Journal of Computational Physics **148,** 199–226 (1999)
Article ID jcph.1998.6117, available online at http://www.idealibrary.com on IDEAL®

A Finite Difference Method for Dispersive Linear Waves with Applications to Simulating Microwave Pulses in Water

Jonathan H. C. Luke

Department of Mathematics and Center for Applied Mathematics and Statistics, New Jersey Institute of Technology, University Heights, Newark, New Jersey 07102

Received August 26, 1997; revised September 17, 1998

The EP_d method, a finite difference method for highly dispersive linear wave equations, is introduced and analyzed. Motivated by the problem of simulating the propagation of microwave pulses through water, the method attempts to relieve the computational burden of resolving fast processes, such as dipole relaxation or oscillation, occurring in a material with dynamic structure. This method, based on a novel differencing scheme for the time step, is considered primarily for problems in one spatial dimension with constant coefficients. It is defined in terms of the solution of an initial value problem for a system of ordinary differential equations that, in an implementation of the method, need be solved only once in a preprocessing step. For certain wave equations of interest (nondispersive systems, the telegrapher's equation, and the Debye model for dielectric media) explicit formulas for the method are presented. The dispersion relation of the method exhibits a high degree of low-wavenumber asymptotic agreement with the dispersion relation of the model to which it is applied. Comparisons with a finite difference time-domain approach and an approach based on Strang splitting demonstrate the potential of the method to substantially reduce the cost of simulating linear waves in dispersive materials. A generalization of the EP_d method for problems with variable coefficients appears to retain many of the advantages seen for constant coefficients.

Key Words: dispersive linear waves; stiff hyperbolic systems; finite difference method; FDTD; electromagnetics waves; dispersion relation; Debye media.

1. INTRODUCTION

Many fields, such as electromagnetics, acoustics, and seismology, use linear hyperbolic systems of partial differential equations (PDEs) as models for wave propagation. Technologies such as radar, sonar, and seismic imaging motivate substantial interest in the numerical simulation of linear wave phenomena. Much of the work on simulating linear waves has

0021-9991/99 $30.00

emphasized nondispersive waves where the frequency components of the waves travel at a common velocity. However, the extension of technologies to new applications and simple scientific curiosity feed a growing interest in simulating dispersive linear waves.

Our interest in numerical methods for dispersive waves stems from the problem of simulating the propagation of microwave pulses through water. Medical applications have stimulated the study of this problem [1] because the electrical properties of biological tissues are to leading order those of water in the microwave regime. In dielectric materials, such as water, propagating electromagnetic waves interact with the material by inducing or aligning molecular dipoles. The macroscopic effect of these dipoles is modeled as a polarization field in the material. Dispersion arises from differences in the response of the molecular dipoles to different driving frequencies. These differences are generally understood in terms of the response of a damped harmonic oscillator to sinusoidal forcing. The Debye model for dielectric material, the main example considered in this paper, assumes this response is dominated by an exponential relaxation process. In the microwave regime, the Debye model for water is typically implemented with a relaxation time of $O(10^{-11})$ s. Microwave pulses with a carrier wave having a temporal period of $O(10^{-9}$–$10^{-10})$ s have received attention due to their technological applications [2]. This $O(10$–$100)$ contrast between the relaxation time and carrier wave period makes the Debye model for water stiff and hence difficult to treat numerically. The primary motivation of the work presented here is to mitigate this stiffness so as to reduce the computational resources needed to effectively simulate problems like the propagation of microwave pulses in water.

The FDTD Method

A widely used approach to numerical simulation of electromagnetic waves is the finite difference time-domain (FDTD) method. Shlager and Schneider [3] give a selective survey of the vast FDTD literature. In its original form [4], the FDTD method applied to numerical solution of Maxwell's equations in nondispersive media. Various distinct extensions of the FDTD method to dispersive materials have been developed. One extension [5] is based on coupling ordinary differential equations (ODEs) that describe the evolution of the polarization to Maxwell's equations. Another extension [6] couples an integral equation (IE) to Maxwell's equations. Although the ODE and the IE formulations are mathematically equivalent, the methods resulting from the discretization of the different formulations are distinct. An analysis of the ODE-based extension [2] provides guidelines for selecting the time and space step to control the amplitude and phase error in a simulation. A similar analysis [7] of the IE-based extension shows that the ODE-based approach gives a dispersion relation that agrees with that of the continuum model better at low wavenumber. For this reason the method of this paper is compared with the ODE-based extension of FDTD in Section 4. An IE-based method having comparable accuracy to the ODE-based approach has been developed recently [8].

Problem Formulation

This paper examines a class of finite difference methods for linear hyperbolic systems of the form

$$\frac{\partial u}{\partial t} + A\frac{\partial u}{\partial x} + \epsilon^{-1}Bu = 0, \tag{1.1}$$

where $u(x,t)$ is an N-vector depending on the spatial variable x and the temporal variable t. The $N \times N$ real matrix A is symmetric, and the $N \times N$ real matrix B has a nonnegative symmetric part. When the positive parameter ϵ is small, (1.1) is stiff and therefore difficult to solve numerically with standard finite difference methods. The method of this paper was developed to produce an efficient approach to simulating solutions of (1.1) by mitigating the effects of stiffness.

To simplify the presentation, we focus attention on the initial value problem ($u(x,0)$ given) with constant coefficients (homogeneous materials) in one spatial dimension. Extensions to inhomogeneous materials are briefly considered in Section 5; material interfaces and two or more spatial dimensions will be treated elsewhere. For notational simplicity, we furthermore assume that the largest and the smallest eigenvalues of A are $+1$ and -1. For many problems in electromagnetics (when scaled with respect to the speed of light), A naturally has this property; extension of the method to the case where A is a general real-symmetric matrix is straightforward.

Stiffness

From the case where the initial data has no spatial dependence, we see that the stiff system of ODEs $\dot{u} + \epsilon^{-1} Bu = 0$ is embedded in (1.1). Thus, we expect to find a stiff ODE solver embedded in a successful numerical method for (1.1). FDTD methods typically make use of implicit time differencing, which is a standard approach of addressing stiff ODEs. These methods, however, typically make little use of the linearity of this stiff system. Embedding the exact method available for linear systems of ODEs into a numerical method for (1.1) would appear to provide an opportunity for improving standard finite difference methods such as FDTD. The method for (1.1) presented in this paper accomplishes such an embedding. The ability to choose larger time steps than with other methods without loss of accuracy is the main advantage. Because the spatial step Δx is usually chosen to be $\Delta x = \Delta t$ to suppress numerical dispersion [2], this increase in the time step allows the simulation to proceed on a substantially coarser grid with significant improvements in efficiency. Further compounding of these gains with more than one spatial dimension is an attractive direction for this work.

Solution Update by Green's Function

Numerical methods for (1.1) usually take the form of a repeated application of an update procedure that approximates $u(x, t + \Delta t)$ given an approximation of $u(x,t)$. The class of finite difference methods considered here can be thought of as a quadrature formula for an exact representation of the update process in terms of the Green's function of the hyperbolic system. The Green's function, an $N \times N$ matrix depending on x and t, satisfies

$$\frac{\partial G}{\partial t}(x,t) + A\frac{\partial G}{\partial x}(x,t) + \epsilon^{-1} BG(x,t) = 0 \qquad \text{for } t > 0, \tag{1.2a}$$

$$G(x,0) = I\delta(x), \tag{1.2b}$$

where I is the $N \times N$ identity matrix and δ is the Dirac delta function. In terms of the Green's function $G(x,t)$, the exact update operator for (1.1) is given by

$$u(x, t + \Delta t) = \int_{x-\Delta t}^{x+\Delta t} G(x - y, \Delta t) u(y,t)\, dy. \tag{1.3}$$

This representation of the solution is used in [9] to obtain estimates on the decay of a propagating pulse; here it is the basis of our numerical method. Assuming an explicit expression for $G(x, t)$ were available, application of a standard quadrature formula to (1.3) would generate a numerical approximation for the update operator. In the next section, we show how such a quadrature formula can be implemented even when no analytic formula for the Green's function is available.

Outline

Section 1 briefly identifies the scientific context of the paper, formulates the linear hyperbolic system whose numerical solution is considered, and introduces the analytical basis for the proposed method (the Green's function representation of the update operator). Section 2 formulates the method and demonstrates an estimate that is needed in its analysis. In Section 3, numerical errors are analyzed through comparison of the dispersion relations for the method and the underlying model. Section 4 examines the special case of the propagation of electromagnetic waves in a Debye material. Comparisons are made with the FDTD method. Section 5 demonstrates, through an example, that the benefits of our numerical method can be extended to the case of smooth, inhomogeneous materials. Section 6 summarizes the paper, discusses the extent to which the method succeeds in mitigating stiffness, and considers issues associated with its generalization. Appendix I presents alternative formulations of the method that are useful in its analysis and that suggest avenues for its extension to other problems. Appendix II derives explicit analytical formulas for the method in some important special cases. Appendix III compares our method to one based on Strang splitting.

2. FORMULATION OF THE EP$_d$ METHOD

In this section we formulate the method of this paper as a quadrature formula for (1.3). This formula is exact for low-degree polynomials (the specific degree depending on a method parameter). Hence, the method is exact for low-degree polynomial initial data. From this characterization the method obtains its designation "EP$_d$," which is a shorthand for "*exact* for *polynomial* initial data of degree d." The form of the method is that the future values of the fields are computed as linear combinations of past values of the fields. The coefficients of these linear combinations are given in terms of the solutions of a system of ODEs. In this sense, the method is based on a novel time differencing scheme.

Numerical Framework

The numerical framework considered is typical of finite difference methods. Specifically, we seek to devise a method that accurately and efficiently approximates the sampled solution U_n^m,

$$U_n^m = u(n\Delta x, m\Delta t) \qquad m = 0, 1, 2, \ldots \text{ and } n = \ldots, -2, -1, 0, 1, 2, \ldots,$$

where u is a solution of (1.1), and Δx and Δt are fixed positive numbers. We indicate a numerical approximation of U_n^m by u_n^m and the grid points $n\Delta x$ and $m\Delta t$ by x_n and t_m. We seek a method in the form of an update procedure taking the approximate solution at one

time level (u_n^m for all n) and giving the approximate solution at the next time level (u_n^{m+1} for all n). Such a procedure with sampled initial data (u_n^0 for all n) completely defines an approximate solution.

We restrict our attention to explicit linear methods with a fixed finite domain of dependence. In particular, we suppose that u_n^{m+1} depends linearly on $\{u_{n+p}^m : -P \le p \le P\}$, where P is a positive integer. Such methods can always be written in the form

$$u_n^{m+1} = \sum_{p=-P}^{P} M_p u_{n+p}^m, \tag{2.1}$$

where each M_p is an $N \times N$ matrix. Thus, the method will be completely specified by the $2P+1$ matrices $M_{-P}, \ldots, M_P$.

