldsymbol{T}\,\boldsymbol{C}^a = \boldsymbol{M}^a$ by Eq. (3.56), and the $\boldsymbol{M}^a$ matrices are symmetric. $\boldsymbol{T}$ is real and symmetric so a necessary and sufficient condition for it to be positive definite is that all its eigenvalues are positive [71]. This requires that $A_a\,A_a < 1$.

Note that our saying the system is hyperbolic is contingent on the evolution of the B^k_a components being hyperbolic. This is discussed further in Sec. 3.6.

3.6 Evolution and constraint equations for A_a and B^k_a

The commutator of the orthonormal basis vectors is

$$[\boldsymbol{e}_\alpha,\, \boldsymbol{e}_\beta] = \nabla_\alpha\, \boldsymbol{e}_\beta - \nabla_\beta\, \boldsymbol{e}_\alpha = (\Gamma^\gamma{}_{\beta\alpha} - \Gamma^\gamma{}_{\alpha\beta})\, \boldsymbol{e}_\gamma. \tag{3.59}$$

Expressed in terms of tetrad components and partial derivatives,

$$[\boldsymbol{e}_\alpha,\, \boldsymbol{e}_\beta] = \left(\lambda^\mu_\alpha \frac{\partial \lambda^\nu_\beta}{\partial x^\mu} - \lambda^\mu_\beta \frac{\partial \lambda^\nu_\alpha}{\partial x^\mu}\right) \frac{\partial}{\partial x^\nu}. \tag{3.60}$$

Expanding Eq. (3.60) and collecting terms gives

$$\begin{aligned} &\left(\lambda^t_\alpha \frac{\partial \lambda^l_\beta}{\partial t} + \lambda^k_\alpha \frac{\partial \lambda^l_\beta}{\partial x^k} - \lambda^t_\beta \frac{\partial \lambda^l_\alpha}{\partial t} - \lambda^k_\beta \frac{\partial \lambda^l_\alpha}{\partial x^k}\right) \frac{\partial}{\partial x^l} + \\ &\left(\lambda^t_\alpha \frac{\partial \lambda^t_\beta}{\partial t} + \lambda^k_\alpha \frac{\partial \lambda^t_\beta}{\partial x^k} - \lambda^t_\beta \frac{\partial \lambda^t_\alpha}{\partial t} - \lambda^k_\beta \frac{\partial \lambda^t_\alpha}{\partial x^k}\right) \frac{\partial}{\partial t}. \end{aligned} \tag{3.61}$$

Doing the same in Eq. (3.59), we get

$$\begin{aligned} &\left[(\Gamma^0{}_{\beta\alpha} - \Gamma^0{}_{\alpha\beta})\, \lambda^k_0 + (\Gamma^c{}_{\beta\alpha} - \Gamma^c{}_{\alpha\beta})\, \lambda^k_c\right] \frac{\partial}{\partial x^k} + \\ &\left[(\Gamma^0{}_{\beta\alpha} - \Gamma^0{}_{\alpha\beta})\, \lambda^t_0 + (\Gamma^c{}_{\beta\alpha} - \Gamma^c{}_{\alpha\beta})\, \lambda^t_c\right] \frac{\partial}{\partial t}. \end{aligned} \tag{3.62}$$

Set Eqs. (3.61) and (3.62) equal, and let the index $\alpha = 0$ and the index $\beta = a$. Simplifying, we obtain evolution equations for B_a^k and A_a:

$$\begin{aligned} D_0\, B_a^k + \frac{B_a^l}{\alpha}\frac{\partial \beta^k}{\partial x^l} &= \frac{1}{\alpha}\left(\frac{\partial}{\partial t} - \mathcal{L}_\beta\right) B_a^k \\ &= -\varepsilon_{abc}\,\omega_b\, B_c^k - K_{ac}\, B_c^k, \end{aligned} \tag{3.63}$$

$$D_0\, A_a = a_a - \varepsilon_{abc}\,\omega_b\, A_c - K_{ac}\, A_c - \frac{B_a^l}{\alpha}\frac{\partial \alpha}{\partial x^l}, \tag{3.64}$$

where $\mathcal{L}_\beta$ in Eq. (3.63) is the Lie derivative. For fixed lapse and shift, the evolution of both the B_a^k and the A_a is just advection along the congruence worldlines and trivially hyperbolic.

The congruence can always be chosen to be orthogonal to the initial hypersurface, so $A_a = 0$ initially. However, A_a will not remain zero during the subsequent evolution unless the condition

$$B_a^l \frac{\partial}{\partial x^l} log\ \alpha = a_a \tag{3.65}$$

is satisfied at all times. For either the fixed or the dynamic gauge conditions on a_a considered here, the evolution of B_a^l and/or a_a is inconsistent with Eq. (3.65), except possibly for very special initial conditions.

Repeating the same process as above, but with the index $\alpha = a$ and the index $\beta = b$, we obtain constraint equations for B_a^k and A_a:

$$\begin{aligned} &\varepsilon_{cab}\, B_a^l \frac{\partial B_b^k}{\partial x^l} = N_{dc}\, B_d^k - \\ &(TrN)\ B_c^k\ +\ \varepsilon_{cab}\ A_a\ (\varepsilon_{bdf}\ \omega_d\ B_f^k\ +\ K_{bd}\ B_d^k), \end{aligned} \tag{3.66}$$

$$\begin{aligned} &\varepsilon_{cab}\, B_a^l \frac{\partial A_b}{\partial x^l} = 2\,\Omega_c + N_{dc}\, A_d - \\ &(TrN)\ A_c\ +\ \varepsilon_{cab}\ A_a\ (-a_b\ +\ \varepsilon_{bdf}\ \omega_d\ A_f\ +\ K_{bd}\ A_d). \end{aligned} \tag{3.67}$$

Note that if $A_a = 0$, the constraint of Eq. (3.67) is satisfied automatically, and Eq. (3.66) can be used to calculate all of the N_{ab} from the B_a^k.

3.7 The Initial Value Problem

Initial conditions for the variables must be chosen so all relevant constraints are satisfied on the initial hypersurface. This is most easily accomplished if the initial tetrad is oriented so the tetrad congruence is orthogonal to the initial hypersurface. Then, $\Omega_a = 0$ and $A_a = 0$ initially. The initial B_a^k are the components of an orthonormal triad of vectors lying in the initial hypersurface, related to the inverse of the spatial metric of the hypersurface by $h^{kl} = B_a^k \, B_a^l$. One way to construct consistent initial conditions is to solve the initial value problem using standard $3 + 1$ methods for the spatial metric and the extrinsic curvature. Construct orthonormal triad fields by a Gramm-Schmidt orthogonalization procedure, orienting the B_1 vector along the x^1 coordinate direction, and B_2 in the $x^1 - x^2$ plane, for instance. Find the N_{ab} from the commutators of the B_a 3-vectors using Eq. (3.66). The K_{ab} are simply the projections of the coordinate components of the extrinsic curvature as found from the $3+1$ initial value problem along the orthonormal triad vectors. This procedure is guaranteed to give consistent initial conditions for the tetrad vectors, as long as the spatial coordinates are basically Cartesian, *ie.*, there are no spatial coordinate singularities. The initial acceleration and angular velocity of the tetrad are arbitrary, except in the context of the Nester gauge. In this case, the initial angular velocity ω_a and the initial $a_a - 2\, n_a$ must have vanishing exterior derivatives. However, there is no guarantee that there are not large twists of the initial triad vectors, possibly leading to large gauge transients in the context of one of the dynamical gauge conditions.

A more elegant and, likely, a better-behaved choice for the initial spatial triad is to require that it satisfy the 3D Nester gauge conditions in the hypersurface. These conditions are that the 3D one-form $\tilde{q}_a$, whose triad components are $\Gamma_a{}^b{}_b$, has zero exterior derivative, and that the trace of the N_{ab} matrix van-

ish, *ie.* $N_{11}+N_{22}+N_{33} = 0$. For simple topologies, the first condition is equivalent to the condition that $2\, n_a = \varepsilon_{abc}\, N_{bc}$ be the gradient of a scalar. Finding a solution for the B^k_a which satisfies these conditions is, in general, a non-trivial elliptic problem (see [55] and [23]). However, the situation is much simpler for conformally flat 3-geometries. Taking Cartesian basis vectors in the conformal geometry, the conformal N_{ab} are zero. If the conformal factor which generates the physical metric is $e^{4\psi}$, the conformal transformation simply rescales the B^k_a by a factor $e^{-2\psi}$. The physical n_a equal $-2\, D_a\, \psi$ (the gradient of a scalar), and the symmetric part of N_{ab} is still zero. The 3D Nester condition is satisfied.

3.8 Discussion

Although the emphasis of the tetrad formalisms in the literature is on evolving spacetimes with physically defined flows, we think a tetrad formalism may be useful for vacuum numerical calculations of black hole spacetimes. The tetrad formulation we have presented, based on that of Estabrook, Robinson, and Wahlquist [24], allows control of the evolution of the timelike congruence through either dynamic or fixed gauge conditions on a_a and ω_a. The lapse and shift we have introduced allow for a completely arbitrary evolution of the coordinates with minimal complication of the formalism. Furthermore, since the variables evolved are defined relative to the orthonormal frames, the metric is the Minkowski metric and there are no nonlinearities in the equations associated with the inverse metric. The system of equations based on coordinate derivatives is symmetrizable hyperbolic, though admittedly more complicated than the quasi-FOSH system based on directional derivatives. Finally, the variables N_{ab}, K_{ab}, a_a and ω_a are all scalars, so derivatives of the shift only appear in the evolution equations for B^k_a. We have successfully implemented this tetrad formalism for 1D colliding gravitational plane waves as described in Chapter

4, where it results in substantially better accuracy and stability compared with our calculations in the $3 + 1$ framework [13].

There are still important issues to be worked out in order to apply this tetrad formalism to 3D black hole codes. First, one must deal with the complications that arise when evolving the congruence as well as the hypersurface. The congruence stays timelike as long as the acceleration is bounded; however, in order for the the hypersurfaces to remain spacelike, the condition $A_a A_a < 1$ must be satisfied at all times. If $A_a A_a = 1$, the system breaks down completely. The hypersurfaces become null, causing the $3 + 1$ lapse to blow up. In addition, the coordinate equations become singular (see Appendix C). $A_a A_a$ can be kept small by an adjustment of the lapse as the calculation proceeds. Such a resetting of the lapse invalidates theorems which bound the growth of the solution based on the symmetrizable hyperbolic structure of the system, but does not affect other advantages of the hyperbolic formulation, such as dealing with boundary conditions.

Second, attaining a stationary solution at late times for general 3D black hole numerical calculations may be problematic in a tetrad formulation. Such solutions are potentially desirable for long-term stability. With bounded acceleration, congruence worldlines just outside the apparent horizon will necessarily cross the horizon and be trapped by the black hole. Worldlines further out may actually accelerate outward relative to a static or stationary observer. While a shift adjusted just right may keep grid stretching under control in spite of this, the adjustment will have to be time-dependent except for very special initial conditions on the congruence. Knowing what these special initial conditions are requires knowing the whole future solution ahead of time, and thus is not practical.

Both of these issues will probably best be addressed by extending the symmetrizable hyperbolic system to allow for a dynamic lapse, to keep $A_a A_a$ small,

and a dynamic shift, to control grid-stretching.

3.9 Acknowledgments

LTB and JMB gratefully acknowledge F. B. Estabrook and H. D. Wahlquist for their contributions to this work. We would also like to thank L. Lindblom, M. Scheel, and D. Meier for their support and insightful discussions during the CalTech visitors program. LTB was supported by the NASA Graduate Student Researchers Program under Grant No. NGT5-50298 for the duration of this research.

Chapter 4

EW TETRAD APPLICATION 1: 1D COLLIDING GRAVITATIONAL PLANE WAVES

4.1 Introduction

The first order symmetric hyperbolic EW tetrad formulation, with Nester gauge and zero shift, is applied to one-dimensional colliding gravitational plane waves of general polarization. The accuracy of the numerical simulations increases 40 to 70 fold over the results presented in Chapter 2, for $3+1$ formulations. This increase in accuracy is correlated with the fact that in the EW formulation, the constraint variable combination excited in the non-linear solution propagates into the grid while the corresponding eigenmode of the characteristic matrix (with the source terms set equal to zero) travels out of the grid. This supports our results in Chapter 2, where we found an increase in accuracy when the speeds of the "constraint" eigenmodes are adjusted so that they are different from the speeds of features in the constraint quantities associated with the dominant source of errors.

4.2 Variables

There are seven contravariant tetrad vector components:

$$\lambda_1^x, \quad \lambda_2^y, \quad \lambda_3^z, \quad \lambda_2^z, \quad \lambda_3^y, \quad \lambda_0^t, \quad \lambda_1^t. \tag{4.1}$$

From Eq. (3.15), we find the following relations, assuming zero shift:

$$\lambda_a^k = B_a^k, \quad \lambda_0^t = \frac{1}{\alpha}, \quad \lambda_1^t = \frac{A_1}{\alpha}, \tag{4.2}$$

where $a = 1,2,3$ and $k = x,y,z$, and $A_1 = \lambda_1^x\, A_x$. We can calculate $g^{\mu\nu}$ from $g^{\mu\nu} = \eta^{\alpha\beta}\, \lambda_\alpha^\mu\, \lambda_\beta^\nu$. Inverting to get $g_{\mu\nu}$, we derive an expression for the coordinate metric,

$$ds^2 = -\alpha^2\, dt^2 + 2\, A_x\, \alpha\, dx\, dt + \left(h_{xx} - A_x^2\right)\, dx^2 + h_{yy}\, dy^2 + h_{zz}\, dz^2 + 2\, h_{yz}\, dy\, dz, \quad (4.3)$$

where $h_{xx} = 1/(\lambda_1^x)^2$, and h_{yy}, h_{zz}, and h_{yz} are complicated functions of λ_2^y, λ_3^z, λ_2^z, and λ_3^y. Using the contravariant tetrad vector components as variables, instead of the coordinate metric coefficients h_{xx}, h_{yy}, and h_{zz}, simplifies the equations considerably. It also eliminates any need to take square roots or inverses of metric coefficient terms. Note that $h_{xx} - A_x^2$ is the metric coefficient of the $t =$ constant hypersurface, whereas h_{xx} is the metric coefficient of the hypersurface which is orthogonal to the timelike congruence of the tetrad.

For convenience, we express Eq. (3.2) for the Ricci rotation coefficients in terms of the contravariant (λ_a^k) and covariant (λ_{ak}) tetrad vector components:

$$2\,\Gamma_{abc} = \lambda_{bk}\left[\lambda_a^l\,\frac{\partial\lambda_c^k}{\partial_{x^l}} - \lambda_c^l\,\frac{\partial\lambda_a^k}{\partial_{x^l}}\right] + \lambda_{ck}\left[\lambda_a^l\,\frac{\partial\lambda_b^k}{\partial_{x^l}} - \lambda_b^l\,\frac{\partial\lambda_a^k}{\partial_{x^l}}\right] - \lambda_{ak}\left[\lambda_b^l\,\frac{\partial\lambda_c^k}{\partial_{x^l}} - \lambda_c^l\,\frac{\partial\lambda_b^k}{\partial_{x^l}}\right]. \quad (4.4)$$

Recall from Chapter 3 that Γ_{abc} are antisymmetric in the first two indices. We can now calculate the dyadic and vector quantities from Eqs. (3.3), (3.5), (3.7), and (3.8), and find the non-zero quantities to be:

$$N_{23},\ N_{32},\ N_{11},\ N_{33} = -N_{22},\ a_1,\ K_{23} = K_{32},\ K_{11},\ K_{22},\ K_{33},\ \omega_1. \quad (4.5)$$

4.3 Tetrad Quasi-Evolution and Quasi-Constraint Equations

The tetrad quasi-evolution equations are derived from Eqs. (3.38), (3.39), and (3.43) (for the Nester gauge). The resulting equations are

$$\begin{aligned} &D_0\, N_{23} - D_1\, K_{33} \\ &= -K_{22}(a_1 + N_{23}) - 2K_{23}(N_{11} + N_{33}) + N_{32}(K_{33} - K_{11}) + 2N_{33}\omega_1, \end{aligned} \quad (4.6)$$

$$D_0 N_{32} + D_1 K_{22} = K_{33}(a_1 - N_{32}) - 2K_{23}(N_{11} - N_{33}) + N_{23}(K_{22} - K_{11}) + 2N_{33}\omega_1, \quad (4.7)$$

$$D_0\, N_{11} + D_1\, \omega_1 = -K_{11}\, N_{11} + (K_{22} - K_{33})\, N_{33} - K_{23}\, (N_{23} + N_{32}) - a_1\, \omega_1, \quad (4.8)$$

$$D_0\, N_{33} + D_1\, K_{23} = (K_{22} - K_{33})\, N_{11} - K_{11}\, N_{33} - a_1\, K_{23} - \omega_1\, (N_{23} + N_{32}), \quad (4.9)$$

$$D_0\, K_{23} + D_1\, N_{33} = N_{11}\, (N_{23} + N_{32}) + \omega_1\, (K_{33} - K_{22})$$
$$+\ N_{33}\, (N_{23} - N_{32} - a_1) - K_{23}\, (K_{11} + K_{22} + K_{33}), \quad (4.10)$$

$$D_0\, K_{11} - D_1\, a_1 = a_1^2 - K_{11}^2 - K_{23}^2 + K_{22}\, K_{33} + N_{33}^2 + N_{23}\, N_{32}, \quad (4.11)$$

$$D_0\, K_{22} + D_1\, N_{32}$$
$$= -K_{22}^2 - K_{23}^2 + K_{11}\, K_{33} - 2\, N_{11}\, N_{33} - N_{33}^2 - a_1\, N_{23} - N_{32}^2 + 2\, K_{23}\, \omega_1, \quad (4.12)$$

$$D_0\, K_{33} - D_1\, N_{23}$$
$$=\ K_{11}\, K_{22} - K_{23}^2 - K_{33}^2 + 2\, N_{11}\, N_{33} - N_{33}^2 + a_1\, N_{32} - N_{23}^2 - 2\, K_{23}\, \omega_1, \quad (4.13)$$

$$D_0\, a_1 - D_1\, K_{11} = 0, \quad (4.14)$$

$$D_0\, \omega_1 + D_1\, N_{11} = -a_1\, N_{11} - K_{11}\, \omega_1. \quad (4.15)$$

The tetrad quasi-constraint equations are:

$$D_1\, (K_{22} + K_{33}) = 2\, K_{23}\, N_{33} + K_{11}\, (N_{32} - N_{23}) + K_{22}\, N_{23} - K_{33}\, N_{32}, \quad (4.16)$$

$$D_1\, (N_{32} - N_{23})$$
$$= K_{11}\, (K_{22} + K_{33}) - K_{23}^2 + K_{22}\, K_{33} + N_{23}\, (N_{32} - N_{23}) - N_{33}^2 - N_{32}^2. \quad (4.17)$$

4.4 True Evolution and Constraint Equations

Applying Eq. (3.55) to the 1D plane wave problem at hand gives the set of coupled partial differential equations which can be numerically integrated. These true evolution equations are

$$D_0\,\mathbf{q} + \boldsymbol{C}^1\,\lambda_1^x\,\partial_x\,\mathbf{q} = \mathbf{S}, \tag{4.18}$$

where

$$\mathbf{q} = (N_{23},\ N_{32},\ N_{11},\ N_{33},\ a_1,\ K_{23},\ K_{11},\ K_{22},\ K_{33},\ \omega_1), \tag{4.19}$$

$$\boldsymbol{C}^1 = \frac{1}{1-A_1^2}\begin{bmatrix} -A_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & -A_1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -A_1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -A_1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -A_1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -A_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -A_1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -A_1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -A_1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -A_1 \end{bmatrix}. \tag{4.20}$$

The source, S, is given in Appendix D. Note that although we are using a non-diagonal metric, the source terms are fairly simple.

The true momentum constraint equation is

$$\begin{aligned}\lambda_1^x\,\partial_x(K_{22}+K_{33}) &= 2\,K_{23}\,N_{33} - K_{11}\,N_{23} + K_{22}\,N_{23} + K_{11}\,N_{32} - K_{33}\,N_{32} \\ &+ A_1\,(K_{22}^2 + K_{23}^2 + K_{33}^2 + K_{22}\,K_{33} + N_{33}^2 + a_1\,N_{23} - a_1\,N_{32} + N_{23}\,N_{32}),\end{aligned} \tag{4.21}$$

and the true energy constraint equation is

$$\lambda_1^x\,\partial_x\,(N_{32}-N_{23}) = K_{11}\,K_{22}-K_{23}^2+K_{11}\,K_{33}+K_{22}\,K_{33}-N_{33}^2-N_{23}^2+N_{23}\,N_{32}-N_{32}^2$$
$$-\,A_1\,(a_1\,K_{22}+a_1\,K_{33}+2\,K_{23}\,N_{33}+K_{22}\,N_{23}-K_{33}\,N_{32}). \tag{4.22}$$

The eigensystem of the characteristic matrix, C^1, is straightforward. It consists of five right-going eigenmodes:

$$N_{23}-K_{33},\quad N_{32}+K_{22},\quad N_{11}+w_1,\quad N_{33}+K_{23},\quad K_{11}-a_1, \tag{4.23}$$

with the degenerate eigenvalue $1/(1+A_1)$, and five left-going eigenmodes:

$$N_{23}+K_{33},\quad N_{32}-K_{22},\quad N_{11}-w_1,\quad N_{33}-K_{23},\quad K_{11}+a_1, \tag{4.24}$$

with the degenerate eigenvalue $-1/(1-A_1)$. The eigenmodes travel along the light cones with local coordinate speeds s_R and s_L:

$$s_R(x,t)=\frac{\lambda_1^x\,\alpha}{1+A_1},\qquad s_L(x,t)=-\frac{\lambda_1^x\,\alpha}{1-A_1}. \tag{4.25}$$

Since the eigensystem has a complete set of eigenvectors with real eigenvalues, the set of partial differential equations is hyperbolic, as defined by LeVeque [48]. Note that the eigenmodes in the EW system have constant coefficients of ± 1 and are thus much simpler than those presented in Chapter 2, Sec. 2.7.4.

Insight is gained by relating the tetrad variables to the $3+1$ variables presented in Chapter 2, Sec. D. Taking linear combinations of the eigenmodes in Eqs. (4.23) and (4.24), we obtain the following eigenmodes of the characteristic matrix:

$$[N_{32}+N_{23}+K_{22}-K_{33}],\quad [N_{32}-N_{23}+K_{22}+K_{33}]\quad \textit{right-going e-modes}, \tag{4.26}$$

$$[N_{32}+N_{23}-(K_{22}-K_{33})],\quad [N_{32}-N_{23}-(K_{22}+K_{33})]\quad \textit{left-going e-modes}. \tag{4.27}$$

Eigenmodes of the characteristic matrix in the modified Bona-Massó formalism with $n=0$ are:

$$\left[\frac{D_y^y-D_z^z}{\sqrt{h_{xx}}}+K_y^y-K_z^z\right],\quad \left[\frac{D_y^y+D_z^z}{\sqrt{h_{xx}}}+K_y^y+K_z^z\right]\quad \textit{right-going e-modes}, \tag{4.28}$$

$$\left[\frac{D_y^y - D_z^z}{\sqrt{h_{xx}}} - (K_y^y - K_z^z)\right], \quad \left[\frac{D_y^y + D_z^z}{\sqrt{h_{xx}}} - (K_y^y + K_z^z)\right] \quad \textit{left-going e-modes.} \quad (4.29)$$

(See Sec.2.7.4).

Recall that in the $3+1$ formulations, the solution to the full nonlinear set of equations, with the lapse equal to one and the shift to zero, has:

$$\frac{D_y^y \pm D_z^z}{\sqrt{h_{xx}}} = (K_y^y \pm K_z^z) \quad \textit{right-going waves of non-linear soln,} \quad (4.30)$$

$$\frac{D_y^y \pm D_z^z}{\sqrt{h_{xx}}} = -(K_y^y \pm K_z^z) \quad \textit{left-going waves of non-linear soln.} \quad (4.31)$$

To obtain a corresponding solution in the EW formulation, we first give the relations between the tetrad and 3+1 variables, derived using a diagonal metric with $A_1 = 0$:

$$N_{23} = -D_y^y/\sqrt{h_{xx}}, \quad N_{32} = D_z^z/\sqrt{h_{xx}}, \quad K_{11} = -K_x^x, \quad K_{22} = -K_y^y, \quad K_{33} = -K_z^z. \quad (4.32)$$

The solution to the nonlinear equations in the EW formulation has:

$$N_{32} + N_{23} = K_{22} - K_{33}, \quad N_{32} - N_{23} = -(K_{22} + K_{33}), \quad N_{33} = K_{23}$$

$$\textit{right-going waves of non-linear soln,} \quad (4.33)$$

$$N_{32} + N_{23} = -(K_{22} - K_{33}), \quad N_{32} - N_{23} = K_{22} + K_{33}, \quad N_{33} = -K_{23}$$

$$\textit{left-going waves of non-linear soln.} \quad (4.34)$$

It is evident that in the $3+1$ formulation, the eigenmodes of the characteristic matrix involving the constraint quantities travel in the same direction as the waves in the full nonlinear equations which involve the constraint quantities. However, in the EW tetrad formulation, the "constraint" eigenmodes of the characteristic matrix travel in the opposite direction from the propagation of the constraint quantities in the full nonlinear solution.

The cross polarized waves, N_{33} and K_{23}, are included in Eqs. (4.33) and (4.34). In the linearized approximation, they contain contributions from the yz parts of the non-diagonal metric.

4.5 Numerical Methods

The true evolution equations, as expressed succintly in Eq. 4.18, are solved using a Strang-split method, as in Chapter 2. Recall that in the Strang-split method, the transport equations (setting the right hand side of Eq. 4.18 to zero), and the integration of the source terms (assuming zero flux) are treated separately. Since the hyperbolic system of equations in Section 4.4 is not in the flux-conservative form of the $3+1$ formulations presented in Chapter 2, the numerical methods developed in that chapter do not apply.

We found it best to evolve the eigenmodes of the hyperbolic system as advection equations with spatially varying velocities. This is done in terms of values at points rather than cell averages. The equations are solved using the method of characteristics [48].

First order accurate expressions for the two average wave speeds over the time step are calculated. We assume that the variables are a linear function of x. The characteristic tracing procedure we use to find the eigenmode at the end of the time step amounts to using a weighted average of the Lax-Wendroff and Beam-Warming schemes. If at a given grid point, we define e_{bw} to be a temporarily updated eigenmode calculated by Beam-Warming, e_{lw} a temporarily updated eigenmode calculated by Lax-Wendroff, and s_{av} the average speed, then the weighted average is:

$$e_{bw} \times \frac{|s_{av}|\, dt}{dx} + e_{lw} \times \left(1 - \frac{|s_{av}|\, dt}{dx}\right). \tag{4.35}$$

The ODE's in the source integration are solved by second order Runge-Kutta. The spatial integrations for the initial and boundary conditions are solved by Predictor-Corrector methods, iterated to convergence.

4.6 Evolution Equations for A_1 and λ_1^x

Evolving the tetrad vector components λ_1^k instead of the metric terms h_{xx}, h_{yy}, and h_{zz} significantly simplifies the equations, and eliminates the need to take inverse metric functions, which can be messy for non-diagonal metrics. Since we have zero shift in this application, the evolution equations for A_1 and λ_1^x are ODE's:

$$\partial_t A_1 = \alpha[-\lambda_1^x \partial_x (\ln \alpha) + a_1 - K_{11} A_1]. \tag{4.36}$$

$$\begin{aligned}
&\partial_t \lambda_1^x = -\alpha K_{11} \lambda_1^x, \\
&\partial_t \lambda_2^z = \alpha [-K_{22} \lambda_2^z + (\omega_1 - K_{23}) \lambda_3^z], \quad \partial_t \lambda_3^y = \alpha [-K_{33} \lambda_3^y - (\omega_1 + K_{23}) \lambda_2^y], \\
&\partial_t \lambda_2^y = \alpha [-K_{22} \lambda_2^y + (\omega_1 - K_{23}) \lambda_3^y], \quad \partial_t \lambda_3^z = \alpha [-K_{33} \lambda_3^z - (\omega_1 + K_{23}) \lambda_2^z].
\end{aligned} \tag{4.37}$$

4.7 Boundary Conditions

All the variables are quadratically extrapolated to a ghost point at each boundary. Then, the constraint quantities $N_{32} - N_{23}$ and $-(K_{22} + K_{33})$ are corrected by numerical integration of the constraint Eqs. 4.22 and 4.21 from the edge physical point to the ghost point (at each boundary). The procedure is exactly as described in Sec. 2.9.2.

Note that the boundary conditions might be improved by using the constraints to reset the boundary values of the incoming "constraint" mode at the edge physical point (at each boundary), as is done in Sec. 5.9.

4.8 Coordinate Conditions

The shift is zero in all the calculations in this chapter. The lapse is set equal to 1 initially. It remains equal to its initial value throughout the numerical simulation.

4.9 Results

4.9.1 Initial Conditions

Our initial conditions for the plane wave problem in the EW formulation (1) allow for general polarization, (2) use the tetrad vector components λ_a^k as variables instead of the metric coefficients, and (3) have a different waveform for the plus polarized waves than our initial conditions in Chapter 2 (compare Figs. 2.1, solid line, and 4.2(a), green line). $N_{23}+N_{32}$, N_{33}, N_{11}, $K_{22}-K_{33}$, K_{23}, K_{11}, a_1, and ω_1 are free initial data. The constrained quantities are $N_{32}-N_{23}$ and $K_{22}+K_{33}$, which must satisfy the energy and momentum constraint equations, and the tetrad vector components, λ_a^k, which are constrained by Eqs. (3.66). The congruence is taken to be orthogonal to the initial hypersurface, so $A_1=0$ initially. We also take $K_{11}=\omega_1=N_{11}=a_1=0$ initially.

Symmetric boundary conditions for the integration of the momentum and energy constraint Eqs. 4.21 and 4.22 are chosen, giving flat space between the two waves initially. This means that $N_{32}-N_{23}$ and $K_{22}+K_{33}$ vanish initially between the waves. Constraint equations for the transverse contravariant tetrad vector components are obtained from Eqs. (3.66). They are:

$$\partial_x \lambda_2^y = \frac{1}{\lambda_1^x}\left[\lambda_2^y N_{23} - \lambda_3^y (N_{11}-N_{33})\right], \tag{4.38}$$

$$\partial_x \lambda_3^z = \frac{1}{\lambda_1^x}\left[-\lambda_3^z N_{32} + \lambda_2^z (N_{11}+N_{33})\right], \tag{4.39}$$

$$\partial_x \lambda_2^z = \frac{1}{\lambda_1^x}\left[\lambda_2^z N_{23} - \lambda_3^z (N_{11}-N_{33})\right], \tag{4.40}$$

$$\partial_x \lambda_3^y = \frac{1}{\lambda_1^x}\left[-\lambda_3^y N_{32} + \lambda_2^y (N_{11}+N_{33})\right]. \tag{4.41}$$

The coordinate x-axis and the tetrad 1-axis, along which the plane waves travel, coincide because of the planar symmetry. The scaling of the $x-$ coordinate is chosen to equal proper distance initially, so that $\lambda_1^x=1$. Because the metric is

explicitly flat at the midpoint of the numerical grid, $\lambda_2^y = \lambda_3^z = 1$, and $\lambda_2^z = \lambda_3^y = 0$ at the center of the grid. Then, Eqs. (4.38), (4.39), (4.40), (4.41), are integrated from the center out, in both directions, to give initial values for λ_a^k.

The variables N_{33}, N_{11}, K_{23}, K_{11}, a_1, and ω_1, and the combinations $N_{23} + N_{32}$ and $K_{22} - K_{33}$ are freely specifiable. The plus polarized waves are given by

$$\frac{1}{2}(N_{23} + N_{32}) = \sum_{i=1}^{2} P_i \cos^2\left[\frac{\pi}{2}\frac{(x - x_{0i})}{w_i}\right] \sin\left[k_i(x - x_{0i}) + \delta_{Pi}\right], \tag{4.42}$$

for $-w_i < (x - x_{0i}) < w_i$, and zero outside that range. One wave is initially on the left and moving to the right, with $[K_{22} - K_{33}]_1 = [N_{23} + N_{32}]_1$. The other wave is initially on the right and moving to the left, with $[K_{22} - K_{33}]_2 = -[N_{23} + N_{32}]_2$. The parameters for $0 \leq x \leq 20$ are $w_i = 4.0$, $k_i = 1.6$, $P_1 = P_2 = 0.128$, $x_{01} = 6.0$, $x_{02} = 14.0$, $\delta_{P1} = 0$, and $\delta_{P2} = \pi$. Graphs of $N_{23} + N_{32}$ and $K_{22} - K_{33}$ for plane (plus) polarized waves at $t = 0$ are shown in green in Fig. 4.2, parts (a) and (b), for $0 \leq x \leq 10$. The initial form of the transverse traceless quantity $(D_y^y - D_z^z)/(2\sqrt{h_{xx}})$ in Chapter 2 is the x derivative of the initial form of $(N_{23} + N_{32})/2$ in Eq. (4.42). The average amplitude of the initial waveform of $(N_{23} + N_{32})/2$ is about 8 percent lower than the average amplitude of the initial waveform of $(D_y^y - D_z^z)/(2\sqrt{h_{xx}})$ in Chapter 2. Recall that the upper limit derived in Sec. 2.10.1 for the amplitude of $(D_y^y - D_z^z)/(2\sqrt{h_{xx}})$ is 0.11, and a value of 0.08 is used for those calculations. An 8 percent reduction of this value is ≈ 0.074.

The cross polarized waves are:

$$N_{33} = -\sum_{i=1}^{2} C_i \cos^2\left[\frac{\pi}{2}\frac{(x - x_{0i})}{w_i}\right] \sin\left[k_i(x - x_{0i}) + \delta_{Ci}\right], \tag{4.43}$$

for $-w_i < (x - x_{0i}) < w_i$, and zero outside that range. One cross polarized wave is initially on the left and moving to the right, with $[K_{23}]_1 = [N_{33}]_1$, and the other is initially on the right and moving to the left, with $[K_{23}]_2 = -[N_{33}]_2$. Graphs of N_{33} and K_{23} at $t = 0$ for $0 \leq x \leq 10$ are shown in green in Fig. 4.5, with $C_1 = C_2 = 0.128$, $\delta_{C1} = 0$, $\delta_{C2} = \pi$.

To create circularly polarized waves, the plus polarized waves of Eq. (4.42) are superimposed with the cross polarized waves of Eq. (4.43). The parameter values are $w_i = 4.0$, $k_i = 1.6$, $x_{01} = 6.0$, $x_{02} = 14.0$, $P_1 = P_2 = C_1 = C_2 = 0.128/\sqrt{2}$, $\delta_{P1} = 0$, $\delta_{P2} = \pi$, $\delta_{C1} = \pi/2$ and $\delta_{C2} = \pi/2$. Graphs of $N_{23} + N_{32}$, $K_{22} - K_{33}$, N_{33}, and K_{23} for circularly polarized waves at $t = 0$ are shown in green in Fig. 4.8, parts (a), (b), (e), and (f), for $0 \leq x \leq 10$.

These initial conditions, for both plane and circularly polarized waves, are symmetric (or antisymmetric) about $x = 10$, (depending on the variables), and symmetry (or antisymmetry) is preserved throughout the evolution. Hence, our figures only show the range $0 \leq x \leq 10$.

4.9.2 Evolution of Physical and Constraint Quantities

Figs. 4.2, 4.4, 4.5, 4.6, 4.8, and 4.10 show the evolution of the physical and constraint quantities for plane polarized waves (with both plus and cross polarizations), and circularly polarized waves (both the plus and cross parts of wave) both before and after the collision. The coordinate speeds of propagation of these quantities for plane (plus) polarization are similar to those for the metric derivatives presented in Chapter 2.

Figs. 4.2, 4.5, and 4.8 show that the results are consistent with the purely right-going waves of Eq. (4.33) prior to the meeting of the right- and left-going waves. Figs. 4.4, 4.6, and 4.10 show that at late times, as the left-going waves are approaching the left edge of the grid, the physical quantities $N_{32} + N_{23}$, $K_{22} - K_{33}$, N_{33}, and K_{23} hold approximately to the form of Eq. (4.34), but the constraint quantities $N_{32} - N_{23}$ and $K_{22} + K_{33}$ do not. This is because after the collision, there is an admixture of left and right going eigenmodes on each side of the grid. The admixture of eigenmodes moving toward the center of the grid after the wave collision is of a higher percentage with the "constraint" eigenmodes than with the "physical" eigenmodes.

4.9.3 *Evolution of* N_{11} *and* ω_1

Although N_{11} and ω_1 are initially zero, they become non-zero in the collision of circularly polarized waves, as shown in Figs. 4.11 and 4.12. Recall that N_{11} describes the twist of the spatial triads along the 1 direction, and ω_1 the angular velocity of the spatial triad with respect to Fermi propagated axes. It is apparent from these figures that N_{11} and ω_1 get large as the waves collide and the spacetime stretches. The twist and angular velocity of the spatial triads, induced by the collision of the waves, propagate towards the edge of the grid as the waves separate and move away from each other.

4.9.4 *Comparison of Accuracy in EW and* $3+1$ *Formulations*

The propagation of the energy and momentum constraint errors, both before and after the collision, and for plane and circularly polarized waves, is shown in Figs. 4.1, 4.3 4.7, and 4.9.

Comparing the green line (at $t = 8$) in Fig. 4.3 with the solid line in Fig. 2.11 gives us a comparison of the accuracy as a function of x of the EW formulation with the $3+1$ formulation in Chapter 2 which gave the highest accuracy. We see that the peak in the $3+1$ formulation is almost two orders of magnitude higher than that in the EW formulation. The 1-norm energy constraint error for this $3+1$ formulation at $t = 12$ and for a 500 grid cell resolution, is approximately 3×10^{-4} (see Fig. 2.3). The 1-norm energy constraint error in the EW formulation at $t = 12$ for plus polarization and a 500 point resolution is 4.9×10^{-6}. This is a 60 fold increase in accuracy.

Table 4.1 shows that the 1-norm energy and momentum constraint errors at $t = 12$, for plus and circularly polarized waves, converge at a rate greater than or equal to second order for three different grid resolutions.

4.9.5 Constraint Error Propagation

As discussed in Sec. 2.10.2, the energy and momentum constraint errors propagate at the same velocity as the "constraint" eigenmodes, since the principal parts of the energy and momentum constraint equations are the derivatives of the constraint quantities.

Fig. 4.1 is consistent with Eq. (4.27), in that the energy and momentum constraint errors are of opposite sign and are travelling to the left, while the constraint quantities in the nonlinear solution are travelling to the right (see Fig. 4.2c and d).

However, this property becomes a numerical disadvantage when the waves are exiting the grid. As can be seen in Fig. 4.3, at $t = 12$ (the red line), the errors are propagating into the grid, and there is a corresponding increase in constraint errors at the boundaries.

With circularly polarized waves, it is somewhat more difficult to determine the direction of propagation of the constraint errors before the collision in Fig. 4.7, since the errors are generated more uniformly. However, one can see that the small peaks in the region $0 \leq x \leq 4$ are moving out of the grid. After the collision, in Fig. 4.9, we see the same behavior as with the plane polarized waves, where the errors appear to be entering the grid around $t = 12$, and getting larger at the boundaries.

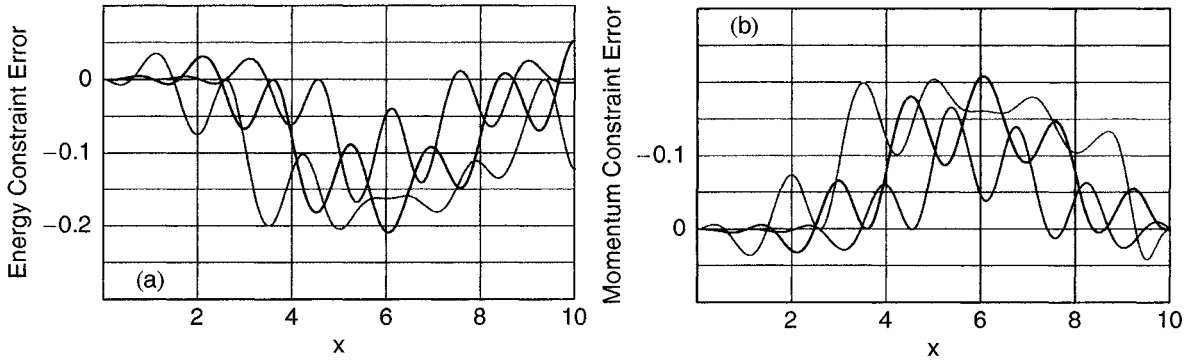


Figure 4.1: Propagation of constraint errors for plane (plus) polarized waves *before* the collision. The green line is at $t = 1$, the blue line at $t = 2$, and the red line at $t = 3$. The vertical axis is scaled by 10^{-5}. Notice the constraint errors are propagating out of the grid.

