SIAM J. NUMER. ANAL.
Vol. 35, No. 1, pp. 31–50, February 1998

003

TIMESTEP ACCELERATION OF WAVEFORM RELAXATION*

B. LEIMKUHLER†

Abstract. Dynamic iteration methods for treating certain classes of linear systems of differential equations are considered. It is shown that the discretized Picard–Lindelöf (waveform relaxation) iteration can be accelerated by solving the defect equations with a larger timestep, or by using a recursive procedure based on a succession of increasing timesteps. A discussion of convergence is presented, including analysis of a discrete smoothing property maintained by symmetric multistep methods applied to linear wave equations. Numerical experiments indicate that the method can speed convergence.

Key words. waveform relaxation, wave equation, multigrid methods

AMS subject classification. 65L05

PII. S003614299528002X

1. Introduction.

Much of modern chemical and physical research relies on the numerical solution of various wave equations. Since these problems are extremely demanding of both storage and cpu-time, new numerical methods and fast algorithms are needed to make optimal use of advanced computers. The *dynamic iteration* or *waveform relaxation* (WR) method [8, 10] is an iterative decoupling scheme for ordinary differential equations which can facilitate concurrent processing of large ODE systems for applications such as VLSI circuit simulation [14] and PDEs [1, 4].

In this article, accelerated dynamic iteration schemes are used to solve systems of linear differential equations, with emphasis on the ODEs arising from discretization of linear wave equations. Although our experiments use finite differences for the spatial derivatives, other spatial discretizations could be used. For time discretization, we use symmetric multistep methods, although other choices may also be appropriate. As is the case for stationary iterative methods applied to spatially discretized elliptic PDEs, it is found that finer fixed step (time) discretizations slow the convergence of the WR iteration, while large timesteps can be used to resolve the slow modes. The idea that is explored here is to use a coarse timestep on the defect equations to speed up convergence of the fine grid iteration. Nevanlinna already pointed out [13, 12] that for general applications of WR it makes sense from an efficiency standpoint to use coarser discretization in the early sweeps (when the iteration error is large), and then to refine the time discretization incrementally near convergence. Our point of view is rather to vary the timestep to resolve different modes present in the solution, using two time stepsizes (or multiple stepsizes). The current article is related to recent work of Horton and Vandewalle [4] and Horton, Vandewalle, and Worley [5], which considered space-time multigrid methods for solving parabolic equations.

The new scheme will be referred to as *timestep acceleration* since it relies on adjustment of the integration timestep to accelerate the dynamic iteration. This approach shares some features of multigrid methods. For the convenience of the

*Received by the editors January 13, 1995; accepted for publication (in revised form) July 31, 1996. This work was supported by NSF grant DMS-9303223.
http://www.siam.org/journals/sinum/35-1/28002.html
†Department of Mathematics, University of Kansas, Lawrence, KS 66045 (leimkuhl@math.ukans.edu).

reader generally familiar with multigrid methods, we outline the algorithm in the abstract setting of solving an unspecified dynamical system as follows.

ACCELERATED WAVEFORM RELAXATION. Given: fine timestep h and an approximation u^0 to the solution with fixed stepsize h.

1. **Smoothing.** Starting from u^0, perform a fixed number of iterations of a smoothing WR iteration with timestep h.
2. **Correction.** Compute the defect (residual) in this solution on the fine time-mesh.
   If the timestep $H = qh$ is sufficiently large, solve the discrete defect equation restricted to the coarse time-mesh directly (i.e., without relaxation).
   Else recursively apply some number of iterations of the algorithm using stepsize H to the defect equation restricted to the coarse time-mesh.
   Next, correct the solution after prolonging onto the fine time-mesh.
3. **Smoothing.** Apply a fixed number of iterations of the fine stepsize smoothing iteration.

(In section 6, we present and analyze a more precisely defined version of this algorithm, TAWR.)

A major barrier to efficient solution of large scale wave equations is the need for small timesteps. Due to the sequential character of standard ODE methods, this effectively reduces the potential for parallel speedups. Compared to standard timestepping schemes, the method discussed here directly addresses this problem by enabling the use of larger timesteps to recover at least a portion of the dynamics. Another important obstacle to computation—particularly in the case of high-dimensional problems—is the necessary storage. The new method actually exacerbates this problem since solution information at many points must be stored. However, in WR based on a block splitting, the storage is naturally segmented according to the decoupling, so the scheme may be appropriate for a parallel computer based on a distributed memory architecture.

Although standard analytical results for multigrid methods or coarse grid acceleration are typically developed for finite-dimensional Hermitian positive definite problems, these can be relaxed to give at least partial convergence results. In fact, proving theoretical convergence for timestep acceleration is easier than for standard multigrid due to the strong smoothing properties of the Picard–Lindelöf operator (it is a contraction on small intervals). Analysis of the behavior of the iteration on special linear model problems is also possible and is briefly discussed here.

The scheme is found to work well in simple numerical experiments with linear wave equations. Although our experiments are conducted in one space dimension, nothing in principle prevents application in higher dimensions (although many practical issues will need to be dealt with).

2. Waveform relaxation. Consider a second-order linear system of differential equations

$$\ddot{u} = Au, \qquad u(0) = u_0, \qquad \dot{u}(0) = \dot{u}_0, \tag{1}$$

where the eigenvalues of the matrix A are assumed to lie in the left half-plane. A special case that we will frequently refer to is the one-dimensional wave equation $U_{tt} = U_{xx}$ discretized with finite differences on the unit square with periodic boundary conditions:

$$A = \frac{1}{\Delta x^2}\begin{bmatrix} -2 & 1 & 0 & \cdots & 0 & 1 \\ 1 & -2 & 1 & & & 0 \\ 0 & 1 & -2 & 1 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & & & 1 & -2 & 1 \\ 1 & 0 & \cdots & 0 & 1 & -2 \end{bmatrix};$$

A is called the discrete Laplacian. Here $u = (u_0(t), \ldots u_{N-1}(t))^t$ is a vector of approximations at nodes x_i, $i = 0, \ldots, N-1$, with $x_{i+1} = x_i + \Delta x$.

Another potential application is to the Schrödinger equation. Discretizing, for example, with finite differences leads to

$$\frac{d}{dt}\Psi = -i(A + V(t))\Psi =: -iB(t)\Psi, \tag{2}$$

where A is the discrete Laplacian. If $v(x,t)$ is the potential energy function of the corresponding classical system, we have $V(t) = \mathrm{diag}(v(x_0,t), v(x_1,t), \ldots, v(x_{N-1},t))$. In simplified settings $v(x,t)$ is time-independent, hence so is V.

The WR method for (1) is based on a splitting $A = A_+ - A_-$, and results in ODE IVPs:

$$\ddot{u}^{l+1} = A_+ u^{l+1} - A_- u^l, \qquad u^{l+1}(0) = u_0, \qquad \dot{u}^{l+1}(0) = \dot{u}_0. \tag{3}$$

For example, we might choose A_+ to be the diagonal of A (Jacobi splitting), a block-diagonal part of A (block-Jacobi splitting), the lower triangular part of A (Gauss–Seidel splitting), etc. Much work involving the discrete Laplacian in elliptic PDEs is based on Gauss–Seidel splitting in red-black ordering. Another useful splitting is the damped Jacobi splitting where $A_+ = \frac{1}{\omega}D$, with D the diagonal of A. The extreme case, $A_+ = 0$, $A_- = -A$, is called the Picard splitting.

When referring to (1) and (3) we will generally limit discussion to the case where A and A_+ are symmetric negative semidefinite matrices.

