

Finite Difference Schemes and Digital Waveguide Networks for the Wave Equation: Stability, Passivity, and Numerical Dispersion

Stefan Bilbao and Julius O. Smith, III, *Member, IEEE*

Abstract—In this paper, some simple families of explicit two-step finite difference methods for solving the wave equation in two and three spatial dimensions are examined. These schemes depend on several free parameters, and can be associated with so-called interpolated digital waveguide meshes. Special attention is paid to the stability properties of these schemes (in particular the bounds on the space-step/time-step ratio) and their relationship with the passivity condition on the related digital waveguide networks. Boundary conditions are also discussed. An analysis of the directional numerical dispersion properties of these schemes is provided, and minimally directionally-dispersive interpolated digital waveguide meshes are constructed.

Index Terms—Digital waveguide networks, finite difference schemes, Von Neumann analysis, waveguide meshes.

I. INTRODUCTION

THE subject of this paper is the analysis of some simple families of finite difference schemes for solving the wave equation in N spatial dimensions

$$\frac{\partial^2 u}{\partial t^2} = c^2 \nabla_{ND}^2 u, \quad \nabla_{ND}^2 \triangleq \sum_{l=1}^N \frac{\partial^2}{\partial x_l^2}. \quad (1)$$

Here, t is time, and $\mathbf{x} = [x_1, \dots, x_N]^T$ is the set of spatial coordinates; c is the wave speed, assumed constant. The solution $u(\mathbf{x}, t)$ is assumed to be defined over $t \geq 0$ and for $\mathbf{x} \in \mathbb{R}^N$. It will be unique provided that two “well behaved” initial conditions $u(\mathbf{x}, 0)$ and $(\partial u / \partial t)|_{\mathbf{x}, t=0}$ are given [1]. The treatment of boundary conditions is deferred to Section V-C.

The focus here is on the cases $N = 2, 3$, for which the numerical solution of (1) has applications in room acoustics [2] and the modeling of sound production in musical instruments [3]. Special attention is paid to the relationship between a particular class of finite difference scheme and so-called interpolated digital waveguide networks [4] or meshes [2], [5], [6], especially with regard to numerical stability properties, and the directional dependence of numerical dispersion. The material in this paper has appeared, in an expanded form, in [7].

Manuscript received September 24, 2001; revised December 3, 2002. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Walter Kellermann.

S. Bilbao is with the Sonic Arts Research Centre and the Department of Music, The Queen’s University, Belfast, Ireland BT7 INN (e-mail: s.bilbao@qub.ac.uk).

J. O. Smith, III, is with the Center for Computer Research in Music and Acoustics, Department of Music, Stanford University, Stanford, CA 94305-8180 USA.

Digital Object Identifier 10.1109/TSA.2003.811535

In Sections II and III, simple families of 2D and 3D explicit difference schemes are introduced, which are suitable for the numerical integration of a physical system described by (1) (among others). Standard Fourier transform techniques for analyzing numerical stability, better known as *Von Neumann analysis*, are also reviewed; this analysis is carried out in Section IV for families of schemes which are dependent on several free parameters. In Section V these same families of schemes are reexamined as implementations of so-called interpolated digital waveguide networks, and certain distinctions between the notion of passivity and standard numerical stability are discussed. It is shown that passivity may be used as a simple sufficient condition for stability; this can be extremely useful in the analysis of boundary conditions, as we will show in Section V-C. In Section VI, returning to the traditional finite difference viewpoint, the spectral analysis of these schemes is further refined in order to show that the problem of directional dependence of numerical dispersion can be dealt with in a straightforward way, and it is shown how the schemes’ free parameters may be adjusted in order to minimize such an effect. Finally, in Section VII, several plots of the numerical phase velocities for these schemes are presented, for various choices of the free parameters.

II. TWO-STEP DIFFERENCE SCHEMES

The wave equation is second order in the time variable, and the simplest *difference approximations* [1] are explicit two-step methods. The solution is assumed to be approximated over an N -dimensional set of grid points by a *grid function*. Updating the scheme at any grid point at any time step requires access to the grid function at the two previous time steps. The spatial separation between the grid points is assumed to be X in all dimensions, and all operations recur periodically at intervals of T seconds ($1/T$ is the *sampling rate*). The difference schemes are of the form

$$v_{\mathbf{m}}^{(n+1)} + B(E_1, \dots, E_N)v_{\mathbf{m}}^{(n)} + v_{\mathbf{m}}^{(n-1)} = 0. \quad (2)$$

Here, v is the grid function, and $\mathbf{m} = [m_1, \dots, m_N]^T \in \mathbb{Z}^N$ is the integer-valued set of coordinates of the spatial grid point at $\mathbf{x} = \mathbf{m}X$. Similarly, the integer time index n indicates an approximation to some model problem at $t = nT$. B is some polynomial function of the spatial unit shift operators E_1, \dots, E_N , and their inverses $E_1^{-1}, \dots, E_N^{-1}$ which are defined by

$$E_l v_{[m_1, \dots, m_l, \dots, m_N]}^{(n)} = v_{[m_1, \dots, m_l+1, \dots, m_N]}^{(n)} \quad (3a)$$

$$E_l^{-1}v_{[m_1, \dots, m_l, \dots, m_N]}^{(n)} = v_{[m_1, \dots, m_{l-1}, \dots, m_N]}^{(n)} \quad (3b)$$

for any $l = 1, \dots, N$.

Clearly, difference scheme (2), if stable, is *nondissipative*, or *lossless*, since it is symmetric with respect to the time index n (the forward and inverse iterations are identical). As yet, nothing has been said about which model system, if any, difference scheme (2) approximates. Depending on the choice of B , it could serve to numerically integrate not only the wave equation (1), but perhaps the equation of motion of a lossless beam (in 1D) or plate (in 2D), or other hybrids such as lossless stiff strings or membranes as well.

A. Von Neumann Analysis

The discrete spatial Fourier transform v of the grid function v is

$$V(\boldsymbol{\beta}) = \frac{1}{(2\pi)^{N/2}} \sum_{\mathbf{m} \in \mathbb{Z}^N} e^{-j\boldsymbol{\beta}^T \mathbf{m} X} v_{\mathbf{m}} X^N \quad (4)$$

and is defined over spatial frequencies $\boldsymbol{\beta}^T = [\beta_1, \dots, \beta_N]$, such that $-\pi/X \leq \beta_1, \dots, \beta_N \leq \pi/X$. By the shift theorem for Fourier transforms

$$E_l v \longleftrightarrow e^{j\beta_l X} V \quad E_l^{-1} v \longleftrightarrow e^{-j\beta_l X} V$$

for any transform pair $v \longleftrightarrow V$, for any $l = 1, \dots, N$. In other words, the shift operators correspond to multiplication by linear phase terms in the spatial frequency domain. The recursion (2) becomes, in the transform domain

$$V^{(n+1)} + B(e^{j\beta_1 X}, \dots, e^{j\beta_N X}) V^{(n)} + V^{(n-1)} = 0. \quad (5)$$

Due to the spatial symmetry of the wave equation, for the schemes addressed in this paper, B is always a function of the operators

$$M_l = (E_l + E_l^{-1})/2 \quad (6)$$

for which

$$M_l v \longleftrightarrow c_l V \quad (7)$$

where

$$c_l \triangleq \cos(\beta_l X). \quad (8)$$

The numerical stability of this scheme can be examined by taking the z transform of (5) to yield the *amplification equation*

$$z + B(e^{j\beta_1 X}, \dots, e^{j\beta_N X}) + z^{-1} = 0. \quad (9)$$

We note that we use the same notation for B as before, though B is now a function of spatial frequency variables instead of shift operators. Its functional form is the same as before. This should cause no confusion. (The same will also be true of the operator F , which will be defined in Section IV.) Scheme (2) is called *stable* (in the restricted sense [1]), if the roots of the amplification equation are bounded by 1 in magnitude, for all spatial frequencies. These roots are simply

$$z_{\pm} = \frac{1}{2} \left(-B \pm \sqrt{B^2 - 4} \right). \quad (10)$$

If B is real, then the stability condition is

$$\max_{\boldsymbol{\beta}} |B| \leq 2 \quad (11)$$

in which case $|z_{\pm}| = 1$ for all $\boldsymbol{\beta}$. The reader is referred to [1] for the details, but the condition can be thought of as generalizing the stability condition for a two-pole digital filter with transfer function $H(z) = 1/(1 + Bz^{-1} + z^{-2})$; here, however, B is not merely a multiplier value, but a function of spatial frequency $\boldsymbol{\beta}$, and condition (11) must hold over all such frequencies.