Method Definition

The matrices defining the EP_d method with $d = 2P$ arise from a quadrature formula for estimating the integral in (1.3) in terms of the available data $u_{n-P}^m, \ldots, u_{n+P}^m$. A characterizing property of this quadrature formula is that it is exact when applied to data obtained by sampling any N-vector-valued polynomial of degree $2P$ or less at $x_{n-P}, \ldots, x_{n+P}$. There are three steps in the construction of this quadrature formula: (1) The unique N-vector-valued polynomial of degree $2P$ or less interpolating the points (x_{n+p}, u_{n+p}^m) with $p = -P, \ldots, P$ is computed. (2) The Green's function is applied to this polynomial to advance the solution through a period of time Δt. (3) The resulting updated solution is evaluated at x_n to give u_n^{m+1}.

The polynomial interpolating (x_{n+p}, u_{n+p}^m) with $p = -P, \ldots, P$ has a representation of the form

$$\sum_{q=0}^{2P} \frac{1}{q!} a_q (x - x_n)^q, \tag{2.2}$$

where the a_q's are N-vectors. Computing the a_q's is a standard problem,

$$a_q = \sum_{p=-P}^{P} D_{qp}(\Delta x) u_{n+p}, \tag{2.3}$$

where the matrix with entries $q!D_{qp}$ is the inverse of the $(2P+1) \times (2P+1)$ Vandermonde matrix V with entries $V_{pq} = (p\Delta x)^q$, where $p = -P, \ldots, P$ and $q = 0, \ldots, 2P$. The qth row of the matrix D (having entries D_{qp}) contains the weights for a standard centered difference approximation of the qth derivative. When needed, we use the convention that $D_{qp} = 0$ for $q > 2P$ and for $|p| > P$.

To propagate the interpolating polynomial (2.2) through an interval of time Δt, we apply the Green's function as indicated in (1.3) to obtain

$$\sum_{q=0}^{2P} \frac{1}{q!} \left[\int_{x_n - \Delta t}^{x_n + \Delta t} G(x - y, \Delta t)(y - x_n)^q \, dy \right] a_q. \tag{2.4}$$

Evaluating at x_n to get u_n^{m+1}, we obtain

$$u_n^{m+1} = \sum_{q=0}^{2P} W_q(\Delta t) a_q, \tag{2.5}$$

where the $N \times N$ matrices $W_0, \ldots, W_{2P}$ are given by

$$W_q(\Delta t) = \frac{1}{q!} \int_{x_n - \Delta t}^{x_n + \Delta t} G(x_n - y, \Delta t)(y - x_n)^q \, dy = \frac{1}{q!} \int_{-\Delta t}^{\Delta t} G(-y, \Delta t) y^q \, dy. \tag{2.6}$$

Combining (2.3) and (2.5) gives an explicit formula for the matrices defining the EP_{2P} method:

$$M_p(\Delta x, \Delta t, P) = \sum_{q=0}^{2P} W_q(\Delta t) D_{qp}(\Delta x). \tag{2.7}$$

We note that (2.7) separates the differencing in space from that in time. Thus, there is considerable latitude to change the spatial differencing to serve special purposes.

Evaluation of the W_q's

The representation of the EP_{2P} method in (2.7) is useful as a practical tool only to the extent that the W_q's can be readily calculated. Because an explicit expression for the Green's function G is not generally available, direct evaluation of the W_q's is usually infeasible. We use an approach based on a system of linear ODEs with constant coefficients satisfied by the W_q's. With the convention $W_{-1} = 0$, we find

$$\dot{W}_q + A W_{q-1} + \epsilon^{-1} B W_q = 0, \tag{2.8a}$$

$$W_0(0) = I, \tag{2.8b}$$

$$W_q(0) = 0 \qquad q \geq 1. \tag{2.8c}$$

This system can be obtained by differentiating (2.6) with respect to Δt, applying the differential equations defining the Green's function (1.2), and integrating by parts. Because G may be singular or even a generalized function, these operations must be performed in terms of the theory of distributions in the most general case.

Because the W_q's satisfy a linear system of differential equations with constant coefficients, they can be represented exactly in terms of the exponential of a constant matrix. Let $\mathcal{M}$ be the matrix of the constant coefficient linear system (2.8a). This block lower triangular matrix contains $(2P+1)^2$ $N \times N$ blocks. Every block entry on the main diagonal is $\epsilon^{-1}B$; every block entry on the principle subdiagonal is A; all remaining block entries are the $N \times N$ zero matrix. It follows that $e^{-\mathcal{M}\Delta t}$ equals

$$\exp\left(- \begin{bmatrix} B & 0 & \cdots & \cdots & 0 \\ \epsilon A & B & \ddots & & \vdots \\ 0 & \epsilon A & \ddots & \ddots & \vdots \\ \vdots & & \ddots & B & 0 \\ 0 & \cdots & 0 & \epsilon A & B \end{bmatrix} \frac{\Delta t}{\epsilon} \right) = \begin{pmatrix} W_0 & 0 & \cdots & \cdots & 0 \\ W_1 & W_0 & \ddots & & \vdots \\ W_2 & W_1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & W_0 & 0 \\ W_{2P} & W_{P-1} & \cdots & W_1 & W_0 \end{pmatrix}. \tag{2.9}$$

Thus, the computation of the W_q's and the computation of $\exp(-\mathcal{M}\Delta t)$ are equivalent. Numerical exponentiation of a matrix is not a wholly resolved problem [10, 11]. Nonetheless, a simplistic analysis of the "scaling and squaring method" gives a sense of how the effort to compute the W_q's increases as $\epsilon \to 0$. This method is based on scaling the argument of the exponential using the formula $\exp(-\mathcal{M}\Delta t) = [\exp(-\mathcal{M}\Delta t/2)]^2$; its iterative use allows computation of the desired exponential from that of a matrix with a small norm through repeated squaring. Because the norm of $\mathcal{M}$ is $O(\epsilon^{-1})$ for small ϵ, the effort to compute $\exp(-\mathcal{M}\Delta t)$ grows like $O(\log_2(1/\epsilon))$. Thus, the effort needed to compute the W_q's grows at a very modest rate as $\epsilon \to 0$.

The matrix $\mathcal{M}$ falls into a class of matrices that Moler and Van Loan [10] identify as problematic for exponentiation; specifically, $\mathcal{M}$ typically has defective eigenvalues of high multiplicity. This difficulty is somewhat mitigated by the block triangular Toeplitz structure which significantly simplifies matrix products. In the case of homogeneous materials considered here, only one set of W_q's needs to be computed to define the EP_{2P} method. Thus, the computation of the W_q's is a very small part of the overall effort. For inhomogeneous materials, where a separate set of W_q's may be needed for each point on the spatial grid, the efficiency of this computation becomes a more significant issue, particularly in several spatial dimensions.

A Bound on the W_q's

Bounds on the size of the W_q's are useful in analysis of the EP_{2P} method. In this paper the size of a matrix is measured by its spectral norm (applied to a matrix M this norm, denoted $\|M\|$, is the square root of the largest eigenvalue of MM^*). Applied to symmetric matrices this norm equals the spectral radius of the matrix.

From (2.8) it follows that $W_0(\Delta t) = \exp(-\epsilon^{-1}B\Delta t)$ and

$$W_q(\Delta t) = -\int_0^{\Delta t} e^{-B(\Delta t - s)/\epsilon} A W_{q-1}(s)\, ds \qquad \text{for } q \geq 1. \tag{2.10}$$

Hence, when $q \geq 1$,

$$\|W_q(\Delta t)\| \leq \int_0^{\Delta t} \|e^{-B(\Delta t - s)/\epsilon}\| \|W_{q-1}(s)\|\, ds \tag{2.11}$$

since $\|A\| = 1$. Moreover, we limit our analysis to the case where B is diagonalizable; that is, we assume that $B = \Gamma_B \Lambda_B \Gamma_B^{-1}$, where Λ_B is diagonal and the columns of Γ_B are linearly independent eigenvectors of B. In this case we have the inequality $\|\exp(-Bs/\epsilon)\| \leq \kappa(\Gamma_B)$ for $s \geq 0$, where $\kappa(\Gamma_B) = \|\Gamma_B\|\|\Gamma_B^{-1}\|$ because the real parts of the eigenvalues of B are nonnegative. This bound applied to (2.11) gives

$$\|W_q(\Delta t)\| \leq \frac{\kappa(\Gamma_B)^{q+1}\Delta t^q}{q!} \qquad \text{for } q \geq 0. \tag{2.12}$$

Two features of this inequality are significant. One, the power series $\sum_q W_q z^q$ converges for all complex values z. Two, the right-hand side of (2.12) is independent of ϵ; that is, the bound is uniform in ϵ. When B is not diagonalizable, a generalization of (2.12) assures convergence of the power series, but the bound is no longer uniform in ϵ when zero is a defective eigenvalue of B.

3. ANALYSIS OF DISPERSION RELATIONS

In this section we investigate the EP_{2P} method by comparing its dispersion relation with that of the model (1.1). This comparison leads to a non-Hermitian eigenvalue perturbation problem. For simplicity, analysis is limited to the more generic cases. Nonetheless, we show that the dispersion relation of the EP_{2P} method generally approximates that of the model better than that of a finite difference methods using finite-order time differencing.

Dispersion Relations

For evolution equations with constant coefficients, the dispersion relation describes how the Fourier components of the initial data evolve in time. Comparison of the dispersion relation of a linear model with that of a numerical method for the model is a standard approach for evaluating finite difference methods. To facilitate comparisons we denote the dispersion relation for (1.1) by ω, that for the EP_{2P} method (2.7) by $\tilde{\omega}$, and that for any finite difference method using finite-order time differencing, such as FDTD, by $\bar{\omega}$.

The model (1.1) has solutions of the form $u(x,t) = ve^{i(kx-\omega t)}$, where v is a constant vector (possibly complex), k is a real scalar (the wavenumber), and ω is a complex scalar (the frequency). Given k, we can find v and ω by solving the eigenvalue problem $(kA - iB)v = \omega v$. Hence, $\omega(k)$ is generally a multivalued function of k with N branches denoted by $\omega_\ell(k)$. A numerical method of the form (2.1) has solutions of the form $u_n^m = \tilde{v}e^{i(k\Delta x n - \tilde{\omega}\Delta t m)}$. Given k, we can find $\tilde{v}$ and $\tilde{\omega}$ by solving the eigenvalue problem $\hat{M}\tilde{v} = e^{-i\tilde{\omega}\Delta t}\tilde{v}$, where

$$\hat{M} = \sum_p M_p\, e^{ik\Delta x p}. \tag{3.1}$$

When the M_p's are given by (2.7), we have (using $D_{qp} = 0$ for $q > 2P$),

$$\hat{M}(k) = \sum_{q=0}^{\infty} W_q \left(\sum_p D_{qp}\, e^{ik\Delta x p} \right) = \sum_{q=0}^{\infty} W_q \hat{D}_q(k). \tag{3.2}$$

Because $\hat{M}(k)$ and $kA - i\epsilon^{-1}B$ are analytic functions of k on the complex plane, it follows [12] that the dispersion relations, $\tilde{\omega}_\ell(k)$ and $\omega_\ell(k)$, are branches of analytic functions of k with only algebraic singularities. When k_s is a point of singularity, the dispersion relation will be $O(|k-k_s|^{1/m})$, where m is an integer satisfying $2 \le m \le N$; hence, the dispersion relation is a continuous function of k. It will become significant in our analysis that these singularities occur only at points where the matrix is not diagonalizable.

