Fourier Mode Analysis of the Multigrid Waveform Relaxation and Time-Parallel Multigrid Methods

Stefan Vandewalle
Graham Horton

CRPC-TR94448
May 1994

Center for Research on Parallel Computation
Rice University
P.O. Box 1892
Houston, TX 77251-1892

FOURIER MODE ANALYSIS OF THE MULTIGRID WAVEFORM RELAXATION AND TIME-PARALLEL MULTIGRID METHODS

STEFAN VANDEWALLE* AND GRAHAM HORTON†

Abstract. The advent of parallel computers has led to the development of new solution algorithms for time-dependent partial differential equations. Two recently developed methods, the multigrid waveform relaxation method and the time-parallel multigrid method, have been designed to solve parabolic partial differential equations on many time-levels simultaneously. This paper compares the convergence properties of these methods, based on the results of an exponential Fourier mode analysis for a model problem.

AMS subject classifications. 65M06, 65M55, 65Y05

Key words. parabolic partial differential equation, multigrid, parallel computing

1. Introduction. Time-dependent partial differential equations (PDEs) are usually solved as a sequence of boundary value problems defined on successive time-levels. The sequential nature of this procedure imposes serious limitations on the obtainable performance of implementations of time-stepping methods on parallel processors or multicomputers. No matter how many time-steps are to be computed, the obtainable degree of parallelism is restricted by the parallelism in the method used to compute the solution in one single time-step. This observation has led to the development of new algorithms that compute the solution on many time-levels, possibly hundreds or thousands, simultaneously. Two such algorithms for solving parabolic partial differential equations have appeared recently in the literature: the multigrid waveform relaxation method, and the time-parallel multigrid method.

The multigrid waveform relaxation method, also called multigrid dynamic iteration method, was developed by Lubich and Ostermann in [13]. It is based on waveform relaxation, a continuous-in-time iterative method for solving large systems of ordinary differential equations. When applied to the system of ordinary differential equations derived by spatial semi-discretization of a parabolic problem, waveform relaxation results in an iterative method that is slowly convergent. Lubich and Ostermann showed that the basic waveform relaxation process can be accelerated by using the multigrid idea, in a very similar way as standard iterative solvers for elliptic problems are accelerated by using multigrid. In [13] they illustrate their theoretical results by computations with a model problem, the heat equation. Later on, the method has been applied successfully to a variety of more complex parabolic problems, like a nonlinear heat-conduction and a chemical reaction-diffusion problem ([17]), and the incompressible Navier-Stokes equations ([14]). The extension of the method to time-periodic differential equations was the subject of [19]. The multigrid waveform relaxation method has been analyzed

* California Institute of Technology, Department of Applied Mathematics, Mail Code 217-50, Pasadena, CA 91125.

† Lehrstuhl für Rechnerstrukturen (IMMD 3), Universität Erlangen – Nürnberg, Martensstr. 3, D-91058 Erlangen, Germany.

for its parallel performance, and timing results obtained on various multicomputers are reported in [18, 20]. Its use on a massively parallel machine of SIMD type is discussed in [12].

The time-parallel multigrid method was developed in a paper by Hackbusch, [6], where it is called parabolic multigrid method with parallel smoothing. An analysis of the method for the one-dimensional heat equation appeared in [4]. The method has been applied to various time-dependent problems, among which are the incompressible Navier-Stokes equations ([5, 8, 9]). Results with a parallel implementation were first reported in [1]. Later on the method was combined with extrapolation which led to a further increase in both accuracy and parallelism ([10]). The results of experiments on multicomputers with large numbers of processors appeared in [8, 9].

It can be shown that both methods, although having been developed independently, are intimately related as multigrid methods on space-time grids ([16, §3.7]). As will be explained in §2, they basically differ only in the choice of the smoother. The inter-grid transfer operators (restriction and prolongation) and the definition of the coarse grid problem are identical. In this paper we will compare the convergence properties of both algorithms, based on an exponential Fourier mode analysis for the one-dimensional heat equation. We refer to the above references for a description of their parallel implementation and for a discussion of their parallel complexity. The Fourier results are presented in §3. They will allow us to investigate the robustness of the methods with respect to the following mesh aspect ratio: $\Delta t/(\Delta x)^2$, with Δt the time-increment and Δx the spatial mesh size. The analysis will assist in understanding some observations reported in earlier papers. In particular, we shall demonstrate that the use of the time-parallel method is restricted to meshes with a large aspect ratio; we shall elucidate the dependence of the multigrid waveform convergence factor on the mesh size, and we shall explain the dependence of the convergence of both methods on the choice of the time-discretization method. We will end in §4 with some concluding remarks, and we will point out a recent research direction, which is based on the insights obtained in the current study.

