SIAM J. NUMER. ANAL.
Vol. 33, No. 2, pp. 456-474, April 1996

003

MULTIGRID WAVEFORM RELAXATION ON SPATIAL FINITE ELEMENT MESHES: THE CONTINUOUS-TIME CASE*

JAN JANSSEN† AND STEFAN VANDEWALLE‡

Abstract. The waveform relaxation method and its multigrid acceleration are studied as solution procedures for the system of ordinary differential equations obtained by finite element discretisation of a linear parabolic initial boundary value problem. The convergence properties of the continuous-time algorithm are theoretically investigated on finite-length and infinite-length time-intervals. In addition, quantitative convergence estimates and numerical results are presented for one-dimensional and two-dimensional model problems.

Key words. parabolic partial differential equations, finite elements, waveform relaxation, dynamic iteration, multigrid

AMS subject classifications. 65F10, 65L05, 65M55, 65M60

1. Introduction.

We consider the numerical solution on a spatial finite element mesh of the following parabolic partial differential equation (PDE):

$$\frac{\partial \mathbf{u}}{\partial t}(x,t) + \mathcal{L}\mathbf{u}(x,t) = \mathbf{f}(x,t), \qquad x \in \Omega\,, \quad t > 0\,, \tag{1.1}$$

with a linear boundary condition and given initial values. In (1.1), $\mathcal{L}$ denotes a linear second-order uniformly strongly elliptic operator with time-independent coefficients, and Ω is a bounded open spatial domain.

The weak formulation of the parabolic initial boundary value problem, supplied with homogeneous Dirichlet boundary conditions, is given as follows: find $\mathbf{u}(\cdot, t) \in H_0^1(\Omega)$ such that

$$\left(\frac{\partial \mathbf{u}}{\partial t}, \mathbf{v}\right) + a(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) \qquad \text{for all } \mathbf{v} \in H_0^1(\Omega)\,,$$

where $a(\cdot, \cdot)$ is the bilinear form corresponding to operator $\mathcal{L}$. In the Galerkin approach, $\mathbf{u}(x, t)$ is approximated, for each t, in a finite-dimensional subspace H_d of the Sobolev space H_0^1. This subspace is spanned by a set of linearly independent basis functions, associated with the points of a discrete mesh, i.e., $H_d = \text{span}\{\varphi_1, \ldots, \varphi_d\}$. The approximation $\bar{u} = \sum_{i=1}^{d} u_i(t)\varphi_i(x)$ is found by solving the following set of equations:

$$\left(\frac{\partial \bar{u}}{\partial t}, \varphi_j\right) + a(\bar{u}, \varphi_j) = (\mathbf{f}, \varphi_j) \qquad \text{for } j = 1, \ldots, d\,.$$

In terms of the mass matrix $B = \{(\varphi_i, \varphi_j)\}$ and the stiffness matrix $A = \{a(\varphi_i, \varphi_j)\}$, we may rewrite these equations in a more standard form, as a system of ordinary differential equations (ODEs)

$$B\dot{u} + Au = f\,, \tag{1.2}$$

*Received by the editors November 23, 1993; accepted for publication (in revised form) May 11, 1994.

†Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3001 Heverlee, Belgium (jan.janssen@cs.kuleuven.ac.be). This text presents research results of the Belgian Incentive Programme "Information Technology", Computer Science of the Future (IT/IF/5), initiated by the Belgian State—Prime Minister's Service—Federal Office for Scientific, Technical and Cultural Affairs. The scientific responsibility is assumed by its authors.

‡California Institute of Technology, Applied Mathematics 217-50, Pasadena, CA 91125 (na.vandewalle@na-net.ornl.gov). This work was supported in part by NSF Cooperative Agreement CCR-9120008.

with $u = [u_1(t), u_2(t), \ldots, u_d(t)]^t$ and $f = [(\mathbf{f}, \varphi_1), (\mathbf{f}, \varphi_2), \ldots, (\mathbf{f}, \varphi_d)]^t$. Matrix B is the Grammian matrix of a set of linearly independent functions and therefore is positive definite symmetric. An initial condition of the form $u(0) = u_0$ is obtained by means of interpolation or projection of the given initial condition for the PDE. A conforming finite element discretisation of (1.1), equipped with more general boundary conditions, also leads to a system of the form (1.2). For further details, we refer to [3] and [18].

The waveform relaxation method, also called the dynamic iteration method, is an iterative technique for solving systems of ODEs. Its background is in electrical network simulation [9, 14]. It differs from most standard iterative techniques in that it is a continuous-time method, iterating with functions, and thereby well suited for parallel computation. It is based on a splitting of B and A, i.e., $B = M_B - N_B$ and $A = M_A - N_A$, and an iteration of the form

$$M_B \dot{u}^{(\nu)} + M_A u^{(\nu)} = N_B \dot{u}^{(\nu-1)} + N_A u^{(\nu-1)} + f \, .$$

The convergence of this waveform relaxation method has been studied exhaustively for systems of the form (1.2), where B is the identity matrix. Such systems arise when (1.1) is discretised using finite differences. In [12] and [13], Miekkala and Nevanlinna formulated the convergence characteristics of the method in terms of the spectral radius of the corresponding waveform relaxation operator, which is of linear Volterra convolution type. An analogous study for the multigrid acceleration of the waveform relaxation method is performed by Lubich and Ostermann in [10]. A survey and a discussion of a parallel implementation of these methods can be found in the book by Vandewalle [19] and in [20] and [21]. We also mention the paper by Miekkala [11], where the convergence properties of the waveform relaxation method are studied for differential-algebraic systems of the form (1.2), where B is possibly singular.

In this paper, we shall concentrate on systems (1.2) with nonsingular B. The presence of the matrix B leads to standard and multigrid waveform relaxation operators that are more general than the corresponding operators in [12] and [10], respectively. In particular, a matrix multiplication operator is added to the operator of linear Volterra convolution type. Our analysis generalises the analysis of the above references: by setting $B = M_B = I$ and $N_B = 0$, their results are regained. This paper is organised as follows. In §2, we study the spectral properties of a general operator consisting of a matrix multiplication part and a linear Volterra convolution part. These results will be used to investigate the convergence of the standard waveform relaxation method (§3) and of its multigrid acceleration (§4). In §3 we allow B and A to be fairly general matrices; we do not restrict the discussion to matrices derived from parabolic PDEs. In §5, we conclude with some specific theoretical results for the heat equation, which are subsequently validated by numerical experiments.

2. Spectral properties of a special operator. It will turn out that both the waveform relaxation iteration and its two-grid acceleration can be written as a successive approximation scheme

$$u^{(\nu)} = \mathcal{H} u^{(\nu-1)} + \varphi \, , \tag{2.1}$$

with $\mathcal{H}$ an operator of the form

$$\mathcal{H} u = H u + \mathcal{H}_c u \, . \tag{2.2}$$

The operator $\mathcal{H}_c$ is a linear Volterra convolution operator with a matrix-valued kernel h_c,

$$\mathcal{H}_c u(t) = h_c \star u(t) = \int_0^t h_c(t-s) u(s) ds \, ,$$

and H is a $d \times d$ complex matrix.

The properties of this operator will be studied in the context of normed linear spaces. We shall consider in particular the spaces of pth power integrable Lebesgue measurable functions $L_p((0, \infty); \mathbb{C}^d)$, or $L_p(0, \infty)$ for short, with the usual mean p-norm, and the space of continuous functions $C([0, T]; \mathbb{C}^d)$, or $C[0, T]$, equipped with the maximum-norm

$$||u||_T = \max_{t \in [0,T]} ||u(t)|| ,$$

where $||\,.\,||$ is any usual $\mathbb{C}^d$ vector-norm. Recall that convergence of the general successive approximation scheme (2.1) is guaranteed if and only if the spectral radius of $\mathcal{H}$ is smaller than 1. The spectral radius of a bounded linear operator in a complex normed linear space is given by

$$\rho(\mathcal{H}) = \lim_{n \to \infty} \sqrt[n]{||\mathcal{H}^n||} . \tag{2.3}$$

It also equals the smallest value of ρ for which $|\lambda| > \rho$ implies that $\lambda - \mathcal{H}$ has a bounded inverse or, equivalently,

$$\rho(\mathcal{H}) = \sup_{\lambda \in \sigma(\mathcal{H})} |\lambda| ,$$

where $\sigma(\mathcal{H})$ denotes the spectrum of $\mathcal{H}$.

