Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 218 (2006) 607–634
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Numerical stability for finite difference approximations of Einstein's equations

G. Calabrese [a,*], I. Hinder [a], S. Husa [b]

[a] School of Mathematics, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[b] Max-Planck-Institut fur GravitationsPhysik, Albert-Einstein-Institut, Am Mühlenberg 1, D-14476, Golm, Germany

Received 18 May 2005; received in revised form 22 December 2005; accepted 27 February 2006
Available online 19 April 2006

Abstract

We extend the notion of numerical stability of finite difference approximations to include hyperbolic systems that are first order in time and second order in space, such as those that appear in numerical relativity and, more generally, in Hamiltonian formulations of field theories. By analyzing the symbol of the second order system, we obtain necessary and sufficient conditions for stability in a discrete norm containing one-sided difference operators. We prove stability for certain toy models and the linearized Nagy–Ortiz–Reula formulation of Einstein's equations.

We also find that, unlike in the fully first order case, standard discretizations of some well-posed problems lead to unstable schemes and that the Courant limits are not always simply related to the characteristic speeds of the continuum problem. Finally, we propose methods for testing stability for second order in space hyperbolic systems.

PACS: 02.70.Bf; 04.25.Dm; 04.20.Ex

Keywords: Numerical relativity; Finite difference methods; Second derivatives; Discrete norms; Numerical stability

1. Introduction

The Einstein equations consist of a set of 10 coupled non-linear second order partial differential equations. In order to perform numerical time evolutions the fully second order system is usually written as a first order in time system. Such systems can be evolved directly [1,2], or a further reduction from second to first spatial order can be performed (see for example [3–6]). Whereas the theory of Cauchy problems for fully first order systems of partial differential equations is understood, in terms of well-posedness at the continuum and the stability of finite difference approximations, the theory of second order in space hyperbolic systems is less well

* Corresponding author. Tel.: +44 2380 593863; fax: +44 2380 595147.
E-mail address: G.Calabrese@soton.ac.uk (G. Calabrese).

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2006.02.027

developed. The recent improvement in the understanding of second order in space formulations of Einstein's equations at the continuum [7–11] has not been matched by developments concerning finite difference approximations of such systems (see however [12,13]). Given that these systems have fewer variables, fewer constraints, and typically smaller errors (see [12] and Appendix B), it is desirable to better appreciate their properties. Note that first order in time hyperbolic systems, which are not necessarily first order in space, also arise naturally in Hamiltonian formulations of field theories.

The standard notion of stability for fully first order systems based on the discrete L_2 norm is unsuitable for analyzing second order in space hyperbolic systems. This can be understood by analogy with the continuum result for the one-dimensional wave equation written in first order in time and second order in space form: $\partial_t\phi(t,x) = \Pi(t,x)$, $\partial_t\Pi(t,x) = \partial_x^2\phi(t,x)$. Consider the family of solutions $\phi(x,t) = \sin(\omega x)\cos(\omega t)$, $\pi(x,t) = -\omega\sin(\omega x)\sin(\omega t)$ generated by the initial data $\phi_0(x) = \sin(\omega x)$, $\pi_0(x) = 0$. By varying ω in the initial data, the L_2 norm of the solution at a fixed time t, $\int_0^{2\pi}(|\phi|^2 + |\Pi|^2)\,dx$, can be made arbitrarily large with respect to the initial data (whose norm is independent of ω), thus contradicting well-posedness of the Cauchy problem in L_2 [14,15]. The introduction of the new variable, $X = \partial_x\phi$, allows the construction of a first order system, the Cauchy problem of which is well-posed in L_2. The original second order problem can then be shown to be well-posed in a norm containing derivatives, namely $\int_0^{2\pi}(|\phi|^2 + |\Pi|^2 + |\partial_x\phi|^2)\,dx$, which corresponds to the L_2 norm of the first order reduction.

In this work, we consider linear constant coefficient Cauchy problems. We use the method of lines to separate the time integration from the spatial discretization. We show that by reducing the discrete system to first order in Fourier space, it is possible to determine stability in physical space with respect to a discrete norm containing one-sided difference operators. This is done by extending the notion of a symmetrizer to the discrete case. We apply these techniques to problems, starting with the wave equation written as a first order in time, second order in space system. We consider both second and fourth order accurate discretizations. A similar but more complicated analysis is done for the Knapp–Walker–Baumgarte (KWB) [16] and Z1 [17] formulations of electromagnetism, and the Nagy–Ortiz–Reula (NOR) [8] formulation of Einstein's equations. We also point out stability issues related to the ADM [18] and Z4 [19] formulations.

In Section 2, we summarize some relevant material from the literature. In Section 3, we introduce the concept of a discrete symmetrizer. We also illustrate the reduction procedure to first order in Fourier space, which can be used for obtaining energy estimates at the continuum. We introduce the analogous idea for the discrete case, and discuss convergence. In Section 4, we apply these techniques to the systems mentioned above. We propose methods in Section 5 for testing stability experimentally both for linear and non-linear systems. We summarize the main results of this paper in Section 6. In Appendix A, we describe the different time integration methods that we consider, and in Appendix B, we compare numerical properties of the wave equation written as a first order system with those of the wave equation written as a first order in time, second order in space system. In Appendix C, we highlight differences in the constraint propagation properties between first and second order systems.

2. Background

Well-posedness, the (local in time) existence of a unique solution which depends continuously on the problem's data, is a fundamental requirement for the successful generation of numerical solutions approximating the solution of a continuum problem. In this section, we review the notion of well-posedness for linear constant coefficient Cauchy problems, as well as the concept of stability for finite difference approximations. We conclude the section by providing a simple sufficient condition for stability of first order fully discrete problems based on the properties of the symbol of the semi-discrete system, which will be extended to discretizations of second order in space problems in the next section.

2.1. Constant coefficient Cauchy problems

In this work, we will be dealing with initial value (or Cauchy) problems of the form

$$\frac{\partial}{\partial t}u(t,x) = P\left(\frac{\partial}{\partial x}\right)u(t,x), \tag{1}$$

$$u(0,x) = f(x), \tag{2}$$

in d spatial dimensions, where $x \in \mathbb{R}^d$, $u = (u^{(1)}, u^{(2)}, \ldots, u^{(m)})^{\mathrm{T}}$ and P is a linear, constant coefficient, differential operator of order p. We consider only the cases $p = 1$ and $p = 2$. Furthermore, we assume that the eigenvalues of the symbol of the differential operator, $\hat{P}(\mathrm{i}\omega)$, which is obtained by replacing $\partial/\partial x_j$ in $P(\partial/\partial x)$ with $\mathrm{i}\omega_j$, for $j = 1, 2, \ldots, d$, have real part uniformly bounded from below and above. We are thus excluding parabolic systems, but we are allowing for systems like the wave equation written as a first order in time, second order in space system. For simplicity we focus on solutions that are 2π-periodic in all spatial coordinate directions. Thus the initial data, $f(x)$, is chosen so that it satisfies this property.

We consider the $p = 1$ case, leaving the $p = 2$ case for the next section. Following Definition 4.1.1 in [20] we say that problem (1) and (2) is well-posed with respect to a norm $\|\cdot\|$ if for every smooth periodic f there is a unique smooth spatially periodic solution and there are constants α and K, independent of f, such that for $t \geqslant 0$

$$\|u(t,\cdot)\| \leqslant K\mathrm{e}^{\alpha t}\|f\|. \tag{3}$$

Exponential growth must be allowed if one wants to treat problems with lower order terms. For first order hyperbolic systems the L_2 norm $\|w\|^2 = \int_0^{2\pi} \ldots \int_0^{2\pi} |w(x)|^2\, \mathrm{d}x_1 \ldots \mathrm{d}x_d$ is usually used in (3). We will see later that the second order systems we study in this work require the use of a different norm.

Taking $f(x) = (2\pi)^{-d/2}\sum_\omega \mathrm{e}^{\mathrm{i}\langle\omega,x\rangle}\hat{f}(\omega)$ the formal solution of (1) and (2) is $u(t,x) = (2\pi)^{-d/2}\sum_\omega \mathrm{e}^{\mathrm{i}\langle\omega,x\rangle} \mathrm{e}^{\hat{P}(\mathrm{i}\omega)t}\hat{f}(\omega)$. It can be shown (Theorem 4.5.1 in [20]) that well-posedness in the L_2 norm is equivalent to there being constants K, α such that, for all ω and for $t \geqslant 0$,

$$|\mathrm{e}^{\hat{P}(\mathrm{i}\omega)t}| \leqslant K\mathrm{e}^{\alpha t}, \tag{4}$$

where $|A| = \sup_{|u|=1}|Au|$ is the matrix (operator) norm of a matrix A.

Well-posedness of the Cauchy problem in the L_2 norm is also equivalent (Theorem 4.5.8 in [20]) to the existence of constants α, $K > 0$ and of Hermitian matrices $\hat{H}(\omega)$ satisfying,[1] for every ω,

$$\begin{aligned} &K^{-1}I \leqslant \hat{H}(\omega) \leqslant KI, \\ &\hat{H}(\omega)\hat{P}(\mathrm{i}\omega) + \hat{P}^*(\mathrm{i}\omega)\hat{H}(\omega) \leqslant 2\alpha\hat{H}(\omega), \end{aligned} \tag{5}$$

where $\hat{P}^*$ represents the Hermitian conjugate of $\hat{P}$. The last inequality gives an energy estimate for each Fourier mode and the estimate in physical space, Eq. (3), follows from Parseval's relation, $\|u(t,\cdot)\|^2 = \sum_\omega|\hat{u}(t,\omega)|^2$. Since the existence of $\hat{H}(\omega)$ is not affected by the addition of a constant matrix to $\hat{P}(\mathrm{i}\omega)$ (Lemma 2.3.5 in [21]), undifferentiated terms on the right-hand side of the equations can be ignored in the analysis of well-posedness. If (5) is satisfied with $\hat{H}\hat{P} + \hat{P}^*\hat{H} = 0$ then $\hat{H}$ is called a *symmetrizer*.

For $p = 1$, system (1) is said to be *strongly hyperbolic* if the corresponding Cauchy problem is well-posed in the L_2 norm (i.e. if $\hat{H}(\omega)$ exists).[2] If $\hat{H}(\omega) = I$, the system is said to be *symmetric hyperbolic*. If $\hat{H}(\omega) = H$ is independent of ω, then we say that the system is *symmetrizable hyperbolic*.[3] In this case the change of variables $\tilde{u} = H^{1/2}u$ brings the system into symmetric hyperbolic form. Finally, well-posedness is not affected by the presence of forcing (inhomogeneous) terms (Theorem 4.7.2 in [20]). For cases where such terms are present, the estimate requires modification.

Note that, in the absence of lower order terms, whereas symmetrizable hyperbolicity guarantees the existence of a conserved energy in physical space, (u, Hu), a strongly hyperbolic system satisfies the estimate $\|u(t,\cdot)\| \leqslant K\|u(0,\cdot)\|$ with a constant $K \geqslant 1$. Furthermore, in the variable coefficient case, well-posedness results require smoothness of the symmetrizer $\hat{H}(x,t,\omega)$ in all arguments [21].

[1] Two Hermitian matrices, A and B, satisfy $A \leqslant B$ if and only if $x^*Ax \leqslant x^*B\,x$ for every x. If a matrix $\hat{H}(\omega)$ satisfies $K^{-1}I \leqslant \hat{H}(\omega) \leqslant KI$ for every ω, we say that $\hat{H}(\omega)$ is *equivalent to the identity matrix*.

[2] For $\partial_t u = A^i\partial_i u$ the symbol is $\hat{P} = \mathrm{i}\omega_i A^i$. The system is said to be weakly hyperbolic if the eigenvalues of $\hat{P}(\mathrm{i}\omega)$ are imaginary. Strong hyperbolicity is equivalent to $\hat{P}(\mathrm{i}\omega)$ being uniformly diagonalizable with imaginary eigenvalues. We define the characteristic speeds in the direction ω_i to be the eigenvalues of $\hat{P}(\mathrm{i}\omega)$ divided by $\mathrm{i}\omega$.

[3] Symmetrizable hyperbolic systems are often also called symmetric hyperbolic.

2.2. Numerical stability

2.2.1. Notation

Our notation and conventions follow closely those of [20]. We introduce a spatial grid $x_j = (x^{(1)}_{j_1}, x^{(2)}_{j_2}, \ldots, x^{(d)}_{j_d}) = (j_1h_1, j_2h_2, \ldots, j_dh_d)$, where $h_r = 2\pi/N_r$ and $j_r = 0,1,\ldots,N_r-1$, and the vector-valued grid function $v_j(t)$ approximating $u(t,x_j)$. Periodicity requires that $v_j = v_{\mathrm{mod}(j,N)}$. The partial derivatives in (1) are approximated using either the *standard second order accurate discretization*

$$\partial_i \to D_{0i}, \quad \partial_i\partial_j \to \begin{cases} D_{0i}D_{0j} & \text{if } i \neq j, \\ D_{+i}D_{-i} & \text{if } i = j \end{cases} \tag{6}$$

or the *standard fourth order accurate discretization*

$$\begin{aligned} &\partial_i \to D^{(4)}_i \equiv D_{0i}\left(1 - \frac{h^2}{6}D_{+i}D_{-i}\right), \\ &\partial_i\partial_j \to \begin{cases} D^{(4)}_i D^{(4)}_j & \text{if } i \neq j, \\ D_{+i}D_{-i}\left(1 - \frac{h^2}{12}D_{+i}D_{-i}\right) & \text{if } i = j, \end{cases} \end{aligned} \tag{7}$$

where $D_+v_j = (v_{j+1} - v_j)/h$, $D_-v_j = (v_j - v_{j-1})/h$, $D_0v_j = (v_{j+1} - v_{j-1})/2h$, and $D_+D_-v_j = (v_{j+1} - 2v_j + v_{j-1})/h^2$. The discretization of ∂_i^2 as in (6) or (7) gives the desired order of local accuracy without requiring a larger stencil. We then integrate the resulting system of $m\prod_{r=1}^{d} N_r$ ordinary differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}v_j(t) = Pv_j(t), \tag{8}$$

$$v_j(0) = f_j, \tag{9}$$

where $f_j = f(x_j)$, with three different time integrators. These are iterative Crank Nicholson (ICN) and third and fourth order Runge–Kutta (3RK and 4RK) methods, which are widely used by numerical relativists (see Appendix A for definitions). Using the fact that the operator P is linear and time independent we can write the fully discrete system in polynomial form (see for example [20])

$$v^{n+1}_j = Qv^n_j = \mathscr{P}(kP)v^n_j, \tag{10}$$

$$v^0_j = f_j, \tag{11}$$

where $k = \lambda h$ is the time step, λ is called the *Courant factor*, and v^n_j represents the grid-function at time $t_n = nk$. This is an explicit, one step, scheme. For ICN we have $\mathscr{P}(x) = 1 + 2\sum_{r=1}^{3}\frac{x^r}{2^r}$, whereas for pth order Runge–Kutta we have $\mathscr{P}(x) = \sum_{r=0}^{p}\frac{x^r}{r!}$.

