Computing 54, 317–330 (1995)

Computing

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

S. Vandewalle, Pasadena, and **G. Horton**, Erlangen

Received August 11, 1994; revised December 9, 1994

Abstract — Zusammenfassung

Fourier Mode Analysis of the Multigrid Waveform Relaxation and Time-Parallel Multigrid Methods. The advent of parallel computers has led to the development of new solution algorithms for time-dependent partial differential equations. Two recently developed methods, multigrid waveform relaxation and time-parallel multigrid, 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.

Fourier-Analyse der Mehrgitter-Wellenformrelaxationsmethode und der zeitparallelen Mehrgittermethode. Die Erscheinung von Parallelrechnern hat zur Entwicklung neuer Lösungsverfahren for zeitabhängige partielle Differentialgleichungen geführt. Zwei der in letzter Zeit entwickelten Verfahren — die Mehrgitter-Wellenformrelaxations-Methode und die zeitparallele Mehrgittermethode — haben zum Ziel, die Lösung zu vielen verschiedenen diskreten Zeitpunkten simultan zu berechnen. In dieser Arbeit wird anhand der Ergebnisse einer Fourier-Analyse für ein Modellproblem das Konvergenzverhalten beider Methoden verglichen.

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. 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 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. Lubich and Ostermann showed that the basic waveform relaxation process can

be accelerated by using the multigrid idea. 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 ([16]), and the incompressible Navier-Stokes equations ([14]). A theoretical convergence analysis of multigrid waveform relaxation for parabolic initial boundary value problems with spatial finite element discretization is given in [11, 12]. The extension of the method to time-periodic differential equations was the subject of [18]. The multigrid waveform relaxation method has been analyzed for its parallel performance, and timing results obtained on various multicomputers are reported in [17, 19]. Its use on a massively parallel machine of SIMD type is discussed in [10].

The time-parallel multigrid method was developed in a paper by Hackbusch, [5], where it is called parabolic multigrid method with parallel smoothing. An analysis of the method for the one-dimensional heat equation appeared in [3]. The method has been applied to various time-dependent problems, among which are the incompressible Navier–Stokes equations ([4, 6, 7]). 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 ([8]). The results of experiments on multicomputers with large numbers of processors appeared in [6, 7].

It can be shown that both methods, although having been developed independently, are intimately related as multigrid methods on space-time grids. As will be explained in §2, they basically differ only in the choice of the smoother. 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. 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. In §4 we report results of some numerical experiments. We end in §5, where we point out a recent research direction, 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 subject to given initial and boundary values,

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

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

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

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 $u_{k,l}$, $k = 1, \ldots, n_x - 1 (n = 1/\Delta x)$ 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 $(\Delta x, \Delta t)$.

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 $\Delta t/(\Delta x)^2$, and with $f_{k,l} = f(x_k, y_l)$, we arrive at following formulae:

$$BDF1: \; -\lambda_h u_{k-1,l} + (2\lambda_h + 1)u_{k,l} - \lambda_h u_{k+1,l} - u_{k,l-1} = \Delta t f_{k,l}; \tag{4}$$

$$BDF2: \; -\lambda_h u_{k-1,l} + (2\lambda_h + 3/2)u_{k,l} - \lambda_h u_{k+1,l} - 2u_{k,l-1} + 1/2u_{k,l-2} = \Delta t f_{k,l}; \tag{5}$$

$$CN: \; -\lambda_h/2u_{k-1,l} + (\lambda_h + 1)u_{k,l} - \lambda_h/2u_{k+1,l} - \lambda_h/2u_{k-1,l-1} + (\lambda_h - 1)u_{k,l-1} - \lambda_h/2u_{k+1,l-1} = \Delta t(f_{k,l} + f_{k,l-1})/2. \tag{6}$$

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 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. 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 weight-

ing (I_h^H) formulae, for example, have stencils whose non-zero values extend in the spatial dimension only.

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 sequential implementation of this time-line solver is particularly simple as it only involves the forward evaluation of first or second order recurrence relations. The parallel implementation involves the use of substructuring and/or cyclic reduction ([10, 19]). 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 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.