2.1. Spectral radius on finite time-intervals.

LEMMA 2.1. *Suppose $h_c \in C[0, T]$ and consider $\mathcal{H}$ as an operator in $C[0, T]$. Then, $\mathcal{H}$ is a bounded operator and $\rho(\mathcal{H}) = \rho(H)$.*

The proof given below is based on a stability result from perturbation theory. An elementary proof can be found in the Appendix.

Proof. Bounding $\mathcal{H}u$ gives

$$\begin{aligned} ||\mathcal{H}u|| &\leq ||Hu|| + ||\mathcal{H}_c u|| \\ &\leq (||H|| + T\,||h_c||)||u|| , \end{aligned}$$

where $||\,.\,||$ denotes (for notational simplicity) both the maximum-norm in $C[0, T]$ and the matrix-norm induced by the $\mathbb{C}^d$ vector-norm. Hence, $\mathcal{H}$ is a bounded operator with $||\mathcal{H}|| \leq ||H|| + T\,||h_c||$.

Since the linear convolution operator $\mathcal{H}_c$ is compact, operator $\mathcal{H}$ is a compact perturbation of H. From [7, Chap. IV, Thm. 5.35], it then follows that

$$\sigma_e(H) = \sigma_e(\mathcal{H}) ,$$

where $\sigma_e(H)$ and $\sigma_e(\mathcal{H})$ denote the essential spectra of the (closed) operators H and $\mathcal{H}$, respectively [7, Chap. IV, §5.6]. We show below that equality also holds for the spectra.

It is easily seen that the spectrum of the matrix multiplication operator H is equal to the spectrum of the matrix H. For any $\lambda \in \sigma(H)$ both the dimension of the null space and the codimension of the range of the operator $H - \lambda$ in $C[0, T]$ are infinite. Hence, the essential spectrum of the matrix multiplication operator H equals the spectrum of H, or

$$\sigma_e(H) = \sigma(H) .$$

It follows that $\sigma_e(\mathcal{H})$ is a finite set, and any point $\mu \in \sigma(\mathcal{H}) \setminus \sigma_e(\mathcal{H})$ must be an isolated eigenvalue of $\mathcal{H}$ [7, Chap. IV, Thm. 5.33]. We will show that there are no such points, i.e., $\sigma(\mathcal{H}) = \sigma_e(\mathcal{H})$. Suppose we have some $u \neq 0$, such that

$$\mathcal{H}u = Hu + \mathcal{H}_c u = \mu u .$$

Since $\mu \notin \sigma_e(\mathcal{H}) = \sigma(H)$, this can be rewritten as

$$(\mu I - H)^{-1}\mathcal{H}_c u = u \,,$$

which means that $(\mu I - H)^{-1}\mathcal{H}_c$ has 1 as an eigenvalue. However, $(\mu I - H)^{-1}\mathcal{H}_c$ is a linear Volterra convolution operator with continuous kernel, whose spectrum equals the singleton $\{0\}$; see, e.g., [8, p. 33]. Hence, $\sigma(\mathcal{H}) = \sigma_e(\mathcal{H})$, and thus $\sigma(\mathcal{H}) = \sigma(H)$, which completes the proof. □

2.2. Spectral radius on infinite time-intervals. The proof of the next lemma is based on a theorem by Paley and Wiener; see, e.g., [4, p. 45] or [15, p. 60]. The theorem deals with the solution of a linear Volterra integral equation, $x + k \star x = f$. Its solution can be expressed in terms of a *resolvent* function r, which is defined by the two equations $r + k \star r = r + r \star k = k$. In particular, $x = f - r \star f$. A necessary and sufficient condition for the boundedness of the resolvent r, and hence, for the boundedness of the solution x, is given in the theorem.

THEOREM 2.2 (Paley-Wiener). *Let $k \in L_1(0, \infty)$. Then the resolvent r of k satisfies $r \in L_1(0, \infty)$ if and only if $\det(I + \mathbf{K}(z)) \neq 0$ for $\mathrm{Re}(z) \geq 0$, where $\mathbf{K}(z)$ denotes the Laplace-transform of k.*

Note that the theorem holds both for scalar and for vector-valued functions.

LEMMA 2.3. *Suppose $h_c \in L_1(0, \infty)$, and consider $\mathcal{H}$ as an operator in $L_p(0, \infty)$ with $1 \leq p \leq \infty$. Then, $\mathcal{H}$ is a bounded operator with spectral radius*

$$\rho(\mathcal{H}) = \sup_{\mathrm{Re}(z) \geq 0} \rho(\mathbf{H}(z)) \tag{2.4}$$

$$= \sup_{\xi \in \mathbb{R}} \rho(\mathbf{H}(i\xi)) \,, \tag{2.5}$$

where $\mathbf{H}(z) = H + \mathbf{H_c}(z)$, and $\mathbf{H_c}(z)$ denotes the Laplace-transform of h_c.

Proof. The boundedness of $\mathcal{H}$ is an immediate consequence of Young's inequality on the convolution operator [16, p. 28].

By definition, the spectral radius of $\mathcal{H}$ is the smallest value of ρ for which $|\lambda| > \rho$ implies that $\lambda - \mathcal{H}$ has a bounded inverse in $L_p(0, \infty)$. Suppose $\lambda \neq 0$, and

$$\lambda u - \mathcal{H}u = (\lambda I - H)\, u - h_c \star u = f \,. \tag{2.6}$$

First, we suppose that λ is not an eigenvalue of H, i.e., $\lambda \notin \sigma(H)$. In that case, (2.6) can be rewritten as

$$u - (\lambda I - H)^{-1} h_c \star u = (\lambda I - H)^{-1} f \,.$$

Applying the Paley-Wiener theorem, we find that u is bounded if and only if

$$\det\big(I - (\lambda I - H)^{-1}\mathbf{H_c}(z)\big) \neq 0 \;\text{ for }\; \mathrm{Re}(z) \geq 0$$

or, equivalently,

$$\det(\lambda I - (H + \mathbf{H_c}(z))) \neq 0 \;\text{ for }\; \mathrm{Re}(z) \geq 0 \,.$$

The set Σ of all λ, with $\lambda \notin \sigma(H)$, that leads to an unbounded solution u, is

$$\Sigma = \bigcup_{\mathrm{Re}(z) \geq 0} \sigma(H + \mathbf{H_c}(z)) \setminus \sigma(H) \,.$$

Define $\bar{\rho}$ as $\sup\{|\lambda| : \lambda \in \Sigma\}$. (Note that $\rho(\mathcal{H}) \geq \bar{\rho}$, with "$\geq$" instead of "$=$" since we did not yet take all possible λ into account.) By the continuity of the eigenvalues of $H + \mathbf{H_c}(z)$ as a function of z, it is clear that

$$\bar{\rho} = \sup_{\mathrm{Re}(z)\geq 0} \rho(H + \mathbf{H_c}(z)) .$$

We still need to consider the λ that are eigenvalues of H. However, because

$$\lim_{z\to\infty} (H + \mathbf{H_c}(z)) = H , \tag{2.7}$$

these eigenvalues are in magnitude smaller than or equal to $\bar{\rho}$. Thus, $\rho(\mathcal{H}) = \bar{\rho}$, and thereby (2.4) follows. The second equality (2.5) is obtained by application of the maximum principle. □

In $L_2(0, \infty)$, an analogous result holds for the norm.

