

**Multigrid Methods for Implicit
Runge–Kutta and Boundary Value
Method Discretizations of Parabolic
PDEs**

Jan Van lent

Stefan Vandewalle

Report TW 369, October 2003



Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Multigrid Methods for Implicit Runge–Kutta and Boundary Value Method Discretizations of Parabolic PDEs

Jan Van lent
Stefan Vandewalle

Report TW 369, October 2003

Department of Computer Science, K.U.Leuven

Abstract

Sophisticated high order time discretization methods, such as implicit Runge–Kutta and boundary value methods, are often disregarded when solving time dependent partial differential equations, despite several appealing properties. This is mainly because it is considered hard to develop efficient methods for the more complex linear systems involved. We show here that for implicit Runge–Kutta and boundary value method discretizations of time dependent parabolic problems, multigrid methods for the elliptic case can be extended in a straightforward way. The key to this approach is the use of a smoother that updates several unknowns at a spatial grid point simultaneously. Combination of the multigrid principle with both time stepping and waveform relaxation techniques are described, together with a convergence analysis. Numerical results are presented for the isotropic heat equation and a general diffusion equation with variable coefficients.

Keywords : parabolic partial differential equations, implicit Runge–Kutta methods, boundary value methods, multigrid, time stepping, waveform relaxation.

AMS(MOS) Classification : Primary : 65M55, Secondary : 65L06.

MULTIGRID METHODS FOR IMPLICIT RUNGE–KUTTA AND BOUNDARY VALUE METHOD DISCRETIZATIONS OF PARABOLIC PDES

JAN VAN LENT^{†‡} AND STEFAN VANDEWALLE[†]

Abstract. Sophisticated high order time discretization methods, such as implicit Runge–Kutta and boundary value methods, are often disregarded when solving time dependent partial differential equations, despite several appealing properties. This is mainly because it is considered hard to develop efficient methods for the more complex linear systems involved. We show here that for implicit Runge–Kutta and boundary value method discretizations of time dependent parabolic problems, multigrid methods for the elliptic case can be extended in a straightforward way. The key to this approach is the use of a smoother that updates several unknowns at a spatial grid point simultaneously. Combination of the multigrid principle with both time stepping and waveform relaxation techniques are described, together with a convergence analysis. Numerical results are presented for the isotropic heat equation and a general diffusion equation with variable coefficients.

Key words. parabolic partial differential equations, implicit Runge–Kutta methods, boundary value methods, multigrid, time stepping, waveform relaxation

AMS subject classifications. 65M55, 65L06

1. Introduction. Large, stiff systems of ordinary differential equations are often solved using low order methods such as the implicit Euler method. This choice is usually dictated by implementation and computational complexity issues. Several more sophisticated time discretization schemes such as implicit Runge–Kutta (IRK) methods [8] and more recently boundary value methods (BVM), block boundary value methods (BBVM) [3, 10, 11] and general linear methods (GLM) [5] have been developed. These methods offer high order together with favorable stability properties, error estimators and other interesting features. Many researchers, however, consider these methods impractical because of the larger, more complex linear systems involved. We illustrate this fact with a few quotes from standard textbooks on numerical methods for ordinary differential equations.

- “... difficult to justify the additional work ...”
(C.W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, 1971, p.40)
- “... such a computational cost ...”
(J.D. Lambert, *Computational Methods in Ordinary Differential Equations*, 1973, p.155)
- “... a considerable proportion ... the solution of the linear system ...”
(J.C. Butcher, *The Numerical Analysis of Ordinary Differential Equations*, 1987, p.214)
- “Compared to the expense of a multistep method, the amount of work per step must be reduced significantly before IRK methods can be competitive.”
(K.E. Brenan, S.L. Campbell, L.R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, 1989, p.77)
- “... daunting computational problem ...”
(J.D. Lambert, *Numerical Methods for Ordinary Differential Equations*, 1991, p.150)

[†]Katholieke Universiteit Leuven, Department of Computer Science, Celestijnenlaan 200A, B-3001 Leuven, Belgium ({Jan.Vanlent, Stefan.Vandewalle}@cs.kuleuven.ac.be)

[‡]Research Assistant of the Fund for Scientific Research - Flanders (Belgium) (F.W.O.-Vlaanderen)

- “Unless these factors can be improved, implicit Runge–Kutta methods are not a feasible option for stiff problems.”
(K. Burrage, *Parallel and Sequential Methods for Ordinary Differential Equations*, 1995, p.48)
- “One of the challenges for implicit Runge–Kutta methods is the development of efficient implementations. . . . usually not competitive with multistep methods . . . Thus, it is important to look for ways to make the iteration process less expensive.”
(U.M. Asher, L.R. Petzold, *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, 1998, p.103)
- “Hence, its calculation faces us with an ordeal altogether of a different kind than that for the explicit method.”
(A. Iserles, *A First Course in the Numerical Analysis of Differential Equations*, 1996, p.41)
- “An efficient solution of this system is the main problem in the implementation of an implicit Runge–Kutta method.”
(E. Hairer, G Wanner, *Solving Ordinary Differential Equations II*, 2002, p.118)

Many methods such as the Butcher transform, banded and sparse matrix solvers and Krylov methods with highly specialized, e.g., circulant, preconditioners have been suggested to handle the large linear systems in an efficient way. We will show here that multigrid methods are very efficient for the stiff systems, arising from the discretization of PDEs, the case many of the above textbooks suggest the implicit Runge–Kutta methods to be especially well suited to.

Some of the results in our paper, especially with respect to multigrid for IRK methods, have already been reported in the literature by other authors, see e.g., [18, 4]. Their results are considerably extended here and cast into the more general framework of time stepping and block time stepping with GLMs. The use of multigrid for BVM and BBVM discretizations is new. This approach enables efficient solution of the BVM and BBVM linear systems, even for very large-scale semi-discrete PDE problems.

In §2 we describe several high order time discretization schemes as well as the finite difference discretization used for the spatial domain. The multigrid iterative method is introduced in §3 and we explain how it can be applied successfully for the systems arising from time stepping and block time stepping schemes. Section 4 shows how the convergence of the iterative methods for time stepping can be analyzed. The main theoretical results of this article are presented in §5 where waveform relaxation methods for block time stepping using GLMs and BBVMs are analyzed. Numerical results for two model problems, an isotropic and an anisotropic diffusion equation, are presented in §6. A number of different IRK and BVM schemes are considered. The corresponding theoretical multigrid convergence rates are evaluated by means of a two-grid Fourier mode analysis and verified by extensive numerical experiments. One of the experiments involves the solution of an implicit linear system with over 23 million unknowns, corresponding to a BVM discretization of the 2D diffusion equation on a 512×512 spatial mesh, with 360 time-steps computed simultaneously. One multigrid iteration step, gaining approximately one extra digit of solution accuracy, takes about 5 minutes on a 1.7 GHz Intel® Pentium® 4 system with 1 GB internal memory. Hence, the results clearly illustrate the effectiveness of the multigrid algorithm for use with implicit time integration methods, even for extremely large systems.

2. Time and Space Discretization Schemes. In this article we consider parabolic equations of the form

$$\frac{\partial u}{\partial t} = \mathcal{L}u + f, \quad (2.1)$$

where \mathcal{L} is an elliptic operator. To allow theoretical analysis and to keep the notation concise we assume that \mathcal{L} is time independent. All algorithms presented here are, however, applicable to the more general case where \mathcal{L} is time dependent. We also only consider discretization in time with a constant time step and finite difference discretizations in space on equidistant, rectangular grids with Dirichlet boundary conditions. The methods presented here can, however, be extended to three-dimensional problems and finite element discretizations on regular or irregular grids with more general boundary conditions.

2.1. Time Discretization. We consider methods for stiff initial value problems of the form

$$\dot{y}(t) = f(t, y(t)), \quad (2.2)$$

where $t \in \Omega_t$, $y(t) \in V$ and $y(0) = y_0 \in V$ is given. The time domain can be bounded ($\Omega_t = [0, T]$) or unbounded ($\Omega_t = \mathbb{R}^+$). The space V can be, for example, \mathbb{R} , in the case of a scalar ODE or \mathbb{R}^m for a system of ODEs. Time dependent PDEs also fit this setting with V a space of functions on some spatial domain Ω .

Linear multistep formulae (LMFe) use solutions at several previous time steps in a formula of the form

$$\sum_{j=-k}^0 \alpha_{k+j} y_{i+j} = \Delta t \sum_{j=-k}^0 \beta_{k+j} f(t_{i+j}, y_{i+j}), \quad (2.3)$$

where $y_i \approx y(t_i) \in V$, $t_i = i\Delta t$ and $\alpha_j, \beta_j, j = 0, \dots, k$ are the parameters of the method. There is an equation for every $i = k, \dots, n$. Since only y_0 is given, the values y_1, \dots, y_{k-1} have to be obtained by other methods. The index j ranges from $-k$ to 0 to indicate the similarity with the boundary value methods formulated below and to emphasize that the set of equations (2.3) for $i = k, \dots, n$ can be described by a banded lower triangular matrix. An important class of methods for stiff initial value problems are the so called backward difference formulae (BDFe) [8, 15]. The implicit Euler method, mentioned in the introduction, is the BDF method of order 1 ($k = 1$).

Implicit Runge-Kutta (IRK) methods [8, 5, 4] use a number of intermediate quantities to advance one time step. These stage values $\tilde{y}_i \in V^s$ are calculated from

$$\tilde{y}_i = \mathbf{1}_s^T y_{i-1} + \Delta t A \tilde{f}(\tilde{t}_i, \tilde{y}_i) \quad (2.4)$$

where $\tilde{t}_i = \mathbf{1}_s t_{i-1} + c\Delta t$, $\mathbf{1}_s = [1 \dots 1]^T \in \mathbb{R}^s$ and \tilde{f} indicates componentwise application of f . The stage values \tilde{y}_i are then used to calculate $y_i \in V$ using

$$y_i = y_{i-1} + \Delta t b^T \tilde{f}(\tilde{t}_i, \tilde{y}_i). \quad (2.5)$$

The parameters of the methods are contained in the matrix $A \in \mathbb{R}^{s \times s}$ and the vectors $b, c \in \mathbb{R}^s$ where s denotes the number of stage values. Some well known methods are the Gauss, Radau and Lobatto methods [8, 5, 4].