2.2.2. Definition of stability

We recall the definition of numerical stability and discuss some necessary and sufficient conditions. The solution of the finite difference scheme (10) and (11) is $v^n = Q^nf$. We introduce the scalar product $(u,v)_h = \sum_j\langle u_j, v_j\rangle h^d$, where $h^d = \prod_{i=1}^{d}h_i$, $j = (j_1, j_2, \ldots, j_d)$ is a multi-index and $\langle u_j, v_j\rangle = \sum_{r=1}^{m}\bar{u}^{(r)}_j v^{(r)}_j$. This allows us to define a norm $\|v\|_h = (v,v)_h^{1/2}$. The approximation (10) and (11) is said to be *stable* with respect to this norm if there exist constants α, K, such that for all h, k, $0 < h \leqslant h_0$, $0 < k \leqslant k_0$, the estimate

$$\|v^n\|_h \leqslant K\mathrm{e}^{\alpha t_n}\|f\|_h \tag{12}$$

holds for all n such that $t_n = nk$ and all initial grid-functions f. This concept of stability is the discrete analogue of (3). It guarantees that the solutions are bounded as $h \to 0$. However, the schemes we consider are at most conditionally stable. By this we mean that there exists a λ_0 such that the above inequality holds if and only if the additional condition $\lambda = k/h \leqslant \lambda_0$ is satisfied.

Theorem 5.1.2 in [20] guarantees that if the scheme (10) and (11) is stable, then the modified scheme

$$v^{n+1}_j = (Q + kR)v^n_j, \tag{13}$$

$$v^0_j = f_j \tag{14}$$

is also stable provided that R is bounded. This will be the case when R represents constant terms (lower order terms) in the continuum problem. Hence for a first order in space system lower order terms can be ignored without affecting stability.

2.2.3. Convergence

Following Theorem 5.1.3 in [20], consistency and stability imply convergence. Assume that the continuum solution u of (1) and (2) is smooth and that the scheme (10) and (11) is stable. Further assume that the scheme and the initial data are consistent. Then, on any finite interval $[0, T]$, the error satisfies

$$\|v^n - u(\cdot, t_n)\|_h \leqslant \mathrm{O}(h^{p_1} + k^{p_2}), \tag{15}$$

i.e. the solutions of the finite difference scheme converge as $h \to 0$ to the solution of the differential equation.[4]

2.2.4. Fourier analysis of stability

For approximations with constant coefficients, Fourier analysis can be used to obtain necessary and sufficient conditions for stability which can be more easily verified than the above definition. We assume that N, the number of grid-points in each direction, is even (the odd case is discussed in Section 2.2.5). If we represent v_j^n by

$$v_j^n = \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{\omega} \mathrm{e}^{\mathrm{i}\langle\omega, x_j\rangle} \hat{v}^n(\omega), \tag{16}$$

where $\omega = (\omega_1, \omega_2, \ldots, \omega_d)$, $\omega_r = -N/2 + 1, \ldots, N/2$, and substitute it into the difference scheme (10) and (11), we obtain

$$\hat{v}^{n+1}(\omega) = \hat{Q}(\xi)\hat{v}^n(\omega), \tag{17}$$

$$\hat{v}^0(\omega) = \hat{f}(\omega), \tag{18}$$

where $\xi_r = \omega_r h = -\pi + 2\pi/N, -\pi + 4\pi/N, \ldots, +\pi$ and $r = 1, 2, \ldots, d$. The $m \times m$ matrix $\hat{Q}(\xi)$ is called the *amplification matrix* of the scheme and is a real polynomial in $\hat{P}$, the symbol of the Fourier transformed semi-discrete problem,

$$\hat{Q}(\xi) = \mathscr{P}(k\hat{P}(\xi)). \tag{19}$$

The matrix $\hat{P}(\xi)$ will play an important role in the next section. It can be readily computed from P in Eq. (8) with the replacements

$$D_{0i} \to \frac{\mathrm{i}}{h} \sin \xi_i, \tag{20}$$

$$D_{+i} D_{-i} \to -\frac{4}{h^2} \sin^2 \frac{\xi_i}{2}. \tag{21}$$

Using the discrete Parseval's relation

$$\|v\|_h^2 = \sum_{\omega} |\hat{v}(\omega)|^2 \tag{22}$$

and the fact that the solution of (17) and (18) is $\hat{v}^n(\omega) = \hat{Q}^n(\xi)\hat{f}(\omega)$ one can show (Theorem 5.2.1 of [20]) that a necessary and sufficient condition for stability with respect to the $\|\cdot\|_h$ norm is given by

$$|\hat{Q}^n(\xi)| \leqslant K\mathrm{e}^{\alpha t_n} \tag{23}$$

for all $h = 2\pi/N \leqslant h_0$, $k \leqslant k_0$, n with $t_n = nk$, and $\omega_r = -N/2 + 1, \ldots, N/2$, $r = 1, 2, \ldots, d$.

A much easier condition to verify is the von Neumann condition, which is only a necessary condition for stability. It corresponds to the requirement that the eigenvalues $z_\nu(\xi)$ of $\hat{Q}(\xi)$ satisfy

[4] Note that the big O in inequality (15) contains higher derivatives of the exact solution. Smoothness of the solution of the continuum problem is not required for convergence. For instance, a weaker condition for fourth order convergence ($p_1 = p_2 = 4$) is that the solution be C^5.

$$|z_v(\xi)| \leqslant e^{\alpha k} \tag{24}$$

for all $h \leqslant h_0$ and $|\xi_r| \leqslant \pi$. However, when the amplification matrix can be uniformly diagonalized (i.e. there exists a non-singular matrix $T(\xi)$ that diagonalizes $\hat{Q}(\xi)$ and satisfies $|T(\xi)||T^{-1}(\xi)| \leqslant C$ with C independent of ξ) then the von Neumann condition is also sufficient for stability. In particular, if $\hat{Q}$ is normal then it can be unitarily (and therefore uniformly) diagonalized, $|T(\xi)| = |T^{-1}(\xi)| = 1$. Since for the time integrators that we consider $\hat{Q}$ is a polynomial in $\hat{P}$, $\hat{Q}$ will be normal if $\hat{P}$ is normal (as would be the case if $\hat{P}$ were Hermitian or anti-Hermitian). Note that if the von Neumann condition is violated then the scheme is not stable in any sense.

It is possible for a discretization to be (conditionally) stable without $\hat{Q}$ being normal (and hence unitarily diagonalizable). This turns out to be the case for most systems considered in this work. In such cases we find it convenient to introduce the norm $|\hat{u}|_{\hat{H}} = \langle \hat{u}, \hat{H}\hat{u}\rangle^{1/2}$ and proceed as follows. Let us assume that $\hat{H}(\xi)$ are Hermitian matrices such that

$$\begin{aligned} &K^{-1}I \leqslant \hat{H}(\xi) \leqslant KI, \\ &|\hat{Q}|_{\hat{H}} \leqslant e^{\alpha k}, \end{aligned} \tag{25}$$

where K is a positive constant. Notice that[5] $|\hat{u}|_{\hat{H}} = |\hat{H}^{1/2}\hat{u}| \leqslant K^{1/2}|\hat{u}|$ and $K^{-1}|A| \leqslant |A|_{\hat{H}} = |\hat{H}^{1/2}A\hat{H}^{-1/2}| \leqslant K|A|$. As a consequence the von Neumann condition is satisfied, $\sigma(\hat{Q}) = \sigma(\hat{H}^{1/2}\hat{Q}\hat{H}^{-1/2}) \leqslant |\hat{H}^{1/2}\hat{Q}\hat{H}^{-1/2}| = |\hat{Q}|_{\hat{H}} \leqslant e^{\alpha k}$, where $\sigma(\hat{Q})$ denotes the spectral radius of $\hat{Q}$. Stability follows from

$$|\hat{Q}^n| \leqslant K|\hat{Q}^n|_{\hat{H}} \leqslant K|\hat{Q}|^n_{\hat{H}} \leqslant K e^{\alpha t_n}. \tag{26}$$

According to the Kreiss Matrix theorem (Section 4.9 of [14]), for a family $\mathscr{F}$ of $m \times m$ matrices A the following two statements are equivalent:

(1) There exists a constant C such that for all $A \in \mathscr{F}$ and all positive integers n

$$|A^n| \leqslant C.$$

(2) There is a constant $K > 0$ and, for each $A \in \mathscr{F}$, a positive definite Hermitian matrix H with the properties

$$K^{-1}I \leqslant H \leqslant KI, \quad A^*HA \leqslant H.$$

This implies that condition (26) is also necessary for stability.

2.2.5. Number of grid points

In this review, we have assumed that the number of grid points in each direction is even. This means that no matter how small the number of grid points is, as long as it is even, the highest frequency $\xi_r = \pi$ is present. To allow for an odd number of grid points one must change the summation range in Eq. (16) to $\omega_r = -(N-1)/2, \ldots, (N-1)/2$, in which case, $|\xi_r|$ never equals π, although it does approach this value as $h \to 0$.

2.3. A sufficient condition for stability

We can now give a simpler sufficient condition for numerical stability. This condition applies to systems which admit a conserved energy in Fourier space and will enable us in Section 3.2 to obtain another condition suitable for the applications. We consider only time integrators such that

$$\hat{Q} = \mathscr{P}(k\hat{P}). \tag{27}$$

The eigenvalues q_v of $\hat{Q}$ are related to the eigenvalues p_v of $\hat{P}$ by $q_v = \mathscr{P}(kp_v)$. This can be seen by using Shur's lemma. Provided that the eigenvalues p_v are imaginary, the inequality $|q_v| \leqslant 1$ is equivalent to $kp_v \leqslant \alpha_0$, where $\alpha_0 = 2$ for ICN, $\sqrt{8}$ for 4RK, $\sqrt{3}$ for 3RK. Hence,

[5] For a positive definite Hermitian matrix H, H^α (for α not necessarily an integer) is defined as $S^*D^\alpha S$ where $H = S^*DS$ and D is the diagonal matrix of positive real eigenvalues.

$$\sigma(k\hat{P}) \leqslant \alpha_0 \tag{28}$$

is equivalent to $\sigma(\hat{Q}) \leqslant 1$. Condition (28) is called *local stability on the imaginary axis* in [22]. Suppose that the time step is such that $\sigma(k\hat{P}) \leqslant \alpha_0$. If we can find Hermitian matrices $\hat{H}(\xi)$ such that

$$K^{-1}I \leqslant \hat{H}(\xi) \leqslant KI, \tag{29}$$

$$\hat{H}(\xi)\hat{P}(\xi) + \hat{P}(\xi)^*\hat{H}(\xi) = 0, \tag{30}$$

we say that $\hat{H}(\xi)$ is a *discrete symmetrizer* of $\hat{P}(\xi)$. The matrices $\hat{H}^{1/2}\hat{P}\hat{H}^{-1/2}$ are anti-Hermitian, hence they can be diagonalized by unitary matrices $S(\xi)$. This implies that the matrices $\hat{H}^{-1/2}(\xi)S(\xi)$ diagonalize $\hat{Q}(\xi)$. The inequality

$$|\hat{Q}|_H = |\hat{H}^{1/2}\hat{Q}\hat{H}^{-1/2}| = |S^{-1}\hat{H}^{1/2}\hat{Q}\hat{H}^{-1/2}S| = \sigma(\hat{Q}) \leqslant 1 \tag{31}$$

guarantees stability. In fact, the amplification matrix can be uniformly diagonalized by $T(\xi) = \hat{H}^{-1/2}(\xi)S(\xi)$.

In applications one would construct a norm (i.e., matrices $\hat{H}(\xi)$ satisfying (29)) which is conserved by the Fourier transformed semi-discrete evolution equations,

$$\frac{\mathrm{d}}{\mathrm{d}t}|\hat{v}|^2_{\hat{H}} = \langle \hat{v}, (\hat{H}\hat{P} + \hat{P}^*\hat{H})\hat{v}\rangle = 0. \tag{32}$$

This implies that condition (30) holds and $\hat{H}(\xi)$ is a discrete symmetrizer.

To construct $\hat{H}$ one can proceed as follows. Assume the existence of a matrix T such that $T^{-1}\hat{P}T = \Lambda$ is diagonal with imaginary elements. Then the quantity $\hat{v}^*\hat{H}\hat{v}$, where $\hat{H} = T^{-1*}DT^{-1}$ and D is a positive definite matrix which commutes with Λ, is conserved by the system $\partial_t\hat{v} = \hat{P}\hat{v}$. Defining the *characteristic variables* of $\hat{P}$ to be $\hat{w} \equiv T^{-1}\hat{v}$ (these are individually conserved: $\partial_t|\hat{w}_i|^2 = 0$), we see that to construct a conserved quantity one can take $\hat{w}^*D\hat{w}$. (For $D = I$ this corresponds to adding the squared absolute values of the characteristic variables.) For $\hat{H}$ to be a symmetrizer it remains to be established that $K^{-1}|\hat{v}|^2 \leqslant \hat{v}^*\hat{H}\hat{v} \leqslant K|\hat{v}|^2$.

3. Stability of first order in time, second order in space systems

What we have done so far applies to fully first order systems. We have shown that if inequalities (28) and (29) and Eq. (32) hold, then the fully discrete scheme is stable and satisfies the estimate (12) with $\alpha = 0$. In this section, we show how this can be extended to second order in space systems. We first look at the continuum problem and then investigate its standard discretization.

3.1. Well-posedness of first order in time and second order in space hyperbolic systems

It is possible for the Cauchy problem of a first order in time and second order in space system of equations to be ill-posed in the L_2 norm, but well-posed in a norm which contains additional derivatives (see Section 1). The system is still useful; for example, a suitable finite difference approximation of the equations can be convergent in the discrete L_2 norm. We analyze the well-posedness of the Cauchy problem for such systems by using the analytical tool of a *reduction to first order*. This will be done in Fourier space, so that the number of additional variables being introduced is minimized [23].