2. Multigrid methods on space-time grids.

2.1. The model problem and its discretization. We shall concentrate on the problem of numerically computing the solution to the one-dimensional heat equation,

$$u_t - \Delta u = f(x,t) \qquad x \in (0,1),\ 0 < t \leq T\ , \tag{1}$$

subject to given initial and boundary values

$$u(x,0) = p(x)\ , \quad x \in (0,1)\ , \tag{2}$$

$$u(0,t) = q(t)\ ,\ u(1,t) = r(t)\ ,\ 0 < t \leq T\ . \tag{3}$$

We discretize this problem on a rectangular *space-time grid*, with spatial mesh size Δx, and constant time-increment Δt. The discretization leads to a large linear system of equations in the unknowns $u_{k,l}$, $k = 1, \ldots, n_x - 1$ ($n_x = 1/\Delta x$) and $l = 1, \ldots, n_t$ ($n_t =$

$T/\Delta t$), that approximate the PDE solution at the grid points (x_k, t_l) with $x_k = k \cdot \Delta x$ and $t_l = l \cdot \Delta t$. The grid will be denoted further by Ω_h, with h standing for the pair $(\Delta x, \Delta t)$ characterizing the size of the grid.

We shall use standard central differences to discretize the spatial operator, and we shall apply various formulae for the time-discretization: the backward Euler or first order backward differentiation method (BDF1), the second order backward differentiation method (BDF2), and the trapezoidal rule or Crank-Nicolson method (CN). With the parameter λ_h defined as $\Delta t/(\Delta x)^2$, we arrive at following formulae. For the BDF1 method:

$$(4) \qquad -\lambda_h u_{k-1,l} + (2\lambda_h + 1) u_{k,l} - \lambda_h u_{k+1,l} - u_{k,l-1} = \Delta t\, f(x_k, t_l) \ ,$$

for the BDF2 method:

$$(5) \quad -\lambda_h u_{k-1,l} + (2\lambda_h + 3/2) u_{k,l} - \lambda_h u_{k+1,l} - 2u_{k,l-1} + 1/2 u_{k,l-2} = \Delta t\, f(x_k, t_l) \ ,$$

and, finally, for the CN method:

$$(6) \qquad \begin{aligned} &-\lambda_h/2\, u_{k-1,l} + (\lambda_h + 1) u_{k,l} - \lambda_h/2\, u_{k+1,l} - \lambda_h/2\, u_{k-1,l-1} + \\ &(\lambda_h - 1) u_{k,l-1} - \lambda_h/2\, u_{k+1,l-1} = \Delta t\, (f(x_k, t_l) + f(x_k, t_{l-1}))/2 \ . \end{aligned}$$

Note that parameter λ_h can be considered as a measure of the degree of anisotropy of the discrete operator. In the case of very small λ_h, the sets of equations are almost decoupled in space. When λ_h equals zero, they correspond to sets of linear recurrences, one per spatial grid point. In the case of a very large λ_h, the sets of BDF1 and BDF2 equations are essentially decoupled in time, and correspond to sets of (almost) independent discrete boundary value problems.

2.2. The multigrid methods. The multigrid waveform relaxation method (in its discrete-time version, [13, p. 227]) and the time-parallel multigrid method can both be considered as multigrid methods operating on Ω_h, a grid extending in space and time. For a given fine grid, both methods solve the same set of equations at the same set of grid points. Both are multigrid algorithms that determine the discrete solution on the entire space-time grid, i.e. on all time-levels, simultaneously. They both use the natural discretization corresponding to (4), (5) and (6) on each level in the multigrid grid hierarchy. Note that, of course, the value of λ_h differs from one grid level to the next. The methods employ a semi-coarsening strategy, with coarsening only in the spatial dimension. The intergrid transfer operators are the standard ones used in combination with semi-coarsening. The linear prolongation (I_H^h) and full weighting (I_h^H) formulae, for example, have stencils whose non-zero values extend in the spatial dimension only,

$$(7) \qquad I_H^h \ : \ \tfrac{1}{2}\big]\ 1 \quad 2 \quad 1\ \big[\quad \text{and} \quad I_h^H \ : \ \tfrac{1}{4}\big[\ 1 \quad 2 \quad 1\ \big] \ .$$

The methods differ in the choice of the smoothing operator. The standard smoother in the multigrid waveform relaxation method is a zebra Gauss-Seidel method, more precisely, a red/black line-relaxation method with lines parallel to the time-axis. (The

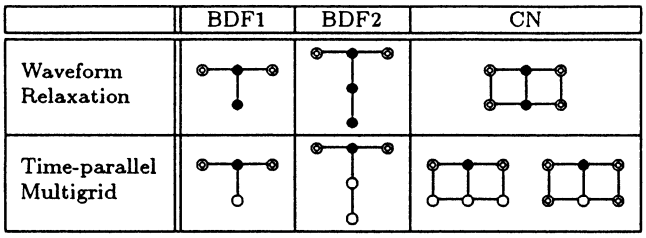

FIG. 1. *Smoothing strategy of multigrid waveform relaxation and time-parallel multigrid methods; values used when updating a 'black' grid point.*

implementation of this time-line solver is particularly simple as it only involves the forward evaluation of first or second order recurrence relations.) The time-parallel multigrid method applies a standard spatial smoother replicated on each time-level. Non-smoothed old values are used whenever values at grid points on previous time-levels are referenced. The time-parallel red/black smoother, for example, consists of one point-wise relaxation step on all red grid points at all time-levels concurrently, followed by a similar operation on all black points. Note that coloring is only w.r.t. the spatial dimension.