Application of Dispersion Relations

In general we want $|\tilde{\omega}_\ell(k) - \omega_\ell(k)|$ to be small. The way that this "smallness" is assessed depends on the application of interest. Our analysis focuses on the situation where appropriate numerical parameters (Δx, Δt, and P) are sought given the model parameters (A, B, and ϵ), initial data ($u_0(x)$), and a certain kind of criterion for acceptable errors in the dispersion relation. We consider criteria that identify a maximum wavenumber of interest k_M (usually from the initial data) and impose a bound (possibly depending on k) on $|\tilde{\omega}_\ell(k) - \omega_\ell(k)|$ for k satisfying $|k| \le k_M$. Subject to this bound, the numerical parameters are then chosen to maximize efficiency (or possibly convenience). The case when this choice

can be made independently of ϵ is of interest, as then we are assured that the performance of the method does not seriously deteriorate as $\epsilon \to 0$.

By choosing a maximum wavenumber of interest, we accept that Fourier components of the initial data with larger wavenumbers may be badly mispropagated. This acceptance assumes that these Fourier components do not grow significantly in the numerical simulation. Thus, we generally require that the method be stable for the full range of wavenumbers representable on the numerical grid. In our formulation this kind of stability corresponds to $\mathrm{Im}(\tilde{\omega}(k)) \le 0$ for $|k| \le \pi/\Delta x$. Because solutions of (1.1) are often decaying in time, a stronger condition may be needed to ensure that the decaying solution of interest is not swamped by high wavenumber components that spuriously decay more slowly. Specific criteria will depend on the problem of interest.

Zero Wavenumber Analysis

When $k = 0$ we are considering the evolution of constant initial data governed by the system of ODEs $\dot{u} + \epsilon^{-1} B u = 0$. Hence, the dispersion relation of a numerical method at $k = 0$ is a reflection of the quality of the ODE solver embedded in the method. Because our method is exact for polynomial data of order $2P$ or less (which includes constant data), we have $\tilde{\omega}_\ell(0) = \omega_\ell(0)$ regardless of the value of Δt and ϵ. In sharp contrast, when finite-order time differencing is used, say of order d, $\bar{\omega}_\ell(0) - \omega_\ell(0) = O(\mu_\ell \Delta t^d / \epsilon^{d+1})$, where μ_ℓ is the ℓth eigenvalue of B. Thus methods based on finite-order time differencing require $\Delta t/\epsilon$ to be small to ensure accuracy at $k = 0$. We note, however, that in many applications it is common for μ_ℓ to vanish for several values of ℓ. For solutions dominated by these branches, finite-order time differencing is not so severely handicapped. An example of this situation, the Debye model, is considered in the next section.

Small Wavenumber Asymptotics

A basic property of many standard finite difference methods applied to the case $B = 0$ is that their dispersion relations substantially agree with the dispersion relation of the hyperbolic system (1.1). This agreement has the form of exact equality of the first few Taylor coefficients of the dispersion relations expanded in k about $k = 0$. For example, $\bar{\omega}(k)$ for the simple scheme

$$\frac{u_n^{m+1} - u_n^m}{\Delta t} + A \frac{u_{n+1}^m - u_{n-1}^m}{2\Delta x} + \epsilon^{-1} B u_n^m = 0 \tag{3.3}$$

agrees with $\omega(k)$ to $O(k^1)$ when $B = 0$ (even though the scheme is unstable). Higher order methods typically give higher order agreement. This agreement is an expression of the exceptional suitability of standard finite difference methods for the case $B = 0$.

As noted earlier, this agreement evaporates for finite difference methods using finite-order time differencing when $B \neq 0$. For the method of (3.3), $\bar{\omega}_\ell(0) - \omega_\ell(0) = i \log(1 - \mu_\ell \Delta t/\epsilon)/\Delta t + i\mu_\ell/\epsilon$. When $\mu_\ell = O(1)$, $\bar{\omega}_\ell(0)$ and $\omega_\ell(0)$ are approximately equal only when $\Delta t/\epsilon$ is small. In the following, we show that the method of (2.7) preserves agreement between all branches of the numerical and model dispersion relation at small wavenumbers when $B \neq 0$; typically, $\tilde{\omega}_\ell(k) - \omega_\ell(k) = O(k^{2P+1})$.

Reformulation of $\omega(k)$

Reformulating the eigenvalue problem for $\omega(k)$ to more closely resemble that for $\tilde{\omega}(k)$ simplifies comparison of these dispersion relations. The exact update operator (AI.5) has

solutions of the form $u(x,t) = ve^{i(kx-\omega t)}$; hence, $\hat{H}v = e^{-i\omega(k)\Delta t}v$, where

$$\hat{H}(k) = \sum_{q=0}^{\infty} W_q (ik)^q. \tag{3.4}$$

Convergence is ensured by (2.12) or the observation that $\hat{H}(k) = \exp[-(iAk + \epsilon^{-1}B)\Delta t]$.

Eigenvalue Perturbation Problem

We have characterized $\omega_\ell(k)$ and $\tilde{\omega}_\ell(k)$ in terms of the eigenvalues of $\hat{H}(k)$ and $\hat{M}(k)$. Our strategy for showing that $|\tilde{\omega}_\ell(k) - \omega_\ell(k)|$ is small is based on demonstrating that $\|\hat{H}(k) - \hat{M}(k)\|$ is small. Assuming B is diagonalizable, we show that for any wavenumber cutoff k_M and any constant $C > 0$ we can ensure

$$\|\hat{H}(k) - \hat{M}(k)\| \le Ck^{2P+1} \qquad \text{for } |k| \le k_M \tag{3.5}$$

with proper choice of Δt and Δx. Moreover, this choice can be made independently of ϵ.

Estimation of the Matrix Perturbation

To begin the derivation of (3.5), we recall that e^z can be written as the sum of its Taylor polynomial of degree $2P$ and a remainder $R_{2P}(z)$—a holomorphic function satisfying $|R_d(z)| \le z^{d+1}e^{|z|}/(d+1)!$ for all complex values of z. Applying this decomposition to the definition of $\hat{D}_q(k)$, found in (3.2), gives

$$\hat{D}_q(k) = \sum_p D_{qp}\, e^{ik\Delta xp} = (ik)^q + \sum_p D_{qp} R_{2P}(ik\Delta xp), \tag{3.6}$$

where we have used the relation $\sum_p D_{qp}(p\Delta x)^\ell = \ell!\delta_{q\ell}$, which follows from the definition of the D_{pq}'s in terms of the inverse of the Vandermonde matrix V. Thus,

$$\hat{M}(k) - \hat{H}(k) = \sum_{q=0}^{2P} W_q \sum_p D_{qp} R_{2P}(ik\Delta xp) - \sum_{q=2P+1}^{\infty} W_q (ik)^q. \tag{3.7}$$

We note that there is a constant K depending only on P such that $\sum_p |D_{qp}| \le K/|\Delta x|^q$. Assuming the CFL condition $\Delta t \le P\Delta x$, we apply the bound on $\|W_q\|$ given in (2.12) and the bound on the remainder R_{2P} to (3.7) to obtain

$$\|\hat{M}(k) - \hat{H}(k)\| \le \frac{\kappa(\Gamma_B)(|k|\Delta xP)^{2P+1}}{(2P+1)!}\left[Ke^{\kappa P + |k|\Delta xP} + \kappa^{2P+1}e^{\kappa P\Delta x|k|}\right]. \tag{3.8}$$

This bound is independent of ϵ and applies for all values of k. Moreover, given $C > 0$ and $k_M < \infty$ we can obtain the estimate (3.5) by appropriate choice of Δx (which constrains the choice of Δt through the hypothesized CFL condition).

Estimation of the Eigenvalue Perturbation

To estimate $|\omega(k) - \tilde{\omega}(k)|$ using (3.8), we make use of a result of matrix perturbation theory that bounds the perturbation in the eigenvalues given a bound on the perturbation in

the matrix (see [13, Thm. 6.3.2]): Let F and E be square matrices (with $\Gamma_F^{-1} F \Gamma_F$ diagonal) and λ^* be an eigenvalue of $F+E$; then there is an eigenvalue λ of F such that

$$|\lambda^* - \lambda| \le \kappa(\Gamma_F)\|E\|, \tag{3.9}$$

where $\kappa(\Gamma_F) = \|\Gamma_F\|\|\Gamma_F^{-1}\|$ is the condition number of the matrix of eigenvectors of F. When F is not diagonalizable, we define $\kappa(\Gamma_F) = \infty$. Application of (3.9) with (3.5) gives

$$|\tilde{\omega}_\ell(k) - \omega_\ell(k)| \le \min(\kappa(\Gamma_{\hat{M}}), \kappa(\Gamma_{\hat{H}}))Ck^{2P+1} \qquad \text{for } |k| \le k_M, \tag{3.10}$$

where $\Gamma_{\hat{M}}$ and $\Gamma_{\hat{H}}$ are the matrices of eigenvectors for $\hat{M}$ and $\hat{H}$, respectively.

Implications for Numerical Errors

When B is diagonalizable, $\hat{M}(0) = \hat{H}(0) = \exp(-B\Delta t)$ is also. Thus, $\kappa(\Gamma_{\hat{M}})$ and $\kappa(\Gamma_{\hat{H}})$ are bounded as functions of k is some neighborhood of $k=0$. Hence, $|\tilde{\omega}(k) - \omega(k)| = O(k^{2P+1})$ for small k. In applications where the components of the solution with larger wavenumbers rapidly decay in time, this result indicates that the long-time solution computed using this numerical method will be highly accurate especially for larger values of P (assuming that the D_{pq}'s have been accurately computed).

Because $C = O(\Delta x^{2P+1})$, we are tempted to characterize the method as $O(\Delta x^{2P+1})$. For many practical purposes this characterization is indeed correct. There are some situations, probably not very significant in applications, where this characterization is technically inaccurate. These exceptional situations arise through the dependence of $\hat{M}$ on Δx. Specifically, there may be wavenumbers for which $\min(\kappa(\Gamma_{\hat{M}}), \kappa(\Gamma_{\hat{H}}))$ is not bounded as $\Delta x \to 0$. Since $\min(\kappa(\Gamma_{\hat{M}}), \kappa(\Gamma_{\hat{H}})) \le \kappa(\Gamma_{\hat{H}})$, this consideration is necessary only at those discrete values of k for which $\hat{H}(k)$ is nondiagonalizable. Let $J_1, J_2, \ldots, J_L$ be open intervals containing the wavenumbers k where $\hat{H}(k)$ is nondiagonalizable. It follows that $(\tilde{\omega}_\ell - \omega_\ell)/\Delta x^{2P+1}$ is uniformly bounded on $[-k_M, k_M] - \cup_{\ell=1}^{L} J_\ell$. The practical implication is that the error in the method is $O(\Delta x^{2P+1})$ when we can neglect the components of the solution having wavenumbers near those making $\hat{H}(k)$ nondiagonalizable. Other methods, such as the FDTD method, suffer a similar diminished performance at these wavenumbers.