Consider system (1) with $p = 2$ and suppose that it can be written in the form

$$\begin{aligned} &\partial_t\boldsymbol{u} = P\boldsymbol{u}, \quad \boldsymbol{u} = \begin{pmatrix} u \\ v \end{pmatrix}, \\ &P = \begin{pmatrix} A^i\partial_i + B & C \\ D^{ij}\partial_i\partial_j + E^i\partial_i + F & G^i\partial_i + J \end{pmatrix}, \end{aligned} \tag{33}$$

where the evolved variables have been split into two types. The column vector u represents those that are differentiated twice (in space) and v represents those that are not. In P a sum over repeated indices is assumed. Not all second order in space systems can be written in this form (for example, $u_t = u_{xx}$). This form is general enough to include all the first order in time, second order in space systems that we have considered that can be reduced to first order in space. Fourier transforming this system, we obtain

$$\partial_t \hat{\boldsymbol{u}} = \hat{P}\hat{\boldsymbol{u}}, \quad \hat{\boldsymbol{u}} = \begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix},$$
$$\hat{P} = \begin{pmatrix} i\omega A^n + B & C \\ -\omega^2 D^{nn} + i\omega E^n + F & i\omega G^n + J \end{pmatrix}, \tag{34}$$

where $M^n \equiv M^i n_i$ and $\omega_i \equiv |\omega|\ n_i$ and $\omega \equiv |\omega|$. We define the *second order principal symbol* to be

$$\hat{P}' = \begin{pmatrix} i\omega A^n & C \\ -\omega^2 D^{nn} & i\omega G^n \end{pmatrix}. \tag{35}$$

We now state the main result of this subsection. If there exists $\hat{H}(\omega) = \hat{H}^*(\omega)$ such that the *energy* $\hat{\boldsymbol{u}}^*\hat{H}\hat{\boldsymbol{u}}$ is conserved by the principal system $\partial_t \hat{\boldsymbol{u}} = \hat{P}'\hat{\boldsymbol{u}}$ and $\hat{H}$ satisfies

$$K^{-1} I_\omega \leqslant \hat{H} \leqslant K I_\omega, \quad I_\omega \equiv \begin{pmatrix} \omega^2 & 0 \\ 0 & 1 \end{pmatrix}, \tag{36}$$

where K is a positive scalar constant, then the solution of (33) satisfies the estimate

$$\|\boldsymbol{u}(t,\cdot)\| \leqslant K e^{\alpha t} \|\boldsymbol{u}(0,\cdot)\|,$$
$$\|\boldsymbol{u}\|^2 \equiv \int |u|^2 + \sum_{i=1}^{d} |\partial_i u|^2 + |v|^2 \, d^d x, \tag{37}$$

and the problem is well-posed in this norm.[6]

The proof proceeds via a pseudo-differential reduction to first order [8]. This involves the introduction of a new variable $\hat{w} = i\omega\hat{u}$. By taking a time derivative of this definition, we obtain the enlarged system in which the second derivative of $\hat{u}$ has been replaced with a first derivative of $\hat{w}$. We reduce the order of the system as much as possible so that any occurrence of $i\omega\hat{u}$ is replaced with $\hat{w}$. This particular first order reduction is

$$\partial_t \hat{\boldsymbol{u}}_R = \hat{P}_R \hat{\boldsymbol{u}}_R, \quad \hat{\boldsymbol{u}}_R = \begin{pmatrix} \hat{u} \\ \hat{w} \\ \hat{v} \end{pmatrix},$$
$$\hat{P}_R = \begin{pmatrix} B & A^n & C \\ 0 & i\omega A^n + B & i\omega C \\ F & i\omega D^{nn} + E^n & i\omega G^n + J \end{pmatrix}. \tag{38}$$

This system is equivalent to the second order system (34) only when the *auxiliary constraints*

$$\hat{C}(t,\omega) \equiv \hat{w}(t,\omega) - i\omega\hat{u}(t,\omega) = 0 \tag{39}$$

are satisfied. It can be shown that $\partial_t \hat{C} = B\hat{C}$ so if these constraints are satisfied initially, then they are satisfied for all time. They are said to be *propagated* by the first order evolution equations.

If this system admits a matrix $\hat{H}_R$ satisfying (5), then the solutions satisfy the estimates

$$|\hat{\boldsymbol{u}}_R(t,\omega)| \leqslant K e^{\alpha t} |\hat{\boldsymbol{u}}_R(0,\omega)|, \tag{40}$$

where $|\hat{\boldsymbol{u}}_R|^2 \equiv |\hat{u}|^2 + |\hat{w}|^2 + |\hat{v}|^2$, for arbitrary initial data and ω. Specifically, the estimate holds for solutions which satisfy the auxiliary constraints and therefore correspond to solutions of the second order system. The uniform estimate in ω of

$$|\hat{u}|^2 + \omega^2|\hat{u}|^2 + |\hat{v}|^2 = |\hat{u}|^2 + \sum_{i=1}^{d} |i\omega_i \hat{u}|^2 + |\hat{v}|^2 \tag{41}$$

implies, by Parseval's relation, the estimate in real space

[6] Note that we made no assumptions regarding the smoothness of the matrix $\hat{H}(\omega)$. In view of generalizations of this work to the variable coefficient case it may be desirable to demand that $T^{-1*}\hat{H}T^{-1}$, where T is defined in Eq. (44), be smooth in all variables.

$$\|\boldsymbol{u}(t,\cdot)\| \leqslant K e^{\alpha t}\|\boldsymbol{u}(0,\cdot)\|,$$
$$\|\boldsymbol{u}\|^2 \equiv \int |u|^2 + \sum_{i=1}^{d} |\partial_i u|^2 + |v|^2 \, d^d x. \tag{42}$$

So the existence of $\hat{H}_R$ for a first order pseudo-differential reduction implies the well-posedness of the second order system with respect to a norm containing derivatives.

We have still to show that we can find an $\hat{H}_R$ for (38). Whether or not this is the case is independent of the *lower order* terms $\hat{P}_R$ contains. A calculation similar to Lemma 2.3.5 in [21] shows that if $\hat{P}(\omega)$ admits an $\hat{H}_R$, then so will $\hat{P}(\omega) + B(\omega)$, where $B(\omega)$ is any matrix which satisfies $|B| + |B^*| \leqslant C$ for C independent of ω. In other words, the terms that are not multiplied by $i\omega$ can be dropped from (38), giving the principal symbol of the first order reduction

$$\hat{P}'_R = \begin{pmatrix} 0 & 0 & 0 \\ 0 & i\omega A^n & i\omega C \\ 0 & i\omega D^{nn} & i\omega G^n \end{pmatrix} \tag{43}$$

without affecting the well-posedness. The principal symbols of the second order system, Eq. (35), and the first order pseudo-differential reduction, Eq. (43), are related by

$$\hat{P}'_R = \begin{pmatrix} 0 & 0 \\ 0 & T\hat{P}'T^{-1} \end{pmatrix}, \quad T \equiv \begin{pmatrix} i\omega & 0 \\ 0 & 1 \end{pmatrix}. \tag{44}$$

(Note that T^{-1} does not exist for $\omega = 0$. However, in this case, $\hat{P}'_R = 0$, and admits the identity as a symmetrizer.) By assumption, there exists $\hat{H}(\omega) = \hat{H}^*(\omega)$ such that $\hat{\boldsymbol{u}}^* \hat{H} \hat{\boldsymbol{u}}$ is conserved by the principal system $\partial_t \hat{\boldsymbol{u}} = \hat{P}' \hat{\boldsymbol{u}}$ and satisfies (36). This $\hat{H}$ satisfies $\hat{H}\hat{P}' + \hat{P}'^* \hat{H} = 0$, and it is straightforward to show that

$$\hat{H}_R \equiv \begin{pmatrix} 1 & 0 \\ 0 & T^{-1*}\hat{H}T^{-1} \end{pmatrix} \tag{45}$$

satisfies $\hat{H}_R = \hat{H}_R^*$ and $\hat{H}_R \hat{P}'_R + \hat{P}'^*_R \hat{H}_R = 0$. Further, by noting that $T^* T = I_\omega$, using (36) one can show that $\hat{H}_R$ satisfies $K^{-1} I \leqslant \hat{H}_R \leqslant K I$. Hence we have found a symmetrizer of $\hat{P}'_R$ and the result has been proved.[7]

To construct $\hat{H}$ one can use the characteristic variables of $\hat{P}'$, as described at the end of Section 2.3. We would like to point out that this analysis did not require that the auxiliary constraint propagation problem be well-posed. These constraints are merely a tool for the analysis of the system. We only need to establish uniqueness of the solution with zero initial data for the auxiliary constraints. In the linear constant coefficient case this result is trivial. When evolving the second order system, these constraints are identically zero at all times. An alternative to the pseudo-differential reduction method is to perform a fully differential reduction by introducing a new variable in physical space for each derivative (see for example [7,11]).

3.2. Stability of discretizations of first order in time and second order in space systems

We now show how the continuum analysis of the previous subsection can be extended to the fully discrete case. The semi-discrete finite difference approximation of (33) can be written as

$$\frac{d}{dt}\boldsymbol{v} = P\boldsymbol{v}, \quad \boldsymbol{v} = \begin{pmatrix} u \\ v \end{pmatrix},$$
$$P = \begin{pmatrix} A^i D_i^{(1)} + B & C \\ D^{ij} D_{ij}^{(2)} + E^i D_i^{(1)} + F & G^i D_i^{(1)} + J \end{pmatrix}, \tag{46}$$

where $D_i^{(1)}$ is a discretization of the first derivative in the i direction and $D_{ij}^{(2)}$ is a discretization of the second derivative in the i and j directions. For example, the standard second order accurate discretization would have

[7] It can also be shown that $\hat{P}'_R$ is diagonalizable with the same eigenvalues as $\hat{P}'$, plus as many zeroes as there are components of u.

$$D_i^{(1)} = D_{0i}, \quad D_{ij}^{(2)} = \begin{cases} D_{0i}D_{0j}, & i \neq j, \\ D_{+i}D_{-i}, & i = j. \end{cases} \tag{47}$$

The principal symbol of the semi-discrete system is

$$\hat{P}' = \begin{pmatrix} A^i\hat{D}_i^{(1)} & C \\ D^{ij}\hat{D}_{ij}^{(2)} & G^i\hat{D}_i^{(1)} \end{pmatrix}, \tag{48}$$

where

$$\hat{D}_i^{(1)} = \frac{\mathrm{i}}{h}\sin\xi_i, \quad \hat{D}_{ij}^{(2)} = \begin{cases} -\frac{1}{h^2}\sin\xi_i\sin\xi_j, & i \neq j, \\ -\frac{4}{h^2}\sin^2\frac{\xi_i}{2}, & i = j \end{cases} \tag{49}$$

for the standard second order discretization. The *pseudo-discrete* first order reduction is obtained by defining

$$\hat{w} \equiv \mathrm{i}\Omega\hat{u}, \quad \Omega^2 = \sum_{i=1}^{d} |\hat{D}_{+i}|^2. \tag{50}$$

The reduced system is

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{v}}_{\mathrm{R}} = \hat{P}_{\mathrm{R}}\hat{\boldsymbol{v}}_{\mathrm{R}}, \quad \hat{\boldsymbol{v}}_{\mathrm{R}} = \begin{pmatrix} \hat{u} \\ \hat{w} \\ \hat{v} \end{pmatrix}, \tag{51}$$

$$\hat{P}_{\mathrm{R}} = \begin{pmatrix} B & (\mathrm{i}\Omega)^{-1}A^i\hat{D}_i^{(1)} & C \\ 0 & A^i\hat{D}_i^{(1)} + B & \mathrm{i}\Omega C \\ F & (\mathrm{i}\Omega)^{-1}(D^{ij}\hat{D}_{ij}^{(2)} + E^i\hat{D}_i^{(1)}) & G^i\hat{D}_i^{(1)} + J \end{pmatrix}. \tag{52}$$

We can show that the discrete auxiliary constraint is preserved by the time integrator. Define $c \equiv (-\mathrm{i}\Omega I\, I\, 0)$, so that the constraint is $c\hat{\boldsymbol{v}}_{\mathrm{R}} = \hat{w} - \mathrm{i}\Omega\hat{u} = 0$. Since $c\hat{P}_{\mathrm{R}}\hat{\boldsymbol{v}}_{\mathrm{R}} = Bc\hat{\boldsymbol{v}}_{\mathrm{R}}$, we have that $c\hat{\boldsymbol{v}}_{\mathrm{R}} = 0$ implies $c\hat{P}_{\mathrm{R}}\hat{\boldsymbol{v}}_{\mathrm{R}} = 0$ and hence $c\hat{P}_{\mathrm{R}}^n\hat{\boldsymbol{v}}_{\mathrm{R}} = 0$ and $c\mathscr{P}(k\hat{P})\hat{\boldsymbol{v}}_{\mathrm{R}} = 0$. Now consider evolving the reduced system with a polynomial time integrator; i.e. $\hat{\boldsymbol{v}}_{\mathrm{R}}^{n+1} = \mathscr{P}(k\hat{P}_{\mathrm{R}})\hat{\boldsymbol{v}}_{\mathrm{R}}^n$. If the auxiliary constraints are satisfied on one time step, then they are satisfied on the next as well, since $c\hat{\boldsymbol{v}}_{\mathrm{R}}^n = 0$ implies $c\hat{\boldsymbol{v}}_{\mathrm{R}}^{n+1} = c\mathscr{P}(k\hat{P})\hat{\boldsymbol{v}}_{\mathrm{R}}^n = 0$. Hence there is a one-to-one correspondence between solutions of the second order fully discrete system and those of the constraint-satisfying reduced system. Note that we have used the fact that the time integrator is a polynomial in $\hat{P}_{\mathrm{R}}$, as is the case for systems with constant coefficients. This result can be extended to the variable coefficient case, where one would have to perform the reduction to first order in physical space by introducing the gridfunctions $X^{(i)} = D_{+i}u$.

Making use of Theorem 5.1.2 of [20], the terms which correspond to the continuum lower order terms can be dropped from $\hat{P}_{\mathrm{R}}$ without affecting the stability of the fully discrete system, provided that $(\mathrm{i}\Omega)^{-1}\hat{D}_i^{(1)}$, $k\hat{D}_i^{(1)}$ and $k\Omega^{-1}\hat{D}_{ij}^{(2)}$ are bounded. This guarantees that the assumptions of the theorem are satisfied. This is true for the second and fourth order accurate standard discretizations.

The result for stability of the fully discrete problem is analogous to that for well-posedness at the continuum. If there exists $\hat{H}(\xi) = \hat{H}^*(\xi)$ such that the energy $\hat{\boldsymbol{v}}^*\hat{H}\hat{\boldsymbol{v}}$ is conserved by the semi-discrete principal system $\partial_t\hat{\boldsymbol{v}} = \hat{P}'\hat{\boldsymbol{v}}$ and $\hat{H}$ satisfies

$$K^{-1}I_\Omega \leqslant \hat{H} \leqslant KI_\Omega, \quad I_\Omega \equiv \begin{pmatrix} \Omega^2 & 0 \\ 0 & 1 \end{pmatrix}, \tag{53}$$

where K is a positive scalar constant, then it is possible to construct a discrete symmetrizer for the first order reduction with no lower order terms. So if, in addition, the principal symbol $\hat{P}'$ satisfies $\sigma(k\hat{P}') \leqslant \alpha_0$, the fully discrete system (including lower order terms) is stable with respect to the norm

$$\|\boldsymbol{v}\|_{h,D_+}^2 \equiv \|u\|_h^2 + \|v\|_h^2 + \sum_{i=1}^{d} \|D_{+i}u\|_h^2, \tag{54}$$

i.e. the solution satisfies the estimate

$$\|\boldsymbol{v}^n\|_{h,D_+} \leqslant K e^{\alpha t_n} \|\boldsymbol{v}^0\|_{h,D_+}. \tag{55}$$

Again, $\hat{H}$ can be constructed from the characteristic variables of $\hat{P}'$, as described at the end of Section 2.3. Note that the matrix $\hat{P}_R$ is not defined for $\Omega = 0$. However, this does not cause any difficulties in the linear constant coefficient case. One can write the space of solutions as a direct sum consisting of constant functions plus a space of solutions with non-trivial Ω, and treat each subspace independently.