LEMMA 2.4. *Suppose $h_c \in L_1(0, \infty)$, and consider $\mathcal{H}$ as an operator in $L_2(0, \infty)$. Denote by $||\,.\,||_2$ the L_2-norm and by $||\,.\,||$ the standard Euclidean vector-norm. Then,*

$$||\mathcal{H}||_2 = \sup_{\mathrm{Re}(z)\geq 0} ||\mathbf{H}(z)|| \tag{2.8}$$

$$= \sup_{\xi\in\mathbb{R}} ||\mathbf{H}(i\xi)|| , \tag{2.9}$$

where $\mathbf{H}(z) = H + \mathbf{H_c}(z)$, and $\mathbf{H_c}(z)$ denotes the Laplace-transform of h_c.

Proof. This result is a consequence of Parseval's formula; see, e.g., [2, p. 8]. □

REMARK 2.1. Consider $\mathcal{H}$ as an operator in $C[0, T]$. From (2.7), we derive

$$\rho(\mathcal{H}) = \rho(H) = \rho(\mathbf{H}(\infty)) ,$$

which means that the spectral radius of $\mathcal{H}$ on finite time-intervals is smaller than the spectral radius of $\mathcal{H}$ on infinite time-intervals.

3. The waveform relaxation method.

3.1. The waveform relaxation operator. Consider the following linear initial value problem:

$$B\dot{u} + Au = f , \text{ with } u(0) = u_0 , \ t > 0 , \tag{3.1}$$

where B and A are complex $d \times d$ matrices, and u and f are $\mathbb{C}^d$-valued functions in time. In the present paper B is a nonsingular matrix. The solution to (3.1) is then formally given by

$$u(t) = e^{-B^{-1}At}u_0 + \int_0^t e^{B^{-1}A(s-t)}B^{-1}f(s)ds . \tag{3.2}$$

By introducing the splittings $B = M_B - N_B$ and $A = M_A - N_A$, the basic continuous-time waveform relaxation iteration can be written as

$$M_B\dot{u}^{(\nu)} + M_Au^{(\nu)} = N_B\dot{u}^{(\nu-1)} + N_Au^{(\nu-1)} + f , \text{ with } u^{(\nu)}(0) = u_0 , \ t > 0 . \tag{3.3}$$

The iteration is usually started by choosing the zeroth iterate $u^{(0)}(t) = u_0$, $t > 0$. We shall always assume M_B to be invertible. Using (3.2), we can rewrite iteration (3.3) as an explicit successive approximation scheme: $u^{(\nu)} = \mathcal{K}u^{(\nu-1)} + \varphi$. The right-hand-side function φ and the *waveform relaxation operator* $\mathcal{K}$ are given by

$$\varphi(t) = e^{-M_B^{-1}M_At}(I - M_B^{-1}N_B)u_0 + \int_0^t e^{M_B^{-1}M_A(s-t)}M_B^{-1}f(s)ds ,$$

$$\mathcal{K}u(t) = M_B^{-1}N_Bu(t) + \mathcal{K}_cu(t) . \tag{3.4}$$

$\mathcal{K}_c$ is a linear Volterra convolution operator with kernel k_c:

$$\mathcal{K}_c u(t) = k_c \star u(t) = \int_0^t k_c(t-s)u(s)ds\,,$$

$$k_c(t) = e^{-M_B^{-1}M_A t}M_B^{-1}(N_A - M_A M_B^{-1}N_B)\,. \tag{3.5}$$

Let $e^{(\nu)}$ be the error of the νth waveform relaxation iterate, i.e., $e^{(\nu)} = u^{(\nu)} - u$. It satisfies $e^{(\nu)} = \mathcal{K}e^{(\nu-1)}$. That is, it is the solution to the differential equation

$$M_B\dot{e}^{(\nu)} + M_A e^{(\nu)} = N_B\dot{e}^{(\nu-1)} + N_A e^{(\nu-1)}\,,\ \text{with } e^{(\nu)}(0) = 0\,,\ t > 0\,. \tag{3.6}$$

If we denote by $\tilde{e}^{(\nu)}(z)$ the Laplace-transform of $e^{(\nu)}$, then we get by Laplace-transforming (3.6) that

$$\tilde{e}^{(\nu)}(z) = \mathbf{K}(z)\tilde{e}^{(\nu-1)}(z) = (M_B^{-1}N_B + \mathbf{K_c}(z))\tilde{e}^{(\nu-1)}(z)\,, \tag{3.7}$$

where $\mathbf{K_c}(z) = (zM_B + M_A)^{-1}(N_A - M_A M_B^{-1}N_B)$ is the Laplace-transform of the kernel k_c, and

$$\mathbf{K}(z) = (zM_B + M_A)^{-1}(zN_B + N_A)\,. \tag{3.8}$$

We shall hereafter refer to $\mathbf{K}(z)$ as the waveform relaxation matrix or the dynamic iteration matrix of operator $\mathcal{K}$.

REMARK 3.1. Assume matrices B and A to be decomposed as $B = -L_B + D_B - U_B$ and $A = -L_A + D_A - U_A$, where D_B and D_A are diagonal matrices, L_B and L_A are strictly lower triangular matrices, and U_B and U_A are strictly upper triangular matrices. The splittings

$$M_B = D_B\,,\ N_B = L_B + U_B\,,\ M_A = D_A\,,\ N_A = L_A + U_A,$$
$$M_B = -L_B + D_B\,,\ N_B = U_B\,,\ M_A = -L_A + D_A\,,\ N_A = U_A$$

define, respectively, the Jacobi and Gauss-Seidel waveform relaxation methods.

3.2. Convergence analysis.

3.2.1. Convergence on finite time-intervals. The spectral radius of the waveform relaxation operator as an operator on finite-length time-intervals is known to be equal to zero when B is the identity matrix [12, p. 461]. That is, convergence of the method is ultimately superlinear. In the current section we shall derive the equivalent formula for the general non-singular B. It turns out that the convergence is ultimately linear and solely dependent on the splitting of B.

THEOREM 3.1. *The waveform relaxation operator $\mathcal{K}$ is a bounded operator in $C[0, T]$ and*

$$\rho(\mathcal{K}) = \rho(M_B^{-1}N_B)\,. \tag{3.9}$$

Proof. Since $k_c \in C[0, T]$, the theorem follows immediately from Lemma 2.1. □

3.2.2. Convergence on infinite time-intervals. The solution to (3.1), given by (3.2), is obviously bounded if and only if all eigenvalues of $B^{-1}A$ have positive real parts. (We assume the right-hand-side function f to be in L_p.) The following lemma deals with the boundedness of the waveform relaxation operator.

LEMMA 3.2. *If all eigenvalues of $B^{-1}A$ have positive real parts, then the following statements are equivalent:*

(i) *$\mathcal{K}$ is a bounded operator on $L_p(0, \infty)$ with $1 \leq p \leq \infty$.*

(ii) *All eigenvalues of $M_B^{-1} M_A$ have positive real parts.*

Proof. The lemma is a direct consequence of [11, Thm. 1]. □

The converse of the lemma is as follows.

LEMMA 3.3. *Suppose all eigenvalues of $M_B^{-1} M_A$ have positive real parts. If $\rho(\mathcal{K}) < 1$, then all eigenvalues of $B^{-1} A$ must have positive real parts too.*

Proof. By inspection of (3.4) and (3.5) we can conclude that, under the assumption of the lemma, $\mathcal{K}$ is bounded in $L_p(0, \infty)$ with $1 \leq p \leq \infty$ and, therefore, $\rho(\mathcal{K}) < \infty$. If $\rho(\mathcal{K}) < 1$, then the waveform relaxation iteration is a convergent successive approximation scheme in L_p. Its fixed point satisfies (3.1) and is therefore given by (3.2). Hence, since the fixed point is in L_p, all eigenvalues of $B^{-1} A$ must have positive real parts. □

THEOREM 3.4. *Assume all eigenvalues of $M_B^{-1} M_A$ have positive real parts, and consider $\mathcal{K}$ as an operator in $L_p(0, \infty)$ with $1 \leq p \leq \infty$. Then,*

$$\rho(\mathcal{K}) = \sup_{\mathrm{Re}(z) \geq 0} \rho(\mathbf{K}(z)) \tag{3.10}$$

$$= \sup_{\xi \in \mathbb{R}} \rho(\mathbf{K}(i\xi)) \,. \tag{3.11}$$

This theorem is a special case of [11, Thm. 2]. Here, we prefer to deduce the theorem from Lemma 2.3, i.e., we give a proof based on the Paley-Wiener theorem.