The WR iteration proceeds as follows: starting from a given *initial waveform* $u^0 = u^0(t)$ (which may be constant), we solve (3) with $l = 0$ as a forced linear system for u^1 over some time interval, say $[0, T]$. (This interval is referred to as the *window*.) The function u^1 then yields a forcing for the next iteration or *sweep*, and the process repeats. In practice, the systems are solved numerically over the entire interval, and the storage of the resulting discrete approximation is an important drawback of the method which may place severe limitations on the size of the time window. On the other hand, we gain in two ways: first, the systems we solve at each iteration can be decoupled into problems of reduced dimension, and second, the decoupled problems can often be solved on separate processors of a parallel computer. An alternative approach would be based on solving the linear equations that result at each step of a standard discretization using a parallel algorithm. However, depending on the computer architecture employed, the flexibility in the choice of window size may reduce the overall communication cost, e.g., by eliminating some of the time spent in initializing the transfer of data between processors.

Preliminary convergence results for WR appear in the paper by Lelarasmee, Ruehli, and Sangiovanni-Vincentelli [8]. Miekkala and Nevanlinna [10, 11] and Nevanlinna [12] have developed an extensive theory for studying WR for linear systems. Lubich and Ostermann proposed to combine the WR method with spatial multigrid schemes [9]. Recent work by Horton and Vandewalle [4] and by Horton, Vandewalle, and Worley [5] has shown that a careful implementation of (spatial) multigrid WR methods for parabolic PDEs can provide excellent parallel speedups. The use of WR for solving hyperbolic partial differential equations and relations to domain decomposition was explored by Bjørhus [1].

3. Mathematical background. In this section we state some elementary results concerning the iteration (3). The reader is directed to the papers of Nevanlinna and Miekkala for basic theory.

The WR method for (1) can be viewed as an iteration $u^{l+1} = \mathcal{S}u^l$ in $L^2([0,T])$. As shown in [10], $\rho(\mathcal{S}) = 0$, implying superlinear convergence. On the other hand, for stiff dissipative linear systems, it makes sense to allow $T \to \infty$, in which case meaningful spectral information is obtained [10]. Since the solution to the equations (1) and (2) does not generally lie in $L^2([0,\infty))$, this approach must be modified. A reasonable practical approach is that taken in [12], where an exponential weighting function $e^{-\alpha t}$ is inserted into the usual L^2 norm. For $\alpha > 0$, the space L^2_α is normed by

$$\|u\|_\alpha := \left[\int_0^\infty |e^{-\alpha t}u(t)|^2 dt\right]^{1/2},$$

and ρ_α refers to spectral radius in that space.

If we take the Laplace transform of (3), we obtain

$$z^2\hat{u}^{l+1} = A_+\hat{u}^{l+1} - A_-\hat{u}^l + \phi(z),$$

where $\phi(z) = zu(0) + \dot{u}(0)$. Define $S(z) = -(z^2 I - A_+)^{-1}A_-$.

The following results are proved by Nevanlinna and Miekkala [10]:

$$\begin{aligned}\rho_\alpha(\mathcal{S}) &= \max_{\mathrm{Re}z \geq \alpha} \rho(S(z)) \\ &= \max_{y \in \mathbf{R}} \rho(S(\alpha + iy)),\end{aligned}$$

which follows from the Paley–Wiener theorem (the second expression follows from a maximum principle after a suitable remapping of the domain) and

$$\|\mathcal{S}\|_\alpha = \max_{y \in \mathbf{R}} |S(\alpha + iy)|,$$

which follows from Parseval's identity.

We now provide some simple estimates for the response of the iteration operator in weighted 2-norm.

First, consider the behavior of the solution operator $\mathcal{L}^{-1} = (\mathcal{D}^2 I - A)^{-1}$ of (1) in the weighted space, where $\mathcal{D} = d/dt$. Examining the spectral radius of the normal matrix $L(z)^{-1} = (z^2 - A)^{-1}$ along the line $\mathrm{Re}z = \alpha$, we find the eigenvalues are:

$$\mu_i(z) = \frac{1}{z^2 - \lambda_i}, \qquad i = 1, 2, \ldots, N,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_N$ are the eigenvalues of A. Hence

$$|\mu_i| = \frac{1}{|\alpha^2 - y^2 + 2\alpha y i - \lambda_i|} = \frac{1}{\sqrt{(y^2 + \alpha^2 + \lambda_i)^2 - 4\alpha^2\lambda_i}}.$$

By maximizing these functions over y, we can compute the moduli of the eigenvalues of solution operator in the weighted space.

THEOREM 3.1. *Define*

$$\bar{\mu}_i := \begin{cases} \frac{1}{2\alpha\sqrt{|\lambda_i|}}, & \alpha < \sqrt{|\lambda_i|}, \\ \frac{1}{\alpha^2 + |\lambda_i|}, & \alpha \geq \sqrt{|\lambda_i|}. \end{cases}$$

Then

$$\rho_\alpha(\mathcal{L}^{-1}) = \max_{i=1,\ldots,N} \bar{\mu}_i.$$

In particular, eigenvalues near zero have the strongest influence. When A is the discrete Laplacian, or any symmetric negative semidefinite matrix with an eigenvalue at zero, we have

$$\rho\left(((\alpha + iy)^2 I - A)^{-1}\right) = \frac{1}{\alpha^2}.$$

We can use this to estimate the norm of the iteration matrix, since

$$|(z^2 I - A_+)^{-1} A_-| \leq \max_{i=1,\ldots N} \bar{\mu}_i |A_-|,$$

where the $\bar{\mu}_i$ are determined from Theorem 3.1 with λ_i the eigenvalues of A_+ rather than A.

Asymptotic ($\alpha = \mathrm{Re} z \to \infty$) estimates for the relation between spectral radius of $S(z)$ and α are given in [7].

Let us consider the wave equations with periodic boundary conditions on the square as a model problem. We will use damped Jacobi iteration with parameter $\omega \in (0, 1]$. In this case, A, A_+, and A_- are all diagonalized by the discrete Fourier transform, so we arrive readily at the eigenvalues η_j, $j = 1 \ldots N$, of $S(\alpha + iy)$:

$$\eta_j = \mu\gamma_j,$$

where

$$\gamma_j = \frac{2}{\Delta x^2}\left(\left(1 - \frac{1}{\omega}\right) - \cos(\theta_j)\right), \qquad \theta_j = 2\pi(j-1)/N,$$

and

$$\mu(z) = \frac{1}{z^2 + \frac{2}{\omega\Delta x^2}}.$$

In this case the spectral radius can be readily computed. We have

$$\max_{y \in \mathbf{R}} |\mu(\alpha + iy)| = \begin{cases} \frac{\Delta x\sqrt{\omega}}{2\sqrt{2}\alpha}, & \alpha < \frac{\sqrt{2}}{\sqrt{\omega}\Delta x}, \\ \frac{1}{\alpha^2 + \frac{2}{\omega\Delta x^2}}, & \alpha \geq \frac{\sqrt{2}}{\sqrt{\omega}\Delta x}, \end{cases}$$

and

$$\rho_\alpha(\mathcal{S}) = \begin{cases} \frac{1}{\alpha\sqrt{2\omega}\Delta x}, & \alpha < \frac{\sqrt{2}}{\sqrt{\omega}\Delta x}, \\ \frac{1}{1+\frac{\alpha^2\omega\Delta x^2}{2}}, & \alpha \geq \frac{\sqrt{2}}{\sqrt{\omega}\Delta x}. \end{cases}$$

We are interested in *moderate* weights α which we define to mean $\alpha < \sqrt{\rho(A)}$. (Intuitively, this corresponds to looking in the time domain on intervals greater than the smallest period of the motion.)