4. NUMERICAL AND MODEL DISPERSION RELATIONS FOR THE DEBYE MODEL

This section focuses on an important example of (1.1), the Debye model for dispersive dielectric media. The Debye model has the form of (1.1) with

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & a & -b \\ 0 & -a & b \end{pmatrix}, \tag{4.1}$$

where a and b are positive constants satisfying $a+b=1$. The dispersion relations for the model, for the EP_d method and for a standard FDTD method, are considered and compared. A similar but abbreviated comparison of the EP_d method with a method based on Strang splitting is found in Appendix III. Our analysis suggests that the EP_d method offers significant improvements in efficiency and accuracy over standard FDTD approaches to the Debye model.

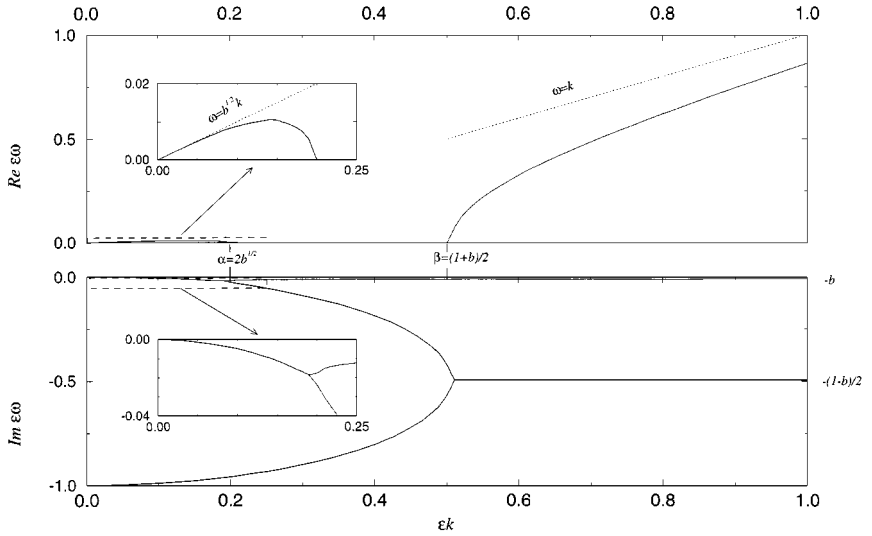

FIG. 1. The real and imaginary parts of the three branches of the dispersion relation for the Debye model with $b = 0.01$.

Dispersion Relation for the Debye Model

The dispersion relation $\omega(k)$ for the nondimensionalized Debye model given in (4.1) satisfies the cubic equation

$$\omega^2(\omega + i/\epsilon) - k^2(\omega + ib/\epsilon) = 0. \tag{4.2}$$

Two branches of $\omega(k)$, denoted $\omega_+(k)$ and $\omega_-(k)$, correspond to forward and backward propagating waves. Since the Debye model is invariant under reversal of the spatial coordinate, $\mathrm{Re}(\omega_+) = -\mathrm{Re}(\omega_-)$ and $\mathrm{Im}(\omega_+) = \mathrm{Im}(\omega_-)$. The third branch, denoted $\omega_0(k)$, whose real part vanishes at all wavenumbers, corresponds to nonpropagating disturbances that decay exponentially in time at a wavenumber-dependent rate.

Figure 1 shows the real and imaginary parts of the dispersion relation with $b = 0.01$, a value approximately corresponding to parameter values used for water in the microwave regime (e.g., in [2] we find $b = 0.0123$). The upper panel of this figure shows only the real part of ω_+. The dispersion relation for the Debye model has a substantial amount of structure including singularities at the points marked α and β. These points, associated with double roots of (4.2), are

$$\alpha = \sqrt{b}[1 - b - b^2 + O(b^3)] \qquad \text{and} \qquad \beta = \frac{1}{2}[1 + b + 2b^2 + O(b^3)]. \tag{4.3}$$

Expanding the three branches of the dispersion relation for small wavenumbers gives

$$\epsilon\omega_\pm = \pm\sqrt{b}(\epsilon k) - i\frac{1-b}{2}(\epsilon k)^2 \mp \frac{(1-5b)(1-b)}{8\sqrt{b}}(\epsilon k)^3 - i\frac{(1-2b)(1-b)}{2}(\epsilon k)^4 + O((\epsilon k)^5), \tag{4.4a}$$

$$\epsilon\omega_0 = -i + i(1-b)(\epsilon k)^2 + i(1-2b)(1-b)(\epsilon k)^4 + O((\epsilon k)^6). \tag{4.4b}$$

The very low wavenumber behavior of the dispersion relation for the propagating branches is that of the convection diffusion equation $w_t \pm \sqrt{b} w_x = \frac{1}{2}\epsilon(1-b)w_{xx}$. Petropoulos [14] observes that this PDE governs the long-time behavior of the Debye model because all components of the solution, except those associated with the low wavenumber portions of the propagating branches, decay with an $O(\epsilon^{-1})$ exponential rate.

For $\alpha < \epsilon k < \beta$ all branches of ω correspond to nonpropagating modes. On this interval the notations $\omega_\pm$ and ω_0 are somewhat inappropriate; the branches are better classified in terms of the rate of decay associated with them. For many applications these modes are of little interest because they do not propagate.

For large wavenumbers, the phase velocity of the propagating branches converges to unity and the associated modes decay like $\exp[-t(1-b)/(2\epsilon)]$. The nonpropagating mode decays like $\exp[-tb/\epsilon]$.

Small Wavenumber Analysis of Method Dispersion Relations

For the EP_d method, the eigenvalue problem, $\hat{M}\tilde{v} = e^{-i\tilde{\omega}\Delta t}\tilde{v}$ with $\hat{M}$ given in (3.1), determines the dispersion relation $\tilde{\omega}$. In the case $P = 1$, emphasized in this section, $\hat{M}$ can be derived explicitly by applying (AII.7) to (3.2). Explicit formulations of the eigenvalue problem for the dispersion relation $\bar{\omega}$ of the FDTD method of [5] are given in [2, 7]. In this section, we restrict our analysis to the case $\Delta t = \Delta x$. This is the usual situation as a unit Courant number is generally used to minimize numerical dispersion [2]. The analysis is presented in terms of the parameter $h = \Delta t/\epsilon$. Discrepancies between the dispersion relations of the model (ω) and a numerical method ($\tilde{\omega}$ or $\bar{\omega}$) indicate the magnitude and the nature of errors that the method introduces. Formulas for the small wavenumber behavior of $\bar{\omega}$ and $\tilde{\omega}$ facilitate comparison with the model dispersion relation ω,

$$\epsilon\bar{\omega}_\pm - \epsilon\omega_\pm = -\frac{(1-b)\sqrt{b}h^2}{24}(\epsilon k)^3 + i\frac{(1-b)(1-3b)h^2}{24}(\epsilon k)^4 + O((\epsilon k)^5), \tag{4.5a}$$

$$\epsilon\bar{\omega}_0 - \epsilon\omega_0 = -i12h^2F_1(h) - i\frac{(1-b)(1-3b)h^2}{12}(\epsilon k)^4 + O((\epsilon k)^6), \tag{4.5b}$$

$$\epsilon\tilde{\omega}_\pm - \epsilon\omega_\pm = -\frac{(1-b)\sqrt{b}h^3}{24}F_2(h)(\epsilon k)^3 - i\frac{(1-b)bh^3}{24}F_3(h)(\epsilon k)^4 + O((\epsilon k)^5), \tag{4.5c}$$

$$\epsilon\tilde{\omega}_0 - \epsilon\omega_0 = i\frac{(1-b)bh^3}{12}(\epsilon k)^4 + O((\epsilon k)^6), \tag{4.5d}$$

where

$$F_1(h) = 12\frac{\log[(2+h)/(2-h)]/h - 1}{h^2}, \tag{4.6a}$$

$$F_2(h) = 24\frac{e^{-h} - \left(1 - h + \frac{1}{2}h^2 - \frac{1}{6}h^3\right)}{h^4}, \tag{4.6b}$$

$$F_3(h) = 12\frac{\left[1 + (1+b)h + \frac{1}{4}h^2\right]e^{-h} - \left[1 + bh - \left(\frac{1}{4} + b\right)h^2 + \frac{1}{2}\left(\frac{1}{6} + b\right)h^3 - \frac{1}{4}bh^4\right]}{bh^4}. \tag{4.6c}$$

To leading order each of the F's equals one: $F_1(h) = 1 + O(h^2)$, $F_2(h) = 1 + O(h)$, and $F_3(h) = 1 + O(h)$. The leading order error in the propagating modes, $|\tilde{\omega}_\pm - \omega_\pm|$ and $|\bar{\omega}_\pm - \omega_\pm|$, is $O(k^3h^3)$ for the EP_d method and $O(k^3h^2)$ for the FDTD method. The leading

order error in the nonpropagating mode, $|\tilde{\omega}_0 - \omega_0|$ and $|\bar{\omega}_0 - \omega_0|$, is $O(k^4h^3)$ for the EP_2 method and $O(k^0h^2)$ for the FDTD method. That the error in the FDTD method enters at such low order in k for the nonpropagating mode indicates that it treats this mode very inaccurately; nonetheless, for many applications this is of little significance because the nonpropagating mode decays very rapidly. Hence, we focus on the performance of the methods for the propagating modes. For the propagating mode, the low wavenumber asymptotics suggest the principle difference between the methods is that the EP_2 method is better than the FDTD method by a factor of h.

Order One Wavenumbers with $P = 1$

To see the extent that this behavior persists for order one values of ϵk, we define $\tilde{C}(\epsilon k, h)$ and $\bar{C}(\epsilon k, h)$ by

$$\mathrm{Re}[\epsilon\tilde{\omega}_{\pm}(\epsilon k, h) - \epsilon\omega_{\pm}(\epsilon k)] = -\frac{(1-b)\sqrt{b}}{24}\tilde{C}(\epsilon k, h)(\epsilon k)^3h^3, \tag{4.7a}$$

$$\mathrm{Re}[\epsilon\bar{\omega}_{\pm}(\epsilon k, h) - \epsilon\omega_{\pm}(\epsilon k)] = -\frac{(1-b)\sqrt{b}}{24}\bar{C}(\epsilon k, h)(\epsilon k)^3h^2. \tag{4.7b}$$

If $\tilde{C}$ and $\bar{C}$ were uniformly bounded for all wavenumbers of interest as $h \to 0$ then the phase error of the EP_2 method would indeed be $O(h^3)$, while the phase error of the FDTD method would be $O(h^2)$. The scaling chosen gives $\tilde{C}(0, 0) = \bar{C}(0, 0) = 1$. Numerical computation of $\tilde{C}$ and $\bar{C}$ for values of h tending to zero indicates that these functions tend to limiting functions, $\tilde{C}(\epsilon k, 0)$ and $\bar{C}(\epsilon k, 0)$, as $h \to 0$. With $a = 0.99$ and $b = 0.01$, numerical approximations of the logarithm of the absolute value of the limiting functions, $\log|\tilde{C}(\epsilon k, 0.02)|$ and $\log|\bar{C}(\epsilon k, 0.1)|$, are plotted in Fig. 2. Different values of h are used in these two plots to obtain reasonable approximations of the limiting curve while minimizing the effects of round-off error. A difference in the nature of round-off errors is expected as $\tilde{C}$ is computed through division by h^3 while $\bar{C}$ is obtained through division by h^2. Computations with values of h ranging from $h = 1.0$ to $h = 0.001$ indicate these both reasonably represent the limiting functions. Plots for smaller values of h are not shown because large numerical errors resulting from round-off distort these plots. The effect of these errors can be seen in Fig. 2 for small values of $k\epsilon$. Numerical sampling and the finite