Proof. Notice that $k_c \in L_1(0, \infty)$ since all eigenvalues of $M_B^{-1} M_A$ have positive real parts. The theorem then follows from Lemma 2.3. □

For a better understanding of the theorem, recall the relation between the Laplace-transforms of successive errors, (3.7). Asymptotically, any "frequency" component of the initial error $\tilde{e}^{(0)}(i\xi)$ converges with the corresponding convergence factor $\rho(\mathbf{K}(i\xi))$. According to (3.11), the spectral radius of the waveform relaxation operator $\rho(\mathcal{K})$ equals the supremum of these factors, taken over all frequencies ξ. That is, the asymptotic convergence behaviour of operator $\mathcal{K}$ is determined by the slowest converging frequency component of the initial error.

Setting $\mathcal{H} = \mathcal{K}$ in Lemma 2.4 yields an analogous result for the L_2-norm.

THEOREM 3.5. *Assume all eigenvalues of $M_B^{-1} M_A$ have positive real parts, and consider $\mathcal{K}$ as an operator in $L_2(0, \infty)$. Denote by $||\,.\,||_2$ the L_2-norm and by $||\,.\,||$ the standard Euclidean vector-norm. Then,*

$$||\mathcal{K}||_2 = \sup_{\mathrm{Re}(z) \geq 0} ||\mathbf{K}(z)|| \tag{3.12}$$

$$= \sup_{\xi \in \mathbb{R}} ||\mathbf{K}(i\xi)|| \,. \tag{3.13}$$

REMARK 3.2. If not all eigenvalues of $M_B^{-1} M_A$ have positive real parts, we can still derive a similar result by using an exponential scaling in the norm. Assuming that for all eigenvalues μ_i of $M_B^{-1} M_A$ it holds that $\mathrm{Re}(\mu_i) + \alpha > 0$, we consider the norm

$$||u||_\alpha = ||e^{-\alpha t} u|| \,, \tag{3.14}$$

where the norm in the right-hand side is a standard mean p-norm. With this change of norm, both Theorems 3.4 and 3.5 apply with the supremum taken over $\mathrm{Re}(z) \geq \alpha$ or, after application of the maximum principle, over the line $z = \alpha + i\xi$.

3.3. Jacobi and Gauss-Seidel waveform relaxation. In the following, $\mathcal{K}^{JAC}$ and $\mathcal{K}^{GS}$ denote the Jacobi and Gauss-Seidel waveform relaxation operators, respectively. Their dynamic iteration matrices are given by

$$\mathbf{K}^{\mathbf{JAC}}(z) = (zD_B + D_A)^{-1}(z(L_B + U_B) + (L_A + U_A)) ,$$
$$\mathbf{K}^{\mathbf{GS}}(z) = (z(D_B - L_B) + (D_A - L_A))^{-1}(zU_B + U_A) .$$

3.3.1. Results on finite time-intervals.

LEMMA 3.6. *Consider $\mathcal{K}^{GS}$ as an operator in $C[0, T]$, and suppose B is Hermitian with positive diagonal elements. Then the Gauss–Seidel waveform relaxation method converges, i.e., $\rho(\mathcal{K}^{GS}) < 1$ if and only if B is positive definite.*

Proof. From Theorem 3.1, we derive $\rho(\mathcal{K}^{GS}) = \rho((D_B - L_B)^{-1}U_B)$. The lemma then follows from [6, p. 71, Cor. 1]. □

COROLLARY 3.7. *For a system of ODEs (3.1), derived from a parabolic PDE by finite element discretisation, the Gauss–Seidel waveform relaxation method is convergent in $C[0, T]$.*

Proof. If (3.1) is derived from a parabolic PDE by finite element discretisation, matrix B is positive definite symmetric. Since the diagonal elements of a positive definite matrix are positive, we can apply Lemma 3.6 to obtain the result. □

3.3.2. Results on infinite time-intervals.

LEMMA 3.8. *Assume all diagonal elements of $D_B^{-1}D_A$ have positive real parts. Let A and B be such that $(zB + A)$ is a consistently ordered matrix for* $\mathrm{Re}(z) \geq 0$. *Then, in $L_p(0, \infty)$ with $1 \leq p \leq \infty$,*

$$\rho(\mathcal{K}^{GS}) = \rho(\mathcal{K}^{JAC})^2 . \tag{3.15}$$

Proof. Observe that, for $\mathrm{Re}(z) \geq 0$,

$$\begin{aligned}\det(\lambda I - \mathbf{K}^{\mathbf{GS}}(z)) &= \det(z(D_B - L_B) + (D_A - L_A))^{-1} \\ &\quad \cdot \det(\lambda(z(D_B - L_B) + (D_A - L_A)) - (zU_B + U_A)) \\ &= \frac{\det(\lambda(z(D_B - L_B) + (D_A - L_A)) - (zU_B + U_A))}{\mathrm{trace}(zD_B + D_A)} . \end{aligned} \tag{3.16}$$

Introducing the shorthands $D^* = zD_B + D_A$, $L^* = zL_B + L_A$, and $U^* = zU_B + U_A$, the numerator of (3.16) becomes

$$\begin{aligned}\det(\lambda D^* - \lambda L^* - U^*) &= \det\left(\lambda D^* - \sqrt{\lambda}\left(\sqrt{\lambda}L^* + \frac{1}{\sqrt{\lambda}}U^*\right)\right) \\ &= (\sqrt{\lambda})^d \det\left(\sqrt{\lambda}D^* - \left(\sqrt{\lambda}L^* + \frac{1}{\sqrt{\lambda}}U^*\right)\right) .\end{aligned} \tag{3.17}$$

Since $zB + A$, $\mathrm{Re}(z) \geq 0$, is a consistently ordered matrix, we can use [23, p. 147, Thm. 3.3] to rewrite (3.17) as

$$(\sqrt{\lambda})^d \det\left(\sqrt{\lambda}D^* - (L^* + U^*)\right) .$$

Hence, λ is an eigenvalue of $\mathbf{K}^{\mathbf{GS}}(z)$ if and only if $\lambda = 0$ or $\sqrt{\lambda}$ is an eigenvalue of $\mathbf{K}^{\mathbf{JAC}}(z)$. The latter means that

$$\rho(\mathbf{K}^{\mathbf{GS}}(z)) = \rho(\mathbf{K}^{\mathbf{JAC}}(z))^2 ,$$

which implies (3.15) by application of Theorem 3.4. □

REMARK 3.3. The first assumption of Lemma 3.8 can be loosened if (3.1) is derived by spatial finite element discretisation from a parabolic PDE. The positive definiteness of B implies that it is sufficient to assume that all diagonal elements of A are positive. The assumption can be dropped completely if the bilinear form $a(.,.)$ is H-elliptic, with H the Sobolev space used in the finite element discretisation of the PDE [3, p. 24].