3.3. Convergence

We briefly discuss convergence of the solution of the discrete problem to that of the continuum problem. We assume that (55) holds. Inserting the exact smooth solution $\boldsymbol{u}(t,x)$ into the scheme $\boldsymbol{v}^{n+1} = Q\boldsymbol{v}^n$ generates truncation errors as inhomogeneous terms in the difference approximation and in the initial data. The error grid-function $\boldsymbol{w}_j^n \equiv \boldsymbol{v}_j^n - \boldsymbol{u}(t_n, x_j)$ satisfies

$$\boldsymbol{w}_j^{n+1} = Q\boldsymbol{w}_j^n + \tilde{\boldsymbol{F}}_j^n, \tag{56}$$

$$\boldsymbol{w}_j^0 = \tilde{\boldsymbol{f}}_j, \tag{57}$$

where $\tilde{\boldsymbol{F}}_j^n = \boldsymbol{\phi}(t_n, x_j)\mathrm{O}(k^{p_1} + h^{p_2})$, and $\tilde{\boldsymbol{f}}_j = \boldsymbol{\psi}(x_j)\mathrm{O}(h^{p_3})$ with $\boldsymbol{\phi}$ smooth. The temporal accuracy of the scheme is p_1 and the spatial accuracy is p_2. The discrete version of Duhamel's principle (see Theorem 5.1.1 in [20]) gives the estimate

$$\|\boldsymbol{w}^n\|_{h,D_+} \leqslant K e^{\alpha t_n} \left(\|\boldsymbol{w}^0\|_{h,D_+} + k \sum_{r=0}^{n-1} \|\tilde{\boldsymbol{F}}^r\|_{h,D_+} \right) \leqslant \mathrm{O}(k^{p_1} + h^{p_2}), \tag{58}$$

provided that the initial data satisfies $\|\boldsymbol{w}^0\|_{h,D_+} \leqslant \mathrm{O}(h^{p_2})$. If $\boldsymbol{\psi}$ is smooth this condition is satisfied and, in particular, it is satisfied for exact initial data.

Inequality (58) implies convergence with respect to the discrete L_2 norm, $\|\boldsymbol{w}\|_h \leqslant \|\boldsymbol{w}\|_{h,D_+}$, despite the scheme being unstable with respect to this norm. Note that pth order convergence is obtained, with $p = \min(p_1, p_2)$ assuming $k = \lambda h$, even though the norm contains first order accurate one-sided difference operators.

4. Applications

In the following subsections, we apply the theoretical tools discussed in Section 3 to different systems. We start with a first order strongly hyperbolic system with no lower order terms. We then investigate three second order in space systems: the wave equation, a generalization of the KWB formulation of Maxwell's equations and the NOR formulation of Einstein's equations. We show that the clear correspondence between strong hyperbolicity and the existence of a discrete symmetrizer which occurs in first order systems with no lower order terms is lost when the standard discretization is used for second order in space systems. Similarly, the simple correspondence between characteristic speeds and the von Neumann condition, Eq. (63), does not hold for second order in space systems. It is convenient to define the following quantities,

$$\chi_q^2 = \sum_{i=1}^{d} \sin^q \frac{\xi_i}{2}, \quad \chi^2 = \sum_{i=1}^{d} \sin^2 \xi_i, \quad \Omega = \frac{2\chi_2}{h}. \tag{59}$$

Note that the maximum of χ_q and χ is $\sqrt{d}$. We also recall that when the eigenvalues of $\hat{P}$ are imaginary,

$$\sigma(k\hat{P}) \leqslant \alpha_0 \quad \Longleftrightarrow \quad \sigma(\hat{Q}) \leqslant 1, \tag{60}$$

where $\alpha_0 = 2$ for ICN, $\sqrt{8}$ for 4RK and $\sqrt{3}$ for 3RK.

4.1. Stability of first order strongly hyperbolic systems

Our first application is a constant coefficient first order system in d spatial dimensions

$$\frac{\partial u}{\partial t} = \sum_{i=1}^{d} A^i \frac{\partial u}{\partial x^i}, \tag{61}$$

where u is a vector valued function of the space–time coordinates. We assume that the system is strongly hyperbolic and that it admits a symmetrizer, i.e., there exists a matrix $\hat{H}(\omega)$ in Fourier space, such that $\hat{H}(\omega)\hat{P}(\mathrm{i}\omega) + \hat{P}^*(\mathrm{i}\omega)\hat{H}(\omega) = 0$, where $\hat{P}(\mathrm{i}\omega) = i\sum_{i=1}^{d}\omega_i A^i$. The discrete symbol associated with the standard second order accurate discretization of this system is

$$\hat{P}_h(\xi) = \frac{\mathrm{i}}{h}\sum_{i=1}^{d} A^i \sin\xi_i = \hat{P}(\mathrm{i}h^{-1}\sin\xi),$$

where we attached the subscript h to the discrete symbol to distinguish it from that of the continuum. We now construct the discrete symmetrizer

$$\hat{H}_h(\xi) \equiv \hat{H}(h^{-1}\sin\xi). \tag{62}$$

Conditions (29) and (30) are satisfied and condition (28) is sufficient for stability. The latter becomes $\sigma(k\hat{P}) = \lambda\chi\sigma(A(n)) \leqslant \alpha_0$, where $A(n) = \sum_{i=1}^{d} n_i A^i$, $n_i = \chi^{-1}\sin\xi_i$, so that $\sum_{i=1}^{d} n_i^2 = 1$. Since this inequality must hold for all ξ_i, and the quantity χ reaches its maximum value $\sqrt{d}$ at $\xi_i = \pm\pi/2$, we obtain the stability condition

$$\lambda \leqslant \frac{\alpha_0}{\sigma(A(n))\sqrt{d}}. \tag{63}$$

In the symmetrizable hyperbolic case one can take the discrete symmetrizer to be the same as that of the continuum (which, by definition, is independent of ω)

$$\hat{H}_h(\xi) = H. \tag{64}$$

This analysis of first order strongly hyperbolic systems shows that if the characteristic speeds depend neither on the direction nor on the dimensionality of the problem, i.e., if $\sigma(A(n))$ depends neither on n nor on d, then the Courant limit has a $1/\sqrt{d}$ dependence. In addition, when the second order accurate centered difference operator D_0 is used to approximate the spatial derivatives, a Courant limit violation would manifest itself as a rapid growth of the mid high frequency mode $|\xi_i| = \frac{\pi}{2} \approx 1.571$. This mode is present if N is a multiple of 4. A similar analysis shows that in the fourth order accurate case the situation differs. The Courant limit is 1.372 times smaller than (63) and above this limit the most rapid growth occurs at a slightly higher frequency, $|\xi_i| = 2\arctan(6^{1/2}/(4 - 6^{1/2}))^{1/2} \approx 1.797$. See also Appendix B.

4.2. First order in time and second order in space wave equation

In this section, we discuss the stability properties of an approximation of the d-dimensional wave equation written as a first order in time and second order in space system

$$\partial_t\phi(t,x) = \Pi(t,x), \tag{65}$$

$$\partial_t\Pi(t,x) = \sum_{i=1}^{d}\partial_i^2\phi(t,x). \tag{66}$$

In Section 1 we pointed out that the Cauchy problem for this system is not well-posed in L_2. One can expect that a direct application of definition (12), which is based on the discrete L_2 norm, to a scheme approximating (65) and (66) would lead to the conclusion that the scheme is unstable. The first order reduction, however, is well-posed in L_2 (it is symmetric hyperbolic), hence the second order system satisfies an energy estimate with respect to

$$\|\boldsymbol{u}(\cdot,t)\|^2 = \int |\phi(x,t)|^2 + |\Pi(x,t)|^2 + \sum_{i=1}^{d}|\partial_i\phi(x,t)|^2\,\mathrm{d}^d x. \tag{67}$$

In this section, we show stability for the standard discretization of this system, both by the pseudo-discrete reduction method given in Section 3.2, and by a direct discrete reduction in physical space. The two methods give equivalent results.

Following the method of lines, we first discretize space and leave time continuous,

$$\frac{d}{dt}\phi_j(t) = \Pi_j(t), \tag{68}$$

$$\frac{d}{dt}\Pi_j(t) = \sum_{i=1}^{d} D_{+i}D_{-i}\phi_j(t). \tag{69}$$

Using the technique described in Section 3.2, we see that the (principal) symbol of the second order semi-discrete problem

$$\hat{P} = \begin{pmatrix} 0 & 1 \\ -\Omega^2 & 0 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} i\Omega & 1 \\ -i\Omega & 1 \end{pmatrix} \tag{70}$$

has purely imaginary eigenvalues $\pm i\Omega$. The matrix T diagonalizes $\hat{P}$. The sum of the squared moduli of the characteristic variables gives the conserved energy (here $D = 1/2I$)

$$\hat{\mathbf{v}}^*(T^{-1})^* D T^{-1}\hat{\mathbf{v}} \equiv |i\Omega\hat{\phi}|^2 + |\hat{\Pi}|^2 = \Omega^2|\hat{\phi}|^2 + |\hat{\Pi}|^2. \tag{71}$$

By taking $K = 1$ in (53) we see that we have numerical stability with respect to the discrete norm

$$\|\mathbf{v}\|^2_{h,D_+} = \sum_j \left(\phi_j^2 + \Pi_j^2 + \sum_{i=1}^{d}(D_{+i}\phi_j)^2\right)h^d, \tag{72}$$

provided that the von Neumann condition

$$\lambda \leqslant \alpha_0/(2\sqrt{d}), \tag{73}$$

which follows from $\sigma(k\hat{P}) = k\Omega = 2\lambda\chi_2 \leqslant \alpha_0$, is satisfied.

We now illustrate a different method for proving stability of this system. A *discrete reduction to first order* can be performed before going to Fourier space. We introduce the quantities

$$X_j^{(i)} = D_{+i}\phi_j \tag{74}$$

and obtain the reduced system

$$\frac{d}{dt}\phi_j(t) = \Pi_j(t), \tag{75}$$

$$\frac{d}{dt}\Pi_j(t) = \sum_{i=1}^{d} D_{-i}X_j^{(i)}(t), \tag{76}$$

$$\frac{d}{dt}X_j^{(i)}(t) = D_{+i}\Pi_j(t). \tag{77}$$

Notice that only if Eq. (74) is identically satisfied is the reduced system equivalent to the original one. It is important to check whether the evolution equations (75)–(77) are compatible with this requirement. Let $C_j^{(i)}(t) \equiv X_j^{(i)} - D_{+i}\phi_j$. If we prescribe initial data such that $C_j^{(i)}(0) = 0$, then at later times $C_j^{(i)}(t) = 0$. This is a consequence of the fact that

$$\frac{d}{dt}C_j^{(i)}(t) = \frac{d}{dt}\left(X_j^{(i)}(t) - D_{+i}\phi_j(t)\right) = 0. \tag{78}$$

There is a one-to-one correspondence between solutions of (68), (69) and those of (74)–(77). Furthermore, one can check that the time integrator does not spoil the propagation of the constraints.

Ignoring lower order terms, the symbol associated with the reduced system (75)–(77) is anti-Hermitian, therefore Eq. (30) is satisfied with $\hat{H} = 1$. The non-trivial eigenvalues of $\hat{P}$ are $\pm i\Omega$, the same as those of the original system (68) and (69). This proves that (73) is a necessary and sufficient condition for stability with respect to the discrete norm (72).

This specific discrete reduction to first order, and the pseudo-discrete reduction to first order described in Section 3.2 give equivalent results.

4.2.1. Fourth order accuracy

In hyperbolic problems a fourth order accurate spatial discretization requires significantly fewer grid-points per wavelength for a given tolerance error (see [20] and Appendix B). The stability proof for the fourth order accurate discretization of the d-dimensional wave equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_j(t) = \Pi_j(t), \tag{79}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Pi_j(t) = \sum_{i=1}^{d} D_{+i}D_{-i}\left(1 - \frac{h^2}{12}D_{+i}D_{-i}\right)\phi_j(t) \tag{80}$$

is similar to the second order accurate case. The discrete symbol and diagonalizing matrix are

$$\hat{P} = \begin{pmatrix} 0 & 1 \\ -\Delta^2 & 0 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} \mathrm{i}\Delta & 1 \\ -\mathrm{i}\Delta & 1 \end{pmatrix}, \tag{81}$$

where $\Delta^2 = \frac{4}{h^2}\sum_{i=1}^{d}\sin^2\frac{\xi_i}{2}(1+\frac{1}{3}\sin^2\frac{\xi_i}{2})$ has purely imaginary eigenvalues $\pm \mathrm{i}\Delta$. Taking $D = 1/2I$ we get the conserved quantity

$$(T^{-1}\hat{\boldsymbol{v}})^* D\hat{T}^{-1}\hat{\boldsymbol{v}} = \Delta^2|\hat{\phi}|^2 + |\hat{\Pi}|^2. \tag{82}$$

Since $\Omega^2 \leqslant \Delta^2 \leqslant \frac{4}{3}\Omega^2$, by taking $K = 4/3$ in (53) we see that we have numerical stability with respect to the norm (72) provided that the principal symbol $\hat{P}$ satisfies $\sigma(k\hat{P}) \leqslant \alpha_0$. This gives a stability limit of $\lambda \leqslant \sqrt{3}\alpha_0/(4\sqrt{d})$.

4.2.2. A note about the D_0-norm and the D_0^2 discretization

Replacing the one sided difference operators D_{+i} with centered difference operators D_{0i} in the norm (72) leads to difficulties, as the D_0-norm does not capture the highest frequency mode. In fact, it is possible to construct a family of solutions of (68) and (69) proportional to $(-1)^j$ for which the D_0-energy estimate fails. For this purpose it is sufficient to consider $\phi_j(t) = (-1)^j\cos(2t/h)$, $\Pi_j(t) = -2/h(-1)^j\sin(2t/h)$, which gives

$$\frac{\|\boldsymbol{v}(t)\|_{h,D_0}}{\|\boldsymbol{v}(0)\|_{h,D_0}} = \left(\cos^2\frac{2t}{h} + \frac{4}{h^2}\sin^2\frac{2t}{h}\right)^{1/2}, \tag{83}$$

where $\|\boldsymbol{v}(t)\|^2_{h,D_0} = \sum_j(\phi_j^2 + \Pi_j^2 + (D_0\phi_j)^2)h$. It it not possible to find constants K and α such that the ratio is bounded by $Ke^{\alpha t}$, independently of the space step h.

