Journal of Computational Physics 229 (2010) 2675–2696

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Finite difference methods for second order in space, first order in time hyperbolic systems and the linear shifted wave equation as a model problem in numerical relativity

M. Chirvasa [a,*], S. Husa [b]

[a] Max-Planck-Institut für Gravitationsphysik, Am Mühlenberg 1, 14475 Potsdam, Germany
[b] Departament de Física, Universitat de les Illes Balears, Cra. Valldemossa Km. 7.5, Palma de Mallorca, E-07122, Spain

ARTICLE INFO

Article history:
Received 13 February 2009
Received in revised form 27 November 2009
Accepted 13 December 2009
Available online 21 December 2009

Keywords:
High order finite differencing
Partial differential equations
Wave equation
Numerical relativity

ABSTRACT

Motivated by the problem of solving the Einstein equations, we discuss high order finite difference discretizations of first order in time, second order in space hyperbolic systems. Particular attention is paid to the case when first order derivatives that can be identified with advection terms are approximated with non-centered finite difference operators. We first derive general properties of these discrete operators, then we extend a known result on numerical stability for such systems to general order of accuracy. As an application we analyze the shifted wave equation, including the behavior of the numerical phase and group speeds at different orders of approximations. Special attention is paid to when the use of off-centered schemes improves the accuracy over the centered schemes.

1. Introduction

Numerical discretization of first order hyperbolic systems of partial differential equations (PDEs) is greatly simplified by a result for the linear constant coefficient case [1]: if the Cauchy problem is well-posed, then the semi-discrete problem (only discretizing space and leaving time continuous) is stable when spatial derivatives are discretized with a centered finite difference operator (CFDO). Furthermore, when using simple Runge–Kutta methods [2] for time integration, for sufficiently small time step the fully-discrete problem is also stable.

Such a result does *not* hold in general for second order systems where first and second spatial derivatives appear [3]! In order to obtain a stable semi-discrete scheme, the second order system needs to have additional properties. In [3], which in the following we refer to as CHH, sufficient conditions for stability of the fully-discrete problem were presented for such systems. Although these conditions are valid for general order centered discretizations, this point has not been explicitly made. One of the results of this article is to make this statement clear, by closing a technical gap related to the boundedness of the lower order terms. The main focus of our work is on a detailed analysis of the numerical properties of discretizations where some first order derivatives are approximated with off-centered finite difference operators and artificial dissipation is added to the equations. The motivation for choosing this more general situation comes from numerical relativity, where it is a common practice to off-center by one point the derivatives corresponding to the Lie advection terms. In numerical simulations of black holes using the BSSN formulation of the Einstein equations [4–6], this procedure of off-centering was found to be essential even for sixth order schemes [7]. Numerical solutions of the Einstein equations are currently quickly expanding our knowledge about the astrophysics of compact binaries (see [8,9] for overviews on what has been achieved since the

* Corresponding author.
E-mail address: mihchi@gmail.com (M. Chirvasa).

0021-9991/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2009.12.016

major breakthroughs in 2005 [10–12]), but a systematic understanding of the underlying numerical techniques has not yet been achieved.

The shifted scalar wave equation serves as a simple but powerful model in numerical relativity [13–17]. The particular case with zero shift and flat background (standard wave equation) has been extensively studied in [18,19] and high order discretization methods have been proposed. In Section 2 we introduce the shifted scalar wave equation as a first order in time, second order in space system, together with a summary of the well-posedness theory for mixed order systems. We show stability for our semi-discrete problem, independent of shift or dissipation terms, while the fully-discrete problem requires artificial dissipation if more than one point is off-centered. Restricting to flat space in one space dimension, Courant limits and numerical phase and group speeds are computed and analyzed in detail. It is shown that increasing the off-centering reduces the Courant limit. However, by increasing the order of approximation while keeping the off-centering fixed, does not necessarily generate lower Courant limits. Regarding the numerical speeds, it is shown that indeed there are cases when off-centering improves the accuracy over the centered scheme. This fact is illustrated also experimentally by the results of some simple numerical tests at the end of Section 5.

Our analysis of the wave equation relies on certain properties of finite difference operators, in particular on their behavior in Fourier space. We introduce these operators in Section 3 together with highlighting some relevant properties. Then, in Section 4 we address the stability method and follow in Section 5 with the analysis of the wave equation. Our results are summarized in Section 6.

2. The shifted wave equation and first order in time second order in space hyperbolic systems

The scalar wave equation in a $d+1$-dimensional spacetime equipped with a Lorentzian metric $g_{\alpha\beta}$ reads

$$g^{\alpha\beta}\partial_\alpha\partial_\beta\Phi = 0. \tag{1}$$

We assume a uniform time slicing for simplicity, $g^{00} = -1$, and perform a $d+1$ split introducing a positive definite d-metric $\gamma^{jl} = g^{jl} + \beta^i\beta^j$, with $i,j = \overline{1,d}$ and a shift vector $\beta^i = g^{0i}$ (see, e.g. [20]). The wave Eq. (1) then becomes

$$\partial_{tt}\Phi = 2\beta^i\partial_i\partial_t\Phi + (\gamma^{jl} - \beta^i\beta^j)\partial_i\partial_j\Phi.$$

The mixed time-space derivatives lead to non-standard behavior as compared to the flat space wave equation with zero shift, and much of the material below will be devoted to their treatment.

We reduce the wave equation to a first order in time, second order in space form by introducing the variable K, in analogy with the York-ADM-system [21] (and other common representations of the Einstein equations),

$$\mathsf{K} = \partial_t\Phi - \beta^j\partial_j\Phi$$

which transforms the wave equation into the first order in time, second order in space system in the way most common in numerical relativity:

$$\partial_t\Phi = \beta^j\partial_j\Phi + \mathsf{K}, \quad \partial_t\mathsf{K} = \gamma^{jl}\partial_{jl}\Phi + \beta^j\partial_j\mathsf{K}. \tag{2}$$

Well-posedness for the Cauchy problem for the system (1) is a standard textbook result both in the original second order form and for reduction to first order symmetric hyperbolic form. In the latter form standard theorems for numerical stability apply [1]. Here we investigate the numerical stability for the first order in time, second order in space system (2), using the methods presented in [3]. In this respect, the appropriate generalization of the shifted wave equation is a linear system of PDEs with constant coefficients of the form [3]:

$$\frac{d}{dt}\mathsf{v}(t,x) = \mathsf{P}\mathsf{v}(t,x), \quad \mathsf{v} = (\mathsf{U},\mathsf{V})^T,$$

with $x \in \mathbb{R}^d$, $\mathsf{U} : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^p$, $\mathsf{V} : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^q$ and

$$\mathsf{P} = \begin{pmatrix} A^j\partial_j + B & C \\ D^{jl}\partial_{jl} + E^j\partial_j + F & G^j\partial_j + J \end{pmatrix}. \tag{3}$$

Note that the state vector v is split into two parts, U are those variables for which only first spatial derivatives appear, while second spatial derivatives of the V-variables do enter the PDE. The well-posedness of the Cauchy problem for first order in time, second order in space systems of PDEs systems has been clarified by [22–26]. We will here recall the presentation in CHH, where the well-posedness of such systems of PDEs is discussed in close analogy with the issue of numerical stability. It is natural to consider 2π-periodic solutions and turn the analysis in Fourier space.

In Fourier space the evolution problem reduces to a system of ordinary differential equations (ODEs) for the Fourier coefficients. By performing a first order reduction in Fourier space, it can be shown that well-posedness is not influenced by lower differential order terms, which we can therefore drop and consider the second-order principal symbol constructed as

$$\hat{\mathsf{P}}' = \begin{pmatrix} i\omega_0 A^n & C \\ -\omega_0^2 D^{nn} & i\omega_0 G^n \end{pmatrix},$$

where $\omega_0 = |\omega|,\ \omega = (\omega_1, \ldots, \omega_d) \in \mathbb{Z}^{\mathbb{d}},\ \mathbb{M}^{\mathbb{n}} = \mathbb{M}^{\mathbb{j}}{}_{\mathbb{j}}$ and $n_j = \omega_j \omega_0^{-1}$.

It is shown in [3] that if there exists a matrix $\hat{\mathsf{H}}(\omega) = \hat{\mathsf{H}}^\dagger(\omega)$ such that $\hat{\mathsf{H}}\hat{\mathsf{P}}' + \hat{\mathsf{P}}'^\dagger\hat{\mathsf{H}} = 0$ and a positive constant K, such that $K^{-1} I_{\omega_0} \leqslant \hat{\mathsf{H}} \leqslant K I_{\omega_0}$ (where $I_{\omega_0} = \mathrm{diag}[\omega_0^2 I_p, I_q]$), then the problem is well-posed in the norms

$$\|\mathsf{v}\|_\partial^2 = \int \sum_{j=1}^{d} |\partial_j \mathsf{U}|^2 + |\mathsf{V}|^2, \quad \|\mathsf{v}\|_{\mathsf{H}}^2 = \sum_{\omega \in \mathbb{Z}^{\mathbb{d}}} \hat{\mathsf{v}}^\dagger \hat{\mathsf{H}} \hat{\mathsf{v}}.$$

In [3] an analysis of numerical stability was performed in analogy with the proof of well-posedness as we have sketched it, and which we will extend to arbitrary approximation order in Section 4. But, before going into stability analysis, we need to discuss some general properties of finite difference operators.

3. Finite difference operators

3.1. Construction and properties in one dimension

Consider a mesh of equidistant points $x_v = vh$, with $v \in \mathbb{Z}$ and h representing the grid spacing. Corresponding to the continuum vector-function $\mathsf{v} : \mathbb{R} \to \mathbb{C} \times \cdots \times \mathbb{C}$ we associate the grid vector-function v by $v : \{x_v,\ v = 0, \pm 1, \pm 2, \ldots\} \to \mathbb{C} \times \cdots \times \mathbb{C}$ and $v_v \equiv v(x_v) = \mathsf{v}(x_v)$.

Using $2n+1$ consecutive points, we want to construct the finite difference operator corresponding to the m-derivative. Let $s \in \{0, 1, \ldots, n\}$ be the offset of these points from symmetry with respect to the center, ($s = 0$ for CFDO) and ϵ the direction of off-centering ($\epsilon = 1$ for off-centering to the right, $\epsilon = -1$ for off-centering to the left).[1] Then the finite difference operator to be constructed will be denoted $D^{(m,n,s,\epsilon)}$. It is a linear combination of shift operators of the form:

$$D^{(m,n,s,\epsilon)} = h^{-m} \sum_{k=-n+\epsilon s}^{n+\epsilon s} \tilde{f}_{m,n,s,\epsilon,k} S^k, \tag{4}$$

where S^k be the shift operator by k points, $S^k v_v = v_{v+k}$. The weights $\tilde{f}_{m,n,s,\epsilon,k}$ can be expressed as the coefficients of y^k in the Taylor expansion of the function

$$f^{m,n,s,\epsilon}(y) = y^{n-\epsilon s} (\ln y)^m$$

around the point $y_0 = 1$ up to the order $(y - y_0)^{2n}$ (see Appendix A for the proof). Using this procedure, one can deduce explicit expressions for the finite difference operators corresponding to the first and second derivative (the relations (60) from Appendix A).

These expressions are fairly complicated, but they can be written in a more convenient form if we make use of the elementary finite difference operators:

$$\begin{aligned} D_\pm v_v &= \pm h^{-1}(v_{v\pm1} - v_v), \\ \delta_0 &= \frac{h}{2}(D_+ + D_-), \\ p &= h(D_+ - D_-) = h^2 D_+ D_-. \end{aligned} \tag{5}$$

Note that the operators δ_0 and p are dimensionless.

Then, a direct but lengthy calculation starting from the definitions (60) leads us to the following expressions for $D^{(1,n)} \equiv D^{(1,n,0,0)},\ D^{(2,n)} \equiv D^{(2,n,0,0)}$, the rest $R^{(n)} \equiv (D^{(1,n)})^2 - D^{(2,n)}$ and $D^{(1,n,s,\epsilon)}$:

$$\begin{aligned} D^{(1,n)} &= h^{-1}\delta_0 \left(1 + \sum_{k=1}^{n-1} c_k p^k\right), \\ D^{(2,n)} &= h^{-2} p \left(1 + \sum_{k=1}^{n-1} d_k p^k\right), \\ R^{(n)} &= h^{-2} \frac{n c_{n-1}}{2} p^{n+1} \sum_{k=0}^{n-1} \frac{c_k}{n+1+k} p^k, \\ D^{(1,n,s,\epsilon)} &= D^{(1,n)} + h^{-1}\left(\delta_0 \sum_{k=1}^{s-1} a_k p^k - \epsilon p \sum_{k=1}^{s-1} b_k p^k\right) p^n, \end{aligned} \tag{6}$$

[1] Though one can simplify the notation by dropping ϵ and considering $s \in \{-n, \ldots, n\}$, it will later turn out useful to separate the sign of s and its absolute value.

Table 1
Leading order truncation errors for the first order discrete derivative.