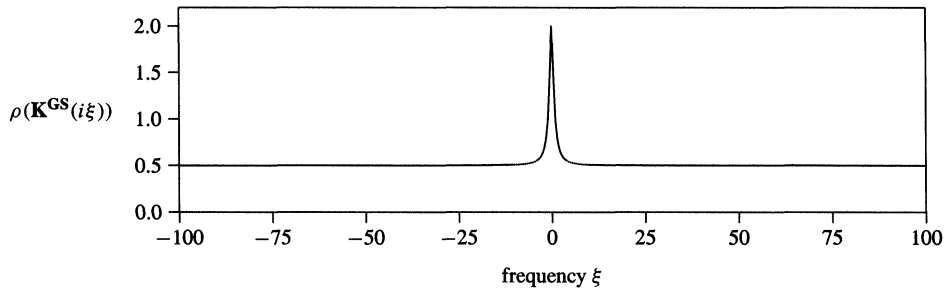

FIG. 3.1. *Spectral radius of the dynamic iteration matrix in function of the frequency* ξ.

3.4. Finite time-interval analysis versus infinite time-interval analysis. In Remark 2.1 we note that the spectral radius of $\mathcal{K}$ as an operator on finite time-intervals is smaller than the spectral radius of $\mathcal{K}$ as an operator on the infinite time-interval. Therefore, it is possible that the waveform relaxation method is convergent on any finite time-interval, but divergent on the infinite time-interval. In a situation like that, computations on a sufficiently long time-interval will at first seem to diverge. Eventually, however, the computations must start to converge. This effect is illustrated in the following example.

Consider the linear initial value problem

$$\begin{pmatrix} 1 & \frac{1}{2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \dot{u_1} \\ \dot{u_2} \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \cos(t) \\ 0 \end{pmatrix},$$

with the initial conditions $u_1(0) = 1$ and $u_2(0) = 0$. The solutions are $u_1(t) = \cos(t)$ and $u_2(t) = \sin(t)$. We solve this problem by using the Gauss-Seidel waveform relaxation method, i.e., we apply the following splittings:

$$M_B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad N_B = \begin{pmatrix} 0 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}, \quad M_A = \begin{pmatrix} \frac{1}{2} & 0 \\ -1 & 1 \end{pmatrix}, \quad N_A = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}.$$

The dynamic iteration matrix is given by

$$\mathbf{K}^{\mathbf{GS}}(z) = \begin{pmatrix} z + \frac{1}{2} & 0 \\ z - 1 & z + 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & -\frac{1}{2}z - 1 \\ 0 & 0 \end{pmatrix}.$$

We have plotted in Fig. 3.1 the spectral radius of the dynamic iteration matrix evaluated along the imaginary axis, i.e., for $z = i\xi$, in function of the frequency ξ.

On finite time-intervals, we can apply the result of Theorem 3.1, which yields

$$\rho(\mathcal{K}^{GS}) = \rho(\mathbf{K}^{\mathbf{GS}}(\infty)) = \frac{1}{2}. \tag{3.18}$$

This result ensures linear convergence on any time-interval of finite length. The spectral radius on infinite time-intervals can be calculated with the aid of (3.11),

$$\rho(\mathcal{K}^{GS}) = \sup_{\xi \in \mathbb{R}} \rho(\mathbf{K}^{\mathbf{GS}}(i\xi)) = 2. \tag{3.19}$$

By that, waveform relaxation is divergent on the interval $(0, \infty)$.

To clarify these results, we have plotted in Fig. 3.2 the error of the second component iterate $e_2^{(\nu)} = u_2^{(\nu)} - u_2$ after 0, 4, 8, and 12 iterations. Roughly speaking, one observes two subintervals with different convergence characteristics, corresponding to the results (3.18)

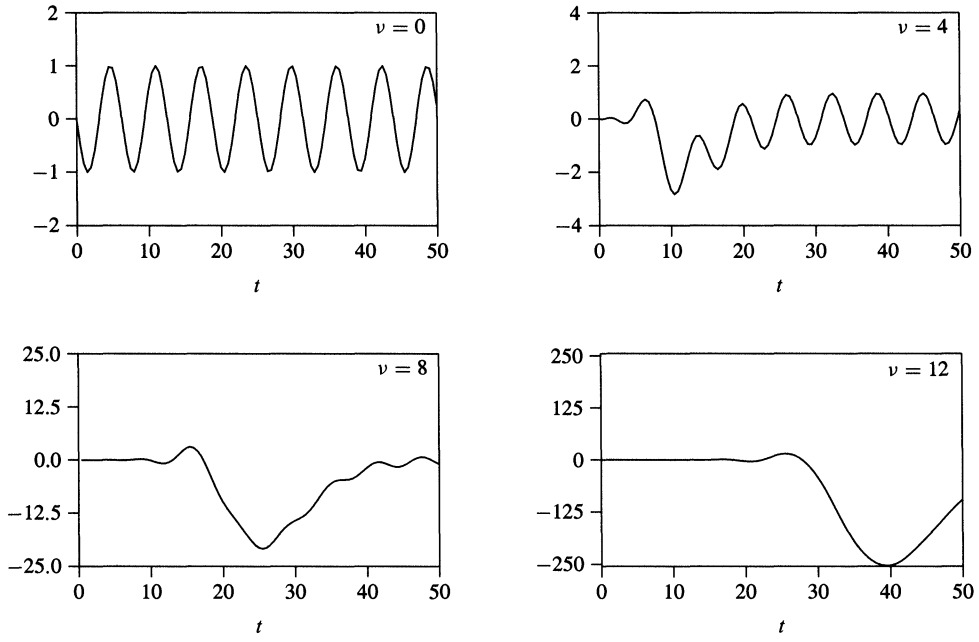

FIG. 3.2. *Error* $e_2^{(\nu)}(t)$ *for* $\nu = 0, 4, 8,$ *and* 12.

and (3.19), respectively. The finite-length interval with small errors extends as more iterations are applied. Consequently, the region of divergent behaviour recedes backward after a large number of iterations. Hence, asymptotically, the convergence behaviour will be dictated by the finite time-interval analysis.

4. The multigrid waveform relaxation method.

4.1. The two-grid waveform relaxation operator. Multigrid is known to be a very efficient solver for elliptic partial differential equations; see, e.g., [5] and [22]. The multigrid principle can be easily extended to time-dependent problems by choosing all the operations in the multigrid cycle as operations on functions [10]. A two-grid cycle for the initial value problem (3.1), derived by finite element discretisation from a parabolic PDE (1.1), is stated below. It is defined on two nested grids Ω_H and Ω_h, with $\Omega_H \subset \Omega_h$, and determines a new iterate $u^{(\nu)}$ from the former waveform $u^{(\nu-1)}$ in three steps: pre-smoothing, coarse grid correction, and post-smoothing. In the following, the subscripts h and H are used to denote fine and coarse grid quantities, respectively.

(i) Pre-smoothing. Set $x^{(0)} = u^{(\nu-1)}$, and perform ν_1 waveform relaxation steps. For $\nu = 1, 2, \ldots, \nu_1$, solve

$$M_{B_h}\dot{x}^{(\nu)} + M_{A_h}x^{(\nu)} = N_{B_h}\dot{x}^{(\nu-1)} + N_{A_h}x^{(\nu-1)} + f_h\,, \tag{4.1}$$

with $x^{(\nu)}(0) = u_0, t > 0$.

(ii) Coarse grid correction. Compute the defect

$$\begin{aligned} d_h &= B_h\dot{x}^{(\nu_1)} + A_h x^{(\nu_1)} - f_h \\ &= N_{B_h}(\dot{x}^{(\nu_1-1)} - \dot{x}^{(\nu_1)}) + N_{A_h}(x^{(\nu_1-1)} - x^{(\nu_1)})\,. \end{aligned}$$

Solve the coarse grid equivalent of the defect equation,

$$B_H\dot{v}_H + A_H v_H = r d_h\,, \quad \text{with } v_H(0) = 0\,, \quad t > 0\,, \tag{4.2}$$

with $r : \Omega_h \to \Omega_H$ the restriction operator. Then interpolate the correction v_H to Ω_h and correct the current approximation

$$\bar{x} = x^{(\nu_1)} - pv_H ,$$

with $p : \Omega_H \to \Omega_h$ the prolongation operator.