It has been suggested that the use of D_0^2 rather than D_+D_- for the second spatial derivatives may improve the stability properties of a second order in space scheme [24,25]. To investigate this we study the wave equation in one space dimension discretized as

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_j(t) = \Pi_j(t), \tag{84}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Pi_j(t) = D_0^2\phi_j(t). \tag{85}$$

The eigenvalues of $k\hat{P}$ are $\pm \mathrm{i}\lambda\sin\xi$, which shows that the von Neumann condition is satisfied as long as $\lambda \leqslant \alpha_0$. Both the stencil and the maximum time step compatible with the von Neumann condition are twice what they are for the D_+D_- discretization. However, for a given spatial resolution the numerical speed of propagation has an error which is four times that of the D_+D_- case (see Appendix B).

So far, we have only shown that the scheme is unstable if $\lambda > \alpha_0$. By looking at the discrete symbol

$$\hat{P}(\xi) = \begin{pmatrix} 0 & 1 \\ -\frac{1}{h^2}\sin^2\xi & 0 \end{pmatrix} \tag{86}$$

we see that there might be a problem for $|\xi| = \pi$. In this case the symbol is not diagonalizable. To explicitly show that the system (84) and (85) is unstable with respect to the norm

$$\|\boldsymbol{v}\|^2_{h,D_+} = \sum_j \left(\phi_j^2 + \Pi_j^2 + (D_+\phi_j)^2\right)h \tag{87}$$

it is sufficient to consider the family of initial data $\phi_j(0) = 0$, $\Pi_j(0) = (-1)^j$, generating the solution $\phi_j(t) = (-1)^j t$, $\Pi_j(t) = (-1)^j$. As $h \to 0$ the ratio

$$\frac{\|\boldsymbol{v}(t)\|_{h,D_+}}{\|\boldsymbol{v}(0)\|_{h,D_+}} = \left(1 + t^2 + \frac{4t^2}{h^2}\right)^{1/2} \tag{88}$$

grows without bound.

Had we chosen the D_0-norm, however, we would have concluded that the scheme satisfies the required estimate. This is because this norm does not capture the highest frequency mode $\phi_j = (-1)^j$. A desirable requirement of a norm is that if a scheme is stable with respect to that norm, then it will remain stable with respect to the same norm when perturbed with lower order terms (independently of how these are discretized). The modified problem

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_j(t) = \Pi_j(t), \tag{89}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Pi_j(t) = D_0^2\phi_j(t) - D_+\phi_j(t) \tag{90}$$

admits the family of exponentially growing solutions $\phi_j(t) = (-1)^j \exp(\sqrt{2/h}t)$, $\Pi_j(t) = (-1)^j\sqrt{2/h}\exp(\sqrt{2/h}t)$ which leads to unbounded growth in the ratio

$$\frac{\|\boldsymbol{v}(t)\|_{h,D_0}}{\|\boldsymbol{v}(0)\|_{h,D_0}} = \exp\left(\sqrt{\frac{2}{h}}t\right). \tag{91}$$

If we want to be able to decide whether a scheme is stable or not just by looking at the principal part of the discrete system, then we must conclude that the D_0-energy is not a suitable energy.

We note that the requirement that stability should not depend on how lower order terms are discretized was crucial. If we restrict ourselves to the perturbation $D_0\phi_j$, then the scheme is still stable with respect to the D_0-energy. If one wants to be able to discretize lower order terms freely, as we do, then one is forced to reject the D_0^2 discretization.

Clearly it is the presence of high frequency modes that makes the D_0^2 discretization unstable with respect to the D_+-norm. The introduction of a mechanism that damps high frequency modes, such as artificial dissipation, may restore stability. In the system

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_j = \Pi_j - \sigma h^3(D_+D_-)^2\phi_j,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Pi_j = D_0^2\phi_j - \sigma h^3(D_+D_-)^2\Pi_j$$

the same family of initial data used to prove instability of (84) and (85) gives $\|\boldsymbol{v}(t)\|_{h,D_+}/\|\boldsymbol{v}(0)\|_{h,D_+} = (1 + t^2 + 4t^2/h^2)^{1/2}\mathrm{e}^{-16\sigma t/h}$, which does not grow without bound.

4.3. The generalized Knapp–Walker–Baumgarte system

We now investigate more complex systems. We adopt the Einstein summation convention. We consider the KWB formulation of Maxwell's equations [16]

$$\partial_t A_i = -E_i, \tag{92}$$

$$\partial_t E_i = -\partial^k\partial_k A_i + \partial_i\Gamma, \tag{93}$$

$$\partial_t \Gamma = 0, \tag{94}$$

and generalize it by introducing $G = \Gamma - r\partial^k A_k$, giving

$$\partial_t A_i = -E_i, \tag{95}$$
$$\partial_t E_i = -\partial^k\partial_k A_i + r\partial_i\partial^k A_k + \partial_i G, \tag{96}$$
$$\partial_t G = r\partial^k E_k. \tag{97}$$

For $r = 0$ we recover (92)–(94) and for $r = 1$ we obtain the Z1 system [17], which was recently introduced as a toy model for the Z4 formulation of general relativity (see Section 4.6). We will show that although the parameter r plays no role at the continuum, at the discrete level it can have a severe impact on the stability properties.

4.3.1. Continuum analysis

If we Fourier transform (95)–(97) and introduce $\hat{\Gamma} = \hat{G} + ri\omega_k\hat{A}_k$ in place of $\hat{G}$ the system simplifies to

$$\partial_t\hat{A}_i = -\hat{E}_i,$$
$$\partial_t\hat{E}_i = \omega^2\hat{A}_i + i\omega_i\hat{\Gamma},$$
$$\partial_t\hat{\Gamma} = 0.$$

The eigenvalues and characteristic variables of the symbol are

$$0, \quad \hat{w}^{(0)} = \hat{\Gamma},$$
$$\pm i\omega, \quad \hat{w}_i^{(\pm)} = \hat{E}_i \mp i\omega\hat{A}_i \pm \hat{\omega}_i\hat{\Gamma},$$

where $\hat{\omega}_i = \omega_i/\omega$ and $\omega^2 = \sum_{k=1}^3 \omega_k^2$. Note that the eigenvalues of the symbol are independent of the parameter r. To construct a conserved energy we take the combination

$$E_C = \frac{1}{2}|\hat{w}_i^{(+)}|^2 + \frac{1}{2}|\hat{w}_i^{(-)}|^2 + a|\hat{w}^{(0)}|^2.$$

To keep the notation compact we omit the sums. We need to check that this conserved quantity is equivalent to[8]

$$|\hat{\boldsymbol{u}}|^2 = |\hat{E}_i|^2 + \omega^2|\hat{A}_i|^2 + |\hat{G}|^2.$$

Since

$$E_C = |\hat{E}_i|^2 + (1+a)|\hat{\Gamma}|^2 + \omega^2|\hat{A}_i|^2 - 2\mathrm{Re}\left(i\omega_i\hat{A}_i\overline{\hat{\Gamma}}\right),$$

we get

$$|\hat{E}_i|^2 + (1+a-\varepsilon_1)|\hat{\Gamma}|^2 + \left(1 - \frac{1}{\varepsilon_1}\right)\omega^2|\hat{A}_i|^2 \leqslant E_C \leqslant |\hat{E}_i|^2 + (1+a+\varepsilon_2)|\hat{\Gamma}|^2 + \left(1 + \frac{1}{\varepsilon_2}\right)\omega^2|\hat{A}_i|^2,$$

where we used the inequality $\pm 2\mathrm{Re}(z_1\bar{z}_2) \leqslant \varepsilon|z_1|^2 + \varepsilon^{-1}|z_2|^2$ for $\varepsilon > 0$. Choosing $a = 3/2$, $\varepsilon_1 = \varepsilon_2^{-1} = 2$ gives

$$K_1^{-1}|\hat{\boldsymbol{u}}|_\Gamma^2 \leqslant E_C \leqslant K_1|\hat{\boldsymbol{u}}|_\Gamma^2,$$

with $K_1 = 3$, where $|\hat{\boldsymbol{u}}|_\Gamma^2 = |\hat{E}_i|^2 + \omega^2|\hat{A}_i|^2 + |\hat{\Gamma}|^2$. Using the inequality

$$(1-\varepsilon)|z_1|^2 + (1-\varepsilon^{-1})|z_2|^2 \leqslant |z_1 + z_2|^2 \leqslant (1+\varepsilon)|z_1|^2 + (1+\varepsilon^{-1})|z_2|^2, \tag{98}$$

with $\varepsilon > 0$, we have that for any r, $|\hat{\boldsymbol{u}}|_\Gamma^2$ is equivalent to $|\hat{\boldsymbol{u}}|^2$, i.e. $K_2^{-1}|\hat{\boldsymbol{u}}|_\Gamma^2 \leqslant |\hat{\boldsymbol{u}}|^2 \leqslant K_2|\hat{\boldsymbol{u}}|_\Gamma^2$. We have the uniform estimate in Fourier space

$$|\hat{\boldsymbol{u}}(t)|^2 \leqslant K_2|\hat{\boldsymbol{u}}(t)|_\Gamma^2 \leqslant K_1K_2E_C(t) = K_1K_2E_C(0) \leqslant K_1^2K_2|\hat{\boldsymbol{u}}(0)|_\Gamma^2 \leqslant K_1^2K_2^2|\hat{\boldsymbol{u}}(0)|^2, \tag{99}$$

which implies the estimate in physical space with respect to the norm

[8] From the results in Section 3 we only need to show that $\hat{H}$ is equivalent to I_ω, see inequality (36), which in this case means that there is no $|\hat{A}_i|^2$ term.

$$\|\boldsymbol{u}\|^2 = \|A_i\|^2 + \|E_i\|^2 + \|\partial_k A_i\|^2 + \|G\|^2, \tag{100}$$

with no restrictions on the parameter r.

4.3.2. Discrete analysis

Consider now the semi-discrete system

$$\partial_t A_i = -E_i, \tag{101}$$

$$\partial_t E_i = -D_{+k}D_{-k}A_i + rD_{ik}^{(2)}A_k + D_{0i}G, \tag{102}$$

$$\partial_t G = rD_{0k}E_k, \tag{103}$$

where $D_{ik}^{(2)}$ is the standard second order accurate approximation of the second partial derivative. The procedure is similar to that at the continuum. We Fourier transform and replace the variable $\hat{G}$ with $\hat{\Gamma} = \hat{G} + r\frac{\mathrm{i}}{h}\sin\xi_k\hat{A}_k$ and obtain

$$\partial_t\hat{A}_i = -\hat{E}_i,$$

$$\partial_t\hat{E}_i = \frac{4}{h^2}\Theta_i^2(\xi)\hat{A}_i + \frac{i}{h}\sin\xi_i\hat{\Gamma},$$

$$\partial_t\hat{\Gamma} = 0,$$

where $\Theta_i^2(\xi) = \sum_{k=1}^3 \sin^2\frac{\xi_k}{2} - r\sin^4\frac{\xi_i}{2}$.

The eigenvalues of the matrix $k\hat{P}(\xi)$ and the corresponding characteristic variables are

$$0, \quad \hat{w}^{(0)} = \hat{\Gamma},$$

$$\pm 2i\Theta_i(\xi)\lambda, \quad \hat{w}_i^{(\pm)} = \hat{E}_i \mp \frac{2\mathrm{i}}{h}\Theta_i(\xi)\hat{A}_i \pm s_i(\xi)\hat{\Gamma},$$

where $2s_i\Theta_i = \sin, \xi_i$. The requirement that $\sigma(k\hat{P}) \leqslant \alpha_0$ imposes the restriction $r \leqslant 1$ on the parameter. If this condition is violated, then the semi-discrete scheme is unstable (and the fully discrete scheme would be unconditionally unstable). Furthermore, for $r = 1$, which corresponds to the Z1 system, the matrix $\hat{P}(\pm\pi, 0, 0)$ (corresponding to the highest frequency in the x direction) is not diagonalizable and one can show that the system admits frequency dependent linearly growing solutions which violate the discrete energy estimate.

Assume $r < 1$. The expression

$$E_C = \frac{1}{2}|\hat{w}_i^{(+)}|^2 + \frac{1}{2}|\hat{w}_i^{(-)}|^2 + a|\hat{\Gamma}|^2 = |\hat{E}_i|^2 + (a + s_i^2)|\hat{\Gamma}|^2 + \frac{4}{h^2}\Theta_i^2|\hat{A}_i|^2 - 2\mathrm{Re}\left(\frac{i}{h}\sin\xi_i\hat{A}_i\overline{\hat{\Gamma}}\right)$$

is conserved. We want to show that it is equivalent to $|\hat{\boldsymbol{u}}|^2 = |\hat{E}_i|^2 + \Omega^2|\hat{A}_i|^2 + |\hat{G}|^2$.

We first show that E_C is equivalent to $|\hat{\boldsymbol{u}}|_\Gamma^2 = |\hat{E}_i|^2 + \Omega^2|\hat{A}_i|^2 + |\hat{\Gamma}|^2$. We distinguish now between two possibilities: $r \leqslant 0$ and $0 < r < 1$. In either case we have that $|s_i| \leqslant 1$. In the first case, using the inequality $\chi_2^2 \leqslant \Theta_i^2 \leqslant (1 - r)\chi_2^2$ we get

$$|\hat{E}_i|^2 + (a - \varepsilon_1)|\hat{\Gamma}|^2 + \left(1 - \frac{1}{\varepsilon_1}\right)\chi_2^2|\hat{A}_i|^2 \leqslant E_C \leqslant |\hat{E}_i|^2 + (a + 1 + \varepsilon_2)|\hat{\Gamma}|^2 + \frac{4}{h^2}\left(1 - r + \frac{1}{\varepsilon_2}\right)\chi_2^2|\hat{A}_i|^2.$$

If we take, for example, $a \geqslant 3$, $\varepsilon_1 = 2$, $\varepsilon_2 = 1/2$, then there exist constants K_1 and K_2 such that $K_1|\hat{\boldsymbol{u}}|_\Gamma^2 \leqslant E_C \leqslant K_2|\hat{\boldsymbol{u}}|_\Gamma^2$.