$\lvert T^{(1,n,s)}\rvert$	$n=1$	$n=2$	$n=3$	$n=4$
$s=0$	$\frac{1}{6}$	$\frac{1}{30}$	$\frac{1}{140}$	$\frac{1}{630}$
$s=1$	$\frac{1}{3}$	$\frac{1}{20}$	$\frac{1}{105}$	$\frac{1}{504}$

where

$$c_k = (-1)^k \frac{(k!)^2}{(2k+1)!}, \quad d_k = \frac{c_k}{k+1},$$
$$a_k = (-1)^n \sum_{j=k}^{s-1} \frac{(-1)^j C_{j+k}^{2k}}{(n-j)C_{2n}^{n+j}}, \quad b_k = (-1)^{n+s} \frac{C_{s+k}^{2k+1}}{2(n+1+k)C_{2n}^{n+s}}. \tag{7}$$

Notice that the coefficients c_k and d_k do not depend on n, while $a_k,\ b_k$ do depend on n and s.

The leading order truncation error of order $2n$ is defined as

$$\left.\frac{dv}{dx}\right|_{x_0} - D^{(1,n,s,\epsilon)} v_0 \equiv T^{(1,n,s,\epsilon)} \left.\frac{d^{2n+1}v}{dx^{2n+1}}\right|_{x_0} h^{2n} + O(h^{2n+1}),$$
$$\left.\frac{d^2v}{d^2x}\right|_{x_0} - D^{(2,n)} \equiv T^{(2,n)} \left.\frac{d^{2n+2}v}{dx^{2n+2}}\right|_{x_0} h^{2n} + O(h^{2n+2}).$$

A direct calculation yields

$$T^{(1,n,s,1)} = T^{(1,n,s,-1)} = T^{(1,n,s)} = (-1)^{s+n} \frac{(n+s)!(n-s)!}{(2n+1)!},$$
$$T^{(2,n)} = d_n = (-1)^n \frac{2(n!)^2}{(2n+2)!}.$$

It is well known that the centered FDO has the smallest leading order truncation error, $|T^{(1,n,0)}| < |T^{(1,n,s)}|$ for $s > 0$ (see also Table 1).

3.2. Fourier representation of difference operators

We assume a finite grid defined by a set of N points,

$$\mathcal{S}_{\underline{x}}(N) = \{x_v = vh, \text{ with } h = 2\pi/N, \ \forall v = 0, \ldots, N-1\}, \tag{8}$$

and consider periodic grid functions $v_v = v_{\mathrm{mod}(v,N)}$, decomposed as

$$v_v = \sum_{\omega \in \mathcal{S}_{\underline{\omega}}(N)} \hat{v}(\omega) b_v(\omega), \tag{9}$$

where

$$b_v(\omega) = (2\pi)^{-1/2} e^{i\omega x_v}, \tag{10}$$
$$\mathcal{S}_{\underline{\omega}}(N) = \begin{cases} \{-N/2+1, \ldots, N/2\}, & \text{if } N \text{ is even}, \\ \{-(N-1)/2, \ldots, (N-1)/2\}, & \text{if } N \text{ is odd}. \end{cases} \tag{11}$$

The set $\mathcal{S}_{\underline{\omega}}(N)$ represents the set of discrete wave numbers, and in the space of periodic grid functions the set $\{b_v(\omega), \omega \in \mathcal{S}_{\underline{\omega}}(N)\}$ forms a orthonormal basis with respect to the scalar product and the associated norm

$$(v,u)_h = \sum_{x_v \in \mathcal{S}_{\underline{x}}(N)} v^\dagger(x_v) u(x_v) V_h, \quad \|v\|_h^2 = (v,v)_h. \tag{12}$$

with $V_h = h$. The quantities $\hat{v}(\omega)$ represent the *discrete Fourier coefficients*. The scalar product satisfies the Parseval relation:

$$(v,u)_h = \sum_{\omega \in \mathcal{S}_{\underline{\omega}}(N)} \hat{v}^\dagger(\omega) \hat{u}(\omega). \tag{13}$$

Let $\xi = \omega h \in \mathcal{S}_{\underline{\xi}}(N)$ with

$$\mathcal{S}_{\underline{\xi}}(N) = \begin{cases} \{-\pi + 2\pi/N, \ldots, \pi\}, & \text{if } N \text{ is even}, \\ \{-\pi + \pi/N, \ldots, \pi - \pi/N\}, & \text{if } N \text{ is odd}. \end{cases} \tag{14}$$

Now apply the shift operator S^k on a basis vector $b_v(\omega)$. This leads to

$$S^k e^{i\omega x_v} = \widehat{S}^k(\xi) e^{i\omega x_v}, \quad \text{with} \quad \widehat{S}^k(\xi) = e^{i\xi k}.$$

The function $\widehat{S}^k(\xi)$ represents the *discrete Fourier symbol* of the shift operator. For any discrete operator $D = \sum_k a^k(h) S^k$ the Fourier symbol is defined by

$$D e^{i\omega x_v} = \widehat{D}(\xi; h) e^{i\omega x_v}, \quad \text{with} \quad \widehat{D}(\xi; h) = \sum_k a^k(h) \widehat{S}^k(\xi),$$

and for a general finite difference operator $D^{(m,n,s,\epsilon)}$ the symbol is

$$\widehat{D}^{(m,n,s,\epsilon)}(\xi; h) = h^{-m} \sum_{k=-n+\epsilon s}^{n+\epsilon s} \tilde{f}_{m,n,s,\epsilon,k} e^{i\xi k}. \tag{15}$$

For the elementary discrete operators (5) we obtain

$$\begin{aligned}
&\widehat{D}_{\pm}(\xi; h) = \pm h^{-1}(e^{\pm i\xi} - 1),\\
&\hat{\delta}_0(\xi) = i\hat{\delta}(\xi), \quad \text{where } \hat{\delta}(\xi) \equiv \sin\xi,\\
&\hat{p}(\xi) = -\hat{\Omega}^2(\xi), \quad \text{where } \hat{\Omega}(\xi) \equiv 2\sin\frac{\xi}{2},
\end{aligned}$$

and it is useful to note that $|\widehat{D}_+(\xi; h)| = |\widehat{D}_-(\xi; h)| = h^{-1}\hat{\Omega}(\xi)$.

The symbols for the first and second order derivative operators are straightforwardly computed using (6),

$$\begin{aligned}
&\widehat{D}^{(1,n)}(\xi; h) = i\hat{d}^{(1,n)}(\xi) h^{-1},\\
&\widehat{D}^{(2,n)}(\xi; h) = -\hat{d}^{(2,n)}(\xi) h^{-2},\\
&\widehat{R}^{(n)}(\xi; h) = \hat{r}^{(n)}(\xi) h^{-2}\\
&\widehat{D}^{(1,n,s,\epsilon)}(\xi; h) = \left(\epsilon \hat{\mathbf{d}}^{(1,n,s)}(\xi) + i\hat{d}^{(1,n,s)}(\xi)\right) h^{-1},
\end{aligned} \tag{16}$$

where we define

$$\begin{aligned}
&\hat{d}^{(1,n)} \equiv \hat{\delta} \sum_{k=0}^{n-1} |c_k| \hat{\Omega}^{2k},\\
&\hat{d}^{(2,n)} \equiv \hat{\Omega}^2 \sum_{k=0}^{n-1} |d_k| \hat{\Omega}^{2k},\\
&\hat{r}^{(n)} = -(\hat{d}^{(1,n)})^2 + \hat{d}^{(2,n)} = \frac{n|c_{n-1}|}{2} \hat{\Omega}^{2(n+1)} \sum_{k=0}^{n-1} \frac{|c_k|}{n+1+k} \hat{\Omega}^{2k},\\
&\hat{\mathbf{d}}^{(1,n,s)} = \hat{\Omega}^{2n+2} \sum_{k=0}^{s-1} (-1)^{n+k} b_k \hat{\Omega}^{2k},\\
&\hat{d}^{(1,n,s)} = \hat{d}^{(1,n)} + \hat{\delta}\hat{\Omega}^{2n} \sum_{k=0}^{s-1} (-1)^{n+k} a_k \hat{\Omega}^{2k}.
\end{aligned} \tag{17}$$

In the following we list a series of particularly relevant properties of the Fourier symbols, further properties are given in Appendix C.

First note that the quantities $\hat{r}^n$ and $\hat{d}^{(2,n)}$ are positive, and even more, from (17) it is easy to check that the following inequalities hold:

$$0 \leqslant \hat{d}^{(2,n)} \leqslant \hat{d}^{(2,n+1)}, \tag{18}$$

$$1 \geqslant \frac{\hat{r}^{(n)}}{\hat{d}^{(2,n)}} \geqslant \frac{\hat{r}^{(n+1)}}{\hat{d}^{(2,n+1)}}, \tag{19}$$

$$C_n^{-1}\hat{\Omega}^2 \leqslant \hat{\Omega}^2 \leqslant \hat{d}^{(2,n)} \leqslant C_n \hat{\Omega}^2, \quad \text{where } C_n \equiv 1 + \sum_{k=1}^{n-1} |d_k| 4^k \geqslant 1. \tag{20}$$

The real part of the Fourier symbol of the first derivative is an even function of the frequency ξ, while the imaginary part is an odd function. The real part of the Fourier symbol of the first derivative also

- vanishes for centered operators ($s = 0$),
- keeps the same sign for all frequencies, in case the operator is one-point off-centered ($s = 1$),
- changes sign for off-centering by more than one point ($s > 1$).

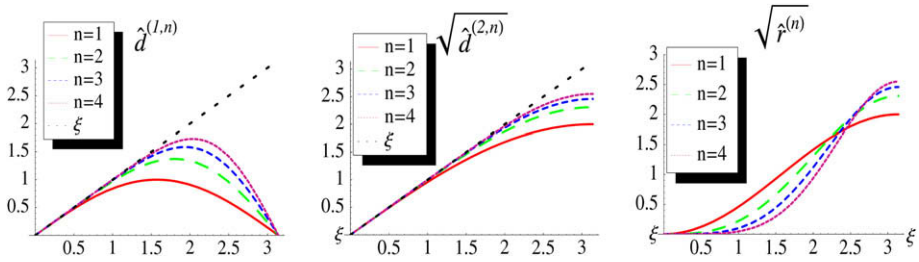

Fig. 1. Fourier symbols as functions of the frequency ξ, for different approximation orders. For increasing order the second derivative becomes more accurate for all frequencies, while the first derivative does not converge for $\xi = \pi$.

The derivatives with respect to ξ of the Fourier functions satisfy:

$$\partial_\xi \hat{d}^{(2,n)} = 2\hat{d}^{(1,n)}, \quad \partial_\xi \hat{r}^{(n)} = 2\frac{(n!)^2}{(2n)!}\hat{\Omega}^{2n}\hat{d}^{(1,n)}, \tag{21}$$

$$\partial_\xi \hat{\mathbf{d}}^{(1,n,s)} = \frac{(-1)^s}{C_{2n}^{n-s}}\sin(s\,\xi)\hat{\Omega}^{2n}, \quad \partial_\xi \hat{d}^{(1,n,s)} = 1 - \frac{(-1)^s}{C_{2n}^{n-s}}\cos(s\,\xi)\hat{\Omega}^{2n}.$$

In the following we show some plots to illustrate how the errors of the Fourier symbols scale with the order of approximation and off-centering. The error is defined in respect to the continuum limit, i.e., $h^{-m}(i\xi)^m$ for $\widehat{D}^{(m,n,s,\epsilon)}$.

Fig. 1 shows the Fourier symbols $\hat{d}^{(1,n)}$, $\hat{d}^{(2,n)}$ and $\hat{r}^{(n)}$ as functions of the frequency ξ for different orders of accuracy. For increasing approximation order, the second derivative becomes more accurate for all frequencies, while the first derivative does not converge to the continuum limit for the highest frequency in the grid, where the symbol is zero. The π-frequency will not be captured also by the off-centered discrete operators associated with the first derivative. In addition, for them, the error scales with the order only at small frequencies.

Fig. 2 shows the scaling of the error for $\hat{d}^{(1,n,s)}$ with the off-centering at fixed order of approximation. In the region of small frequencies, off-centering increases the error. At larger frequencies this behavior changes. For each s, there are exactly s frequencies in $(0,\pi)$ where the error cancels. However, for $s \geqslant 2$ there are large intervals where the error overcomes by far the error when $s = 0$. For $s = 1$ we observe that while at small frequencies, the error is slightly larger than for $s = 0$, for each order $2n$, there is a frequency, $\xi^{(n)}$, beyond which the error is smaller than for the case $s = 0$. This frequency can be computed numerically, e.g. $\xi^{(1)} = 1.3787$, $\xi^{(2)} = 1.0036$, $\xi^{(3)} = 0.8234$, $\xi^{(4)} = 0.7136$.

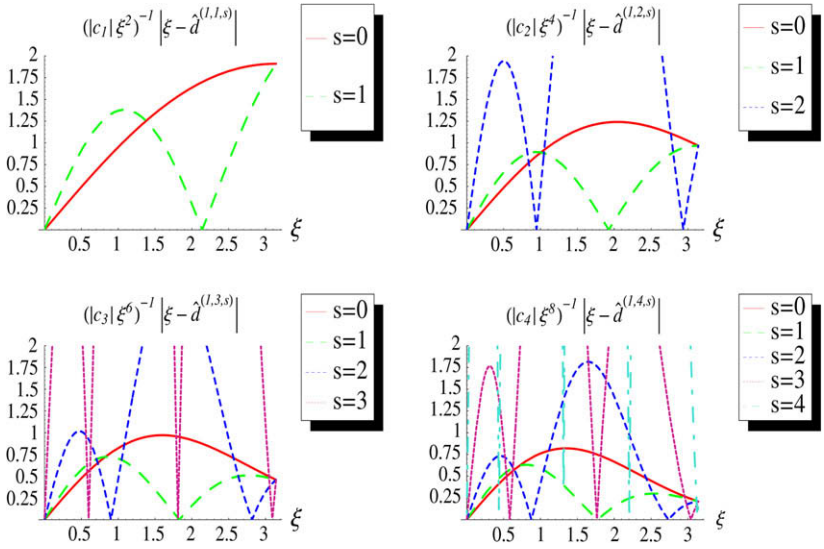

Fig. 2. Errors for $\hat{d}^{(1,n,s)}$ at fixed order but different off-centerings are shown, scaled with $|c_n|\xi^{2n}$ $(n = 1,2,3,4)$. For small ξ the curves are straight lines with the slope $|c_n T^{(1,n,s)}|$, and larger s increases the error. At higher frequencies this behavior changes, but for $s \geqslant 2$ large intervals appear where the error overcomes by far the error when $s = 0$. For $s = 1$ we observe that while at small frequencies, the error is slightly larger than for $s = 0$, for each order there is a frequency $\xi^{(n)}$ beyond which the error is smaller than for $s = 0$.