It is also useful to introduce the *symbol* [1] of the difference scheme; writing $z = e^{sT}$, for some complex frequency $s = \sigma + j\omega T$, then the symbol $P^d(s, \boldsymbol{\beta})$ is simply the left side of the amplification equation (9) above, normalized by the factor $1/T^2$, i.e.,

$$P^d(s, \boldsymbol{\beta}) \triangleq \frac{1}{T^2} (e^{sT} + e^{-sT} + B). \quad (12)$$

The symbol of a difference equation indicates its behavior for a single plane-wave solution of the form $v = e^{snT + j\boldsymbol{\beta}^T \mathbf{m} X}$ and is to be compared with the analogous symbol $P(s, \boldsymbol{\beta})$ for the model problem the scheme approximates. For the system of interest here, namely, the wave equation (1), the symbol P is obtained by examining a plane-wave solution of the form $u = e^{st + j\boldsymbol{\beta}^T \mathbf{x}}$, in which case we find

$$P u = 0 \quad \text{with} \quad P \triangleq s^2 + c^2 |\boldsymbol{\beta}|^2 \quad (13)$$

where $|\boldsymbol{\beta}|^2$ is simply the squared magnitude of the vector $\boldsymbol{\beta} = [\beta_1, \dots, \beta_N]_T$ of spatial frequencies.

III. EXPLICIT DIFFERENCE SCHEMES FOR THE WAVE EQUATION IN 2D AND 3D

Now consider (1) for $N = 2$. In order to approximate its solution on a grid, the following basic difference approximations to the 2D Laplacian may be used

$$\delta_{2D,+}^2 = \frac{2}{X^2} (M_1 + M_2 - 2) = \nabla_{2D}^2 + O(X^2) \quad (14a)$$

$$\delta_{2D,\times}^2 = \frac{2}{X^2} (M_1 M_2 - 1) = \nabla_{2D}^2 + O(X^2). \quad (14b)$$

(Even though the operators M_1 and M_2 , defined in (6), are applied to the continuous function u , we use the same notation as before; this should cause no confusion.) $\delta_{2D,+}^2$ is an approximation based on neighboring points to the north, south, east, and west at a distance of X , and $\delta_{2D,\times}^2$ employs points at a distance of $\sqrt{2}X$ from the current update point, in the diagonal directions. It should be clear that since both difference operators are second-order accurate [1] approximations to ∇_{2D}^2 , then any linear combination $p\delta_{2D,+}^2 + (1-p)\delta_{2D,\times}^2$ will be as well, for any p (assumed real). The second-order, centered, time-difference operator is defined by

$$\begin{aligned} \delta_t^2 u(\mathbf{x}, t) &= \frac{1}{T^2} [u(\mathbf{x}, t+T) - 2u(\mathbf{x}, t) + u(\mathbf{x}, t-T)] \\ &= \frac{\partial^2 u}{\partial t^2} + O(T^2) \end{aligned} \quad (15)$$

and a simple family of finite difference schemes can be obtained by writing (1) as

$$\delta_t^2 u = c^2 [p\delta_{2D,+}^2 + (1-p)\delta_{2D,\times}^2] u + O(X^2, T^2) \quad (16)$$

and replacing u by a grid function v . Updating the grid function at a given grid point requires access to previous values of the grid function at locations at most one grid point away in either the x_1 or x_2 directions (or both, along the diagonal).

For $N = 3$, let us consider three types of approximations to ∇_{3D}^2

$$\delta_{3D,+}^2 = \frac{2}{X^2} (M_1 + M_2 + M_3 - 3) \quad (17a)$$

$$\delta_{3D,*}^2 = \frac{1}{X^2} (M_1M_2 + M_1M_3 + M_2M_3 - 3) \quad (17b)$$

$$\delta_{3D,\times}^2 = \frac{2}{X^2} (M_1M_2M_3 - 1). \quad (17c)$$

These approximations make use of grid points which are X , $\sqrt{2}X$, and $\sqrt{3}X$, respectively, away from the current update point; all are accurate to $O(X^2)$, and thus any linear combination of them will be as well, i.e.,

$$(p\delta_{3D,+}^2 + q\delta_{3D,*}^2 + (1-p-q)\delta_{3D,\times}^2) u = \nabla_{3D}^2 u + O(X^2)$$

for any real p and q . Using the second-order time-differencing operator (15), the 3D wave equation may then be rewritten as

$$\delta_t^2 u = c^2 [p\delta_{3D,+}^2 + q\delta_{3D,*}^2 + (1-p-q)\delta_{3D,\times}^2] u + O(X^2, T^2). \quad (18)$$

When u is replaced by the grid function v , both the 2D and 3D difference schemes (16) and (18) can be written in the form of (2), after multiplying through by a factor of T^2 .

It will be useful to write the difference schemes (16) and (18) as they will be implemented by expanding the spatial differencing operators. In 2D

$$v_{\mathbf{m}}^{(n+1)} = \sum_{r_1=-1}^1 \sum_{r_2=-1}^1 a_{r_1, r_2} E_1^{r_1} E_2^{r_2} v_{\mathbf{m}}^{(n)} - v_{\mathbf{m}}^{(n-1)} \quad (19)$$

where

$$a_{r_1, r_2} = \begin{cases} 2 - 2\lambda^2(1+p), & |r_1| + |r_2| = 0 \\ p\lambda^2, & |r_1| + |r_2| = 1 \\ (1-p)\lambda^2/2, & |r_1| + |r_2| = 2 \end{cases} \quad (20)$$

where the dimensionless quantity λ defined by

$$\lambda = \frac{cT}{X} \quad (21)$$

plays an important role in the numerical stability analysis to follow.

Similarly, in 3D, (18) can be rewritten as

$$v_{\mathbf{m}}^{(n+1)} = \sum_{r_1=-1}^1 \sum_{r_2=-1}^1 \sum_{r_3=-1}^1 a_{r_1, r_2, r_3} E_1^{r_1} E_2^{r_2} E_3^{r_3} v_{\mathbf{m}}^{(n)} - v_{\mathbf{m}}^{(n-1)} \quad (22)$$

with

$$a_{r_1, r_2, r_3} = \begin{cases} 2 - \lambda^2(4p + q + 2), & |r_1| + |r_2| + |r_3| = 0 \\ p\lambda^2, & |r_1| + |r_2| + |r_3| = 1 \\ q\lambda^2/4, & |r_1| + |r_2| + |r_3| = 2 \\ (1-p-q)\lambda^2/4, & |r_1| + |r_2| + |r_3| = 3 \end{cases} \quad (23)$$

where λ is defined as in (21).

A simple class of difference schemes of the form (2) which approximates the wave equation (1) to second-order accuracy is thus established. Any member of this class will solve the wave equation numerically provided it is numerically stable—the necessary stability analysis is performed in Section IV.

A. Transformed Laplacian Operators

The Laplacian operators of (14) and (17), interpreted as operating on a grid function v , transform to the spatial frequency domain as

$$\delta_{2D,+}^2 v \longleftrightarrow \frac{2}{X^2} (c_1 + c_2 - 2) V \quad (24a)$$

$$\delta_{2D,\times}^2 v \longleftrightarrow \frac{2}{X^2} (c_1 c_2 - 1) V \quad (24b)$$

in 2D, and

$$\delta_{3D,+}^2 v \longleftrightarrow \frac{2}{X^2} (c_1 + c_2 + c_3 - 3) V \quad (25a)$$

$$\delta_{3D,*}^2 v \longleftrightarrow \frac{1}{X^2} (c_1 c_2 + c_1 c_3 + c_2 c_3 - 3) V \quad (25b)$$

$$\delta_{3D,\times}^2 v \longleftrightarrow \frac{2}{X^2} (c_1 c_2 c_3 - 1) V \quad (25c)$$

in 3D, where c_i is given by (8).

Notice that because these approximations to the Laplacian are symmetric and make use of points that are at most one grid point away in any or all of the spatial coordinate directions, their transforms are *multilinear* [8] functions of the cosines c_i of the components of the spatial frequency variable (see the Appendix for a definition of multilinearity).

IV. STABILITY

Difference schemes (16) and (18), when written in terms of the grid function v , and normalized by the factor T^2 , have the form (2), with

$$B = -2 - 2\lambda^2 F \quad (26)$$

and the operator F is defined, for schemes (16) and (18), by

$$F = \begin{cases} F_{2D} = (p\delta_{2D,+}^2 + (1-p)\delta_{2D,\times}^2) X^2/2 \\ F_{3D} = (p\delta_{3D,+}^2 + q\delta_{3D,*}^2 \\ \quad + (1-p-q)\delta_{3D,\times}^2) X^2/2. \end{cases} \quad (27)$$

Notice that the stability condition (11) can then be rewritten as

$$\max_{\beta} |B| \leq 2 \iff \max_{\beta} |1 + \lambda^2 F| \leq 1$$

and if F is independent of λ , it is straightforward to arrive at the equivalent pair of conditions

$$\max_{\beta} F \leq 0 \quad \lambda^2 \leq -\frac{2}{\min_{\beta} F}. \quad (28)$$

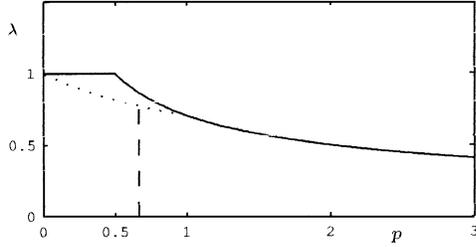


Fig. 1. Upper bounds on λ for the 2D scheme (16), plotted against the free parameter p : the maximum value of λ for Von Neumann stability (solid line), and the maximum value for a passive waveguide mesh implementation (dotted line); there is a passive mesh structure only for $0 \leq p \leq 1$. Choices of parameters for the optimally direction-independent scheme are plotted as a dashed line.