For the case $0 < r < 1$, using the inequality $(1 - r)\chi_2^2 \leqslant \Theta_i^2 \leqslant \chi_2^2$ we get

$$|\hat{E}_i|^2 + (a - \varepsilon_1)|\hat{\Gamma}|^2 + \left(1 - r - \frac{1}{\varepsilon_1}\right)\chi_2^2|\hat{A}_i|^2 \leqslant E_C \leqslant |\hat{E}_i|^2 + (a + 1 + \varepsilon_2)|\hat{\Gamma}|^2 + \frac{4}{h^2}\left(1 + \frac{1}{\varepsilon_2}\right)\chi_2^2|\hat{A}_i|^2.$$

If we choose $a > \varepsilon_1 > 1/(1 - r)$ we have the equivalence to $|\hat{\boldsymbol{u}}|_\Gamma^2$. On the other hand, using

$$\frac{1}{h}|\sin\xi_k| \leqslant |\Omega|, \tag{104}$$

one can show that the norms $|\hat{\boldsymbol{u}}|_\Gamma^2$ and $|\hat{\boldsymbol{u}}|^2$ are equivalent. This proves stability with respect to the norm

$$(\|A_i\|_h^2 + \|E_i\|_h^2 + \|D_{+k}A_i\|_h^2 + \|G\|_h^2)^{1/2}. \tag{105}$$

Note that the Cauchy problem for the continuum system is well-posed for all values of r, but the discrete system is stable only for $r < 1$. For $r \leqslant 1/2$ the von Neumann condition gives a Courant limit of $\lambda \leqslant \alpha_0/(2\sqrt{3-r})$. Moreover, the numerical speeds of propagation depend on r.

4.4. The Nagy–Ortiz–Reula system

The NOR formulation of Einstein's equations linearized about Minkowski space with zero shift and densitized lapse ($\alpha = \det(\gamma_{ij})^{1/2}$) has the form

$$\partial_t\gamma_{ij} = -2K_{ij}, \tag{106}$$

$$\partial_t K_{ij} = -\frac{1}{2}\partial^k\partial_k\gamma_{ij} + \frac{r}{2}\partial_i\partial_j\tau + \partial_{(i}f_{j)}, \tag{107}$$

$$\partial_t f_i = r\partial_i K, \tag{108}$$

where $\tau = \delta^{kl}\gamma_{kl}$. This system corresponds to the one in [10] with the choice of parameters $a = b = \sigma = 1$, $c = 0$ and $\rho = r + 2$. It is obtained from the ADM system with densitized lapse by introducing the variables $f_i = \partial_j\gamma_{ij} - \partial_i\tau$, which are used in the evolution equations for the K_{ij} variables, and adding the momentum constraint to the time derivative of the new variables.

4.4.1. Continuum analysis

We Fourier transform the system and introduce $\hat{\Gamma}_i = \hat{f}_i + \frac{r}{2}\mathrm{i}\omega_i\hat{\tau}$, obtaining

$$\partial_t\hat{\gamma}_{ij} = -2\hat{K}_{ij},$$
$$\partial_t\hat{K}_{ij} = \frac{1}{2}\omega^2\hat{\gamma}_{ij} + i\omega_{(i}\hat{\Gamma}_{j)},$$
$$\partial_t\hat{\Gamma}_i = 0.$$

The eigenvalues and characteristic variables associated with the symbol are

$$0, \quad \hat{w}_i^{(0)} = \hat{\Gamma}_i,$$
$$\pm\,\mathrm{i}\omega, \quad \hat{w}_{ij}^{(\pm)} = \hat{K}_{ij} \mp \frac{1}{2}\mathrm{i}\omega\hat{\gamma}_{ij} \pm \hat{\omega}_{(i}\hat{\Gamma}_{j)}.$$

Proceeding in the usual manner we construct a conserved quantity and show that it is equivalent to

$$|\hat{\boldsymbol{u}}|^2 = |\hat{K}_{ij}|^2 + \omega^2|\hat{\gamma}_{ij}|^2 + |\hat{f}_i|^2.$$

We have

$$E_C = \frac{1}{2}|\hat{w}_{ij}^{(+)}|^2 + \frac{1}{2}|\hat{w}_{ij}^{(-)}|^2 + a|\hat{w}_i^{(0)}|^2 = |\hat{K}_{ij}|^2 + |\hat{\omega}_{(i}\hat{\Gamma}_{j)}|^2 + \frac{1}{4}\omega^2|\hat{\gamma}_{ij}|^2 - \mathrm{Re}\left(\mathrm{i}\omega_i\hat{\gamma}_{ij}\overline{\hat{\Gamma}_j}\right) + a|\hat{\Gamma}_j|^2.$$

Since

$$0 \leqslant |\hat{\omega}_{(i}\hat{\Gamma}_{j)}|^2 \leqslant |\hat{\omega}_i\hat{\Gamma}_j|^2 \leqslant |\hat{\Gamma}_i|^2 - \frac{\omega^2}{\varepsilon_1}|\hat{\gamma}_{ij}|^2 - \varepsilon_1|\hat{\Gamma}_i|^2 \leqslant -2\mathrm{Re}\left(\mathrm{i}\omega_i\hat{\gamma}_{ij}\overline{\hat{\Gamma}_j}\right) \leqslant \frac{\omega^2}{\varepsilon_2}|\hat{\gamma}_{ij}|^2 + \varepsilon_2|\hat{\Gamma}_i|^2,$$

we obtain the equivalence with $|\hat{\boldsymbol{u}}|_\Gamma^2$,

$$|\hat{K}_{ij}|^2 + \frac{1}{4}\left(1 - \frac{1}{\varepsilon_1}\right)\omega^2|\hat{\gamma}_{ij}|^2 + (a - \varepsilon_1)|\hat{\Gamma}_i|^2 \leqslant E_C \leqslant |\hat{K}_{ij}|^2 + \frac{1}{4}\left(1 + \frac{1}{\varepsilon_2}\right)\omega^2|\hat{\gamma}_{ij}|^2 + (1 + a + \varepsilon_2)|\hat{\Gamma}_i|^2,$$

by choosing $a = 3$, $\varepsilon_1 = 2$, $\varepsilon_2 = 1$. Finally, noting that $|\hat{\tau}|^2 \leqslant 3|\hat{\gamma}_{ij}|^2$ one can show that $|\hat{\boldsymbol{u}}|_\Gamma^2$ and $|\hat{\boldsymbol{u}}|^2$ are equivalent.

4.4.2. Discrete analysis

We consider the standard second order accurate discretization of system (106)–(108). The semi-discrete system is

$$\partial_t \gamma_{ij} = -2K_{ij}, \tag{109}$$

$$\partial_t K_{ij} = -\frac{1}{2} D_{+k} D_{-k} \gamma_{ij} + \frac{r}{2} D^{(2)}_{ij} \tau + D_{0(i} f_{j)}, \tag{110}$$

$$\partial_t f_i = r D_{0i} K. \tag{111}$$

Taking the Fourier transform and introducing $\hat{\Gamma}_i = \hat{f}_i + \frac{r\mathrm{i}}{2h} \sin \xi_i \hat{\tau}$ gives

$$\partial_t \hat{\gamma}_{ij} = -2\hat{K}_{ij},$$
$$\partial_t \hat{K}_{ij} = \frac{1}{2} \Omega^2 \hat{\gamma}_{ij} + \frac{r}{2} \hat{\Delta}_{ij} \hat{\tau} + \frac{\mathrm{i}}{h} \sin \xi_{(i} \hat{\Gamma}_{j)},$$
$$\partial_t \hat{\Gamma}_i = 0,$$

where

$$\hat{\Delta}_{ij} = \begin{cases} 0, & i \neq j, \\ -\frac{4}{h^2} \sin^4 \frac{\xi_i}{2} & i = j. \end{cases}$$

The eigenvalues of $k\hat{P}$ and the corresponding characteristic variables are

$$0, \quad \hat{w}^{(0)}_i = \hat{\Gamma}_i,$$

$$\pm 2\mathrm{i}\Theta\lambda, \quad \hat{w}^{(\pm)} = \hat{K} \mp \frac{\mathrm{i}}{h} \Theta \hat{\tau} \pm \frac{\sin \xi_i}{2\Theta} \hat{\Gamma}_i,$$

$$\pm 2\mathrm{i}\chi_2\lambda, \quad \hat{w}^{(\pm)}_{ij} = \hat{K}_{ij} \mp \frac{1}{2} \mathrm{i}\Omega \hat{\gamma}_{ij} \pm \frac{\sin \xi_{(i} \hat{\Gamma}_{j)}}{2\chi_2}, \quad i \neq j,$$

$$\hat{w}^{(\pm)}_i = \left(\widetilde{K}_{ii} \mp \frac{1}{2} \mathrm{i}\Omega \widetilde{\gamma}_{ii} \pm \frac{\sin \xi_i \widetilde{\Gamma}_i}{2\chi_2} \right)^{\mathrm{TF}},$$

where $\Theta^2 = \chi_2^2 - r\sum_{k=1}^3 \sigma_i^4$, $\sigma_i^4 = \sin^4 \frac{\xi_i}{2}$, $\sigma_i^4 \widetilde{K}_{ii} = \hat{K}_{ii}$, $\sigma_i^4 \widetilde{\gamma}_{ii} = \hat{\gamma}_{ii}$, $\sigma_i^4 \widetilde{\Gamma}_i = \hat{\Gamma}_i$, and $A^{\mathrm{TF}}_{ij} = (A_{ij} - \delta_{ij} A/3)$.

Note that stability demands that $r < 1$ ($\rho < 3$). Furthermore, the von Neumann condition depends on the value of this parameter. Explicitly, this is

$$\lambda \leqslant \frac{\alpha_0}{2 \max\limits_{|\xi_i| \leqslant \pi} \{\Theta, \chi_2\}}$$

and its dependence on r is illustrated in Fig. 1. This is in contrast to the fact that at the continuum r has no influence on the characteristic speeds or the hyperbolicity of the system.

We now restrict ourselves to the case $r = 0$ and prove numerical stability. In this case the characteristic variables associated with the non-trivial eigenvalues are

$$\hat{w}^{(\pm)}_{ij} = \hat{K}_{ij} \mp \frac{1}{2} \mathrm{i}\Omega \hat{\gamma}_{ij} \pm \frac{\sin \xi_{(i} \hat{\Gamma}_{j)}}{2\chi_2}. \tag{112}$$

A conserved quantity is

$$E_C = \frac{1}{2} |\hat{w}^{(+)}_{ij}|^2 + \frac{1}{2} |\hat{w}^{(-)}_{ij}|^2 + a|\hat{w}^{(0)}_i|^2 = |\hat{K}_{ij}|^2 + |s_{(i} \hat{\Gamma}_{j)}|^2 + \frac{\Omega^2}{4} |\hat{\gamma}_{ij}|^2 - \mathrm{Re}\left(\frac{\mathrm{i}}{h} \sin \xi_i \hat{\gamma}_{ij} \overline{\hat{\Gamma}_j} \right) + a|\hat{\Gamma}_i|^2,$$

where $2\chi_2 s_i = \sin \xi_i$.

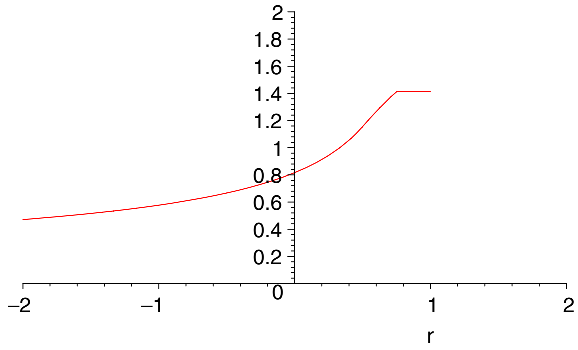

Fig. 1. The von Neumann condition for the second order accurate discretization of the NOR system in 3D using 4RK as a function of the parameter r. For $r > 1$ the scheme is unconditionally unstable.

Since

$$|s_i| \leqslant 1,$$
$$0 \leqslant |s_{(i}\hat{\Gamma}_{j)}|^2 \leqslant |s_i\hat{\Gamma}_j|^2 \leqslant |\hat{\Gamma}_i|^2 - \frac{4}{\varepsilon_1 h^2}\chi_2^2|\hat{\gamma}_{ij}|^2 - \varepsilon_1|\hat{\Gamma}_i|^2$$
$$\leqslant -2\mathrm{Re}\left(\frac{i}{h}\sin\xi_i\hat{\gamma}_{ij}\overline{\hat{\Gamma}_j}\right) \leqslant \frac{4}{\varepsilon_2 h^2}\chi_2^2|\hat{\gamma}_{ij}|^2 + \varepsilon_2|\hat{\Gamma}_i|^2,$$

we have the equivalence with $|\hat{\boldsymbol{u}}|^2_{\Gamma}$. Inequality (104) guarantees the equivalence of the latter with $|\hat{\boldsymbol{u}}|^2$. This completes the proof of stability with respect to the norm

$$\left(\|\gamma_{ij}\|_h^2 + \|K_{ij}\|_h^2 + \|D_{+k}\gamma_{ij}\|_h^2 + \|f_i\|_h^2\right)^{1/2}. \tag{113}$$

4.5. The ADM system

With a densitized lapse function, $\alpha = \det(\gamma_{ij})^{1/2}$, the ADM equations linearized about the Minkowski solution in Cartesian coordinates take the form

$$\partial_t\gamma_{ij} = -2K_{ij}, \tag{114}$$
$$\partial_t K_{ij} = \partial_k\partial_{(i}\gamma_{j)k} - \frac{1}{2}\partial^k\partial_k\gamma_{ij} - \partial_i\partial_j\tau. \tag{115}$$

The symbol $\hat{P}(\mathrm{i}\omega)$ of (114) and (115) is not diagonalizable and neither is that of its differential nor its pseudo-differential reduction. The family of solutions in which the only non-vanishing components are $\gamma_{1A} = \sin(\omega x)t$, $K_{1A} = -\sin(\omega x)/2$, where $A = 2, 3$, can be used to explicitly show instability. It gives

$$\frac{\|\boldsymbol{u}(t,\cdot)\|}{\|\boldsymbol{u}(0,\cdot)\|} = \left(1 + 4t^2 + 4\omega^2 t^2\right)^{1/2}, \tag{116}$$

where $\|u(t,\cdot)\|^2 = \|\ \gamma_{ij}(t,\cdot)\|^2 + \|K_{ij}(t,\cdot)\| + \|\ \partial_k\gamma_{ij}(t,\cdot)\ \|^2$. The ratio cannot be bounded by $Ke^{\alpha t}$ with K and α independent of ω.