3.3. Generalization to d-dimensions

In this work we will use the common straightforward generalization of finite difference operators from one to d dimensions. We extend first derivatives in a particular coordinate direction in the trivial way, and second derivative operators in the jl-directions are defined as

$$D^{(2,n)}_{jl} = \begin{cases} D^{(1,n)}_j D^{(1,n)}_l, & \text{for } j \neq l, \\ D^{(2,n)}_j, & \text{for } j = l. \end{cases} \tag{22}$$

The Fourier symbols for the first and second derivative operators take the form

$$\begin{aligned} \widehat{D}^{(1,n,s,\epsilon)}_j &= h_j^{-1}\left(\epsilon_j \hat{\mathbf{d}}^{(1,n,s_j)}(\xi_j) + i\hat{d}^{(1,n,s_j)}(\xi_j)\right), \\ \widehat{D}^{(1,n)}_j &= h_j^{-1} i\hat{d}^{(1,n)}_j, \\ \widehat{D}^{(2,n)}_{jl} &= h_j^{-1}h_l^{-1}\begin{cases} -\hat{d}^{(1,n)}_j \hat{d}^{(1,n)}_l & j \neq l, \\ -\hat{d}^{(2,n)}_j & j = l. \end{cases} \end{aligned} \tag{23}$$

In order to simplify notation, we will also use the following convention: any function in frequencies, and possible, grid spacings, $\hat{f}(\xi_{i_1},\ldots,\xi_{i_r};h_{i_1},\ldots,h_{i_r})$ will be referred to by $\hat{f}_{i_1\cdots i_r}$, and by $\hat{f}$ in case it depends on all frequencies. More detailed definitions for the d-dimensional case are given in Appendix B.

3.4. Dissipation operators

In order to achieve numerical stability for problems that go beyond the linear constant coefficient case, it is common practice to add artificial dissipation to the right-hand-sides of the time evolution equations. In this work we only deal with the constant coefficient problem, but in (5.2) we will also use dissipation to stabilize numerical schemes which would be unstable otherwise.

Dissipation terms are typically chosen to converge away fast enough so as not to change the convergence order of the scheme. Here we use the Kreiss–Oliger dissipation operator $\mathcal{D}^{(2m)}$ of order $2m$ [1] and its Fourier representation $\widehat{\mathcal{D}}^{(2m)}$,

$$\mathcal{D}^{(2m)} = -\frac{(-1)^m}{2^{2m}}\sum_{j=1}^{d}\sigma_j h_j^{2m-1}(D_{+j})^m(D_{-j})^m, \quad \widehat{\mathcal{D}}^{(2m)} = -\frac{1}{2^{2m}}\sum_{j=1}^{d}\frac{\sigma_j}{h_j}\hat{\Omega}_j^{2m}, \tag{24}$$

for a $2m-2$ accurate scheme, where the parameters $\sigma_j \geqslant 0$ regulate the strength of the dissipation. Using this form of numerical dissipation, theorems can be proved concerning the numerical stability of non-constant-coefficient hyperbolic PDEs [1]. Note that it is more common to have the dissipation parameters σ_j not depend on the direction or other parameters of the system.

4. Numerical stability for first order in time, second order in space hyperbolic systems

We now turn to the analysis of numerical stability for the system (3), following [3]. This problem is greatly simplified by adopting the method-of-lines approach where initially time is kept continuous and only space is discretized (i.e. the semi-discrete problem). Then the discrete system to be analyzed becomes

$$\begin{aligned} &\frac{d}{dt}v = Pv, \quad v = (U,V)^T, \\ &P = \begin{pmatrix} A^j D^{(1,n)}_j + B & C \\ D^{jl}D^{(2,n)}_{jl} + E^j D^{(1,n)}_j + F & G^j D^{(1,n)}_j + J \end{pmatrix}. \end{aligned} \tag{25}$$

We consider periodic grid functions in each direction, and Fourier transform the system as discussed in Appendix B. Then a first order reduction is performed by introducing the variable $\hat{w}$,

$$\hat{w} \equiv i\Omega_0\hat{u}, \quad \Omega_0^2 = \sum_{j=1}^{d}|\widehat{D}_{+j}|^2, \tag{26}$$

where $\widehat{D}_{+j}$ is the Fourier symbol of the forward finite difference operator in the j-direction, D_{+j}. The case $\Omega_0 = 0$ (which corresponds to zero frequencies in all directions) does not play any role in the stability analysis,[2] so we define $\mathcal{S}^*_{\underline{\xi}}(N) = \mathcal{S}_{\underline{\xi}}(N) - \mathbf{0}^d$ and assume $\xi \in \mathcal{S}^*_{\underline{\xi}}(N)$.

[2] The zero frequency vector corresponds to a term constant in space.

By (26) we obtain the following system of ODEs

$$\frac{d}{dt}\hat{v}_R = \widehat{P}_R\hat{v}_R \quad \text{with} \quad \hat{v}_R = (\hat{u},\hat{w},\hat{v})^T,$$
$$\widehat{P}_R = \begin{pmatrix} B & (i\Omega_0)^{-1}A^j\widehat{D}_j^{(1,n)} & C \\ 0 & A^j\widehat{D}_j^{(1,n)} + B & i\Omega_0 C \\ F & (i\Omega_0)^{-1}\left(D^{jl}\widehat{D}_{jl}^{(2,n)} + E^j\widehat{D}_j^{(1,n)}\right) & G^j\widehat{D}_j^{(1,n)} + J \end{pmatrix}. \tag{27}$$

Using the Theorem 5.1.2 of [1] CHH show that the terms which correspond to the continuum lower order terms can be dropped from $\widehat{P}_R$ without affecting the stability analysis if

$$(i\Omega_0)^{-1}\widehat{D}_j^{(1,n)}, \quad k\widehat{D}_j^{(1,n)}, \quad k\Omega_0^{-1}\widehat{D}_{jl}^{(2,n)} \tag{28}$$

are bounded for all frequencies $\xi \in \mathcal{S}_{\underline{\xi}}^*(N)$. In the relations (28), k represents the time step. We will show in Lemma 4.1 that this is indeed the case for any order of accuracy $2n$.

Having proved this, the rest of the discussion in CHH applies. The problem now reduces to the analysis of a first order system with the principal part:

$$\widehat{P}'_R = \begin{pmatrix} A^j\widehat{D}_j^{(1,n)} & i\Omega_0 C \\ (i\Omega_0)^{-1}D^{jl}\widehat{D}_{jl}^{(2,n)} & G^j\widehat{D}_j^{(1,n)} \end{pmatrix}. \tag{29}$$

For this type of system, sufficient conditions for stability have been deduced in [1]. These conditions have been exploited in CHH to analyze the stability of the second order system. By introducing the so-called *second-order principal symbol* of the semi-discrete system,

$$\widehat{P}' = \begin{pmatrix} A^j\widehat{D}_j^{(1,n)} & C \\ D^{jl}\widehat{D}_{jl}^{(2,n)} & G^j\widehat{D}_j^{(1,n)} \end{pmatrix}, \tag{30}$$

and assuming that the time integration is done using one-step explicit schemes, CHH show that the following conditions are sufficient for stability:

Condition 1: There exists a hermitian matrix $\widehat{H}(\xi,h)$ such that

$$K^{-1}I_{\Omega_0} \leqslant \widehat{H} \leqslant KI_{\Omega_0}, \quad I_{\Omega_0} = \mathrm{diag}[\Omega_0^2, I_{qN}],$$
$$\widehat{H}\widehat{P}' + \widehat{P}'^{\dagger}\widehat{H} = 0, \tag{31}$$

for some positive constant K.

Condition 2: The eigenvalues of $k\widehat{P}'$ have non-positive real parts and

$$\sigma(k\widehat{P}\prime) \leqslant \alpha_0, \tag{32}$$

where $\sigma(k\widehat{P}')$ is the maximum spectral radius of $k\widehat{P}'$ and α_0 is a constant specific to the time integrator.

Remarks

- The condition (31) implies that the semi-discrete problem is stable with respect to the norms $D_\pm$ defined as:

$$\|v\|_{h,D_\pm}^2 = \sum_{i=1}^{d}\|D_{\pm i}U\|_h^2 + \|V\|_h^2, \tag{33}$$

  where $\|.\|_h^2$ is the d-dimensional analog of (12).
- The semi-discrete problem is stable also in the norm $\|v\|_{h,H}^2$ defined by:

$$\|v\|_{h,H}^2 = \sum_{\omega\in\mathcal{S}_{\underline{\omega}}(N)}\hat{v}^\dagger\widehat{H}\hat{v}. \tag{34}$$

  This norm is conserved by the principal symbol of the evolution system, $\|v(t,.)\|_{h,H} = \|v(0,.)\|_{h,H}$.

- The constant α_0 in (32) denotes the radius of local stability on the imaginary axis (R_{lsia}) in case the eigenvalues of $k\widehat{P}'$ are purely imaginary, and the radius of local stability (R_{ls}), otherwise.[3]
- In case all the grid spacings are equal, $h_1 = \cdots = h_d = h$, and we introduce the Courant factor $\lambda = k/h$, then the relation (32) provides the Courant limit:

$$\lambda \leqslant \frac{\alpha_0}{\sigma(h\widehat{P}\prime)}. \tag{35}$$

- If the right-hand-side of the system (25) is modified by adding artificial dissipation (using the operator $\mathcal{D}^{(2m)}$ defined in (24)) and/or by adding shift advection terms of the form $I\beta^j D_j^{(1,n,s_j,\epsilon)}$ (where $D_j^{(1,n,s_j,\epsilon)}$ is the non-centered FDO in the j-direction constructed from (6)), these modifications only have effect on the diagonal entries of the principal part. The new system will have different eigenvalues than $\widehat{P}'$ but the same set of eigenvectors. The symmetrizer will not depend on the way we discretize the advection terms, nor on the dissipation operator. The stability **Conditions 1–2** remain valid if

$$(i\Omega_0)^{-1}\widehat{D}_j^{(1,n,s_j,\epsilon)}, \quad (i\Omega_0)^{-1}\widehat{\mathcal{D}}^{(2m)} \tag{36}$$

are bounded and this will be shown below together with the boundedness of the terms (28).

Lemma 4.1. *The following quantities are bounded for all frequencies* $\xi \in \mathcal{S}^*_{\underline{\xi}}(N)$*:*

$$(i\Omega_0)^{-1}\widehat{D}_j^{(1,n)}, k\widehat{D}_j^{(1,n)}, \quad k\Omega_0^{-1}\widehat{D}_{jl}^{(2,n)}, \quad (i\Omega_0)^{-1}\widehat{D}_j^{(1,n,s_j,\epsilon)}, \quad (i\Omega_0)^{-1}\widehat{\mathcal{D}}^{(2m)}. \tag{37}$$

Making use of the relations (23) and (24), the proof reduces to showing the boundedness of

$$\hat{\Omega}_0^{-1}\hat{d}_j^{(1,n)}, \quad \hat{d}_j^{(1,n)}, \quad \hat{\Omega}_0^{-1}\hat{d}_j^{(2,n)}, \quad \hat{\Omega}_0^{-1}(\hat{d}_j^{(1,n)})^2, \quad \hat{\Omega}_0^{-1}\hat{d}_j^{(1,n,s_j)}, \quad \hat{\Omega}_0^{-1}\hat{\mathbf{d}}_j^{(1,n,s_j)}, \quad \hat{\Omega}_0^{-1}\hat{\Omega}_j^{2m}.$$

From the relations (17) we observe that each of these quantities can be written formally as a product $\hat{\Omega}_0^{-1}\hat{\Omega}_j F(\xi_j)$, with $F(\xi_j)$ a continuous and bounded function in $(-\pi,\pi]$. Since $\hat{\Omega}_0^{-1}\hat{\Omega}_j$ is bounded for all $\Omega_j \in (-2,2]$, $j = 1,\ldots,d$ but not all zero in the same time, we obtain the desired result.

5. Application: scalar wave equation

5.1. Semi-discrete problem

The system (2) is discretized assuming, for simplicity, that the grid spacings are equal ($h_1 = \cdots = h_d = h$). The case $h_i \neq h_j$ for some directions i and j does not introduce further complications in the following analysis.