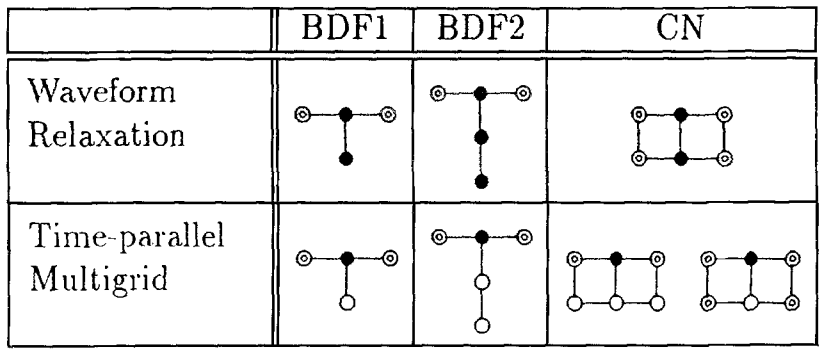

Figure 1. Smoothing strategy of multigrid waveform relaxation and time-parallel multigrid

The implementation of the 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. Two different strategies exist when using the time-parallel CN 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.

Note also that, in the limiting case of $n_t = 1$, both algorithms are identical, and

correspond to a standard multigrid method within a standard implicit time-stepping scheme.

2.3 The Two-Grid Iteration Matrix

In this section we shall consider the two-grid variants of the methods presented above. 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. The convergence of the method is characterized by the two-grid iteration matrix. This matrix, of dimension $(n_x - 1)n_t$, is given by

$$M_h^H = S_h^{\nu_2} \left(I_h - I_H^h L_H^{-1} I_h^H L_h\right) S_h^{\nu_1}, \tag{7}$$

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 Δx and Δt.

Assume that the unknown solution vector is ordered as

$$(u_{1,1}, u_{2,1}, \cdots, u_{n_x-1,1}, u_{1,2}, u_{2,2}, \cdots, u_{n_x-1,2}, \cdots, u_{1,n_t}, u_{2,n_t}, \cdots, u_{n_x-1,n_t})^T.$$

It is then found for both algorithms that M_h^H is a lower triangular block-Toeplitz matrix,

$$M_h^H = \begin{pmatrix} M_{h,1}^H & & & & \\ M_{h,2}^H & M_{h,1}^H & & & \\ M_{h,3}^H & M_{h,2}^H & M_{h,1}^H & & \\ \cdot & \cdot & \cdot & \cdot & \\ M_{h,n_t}^H & M_{h,n_{t-1}}^H & \cdot & \cdot & M_{h,1}^H \end{pmatrix} \tag{8}$$

Its blocks are of dimension $n_x - 1$. The Toeplitz nature follows from the fact that $u_{:,m}^{(\nu)}$, the ν'th iterate on time-level m, depends on $u_{:,l}^{(\nu-1)}$, $l \le m$ in the same way as $u_{:,n}^{(\nu)}$ depends on $u_{:,l}^{(\nu-1)}$, $l \le n$. The exact form of the matrices $M_{h,j}^H$ is not important for the present study, except for the matrix $M_{h,1}^H$. By considering the limiting case of $n_t = 1$, it is clear that $M_{h,1}^H$ equals the classical two-grid iteration matrix encountered in a standard implicit time-stepping scheme. As such, the block-diagonal of the waveform relaxation iteration matrix equals the block-diagonal of the time-parallel iteration matrix. Hence, the spectral radii of the two iteration matrices are identical. Both methods will converge whenever the two-grid method in the standard implicit time-stepping scheme converges.

Actual computations do show a very different convergence behaviour for both methods, see e.g. §4. Obviously, the spectral radius analysis does not explain these differences. This has to do with the strong non-normality of the iteration matrices. An alternative would be to look for norm estimates. However, tight bounds for the norms of the M_h^H-matrices have not yet been found. Here, we

shall use a classical Fourier mode analysis instead. This extremely simple tool will provide insight into the observed convergence behaviour.

3. Two-Grid Fourier Mode Analysis

3.1 Introduction

The properties of M_h^H are often determined in the frequency domain, by an *exponential Fourier mode analyis* ([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 for these problems multiplication with matrix M_h^H leaves certain linear spaces of exponential Fourier modes invariant. More precisely, it can be shown that in the present case 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 smoother, the identity, the restriction, the prolongation, and the fine and coarse grid PDE operators.