In the standard theory, one uses the value of ω in the damped Jacobi splitting to enhance a *smoothing property*: a damping in the iteration of the modes corresponding to larger eigenvalues. However, the important consideration for timestep acceleration is not the way in which the smoother acts on fast "spatial" modes, but rather the response of the smoother to high frequency forcings. In fact, the real smoothing property we are interested in has to do with the shape of the graph of $\rho(S(\alpha + iy))$ as a function of y. For example, when a damped Jacobi splitting is applied to solving the semidiscrete wave equation, we find that the spectral radius of S achieves its maximum on $\mathrm{Re}z = \alpha$ at the point (if $\alpha < \frac{\sqrt{2}}{\sqrt{\omega}\Delta x}$)

$$y^2 + \alpha^2 - \frac{2}{\omega\Delta x^2} = 0$$

(or at $y = 0$ if $\alpha > \frac{\sqrt{2}}{\sqrt{\omega}\Delta x}$). The maximum is typically achieved well away from $y = 0$.

For the Picard splitting, it is easy to see rather that the maximum occurs at $y = 0$. In this case, we say that the iteration has a smoothing property with respect to high frequency forcings. It is not necessary to use a slowly converging splitting such as the Picard splitting to obtain a good smoothing property. A typical feature of a good splitting for this purpose is that A_+ would have an eigenvalue at or near the origin. Thus the smoothing property of a block-Jacobi splitting of the discrete Laplacian would improve with the block size.

To illustrate this smoothing concept, consider the time-dependent Schrödinger equation (2). The iteration becomes

$$\frac{d}{dt}\Psi^{l+1} = -iB_+\Psi^{l+1} + iB_-\Psi^l.$$

We could again use a Jacobi or damped Jacobi splitting, but in practice, a more useful choice might be

$$B_+ = -(A + V(t_0)), \qquad B_- = (V(t) - V(t_0)),$$

or more simply,

$$B_+ = -A, \qquad B_- = V(t).$$

After time discretization, these choices will lead to equations at each timestep which can be efficiently solved by, for example, using a parallel implementation of the *fast Fourier transform* (FFT).

Still another possibility is to work directly in the Fourier coefficients. Let $QAQ^H = \Lambda$, where

$$Q = (q_{mn}), \qquad q_{mn} = e^{2\pi i(m-1)(n-1)/N},$$

and

$$\Lambda = -2I + 2\mathrm{diag}(1, \cos(2\pi/N), \cos(4\pi/N), \ldots, \cos(2(N-1)\pi/N)).$$

Set $Q^H\bar{\Psi} = \Psi$ so that the equations become

$$\frac{d}{dt}\bar{\Psi} = -i(\Lambda + QV(t)Q^H)\bar{\Psi}.$$

We can then apply a Jacobi splitting to this problem. One finds that the diagonal of $QV(t)Q^H$ is dI where $d = v(x_0,t) + v(x_1,t) + \cdots + v(x_{N-1},t)$, so

$$B_+ = -(\Lambda + dI), \qquad B_- = (QV(t)Q^H + dI). \tag{4}$$

The computation B_-w can be implemented efficiently using the FFT.

The symbol of the WR iteration operator $\mathcal{R}$ for the Schrödinger equation with $V(t) =$ constant is

$$R(z) = (zI + iB_+)^{-1}iB_-,$$

and, trivially,

$$\|\mathcal{R}\|_\alpha \le \frac{1}{\alpha}|B_-|.$$

Using the Laplacian+potential splitting or one of its cousins can be expected to yield a good smoothing property with respect to high frequency forcings.

4. Discretization. In this section, we focus on (1) and apply a discrete transform as in [11] to analyze the symmetric multistep methods [3, 6] commonly used for integrating oscillatory problems.

Multistep methods construct an approximating sequence $\{u_n\}$ to $\{u(t_n)\}$ at successive time points $t_n = nh$. We use $\{u_n^k\}$ to refer to the numerical solution generated at the kth sweep of WR. Symmetric multistep methods for $\ddot{u} = f(t,u)$ take the form

$$\sum_{i=0}^{k} \alpha_i u_{n-i} = h^2 \sum_{i=0}^{k} \beta_i f(t_{n-i}, u_{n-i}), \tag{5}$$

with $(\alpha_0, \alpha_1, \ldots, \alpha_k)$ and $(\beta_0, \beta_1, \ldots, \beta_k)$ "palindromic" sequences, for which $\alpha_0 = \alpha_k$, $\beta_0 = \beta_k$, $\alpha_1 = \alpha_{k-1}$, $\beta_1 = \beta_{k-1}$, etc. These methods are used for integration of second-order oscillatory problems. An important feature of this class of methods is their time-reversibility. Note that the multistep methods require k *starting values*.

In discretizing dissipative problems it is sensible to replace the space L^2 by l_h^2 with norm $\|\{u_n\}\|_h = (h\sum |u_n|^2)^{1/2}$. For our investigations, we use the weighted space with norm

$$\|\{u_n\}\|_{h,\alpha} = \left(h\sum |e^{-nh\alpha}u_n|^2\right)^{1/2},$$

which can be viewed as a discretization of the L^2_α norm.

Following the usual practice we define operators a and b on sequences by

$$a\{u_n\} = \left\{\sum_{i=0}^{k} \alpha_i u_{n-i}\right\}, \qquad b\{u_n\} = \left\{\sum_{i=0}^{k} \beta_i u_{n-i}\right\}.$$

We also use the symbols a and b to refer to the corresponding characteristic polynomials:

$$a(\zeta) = \sum_{i=0}^{k} \alpha_i \zeta^{k-i}, \qquad b(\zeta) = \sum_{i=0}^{k} \beta_i \zeta^{k-i}.$$

To preserve the intuitive correspondence of results from the continuous time to discrete worlds, define a discrete transform which takes $\{u_n\}$ to $\hat{u}(z)$ by $\hat{u}(z) = h\sum_0^\infty e^{-nhz}u_n$ (essentially a discretization of the Laplace transform, and equivalent to the discrete Laplace or ζ transform).

Applying (5) to the linear problem (3) and computing the discrete transform, we find that

$$\hat{u}^{l+1} = S_h(z)\hat{u}^l + \phi,$$

where

$$S_h(z) = -\left(\frac{1}{h^2}\frac{a(e^{hz})}{b(e^{hz})}I - A_+\right)^{-1} A_-,$$

and ϕ includes the effects due to the k starting values. We are going to assume that these starting values are exact (for the unsplit discrete problem) so they do not affect the convergence of the iteration.

For example, if $k = 2m$, we find

$$S_h(z) = -\left(P_h(z)I - A_+\right)^{-1} A_-,$$

where

$$P_h(z) = \frac{1}{h^2}\frac{\sum_{j=0}^{m-1}\alpha_j \cosh(jhz) + \alpha_m}{\sum_{j=0}^{m-1}\beta_j \cosh(jhz) + \beta_m}.$$

In the general case, discrete versions of the Paley–Wiener theorem and Parseval's identity give (after modifying results in [11] to take into account the exponential weight):

$$\rho_{h,\alpha}(\mathcal{S}_h) = \max_{\mathrm{Re}z \geq \alpha} \rho(S_h(z))$$

and

$$\|\mathcal{S}_h\|_{h,\alpha} = \max_{y\in\mathbf{R}} |S_h(\alpha + iy)|.$$

In order for the discretized operator to be bounded, we evidently need to require

$$a(e^{hz}) - h^2\lambda b(e^{hz}) \neq 0, \qquad \mathrm{Re}z \geq \alpha$$

for any $\lambda \in \sigma(A_+)$.