The implementation of the different red/black smoothers is illustrated in Fig. 1. Depicted are the discretization molecules of the BDF1, BDF2, and CN methods. The pictures show what values are used at neighboring grid points when updating the approximation at a 'black' grid point. A " ○ " indicates a non-smoothed value. The symbol " ◎ " corresponds to a value calculated in the 'red'-step. Finally, a " ● " indicates a value that is to be updated in the current 'black' phase. Note that two different strategies exist when using the Crank-Nicolson formula. They depend on whether the updated red values are communicated between time-levels only after the combined 'red/black' step (strategy A: left picture), or immediately after the 'red' step (strategy B: right picture).

Further inspection reveals an interesting relation between the smoothers of both multigrid methods. While the red/black waveform smoother solves the system of equations at a time-line exactly, the time-parallel red/black smoother solves them approximately, by doing one Jacobi relaxation step.

3. Two-grid Fourier mode analysis.

3.1. Introduction. In this section we shall analyze the two-grid variants of the multigrid methods presented in the previous section. The two-grid method makes use of one additional grid, Ω_H, derived from Ω_h by doubling the mesh size in the space dimension ($H = (2\Delta x, \Delta t)$). The convergence of the method is characterized by the so-called two-grid iteration matrix. This matrix is given by

$$M_h^H = S_h^{\nu_2} \left(I_h - I_H^h L_H^{-1} I_h^H L_h \right) S_h^{\nu_1} , \tag{8}$$

where S_h is the smoothing operator on Ω_h; ν_1 and ν_2 are the numbers of pre- and post-smoothing iterations; I_h, I_H^h, I_h^H, are the identity, prolongation, and restriction operators. L_H and L_h are discretized differential operators on Ω_H and Ω_h. It can be

shown that the entries of M_h^H depend on λ_h only, and not on the particular values of the discretization parameters Δx and Δt.

The properties of the two-grid iteration matrix are often determined in the frequency domain, by an *exponential Fourier mode analysis* ([2]). This analysis can be regarded as an analysis for special model problems, namely those with periodic boundary conditions or those on infinite domains. This analysis shows that multiplication with matrix M_h^H leaves certain linear spaces of exponential Fourier modes invariant. More precisely, it can be shown that, in the case of semi-coarsening with coarsening in a single dimension only, M_h^H is equivalent to a block-diagonal matrix, whose diagonal blocks are matrices of rank at most two. The general expression for the diagonal blocks is called the *Fourier mode symbol* of the two-grid operator. It is known to be

$$\hat{M}_h^H(\theta) = \hat{S}_h^{\nu_2}(\theta) \left(\hat{I}_h - \hat{I}_H^h(\theta)\, \hat{L}_H^{-1}(\theta)\, \hat{I}_h^H(\theta)\, \hat{L}_h(\theta) \right) \hat{S}_h^{\nu_1}(\theta) , \tag{9}$$

where $\hat{S}_h(\theta)$, $\hat{I}_h$, $\hat{I}_h^H(\theta)$, $\hat{I}_H^h(\theta)$, $\hat{L}_h(\theta)$, and $\hat{L}_H(\theta)$ denote the symbols of the smoothing operator, identity operator, restriction operator, prolongation operator, fine grid PDE operator, and coarse grid PDE operator. The formulae of the symbols will be given in §3.2.

The quality of a smoother is expressed by its *smoothing factor*, defined in reference [22, p. 149] as

$$\mu = \max\{\kappa(Q(\theta)\hat{S}_h(\theta)) : \theta \in \Theta_{\tilde{s}}\} . \tag{10}$$

Here, $\kappa(\cdot)$ stands for the matrix spectral radius. Matrix $Q(\theta)$ is a projection matrix, expressing projection onto the space of 'high' frequencies. Set $\Theta_{\tilde{s}}$ is a set of frequencies; it will be made more specific in §3.2. The convergence of the entire two-grid cycle, consisting of smoothing and coarse grid correction, is characterized by the *Fourier mode spectral radius* and *spectral norm*,

$$\rho = \max\{\kappa(\hat{M}_h^H(\theta)) : \theta \in \Theta_{\tilde{s}}\} , \quad \sigma = \max\{||\hat{M}_h^H(\theta)||_2 : \theta \in \Theta_{\tilde{s}}\} . \tag{11}$$

$||\,.\,||_2$ stands for the Euclidean matrix norm. The value of ρ usually shows very good agreement with actual convergence factors obtained on Ω_h. It corresponds to the asymptotic value of the convergence factor. The spectral norm is a non-asymptotic bound for the convergence factor.