The quality of a smoother is expressed by its *smoothing factor*, see e.g. [21, p. 149],

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

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 §3.3. The convergence of the entire two-grid cycle is characterized by the *Fourier mode spectral radius* and *spectral norm*,

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

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

3.2 Fourier Analysis Formulae

Our Fourier analysis will follow a slightly different track than the classical analysis explained in [2, 15]. Rather than using the definitions (for $\Theta_{\tilde{s}}$ in particular), which aim at the limit of small mesh size, we follow the guidelines laid out in [21, 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.

The exponential Fourier mode $\psi_h(\theta)$ with frequency $\theta \in \Theta_h$ is given by its components $\psi_{h,j}(\theta) = e^{ij\cdot\theta}$, where "$\cdot$" denotes the usual $\mathbb{R}^2$ inner-product; i is the imaginary unit, and j ranges over the index set J. With α standing for x or t, sets J and Θ_h are given by

$$J = \{(j_x, j_t): j_\alpha = 0, 1, \ldots, n_\alpha - 1\}, \tag{12}$$

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

where n_x and n_y are assumed to be even. In the case of semi-coarsening in space, 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 [21, p. 109] and the rationale given in [20, p. 107]. The closely related set $\Theta_{\bar{s}}$ is defined as

$$\Theta_{\bar{s}} = \Theta_h \cap ([-\pi/2, \pi/2) \times (-\pi, \pi]). \tag{15}$$

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 $\psi_h(\theta)$ and $\psi_h(\bar{\theta})$ is left invariant by any of the multigrid operators. The associated transformation matrices are the Fourier mode symbols of these operators. The projection 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.

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

$$\sum_{m \in i} s_m u_{(k,l)+m} = \Delta t \hat{f}_{k,l}. \tag{16}$$

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 corresponding to the "○"-values in Fig. 1; $I_{1/2}$ and I_1 correspond to the "◎"-values and "●"-values respectively. For the time-parallel BDF1 method, for example, we have $I_0 = \{(0, -1)\}$, $I_{1/2} = \{(-1, 0), (1, 0)\}$, and $I_1 = \{(0, 0)\}$. It is readily verified that the red/black operators map $\psi_h(\theta)$ into $\alpha(\theta)\psi_h(\theta)$ at the 'red' points, and into $\beta(\theta)\psi_h(\theta)$ at the 'black' points, with

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

$$\beta(\theta) = -\frac{\sum_{m \in I_0} s_m e^{im\cdot\theta} + \sum_{m \in I_{1/2}} \alpha(\theta) s_m e^{im\cdot\theta}}{\sum_{m \in I_1} s_m e^{im\cdot\theta}}. \tag{17}$$

A short calculation leads to the smoother symbol,

$$\hat{S}_h(\theta) = \frac{1}{2}\begin{pmatrix} \alpha(\theta) + \beta(\theta) & \alpha(\bar{\theta}) - \beta(\bar{\theta}) \\ \alpha(\theta) - \beta(\theta) & \alpha(\bar{\theta}) + \beta(\bar{\theta}) \end{pmatrix}. \tag{18}$$

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

$$\hat{L}_h(\theta) = \operatorname{diag}\big(\tilde{L}_h(\theta), \tilde{L}_h(\bar{\theta})\big) \quad \text{with} \quad \tilde{L}_h(\theta) = \Sigma_{m \in I}\, s_m e^{im\cdot\theta}. \tag{19}$$

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

$$\hat{I}_h^H(\theta) = \tfrac{1}{2}(1 + \cos(\theta_x)\ 1 - \cos(\theta_x)) \quad \text{and} \quad \hat{I}_H^h(\theta) = \big(\hat{I}_h^H(\theta)\big)^T. \tag{20}$$

3.3 Results

We have calculated the Fourier mode smoothing factor, and the two-grid Fourier mode spectral radius and spectral norm for both multigrid methods considered in this paper. Spectral radius and spectral norm are calculated for $\nu_1 = \nu_2 = 1$. Throughout the calculations we have used $n_x = n_t = 128$. The results are graphically depicted as functions of λ_h in Fig. 2. Of course, in an actual computation with an implicit time-discretization scheme one is mainly interested in the large $\Delta t/(\Delta x)^2$ case. Meshes with small $\Delta t/(\Delta x)^2$ do arise however on the coarser levels in the multigrid algorithm.