Replacing α by β yields the structure for B_e . Using the partitions $A_e = [a_0|A]$ and $B_e = [b_0|B]$, where we split off the first columns, we can rewrite this as a system for the unknowns $y \in V^n$. We get

$$Ay = \Delta t B f(t, y) + g_0, \quad (2.12)$$

where $g_0 = -a_0 y_0 + \Delta t b_0 f(t_0, y_0)$ contains the initial condition.

We conclude this part on time discretization schemes with a method that can form an extra step in the process of transforming the continuous problem (2.2) to a discrete problem. Dividing the time interval into n subintervals leads the sequence of subproblems

$$\begin{aligned} \dot{y}_i(t) &= f(t, y_i), & y_i &: [t_{i-1}, t_i] \rightarrow V, & i=1, \dots, n, \\ y_1(t_0) &= y_0, & y_i(t_{i-1}) &= y_{i-1}(t_{i-1}), & i=2, \dots, n, \end{aligned}$$

where the final value of one subinterval is used as the initial condition on the next. This procedure is called time windowing or block time stepping.

Applying time windowing and a k -step BVM using s time steps on each of the subintervals yields a *block boundary value method* (BBVM) [3, 11]

$$Ay_i + A_0 y_{i-1} = \Delta t B f(t_i, y_i) + \Delta t B_0 f(t_{i-1}, y_{i-1}), \quad (2.13)$$

where $t_i \in \mathbb{R}^s$, $y_i \in V^s$, $A_0 = a_0 e_s^T$, $B_0 = b_0 e_s^T \in \mathbb{R}^{s \times s}$ and $e_s = [0 \cdots 0 1]^T \in \mathbb{R}^s$. This approach allows matrices A and $B \in \mathbb{R}^{s \times s}$ of much lower dimension. When $s \approx k$ these matrices can no longer be considered banded and the Toeplitz character is lost.

Remark 2.1. It is straightforward to formulate IRK methods as BBVMs and GLMs as so called BBVMs with memory. The main difference in the latter case is that the matrices A_0 and B_0 are general. Using, for example,

$$y_{\text{BBVM},i} = \begin{bmatrix} \tilde{y}_{\text{GLM},i} \\ y_{\text{GLM},i} \end{bmatrix}$$

we get

$$A = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \quad A_0 = \begin{bmatrix} 0 & -C \\ 0 & -D \end{bmatrix}, \quad B = \begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Remark 2.2. It is also possible to formulate BBVMs as GLMs. Using, for example,

$$\tilde{y}_{\text{GLM},i} = y_{\text{GLM},i} = \begin{bmatrix} y_{\text{BBVM},i} \\ y_{\text{BBVM},i-1} \end{bmatrix}$$

we get

$$A = B = \begin{bmatrix} A^{-1}B & A^{-1}B_0 \\ 0 & 0 \end{bmatrix}, \quad C = D = \begin{bmatrix} -A^{-1}A_0 & 0 \\ I & 0 \end{bmatrix}.$$

Remark 2.3. It is important to note that Δt has a different interpretation in GLMs and BBVMs. For boundary value methods it is the distance between stage values, for GLMs it is the length of one time step, which contains s stage values.

For more information on BBVMs and GLMs we refer to [11, 3, 5].

2.2. Space Discretization. We consider equations of the form

$$\mathcal{L}u = (au_x)_x + (bu_y)_y + cu = f, \quad (2.14)$$

where u , a , b , c and f are functions defined on a spatial domain. More specifically we consider only the unit square $\Omega = [0, 1]^2$ and define the equidistant rectangular grid

$$\begin{aligned} x_i &= i\Delta x, & i &= 0, \dots, n_x, & \Delta x &= n_x^{-1}, \\ y_j &= j\Delta y, & j &= 0, \dots, n_y, & \Delta y &= n_y^{-1}. \end{aligned}$$

Using central differences we can approximate equation (2.14) by

$$c_{i,j}^{0,0}u_{i,j} + c_{i,j}^{-1,0}u_{i-1,j} + c_{i,j}^{1,0}u_{i+1,j} + c_{i,j}^{0,-1}u_{i,j-1} + c_{i,j}^{0,1}u_{i,j+1} = f_{i,j},$$

where $i = 1, \dots, n_x - 1$, $j = 1, \dots, n_y - 1$, $u_{i,j} \approx u(x_i, y_j)$ and

$$\begin{aligned} c_{i,j}^{\pm 1,0} &= \Delta x^{-2}a(x_i \pm \Delta x/2, y_j), & c_{i,j}^{0,\pm 1} &= \Delta y^{-2}b(x_i, y_j \pm \Delta y/2), \\ c_{i,j}^{0,0} &= c(x_i, y_j) - c_{i,j}^{-1,0} - c_{i,j}^{1,0} - c_{i,j}^{0,-1} - c_{i,j}^{0,1}. \end{aligned}$$

Packing the unknowns at the interior points of the grid into a vector of dimension $m = (n_x - 1)(n_y - 1)$ using, for example, the following ordering

$$u_k = u_{i,j}, \quad k = (i - 1)(n_y - 1) + j,$$

the discrete system can be represented in the matrix form

$$Lu = f, \quad (2.15)$$

with $u, f \in \mathbb{R}^m$ and $L \in \mathbb{R}^{m \times m}$. The boundary conditions are incorporated in f . The matrix L is sparse and has a block tridiagonal structure with tridiagonal subblocks.

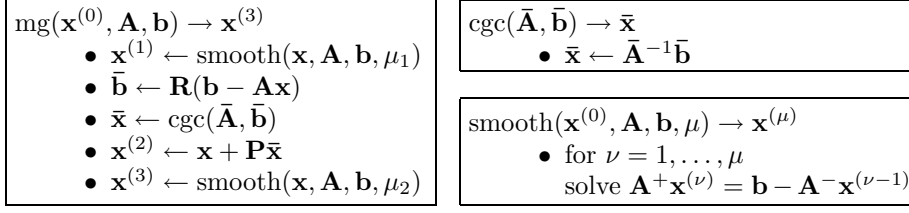
3. Time Integration: Time Stepping and Waveform Relaxation.

3.1. Introduction: Classical Iterative Methods and Multigrid. It is well known that for systems of equations involving large sparse matrices, such as the ones derived from discretizing PDEs, direct methods are often inefficient. Many iterative methods, on the other hand, deal very well with this kind of problems. Two very simple iterative schemes are the Jacobi and the Gauss–Seidel methods. Both can be described using a splitting of the system matrix. Using $L = L^+ + L^-$ and equation (2.15) the iteration scheme becomes

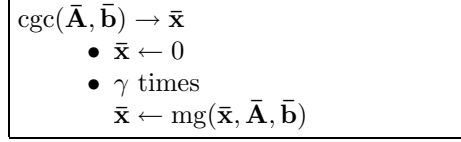
$$L^+u^{(\nu)} + L^-u^{(\nu-1)} = f.$$

The matrix L^+ is chosen such that this system can be easily solved. For the Jacobi method it contains only the diagonal of L . For the Gauss–Seidel method it is the lower triangular part of L . Unlike for the Jacobi method, the ordering of the unknowns is relevant for the Gauss–Seidel method. Common orderings on 2D grids are lexicographical and red-black. For more variations and other iterative methods of this kind we refer to [25]. A description of methods appropriate for time dependent parabolic equations with spatial anisotropy can be found in [26].

A Jacobi or Gauss–Seidel iteration step can be implemented very efficiently in the case of sparse matrices. The implementation is even simpler when uniform grids



Algorithm 3.1: Two-grid iteration for $\mathbf{A}\mathbf{x} = \mathbf{b}$, using the splitting $\mathbf{A} = \mathbf{A}^+ + \mathbf{A}^-$.



Algorithm 3.2: Coarse grid correction for multigrid iteration. V-cycle: $\gamma = 1$, W-cycle: $\gamma = 2$.

are used. Unfortunately these simple splitting methods converge very slowly. They do, however, form the basis of the very efficient multigrid methods discussed next.

Figure 3.1 gives a schematic overview of a two-grid iteration for a linear system of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$. The quantities \mathbf{x} and \mathbf{b} are elements of a vector space approximating functions on a spatial domain. The operator \mathbf{A} is a discrete approximation to an elliptic operator. The barred symbols indicate quantities and operators on a coarse grid. Restriction \mathbf{R} and prolongation \mathbf{P} are the intergrid operators, transferring quantities from the fine to the coarse level and vice versa. The whole iteration can be written as $\mathbf{x}^{(3)} = \mathbf{M}\mathbf{x}^{(0)} + \mathbf{q}$ or equivalently $\mathbf{e}^{(3)} = \mathbf{M}\mathbf{e}^{(0)}$ for the errors $\mathbf{e}^{(\cdot)} = \mathbf{x}^{(\cdot)} - \mathbf{x}$. It is straightforward to derive the following form for the two-grid iteration operator

$$\mathbf{M} = \mathbf{K}^{\mu_2}(\mathbf{I} - \mathbf{P}\bar{\mathbf{A}}^{-1}\mathbf{R}\mathbf{A})\mathbf{K}^{\mu_1}$$

$$\mathbf{K} = -\mathbf{A}^{+^{-1}}\mathbf{A}^-$$

In the two-grid iteration the equation $\bar{\mathbf{A}}\bar{\mathbf{x}} = \bar{\mathbf{b}}$ on the coarse grid is solved exactly. We can, however, solve this equation using the same iterative algorithm. Figure 3.2 shows how this can be extended to a general multigrid iteration by applying this idea recursively. Discretizations on a hierarchy of increasingly coarser grids are used. The problem on the coarsest level can be solved exactly or approximately by applying a few smoothing steps.

For clarity we consider only linear problems. Nonlinear problems can be solved by embedding the multigrid algorithm as a linear solver in a Newton iteration or by using a nonlinear multigrid scheme [25]. Here, we use the multigrid method as a solver, not as a preconditioner for a Krylov subspace method [33]. Finally, we will only consider the standard geometric multigrid procedure for discretizations on regular, rectangular grids. For irregular grids and other multigrid techniques, such as the algebraic multigrid method (AMG), we refer to [25].

The standard geometric multigrid method for discretization (2.15) of the elliptic equation (2.14) fits the framework described here with

$$\mathbf{x} = u, \quad \mathbf{b} = f,$$

$$\mathbf{A} = L, \quad \mathbf{A}^+ = L^+, \quad \mathbf{A}^- = L^-,$$

where $\mathbf{x}, \mathbf{b} \in \mathbb{R}^m$. For a given coarsest grid, the hierarchy of grids is built up by successively doubling the number of intervals in both the x - and the y -direction. $\mathbf{R} = R$ and $\mathbf{P} = P$ are the standard full weighting restriction and bilinear interpolation operators. For details we refer to [25].