3.2. Fourier analysis formulae. Our Fourier analysis will follow a slightly different track than the classical mode analysis explained in [2, 3, 15]. Rather than using the classical definitions (for $\Theta_{\tilde{s}}$ in particular), which aim at the limit of small mesh size, we follow the guidelines laid out in [22, Ch. 7]. The latter gives *exact* results for model problems with periodic boundaries, and, more importantly, it clearly brings out the dependence of the convergence on the values of Δx and Δt. Most Fourier analysis formulae used in our study are well-known. However, the precise formulae are not always easily located in the literature, or their formulation is sometimes different from the one required for our purposes. For the reader's convenience, we shall therefore quickly overview the formulae used to obtain the results reported in this paper.

The exponential Fourier mode $\psi_h(\theta)$ with frequency θ is given by $\psi_h(\theta) = e^{ij\cdot\theta}$, where "$\cdot$" denotes the usual $\mathbb{R}^2$ inner-product; i is the imaginary unit, and

$$j = (j_x, j_t),\ j_\alpha = 0, 1, \ldots, n_\alpha - 1 \quad (\alpha = x, t)\ , \tag{12}$$

$$\theta \in \Theta_h = \{(\theta_x, \theta_t) : \theta_\alpha = 2\pi k_\alpha / n_\alpha,\ k_\alpha = -n_\alpha/2 + 1,\ -n_\alpha/2 + 2, \ldots, n_\alpha/2\}\ . \tag{13}$$

In the case of semi-coarsening in the spatial dimension, the set of 'low' frequencies, Θ_s, corresponding to 'smooth' Fourier modes, and the set of 'high' frequencies, Θ_r, corresponding to 'rough' Fourier modes, are given by

$$\Theta_s = \Theta_h \cap ((-\pi/2, \pi/2) \times (-\pi, \pi]) \quad \text{and} \quad \Theta_r = \Theta_h \setminus \Theta_s\ . \tag{14}$$

We do not include $(\pi/2, \theta_t)$ and $(-\pi/2, \theta_t)$ in the set of low frequencies, following [22, p. 109] and the rationale given in [21, p. 107]. The closely related set $\Theta_{\bar{s}}$ is defined as $\Theta_{\bar{s}} \equiv \Theta_h \cap ([-\pi/2, \pi/2) \times (-\pi, \pi])$. To any frequency $\theta \in \Theta_{\bar{s}}$ there corresponds a unique $\bar{\theta} \in \Theta_h \setminus \Theta_{\bar{s}}$ given by $\bar{\theta} = \theta - (\text{sign}(\theta_x)\pi, 0)$. It can be verified that the space spanned by $\psi(\theta)$ and $\psi(\bar{\theta})$ is left invariant by any of the multigrid operators. The associated two by two transformation matrices are the Fourier mode symbols of these operators.

In order to derive the the symbol of the smoother, we rewrite (4), (5) and (6) as

$$\textstyle\sum_{m \in I} s_m u_{(k,l)+m} = \Delta t\ f_{k,l}\ . \tag{15}$$

Set I is a set of pairs of integers, corresponding to the non-zero coefficients in the equations. Let I be partitioned into three sets: I_0, $I_{1/2}$ and I_1. I_0 is the subset of indices corresponding to the "$\circ$"-values in Fig. 1; $I_{1/2}$ and I_1 correspond to the "$\circledcirc$"-values and "$\bullet$"-values respectively. For the time-parallel BDF1-method, for example, we find that $I_0 = \{(0, -1)\}$, $I_{1/2} = \{(-1, 0), (1, 0)\}$, and $I_1 = \{(0, 0)\}$. It is readily verified that the red/black operators map function $\psi_h(\theta)$ into $\alpha(\theta)\psi_h(\theta)$ at the 'red' grid points, and into $\beta(\theta)\psi_h(\theta)$ at the 'black' grid points, with

$$\alpha(\theta) = -\frac{\sum_{m \in I_0 \cup I_{1/2}} s_m e^{im\cdot\theta}}{\sum_{m \in I_1} s_m e^{im\cdot\theta}}\ , \quad \beta(\theta) = -\frac{\sum_{m \in I_0} s_m e^{im\cdot\theta} + \sum_{m \in I_{1/2}} \alpha(\theta) s_m e^{im\cdot\theta}}{\sum_{m \in I_1} s_m e^{im\cdot\theta}}\ . \tag{16}$$

Note that $\beta(\theta) = \alpha^2(\theta)$, in the case of the waveform relaxation method. A short calculation leads to the smoother symbol,

$$\hat{S}_h(\theta) = \tfrac{1}{2} \begin{pmatrix} \alpha(\theta) + \beta(\theta) & \alpha(\bar{\theta}) - \beta(\bar{\theta}) \\ \alpha(\theta) - \beta(\theta) & \alpha(\bar{\theta}) + \beta(\bar{\theta}) \end{pmatrix}\ . \tag{17}$$

Matrix $Q(\theta)$, in (10), is a two by two matrix of the form $\text{diag}(\delta(\theta), 1)$ with $\delta(\theta) = 1$ if $\theta = -\pi/2$, and $\delta(\theta) = 0$ otherwise.