(iii) Post-smoothing. Perform ν_2 iterations of type (4.1), starting with $x^{(0)} = \bar{x}$, and set $u^{(\nu)} = x^{(\nu_2)}$.

Since (4.2) is formally equal to (3.1), this two-grid cycle can be applied in a recursive way to obtain a multigrid cycle.

The two-grid cycle can be written as an explicit successive approximation scheme: $u^{(\nu)} = \mathcal{M}u^{(\nu-1)} + \varphi$. The *two-grid waveform relaxation operator* $\mathcal{M}$ is given by

$$\mathcal{M}u(t) = \mathcal{K}^{\nu_2}\,\mathcal{C}\,\mathcal{K}^{\nu_1}\,u(t) , \tag{4.3}$$

with $\mathcal{K}$ the standard waveform relaxation operator (3.4), and $\mathcal{C}$ the two-grid correction waveform operator

$$\mathcal{C}u(t) = (I - pB_H^{-1}rB_h)u(t) + \mathcal{C}_c u(t) .$$

The operator $\mathcal{C}_c$ is also of linear Volterra convolution type. Its matrix-valued kernel equals $c_c(t) = pe^{-B_H^{-1}A_H t}B_H^{-1}(A_H B_H^{-1}rB_h - rA_h)$. By rearranging terms in (4.3), we may rewrite $\mathcal{M}$ in the general form (2.2):

$$\mathcal{M}u(t) = (M_{B_h}^{-1}N_{B_h})^{\nu_2}(I - pB_H^{-1}rB_h)(M_{B_h}^{-1}N_{B_h})^{\nu_1}u(t) + \mathcal{M}_c u(t) .$$

Operator $\mathcal{M}_c$ is a linear combination of products of linear Volterra convolution operators $\mathcal{K}_c$ and $\mathcal{C}_c$. Therefore, it is itself of linear Volterra convolution type. We shall denote its kernel by $m_c(t)$ and the Laplace-transform of $m_c(t)$ by $\mathbf{M_c}(z)$. The precise expressions for $m_c(t)$ and $\mathbf{M_c}(z)$ are rather complicated and, since they are not required further on, omitted.

Let $e^{(\nu)}$ be the error of the νth two-grid waveform iterate, i.e., $e^{(\nu)} = u^{(\nu)} - u$. It satisfies $e^{(\nu)} = \mathcal{M}e^{(\nu-1)}$. Laplace-transforming this relation yields

$$\begin{aligned}\tilde{e}^{(\nu)}(z) &= \left((M_{B_h}^{-1}N_{B_h})^{\nu_2}(I - pB_H^{-1}rB_h)(M_{B_h}^{-1}N_{B_h})^{\nu_1} + \mathbf{M_c}(z)\right)\tilde{e}^{(\nu-1)}(z) \\ &= \mathbf{M}(z)\tilde{e}^{(\nu-1)}(z) .\end{aligned}$$

By Laplace-transforming the equations of the two-grid cycle, we find the following equivalent expression for the two-grid dynamic iteration matrix $\mathbf{M}(z)$:

$$\mathbf{M}(z) = \mathbf{K}^{\nu_2}(z)(I - p(zB_H + A_H)^{-1}r(zB_h + A_h))\mathbf{K}^{\nu_1}(z) , \tag{4.4}$$

$$\mathbf{K}(z) = (zM_{B_h} + M_{A_h})^{-1}(zN_{B_h} + N_{A_h}) . \tag{4.5}$$

REMARK 4.1. In the case of a Gauss-Seidel (or Jacobi) splitting of A_h and B_h, $\mathbf{K}(z)$ and $\mathbf{M}(z)$ are, respectively, the Gauss-Seidel (or Jacobi) iteration matrix and the two-grid iteration matrix for the linear system constructed by finite element discretisation of the elliptic problem $z\mathbf{u} + \mathcal{L}\mathbf{u} = \mathbf{f}$.

4.2. Convergence analysis.

4.2.1. Convergence on finite time-intervals. The spectral radius of the two-grid waveform relaxation operator in the case of a finite difference discretisation is known to be zero on finite time-intervals [19, Thm. 3.4.1]. In the following theorem, which is the multigrid

waveform analogue of Theorem 3.1, we will state the equivalent formula for general non-singular B.

THEOREM 4.1. *The two-grid waveform relaxation operator $\mathcal{M}$ is a bounded operator in $C[0, T]$ and*

$$\rho(\mathcal{M}) = \rho\left((M_{B_h}^{-1}N_{B_h})^{\nu_2}(I - pB_H^{-1}rB_h)(M_{B_h}^{-1}N_{B_h})^{\nu_1}\right) . \tag{4.6}$$

Proof. Since both k_c and c_c are continuous on $[0, T]$, we have $m_c \in C[0, T]$. Consequently, the theorem follows from Lemma 2.1. □

REMARK 4.2. The spectral radius of $\mathcal{M}$ equals the spectral radius of the standard two-grid operator for the trivial elliptic problem $\mathcal{L}\mathbf{u} = \mathbf{f}$, where $\mathcal{L}$ is the identity operator, discretised on a finite element mesh.

4.2.2. Convergence on infinite time-intervals. In [10, pp. 219-220], Lubich and Ostermann examined the multigrid waveform relaxation method for the finite difference case, i.e., for systems of ODEs (3.1) with $B = I$. We shall extend their results to initial value problems (3.1), derived from (1.1) by finite element discretisation.

We assume that

$$\text{all eigenvalues of } B_H^{-1}A_H \text{ and } M_{B_h}^{-1}M_{A_h} \text{ have positive real parts.} \tag{4.7}$$

REMARK 4.3. Notice that condition (4.7) is satisfied if we assume the boundedness of the analytical solution of (3.1) on Ω_H and the boundedness of the standard waveform relaxation operator $\mathcal{K}$.

THEOREM 4.2. *Assume* (4.7), *and consider $\mathcal{M}$ as an operator in $L_p(0, \infty)$ with $1 \le p \le \infty$. Then, $\mathcal{M}$ is a bounded operator with spectral radius*

$$\rho(\mathcal{M}) = \sup_{\mathrm{Re}(z)\ge 0} \rho(\mathbf{M}(z)) \tag{4.8}$$

$$= \sup_{\xi\in\mathbb{R}} \rho(\mathbf{M}(i\xi)) . \tag{4.9}$$

Proof. It is easily verified that

$$\mathbf{M_c}(z) = \mathbf{M}(z) - \lim_{z\to\infty} \mathbf{M}(z) .$$

From (4.7), the entries of $\mathbf{M_c}(z)$ are rational functions of z vanishing at infinity, all of whose poles have negative real part. This implies that $m_c \in L_1(0, \infty)$ by an inverse Laplace-transformation argument. The theorem then follows from Lemma 2.3. □

Applying Lemma 2.4 yields the following result for the L_2-norm of $\mathcal{M}$.

THEOREM 4.3. *Assume* (4.7), *and consider $\mathcal{M}$ as an operator in $L_2(0, \infty)$. Denote by $||\,.\,||_2$ the L_2-norm and by $||\,.\,||$ the standard Euclidean vector-norm. Then,*

$$||\mathcal{M}||_2 = \sup_{\mathrm{Re}(z)\ge 0} ||\mathbf{M}(z)|| \tag{4.10}$$

$$= \sup_{\xi\in\mathbb{R}} ||\mathbf{M}(i\xi)|| . \tag{4.11}$$

REMARK 4.4. Suppose that (4.7) is not satisfied, but $\mathrm{Re}(\mu_i) + \alpha > 0$ and $\mathrm{Re}(\nu_j) + \alpha > 0$ for all eigenvalues μ_i and ν_j of $M_{B_h}^{-1}M_{A_h}$ and $B_H^{-1}A_H$, respectively. Using the exponentially scaled norm (3.14), Theorems 4.2 and 4.3 hold when the supremum is taken over $\mathrm{Re}(z) \ge \alpha$ or over the line $z = \alpha + i\xi$.