By changing the order in which the discretization and solution methods are applied, a whole variety of multigrid based solution methods for time dependent problems can be derived. We now formulate the two families of methods we will focus on in the remainder of this article.

3.2. Iterative Methods for Time Stepping Schemes. If we apply any of the time discretization schemes discussed before, to a time dependent partial differential equation, we get a so called time stepping method. In the case of a parabolic initial value problem, the elliptic subproblems can be solved using a multigrid algorithm. One smoothing step consists of a loop over all spatial grid points, updating all unknowns for the considered time step simultaneously. For LMF methods there is only one unknown per grid point and only scalar computations are required. A dense, s -dimensional system has to be solved at each grid point in the IRK, GLM and BBVM cases. The BVM case cannot be considered a time stepping method since the values at all time steps have to be obtained from a system of equations that is banded, but not triangular or block triangular.

Applying to equation (2.1) first a time discretization from §2.1 and then the space discretization and the splitting of the discretized elliptic operator introduced in sections 2.2 and 3.1, we get the following iterative methods, which can all be used as the smoother in a multigrid scheme.

Using a LMF (2.3) results in

$$\begin{aligned} \sum_{j=-k}^0 \alpha_{k+j} u_{i+j}^{(\nu)} = \Delta t \sum_{j=-k}^0 \beta_{k+j} L^+ u_{i+j}^{(\nu)} + \\ \Delta t \sum_{j=-k}^0 \beta_{k+j} L^- u_{i+j}^{(\nu-1)} + \Delta t \sum_{j=-k}^0 \beta_{k+j} f_{i+j}. \end{aligned} \quad (3.1)$$

Note that additional methods have to be provided to approximate the initial values $u_i \in \mathbb{R}^m, i = 1, \dots, k-1$ to turn this into a practical method.

When we apply a BVM (2.8), we get

$$\begin{aligned} \sum_{j=-k_1}^{k_2} \alpha_{k_1+j} u_{i+j}^{(\nu)} = \Delta t \sum_{j=-k_1}^{k_2} \beta_{k_1+j} L^+ u_{i+j}^{(\nu)} + \\ \Delta t \sum_{j=-k_1}^{k_2} \beta_{k_1+j} L^- u_{i+j}^{(\nu-1)} + \Delta t \sum_{j=-k_1}^{k_2} \beta_{k_1+j} f_{i+j}. \end{aligned} \quad (3.2)$$

The initial and final conditions can be derived in the same way from (2.9) and (2.10). This type of method cannot really be called a time stepping method since the unknowns at all time steps have to be considered simultaneously.

For a GLM described by (2.6) and (2.7), the smoother becomes

$$\begin{aligned} \tilde{u}_i^{(\nu)} &= (C \otimes I_m)u_{i-1}^{(\nu)} + \Delta t(A \otimes L^+) \tilde{u}_i^{(\nu)} \\ &\quad + \Delta t(A \otimes L^-) \tilde{u}_i^{(\nu-1)} + \Delta t(A \otimes I_m) \tilde{f}_i, \end{aligned} \quad (3.3)$$

$$\begin{aligned} u_i^{(\nu)} &= (D \otimes I_m)u_{i-1}^{(\nu)} + \Delta t(B \otimes L^+) \tilde{u}_i^{(\nu)} \\ &\quad + \Delta t(B \otimes L^-) \tilde{u}_i^{(\nu-1)} + \Delta t(B \otimes I_m) \tilde{f}_i, \end{aligned} \quad (3.4)$$

where “ \otimes ” denotes the matrix Kronecker product.

Discretizing time using a BBVM (2.13) results in the iterative method

$$\begin{aligned} (A \otimes I_m)u_i^{(\nu)} + (A_0 \otimes I_m)u_{i-1}^{(\nu)} &= \\ \Delta t(B \otimes L^+)u_i^{(\nu)} + \Delta t(B \otimes L^-)u_i^{(\nu-1)} + \Delta t(B \otimes I_m)f_i &+ \\ \Delta t(B_0 \otimes L^+)u_{i-1}^{(\nu)} + \Delta t(B_0 \otimes L^-)u_{i-1}^{(\nu-1)} + \Delta t(B_0 \otimes I_m)f_{i-1}. \end{aligned} \quad (3.5)$$

The BVM discretization corresponds to taking the number of subintervals $n = 1$. When the dimension of the matrices A and B becomes large, it is again no longer appropriate to call this type of discretization time stepping.

3.3. Waveform Relaxation Methods for Block Time Stepping Schemes.

An interesting modification of the time stepping schemes is to collect the equations associated with a consecutive set of time steps together into one big system of equations [34, 24, 9]. Combining four BDF1 time steps, for example, results in a system of equations of the form

$$(A \otimes I_m)u = \Delta t(B \otimes L)u + \Delta t(B \otimes I_m)f + g_0,$$

where $u, f, g_0 \in \mathbb{R}^{4m}$, $B = I_4$ and

$$A = \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{bmatrix}.$$

Information about the previous time step is incorporated in g_0 as in (2.12). The splitting $L = L^+ + L^-$ leads to an iteration of the form

$$(A \otimes I_m)u^{(\nu)} = \Delta t(B \otimes L^+)u^{(\nu)} + \Delta t(B \otimes L^-)u^{(\nu-1)} + \Delta t(B \otimes I_m)f + g_0.$$

This can be seen as a collective update for all the unknowns on the set of time levels considered and can be executed by solving a series of small systems of equations, one 4 by 4 system at each grid point in this particular example. We call this type of method a block time stepping scheme and the associated iteration a discrete waveform relaxation method. As mentioned in the previous section, “block time stepping” is the only option for BVMs. Similar extensions are possible for any other LMFs as well as for GLMs and BBVMs. For LMFs the subsystems are lower triangular and characterized by matrices A and B of the form

$$A = \begin{bmatrix} \alpha_k & & & & \\ \vdots & \ddots & & & \\ \alpha_0 & \dots & \alpha_k & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_0 & \dots & \alpha_k \end{bmatrix}.$$

For BVMs they are banded of the form (2.11) and in the case of GLMs or BBVMs they are block lower triangular. For all these cases, the systems that arise when using a discrete waveform relaxation method, can easily be solved by direct methods.

Another way of deriving the methods in this section is by first transforming (2.1) into the system of ODEs

$$\dot{u} = Lu + f \quad (3.6)$$

by discretizing only the spatial domain. This fits the framework outlined in §3.1 with $u, f : \Omega_t \rightarrow R^m$ and the operator $\mathbf{A} = \frac{d}{dt} - L$. The splitting $\mathbf{A}^+ = \frac{d}{dt} - L^+$, $\mathbf{A}^- = -L^-$ gives us the smoother

$$\dot{u}^{(\nu)} = L^+ u^{(\nu)} + L^- u^{(\nu-1)} + f \quad (3.7)$$

and thus a multigrid algorithm. This algorithm can be defined on a time interval that is finite, $\Omega_t = [0, T]$ or infinite, $\Omega_t = [0, \infty)$. In this continuous waveform relaxation method, one smoothing step consists of the solution of m scalar ODEs, one for each grid point. Applying any of the time discretization schemes to solve these, brings us back to the discrete waveform relaxation methods discussed before.

3.4. Remarks. The equations describing methods that update all time steps simultaneously are exactly the same as those of the previous section. Only the order in which the updates of the unknowns are performed is different. In the time stepping case the ODE solver forms the outer loop and a multigrid solver is used within each time step. In the waveform relaxation case, the multigrid solver forms the outer loop and the ODE solver is used within its smoother. Summarizing we can say that a time stepping method loops over time steps, then multigrid iterations, then multigrid levels, then smoothing steps, and finally grid points. A waveform relaxation method loops over multigrid iterations, then multigrid levels, then smoothing steps, then grid points, and finally time steps.

The approach discussed here has to be contrasted with other approaches for the solution of systems arising from the discretization of general ODEs. One approach to solving systems with matrices of the form $A \otimes I_m - \Delta t B \otimes L$, obtained by using BVMs, is to use circulant approximations for the matrices A and B to precondition a Krylov subspace iteration (see for example [1, 6, 2, 12]). This essentially comes down to splitting the matrices A and B , whereas we split the matrix L . The approaches are of course not mutually exclusive, but for the subsystems that have to be solved at each grid point in our multigrid smoothers, direct methods are hard to beat. Finally, we note that smoothers derived by directly applying the Jacobi or Gauss-Seidel methods to the system matrix, give very unsatisfactory results.

4. Convergence Analysis for the Time Stepping Case. In this section we present a convergence analysis of the iterative methods for time stepping schemes discussed in the previous sections. We use the asymptotic convergence factor, i.e., the spectral radius of the iteration operator, as a measure for convergence. In the time stepping case the iterates are vectors and the iteration operator is a matrix. The convergence factors can therefore be derived using only linear algebra manipulations.

4.1. Classical Iterative Methods. For the LMF case, we obtain from (4.1) the error iteration

$$\alpha_k I_m e_i^{(\nu)} = \Delta t \beta_k L^+ e_i^{(\nu)} + \Delta t \beta_k L^- e_i^{(\nu-1)}, \quad (4.1)$$

where the iteration errors are $e_i^{(\nu)} = u_i^{(\nu)} - u_i \in \mathbb{C}^m$. Let $\mathcal{K}_{\Delta t}$ be the corresponding iteration operator. The convergence factor we are looking for is

$$\begin{aligned} \rho(\mathcal{K}_{\Delta t}) &= \rho\left((\alpha_k I_m - \Delta t \beta_k L^+)^{-1} \Delta t \beta_k L^-\right) \\ &= \rho\left(\left(\frac{1}{\Delta t} \frac{\alpha_k}{\beta_k} I_m - L^+\right)^{-1} L^-\right), \end{aligned}$$

where we assumed that $\beta_k \neq 0$ and $\frac{1}{\Delta t} \frac{\alpha_k}{\beta_k} \notin \sigma(L^+)$. Introducing the operator

$$K(z) = (zI_m - L^+)^{-1} L^- \quad (4.2)$$

we can write this as

$$\rho(\mathcal{K}_{\Delta t}) = \rho\left(K\left(\frac{1}{\Delta t} \frac{\alpha_k}{\beta_k}\right)\right).$$

We can now state the following theorem.