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; both ρ and σ are zero since the multigrid method is an exact solver. This argument, however, does not hold for the CN method. Instead, it can be shown that any mode $\psi_h(\theta)$ with $\theta_t = \pi$ becomes an eigenfunction of the smoother, with eigenvalue 1, when λ_h goes to infinity. Moreover, in strategy A a maximum of 5/4 is obtained for $\kappa(Q(\theta)\hat{S}_h(\theta))$ at $\theta = (\pi/3, 0)$, when $\lambda_h = \infty$.

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 parentheses involves the errors at nearby grid points. Therefore, for very small λ_h, the smoother merely shifts any error forward in time. Fourier mode $\psi_h(\theta)$ with $\theta_t = \pi$ becomes an eigenfunction with eigenvalue -1, when $\lambda_h = 0$. It can be

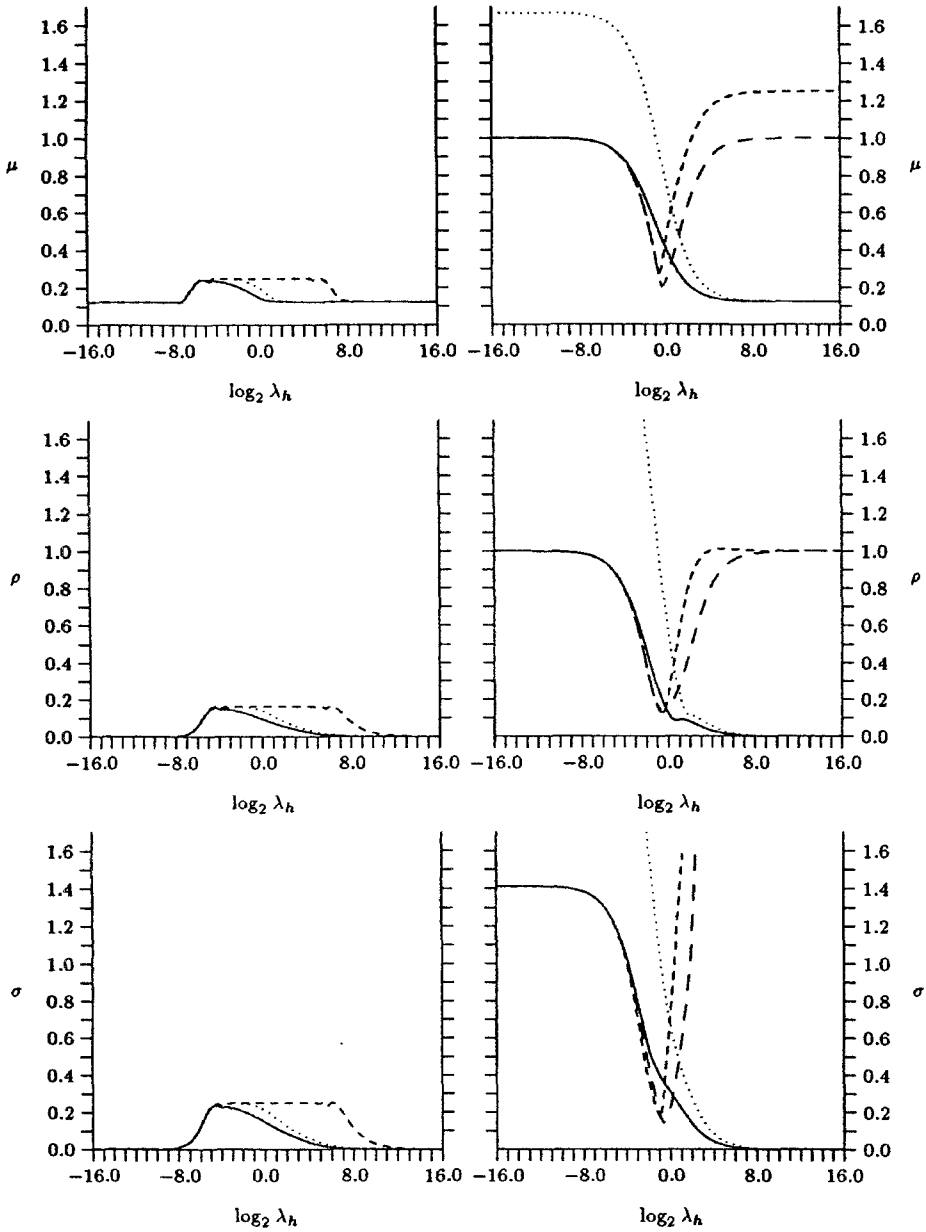

Figure 2. Smoothing factor (μ), two-grid spectral radius (ρ), and two-grid spectral norm (σ) of multigrid waveform relaxation method (left) and time-parallel multigrid method (right). Solid line: BDF1 method, dotted line: BDF2 method, dashed line: CN method (time-parallel method: strategy A: short dashed line, strategy B: long dashed line)

verified in a similar way that, for $\lambda_h = 0$, any $\psi_h(\theta)$ with $\theta_t = \pi$ is an eigenfunction of the BDF2 smoother, with eigenvalue $-5/3$.

Finally, the time-parallel results have been recomputed with $n_x = n_t = 512$. No significant difference was found with the results for the $n_x = n_t = 128$ case.

3.3.2 The Two-Grid Waveform Relaxation Method

The waveform BDF results for large λ_h are similar to the time-parallel BDF results. Since the problems are nearly decoupled in time, the typical μ-, ρ-, and σ-values for the corresponding elliptic problem are found. No problems are experienced with the CN method, and small λ_h-values do not seem to hinder convergence. All high frequency or oscillatory error components are satisfactorily damped by the red/black time-line smoother.