It thus suffices to find the maximum and minimum values of the function F . Using the definitions of the transformed Laplacians from (24) and (25), F for schemes (16) and (18) is given by

$$\begin{aligned} F_{2D} &= p(c_1 + c_2) + (1-p)c_1c_2 - 1 - p \\ F_{3D} &= p(c_1 + c_2 + c_3) + \frac{q}{2}(c_1c_2 + c_1c_3 + c_2c_3) \\ &\quad + (1-p-q)c_1c_2c_3 - 2p - q/2 - 1. \end{aligned}$$

The functions F_{ND} above, for $N = 2, 3$ are easily seen to be multilinear in the variables c_l , $l = 1, \dots, N$, which take on values between -1 and 1 . These functions are thus defined over the ND hypercube

$$\Omega_N = \{\mathbf{c} = (c_1, \dots, c_N) \in \mathbb{R}^N \mid -1 \leq c_1, \dots, c_N \leq 1\}. \quad (29)$$

It is simple to show that a multilinear function defined over a hypercube whose sides are aligned with the coordinate axes will take on its extreme values at the hypercube corners (see the Appendix). Since the domain Ω_N is such a hypercube in \mathbb{R}^N , all extrema of F must then occur at its 2^N corners, i.e., over the points in Ω_N^c defined by

$$\Omega_N^c = \{\mathbf{c} \in \Omega_N \mid |c_1| = \dots = |c_N| = 1\}. \quad (30)$$

The stability conditions (28) can be rephrased as

$$\max_{\mathbf{c} \in \Omega_N^c} F(\mathbf{c}) \leq 0 \quad \lambda^2 \leq -\frac{2}{\min_{\mathbf{c} \in \Omega_N^c} F(\mathbf{c})}. \quad (31)$$

Thus, in 2D, the candidates for the global extrema of F_{2D} are

$$\begin{aligned} \mathbf{c} = (1, 1): & \quad F_{2D} = 0 \\ \mathbf{c} = (1, -1): & \quad F_{2D} = -2 \\ \mathbf{c} = (-1, -1): & \quad F_{2D} = -4p \end{aligned}$$

where the symmetry of F_{2D} with respect to c_1 and c_2 permits dropping the evaluation at $\mathbf{c} = (-1, 1)$. The stability conditions (31) then reduce to simply

$$p \geq 0 \quad \lambda^2 \leq \min\left(1, \frac{1}{2p}\right). \quad (32)$$

The maximum value of λ is plotted as a function of p in Fig. 1.

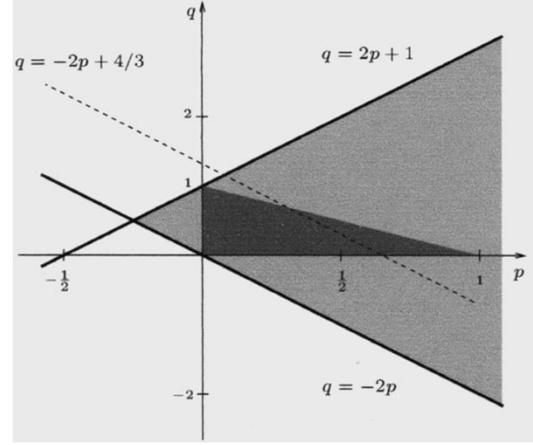


Fig. 2. Stability region for (18) in the (p, q) plane. The maximum value of λ for stability over this region is given in (33). The sub-region over which a passive waveguide mesh implementation exists is shown in dark grey, with the bound on λ given in (44). Choices of parameters for an optimally direction-independent scheme are plotted as a dashed line.

In 3D, the possible global extrema of F_{3D} (again employing the symmetry of F_{3D}) are

$$\begin{aligned} \mathbf{c} = (1, 1, 1): & \quad F_{3D} = 0 \\ \mathbf{c} = (1, 1, -1): & \quad F_{3D} = -2 \\ \mathbf{c} = (1, -1, -1): & \quad F_{3D} = -4p - 2q \\ \mathbf{c} = (-1, -1, -1): & \quad F_{3D} = -4p + 2q - 2 \end{aligned}$$

and the stability conditions are thus

$$-2p \leq q \leq 2p + 1 \quad \lambda^2 \leq \min\left(1, \frac{1}{2p+q}, \frac{1}{2p-q+1}\right). \quad (33)$$

The stability region is plotted in the (p, q) plane in Fig. 2.

V. DIGITAL WAVEGUIDE NETWORKS

In this section, a class of finite difference schemes based on the use of *digital waveguide networks* (DWNs) [4] is considered. DWNs are known to yield efficient and numerically robust computational models for wind, string, and brass musical instruments [9], [10]. More recently they have been applied also to the modeling of acoustic membranes and spaces [3], [2], [11]–[14].

A. Background

A DWN is defined as a completely general collection of *bidirectional delay lines* terminated on *scattering junctions* [4]. See Fig. 3 for a graphical representation of such a network and these two important component types. The signals stored in this discrete-time network are referred to as *waves*.

Each bidirectional delay line or *digital waveguide* can be conceived of as a discrete-time transmission line segment or acoustic tube, transporting digital wave signals in opposite directions at a fixed sample rate. Referring to the enlarged view of a digital waveguide of delay T shown in Fig. 3(a), where the two *input wave variables* are the discrete time signals $v_A^{(n)}$ –

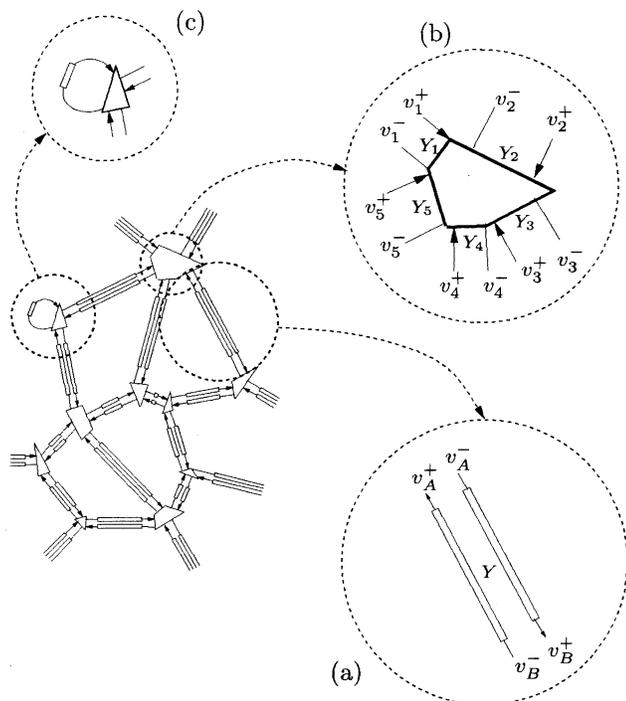


Fig. 3. Portion of a digital waveguide network, and enlarged views of its principal components: (a) a bidirectional delay line, of delay duration T and admittance Y (accepting two waves v_A^- and v_B^- output from scattering junctions, delaying them, and producing two waves v_A^+ and v_B^+ each of which is then incident on a scattering junction); (b) a scattering junction connected to five waveguides of admittances Y_1, \dots, Y_5 , (accepting, in this case, five input waves v_1^+, \dots, v_5^+ , and yielding five output waves v_1^-, \dots, v_5^-); and (c) a self-loop.

and $v_B^{(n)-}$ and the two *output wave variables* are $v_A^{(n)+}$ and $v_B^{(n)+}$, it is easy to see that

$$v_A^{(n)+} = v_B^{(n-1)-} \quad v_B^{(n)+} = v_A^{(n-1)-}. \quad (34)$$

Here the wave signals are indexed by integer n , indicating that they take on values at times $t = nT$. (We note here that in a DWN, the individual waveguides are not all necessarily of the same delay length; they need only be multiples of a common unit delay. In the DWN's for the 2D and 3D wave equation which we will shortly examine, however, all delays will be of duration T .) One can conclude, trivially, from (34), that

$$Y \cdot \left[\left(v_A^{(n)+} \right)^2 + \left(v_B^{(n)+} \right)^2 \right] = Y \cdot \left[\left(v_A^{(n-1)-} \right)^2 + \left(v_B^{(n-1)-} \right)^2 \right] \quad (35)$$

or, in other words, that the sum of the squares of the input waves is the same as that of the output waves; the weighting by the waveguide *admittance* $Y > 0$ is not necessary here, but becomes significant when scattering is introduced.¹