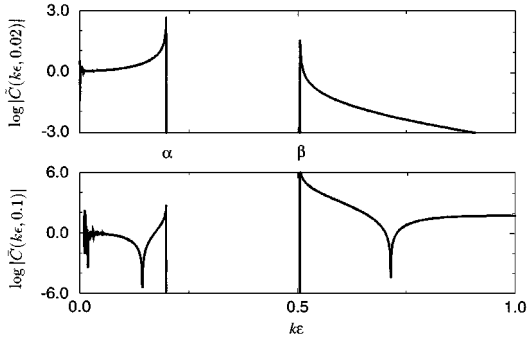

FIG. 2. Numerical approximations of the limit functions, $\log|\tilde{C}(k\epsilon, 0)$ and $\log|\bar{C}(k\epsilon, 0)|$, showing the nature of the phase error for small values of h.

values of h make the cusps shown in Fig. 2 finite; numerical evidence suggests that actually $\tilde{C}(\alpha, 0) = \tilde{C}(\beta, 0) = \bar{C}(\alpha, 0) = \bar{C}(\beta, 0) = \infty$. Thus, neither $\tilde{C}$ nor $\bar{C}$ is uniformly bounded. Nonetheless, Fig. 2 suggests that $\tilde{C}$ and $\bar{C}$ are uniformly bounded on $[0, 1] - (J_\alpha \cup J_\beta)$, where J_α and J_β are open intervals containing the points of singularity α and β, respectively. The nonuniformity of convergence near α and β arises from differences between the location of the singularities in dispersion relations for the model and the numerical methods. Figure 2 suggests that this mislocation is less severe for the EP_2 method than for the FDTD method. Moreover, $\tilde{C}$ is quite small for larger values of ϵk. The analogous study of the imaginary part yields similar results. The practical conclusion of these considerations is that, applied to the Debye model with waves not having a significant amount of energy in wavenumbers near α/ϵ or β/ϵ, the phase error in the EP_2 method is $O(h^3)$ while that in the FDTD method is $O(h^2)$.

Phase Error for Various Values of P

Seeking to confirm the general analysis of Section 3 for the Debye model with $P > 1$, we compute the dispersion relations ω and $\tilde{\omega}$ by numerical solution of the eigenvalues problems for $\hat{H}$ and $\hat{M}$ with $a = 0.99$ and $b = 0.01$. The comparison of the numerically computed values of ω and $\tilde{\omega}$ is made somewhat difficult by the extreme smallness of $\tilde{\omega} - \omega$ for smaller values of h; that is, for larger values of P and modest values of h, $\tilde{\omega} - \omega$ vanishes to machine accuracy (standard double precision). Of course, the difficulty of distinguishing $\tilde{\omega} - \omega$ from zero in finite precision arithmetic reflects the very high level of accuracy of the EP_d method. With $\epsilon k < \alpha$, no meaningful difference between $\tilde{\omega}$ and ω is found for values of $h < 5.0$. In this wavenumber regime, solutions generated by the EP_d method using as few as six grid points per wavelength are expected to proceed at near machine accuracy.

For $\epsilon k > \beta$, a small discrepancy between $\tilde{\omega}$ and ω can be found numerically. Figure 3 is a log–log plot of the real part of $\tilde{\omega} - \omega$ versus h with $\epsilon k = 1.0$. We examine this graph to determine the degree to which it agrees with the result of the previous section that the EP_{2P} method has $O(h^{2P+1})$ error for small h. We suppose that $\tilde{\omega} - \omega$ goes to zero as h^{η_P} and use the values of $\tilde{\omega} - \omega$ computed for $h \in [0.2, 1.0]$ to estimate η_P; we find $\eta_1 = 3.96$, $\eta_2 = 5.92$, $\eta_3 = 7.87$, $\eta_4 = 9.83$, and $\eta_5 = 11.8$. For $P > 1$, these values indicate rates of convergence almost one order better than $O(h^{2P+1})$. We do not interpret these results as saying that the method performs better than the theoretical predictions. Rather, we believe that the rate of convergence decreases when $\tilde{\omega} - \omega$ is examined for smaller

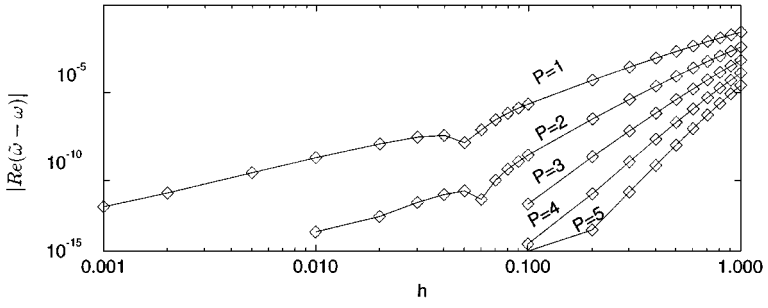

FIG. 3. Numerically computed values of the phase error, $|\mathrm{Re}(\tilde{\omega} - \omega)|$, versus h for several values of P and $\epsilon k = 1$. For $h < 0.1$ and $P \geq 2$, the computed error is virtually zero to machine accuracy.

values of h. For $P=1$, this hypothesis has been verified. For $P>1$, verification has not been performed due to our inability to distinguish $\tilde{\omega}-\omega$ from zero in double precision arithmetic. Figure 2 foreshadowed this result, as the coefficient $\tilde{C}(1, 0.02)$ seems to be very small. That the attempt to confirm the theoretical error estimates for the EP_d method numerically falls prey to round-off error in this way demonstrates that for larger values of P the method can be expected to perform near machine accuracy. Actual errors expected in computations using the method should generally be dominated by round-off error rather than discretization error.

Performance for $\Delta t/\epsilon \gg 1$

One potential application for the method of this paper is numerical computation in the regime where $\Delta t/\epsilon \gg 1$. Numerical experiments with the Debye model ($a=0.99$ and $b=0.01$) indicate roughly the same accuracy for the FDTD method and for the EP_2 method in the limit $\epsilon \to 0$ with Δt held constant. This result is consistent with the low wavenumber analysis of the dispersion relation obtain from (4.5a) and (4.5c) in the limit $\epsilon \to 0$:

$$\bar{\omega}_{\pm} - \omega_{\pm} = \frac{1}{4}(\tilde{\omega}_{\pm} - \omega_{\pm}) = -\frac{(1-b)\sqrt{b}\Delta t^2}{24}k^3. \tag{4.8}$$

Both methods have $O(\Delta t^2)$ errors for the Debye model when $\epsilon \to 0$. A similar analysis for simple Strang splitting indicates large $O(\Delta t)$ errors (see Appendix III). Although this example suggests the EP_d method offers no advantages over the FDTD method when $\Delta t/\epsilon \gg 1$, a closer examination reveals situations where the EP_d method gives superior performance in this regime.

The key to identifying these situations is the observation that the errors reported in (4.8) appear to be due to grid dispersion associated with a lack of alignment of the characteristics of the limiting model with the grid $\{(x_n, t_m)\}$; that is, we cannot expect any method using the EP_2 or basic FDTD stencils to give significantly better results. To strip away the effects of grid dispersion, we make comparisons in a case where the limiting characteristics are aligned with the mesh. Such a case is the Debye model with $a=3/4$ and $b=1/4$; the limiting speed is $\sqrt{b}=1/2$. With $\Delta t = 2\Delta x$ and $P=2$, the characteristics of the limiting solution are aligned with the grid, and numerical computations of the update matrices for the EP_4 method strongly indicate that they converge to the update matrices of an exact method when $\epsilon \to 0$ and Δt and Δx are fixed. Moreover, numerical computation of the dispersion relation shows that the method is stable for $\epsilon \geq 0$.

Aligning the grid with the characteristics of the FDTD method without violating the CFL condition may require widening the spatial stencil, which can be done in more than one way. One such extension of the FDTD method has the form

$$\frac{u_n^{m+1} - u_n^{m-1}}{2\Delta t} + A\sum_j w_j u_{n+j}^m + B\frac{u_n^{m+1} + u_n^{m-1}}{2\epsilon} = 0, \tag{4.9}$$

where the w_j's are scalars defining a finite difference approximation for the spatial derivative. The FDTD method is usually implemented on a staggered grid to take advantage of the structure of A and B; however, this transformation to a staggered grid for the sake of efficiency does not affect the analytical properties of the method.

In the limit $\epsilon \to 0$, the general solution of the Debye model is

$$u(x,t) = (\sqrt{b}, b, a)^t f(x - \sqrt{b}t) + (-\sqrt{b}, b, a)^t g(x + \sqrt{b}t). \tag{4.10}$$

Assuming that $\Delta t = 2\Delta x$, it can be shown that the limiting form of (4.9) supports these solutions if and only if $w_1 = 1/2\Delta x$, $w_{-1} = -1/2\Delta x$, and $w_j = 0$ for $j \notin \{-1, 1\}$. Hence, within this extended FDTD framework, there is a unique set of w_j's that gives an exact method in the limit $\epsilon \to 0$. With these w_j's, however, the method is unstable for $\epsilon > 0$. The origin of the instability is that with these w_j's waves travel at a speed of no more than $1/2$ in the numerical computations, while the maximum signal speed of the model is 1. Hence, it is impossible for the numerical computations to converge to the solution of the hyperbolic system no matter how fine the mesh is made.

For the Debye model with $a = 3/4$ and $b = 1/4$, we find that both the EP_d and FDTD approaches can provide methods that converge to an exact method in the limit $\epsilon \to 0$. For $\epsilon > 0$, the EP_d method is stable, but the FDTD method is unstable. The failure of the FDTD approach to provide a stable method converging to an exact method in the limit $\epsilon \to 0$ seems to stem from the rigidity of having the w_j's set independently of ϵ. In any case, this example demonstrate that there are situations where the EP_d method offers significant advantages over the FDTD method for computation in the regime $\Delta t/\epsilon \gg 1$.

5. EXTENSION FOR MATERIALS WITH SMOOTH INHOMOGENEITIES

In the applications to electromagnetics motivating this work, the electrical properties of materials are generally encoded in the matrix B, and these properties typically vary in space. Hence, we are interested in extending the EP_d method to the case of inhomogeneous materials where B depends on x. Here we limit attention to situations where B is a relatively smooth function of x.

When B is a constant matrix and the initial data are a polynomial, we have seen that the solutions of (1.1) are polynomials with time-dependent coefficients. Because this structure is lost when B depends on x, there is some ambiguity in generalizing EP_d to the inhomogeneous case. One approach is to use the method that advances polynomial data of degree d exactly for a single time step. A second approach is based on having solutions of the form of power series with time-dependent coefficients when B is analytic. The approach we consider here is based on the recursion relation given in (AI.7) and (AI.9).