We now consider an example. Ignoring rounding error, the popular leapfrog method for second-order systems is equivalent to Störmer's rule (also known as the Verlet method), a symmetric two-step method with $\alpha_0 = \alpha_2 = 1$, $\alpha_1 = -2$, and

$\beta_0 = \beta_2 = 0$, $\beta_1 = 1$. Applying this scheme to the WR iteration for the linear problem and taking the discrete transform gives

$$S_h(z) = -\left(\frac{2}{h^2}(\cosh(hz) - 1)I - A_+\right)^{-1} A_-.$$

The function

$$P_h^{\text{s.r.}}(z) = \frac{2}{h^2}(\cosh(hz) - 1) \tag{6}$$

is an $O(h^2 z^4)$ approximation to z^2. The poles of the transformed discrete iteration operator satisfy

$$z = \frac{1}{h}\cosh^{-1}\left(1 + \frac{h^2\lambda}{2}\right),$$

with λ an eigenvalue of A_+. Explicit multistep schemes are always conditionally stable, meaning that the stability of the schemes will depend on the stepsize being restricted roughly in inverse relation to the square root of the spectral radius of A_+. For the Störmer method, the stability condition is that $\lambda < 0$ and $-h^2\lambda \leq 2$, which is also the condition that the poles of S_h remain on the imaginary axis. The function $\operatorname{Im}\cosh(\xi)$ is monotone in the real variable ξ on $[-1, 1]$, hence the *ordering* of the poles is preserved along the imaginary axis.

Another popular second-order method is the (implicit) trapezoidal rule which has transform

$$P_h^{\text{t.r.}}(z) = \left(\frac{2}{h}\frac{e^{hz} - 1}{e^{hz} + 1}\right)^2. \tag{7}$$

4.1. Decay of the discrete symbol. Theory due to Miekkala and Nevanlinna [11] compares the convergence of the discrete iteration in l_h^2 to that of the continuous time iteration for dissipative problems and for methods that are not weakly stable. We need to modify this mechanism to cover convergence for stable methods for second-order differential equations in the weighted spaces rather than L^2 and l_h^2. In what follows, it is assumed that the k starting values are held fixed as we iterate. These could also be obtained by some convergent process, but this does not seem a meaningful generalization.

In the case of the discretized iteration we need to examine the images of segments $\alpha + iy$, $-\infty < y < \infty$, under P_h. The situation for $\alpha = 1$ is representative, and in Figure 1 we see the images of this line for the trapezoidal and Störmer discretizations, for various values of the stepsize.

Putting real α into each of the functions $P_h^{\text{t.r.}}$, $P_h^{\text{s.r.}}$, and z^2 one can show that for sufficiently small $h\alpha$

$$P_h^{\text{t.r.}}(\alpha) < \alpha^2 < P_h^{\text{s.r.}}(\alpha), \qquad \alpha > 0,$$

which means, somewhat surprisingly, that in the neighborhood of $y = 0$, the Störmer discretization actually leads to a slightly *more stable* overall iteration than that generated by the trapezoidal rule.

The situation for large h is more dramatic. For nonstiff problems with eigenvalues λ very near the origin in the complex plane, large steps should be possible and one

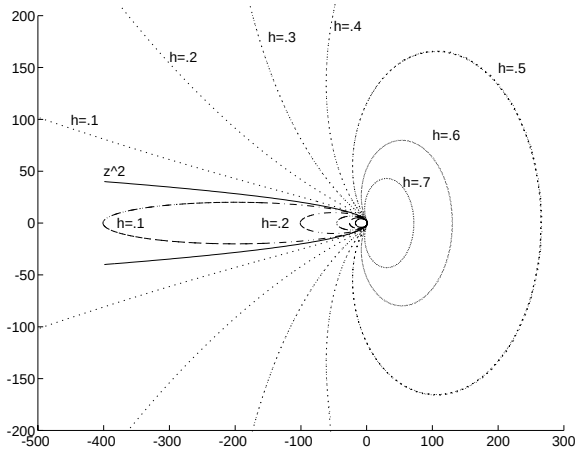

FIG. 1. *Approximation of z^2 by Störmer (dashed lines) and trapezoidal rule (dotted lines).*

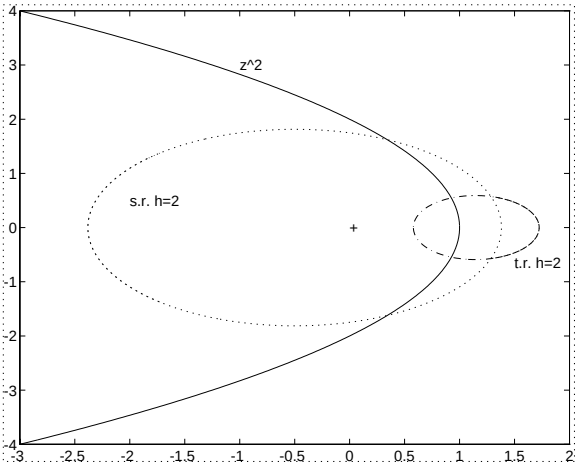

FIG. 2. *Large timestep comparison of images of $1 + iy$.*

might suppose that the Störmer and trapezoidal rule discretizations would behave similarly with respect to WR convergence. In fact, this is not the case and it turns out that the Störmer method yields a much more stable WR iteration than the trapezoidal rule over comparable time intervals. Figure 2 shows the image of $1 + iy$, $-2 < y < 2$, when $h = 2$, under $P_h^{\text{t.r.}}$, $P_h^{\text{s.r.}}$, and z^2. Figure 2 also indicates that results such as Proposition 9 of [9] and Theorem 3.4 of [11] which bound the spectral radius of the discrete iteration in terms of that of the continuous iteration for A-stable multistep methods will not typically hold in our setting.

The problem with generalizing the results of [11] is that they were based on the strengthened stability assumption that the stability region includes a disk on the negative real axis touching the origin, whereas many of the symmetric methods we consider (e.g., Störmer's rule) do not satisfy this condition.

We will use the exponential weight to correct for the weak stability of the method. For $\gamma > 0$, define the γ-stability region Ω_γ of the method as the set of all $\mu \in \bar{C}$ such that the roots ζ_i of $a(\zeta) - \mu^2 b(\zeta) = 0$ lie in the disk $|\zeta| \leq e^\gamma$ and are simple on the

boundary. The iteration operator $\mathcal{S}_h$ is bounded in $l^2_{h,\alpha}$ if $\sqrt{\lambda} \in 1/h\text{int}\Omega_{\alpha h}$ for all $\lambda \in \sigma(A_+)$.

Now observe that

$$\begin{aligned} S_h(z) &= (P_h(z)I - A_+)^{-1}A_- \\ &= S(\sqrt{P_h(z)}). \end{aligned}$$

We can directly relate the spectral radius of the discrete iteration to that of the continuous time iteration. In fact,

$$\begin{aligned} \rho_{\alpha,h}(\mathcal{S}_h) &= \sup_{\text{Re}z \geq \alpha} \rho\left(S\left(\sqrt{P_h(z)}\right)\right) \\ &= \sup_{\text{Re}z \geq \alpha} \rho\left(S\left(\frac{1}{h}\sqrt{\frac{a(e^{hz})}{b(e^{hz})}}\right)\right). \end{aligned}$$

Let the notation bdyW be used to indicate the boundary of the set W. Since $\{\sqrt{\frac{a(e^{hz})}{b(e^{hz})}}|\text{Re}z = \alpha\} = \text{bdy}\Omega_{h\alpha}$, we have, analogous to a result in [11], the following theorem.

THEOREM 4.1. *Suppose $\sigma(hA_+) \subset \text{int}\Omega_{\alpha h}$; then*

$$\rho_{\alpha,h}(\mathcal{S}_h) = \sup\{\rho(S(z))|hz \in \bar{C} \setminus \text{int}\Omega_{\alpha h}\}$$

and

$$\|\mathcal{S}_h\|_{\alpha,h} = \sup\{|S(z)||hz \in \text{bdy}\Omega_{\alpha h}\}.$$

THEOREM 4.2. *If the dynamic iteration converges in L^2_α and the symmetric multistep method is irreducible and convergent, then the discretized iteration converges in $l^2_{\alpha,h}$ for sufficiently small h and*

$$\rho_{\alpha,h}(\mathcal{S}_h) = \rho_\alpha(\mathcal{S}) + O(h).$$

We will outline a proof of this result, since the reasoning is somewhat different than that used in [11].