A set of M wave variables v_1^+, \dots, v_M^+ incident on a scattering junction [see Fig. 3(b)] is then scattered instantaneously² according to the equation

$$v_k^- = v - v_k^+ \quad (36)$$

where v , the *junction variable* is not a wave, and is defined by

$$v = \sum_{j=1}^M \alpha_j v_j^+ \quad \text{for } \alpha_j \triangleq \frac{2Y_j}{\sum_{q=1}^M Y_q}, \quad j = 1, \dots, M. \quad (37)$$

The constants $Y_j > 0, j = 1, \dots, M$ are the admittances associated with the M adjoining waveguides, and the $\alpha_j, j = 1, \dots, M$, also positive, are known as the *partial transmission coefficients*. The scattering equation above is a statement of Kirchoff's Laws for a *parallel* connection of transmission lines propagating *voltage* waves. As such, it is possible to show that

$$\sum_{k=1}^M Y_k \cdot (v_k^-)^2 = \sum_{k=1}^M Y_k \cdot (v_k^+)^2 \quad (38)$$

and thus the sum of squares of the wave signals, weighted by the admittances, is preserved by the scattering operation. It is also, of course, possible to define a *series* scattering junction with respect to waveguide *impedances*. Similarly, *dual* wave variables can also be treated [15].

A key property of a DWN made up of scattering junctions and bidirectional delay lines is that, provided the immittances³ of all waveguides are nonnegative, then the network must be *lossless* [16] as a whole. Indeed, it is possible to conclude, from (35) and (38) that for a closed network made up of J unit-delay digital waveguides, the j th of which contains two signal samples $x_{j,1}$ and $x_{j,2}$ and has admittance Y_j , that

$$E \triangleq \sum_{j=1}^J Y_j \cdot (x_{j,1}^2 + x_{j,2}^2) = \text{constant}. \quad (39)$$

Thus, the preservation of this positive definite energy measure serves as a stability condition on the network as a whole. By appealing to classical network theory [16], it is possible to introduce other lossless elements such as discrete time inductors, capacitors, transformers and gyrators, and also lossy elements such as resistors, which leave the network more generally *passive*, i.e., an energy measure such as E as defined in (39) must be nonincreasing with n .

Though terms such as “lossless” and “passive” are used here, it is worth being a bit more specific; a network made up of elements, each of which can be shown to be lossless or passive individually, will be called, following Belevitch [16], *concretely lossless* or *concretely passive*. This is the property which interests us, as it leads to a simple stability test, namely, checking the positivity of the network immittances. A network may, however, be passive as a whole, even if some of its components are not.

²When the time index n is omitted in an expression, it holds for all n .

¹The grid wave variables v are here assumed to be proportional to a “force-like” variable such as voltage, pressure, etc., as opposed to a “flow-like” variable such as current, velocity, etc.

³An *immittance* is defined as either an admittance or an impedance [16]. In acoustic modeling applications, one normally works with the wave impedance of an acoustic tube or a vibrating string.

For example, a series connection of a two resistors of resistances R and $-R$ is exactly lossless, even though it contains an active element. In this case, the network is called *abstractly lossless* or *abstractly passive*. As will be shown later in this section, it is indeed possible for a DWN for the wave equation to be abstractly lossless while not concretely lossless. For simplicity, “lossless” and “passive” will henceforth mean “concretely lossless” and “concretely passive.”

It should be noted that the DWN, when applied to numerical integration problems, bears a very strong resemblance to filter designs such as *wave digital filters* (WDFs) [17], [18], and other simulation methods based on scattering such as the *transmission line matrix* method (TLM) [19], [20], [21] and *multidimensional wave digital filter* (MDWDF) methods [22], [23]. In lumped wave digital networks, for example, scattering is defined with respect to *port resistances* [18] instead of admittances, but the scattering equation for a parallel connection (a *parallel adaptor* [18], in WDF terminology) is identical to (36); the unit-delay waveguide, though conceived of here as a transmission line segment, also exists in the lumped WDF framework, and has been called a *unit element* [17]. In fact, the DWN’s presented in this section can be viewed as wave digital networks, where the waveguide admittances are taken as port resistances at the scattering junctions. For all the methods mentioned above, scattering according to an equation such as (36) is the key operation, and passivity is the crucial attribute, regardless of whether the application is filtering or simulation.

B. DWNs for the 2D and 3D Wave Equation

The DWN shown in Fig. 3 is unstructured; the scattering junctions are not associated with particular spatial locations. If they occur in a regular, grid-like arrangement, the network is often referred to as a *waveguide mesh* [2], [5], [6], and it begins to be possible to associate the behavior of such a network with a numerical integration method. Passivity, in the context of numerical integration, can be viewed as a sufficient condition for numerical stability, and it is much simpler to verify than numerical stability of the type discussed in Section IV—as mentioned in the previous section, nonnegativity of the network element values (the waveguide immittances) is necessary and sufficient for passivity. The same condition also extends readily to more complex systems having spatially varying coefficients, passive nonlinearities, and certain types of time-variation [9], [14]; the frequency-domain Von Neumann analysis does not apply to such systems. Boundary conditions are also particularly easy to deal with, as we will see shortly in Section V-C; the termination of a passive network by passive lumped discrete circuit elements must behave passively as a whole. Checking the numerical stability of a particular boundary condition coupled with a given finite difference scheme can be extremely involved otherwise, though analysis techniques such as the so-called *GKSO theory* [1], [24] and the *energy method* [24] can be used.

Waveguide networks suitable for the numerical integration of the 2D and 3D wave equation are shown in Fig. 4; for the 2D network, each scattering junction is connected to its nearest neighbors (a distance X away) by waveguides of admittance Y_1 , and to neighbors in the diagonal directions (a distance $\sqrt{2}X$ away) by waveguides of admittance Y_2 . At each junction, there is, ad-

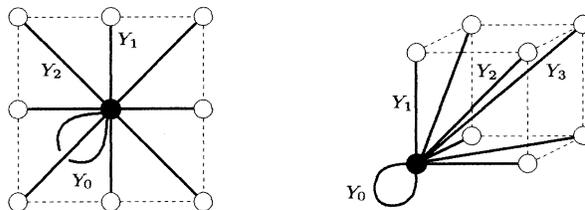


Fig. 4. Waveguide networks for the 2D and 3D wave equation, at left and right respectively. Waveguide connections from a given junction (black) to its neighbors are indicated by thick black lines; in the representation of the 3D mesh, only connections to points in a single neighboring octant are shown. Admittances of waveguides are indicated for one representative of each type in both cases.

ditionally, a *self-loop* of admittance Y_0 ; a self-loop is simply a waveguide whose ends are both connected to the same junction, where one of the two signal paths has been dropped, as it is redundant [see Fig. 3(c)]. It is equivalent in the unit-delay case to a *wave digital capacitor* [17], as well as a *capacitive stub* [20] in the TLM framework. In 3D, there are three types of connecting waveguides, of admittances Y_1 , Y_2 and Y_3 , running between any junction and its neighbors at distances of X , $\sqrt{2}X$ and $\sqrt{3}X$, respectively. Again, there are self-loops of admittance Y_0 at every junction. For either the 2D or the 3D mesh, the partial transmission coefficients can be written as per (37), in terms of the waveguide admittances Y_l , $l = 0, \dots, N$, for $N = 2, 3$ by

$$\alpha_{r_1, \dots, r_N} = \frac{2Y_l}{Y_{J, ND}} \quad \text{for } \sum_{k=1}^N |r_k| = l$$

where $Y_{J, ND}$ is the total *junction admittance* at a given scattering junction, and is equal to the sum of the admittances of all waveguides connected to the junction. Thus

$$Y_{J, 2D} = Y_0 + 4Y_1 + 4Y_2$$

$$Y_{J, 3D} = Y_0 + 6Y_1 + 12Y_2 + 8Y_3.$$

In order for the network to be passive, it is required that all of these admittances (or, equivalently, the partial transmission coefficients⁴) be positive, in either the 2D or 3D cases. This simple criterion immediately implies numerical stability, for as mentioned in Section V-A, there is a direct physical interpretation of the sum of the squares of all the signal values in the bidirectional delay lines as an energy; this quantity will be non-increasing as time progresses if the network is passive and there are no sources.