We construct the semi-discrete system corresponding to (2) by:

$$\begin{aligned}
\frac{d}{dt}\Phi &= \sum_{j=1}^{d}\beta^j D_j^{(1,n,s_j,\epsilon_j)}\Phi + K,\\
\frac{d}{dt}K &= \gamma^{jl}D_{jl}^{(2,n)}\Phi + \sum_{j=1}^{d}\beta^j D_j^{(1,n,s_j,\epsilon_j)}K.
\end{aligned} \tag{38}$$

This way of discretizing the first order derivative terms, which correspond to advection along the shift vector β^i, with off-centered derivatives has become customary in numerical relativity (see, e.g. [7,27,28]).

We define the shorthand quantity $\hat{\Delta}$ as

$$\hat{\Delta} \equiv \sqrt{-\gamma^{jl}\widehat{D}_{jl}^{(2,n)}} = h^{-1}\sqrt{\gamma^{jl}\hat{d}_j^{(1,n)}\hat{d}_j^{(1,n)} + \sum_{j=1}^{d}\gamma^{jj}\hat{r}_j^{(n)}}.$$

Then the discrete symbol, the diagonalizing matrix and the eigenvalues can be written as

$$\widehat{P}' = \begin{pmatrix} \beta^j\widehat{D}_j^{(1,n,s,\epsilon)} & 1 \\ -\hat{\Delta}^2 & \beta^j\widehat{D}_j^{(1,n,s,\epsilon)} \end{pmatrix}, \tag{39}$$

$$\widehat{T}^{-1} = \begin{pmatrix} i\hat{\Delta} & 1 \\ -i\hat{\Delta} & 1 \end{pmatrix}, \quad \hat{\Lambda}_{\pm} = \sum_{j=1}^{d}\beta^j\widehat{D}_j^{(1,n,s_j,\epsilon_j)} \pm i\hat{\Delta}. \tag{40}$$

[3] For the classical fourth order Runge–Kutta, $R_{lsia} = \sqrt{8} = 2.83$ and $R_{ls} = 2.61$.

Because $\hat{r}_j^{(n)} \geqslant 0$, according to (17), and the matrix γ^{jl} is positive definite, the quantity $\hat{\Delta}$ is real and $\hat{\Delta} \geqslant 0$ with equality only when all ξ_j are zero. Thus

$$\widehat{H} \equiv \frac{1}{2}\widehat{T}^{-1\dagger}\widehat{T}^{-1} = \begin{pmatrix} \hat{\Delta}^2 & 0 \\ 0 & 1 \end{pmatrix}$$

is a symmetrizer for the system (38). We observe that the symmetrizer does not depend on the diagonal entries of the symbol $\widehat{P}'$, e.g. does not depend on the way we advect the shift terms.

We still have to prove that there exists a constant $K \geqslant 1$ such that

$$K^{-1}\Omega_0^2 \leqslant \hat{\Delta}^2 \leqslant K\Omega_0^2. \tag{41}$$

The positivity of the matrix γ^{jl} implies the existence of a constant $c_1 > 0$ such that

$$c_1 \leqslant \min \gamma^{jj} \quad \text{and} \quad \gamma^{jl}y_j y_l \geqslant c_1|y|^2, \quad \forall y = (y_1, \ldots, y_d) \in \mathbb{R}^{\mathbb{d}}. \tag{42}$$

Furthermore, because $|\gamma^{jl}| < \infty$ there also exists a constant $c_2 > 0$ such that

$$c_2 \geqslant \max \gamma^{jj} \quad \text{and} \quad \gamma^{jl}y_j y_l \leqslant c_2|y|^2, \quad \forall y = (y_1, \ldots, y_d) \in \mathbb{R}^{\mathbb{d}}. \tag{43}$$

Using (42) and the inequalities (20) we obtain

$$h^2\hat{\Delta}^2 \geqslant (\min \gamma^{jj})\sum_{j=1}^{d}\hat{r}_j^{(n)} + c_1\sum_{j=1}^{d}(\hat{d}_j^{(1,n)})^2 \geqslant c_1\sum_{j=1}^{d}\hat{d}_j^{(2,n)} \geqslant c_1\hat{\Omega}_0^2.$$

On the other hand, by (43) and again (20) we have that

$$h^2\hat{\Delta}^2 \leqslant (\max \gamma^{jj})\sum_{j=1}^{d}\hat{r}_j^{(n)} + c_2\sum_{j=1}^{d}(\hat{d}_j^{(1,n)})^2 \leqslant c_2\sum_{j=1}^{d}\hat{d}_j^{(2,n)} \leqslant c_2C_n\hat{\Omega}_0^2.$$

We chose $K = \max\{c_1^{-1}, (c_2C_n), 1\}$ and obtain the relation (41).

The conserved discrete quantity in physical space associated to $\widehat{H}$, i.e. the norm $\|v\|_{h,H}$ defined in (34), is

$$\|v\|_{h,H}^2 = \frac{1}{h^2}\left[\sum_{j=1}^{d}\gamma^{jj}\sum_{k=1}^{n}|d_{k-1}|\|(hD_{+j})^k\Phi\|_h^2 + \sum_{j\neq l}\gamma^{jl}\|hD_l^{(1,n)}\Phi\|_h^2\right] + \|K\|_h^2,$$

where $v = (\Phi^T, K^T)^T$. Having proved the existence of a symmetrizer we have proved that the semi-discrete problem is stable with respect to the norms D_+ and H. Note again that the stability property does in particular not depend on how the shift terms are discretized.

5.2. Courant limits and the role of dissipation

In order for the fully-discrete problem to be stable we impose the non-positivity condition on the real part of the eigenvalues and restrict the Courant factor λ according to the inequality (35):

$$Re(\hat{\Lambda}_\pm) \leqslant 0, \tag{44}$$

$$\lambda \leqslant \frac{\alpha_0}{\max_{\xi\in S_{\vec{\xi}}}\left|h\hat{\Lambda}_\pm\right|}. \tag{45}$$

From (40) we have

$$hRe(\hat{\Lambda}_\pm) = \sum_{j=1}^{d}\beta^j\epsilon_j\hat{\mathbf{d}}_j^{(1,n,s_j)}. \tag{46}$$

The relation (44) has to hold for all frequencies $\xi_j \in (-\pi, \pi)$. Because each term j in the sum (46) can be canceled individually at $\xi_j = 0$, the non-positivity condition has to applied for each term. The problem reduces to the study of the one-dimensional case,

$$\beta\epsilon\hat{\mathbf{d}}^{(1,n,s)}(\xi) \leqslant 0 \quad \text{with} \quad \xi \in (-\pi, \pi]. \tag{47}$$

Because $\hat{\mathbf{d}}^{(1,n,s)}$ is zero for $s = 0$, negative for $s = 1$ and changes sign for $s \geqslant 2$, it is clear that the condition holds for centered and one-point upwinded ($\epsilon = \text{sign}\beta$) schemes and is violated in all the other cases.

However, the condition can be reestablished if appropriate artificial dissipation is added to the system. By using the Kreiss–Oliger dissipation operator (24), the condition (47) changes to

Table 2
Formulas for the dissipation parameters $\bar{\sigma}_{\pm}^{(n,s)}$ when $s = 1, 2, 3$. The quantity $2^{2(n+1)}|\beta^j|\bar{\sigma}_{\pm}^{(n,s)}$ ($\pm$ stands for upwind/downwind) represents the minimum dissipation that one has to add to make the numerical scheme stable.

	$\bar{\sigma}_{+}^{(n,s)}$	$\bar{\sigma}_{-}^{(n,s)}$
$s = 1$	$-\frac{1}{2C_{2n}^{n+1}}\frac{1}{n+1}$	$\frac{1}{2C_{2n}^{n+1}}\frac{1}{n+1}$
$s = 2$	$\frac{1}{2C_{2n}^{n+2}}\frac{2}{n+1}$	$\frac{1}{2C_{2n}^{n+2}}\frac{2n}{n^2+3n+2}$
$s = 3$	$\frac{1}{2C_{2n}^{n+3}}\frac{n(n+4)}{(n+1)(n+2)^2}$	$\frac{1}{2C_{2n}^{n+3}}\frac{3}{n+1}$

$$\beta\epsilon\hat{\mathbf{d}}^{(1,n,s)}(\xi) - \frac{1}{2^{2(n+1)}}\sigma\hat{\Omega}^{2(n+1)}(\xi) \leqslant 0 \quad \text{with} \quad \xi \in (-\pi, \pi] \tag{48}$$

This imposes a lower limit on the dissipation parameter σ:

$$\sigma \geqslant \sigma_{\min}(\beta, n, s, \epsilon) = \begin{cases} 2^{2(n+1)}|\beta|\bar{\sigma}_{+}^{(n,s)}, & \epsilon = \text{sign}\beta \text{ (upwind)}, \\ 2^{2(n+1)}|\beta|\bar{\sigma}_{-}^{(n,s)}, & \epsilon = -\text{sign}\beta \text{ (downwind)}, \end{cases} \tag{49}$$

where we have denoted

$$\bar{\sigma}_{+}^{(n,s)} \equiv \max_{\hat{\Omega}\in(0,2]} \frac{\hat{\mathbf{d}}^{(1,n,s)}}{\hat{\Omega}^{2(n+1)}}, \quad \bar{\sigma}_{-}^{(n,s)} \equiv -\min_{\hat{\Omega}\in(0,2]} \frac{\hat{\mathbf{d}}^{(1,n,s)}}{\hat{\Omega}^{2(n+1)}} \tag{50}$$

and used the fact that $\hat{\mathbf{d}}^{(1,n,s)}$ is, according to (17), a sum over powers of $\hat{\Omega}$.

In Table 2 we give the formulas for $\bar{\sigma}_{\pm}^{(n,s)}$ for $s = 1, 2, 3$. We remark that for all $n \geqslant 1$, $\bar{\sigma}_{+}^{(n,1)} < 0$, $\bar{\sigma}_{-}^{(n,1)} > 0$ and $\bar{\sigma}_{\pm}^{(n,s)} > 0$ for all $s > 1$.

This means that when using one-point upwinded stencils, we can add "negative" dissipation and still obtain a stable scheme. In fact, the following situations are equivalent:

- Upwind one point and add dissipation with $\sigma = 2^{2(n+1)}|\beta|\bar{\sigma}_{+}^{(n,1)} < 0$.
- Downwind one point and add dissipation with $\sigma = 2^{2(n+1)}|\beta|\bar{\sigma}_{-}^{(n,1)} > 0$.
- Use the CFDO operator $1/2(D^{(1,n,s,1)} + D^{(1,n,s,-1)})$ (constructed with $2(n+s)+1$ points), and do not add dissipation, $\sigma = 0$.

In any of the above three situations, the real part of the eigenvalues is zero, so if the Courant limit is small enough then we obtain stability on the imaginary axis. Now, coming back to the d-dimensional case, it is obvious that the dissipation parameters in (24), σ_j, have to be chosen according to the value of the shift and the type of off-centering in the j-direction ($\sigma_j \geqslant \sigma_{\min}(\beta_j, n, s_j, \epsilon_j)$). It can shown that, by choosing exactly $\sigma_j = \sigma_{\min}(\beta_j, n, s_j, \epsilon_j)$ the Courant limit is maximized. However, to compute it explicitly, (as a function of shifts, order of approximation, and off-centerings) is not easy in the general case.

In the particular case of a flat $d+1$-metric with zero shift, the Courant limit is easy to write down:

$$\lambda \leqslant \frac{\alpha_0}{2\sqrt{dC_n}},$$

where C_n is given in (20) and $\alpha_0 = R_{lsia}$. In the general case, the Courant limit has to be evaluated numerically.

For the 1-D wave equation with shift $\beta > 0$ with upwind discretization of the advection term and adding the minimal amount of dissipation if necessary, the limit of the Courant factor is given by

$$\lambda^{(n,s)}(\beta) \equiv \frac{\alpha_0}{\max_{\xi\in(\pi,\pi]}\left|\beta\hat{\Omega}^{2(n+1)}\left(\frac{\hat{\mathbf{d}}^{(1,n,s)}}{\hat{\Omega}^{2(n+1)}} - \sigma_{+}^{(n,s)}\right) + i\left(\beta\hat{d}^{(1,n,s)} + \sqrt{\hat{d}^{(2,n)}}\right)\right|}.$$

We compare the Courant limits for different orders of approximations at fixed advection stencil in Fig. 3, and the Courant limit at fixed order of approximation for different advection stencils in Fig. 4.

In Fig. 3 we see that if $s = 0$, the higher the order of approximation, the lower the Courant limit. For $s \geqslant 1$, this is not true anymore beyond a certain value of the shift. For large shifts we observe that increasing the order of approximation, actually *decreases* the Courant limit.

Comparing different stencils in Fig. 4, we observe that advecting more points decreases the Courant limit, and there is a significant drop in the Courant factor between $s = 1$ and $s = 2$, for all orders of approximation.

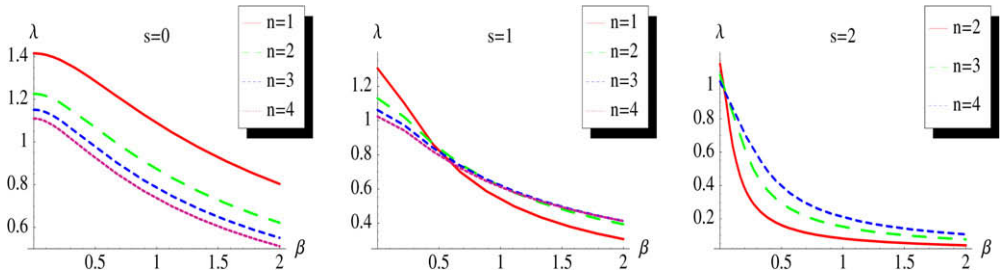

Fig. 3. Courant limit as a function of β, for different orders of approximation at fixed off-centering s. For $s = 0$ (left) no dissipation is needed, ($\sigma = 0$), and we are in the regime of local stability on the imaginary axis ($\alpha_0 = 2.83$). For $s = 1$ (middle), again no dissipation is needed ($\sigma = 0$), but now we are in the regime of local stability ($\alpha_0 = 2.61$). For $s = 2$ (right) dissipation is required and we add the minimum amount in order to attain stability.