It can be shown that any $\psi_h(\theta)$ is an eigenvector of L_h, whose symbol is given by

$$\hat{L}_h(\theta) = \text{diag}\,(\,\tilde{L}_h(\theta)\,,\,\tilde{L}_h(\bar{\theta})\,) \quad \text{with} \quad \tilde{L}_h(\theta) = \textstyle\sum_{m \in I} s_m e^{im\cdot\theta}\ . \tag{18}$$

The symbol of the coarse grid operator L_H is a scalar, similar to $\tilde{L}_h(\theta)$, but with λ_h replaced by λ_H, and with $\theta = (\theta_x, \theta_t)$ replaced by $(2\theta_x, \theta_t)$. Finally, the symbols of the full weighting restriction operator and of the linear interpolation operator are given by

$$\hat{I}_h^H(\theta) = \tfrac{1}{2}\,(\,1 + \cos(\theta_x)\ \ 1 - \cos(\theta_x)\,) \quad \text{and} \quad \hat{I}_H^h(\theta) = \left(\hat{I}_h^H(\theta)\right)^T\ . \tag{19}$$

3.3. Results. We have calculated the Fourier mode smoothing factor, and the two-grid Fourier mode spectral radius and spectral norm of the waveform relaxation method and of the time-parallel method. (Spectral radius and spectral norm are calculated for $\nu_1 = \nu_2 = 1$.) The results are graphically depicted as functions of the parameter λ_h in Fig. 2. Throughout the calculations we have used $n_x = n_t = 128$. Qualitatively, the form of the curves can be understood as explained below. We will start with the time-parallel method.

3.3.1. The time-parallel two-grid method. For large values of λ_h, the equations derived with the BDF1 method and with the BDF2 method are almost decoupled in time. They represent a set of discrete elliptic problems, one on each time-level. Hence, μ, ρ and σ reflect well-known values for the one-dimensional Poisson equation. In particular, the smoothing factor is 0.125, and both ρ and σ are zero since the multigrid method is, in this case, an exact solver. This argument, however, does not hold for the Crank-Nicolson method. Instead, it can be shown that any Fourier mode $\psi_h(\theta)$ with $\theta_t = \pi$ becomes an eigenfunction of the red/black smoother, with eigenvalue 1, when λ_h goes to infinity. Moreover, in strategy A of the smoother a maximum of 5/4 is obtained for $\kappa(\, Q(\theta)\hat{S}_h(\theta)\,)$ at $\theta = (\pi/3, 0)$, when $\lambda_h = \infty$. As a result, the algorithm with the Crank-Nicolson discretization converges very slowly or it may even diverge.

The time-parallel smoother fails completely for small values of λ_h. In the case of the BDF1 or CN discretization the equations satisfied by the error $e_{k,l} = u_{k,l} - \bar{u}_{k,l}$, with $u_{k,l}$ the exact discrete solution and $\bar{u}_{k,l}$ the current approximation, are of the form

$$e_{k,l} = e_{k,l-1} + \lambda_h \, (\cdots) \, ,$$

where the expression inside the brackets involves the errors at nearby grid points. Therefore, for very small λ_h, the smoother merely shifts any error forward in time by one time-step. This leads to a value of at least 1 for μ. It can be verified in a similar way that any $\psi_h(\theta)$ with $\theta = (\theta_x, \pi)$ is an eigenfunction of the BDF2 smoother, with eigenvalue $-5/3$, when $\lambda_h = 0$.

3.3.2. The two-grid waveform relaxation method. The waveform relaxation diagrams in Fig. 2 can be understood by relating the Fourier analysis results to convergence results for waveform relaxation methods, published in [13, 16, 19]. The Fourier analysis is an approximate analysis for initial boundary value problem (1). It becomes an exact analysis when the boundary conditions and the initial condition are replaced by a spatial periodicity and a time-periodicity condition. Hence, the Fourier results should match the convergence formulae presented for the discrete-time time-periodic multigrid waveform relaxation method in [19]. We recall that the spectral radius of the waveform relaxation operator is given by

$$\rho = \max_{z \in \Sigma} \kappa(\, M(z) \,) \, . \tag{20}$$

$M(z)$ is a matrix function of a complex variable z. It vanishes at infinity. For $z = 0$, it is the multigrid iteration matrix for the elliptic part of the parabolic equation. Set