A DWN can, if its topology and immittance values are set in a particular way (to be discussed shortly), be viewed as a wave implementation of a finite difference scheme, usually of the *finite-difference time domain* (FDTD) variety [25], [26]. That is, the scattering and shifting/delaying operations applied to wave variables can always be reduced to an equivalent finite difference scheme in the physical junction variables (sampled voltages and currents in the electrical setting, or pressures and volume velocities in acoustic tubes, etc.). The 2D and 3D waveguide networks

⁴In these DWN’s for the wave equation, it is the partial transmission coefficients which are of importance; the network admittances may all be scaled by a common factor without affecting the calculation. In a simulation of the full system of conservation laws (from which the wave equation is derived), the admittance values regain their importance. See [7] for further discussion.

shown in Fig. 4 can be made to behave exactly according to the difference schemes (16) and (18), respectively. This equivalence for the case of the 2D mesh can be shown by tracing the flow of signals backward through the network.

First, let us consider the operation of the 2D mesh shown at left in Fig. 4 at the junction at a point with coordinates $\mathbf{x} = \mathbf{m}X$, $\mathbf{m} = (m_1, m_2)$ for m_1, m_2 integer. At any given time step $n+1$, the junction accepts nine waves, which will be called $v_{\mathbf{m}, \mathbf{r}}^{(n+1)+}$. The index $\mathbf{r} = (r_1, r_2)$ refers to the wave approaching the junction along the vector \mathbf{r} in the (x_1, x_2) plane, for $r_1, r_2 = -1, 0, 1$ ($\mathbf{r} = \mathbf{0}$ indicates the wave approaching from the self-loop). The junction variable $v_{\mathbf{m}}^{(n+1)}$ can be expressed, from (37), as

$$v_{\mathbf{m}}^{(n+1)} = \sum_{r_1=-1}^1 \sum_{r_2=-1}^1 v_{\mathbf{m}, \mathbf{r}}^{(n+1)+}. \quad (40)$$

It should be clear, from (34), that

$$v_{\mathbf{m}, \mathbf{r}}^{(n+1)+} = E_1^{r_1} E_2^{r_2} v_{\mathbf{m}, -\mathbf{r}}^{(n)-} \quad (41)$$

where E_1 and E_2 and their inverses are the shift operators as defined in (3) (E_1^0 and E_2^0 are simply identity operators). That is, a wave incident on a junction at the current time step was emitted from an adjacent junction at the previous time step. Furthermore, from (36)

$$\begin{aligned} v_{\mathbf{m}, -\mathbf{r}}^{(n)-} &= v_{\mathbf{m}}^{(n)} - v_{\mathbf{m}, -\mathbf{r}}^{(n)+} \\ &= v_{\mathbf{m}}^{(n)} - E^{-r_1} E^{-r_2} v_{\mathbf{m}, \mathbf{r}}^{(n-1)-} \end{aligned}$$

and by inserting this expression back in (41) and (40), and using (36) and (37) again

$$v_{\mathbf{m}}^{(n+1)} = \sum_{r_1=-1}^1 \sum_{r_2=-1}^1 \alpha_{r_1, r_2} E^{r_1} E^{r_2} v_{\mathbf{m}}^{(n)} - v_{\mathbf{m}}^{(n-1)} \quad (42)$$

which is identical to (19) under the identification of α_{r_1, r_2} with a_{r_1, r_2} , as defined in (20). (That is, the DWN generates a difference scheme identical to that obtained by applying finite differences.) Using similar manipulations, it is possible to identify the 3D waveguide mesh in Fig. 4 with the difference scheme (22), if $\alpha_{r_1, r_2, r_3} = a_{r_1, r_2, r_3}$, with a_{r_1, r_2, r_3} as defined in (23).

The important thing here is that for these networks to be made up of passive transmission line segments, the partial transmission coefficients must always be *nonnegative*, due to the nonnegativity of the waveguide admittances themselves. (Waveguides of zero admittance are formally allowed, although any such waveguide can be dropped from the network entirely.) Applying these nonnegativity conditions to the α parameters for the 2D scheme, from (20), we get

$$0 \leq p \leq 1 \quad \lambda^2 \leq \frac{1}{1+p} \quad (43)$$

and in 3D, from (23)

$$p, q \geq 0 \quad p+q \leq 1 \quad \lambda^2 \leq \frac{2}{4p+q+2}. \quad (44)$$

These conditions on p , q and λ are substantially more restrictive than the stability conditions from (32) and (33). The passivity regions described by (43) and (44) are plotted in Figs. 1 and 2, respectively.

The passivity conditions above can thus be viewed as trivially verifiable sufficient stability conditions for the difference schemes (16) and (18); they are not, however, necessary. The stability of schemes which do not satisfy the passivity conditions must be verified by the approach outlined in the previous section; for more involved difference schemes, this can become quite difficult. A DWN can behave in a stable manner even when certain components of the network are active (though in these configurations, power-normalization [4] of wave quantities is not possible)—as per the discussion in the second-last paragraph of Section V-A, such a network is abstractly, but not concretely passive [16]. It is worth mentioning, however, that it is by no means impossible that there exist concretely passive network representations corresponding to the stable difference schemes mentioned here; such a representation will necessarily have a different topology from the structures shown in Fig. 4. For an interesting example of alternate topologies for the same difference scheme, the reader is referred to the discussion of triangular and hexagonal waveguide meshes in Appendix A of [7].

C. Boundary Conditions in the DWN

In order to see more clearly the interest in a DWN representation for a finite difference scheme, it is useful to examine the problem of boundary termination. Consider the 2D wave equation, and a boundary along the x_1 axis, at $x_2 = 0$. There are two commonly encountered lossless boundary conditions [1]

$$u(x_1, 0) = 0 \quad \left. \frac{\partial u}{\partial x_2} \right|_{x_2=0} = 0.$$

The first, or Dirichlet, condition describes the boundary condition for a membrane terminated rigidly along $x_2 = 0$, where u is the transverse displacement of the membrane. The second, or Neumann, condition serves to describe the termination of a 2D acoustic space in a hard boundary or wall, when u is a pressure variable.

Because updating the difference schemes (16) requires access to grid variables at neighboring points, they may not be used directly at grid points on a boundary. For the Dirichlet condition, this is not problematic; if the grid variables are located at points on the boundary itself, they may be set permanently to zero, and need not be used in the updating of the grid function over internal points. The DWN implementation at such a grid point follows simply by cutting all waveguide connections to the exterior, and by scattering according to (36), where v is set to zero—see Fig. 5, at left. In fact, it is easy to show that this corresponds to a “short-circuit” termination of a parallel connection of voltage (or pressure) waveguides, and is thus lossless [17]; waves incident on the boundary junction are reflected with sign inversion. Thus, the weighted sum of squares of the signals in the network remains constant even when such a boundary condition is applied.

The Neumann condition is slightly more complex, though we may again associate a particular termination of the difference

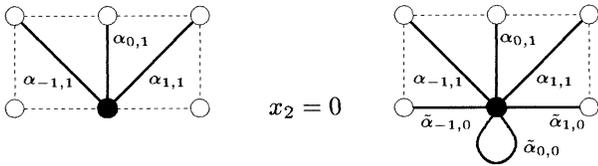


Fig. 5. Termination of the 2D waveguide mesh at a southern boundary, with $x_2 = 0$. At left, termination corresponding to the Dirichlet boundary condition $u = 0$, where the boundary junction is short-circuited, and at right, to the Neumann condition, where partial transmission coefficients of waveguides along the boundary and self-loops are modified by identification with the coefficients (47) of a boundary difference scheme (46).

scheme (16) with a lossless termination of the DWN. First consider the difference scheme (16) in its expanded form (19) at a southern boundary, i.e., where the grid variable $v_{m_1,0}^{(n+1)}$ is to be determined from its neighbors at time steps n and $n-1$. Thus,

$$v_{m_1,0}^{(n+1)} = \sum_{r_1=-1}^1 \sum_{r_2=-1}^1 a_{r_1,r_2} E_1^{r_1} E_2^{r_2} v_{m_1,0}^{(n)} - v_{m_1,0}^{(n-1)}. \quad (45)$$

Obviously, the variables $v_{m_1,-1}$, $v_{m_1\pm 1,-1}$ which will be required by this equation are beyond the mesh boundary, and thus the above update cannot be used at such points. Using the approximation

$$\left. \frac{\partial u}{\partial x_2} \right|_{x_1=m_1 X, x_2=0} \approx (u(m_1 X, 0) - u(m_1 X, -X)) / X$$

the Neumann boundary condition becomes, in terms of the grid function v

$$v_{m_1,0} - v_{m_1,-1} = 0.$$

The boundary difference scheme (45) becomes

$$v_{m_1,0}^{(n+1)} = \sum_{r_1=-1}^1 \sum_{r_2=-1}^1 \tilde{a}_{r_1,r_2} E_1^{r_1} E_2^{r_2} v_{m_1,0}^{(n)} - v_{m_1,0}^{(n-1)} \quad (46)$$

where

$$\begin{aligned} \tilde{a}_{0,0} &= a_{0,0} + a_{-1,0} \\ \tilde{a}_{0,-1} &= \tilde{a}_{1,-1} = \tilde{a}_{-1,-1} = 0 \\ \tilde{a}_{1,0} &= a_{1,0} + a_{1,-1} \\ \tilde{a}_{-1,0} &= a_{-1,0} + a_{-1,-1} \\ \tilde{a}_{0,1} &= a_{0,1} \\ \tilde{a}_{1,1} &= a_{1,1} \\ \tilde{a}_{-1,1} &= a_{-1,1}. \end{aligned} \quad (47)$$