To see that the second order accurate standard discretization is unstable we take $\gamma_{1A} = (-1)^j t$ and $K_{1A} = (-1)^{j+1}/2$. As in the continuum, the ratio

$$\frac{\|\boldsymbol{v}(t)\|_{h,D_+}}{\|\boldsymbol{v}(0)\|_{h,D_+}} = \left(1 + 4t^2 + 16\frac{t^2}{h^2}\right)^{1/2} \tag{117}$$

cannot be bounded. We can nevertheless compute the von Neumann condition, which is given by

$$\lambda \leqslant \frac{\sqrt{3}\alpha_0}{2\sqrt{7d}}. \tag{118}$$

In [26] stability tests were done with the non-linear version of this formulation. The domain used consisted of a thin channel, with an even number N of grid points in one spatial direction and three grid points in the other two directions. By taking this into account we see that modes corresponding to the frequencies $\xi_1 = \pi$, and $\xi_2 = \xi_3 = 2\pi/3$ grow exponentially if $\lambda > 0.4163$. Figure 2 in [26] confirms that with a Courant factor of $\lambda = 0.5$ there is a violation of the von Neumann condition.[9]

Although the symbol associated with the continuum system (114) and (115) has four Jordan blocks of size two for any ω, interestingly, the symbol associated with the semi-discrete problem obtained with the standard second order accurate discretization can have rather different properties. For Fourier modes traveling in directions parallel to the axis the continuum result still holds. However, for Fourier modes not parallel to any of the axis, we found that the symbol may have fewer Jordan blocks. For some Fourier frequencies we even noticed that the symbol is diagonalizable. There is no conflict between this observation and the fact that the continuum problem is ill-posed. As shown at the beginning of this subsection the discrete initial value problem for the ADM system is also ill-posed. In the limit of high resolution, $h \to 0$ ($\xi \to 0$ and ω fixed), the discrete symbol is a perturbation of the continuum one[10]

$$\hat{P}_d = \hat{P}_c + \mathrm{O}(h^2).$$

Even though for some frequencies $\hat{P}_d$ is diagonalizable, the characteristic variables become degenerate in the limit $h \to 0$, which implies that the discrete symmetrizer becomes unbounded (it is not possible to find a K, independent of h, satisfying inequality (53)).

4.6. The Z4 system

The same family of solutions that was used to show instability of the discretized ADM equations can be used for the standard discretization of the linearized $Z4$ system [19]

$$\begin{aligned}
\partial_t \alpha &= -f(K - m\Theta), \\
\partial_t \gamma_{ij} &= -2K_{ij}, \\
\partial_t K_{ij} &= -\partial_i\partial_j\alpha - \frac{1}{2}\partial_k\partial_k\gamma_{ij} + \partial_k\partial_{(i}\gamma_{j)k} - \frac{1}{2}\partial_i\partial_j\tau + 2\partial_{(i}Z_{j)}, \\
\partial_t \Theta &= \frac{1}{2}(\partial_k\partial_l\gamma_{kl} - \partial_k\partial_k\tau) + \partial_k Z_k, \\
\partial_t Z_i &= \partial_k K_{ik} - \partial_i K + \partial_i\Theta,
\end{aligned}$$

for any values of the parameters f and m. This instability, however, is not present if the D_0^2 discretization is used as in [24], in conjunction with the D_0-norm. Furthermore, it is possible that artificial dissipation may cure this instability of the standard discretization, at least for $0 < f \neq 1$ or $1 = f = m/2$, since in this case the continuum Cauchy problem is well-posed. Note that while we use the same family of solutions that was used to show instability for the ADM case, the two cases are very different: While the ADM instability is due to the lack of well-posedness of the continuum equations, the problem with the Z4 system arises purely at the discrete level, and can be traced back to the difference in structure between the principal symbols of the pseudodifferential first order reductions of the continuum and discrete equations, see Eqs. (43) and (52). For second order in space systems diagonalizability of the discrete symbol is not implied by diagonalizability of the continuum symbol.

The ADM and Z4 examples suggest a simple criterion that can be used to rule out certain schemes. Any first order in time, second order in space system of PDEs which gives rise to an ill-posed problem when the first order and mixed second order spatial derivatives are dropped will result in an unstable scheme if the standard discretization is used and no artificial dissipation is added. This is a consequence of the fact that grid

[9] A one-dimensional von Neumann analysis gives the limit (118) with $d = 1$ and $\alpha_0 = 2$, which corresponds to 0.655. However, this would not capture the fact that there could be exponentially growing modes with non-trivial dependence in the two thin directions.

[10] Note that in general by perturbing a non-diagonalizable matrix one obtains a diagonalizable matrix, so the diagonalizability of the discrete ADM symbol for some frequencies should not be so surprising.

modes with the highest frequency belong to the kernel of the D_0 operator. Although the D_0^2 discretization gives stable schemes with respect to the D_0-norm, provided that the continuum problem is well-posed, it suffers from the limitations described in Section 4.2.2.

5. Testing stability

When dealing with variable coefficient or non-linear problems it can be difficult, if not impossible, to prove stability with respect to a certain norm. Numerical experiments are often the only option. Given a discretization of the linear initial value problem (1) and (2), a stability test should be aimed at establishing the existence of the constants α and K, independent of the initial data and for all $h \leqslant h_0$ (and possibly $k \leqslant \lambda_0 h$), by computing the ratio between a suitable discrete norm at time-step $t_n = nk$ and its initial value,

$$\frac{\|v^n\|}{\|v^0\|} \leqslant K\mathrm{e}^{\alpha t_n}. \tag{119}$$

Although it is not possible to infer stability by examining a finite number of numerical experiments (one would have to explore the entire set $h \leqslant h_0$ that appears in the definition of stability), it is usually not difficult to spot a trend of behavior as the resolution is increased. To ensure that a wide range of frequencies is excited, random initial data can be used [27], as no smoothness assumptions are used in the definition of stability.

In the examples of first order in time, second order in space hyperbolic systems for which we are able to determine stability, we use a norm which is the discrete version of the continuum one. The derivatives are approximated using the one-sided operators D_+ (or, equivalently, D_-) rather than D_0. For the NOR system, for example, we use the square root of the expression

$$\sum_{i,j=1}^{3} \|\gamma_{ij}\|_h^2 + \sum_{i,j=1}^{3} \|K_{ij}\|_h^2 + \sum_{k,i,j=1}^{3} \|D_{+k}\gamma_{ij}\|_h^2 + \sum_{i=1}^{3} \|f_i\|_h^2.$$

If, as we vary the initial data and the resolution, the experiments indicate that the constants α and K in (119) exist, then one would conclude that the scheme appears to be stable. If not, the scheme appears to be unstable.

In the non-linear case, if the problem has a sufficiently smooth solution u_0, then to first approximation the error equation can be linearized about u_0 and convergence follows if the linearized equation is stable (Section 5.5 in [20]). Establishing stability experimentally using the linearized equations would not be very practical. However, convergence to a known exact solution can be tested directly and it avoids many complications. Rather than testing for stability, one could test convergence in a more demanding way: initial data can be chosen which is not smooth, but is accurate to the correct order in the appropriate norm. For instance, for the NOR system, one would use the square root of

$$\sum_{i,j=1}^{3} \|\delta\gamma_{ij}\|_h^2 + \sum_{i,j=1}^{3} \|\delta K_{ij}\|_h^2 + \sum_{k,i,j=1}^{3} \|D_{+k}(\delta\gamma_{ij})\|_h^2 + \sum_{i=1}^{3} \|\delta f_i\|_h^2,$$

where $\delta v = v - u_0$, and one could add random noise to the initial data with amplitude h^p for the K_{ij} and f_i variables and h^{p+1} for the γ_{ij} variables. The scheme is convergent around the solution u_0 if the D_+-norm of the error at time T is of order h^p. In particular, this implies that for a convergent scheme the discrete L_2 norm of the error is of order h^p if the D_+-norm of the initial error is of order h^p.

Finally, we note that the notion of *robust stability* introduced in [27] does not imply nor follows from the concept of numerical stability investigated in this paper.

5.1. Numerical tests

We have performed numerical tests to complement the analytical stability results of Section 4.

For each run, the numerical grid has dimensions $50\rho \times 4 \times 4$, where $\rho = 1, 2, 4, 8$ parameterizes the resolution, and we impose periodic boundary conditions. The coordinate domain is $x, y, z \in [-0.5, 0.5)$. The time integrator is RK4 with Courant factor $\lambda = 0.5$. We choose random noise of order unity as initial data (except for the Z4 tests, see below) so that many discrete Fourier modes are present in the initial data. Empirically, we find that using smooth initial data in the constant coefficient problems of this paper can make it difficult to observe an instability. This was also noticed in the non-linear case in [28].

Fig. 2 shows the results for the ADM system. The apparent trend is that as the resolution is increased, and higher frequency Fourier modes are present in the initial data, the ratio of the D_+-norm of the solution to its initial value at any given time increases. It appears that there is no K,α such that this quantity can be bounded by a function $Ke^{\alpha t_n}$, and this indicates that the system is unstable.

In Fig. 3, we show the results of the stability test for the linearized NOR system. The results suggest that the ratio of the D_+-norm of the solution to its initial value remains bounded, and hence that the system is stable. This reflects the analytic result that we proved in Section 4.

Showing the instability of the Z4 system was more complicated. In this case, it was not sufficient to use random initial data of order unity in all variables. When this was attempted, the ratio of the D_+-norm to its initial value remained bounded. In order to numerically demonstrate the instability, we used knowledge of the exact solution that violates the estimate. Random data of order unity was given to the variables K_{22} and K_{33} and the remaining variables were set to zero. The test results for the linearized Z4 system are shown in Fig. 4, and confirm that this system is unstable.

When artificial dissipation with $\sigma = 0.02$ is used, the linearized Z4 system tested with the same initial data shows no sign of instability. See Fig. 5.

The example of the Z4 system shows that numerical testing of stability is not always straightforward, and that schemes which appear stable for simple test cases may in fact be unstable. All tests were done using the standard second order accurate discretization.

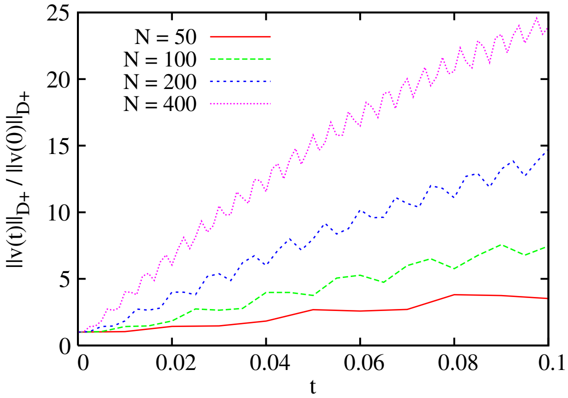

Fig. 2. Linearized ADM stability test.

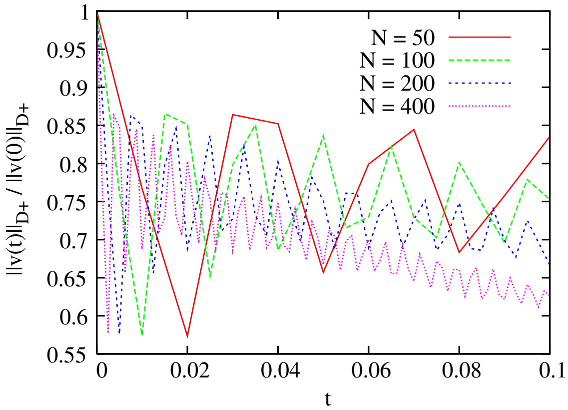

Fig. 3. Linearized NOR stability test, $r = 0$.

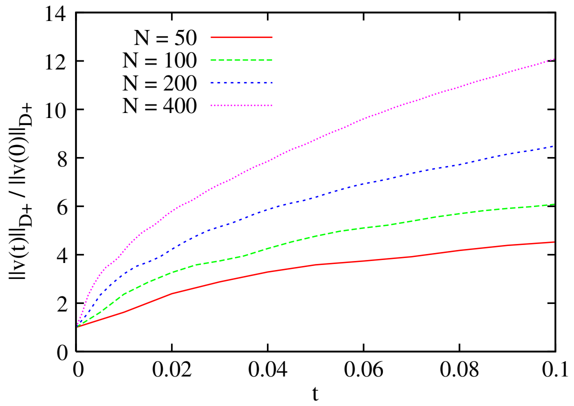

Fig. 4. Linearized Z4 stability test, $f = 1$, $m = 2$. Initial data consists of random values in K_{22} and K_{33}, and all other variables are zero.

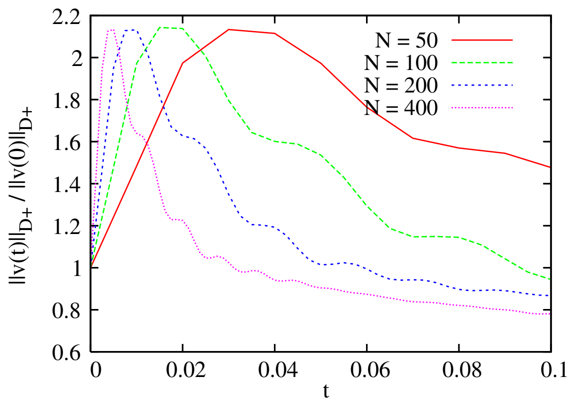

Fig. 5. Linearized Z4 stability test with dissipation $\sigma = 0.02$, $f = 1$, $m = 2$. Initial data consists of random values in K_{22} and K_{33}, and all other variables are zero.

6. Discussion

In this work, we extended the notion of numerical stability of finite difference approximations to include hyperbolic systems that are first order in time and second order in space. We considered the standard discretization of the wave equation, a generalization of the KWB formulation of electromagnetism and the NOR formulation of Einstein's equations linearized about the Minkowski solution. By analyzing the symbol of the second order system, and constructing a discrete symmetrizer, we were able to prove stability in a discrete norm containing one-sided difference operators, provided that the von Neumann condition is satisfied. Consistency and stability with respect to the D_+-norm imply convergence with respect to the discrete L_2 norm. We also found that in some cases ($r \geqslant 1$ in the NOR and generalized KWB systems, and Z4) standard discretizations of well-posed continuum problems can lead to unconditionally unstable schemes. This is closely related to the instability of the fully second order shifted wave equation investigated in [29], but our examples contain no shift terms.

Our analysis of discretizations of first order in time hyperbolic systems shows that in the first order in space case there is a clear correspondence between strong hyperbolicity and numerical stability, and between characteristic speeds and Courant limits. See inequality (63) and Eq. (64). In the second order in space case, on the other hand, the mixing of $D_\pm$ and D_0 operators breaks this correspondence. To restore the correspondence one could use the D_0^2 discretization, however, as discussed in Section 4.2.2, this can lead to difficulties.

In Section 4.6, we propose a simple criterion that can be used to rule out certain schemes when the standard discretization is used and no artificial dissipation is added. This criterion detects schemes in which the highest frequency mode grows faster as the resolution is increased.