5. Model problem analysis and numerical results. In this section, we conclude with some experimental results for two model problems.

First, we consider the one-dimensional heat equation

$$\frac{\partial \mathbf{u}}{\partial t} - \Delta_1 \mathbf{u} = 0\,, \quad x \in [0,1]\,, \quad t \in [0,1]\,, \tag{5.1}$$

with homogeneous Dirichlet boundary conditions and an initial condition chosen such that the analytical solution is given by $\mathbf{u}(x,t) = \sin(\pi x)\exp(-\pi^2 t)$. The equation is discretised using linear, quadratic, or cubic basis functions on a discrete grid with mesh size h, i.e., $\Omega_h = \{x_i = ih \mid 0 \le i \le 1/h\}$.

Next, the convergence of the (multigrid) waveform relaxation method is investigated for the two-dimensional heat equation

$$\frac{\partial \mathbf{u}}{\partial t} - \Delta_2 \mathbf{u} = 0\,, \quad (x,y) \in [0,1]\times[0,1]\,, \quad t \in [0,1]\,, \tag{5.2}$$

completed with Dirichlet boundary conditions and an initial condition. The analytical solution is given by $\mathbf{u}(x,y,t) = 1 + \sin(\pi x/2)\sin(\pi y/2)\exp(-\pi^2 t/2)$. The equation is discretised using linear basis functions (triangular finite elements) and bilinear basis functions (rectangular finite elements) on a discrete mesh with equal mesh size h in the x-direction and the y-direction:

$$\Omega_h = \{(x_i = ih,\, y_j = jh) \mid 0 \le i,\, j \le 1/h\}\,.$$

The finite element discretisation of both (5.1) and (5.2) leads to a system of ODEs of the form (1.2). For all subsequent choices of the basis functions, all the eigenvalues of $B^{-1}A$ and $M_B^{-1}M_A$ (for Jacobi and Gauss-Seidel splittings) have positive real parts, i.e., the conditions for applicability of Theorems 3.4 and 4.2 are satisfied.

5.1. A one-dimensional model problem.

5.1.1. Theoretical results. To determine the spectral radius of $\mathcal{K}^{JAC}$ and $\mathcal{K}^{GS}$ on infinite time-intervals, the spectral radii of $\mathbf{K}^{\mathbf{JAC}}(z)$ and $\mathbf{K}^{\mathbf{GS}}(z)$ are to be calculated for every value of z along the imaginary axis. This is generally a very difficult task. However, for our model problem (5.1), discretised using linear basis functions, we have the following result.

LEMMA 5.1. *The Jacobi and Gauss–Seidel waveform relaxation operators $\mathcal{K}^{JAC}$ and $\mathcal{K}^{GS}$, respectively, for the one-dimensional heat equation, studied in $L_p(0,\infty)$ with $1 \le p \le \infty$, discretised using linear basis functions, satisfy the following formulae, valid for small h,*

$$\rho(\mathcal{K}^{JAC}) \approx 1 - \pi^2 h^2/2 \;\; \textit{and} \;\; \rho(\mathcal{K}^{GS}) \approx 1 - \pi^2 h^2\,. \tag{5.3}$$

Proof. Discretising (5.1) with linear basis functions yields

$$\rho(\mathbf{K}^{\mathbf{JAC}}(z)) = \left|\frac{-2zh^2 + 12}{4zh^2 + 12}\right| \cos(\pi h)\,.$$

As a consequence of Theorem 3.4 and equation (3.11), we find that

$$\begin{aligned}\rho(\mathcal{K}^{JAC}) = \sup_{\xi\in\mathbb{R}} \left|\frac{-2i\xi h^2 + 12}{4i\xi h^2 + 12}\right| \cos(\pi h) \;&=\; \cos(\pi h)\\ &\approx 1 - \pi^2 h^2/2\,.\end{aligned}$$

Since the assumptions of Lemma 3.8 are satisfied, the second formula of (5.3) follows immediately from the first one by application of (3.15). □

Next, we consider the multigrid waveform relaxation method. We assume that the coarse grid Ω_H is derived from the fine grid Ω_h by standard coarsening ($H = 2h$). For the prolongation operator $p : \Omega_H \to \Omega_h$, we use the piecewise linear interpolation operator; see, e.g., [5, p. 22]. The restriction operator $r : \Omega_h \to \Omega_H$ is defined by the transpose of the prolongation operator: $r = p^t$ [22, pp. 70–71].

LEMMA 5.2. *The two-grid operator $\mathcal{M}$ for the one-dimensional heat equation, studied in $L_p(0, \infty)$ with $1 \le p \le \infty$, discretised using linear basis functions, with red/black Gauss–Seidel smoothing and with the prolongation and restriction operator defined as above, satisfies*

$$\rho(\mathcal{M}) \le \sqrt{3}\sqrt{\eta_0(2\nu - 1)}\,, \quad \text{with } \eta_0(\nu) = \frac{\nu^\nu}{(\nu+1)^{\nu+1}}\,, \tag{5.4}$$

for $\nu = \nu_1 + \nu_2 \ge 1$.

Proof. The proof is a generalisation of the model problem analysis in [5, p. 25] and the proof of [10, Prop. 5]. Writing

$$\mathbf{M}(z) = \mathbf{K}^{(\nu_2)}(z)\mathbf{C}(z)\mathbf{K}^{(\nu_1)}(z)\,, \quad \text{with } \mathbf{C}(z) = I - p(zB_H + A_H)^{-1}r(zB_h + A_h)\,,$$

we have to study

$$\rho(\mathbf{M}(z)) = \rho\left(\mathbf{C}(z)\mathbf{K}^{(\nu)}(z)\right)\,, \quad \nu = \nu_1 + \nu_2\,.$$

Let e_m ($m = 1, 2, \ldots, N-1$), with $N = 1/h$, denote the eigenvectors of B_h and A_h,

$$e_m(x) = \sqrt{2h}\sin(m\pi x)\,, \quad x = kh\,, \quad 1 \le k \le N-1\,.$$

Both $\mathbf{C}(z)$ and $\mathbf{K}^{(\nu)}(z)$ leave the subspace spanned by (e_m, e_{N-m}) invariant. Their restrictions with respect to this basis have the matrix representations

$$\mathbf{C_m}(z) = \begin{bmatrix} s_m^2 & c_m^2 \\ s_m^2 & c_m^2 \end{bmatrix} + \frac{3zh^2}{zh^2(1+2(c_m^2 - s_m^2)^2) + 12c_m^2 s_m^2} \begin{bmatrix} c_m^2 s_m^2(1-2c_m^2) & -c_m^4(1-2s_m^2) \\ -s_m^4(1-2c_m^2) & c_m^2 s_m^2(1-2s_m^2) \end{bmatrix},$$

$$\mathbf{K_m}^{(\nu)}(z) = (c_m^2 - s_m^2)^{2\nu-1}\left(\frac{1 - zh^2/6}{1 + zh^2/3}\right)^{2\nu} \cdot \left(\begin{bmatrix} c_m^2 & c_m^2 \\ -s_m^2 & -s_m^2 \end{bmatrix} + \frac{3zh^2/6}{2(1-zh^2/6)}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}\right),$$

with $c_m = \cos(mh\pi/2)$, $s_m = \sin(mh\pi/2)$, and $m = 1, \ldots, N/2 - 1$. (We may omit the degenerate case $m = N/2$.) We then have, for $\nu \ge 1$ and $Z = zh^2$,

$$\mathbf{M_m}(Z) = \mathbf{C_m}(Z)\mathbf{K_m}^{(\nu)}(Z) = (c_m^2 - s_m^2)^{2\nu-1}\frac{(1-Z/6)^{2\nu-1}}{(1+Z/3)^{2\nu}}\frac{Z}{4}\left((s_m^2 - c_m^2)\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{3Z(1-4c_m^2 s_m^2)}{Z(1+2(c_m^2-s_m^2)^2) + 12c_m^2 s_m^2}\begin{bmatrix} c_m^2 & c_m^2 \\ -s_m^2 & -s_m^2 \end{bmatrix}\right).$$

The spectral radius of $\mathbf{M_m}(Z)$ equals

$$\rho(\mathbf{M_m}(Z)) = \left|(c_m^2 - s_m^2)^{2\nu}\frac{(1-Z/6)^{2\nu-1}}{(1+Z/3)^{2\nu}}\frac{Z}{4}\left(2 - \frac{3Z(1-4c_m^2 s_m^2)}{Z(1+2(c_m^2-s_m^2)^2) + 12c_m^2 s_m^2}\right)\right|.$$

TABLE 5.1
Values of the upper bound for $\rho(\mathcal{M})$ as a function of $\nu = \nu_1 + \nu_2$.