This modified update at the boundary grid points can be interpreted as a DWN termination of the type shown at right in Fig. 5, where the partial transmission coefficients α are modified to $\tilde{\alpha}$, and identified with the \tilde{a} given above. Note that the partial transmission coefficients in waveguides leading to the mesh interior are unchanged, and thus scattering at internal junctions is carried out using the α parameters, and not $\tilde{\alpha}$. $\tilde{\alpha}_{-1,0}$, $\tilde{\alpha}_{1,0}$, the transmission coefficients in waveguides connecting junctions on the boundary, and $\tilde{\alpha}_{0,0}$, that of the self-loop are now given from (47) and (20), in terms of p and λ , by

$$\begin{aligned} \tilde{\alpha}_{-1,0} &= \tilde{\alpha}_{1,0} = \lambda^2(1+p)/2 \\ \tilde{\alpha}_{0,0} &= 2 - \lambda^2(2+p). \end{aligned}$$

Again, the network will be lossless if these parameters are non-negative, giving the additional conditions

$$p \geq -1 \quad \lambda^2 \leq \frac{1}{1+p/2}.$$

These conditions are clearly respected for any choices of λ and p which satisfy the passivity conditions given in (43), which must hold over the mesh interior. Thus, this particular boundary termination of the finite difference scheme does not compromise the passivity of the DWN as a whole, and the difference scheme remains exactly lossless in the absence of round-off errors.

The above analysis obviously can be applied to any edge of a rectangular domain, and extends easily to the 3D case as well. What we have shown above is that, regardless of whether one chooses a finite difference scheme or a DWN in the actual computer implementation, the DWN representation gives simple conditions for numerical stability; it should be mentioned, however, that the addition of numerical boundary conditions can be analyzed as such only when the DWN is passive over the interior, i.e., under restricted conditions such as (43) and not (32). The boundary termination for more general DWNs has been explored in great detail in [7].

VI. DIRECTIONAL DISPERSION

It is of interest to examine now the symbols of the difference schemes (16) and (18). For the 2D difference scheme, from (12), using the definition of f_{2D} from (27), the symbol is

$$P_{2D}^d(s, \beta) = \frac{1}{T^2} (e^{sT} + e^{-sT} - 2 - 2\lambda^2 F_{2D})$$

and, expanding e^{sT} in a Taylor series in sT and the cosines c_1 and c_2 in $F_{2D}(c)$ in a series in βX

$$P_{2D}^d = P + \frac{1}{12} s^4 T^2 - \frac{1}{12} c^2 (\beta_1^4 + \beta_2^4 + 6(1-p)\beta_1^2\beta_2^2) X^2 + O(T^4, X^4).$$

The first term in the expansion, $P(s, \beta)$, as defined previously in (13), is simply the symbol for the 2D wave equation itself, and indicates consistency of the difference scheme with the problem; the higher order terms in β_1 and β_2 give rise to directionally dependent numerical dispersion in the scheme.

For the special choice of $p = 2/3$, this expression simplifies to

$$P_{2D}^d = P + \frac{1}{12} (s^4 T^2 - c^2 |\beta|^4 X^2) + O(T^4, X^4)$$

and directional dependence of the scheme is confined to higher order terms. That is, the symbol for the difference scheme is dependent only on $|\beta|^4$ (and not the individual components of β) to second order. This special choice of p is plotted as a dashed line in Fig. 1. The scheme is still second-order accurate in T and X ; if it were possible to choose $X/T = c$, or, in other words, $\lambda = 1$, then it would be further possible to extract a factor of $P_{2D}(s, \beta)$ as defined in (13) from the second order terms, giving a fourth order accurate scheme. Because $p = 2/3$, however, the stability bound from (32), giving $\lambda \leq \sqrt{3/4}$, prohibits such a choice of λ .

Similarly, in 3D, it is possible to write

$$\begin{aligned} P_{3D}^d &= \frac{1}{T^2} (e^{sT} + e^{-sT} - 2 - 2\lambda^2 F_{3D}) \\ &= P + \frac{1}{12} s^4 T^2 - \frac{1}{12} c^2 (\beta_1^4 + \beta_2^4 + \beta_3^4) X^2 \\ &\quad - \frac{(1-p-q/2)}{2} c^2 (\beta_1^2 \beta_2^2 + \beta_1^2 \beta_3^2 + \beta_2^2 \beta_3^2) X^2 \\ &\quad + O(T^4, X^4) \end{aligned}$$

and thus, for $q = -2p + 4/3$, we have

$$P_{3D}^d = P + \frac{1}{12} (s^4 T^2 - c^2 |\boldsymbol{\beta}|^4 X^2) + O(T^4, X^4).$$

This special one-parameter family of schemes is plotted as a dashed line in Fig. 2.

Reducing the directional dependence of numerical dispersion may be useful, because in this case, frequency-warping techniques [13], [27] may be used to correct dispersion error; this can be done only if the numerical dispersion is nearly isotropic. Plots of this dispersion error are presented in the next section.

It is also worth noting that other sampling geometries also provide opportunities for the reduction of the directional dispersion; we call attention to, in particular, the triangular and hexagonal geometries in 2D (discussed at length in [3], [7], [12], and [28]). Certain numerical integration structures can be viewed as waveguide meshes or difference schemes, just as the families mentioned in this paper, though it is difficult to directly compare their efficiencies with those of the rectangular grid structures in this paper without making certain assumptions about the spatial bandwidth of the grid function (i.e., the shape of the spectral region it occupies). There are other distinctions as well: triangular and hexagonal sampling patterns do not lend themselves to a simple treatment of boundary conditions over a rectangular region, and the hexagonal scheme may exhibit *parasitic modes* [1], [7]. In 3D, there are quite a few new sampling possibilities, notably the tetrahedral grid [11], [7]. It is possible, nonetheless, to perform the same type of Taylor analysis as above for all of these schemes, though we do not pursue this direction here (and furthermore, these schemes do not contain any free parameters to be optimized, except for λ).

VII. NUMERICAL PHASE VELOCITIES

For an exponential plane wave solution of the form $u = e^{st+j\boldsymbol{\beta}^T \mathbf{x}}$, the wave equation (1) can be written as

$$P(s, \boldsymbol{\beta})u = 0 \quad \text{with} \quad P(s, \boldsymbol{\beta}) = s^2 + c^2 |\boldsymbol{\beta}|^2.$$

For nontrivial solutions, $P(s, \boldsymbol{\beta}) = 0$, and so

$$s = j\omega = \pm jc|\boldsymbol{\beta}|.$$

The phase velocity for the wave equation is defined (disregarding the sign ambiguity above) by

$$v_\phi \triangleq \left| \frac{\omega}{|\boldsymbol{\beta}|} \right| = c.$$

Thus, wave propagation is dispersionless, and speeds are independent of direction.

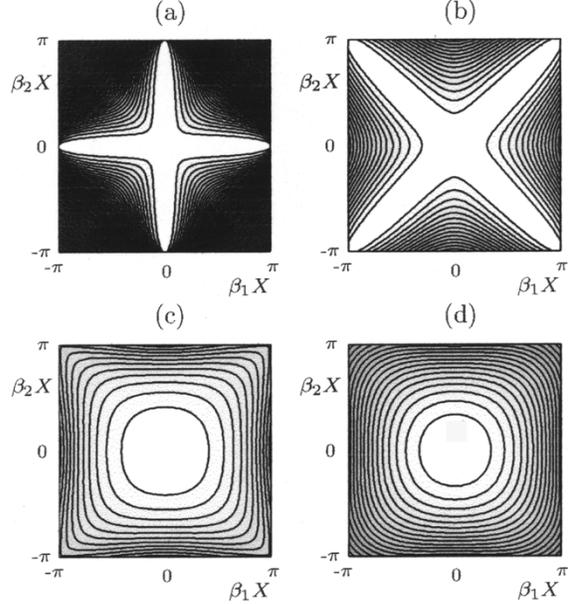


Fig. 6. Relative numerical phase velocities v_{rel} for difference scheme (16), for various values of p and λ ; v_{rel} takes on a value of 1 at $\boldsymbol{\beta} = 0$, and successive deviations of 2% are plotted as contours. In (a) and (b) are shown contours for a simple rectilinear scheme with $p = 0$ and $\lambda = 1$, and $p = 1$ and $\lambda = \sqrt{1/2}$ respectively, in (c) the maximally direction-independent scheme with $p = 2/3$ at the Von Neumann stability bound $\lambda = \sqrt{3/4}$, and in (d) the same scheme with $p = 2/3$ at the passivity bound $\lambda = \sqrt{3/5}$. All plots are over the entire spatial frequency spectrum, i.e., with $-\pi/X \leq \beta_1, \beta_2 \leq \pi/X$.