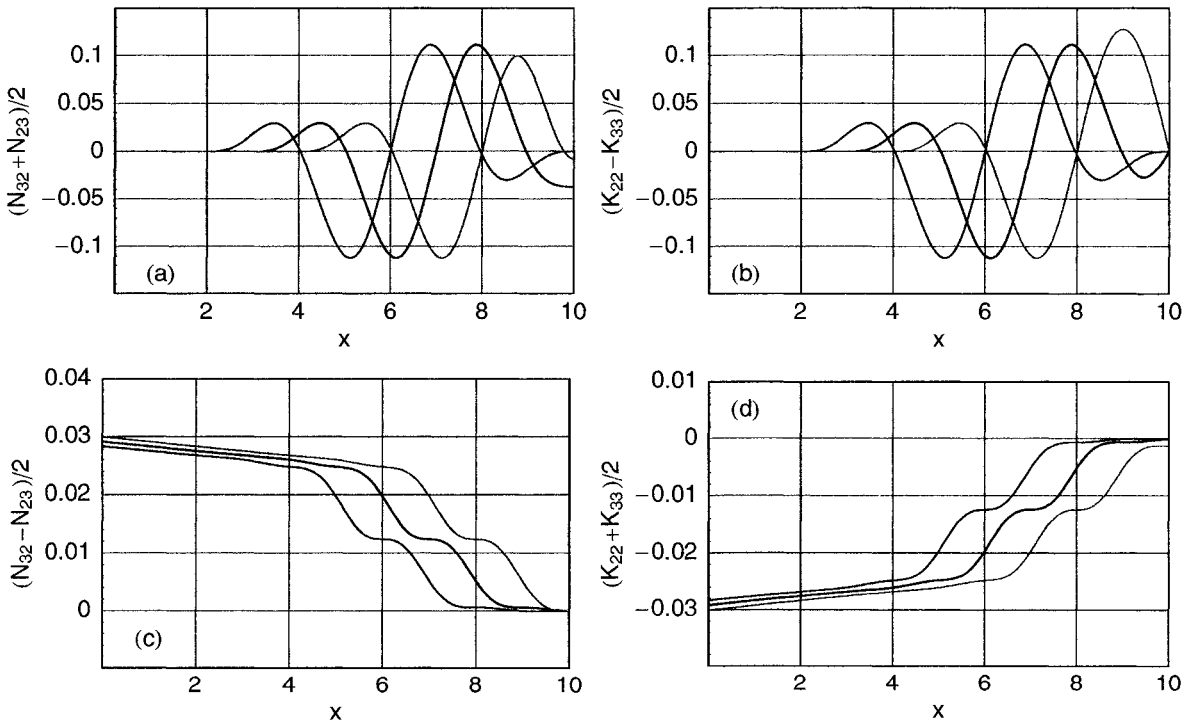


Figure 4.2: Evolution of physical and constraint quantities for plane (plus) polarized waves *before* the collision. The green line is at $t = 0$, the blue line at $t = 1$, and the red line at $t = 2$. Evaluated with a grid resolution of 500 points. Note that $x = 10$ is the center of the grid.

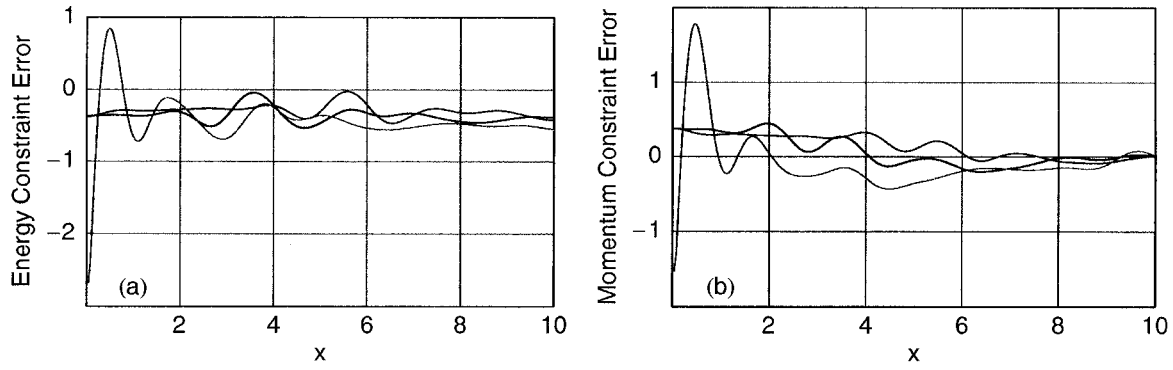


Figure 4.3: Propagation of constraint errors for plane (plus) polarized waves *after* the collision. The green line is at $t = 8$, the blue line at $t = 10$, and the red line at $t = 12$. The vertical axis is scaled by 10^{-5}.

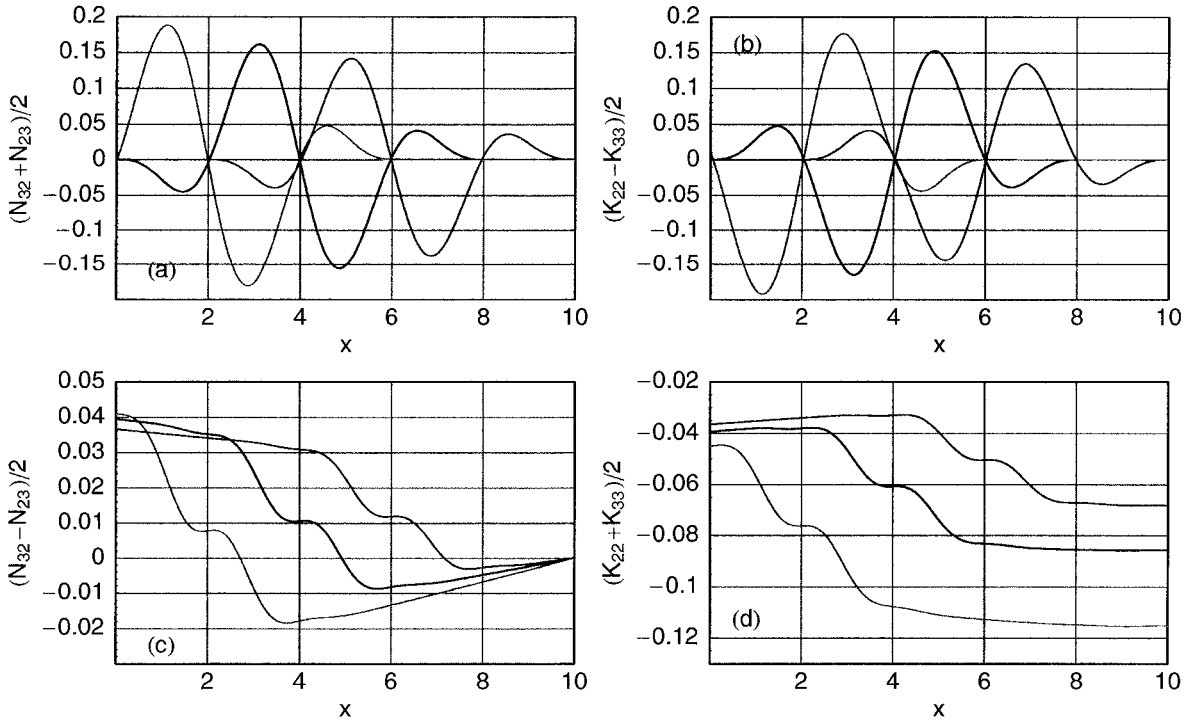


Figure 4.4: Evolution of physical and constraint quantities for plane (plus) polarized waves *after* the collision. The green line is at $t = 8$, the blue line at $t = 10$, and the red line at $t = 12$. Evaluated with a grid resolution of 500 points.

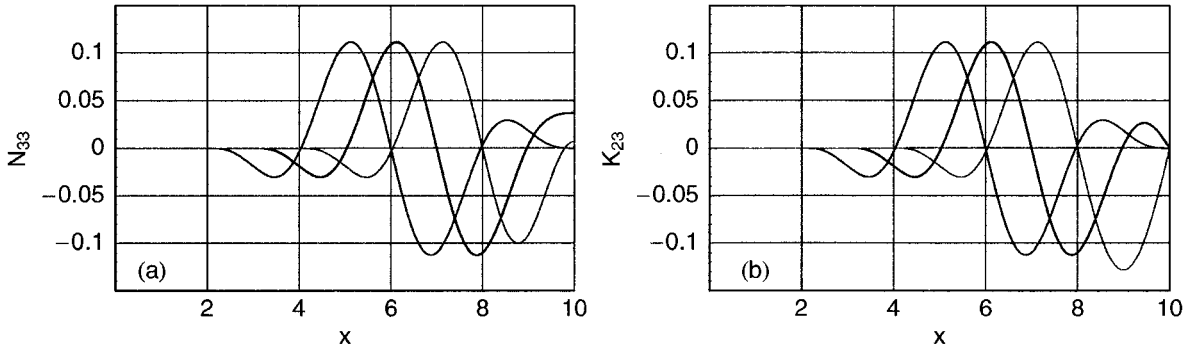


Figure 4.5: Evolution of plane (cross) polarized waves *before* the collision. The green line is at $t = 0$, the blue line at $t = 1$, and the red line at $t = 2$. Evaluated with a grid resolution of 500 points.

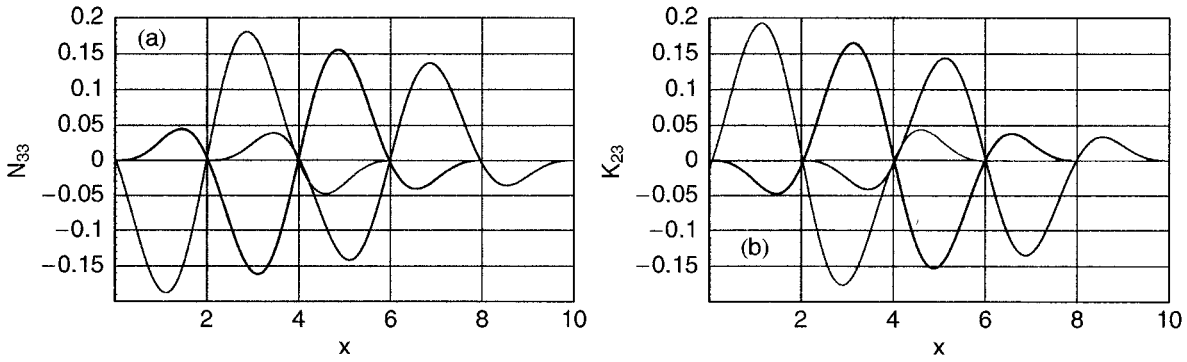


Figure 4.6: Evolution of plane (cross) polarized waves *after* the collision. The green line is at $t = 8$, the blue line at $t = 10$, and the red line at $t = 12$. Evaluated with a grid resolution of 500 points.

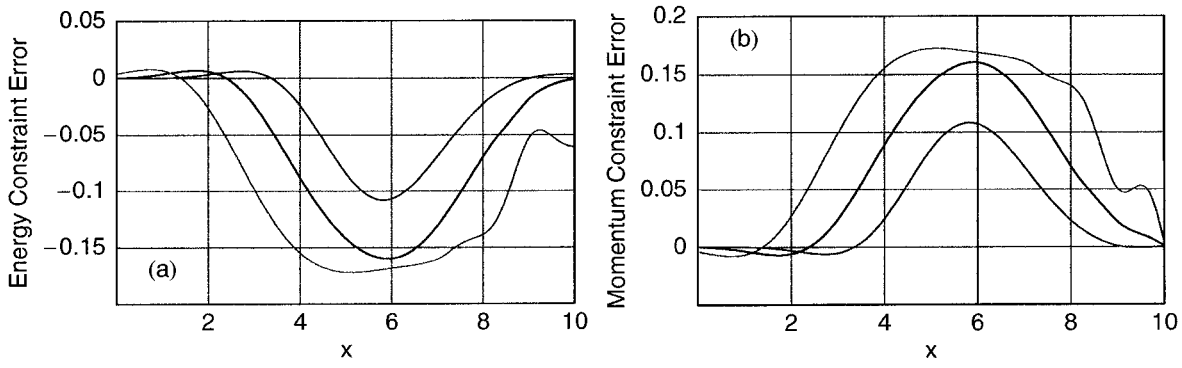


Figure 4.7: Propagation of constraint errors with circular polarization *before* the collision. The green line is at $t = 1$, the blue line at $t = 2$, and the red line at $t = 3$. The vertical axis is scaled by 10^{-5}.

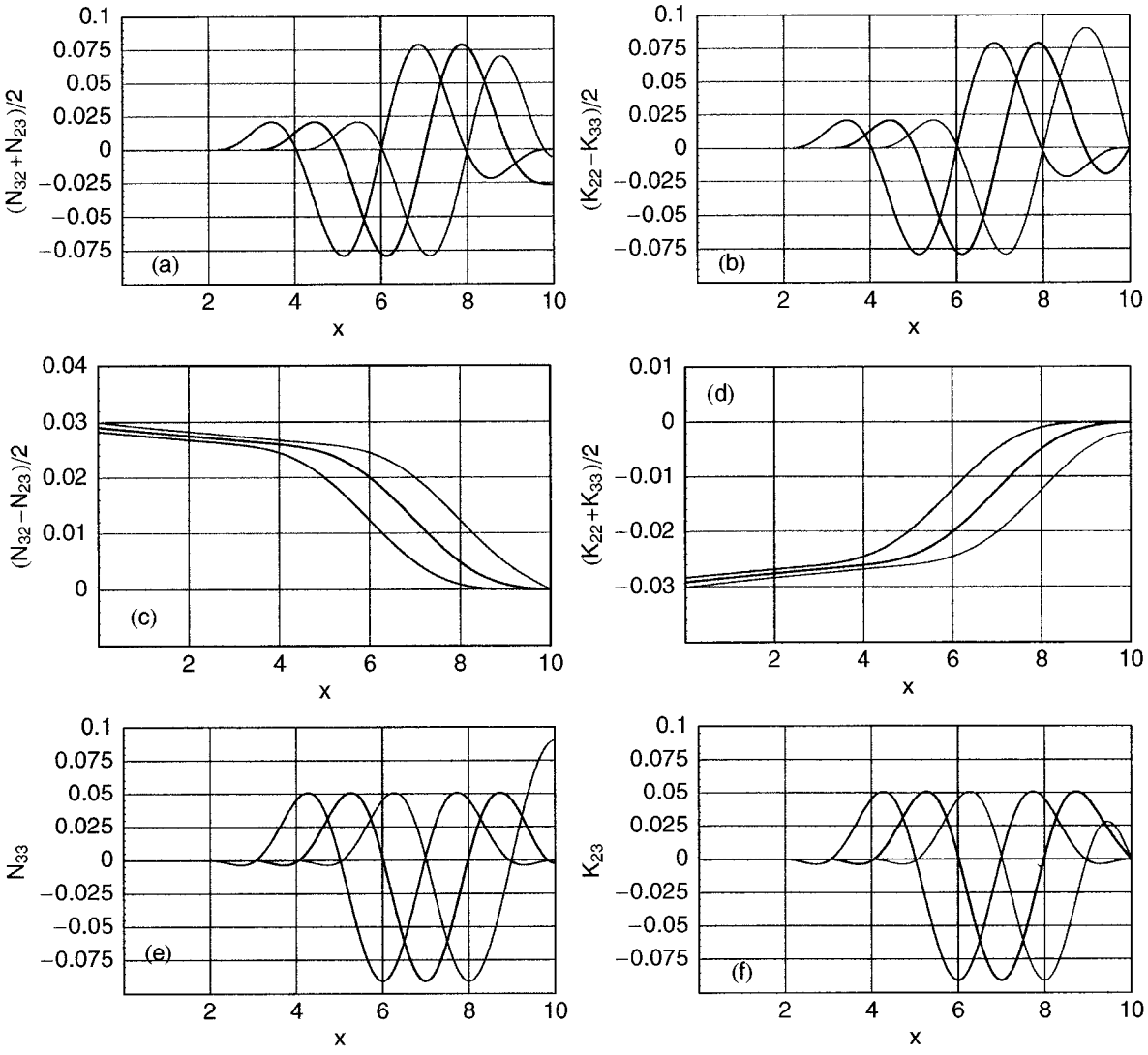


Figure 4.8: Evolution of physical (plus and cross parts of the wave) and constraint quantities with circular polarization *before* the collision. The green line is at $t = 0$, the blue line at $t = 1$, and the red line at $t = 2$.

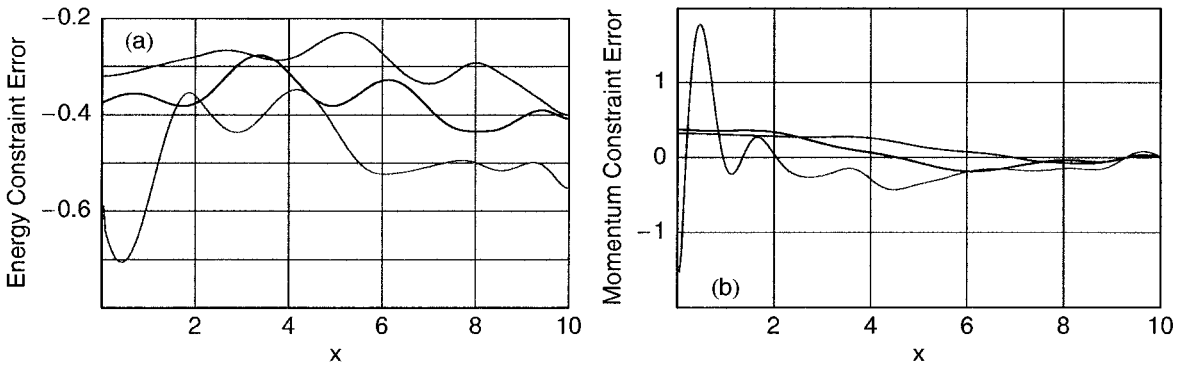


Figure 4.9: Propagation of constraint errors with circular polarization *after* the collision. The green line is at $t = 8$, the blue line at $t = 10$, and the red line at $t = 12$. The vertical axis is scaled by 10^{-5}.

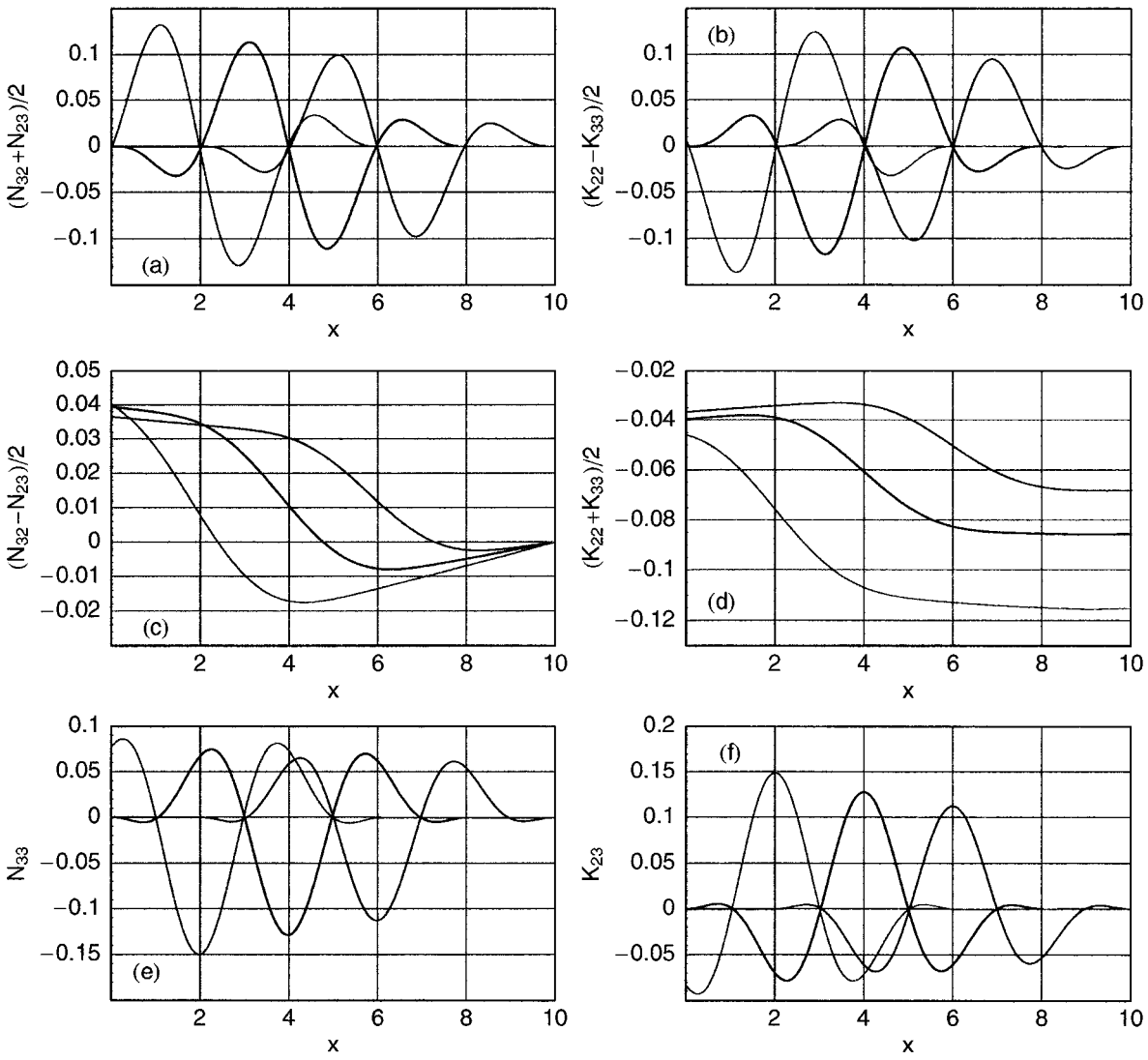


Figure 4.10: Evolution of physical (plus and cross parts of the wave) and constraint quantities with circular polarization *after* the collision. The green line is at $t = 8$, the blue line at $t = 10$, and the red line at $t = 12$. Evaluated with a grid resolution of 500 points.

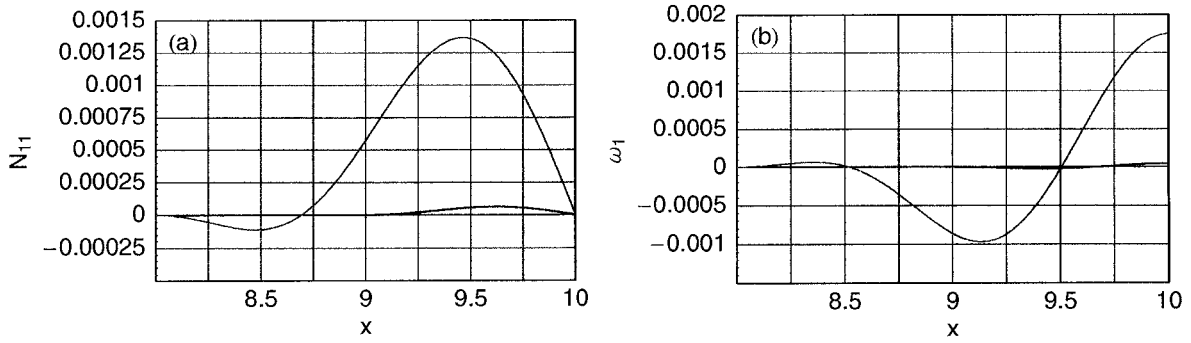


Figure 4.11: Evolution of (a) N_{11} and (b) ω_1 *before* the collision of circularly polarized waves. The green line is at $t = 0$, the blue line at $t = 1$, and the red line at $t = 2$. Evaluated with a grid resolution of 500 points. Note that the graph has been expanded to show the region $8 \leq x \leq 10$. N_{11} and ω_1 are zero at these times for $x < 8$.

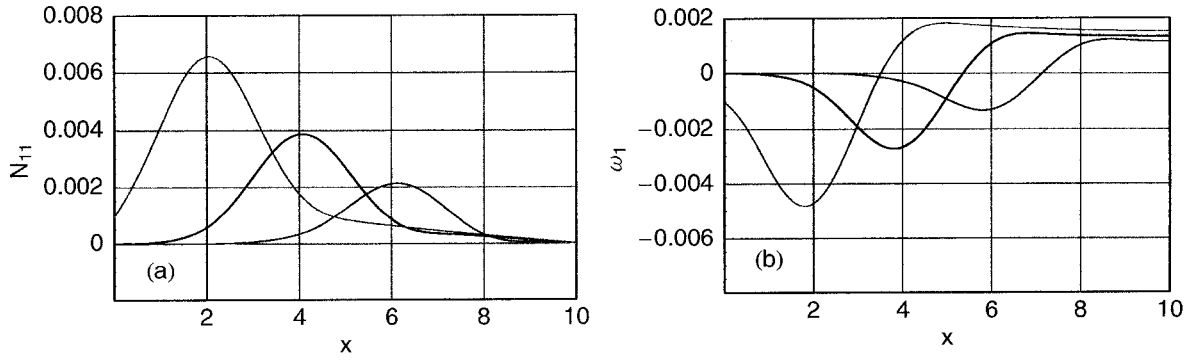


Figure 4.12: Evolution of (a) N_{11} and (b) ω_1 *after* the collision of circularly polarized waves. The green line is at $t = 8$, the blue line at $t = 10$, and the red line at $t = 12$. Evaluated with a grid resolution of 500 points.

Table 4.1: Convergence of energy and momentum constraint errors at $t = 12$, for both plus and circularly polarized waves. We take the 1-norms of the results obtained using the indicated grid resolutions.

Grid resolutions	Energy constraint		Momemtum constraint	
	plus	circular	plus	circular
501/1001	6.83	7.61	5.73	7.01
1001/2001	5.69	7.30	4.31	5.13
2001/4001	4.80	6.44	3.93	4.31

4.10 Discussion

The most striking result of this chapter is that presented in Sec. 4.9.4; namely, that the constraint errors in the simulation of plane (plus polarized) waves using the EW tetrad formulation are 40 to 70 times lower than the constraint errors shown in Fig. 2.3, obtained in comparable simulations using $3 + 1$ formulations. This result supports one of the main conclusions of Chapter 2, that in cases where the dominant source of errors is associated with features in the nonlinear solution for the constraint quantities, accuracy can be significantly increased by changing the speeds of the "constraint" eigenmodes to be different from the speeds at which these features propagate. In the EW formulation, the "constraint" eigenmodes travel outward and exit the grid as the nonlinear waves (and thus the constraint quantities) move into the center of the grid and collide. Thus, the propagation of the constraint errors is as far removed as possible from the dominant source of errors during the collision of the waves, and there is a consequent increase in accuracy. A similar, large improvement in accuracy can be obtained using the $3 + 1$ ADM and reset-BM formulations presented in Chapter 2, with a value of the energy constraint coefficient, n, equal to 2 [12].

Chapter 5

EW TETRAD APPLICATION 2: SCHWARZSCHILD BLACK HOLES IN SPHERICAL SYMMETRY

5.1 Introduction

The most basic test of any code for vacuum black hole numerical relativity is the Schwarzschild black hole in spherically symmetric spacetime. The challenge of this testbed is the black hole singularity, and the extreme spacetime warping in the proximity of the event horizon. The most succcessful approach to evolving a numerical grid in such a warped and physically singular spacetime has been to actually take advantage of the special nature of the event horizion. The event horizion is like a one-way membrane, through which information can enter (in the direction of the singularity), but not escape. Numerical relativists use this property to their advantage by truncating the numerical grid just inside the event horizon. The idea for this approach, which is called *excision*, is attributed to a suggestion by Unruh in 1984 [72]. The first application of excision to vacuum numerical relativity was accomplished by Seidel and Suen in 1992 [65] in spherically symmetric Schwarzschild spacetime. Since then, excision has been implemented successfully for $3+1$ formulations of the Einstein equations, both in spherically symmetric [7, 53, 62, 63, 35, 40, 38, 43] and in three-dimensional [39, 1, 3, 80, 64] black hole spacetimes.

To be useful for obtaining merger waveforms for binary black holes, a $3D$ numerical relativity code must run accurately and stably for at least $100M$. This was accomplished for the first time with the non-hyperbolic Baumgarte-

Shapiro-Shibata-Nakamura (BSSN) scheme [66, 14], which is based on a conformal decomposition of the metric. Using excision and a dynamic gauge condition, Alcubierre, Brügmann, Pollney, Seidel, and Takahashi [3] were able to accurately evolve 3D black holes to a few hundred M.

With excision comes the question of boundary conditions at the inner edge of the grid. This very question is a prime motivation for using hyperbolic formulations of the Einstein equations. All the information in hyperbolic formulations propagate from hypersurface to hypersurface along characteristics. If these characteristics lie along or within the light cone at the excision boundary of the numerical grid, then the information will flow out of the numerical grid and into the black hole. This is a great advantage because it eliminates the need for boundary conditions at the inner edge, and prevents the transmission of spurious signals into the grid.

Kidder, Scheel, and Teukolsky [39] have performed an exhaustive study of a parameterized set of hyperbolic formulations in a $3D$ black hole code, with excision. The parameters specify (1) multiples of the constraint equations which are added to the evolution equations, and (2) re-definition of the variables. They found that by changing these parameters, they could lengthen the stability of their runs. In a follow-up work [50], Lindblom and Scheel report that the primary factor which is affected by these parameters is the growth rate of analytic, but non-physical solutions to the Einstein equations. These non-physical solutions can be constraint violating or constraint satisfying, but in either case, if they grow exponentially, they are unstable. Lindblom and Scheel characterize the growth rates of these non-physical solutions via the energy norm. Finally, in 2002, Scheel, Kidder, Lindblom, Pfeiffer, and Teukolsky [64] achieve optimal stability by fine-tuning their parameter space according to knowledge obtained from the energy norms. Using a fully general $3D$ code and fixed coordinate conditions, they evolve a single static Schwarzschild black hole to $8000M$.

Yo, Baumgarte, and Shapiro [80] experiment with adding constraints to the evolution equations in the BSSN formulation. With a single, nonrotating black hole, excision, and no symmetry assumptions, they report stable and accurate running times to $1500M$.

In this chapter, we use both excision and a hyperbolic formulation (with information propagating out of the boundary at the inner edge) in our application of the EW tetrad formulation to the spherically symmetric Schwarzschild black hole. The EW spherically symmetric equations with Nester and Lorentz tetrad gauges are derived, coded, and tested, with both stationary and constant mean curvature (CMC) initial conditions. We have considered fixed, dynamic, and reset coordinate conditions, but have not yet implemented constraint addition in a generalizable way. Although accurate evolutions have been achieved with time independent initial conditions to $t = 40M$ with reset coordinate conditions, the desired long-term stability has not been achieved. For constant mean curvature initial conditions, the evolution of the congruence in the Nester gauge exhibits pathological behavior, a result which is not salvagable by modification of the EW equations or the coordinate conditions. Whether or not the Nester tetrad gauge condition can be used for 3D numerical relativity will only be determined by testing it with a greater range of initial conditions. The congruence evolution with the Lorentz gauge and CMC initial conditions is well-behaved, but also should be tested with a wider variety of initial conditions before any definitive statement can be made.

Our numerical simulations of colliding gravitational plane waves, presented in Chapters 2 and 4, show significant improvement in accuracy when the speeds of the "constraint" eigenmodes are adjusted to be different from the speeds of the dominant sources of errors, namely, features in the constraint quantities which are propagating at light speed. For our plane wave calculations, the energy and momentum constraint errors propagate at the speed of the "con-

straint" eigenmodes, which can be adjusted in the $3+1$ ADM or BM-reset formulations to be different from light speed by adding certain multiples of the energy constraint equation to the evolution equations. For the EW tetrad formulation applied to 1D nonlinear colliding plane waves, the "constraint" eigenmodes and constraint quantities are inherently travelling in opposite directions, and a large improvement in accuracy results.

For spherically symmetric spacetimes, the dominant source of errors is from the source terms, and thus does not propagate at characteristic speeds. Our results in this chapter indicate that there is a localized region of instability, right near the event horizon of the black hole. If the speeds of the "constraint" eigenmodes are changed so that any new errors propagating at these speeds are advected away from the dominant source of errors (the event horizon), then we expect an increase in accuracy and stability. We find that when we change the speeds of the "constraint" eigenmodes by adding constraint equations to the evolution equations, so that the "constraint" eigenmode speeds are relatively large in the vicinity of the event horizon, then stable and accurate numerical evolutions result, to at least $500M$. Since adding the constraint equations to the evolution equations changes the source terms as well as the characteristic speeds, more tests are required in order to determine exactly what is happening to cause such a large improvement over our results without constraint addition. Furthermore, the method of constraint addition presented in this thesis is specific to the spherically symmetric equations; we have not yet ascertained whether or not it is generalizable to $3D$. This is an important area of investigation for future work. Bardeen [12] has achieved similar long-term stability with a hypersurface orthogonal Einstein-Bianchi tetrad formulation, related to that presented by Choquet-Bruhat and York [22], in which the "constraint" eigenmodes are also strongly advected away from the event horizon and out of the numerical grid.

5.2 Variables

The 3D EW formalism presented in Chapter 3 is applied to 1D spherically symmetric Schwarzschild black holes. This application requires the construction of a globally well-behaved orthonormal tetrad field in spherical symmetry. Spherically symmetric spacetime is described by a series of nested two-spheres, so the most natural basis vectors for this spacetime are either orthogonal or tangential to the two-spheres. We choose our timelike congruence to be spherically symmetric and therefore orthogonal to the two-spheres. The timelike basis vectors, $\boldsymbol{e}_0$, point along this congruence. The spacelike basis vectors which are optimal for spherical symmetry are $\boldsymbol{e}_{\hat{r}}$, which is orthogonal both to $\boldsymbol{e}_0$ and to the two-spheres, and $\{\boldsymbol{e}_{\hat{\theta}},\ \boldsymbol{e}_{\hat{\phi}}\}$, which are tangent to the two-spheres. These "spherical" spacelike basis vectors are unsuitable for direct application to the EW tetrad formalism, however, because the degeneracies in $\{\boldsymbol{e}_{\hat{\theta}},\ \boldsymbol{e}_{\hat{\phi}}\}$ at the poles cause a singularity in the twist of the congruence worldlines. We construct a field of well-behaved Cartesian tetrads from the spherical tetrads by the same rotation from the spherical tetrad that applies in flat space, namely:

$$\boldsymbol{e}_1 = \cos\phi\ \sin\theta\ \boldsymbol{e}_{\hat{r}} + \cos\phi\ \cos\theta\ \boldsymbol{e}_{\hat{\theta}} - \sin\phi\ \boldsymbol{e}_{\hat{\phi}}, \tag{5.1}$$

$$\boldsymbol{e}_2 = \sin\phi\ \sin\theta\ \boldsymbol{e}_{\hat{r}} + \sin\phi\ \cos\theta\ \boldsymbol{e}_{\hat{\theta}} + \cos\phi\ \boldsymbol{e}_{\hat{\phi}}, \tag{5.2}$$

$$\boldsymbol{e}_3 = \cos\theta\ \boldsymbol{e}_{\hat{r}} - \sin\theta\ \boldsymbol{e}_{\hat{\theta}}. \tag{5.3}$$

The inverse transformation is:

$$\boldsymbol{e}_{\hat{r}} = \cos\phi\ \sin\theta\ \boldsymbol{e}_1 + \sin\phi\ \sin\theta\ \boldsymbol{e}_2 + \cos\theta\ \boldsymbol{e}_3, \tag{5.4}$$

$$\boldsymbol{e}_{\hat{\theta}} = \cos\phi\ \cos\theta\ \boldsymbol{e}_1 + \sin\phi\ \cos\theta\ \boldsymbol{e}_2 - \sin\theta\ \boldsymbol{e}_3, \tag{5.5}$$

$$\boldsymbol{e}_{\hat{\phi}} = -\sin\phi\ \boldsymbol{e}_1 + \cos\phi\ \boldsymbol{e}_2. \tag{5.6}$$

So, both the Cartesian and spherical spatial triads are orthogonal to $\boldsymbol{e}_0$, and can be transformed back and forth via spatial rotations in their local 3-space.

As in any Cauchy formulation for numerical relativity, the spacetime is sliced into a set of hypersurfaces on which the numerical integration takes place. As described in Chapter 3, the spatial triad vectors are decomposed into a timelike component parallel to the congruence, and a spacelike component tangent to the hypersurface:

$$\boldsymbol{e}_a = A_a\, \boldsymbol{e}_0 + B_a^k\, \frac{\partial}{\partial\, x^k}. \tag{5.7}$$

We know that $\boldsymbol{e}_{\hat{\theta}}$ and $\boldsymbol{e}_{\hat{\phi}}$ are tangent to the hypersurfaces, so $A_{\hat{\theta}} = A_{\hat{\phi}} = 0$. Thus, A_a only has one degree of freedom, in the $\boldsymbol{e}_{\hat{r}}$ direction. $A_{\hat{r}}$ is the velocity of a tetrad observer in the $\boldsymbol{e}_{\hat{r}}$ direction with respect to an observer at rest in the $t =$ constant hypersurface. We can now write

$$\boldsymbol{e}_{\hat{r}} = A_{\hat{r}}\, \boldsymbol{e}_0 + B_{\hat{r}}^r\ \partial_r, \qquad \boldsymbol{e}_{\hat{\theta}} = B_{\hat{\theta}}^{\theta}\ \partial_\theta, \qquad \boldsymbol{e}_{\hat{\phi}} = B_{\hat{\phi}}^{\phi}\ \partial_\phi, \tag{5.8}$$

where r is the radial coordinate in the hypersurface. The metric in the hypersurface is perpendicular to the two-spheres, so $\boldsymbol{B}_{\hat{r}} \propto \boldsymbol{e}_r$. We call the proportionality factor $e^{-\lambda}$, with λ an arbitrary function of r. $\boldsymbol{B}_{\hat{\theta}}$ and $\boldsymbol{B}_{\hat{\phi}}$ are unit vectors aligned with $\boldsymbol{e}_{\hat{\theta}}$ and $\boldsymbol{e}_{\hat{\phi}}$; their normalization factors are defined by the metric. Hence, we have

$$B_{\hat{r}}^r = e^{-\lambda}, \qquad B_{\hat{\theta}}^{\theta} = \frac{1}{R}, \qquad B_{\hat{\phi}}^{\phi} = \frac{1}{R\, \sin\theta}, \tag{5.9}$$

with R the circumferential radius of the black hole.

Eq. (3.12) gives the directional derivatives along the spatial triad directions as

$$D_a = A_a\, D_0 + B_a^k\, \frac{\partial}{\partial\, x^k}. \tag{5.10}$$

In the Cartesian bases, this gives

$$D_1 = \cos\phi\ \sin\theta\ D_{\hat{r}} + \cos\phi\ \cos\theta\ D_{\hat{\theta}} - \sin\phi\ D_{\hat{\phi}}, \tag{5.11}$$

$$D_2 = \sin\phi\ \sin\theta\ D_{\hat{r}} + \sin\phi\ \cos\theta\ D_{\hat{\theta}} + \cos\phi\ D_{\hat{\phi}}, \tag{5.12}$$

$$D_3 = \cos\theta\ D_{\hat{r}} - \sin\theta\ D_{\hat{\theta}}, \tag{5.13}$$

with

$$D_{\hat{r}} = e^{-\lambda}\ \partial_r + A_{\hat{r}}\ D_0, \qquad D_{\hat{\theta}} = \frac{1}{R}\ \partial_\theta, \qquad D_{\hat{\phi}} = \frac{1}{R\ \sin\theta}\ \partial_\phi. \tag{5.14}$$

Recall that D_0 is the directional derivative along the congruence as defined in Eq. (3.13), with α equal to the lapse, and β^r the shift:

$$D_0 = \frac{1}{\alpha}\ (\partial_t - \beta^r\ \partial_r). \tag{5.15}$$

The twenty-four connection coefficients are derived from the commutators of the Cartesian spacetime tetrad directional derivatives given in Eqs. (5.11), (5.12), and (5.13). Recall that in the EW formulation, a space-time split with respect to the congruence has been made so that the connection coefficients are labelled with spatial triad indices. First consider the primary connection coefficients, N_{ab} and K_{ab}. It is a special property of the Cartesian tetrads that the combinations $(N_{ab} + N_{ba})$ are zero. Hence, N_{ab} is anti-symmetric:

$$N_{11} = N_{22} = N_{33} = 0, \tag{5.16}$$

$$N_{12} = -N_{21}, \qquad N_{23} = -N_{32}, \qquad N_{31} = -N_{13}. \tag{5.17}$$

Recall from Chapter 3 that we represent the antisymmetric part of N_{ba} by $n_a \equiv \varepsilon_{abc}\ N_{bc}/2$. Consequently,

$$n_1 = N_{23}, \qquad n_2 = N_{31}, \qquad n_3 = N_{12}. \tag{5.18}$$

After calculating n_a from the connection coefficients, we find that they are the Cartesian components of a vector $\boldsymbol{n}$ in the radial direction such that $\boldsymbol{n} = n_{\hat{r}}\ \boldsymbol{e}_{\hat{r}}$. Explicitly,

$$n_1 = \cos\phi\ \sin\theta\ n_{\hat{r}}, \qquad n_2 = \sin\phi\ \sin\theta\ n_{\hat{r}}, \qquad n_3 = \cos\theta\ n_{\hat{r}}, \tag{5.19}$$

where

$$n_{\hat{r}} = \frac{1 - D_{\hat{r}}\ R}{R}. \tag{5.20}$$

The diagonal components, N_{11}, N_{22}, and N_{33}, are calculated to be zero. Because we have chosen the rotations to and from Cartesian and spherical orthonormal tetrads to be the same as in flat space, there are no twists of the spatial triads as one moves through successive two-spheres in the $\boldsymbol{e}_{\hat{r}}$ direction.