Let $\beta_1, \beta_2 \ldots, \beta_{k-1}$ be the k zeros of a, with $\beta_1 = 1$ being the principle root counted with multiplicity two. For simplicity, assume that these zeros all lie on the unit circle S^1 and that they are ordered counterclockwise about the unit circle; thus $\beta_j = e^{i\theta_j}$, $\theta_j \in [0, 2\pi]$, $\theta_j < \theta_{j+1}$. (It would not be difficult to treat the case where some zeros lie *inside* the unit circle.) From consistency, we must have that $\beta_1 = 1$ is a double root, while all of the other roots are simple.

We can view $|a(e^{h\alpha}w)|^2$ as a function of w on S^1. For $\alpha = 0$, it has $k-1$ minima at the zeros of a; for h sufficiently small and $\alpha > 0$, it has $k-1$ local minima located near the points β_i. We can expand a in Taylor's series about the β_i to obtain

$$\begin{aligned} a(e^{h\alpha}w) &= a(\beta_i) + a'(\beta_i)(e^{h\alpha}w - \beta_i) + O((e^{h\alpha}w - \beta_i)^2) \\ &= a'(\beta_i)(e^{h\alpha}w - \beta_i) + O(e^{h\alpha}w - \beta_i)^2. \end{aligned}$$

Only $\beta_1 = 1$ is a multiple root of a, hence $a'(\beta_i) \neq 0$ for $i = 2, \ldots, k-1$. This means that, for $e^{h\alpha}w$ in the vicinity of β_i, $i = 2, \ldots, k-1$, we must have $a(e^{h\alpha}w) = O(h)$; such a relation must also hold at the local minimum.

Using this, we can prove a small lemma which shows that the spectral radius is determined for small h by the approximation property of the principle root.

LEMMA 4.3. *For h sufficiently small,*

$$\max_{y\geq 0} \rho(S_h(\alpha+iy)) = \max_{0\leq y\leq \theta_2/h} \rho(S_h(\alpha+iy)),$$

and a similar result holds for $\|S_h\|$.

Proof. By symmetry, $\theta_{k-1} = 2\pi - \theta_2$. Denote $I_h = [\theta_2/h, \theta_{k-1}/h]$. For h sufficiently small, the global minimum of $|a(e^{h\alpha}e^{ihy})|^2$ on I_h must occur at one of its local minima over that interval or at the endpoints. Since $b(e^{hz})$ can be uniformly bounded in any bounded region, it is straightforward to see that the quantity $\mu_h(y)$ defined by

$$\mu_h(y) := \frac{1}{h}\sqrt{\frac{a(e^{h(\alpha+iy)})}{b(e^{h(\alpha+iy)})}}$$

satisfies

$$\min_{y\in I_h} |\mu_h(y)| \sim h^{-1/2},$$

and hence that

$$\max_{y\in I_h} \rho(S_h(\alpha+iy)) = O(h).$$

Due to symmetry, the behavior of ρ is the same on the intervals $[\theta_{k-1}/h, 2\pi/h]$ and $[0, \theta_2/h]$; in other words, we need only look in the latter subinterval for the maximizing value. ∎

Given $\bar{y} > 0$, consistency implies that

$$\max_{0\leq y\leq \bar{y}} |(\alpha+iy) - \mu_h(y)| = O(h).$$

Choose a large enough rectangular neighborhood N of the origin so that, e.g., $\rho(S(z)) < \rho_\alpha(\mathcal{S})/2$ for z outside N. Now $|\mu_h(\theta_2/h)| \sim h^{-1/2}$; thus for h sufficiently small, the curve $\Gamma_h := \{\mu_h(y) : 0 \leq y \leq \theta_2/h\}$ leaves N. After leaving N, it cannot reenter N (or $|\mu_h(y)^2|$ would have another local minimum). Within N, the curve Γ_h will approximate the line segment $\alpha + iy$ to $O(h)$. Thus, for h sufficiently small, the maximum value of $\rho(S_h(z))$ will occur when $\mu_h(y)$ lies within N, and since this point lies within $O(h)$ of $\alpha+iy$, we can see that asymptotically, the spectral radius of $\mathcal{S}_h$ in the weighted space can differ by only $O(h)$ from that of $\mathcal{S}$. This concludes the proof of the theorem.

5. Aliasing effects. Consider the transformed discrete iterator $S_h(z) = -(P_h(z) - A_+)^{-1}A_-$ on the vertical line $\alpha + i\nu$, $\nu \in \mathbf{R}$. The degree to which an eigenvalue $-\omega^2$ of A_+ has an impact on the solution at frequency ν depends inversely on the separation between $P_h(\alpha+i\nu)$ and $-\omega^2$. Those frequencies ν for which $P_h(\alpha+i\nu)$ lies far from the spectrum of A_+ will be only weakly propagated by the iteration.

For any multistep method, the function $P_h(\alpha+i\nu)$ is actually periodic in ν with period $2\pi/h$. This aliasing effect means that high frequencies can be excited with large stepsizes. Frequencies $\nu_0 + 2k\pi/h$ all give the same response. Actually, the situation is even somewhat worse due to the symmetry about the real axis: the response to

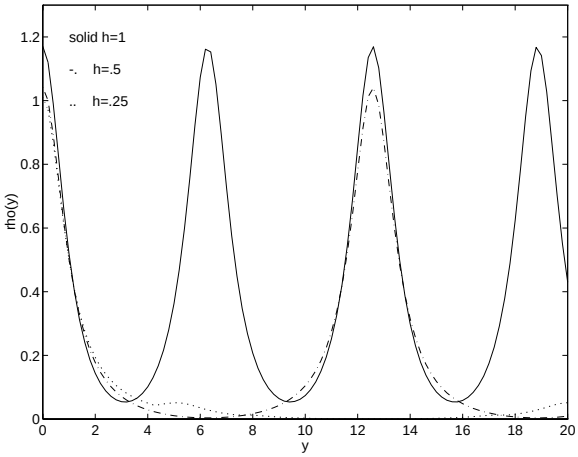

FIG. 3. *Spectral radius of* $L(1+iy)$*, wave equation,* $N = 32$.

$-\nu_0+2\pi/h$ will be the same as the response to ν_0. Of course, if there are no frequencies present in the forcing function above, say, π/h, then these anomalous excitations do no harm.

We will illustrate with the wave equation example. We first look at the response of the discrete solution operator for the (unsplit) $N = 32$ spatially discretized wave equation along the line $1 + iy$. The curves shown in Figure 3 show the spectral radii (hence also the norm) of $L(1 + iy)^{-1}$ versus y for $h = 1, .5, .25$. The maximum value is achieved near $y = 0$, as expected from Theorems 3.1 and 4.2. An increase in the stepsize h provides accuracy for small y while introducing some extraneous excitation at $2\pi k/h$, $k \in \mathbf{Z}$.

We next examine the spectral radius of the transformed discrete iteration operator $S(z)$ for the Jacobi splitting of the ($N = 32$) spatially discretized wave equation along the line $1 + iy$. The curves shown in Figure 4 show the spectral radii versus y for h in the progression $h = 1, .5, .25, .125$. As h decreases, the spectral radius has increasing maxima achieved at increasing values of y.

Using a large stepsize to resolve the small y response will not apparently improve the convergence of the small step Jacobi iteration, since the large stepsize solution operator does not even act on the high frequencies (where the spectral radius is large). The exception to this will be in the case when h is so small that the "coarse" grid is not coarse at all (in which case, little is gained through iteration). Moreover, unless $\rho(S_h(z))$ is small outside of an interval about the origin of length roughly $2\pi/H$, the artifacts introduced at the high frequencies on the coarse grid will not be damped out.

To see a substantial improvement, our iteration operator should be designed to achieve its maximum at or near $y = 0$. As mentioned previously, for the wave equation, a natural (if slow) choice is standard Picard iteration.