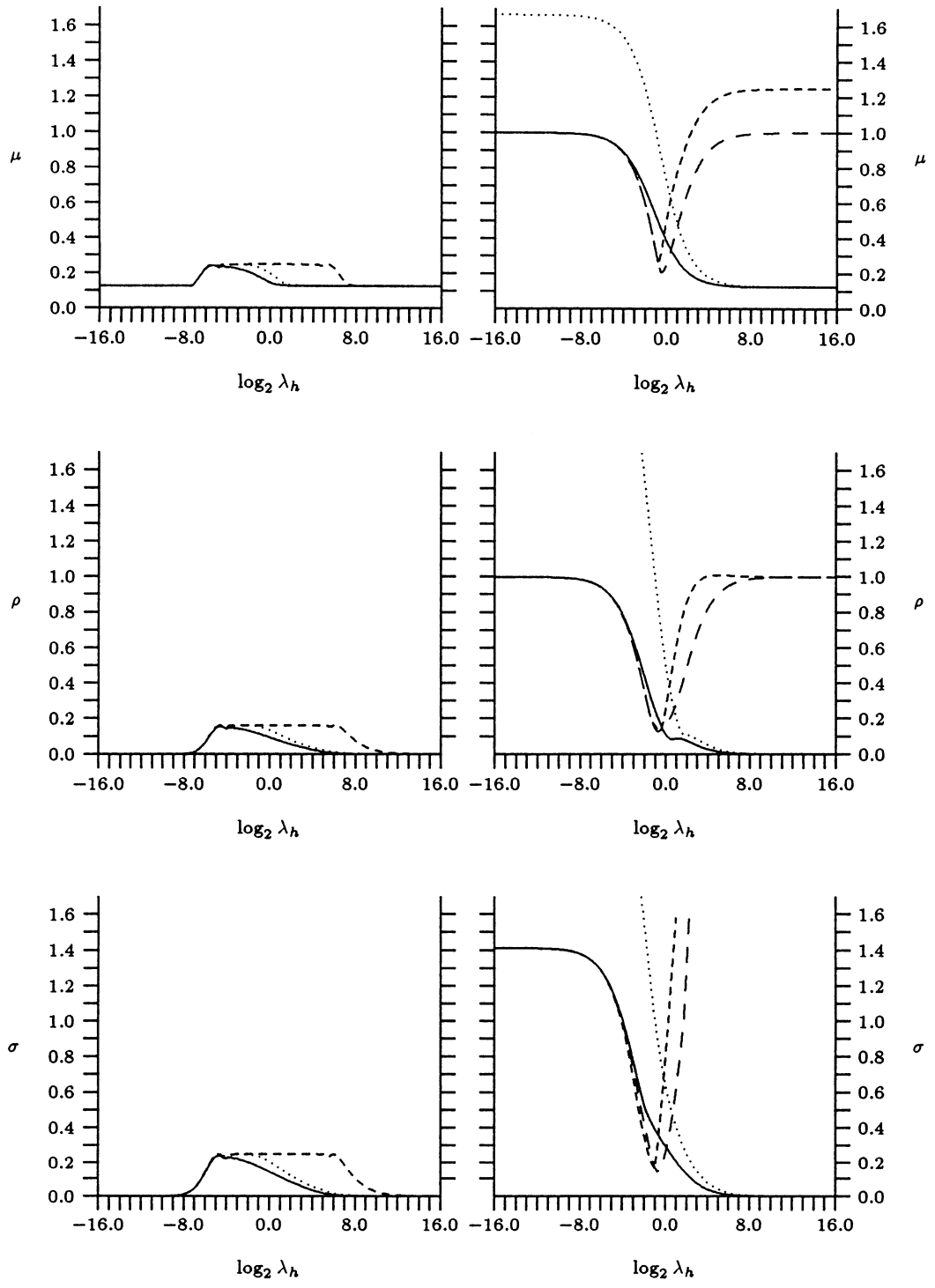

FIG. 2. *Smoothing factor (μ), two-grid spectral radius (ρ), and two-grid spectral norm (σ) of multigrid waveform relaxation method (left) and time-parallel multigrid method (right). Solid line: backward Euler method, dotted line: second order backward differentiation method, dashed line: Crank-Nicolson method (time-parallel method: strategy A: short dashed line, strategy B: long dashed line).*

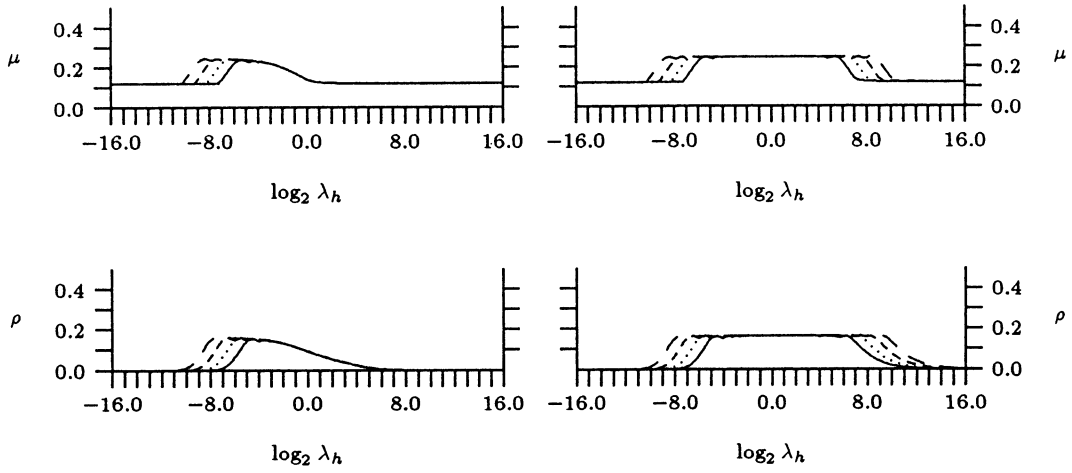

FIG. 3. *Smoothing factor (μ) and spectral radius (ρ) of the multigrid waveform relaxation method with BDF1 discretization (left) or Crank-Nicolson discretization (right). Dependence on (n_x, n_t); solid line: (128,128), dotted line: (256,256), short dashed line: (512,512), long dashed line: (1024,1024).*