For the numerical scheme (2), there are, in the spatial frequency domain, two amplification factors $z_{\pm} = e^{\pm sT}$, defined by (10). For stable schemes (i.e., for $|B| \leq 2$), these factors have unit magnitude and can be written as

$$z_{\pm} = e^{\pm j\omega T} = \frac{1}{2} \left(-B \pm \sqrt{B^2 - 4} \right).$$

The numerical phase velocity, defined by

$$v_\phi^d \triangleq \left| \frac{\omega}{|\boldsymbol{\beta}|} \right| \quad (48)$$

can thus be written, using (26), as

$$v_\phi^d = \frac{\cos^{-1}(-B/2)}{T|\boldsymbol{\beta}|}.$$

since $\cos(\omega T) = -B/2$ from (9). For the schemes defined by (16) and (18), the phase velocity of the scheme relative to that of the model system is then

$$v_{rel} = v_\phi^d / v_\phi = \frac{\cos^{-1}(1 + \lambda^2 F_{ND})}{X\lambda|\boldsymbol{\beta}|}.$$

It is useful at this point to examine plots of this quantity for various values of the scheme free parameters; such plots are provided for the 2D scheme in Fig. 6, and for the 3D scheme in Fig. 7.

By consistency of the schemes with the wave equation, it is always true that $v_{rel} = 1$, for $\boldsymbol{\beta} = 0$; depending on the choices of p and λ (in 2D) or p, q and λ (in 3D), there are gross differences between the behaviors of these schemes away from the

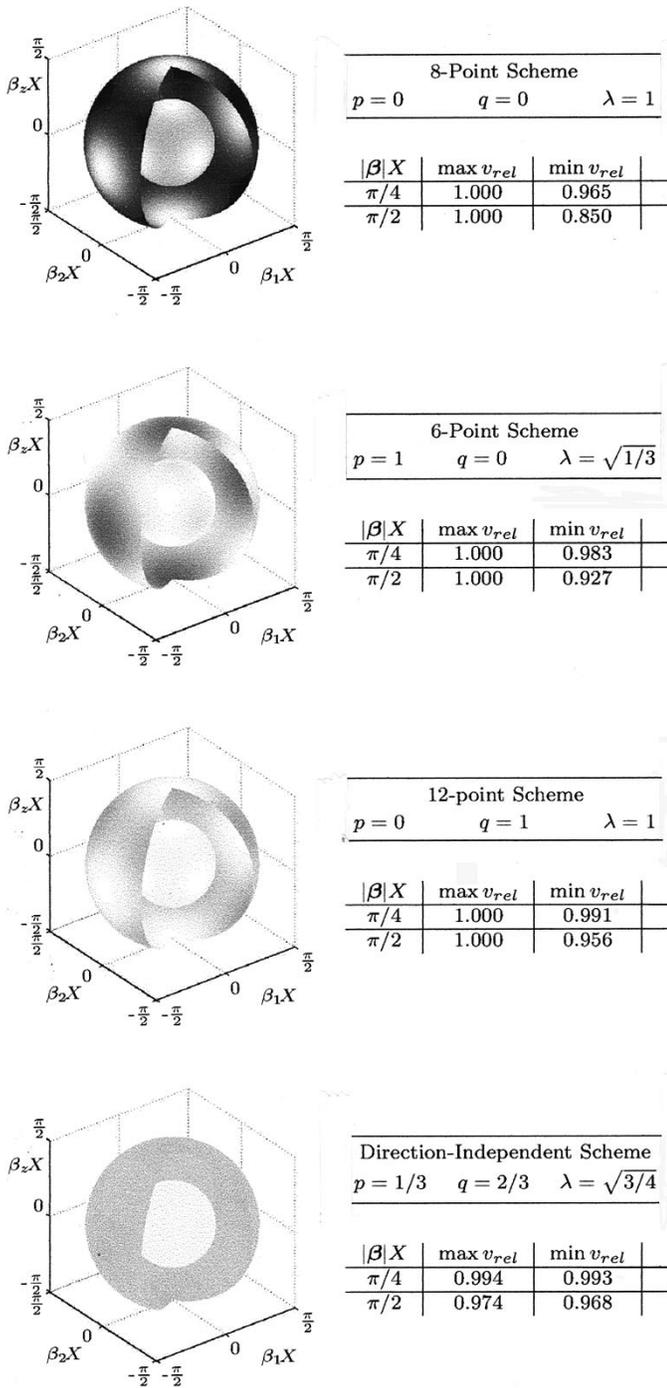


Fig. 7. Plots of v_{rel} for the 3D family of schemes (18), for several values of the parameters p , q and λ (as indicated). v_{rel} is plotted over the two surfaces $|\beta|X = \pi/4$ and $|\beta|X = \pi/2$; white corresponds to $v_{rel} = 1$, indicating correct wave propagation velocity, and black corresponds to the maximal error over the four cases shown, i.e., $v_{rel} = 0.850$. For each surface, the maximum and minimum values of v_{rel} are indicated. The first three plots correspond to simple special cases, for which difference scheme (18) makes use of eight, six, or twelve neighboring points; λ is chosen at the stability bound from (33). Maximum and minimum values of v_{rel} over the two surfaces are given in the adjacent tables.

long wavelength limit. In particular, for certain choices of the parameter values, there are spatial directions for which the schemes are *dispersionless*, i.e., the phase velocity of

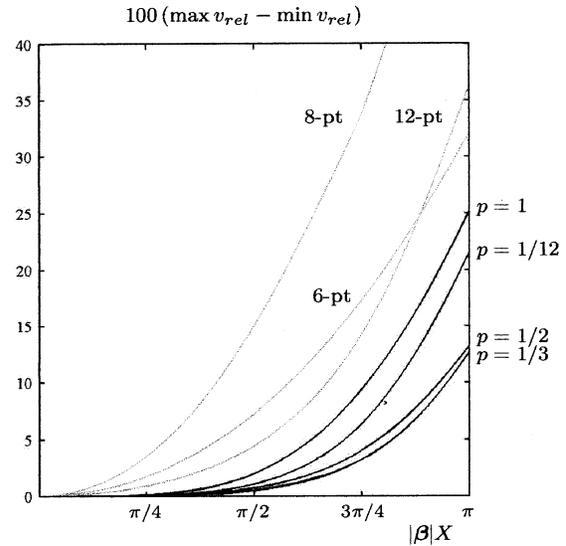


Fig. 8. Plot of the maximum percentage difference between the maximum and minimum of v_{rel} as a function of $|\beta|X$. In black, directionally independent schemes as discussed in Section VI, for various values of p , and in grey, the three simple schemes shown in Fig. 7 (parameter values indicated).

the scheme is exactly equal to c ; this will be true only at the maximum allowed value of λ . In these cases, however, the dispersion of the scheme can be quite large in other directions. For the optimally direction-independent parameter choices discussed in Section VI, there is no dispersionless direction, but the dispersion is more evenly distributed among all the propagation directions—see the bottom two plots in Fig. 6 and the bottom plot in Fig. 7. Also note (from the bottom two plots in Fig. 6 for example) that the numerical dispersion is lessened as λ approaches the Von Neumann stability limit, although it does become more directional.

To emphasize the directional independence of the schemes discussed in Section VI, plots of the maximum percentage difference between the maximum and minimum of v_{rel} as a function of $|\beta|$ are shown in Fig. 8, for various members of the 3D family of schemes.

VIII. CONCLUSIONS

An analysis of the stability of simple families of two-step explicit finite difference schemes for the wave equation has been presented here. These schemes were dependent on several free parameters, and the stability of these schemes can be analyzed in a fully general manner; explicit bounds on the time-step/space-step ratio λ are determined. These schemes, when viewed as digital waveguide networks, are subject to a passivity condition, which is distinct from the stability condition—passivity was found to be a sufficient but not necessary condition for numerical stability which is easy to verify. This ease of verification also holds when numerical boundary conditions are applied. A means of reducing the directional dependence of numerical dispersion, based on a Taylor series expansion of the scheme, was also presented, as were plots of numerical phase velocities for various choices of the scheme parameters.

APPENDIX

In this brief appendix, we sketch a simple proof that a multilinear function defined over a hypercube must take on its extreme values at one of the corners of the region. Let

$$F_N(\mathbf{c}_N) = \sum_{\mathbf{m} \in \mathbb{B}_N} a_{\mathbf{m}} \prod_{j=1}^N c_j^{m_j}, \quad \mathbf{c}_N \in \Omega_N$$

be such a function, where Ω_N is as defined in (29), $\mathbf{c}_N = (c_1, c_2, \dots, c_N)$, and $a_{\mathbf{m}}$ are some real constants, indexed with respect to $\mathbf{m} \in \mathbb{B}_N$, the set of all binary N -tuples. $F_N(\mathbf{c}_N)$ is clearly linear with respect to any single variable c_j , $j = 1, \dots, N$.