A spherically symmetric congruence is orthogonal to the two-spheres. Thus, the vorticity of the congruence,

$$\Omega_a \equiv \frac{1}{2}\,\varepsilon_{abc}\,K_{bc} \tag{5.21}$$

is zero initially, which makes K_{ab} symmetric initially. Because the congruence worldlines remain orthogonal to the two-spheres, although not necessarily to the hypersurfaces, Ω_a continues to be zero throughout the numerical calculation. Hence,

$$K_{12} = K_{21}, \qquad K_{23} = K_{32}, \qquad K_{31} = K_{13}. \tag{5.22}$$

When calculating the Cartesian components of the tensor K_{ab}, we find that the only degrees of freedom are what we call K_R and K_T, so that

$$\begin{aligned}
K_{12} &= \cos\phi\ \sin\phi\ \sin^2\theta\ (K_R - K_T),\\
K_{23} &= \cos\theta\ \sin\theta\ \sin\phi\ (K_R - K_T),\\
K_{31} &= \cos\theta\ \sin\theta\ \cos\phi\ (K_R - K_T),\\
K_{11} &= \cos^2\phi\ \sin^2\theta\ K_R + (\sin^2\phi + \cos^2\theta\ \cos^2\phi)\ K_T,\\
K_{22} &= \sin^2\phi\ \sin^2\theta\ K_R + (\cos^2\phi + \sin^2\phi\ \cos^2\theta)\ K_T,\\
K_{33} &= \cos^2\theta\ K_R + \sin^2\theta\ K_T.
\end{aligned} \tag{5.23}$$

The variables K_R and K_T are defined as

$$K_R \equiv K_{\hat{r}\hat{r}} \equiv \boldsymbol{e}_{\hat{r}} \cdot \nabla_{\boldsymbol{e}_{\hat{r}}}\, \boldsymbol{e}_0, \tag{5.24}$$

$$K_T \equiv K_{\hat{\theta}\hat{\theta}} \equiv \boldsymbol{e}_{\hat{\theta}} \cdot \nabla_{\boldsymbol{e}_{\hat{\theta}}}\, \boldsymbol{e}_0 = K_{\hat{\phi}\hat{\phi}} \equiv \boldsymbol{e}_{\hat{\phi}} \cdot \nabla_{\boldsymbol{e}_{\hat{\phi}}}\, \boldsymbol{e}_0, \tag{5.25}$$

and are the extrinsic curvature components of the 3-space which is locally orthogonal to the congruence worldlines.

Next, consider the gauge connection coefficients, ω_a and a_a. Because the spatial triads are Fermi-propagated along the congruence worldlines, which are orthogonal to the two-spheres, the angular velocity vector components, ω_a, are zero. The quantities a_a are the Cartesian components of the acceleration four-vector, which is directed along $\boldsymbol{e}_{\hat{r}}$. So,

$$\omega_1 = \omega_2 = \omega_3 = 0, \tag{5.26}$$

$$a_1 = \cos\phi \ \sin\theta \ a_{\hat{r}}, \qquad a_2 = \sin\phi \ \sin\theta \ a_{\hat{r}}, \qquad a_3 = \cos\theta \ a_{\hat{r}}. \tag{5.27}$$

The EW formulation in spherical symmetry is streamlined by using K_R, K_T, $n_{\hat{r}}$, $a_{\hat{r}}$, $B^r_{\hat{r}} = e^{-\lambda}$, $B^\theta_{\hat{\theta}} = 1/R$, and $A_{\hat{r}}$ as variables instead of the Cartesian components. This results in a reduction of the total number of variables from thirty-six to seven. The lapse function and shift vector are not evolved dynamically in this application.

The spacetime metric is obtained by first calculating $g^{\mu\nu} = \eta^{\alpha\beta} \ \lambda^\mu_\alpha \ \lambda^\nu_\beta$. The $g^{\mu\nu}$ matrix is then inverted to give $g_{\mu\nu}$. The resulting metric is:

$$ds^2 = [-\alpha^2 + \beta^2 \ e^{2\lambda} \ (1 - A^2_{\hat{r}}) + 2 \ e^{\lambda} \ \alpha \ \beta \ A_{\hat{r}}] \ dt^2 + 2 \ e^{\lambda} \ [\alpha \ A_{\hat{r}} + \beta \ e^{\lambda} \ (1 - A^2_{\hat{r}})] \ dr \ dt$$
$$+ e^{2\lambda} \ (1 - A^2_{\hat{r}}) \ dr^2 + R^2 \ d\theta^2 + R^2 \ \sin^2\theta \ d\phi^2. \tag{5.28}$$

The spatial metric of the t = constant hypersurface is

$$dl^2 = e^{2\lambda} \ (1 - A^2_{\hat{r}}) \ dr^2 + R^2 \ d\theta^2 + R^2 \ \sin^2\theta \ d\phi^2. \tag{5.29}$$

5.3 Tetrad Quasi-Evolution and Quasi-Constraint Equations

Since the Einstein equations are covariant under rotations of the tetrad, the quasi-evolution equations for K_R, K_T, and $n_{\hat{r}}$, and the quasi-constraint equations for K_T and $n_{\hat{r}}$, can be derived directly from the Einstein equations in the

$\boldsymbol{e}_{\hat{r}}$, $\boldsymbol{e}_{\hat{\theta}}$, $\boldsymbol{e}_{\hat{\phi}}$ basis. The resulting quasi-evolution equations are

$$D_0\, K_R - D_{\hat{r}}\, a_{\hat{r}} = S\_K_R, \tag{5.30}$$

$$D_0\, K_T - D_{\hat{r}}\, n_{\hat{r}} = S\_K_T, \tag{5.31}$$

$$D_0\, n_{\hat{r}} - D_{\hat{r}}\, K_T = S\_n_{\hat{r}}, \tag{5.32}$$

with

$$S\_K_R = a_{\hat{r}}^2 - n_{\hat{r}}^2 - K_R^2 + K_T^2 + \frac{2\, n_{\hat{r}}}{R},$$
$$S\_K_T = \frac{a_{\hat{r}} + n_{\hat{r}}}{R} + K_R\, K_T - K_T^2 - n_{\hat{r}}^2 - a_{\hat{r}}\, n_{\hat{r}},$$
$$S\_n_{\hat{r}} = \frac{K_T - K_R}{R} - a_{\hat{r}}\, K_T + n_{\hat{r}}\, (K_R - 2\, K_T).$$

The momentum and energy quasi-constraint equations are, respectively,

$$D_{\hat{r}}\, K_T = \frac{(K_R - K_T)\,(1 - R\, n_{\hat{r}})}{R}, \tag{5.33}$$

$$D_{\hat{r}}\, n_{\hat{r}} = \frac{3\, n_{\hat{r}}^2}{2} - \frac{2\, n_{\hat{r}}}{R} - \frac{K_T}{2}\,(2\, K_R + K_T). \tag{5.34}$$

The gauge quasi-evolution equations for $a_{\hat{r}}$ are not covariant under rotation, and must be derived from Eqs. (3.43) and (3.46) for the Nester and Lorentz gauges, respectively, in a Cartesian basis. Then, using Eqs. (5.11), (5.12), (5.13), and (5.23), we obtain equations in the $\{\boldsymbol{e}_{\hat{r}},\ \boldsymbol{e}_{\hat{\theta}},\ \boldsymbol{e}_{\hat{\phi}}\}$ basis. Both the Nester and Lorentz gauges give quasi-evolution equations for $a_{\hat{r}}$ of the form

$$D_0\, a_{\hat{r}} - D_{\hat{r}}\, K_R = S\_a_{\hat{r}}. \tag{5.35}$$

The source, $S\_a_{\hat{r}}$ depends on the gauge. For the Nester gauge,

$$S\_a_{\hat{r}} = \frac{2\,(K_R - K_T)}{R}, \tag{5.36}$$

and for the Lorentz gauge,

$$S\_a_{\hat{r}} = \frac{2\,(K_R - K_T)}{R} - 2\,(K_R\, n_{\hat{r}} + K_T\, a_{\hat{r}}). \tag{5.37}$$

The quasi-constraint equation for the Nester gauge, Eq. (3.44), is trivially zero because in spherical symmetry, the curl of a radial vector is zero.

5.4 True Evolution Equations and their Hyperbolic Structure

In order to evolve K_R, K_T, $n_{\hat{r}}$, and $a_{\hat{r}}$ numerically, the evolution equations must be expressed as partial derivatives along the coordinate directions r and t. To obtain these true evolution equations, Eq. (5.14) is substituted into the quasi-evolution equations in Sec.5.3. This substitution introduces an additional D_0 term into each equation, which is then collected with the D_0 term already there. A set of equations in the same form as Eq. (3.52) results. Multiplication by the inverse of the $\boldsymbol{T}$ matrix (which is the coefficient in front of D_0) yields

$$D_0\,\mathbf{q} + \boldsymbol{C}^{\hat{r}}\,B^r_{\hat{r}}\,\partial_r\,\mathbf{q} = \mathbf{S}. \tag{5.38}$$

In this equation,

$$\mathbf{q} = \begin{pmatrix} K_R \\ a_{\hat{r}} \\ K_T \\ n_{\hat{r}} \end{pmatrix}, \qquad \boldsymbol{C}^{\hat{r}} = -\frac{1}{1-A_{\hat{r}}^2} \begin{pmatrix} A_{\hat{r}} & 1 & 0 & 0 \\ 1 & A_{\hat{r}} & 0 & 0 \\ 0 & 0 & A_{\hat{r}} & 1 \\ 0 & 0 & 1 & A_{\hat{r}}, \end{pmatrix},$$

and

$$\mathbf{S} = \frac{1}{1-A_{\hat{r}}^2} \begin{pmatrix} S\_K_R + A_{\hat{r}}\,S\_a_{\hat{r}} \\ S\_a_{\hat{r}} + A_{\hat{r}}\,S\_K_R \\ S\_K_T + A_{\hat{r}}\,S\_n_{\hat{r}} \\ S\_n_{\hat{r}} + A_{\hat{r}}\,S\_K_T \end{pmatrix}.$$

The eigensystem of the characteristic matrix, $\boldsymbol{C}^{\hat{r}}$, consists of four eigenmodes:

$$a_{\hat{r}} + K_R, \quad a_{\hat{r}} - K_R, \quad n_{\hat{r}} + K_T, \quad n_{\hat{r}} - K_T, \tag{5.41}$$

and eigenvalues $1/(1+A_{\hat{r}})$ and $-1/(1-A_{\hat{r}})$. The complete set of eigenvectors with real eigenvalues verifies the hyperbolicity of the system.

Four advection equations describe the propagation of the eigenmodes along the light cones with speeds s_1 and s_2:

$$s_1(r,t) = \frac{e^{-\lambda}\,\alpha}{1+A_{\hat{r}}} - \beta^r, \tag{5.42}$$

$$s_2(r,t) = -\frac{e^{-\lambda}\,\alpha}{1-A_{\hat{r}}} - \beta^r. \tag{5.43}$$

These speeds are the local coordinate speeds of light. (The lapse and the shift appear in the eigenvalues when the D_0 operator is decomposed into its partial derivatives.) The advection equations are:

$$\partial_t\,(a_{\hat{r}} - K_R) + s_1\;\partial_r\,(a_{\hat{r}} - K_R) = \frac{\alpha}{1+A_{\hat{r}}}\,(S\_a_{\hat{r}} - S\_K_R), \tag{5.44}$$

$$\partial_t\,(n_{\hat{r}} - K_T) + s_1\;\partial_r\,(n_{\hat{r}} - K_T) = \frac{\alpha}{1+A_{\hat{r}}}\,(S\_n_{\hat{r}} - S\_K_T), \tag{5.45}$$

$$\partial_t\,(a_{\hat{r}} + K_R) + s_2\;\partial_r\,(a_{\hat{r}} + K_R) = \frac{\alpha}{1-A_{\hat{r}}}\,(S\_a_{\hat{r}} + S\_K_R), \tag{5.46}$$

$$\partial_t\,(n_{\hat{r}} + K_T) + s_2\;\partial_r\,(n_{\hat{r}} + K_T) = \frac{\alpha}{1-A_{\hat{r}}}\,(S\_n_{\hat{r}} + S\_K_T). \tag{5.47}$$

The two eigenmodes involving K_R and $a_{\hat{r}}$, are "longitudinal" modes, since they are constructed from components of the extrinsic curvature tensor and acceleration vector projected along $\boldsymbol{e}_{\hat{r}}$. The two involving K_T and $n_{\hat{r}}$, are "constraint" modes, since K_T and $n_{\hat{r}}$ are the variables which appear in the principal parts of the constraint equations.

5.5 *True Constraint Equations*

The tetrad constraint Eqs. (5.33) and (5.34) are not true constraint equations because they contain D_0 (see the definition of $D_{\hat{r}}$ in Eq. (5.14)). The true constraint equations are obtained by eliminating the D_0 terms with Eq. (5.38), and solving for $\partial_r\ K_T$ and $\partial_r\ n_{\hat{r}}$. The resulting momentum and energy true constraint equations are, respectively,

$$\partial_r\ K_T = e^{\lambda}\left[(K_R - K_T)\left(\frac{1}{R} - n_{\hat{r}}\right) - \frac{A_{\hat{r}}}{2}\left(\frac{2(a_{\hat{r}} - n_{\hat{r}})}{R} - 3K_T^2 - 2a_{\hat{r}}n_{\hat{r}} + n_{\hat{r}}^2\right)\right], \tag{5.48}$$

$$\partial_r\, n_{\hat{r}} = \; -e^{\lambda}\left[\frac{2\, n_{\hat{r}}}{R} + K_R\, K_T + \frac{K_T^2}{2} - \frac{3\, n_{\hat{r}}^2}{2} - A_{\hat{r}}\, K_T\, (a_{\hat{r}} + n_{\hat{r}})\right] \tag{5.49}$$

These equations are used in calculating the initial conditions, boundary conditions, and as an accuracy check for the numerical evolution.

5.6 Evolution and Constraint Equations for $B_{\hat{r}}$, $B_{\hat{\theta}}$, and $A_{\hat{r}}$

Evolution equations for the variables which determine the tetrad vector components are also required. Those for the non-zero components of $\boldsymbol{B}_{\hat{r}}$ and $\boldsymbol{B}_{\hat{\theta}}$ are obtained from Eq. (3.63). They are:

$$(\partial_t - \mathcal{L}_{\boldsymbol{\beta}})e^{-\lambda} = -\, e^{-\lambda}\, \alpha\, K_R \;\Rightarrow D_0\, \lambda = K_R + \frac{\partial_r\, \beta^r}{\alpha}, \tag{5.50}$$

$$D_0\, R = R\, K_T. \tag{5.51}$$

The evolution equation for $A_{\hat{r}}$, from Eq. (3.64), is

$$D_0\, A_{\hat{r}} = a_{\hat{r}} - K_R\, A_{\hat{r}} - e^{-\lambda}\, \partial_r(\ln \alpha). \tag{5.52}$$

The only non-zero constraint equation derived from Eqs. (3.66) and (3.67) is that which has already been presented in Eq. (5.20), namely,

$$\partial_r\, R = e^{\lambda}\, (1 - R\, n_{\hat{r}} - R\, A_{\hat{r}}\, K_T). \tag{5.53}$$

In addition to the energy and momentum constraint equations, this constraint on $\partial_r\, R$ is used to check accuracy of the numerical evolution.

5.7 The Initial Value Problem (IVP)

The initial value problem consists of finding values for $A_{\hat{r}}$, K_T, $n_{\hat{r}}$, R, λ, K_R, and $a_{\hat{r}}$ with which to begin the numerical evolution. We take the congruence orthogonal to the initial hypersurface, so $A_{\hat{r}} = 0$ initially. K_T, $n_{\hat{r}}$, and R are constrained by Eqs. (5.48), (5.49), and (5.53), which, when $A_{\hat{r}} = 0$, reduce to

$$\partial_r\, K_T = e^{\lambda}\, (K_R - K_T)\left(\frac{1}{R} - n_{\hat{r}}\right), \tag{5.54}$$

$$\partial_r\, n_{\hat{r}} = \; -e^{\lambda}\left[\frac{2\, n_{\hat{r}}}{R} + K_R\, K_T + \frac{K_T^2}{2} - \frac{3\, n_{\hat{r}}^2}{2}\right], \tag{5.55}$$

$$\partial_r\, R = e^{\lambda}\,(1 - R\, n_{\hat{r}}). \tag{5.56}$$

For the special case of spherical symmetry, one can use the first integral of the constraint equations,

$$R_{;\mu}\, R^{;\mu} = (e^{-\lambda}\,\partial_r\, R)^2 - (D_0\, R)^2 = (1 - R\, n_{\hat{r}})^2 - (R\, K_T)^2 \equiv 1 - \frac{2M}{R}, \tag{5.57}$$

to obtain an analytic expression for $n_{\hat{r}}$. This expression is:

$$n_{\hat{r}} = \frac{1 - \sqrt{1 - \frac{2\,M}{R} + R^2\, K_T^2}}{R}. \tag{5.58}$$

In vacuum, M is a constant; namely, the total mass of the black hole. The only requirement on λ is that it be a function of r, so $\lambda(r)$ is a completely free variable. This leaves K_R and $a_{\hat{r}}$ to be determined.

5.7.1 Time Independent IVP

The simplest and most basic test for the EW formulation in Schwarzschild spacetime is a time independent IVP. Schwarzschild spacetime has a time Killing vector field, which means that a coordinate system can be found in which the metric is time independent. However, in the EW tetrad formulation, the variables are related to the congruence, not to the metric in the t = constant hypersurface. In order for the EW variables to be time independent, the evolution of the congruence has to be stationary. So, the first task is to demonstrate the possibility of finding an initial K_R and $a_{\hat{r}}$ such that the tetrad congruence is stationary.

We start by setting the partial time derivatives for K_R and $a_{\hat{r}}$ to zero in Eqs. (5.44) and (5.46). Combining the resulting equations in order to solve for $\partial_r\, K_R$, we get

$$\partial_r\, K_R = \frac{e^{\lambda}\left(\frac{\beta^r\, e^{\lambda}}{\alpha}\, S\_K_R - S\_a_{\hat{r}}\right)}{1 - \frac{\beta^{r\,2}\, e^{2\,\lambda}}{\alpha^2}}, \tag{5.59}$$

where $S\_a_{\hat{r}}$ is dependent on whether the Nester or Lorentz gauge is used. We choose a shift consistent with a time independent evolution by setting $\partial_t\ R$ to zero in Eq. (5.51), to give

$$\beta^r = -\,e^{-\lambda}\,\alpha\,\left(\frac{R\,K_T}{1-R\,n_{\hat{r}}}\right). \tag{5.60}$$

Plugging this expression for the shift into Eq. (5.59), we obtain a differential equation for K_R in terms of the other variables:

$$\partial_r\ K_R = -\,\frac{e^{\lambda}\,(1-R\,n_{\hat{r}})\,[(R\,K_T)\,S\_K_R + (1-R\,n_{\hat{r}})\,S\_a_{\hat{r}}]}{(1-R\,n_{\hat{r}})^2 - (R\,K_T)^2}. \tag{5.61}$$

Eq. (5.61) is singular when $(1-R\,n_{\hat{r}})^2 - (R\,K_T)^2 = 0$, which occurs at the black hole horizon (see Eq. (5.57)). A regular solution on the horizon is required for numerical integration, and is satisfied by the imposition of a constraint on the value of K_R there. This constraint is obtained by setting the numerator of Eq. (5.61) to zero at $R = 2M$, giving $0/0$. The value of K_T on the horizon is a free parameter, the sign of which is chosen so that the timelike trajectory points into the horizon of the black hole. Using L'Hôpital's rule to solve for the right hand side at $R = 2M$, Eq. (5.61) can then be integrated starting at the horizon, satisfying the constraint initially, to get a one-parameter family of solutions.

An algebraic expression for $a_{\hat{r}}$ can be found by setting $\partial_t\ n_{\hat{r}}$ to zero in the true evolution equation for $n_{\hat{r}}$, and substituting the true energy constraint equation into $\partial_r\ n_{\hat{r}}$. The result is:

$$a_{\hat{r}} = \frac{M + R^3\,K_R\,K_T}{R^2\,(1-R\,n_{\hat{r}})}. \tag{5.62}$$

With $a_{\hat{r}}$ satisfying the stationary solution, the derivative of the shift can be evaluated analytically, and is

$$\partial_r\,(e^{\lambda}\,\beta^r) = -\alpha\,K_R\,e^{\lambda}. \tag{5.63}$$

Finally, the lapse is obtained from Eq. (5.52) with $A_{\hat{r}} = 0$:

$$\partial_r\,(\ln\alpha) = e^{\lambda}\,a_{\hat{r}}. \tag{5.64}$$

Substitution of the lapse and shift, as well as Eqs. (5.54), (5.55), (5.56), (5.58), and (5.62), into the true evolution equations for K_T, $n_{\hat{r}}$, and λ, gives $\partial_t K_T = \partial_t n_{\hat{r}} = \partial_t \lambda = 0$. Thus, for a particular choice of K_T on the horizon, one can obtain K_R, $a_{\hat{r}}$, α, and β^r so that the evolution of all the variables is time independent.

It is instructive to discuss the analytic solutions for the asymptotic behavior of K_T and K_R. Another way to write the momentum constraint Eq. (5.54) is

$$\partial_r K_T = (K_R - K_T) \frac{\partial_r R}{R}. \tag{5.65}$$

If we assume that K_R approaches a constant at large R, then we can write Eq. (5.65) as

$$-\partial_r \ln(K_R - K_T) = \partial_r \ln(R) \Rightarrow K_R - K_T = \frac{C}{R} \xrightarrow[R\to\infty]{} 0, \tag{5.66}$$

with C a constant. This implies that K_R and K_T approach the same limiting constant for large R (which is also large r). We check to see if our assumption of constant K_R at large R is consistent with Eq. (5.61) by evaluating the asymptotic limit of the right hand side of Eq. (5.61), to see if it is zero. If $lim_{R\to\infty}\,(K_R) = lim_{R\to\infty}\,(K_T) = C$, then $lim_{R\to\infty}\,(a_{\hat{r}}) = |K_T|$ and $lim_{R\to\infty}\,(1 - R n_{\hat{r}}) = R|K_T|$. Plugging these values into Eq. (5.61), we find that $lim_{R\to\infty}\,(S\_a_{\hat{r}}) = lim_{R\to\infty}\,(R\,K_T\,S\_K_R) = 0$. Thus,

$$\partial_r K_R \xrightarrow[R\to\infty]{} 0. \tag{5.67}$$

Our assumption of asymptotically constant K_R is consistent.

If the value of K_T on the horizon K_{TH}, equals a critical value, K_{THC}, then both K_R and K_T asymptotically approach zero, and the hypersurface is asymptotically flat. For $K_{THC} < K_{TH} < 0$, K_T goes from negative on the horizon to positive at large r, and the slice is asymptotically hyperbolic.

5.7.2 Constant Mean Curvature (CMC) Slice

A more general, time dependent problem can be tested with different initial choices for K_R and $a_{\hat{r}}$. We have tried a constant mean curvature initial slice, so that ($Tr\ K = K_R + 2\ K_T = K_0$) initially, with K_0 a constant. We choose $a_{\hat{r}}$, α, and β^r according to Eqs. (5.62), (5.60), and (5.64), so that they satisfy the stationary solution initially. However, because of the constant mean curvature initial choice for K_R, $a_{\hat{r}}$ will evolve according to Eqs. (5.44) and (5.46). The question we pose is: if the initial conditions are reasonably close to a stationary solution, will the evolution become dynamic and run away, or will it eventually settle down to a time independent solution?

5.8 Coordinate Conditions

Conditions on the lapse and shift have been developed to keep the coordinates and spacelike hypersurfaces well-behaved throughout the numerical evolution, especially near the event horizon of the black hole. The lapse can be used to control the velocity of the tetrads relative to the hypersurfaces, and must be chosen in the spherically symmetric EW formulation to prevent a coordinate singularity at $A_{\hat{r}} = 1$. The shift controls the stretching of the coordinate grid. Three classes of coordinate conditions are considered in this chapter, those which (1) fix the lapse and shift at their initial values, (2) evolve the lapse according to equations which preserve the hyperbolicity of the system (keeping the shift fixed at its initial value), and (3) reset both the lapse and shift at frequent intervals throughout the calculation.

5.8.1 Fixed Conditions for α and β^r.

The shift and lapse are calculated initially from Eqs. (5.60) and (5.64), respectively. They are then fixed at these values throughout the subsequent numeri-

cal evolution. However, even with stationary initial conditions, they are problematic. As numerical errors inevitably cause the solution to deviate from its initial value, fixed coordinate conditions are not able to compensate for these changes, so the errors will tend to get out of hand.

For example, the congruence becomes superluminal when the advection speed $s_2 = e^{-\lambda}\ \alpha/(1 + A_{\hat{r}}) - \beta^r$ becomes positive. The coordinate speed of the congruence relative to an observer at constant R is β, and the coordinate speed of a light signal relative to the congruence is $e^{-\lambda}\ \alpha/(1 + A_{\hat{r}})$. Theoretically, the coordinates remain at constant R in a time independent problem, so $s_2 = 0$ at the event horizon. But because of numerical errors, the coordinates do not stay at constant R, which means that the congruence may become superluminal inside the event horizon.

5.8.2 Dynamic Evolution for the Lapse.

A dynamic evolution of the lapse similar to the modified Bona-Massó condition in Alcubierre *et al.* [4] has been considered. Let $f(\alpha)$ be an arbitrary function of the lapse, $h = \ln \alpha$, $g_{\hat{r}} = e^{-\lambda}\ \partial_r\ (\ln \alpha)$, and K_0 a positive number. Then the hyperbolic system with dynamic lapse is identical to the original system presented in Sec. 5.4, except for the addition of the two equations:

$$D_0\ h = f\ (K_R + 2\ K_T - K_0), \tag{5.68}$$

$$D_0\ [g_{\hat{r}} - f\ (a_{\hat{r}} + 2\ n_{\hat{r}} - A_{\hat{r}}\ (K_R + 2\ K_T))] = S, \tag{5.69}$$

where S is the source term, the specifics of which are not necessary here. These equations do not affect the hyperbolicity of the original system, since the flux terms have trivial speeds of $-\beta^r$. The shift remains fixed at the value obtained initially from Eq. (5.60). The form of the dynamic lapse equation, generalized

to 3D, is:

$$D_0\, g_a - f\, B_a^k\, \frac{\partial TrK}{\partial_{x^k}} = (TrK - K_0)\, B_a^k\, \frac{\partial f}{\partial_{x^k}} - (\varepsilon_{abc}\, \omega_b + K_{ac})\, g_c + g_a\, f\, (TrK - K_0).$$

5.8.3 Resetting Conditions.

The lapse and the shift can be periodically updated between time steps according to specified conditions. During a given time step, the lapse and shift are held fixed at the values they have at the start of the time step. While it is essential to accuracy and stability that the system be hyperbolic during the time steps [13], it is not critical that the overall evolution be hyperbolic, only well-posed [60]. In other words, the resetting conditions do not need to preserve the hyperbolicity of the system. This gives a much greater range of choice for resetting conditions than for the dynamic coordinate conditions described above.

The lapse is reset to prevent a coordinate singularity at $A_{\hat{r}} = 1$. This is achieved by resetting the lapse according to (Eq. 5.64) at frequent intervals throughout the numerical evolution, to keep $A_{\hat{r}}$ small.

To reset the shift, we use a minimal strain condition, which was first introduced by Smarr and York [70] for $3 + 1$ formulations. The minimal strain condition minimizes the time derivative of the spatial metric so that if the spatial metric is allowed to relax to a stationary metric, it will. In order for this to happen in $3 + 1$ formulations, the lapse condition must be such that it results in an asymptotically stationary evolution of the hypersurface. In the EW tetrad formulation, the congruence as well as the hypersurface must approach stationarity.

For $3 + 1$ formulations, Smarr and York [70] minimize the action

$$I_1 = \int \left[\frac{\partial h_{ij}}{\partial t}\, \frac{\partial h_{kl}}{\partial t}\, h^{ik}\, h^{jl}\right] \sqrt{h}\, d^3x \tag{5.71}$$

with respect to variations in the shift. In the expression for I_1, h_{ij} is the spatial metric of the $t =$ constant hypersurface and h is the determinant of this metric.

The shift appears in this action through the Lie derivative, because in the $3+1$ formulation,

$$\frac{\partial h_{ij}}{\partial t} = -2\,\alpha\,K_{ij} + \mathcal{L}_{\boldsymbol{\beta}}\,h_{ij}, \tag{5.72}$$

where K_{ij} is the extrinsic curvature of the t = constant hypersurface. So, the action is the square of the time derivative of the spatial metric tensor integrated over the t = constant hypersurface.

To apply this minimal strain condition to the EW tetrad formulation, we use a spatial metric with respect to which the $\boldsymbol{B}_a$ vectors are mutually orthonormal. This spatial metric is constructed from the one-forms $\boldsymbol{B}^a$ dual to the $\boldsymbol{B}_a$ vectors. The B^a_k matrix is obtained simply by taking the inverse of the B^k_a matrix. Thus we have

$$h_{kl} = B_{kb}\,B^b_l, \qquad h^{kl} = B^k_b\,B^{lb}, \qquad B^k_a\,B^b_k = \delta^b_a, \tag{5.73}$$

and

$$dl^2 = h_{kl}\,dx^k\,dx^l. \tag{5.74}$$

Note that the tetrad spatial indices are trivially raised and lowered using the Euclidean spatial metric. If one decomposes the metric of the t = constant hypersurface along $\boldsymbol{e}_0$ and $\boldsymbol{e}_a$, then the metric described by Eqs. (5.73) and (5.74) lies in the 3-space spanned by $\boldsymbol{e}_a$. Substitution of the expressions in Eq. (5.73) (for a diagonal metric) into Eq. (5.71) gives an action for the minimal strain condition in the EW formulation:

$$I_2 = \int h_{ij}\,\frac{\partial\,B^i_a}{\partial t}\,\frac{\partial\,B^{aj}}{\partial t}\,\sqrt{h}\,d^3x. \tag{5.75}$$

These equations are applied to the EW spherically symmetric problem. The first step is to construct a spatial metric as prescribed by Eqs. (5.73) and (5.74):

$$dl^2 = e^{2\lambda}\,dr^2 + R^2\,d\theta^2 + R^2\,\sin^2\theta\,d\phi^2. \tag{5.76}$$

Since the lapse resetting condition keeps $A_{\hat{r}}$ small, Eq. (5.76) actually approximates the true metric of the t = constant hypersurface, Eq. (5.29). An expansion of Eq. (5.75) using this spatial metric gives

$$I_2 = \int dr\ Q, \tag{5.77}$$

where

$$Q = e^{\lambda}\ R^2\ \sin\theta\ [(\partial_t\ \lambda)^2 + \frac{2}{R^2}\ (\partial_t\ R)^2]. \tag{5.78}$$

Now, let $(\beta^r + \varepsilon)$ represent a small variation from β^r. The action I_2 is minimized with respect to variations in the shift if

$$\delta I_2 = \int_a^b dr \left(\frac{\partial Q}{\partial(\beta^r)'}\ \varepsilon' + \frac{\partial Q}{\partial \beta^r}\ \varepsilon \right) = 0, \tag{5.79}$$

where a and b are the inner and outer grid points, respectively, and $\{\ ' \ \Rightarrow \partial_r\}$. Integrating by parts,

$$\delta I_2 = \left[\varepsilon\ \frac{\partial Q}{\partial(\beta^r)'} \right]_a^b + \int_a^b \varepsilon(r)\ dr \left[\frac{\partial Q}{\partial \beta^r} - \frac{d}{dr}\ \frac{\partial Q}{\partial(\beta^r)'} \right] = 0. \tag{5.80}$$

The action is minimized if

$$\frac{d}{dr}\ \frac{\partial Q}{\partial(\beta^r)'} - \frac{\partial Q}{\partial \beta^r} = 0 \tag{5.81}$$

is satisfied, and $\partial Q / \partial(\beta^r)'$ vanishes at the boundaries. The result is a set of relatively simple coupled first order differential equations for the shift:

$$y \equiv \partial_t \lambda = \beta^r\ \partial_r \lambda + \alpha\ K_R + \partial_r\ \beta^r, \tag{5.82}$$

$$\partial_r\ y = 2\ e^{\lambda} \left(\frac{1}{R} - n_{\hat{r}} - A_{\hat{r}}\ K_T \right) \left[\alpha\ K_T - y + \beta^r\ e^{\lambda} \left(\frac{1}{R} - n_{\hat{r}} - A_{\hat{r}}\ K_T \right) \right). \tag{5.83}$$

Elliptic equations are unstable in 3D with boundary conditions at only one boundary; hence, one condition is imposed at both the inner and outer boundaries. Strict minimization requires $\partial_t\ \lambda = 0$ at each boundary, since

$$\frac{\partial Q}{\partial(\beta^r)'} = 2\ e^{\lambda}\ R^2\ \sin\theta\ (\partial_t \lambda). \tag{5.84}$$

However, we have found it more important to impose $\partial_t\ R = 0$ at the boundaries, especially with non-stationary initial conditions, in order to keep the inner edge of the grid inside the event horizon and to keep the edges of the grid stationary.

5.9 *Boundary Conditions*

There are two boundaries in the black hole problem. One is at the outer edge of the grid, and the other is a bit inside the event horizon located at $R = 2\ M$. In the 3D non-spherically symmetric problem, one desires to place the outer boundary as far away from the event horizon as possible, in order to calculate the gravitational waves in an asymptotically flat spacetime. However, in practice, the outer boundary cannot be placed very far out, because of the high computational costs in 3D. We choose our outer boundary to be where the fundamental variables have reached nearly constant values in the time independent solution, around $R = 20M$.

The choice of K_T on the horizon (K_{TH}) affects the speeds s_1, s_2, and $-\beta$ at the inner and outer boundaries. These speeds are plotted in Fig. 5.1 for the stationary IVP, with $K_{TH} = -0.19$ for the Nester gauge, and $K_{TH} = -0.27$ for the Lorentz gauge.

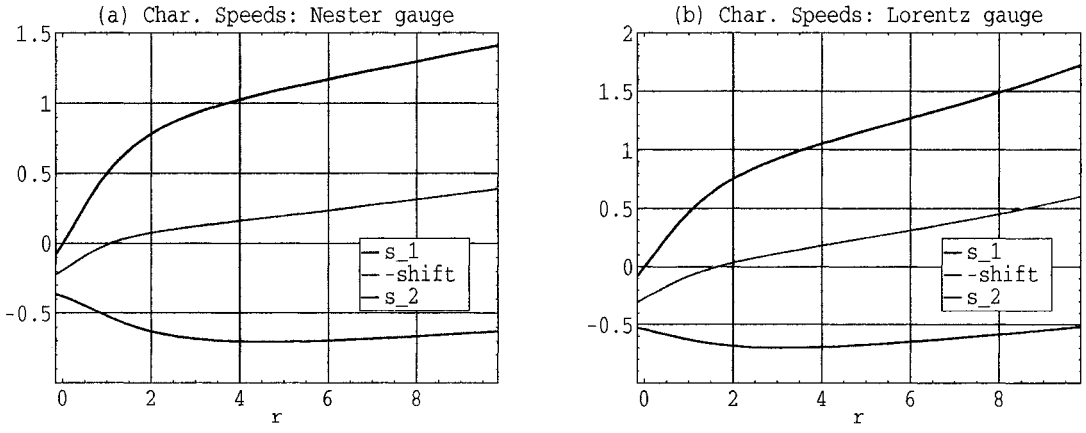


Figure 5.1: Characteristic speeds at $t = 0$, stationary IVP. $K_{TH} = -0.19$ for Nester gauge, and $K_{TH} = -0.27$ for Lorentz gauge

The inner boundary is located just inside the event horizon (which is at

$r = 0$). With $K_{TH} < 0$, all the characteristic speeds are negative at the inner edge. This means that all of the characteristics point into the black hole, and away from the grid. If the shift condition is such that the inner boundary is kept inside the event horizon throughout the numerical evolution, then no information can propagate into the computational grid through the inner boundary. This is the basis for excising the entire region interior to the inner boundary from the computational domain. In addition, since the characteristics of our hyperbolic set of equations lie either on or within the local light cone, those that originate inside the event horizon never cross the horizon. Hence, all the information needed to update the variables at the inner boundary can be obtained from values at upwind points, and no inner boundary condition is required for the evolution.

At the outer boundary, speeds s_1 and $-\beta^r$ are positive, and s_2 is negative. This means that the five eigenmodes with speeds s_1 or $-\beta^r$ propagate along characteristics pointing out of the grid, and do not require boundary conditions. However, the two eigvectors with speeds s_2 are propagating into the grid at the outer boundary. These eigenmodes are $(a_{\hat{r}} + K_R)$, which we call a longitudinal mode, and $(n_{\hat{r}} + K_T)$, which we call a "constraint" mode. Proper boundary conditions are required to prevent spurious information from entering the grid. Our procedure for implementing these boundary conditions are as follows. At the end of each time step, K_T and $n_{\hat{r}}$ in the outermost physical point are corrected for the incoming "constraint" eigenmode, $K_T + n_{\hat{r}}$, by integrating this mode from the closest interior point using a linear combination of the momentum and energy constraint equations. Then, K_T and $n_{\hat{r}}$ in the ghost point at the outer boundary are updated by integrating the energy and momentum constraint equations from the newly corrected outermost physical point. This gives values for K_T and $n_{\hat{r}}$ in the ghost point which satisfy both the incoming and outgoing "constraint" eigenmodes there. For stationary initial conditions,

the longitudinal variables, K_R and $a_{\hat{r}}$, are set to their stationary values at the ghost point, to give boundary conditions for the incoming longitudinal mode, $K_R + a_{\hat{r}}$. For CMC initial conditions, we quadratically extrapolate the longitudinal variables to the ghost point.

An alternative way to impose constraint preserving boundary conditions at the outer boundary has been proposed by Frittelli and Gomez [33], for symmetric hyperbolic $3+1$ systems, with zero shift. By projecting the Einstein tensor along the normal to the boundary, the components containing second derivatives across the boundary are eliminated. The boundary conditions are obtained by requiring the projected Einstein tensor to vanish. In this way, constraint violating solutions are prevented from entering the grid at the boundary.

As has already been mentioned in Sec. 5.8.3, numerically solving the minimal strain elliptic condition for the shift requires boundary conditions at both the inner and outer edges. Our boundary condition $\partial_t R = 0$ is satisfied simply by setting the shift equal to

$$\beta^r = -e^{-\lambda}\,\alpha\,\left(\frac{R\,K_T}{1 - R\,n_{\hat{r}} - R\,A_{\hat{r}}\,K_T}\right) \tag{5.85}$$

at the innermost point, which requires no ghost points. If R is initially less than $2M$ at the inner boundary, and $\partial_t\, R = 0$ there, the R remains inside the event horizon.

5.10 Numerical Methods

All the numerical methods used in the codes are designed to be second order accurate in both space and time.

Both split and unsplit schemes [48] were tested with fixed coordinate conditions, stationary initial conditions, and the Nester tetrad gauge (see LeVeque [48]). In the split schemes, the hyperbolic terms (setting the source terms to

zero) and the source terms (setting the fluxes to zero) are treated separately. In the unsplit method (which is called the method of lines [48]), the terms involving spatial derivatives are first computed using second order finite differencing. Then, the equations are treated as a large set of ODE's in time, and solved using a fourth order Runge-Kutta [59] stable numerical method. An unsplit numerical method was thought to be appropriate for our problem since we have smooth solutions. However, in practice, the results are not better than with a split method. Further, the unsplit method requires boundary conditions at both the inner and outer boundaries to solve the second order finite difference approximation to the spatial derivatives.

A second order Strang split method, in which the hyperbolic terms and the source terms are treated separately, was chosen (see Chapter 2 for a full description). In the transport step of the Strang split method, we solve homogeneous, variable coefficient advection equations (the left hand sides of Eqs. 5.44, 5.45, 5.46, 5.47, 5.50, 5.51, and 5.52), with constant eigenvectors. These can be solved easily with the method of characteristic tracing [48]. First order accurate expressions for the average speeds over the time step are calculated. The eigenmodes are then updated at each grid point in the following way. For eigenmodes with speeds s_1 and $-\beta$, Beam-Warming is used until the speeds change sign. Then, what we call a weighted average of Beam-Warming and Lax-Wendroff is applied, until the outermost physical point, at which pure Lax-Wendroff is used. If at a given grid point, we define e_{bw} to be a temporarily updated eigenmode calculated by Beam-Warming, e_{lw} a temporarily updated eigenmode calculated by Lax-Wendroff, and s_{av} the average speed, then the weighted average is:

$$e_{bw} \times \frac{|s_{av}|\, dt}{dx} + e_{lw} \times \left(1 - \frac{|s_{av}|\, dt}{dx}\right). \tag{5.86}$$

For the eigenmodes with speed s_1, which is negative all across the grid, Beam-

Warming is used at the innermost physical point, the weighted average is applied across the grid until the outermost physical point, then Lax-Wendroff is used.

All these advection equations are coupled via the source terms. When handling the source integration during the time step, the flux terms are ignored, as is standard for Strang splitting. The resulting ordinary differential equations (in time) are integrated with second order Runge-Kutta.

The integration of the spatial ODE's during the initial condition and boundary condition subroutines is implemented with Predictor-Corrector schemes [59], iterated until the difference between the predicted and corrected values are less than machine precision.

5.11 Results

There are several constraints in the 1D Schwarzschild problem, imposed on the initial time slice only. These are the true momentum and energy constraint Eqs. (5.48) and (5.49), the constraint on R Eq. (5.53), and the mass constraint Eq. (5.57). The numerical evaluation of the left-hand side minus the right-hand side of these equations at different times in the numerical evolution gives the constraint errors. If the numerical results converge to the analytic solution, then the constraint errors exhibit second order convergence to zero. The energy, momentum, and R constraints are the most sensitive error indicators since they involve derivatives.

Our convergence plots show the constraint errors as a function of r, at times $t = 20M$ and $t = 40M$. Such plots show the convergence of individual features of the constraint errors, and how these features change with time. The results from two grid resolutions differing by a factor of two are shown on the same graph. As the finer grid results are scaled by a factor of four, they will overlay

the coarser grid results if they converge to second order.

With a time independent IVP, the errors in the longitudinal and metric variables can be calculated as a percentage difference from their stationary initial values. If the numerical solution converges to the analytic solution, these errors will also vanish as $(\triangle r)^2$. These convergence plots are presented as described for the constraint errors.

The stability of the code is evaluated by establishing whether or not the growth of the constraint errors is exponential. We do this by plotting the 1-norms of the constraint errors versus time.

When evaluating the viability of the code, it is also important to plot the numerical evolution of the primary variables K_R, $a_{\hat{r}}$, K_T, and $n_{\hat{r}}$. That way, we can understand why there might be poor accuracy in particular places along the grid. For constant mean curvature initial conditions, we can see if the evolution of the variables becomes pathological or if it approaches a stationary solution.