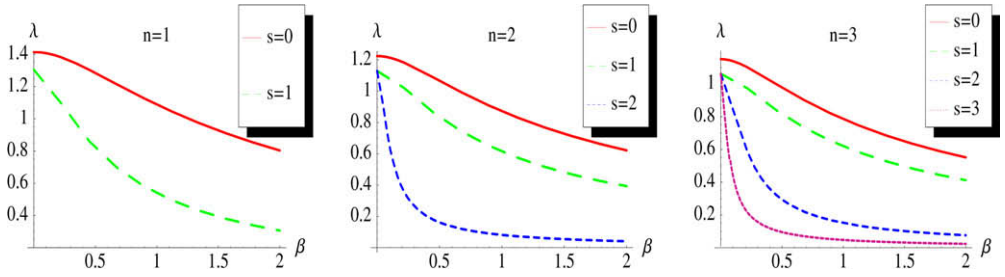

Fig. 4. Courant limit as a function of β for different advection stencils at fixed order of spatial accuracy. From left to right: Courant limits at approximation orders 2, 4, 6. As in Fig. 3 the Courant limit calculation takes into account whether we are in the regime of local stability on the imaginary axis (the case $s = 0$), or only local stability (for $s \geqslant 1$), and the minimal amount of Kreiss–Oliger dissipation is added for $s \geqslant 2$.

5.3. Phase and group speeds

For the wave equation the continuum phase and group speeds are:

$$\hat{v}_p = \hat{v}_g = \beta^n \pm \sqrt{\gamma^{nn}}, \quad \text{where} \quad \hat{v}_p \equiv \frac{\hat{\Lambda}}{\omega_0}, \quad \hat{v}_g \equiv n^j \frac{d}{d\omega_j}\hat{\Lambda},$$

where $\beta^n = \beta^j n_j$ and $\gamma^{nn} = \gamma^{jl} n_j n_l$. The discrete speeds can be defined in a similar manner from the discrete eigenvalues $\hat{\Lambda}$. This would lead to complex speeds in case of off-centered schemes which are difficult to investigate [29]. In the following, we assume that the real part of the discrete eigenvalues has been canceled by adding appropriate artificial dissipation terms (positive or negative). Otherwise, our investigation remains valid only at small frequencies, where the damping/amplification effect introduced by the real part is dominated by the dispersion effect associated to the imaginary part of the eigenvalues. With these remarks, we define the discrete speeds by:

$$\hat{v}_p \equiv \frac{\hat{\Lambda}^{lm}}{\omega_0} = \left(\sum_{j=1}^{d} \beta^j n_j \frac{\hat{d}_j^{(1,n,s_j)}}{\xi_j}\right) \pm \frac{(h\hat{\Delta})(\xi)}{\xi_0},$$

$$\hat{v}_g \equiv n^j \frac{d}{d\xi_j}(h\hat{\Lambda}^{lm}) = \sum_{j=1}^{d} \left[\beta^j n_j^2 \frac{\partial \hat{d}_j^{(1,n,s_j)}}{\partial \xi_j} \pm \frac{\partial (h\hat{\Delta})(\xi)}{\partial \xi_j} n_j\right].$$

We also restrict attention to the one-dimensional case. Because $\pm$ speeds interchange when ξ changes sign, it is enough to consider only the "+" speed over the whole spectrum $\xi \in (-\pi, \pi]$. Also because we will compare speeds at different orders of approximation or at different stencils, we attach the superscript (n,s) (or only (n) in case $s = 0$), to the symbols representing the discrete speeds and the corresponding errors:

$$\hat{v}_p^{(n,s)}(\xi) = \frac{1}{\xi}\left(\beta\hat{d}^{(1,n,s)} + \sqrt{\hat{d}^{(2,n)}}\right),$$

$$\hat{v}_g^{(n,s)}(\xi) = \frac{d}{d\xi}\left(\beta\hat{d}^{(1,n,s)} + \sqrt{\hat{d}^{(2,n)}}\right).$$

The continuum limits for both, phase and group speeds are $\beta + 1$ for $\xi > 0$ and $\beta - 1$ for $\xi < 0$. We will analyze the behavior of the speed errors defined as

$$\hat{\epsilon}_p^{(n,s)} \equiv \beta\left(\frac{\hat{d}^{(1,n,s)}}{\xi} - 1\right) + \left(\frac{\sqrt{\hat{d}^{(2,n)}}}{\xi} - \text{sign}\xi\right), \tag{51}$$

$$\hat{\epsilon}_g^{(n,s)} \equiv \beta\left(\frac{d}{d\xi}\hat{d}^{(1,n,s)} - 1\right) + \left(\frac{d}{d\xi}\sqrt{\hat{d}^{(2,n)}} - \text{sign}\xi\right). \tag{52}$$

We will also assume $\beta \geqslant 0$ without restricting generality, if $\beta \to -\beta$, then $\hat{\epsilon}_{p,g}^{(n,s)}(\xi) \to -\hat{\epsilon}_{p,g}^{(n,s)}(-\xi)$.

5.3.1. Small frequencies

When $\xi \simeq 0$ one can show that the phase and group speed errors satisfy

$$\hat{\epsilon}_p^{(n,s)} = -|c_n|\left[(-1)^s\frac{(n+s)!(n-s)!}{(n!)^2}\beta + \frac{\text{sign}\xi}{2(n+1)}\right]\xi^{2n} + O(\xi^{2n+2}),$$

$$\hat{\epsilon}_g^{(n,s)} = -(2n+1)|c_n|\left[(-1)^s\frac{(n+s)!(n-s)!}{(n!)^2}\beta + \frac{\text{sign}\xi}{2(n+1)}\right]\xi^{2n} + O(\xi^{2n+2}).$$

Because the errors scale with ξ^{2n}, it is obvious that for small enough frequencies higher order approximations will improve the phase and group errors for all the values of the shift and for all advection stencils.

If we keep the order fixed and compare the speeds corresponding to an off-centering by $s \geqslant 1$-points with the ones corresponding to the centered scheme, $s = 0$, then one can easily show that the off-centered scheme improves over the centered one

- the "+" numerical speeds ($\xi > 0$) if s is odd and β is small enough;
- the "−" numerical speeds ($\xi < 0$) if s is even and β is small enough where small enough means

$$\beta < \frac{1}{(n+1)}\frac{1}{\frac{(n+s)!(n-s)!}{(n!)^2} - 1}.$$

Obs. For $s = 1$, the above inequality becomes $\beta < \frac{n}{(n+1)}$. Also notice that with increasing s the above limit on β decreases.

In the next subsection we will analyze the behavior for the whole spectrum in some more detail.

5.3.2. Comparison with wave equation written in first order form

If the wave equation is written in first order form (approximating the first derivatives with the corresponding CFDO), then the eigenvalues become $(h\hat{\Lambda}_\pm)(\xi) = i(\beta \pm 1)\hat{d}^{(1,n)}$. For $\xi \simeq 0$ one then gets

$$\hat{\epsilon}_p^{(n)} = -(\beta + \text{sign}\xi)|c_n|\xi^{2n} + O(\xi^{2n+2}),$$

$$\hat{\epsilon}_g^{(n)} = -(\beta + \text{sign}\xi)(2n+1)|c_n|\xi^{2n} + O(\xi^{2n+2}).$$

We notice that for a given order, the second order system discretized with CFDO, has smaller phase and group errors then the first order one (for both eigenvalues), if and only if $|\beta| \leqslant \frac{2n+3}{4(n+1)}$. If $|\beta|$ is not in this interval then one pair of speeds (phase and group) is better approximated by the second order system, while the other one is better approximated by the first order system.

5.3.3. Scaling of the speed errors with the order of approximation

Lemma 5.1. *If $\beta = 0$, then increasing the order of approximation decreases the phase and group speed errors for all frequencies.*

To prove this we make use of the relations (21) in the definitions of the speeds and obtain

$$\hat{v}_p^{(n)} = \frac{\sqrt{\hat{d}^{(2,n)}}}{\xi}, \quad \hat{\epsilon}_p^{(n)} = \frac{\sqrt{\hat{d}^{(2,n)}}}{\xi} - \text{sign}\xi,$$

$$\hat{v}_g^{(n)} = \frac{\hat{d}^{(1,n)}}{\sqrt{\hat{d}^{(2,n)}}}, \quad \hat{\epsilon}_g^{(n)} = \frac{\hat{d}^{(1,n)}}{\sqrt{\hat{d}^{(2,n)}}} - \text{sign}\xi.$$

Using the inequalities (18) and (19) one can easily show that $|\hat{\epsilon}_p^{(n+1)}| < |\hat{\epsilon}_p^{(n)}|$ and $|\hat{\epsilon}_g^{(n+1)}| < |\hat{\epsilon}_g^{(n)}|$ for all frequencies. The situation is illustrated in Fig. 5 where we plot the speeds $\hat{v}_p^{(n)}$ and $\hat{v}_g^{(n)}$ versus ξ.

If $\beta \neq 0$ then it is not true anymore that higher order approximations improve the numerical speeds for all frequencies (not even for the case of using CFDO). Though one can go into details and determine the regions in the spectrum where the scaling with order fails, we restrict ourselves to illustrating this situation by plotting the numerical speeds versus frequency at a particular value of the shift. In Fig. 6 we show the numerical speeds at different orders of approximation with the same advection stencil when $\beta = 0.5$.

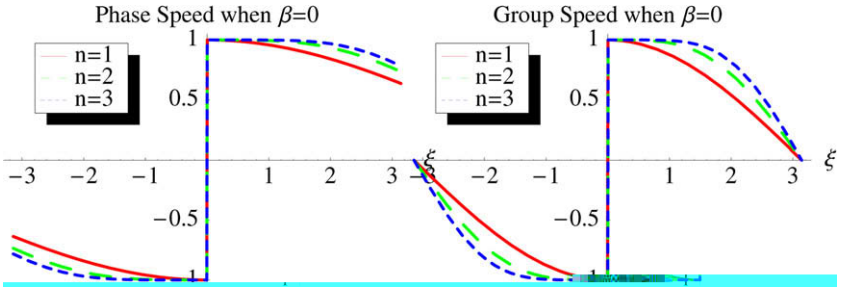

Fig. 5. Phase and group speeds for $\beta = 0$. The higher the order of the approximation, the more accurate the phase and group speeds are for all frequencies.

5.3.4. Scaling of the speed errors with off-centering

The next question we want to answer is what happens with the numerical speed errors, if we keep the order of approximation fixed and vary the off-centering of the first derivative. For example, we illustrate this situation in Fig. 7, when $\beta = 0.5$. As one can see already in these plots, although we know that off-centering increases the error of the finite difference operator, it is not necessary that the numerical speeds will follow the same pattern, e.g. for this value of the shift, the "+" speed seems more accurate with $s = 1$ than with $s = 0$ for all the spectrum. In the following we will determine the regions in the (ξ, β)-plane where off-centering improves the numerical speed errors over the centered scheme.

5.3.4.1. Phase speeds. Imposing $|\hat{\epsilon}^{(n,s)}_{p\pm}| < |\hat{\epsilon}^{(n,0)}_{p\pm}|$ and using the definition (51) yields the inequality

$$f_1^{(n,s)}(\xi) f_2^{(n,s)}(\xi)\left(\beta - \beta^{(n,s)}(\xi)\right) < 0, \tag{53}$$

where

$$\begin{aligned} f_1^{(n,s)}(\xi) &\equiv \hat{d}^{(1,n,s)} - \hat{d}^{(1,n,0)}, \\ f_2^{(n,s)}(\xi) &\equiv \hat{d}^{(1,n,s)} + \hat{d}^{(1,n,0)} - 2\xi, \\ g^{(n)}(\xi) &\equiv 2\left(|\xi| - \sqrt{\hat{d}^{(2,n)}}\right), \\ \beta^{(n,s)}(\xi) &\equiv \frac{g^{(n)}(\xi)}{f_2^{(n,s)}(\xi)}. \end{aligned} \tag{54}$$

We have $g^{(n)}(\xi) > 0$ but $f_{1,2}^{(n,s)}$ can change sign over the spectrum. The inequality (53) holds at a given frequency ξ, if $\beta > \beta^{(n,s)}(\xi)$ and $\text{sign} f_{1,2}^{(n,s)}(\xi) < 0$ or $\beta < \beta^{(n,s)}(\xi)$ and $\text{sign} f_{1,2}^{(n,s)}(\xi) > 0$. In general, the regions in (ξ, β) plane where at fixed order of approximation, off-centering by s points improves the accuracy of the phase speed, are difficult to determine analytically

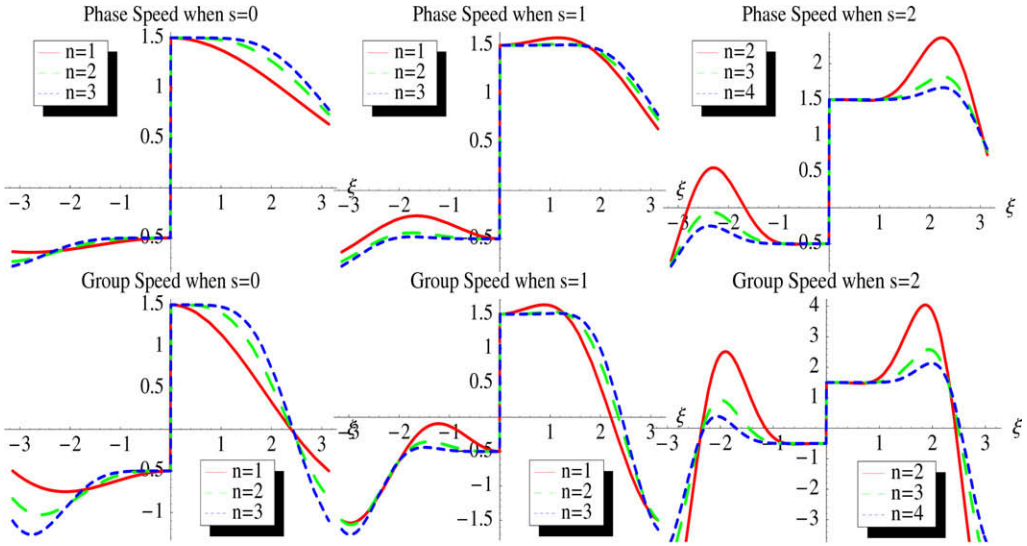

Fig. 6. The phase and group speeds at different orders of approximation but keeping the level of off-centering fixed, for $\beta = 0.5$.