An Extension for Inhomogeneous Materials

For simplicity we consider only the case $P = 1$; that is, for each point x_n in the spatial grid, we seek matrices $M_{-1}(n)$, $M_0(n)$, and $M_1(n)$ that are to be used to update the solution at that gridpoint. That a separate set of update matrices may be needed for each gridpoint implies that memory limitations and the cost of generating the update matrices will be significant considerations in implementations. To generate the update matrices for a mesh S_0 with grid spacing Δx we use a collection of refined meshes $S_1, S_2, \ldots, S_T$, each having grid spacings one-half of the one before. On the finest mesh S_T update matrices are generated at each gridpoint x_j assuming that $B(x) = B(x_j)$; that is, the formulas for generating update matrices in the homogeneous case, (2.7) and (2.9), are used with the material properties found at x_j. The recursion formula given in (AI.7) and (AI.9) is then applied to generate the update matrices for the coarser meshes including the mesh of interest S_0.

We note a few of the properties of the method generated by this algorithm. If B is a constant the algorithm produces the original EP_2 update matrices as defined by (2.8) and (2.9) on the coarsest grid S_0. In general, the update matrices on the coarse grid $M_p(n)$ depend on the values of $B(x)$ for values of $x \in \{x_n + \ell\Delta x/2^T : \ell = -2^T, 1-2^T, \ldots, 2^T-1, 2^T\}$. The computational effort needed to generate the update matrices on the coarse grid is proportional to 2^T. The computation can be arranged to avoid storing the update matrices on the fine grids; hence, the memory required to generate the method is essentially that needed to store the update matrices for the coarse grid. We limit our analysis of this algorithm to a numerical example.

A Numerical Example

In our example, a pulse is scattered from a material inhomogeneity surrounded by free space. We study the error in the transmitted wave at a certain point in space-time as a function of the numerical parameters, the grid spacing Δx and the truncation level T. As before, $\Delta t = \Delta x$. The initial data in the experiment are $u(x, 0) = 256(x-1)^4(x-2)^4(1, 1, 0)$ for $x \in [1, 2]$ and $u(x, 0) = 0$ for $x \notin [1, 2]$. These initial data give rise to a purely rightward propagating pulse which impinges on an inhomogeneity supported on the interval [3, 4]. The inhomogeneity is a Debye material ($a = 0.99$ and $b = 0.01$) where spatial variations enter through ϵ: $1/\epsilon(x) = 2560(x-3)^4(x-4)^4$.

We determined the numerically computed estimates of the first component of $u(5, 2.7)$ for 45 combinations of the numerical parameters T and Δx. Specifically, we use all combinations where $T \in \{4, 6, 8, 10, 12\}$ and $\Delta x \in \{0.1 \times 2^{-n} \mid n = 0, 1, 2, 3, 4, 5, 6, 7, 8\}$. The numerically computed values of the first component of $u(5, 2.7)$ are very nearly given by

$$0.183102359 + \frac{0.212}{2^T}\Delta x - 0.00325\Delta x^2 - 0.0819\Delta x^3 - 4.6\Delta x^4. \tag{5.1}$$

Using the FDTD method for the same problem and values of Δx results in numerically computed estimates of first component of $u(5, 2.7)$ very nearly given by

$$0.183102359 - 2.56\Delta x^2 + 2.04\Delta x^3 - 77.9\Delta x^4. \tag{5.2}$$

Figure 4 shows both the numerically computed values and curves given by (5.1) and (5.2). For the purposes of our discussion these formulas are an excellent characterization of the numerically computed values.

Interpretation of the Numerical Experiments

The main result from these experiments is that for practical values of Δx, the estimate of the first component of $u(5, 2.7)$ given by the extension of the EP_d method is far better than the estimate given by the FDTD method. As shown in Fig. 4, the accuracy of the extended EP_d method is better than that of the FDTD method when $\Delta x > 0.01$ even when $T = 4$. For the larger values of T the extended EP_d method gives a more accurate estimate than the FDTD method for $\Delta x = 0.000390625$—the smallest value of Δx considered in this experiment. The practical point here is illustrated by the observation that using the EP_d method with $T = 12$ and $\Delta x = 0.1$ the solution is accurate to three decimal places, while the FDTD method requires a mesh refined by a factor of eight to obtain the same accuracy.

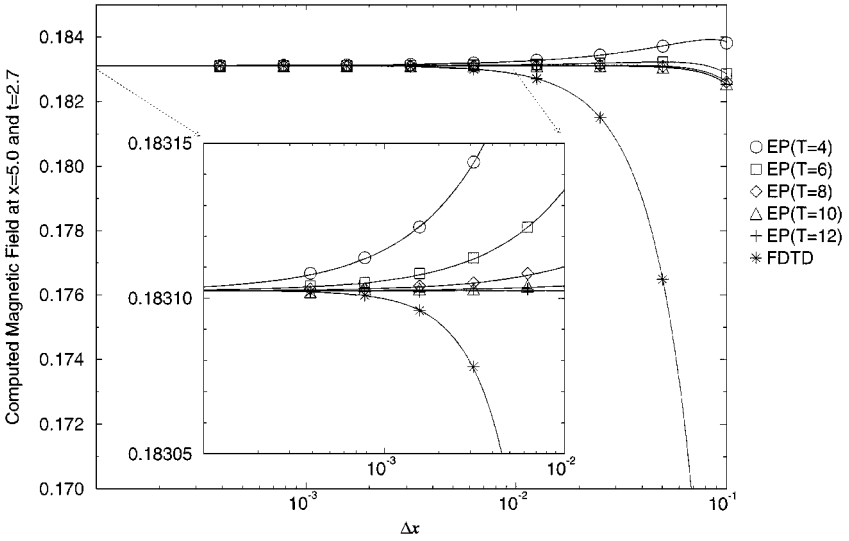

FIG. 4. Convergence of the numerically computed magnetic field values at $x = 5.0$ and $t = 2.7$ for extensions of the EP_d method and the FDTD method. The solid lines are the fitted formulas given in (5.1) and (5.2).

The significant difference in the accuracy of the methods is reflected in the three-orders-of-magnitude disparity between the coefficients of the quadratic terms in (5.1) and (5.2). Although this analysis examines but a single point in the solution of a single example, it suggests that the extended EP_d method potentially offers substantial advantages for practical computations which are typically run on grids having less than one hundred gridpoints per wavelength.

Comparisons of Resource Requirements

The advantages of using coarser grids than the FDTD method normally allows are partially offset by the additional resource requirements of the extended EP_d method. The amount of memory used per gridpoint is larger for the extended EP_d method due to the need to store separate update matrices for each gridpoint in the inhomogeneity. The number of arithmetic operations needed to update the solution at a gridpoint is greater by roughly a factor of two. Most significantly, the cost of generating the update matrices is substantial for the extended EP_d method (more than 27×2^T matrix multiplications per gridpoint) but negligible for the FDTD method. But we emphasize that this cost is paid only once, and not at each time step. Moreover, once the method has been generated for a given material configuration, it can be used repeatedly for different initial data. Nonetheless, more efficient algorithms for generating the update matrices are needed.

Further Extensions

Because the derivation of the recursion relation given in (AI.7) and (AI.9) assumes that the fields have a polynomial structure in space, we do not expect this extension to perform well in situations where the spatial structure of the fields is not approximated well by a lower order polynomial. This occurs when there are sharp gradients in the material properties. In the extreme case of a jump discontinuity in the material properties, the extension of the EP_d method considered here fails to converge when $\Delta x \to 0$. The development of alternative extensions is needed for the treatment of material interfaces.

The extension of the basic EP_d method to hyperbolic systems in several spatial variables is straightforward. However, the simple fact that the number of gridpoints neighboring a given point scales as the radius of the neighborhood to the power of the spatial dimension magnifies some of its disadvantages in more than one spatial dimension. Because the number of update matrices associated with a gridpoint equals the number of neighboring gridpoints, the memory requirements for storing these matrices can be substantial. Moreover, since the update procedure involves multiplying each of these update matrices by a vector, the computational cost of applying the method grows rapidly with the stencil radius. Of course, the hope is that the ability to compute accurate solutions on meshes with dramatically fewer grid points will overcome the somewhat high cost in memory and computation associated with each gridpoint.

6. SUMMARY AND COMMENTS

Summary

The reported need [2] for very fine meshes to resolve microwave pulses propagating in water motivated the development of the EP_d method for the stiff dispersive waves equation (1.1). From the point of view that fine meshes are needed because of the stiff system of ODEs, $\dot{u} + \epsilon^{-1} Bu = 0$, embedded in the problem, we are naturally drawn to consider how to incorporate the exact solution of this system of ODEs, $\exp(-Bt/\epsilon)u(0)$, into a method for (1.1). Our development of the EP_d method stems from considering the action of the Green's function on interpolating polynomials. Four alternative formulations of the EP_d method (Appendix I) provide information useful for its analysis and suggest approaches for its generalization. A defining characteristic of the EP_d method is that it is exact when applied to polynomial initial data of degree d. Our general analysis of the PE_d method is based on comparing its dispersion relation with that of the model solution to be simulated numerically. This comparison shows that in the generic case the EP_{2P} method has errors that are $O(\Delta x^{2P+1} k^{2P+1})$ for Δx and k small. In the case of the Debye model, a fuller analysis, including comparisons with the FDTD method and Strang splitting (Appendix III), is available owing in part to explicit represtation of the EP_2 method (Appendix II). This analysis demonstrates advantages of the EP_d method over both the FDTD method and Strang splitting for both large and small values of $h = \Delta t/\epsilon$ and confirms the expectation that very high accuracy can be achieved on practical meshes particularly with $P > 1$. For media with smooth inhomogeneities (B a smooth function of x), an extension of the EP_2 method is defined and compared to the FDTD method in an example. For meshes with practical grid spacing, the extension gives significantly better performance than the FDTD method.

Mitigation of Stiffness

The EP_d method is intended to mitigate the effects of stiffness in numerical simulations of the solutions of (1.1) when ϵ is small. For methods such as the FDTD method, obtaining accurately computed solutions often requires that the time step Δt be small compared with ϵ. This restriction on Δt easily leads to a prohibitively large requirement for computational resources when the desired computation time is large compared to ϵ. Our detailed considerations are limited to the Debye model. In the regime $\Delta t/\epsilon \ll 1$, the EP_2 method relieves the effects of stiffness, allowing larger time steps than for the FDTD method; that

is, even though we suppose $\Delta t \ll \epsilon$, the time step need not be as small as for the FDTD method to achieve the same level of accuracy. In the regime $\Delta t/\epsilon \gg 1$, the error for both the FDTD method and the EP_2 method are bounded independently of ϵ with $\Delta t = \Delta x$ fixed. However, in an example where the grid is aligned with the characteristics of the limiting wave equation, we show that the error for the EP_d method can go to zero with ϵ. This result is not possible with the FDTD method without loss of stability for $\epsilon > 0$. Hence, we have shown that there are situations where the EP_d method gives better performance than the FDTD method when $\Delta t \gg \epsilon$. A more complete characterization of these situations and a fuller assessment the quality of the EP_d method in these situations are needed.