Time

When presenting the results, it is requisite to convert the coordinate time of the numerical results to a physically meaningful time. The reference for Schwarzschild coordinates is a frame infinitely far from the black hole source; hence, Schwarzschild time (t_S) is ideal for representing the time that would be measured by a gravitational wave detector on or in orbit around Earth. The procedure for this conversion is quite straightforward for the stationary IVP, and is as follows. The spacetime metric in Schwarzschild coordinates is

$$ds^2 = -\left(1 - \frac{2\,M}{R}\right)\, dt_S^2 + \frac{1}{\left(1 - \frac{2\,M}{R}\right)} dR^2 + R^2\, d\Omega^2, \tag{5.87}$$

where $d\Omega^2 = R^2\ d\theta^2 + R^2\ \sin^2\theta\ d\phi^2$. The proper time for a stationary observer at radius R during an interval dt_S of Schwarzschild time is

$$d\tau = \sqrt{1 - \frac{2\ M}{R}}\ dt_S. \tag{5.88}$$

Since $(1 - 2M/R) < 1$ everywhere outside the event horizon, $d\tau < dt_S$. This represents the fact that clocks at rest run slower inside the gravitational potential well of the black hole than outside. Now, we can also calculate the proper time for a stationary observer at radius R in the coordinate system of our EW formulation, using the metric given in Eq. (5.29). The result for the time independent IVP is

$$d\tau = \sqrt{\alpha^2 - \beta^2\ e^{2\lambda}}\ dt. \tag{5.89}$$

Combining Eqs. (5.88) and (5.89), we get

$$dt_S = \sqrt{\frac{\alpha^2 - \beta^2\ e^{2\lambda}}{1 - \frac{2\ M}{R}}}\ dt. \tag{5.90}$$

Let $\Gamma = 1 - R\ n_{\hat{r}}$. Then we find that

$$e^{-\lambda}\ \partial_r \ln \Gamma = a_{\hat{r}} = e^{-\lambda}\ \partial_r \ln \alpha. \tag{5.91}$$

Plugging this result into Eq. (5.90) gives

$$dt_S = \sqrt{\frac{(1 - R\ n_{\hat{r}})^2 - \beta^2\ e^{2\lambda}}{1 - \frac{2\ M}{R}}}\ dt. \tag{5.92}$$

Then using the shift obtained from setting $\partial_t\ R = 0$ in the time independent problem, Eq. (5.60), we get

$$dt_S = \sqrt{\frac{(1 - R\ n_{\hat{r}})^2 - (R\ K_T)^2}{1 - \frac{2\ M}{R}}}\ dt. \tag{5.93}$$

But $(1 - R\ n_{\hat{r}})^2 - (R\ K_T)^2 \equiv 1 - 2M/R$ according to the first integral of the constraint equations, Eq. (5.57), so

$$dt_S = dt \quad \text{everywhere for stationary IVP.} \tag{5.94}$$

The procedure to find the relation between Schwarzschild time and coordinate time is somewhat more complicated for CMC initial conditions. This is because constant R does not imply constant r. Since $A_{\hat{r}}$ is kept close to zero by resetting, we derive this relation with the assumption that $A_{\hat{r}} = 0$. The proper time given in Eq. 5.89 is for an observer at constant r, θ, and ϕ. The proper time for non-constant R obtained from the Schwarzschild metric, Eq. 5.87, is

$$d\tau = \sqrt{\left(1 - \frac{2\,M}{R}\right)\, dt_S^2 - \frac{1}{1 - \frac{2\,M}{R}}\, dR^2}. \tag{5.95}$$

Re-arranging terms,

$$dt_S^2 = \frac{1}{1 - \frac{2\,M}{R}} \left[d\tau^2 + \frac{1}{1 - \frac{2\,M}{R}}\, dR^2 \right]. \tag{5.96}$$

From this we can derive a relation between the change in Schwarzschild time and the change in coordinate time:

$$\left(\frac{dt_S}{dt}\right)^2 = \frac{1}{1 - \frac{2\,M}{R}} \left[\left(\frac{d\tau}{dt}\right)^2 + \frac{1}{1 - \frac{2\,M}{R}} \left(\frac{dR}{dt}\right)^2 \right]. \tag{5.97}$$

We can now plug in expressions for $d\tau/dt$ and dR/dt from Eqs. 5.89 and 5.51, respectively. After simplication, the resulting expression is

$$\frac{dt_S}{dt} = \frac{1}{1 - \frac{2\,M}{R}} \left[\alpha(1 - R\, n_{\hat{r}}) + \beta\, e^{\lambda}\, (R\, K_T) \right]. \tag{5.98}$$

If we again use the shift obtained from setting $\partial_t\, R = 0$, Eq. (5.60), then

$$\frac{dt_S}{dt} = \frac{\alpha}{1 - R\, n_{\hat{r}}}. \tag{5.99}$$

Recall that we impose $\partial_t\, R = 0$ as our boundary conditions for the minimal strain resetting condition. So,

$$t_S = \int \frac{\alpha}{1 - R\, n_{\hat{r}}}\, dt \qquad \text{at the outer edge of the grid for CMC IVP.} \tag{5.100}$$

Since we are using geometrized units with $G = c = 1$, time has the same units as mass. Since $2M = 1$ in our calculations, all the times are multiplied by $2M$.

5.11.1 Stationary IVP

Testing the code using a stationary solution as the initial condition is the first step towards determining the viability of the code. Ideally, as the numerical evolution proceeds, all the variables stay constant. This does not occur in practice, however, because numerical errors act as perturbations to the stationary solution. These perturbations are likely to excite unstable modes of the analytic solution, some of which may be constraint violating modes (CVMs). In Sec. 5.11.1 below, the equations are linearized and studied, in order to determine if such unstable modes exist in the analytic solution. Numerical errors can also cause unstable modes to develop which are not a solution of the continuum equations. This kind of numerical error can be detected if there are features in the numerical solution which do not exhibit second order convergence to the stationary solution as the grid is refined.

With the stationary IVP, the only free parameter is K_T on the horizon (K_{TH}). We find that the best results for numerical calculations are obtained for $K_{THC} < K_{TH} < 0$, which give an asymptotically hyperbolic slice. For these values of K_{TH}, the congruence worldlines point out of the grid at both boundaries, resulting in an outgoing propagation of all seven eigenmodes at the inner boundary and five out of seven eigenmodes at the outer boundary.

Standard conditions for results with stationary IVP have the grid ranging from $-0.16 \leq r \leq 9.84$. Our initial grid is uniform in r, which we take equal to the proper radius by setting $\lambda = 0$. The event horizon is arbitrarily located at $r = 0$. For convergence studies, the coarse grid spacing is $dr = 0.02$, and the fine grid spacing is $dr = 0.01$. The time steps are such that $dt/dr = 1/2$. With all coordinate conditions, $\ln\alpha$, $\partial_r \ln\alpha$, β^r, and $\partial_r\beta^r$ are used as the (auxiliary) variables, where α is the lapse and β^r is the shift. With reset coordinate conditions, $\ln\alpha$, $\partial_r \ln\alpha$, β^r, and $\partial_r\beta^r$ are reset at every $dt = 0.01$ for both grid

resolutions in the convergence studies (with the shift and its derivative reset according to the minimal strain condition). The initial horizon values are: (1) $K_{TH} = -0.19$, $\alpha_H = 0.19$ (for the Nester gauge), and (2) $K_{TH} = -0.27$, $\alpha_H = 0.27$ (for the Lorentz gauge). Any deviations from the standard conditions are noted in the figure captions.

The true solution to the time independent problem using the Nester gauge is shown in Fig. 5.2. That with the Lorentz gauge is shown in Fig. 5.3. Recall that λ and $A_{\hat{r}}$ are zero initially, so are not shown.

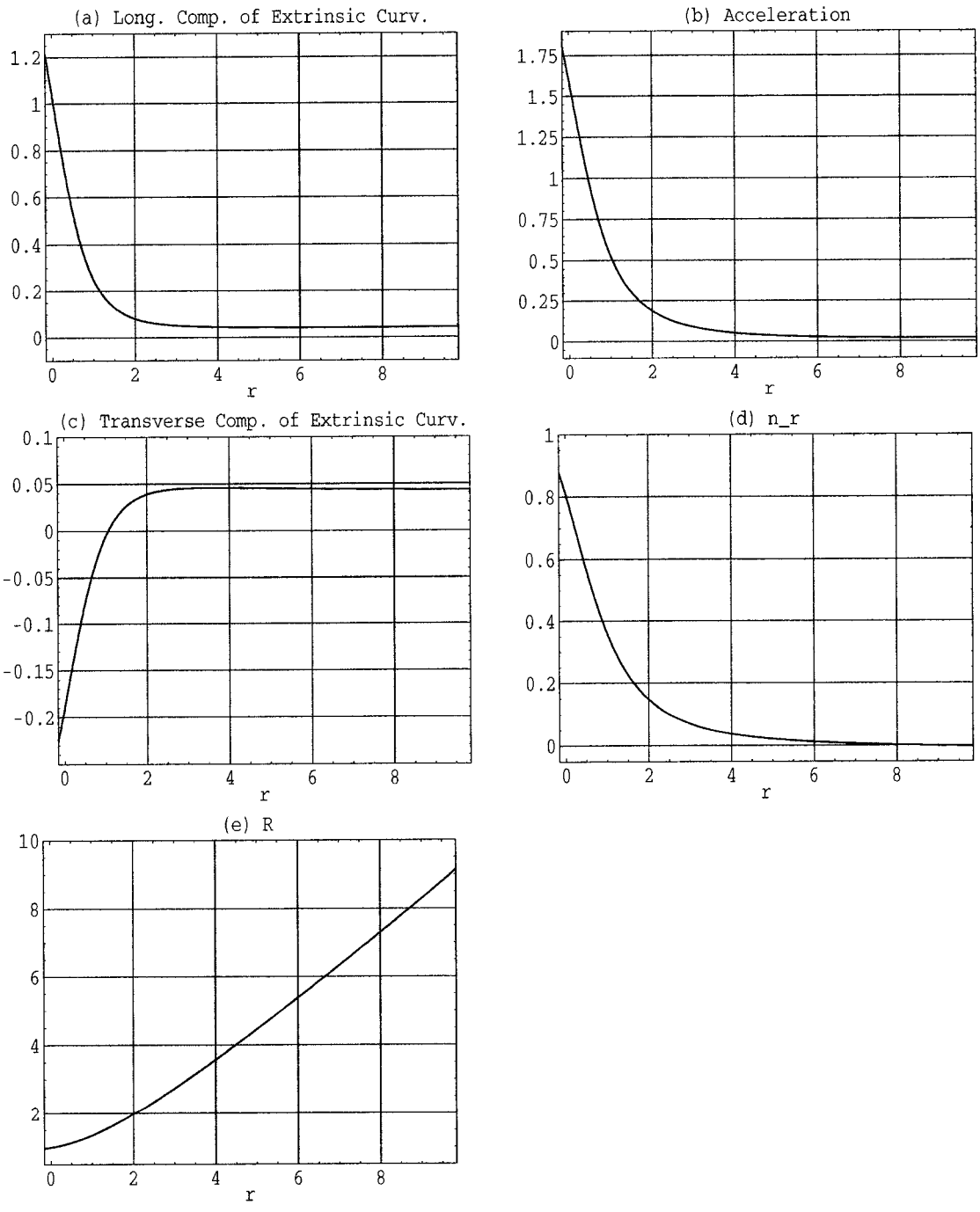


Figure 5.2: Time independent solution for Nester gauge.

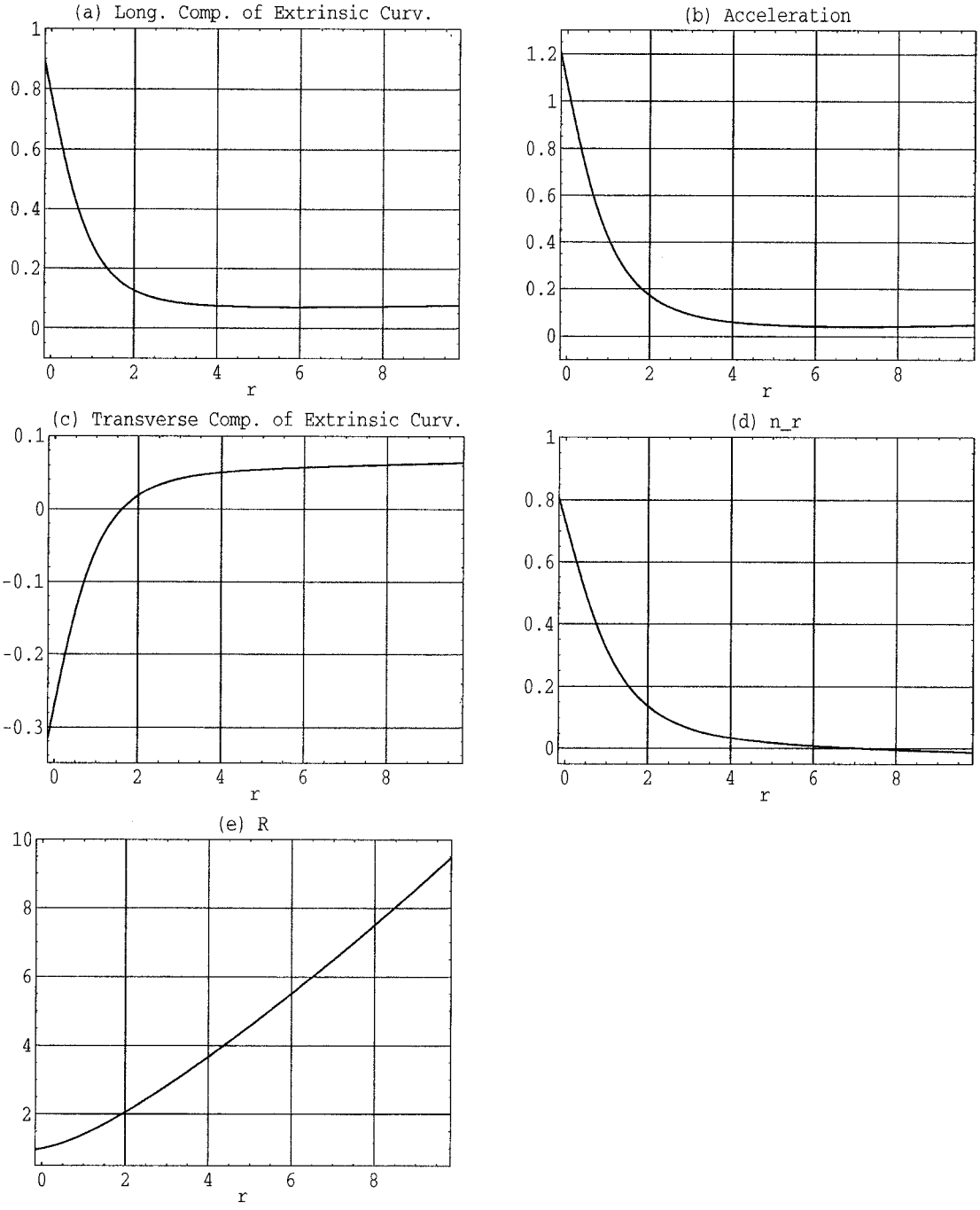


Figure 5.3: Time independent solution for Lorentz gauge.

Fixed Coordinate Conditions

With a fixed lapse and shift, the errors in the metric variables, λ, and $A_{\hat{r}}$, grow rapidly at the event horizon between times $t_S = 20M$ and $t_S = 40M$. This can be seen by comparing the convergence plots in Figs. 5.4 and 5.5. An unchecked growth in $A_{\hat{r}}$ is disastrous, because if $A_{\hat{r}}$ approaches 1, a coordinate singularity occurs. At least second order convergence is seen in these plots to $t_S = 40M$, since the results from $dr = 0.01$ (a blue line) overlay those from $dr = 0.02$ (a red line), when scaled by a factor of 4.

The growth of the constraint errors is shown in Fig. 5.6, to $t_S = 80M$. There is clearly exponential growth until at least $t_S = 20M$, and then an oscillatory behavior which is poorly understood. To try to understand whether this exponential growth is due to numerical errors or unstable analytic modes, we plot convergence of the constraint and longitudinal errors as a function of r at times $t_S = 20M$ and $t_S = 40M$ in Figs. 5.7, and 5.8, respectively. We see that convergence is quite good at both times, except near or inside the event horizion. This suggests that (1) there are numerical problems in our handling of the eigenmode propagations around the event horizon, and (2) there are both constraint violating and longitudinal analytic unstable modes, which show good second order convergence to the true solution. To confirm the presence of unstable analytic modes with fixed coordinate conditions, we perform a linear stablity analysis, described next.

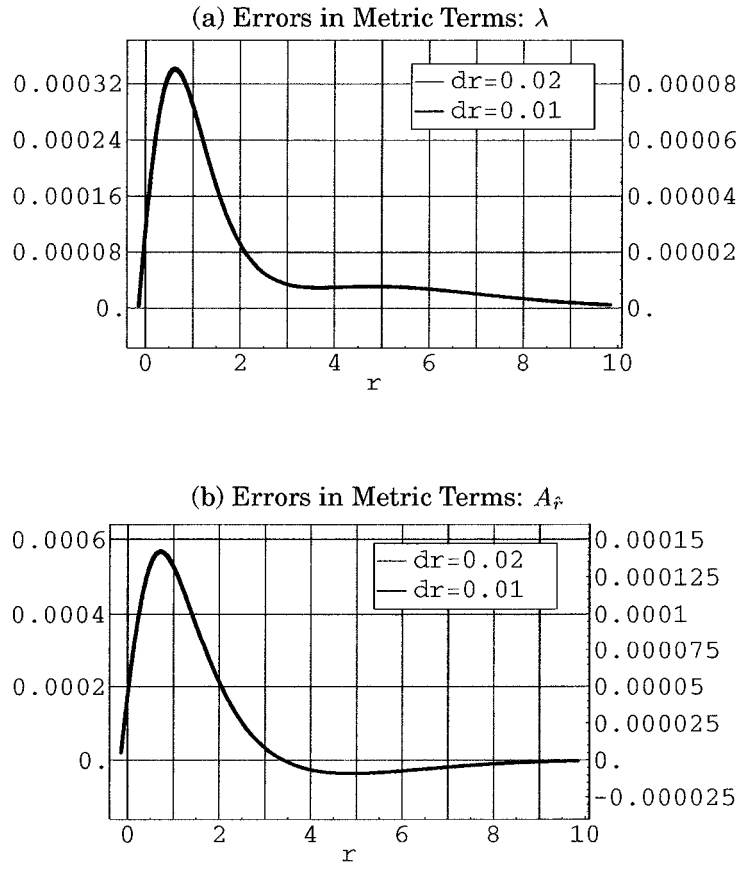


Figure 5.4: Convergence of errors in (a) λ and (b) $A_{\hat{r}}$ at $t_S = 20M$, for stationary IVP, Nester gauge, and fixed coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

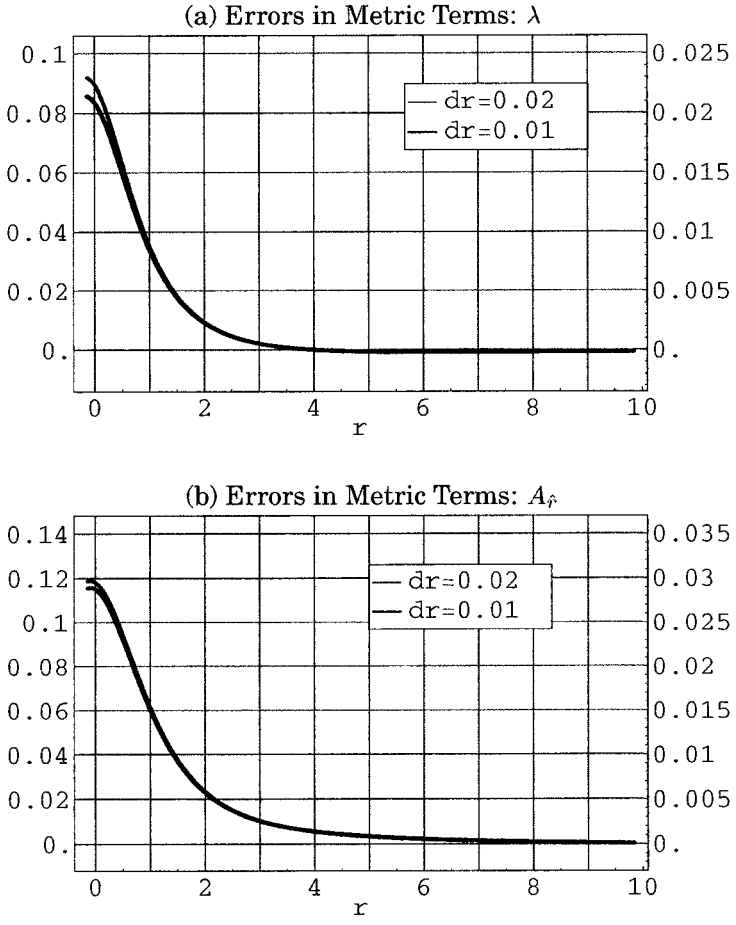


Figure 5.5: Convergence of errors in (a) λ and (b) $A_{\hat{r}}$ at $t_S = 40M$, for stationary IVP, Nester gauge, and fixed coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

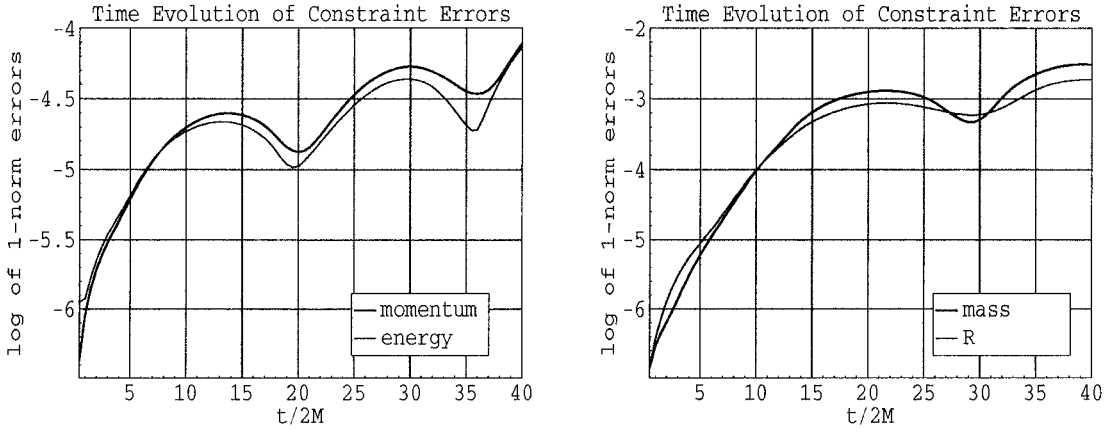


Figure 5.6: For stationary IVP, Nester gauge, and fixed coordinate conditions. Evaluated with a grid resolution $dr = 0.01$.

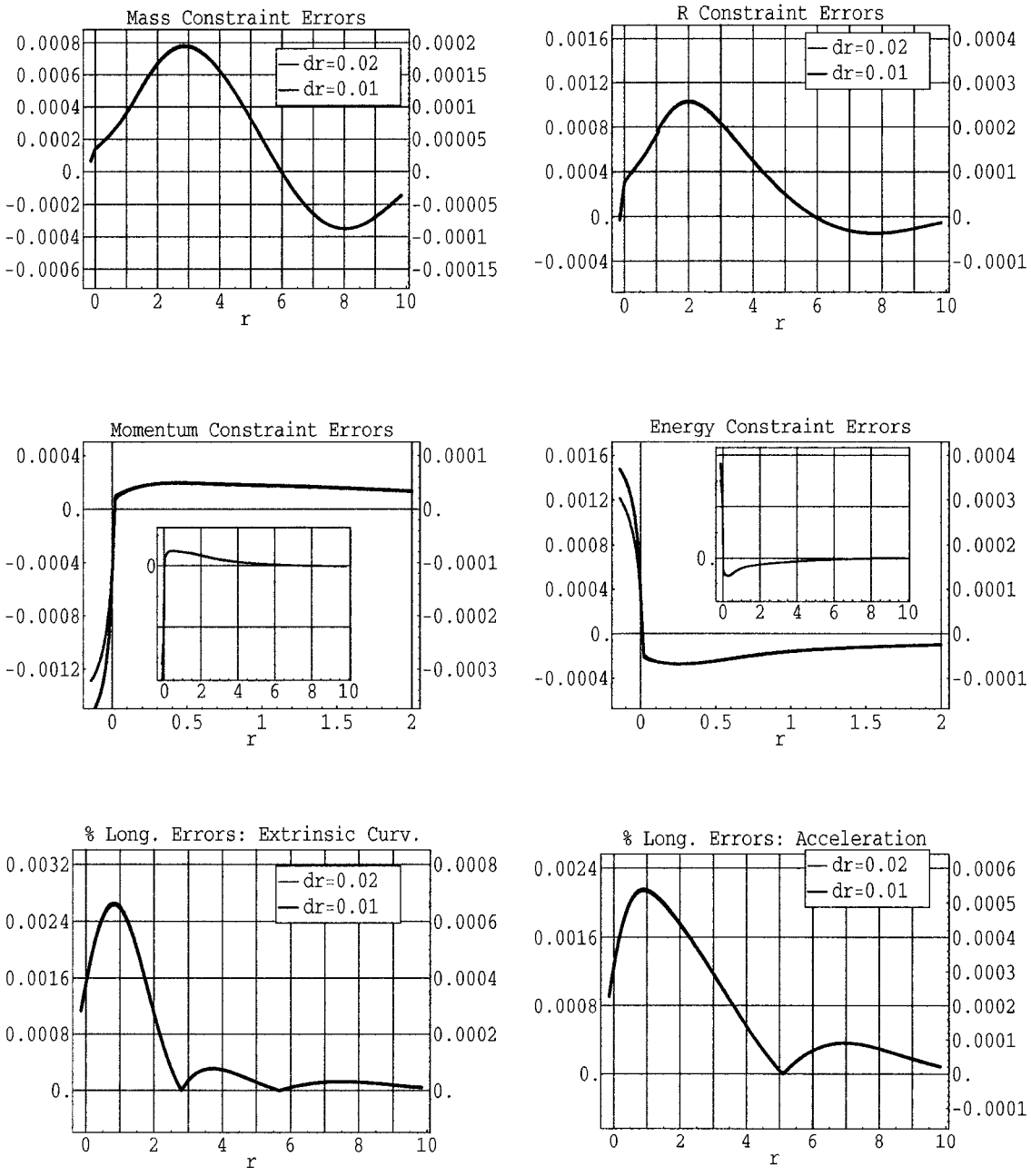


Figure 5.7: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 20M$, for stationary IVP, Nester gauge, and fixed coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

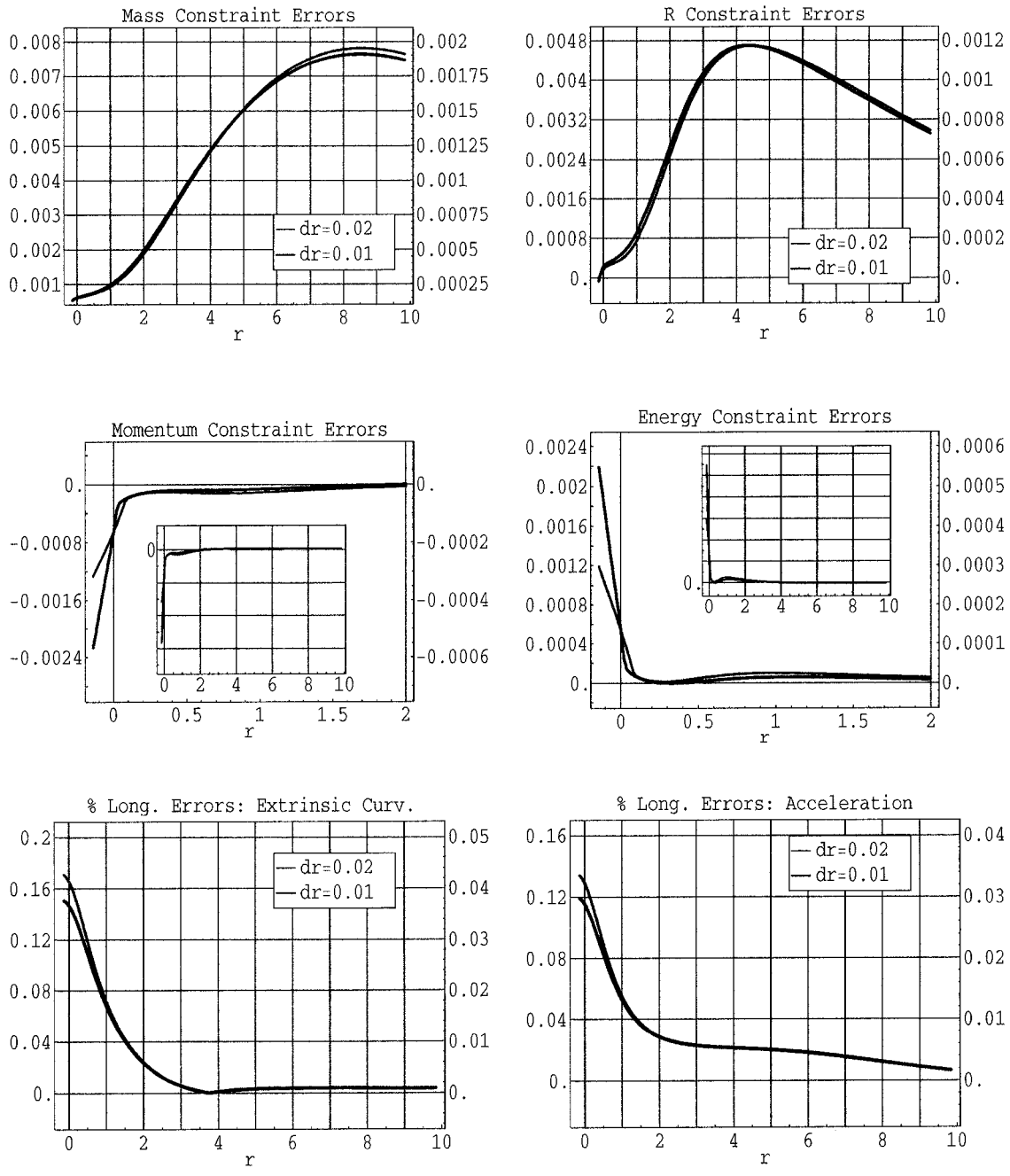


Figure 5.8: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 40M$, for stationary IVP, Nester gauge, and fixed coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

• **Linearization**

A linear stability analysis is performed, in order to study the nature of the growth rate of the eigenmodes in the simplest and cleanest possible way. The equations (with the Nester gauge) are linearized about the stationary solution. The solution is then analyzed, both analytically and numerically, to determine whether or not the perturbations about the stationary solution develop into exponentially growing modes.

Let $\boldsymbol{q}$ be the vector of variables, $\boldsymbol{q}_0$ the steady state solution, and $\boldsymbol{q}_1$ the perturbation. Then,

$$\boldsymbol{q} = \boldsymbol{q}_0 + \xi\, \boldsymbol{q}_1. \tag{5.101}$$

Substituting this into the true constraint Eqs. (5.48) and (5.49), and the true evolution Eqs. (5.44), (5.45), (5.46), and (5.47), and keeping only terms first order in ξ, we obtain the linearized true evolution equations:

$$\partial_t\,(a_{\hat{r}1} - K_{R1}) + (\alpha - \beta^r)\;\partial_r\,(a_{\hat{r}1} - K_{R1}) =$$
$$2\,\alpha\left[\frac{K_{R1} - K_{T1} - n_{\hat{r}1}}{R_0} - a_{\hat{r}0}\,a_{\hat{r}1} + K_{R0}\,K_{R1} - K_{T0}\,K_{T1} + n_{\hat{r}0}\,n_{\hat{r}1}\right], \tag{5.102}$$

$$\partial_t\,(n_{\hat{r}1} - K_{T1}) + (\alpha - \beta^r)\;\partial_r\,(n_{\hat{r}1} - K_{T1}) = \alpha(a_{\hat{r}0} + n_{\hat{r}0} + K_{R0} - K_{T0})(n_{\hat{r}1} - K_{T1})$$
$$+\;\alpha\left[(a_{\hat{r}1} + n_{\hat{r}1} + K_{R1} - K_{T1})(n_{\hat{r}0} - K_{T0} - \frac{1}{R_0})\right], \tag{5.103}$$

$$\partial_t\,(a_{\hat{r}1} + K_{R1}) - (\alpha + \beta^r)\;\partial_r\,(a_{\hat{r}1} + K_{R1}) =$$
$$2\,\alpha\left[\frac{K_{R1} - K_{T1} + n_{\hat{r}1}}{R_0} + a_{\hat{r}0}\,a_{\hat{r}1} - K_{R0}\,K_{R1} + K_{T0}\,K_{T1} - n_{\hat{r}0}\,n_{\hat{r}1}\right], \tag{5.104}$$

$$\partial_t\,(n_{\hat{r}1} + K_{T1}) - (\alpha + \beta^r)\;\partial_r\,(n_{\hat{r}1} + K_{T1}) = \alpha(-a_{\hat{r}0} - n_{\hat{r}0} + K_{R0} - K_{T0})(n_{\hat{r}1} + K_{T1})$$
$$-\;\alpha\left[(a_{\hat{r}1} + n_{\hat{r}1} - K_{R1} + K_{T1})(n_{\hat{r}0} + K_{T0} - \frac{1}{R_0})\right]. \tag{5.105}$$

The linearized true constraint equations are

$$\partial_r K_{T1} = n_{\hat{r}0}\,(K_{T1} - K_{R1}) + n_{\hat{r}1}\,(K_{T0} - K_{R0}) + \frac{K_{R1} - K_{T1}}{R_0}, \tag{5.106}$$

$$\partial_r n_{\hat{r}1} = -K_{R0}\,K_{T1} - K_{T0}\,(K_{R1} + K_{T1}) + 3\,n_{\hat{r}0}\,n_{\hat{r}1} - \frac{2\,n_{\hat{r}1}}{R_0}. \tag{5.107}$$

Here, the metric perturbations λ_1, R_1, and $A_{\hat{r}1}$ have been set to zero. This is done so that the contributions of perturbations in the fundamental variables to the development of unstable modes can be isolated. The code for the full non-linear equations, with conditions as described in Sec. 5.11.1 for the Nester gauge, has been modified to accommodate the linearized equations. The boundary conditions are implemented in the same way as described in Sec. 5.9, using the linearized constraint equations. The initial perturbation is set equal to $-10^{-6}\times$ the steady state solution. The code has been tested for exponential growth of the perturbations by plotting $a_{\hat{r}1}$, $n_{\hat{r}1}$, K_{R1}, and K_{T1} versus time in Fig. 5.9.

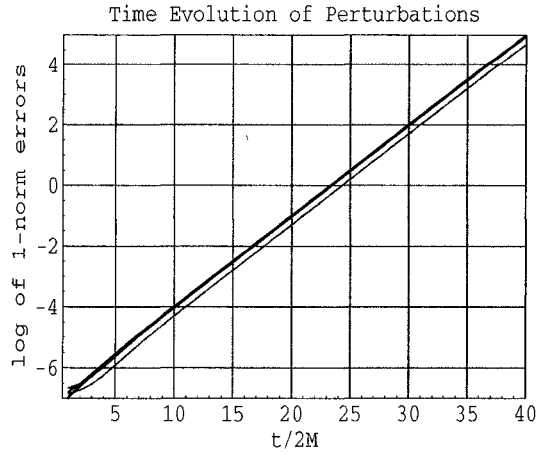


Figure 5.9: Exponential growth of 1-norms of perturbations in numerical solution to linearized equations, for stationary IVP, Nester gauge, and fixed coordinate conditions. Calculated with $dr = 0.01$. The red line is K_{R1}, and the green, blue, and black lines are $a_{\hat{r}1}$, K_{T1}, and $n_{\hat{r}1}$, respectively.

We find that an unstable mode develops in the numerical solution for all

four fundametal variables. The shape of this mode as a function of r is shown at times $t1 = 24M$ and $t2 = 40M$ in Fig. 5.10. The results at the later time, shown in blue, have been re-scaled by a factor of 250, and they exactly coincide with the results at the earlier time (plotted in red). This indicates that the unstable mode does not propagate, but grows in place. Because the perturbations grow exponentially in place, the growth rate τ can be determined from the numerical results in Fig. 5.10 by the equation

$$\boldsymbol{q}_1(r, t_2) = \boldsymbol{q}_1(r, t_1)\, e^{\tau\, \Delta t}. \tag{5.108}$$

In this equation, t_1 and t_2 are the two times at which the perturbations are plotted (versus proper distance, r), and $\Delta t = t_2 - t_1 = 16M$. The numerically obtained growth rate is: $\tau \approx 0.345/M$. The results in these figures are independent of the type of numerical method used in the transport step (pure Beam-Warming versus a weighted average of Lax-Wendroff and Beam-Warming have been tried); hence, the unstable mode is not from numerical problems, but is a non-physical, analytic solution to the linearized equations.

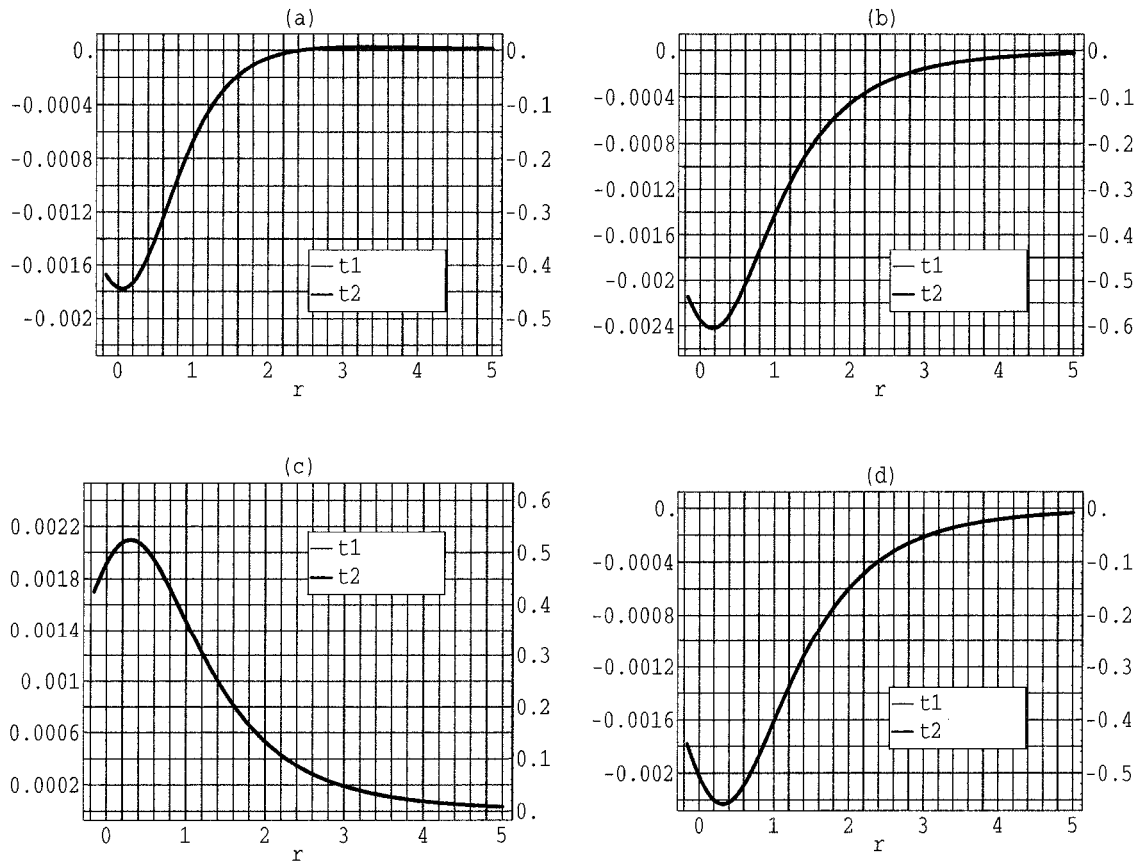


Figure 5.10: Shape of perturbations in numerical solution to linearized equations, and times $t1 = 24M$ and $t2 = 40M$. For stationary IVP, Nester gauge, and fixed coordinate conditions, and calculated with $dr = 0.01$. Shown are (a) $K_{\hat{R}1}$, (b) $a_{\hat{r}1}$, (c) K_{T1}, and (d) $n_{\hat{r}1}$. Notice that the blue scale is $250\times$ the red scale.

The linearized Eqs. (5.102), (5.103), (5.104), and (5.105) can now be analyzed straightforwardly. Assuming an exponential time dependence of the perturbations, we are left with four coupled linear ordinary differential equations. If we make the approximation that the radial derivatives of the perturbations $a_{\hat{r}1}$, $n_{\hat{r}1}$, K_{R1}, and K_{T1} vanish simultaneously in the numerical plots at approximately $r = 0.2$ (which is a rough approximation, as can be seen by examining the plots), then we can set the radial derivatives to zero in the ODE's. Doing so results in an algebraic set of equations, the eigenmodes of which are found by substituting the values of the unperturbed quantities at $r = 0.2$ into the equations. An eigenmode with a positive eigenvalues indicates an analytic unstable mode, with the growth rate equal to the eigenvalue. For this problem, the eigenvalue equation is

$$\boldsymbol{A} \cdot \boldsymbol{q}_1(t_1) = \tau\, \boldsymbol{q}_1(t_1), \tag{5.109}$$

with

$$\boldsymbol{A} = \begin{pmatrix} -0.4167 & 0.6726 & -0.07261 & 0.1238 \\ 0.4870 & 0 & -0.4870 & 0 \\ -0.03631 & 0.06192 & 0.2810 & -0.4560 \\ -0.06192 & 0.036305 & -0.4560 & 0.2810 \end{pmatrix}, \tag{5.110}$$

$$\boldsymbol{q}_1(t_1) = \begin{pmatrix} K_{R1}(t_1) \\ a_{\hat{r}1}(t_1) \\ K_{T1}(t_1) \\ n_{\hat{r}1}(t_1) \end{pmatrix}, \tag{5.111}$$

and eigenvalues, τ. The computed eigenvalues are approximately -0.78, 0.71, 0.43, and -0.21. This means that analytically, there are two growing modes and two decaying modes. The mode with growth rate $\tau \approx 0.71 = 0.355/M$ confirms the numerical result.