We now turn to the Schrödinger equation and consider the splitting (4) for $V(t) =$ constant. In case V is not too large, we would expect here that the maximum of $\rho(R(z))$ is achieved near $y = 0$ (A_+ has an eigenvalue near the origin) and that substantial improvement may be possible by exploiting a coarse timestep solution.

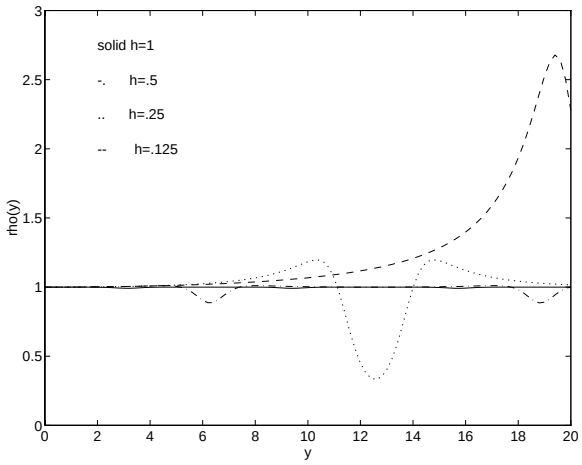

FIG. 4. *Spectral radius of* $S(1+iy)$*, wave equation,* $N = 32$.

6. Timestep acceleration of WR. The examples of the previous section suggest that an approach in which different time-meshes are used at each sweep could be successful. The goal, as for standard multigrid, is to iterate on successively coarser grids, thus resolving those components of the residual that are most difficult to obtain on the fine grid. We envision ultimately combining spatial multigrid with this timestep acceleration scheme. For the formulation and analysis of standard multigrid methods in the context of elliptic PDEs, the reader is referred to [2].

We will now define the steps of the algorithm described in the introduction. Let b and a represent the operators which define the discretization. We use h to represent the fine timestep, and H to represent the coarse timestep. Normally, in solving elliptic PDEs, we use $H = 2h$. In our case, this choice may or may not be appropriate; for the purposes of discussing an algorithm, we assume that the stepsize changes by a common factor at each iteration, but this is perhaps not essential in practice. Let I_h^H and I_H^h represent *restriction* and *prolongation* operators, respectively, which act between the fine and coarse time-meshes; thus $I_h^H : l^2_{h,\alpha} \to l^2_{H,\alpha}$ and $I_H^h : l^2_{H,\alpha} \to l^2_{h,\alpha}$.

Note that whether we wish to solve problems with or without forcing, description for a forced problem permits an easy recursive definition.

ALGORITHM TAWR(h). Given: fine timestep h, a sequence $\{f_n\} \in l^2_{h,\alpha}$, an approximation u^0 to the solution with fixed stepsize h, and a splitting $A = A_+ - A_-$. The following algorithm solves

$$(aI - h^2 bA)\{u_n\} = \{f_n\}$$

subject to k prescribed starting values θ_i, $i = 0, \ldots, k-1$.

1. Small timestep presmoothing. Starting from u^0, perform ν sweeps of WR iteration with timestep h:

$$(aI - h^2 bA_+)\{u_n^{l+1}\} = -h^2 bA_-\{u_n^l\} + \{f_n\}, \qquad l = 0, 1, \ldots, \nu - 1, \tag{8}$$

where $u_i^{l+1} = \theta_i$, $i = 0, \ldots, k-1$.

2. Large timestep correction. Compute the defect $\{d_n\}$ satisfying $d_n = 0$, $n = 0, \ldots, k-1$, and

$$d_n = (aI - h^2 bA)u_n^\nu - f_n, \qquad n \geq k,$$

from the formula

$$d_n = h^2 b A_-(u_n^\nu - u_n^{\nu-1}).$$

If the timestep $H = qh$ is sufficiently large, solve

$$(aI - H^2 bA)\{v_n\} = I_h^H\{d_n\} \tag{9}$$

directly (i.e., without relaxation) using zeros for starting values.
Else apply μ iterations of **TAWR(H)** to (9), using zeros for starting values.
Next, correct:

$$\{\bar{u}_n\} = \{u_n^\nu\} - I_H^h\{v_n\}.$$

3. Small timestep postsmoothing. Apply ν iterations of the fine mesh smoothing operation (8).

Notes:

- For $\mu = 1$ this is the V-cycle; for $\mu \geq 2$, it is called the W-cycle.
- Different numbers of smoothing steps could be used in the pre- and postsmoothings.
- To solve the problem $\mathcal{L}u = f$ using timestep acceleration, we first compute $\{f_n\} := h^2 b\{f(t_n)\}$.

7. Convergence analysis. In this section we present an elementary general convergence result regarding two-grid acceleration. This result could be easily extended to the full timestep acceleration iteration. The iteration operator in the two stepsize case can be written as $\mathcal{S}_h^\nu \mathcal{C}_{h,H} \mathcal{S}_h^\nu$, where $\mathcal{S}_h$ represents the smoothing sweep and $\mathcal{C}_{h,H}$ represents the coarse grid correction. In general, it is enough to understand $\mathcal{M} = \mathcal{C}_{h,H}\mathcal{S}_h^\nu$. The operator $\mathcal{C}_{h,H}$ can be written as

$$\mathcal{C}_{h,H} = I - p(aI - H^2 bA)^{-1} r(aI - h^2 bA),$$

where we have denoted the prolongation and restriction operators by p and r, respectively. On the other hand,

$$\mathcal{S}_h = (aI - h^2 bA_+)h^2 bA_-.$$

It is enough to show that

$$\mathcal{M}_1 := \mathcal{C}_{h,H}(aI - h^2 bA)^{-1}$$

is $O(h^{-2})$ in $l^2_{\alpha,h}$, while the norm of

$$\mathcal{M}_2 := (aI - h^2 bA)\mathcal{S}_h^\nu$$

is $O(h^2\phi(\nu))$, where ϕ tends to zero as $\nu \to 0$. In fact we anticipate that the situation is often rather better than this result would indicate, but this approach allows us to state a quite general convergence result.

Set $\mathcal{R}_h := (aI - h^2 bA)$. Note

$$\mathcal{C}_{h,H}(aI - h^2 bA)^{-1} = \mathcal{R}_h^{-1} - p\mathcal{R}_H^{-1} r.$$

Based on the relation $\mathcal{R}_h^{-1} h^2 b = \mathcal{L}_h^{-1}$, the fact that b is a bounded operator, and the theorems of the last section, we have the following lemma.

LEMMA 7.1. *Suppose consistent, stable linear multistep is used and the restriction and prolongation operators are bounded operators. Then* $\|\mathcal{M}_1\|_{\alpha,h} < Ch^{-2}\|\mathcal{R}^{-1}\|_\alpha$ *for all* h *sufficiently small.*

The proof follows since (i) $\|\mathcal{L}_h^{-1}\|_{\alpha,h} = \|\mathcal{L}^{-1}\|_\alpha + O(h)$, (ii) the same thing holds for h replaced by H and $H = qh$, and (iii) the restriction and prolongation operators are bounded. ∎

For the smoothing, we have

$$\begin{aligned} h^{-2}b^{-1}\mathcal{M}_2 &= h^{-2}b^{-1}(aI - h^2bA)(aI - h^2bA_+)^{-1}h^2bA_-\mathcal{S}_h^{\nu-1} \\ &= h^{-2}b^{-1}(aI - h^2b(A_+ - A_-))(aI - h^2bA_+)^{-1}h^2bA_-\mathcal{S}_h^{\nu-1} \\ &= A_-(I - (aI - h^2bA_+)^{-1}h^2bA_-)\mathcal{S}_h^{\nu-1} \\ &= A_-(I - \mathcal{S}_h)\mathcal{S}_h^{\nu-1}. \end{aligned}$$

We therefore have

$$\|\mathcal{M}_2\|_{\alpha,h} \le h^2\|b\|_{\alpha,h}|A_-|(1 + \|\mathcal{S}_h\|_{\alpha,h})\|\mathcal{S}_h\|_{\alpha,h}^\nu.$$

This converges to zero provided

$$\rho_{\alpha,h}(\mathcal{S}_h) < \gamma < 1.$$

Thus we can state the following theorem.