Contrary to the case of the time-parallel method, the waveform curves do depend strongly on the values of n_x and n_t. This is shown for μ and ρ in Fig. 3. (The diagrams for σ are similar to the ones for ρ). This is due to the nature of the functions that are maximized in (10) and (11). For the CN method, $\kappa(Q(\theta)\hat{S}_h(\theta))$ 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.

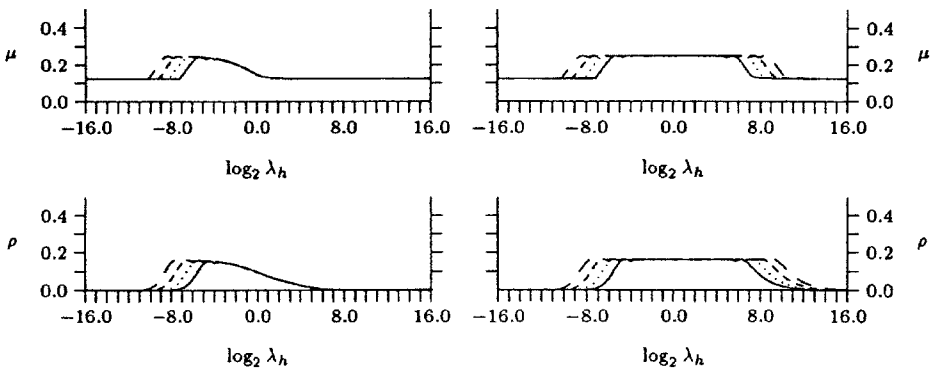

Figure 3. Smoothing factor (μ) and spectral radius (ρ) of two-grid waveform method with BDF1 (left) or CN discretization (right). Dependence on (n_x, n_t); solid line: (128,128), dotted line: (256,256), short dashed line: (512,512), long dashed line: (1024,1024)

3.3.3 Conclusion

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 depends in a non-trivial way on the meshsize. The time-parallel two-grid method, on the contrary, lacks robustness. Its convergence is satisfactory only with the BDF

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. [7], and with an analysis for the BDF1 method in [5]. There it was already pointed out that the time-parallel BDF1 method should work well only for large λ_h.

Finally, recall that the Fourier analysis is an exact analysis when the boundary and initial condition are replaced by space- and time-periodicity conditions. Hence, the Fourier results match the analytical convergence formulae presented for the discrete-time time-periodic two-grid waveform relaxation method in [18]. There, it was shown that the spectral radius of the latter method is bounded by a constant, smaller than one, independent of Δx and Δt (Cor. 4.3 and Cor. 4.5). This constant was given to be 0.162, which is exactly the height of the plateau in the ρ-curves of Fig. 3.