Resetting Coordinate Conditions–Nester gauge

Resetting the lapse keeps $A_{\hat{r}}$ small, as can be seen by comparing Figs. 5.11b and 5.12b for resetting coordinate conditions with Figs. 5.4b and 5.5b for fixed coordinate conditions. In this way, the coordinate system is prevented from developing singular behavior. Comparing parts (a) of these figures, we see that λ grows much less rapidly between times $20M$ and $40M$ with reset coordinate conditions than it does with fixed coordinate conditions. Convergence with reset conditions at $t_S = 20M$ is not as good as with fixed conditions, but certainly does not worsen by $t_S = 40M$.

The growth of the constraint errors is shown in Fig. 5.13, to $t_S = 140M$. As with fixed conditions, there is clear exponential growth of the 1-norms to about $t_S = 30M$, and then there are oscillations. Different numerical methods give similar behavior [12], but with different amplitudes, periods, and growth of oscillations, especially after about $t_S = 40M$. We conclude that the behavior of the constraint errors is complicated and not well understood after this time.

Convergence plots of the constraint and longitudinal errors as a function of r are given in Figs. 5.14 and 5.15. At $t_S = 20M$, the magnitudes and shapes of the errors are similar for fixed and reset coordinate conditions, except for the dominant peak of the longitudinal errors. By $t_S = 40M$, the errors in the longitudinal variables have increased substantially at or inside the event horizon with fixed coordinate conditions whereas with reset coordinate conditions, the growth of these errors is suppressed.

Convergence breaks down around $40M < t_S \leq 60M$. (Second order convergence must be demonstrated in all of the constraints to have valid results). Thus, we believe that the error estimates shown in Fig. 5.13 are reliable to about $t_S = 40M$.

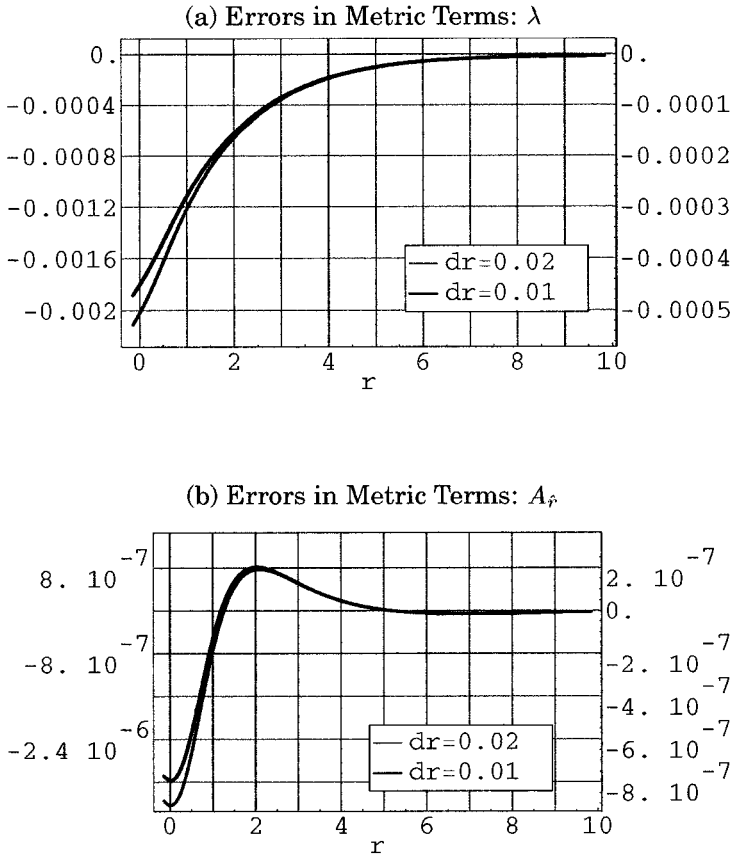


Figure 5.11: Convergence of errors in (a) λ and (b) $A_{\hat{r}}$ at $t_S = 20M$, for stationary IVP, Nester gauge, and reset coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

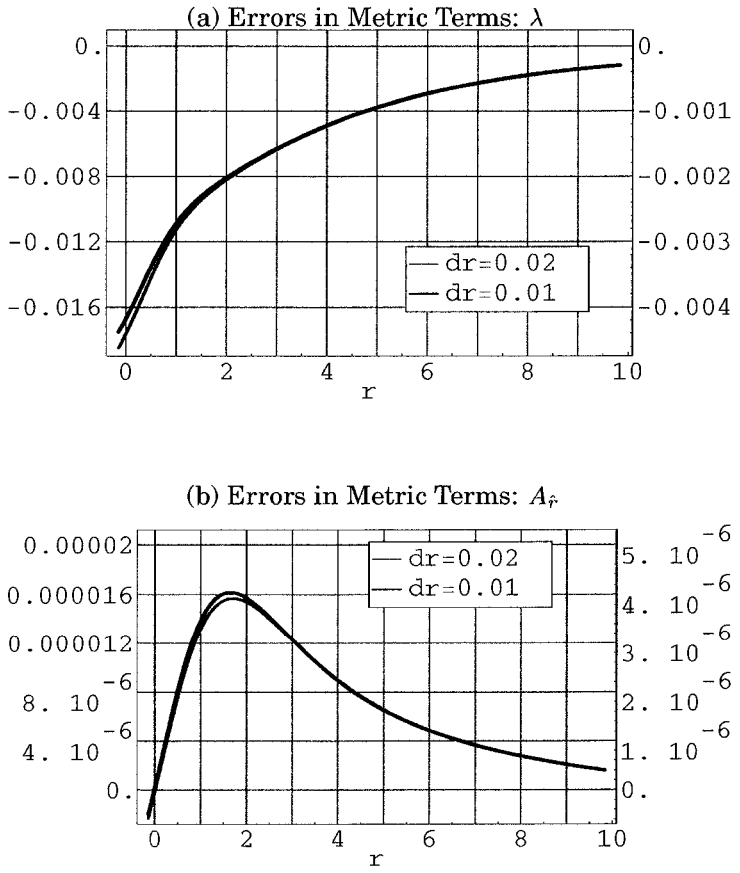


Figure 5.12: Convergence of errors in (a) λ and (b) $A_{\hat{r}}$ at $t_S = 40M$, for stationary IVP, Nester gauge, and reset coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

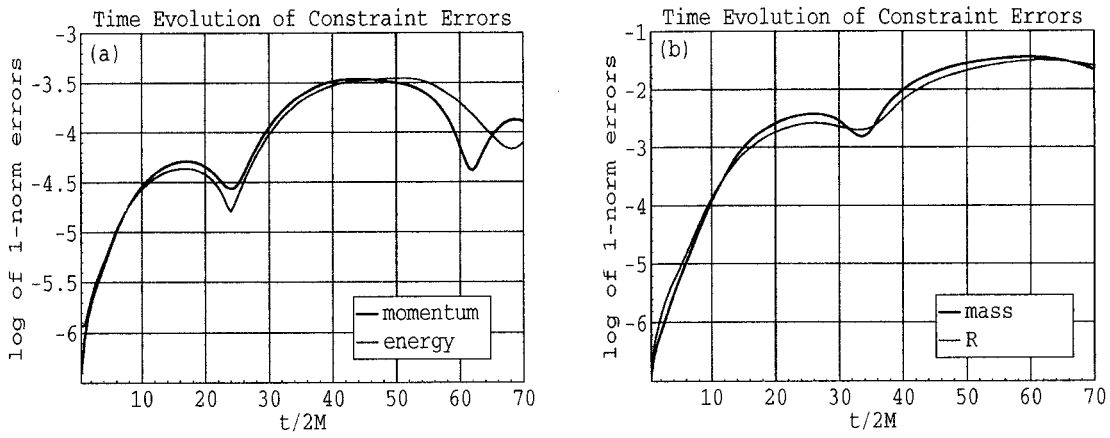


Figure 5.13: For stationary IVP, Nester gauge, and reset coordinate conditions. Evaluated with a grid resolution $dr = 0.01$. $\ln \alpha$, $\partial_r \ln \alpha$, β^r, and $\partial_r \beta^r$ are reset every $dt = 0.005$.

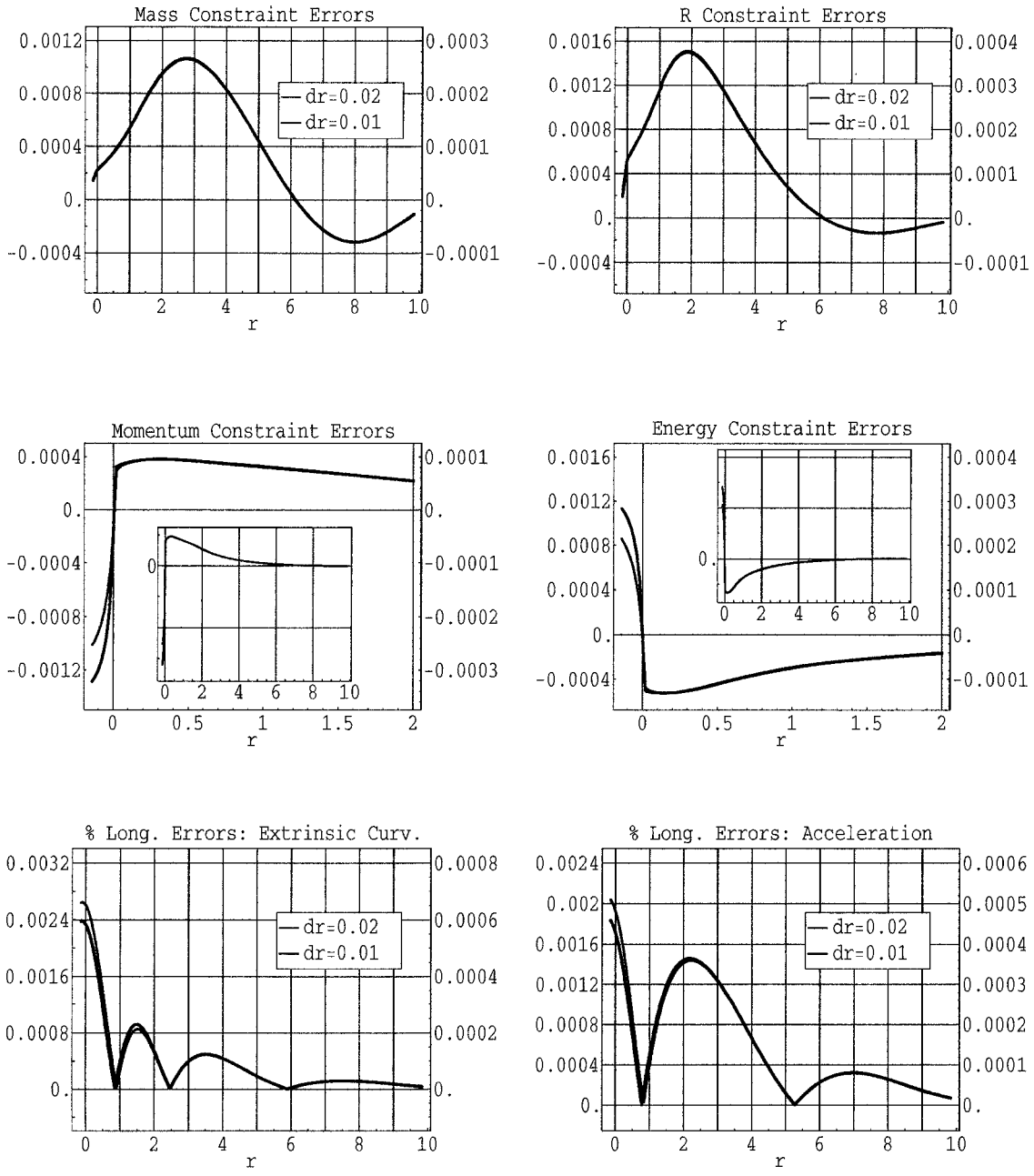


Figure 5.14: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 20M$, for stationary IVP, Nester gauge, and reset coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

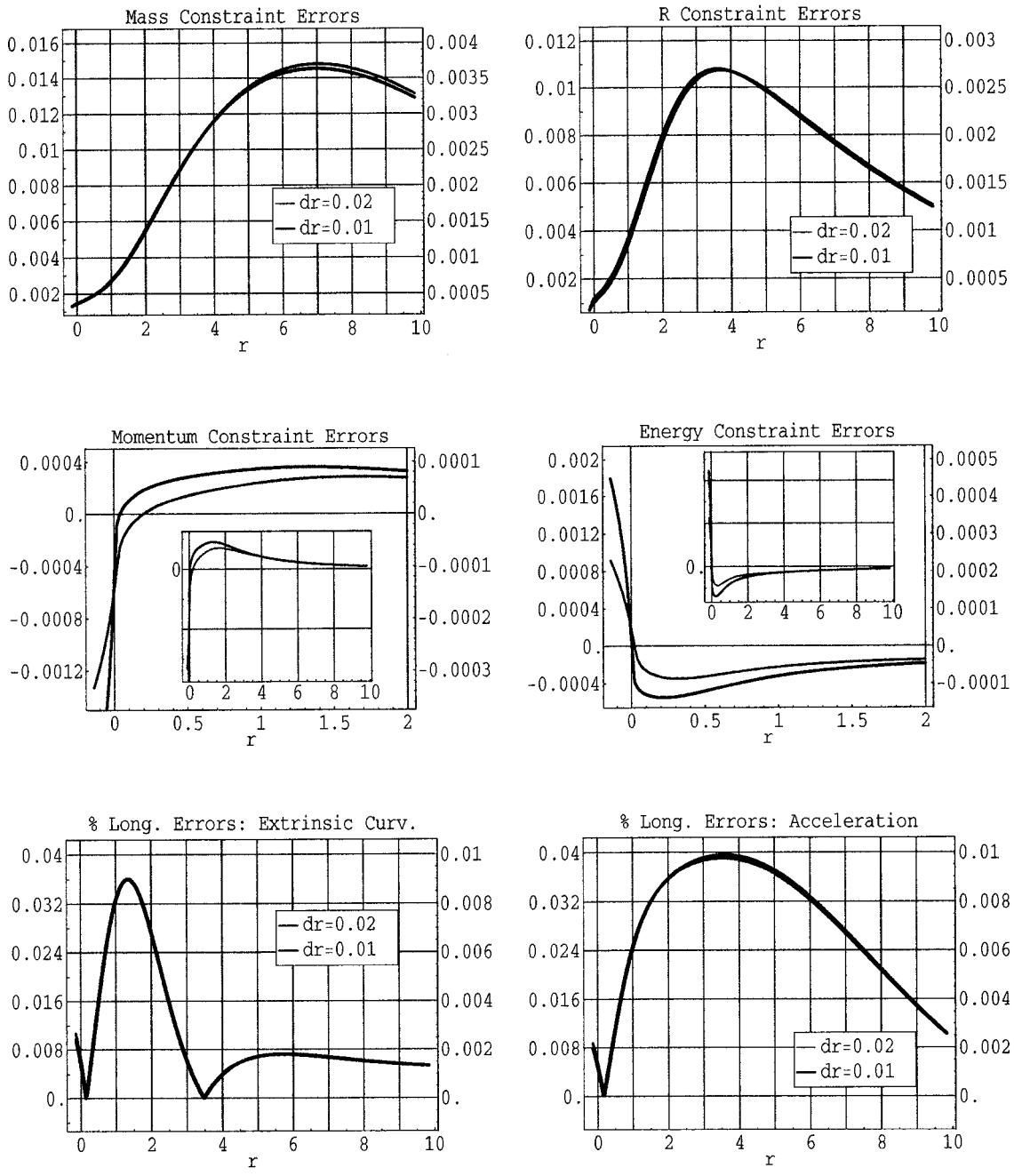


Figure 5.15: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 40M$, for stationary IVP, Nester gauge, and reset coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

For the Nester gauge, $K_{THC} = -0.25$. $K_{TH} = -0.19$ is chosen for these figures because that is about where optimal results are obtained. The increase in errors with a different value of K_{TH} is seen by comparing Figs. 5.15 (at $K_{TH} = -0.19$) and 5.16 (at $K_{TH} = -0.22$).

The sensitivity of the results to the placement of the inner and outer boundaries is tested in Figs. 5.17, 5.18, 5.19, and 5.20. Extending the outer boundary as in Fig. 5.17 makes no significant difference at all. Moving the location of the inner boundary does not make a noticeable difference in the magnitude or convergence of any of the constraint errors. However, it does produce a pattern in the magnitude of the longitudinal errors at the inner edge. As the inner boundary is moved closer to the event horizon, the longitudinal errors at the inner boundary rise (compare Figs. 5.14 and 5.18). Likewise, as the inner boundary is moved away from the event horizon, the longitudinal errors at the inner boundary decrease (see Figs. 5.19 and 5.20).

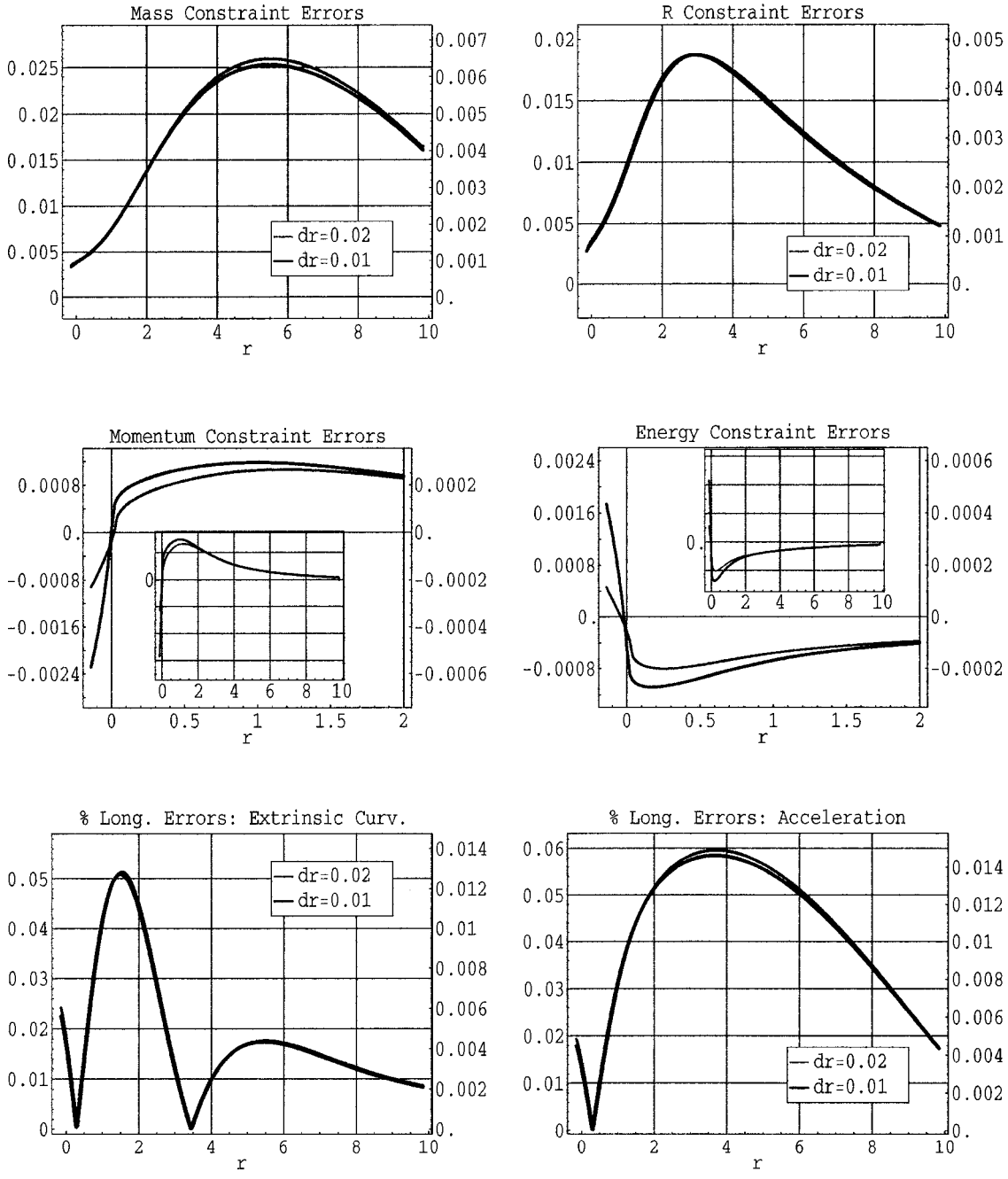


Figure 5.16: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 40M$, for stationary IVP, Nester gauge, and reset coordinate conditions. The initial horizon values are: $K_{TH} = -0.22$, $\alpha_H = 0.22$. Notice that the red scale is $4\times$ the blue scale.

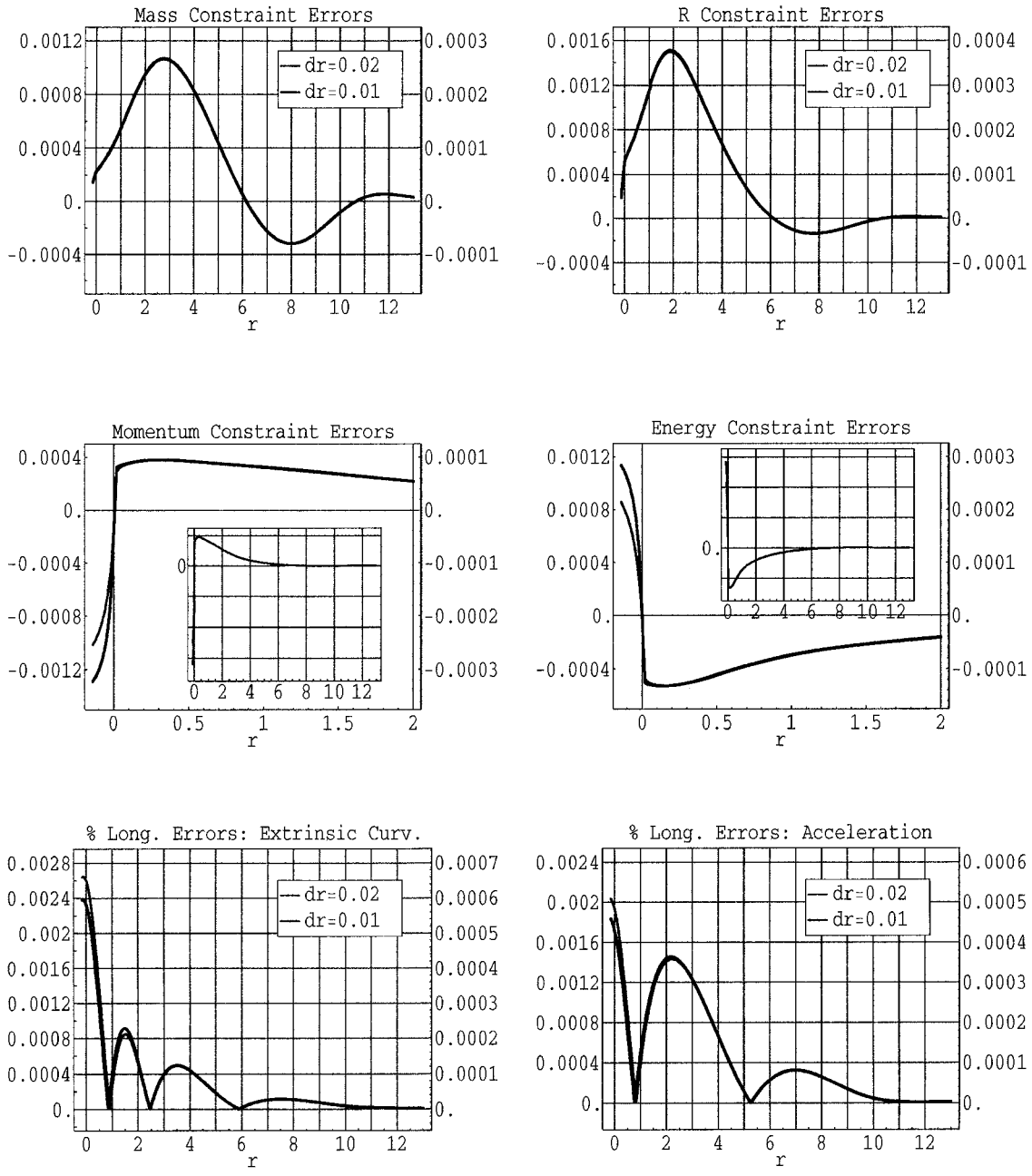


Figure 5.17: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 20M$, for stationary IVP, Nester gauge, and reset coordinate conditions. This run tests the sensitivity of the code to the location of the outer edge, which has been extended so that $-0.16 \leq r \leq 13.0$. Notice that the red scale is $4\times$ the blue scale.

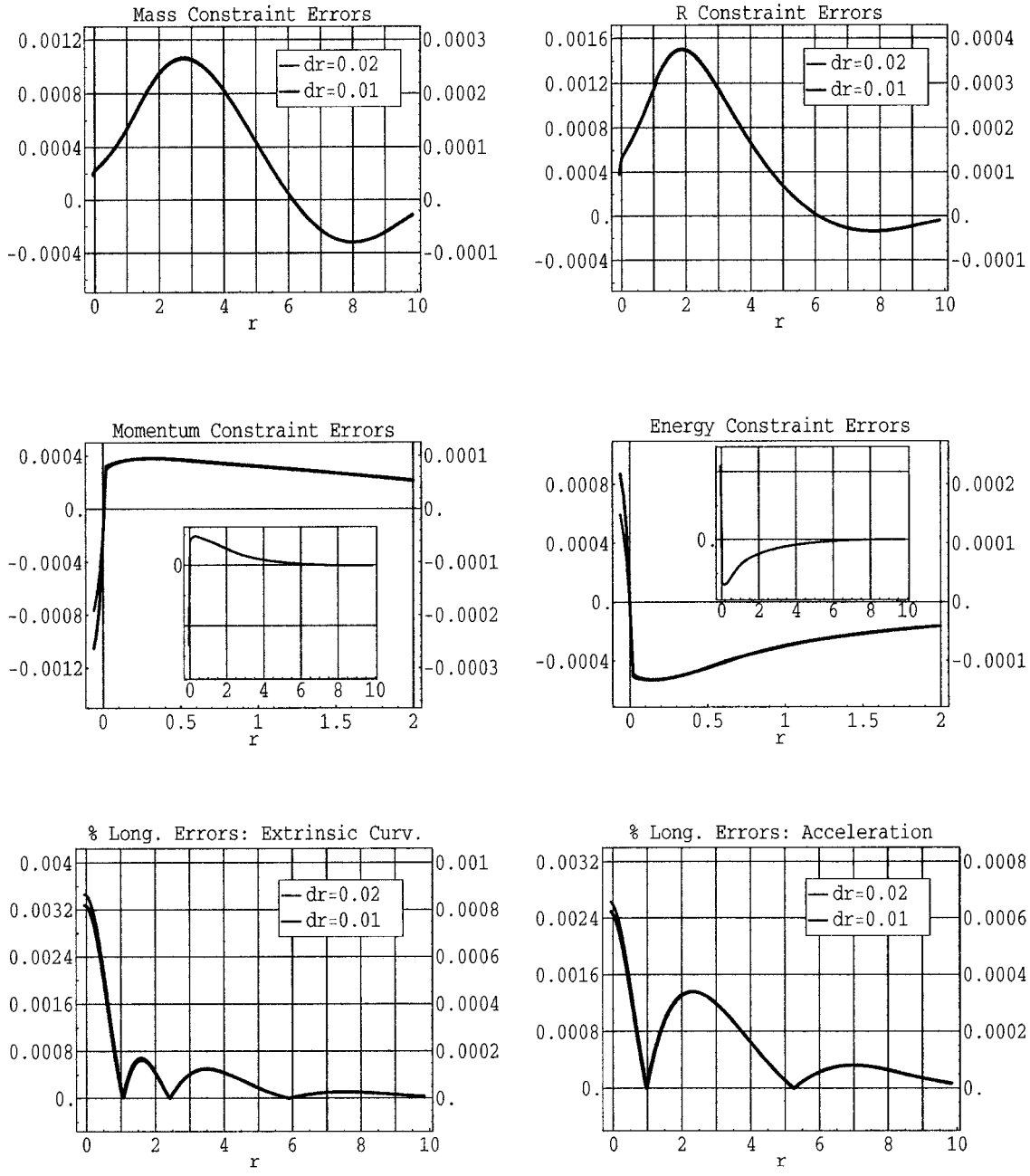


Figure 5.18: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 20M$, for stationary IVP, Nester gauge, and reset coordinate conditions. This run tests the sensitivity of the code to the location of the inner edge, which has been brought in so that $-0.08 \leq r \leq 9.84$. Notice that the red scale is $4\times$ the blue scale.

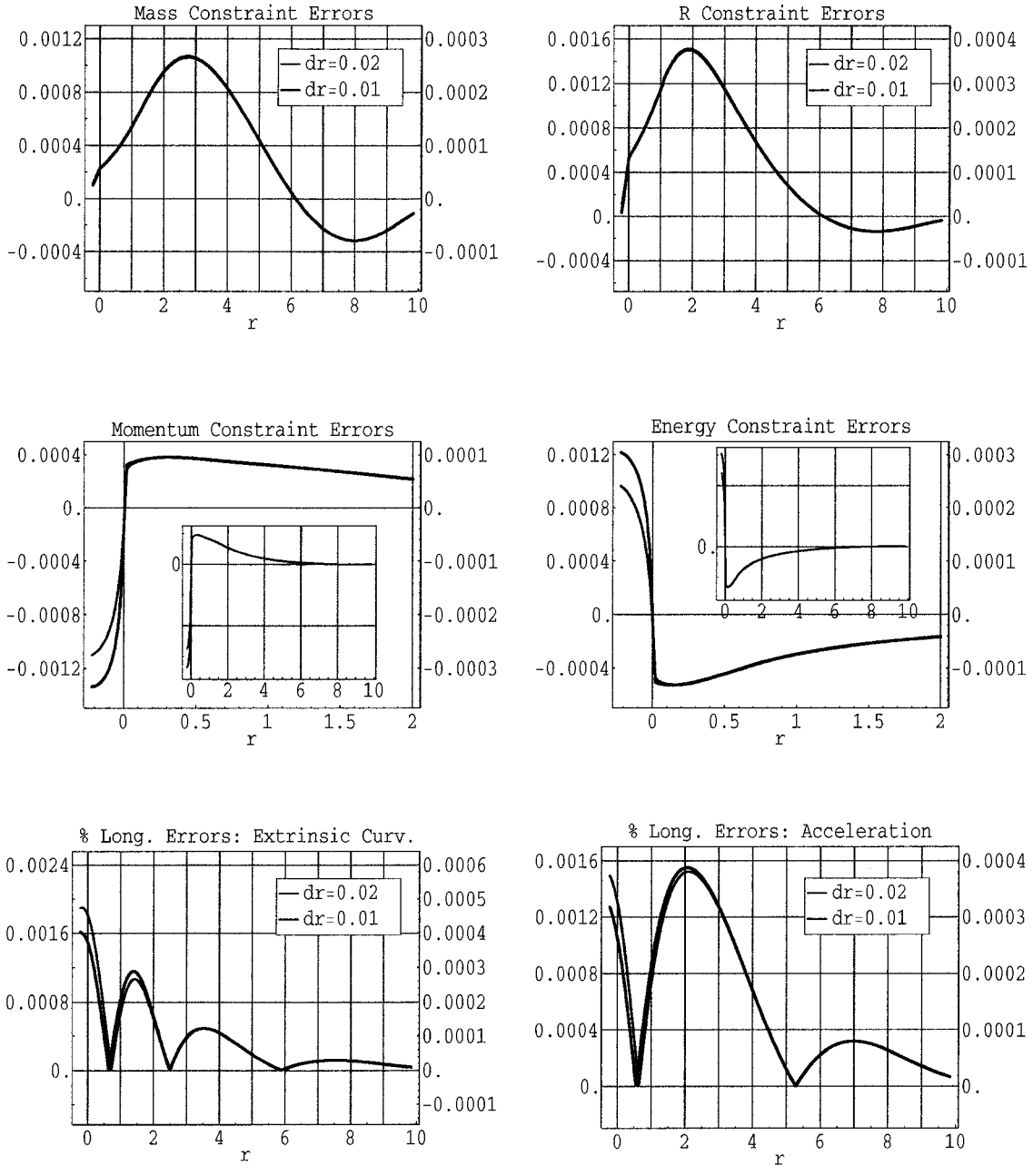


Figure 5.19: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 20M$, for stationary IVP, Nester gauge, and reset coordinate conditions. This run tests the sensitivity of the code to the location of the inner edge, which has been extended so that $-0.24 \leq r \leq 9.84$. Notice that the red scale is $4\times$ the blue scale.

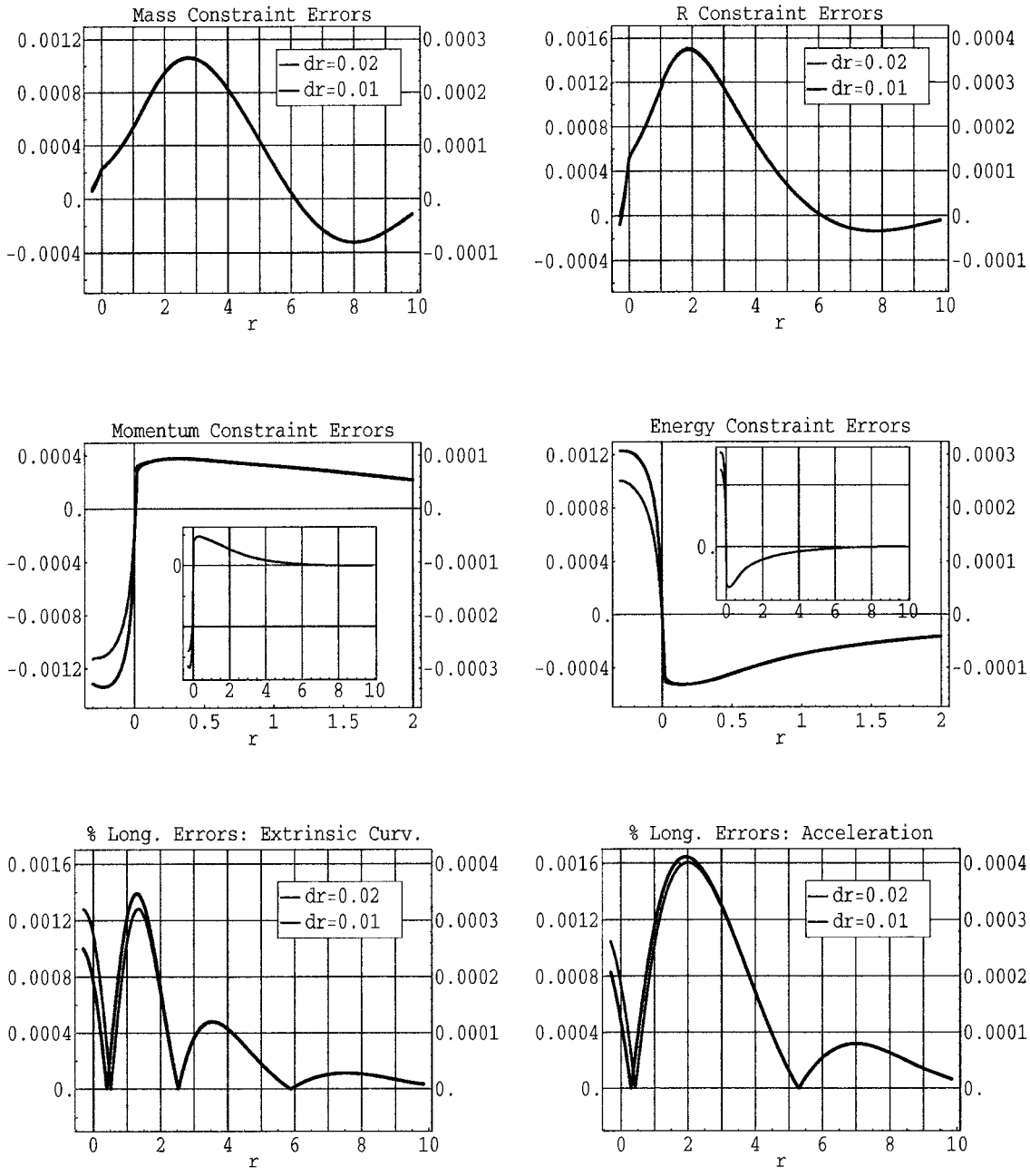


Figure 5.20: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 20M$, for stationary IVP, Nester gauge, and reset coordinate conditions. This run tests the sensitivity of the code to the location of the inner edge, which has been extended so that $-0.32 \leq r \leq 9.84$. Notice that the red scale is $4\times$ the blue scale.

Resetting Coordinate Conditions–Lorentz gauge

The growth of both $A_{\hat{r}}$ and λ between times $20M$ and $40M$ is much greater with the Lorentz gauge than with the Nester gauge, with reset coordinate conditions. This can be seen by comparing Figs. 5.21 and 5.22 with Figs. 5.11 and 5.12, respectably. The growth with the Lorentz gauge occurs markedly at the inner boundary. The second order convergence to the true solution, however, is similar with both gauges.

The growth of the 1-norms of the constraint errors is shown in Fig. 5.23, to $t_S = 60M$. There is clear exponential growth, as with the Nester gauge. Hence, the code is unstable and cannot be run for long times. At $t_S = 60M$, the 1-norm constraint errors are at least $10\times$ larger with the Lorentz gauge than with the Nester gauge.

Convergence plots of the constraint and longitudinal errors as a function of r are given in Figs. 5.24 and 5.25. As with the Nester gauge, the errors show good, second order convergence outside the event horizon to approximately $t_S = 40M$. Inside and just outside the event horizon (to $r \approx 0.5$ for the Lorentz gauge and to $r \approx 1.5$ for the Nester gauge), second order convergence of the energy and momentum constraint errors degrades significantly by $t_S = 40M$. The magnitudes of all the errors are considerably larger with the Lorentz gauge than with the Nester gauge at $t_S = 40M$. This is especially true for the energy and momentum constraint errors, which rise to a peak 15 to $20\times$ higher for the Lorentz gauge than for the Nester gauge just outside the event horizon.