Σ is a subset of the complex plane. For a linear multi-step time-discretization method defined by characteristic polynomials $a(\xi)$ and $b(\xi)$, it is given by (see [19, (4.13)])

$$\Sigma = \{\tfrac{1}{\Delta t}\tfrac{a}{b}(\xi^m) \text{ with } \xi = e^{i(2\pi/n_t)}\ ,\ \ m = 0, 1, \ldots, n_t - 1\}\ . \tag{21}$$

It is a set of points on the boundary of the method's stability region, scaled by $1/\Delta t$. Σ always includes the complex origin. (Note that expressions similar to (20) hold for the waveform relaxation smoothing factor and spectral norm.)

Consider a problem on a grid of fixed size n_x by n_t. With *increasing* Δt, the set Σ is contracted towards the origin. With *decreasing* Δt, all points of Σ move towards infinity, except for the point $\{0\}$. Consequently, both for very large λ_h and for very small λ_h, formula (20) simplifies to $\rho \simeq \kappa(M(0))$, with equality when $\lambda_h = 0$ or $\lambda_h = \infty$. Thus, at the extremal values of λ_h the spectral radius of the multigrid waveform relaxation method approaches the spectral radius of the multigrid iteration matrix. This explains the (almost) zero values found for ρ in Fig. 2 for large and small λ_h. An identical argument explains the value zero found for σ, and the value 0.125 for μ.

It is shown for μ and ρ in Fig. 3 that the curves depend on the values of n_x and n_t. (The diagrams for σ are similar to the ones for ρ.) This is due to the nature of the functions that are maximized in formulae (10) and (11). For the Crank-Nicolson method for instance, $\kappa(Q(\theta)\hat{S}_h(\theta))$ has a maximum of 1/4 at $\theta = (0, \pm 2\arctan(\lambda_h))$. The maximum is smooth as a function of θ_x, but very sharp as a function of θ_t. For small and large values of λ_h the maximum is very close to $\theta_t = 0$ or $\theta_t = \pm\pi$, and therefore 'missed' on grids with a moderate value of n_t. A similar maximum is found very close to $\theta_t = 0$ for small λ_h-values in the BDF1 and the BDF2 methods. (There is no similar maximum for large λ_h though.)

A more intuitive and qualitative understanding of the results in Fig. 3 is again possible by relating the Fourier analysis results to the waveform convergence theory

outlined above. For a fixed λ_h and increasing n_t, the set of points Σ becomes more and more dense, ultimately equaling the (continuous) scaled boundary of the stability region. As such, ρ determined by the Fourier analysis approaches a limiting value. This limit is the spectral radius of the discrete-time initial value waveform relaxation operator, [13, p. 20]. For the Crank-Nicolson method, set Σ becomes the imaginary axis, and the limit is independent of λ_h. The height of the plateau observed in the ρ-curves for the Crank-Nicolson method is given by 0.1624, which is, indeed, exactly the value found by the initial value waveform relaxation model problem analysis in [13]. For the BDF1 and BDF2 methods, the limiting continuous Σ depends on the particular value of Δt, or, equivalenty, when n_x is fixed, on the value of λ_h. This shows up clearly in Fig. 3. For very small Δt, the scaled stability region boundary of the BDF methods expands and approaches the imaginary axis. As a result, for small λ_h, the Fourier mode spectral radius of the BDF1 and the BDF2 methods will resemble the values obtained for the Crank-Nicolson method.

4. Concluding remarks. Obviously, the small difference in the algorithms – the use of an exact time-line smoother versus the use of an approximate time-line smoother – has a dramatic effect on their convergence properties. In particular, robustness with respect to λ_h is strongly affected. This robustness is important, since the value of λ_h changes considerably over the multigrid levels. More precisely, λ_h is divided by four from one grid to the next coarser grid. Within a multigrid iteration on a sufficiently large number of grid levels, very small values of λ_h may therefore be encountered.

The waveform relaxation two-grid cycle is robust for the three discretization schemes considered. The smoothing factor, spectral radius and spectral norm are bounded well below one, for every value of λ_h. Their actual values depend in a non-trivial way on the mesh-size. This was already observed in numerical experiments reported in [16, §3.5]. The time-parallel two-grid method, on the contrary, is not robust for any of the methods considered. Its convergence is satisfactory only with the BDF1 and BDF2 methods, and even then only in the case of large λ_h. This is consistent with numerical experiments reported in the literature, see e.g. [16, p.229], [9], and with a theoretical analysis for the BDF1 method in [6]. There it was already pointed out that the time-parallel BDF1 method should work well only for large λ_h.