THEOREM 7.2. *If the undiscretized smoothing iteration is convergent* ($\rho_\alpha(\mathcal{S}) < 1$)*, a consistent, stable linear multistep is used, the restriction and prolongation are bounded operators between* $l^2_{h,\alpha}$ *and* $l^2_{H,\alpha}$*, and enough smoothing iterations are performed, then the timestep-accelerated WR algorithm converges.*

Because of the strong contractivity of the Picard operator on small time intervals, it would be straightforward to extend this result to the full multiple-mesh recursive acceleration scheme. On the other hand, besides proving asymptotic convergence, this simplified approach provides no practical estimates of convergence.

7.1. Treatment of model problems. A key observation is that two modes are coupled via restriction. It is possible to write a formula for the "symbol" of the iteration operator as a 2×2 matrix operating on the pair of coupled modes e^{hnz} and $e^{hn(z+i\pi)}$. As an example, taking the operators

$$r = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}$$

(full weighting restriction) and $p = 2r^*$ (piecewise linear interpolation), and assuming any symmetric multistep method (a, b), then we find that the action of M on the pair of modes is given by

$$\hat{M}(z) = \hat{C}(z)\hat{S}(z)^\nu,$$

where

$$\begin{aligned} \hat{C}(z) = I_{2N} - \frac{1}{4}\begin{bmatrix} (1+\cosh(hz))^2 & 1-\cosh^2(hz) \\ 1-\cosh^2(hz) & (1-\cosh(hz))^2 \end{bmatrix} \\ \otimes \left(\frac{1}{H^2}\frac{a(e^{Hz})}{b(e^{Hz})}I - A\right)^{-1}\left(\frac{1}{h^2}\frac{a(e^{hz})}{b(e^{hz})}I - A\right) \end{aligned}$$

and

$$\hat{S}(z) = I_2 \otimes \left(\frac{1}{h^2} \frac{a(e^{hz})}{b(e^{hz})} I - A_+ \right)^{-1} A_-.$$

($\otimes$ is the Kronecker product.)

Now for Jacobi or Picard iteration on the wave equation, for example, the matrix $\hat{M}$ is easily reduced to a diagonal matrix of 2×2 blocks, so the asymptotic convergence behavior can be determined relatively easily. For red-black Gauss–Seidel iteration on the square, we get a further pairing of the spatial modes, so $\hat{M}$ actually is reduced to 4×4 blocks. By studying the spectra of these blocks, various ODE discretizations could be compared for their effect on asymptotic rate of convergence, as could other choices of restriction/prolongation.

Note that besides the restriction and prolongation having an adjoint relationship ($r = cp^*$), if we choose the second-order, two-step discretization

$$u_{n-1} - 2u_n + u_{n+1} = \frac{h^2}{6}(f_{n-1} + 4f_n + f_{n+1}), \tag{10}$$

we find that $r\mathcal{R}_h p = \mathcal{R}_H$ also. This appears to be the only consistent, stable two-step scheme for which this property holds with the given r and p. These resemble the conditions for "variational form"; however, the operator $\mathcal{R}$ and its discretization are not self-adjoint in our setting, so we *do not have* the space decomposition

$$l^2_{h,\alpha} = N(r\mathcal{R}_h) \oplus R(p), \tag{11}$$

and the standard theoretical results cannot be directly applied.

8. Numerical experiments. We performed experiments using the two-grid iteration on linear wave equations. We found that the performance improvements were sensitive to many factors, including timestep, time window length, and splitting. Unfortunately, we cannot expect to have complete flexibility in the choice of the time interval or "window," as this may be determined from a storage or communication limitation. Similarly, the timestep is typically chosen for accuracy reasons.

Consider the standard one-dimensional wave equation (1), $N = 16$, using Jacobi iteration for the smoothing. We used the discretization (10) together with full weighting restriction and piecewise linear interpolation. We did not anticipate very good behavior since the smoothing property is relatively weak for this splitting (fast modes are not very strongly damped). Indeed, this is what we observed. For most values of the stepsize, the two-grid acceleration improved convergence, but not by very much. (In some cases performance was even slightly degraded.) In each of the figures the two-norm of the error is graphed as a function of the sweep number s and the timestep n. In Figure 5 the error in Jacobi WR is indicated for stepsize $h = .025$. Figure 6 shows the mild improvement in the error when a coarse grid correction is applied at each Jacobi WR sweep.

We next examined a modified wave equation of the form

$$\ddot{u} = (A + \mu I)u,$$

where A is the discrete Laplacian and μ is a scalar parameter. We used "Laplacian splitting" into A and μI. It is easy to see that this splitting possesses a strong "smoothing property." We first chose $\mu = 50$, which, with 16 meshpoints, means that we had a substantial perturbation of the discrete Laplacian. Initial data excited

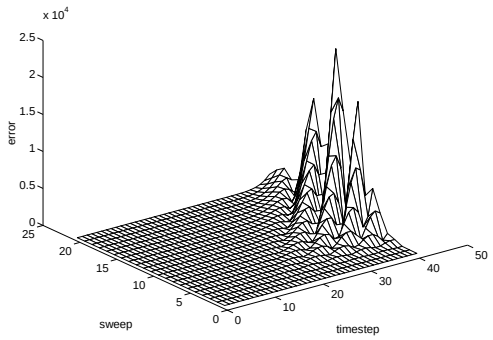

FIG. 5. *Errors in Jacobi WR, $h = .025$, 40 steps without correction.*

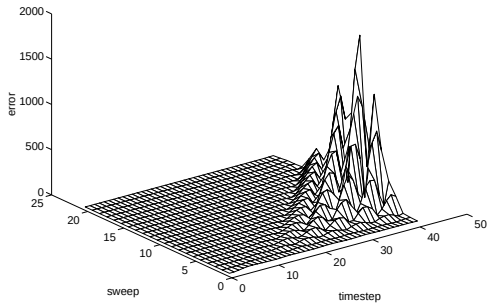

FIG. 6. *Errors with coarse grid correction, $h = .025$, 40 steps, showing poor acceleration. The benefit of coarse grid correction is diminished by the poor smoothing property of the Jacobi smoother.*

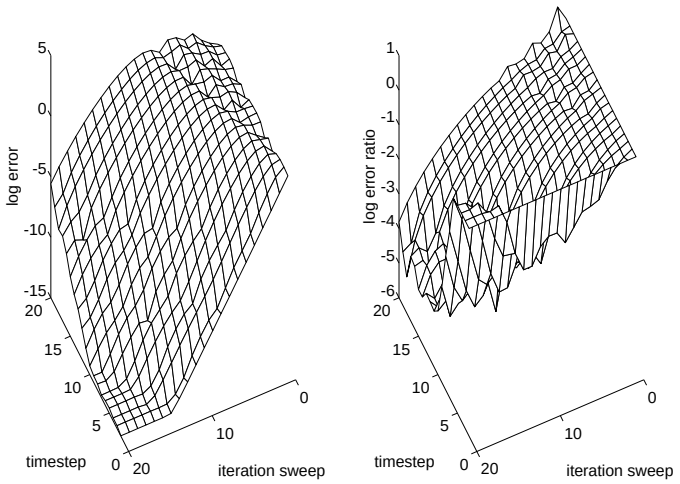

FIG. 7. (a) *log error and* (b) *log error ratio, $\mu = 50$, $h = .1$, 20 steps. The splitting provides a strong smoothing property, and a substantial improvement is possible with the two-grid iteration.*

the first two eigenfunctions of the Laplacian (slow modes), although this choice was not critical to the results we obtained. Twenty timesteps of size $h = .1$ were used. In this case, the coarse grid corrections offer substantial improvement, as shown in Figure 7. The left figure shows the $\log_{10}$ error versus timestep and sweep number; on the right we have shown the $\log_{10}$ ratio of the errors with and without the two-grid acceleration.