THEOREM 4.1. *Let $\mathcal{K}_{\Delta t}$ be the iteration operator of the iteration (3.1) used in the LMF time stepping method and let $\Sigma = \left\{\frac{\alpha_k}{\beta_k}\right\}$. If $\frac{1}{\Delta t}\Sigma \cap \sigma(L^+) = \emptyset$, then*

$$\rho(\mathcal{K}_{\Delta t}) = [\rho(K(z))]_{z \in \frac{1}{\Delta t}\Sigma}. \quad (4.3)$$

Note that for the Gauss–Seidel and Jacobi methods, the operator $K(z)$ is the iteration matrix for the equivalent splitting method applied to

$$zu = Lu + f \quad (4.4)$$

which is the Laplace transform of the system of ODEs (3.6).

For block BVMs (with standard BVMs and LMFe as special cases), we get

$$(A \otimes I_m) e_i^{(\nu)} = \Delta t (B \otimes L^+) e_i^{(\nu)} + \Delta t (B \otimes L^-) e_i^{(\nu-1)}, \quad (4.5)$$

where $e_i^{(\nu)} = u_i^{(\nu)} - u_i \in \mathbb{C}^{sm}$. Let $\mathcal{K}_{\Delta t}$ be the corresponding iteration operator. Assuming that B^{-1} exists and that $\frac{1}{\Delta t} \sigma(B^{-1}A) \cap \sigma(L^+) = \emptyset$, we get

$$\begin{aligned} \rho(\mathcal{K}_{\Delta t}) &= \rho\left((A \otimes I_m - \Delta t B \otimes L^+)^{-1} \Delta t (B \otimes L^-)\right) \\ &= \rho\left(\left(\frac{1}{\Delta t} B^{-1} A \otimes I_m - I_s \otimes L^+\right)^{-1} (I_s \otimes L^-)\right), \end{aligned}$$

which we can write as

$$\rho(\mathcal{K}_{\Delta t}) = \rho\left(K\left(\frac{1}{\Delta t} B^{-1} A\right)\right)$$

This leads us to the following theorem.

THEOREM 4.2. *Let $\mathcal{K}_{\Delta t}$ be the iteration operator of the iteration (3.5) used in the BBVM time stepping method and let $\Sigma = \sigma(B^{-1}A)$. If $\frac{1}{\Delta t}\Sigma \cap \sigma(L^+) = \emptyset$, then*

$$\rho(\mathcal{K}_{\Delta t}) = \max_{z \in \frac{1}{\Delta t}\Sigma} \rho(K(z)). \quad (4.6)$$

By choosing Σ an appropriate subset of \mathbb{C} , equation (4.6) will apply to all other time discretization methods considered here (max is replaced by sup when Σ is an infinite set).

In the IRK/GLM time stepping case, the iterative method is used to find the solutions of the equations for the stage values. The values u_i are calculated directly from \tilde{u}_i and do not take part in the iteration. The iteration for the errors $\tilde{e}_i^{(\nu)} = \tilde{u}_i^{(\nu)} - \tilde{u}_i \in \mathbb{C}^{sm}$ becomes

$$\tilde{e}_i^{(\nu)} = \Delta t(A \otimes L^+) \tilde{e}_i^{(\nu)} + \Delta t(A \otimes L^-) \tilde{e}_i^{(\nu-1)} \quad (4.7)$$

The convergence factor of the corresponding iteration operator $\mathcal{K}_{\Delta t}$ can be written as

$$\begin{aligned} \rho(\mathcal{K}_{\Delta t}) &= \rho\left((I_s \otimes I_m - \Delta t A \otimes L^+)^{-1} (\Delta t A \otimes L^-)\right) \\ &= \rho\left(K\left(\frac{1}{\Delta t} A^{-1}\right)\right) \end{aligned}$$

where we assume that A^{-1} exists and $\frac{1}{\Delta t} \sigma(A^{-1}) \cap \sigma(L^+) = \phi$.

THEOREM 4.3. *Let $\mathcal{K}_{\Delta t}$ be the iteration operator of the iteration (3.3), used for the stage values of the GLM time stepping method and let $\Sigma = \sigma(A^{-1})$. If $\frac{1}{\Delta t} \Sigma \cap \sigma(L^+) = \phi$, then*

$$\rho(\mathcal{K}_{\Delta t}) = \max_{z \in \frac{1}{\Delta t} \Sigma} \rho(K(z)). \quad (4.8)$$

4.2. Multigrid. A standard approach for the convergence analysis of geometric multigrid methods is to consider a two-grid method applied to an equation with constant coefficients and periodic boundary conditions for the spatial domain. These simplifications make it possible to do a Fourier mode analysis. This so called two-grid local Fourier mode analysis provides an indication of what convergence to expect for the multigrid iteration. For more information on convergence analysis of multigrid methods we refer to [29, 25, 33]. A three-grid analysis can be found in [32]. As an example we show how the analysis for classical iterative methods can be extended to a two-grid time stepping method using a BBVM time discretization. The results for the other time stepping schemes are analogous.

We first derive the equivalent to the operator $K(z)$ by applying the framework in §3.1 to equation (4.4). Using $\mathbf{A} = zI_m - L$, $\mathbf{A}^+ = zI_m - L^+$, $\mathbf{A}^- = -L^-$, $\mathbf{R} = R$ and $\mathbf{P} = P$ we obtain the two-grid iteration matrix

$$M(z) = K(z)^{\mu_2} (I - P(zI_m - \bar{L}))^{-1} R(zI_m - L) K(z)^{\mu_1}, \quad (4.9)$$

where $K(z)$ is defined as before.

Doing the same for $\mathbf{A} = A \otimes I_m - \Delta t B \otimes L$, $\mathbf{A}^+ = A \otimes I_m - \Delta t B \otimes L^+$, $\mathbf{A}^- = -\Delta t B \otimes L^-$, $\mathbf{R} = I_s \otimes R$ and $\mathbf{P} = I_s \otimes P$, we derive that the iteration operator for the two-grid method using (3.5) as smoother can be written as

$$\mathcal{M}_{\Delta t} = M\left(\frac{1}{\Delta t} B^{-1} A\right)$$

The results from the previous section therefore hold for the two-grid iteration if we replace $K(z)$ by $M(z)$ and extend the condition on the set Σ to

$$\frac{1}{\Delta t} \Sigma \cap (\sigma(L^+) \cup \sigma(\bar{L})) = \phi.$$

5. Convergence Analysis for the Block Time Stepping Case. Waveform relaxation schemes were first introduced and analyzed in a qualitative way in the electrical engineering literature [16, 31, 30]. Here we follow the more quantitative convergence analysis introduced in [20, 21]. This analysis was extended to multigrid waveform relaxation [18, 27], IRK time discretizations [18, 4], time periodic problems [27] and space discretizations using finite elements [13, 14]. In this section we present an analysis for GLMs and BBVMs. Notation and presentation are mainly taken from the last two references, except that for clarity we will not consider finite element discretizations here. As usual we consider methods defined on infinite time intervals since this provides results that correspond better to numerically observed behavior. For further motivation of this approach we refer to [20, 27, 13]. Weighted norms or pseudospectra can also be used to obtain meaningful information about convergence. In [19] the convergence analysis on infinite time intervals is shown to be a limiting case of the pseudospectral analysis.

5.1. Preliminaries: The Stability Domain and Its Complement. Stability domains of time discretization schemes play an important part in the convergence analysis of discrete waveform relaxation methods. We recall the definitions of the stability domains of the methods considered here. For each method we also introduce a set $\Sigma \subset \mathbb{C}$ that we will later encounter in the convergence analysis. For the case of LMF a close relation exists between Σ and the stability domain. We show that the same relation holds for GLMs (and therefore also for IRK methods), BVMs and BBVMs. As a bonus, this relation provides us with a convenient way to calculate points on a curve containing the boundary of the stability region.

The linear stability of a time discretization scheme is studied using the Dahlquist test equation [8]

$$y' = \lambda y, \quad y(0) = y_0. \quad (5.1)$$

The stability domain S consists of all $z = \Delta t \lambda \in \mathbb{C}$ for which the discrete solution is bounded. The following condition will be used to define the stability regions.

DEFINITION 5.1. *A polynomial satisfies the root condition if all of its zeros lie in the closed unit disc and all of its multiple zeros lie in the open unit disc [8, 5].*

For a LMF (2.3) we define the polynomials

$$a(w) = \sum_{j=0}^k \alpha_j w^j \quad \text{and} \quad b(w) = \sum_{j=0}^k \beta_j w^j, \quad (5.2)$$

and make the common assumptions of irreducibility, consistency and zero stability [7, 5]. Furthermore, we only consider implicit methods ($\beta_k \neq 0$). We can use the same definitions for BVMs [3]. Again, the methods are assumed to be irreducible, consistent and zero stable. BVMs are implicit by nature.

By applying each of the schemes in section 2.1 to the test equation (5.1), we obtain the following definitions.

DEFINITION 5.2. *The stability domain S of a LMF or BVM consists of all $z \in \mathbb{C}$ for which the polynomial $a(w) - zb(w)$ satisfies the root condition.*

DEFINITION 5.3. *The stability domain S of a GLM consists of all $z \in \mathbb{C}$ for which the characteristic polynomial of the matrix $D + B(z^{-1}I_s - A)^{-1}C$ satisfies the root condition [8, 5].*

DEFINITION 5.4. *The stability domain S of a BBVM consists of all $z \in \mathbb{C}$ for which the characteristic polynomial of the matrix $-(A - zB)^{-1}(A_0 - zB_0)$ satisfies the root condition [11, 3].*

In the convergence analysis of LMFe and BVMs we encounter the set

$$\Sigma = \{z \in \mathbb{C} : z = \frac{a}{b}(w), |w| \geq 1\}. \quad (5.3)$$

It is clear that the set Σ is the complement in \mathbb{C} of the interior of S , i.e.,

$$\Sigma = \mathbb{C} \setminus \text{int } S. \quad (5.4)$$

We now introduce the sets Σ that we encounter in the convergence analysis of GLMs and BBVMs and show that (5.4) still holds.

LEMMA 5.5. *Let S be the stability domain of a GLM as defined by Definition 5.3 and let the set Σ be defined as*

$$\Sigma = \{z \in \mathbb{C} : z^{-1} \in \sigma(A + C(wI_r - D)^{-1}B), |w| \geq 1\}, \quad (5.5)$$

then relation (5.4) is satisfied.

Proof. We first prove the following lemma.