The time-parallel method was introduced for reason of its superior parallel complexity, enabling a very efficient parallelization across time. For this, numerical robustness was sacrificed. Based on the insights obtained in the current study, different strategies could be put forward to address the problem of poor convergence of the time-parallel method for small values of λ_h. These methods may be based on the use of more efficient smoothers, different coarsening strategies, or the use of other discretization methods. The possible combinations are numerous, and most remain to be investigated. The analysis has already led the authors to develop another space-time multigrid method for parabolic problems ([11]). The algorithm uses a pointwise red/black smoother, and it coarsens the grid in both space and time. It retains the superior parallel efficiency of the time-parallel multigrid method, and partially recovers the robustness of the waveform relaxation method.

REFERENCES

[1] P. Bastian, J. Burmeister, and G. Horton. Implementation of a parallel multigrid method for parabolic partial differential equations. In Hackbusch [7], pages 18–27.

[2] A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Math. Comp.*, 31:333–390, 1977.

[3] A. Brandt. Multigrid techniques: 1984 guide, with application to fluid dynamics. GMD Studien 85, GMD-AIW, Postfach 1240, D-5205 St.-Augustin, Germany, 1984.

[4] J. Burmeister. *Paralleles lösen diskreter parabolischer Probleme mit Mehrgittertechniken.* Diplomarbeit, Universität Kiel, 1985.

[5] J. Burmeister and G. Horton. Time-parallel multigrid solution of the Navier-Stokes equations. In W. Hackbusch and U. Trottenberg, editors, *Multigrid methods III (Proceedings of the third European Multigrid Conference, Bonn, 1990)*, number 98 in ISNM, pages 155–166, Basel, 1991. Birkhaüser Verlag.

[6] W. Hackbusch. Parabolic multi-grid methods. In R. Glowinski and J.-L. Lions, editors, *Computing Methods in Applied Sciences and Engineering VI*, pages 189–197, Amsterdam, 1984. North Holland.

[7] W. Hackbusch, editor. *Parallel Algorithms for PDEs (Proceedings of the 6th GAMM Seminar Kiel, January 19-21, 1990)*, Wiesbaden, 1990. Vieweg Verlag.

[8] G. Horton. Time-parallel multigrid solution of the Navier-Stokes equations. In C. Brebbia, editor, *Applications of Supercomputers in Engineering.* Elsevier, August 1991.

[9] G. Horton. The time-parallel multigrid method. *Communic. in Appl. Num. Meth.*, 8:585–595, 1992.

[10] G. Horton and R. Knirsch. A time-parallel multigrid-extrapolation method for parabolic partial differential equations. *Parallel Computing*, 18:21–29, 1992.

[11] G. Horton and S. Vandewalle. A space-time multigrid method for parabolic P.D.E.s. Technical Report IMMD 3, 6/93, Universität Erlangen-Nürnberg, Martensstrasse 3, D-91058 Erlangen, Germany, July 1993. (to appear in SIAM J. Sci. Comput.).

[12] G. Horton, S. Vandewalle, and P. Worley. An algorithm with polylog parallel complexity for solving parabolic partial differential equations. Technical Report IMMD 3, 8/93, Universität Erlangen-Nürnberg, Martensstrasse 3, D-91058 Erlangen, Germany, July 1993. (to appear in SIAM J. Sci. Comput.).

[13] C. Lubich and A. Ostermann. Multigrid dynamic iteration for parabolic equations. *BIT*, 27:216–234, 1987.

[14] C.W. Oosterlee and P. Wesseling. Multigrid schemes for time-dependent incompressible Navier-Stokes equations. *Impact of Computing in Science and Engineering*, 5:153–175, 1993.

[15] K. Stüben and U. Trottenberg. Multigrid methods: fundamental algorithms, model problem analysis and applications. In W. Hackbusch and U. Trottenberg, editors, *Multigrid Methods*, number 960 in Lecture Notes in Mathematics, pages 1–176, Berlin, 1982. Springer Verlag.

[16] S. Vandewalle. *Parallel Multigrid Waveform Relaxation for Parabolic Problems.* B.G. Teubner Verlag, Stuttgart, 1993.

[17] S. Vandewalle and R. Piessens. Numerical experiments with nonlinear multigrid waveform relaxation on a parallel processor. *Applied Numerical Mathematics*, 8(2):149–161, 1991.

[18] S. Vandewalle and R. Piessens. Efficient parallel algorithms for solving initial-boundary value and time-periodic parabolic partial differential equations. *SIAM J. Sci. Stat. Comput.*, 13(6):1330–1346, November 1992.

[19] S. Vandewalle and R. Piessens. On dynamic iteration methods for solving time-periodic differential equations. *SIAM J. Num. Anal.*, 30(1):286–303, February 1993.

[20] S. Vandewalle and E. Van de Velde. Space-time concurrent multigrid waveform relaxation. Technical Report CRPC-93-2, Center for Research on Parallel Computation, California Institute of Technology, April 1993. (to appear in Annals of Numerical Mathematics).

[21] P. Wesseling. A survey of fourier smoothing analysis results. In Hackbusch [7], pages 105–127.

[22] P. Wesseling. *An introduction to multigrid methods.* J. Wiley, Chichester, England, 1992.