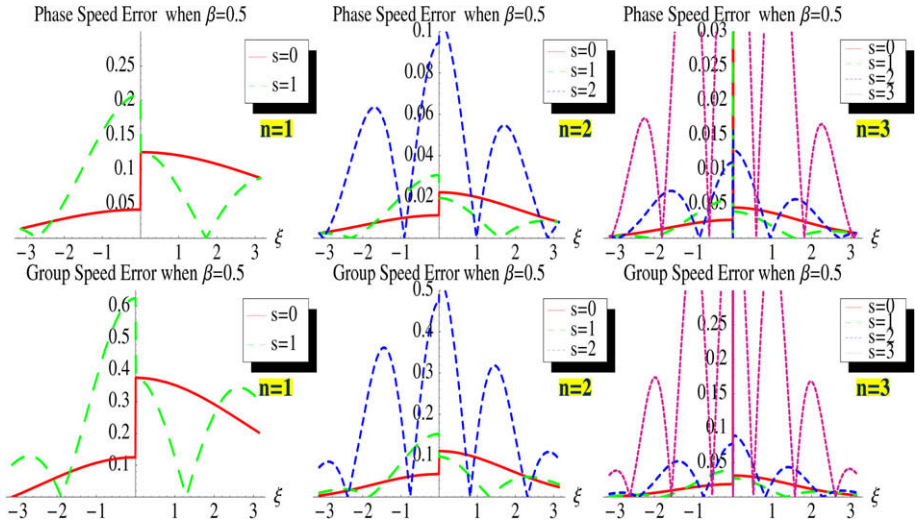

Fig. 7. The phase and group speeds errors (scaled with ξ^{2n}) are shown at different advection stencils for $\beta = 0.5$.

and we restrict ourselves to a numerical evaluation (Fig. 8). What we see in the plots is that if s is odd (even) then for sufficiently small β, the "+" ("−") speed has smaller error compared with the case of CFDO in some intervals of the spectrum that include the small frequency range. However these regions become narrower with increasing the off-centering, such that for $s = 1$ we have the strongest effect. We analyze this case in more detail below.

If $\boldsymbol{s} = 1$, the functions $f_{1,2}^{(n,s)}$ and $\beta^{(n,s)}$, defined in (54) become

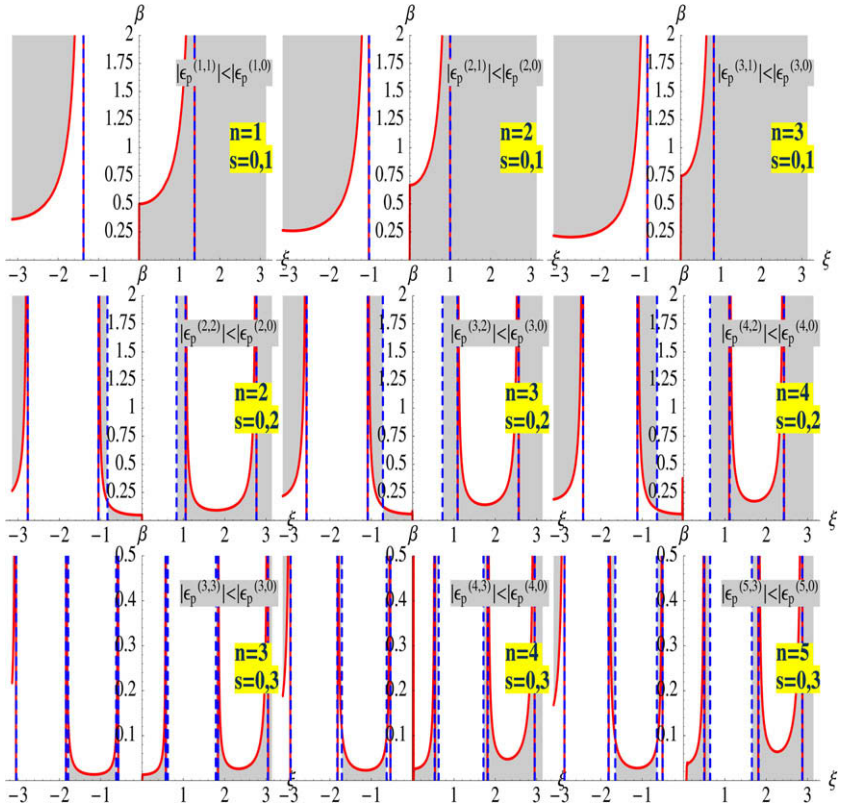

Fig. 8. Shown are the regions where advected stencils improve the phase speed error over the centered scheme. The regions are delimited by the quantity $\beta^{(n,s)}$ and the zeros of the function $f_1^{(n,s)}$ as defined in (54).

$$
\begin{aligned}
f_1^{(n,1)}(\xi) &= \frac{|c_{n-1}|}{2}(\sin\xi)\hat{\Omega}^{2n},\\
f_2^{(n,1)}(\xi) &= \hat{\delta}\frac{|c_{n-1}|}{2}\hat{\Omega}^{2n} + 2(\hat{d}^{(1,n)} - \xi),\\
\beta^{(n,1)}(\xi) &= \frac{2\left(|\xi| - \sqrt{\hat{d}^{(2,n)}}\right)}{f_2^{(n,1)}(\xi)}.
\end{aligned}
\tag{55}
$$

We have $\operatorname{sign} f_1^{(n,1)}(\xi) = \operatorname{sign}\xi$ and $f_1^{(n,1)}(\pm\pi) = 0$.

Then the inequality $|\hat{\epsilon}_p^{(n,1)}| < |\hat{\epsilon}_p^{(n,0)}|$ holds

- for $\xi > 0$ if $\beta^{(n,1)}(\xi) < 0$ or $0 < \beta < \beta^{(n,1)}(\xi)$,
- for $\xi < 0$ if $\beta > \beta^{(n,1)}(\xi) > 0$.

The limits of $\beta^{(n,1)}(\xi)$ in 0 and π are

$$
\begin{aligned}
&\lim_{\xi \searrow 0} \beta^{(n,1)}(\xi) = -\lim_{\xi \nearrow 0} \beta^{(n,1)}(\xi) = \frac{n}{n+1},\\
&\beta^{(n,1)}(\pi) = -1 + 2\frac{\sqrt{C_n}}{\pi} < 0.
\end{aligned}
$$

It can be shown that the equation $\beta = \beta^{(n,1)}(\xi)$ has at most one solution in each of the branches $\xi > 0$ and $\xi < 0$, that we will denote by $\xi^{\pm}$. It turns out that $|\hat{\epsilon}_p^{(n,1)}| < |\hat{\epsilon}_p^{(n,0)}|$ holds if

- $\beta < 1 - 2\frac{\sqrt{C_n}}{\pi}$ and $\xi \in (0, \pi)$,
- $1 - 2\frac{\sqrt{C_n}}{\pi} < \beta < \frac{n}{n+1}$ and $\xi \in (-\pi, \xi^-) \cup (0, \pi)$,
- $\beta > \frac{n}{n+1}$ and $\xi \in (-\pi, \xi^-) \cup (\xi^+, \pi)$.

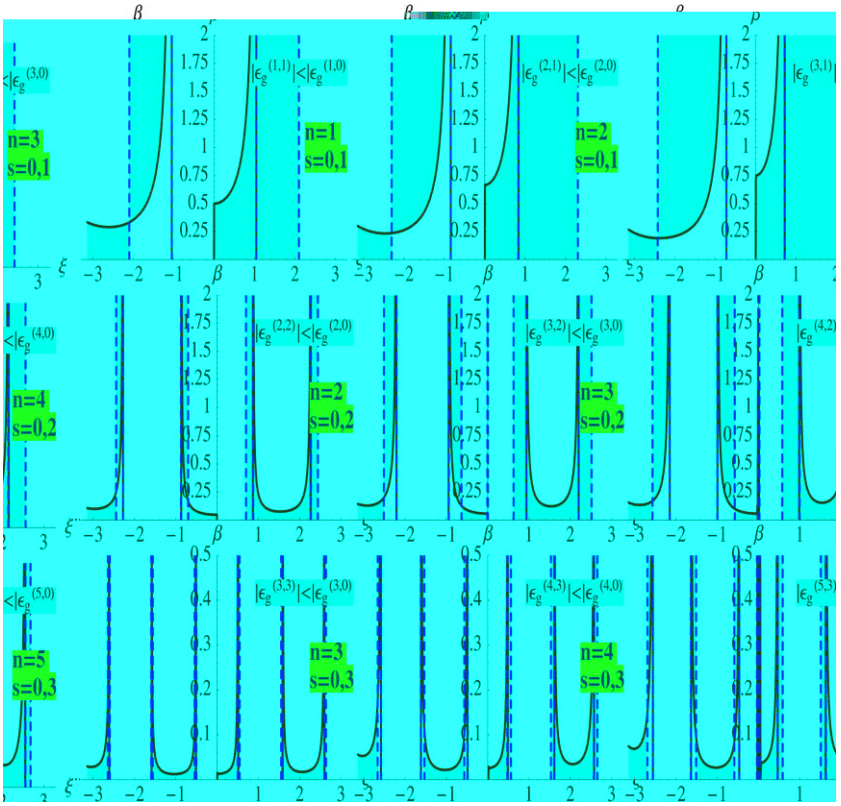

Fig. 9. Shown are the regions where advected stencils improve the group speed error over the centered scheme. The regions are delimited by the quantity $\beta^{(n,s)}$ and the zeros of the function $F_1^{(n,s)}$ defined in (56).

At a given order of approximation, $2n$, for sufficiently small $\beta < \frac{n}{n+1}$ the "+" speed has smaller error in the case when we advect one point than in the case when we use CFDO, for all frequencies $0 < \xi \leqslant \pi$, but the "−" speed will have larger error, at least for small and mid frequencies.

If $\beta > \frac{n}{n+1}$ then for both ± speeds, in the regime of small frequencies, the CFDO give less error than one-point advected scheme, while for mid and high frequencies the situation reverses. The interval of small frequencies where CFDO are better than advected scheme shrinks with increasing the order of approximation.

5.3.4.2. Group speeds. Imposing $|\hat{\epsilon}_g^{(n,s)}| < |\hat{\epsilon}_g^{(n,0)}|$ and using the definition (52) yields the inequality

$$F_1^{(n,s)}(\xi)F_2^{(n,s)}(\xi)\Big(\beta - \beta^{(n,s)}(\xi)\Big) < 0,$$

where

$$\begin{aligned}
&F_1^{(n,s)}(\xi) \equiv \partial_\xi f_1^{(n,s)}(\xi),\\
&F_2^{(n,s)}(\xi) \equiv \partial_\xi f_2^{(n,s)}(\xi),\\
&G^{(n)}(\xi) \equiv \partial_\xi g^{(n)}(\xi),\\
&\beta^{(n,s)}(\xi) \equiv \frac{G^{(n)}(\xi)}{F_2^{(n,s)}(\xi)},
\end{aligned} \tag{56}$$

and $f_{1,2}^{(n,s)}$ and $g^{(n)}$ are given by (54). It is easy to see that $G^{(n)}(\xi) = -G^{(n)}(-\xi)$. However the signs of $F_{1,2}^{(n,s)}(\xi)$ are more difficult to determine. As in the case of phase speeds analysis, we determine graphically (see Fig. 9) the regions in (ξ, β) plane where at fixed order of approximation, off-centering by s points improves the accuracy of the group speed. We see the same qualitative behavior as for the phase speeds, in the sense that for sufficiently small β, the "+" ("−") speed has smaller error compared with the case of CFDO at least at small frequencies, and off-centering decreases the extent of these regions in (ξ, β)-space.