LEMMA 5.6. *For all matrices $A \in \mathbb{C}^{s \times s}$, $B \in \mathbb{C}^{r \times s}$, $C \in \mathbb{C}^{s \times r}$, $D \in \mathbb{C}^{r \times r}$ and scalars $p, q \in \mathbb{C}$ the following statements are equivalent*

1. $q \in \sigma(A + C(pI_r - D)^{-1}B)$
2. $p \in \sigma(D + B(qI_s - A)^{-1}C)$

assuming that the inverses exist ($p \notin \sigma(D)$, $q \notin \sigma(A)$).

Proof. Because of symmetry, it is enough to prove that the first statement implies the second. Using the definition of an eigenvalue, the first statement results in the following sequence of implications.

$$\begin{aligned} \exists x \neq 0 : C(pI_r - D)^{-1}Bx + Ax = qx & \quad \wedge \quad Bx \neq 0 \\ \exists x \neq 0 : C(pI_r - D)^{-1}Bx = (qI_s - A)x & \quad \wedge \quad Bx \neq 0 \\ \exists x \neq 0 : (qI_s - A)^{-1}C(pI_r - D)^{-1}Bx = x & \quad \wedge \quad Bx \neq 0 \\ \exists x \neq 0 : B(qI_s - A)^{-1}C(pI_r - D)^{-1}Bx = Bx & \quad \wedge \quad Bx \neq 0 \\ \exists y \neq 0 : B(qI_s - A)^{-1}Cy = (pI_r - D)y & \quad \wedge \quad Cy \neq 0 \\ \exists y \neq 0 : B(qI_s - A)^{-1}Cy + Dy = py & \quad \wedge \quad Cy \neq 0 \end{aligned}$$

By the definition of an eigenvalue, this yields the second statement. \square

Using Lemma 5.6 we see that $z \in \Sigma$ if and only if there exists a w such that

$$w \in \sigma(D + B(z^{-1}I_s - A)^{-1}C) \quad \text{and} \quad |w| \geq 1$$

or equivalently $z \notin \text{int } S$. \square

Remark 5.7. Lemma 5.6 can also be proven by applying the identity

$$\begin{aligned} \det \left(\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \right) &= \det(M_{11}) \det(M_{22} - M_{21}M_{11}^{-1}M_{12}) \\ &= \det(M_{11} - M_{12}M_{22}^{-1}M_{21}) \det(M_{22}) \end{aligned}$$

to the matrix

$$\begin{bmatrix} qI_s - A & -C \\ -B & pI_r - D \end{bmatrix}.$$

LEMMA 5.8. *Let S be the stability domain of a BBVM as defined by Definition 5.4 and let the set Σ be defined as*

$$\Sigma = \{z \in \mathbb{C} : z \in \sigma((B + w^{-1}B_0)^{-1}(A + w^{-1}A_0)), |w| \geq 1\}, \quad (5.6)$$

then relation (5.4) is satisfied.

Proof. We first prove the following lemma.

LEMMA 5.9. *For all matrices $A, B, A_0, B_0 \in \mathbb{C}^{s \times s}$ and scalars $p, q \in \mathbb{C}$ the following statements are equivalent*

1. $q \in \sigma((B_0 + pB)^{-1}(A_0 + pA))$
2. $p \in \sigma(-(A - qB)^{-1}(A_0 - qB_0))$

assuming that the inverses exist.

Proof. Because of symmetry, it is enough to prove that the first statement implies the second. Using the definition of an eigenvalue the first statement results in the following sequence of implications

$$\begin{aligned} \exists x \neq 0 : (B_0 + pB)^{-1}(A_0 + pA)x &= qx \\ \exists x \neq 0 : (A_0 + pA)x &= q(B_0 + pB)x \\ \exists x \neq 0 : (A_0 - qB_0)x &= -p(A - qb)x \\ \exists x \neq 0 : -(A - qb)^{-1}(A_0 - qB_0)x &= px \end{aligned}$$

By the definition of an eigenvalue, this yields the second statement. \square

Using Lemma 5.9 we see that $z \in \Sigma$ if and only if there exists a w such that

$$w \in \sigma(-(A - zB)^{-1}(A_0 - zB_0)) \quad \text{and} \quad |w| \geq 1$$

or equivalently $z \notin \text{int } S$. \square

Remark 5.10. Taking $w = e^{\epsilon + i\theta}$ with $\epsilon > 0$ in the definition of Σ provides a convenient way to calculate points outside of S . If Σ is used in a formula of the form (4.6) and if all poles of $K(z)$ lie in the interior of S , then the maximum (supremum) can be taken over the curve defined by $\epsilon = 0$ instead of over the whole set Σ .

5.2. Convergence Analysis of Discrete Waveform Relaxation. We first state without proof two lemma's regarding the convergence of convolution operators. In what follows $\mathcal{H}_{\Delta t}$ is a discrete linear convolution operator with a matrix-valued kernel $h_{\Delta t}$, i.e.,

$$(\mathcal{H}_{\Delta t}e)_i = \sum_{j=0}^i h_{i-j}e_j,$$

where $e_i \in \mathbb{C}^d$ and $h_i \in \mathbb{C}^{d \times d}$.

A first lemma relates the convergence factor of a discrete convolution operator acting on finite sequences to the convergence factor of the first term of the discrete Laplace transform of its kernel (see [21, 14]).

LEMMA 5.11. *Consider $\mathcal{H}_{\Delta t}$ as an operator in $l_p(n+1; \mathbb{C}^d)$, with $1 \leq p \leq \infty$ and n finite. Then, $\mathcal{H}_{\Delta t}$ is bounded and*

$$\rho(\mathcal{H}_{\Delta t}) = \rho(h_0) = \rho(H_{\Delta t}(\infty)), \quad (5.7)$$

where $H_{\Delta t}(w) = \sum_{i=0}^n h_i w^{-i}$ denotes the discrete Laplace transform of $h_{\Delta t}$. For discrete convolution operators acting on infinite sequences we get the following lemma (see [21, 14]).

LEMMA 5.12. *Consider $H_{\Delta t}$ as an operator in $l_p(\infty; \mathbb{C}^d)$ with $1 \leq p \leq \infty$, and suppose $h_{\Delta t} \in l_1(\infty; \mathbb{C}^{d \times d})$. Then, $\mathcal{H}_{\Delta t}$ is bounded and*

$$\rho(\mathcal{H}_{\Delta t}) = \sup_{|w| \geq 1} \rho(H_{\Delta t}(w)) = \sup_{|w|=1} \rho(H_{\Delta t}(w)), \quad (5.8)$$

where $H_{\Delta t}(w) = \sum_{i=0}^{\infty} h_i w^{-i}$ denotes the discrete Laplace transform of $h_{\Delta t}$.

We will now use Lemma 5.12 to derive convergence factors for the discrete wave-form relaxation method on infinite time intervals using the LMF, BVM, GLM and BBVM time discretization schemes. By taking $|w| \rightarrow \infty$, we can then see that the results for finite time intervals are the same as for time stepping.

To apply Lemma 5.12 we will have to prove the boundedness of the convolution kernels involved. To this end we will use a matrix version of Wiener's inversion theorem.

THEOREM 5.13 (Wiener's inversion theorem). *Given a matrix valued sequence $A_{\Delta t}$ such that $A_{\Delta t} \in l_1(\infty; \mathbb{C}^{d \times d})$, and assume that*

$$\det \sum_{i=0}^{\infty} A_i w^{-i} \neq 0$$

for $|w| \geq 1$. Setting $\sum_{i=0}^{\infty} B_i w^{-i} = (\sum_{i=0}^{\infty} A_i w^{-i})^{-1}$, we have $B_{\Delta t} \in l_1(\infty; \mathbb{C}^{d \times d})$. For a proof we refer to [17] or [21].

5.2.1. LMF and BVM. From equation (3.2) we find the error iteration

$$\sum_{j=-k_1}^{k_2} \alpha_{k_1+j} e_{i+j}^{(\nu)} = \Delta t \sum_{j=-k_1}^{k_2} \beta_{k_1+j} L^+ e_{i+j}^{(\nu)} + \Delta t \sum_{j=-k_1}^{k_2} \beta_{k_1+j} L^- e_{i+j}^{(\nu-1)}.$$

By discrete Laplace transforming these equations and assuming that $e_0 = \dots = e_{k_1} = 0$ we get

$$(a(w)I_m - \Delta t b(w)L^+) e^{(\nu)}(w) = \Delta t b(w)L^- e^{(\nu-1)}(w),$$

where $e^{(\nu)}(w) = \sum_{i=0}^{\infty} e_i^{(\nu)} w^{-i}$ and a and b are the stability polynomials as defined by (5.2). The discrete Laplace transform of the operator $\mathcal{K}_{\Delta t}$, described by (3.1) is therefore

$$\begin{aligned} K_{\Delta t}(w) &= (a(w)I_m - \Delta t b(w)L^+)^{-1} \Delta t b(w)L^- \\ &= K \left(\frac{1}{\Delta t} \frac{a}{b}(w) \right). \end{aligned} \quad (5.9)$$

The following lemma is proven in [21, 14]. We omit the proof since it is very similar to the ones given below for GLMs and BBVMs. The stability domain S is the one defined by Definition 5.2.

LEMMA 5.14. *If $\sigma(\Delta t L^+) \subset \text{int } S$, then $\mathcal{K}_{\Delta t}$ is bounded in $l_p(\infty; \mathbb{C}^m)$.*

Combining lemmas 5.12 and 5.14 with equations (5.9) and (5.3) results in the following theorem.

THEOREM 5.15. *Consider $\mathcal{K}_{\Delta t}$ as an operator in $l_p(\infty; \mathbb{C}^m)$ with $1 \leq p$, and assume $\sigma(\Delta t L^+) \subset \text{int } S$. Then,*

$$\rho(\mathcal{K}_{\Delta t}) = \sup_{z \in \frac{1}{\Delta t} \Sigma} \rho(K(z)) = \sup_{z \in \frac{1}{\Delta t} \partial \Sigma} \rho(K(z)), \quad (5.10)$$

For LMFe we can take $|w| \rightarrow \infty$ and retrieve the results for time stepping and discrete waveform relaxation on finite intervals. This is not possible for general BVMs, but taking $|w| \rightarrow \infty$ in the definition of Σ for BBVMs does provide the expected results.