Pick an arbitrary point $\mathbf{c}_N = \hat{\mathbf{c}}_N \in \Omega_N$, and suppose that $F_N(\hat{\mathbf{c}}_N) = \hat{F}_N$. Fix $(c_1, \dots, c_{N-1}) = (\hat{c}_1, \dots, \hat{c}_{N-1})$. Then $F_N(\hat{c}_1, \dots, \hat{c}_{N-1}, c_N)$ is a linear function of c_N , over $c_N \in [-1, 1]$, and takes on the value \hat{F}_N for $c_N = \hat{c}_N$. Clearly then, this function must take on its maximum \hat{F}_{N-1} at $c_N = q_N$, where $q_N = 1$ or -1 , and thus $\hat{F}_N \leq \hat{F}_{N-1}$.

Now, define $F_{N-1}(c_{N-1}) \triangleq F_N(c_{N-1}, q_N)$ over $c_{N-1} \in \Omega_{N-1}$. F_{N-1} is also multilinear, and repeating the above procedure $N - 1$ times, we obtain a sequence of values

$$\hat{F}_N \leq \hat{F}_{N-1} \leq \dots \leq \hat{F}_1 \quad (49)$$

where $\hat{F}_1 = F_N(q_1, \dots, q_N)$, the value of F_N at coordinates $(q_1, \dots, q_N) \in \Omega_N^c$, the set of the 2^N corners of Ω_N , as defined in (30). Since the point $\hat{\mathbf{c}}_N$ was arbitrary, a sequence such as (49) and a set of values (q_1, \dots, q_N) always exists, and it is clear that the maximum of F_N must occur in Ω_N^c . The proof that the minimum must also occur over Ω_N^c is similar.

REFERENCES

- [1] J. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*. Pacific Grove, CA: Wadsworth and Brooks/Cole, 1989.
- [2] L. Savioja, T. Rinne, and T. Takala, "Simulation of room acoustics with a 3-D finite-difference mesh," in *Proc. Int. Computer Music Conf.*, Århus, Denmark, Sept. 1994, pp. 463–466.
- [3] S. A. VanDuyne and J. O. Smith, III, "Physical modeling with the 2D digital waveguide mesh," in *Proc. Int. Computer Music Conf.*, Tokyo, Japan, 1993, pp. 40–47.
- [4] J. O. Smith, III, "Music applications of digital waveguides," Center for Computer Research in Music and Acoustics, Dept. Music, Stanford Univ., Stanford, CA, Tech. Rep. STAN-M-39, 1987.
- [5] S. A. VanDuyne and J. O. Smith, III, "A simplified approach to modeling dispersion caused by stiffness in strings and plates," in *Proc. Int. Computer Music Conf.*, Århus, Denmark, Sept. 1994.
- [6] S. A. VanDuyne, J. R. Pierce, and J. O. Smith, III, "Travelling wave implementation of a lossless mode-coupling filter and the wave digital hammer," in *Proc. Int. Computer Music Conf.*, Århus, Denmark, Sept. 1994, pp. 411–418.
- [7] S. Bilbao, "Wave and scattering methods for the numerical integration of partial differential equations," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2001. [Online]. Available: <http://www-ccrma.stanford.edu/~bilbao/>.
- [8] R. Abraham, *Linear and Multilinear Algebra*. New York: W. A. Benjamin, 1966.
- [9] J. O. Smith, III, "Efficient simulation of the reed-bore and bow-string mechanisms," in *Music Applications of Digital WaveGuides*, 1987, pp. 29–34. Tech. Rep. STAN-M-39, Center for Computer Research in Music and Acoustics (CCRMA), Dept. Music, Stanford Univ., Stanford, CA.

- [10] —, "Principles of digital waveguide models of musical instruments," in *Applications of Digital Signal Processing to Audio and Acoustics*, M. Kahrs and K. Brandenburg, Eds. Norwell, MA: Kluwer, 1998, pp. 417–466. [Online]. Available: <http://www.wkap.nl/book.htm/0-7923-8130-0>.
- [11] S. A. VanDuyne and J. O. Smith, III, "The 3D tetrahedral digital waveguide mesh with musical applications," in *Proc. Int. Computer Music Conf.*, Hong Kong, Aug. 18–21, 1996, pp. 9–16.
- [12] F. Fontana and D. Rocchesso, "Physical modeling of membranes for percussive instruments," *Acustica United with Acta Acustica*, vol. 84, pp. 529–542, May/June 1998.
- [13] L. Savioja and V. Välimäki, "Reducing the dispersion error in the digital waveguide mesh using interpolation and frequency-warping techniques," *IEEE Trans. Speech Audio Processing*, vol. 8, pp. 184–194, Mar. 2000.
- [14] J. O. Smith, III and D. Rocchesso. (1997, Dec.) Aspects of digital waveguide networks for acoustic modeling applications. [Online]. Available: <http://www-ccrma.stanford.edu/~jos/wgj/>.
- [15] J. O. Smith, III (2002, May) Digital waveguide modeling of musical instruments. [Online]. Available: <http://www-ccrma.stanford.edu/~jos/waveguide/>.
- [16] V. Belevitch, *Classical Network Theory*. San Francisco, CA: Holden Day, 1968.
- [17] A. Fettweis, "Wave digital filters: Theory and practice," *Proc. IEEE*, vol. 74, pp. 270–327, Feb. 1986.
- [18] —, "Digital filters related to classical structures," *AEU: Archiv für Elektronik und Übertragungstechnik*, vol. 25, pp. 79–89, Feb. 1971. (see also U.S. Patent 3,967,099, 1976, now expired).
- [19] P. B. Johns and R. L. Beurle, "Numerical solution of 2-dimensional scattering problems using a transmission-line matrix," *Proc. Inst. Elect. Eng.*, vol. 118, pp. 1203–1208, Sept. 1971.
- [20] C. Christopoulos, *The Transmission-Line Modeling Method*. New York: IEEE Press, 1995.
- [21] W. J. R. Hoefer, *The Electromagnetic Wave Simulator*. New York: Chichester, 1991.
- [22] A. Fettweis and G. Nitsche, "Transformation approach to numerically integrating PDE's by means of WDF principles," *Multidimensional Systems Sig. Process.*, vol. 2, no. 2, pp. 127–159, May 1991.
- [23] —, "Numerical integration of partial differential equations using principles of multidimensional wave digital filters," *J. VLSI Signal. Process.*, vol. 3, no. 1–2, pp. 7–24, June 1991.
- [24] B. Gustaffson, H.-O. Kreiss, and J. Oliger, *Time Dependent Problems and Difference Methods*. New York: Wiley, 1995.
- [25] A. Taflove, *Computational Electrodynamics*. Norwood, MA: Artech House, 1995.
- [26] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, vol. 14, pp. 302–7, 1966.
- [27] L. Savioja and V. Välimäki, "Reduction of the dispersion error in the interpolated digital waveguide mesh using frequency warping," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, vol. 2, Mar. 15–19, 1999, pp. 973–976. [Online]. Available: <http://www.acoustics.hut.fi/~vpv/publications/icassp99-warp.htm>.
- [28] F. Fontana and D. Rocchesso, "Signal-theoretic characterization of waveguide mesh geometries for models of two-dimensional wave propagation in elastic media," *IEEE Trans. Speech Audio Processing*, vol. 9, pp. 152–61, Feb. 2001.



Stefan Bilbao received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 2001.

He is currently a Lecturer in the Department of Music, at The Queen's University, Belfast, Ireland, and is also associated with the Sonic Arts Research Centre. His research interests include numerical methods and digital filter design applied to musical sound synthesis, and real-time sound synthesis methods.



Julius O. Smith (M'76) received the B.S.E.E. degree from Rice University, Houston, TX, in 1975. He received the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1978 and 1983, respectively. His Ph.D. research involved the application of digital signal processing and system identification techniques to the modeling and synthesis of the violin, clarinet, reverberant spaces, and other musical systems.

From 1975 to 1977, he was with the Signal Processing Department at ESL, Sunnyvale, CA, working on systems for digital communications. From 1982 to 1986, he was with the Adaptive Systems Department at Systems Control Technology, Palo Alto, CA, where he worked in the areas of adaptive filtering and spectral estimation. From 1986 to 1991, he was with NeXT Computer, Inc., responsible for sound, music, and signal processing software for the NeXT computer workstation. Since then he has been an Associate Professor at the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford teaching courses in signal processing and music technology, and pursuing research in signal processing techniques applied to music and audio.