In case $s = 1$, the relations (56) become

$$\begin{aligned}
F_1^{(n,1)}(\xi) &= \frac{(n+1)|c_{n-1}|}{2}\left(\frac{n}{n+1} + \cos\xi\right)\hat{\Omega}^{2n}\\
F_2^{(n,1)}(\xi) &= \frac{(n+1)|c_{n-1}|}{2}\left(-\frac{n}{n+1} + \cos\xi\right)\hat{\Omega}^{2n},\\
\beta^{(n,1)}(\xi) &= \frac{2\Big(\operatorname{sign}\xi - \hat{d}^{(1,n)}/\sqrt{\hat{d}^{(2,n)}}\Big)}{F_2^{(n,1)}(\xi)}.
\end{aligned} \tag{57}$$

By analyzing the monotony of these functions using the properties from Section 3.2, the following result can be formulated.

At a given order of approximation $2n$, for sufficiently small $\beta < \frac{n}{n+1}$, the "+" group speed has smaller error in the case when we advect one point than in the case when we use CFDO for all frequencies $0 < \xi < \pi - \arccos\frac{n}{n+1}$, (in the case of phase speed this was the whole range $(0, \pi)$!), but the "−" speed will have larger error, at least for small and mid frequencies.

If $\beta > \frac{n}{n+1}$ then for both ± speeds, in the regime of small frequencies, the CFDO give less error than one-point advected scheme, while for mid and high frequencies, the situation reverses. The interval of small frequencies where CFDO are better than advected scheme narrows with increasing the order of approximation.

5.3.5. Centered versus one-point upwinded scheme, numerically

In Section 5.3.4 we showed that when $0 < \beta \leqslant \frac{n}{n+1}$ the numerical "+" speeds are better approximated with one-point off-centered schemes than with centered schemes at least up to very high frequencies in the grid. In this section we show some simple numerical tests to illustrate this fact. We chose l-periodic initial data:

$$\begin{aligned}
&\Phi(0,x) = e^{-(2\pi l\tau^2)^{-1}\sin^2\left(\frac{\pi}{l}x - \frac{\pi}{2}\right)},\\
&K(0,x) = a\partial_x\Phi(0,x), \quad x \in [0,l).
\end{aligned} \tag{58}$$

The parameter $a \in [-1, 1]$ sets the amplitude of the "±" components,

$$C_\pm = (a \pm 1)\partial_x\Phi.$$

When $a = 1(-1)$ the signal is purely "left" ("right") going and when $a = 0$, the signal is equally distributed between both modes.

We choose a grid with $N = 101$ points and resolution $h = 0.01$, the width of the grid is $l = Nh = 1.01$. Also we chose $\tau = 0.1$, $\beta = 0.5$ and we integrate the wave equation using fourth order FDOs for space derivatives and the fourth order Runge–Kutta as time integrator.

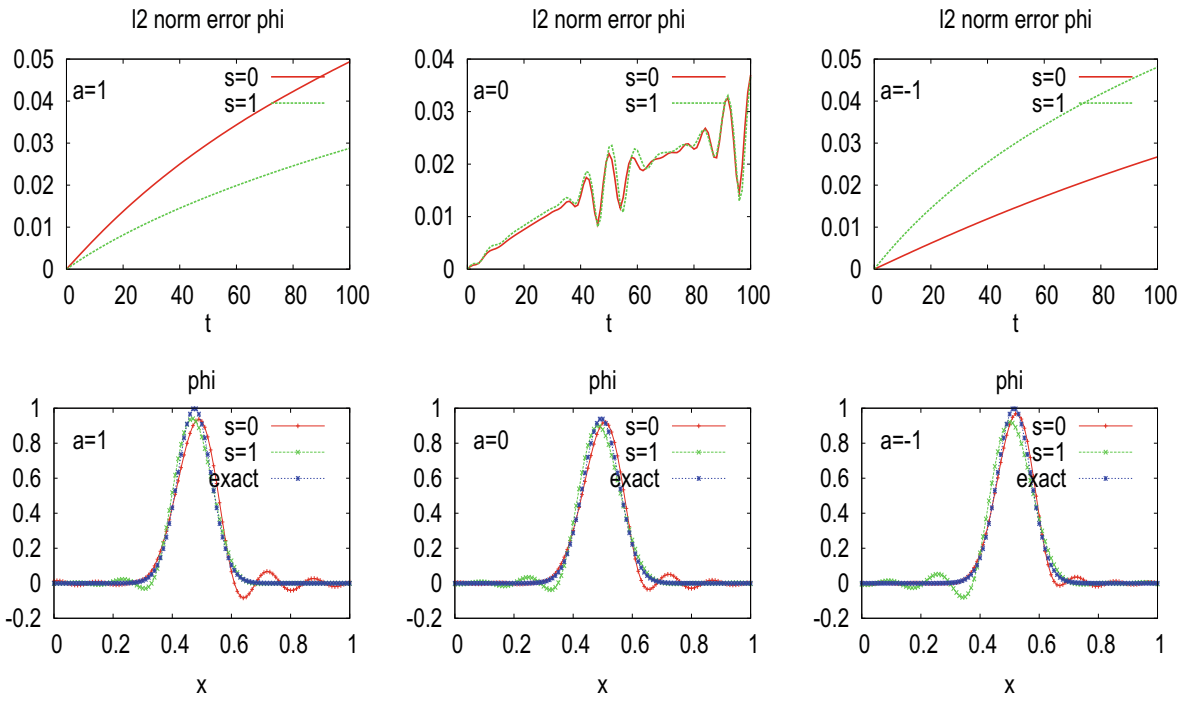

Fig. 10. The top panel shows the errors in the l^2 norm, the bottom panel snapshots of Φ at $t = 99CT$. The three cases from left to right are a purely left-going signal ($a = 1$), a signal with equal amplitudes of left and right going modes ($a = 0$), and a purely right-going signal ($a = -1$). Red lines mark the centered scheme, green the one-point advected stencil, blue the exact solution. For $a = 1$ upwinding is more accurate, but the situation reverses if the signal is right-going, for $a = 0$ both schemes yield similar accuracy. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

We let $a \in \{1, 0, -1\}$ and for each value of a we look at the errors for the main variables when $s = 0, 1$ (see Fig. 10). The numerical results show that, indeed, when the signal is "left" going, the upwinded scheme has less error than the centered scheme, while when the signal is going "right", the centered scheme is to be preferred.

6. Conclusions

In this paper we have investigated several aspects related to the discretization of the initial value problem for first order in time and second order in space systems of differential equations, using high order finite difference operators. Special attention has been paid to the situation when some of the first derivatives are approximated with off-centered discrete operators as is customary for treating black hole spacetimes in numerical relativity. Our investigation has been divided into three parts: (a) We started with an analysis of certain properties of the finite difference operators (Section 3). (b) Using these properties we have extended the validity of an existing stability method (Section 4). (c) We analyzed the stability and the numerical speeds in the case of the scalar wave equation (Section 5).

In the following we will give a brief overview of the results.

6.1. Analysis of first and second order discrete derivative operators

A set of mathematical properties have been deduced for the Fourier symbols associated with the second order centered and first order (not necessarily centered) discrete derivatives. They are in the form of inequalities, recursive and differential relations for the Fourier symbols at different orders of approximation or at different off-centerings. Here we mention two of them.

While first derivatives do not converge in the limit $n \to \infty$ at the maximum grid frequency ($\xi = \pi$), second derivatives do converge at all frequencies (that is the highest frequency in the grid will not be captured by the first order derivative, regardless of the order of approximation or the off-centering, while the second centered derivative can "see" it and approximates it better with increasing order).

For first order derivatives, increasing the off-centering (s) at a fixed order of approximation increases the error of the derivative at small frequencies. At larger frequencies this behavior changes, e.g. for $s = 1$, beyond a certain frequency $\xi^{(n)}$, the error is smaller than for the case $s = 0$. As a consequence, off-centering of the first order derivative in the case of the advection equation, increases the error at small frequencies, while at high frequencies, this situation can reverse.

6.2. Generalization of a stability analysis method

In [3], necessary and sufficient conditions for stability have been deduced assuming that (3) is discretized using 2nd or 4th order CFDOs and integrated in time using a time integrator locally stable on the imaginary axis. The validity of this stability method is extended here to $2n$-order spatial accuracy, including also the case where some derivatives are approximated with non-centered FDOs and dissipation is added to the system. It is pointed out that neither adding artificial dissipation nor shift advection terms affects the eigenvectors of the discrete symmetrizer, and thus the conditions for semi-discrete numerical stability. The Courant limit will of course be affected in general.

6.3. Application: scalar shifted wave equation

The stability method presented in Section 4 is applied to the case of the wave equation on a curved background in $d+1$ dimensions.

In the case of 1-D shifted wave equation in flat spacetime, the Courant limits and the numerical speeds have been analyzed in detail in respect to the order of approximation and off-centering of the first derivative.

6.3.1. Courant limits

Off-centerings by more than one point require dissipation for stability. In these cases, the minimal Kreiss–Oliger dissipation needed for stability has been computed and found to be proportional to the shift β. For centered schemes, higher order approximations have lower Courant limits. Interestingly, this does not hold for off-centered schemes (when adding just dissipation to be in the local stability regime) — for large enough shift, the Courant limit is actually larger for higher order schemes. Off-centering generally reduces the Courant limit drastically, except for at least fourth order accurate schemes, when only one-point off-centering is used: for higher than fourth order schemes one-point off-centering only leads to a minor reduction of the CFL factor.

6.3.2. Numerical speeds

Without shift, higher order approximations always result in more accurate numerical speeds, with nonzero shift this is not generally true at higher frequencies.

Although the truncation error for the first order derivative increases with the off-centering, the mixing with the second order discrete derivative in the scheme, causes upwinded stencils to give a higher overall accuracy in some situations.

More precisely, it is shown that advecting shift terms by an odd (even) number of points reduces the errors of the "+" ("−") numerical speeds in some intervals of the spectrum that include the small frequency range, if the shift is not too large. The extent of the regions in the (frequency, shift)-parameter space where this improvement appears, decreases with off-centering, in such a way that for $s=1$ one gets the strongest effect.

Thus, at a given order $2n$, if the shift satisfies $0<\beta<\frac{n}{n+1}$, then off-centering by one point has in comparison with the centered scheme, better "+" phase speed error for all frequencies, and better "+" group speed error for all frequencies up to a very high frequency in the grid, $\pi-\arccos\frac{n}{n+1}$.

If the wave equation is written in first order form and discretized using CFDO, then for a given order of approximation, the second order system discretized also with CFDO has smaller phase and group errors than the first order one, if $|\beta|\leqslant\frac{2n+3}{4(n+1)}$. If $|\beta|$ is not in this interval then one pair of speeds (phase and group) is better approximated by the second order system, while the other one is better approximated by the first order system.

A detailed understanding of finite difference algorithms for first order in time, second order in space systems, in particular as applied to the Einstein equations, will require significant further work. Already for the shifted wave equation, it will be interesting to study the errors in the multidimensional case, e.g. when the wave propagates in a direction that is not aligned with the grid. A similar analysis for the full Einstein equations will require a substantial use of computer algebra methods. We also point out that for the Einstein equations much of the complications come from the nonlinear source terms, which are beyond the scope of our present analysis.

Acknowledgments

M. Chirvasa would like to thank Bela Szilagyi for inspiring discussions at the early stages of this work, and we thank Jeffrey Winicour and Gerhard Zumbusch for helpful comments on our manuscript. S. Husa has been in supported in part as a VESF fellow of the European Gravitational Observatory (EGO), by DFG Grant SFB/Transregio 7 "Gravitational Wave Astronomy", by DAAD Grant D/07/13385 and Grant FPA-2007-60220 from the Spanish Ministerio de Educación y Ciencia.

Appendix A. Explicit expressions for finite difference operators

Explicit formulas for general finite difference operators in one dimension can be constructed in a surprisingly simple way by use of the in the following lemma.

Lemma A.1. *Consider a general finite difference operator $D^{(m,n,s,\epsilon)}$, which is written as a linear combination of shift operators of the form (4)*

$$D^{(m,n,s,\epsilon)} = h^{-m} \sum_{k=-n+\epsilon 0 s}^{n+\epsilon s} \tilde{f}_{m,n,s,\epsilon,k} S^k.$$

Then the coefficients $\tilde{f}_{m,n,s,\epsilon,k}$ are the coefficients of y^k in the Taylor expansion of the function

$$f^{m,n,s,\epsilon}(y) = y^{n-\epsilon s} (\ln y)^m$$

around the point $y_0 = 1$ up to the order $(y-y_0)^{2n}$. In general, the accuracy of this operator will be $2n+1-m$.

To prove the lemma, it is enough to consider the scalar function $\mathsf{v}(x) = e^{i\omega x}$ with $\omega \in \mathbb{C}$ and the associated grid function. By applying accordingly the differential and discrete operators we get:

$$\begin{aligned}
&\partial^m \mathsf{v}(x_v) = (i\omega)^m e^{i\omega h v},\\
&D^{(m,n,s,\epsilon)} v_v = h^{-m} \sum_{k=-n+\epsilon s}^{n+\epsilon s} \tilde{f}_{m,n,s,\epsilon,k} e^{i\omega h (v+k)},\\
&\partial^m \mathsf{v}(x_v) = D^{(m,n,s,\epsilon)} v_v + \mathcal{O}(h^q).
\end{aligned} \tag{59}$$

Introduce $y = e^{i\omega h}$ which gives $i\omega = h^{-1} \ln y$. The relations (59) lead to

$$y^{n-\epsilon s} (\ln y)^m = \sum_{k=-n+\epsilon s}^{n+\epsilon s} \tilde{f}_{m,n,s,\epsilon,k} y^{k+n-\epsilon s} + \mathcal{O}(h^{q+m}) y^{-v+n-\epsilon s}.$$

In the limit $h \to 0,\ y \to 1$. The function $y^{n-\epsilon s}(\ln y)^m$ is now Taylor expanded around the point $y_0 = 1$ up to $(y-y_0)^{2n}$ and the coefficients $\tilde{f}_{m,n,s,\epsilon,k}$ are identified. What remains is:

$$\mathcal{O}((y-1)^{2n+1}) = \mathcal{O}(h^{q+m}) y^{-v+n-\epsilon s}.$$

After replacing y with its definition and taking the limit $h \to 0$ we obtain $q = 2n+1-m$ and the lemma is proved.