5.2.2. IRK and GLM. From equations (3.3) and (3.4) we derive that the error iteration for the discrete waveform relaxation method using a GLM can be described by

$$\tilde{e}_i^{(\nu)} = (C \otimes I_m) e_{i-1}^{(\nu)} + \Delta t (A \otimes L^+) \tilde{e}_i^{(\nu)} + \Delta t (A \otimes L^-) \tilde{e}_i^{(\nu-1)}, \quad (5.11)$$

$$e_i^{(\nu)} = (D \otimes I_m) e_{i-1}^{(\nu)} + \Delta t (B \otimes L^+) \tilde{e}_i^{(\nu)} + \Delta t (B \otimes L^-) \tilde{e}_i^{(\nu-1)}. \quad (5.12)$$

Applying the discrete Laplace transform to these equations results in

$$P(w) \hat{e}^{(\nu)}(w) = Q(w) \hat{e}^{(\nu-1)}(w) \quad (5.13)$$

with

$$\begin{aligned} \hat{e}^{(\nu)}(w) &= \begin{bmatrix} \tilde{e}^{(\nu)}(w) & e^{(\nu)}(w) \end{bmatrix}^T, \\ P(w) &= \begin{bmatrix} I_s \otimes I_m - \Delta t A \otimes L^+ & -w^{-1} C \otimes I_m \\ -\Delta t B \otimes L^+ & I_r \otimes I_m - w^{-1} D \otimes I_m \end{bmatrix}, \\ Q(w) &= \begin{bmatrix} \Delta t A \otimes L^- & 0 \\ \Delta t B \otimes L^- & 0 \end{bmatrix}. \end{aligned}$$

By eliminating $e^{(\nu)}(w)$ from (5.13) we get the following iteration for $\tilde{e}^{(\nu)}(w)$.

$$\begin{aligned} & [I_s \otimes I_m - (\Delta t A + w^{-1} C (I_r - w^{-1} D)^{-1} \Delta t B) \otimes L^+] \tilde{e}^{(\nu)}(w) \\ &= [(\Delta t A + w^{-1} C (I_r - w^{-1} D)^{-1} \Delta t B) \otimes L^-] \tilde{e}^{(\nu-1)}(w) \end{aligned}$$

If we denote by $\mathcal{K}_{\Delta t}$ the operator mapping $\tilde{e}^{(\nu-1)}$ to $\tilde{e}^{(\nu)}$, then its discrete Laplace transform is

$$K_{\Delta t}(w) = K \left(\frac{1}{\Delta t} (A + w^{-1} C (I_r - w^{-1} D)^{-1} B)^{-1} \right) \quad (5.14)$$

The following lemma provides a sufficient condition for $\mathcal{K}_{\Delta t}$ to be bounded. The stability domain S is the one defined by Definition 5.3.

LEMMA 5.16. *If $\sigma(\Delta t L^+) \subset \text{int } S$, then $\mathcal{K}_{\Delta t}$ is bounded in $l_p(\infty; \mathbb{C}^{sm})$.*

Proof. It is sufficient to prove that $P(w)^{-1} Q(w)$ is the discrete Laplace transform of an l_1 sequence. It is clear that both $P(w)$ and $Q(w)$ are discrete Laplace transforms of l_1 -sequences. By Wiener's Theorem we have that $P(w)^{-1}$ is the Laplace transform of an l_1 -sequence if

$$\det(P(w)) \neq 0 \quad \text{for } |w| \geq 1. \quad (5.15)$$

If this is true then $P(w)^{-1} Q(w)$ is the discrete Laplace transform of the convolution of two l_1 -sequences and therefore the operator corresponding to (5.11) and (5.12) is bounded. Condition (5.15) can only hold if the condition

$$\det \left(\begin{bmatrix} I_s - zA & -w^{-1}C \\ -zB & I_r - w^{-1}D \end{bmatrix} \right) \neq 0 \quad (5.16)$$

holds for all $z \in \sigma(\Delta t L^+)$ and all $|w| \geq 1$. By considering the LU factorization of the matrix we see that condition (5.16) is equivalent to the conditions

$$\det(I_s - zA) \neq 0, \quad (5.17)$$

$$\det(I_r - w^{-1}D - zB(I_s - zA)^{-1}w^{-1}C) \neq 0 \quad (5.18)$$

For z to be in $\text{int } S$, the matrix $D + B(z^{-1}I_s - A)^{-1}C$ has to be well defined and therefore condition (5.17) is satisfied. Suppose equality holds for some $z \in \sigma(\Delta t L^+)$ in (5.18). This would mean that

$$w \in \sigma(D + zB(I_s - zA)^{-1}C)$$

with $|w| \geq 1$ and therefore $z \notin \text{int } S$. This contradicts the assumption of the lemma. Hence, (5.18) is satisfied. \square

Combining Lemmas 5.12 and 5.16 with equations (5.14) and (5.5) results in the following theorem.

THEOREM 5.17. *Consider $\mathcal{K}_{\Delta t}$ as an operator in $l_p(\infty; \mathbb{C}^{sm})$ with $1 \leq p \leq \infty$, and assume $\sigma(\Delta t L^+) \subset \text{int } S$. Then,*

$$\rho(\mathcal{K}_{\Delta t}) = \sup_{z \in \frac{1}{\Delta t} \Sigma} \rho(K(z)) = \sup_{z \in \frac{1}{\Delta t} \partial \Sigma} \rho(K(z)), \quad (5.19)$$

5.2.3. BBVM. Using equation (3.5) and $e_i^{(\nu)} = u_i^{(\nu)} - u_i \in \mathbb{C}^{sm}$ we find the error iteration

$$\begin{aligned} (A \otimes I_m) e_i^{(\nu)} + (A_0 \otimes I_m) e_{i-1}^{(\nu)} &= \\ \Delta t (B \otimes L^+) e_i^{(\nu)} + \Delta t (B \otimes L^-) e_i^{(\nu-1)} &+ \\ \Delta t (B_0 \otimes L^+) e_{i-1}^{(\nu)} + \Delta t (B_0 \otimes L^-) e_{i-1}^{(\nu-1)}. & \end{aligned}$$

Applying the discrete Laplace transform to these equations results in

$$\begin{aligned} [(A + w^{-1}A_0) \otimes I_m] e^{(\nu)}(w) &= [\Delta t (B + w^{-1}B_0) \otimes L^+] e^{(\nu)}(w) + \\ &[\Delta t (B + w^{-1}B_0) \otimes L^-] e^{(\nu-1)}(w). \end{aligned} \quad (5.20)$$

The discrete Laplace transform of the operator $\mathcal{K}_{\Delta t}$ described by equations (3.5) is therefore

$$K_{\Delta t}(w) = K \left(\frac{1}{\Delta t} (B + w^{-1}B_0)^{-1} (A + w^{-1}A_0) \right).$$

In order to apply Lemma 5.12 the operator $\mathcal{K}_{\Delta t}$ must be bounded. A sufficient condition is provided by the following lemma. The stability domain S is the one defined by Definition 5.4.

LEMMA 5.18. *If $\sigma(\Delta t L^+) \subset \text{int } S$, then $\mathcal{K}_{\Delta t}$ is bounded in $l_p(\infty; \mathbb{C}^{sm})$.*

Proof. It is sufficient to prove that $K_{\Delta t}(w)$ is the discrete Laplace transform of an l_1 -sequence. Following the same reasoning as for Lemma 5.16 we arrive at the condition

$$\det((A + w^{-1}A_0) \otimes I_m - \Delta t (B + w^{-1}B_0) \otimes L^+) \neq 0, \quad (5.21)$$

for $|w| \geq 1$. This condition can only hold if

$$\det((A + w^{-1}A_0) - z(B + w^{-1}B_0)) \neq 0 \quad (5.22)$$

for $z \in \sigma(\Delta t L^+)$. If we assume that equality holds then using Lemma 5.9 we get

$$w \in \sigma(-(A - zB)^{-1}(A_0 - zB_0))$$

Because $|w| \geq 1$ this would mean that $z \notin \text{int } S$ which contradicts the assumption of the lemma and therefore (5.22) is true. \square

Combining Lemmas 5.12 and 5.18 with equations (5.20) and (5.6) results in the following theorem.

THEOREM 5.19. *Consider $\mathcal{K}_{\Delta t}$ as an operator in $l_p(\infty; \mathbb{C}^{sm})$ with $1 \leq p \leq \infty$, and assume $\sigma(\Delta t L^+) \subset \text{int } S$. Then,*

$$\rho(\mathcal{K}_{\Delta t}) = \sup_{z \in \frac{1}{\Delta t} \Sigma} \rho(K(z)) = \sup_{z \in \frac{1}{\Delta t} \partial \Sigma} \rho(K(z)), \quad (5.23)$$

5.3. Multigrid Convergence Analysis. As in the time stepping case the results for the two-grid iteration can be found by substituting $M(z)$ for $K(z)$. The condition on the set Σ becomes $(\sigma(\Delta t L^+) \cup \sigma(\bar{L})) \subset \text{int } S$. For proofs we refer to [18, 27, 13, 14].

5.4. Relation to a Continuous in Time Method. The asymptotic convergence factor for the continuous multigrid waveform relaxation method on infinite intervals using (3.7) as smoother, is given by

$$\rho = \sup_{z \in \mathbb{C}^+} \rho(M(z)) = \sup_{z \in i\mathbb{R}} \rho(M(z)),$$

assuming that $(\sigma(L^+) \cup \sigma(\bar{L})) \subset \mathbb{C}^-$ [18, 27, 13]. For LMFe, GLMs and BBVMs that are A -stable and for A_{k_1, k_2} -stable BVMs, i.e., when $\Sigma \subset \mathbb{C}^+$ ($\mathbb{C}^- \subset \text{int } S$), the convergence factor of discrete waveform relaxation is therefore bounded by that of continuous waveform relaxation.

5.5. Summary. The asymptotic convergence factor of the multigrid methods for time dependent problems described before, can be found from

$$\rho = \sup_{z \in \frac{1}{\Delta t} \partial \Sigma} \rho(M(z)). \quad (5.24)$$

The appropriate sets Σ for different time discretization schemes are listed in Table 5.1. The first row refers to time stepping and discrete waveform relation on finite intervals, the second row refers to discrete waveform relation on infinite intervals. Note that except for the BVM case the first row can be derived by taking $|w| \rightarrow \infty$. The function $\rho(M(z))$ can be estimated by a local mode Fourier analysis of the two-grid iteration applied to the elliptic equation

$$zu = \mathcal{L}u + f.$$

LMF	GLM	BVM	BBVM
$\frac{\alpha_k}{\beta_k}$	$\sigma(A^{-1})$	$\sigma(B^{-1}A)$	$\sigma(B^{-1}A)$
$\frac{a(w)}{b(w)}$	$\sigma((A + C(wI_r - D)^{-1}B)^{-1})$	$\frac{a(w)}{b(w)}$	$\sigma((B_0 + wB)^{-1}(A_0 + wA))$

TABLE 5.1

The sets Σ to be maximized over in equation (4.6).