Extensions of the EP_d Method

Our presentation focuses on the EP_d method for homogeneous materials in one spatial dimension. Naturally, these considerations are intended to form the foundation for extensions treating inhomogeneities, interfaces, and several spatial dimensions. One of several possible extensions to smoothly varying inhomogeneous materials is introduced in Section 5. Although analysis is limited to a numerical example, the potential of the extension to provide highly accurate solutions on relatively coarse grids is clear. More general analysis of the possible extensions to inhomogeneous materials is needed. Equally important is the need for algorithms that efficiently generate the update matrices for these extensions. Although interpolation by multivariate polynomials presents some complications, generalization of the EP_d method to several spatial dimensions is straightforward. For homogeneous materials, much of the analysis of Section 3 generalizes to several spatial dimensions as well. For large gradients in the material properties including material interfaces the proper generization of the EP_d method is less obvious. Seeking a quadrature formula for the representation of the update operator in terms of the Green's function seems an attractive approach. The major change is that we can no longer expect the spatial structure of the fields to be locally approximated by polynomials. To be acceptable, an alternative representation of this spatial structure must give rise to an accurate quadrature formula that can be incorporated into the general method without undue difficulty or cost. With the ability to handle both smooth and discontinuous inhomogeneities in several spatial dimensions, the EP_d method could serve as the basis for a general purpose code for simulating linear waves.

Computational Resources

The generization of the EP_d method to inhomogeneous materials in several spatial dimensions places two substantial burdens on computational resources: (1) Generation of the update matrices requires a very significant amount of arithmetic. (2) Storing the update matrices during a simulation is the dominant memory requirement of the method. In this paper the main approach to generating update matrices is based on (2.9), which requires the exponentiation of a large matrix. Even for inhomogeneous materials the matrix is highly structured. Exploiting this structure in the matrix exponentiation seems a very promising approach to easing the burden of computing the update matrices. The alternative approach to generating the update matrices, based on the recursion formula given in (AI.7) and (AI.9), has the unpleasant property that the amount of arithmetic needed roughly doubles with each application of that formula. The potential for dramatically improving the efficiency of this algorithm is obvious. The memory needed to store the update matrices is substantial. For electromagnetics problems in three spatial dimensions the memory requirements for the

update matrices could easily exceed those for the fields by a factor of one hundred. In contrast, the storage requirements of the FDTD method are often little more than the memory needed for the fields. Maintaining memory parity between the methods requires that the grid spacing of the EP_d method must be about a factor of five larger than that of the FDTD method. We note, however, that for many problems the memory requirements of the EP_d method can be substantially reduced. When there are large regions in the simulation domain where the material properties are constant, most of the gridpoints within such a region can share a single set of update matrices. Taking advantage of this structure can dramatically reduce both the cost of generating the update matrices and the memory needed to store them. Issues concerning computational resources will necessarily play an important role in the future development of the EP_d method.

APPENDIX I

Alternative Formulations of the EP_{2P} Method

Four alternative formulations of the EP_{2P} method provide useful information for its analysis and paths for its extension.

Local Polynomial Interpolation

This procedural formulation closely resembles that of the previous section except instead of using the Green's function to propagate the interpolating polynomial we use a system of ODEs describing the evolution of its coefficients. We use three steps to determine u_n^{m+1} from $u_{n-P}^m, \ldots, u_{n+P}^m$: (1) Find the interpolating polynomial (2.2). (2) Use this polynomial as initial data and solve (1.1) exactly. (3) Evaluate this solution at $t = \Delta t$ and $x = x_n$ to obtain u_n^{m+1}. The second step is accomplished using the observation that with polynomial initial data the solution of (1.1) is a polynomial whose coefficients evolve according to the system of ODEs $\dot{a}_q + A a_{q+1} + \epsilon^{-1} B a_q = 0$.

Limit of Finite Difference Methods

Truncation of Taylor series is a standard approach to generating finite difference methods. We obtain (AI.1a) by truncating the Taylor series for $u(x, t + \Delta t)$ after $T + 1$ terms and using (1.1) to substitute space derivatives for time derivatives. Expanding $(A\partial/\partial x + B)^\ell$ as a polynomial in $\partial/\partial x$ gives (AI.1b). Reversing the order of summation yields (AI.1c):

$$\sum_{\ell=0}^{T} \frac{1}{\ell!} \frac{\partial^\ell u}{\partial t^\ell}(x,t)\Delta t^\ell = \sum_{\ell=0}^{T} \frac{(-1)^\ell}{\ell!} \left(A\frac{\partial}{\partial x} + \epsilon^{-1} B \right)^\ell u(x,t)\Delta t^\ell \tag{AI.1a}$$

$$= \sum_{\ell=0}^{T} \frac{(-\Delta t)^\ell}{\ell!} \sum_{q=0}^{\ell} Z_{\ell q} \frac{\partial^q u}{\partial x^q}(x,t) \tag{AI.1b}$$

$$= \sum_{q=0}^{T} \left(\sum_{\ell=q}^{T} \frac{(-\Delta t)^\ell}{\ell!} Z_{\ell q} \right) \frac{\partial^q u}{\partial x^q}(x,t). \tag{AI.1c}$$

Replacing the spatial derivatives by the finite difference approximations represented by D_{qp}

gives a family of finite difference methods indexed by T and P:

$$u_n^{m+1} = \sum_{q=0}^{\min(T,2P)} \left(\sum_{\ell=0}^{T} \frac{(-\Delta t)^\ell}{\ell!} Z_{\ell q} \right) \sum_{p=-P}^{P} D_{qp} u_{n+p}^m. \tag{AI.2}$$

The primary result concerning this family of methods is that for a fixed value of P the method given in (AI.2) converges to the EP_{2P} method as $T \to \infty$ (with the corresponding value of P). A derivation of this result can be based on the following properties of the $Z'_{\ell q}$s:

$$Z_{00} = I, \tag{AI.3a}$$

$$Z_{\ell q} = 0 \qquad q < 0 \text{ or } q > \ell, \tag{AI.3b}$$

$$Z_{\ell+1 q} = Z_{\ell q-1} A + \epsilon^{-1} Z_{\ell q} B = A Z_{\ell q-1} + \epsilon^{-1} B Z_{\ell q}. \tag{AI.3c}$$

From these relations, $Z_{\ell 0}, \ldots, Z_{\ell 2P}$ can be identified with the first block column of $\mathcal{M}^q$. This identification, the representation of the W_q's as the first block column of $\exp(-\mathcal{M}\Delta t)$ given in (2.9), and the Taylor series representation of this exponential give the Taylor representation of $W_q(\Delta t)$:

$$W_q(\Delta t) = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{\ell!} Z_{\ell q} \Delta t^\ell. \tag{AI.4}$$

Applying (AI.4) to (AI.2) and comparing with (2.7) confirm that the method defined in (AI.2) converges to the method defined in the previous section when $T \to \infty$. This characterization of our method allows us to consider it a finite difference method with a spectrally accurate time-differencing scheme.

Exact Update Operator as an Exponential

The EP_{2P} method can be derived from the formal exponential representation of the update operator. The exact update operator for (1.1), in terms of an exponential, is formally given by

$$u(x, t + \Delta t) = e^{-(A(\partial/\partial x) + \epsilon^{-1} B)\Delta t} u(x, t) \tag{AI.5a}$$

$$= \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{\ell!} \left(A \frac{\partial}{\partial x} + \epsilon^{-1} B \right)^\ell \Delta t^\ell u(x, t) \tag{AI.5b}$$

$$= \sum_{q=0}^{\infty} W_q(\Delta t) \frac{\partial^q u}{\partial x^q}(x, t), \tag{AI.5c}$$

where we have made use of (AI.4) to obtain (AI.5c). The EP_{2P} method can be obtained from this representation of the exact update operator by replacing the spatial derivatives $(\partial^q/\partial x^q)$ in (AI.5c) by finite difference approximations given in terms of the D_{qp}'s. This formulation gives a simple characterization of the W_q's as the Taylor coefficients of $\exp[-(Az + \epsilon^{-1} B)\Delta t]$ expanded in z about $z = 0$.

Multigrid Interpretation

A multigrid formulation of the EP_{2P} method can be developed based on its characterization as the only method of the form (2.1) that is exact when applied to polynomials of degree $2P$ or less. For simplicity, we restrict our attention to the case $P = 1$. We show how to compute $M_p(2\Delta x, 2\Delta t)$ for $p = -1, 0, 1$ given $M_p(\Delta x, \Delta t)$ for $p = -1, 0, 1$. Iterative use of this procedure allows us to derive the method for a coarse grid given the method on a fine grid. The point, of course, is that ordinary finite difference methods give accurate approximations of the EP_{2P} method on sufficiently fine grids.

Repeated use of (2.1) relates u_n^{m+2} to $u_{n-2}^m, u_{n-1}^m, u_n^m, u_{n+1}^m, u_{n+2}^m$,

$$u_n^{m+2} = K_{-2}u_{n-2}^m + K_{-1}u_{n-1}^m + K_0u_n^m + K_1u_{n+1}^m + K_2u_{n+2}^m, \tag{AI.6}$$

where

$$K_p = \sum_\ell M_{p-\ell}(\Delta x, \Delta t)M_\ell(\Delta x, \Delta t) \tag{AI.7}$$

and $M_p = 0$ for $|p| > P = 1$. The update formula (AI.6) is exact when the five values $u_{n-2}^m, \ldots, u_{n+2}^m$ are obtained by sampling a quadratic polynomial. Moreover, in this case u_{n-1}^m and u_{n+1}^m are related to the other three sampled values:

$$u_{n\pm1} = \frac{3}{4}u_n + \frac{3}{8}u_{n\pm2} - \frac{1}{8}u_{n\mp2}. \tag{AI.8}$$

Elimination of $u_{n\pm1}^m$ from (AI.6) gives a method for computing u_n^{m+2} from the three values $u_{n-2}^m, u_n^m, u_{n+2}^m$ which is exact when the data is obtained from quadratic polynomials. Since the method of interest is the only method with this property, we have

$$M_0(2\Delta x, 2\Delta t) = \frac{3}{4}K_{-1} + K_0 + \frac{3}{4}K_1, \tag{AI.9a}$$

$$M_{\pm1}(2\Delta x, 2\Delta t) = K_{\pm2} + \frac{3}{8}K_{\pm1} - \frac{1}{8}K_{\mp1}. \tag{AI.9b}$$

The case of general P proceeds similarly. The computations become increasingly tedious as P increases, but they are easily automated. We note that this formulation is closely related to an addition formula for the W_q's:

$$W_q(t_1 + t_2) = \sum_{\ell=0}^{q} W_\ell(t_1)W_{q-\ell}(t_2). \tag{AI.10}$$

APPENDIX II

Analytic Expression of the EP_{2P} Method in Special Cases

In three important cases, the matrices defining the EP_{2P} method can be computed analytically using (AI.4): the case where A and B commute, the telegrapher's equation, and the Debye model for electromagnetic propagation in dielectric materials. For each case, we derive explicit analytic formulas for the matrices W_0, W_1, and W_2. These formulas are sufficient to define the method in the case $P = 1$, and they allow for more complete analysis of the method in these special cases.