We also discuss stability tests for second order in space systems. These tests should be aimed at establishing the existence, for sufficiently small h, of the constants K and α that appear in the definition of stability with respect to the D_+-norm. In the non-linear case the situation is more complicated. In this case we suggest, when an exact smooth solution of the continuum problem is available, to do convergence tests with initial data given by that of the continuum problem plus random noise of order h^p with respect to the D_+-norm (see Section 5).

Although our analysis was restricted to the constant coefficient case, we expect that for the variable coefficient case generalizations of results similar to those presented in Section 6.6 of [20] for first order hyperbolic systems, where artificial dissipation plays an important role, might apply.

Acknowledgments

We thank Carsten Gundlach and Olivier Sarbach for helpful discussions and suggestions. This research was supported by a Marie Curie Intra-European Fellowship within the 6th European Community Framework Program.

Appendix A. Time integrators

In this work, we restrict our attention to the following three time integrators: third and fourth order Runge–Kutta, and iterative Crank–Nicholson [30]. Given a system of ordinary differential equations, $\mathrm{d}y/\mathrm{d}t = f(t, y(t))$, these integrators are defined as

3RK

$$\begin{aligned} k_1 &= kf(t_n, y^n), \\ k_2 &= kf(t_n + k/2, y^n + k_1/2), \\ k_3 &= kf(t_n + 3k/4, y^n + 3k_2/4), \\ y^{n+1} &= y^n + (2k_1 + 3k_2 + 4k_3)/9. \end{aligned}$$

4RK

$$\begin{aligned} k_1 &= kf(t_n, y^n), \\ k_2 &= kf(t_n + k/2, y^n + k_1/2), \\ k_3 &= kf(t_n + k/2, y^n + k_2/2), \\ k_4 &= kf(t_n + k, y^n + k_3), \\ y^{n+1} &= y^n + (k_1 + 2k_2 + 2k_3 + k_4)/6. \end{aligned}$$

ICN

$$\begin{aligned} k_1 &= kf(t_n, y^n), \\ k_2 &= kf(t_n + k/2, y^n + k_1/2), \\ k_3 &= kf(t_n + k/2, y^n + k_2/2), \\ y^{n+1} &= y^n + k_3. \end{aligned}$$

Appendix B. Some numerical properties of first and second order systems

In this section, we assume that the time integrator is one of those discussed in Appendix A. We consider standard second and fourth order accurate discretizations of the following two toy model problems

$$u_t = u_x, \tag{B.1}$$

and

$$\phi_t = \Pi, \quad \Pi_t = \phi_{xx}. \tag{B.2}$$

Eq. (B.1) arises in the full reduction to first order of $\phi_{tt} = \phi_{xx}$, while (B.2) represents its reduction in time. If we denote by $\lambda(\xi)$ an eigenvalue of the discrete symbol, the corresponding phase and group velocities are given by

$$v_{\rm p} = {\rm i}\frac{\lambda(\xi)}{\omega},$$
$$v_{\rm g} = {\rm i}\frac{\rm d}{{\rm d}\omega}\lambda(\xi),$$

where $\xi = \omega h$. In the following table we compute the numerical phase velocities, $v_{\rm p}$, group velocities, $v_{\rm g}$, the Courant limits (C.l.), the frequencies of undamped modes (u.m.) and of the first unstable mode (f.u.m.) for the two systems. The numerical phase and group velocities are plotted in Fig. B.1 as a function of ξ.

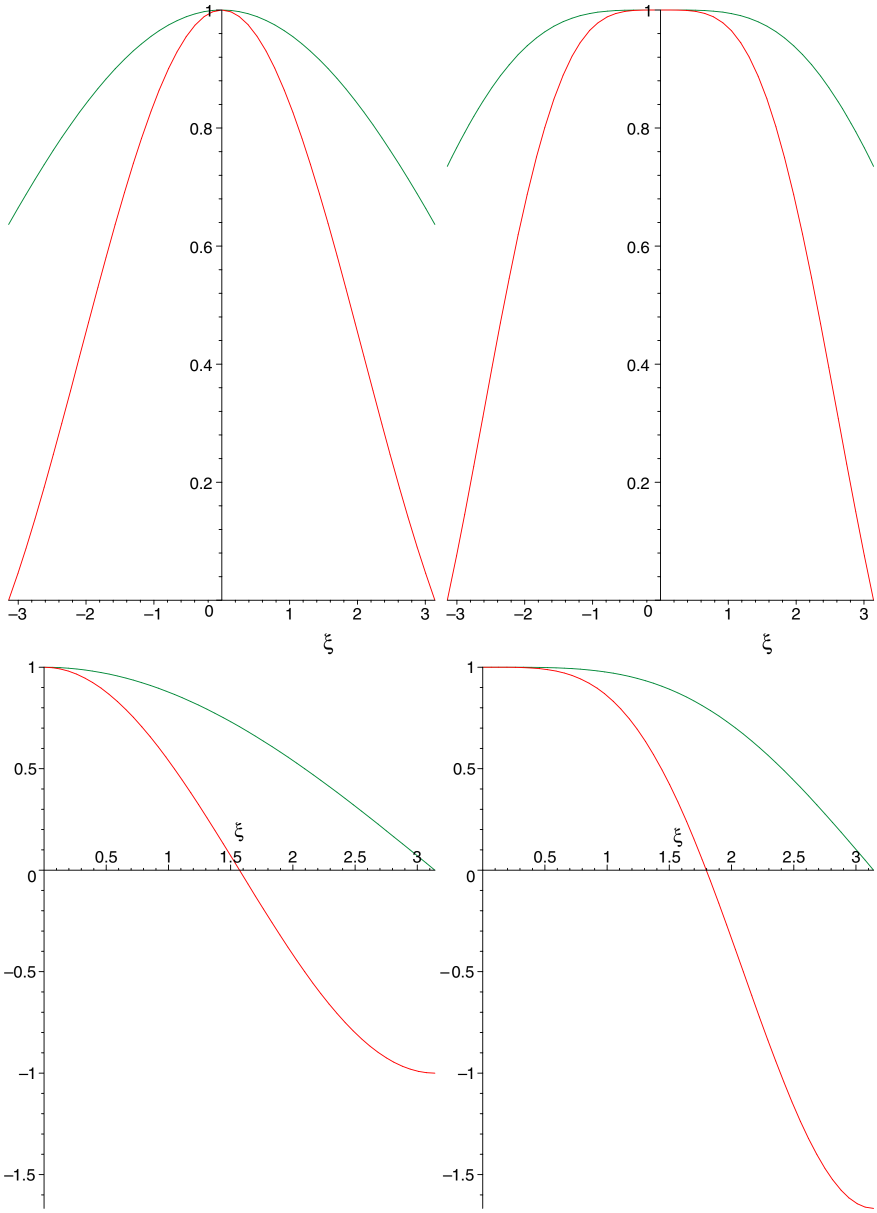

Fig. B.1. The phase (top) and group (bottom) velocities for the second (left) and fourth (right) order standard approximation of the advective Eq. (B.1) (red) and the wave Eq. (B.2) (green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In the table we used $\Delta^2 = 1 + \frac{1}{3}\sin^2\frac{\xi}{2}$. The exact continuum phase and group velocity is 1. The Taylor expansion of the numerical velocities gives an idea of the magnitude of the error, provided that enough grid-points per wave length are used. The table shows that in the second order accurate case the phase error for the wave equation is 4 times smaller than for the advective equation, and that this improvement in accuracy is even stronger for the fourth order accurate discretization.

Furthermore, the standard discretizations of fully first order hyperbolic systems have numerical phase velocities that vanish at the highest frequencies and numerical group velocities with the opposite sign to the continuum one. In numerical relativity simulations involving black holes which make use of the excision technique to handle the singularity one can expect to see numerical high frequency solutions escaping from the black hole, if a first order formulation combined with the standard discretization is used, unless artificial dissipation is added to the scheme.

Finally, whereas for (B.1) the transition from second order accuracy to fourth order implies the reduction of the Courant limit by a factor of 1.372, for the second order in space system (B.2), this transition requires a Courant limit $2/\sqrt{3} \approx 1.155$ times smaller. This indicates that there is an even higher gain in going to fourth order accuracy for second order in space formulations.

	Second order accurate		Fourth order accurate	
	Advective	Wave	Advective	Wave
v_p	$\frac{\sin\xi}{\xi} \approx 1 - \frac{\xi^2}{6} + O(\xi^4)$	$\frac{2}{\xi}\sin\frac{\xi}{2} \approx 1 - \frac{\xi^2}{24} + O(\xi^4)$	$\frac{\sin\xi}{\xi}(1 + \frac{2}{3}\sin^2\frac{\xi}{2}) \approx 1 - \frac{\xi^4}{30} + O(\xi^6)$	$\frac{2}{\xi}\sin\frac{\xi}{2}\Delta \approx 1 - \frac{\xi^4}{180} + O(\xi^6)$
v_g	$\cos\xi \approx 1 - \frac{\xi^2}{2} + O(\xi^4)$	$\cos\frac{\xi}{2} \approx 1 - \frac{\xi^2}{8} + O(\xi^4)$	$1 - \frac{8}{3}\sin^4\frac{\xi}{2} \approx 1 - \frac{\xi^4}{6} + O(\xi^6)$	$\cos\frac{\xi}{2}(1 + \frac{2}{3}\sin^2\frac{\xi}{2})/\Delta \approx 1 - \frac{\xi^4}{36} + O(\xi^6)$
C.l.	α_0	$\alpha_0/2$	$\alpha_0/1.372$	$\frac{\sqrt{3}}{4}\alpha_0 \approx \alpha_0/2.309$
u.m.	$0,\pi$	0	$0,\pi$	0
f.u.m.	$\pm\frac{\pi}{2} \approx \pm 1.571$	π	$\pm 2\arctan\left(\frac{6^{1/4}}{\sqrt{4-\sqrt{6}}}\right) \approx \pm 1.797$	π

Appendix C. Discrete constraint propagation

When simulating systems such as Maxwell's or Einstein's equations, one has to take into account that the data has to satisfy initial data constraints. The evolution equations guarantee that if these constraints are satisfied initially, then they will be satisfied at later times. In this appendix, we show that even in the constant coefficient case, when using standard discretizations of second order in space systems, the discrete constraints do not propagate exactly. Initial data which satisfy the discrete constraints do not lead to constraint satisfying solutions.

As an example, we consider the ADM Eqs. (114) and (115) with constraints

$$C \equiv \frac{1}{2}(\partial_i\partial_j\gamma_{ij} - \partial_i\partial_i\tau) = 0, \quad C_i \equiv \partial_j K_{ij} - \partial_i K = 0.$$

For simplicity we confine ourselves to solutions which depend only on one space coordinate. The discretized constraints are

$$C \equiv -\frac{1}{2}D_+D_-\gamma_{AA} = 0, \quad C_1 \equiv -D_0 K_{AA} = 0,$$
$$C_A \equiv D_0 K_{1A} = 0,$$

where $A = 2, 3$.

The time derivative of the first constraint cannot be expressed in terms of finite difference combinations of the constraints

$$\frac{\mathrm{d}}{\mathrm{d}t}C = D_+D_-K_{AA} \neq -D_0 C_1.$$

This is to be contrasted with the fact that in the constant coefficient case, the discrete constraints of a first order reduction would propagate as in the continuum, with partial derivatives replaced by D_0 operators. Furthermore, this issue would not be present if one used D_0^2 to approximate the second derivatives.

References

[1] M. Shibata, T. Nakamura, Phys. Rev. D 52 (1995) 5428.
[2] T. Baumgarte, S. Shapiro, Phys. Rev. D 59 (1999) 024007.
[3] S. Frittelli, O. Reula, Phys. Rev. Lett. 76 (1996) 4667.
[4] S.D. Hern, Ph.D. Thesis, University of Cambridge, 1999, gr-qc/0004036.
[5] A. Anderson, J.W. York Jr., Phys. Rev. Lett. 82 (1999) 4384.
[6] L.E. Kidder, M.A. Scheel, S.A. Teukolsky, Phys. Rev. D 64 (2001) 064017.
[7] O. Sarbach, G. Calabrese, J. Pullin, M. Tiglio, Phys. Rev. D 66 (2002) 064002.
[8] G. Nagy, O. Ortiz, O. Reula, Phys. Rev. D 70 (2004) 044012.
[9] C. Gundlach, J.M. Martin-Garcia, Phys. Rev. D 70 (2004) 044031.
[10] C. Gundlach, J.M. Martin-Garcia, Phys. Rev. D 70 (2004) 044032.
[11] H. Beyer, O. Sarbach, Phys. Rev. D 70 (2004) 104004.
[12] H. Kreiss, N. Petersson, J. Yström, SIAM J. Numer. Anal. 40 (2002) 1940–1967.
[13] B. Szilágyi, H.-O. Kreiss, J. Winicour, Phys. Rev. D 71 (2005) 104035.
[14] R.D. Richtmyer, K. Morton, Difference Methods for Initial Value Problems, Interscience Publisher, New York, 1967.
[15] S. Frittelli, R. Gómez, J. Math. Phys. 41 (2000) 5535.
[16] A. Knapp, E. Walker, T. Baumgarte, Phys. Rev. D 65 (2002) 064031.
[17] C. Nunn, M.phil. Thesis, University of Southampton, 2005.
[18] R. Arnowitt, S. Deser, C. Misner, in: L. Witten (Ed.), Gravitation: An Introduction to Current Research, Wiley, New York, 1962.
[19] C. Bona, T. Ledvinka, C. Palenzuela, M. Žáček, Phys. Rev. D 67 (2003) 104005.
[20] B. Gustafsson, H.-O. Kreiss, J. Oliger, Time Dependent Problems and Difference Methods, Wiley, New York, 1995.
[21] H.-O. Kreiss, J. Lorenz, Initial-Boundary Value Problems and the Navier–Stokes Equations, Academic Press, Boston, 1989.
[22] H.-O. Kreiss, G. Scherer, SIAM J. Numer. Anal. 29 (3) (1992) 640–646.
[23] H.-O. Kreiss, O.E. OrtizLecture Notes in Physics, vol. 604, Springer, New York, 2002.
[24] C. Bona, T. Ledvinka, C. Palenzuela, M. Žáček, Phys. Rev. D 69 (2004) 064036.
[25] M. Babiuc, B. Szilágyi, J. Winicour, gr-qc/0404092.
[26] M. Alcubierre et al., Class. Quantum Grav. 21 (2004) 589.
[27] B. Szilágyi, R. Gómez, N.T. Bishop, J. Winicour, Phys. Rev. D 62 (2000) 104006.
[28] G. Calabrese, J. Pullin, O. Sarbach, M. Tiglio, Phys. Rev. D 66 (2002) 064011.
[29] G. Calabrese, Phys. Rev. D 71 (2005) 027501.
[30] S.A. Teukolsky, Phys. Rev. D 61 (2000) 087501.