4. A Numerical Experiment

We discretized (1) with homogeneous Dirichlet boundary conditions and a zero initial condition. A random initial approximation to the solution was chosen so as to excite all possible error components. We shall report convergence factors based on monitoring the two-norm of the residual. The convergence factors are averaged over the first 30 iterations, or, in the case of fast convergence, over iteration one up to the iteration where the initial residual is reduced in norm by a factor 10^{-10}. This average was found to be a good measure of convergence speed. Yet, one should bear in mind that the actual iteration to iteration convergence factors may differ significantly from the averaged ones, especially in the initial iterations and in the last few iterations before convergence.

In Table 1 we present results obtained with a two-grid $(1,1)$-cycle, i.e., a two-grid cycle with one pre-smoothing step and one post-smoothing step, on a 128 by 128 grid. The Fourier analysis of the previous section does not strictly apply to the initial boundary value problem we consider here. Yet, it does appear to predict the convergence rather well. The results in Table 1 are qualitatively very similar to the ones plotted in Fig. 2. Apparently, the difference w.r.t. robustness is not restricted to problems with periodic boundary and initial conditions.

Table 1. Convergence factors for two-grid $(1,1)$-cycle ($n_x = n_t = 128$)

	Multigrid waveform relaxation						Time-parallel multigrid					
$\log_2(\lambda_h)$	–6	–3	0	3	6	9	–6	–3	0	3	6	9
BDF1	.02	.11	.08	.04	.01	.00	.92	.60	.09	.04	.00	.00
BDF2	.02	.12	.12	.05	.01	.00	2.6	1.9	.44	.08	.01	.00
CN	.02	.12	.15	.14	.09	.04	.92	.58	.15	.58	.84	.88

For parallel computing purposes, the use of a V-cycle is particularly interesting, since it often combines good parallel and numerical efficiencies. We report V(1, 1)-cycle results in Table 2. (Note that the averaged V-cycle convergence factors are sometimes smaller than the corresponding two-grid ones. This is due to the variability of the convergence factors over different iterations). The waveform V-cycles perform very well for the three time-discretization methods considered, and they are robust across λ_h. The time-parallel V-cycles perform poorly, indicating that the use of an F-cycle or W-cycle may be more appropriate.

Table 2. Convergence factors for V(1, 1)-cycle ($n_x = n_t = 128$)

	Multigrid waveform relaxation						Time-parallel multigrid					
$\log_2(\lambda_h)$	–6	–3	0	3	6	9	–6	–3	0	3	6	9
BDF1	.02	.10	.09	.09	.09	.07	.89	.72	.47	.43	.38	.14
BDF2	.02	.11	.11	.11	.10	.08	4.3	3.2	.84	.58	.52	.23
CN	.02	.11	.13	.14	.14	.12	.88	.71	.41	.63	.83	.89

Finally, in Table 3 we present V(1, 1)-cycle results for a mesh with $n_x = 128$ and $n_t = 8$. While the peformance of the V-cycle time-parallel methods is unsatisfactory on meshes with large numbers of time-levels, these methods do appear to work rather well when the number of time-levels is small. Apparently, the actual convergence factor behaves like the spectral radius of (8), i.e. $\rho(M_{h,1}^H)$, the spectral radius of the two-grid iteration operator for the elliptic problem with operator $1/\Delta t\, u - \Delta u$.

Table 3. Convergence factors for V(1, 1)-cycle ($n_x = 128$, $n_t = 8$)

	Multigrid waveform relaxation						Time-parallel multigrid					
$\log_2(\lambda_h)$	–6	–3	0	3	6	9	–6	–3	0	3	6	9
BDF1	.00	.02	.07	.09	.09	.07	.02	.05	.08	.14	.14	.11
BDF2	.00	.01	.07	.09	.09	.08	.01	.04	.09	.15	.17	.15
CN	.00	.01	.08	.10	.10	.10	.01	.03	.08	.15	.14	.10

5. Concluding Remarks

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.

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 could 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 led the authors to develop another space-time multigrid method for parabolic problems ([9]). 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.