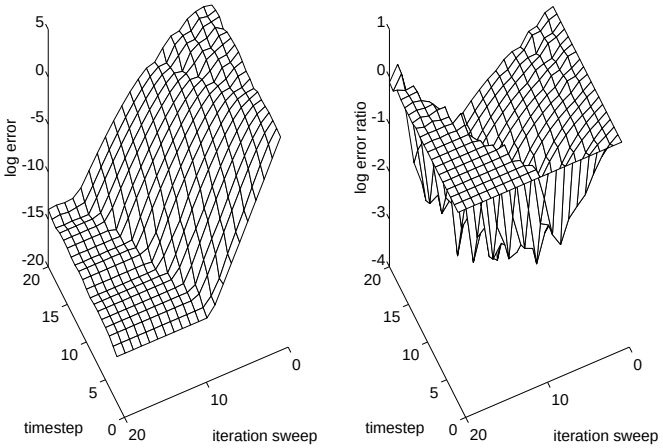

FIG. 8. (a) *log error and* (b) *log error ratio,* $\mu = 10$*,* $h = .1$*,* 20 *steps.*

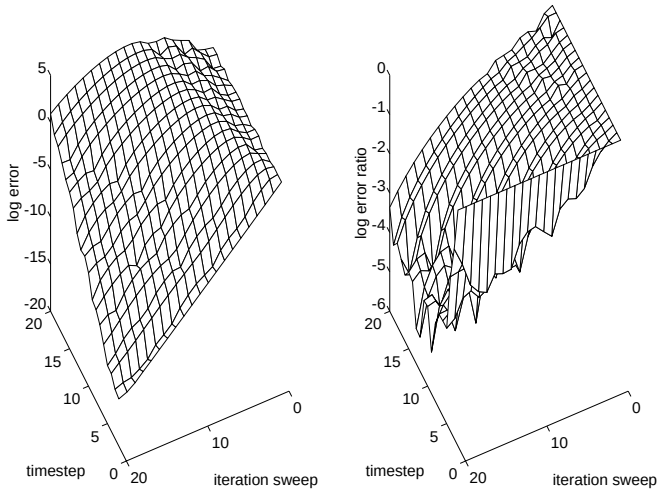

FIG. 9. (a) *log error and* (b) *log error ratio,* $\mu = 100$*,* $h = .1$*,* 20 *steps.*

The improvement evidenced here is as much as a factor of 10^5 in 20 sweeps, or a little under a factor of 2 per sweep on average.

The improvement is evident until the error reaches the level of roundoff. At the weaker perturbation $\mu = 10$, the effect somewhat diminished (Figure 8). If we instead increased the strength of the applied field ($\mu = 100$), the coarse grid correction continued to offer substantial acceleration; this is indicated in Figure 9. At larger or smaller timesteps, the improvement slightly diminished. A linear acceleration effect was observed on longer time intervals (Figure 10).

Although these experiments suggest that time-mesh coarsening accelerations hold promise for improving the parallel WR algorithm, they certainly do not settle all the issues. In particular, we do not have an easy and robust mechanism for determining what splittings will benefit from acceleration, or for determining various parameters such as number of smoothing sweeps, optimal coarsening, etc. We also have not yet experimented with the use of more than two levels of time-mesh acceleration.

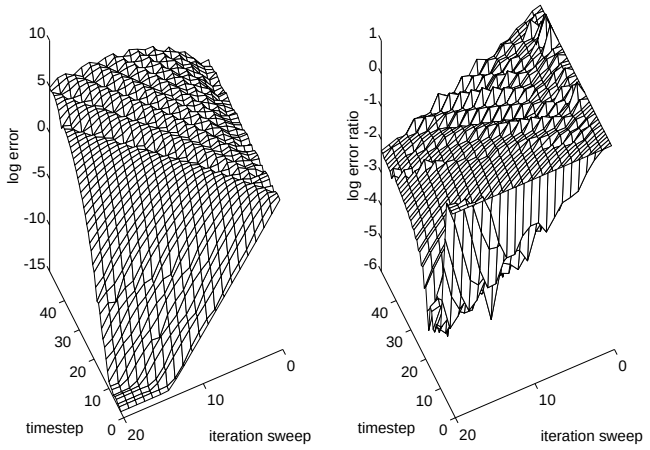

FIG. 10. (a) *log error and* (b) *log error ratio,* $\mu = 50$*,* $h = .1$*,* 50 *steps. On longer time intervals, the convergence acceleration factor (ratio of errors with and without acceleration) becomes approximately linear in the sweep.*

Acknowledgments. The author is indebted to Pawel Szeptycki for several helpful discussions during the early stages of this work. Stefan Vandewalle read a preliminary version of the manuscript and contributed several very useful comments. The computers of the Kansas Institute for Theoretical and Computational Science were used for the numerical experiments.

REFERENCES

[1] M. BJØRHUS, *On Domain Decomposition, Subdomain Iteration, and Waveform Relaxation*, Ph.D. thesis, Norwegian Institute of Technology, University of Trondheim, Norway, 1995.
[2] W. HACKBUSCH, *Multi-Grid Methods and Applications*, Springer-Verlag, New York, 1985.
[3] E. HAIRER, S. NØRSETT, AND G. WANNER, *Solving Ordinary Differential Equations*, I, Springer-Verlag, New York, 1987.
[4] G. HORTON AND S. VANDEWALLE, *A spacetime multigrid method for parabolic PDEs*, SIAM J. Sci. Comput., 16 (1995), pp. 848–864.
[5] G. HORTON, S. VANDEWALLE, AND P. WORLEY, *An algorithm with polylog parallel complexity for solving parabolic partial differential equations*, SIAM J. Sci. Comput., 16 (1995), pp. 531–541.
[6] J. LAMBERT AND D. WATSON, *Symmetric multistep methods*, J. Inst. Math. Appl., 18 (1976), pp. 189–202.
[7] B. LEIMKUHLER, *Estimating waveform relaxation convergence*, SIAM J. Sci. Comput., 14 (1993), pp. 872–889.
[8] E. LELARASMEE, A.E. RUEHLI, AND A.L. SANGIOVANNI-VINCENTELLI, *The waveform relaxation method for time-domain analysis of large scale integrated circuits*, IEEE Trans. CAD IC Sys., 3 (1982), pp. 131–145.
[9] CH. LUBICH AND A. OSTERMANN, *Multigrid dynamic iteration for parabolic equations*, BIT, 27 (1987), pp. 216–234.
[10] U. MIEKKALA AND O. NEVANLINNA, *Convergence of dynamic iteration methods for initial value problems*, SIAM J. Sci. Comput., 8 (1987), pp. 459–482.
[11] U. MIEKKALA AND O. NEVANLINNA, *Sets of convergence and stability regions*, BIT, 27 (1987), pp. 554–584.
[12] O. NEVANLINNA, *Remarks on Picard Lindelöf iteration*, BIT, 29 (1989), pp. 328–346.
[13] O. NEVANLINNA, *Power bounded prolongations and Picard-Lindelöf iteration*, Numer. Math., 58 (1990), pp. 479–501.
[14] A.L. SANGIOVANNI-VINCENTELLI AND J.K. WHITE, *Partitioning algorithms and parallel implementation of waveform relaxation algorithms for circuit simulation*, in Proc. Internat. Symp. Circ. Sys. (Kyoto, Japan), IEEE, Piscataway, NJ, 1985.