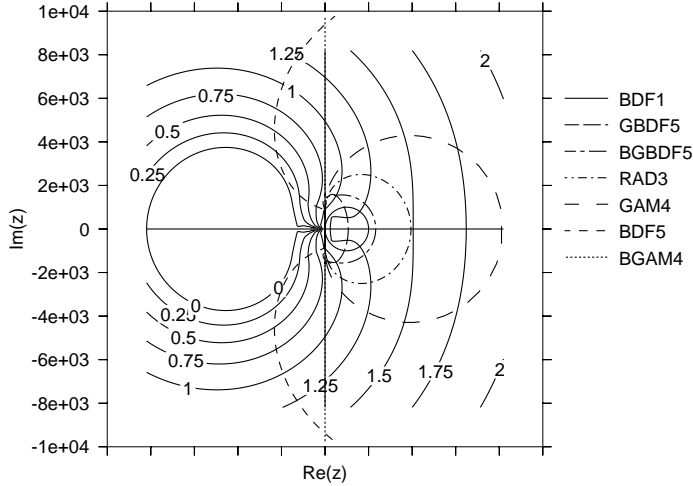


FIG. 6.1. Contour lines of the convergence rate $R(z) = -\log_{10} \rho(M(z))$ and boundaries of the sets Σ for discrete waveform relaxation on infinite time domains. The boundary for the BGAM4 discretization coincides with the imaginary axis.

6. Numerical Results. The time discretization schemes used for the numerical experiments are implicit Euler (=BDF1=GBDF1=RadauIIA1), the backward difference formula with 5 steps (BDF5), the Radau IIA implicit Runge–Kutta method with 3 stages (RAD3), the generalized Adams methods with 4 steps (GAM4) and the generalized backward difference formula with 5 steps (GBDF5). All but the implicit Euler method are fifth order methods. The GAM4 and GBDF5 boundary value methods were also used as block BVMs. The coefficients for the discretization schemes used, can be found in the Appendix.

We prefer to report convergence rates instead of convergence factors. The convergence rate R is defined by

$$R = -\log_{10}(\rho)$$

and can be interpreted as the average number of additional correct digits obtained per iteration. The asymptotic convergence rate can be estimated using

$$\rho_n^m = \frac{\|e_n\|_\infty}{\|e_{n-m}\|_\infty}, \quad R_n^m = -\frac{\log_{10}(\rho_n^m)}{m}, \quad (6.1)$$

where e_i is the error after the i -th iteration. We used the average of R_i^1 for $i = 10, \dots, 20$ for all numerical experiments reported here. This corresponds to taking the geometric mean of ρ_i^1 or taking $n = 20$, $m = 10$ in (6.1).

6.1. Heat Equation. As a first example we consider the heat equation

$$u_t = u_{xx} + u_{yy} + f \quad (6.2)$$

	BDF1	BDF5	GBDF5	GAM4	RAD3	BGBDF5	BGAM4
TS	1.29	1.32	-	-	1.16	0.92	1.08
WR	1.13	0.06	0.90	0.86	0.89	0.85	0.79

TABLE 6.1

Convergence rates obtained by a two-grid local Fourier analysis (heat equation).

	BDF1	BDF5	GBDF5	GAM4	RAD3	BGBDF5	BGAM4
TS	1,1	1,5	-	-	1,3	1,5	1,4
WR	1020,1	1020,1	1020,1	1020,1	360,3	204,5	255,4

TABLE 6.2

Number of time steps and number of unknowns per time step for the time discretizations used in the numerical experiments.

on the unit square. The problem is discretized on a grid with $\Delta x = \Delta y = 2^{-5}$. The two-grid analysis was conducted for full weighting restriction, bilinear interpolation and 1 pre- and 1 postsmoothing red-black Gauss–Seidel step. Time integration is over $\Omega_t = [0, 1]$, i.e., $\Delta t = 3 \cdot 10^{-3}$ for the RAD3 discretization and $\Delta t = 10^{-3}$ for all others. Figure 6.1 shows contour lines of the predicted convergence rate $R(z)$ together with the boundaries of the regions Σ of the discretization schemes used in waveform relaxation on infinite time domains. We expect convergence for the BDF5 discretization to be slow since the boundary of the corresponding set Σ nearly touches the line $R = 0$. If these lines were to touch or intersect, convergence would no longer be guaranteed. Note that all but the BDF5 and BGBDF5 method are A-stable. The BGBDF5 method is, however, very close to being A-stable and we expect very good convergence. The results are presented in Table 6.1. The first row contains the results for time stepping or discrete waveform relaxation on a finite domain. These results are obtained by maximizing $R(z)$ over Σ , which, in this case, is a discrete set of points. The BVMs (GAM4 and GBDF5) cannot be used in a time stepping sense. The second row of Table 6.1 contains the results for discrete waveform relaxation on infinite time domains. The results were obtained by maximizing over a number of points on the curves depicted in Figure 6.1. The points were found by taking $w = \exp(\frac{2\pi ik}{n})$, $k = 0, \dots, n-1$, $n = 200$ in the definitions of the sets Σ . Since $\Sigma = \mathbb{C}^+$ for the BGAM4 discretization the last result in the table provides a good approximation to the convergence rate for continuous waveform relaxation. From these results we expect good convergence for all methods except waveform relaxation with BDF5 time discretization.

Decreasing Δt corresponds to a proportional increase in the size of the sets Σ in Figure 6.1. Similarly when Δx and Δy are decreased simultaneously the contour lines of $\rho(M(z))$ have approximately the same shape, but on a larger scale (the dependency is quadratic in this case). This can be explained by observing that scaling L , L^+ and L^- in the formulas for $K(z)$ and $M(z)$ is equivalent to the inverse scaling of z . Decreasing Δx and Δy will of course also increase the size of all the matrices involved, but this can be neglected for fine discretizations (large matrices). Figure 6.1 can thus be used to estimate the convergence for other discretization resolutions by changing the relative size of the curves depicted.

For the numerical experiments the initial conditions, the Dirichlet boundary conditions and the source term f are chosen such that the exact solution is $u = 0$. We use a standard W-cycle geometric multigrid algorithm with 5 levels, full weighting restriction, bilinear interpolation and 1 pre- and 1 postsmoothing red-black Gauss–

	BDF1	BDF5	GBDF5	GAM4	RAD3	BGBDF5	BGAM4
TS	1.18	1.11	-	-	1.33	1.11	1.10
WR	0.92	0.03	0.87	0.87	0.87	0.87	0.88

TABLE 6.3

Convergence rates obtained from numerical experiments (heat equation).

	BDF1	BDF5	GBDF5	GAM4	RAD3	BGBDF5	BGAM4
TS	1.50	1.34	-	-	1.91	1.34	1.47
WR	1.22	0.69	1.22	1.22	1.22	1.22	1.22

TABLE 6.4

Convergence rates for the MGS iteration (diffusion equation with variable coefficients).

Seidel step. The spatial domain is discretized using finite differences. The grid sizes are $\Delta x = \Delta y = 2^{-1}$ on the coarsest grid and $\Delta x = \Delta y = 2^{-5}$ on the finest grid. As in the analysis the time steps are $\Delta t = 3 \cdot 10^{-3}$ for RAD3 and $\Delta t = 10^{-3}$ otherwise. The number of time steps for each method was chosen such that the number of unknowns per spatial grid point was minimal in the time stepping case and the same for all methods in the waveform relaxation case. Table 6.2 shows the number of time steps and the number of unknowns per time step for each method. All unknowns were set to 1 for the initial approximation.

The results are presented in Table 6.3. It is clear that they correspond very well to the results obtained from the theoretical analysis. The theoretical convergence rates are not strictly lower bounds (upper bounds on the convergence factors), since they are asymptotic convergence rates obtained using a local Fourier mode two-grid approximation. Similar results with poor convergence for the BDF5 method and convergence rates close to 1 for the other methods were obtained for waveform relaxation on grids of dimensions $2^5 \times 2^5 \times 10^4$, $2^8 \times 2^8 \times 360$ and $2^9 \times 2^9 \times 60$. We can conclude that, for this problem, multigrid methods are very efficient even when using high order time discretization schemes.

6.2. Diffusion Equation with variable coefficients. The next example is a more general diffusion equation of the form

$$\begin{aligned} u_t &= (au_x)_x + (bu_y)_y + f, \\ a(x, y, t) &= \exp(10(x - y)), \\ b(x, y, t) &= \exp(-10(x - y)). \end{aligned}$$

This is a so called anisotropic problem since the diffusion coefficients in a certain point can be very different depending on the direction of diffusion. Standard multigrid methods do not handle this type of problems well and yield very slow convergence ($R = 0.1$). However, just like we did with the original multigrid algorithm, we can extend multigrid methods developed for time independent anisotropic problems to time dependent ones. We used here the “multigrid as smoother” (MGS) method which uses the same simple smoothers as before, together with an extended hierarchy of coarse grids. For more information on multigrid for anisotropic problems we refer to [25, 26, 23, 28, 22].

The same discretizations, initial and boundary conditions and initial approximation were used as for the previous example. The multigrid algorithm uses full weighting restriction, bilinear and linear interpolation, 1 pre- and 1 postsmoothing step. The smoothing consists of a multigrid semicoarsening step in the x -direction

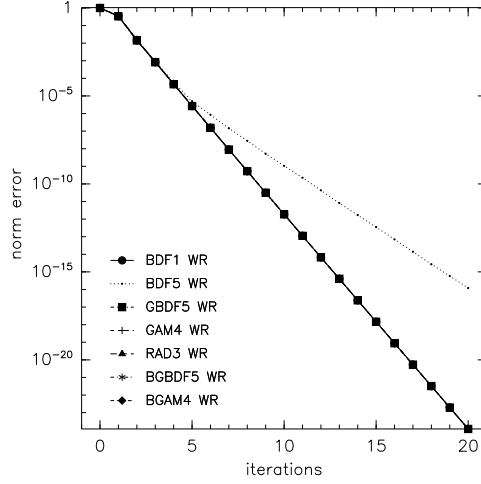


FIG. 6.2. *Convergence of MGS method (diffusion equation with variable coefficients).*

followed by one in the y -direction both using only 1 presmoothing red-black Gauss–Seidel step. A hierarchy of 5×5 grids is used. The results are presented in Table 6.4. The convergence history of the MGS waveform relaxation method is visualized in Figure 6.2 for different time discretization schemes. Except for the BDF5 method, the convergence histories are indistinguishable. We see once more that it is possible to obtain very efficient multigrid methods even when high order time discretization schemes are used.