Corollary A.2. *The FDOs associated with the first and second derivative are*

$$\begin{aligned}
&D^{(1,n,s,\epsilon)} = h^{-1} \sum_{k=-n+\epsilon s}^{n+\epsilon s} \alpha_{n,s,\epsilon,k} S^k,\\
&D^{(1,n)} \equiv D^{(1,n,0,0)} = h^{-1} \sum_{k=1}^{n} \frac{k\beta_{n,k}}{2} \left(S^k - S^{-k}\right),\\
&D^{(2,n)} \equiv D^{(2,n,0,0)} = h^{-2} \sum_{k=0}^{n} \beta_{n,k} \left(S^k + S^{-k}\right),
\end{aligned} \tag{60}$$

where

$$\alpha_{n,s,\epsilon,k} = \begin{cases} \frac{(-1)^{k+1}(n+s)!(n-s)!}{k(n+\epsilon s-k)!(n-\epsilon s+k)!}, & k \neq 0,\\ \epsilon(H_{n-s} - H_{n+s}), & k = 0, \end{cases} \quad \text{and} \quad \beta_{n,k} = \begin{cases} \frac{2(-1)^{k+1}(n!)^2}{k^2(n+k)!(n-k)!} & k \geqslant 1,\\ -\sum_{k=1}^{n} \beta_{n,k} & k = 0. \end{cases}$$

In the relations above, $H_n = \sum_{j=1}^{n} j^{-1}$ is the harmonic number. Note that $k\beta_{n,k} = 2\alpha_{n,0,0,k}$ for $k \geqslant 1$.

Appendix B. Finite difference operators in *d*-dimensions

Consider a d-dimensional grid defined by the set of points $x_v = (v_1 h_1, \ldots, v_d h_d)$, where $v = (v_1, \ldots, v_d)$ is a multiple index, $v_j \in \mathbb{Z}$ and h_j represents the grid spacing in j-direction $(j = 1, \ldots, d)$. Corresponding to the continuum vector-function $\mathsf{v} : \mathbb{R}^d \to \mathbb{C} \times \cdots \times \mathbb{C}$ we associate the grid vector-function v such that $v_{(v)} \equiv v(x_{(v)}) = \mathsf{v}(x_{(v)})$.

The shift operator by k-points in the j-direction, S_j^k is defined by

$$S_j^k v_{(v_1,\ldots,v_d)} = v_{(v_1,\ldots,v_j+k,\ldots,v_d)}. \tag{61}$$

A discrete operator D_j acting in the j-direction is constructed as a linear combination of the shift operators defined in (61) using the same weights as the corresponding one dimensional operator D; an operator $D_{j_1 \cdots j_r}$ acting in $j_1 \cdots j_r$-directions is constructed as a composition of one-directional operators $\{D_{j_1}^1, \ldots, D_{j_r}^r\}$:

$$D_j = \sum_k a^k(h_j) S_j^k \quad \text{where} \quad D = \sum_k a^k(h) S^k, \tag{62}$$

$$D_{j_1 \cdots j_r} = a(h_{j_1}, \ldots h_{j_r}) D_{j_1}^1 \cdots D_{j_r}^r. \tag{63}$$

In order to represent the functions in Fourier space, we consider only grid function which are periodic in each direction and limit the grid to having a finite number of points, N_j for the direction $j, j = 1, \ldots, d$. We introduce

$$
\begin{aligned}
&h = (h_1, \ldots, h_d), \quad N = (N_1, \ldots, N_d),\\
&b_v(\omega) = (2\pi)^{-d/2} e^{i\omega x_v},\\
&V_h = h_1 \cdots h_d,\\
&\mathcal{S}_{\underline{x}}(N) = \mathcal{S}_{\underline{x}}(N_1) \times \cdots \times \mathcal{S}_{\underline{x}}(N_d),\\
&\mathcal{S}_{\underline{\omega}}(N) = \mathcal{S}_{\underline{\omega}}(N_1) \times \cdots \times \mathcal{S}_{\underline{\omega}}(N_d),\\
&\mathcal{S}_{\underline{\xi}}(N) = \mathcal{S}_{\underline{\xi}}(N_1) \times \cdots \times \mathcal{S}_{\underline{\xi}}(N_d).
\end{aligned}
$$

In the relations above, $\mathcal{S}_{\underline{x}}(N_j \in \mathbb{N})$, $\mathcal{S}_{\underline{\omega}}(N_j \in \mathbb{N})$ and $\mathcal{S}_{\underline{\xi}}(N_j \in \mathbb{N})$ have been defined in (8), (11) and (14), respectively. Then, the formulas for the Fourier decomposition (9), scalar product and a norm (12) and Parseval relation (13) are valid also in d-dimensions.

Let $\xi \in \mathcal{S}_{\underline{\xi}}(N)$ and apply the shift operator by k-points in the j-direction S_j^k on a basis vector $b_v(\omega)$. This leads to

$$
S_j^k b_v(\omega) = \widehat{S}^k(\xi_j) b_v(\omega), \quad \text{with} \quad \widehat{S}^k(\xi) = e^{i\xi k}.
$$

Then the Fourier symbols for the operators D_j and $D_{j_1 \ldots j_r}$ from (62), (63) are defined by:

$$
\begin{aligned}
&D_j b_v(\omega) = \widehat{D}(\xi_j; h_j) b_v(\omega),\\
&D_{i_1 \cdots i_r} b_v(\omega) = \widehat{D}(\xi_{i_1}, \ldots, \xi_{i_r}; h_{i_1}, \ldots, h_{i_r}) b_v(\omega),\\
&\quad \text{with}\\
&\widehat{D}(\xi_j; h_j) = \sum_k a^k(h_j) \widehat{S}^k(\xi_j),\\
&\widehat{D}(\xi_{i_1}, \ldots, \xi_{i_r}; h_{i_1}, \ldots, h_{i_r}) = a(h_{i_1}, \ldots, h_{i_r}) \widehat{D}^1(\xi_{i_1}) \cdots \widehat{D}^r(\xi_{i_r}).
\end{aligned}
$$

Appendix C. Further properties of the Fourier symbols

1. **Recurrence relations**:

$$
\begin{aligned}
&\hat{d}^{(1,n+1)} = \hat{d}^{(1,n)} + \hat{\delta} |c_n| \hat{\Omega}^{2n},\\
&\hat{d}^{(2,n+1)} = \hat{d}^{(2,n)} + |d_n| \hat{\Omega}^{2n+2},\\
&\hat{d}^{(1,n,s)} = \hat{d}^{(1,n,s-1)} - \frac{(-1)^s}{(n+s) C_{2n}^{n+s}} \hat{\Omega}^{2n+1} \cos \frac{(2s-1)\xi}{2},\\
&\hat{\mathbf{d}}^{(1,n,s)} = \hat{\mathbf{d}}^{(1,n,s-1)} - \frac{(-1)^s}{(n+s) C_{2n}^{n+s}} \hat{\Omega}^{2n+1} \sin \frac{(2s-1)\xi}{2}.
\end{aligned}
$$

2. **Small frequency behavior**:

$$
\begin{aligned}
&\hat{d}^{(1,n,s)} \simeq \xi \left[1 - (-1)^n T^{(1,n,s)} \xi^{2n} \right],\\
&\hat{\mathbf{d}}^{(1,n,s)} \simeq (-1)^n s \frac{2n+1}{2n+2} T^{(1,n,s)} \xi^{2n+2},\\
&\sqrt{\hat{d}^{(2,n)}} \simeq \xi \left[1 - \frac{(-1)^n}{2} T^{(2,n)} \xi^{2n} \right].
\end{aligned}
$$

3. **Sums**

$$
\begin{aligned}
&\sum_{k=0}^{\infty} |c_k| x^{2k} = \frac{\arcsin \frac{x}{2}}{\frac{x}{2} \sqrt{1 - \left(\frac{x}{2}\right)^2}},\\
&\sum_{k=0}^{\infty} |d_k| x^{2k} = \left(\frac{\arcsin \frac{x}{2}}{\frac{x}{2}} \right)^2.
\end{aligned}
$$

4. **Limits** $n \to \infty$:

$$\lim_{n\to\infty} \hat{d}^{(2,n)}(\xi) = \xi^2, \quad \forall \xi \in [0,\pi],$$
$$\lim_{n\to\infty} \hat{d}^{(1,n,s)}(\xi) = \xi, \quad \forall \xi \in [0,\pi),$$
$$\lim_{n\to\infty} \hat{\mathbf{d}}^{(1,n,s)}(\xi) = \mathbf{0}, \quad \forall \xi \in [0,\pi],$$
$$\lim_{n\to\infty} \hat{r}^{(n)}(\xi) = 0, \quad \forall \xi \in [0,\pi).$$

5. **The** $D^{(2,n)}$**-norm** defined by

$$\|v\|^2_{h,D^{(2,n)}} = \frac{1}{h^2}\sum_{i=1}^{d}\sum_{k=1}^{n} |d_{k-1}|\|(hD_{+i})^k u\|^2_h + \|v\|^2_h$$

is equivalent with the D_+ norm. This norm has been used to prove strong stability of the initial boundary value problem for the wave equation in [16] for the second and fourth order accuracy case.

References

[1] B. Gustafsson, H.O. Kreiss, J. Oliger, Time Dependent Problems and Difference Methods, John Wiley & Sons, New York, 1995.
[2] E. Hairer, S.P. Norsett, G. Wanner, Solving Ordinary Differential Equations. I: Nonstiff Problems, Springer-Verlag, Berlin, 1987.
[3] G. Calabrese, I. Hinder, S. Husa, J. Comput. Phys. 218 (2006) 607.
[4] M. Shibata, T. Nakamura, Phys. Rev. D 52 (1995) 5428.
[5] T.W. Baumgarte, S.L. Shapiro, Phys. Rev. D 59 (1999) 024007.
[6] C. Gundlach, J.M. Martin-Garcia, Phys. Rev. D 74 (2006) 024016.
[7] S. Husa, J.A. Gonzalez, M. Hannam, B. Brügmann, U. Sperhake, Class. Quant. Grav. 25 (2008) 105006.
[8] F. Pretorius, <arXiv:0710.1338[gr-qc]>.
[9] M. Hannam, Class. Quant. Grav. 26 (2009) 114001. <arXiv:0901.2931[gr-qc]>.
[10] F. Pretorius, Phys. Rev. Lett. 95 (2005) 121101.
[11] M. Campanelli, C.O. Lousto, P. Marronetti, Y. Zlochower, Phys. Rev. Lett. 96 (2006) 111101.
[12] J.G. Baker, J. Centrella, D.I. Choi, M. Koppitz, J. van Meter, Phys. Rev. Lett. 96 (2006) 111102.
[13] B. Szilágyi, H.-O. Kreiss, J. Winicour, Phys. Rev. D 71 (2005) 104035.
[14] M. Babiuc, B. Szilagyi, J. Winicour, Lect. Notes Phys. 692 (2006) 251.
[15] M.C. Babiuc, B. Szilagyi, J. Winicour, Class. Quant. Grav. 23 (2006) S319.
[16] G. Calabrese, C. Gundlach, Class. Quant. Grav. 23 (2006) S343–S367.
[17] M. Motamed, M. Babiuc, B. Szilágyi, H.-O. Kreiss, J. Winicour, Phys. Rev. D 73 (2006) 124008.
[18] G. Cohen, Higher-Order Numerical Methods for Transient Wave Equations, Springer, 2001.
[19] L. Anné, P. Joly, Q.H. Tran, Construction and analysis of higher order finite difference schemes for the 1D wave equation, Comput. Geosci. 4 (3) (2009) 207–249.
[20] R.M. Wald, General Relativity, University Chicago Press, 1984.
[21] J.W. York, in: L. Smarr (Ed.), Sources of Gravitational Radiation, Cambridge University Press, 1979.
[22] O. Sarbach, G. Calabrese, J. Pullin, M. Tiglio, Phys. Rev. D 66 (2002) 064002.
[23] G. Nagy, O. Ortiz, O. Reula, Phys. Rev. D 70 (2004) 044012.
[24] C. Gundlach, J.M. Martin-Garcia, Phys. Rev. D 70 (2004) 044031.
[25] C. Gundlach, J.M. Martin-Garcia, Phys. Rev. D 70 (2004) 044032.
[26] H. Beyer, O. Sarbach, Phys. Rev. D 70 (2004) 104004.
[27] M. Alcubierre, B. Brügmann, Phys. Rev. D 63 (2001) 104006.
[28] Y. Zlochower, J.G. Baker, M. Campanelli, C.O. Lousto, Phys. Rev. D 72 (2005) 024021.
[29] L.N. Trefethen, SIAM Rev. 24 (2) (1982) 113–136.