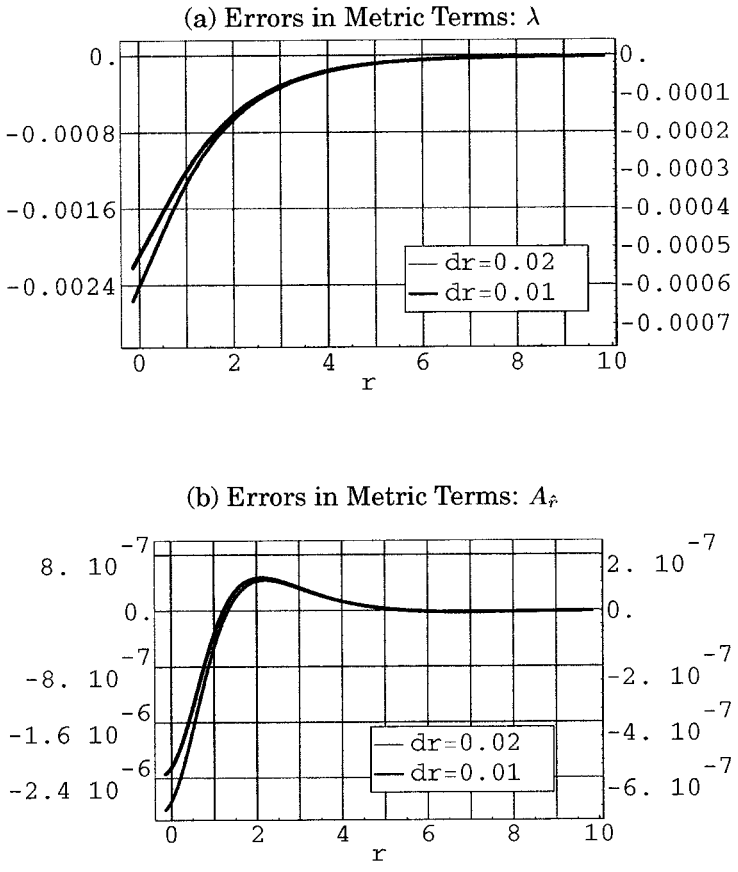


Figure 5.21: Convergence of errors in (a) λ and (b) $A_{\hat{r}}$ at $t_S = 20M$, for stationary IVP, Lorentz gauge, and reset coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

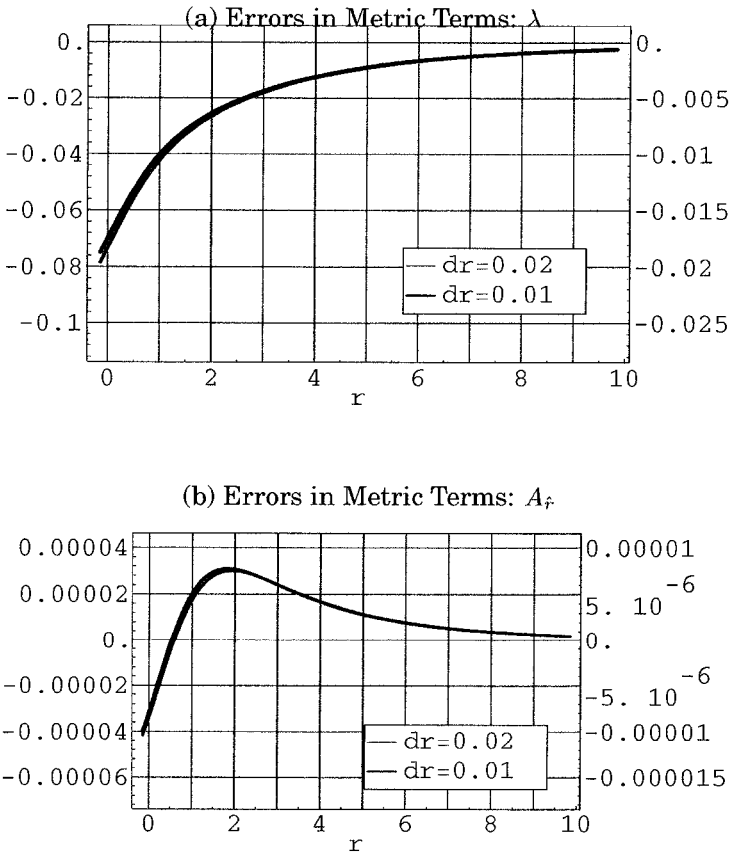


Figure 5.22: Convergence of errors in (a) λ and (b) $A_{\hat{r}}$ at $t_S = 40M$, for stationary IVP, Lorentz gauge, and reset coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

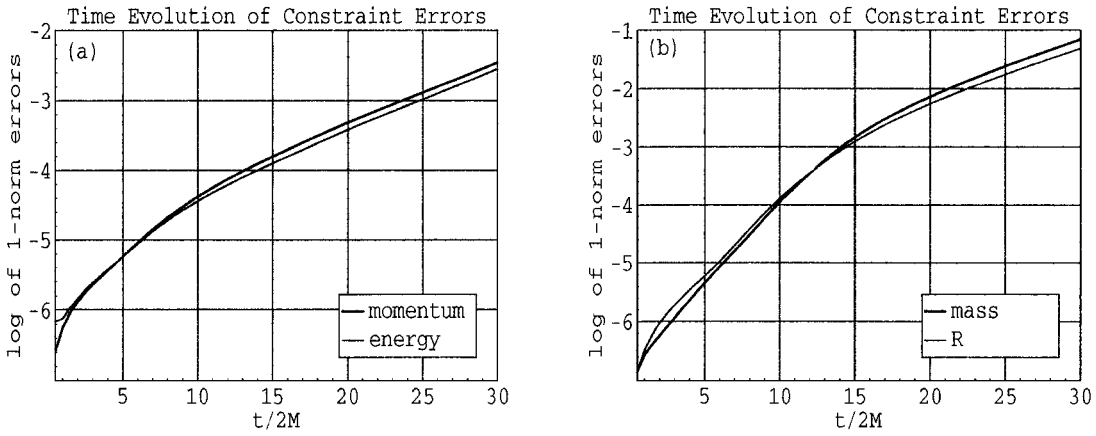


Figure 5.23: For stationary IVP, Lorentz gauge, and reset coordinate conditions. Evaluated with a grid resolution $dr = 0.01$. $\ln \alpha$, $\partial_r \ln \alpha$, β^r, and $\partial_r \beta^r$ are reset at every $dt = 0.005$.

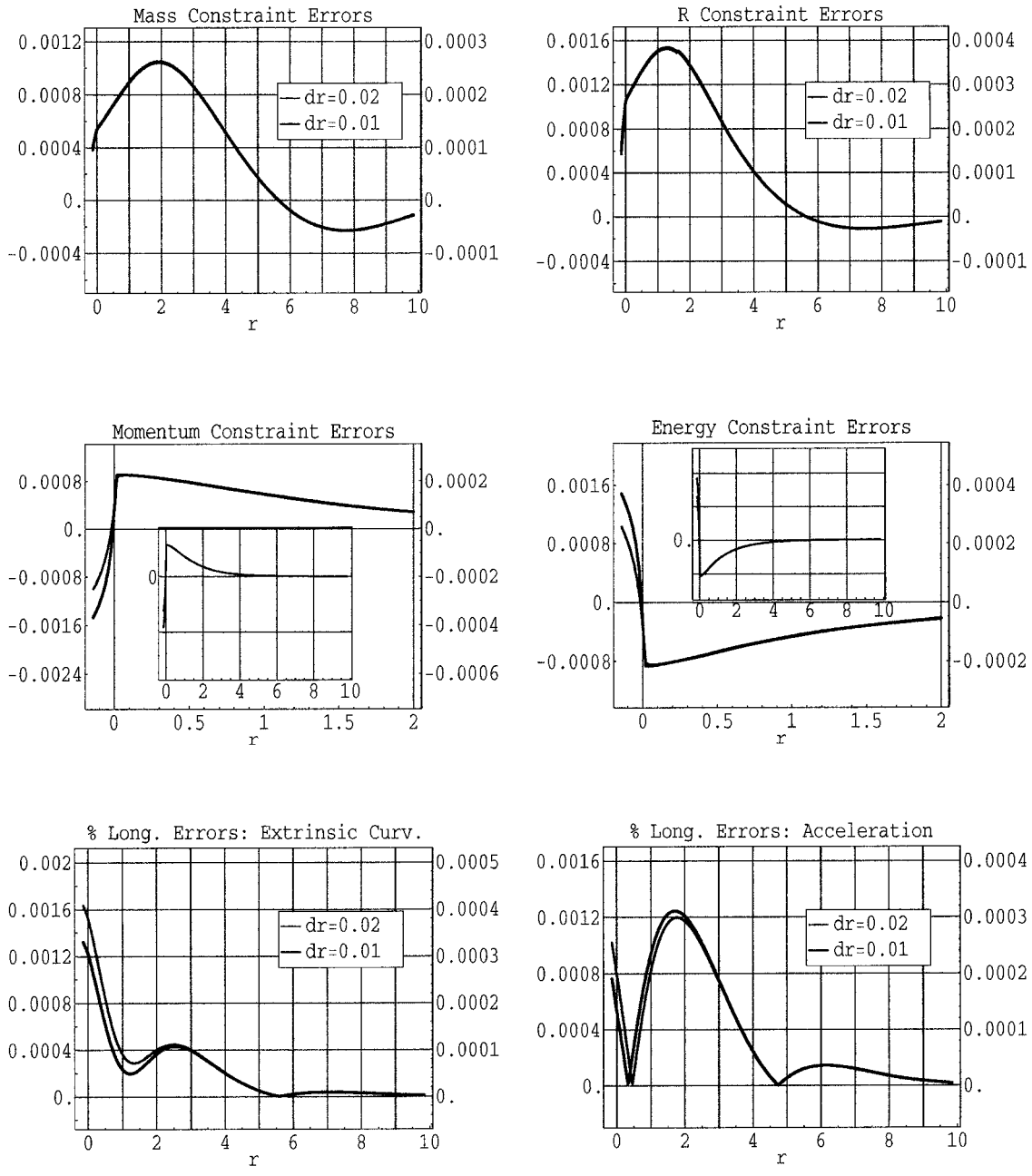


Figure 5.24: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 20M$, for stationary IVP, Lorentz gauge, and reset coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

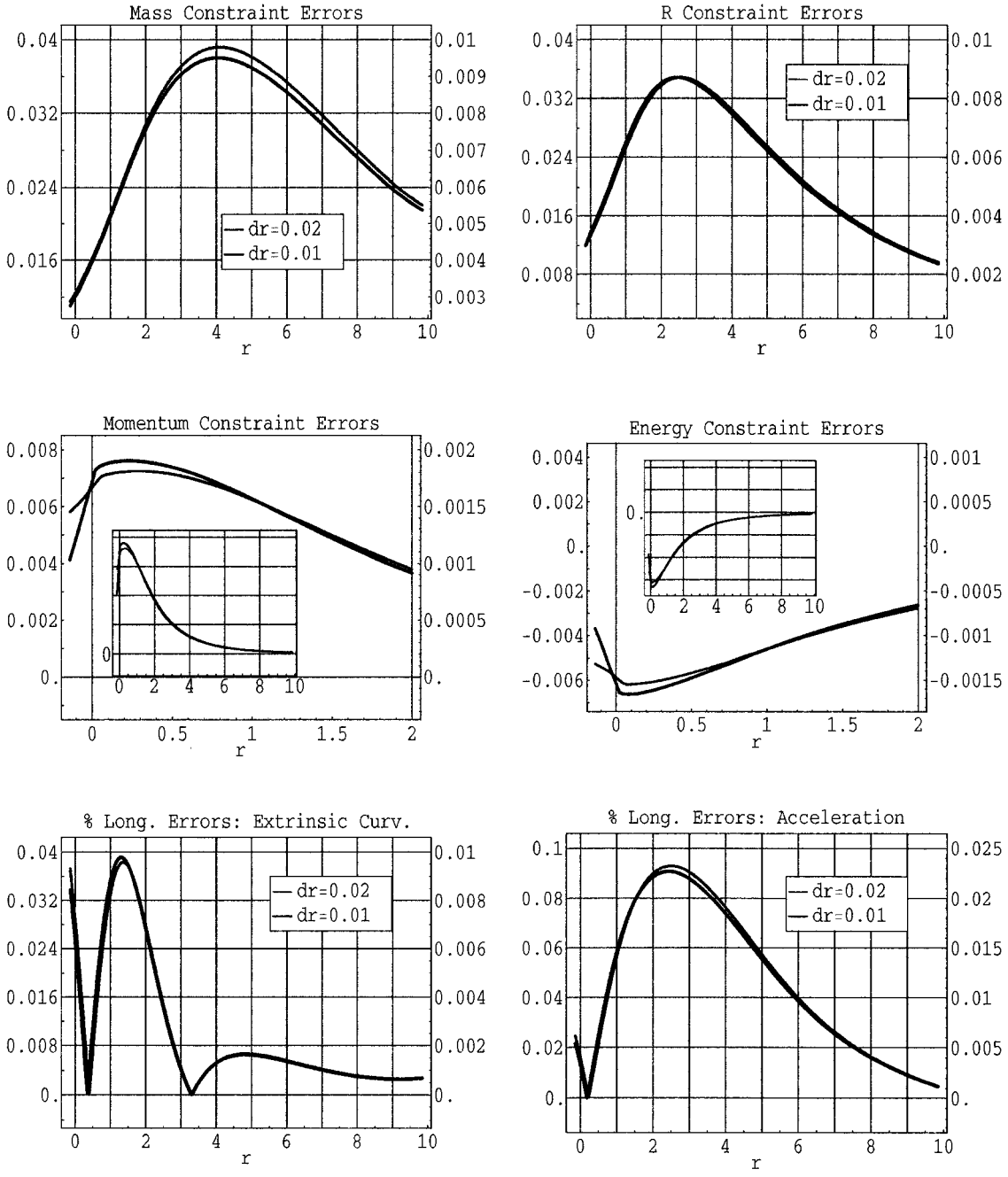


Figure 5.25: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 40M$, for stationary IVP, Lorentz gauge, and reset coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

Sum of fixed vs. reset results with stationary IVP

Exponential growth of the constraint error 1-norms in all our studies indicates the presence of unstable modes. Second order convergence outside the event horizon to about $t_S = 40M$ suggests that these unstable modes are actual analytic (though non-physical) solutions to the continuum equations. A linearization study provides further evidence for the existence of analytic unstable modes with fixed coordinate conditions. The linear stability analysis gives two analytic exponentially growing modes, one of which has the same growth rate as the mode in our numerical solution. Since this mode is not propagating, it most likely originates from the source terms of the equations.

These analytic non-physical modes are not the whole story, however. The lack of second order convergence inside the event horizon, especially for the momentum and energy constraint errors, indicates that our numerical methods are insufficient to handle the propagating physical and non-physical modes in the vicinity of the event horizon.

The unstable analytic modes and numerical problems are evident with fixed and resetting coordinate conditions, and with Lorentz and Nester gauges. Loss of second order convergence outside the event horizon, which we think is due in large part to unstable analytic modes, does not occur until after $t_S = 40M$. Hence, our results outside the event horizon are good up to this time. $40M$ is not long enough, however, to be useful for 3D binary black hole calculations.

Resetting coordinate conditions are required to limit the growth of A_i which, if unchecked, results in a coordinate singularity. Although the errors are bigger with the Lorentz gauge than with the Nester gauge for stationary initial conditions, we do not see any evidence at this point that one gauge is preferable over the other.

5.11.2 Constant Mean Curvature (CMC) initial conditions

A time dependent problem is tested by using constant mean curvature initial conditions so that $Tr\ K = K_0$ initially. Although $a_{\hat{r}}$ is chosen to satisfy the stationary solution initially, it will not stay stationary because of the CMC initial slice. We want to see if the evolution runs away, or settles down to a time independent solution.

The CMC code is the same as the stationary IVP code, except for the following. The initial condition on K_R satisfies the constant mean curvature condition, as described in Sec. 5.7.2. The lapse is normalized at each resetting so that $\alpha = (1 - R\ n_{\hat{r}})$ at the outermost ghost grid point. When the lapse is normalized in this way, then Schwarzschild time equals coordinate time at the outer edge of the grid, as explained in Sec. 5.11. Furthermore, the longitudinal variables are quadratically extropolated to the ghost point to give values for the incoming and outgoing longitudinal eigenmodes, rather than set to a stationary value. With CMC initial conditions, there are two free parameters: K_{TH} and K_0. The initial values used are: (1) $K_{TH} = -0.25$, $K_0 = 0.3$ (for the Nester gauge), and (2) $K_{TH} = -0.3$, $K_0 = 0.4$ (for the Lorentz gauge).

As in the stationary IVP figures, the grid ranges from $-0.16 \leq r \leq 9.84$. Our initial grid is uniform in r, which we take equal to the proper radius by setting $\lambda = 0$, and the event horizon is arbitrarily located at $r = 0$. The time steps are fixed such that $dt/dr = 1/4$. For convergence studies, the coarse grid spacing is $dr = 0.02$, and the fine grid spacing is $dr = 0.01$. The $\ln\alpha$, $\partial_r \ln\alpha$, β^r, and $\partial_r\beta^r$ are reset at every $dt = 0.01$ for both grid resolutions (with the shift and its derivative reset according to the minimal strain condition).

Any deviations from the standard conditions are noted in the Figure captions.

Nester

Figs. 5.26 and 5.27 show the time evolution of the primary and metric variables. With the Nester gauge, the solution does not settle down to a stationary state, but rather develops pathological behavior by time $t = 24M$. This is an important result because it tells us that the EW tetrad formulation with the Nester gauge is not viable with non-stationary, constant mean curvature initial conditions. This result cannot be improved by changing the coordinate conditions or by adding constraints to the equations; it is a statement only about the Nester gauge condition.

Fig. 5.28 shows that the constraint 1-norm errors grow exponentially. These constraint errors converge to second order both inside and outside of the event horizon at $t = 16M$, as seen in Fig. 5.29, indicating that the 1-norm errors in Fig. 5.28 are valid to at least this time.

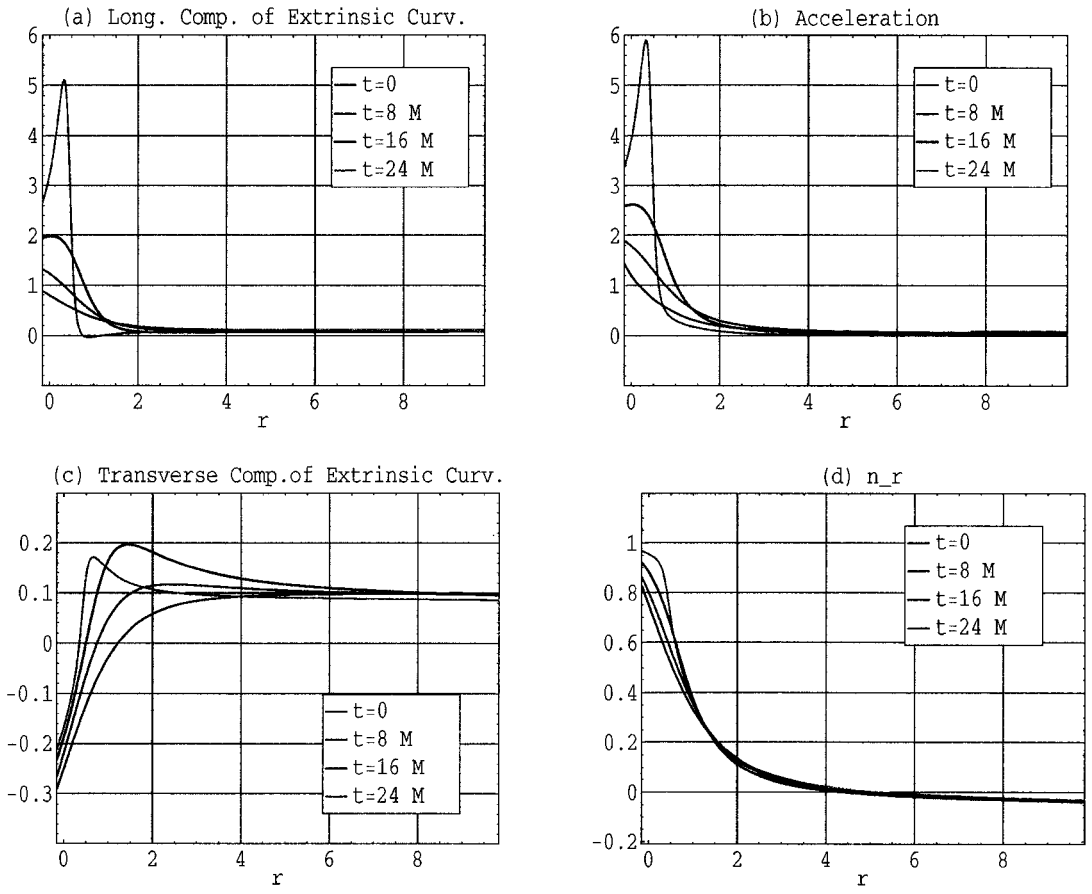


Figure 5.26: Time evolution of primary variables, with CMC initial conditions, Nester gauge, and reset coordinate conditions. Calculated with grid spacing $dr = 0.02$. $\ln \alpha$, $\partial_r \ln \alpha$, β^r, and $\partial_r \beta^r$ are reset every $dt = 0.01$.

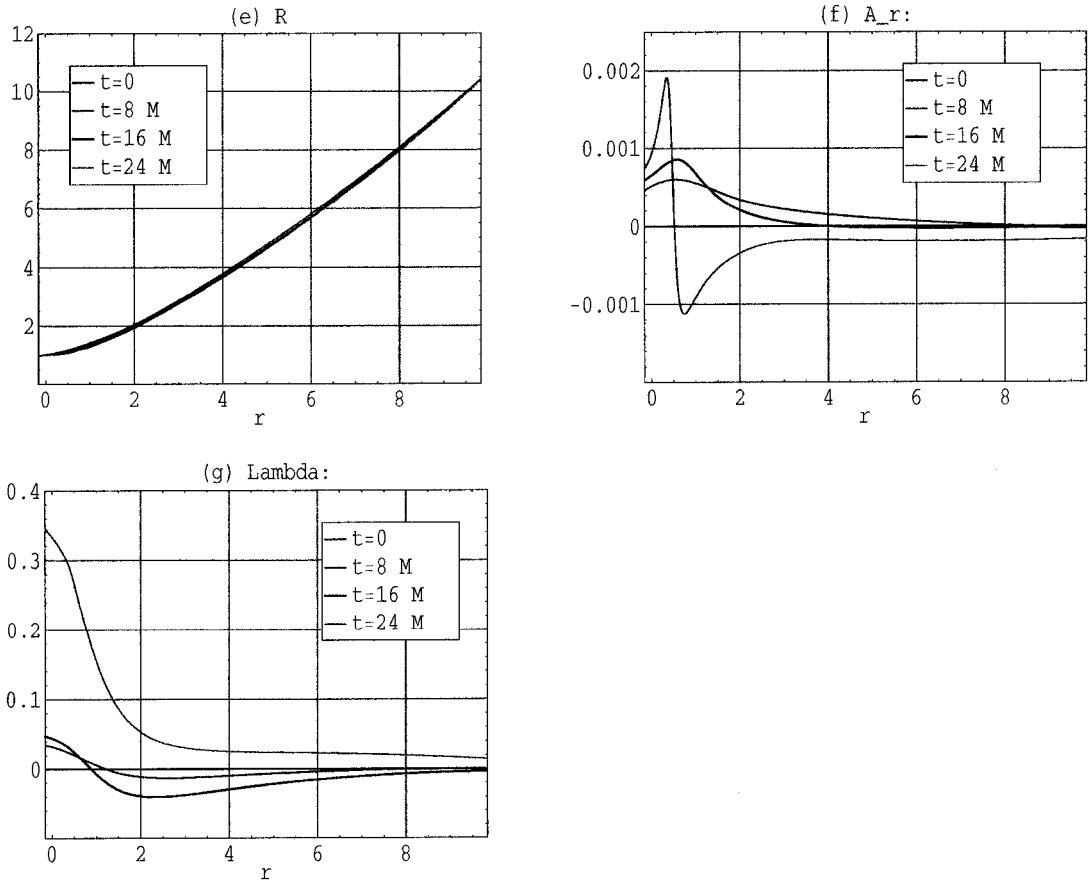


Figure 5.27: Time evolution of metric variables, with CMC initial conditions, Nester gauge, and reset coordinate conditions. Calculated with grid spacing $dr = 0.02$. $\ln \alpha$, $\partial_r \ln \alpha$, β^r, and $\partial_r \beta^r$ are reset every $dt = 0.01$.

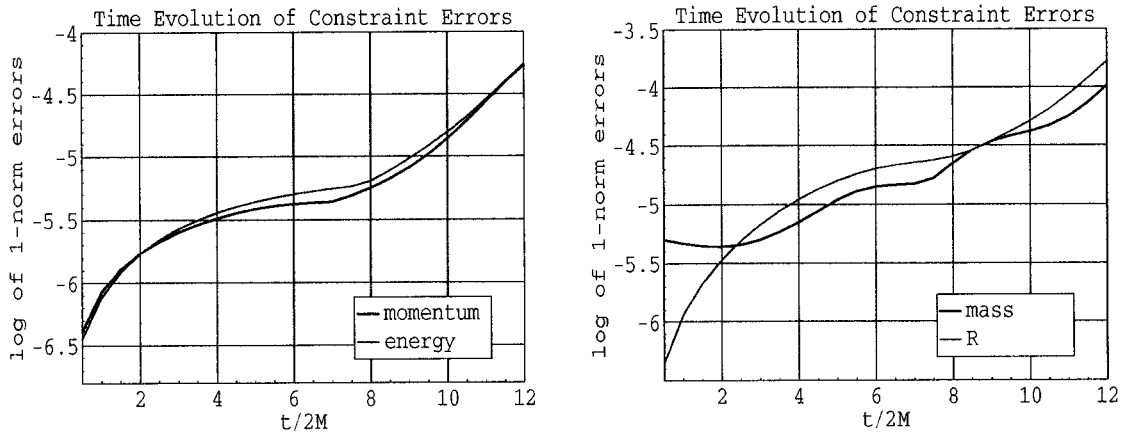


Figure 5.28: For CMC IVP, Nester gauge, and reset coordinate conditions. Evaluated with a grid resolution $dr = 0.01$. $\ln\alpha$, $\partial_r \ln\alpha$, β^r, and $\partial_r\beta^r$ are reset every $dt = 0.005$.

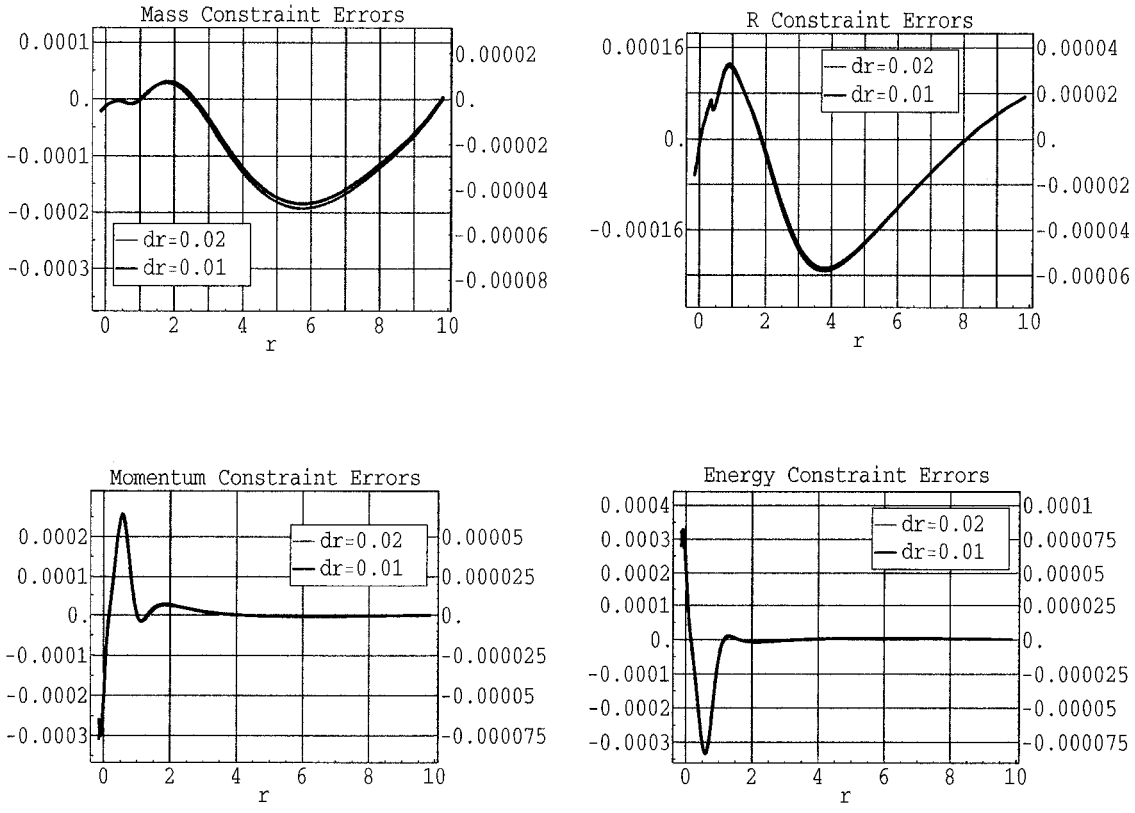


Figure 5.29: Convergence of constraint errors at $t = 16M$, for CMC IVP, Nester gauge, and reset coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

Lorentz

In contrast to the Nester gauge with CMC initial conditions, the Lorentz gauge gives an asymptotically time independent solution, as seen in Figs. 5.30 and 5.31. However, the solution with the Lorentz gauge is unstable, since the 1-norm constraint errors grow exponentially, shown in Fig. 5.32.

The convergence of all the constraint errors at $t = 16M$, plotted in Fig. 5.33, is worse than with the Nester gauge at this time, and second order convergence is lost in the mass and R constraint errors.

The differences between the CMC codes with Nester and Lorentz gauges are (1) the initial values for K_{TH} and K_0, and (2) the source terms for $a_{\hat{r}}$. (The initial conditions as well as the source terms are different for the stationary IVP because $S\_a_{\hat{r}}$ in the time independent solution for K_R, Eq. 5.61, depends on the gauge.) If the values of K_{TH} and K_0 for the Nester gauge code are used for the Lorentz gauge code, making the initial slices exactly the same, then the loss of convergence at $t = 16M$ is even worse than the results in Fig. 5.33. So, it is fair to conclude that the difference in $S\_a_{\hat{r}}$ for the two gauges has a large effect on the solution with CMC initial conditions. When $S\_a_{\hat{r}}$ is from the Nester gauge, a pathological solution results. Even though the evolution settles down to a stationary state when $S\_a_{\hat{r}}$ is from the Lorentz gauge, the errors are larger, and exhibit poor convergence.

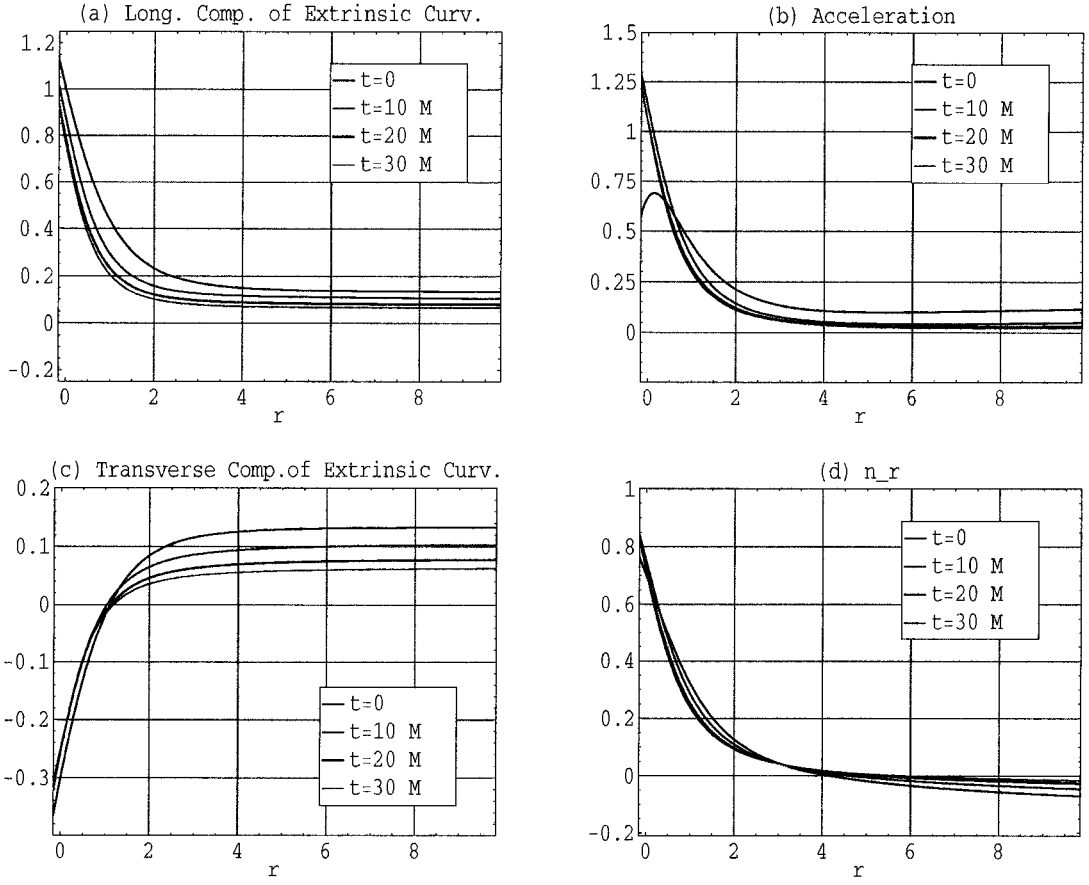


Figure 5.30: Time evolution of primary variables, with CMC initial conditions, Lorentz gauge, and reset coordinate conditions. Calculated with grid spacing $dr = 0.02$. $\ln \alpha$, $\partial_r \ln \alpha$, β^r, and $\partial_r \beta^r$ are reset every $dt = 0.01$.

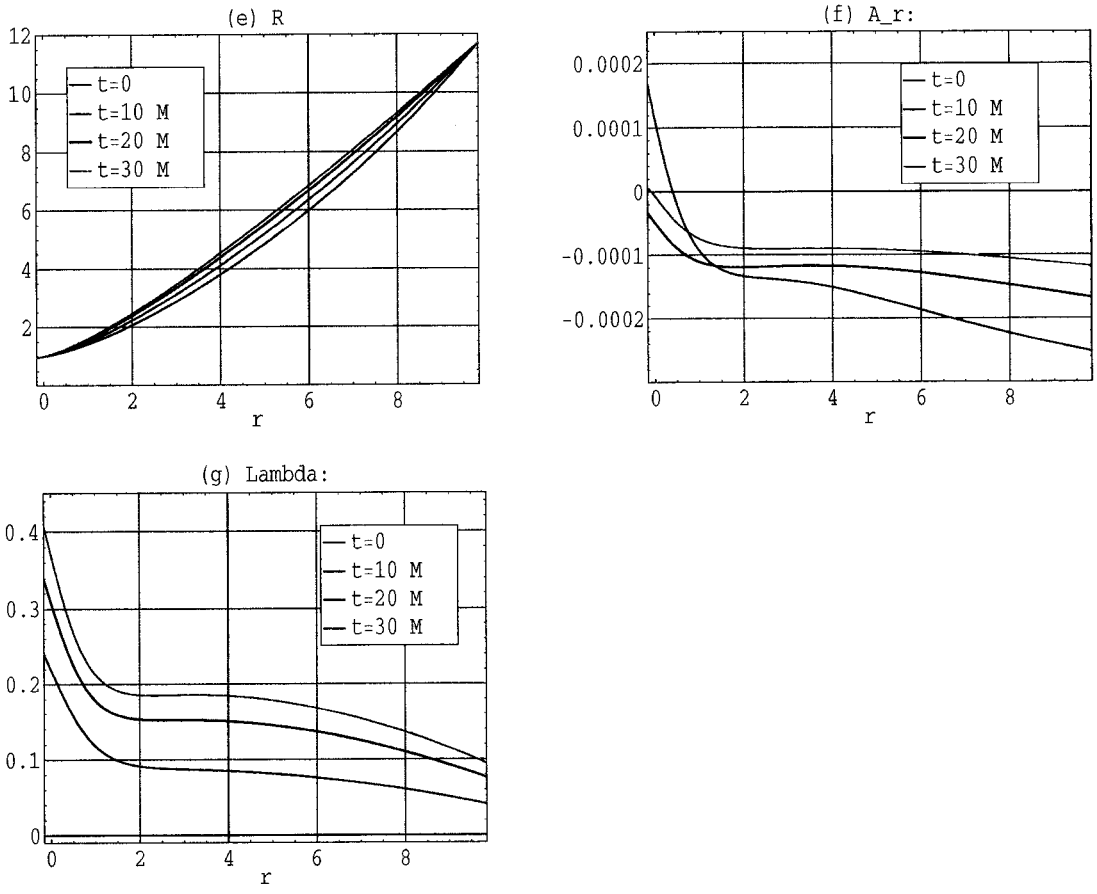


Figure 5.31: Time evolution of metric variables, with CMC initial conditions, Lorentz gauge, and reset coordinate conditions. Calculated with grid spacing $dr = 0.02$. $\ln \alpha$, $\partial_r \ln \alpha$, β^r, and $\partial_r \beta^r$ are reset every $dt = 0.01$.

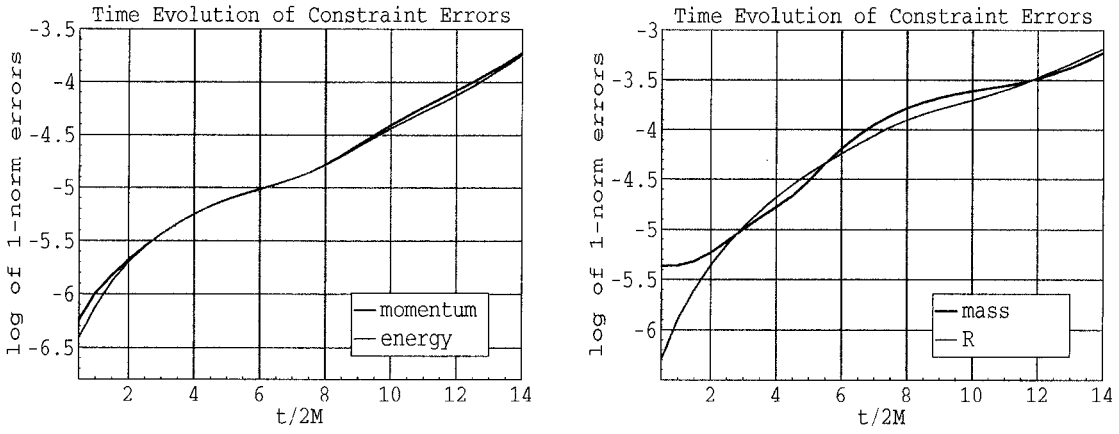


Figure 5.32: For CMC IVP, Lorentz gauge, and reset coordinate conditions. Evaluated with a grid resolution $dr = 0.01$. $\ln\alpha$, $\partial_r \ln\alpha$, β^r, and $\partial_r \beta^r$ are reset every $dt = 0.005$.

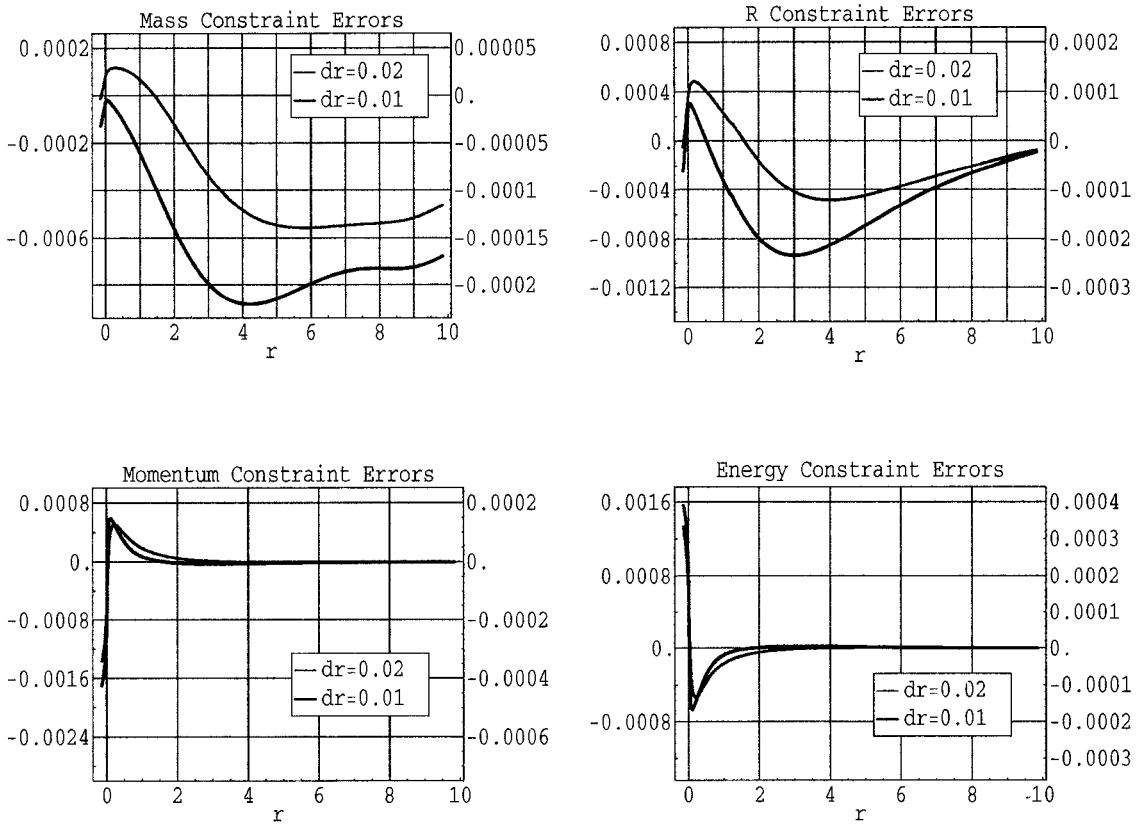


Figure 5.33: Convergence of constraint errors at $t = 16M$, for CMC IVP, Lorentz gauge, and reset coordinate conditions. Notice that the red scale is $4\times$ the blue scale.

5.11.3 Constraint Addition

The energy and momentum constraint equations are added to the spherically symmetric advection equations involving $n_{\hat{r}}$ and K_T. The motivation for this procedure is to change the speeds of the "constraint" eigenmodes so that any generated constraint violating errors which travel at the same speed as the "constraint" eigenmodes will move quickly through the region of instability near the event horizon and advect out of the grid.

Let S_mc equal the right hand side of the true momentum constraint Eq. (5.48), and let S_ec equal the right hand side of the true energy constraint Eq. (5.49). An arbitrary multiple p_1 of $\partial_r\,(n_{\hat{r}} - K_T) = S\_ec - S\_mc$ is added to advection Eq. (5.45), and an arbitrary multiple p_2 of $\partial_r\,(n_{\hat{r}} + K_T) = S\_ec + S\_mc$ to advection Eq. (5.47). This procedure has no bearing on the physics of the problem; it is simply a different way of expressing the equations. The advection equations for the constraint quantities are now:

$$\partial_t\,(n_{\hat{r}} - K_T) + s_3\;\partial_r\,(n_{\hat{r}} - K_T) = \frac{\alpha}{1 + A_{\hat{r}}}\,(S\_n_{\hat{r}} - S\_K_T) + p_1\,(S\_ec - S\_mc), \quad (5.112)$$

$$\partial_t\,(n_{\hat{r}} + K_T) + s_4\;\partial_r\,(n_{\hat{r}} + K_T) = \frac{\alpha}{1 - A_{\hat{r}}}\,(S\_n_{\hat{r}} + S\_K_T) + p_2\,(S\_ec + S\_mc), \quad (5.113)$$

where $s_3 = (s_1 + p_1)$ and $s_4 = (s_2 + p_2)$. It is evident that in the process of changing the speeds of advection of the "constraint" eigenmodes, the source terms are also modified. The way in which the constraints are added is "ad hoc" because it is done solely in the context of the spherically symmetric problem.

Stationary IVP–Lorentz

For the stationary IVP and Lorentz gauge, parameter values of $p_1 = -1.5$ and $p_2 = -1.8$ shift speeds s_3 and s_4 in the negative direction from their original values, s_1 and s_2, as shown in Fig. 5.34.