Commuting Matrices

When A and B commute, $AB = BA$, we have

$$e^{-(Az+B)\Delta t/\epsilon} = e^{-B\Delta t/\epsilon} e^{-\Delta t Az}. \tag{AII.1}$$

It follows from (AI.5) that

$$W_q(\Delta t) = \frac{(-1)^q}{q!} e^{-B\Delta t/\epsilon} A^q \Delta t^q. \tag{AII.2}$$

The Telegrapher's Equation

The telegrapher's equation is equivalent to (1.1) with

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \tag{AII.3}$$

where a and b are nonnegative constants satisfying $a + b = 1$. Letting $\mu = (a - b)/2$, we have $B = \frac{1}{2}I + \mu H$, where H is the 2×2 diagonal matrix with diagonal elements 1 and -1. From $A^2 = H^2 = I$ and $AH + HA = 0$ we find

$$Z_{\ell 0} = \exp\left(-\frac{\Delta t}{2\epsilon}\right)\left(\frac{\mu}{\epsilon}\right)^{\ell} H^{\ell}, \tag{AII.4a}$$

$$Z_{\ell 1} = \exp\left(-\frac{\Delta t}{2\epsilon}\right)\left(\frac{\mu}{\epsilon}\right)^{\ell-1} A\chi_o(\ell), \tag{AII.4b}$$

$$Z_{\ell 2} = \frac{1}{2}\exp\left(-\frac{\Delta t}{2\epsilon}\right)\left(\frac{\mu}{\epsilon}\right)^{\ell-2} H^{\ell-2}(\ell - \chi_o(\ell)), \tag{AII.4c}$$

where χ_o is the indicator function for the odd integers ($\chi_o(q) = 1$ when q is odd integer and $\chi_o(q) = 0$ otherwise). Let $\rho = \mu\Delta t/\epsilon$; summation of (AI.4) yields

$$W_0(\Delta t) = \exp\left(-\frac{\Delta t}{2\epsilon}\right)[(\cosh\rho)I - (\sinh\rho)H], \tag{AII.5a}$$

$$W_1(\Delta t) = -\frac{\epsilon}{\mu}\exp\left(-\frac{\Delta t}{2\epsilon}\right)(\sinh\rho)A, \tag{AII.5b}$$

$$W_2(\Delta t) = \frac{\epsilon^2}{2\mu^2}\exp\left(-\frac{\Delta t}{2\epsilon}\right)[(\rho\sinh\rho)I + (\sinh\rho - \rho\cosh\rho)H]. \tag{AII.5c}$$

The Debye Model

The Debye model is defined in (4.1). It is easy to confirm that $A^3 = A$, $BAB = 0$, $BA^2B = aB$, and $B^2 = B$. We note that the relation $BAB = 0$ seems to be characteristic of models of dielectric materials. From these relations it follows from (AI.3) that

$$\epsilon^q Z_{q0} = B \qquad q \geq 1, \tag{AII.6a}$$

$$\epsilon^q Z_{q1} = \epsilon(AB + BA) \qquad q \geq 2, \tag{AII.6b}$$

$$\epsilon^q Z_{q2} = \epsilon^2(ABA + A^2B + BA^2 + (q-3)aB) \qquad q \geq 3. \tag{AII.6c}$$

Direct summation of (AI.4) yields

$$W_0(\Delta t) = I + B(e^{-\Delta t/\epsilon} - 1), \tag{AII.7a}$$

$$W_1(\Delta t) = -\Delta t A + \epsilon(AB + BA)\left(e^{-\Delta t/\epsilon} - 1 + \frac{\Delta t}{\epsilon}\right), \tag{AII.7b}$$

$$W_2(\Delta t) = \frac{1}{2}\Delta t^2 A^2 + \epsilon^2(A^2B + ABA + BA^2)\left[e^{-\Delta t/\epsilon} - 1 + \frac{\Delta t}{\epsilon} - \frac{\Delta t^2}{2\epsilon^2}\right] + \epsilon^2 aB\left[-e^{-\Delta t/\epsilon}\left(\frac{\Delta t}{\epsilon} + 3\right) + 3 - \frac{2\Delta t}{\epsilon} + \frac{\Delta t^2}{2\epsilon^2}\right]. \tag{AII.7c}$$

It is interesting that W_0, W_1, and W_2 have nontrivial limiting values when $\epsilon \to 0^+$.

APPENDIX III

Comparison with Strang Splitting

When $AB \neq BA$ much of the computational difficulty associated with (1.1) can be viewed as arising from the failure of the identity $\exp[-(A\partial/\partial x + \epsilon^{-1}B)\Delta t] = \exp[-A\partial/\partial x \Delta t]\exp[-\epsilon^{-1}B\Delta t]$; that is, the exact update operator for $u_t + Au_x + \epsilon^{-1}Bu = 0$ is not the update operator for $u_t + \epsilon^{-1}Bu = 0$ followed by that for $u_t + Au_x = 0$. In such situations, Strang splitting, which in the simplest case approximates $\exp[-(A\partial/\partial x + \epsilon^{-1}B)\Delta t]$ by $\exp[-\epsilon^{-1}B\Delta t/2]\exp[-A\partial/\partial x \Delta t]\exp[-\epsilon^{-1}B\Delta t/2]$ to third-order accuracy, is a natural approach. This approach suggests that the update operator for $u_t + Au_x + \epsilon^{-1}Bu = 0$ is approximated well by (1) applying the update operator for $u_t + \epsilon^{-1}Bu = 0$ for a half time step, (2) applying the update operator for $u_t + Au_x = 0$ for full time step, and (3) applying the update operator for $u_t + \epsilon^{-1}Bu = 0$ again for a half time step. A special attraction of this approach is that the EP_d method is exact for both $u_t + Au_x = 0$ and $u_t + \epsilon^{-1}Bu = 0$ in many cases including the Debye model. In these case an exact implementation of Strang splitting in the form (2.7) can be given with

$$W_\ell = \frac{(-1)^\ell}{\ell!} e^{-B\Delta t/2\epsilon} A^\ell e^{-B\Delta t/2\epsilon}. \tag{AIII.1}$$

For this version of Strang splitting applied to the Debye model, we derive the low wavenumber asymptotics of the dispersion relation of the method $\bar{\omega}$:

$$\epsilon\bar{\omega}_\pm = \sqrt{b}(\epsilon k) - i\frac{1-b}{2}\left[\frac{h\cosh(h/2)}{2\sinh(h/2)}\right](\epsilon k)^2 + O((\epsilon k)^3) \tag{AIII.2a}$$

$$\epsilon\bar{\omega}_0 = -i + i(1-b)\left[\frac{h\cosh(h/2)}{2\sinh(h/2)}\right](\epsilon k)^2 + O((\epsilon k)^4). \tag{AIII.2b}$$

We compare this with the low wavenumber expansion for ω given in (4.4):

$$\epsilon\bar{\omega}_\pm - \epsilon\omega_\pm = -i\frac{1-b}{24}[h^2 + O(h^4)](\epsilon k)^2 + O((\epsilon k)^3), \tag{AIII.3a}$$

$$\epsilon\bar{\omega}_0 - \epsilon\omega_0 = i\frac{1-b}{12}[h^2 + O(h^4)](\epsilon k)^2 + O((\epsilon k)^4). \tag{AIII.3b}$$

Hence, Strang splitting is at best $O(h^2)$ for small h. Moreover, the error is $O(k^2)$ rather than $O(k^3)$ as shown in (4.5c) and (4.5d) for the EP_2 method and as shown in (4.5a) for the propagating branches of the FDTD method. However, comparing (4.5b) and (AIII.3b) shows that Strang splitting treats the nonpropagating branches somewhat better that the FDTD method.

In the limit $\epsilon \to 0$, we find the dispersion relation for Strang splitting is

$$\bar{\omega}_{\pm} = \sqrt{b}k - i\frac{1-b}{4}\Delta t k^2 \qquad \text{and} \qquad \bar{\omega}_0 = -i + i\frac{1-b}{2}\Delta t k^2. \qquad \text{(AIII.4)}$$

Thus, in the small ϵ limit, the error for Strang splitting is no better than $O(\Delta t)$. These results are consistent with our numerical experience with broadband pulses in the regime where $\epsilon \ll 1$. In this regime, we find that the dominant error in solutions computed with the EP_2 and FDTD methods is an $O(\Delta t^2)$ phase error arising from grid dispersion while the dominant error in solutions computed with Strang splitting is an amplitude error that filters out the high wavenumber components of the solution. Even qualitatively correct numerical solutions can not be obtained in this regime using this simplest version of Strang splitting without hundreds of gridpoints per wavelength. Despite being exact when applied to constant initial data, the simplest version of Strang splitting is generally inferior to both the EP_2 and FDTD methods.

ACKNOWLEDGMENTS

The author is grateful to the two anonymous referees for the careful reviews and insightful comments. This work was supported by a grant (F49620-96-1-0039) from the Air Force Office of Scientific Research, Air Force Material Command, USAF. Computing equipment obtained through grants (DMS-9305665 and DMS-9407196) from the National Science Foundation was utilized. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation therein. The views and conclusions contained herein are those of the author and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research or the U.S. Government.

REFERENCES

1. R. A. Albanese, J. Penn, and R. Medina, Short-rise-time microwave pulse propagation through dispersive biological tissue, *J. Opt. Soc. Amer. A* **6**, 1441 (1989).
2. P. G. Petropoulos, Stability and phase error analysis of FD-TD in dispersive dielectrics, *IEEE Trans. Antennas Propagat.* **42**, 62 (1994).
3. K. L. Shlager and J. B. Schneider, A selective survey of the finite-difference time-domain literature, *IEEE Antennas Propagat. Mag.* **37**, 39 (1995).
4. K. S. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Trans. Antennas Propagat.* **14**, 302 (1966).
5. R. M. Joseph, S. C. Hagness, and A. Taflove, Direct time integration of Maxwell's equations in linear dispersive media with adsorption for scattering and propagation of femtosecond electromagnetic pulses, *Opt. Lett.* **16**, 1412 (1991).
6. R. Luebbers, F. P. Hunsberger, K. S. Kunz, R. B. Standler, and M. Schneider, A frequency-dependent finite-difference time-domain formulation for dispersive materials, *IEEE Trans. Electromagn. Compat.* **32**, 222 (1990).

7. C. Hile, J. H. C. Luke, and E. Gordon, Error analysis of finite-difference time-domain methods for pulse propagation in Debye materials, submitted for publication.
8. D. F. Kelly and R. J. Luebbers, Piecewise linear recursive convolution for dispersive media using FDTD, *IEEE Trans. Antennas Propagat.* **44**, 792 (1996).
9. T. M. Roberts and P. G. Petropoulos, Asymptotics and energy estimates for electromagnetic pulses in dispersive media, *J. Opt. Soc. Amer. A* **13**, 1204 (1996).
10. C. Moler and C. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, *SIAM Rev.* **20**, 801 (1978).
11. G. H. Golub and C. F. Van Loan, *Matrix Computations* (The Johns Hopkins Univ. Press, Baltimore/London, 1996), 3rd ed., p. 572.
12. T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin/Heidelberg/New York, 1976), p. 63.
13. R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge Univ. Press, Cambridge/New York/Melbourne, 1985), p. 365.
14. P. G. Petropoulos, The wave hierarchy for propagation in relaxing dielectrics, *Wave Motion* **21**, 253 (1995).