Appendix A. Coefficient Matrices.

A.1. Implicit Runge–Kutta methods. We only considered the RadauIIA method with 3 stages.

RAD3.

$$A = \begin{bmatrix} \frac{88-7\sqrt{6}}{360} & \frac{296-169\sqrt{6}}{1800} & \frac{-2+3\sqrt{6}}{225} \\ \frac{296+169\sqrt{6}}{1800} & \frac{88+7\sqrt{6}}{360} & \frac{-2-3\sqrt{6}}{225} \\ \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} \end{bmatrix}, \quad b = \begin{bmatrix} \frac{16-\sqrt{6}}{36} \\ \frac{16+\sqrt{6}}{36} \\ \frac{1}{9} \end{bmatrix}, \quad c = \begin{bmatrix} \frac{4-\sqrt{6}}{10} \\ \frac{4+\sqrt{6}}{10} \\ 1 \end{bmatrix}$$

A.2. Boundary Value Methods. The coefficient matrices A_e and B_e for the LMFe and BBVMs we used are given below. The boldface rows correspond to the coefficients α_j and β_j of the basic method (2.8). The rows above correspond to the initial conditions and the rows below to the final conditions. The coefficient matrices for $s > k$ are obtained by repeating the boldface rows as in (2.11).

BDF1.

$$A_e = [-1 \mid \mathbf{1}], \quad B_e = [\mathbf{0} \mid \mathbf{1}]$$

GAM4.

$$A_e = \left[\begin{array}{c|cccc} -1 & & & & \\ \hline -1 & & & & \\ & 1 & & & \\ & & -1 & & \\ & & & 1 & \\ & & & & -1 & 1 \end{array} \right], \quad B_e = \frac{1}{720} \left[\begin{array}{c|cccc} 251 & 646 & -264 & 106 & -19 \\ \hline -19 & \mathbf{346} & \mathbf{456} & -\mathbf{74} & \mathbf{11} \\ & 11 & -74 & 456 & 346 & -19 \\ -19 & 106 & -264 & 646 & 251 \end{array} \right]$$

GBDF5.

$$A_e = \frac{1}{60} \begin{bmatrix} -12 & -65 & 120 & -60 & 20 & -3 \\ 3 & -30 & -20 & 60 & -15 & 2 \\ -2 & 15 & -60 & 20 & 30 & -3 \\ 3 & -20 & 60 & -120 & 65 & 12 \\ -12 & 75 & -200 & 300 & -300 & 137 \end{bmatrix}, \quad B_e = \begin{bmatrix} 0 & 1 & & & & \\ 0 & & 1 & & & \\ 0 & & & 1 & & \\ 0 & & & & 1 & \\ 0 & & & & & 1 \end{bmatrix}$$

BDF5. The last rows of the previous matrices contain the coefficients of the BDF5 method.

RAD3. The RadauIIA methods can easily be formulated as a BBVM. It is important to note that the points within a time step are not equidistant as in the previous BVMs and that the time step Δt has a different interpretation.

$$A_e = \begin{bmatrix} -1 & 1 & & & \\ -1 & & 1 & & \\ -1 & & & 1 & \end{bmatrix}, \quad B_e = \begin{bmatrix} 0 & \frac{88-7\sqrt{6}}{360} & \frac{296-169\sqrt{6}}{1800} & \frac{-2+3\sqrt{6}}{225} \\ 0 & \frac{296+169\sqrt{6}}{1800} & \frac{88+7\sqrt{6}}{360} & \frac{-2-3\sqrt{6}}{225} \\ 0 & \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} \end{bmatrix}$$

Acknowledgements. Some of the results were verified using a version of the Fourier analysis code `LFA00_2D_scalar` [32] extended to complex numbers in collaboration with its author.

REFERENCES

- [1] D. BERTACCINI, *A circulant preconditioner for the systems of LMF-based ODE codes*, SIAM J. Sci. Comput., 22 (2000), pp. 767–786.
- [2] D. BERTACCINI AND M. K. NG, *Block $\{\omega\}$ -circulant preconditioners for the systems of differential equations*, Calcolo, 40 (2003), pp. 71–90.
- [3] L. BRUGNANO AND D. TRIGIANTE, *Solving differential problems by multistep initial and boundary value methods*, vol. 6 of Stability and Control: Theory, Methods and Applications, Gordon and Breach Science Publishers, Amsterdam, 1998.
- [4] K. BURRAGE, *Parallel and sequential methods for ordinary differential equations*, Numerical Mathematics and Scientific Computation, The Clarendon Press Oxford University Press, New York, 1995. Oxford Science Publications.
- [5] J. C. BUTCHER, *The numerical analysis of ordinary differential equations*, A Wiley-Interscience Publication, John Wiley & Sons Ltd., Chichester, 1987. Runge-Kutta and general linear methods.
- [6] R. H. CHAN, M. K. NG, AND X.-Q. JIN, *Strang-type preconditioners for systems of LMF-based ODE codes*, IMA J. Numer. Anal., 21 (2001), pp. 451–462.
- [7] E. HAIRER, S. P. NØRSETT, AND G. WANNER, *Solving ordinary differential equations. I*, vol. 8 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 1993. Nonstiff problems.
- [8] E. HAIRER AND G. WANNER, *Solving ordinary differential equations. II*, vol. 14 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 1996. Stiff and differential-algebraic problems.
- [9] G. HORTON AND S. VANDEWALLE, *A space-time multigrid method for parabolic partial differential equations*, SIAM J. Sci. Comput., 16 (1995), pp. 848–864.
- [10] F. IAVERNARO AND F. MAZZIA, *Solving ordinary differential equations by generalized Adams methods: properties and implementation techniques*, Appl. Numer. Math., 28 (1998), pp. 107–126.
- [11] ———, *Block-boundary value methods for the solution of ordinary differential equations*, SIAM J. Sci. Comput., 21 (1999), pp. 323–339.
- [12] F. IAVERNARO AND D. TRIGIANTE, *Preconditioning and conditioning of systems arising from boundary value methods*, Nonlinear Dyn. Syst. Theory, 1 (2001), pp. 59–79.
- [13] J. JANSSEN AND S. VANDEWALLE, *Multigrid waveform relaxation of spatial finite element meshes: the continuous-time case*, SIAM J. Numer. Anal., 33 (1996), pp. 456–474.
- [14] ———, *Multigrid waveform relaxation on spatial finite element meshes: the discrete-time case*, SIAM J. Sci. Comput., 17 (1996), pp. 133–155.
- [15] J. D. LAMBERT, *Computational methods in ordinary differential equations*, John Wiley & Sons, London-New York-Sydney, 1973.

- [16] E. LELARSMEE, A. RUEHLI, AND A. SANGIOVANNI-VINCENTELLI, *The waveform relaxation method for time-domain analysis of large-scale integrated circuits*, IEEE Trans. CAD Integrated Circuits and Systems, 1 (1982), pp. 131–145.
- [17] C. LUBICH, *On the stability of linear multistep methods for Volterra convolution equations*, IMA J. Numer. Anal., 3 (1983), pp. 439–465.
- [18] C. LUBICH AND A. OSTERMANN, *Multigrid dynamic iteration for parabolic equations*, BIT, 27 (1987), pp. 216–234.
- [19] A. LUMSDAINE AND D. WU, *Spectra and pseudospectra of waveform relaxation operators*, SIAM J. Sci. Comput., 18 (1997), pp. 286–304.
- [20] U. MIEKKALA AND O. NEVANLINNA, *Convergence of dynamic iteration methods for initial value problem*, SIAM J. Sci. Statist. Comput., 8 (1987), pp. 459–482.
- [21] ———, *Sets of convergence and stability regions*, BIT, 27 (1987), pp. 554–584.
- [22] C. W. OOSTERLEE, *The convergence of parallel multiblock multigrid methods*, Appl. Numer. Math., 19 (1995), pp. 115–128.
- [23] C. W. OOSTERLEE AND P. WESSELING, *On the robustness of a multiple semi-coarsened grid method*, Z. Angew. Math. Mech., 75 (1995), pp. 251–257.
- [24] J. H. SALTZ AND V. K. NAIK, *Towards developing robust algorithms for solving partial differential equations on MIMD machines*, Parallel Comput., 6 (1988), pp. 19–44.
- [25] U. TROTTEBERG, C. W. OOSTERLEE, AND A. SCHÜLLER, *Multigrid*, Academic Press Inc., San Diego, CA, 2001.
- [26] J. VAN LENT AND S. VANDEWALLE, *Multigrid waveform relaxation for anisotropic partial differential equations*, Numer. Algorithms, 31 (2002), pp. 361–380.
- [27] S. VANDEWALLE, *Parallel multigrid waveform relaxation for parabolic problems*, Teubner Skripten zur Numerik., B. G. Teubner, Stuttgart, 1993.
- [28] T. WASHIO AND C. W. OOSTERLEE, *Flexible multiple semicoarsening for three-dimensional singularly perturbed problems*, SIAM J. Sci. Comput., 19 (1998), pp. 1646–1666.
- [29] P. WESSELING, *An introduction to multigrid methods*, Pure and Applied Mathematics, John Wiley & Sons Ltd., Chichester, 1992.
- [30] J. WHITE AND A. SANGIOVANNI-VINCENTELLI, *Relaxation Techniques for the Simulation of VLSI Circuits*, Kluwer Academic Publishers, Boston, 1987.
- [31] J. WHITE, A. SANGIOVANNI-VINCENTELLI, F. ODEH, AND A. RUEHLI, *Waveform relaxation theory and practice*, Trans. Soc. for Comp. Simulation, 2 (1985).
- [32] R. WIENANDS AND C. W. OOSTERLEE, *On three-grid Fourier analysis for multigrid*, SIAM J. Sci. Comput., 23 (2001), pp. 651–671.
- [33] R. WIENANDS, C. W. OOSTERLEE, AND T. WASHIO, *Fourier analysis of GMRES(m) preconditioned by multigrid*, SIAM J. Sci. Comput., 22 (2000), pp. 582–603.
- [34] D. E. WOMBLE, *A time-stepping algorithm for parallel computers*, SIAM J. Sci. Statist. Comput., 11 (1990), pp. 824–837.