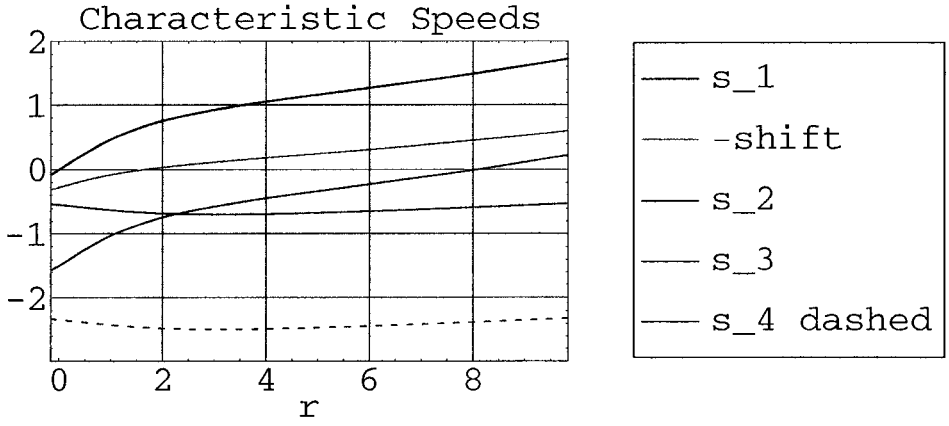


Figure 5.34: Characteristic speeds at $t = 0$, stationary IVP, Lorentz gauge, and ad hoc constraint addition. Speeds $s_3 = s_1 - 1.5$, $s_4 = s_2 - 1.8$.

Notice that s_3 (the solid black line) crosses through zero at $r \approx 8$, whereas without constraint addition, it passes through zero at $r = 0$ ($s_3 = s_1$ without constraint addition). Because both s_3 and s_4 are strongly negative at and inside the event horizon, any errors travelling with these speeds will move quickly out of the grid.

The resulting stability, accuracy, and convergence of the code is tested. We find rock-solid stability to at least $500M$ for the momentum, energy, and R constraint errors, as is seen in Fig. 5.35. The 1-norm mass errors grow slowly. Excellent second order convergence of the constraint, longitudinal, and metric errors at $100M$ and $200M$ is shown in Figs. 5.36, 5.37, 5.38, and 5.39. The second order convergence holds even inside the event horizon at these late times. The energy and momentum constraint errors, and the metric errors in A_r show little to no growth from $100M$ to $200M$. The mass and longitudinal errors, and the metric errors in λ, grow slowly in place.

To test the idea that the dramatic increase in accuracy and stability is at least in part due to moving the "constraint" eigenmodes quickly through the

region of instability around the event horizon, we changed the parameter p_1 so that s_3 passes through zero at $r \approx 1$. The speeds for this test are shown in Fig. 5.40. As expected, the 1-norm errors grow exponentially, as seen in Fig. 5.41.

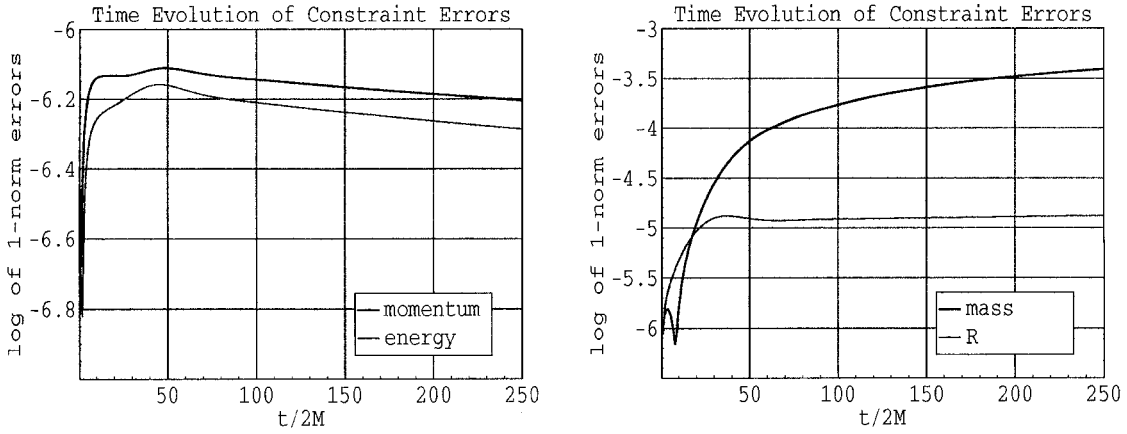


Figure 5.35: For stationary IVP, Lorentz gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -1.5$ and $p_2 = -1.8$. Evaluated with a grid resolution $dr = 0.01$. $\ln \alpha$, $\partial_r \ln \alpha$, β^r, and $\partial_r \beta^r$ are reset every $dt = 0.005$.

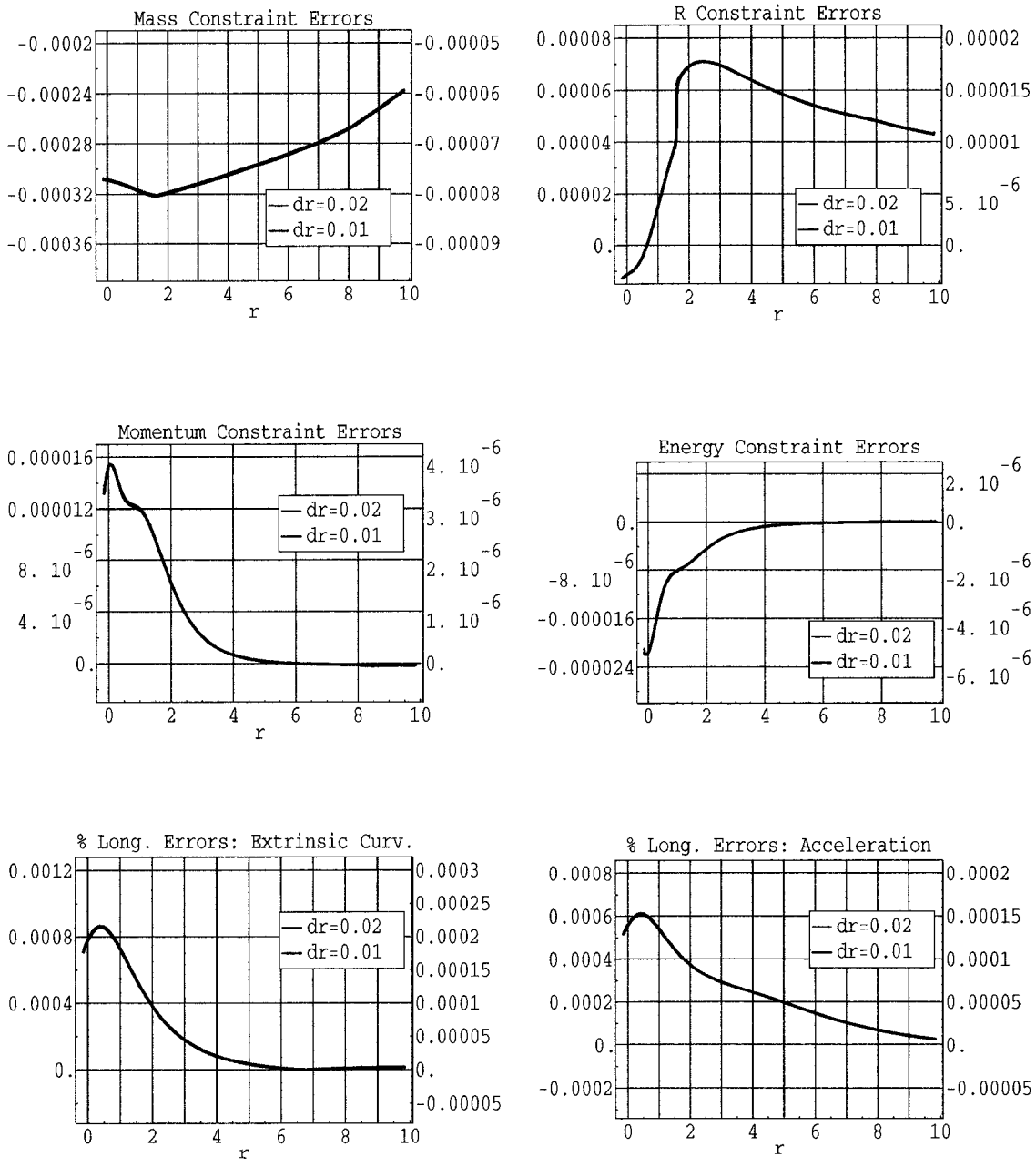


Figure 5.36: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 100M$, for stationary IVP, Lorentz gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -1.5$ and $p_2 = -1.8$. Notice that the red scale is $4\times$ the blue scale.

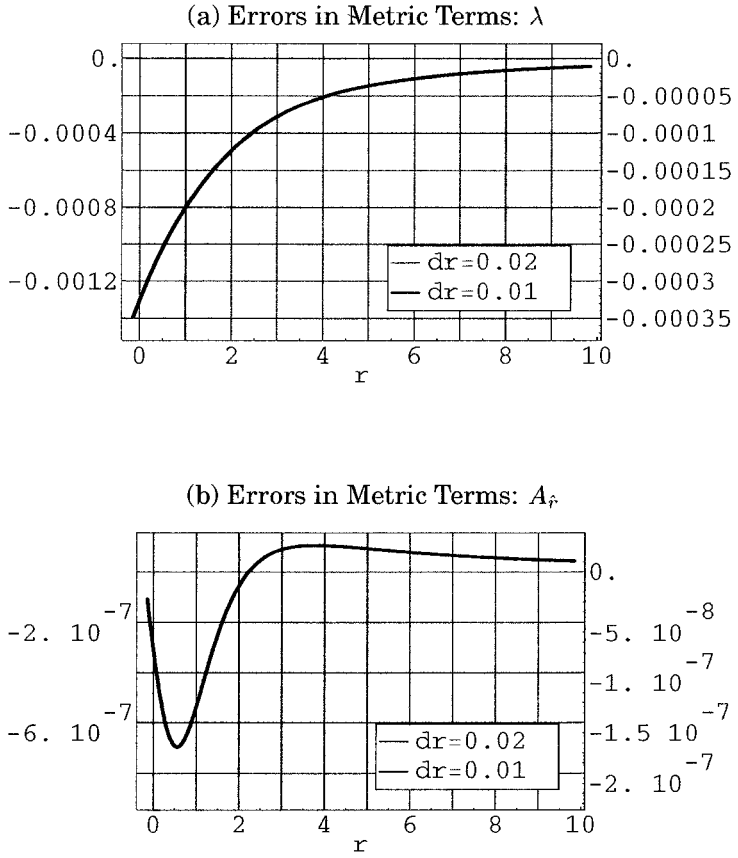


Figure 5.37: Convergence of errors in (a) λ and (b) $A_{\hat{r}}$ at $t_S = 100M$, for stationary IVP, Lorentz gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -1.5$ and $p_2 = -1.8$. Notice that the red scale is $4\times$ the blue scale.

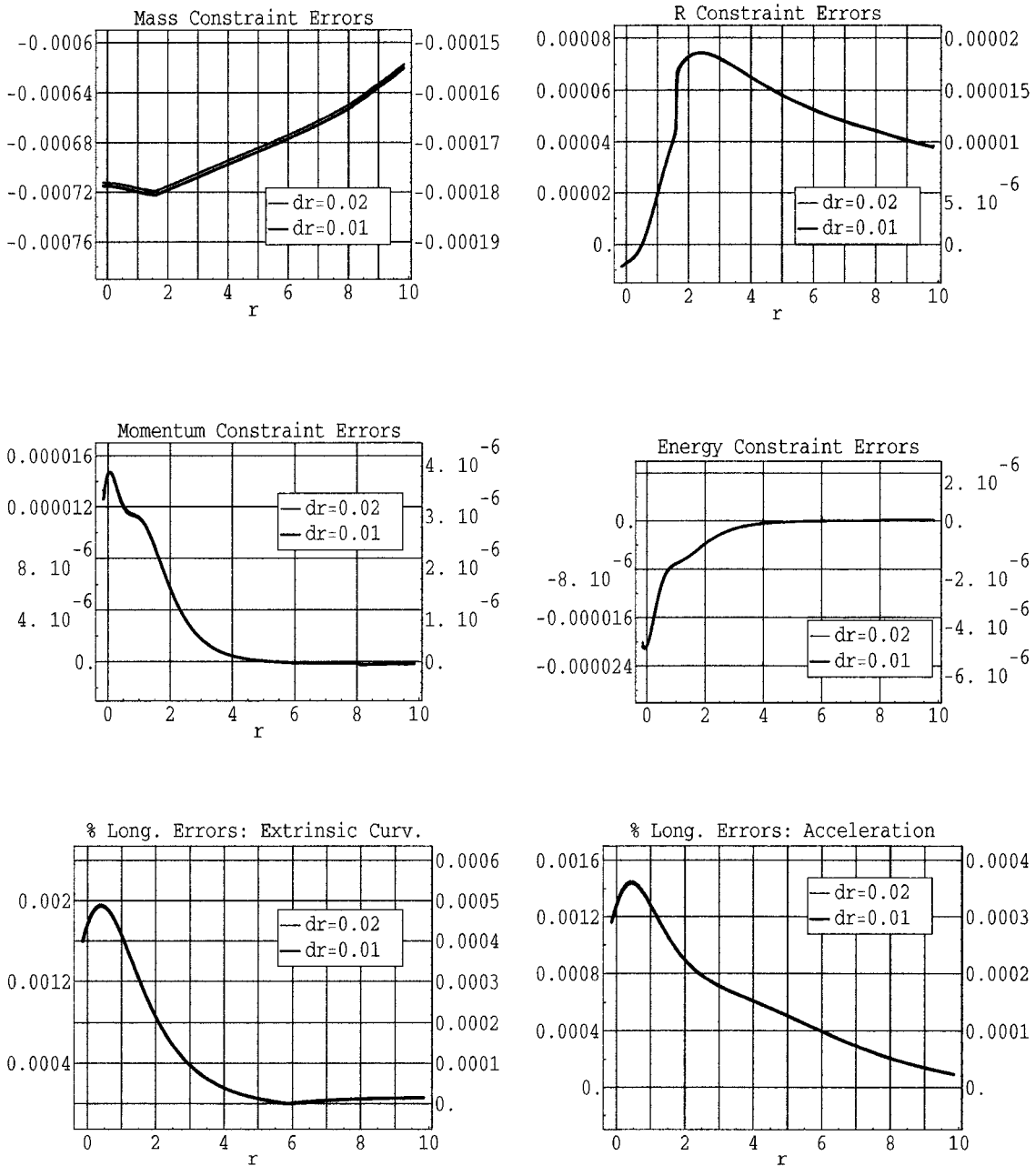


Figure 5.38: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 200M$, for stationary IVP, Lorentz gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -1.5$ and $p_2 = -1.8$. Notice that the red scale is $4\times$ the blue scale.

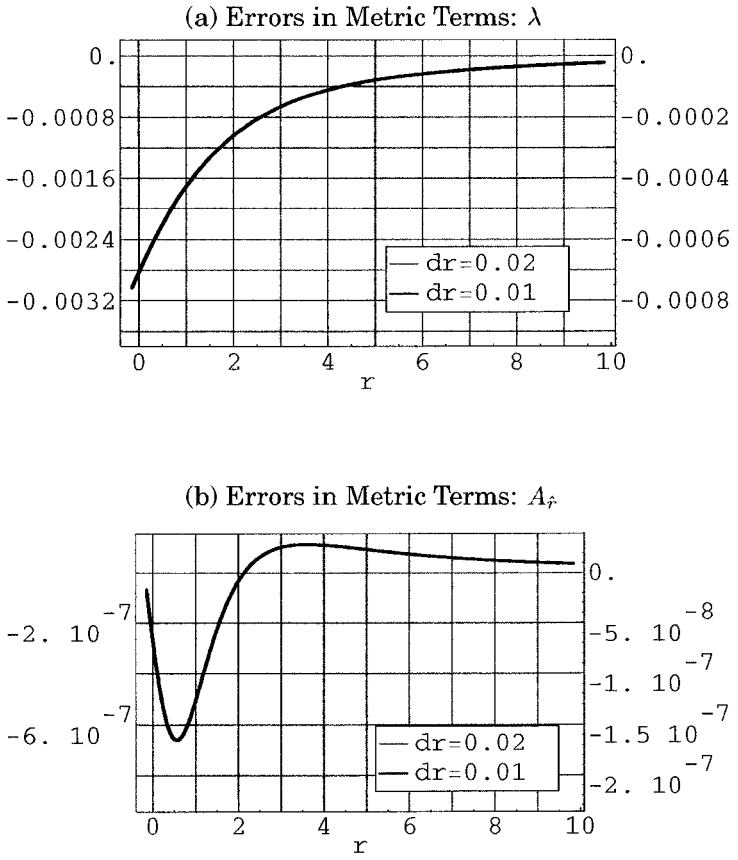


Figure 5.39: Convergence of errors in (a) λ and (b) $A_{\hat{r}}$ at $t_S = 200M$, for stationary IVP, Lorentz gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -1.5$ and $p_2 = -1.8$. Notice that the red scale is $4\times$ the blue scale.

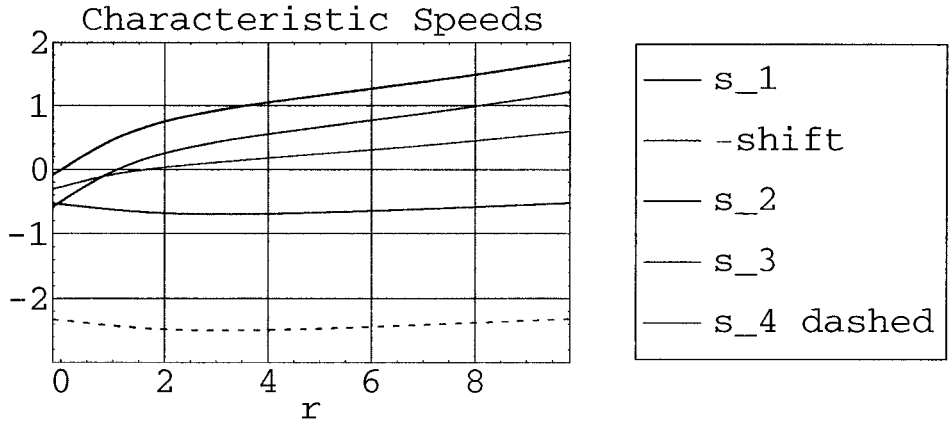


Figure 5.40: Characteristic speeds at $t = 0$, stationary IVP, Lorentz gauge, and ad hoc constraint addition. Speeds $s_3 = s_1 - 0.5$, $s_4 = s_2 - 1.8$.

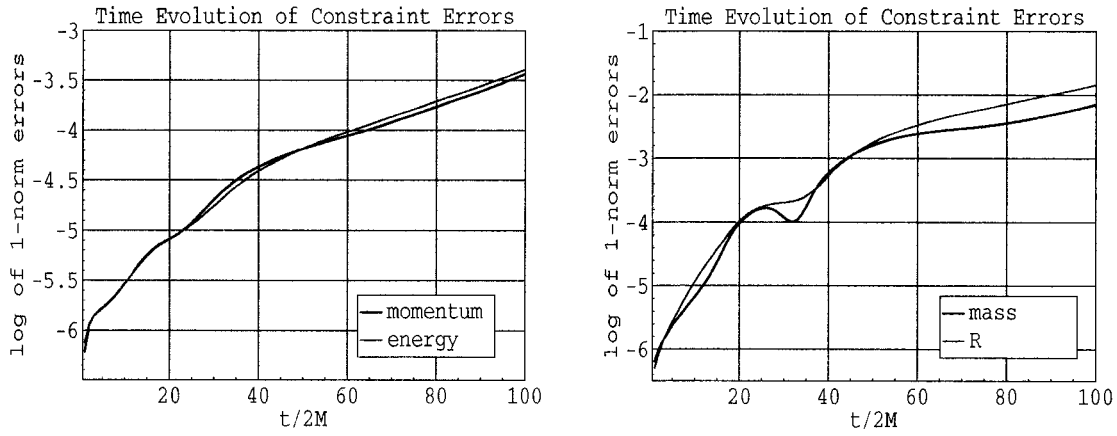


Figure 5.41: For stationary IVP, Lorentz gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -.5$ and $p_2 = -1.8$. Evaluated with a grid resolution $dr = 0.01$. $\ln \alpha$, $\partial_r \ln \alpha$, β^r, and $\partial_r \beta^r$ are reset every $dt = 0.005$.

Stationary IVP–Nester

With the Nester gauge and stationary initial conditions, parameter values of $p_1 = -1.3$ and $p_2 = -1.8$ are used, giving the speeds shown in Fig. 5.42. As with the Lorentz gauge, s_3 passes through zero at $r \approx 8$.

After an initial rise, the Nester gauge gives rock-solid stable evolution for all the constraint 1-norm errors to at least $500M$, shown in Fig. 5.43. Excellent second order convergence of the constraint, longitudinal, and metric errors at $200M$ is shown in Figs. 5.44 and 5.45.

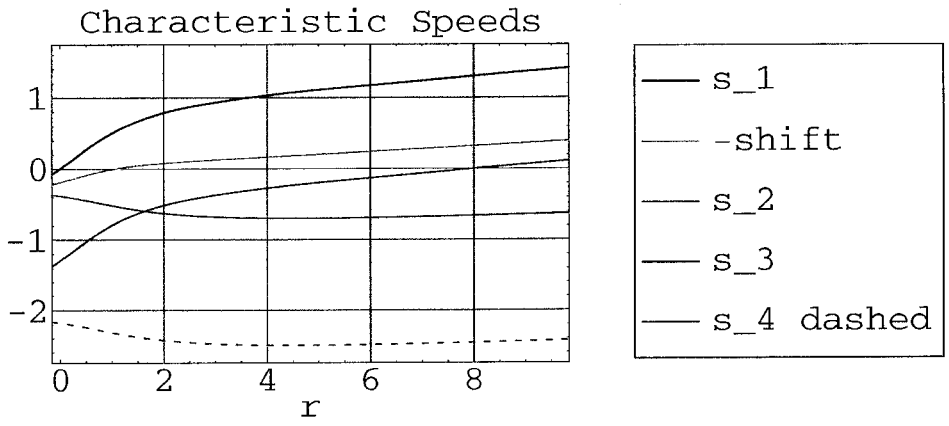


Figure 5.42: Characteristic speeds at $t = 0$, stationary IVP, Nester gauge, and ad hoc constraint addition. Speeds $s_3 = s_1 - 1.3$, $s_4 = s_2 - 1.8$.

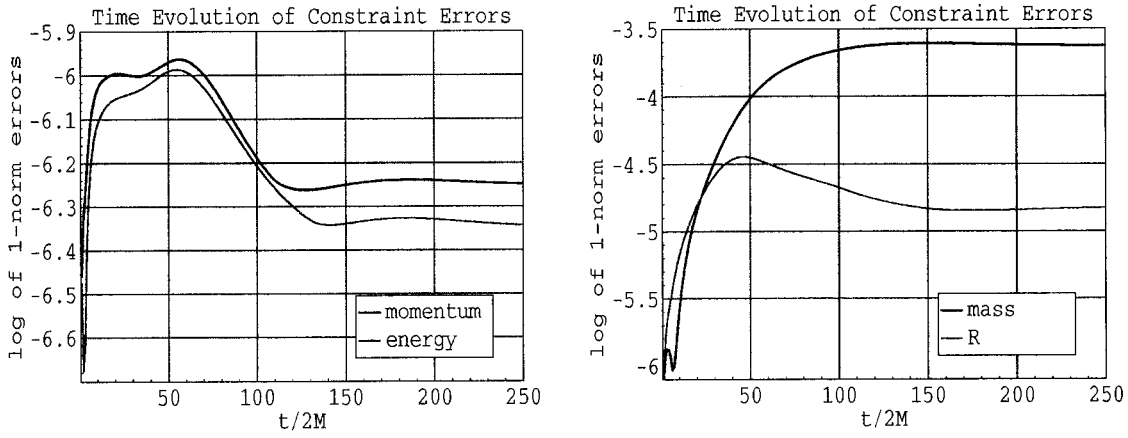


Figure 5.43: For stationary IVP, Nester gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -1.3$ and $p_2 = -1.8$. Evaluated with a grid resolution $dr = 0.01$. $\ln\alpha$, $\partial_r \ln\alpha$, β^r, and $\partial_r\beta^r$ are reset every $dt = 0.005$.

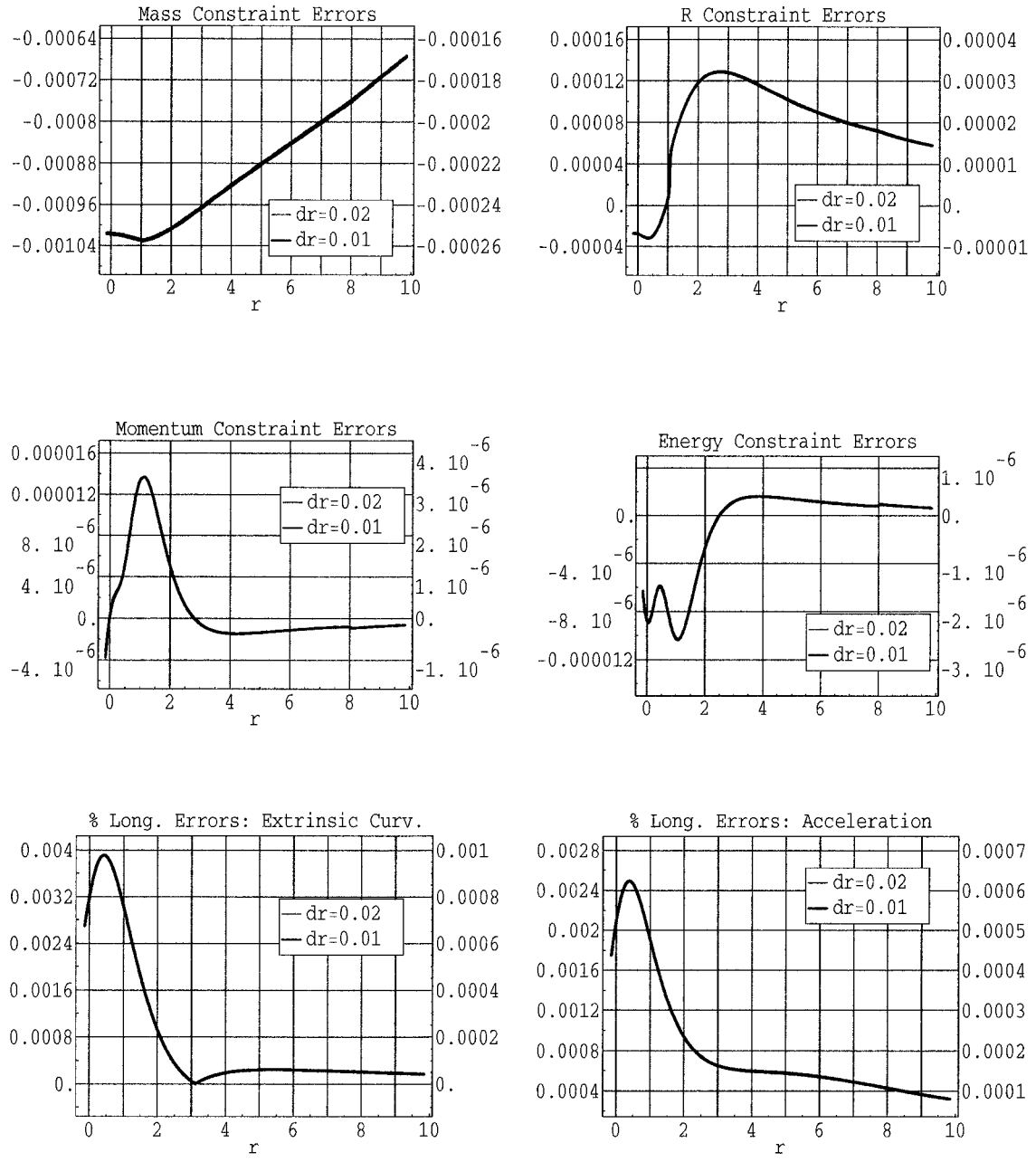


Figure 5.44: Convergence of constraint and longitudinal errors (absolute values) at $t_S = 200M$, for stationary IVP, Nester gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -1.3$ and $p_2 = -1.8$. Notice that the red scale is $4\times$ the blue scale.

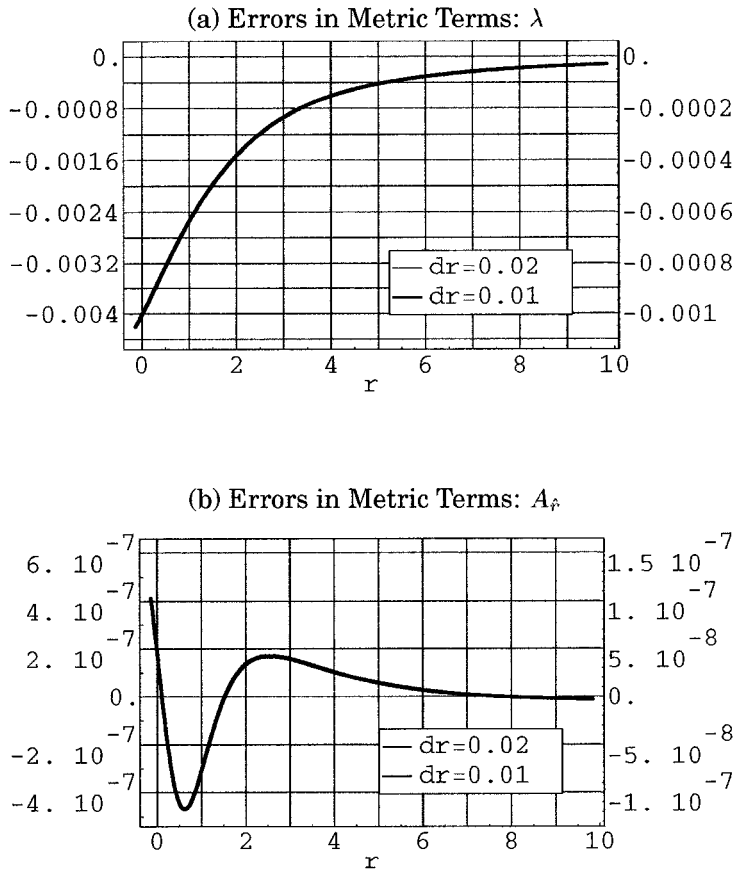


Figure 5.45: Convergence of errors in (a) λ and (b) $A_{\hat{r}}$ at $t_S = 200M$, for stationary IVP, Nester gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -1.3$ and $p_2 = -1.8$. Notice that the red scale is $4\times$ the blue scale.

CMC IVP–Lorentz

With CMC initial conditions, Lorentz gauge, and $p_1 = p_2 = -2.5$, the speeds are as shown in Fig. 5.46. The momentum, energy, and R constraint 1-norm errors are stable to at least $500M$, as seen in Fig. 5.47. As with stationary initial conditions, the mass constraint 1-norm errors grow slowly. The convergence of

the errors at $200M$ is shown in Fig. 5.48. The momentum and energy constraint errors converge to at least second order outside the event horizon. The mass constraint errors converge to ≈ 1.6 order across the grid. Second order (or greater) convergence breaks down for the R constraint errors when $r > 4$.

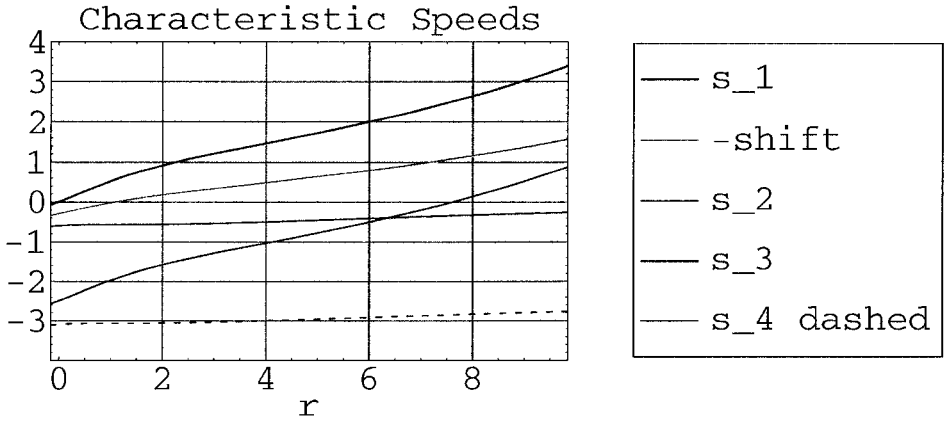


Figure 5.46: Characteristic speeds at $t = 0$, CMC IVP, Lorentz gauge, and ad hoc constraint addition. Speeds $s_3 = s_1 - 2.5$, $s_4 = s_2 - 2.5$.

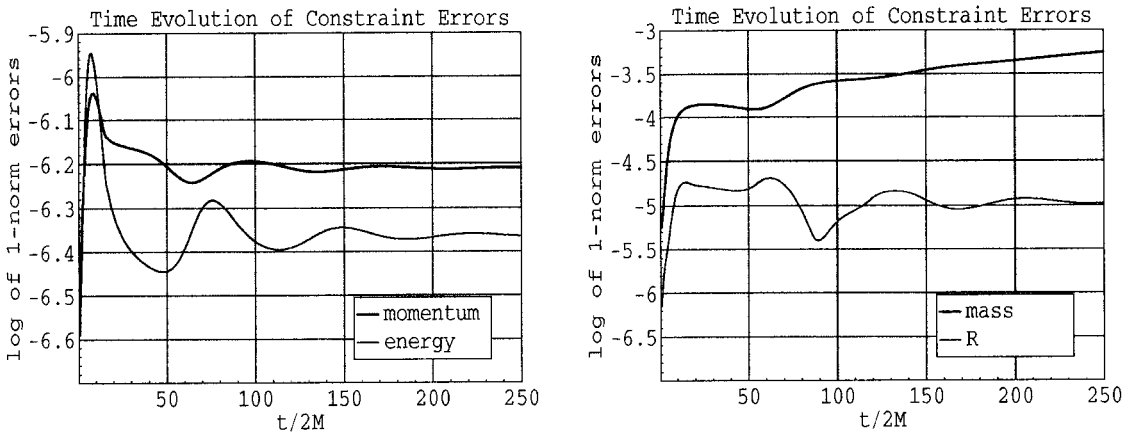


Figure 5.47: For CMC IVP, Lorentz gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -2.5$ and $p_2 = -2.5$. Evaluated with a grid resolution $dr = 0.01$. $\ln\alpha$, $\partial_r \ln\alpha$, β^r, and $\partial_r \beta^r$ are reset every $dt = 0.005$.

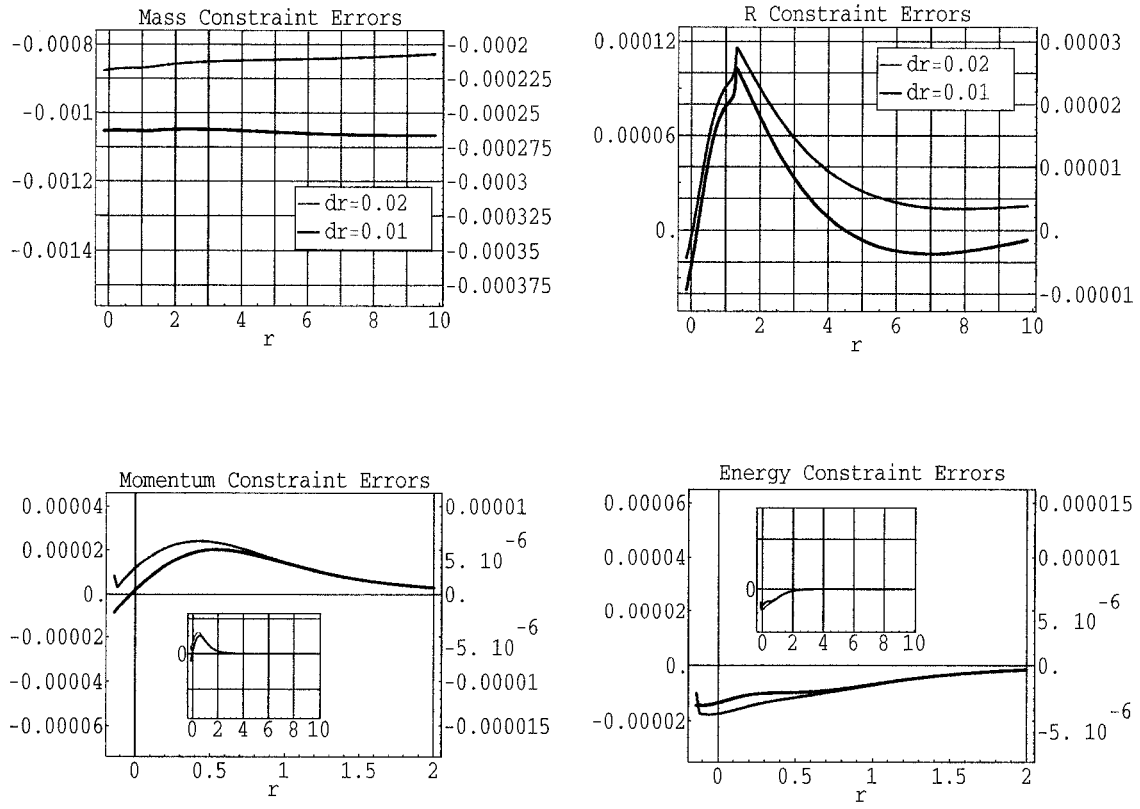


Figure 5.48: Convergence of constraint errors at $t = 200M$, for CMC IVP, Lorentz gauge, reset coordinate conditions, and ad hoc constraint addition. The parameters $p_1 = -2.5$ and $p_2 = -2.5$. Notice that the red scale is $4\times$ the blue scale.

5.11.4 Constraint Addition–Future Work

Further tests are needed in order to understand the exact causes of the dramatic improvement in stability and accuracy obtained with our ad hoc constraint addition, because this constraint addition changes the source terms of the nonlinear equations as well as the characteristic speeds of the eigensystem. The preliminary results presented in this chapter suggest that changing the speeds of the "constraint" eigenmodes causes errors associated with these speeds to propagate quickly through the region of instability around the event horizon, preventing them from accumulating. To verify this, we need to determine the speeds at which the constraint errors propagate. This can be done from evolution equations for the constraints. Further, the effect of the source terms on the growth of the instability must be ascertained. After the reasons for the large improvement obtained with the ad hoc constraint addition have been understood, we will have to see whether or not these improvements can be successfully implemented in the 3D tetrad equations.

5.12 Discussion

These experiments in spherically symmetric Schwarzschild spacetime indicate that resetting the lapse and shift give the most promising approach for stable evolutions of the EW tetrad equations. With stationary initial conditions, both the Nester and Lorentz tetrad gauge conditions give reliable results to about $t_S = 40M$, after which the second order convergence of the constraint errors breaks down. With both gauges, the growth of the 1-norm constraint errors is exponential, and therefore unstable. With constant mean curvature initial conditions, the numerical evolution using the Nester gauge is pathological, while with the Lorentz gauge, the evolution approaches a stationary solution. The constraint errors grow exponentially in both cases.

The amplitudes of the spikes in the energy and momentum constraint errors inside the event horizon are dependent on the numerical methods used in solving the transport step of the Strang splitting. Since this feature spans such a tiny fraction of the grid, it contributes negligibly to the 1-norm errors. The shape of the energy and momentum constraint errors versus r outside the event horizon appears to be independent of numerical method, and converges to second order until around $t_S = 40M$. These results suggest that the feature outside the horizon is a non-physical but true solution to the equations. The linearization study shows that there is a true unstable mode with fixed coordinate conditions, which is due to a growing perturbation of the analytic solution. This mode exists even though the metric perturbations are set to zero.

The EW formalism, as presented in this dissertation, is not as stable as is necessary for $3D$ black hole calculations, even with stationary initial conditions. It appears that the most promising way to achieve long-term stability is through adding multiples of the constraint equations to the evolution equations, as has already proved successful in $3D$ applications using $3+1$ formualations [39, 80]. An ad hoc test of this strategy, which changes both the speeds and the source terms of the advection equations involving the "constraint" eigenmodes, has been tried. The results are rock solid stability to at least $t_S = 500M$ for the Nester and Lorentz gauges with stationary initial conditions, and for the Lorentz gauge with constant mean curvature initial conditions. It remains for future work to investigate the exact cause of this dramatically improved accuracy and stability, and whether a similar approach can be successfully applied to the 3D EW equations.

Chapter 6

CONCLUSION

This dissertation has presented a tetrad approach for numerical relativity. We have learned that (1) the Nester gauge is probably not an option for general initial conditions, (2) resetting the lapse to keep A_a small, and resetting the shift according to the minimal strain condition work well, and (3) the EW tetrad equation formulation has some intrinsic advantages over the $3+1$ formulations, but needs to be modified in order to yield long-term stable numerical simulations.

The evolution of the coordinates is tied to that of the congruence in the EW tetrad formulation. We hoped that the Nester and/or Lorentz gauge conditions would give a well-behaved evolution of the congruence, and therefore guide us in choosing coordinate conditions. We discovered that this would not be possible with the Nester gauge, since with constant mean curvature initial conditions, the acceleration of the congruence develops pathological behavior. A wider range of non-stationary initial conditions needs to be tested, however, before definitively eliminating the Nester gauge from consideration for numerical relativity. With the Lorentz gauge, the evolution of the congruence with constant mean curvature initial conditions is well-behaved, and asymptotically approaches a stationary solution. Thus, the Lorentz gauge is a candidate for numerical simulations with the EW tetrad formulation. Another possibility, which is not tested in this thesis, is a hypersurface orthogonal gauge.