Acknowledgement

The work of the first author was supported in part by the NSF under Cooperative Agreement No. CCR-9120008.

References

[1] Bastian, P., Burmeister, J., Horton, G.: Implementation of a parallel multigrid method for parabolic partial differential equations. In: Parallel algorithms for PDEs (Proceedings of the 6th GAMM Seminar Kiel, January 19–21, 1990) (Hackbusch, W., ed.), pp. 18–27. Wiesbaden: Vieweg 1990.

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

[3] Burmeister, J.: Paralleles Lösen diskreter parabolischer Probleme mit Mehrgittertechniken. Diplomarbeit, Universität Kiel, 1985.

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

[5] Hackbusch, W.: Parabolic multi-grid methods. In: Computing methods in applied sciences and engineering VI (Glowinski, R., Lions, J.-L. eds.), pp. 189–197. Amsterdam: North-Holland 1984.

[6] Horton, G.: Time-parallel multigrid solution of the Navier-Stokes equations. In: Applications of supercomputers in engineering (Brebbia, C., ed.). Amsterdam: Elsevier 1991.

[7] Horton, G.: The time-parallel multigrid method. Comm. Appl. Numer. Meth. *8*, 585–595 (1992).

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

[9] Horton, G., Vandewalle, S.: 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.).

[10] Horton, G., Vandewalle, S., Worley, P.: 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.).
[11] Janssen, J., Vandewalle, S.: Multigrid waveform relaxation on spatial finite element meshes: the continuous-time case. Technical Report TW 201, Katholieke Universiteit Leuven, Department of Computer Science, Celestijnenlaan 200A, B-3001 Leuven, Belgium, November 1993 (to appear in SIAM J. Num. Anal.).
[12] Janssen, J., Vandewalle, S.: Multigrid waveform relaxation on spatial finite element meshes: the discrete-time case. Technical Report CRPC-94-8, California Institute of Technology, Pasadena, CA 91125, May 1994 (submitted to SIAM J. Sci. Comput.).
[13] Lubich, C., Ostermann, A.: Multigrid dynamic iteration for parabolic equations. BIT *27*, 216–234 (1987).
[14] Oosterlee, C., Wesseling, P.: Multigrid schemes for time-dependent incompressible Navier-Stokes equations. Impact Comput. Sci. Eng. *5*, 153–175 (1993).
[15] Stüben, K., Trottenberg, U.: Multigrid methods: fundamental algorithms, model problem analysis and applications. In: Multigrid methods (Hackbusch, W., Trottenberg, U., eds.), pp. 1–176. Berlin, Heidelberg, New York: Springer 1982 (Lecture Notes in Mathematics vol. 960).
[16] Vandewalle, S., Piessens, R.: Numerical experiments with nonlinear multigrid waveform relaxation on a parallel processor. Appl. Numer. Math. *8*, 149–161 (1991).
[17] Vandewalle, S., Piessens, R.: Efficient parallel algorithms for solving initial-boundary value and time-periodic parabolic partial differential equations. SIAM J. Sci. Stat. Comput. *13*, 1330–1346 (1992).
[18] Vandewalle, S., Piessens, R.: On dynamic iteration methods for solving time-periodic differential equations. SIAM J. Numer. Anal. *30*, 286–303 (1993).
[19] Vandewalle, S., Van de Velde, E.: Space-time concurrent multigrid waveform relaxation. Ann. Numer. Math. *1*, 347–363 (1994).
[20] Wesseling, P.: A survey of Fourier smoothing analysis results. In: Parallel algorithms for PDEs (Proceedings of the 6th GAMM Seminar Kiel, January 19–21, 1990) (Hackbusch, W., ed.), pp. 105–127. Wiesbaden: Vieweg 1990.
[21] Wesseling, P.: An introduction to multigrid methods. Chichester: J. Wiley, 1992.

Dr. S. Vandewalle
Applied Mathematics 217–50
California Institute of Technology
Pasadena, CA 91125
U.S.A.
stefan@ama.caltech.edu

Dr. G. Horton
Lehrstuhl für Rechnerstrukturen (IMMD 3)
Universität Erlangen - Nürnberg
Martensstrasse 3
D-91058 Erlangen
Germany
graham@informatik.uni-erlangen.de