Two key elements required of coordinate conditions in the EW tetrad formulation are: (1) keeping A_a small, and (2) keeping the excision boundary inside

the event horizon of the black hole. The dynamic lapse we tried in Sec. 5.8.2 did not keep A_a small; therefore we abandoned that approach. Resetting the lapse and shift, as described in Sec. 5.8.3, succeeds in both respects, and is generalizable to $3D$. However, the equations used to reset the lapse and the shift are elliptic equations, which are computationally expensive to solve with finite differencing in $3D$. Others [2, 3, 51] have dealt with a similar problem in $3+1$ formulations by deriving hyperbolic coordinate conditions, whose solutions relax to the solutions of the elliptic conditions. This will be an approach to try when implementing our resetting conditions in $3D$.

The EW tetrad equations themselves are attractive for numerical relativity, for the following reasons. Expressed as partial differential equations (PDE's) for numerical integration, the equations are symmetrizable hyperbolic, a condition for stability according to Lindblom and Scheel [50]. The hyperbolic structure of the PDE's is such that boundary conditions are not required at the excision boundary of the black hole. Further, the EW tetrad equations have some significant advantages over comparable first order symmetric hyperbolic $3+1$ systems. First, there are no complicated inverse metric terms required for raising or lowering tetrad indices, since the metric in the orthonormal tetrad frame is the Minkowski metric. Additionally, the EW variables are naturally related to the "mixed" coordinate components which we find give greater accuracy in Chapter 2. Finally, the EW variables can be used with a non-diagonal metric without the complications that arise when using the "mixed" variables of the $3+1$ formulations, as shown in Chapter 4.

Even with these advantages, it is evident from this dissertation that the EW equations must be modified in order to gain the long-term stability that is required for $3D$ numerical calculations. The most promising approach is to add multiples of the constraint equations to the evolution equations in a way which preserves the symmetrizable hyperbolicity of the system, as implemented in

Chapter 2 and in [39] for $3+1$ formulations. Before implementing this approach in the context of the EW tetrad formulation, however, it is important to have some understanding regarding how to modify the equations. It can be inferred from the work presented in this dissertation that an important strategy is to change the speed of propagation of the constraint errors, so that they (1) are not in step with a propagating source of error in the nonlinear solution (for nonlinear plane wave spacetimes), or (2) travel quickly away from a region of instability (as around the event horizon of the black hole). The second point is suggested by experiments presented in Sec. 5.11.3 where adding the constraint equations in a way specific to the spherically symmetric EW tetrad equations results in long-term stability to at least $500M$. Bardeen [12] also found results which support this point, with a hypersurface orthogonal gauge in the Einstein-Bianchi system. To fully establish point (2) for the EW tetrad system, more work needs to be done to determine the regions of instability due to source terms, and the speeds of propagation of the constraint errors. Then, it remains to develop the appropriate modifications to the $3D$ EW tetrad equations.

Our approach to boundary conditions is the same for our $3+1$ and tetrad simulations. We find it important to correct the boundary values according to the energy and momentum constraint equations, in order to prevent spurious, constraint-violating signals from entering the numerical domain.

Numerical methods used in the transport step of the Strang splitting are quite different for our $3+1$ and tetrad calculations. In the EW tetrad formulation, the equations are variable coefficient advection equations, in non-conservative form. In the $3+1$ formulation, on the other hand, the principal part of the equations is a conservation law, with the characteristic matrix spatially varying. A "flux-based wave decomposition" numerical method (see also [49, 11]) is developed for the $3+1$ computations in order to obtain second order convergence.

An area of investigation for future work is how to numerically integrate the multi-dimensional EW tetrad equations. A common practice in $3+1$ simulations is to perform what is called a "locally one-dimensional" split. This means that within one time evolution step, numerical integration of the homogeneous transport equations is executed along one coordinate spatial direction at a time. The EW PDE's, as given in Eq. C.2, can be split either along the coordinate directions, as is done in $3+1$, or along the tetrad directions. For an arbitrary coordinate direction, the eigenvectors (given in Appendix C) depend on B_a^k and A_a. If one splits along tetrad directions, however, the eigenvectors depend only on $A_a\ A_a$. These eigenvectors approximate constant coefficient eigenvectors if a hypersurface orthogonal gauge is used, or if $A_a\ A_a$ is kept small. The disadvantage of splitting along tetrad directions is that *nine* locally one-dimensional splits per time step are required, and this is likely to cause numerical problems. Significant effort is needed to understand the best way to integrate the EW tetrad equations in $3D$.

In conclusion, a novel, tetrad approach to numerical relativity has been developed for $3D$ calculations, and applied to $1D$ plane wave and black hole spacetimes. Some of the analytic and numerical results are promising, and encourage us to continue in our efforts to implement this tetrad approach in full $3D$ simulations. It is clear, however, that the equations need to be modified in order to gain long-term stability for black hole spacetimes. Insight has been developed into strategies for modifying the equations which dramatically improve accuracy and stability in $1D$. We look forward to developing and testing these strategies for application to $3D$.

BIBLIOGRAPHY

[1] M. Alcubierre and B. Brügmann. *Phys. Rev. D*, 63:104006, 2001.

[2] M. Alcubierre, B. Brügmann, P. Diener, M. Koppitz, D. Pollney, E. Seidel, and R. Takahashi. *Phys. Rev. D*, 67:084023, 2003.

[3] M. Alcubierre, B. Brügmann, D. Pollney, E. Seidel, and R. Takahashi. *Phys. Rev. D*, 64:061501, 2001.

[4] M. Alcubierre, A. Corichi, J. A. Gonzalez, D. Nunez, and M. Salgado. *Phys. Rev. D*, 67:104021, 2003.

[5] A. Anderson and J. W. York. *Phys. Rev. Lett.*, 82:4384, 1999.

[6] P. Anninos, K. Camarda, J. Massó, E. Seidel, W.-M. Suen, and J. Towns. *Phys. Rev. D*, 52:2059, 1995.

[7] P. Anninos, G. Daues, J. Massó, E. Seidel, and W.-M. Suen. *Phys. Rev. D*, 51:5562, 1995.

[8] R. Arnowitt, S. Deser, and C. W. Misner. *Gravitation: An Introduction to Current Research*. Wiley, New York, 1962. pp. 227–265.

[9] A. Ashtekar. *Phys. Rev. Lett.*, 57:2244, 1986.

[10] A. Ashtekar. *Phys. Rev. D*, 36:1587, 1987.

[11] D. S. Bale, R. J. LeVeque, S. Mitran, and J. A. Rossmanith. *SIAM J. Sci. Comp.*, 24:955, 2003.

[12] J. M. Bardeen. private communication.

[13] J. M. Bardeen and L. T. Buchman. *Phys. Rev. D*, 65:064037, 2002.

[14] T. W. Baumgarte and S. L. Shapiro. *Phys. Rev. D*, 59:024007, 1999.

[15] T. W. Baumgarte and S. L. Shapiro. *Phys. Rept.*, 376:41, 2003.

[16] C. Bona, J. Massó, E. Seidel, and J. Stela. *Phys. Rev. D*, 56:3405, 1997.

[17] C. Bona, J. Massó, E. Seidel, and P. Walker, 1998. gr-qc/9804052.

[18] L. T. Buchman and J. M. Bardeen. *Phys. Rev. D*, 67:084017, 2003.

[19] M. W. Choptuik. *Phys. Rev. Lett.*, 70:9, 1993.

[20] Y. Choquet-Bruhat and J. W. York Jr. *C. R. Acad. Sci.*, A321(Ser. I:Math):1089, 1995.

[21] Y. Choquet-Bruhat and T. Ruggeri. *Commun. Math. Phys.*, 89:269, 1983.

[22] Y. Choquet-Bruhat and J. W. York. gr-qc/0202014.

[23] A. Dimakis and F. Müller-Hoissen. *Phys. Lett. A*, 142:73, 1989.

[24] F. B. Estabrook, R. S. Robinson, and H. D. Wahlquist. *Class. Quantum Grav.*, 14:1237, 1997.

[25] F. B. Estabrook and H. D. Wahlquist. unpublished.

[26] F. B. Estabrook and H. D. Wahlquist. *J. Math. Phys.*, 5:1629, 1964.

[27] J. Balakrishna *et al. Class. Quantum Grav.*, 13:L135, 1996.

[28] M. Alcubierre *et al. Phys. Rev. D*, 62:044034, 2000.

[29] É. É. Flanagan. gr-qc/9804024.

[30] É. É. Flanagan and S. A. Hughes. *Phys. Rev. D*, 57:4535, 1998.

[31] H. Friedrich. *Class. Quantum Grav.*, 13:1451, 1996.

[32] H. Friedrich and A. D. Rendall. *Lect. Notes Phys.*, 540:127, 2000.

[33] S. Frittelli and R. Gómez, 2003. gr-qc/0310064.

[34] S. Frittelli and O. A. Reula. *Phys. Rev. Lett.*, 76:4667, 1996.

[35] C. Gundlach and P. Walker. *Class. Quantum Grav.*, 16:991, 1999.

[36] M. S. Iriondo, E. O. Leguizamón, and O. A. Reula. *Adv. Theor. Math. Phys.*, 2:1075, 1998.

[37] R. T. Jantzen, P. Carini, and D. Bini. Understanding spacetime splittings and their relationships or gravitoelectromagnetism: the user manual, June 2001. http://www34.homepage.villanova.edu/robert.jantzen/gem/.

[38] B. Kelly, P. Laguna, K. Lockitch, J. Pullin, E. Schnetter, D. Shoemaker, and M. Tiglio. *Phys. Rev. D*, 64:084013, 2001.

[39] L. E. Kidder, M. A. Scheel, and S. A. Teukolsky. *Phys. Rev. D*, 64:064017, 2001.

[40] L. E. Kidder, M. A. Scheel, S. A. Teukolsky, E. D. Carlson, and G. B. Cook. *Phys. Rev. D*, 62:084032, 2000.

[41] L. Lehner. gr-qc/0202055.

[42] L. Lehner. *Class. Quantum Grav.*, 18:R25–R86, 2001.

[43] L. Lehner, M. Huq, M. Anderson, E. Bonning, D. Schaefer, and R. Matzner. *Phys..*, 62:044037, 2000.

[44] R. J. LeVeque. private discussion.

[45] R. J. LeVeque. Clawpack software. http://www.amath.washington.edu/ claw.

[46] R. J. LeVeque. *Numerical Methods for Conservation Laws.* Birkhäuser Verlag, Basel, 1992.

[47] R. J. LeVeque. *J. Comput. Phys.*, 131:327, 1997.

[48] R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems.* Cambridge University Press, 2002.

[49] R. J. LeVeque and M. Pelanti. *J. Comput. Phys.*, 172:572, 2001.

[50] L. Lindblom and M. A. Scheel. *Phys. Rev. D*, 66:084014, 2002.

[51] L. Lindblom and M. A. Scheel. *Phys. Rev. D*, 67:124005, 2003.

[52] R. W. MacCormack. Proc. of the second int. conf. on numerical methods in fluid dynamics. In M. Holt, editor, *Lecture Notes in Physics*, volume 8, Berlin, 1971. Springer-Verlag.

[53] R. Marsa and M. Choptuik. *Phys. Rev. D*, 54:4929, 1996.

[54] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation.* W. H. Freeman and Company, 1973.

[55] J. M. Nester. *J. Math. Phys.*, 30:624, 1989.

[56] J. M. Nester. *J. Math. Phys.*, 33:910, 1992.

[57] E. T. Newman and R. Penrose. *J. Math. Phys.*, 3:566, 1962. erratum **4** 998(E) (1963).

[58] R. Penrose. *Proc. R. Soc. A*, 284:159, 1965.

[59] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes*. Cambridge University Press, Cambridge, England, 1986.

[60] O. A. Reula. Hyperbolic methods for Einstein's equations, 1998. http://www.livingreviews.org/lrr-1998-3.

[61] L. Rezzolla, A. M. Abrahams, R. A. Matzner, M. E. Rupright, and S. L. Shapiro. *Phys. Rev. D*, 59:064001, 1999.

[62] M. A. Scheel, T. W. Baumgarte, G. B. Cook, S. L. Shapiro, and S. A. Teukolsky. *Phys. Rev. D*, 56:6320, 1997.

[63] M. A. Scheel, T. W. Baumgarte, G. B. Cook, S. L. Shapiro, and S. A. Teukolsky. *Phys. Rev. D*, 58:044020, 1998.

[64] M. A. Scheel, L. E. Kidder, L. Lindblom, H. P. Pfeiffer, and S. A. Teukolsky. *Phys. Rev. D*, 66:124005, 2002.

[65] E. Seidel and W.-M. Suen. *Phys. Rev. Lett.*, 69:1845, 1992.

[66] M. Shibata and T. Nakamura. *Phys. Rev. D*, 52:5428, 1995.

[67] H. Shinkai and G. Yoneda. gr-qc/0209111.

[68] H. Shinkai and G. Yoneda. *Class. Quantum Grav.*, 17:4799, 2000.

[69] H. Shinkai and G. Yoneda, 2001. gr-qc/0103031.

[70] L. Smarr and J. W. York. *Phys. Rev. D*, 17:2529, 1978.

[71] G. Strang. *Linear Algebra and Its Applications*. Academic Press, New York, 1976. p. 238.

[72] J. Thornburg. *Class. Quantum Grav.*, 4:1119, 1987.

[73] Kip S. Thorne. *Black Holes and Time Warps*. W. W. Norton and Company, New York, 1994.

[74] H. van Elst and C. Uggla. *Class. Quantum Grav.*, 14:2673, 1997.

[75] M. H. P. M. van Putten. gr-qc/0203076.

[76] M. H. P. M. van Putten. *Phys. Rev. D*, 55:4705, 1997.

[77] M. H. P. M. van Putten and D. M. Eardley. *Phys. Rev. D*, 53:3056, 1996.

[78] H. D. Wahlquist. *J. Math. Phys.*, 33:304, 1992. Appendices A and B.

[79] R. M. Wald. *General Relativity*. The University of Chicago Press, 1984.

[80] H.-J. Yo, T. W. Baumgarte, and S. L. Shapiro. *Phys. Rev. D*, 66:084026, 2002.

[81] G. Yoneda and H. Shinkai. *Phys. Rev. Lett.*, 82:263, 1999.

[82] G. Yoneda and H. Shinkai. *Int. J. Mod. Phys. D*, 9:13, 2000.

[83] G. Yoneda and H. Shinkai. *Class. Quantum Grav.*, 18:441, 2001.

[84] G. Yoneda and H. A. Shinkai. *Phys. Rev. D*, 63:124019, 2001.

[85] U. Yurtsever. *Phys. Rev. D*, 37:2790, 1988.

[86] U. Yurtsever. *Phys. Rev. D*, 38:1731, 1988.

Appendix A

SECOND ORDER ACCURACY OF FLUX-BASED WAVE DECOMPOSITION

In the wave propagation stage of solving for the time evolution of a hyperbolic system, the equations being solved are exact conservation laws (Eq. (2.54)). The integral form of the conservation law can be used to update the average $\mathbf{q}$'s in the ith cell,

$$\mathbf{q}_i^{n+1} \equiv \frac{1}{\triangle x}\int_{cell\ i} \mathbf{q}^{n+1} dx = \mathbf{q}_i^n - \frac{1}{\triangle x}\left[\int_{t_n}^{t_{n+1}} \mathbf{F}_{i+\frac{1}{2}} dt - \int_{t_n}^{t_{n+1}} \mathbf{F}_{i-\frac{1}{2}} dt\right], \qquad \text{(A.1)}$$

where $i \pm \frac{1}{2}$ denotes values on the cell boundaries. We base our discussions of accuracy on Taylor series expansions in both t and x, assuming that the cell size $\triangle x$ and the time step $\triangle t = t_{n+1} - t_n$ are the same order. Time and spatial derivatives are related by the wave propagation equation, Eq. (2.54), and $\mathbf{F}(\mathbf{q}, x) = \mathbf{A}(x)\,\mathbf{q}$. Second order accuracy means that the error in one time step in $\mathbf{q}_i^{n+1}$ decreases faster than $(\triangle t)^2$ as $\triangle t$ and $\triangle x$ go to zero.

Expanding $\mathbf{F}$ in a Taylor series in time,

$$\int_{t_n}^{t_{n+1}} \mathbf{F}_{i\pm\frac{1}{2}} dt \approx \triangle t\ \mathbf{F}^n_{i\pm\frac{1}{2}} + \frac{1}{2}\triangle t^2\ \partial_t \mathbf{F}^n_{i\pm\frac{1}{2}} + \frac{1}{6}\triangle t^3\ \partial_t^2 \mathbf{F}^n_{i\pm\frac{1}{2}}. \qquad \text{(A.2)}$$

In order that $\mathbf{q}_i^{n+1}$ be accurate to second order while neglecting the last term in Eq. (A.2), $\partial_t^2 \mathbf{F}$ must be continuous in x. From the evolution Eq. (2.55) for $\mathbf{F}$, assuming the dependence of $\mathbf{F}$ on $\mathbf{q}$ is analytic, second order continuity in $\partial_t^2 \mathbf{F}$ is guaranteed by second order continuity in $\partial_x^2 \mathbf{F}$. The resulting equation is

$$\int_{t_n}^{t_{n+1}} \mathbf{F}_{i\pm\frac{1}{2}} dt \approx \triangle t\ \mathbf{F}^n_{i\pm\frac{1}{2}} - \frac{1}{2}\triangle t^2\ \mathbf{A}_{i\pm\frac{1}{2}} (\partial_x \mathbf{F})^n_{i\pm\frac{1}{2}}. \qquad \text{(A.3)}$$

Let $\mathbf{F}_i$ be the cell-centered values of $\mathbf{F}$, which differ from $\mathbf{F}(\mathbf{q}_i, x_i)$ by a term of order $\triangle x^2\ \partial_x^2 \mathbf{F}_i$. Then expanding $\mathbf{F}$ in a Taylor series in x about the cell-center value,

$$\mathbf{F}^n_{i\pm\frac{1}{2}} \approx \frac{1}{2}(\mathbf{F}^n_i + \mathbf{F}^n_{i\pm1}) - \frac{1}{8}\triangle x^2\ (\partial_x^2\mathbf{F})^n_i \equiv \mathbf{F}^n_i \pm \frac{1}{2}\triangle \mathbf{F}^n_{i\pm\frac{1}{2}} - \frac{1}{8}\triangle x^2\ (\partial_x^2\mathbf{F})^n_i. \quad \text{(A.4)}$$

Since the second-order error terms are the *same* for $\mathbf{F}_{i+\frac{1}{2}}$ and $\mathbf{F}_{i-\frac{1}{2}}$, and change the same way when $\mathbf{F}_i$ and $\mathbf{F}_{i+1}$ are replaced by $\mathbf{F}(\mathbf{q}_i, x_i)$ and $\mathbf{F}(\mathbf{q}_{i+1}, x_{i+1})$, they cancel in the difference of the flux integrals in Eq. (A.1) and can be omitted, *if* $\partial_x^2\mathbf{F}$ is continuous.

To obtain a second-order accurate contribution to $\mathbf{q}_i^{n+1}$ from the second term in Eq. (A.3), it is sufficient that $\mathbf{A}_{i\pm\frac{1}{2}}$ and $(\partial_x\mathbf{F})^n_{i\pm\frac{1}{2}}$ be accurate to first order. In Method **II**, $\mathbf{A}_{i\pm\frac{1}{2}}$ is approximated to first order by $\frac{1}{2}(\mathbf{A}_i + \mathbf{A}_{i\pm1})$. A Taylor expansion of $(\partial_x\mathbf{F})^n_{i\pm\frac{1}{2}}$ gives, to first order,

$$(\partial_x\mathbf{F})^n_{i\pm\frac{1}{2}} \approx \frac{\triangle\mathbf{F}^n_{i\pm\frac{1}{2}}}{\triangle x}. \quad \text{(A.5)}$$

Flux-based wave decomposition consists of decomposing the flux differences $\triangle\mathbf{F}^n_{i+\frac{1}{2}}$ into a sum over eigenvectors of the characteristic matrix, and similarly for $\triangle\mathbf{F}^n_{i-\frac{1}{2}}$. Different versions evaluate these eigenvectors at different locations. From the point of view of smooth solutions, the preferred location is to base the decomposition of $\mathbf{A}_{i\pm\frac{1}{2}}$ at the respective cell boundaries. The decomposition of $\triangle\mathbf{F}^n_{i+\frac{1}{2}}$ into a sum of right eigenvectors of the matrix $\mathbf{A}_{i+\frac{1}{2}}$ can be written

$$\triangle\mathbf{F}^n_{i+\frac{1}{2}} = \mathbf{R}_{i+\frac{1}{2}}\mathbf{\Gamma}_{i+\frac{1}{2}}, \quad \text{(A.6)}$$

where $\mathbf{R}_{i+\frac{1}{2}}$ is a matrix whose columns are the right eigenvectors $\mathbf{r}^p_{i+\frac{1}{2}}$, and the column vector $\mathbf{\Gamma}_{i+\frac{1}{2}}$ consists of the coefficients of the decomposition, $\gamma^p_{i+\frac{1}{2}}$. Then

$$\mathbf{A}_{i+\frac{1}{2}} \triangle \mathbf{F}^n_{i+\frac{1}{2}} = \mathbf{R}_{i+\frac{1}{2}}\mathbf{\Lambda}_{i+\frac{1}{2}}\mathbf{\Gamma}_{i+\frac{1}{2}}, \quad \text{(A.7)}$$

where $\mathbf{\Lambda}_{i+\frac{1}{2}}$ is the diagonal matrix of eigenvalues of $\mathbf{A}_{i+\frac{1}{2}}$. Let $\lambda^p_{i+\frac{1}{2}}$ be the pth eigenvalue, or wavespeed, and let $\mathbf{W}^p_{i+\frac{1}{2}} \equiv \gamma^p_{i+\frac{1}{2}} \mathbf{r}^p_{i+\frac{1}{2}}$, where $\mathbf{r}^p_{i+\frac{1}{2}}$ is the pth eigenvector, *ie.*, the pth column of $\mathbf{R}_{i+\frac{1}{2}}$. Equation (A.7) can be written

$$\mathbf{A}_{i+\frac{1}{2}} \triangle \mathbf{F}^n_{i+\frac{1}{2}} = \sum_p \lambda^p_{i+\frac{1}{2}} \mathbf{W}^p_{i+\frac{1}{2}}. \tag{A.8}$$

Combining Eqs. (A.3)–(A.5) and (A.8), and noting that $\triangle \mathbf{F}^n_{i+\frac{1}{2}} = \sum_p \mathbf{W}^p_{i+\frac{1}{2}}$, gives

$$\int_{t_n}^{t_{n+1}} \mathbf{F}_{i+\frac{1}{2}} dt \approx \triangle t \left[\mathbf{F}^n_i + \frac{1}{2} \sum_p \mathbf{W}^p_{i+\frac{1}{2}} - \frac{1}{2} \frac{\triangle t}{\triangle x} \sum_p \lambda^p_{i+\frac{1}{2}} \mathbf{W}^p_{i+\frac{1}{2}} \right], \tag{A.9}$$

and a similar expression for the flux integral at $i - \frac{1}{2}$. In Eq. (A.1), these give for $\mathbf{q}^{n+1}_i$ an expression which is easily shown to be equivalent to the substitution of Eqs. (2.58) into Eq. (2.57). Thus, we have shown that Method **II** of Sec. 2.9.2 is second-order accurate for $\partial^2_x \mathbf{F}$ continuous.

Method **I** is based on expressions similar to Eq. (A.9) for the flux integrals, but the eigenvectors in $\mathbf{W}_{i+\frac{1}{2}}$, for instance, are derived from $\mathbf{A}_i$ for the left-going eigenmodes at the $i + \frac{1}{2}$ interface and derived from $\mathbf{A}_{i+1}$ for the right-going eigenmodes. In the last term in Eq. (A.9), the eigenvalues $\lambda^p_{i+\frac{1}{2}}$ are approximated to first-order accuracy as the average of the eigenvalues associated with the adjacent cells, but the $\mathbf{W}_{i+\frac{1}{2}}$ are only zeroth-order accurate. The first-order corrections to the eigenvectors are the same at the $i \pm \frac{1}{2}$ interfaces, if the same eigenmodes are left and right-going at each interface and the first derivatives of the eigenvectors are continuous. (Recall that in Method **I**, the sign of the averaged eigenvalue at a given interface determines the direction of the eigenmode at that interface.) The second condition normally follows from continuity of $\partial_x \mathbf{A}$, but may fail if some of the eigenmodes are nearly degenerate. With this qualification, the first-order corrections in this term will cancel between the two interfaces, as long as the same eigenmodes are left- and right-going at each interface. When the averaged eigenvalues at the cell interfaces do change

sign, continuity of the first derivatives of the eigenvalues implied by continuity of $\partial_x \mathbf{A}$ means $\lambda^p_{i+\frac{1}{2}}$ and $\lambda^p_{i-\frac{1}{2}}$ are both of order $\triangle x$. The last term in Eq. (A.9), which becomes the last term in Eq. (2.58), is then second-order accurate by itself. The terms first-order in $\triangle t$ in Eqs. (2.57) and (2.58) will differ in second order from their values in Method **II**, but combine by construction to give

$$\sum_L \mathbf{W}^L_{i+\frac{1}{2}} + \frac{1}{2}\left[\sum_R \mathbf{W}^R_{i+\frac{1}{2}} - \sum_L \mathbf{W}^L_{i+\frac{1}{2}}\right] = \frac{1}{2}\left[\sum_R \mathbf{W}^R_{i+\frac{1}{2}} + \sum_L \mathbf{W}^L_{i+\frac{1}{2}}\right] \equiv \frac{1}{2} \triangle \mathbf{F}^n_{i+\frac{1}{2}} \tag{A.10}$$

and

$$\sum_R \mathbf{W}^R_{i-\frac{1}{2}} - \frac{1}{2}\left[\sum_R \mathbf{W}^R_{i-\frac{1}{2}} - \sum_L \mathbf{W}^L_{i-\frac{1}{2}}\right] = \frac{1}{2} \triangle \mathbf{F}^n_{i-\frac{1}{2}}. \tag{A.11}$$

The result for $\mathbf{q}^{n+1}_i$ is the same as Method **II** through second order, as long as the eigenvectors are smooth as discussed above. While eigenmodes are degenerate for many hyperbolic formulations of the Einstein equations, the eigenvectors can be chosen with the required smoothness.

Both methods, when the smoothness conditions are satisfied, are equivalent through second order to the Lax-Wendroff finite difference scheme presented in Sec. 2.9.2.

Appendix B

SOURCE TERMS OF EW TETRAD EVOLUTION AND QUASI CONSTRAINT EQUATIONS

The tetrad energy quasi-constraint equation, derived from $G_{00} = 0$, is

$$2\,D_a\,n_a = -2\,\omega_a\,\Omega_a + N_{cd}\,N_{cd} + \frac{1}{2}(K_{cd}\,K_{dc} - (TrK)^2 - N_{cd}\,N_{dc} - (TrN)^2). \quad \text{(B.1)}$$

The tetrad momentum quasi-constraint equations, obtained from $G_{0a} = 0$, are

$$D_1\,(K_{22} + K_{33}) - D_2\,K_{12} - D_3\,K_{13} = 2\,\varepsilon_{1bc}\,a_b\,\Omega_c + \varepsilon_{1bc}\,K_{bd}\,N_{dc} - 2\,K_{1c}\,n_c, \quad \text{(B.2)}$$

$$D_2\,(K_{11} + K_{33}) - D_1\,K_{21} - D_3\,K_{23} = 2\,\varepsilon_{2bc}\,a_b\,\Omega_c + \varepsilon_{2bc}\,K_{bd}\,N_{dc} - 2\,K_{2c}\,n_c, \quad \text{(B.3)}$$

$$D_3\,(K_{11} + K_{22}) - D_1\,K_{31} - D_2\,K_{32} = 2\,\varepsilon_{3bc}\,a_b\,\Omega_c + \varepsilon_{3bc}\,K_{bd}\,N_{dc} - 2\,K_{3c}\,n_c. \quad \text{(B.4)}$$

Analogous quasi-constraint equations for Ω_a and N_{ab} are

$$2\,D_a\,\Omega_a = 2\,a_a\,\Omega_a + 4\,n_a\,\Omega_a, \quad \text{(B.5)}$$

$$\begin{aligned} &D_1\,(N_{22} + N_{33}) - D_2\,N_{12} - D_3\,N_{13} \\ &= -2\,\varepsilon_{1bc}\,\omega_b\,\Omega_c + 2\,K_{1c}\,\Omega_c - \varepsilon_{1bc}\,N_{1b}\,N_{c1} + 2\,n_1\,(N_{22} + N_{33}), \end{aligned} \quad \text{(B.6)}$$

$$\begin{aligned} &D_2\,(N_{11} + N_{33}) - D_1\,N_{21} - D_3\,N_{23} \\ &= -2\,\varepsilon_{2bc}\,\omega_b\,\Omega_c + 2\,K_{2c}\,\Omega_c - \varepsilon_{2bc}\,N_{2b}\,N_{c2} + 2\,n_2\,(N_{11} + N_{33}), \end{aligned} \quad \text{(B.7)}$$

$$\begin{aligned} &D_3\,(N_{11} + N_{22}) - D_1\,N_{31} - D_2\,N_{32} \\ &= -2\,\varepsilon_{3bc}\,\omega_b\,\Omega_c + 2\,K_{3c}\,\Omega_c - \varepsilon_{3bc}\,N_{3b}\,N_{c3} + 2\,n_3\,(N_{11} + N_{22}). \end{aligned} \quad \text{(B.8)}$$

To obtain the true constraint equations from these quasi-constraint equations, Eq. (3.12) must be substituted for each appearance of D_a. Then Eq. (3.55) must be used to eliminate the terms involving D_0 that have been introduced. The true constraints are only needed to check accuracy as the solution evolves, once consistent initial data is obtained.

The sources for K_{ab} and N_{ab} evolution Eqs. (3.38) and (3.39) are

$$S\_K_{ab} = a_a\, a_b + \varepsilon_{bcd}\,(-N_{ac}\, a_d + K_{ac}\,\omega_d) + \varepsilon_{acd}\, K_{cb}\,\omega_d$$
$$+ \frac{1}{2}\,\varepsilon_{adf}\,\varepsilon_{bce}\,(K_{dc}\,K_{fe} - N_{dc}\,N_{fe}) + (TrN)\,N_{ab} - N_{ca}\,N_{cb} - K_{ac}\,K_{cb} + 2\,\omega_b\,\Omega_a, \tag{B.9}$$

$$S\_N_{ab} = -a_a\,\omega_b + \varepsilon_{bcd}\,(K_{ac}\,a_d + N_{ac}\,\omega_d) + \varepsilon_{acd}\,N_{cb}\,\omega_d$$
$$+ \varepsilon_{adf}\,\varepsilon_{bce}\,N_{dc}\,K_{fe} - (TrN)\,K_{ab} + N_{ca}\,K_{cb} - K_{ac}\,N_{cb} + 2\,a_b\,\Omega_a. \tag{B.10}$$

Appendix C

DETAILS OF EW COORDINATE EQUATIONS

In this appendix, we expand Eq. (3.55). Because of the block diagonal structure of the C^a and M^a matrices, it is only necessary to work with one set of eight variables. We will let the free index $d = 1,\ 2,$ or 3 (for the first, second, and third set of eight variables in the vector $\mathbf{q}$), *ie.*

$$\mathbf{q}_d = (N_{1d},\ N_{2d},\ N_{3d},\ a_d,\ K_{1d},\ K_{2d},\ K_{3d},\ \omega_d). \tag{C.1}$$

So as to include a non-zero shift with minimal notation, we use D_0 to represent the partial derivatives in Eq. (3.13). Then we can write Eq. (3.55) as

$$D_0\ \mathbf{q}_d + [\boldsymbol{C}^1_{block}\ B^k_1 + \boldsymbol{C}^2_{block}\ B^k_2 + \boldsymbol{C}^3_{block}\ B^k_3]\ \frac{\partial \mathbf{q}_d}{\partial x^k} = \mathbf{S}_d, \tag{C.2}$$

where

$$\boldsymbol{C}^1_{block} = \frac{1}{1 - A_a\ A_a} \begin{bmatrix} A_1 & -A_2 & -A_3 & 0 & 0 & 0 & 0 & -1 \\ A_2 & A_1 & 0 & -A_3 & 0 & 0 & 1 & 0 \\ A_3 & 0 & A_1 & A_2 & 0 & -1 & 0 & 0 \\ 0 & A_3 & -A_2 & A_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & A_1 & -A_2 & -A_3 & 0 \\ 0 & 0 & -1 & 0 & A_2 & A_1 & 0 & -A_3 \\ 0 & 1 & 0 & 0 & A_3 & 0 & A_1 & A_2 \\ -1 & 0 & 0 & 0 & 0 & A_3 & -A_2 & A_1 \end{bmatrix}, \tag{C.3}$$

$$C^2_{block} = \frac{1}{1 - A_a\, A_a} \begin{bmatrix} A_2 & A_1 & 0 & A_3 & 0 & 0 & -1 & 0 \\ -A_1 & A_2 & -A_3 & 0 & 0 & 0 & 0 & -1 \\ 0 & A_3 & A_2 & -A_1 & 1 & 0 & 0 & 0 \\ -A_3 & 0 & A_1 & A_2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & A_2 & A_1 & 0 & A_3 \\ 0 & 0 & 0 & 1 & -A_1 & A_2 & -A_3 & 0 \\ -1 & 0 & 0 & 0 & 0 & A_3 & A_2 & -A_1 \\ 0 & -1 & 0 & 0 & -A_3 & 0 & A_1 & A_2 \end{bmatrix}, \tag{C.4}$$

$$C^3_{block} = \frac{1}{1 - A_a\, A_a} \begin{bmatrix} A_3 & 0 & A_1 & -A_2 & 0 & 1 & 0 & 0 \\ 0 & A_3 & A_2 & A_1 & -1 & 0 & 0 & 0 \\ -A_1 & -A_2 & A_3 & 0 & 0 & 0 & 0 & -1 \\ A_2 & -A_1 & 0 & A_3 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & A_3 & 0 & A_1 & -A_2 \\ 1 & 0 & 0 & 0 & 0 & A_3 & A_2 & A_1 \\ 0 & 0 & 0 & 1 & -A_1 & -A_2 & A_3 & 0 \\ 0 & 0 & -1 & 0 & A_2 & -A_1 & 0 & A_3 \end{bmatrix}, \tag{C.5}$$

and

$$\mathbf{S}_d = \frac{1}{1 - A_a\, A_a} \begin{bmatrix} A_3\, S\_K_{2d} - A_2\, S\_K_{3d} + S\_N_{1d} - A_1\, S\_\omega_d \\ -A_3\, S\_K_{1d} + A_1\, S\_K_{3d} + S\_N_{2d} - A_2\, S\_\omega_d \\ A_2\, S\_K_{1d} - A_1\, S\_K_{2d} + S\_N_{3d} - A_3\, S\_\omega_d \\ S\_a_d + A_1\, S\_K_{1d} + A_2\, S\_K_{2d} + A_3\, S\_K_{3d} \\ A_1\, S\_a_d + S\_K_{1d} - A_3\, S\_N_{2d} + A_2\, S\_N_{3d} \\ A_2\, S\_a_d + S\_K_{2d} + A_3\, S\_N_{1d} - A_1\, S\_N_{3d} \\ A_3\, S\_a_d + S\_K_{3d} - A_2\, S\_N_{1d} + A_1\, S\_N_{2d} \\ -A_1\, S\_N_{1d} - A_2\, S\_N_{2d} - A_3\, S\_N_{3d} + S\_\omega_d \end{bmatrix}. \tag{C.6}$$

The expressions for S_N_{1d}, S_N_{2d}, S_N_{3d}, S_K_{1d}, S_K_{2d}, S_K_{3d} in Eq. (C.6) are obtained from Eqs. (B.9) and (B.10). Those for S_a_d and S_ω_d are from either Eqs. (3.43) or Eqs. (3.46).

The eight eigenvectors of the arbitrary linear combination $b_a\, C^a_{block}$ consist of four pairs of left and right-propagating modes. Each pair only involves one of the four variables K_{1d}, K_{2d}, K_{3d}, ω_d. For propagation in the x^k coordinate direction, $b_a = B^k_a$. As in Sec. 3.5, we simplify notation by using $\boldsymbol{b}$ and $\boldsymbol{A}$ to denote the 3-vectors in the spatial orthonormal frame with components b_a and A_a. Then the ω_d eigenvectors, normalized so $\omega_d = \boldsymbol{b}\cdot\boldsymbol{b} = |\boldsymbol{b}|^2$, have $N_{bd} = [\boldsymbol{b}\times(\boldsymbol{b}\times\boldsymbol{A})]_b \pm b_b\sqrt{|\boldsymbol{b}|^2 - |\boldsymbol{b}\times\boldsymbol{A}|^2}$, and $a_d = 0$. The K_{ad} eigenvectors, normalized so $K_{ad} = |\boldsymbol{b}|^2$, have $N_{bd} = \varepsilon_{abc}\left([\boldsymbol{b}\times(\boldsymbol{b}\times\boldsymbol{A})]_c \pm b_c\sqrt{|\boldsymbol{b}|^2 - |\boldsymbol{b}\times\boldsymbol{A}|^2}\right)$ and $a_d = -\left([\boldsymbol{b}\times(\boldsymbol{b}\times\boldsymbol{A})]_a \pm b_a\sqrt{|\boldsymbol{b}|^2 - |\boldsymbol{b}\times\boldsymbol{A}|^2}\right)$. The upper sign on the square root corresponds to the upper sign in Eq. (3.58) for the eigenvalues. Note that N_{ad} is zero in the eigenvectors.

Appendix D

EW PLANE WAVE PDE SOURCE

$$\mathbf{S} = \frac{1}{1 - A_1\, A_1} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & A_1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -A_1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -A_1 \\ 0 & 0 & 0 & 1 & 0 & -A_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & A_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -A_1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -A_1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ A_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -A_1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times$$

$$\begin{bmatrix} -a_1\, K_{22} - 2\, K_{23}\, N_{11} - K_{22}\, N_{23} + N_{32}(K_{33} - K_{11}) + 2N_{33}(\omega_1 - K_{23}) \\ a_1\, K_{33} - 2\, K_{23}\, N_{11} + N_{23}(K_{22} - K_{11}) - K_{33}\, N_{32} + 2N_{33}(K_{23} + \omega_1) \\ -K_{11}\, N_{11} - K_{23}(N_{23} + N_{32}) + N_{33}(K_{22} - K_{33}) - a_1\, \omega_1 \\ -a_1\, K_{23} + N_{11}(K_{22} - K_{33}) - K_{11}\, N_{33} - \omega_1(N_{23} + N_{32}) \\ 0 \\ -K_{23}\, TrK + N_{11}(N_{23} + N_{32}) + N_{33}(-a_1 + N_{23} - N_{32}) + \omega_1(K_{33} - K_{22}) \\ a_1^2 - K_{11}^2 - K_{23}^2 + K_{22}\, K_{33} + N_{23}\, N_{32} + N_{33}^2 \\ -K_{22}^2 - K_{23}^2 + K_{11}\, K_{33} - a_1\, N_{23} - N_{32}^2 - 2\, N_{11}\, N_{33} - N_{33}^2 + 2\, K_{23}\, \omega_1 \\ K_{11}\, K_{22} + 2\, N_{11}\, N_{33} - K_{23}^2 - K_{33}^2 - N_{23}^2 - N_{33}^2 + a_1\, N_{32} - 2\, K_{23}\, \omega_1 \\ -a_1\, N_{11} - K_{11}\, \omega_1. \end{bmatrix} \quad \text{(D.1)}$$

VITA

Luisa T. Buchman

Education:

University of Washington

Astronomy: PhD., 2003 (expected), M.S., 1996

Physics: M.S., 2003 (expected)

Harvard-Radcliffe College

Biochemical Sciences: A.B. *cum laude*, 1982

Employment:

University of Washington (1994-2000)

Depts. of Physics and Astronomy: Teaching Assistant

Jet Propulsion Laboratory (1989-1994)

Guidance and Control Section: Member of the Technical Staff

University of California, Los Angeles (1987-1989)

Dept. of Chemistry and Biochemistry: Staff Research Associate

Harvard University (1982-1985)

Dept. of Biochemistry: Research Assistant

Awards and Fellowships:

National Research Council Research Associateship Award (2004-2007)

Graduate Student Researchers Program Fellowship (2000-2003)

NASA Space Grant Award and NSF travel grant (1998)

NASA Group Achievement Award–Cassini (1998)

Achievement Rewards for College Scientists Fellowship (1994-1997)

Publications:

Refereed Journals

1. *A Hyperbolic Tetrad Formulation of the Einstein Equations for Numerical Relativity*, L. T. Buchman and J. M. Bardeen, Phys. Rev. D **67**, 084017 (2003).

2. *Numerical Tests of Evolution Systems, Gauge Conditions, and Boundary Conditions for 1D Colliding Gravitational Plane Waves*, J. M. Bardeen and L. T. Buchman, Phys. Rev. D **65**, 064037 (2002).

3. *Hydrophobic Organization of Membrane Proteins*, D. C. Rees, L. DeAntonio*, D. Eisenberg, Science **245**, 510 (1989).

Proceedings

1. *Topographic / photometric model of small bodies*, M. Pomerantz, L. DeAntonio*, C. C. Cheng, Proceedings of the SPIE–The International Society for Optical Engineering **2466**, 23 (1995).

2. *Detection of Planetary Terrain Features*, S. Udomkesmalee and L. DeAntonio*, Proceedings of the SPIE–The International Society for Optical Engineering, **2466**, 31 (1995).

3. *Star Tracker Based, All-Sky, Autonomous Attitude Determination*, L. DeAntonio*, S. Udomkesmalee, J. Alexander, R. Blue, E. Dennison, G. Sevaston, M. Scholl, Proceedings of the SPIE–Space Guidance, Control, and Tracking, **1949**, 204 (1993).

---

*under my former name, DeAntonio