ν	1	2	3	4
$\sqrt{3}\sqrt{\eta_0(2\nu-1)}$	0.866	0.563	0.448	0.384

TABLE 5.2
Numerical values of $\rho(\mathcal{M})$, $\nu = 2$.

h	1/8	1/16	1/32	1/64
$\rho_{\text{num}}(\mathcal{M})$	0.217	0.263	0.276	0.280

Thus,

$$\rho(\mathcal{M}) = \sup_{\text{Re}(Z)\geq 0} \rho(\mathbf{M}(Z)) = \sup_{\text{Re}(Z)\geq 0} \max_{m} \rho(\mathbf{M}_{\mathbf{m}}(Z)) \tag{5.5}$$
$$\leq \sup_{\xi\in\mathbb{R}} \frac{3}{2^{2\nu}} \left| \frac{(6-i\xi)^{2\nu-1} i\xi}{(3+i\xi)^{2\nu}} \right|$$
$$= \sqrt{3}\sqrt{\eta_0(2\nu-1)}\,,$$

where the latter supremum is attained for $\xi = \pm 3\sqrt{2}/\sqrt{3\nu-2}$. □

The lemma states that the spectral radius of the two-grid operator $\mathcal{M}$ can be bounded by a constant, independent of h. Some values of the bound are given in Table 5.1.

Since the bound in Lemma 5.2 is not optimal, we numerically computed the spectral radius of the two-grid operator by evaluation of (5.5), for $\nu = 2$ and for several values of h. These results are reported in Table 5.2.

5.1.2. Numerical results. We discretised problem (5.1) using linear basis functions on a spatial finite element mesh with mesh size h = 1/8, 1/16, 1/32, and 1/64. The resulting system was solved using Gauss-Seidel waveform relaxation and multigrid waveform relaxation. In the latter method we applied standard V- and W-cycles, with one pre-smoothing and one post-smoothing step of red/black Gauss-Seidel waveform relaxation type, standard coarsening down to a coarse grid with mesh size $h = 1/2$, linear interpolation, and the corresponding restriction. The effect of time-discretisation on the convergence properties will be a subject of further studies. Here, the Crank-Nicolson method is used for time-discretisation of the iteration schemes. To approximate the continuous-time convergence results, we take a small time-step, e.g., 1/1000.

The νth iteration convergence factor is determined by calculating the l_2-norm of the discrete defect of the approximation and by dividing the result for successive iterates. After a sufficiently large number of iterations, this factor takes a more or less constant value. The *averaged convergence factor* is then defined as the geometric average of these iteration convergence factors over the region of nearly constant behaviour.

In Table 5.3, we have reported the observed averaged convergence factors for the one-dimensional heat equation, spatially discretised with linear elements. Even though the time-interval in this experiment is finite, the measured waveform relaxation convergence factors closely match the ones that can be obtained by evaluation of the infinite interval theoretical formula (5.3). For a discussion of this phenomenon, we refer to [19, §3.2.4, §3.5]. The multigrid convergence factors are clearly bounded above by a constant less than 1, independent of h. We have pictured successive iterates, $u^{(\nu)}$, evaluated at $x = 1/2$, in Fig. 5.1. For a mesh size $h = 1/32$, the Gauss-Seidel method is very slowly converging. One iteration of

TABLE 5.3
Averaged convergence factors for the one-dimensional heat equation, linear basis functions.

h	1/8	1/16	1/32	1/64
Gauss-Seidel	0.841	0.959	0.990	0.997
V-cycle	0.229	0.300	0.326	0.331
W-cycle	0.210	0.254	0.265	0.267

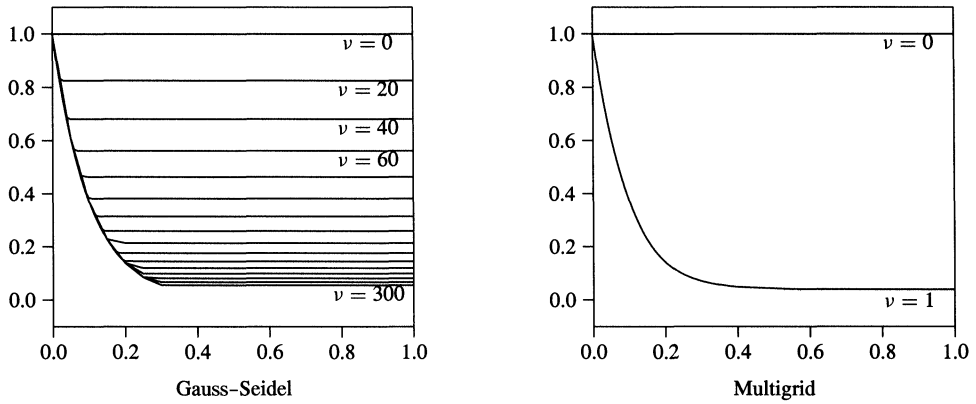

FIG. 5.1. *Successive iterates* $u^{(\nu)}(1/2, t)$, *obtained with (multigrid) waveform relaxation.*

TABLE 5.4
Averaged convergence factors for the one-dimensional heat equation, quadratic basis functions.

h	1/8	1/16	1/32	1/64
Gauss-Seidel	0.967	0.992	0.998	0.999
V-cycle	0.202	0.285	0.316	0.325
W-cycle	0.201	0.276	0.301	0.309

TABLE 5.5
Averaged convergence factors for the one-dimensional heat equation, cubic basis functions.

h	1/8	1/16	1/32	1/64
Gauss-Seidel	0.862	0.965	0.991	0.998
V-cycle	0.191	0.234	0.236	0.237
W-cycle	0.184	0.218	0.213	0.210

the multigrid method suffices, to get an approximation that can no longer be distinguished graphically from subsequent iterates.

In Tables 5.4 and 5.5, we report the averaged convergence factors for the one-dimensional heat equation, discretised using quadratic and cubic basis functions, respectively. We consider both Gauss-Seidel waveform relaxation and multigrid waveform relaxation. In the quadratic case, we use three-colour Gauss-Seidel waveform relaxation as a smoother: two colours are needed to decouple the unknowns corresponding to integer nodes, while the third colour is used to update the unknowns at half-integer nodes. In the cubic case, the even unknowns (and their derivatives) are decoupled from the odd unknowns (and their derivatives) using red/black Gauss-Seidel waveform relaxation smoothing. The other multigrid assumptions are identical to those of the linear case. Notice that the Gauss-Seidel waveform relaxation convergence factors seem to satisfy a relation of the form $\rho(\mathcal{K}^{GS}) \approx 1 - O(h^2)$, although no explicit

TABLE 5.6
Averaged convergence factors for the two-dimensional heat equation, linear basis functions.

h	1/4	1/8	1/16	1/32
Gauss-Seidel	0.478	0.845	0.960	0.990
V-cycle	0.135	0.335	0.437	0.470
W-cycle	0.135	0.304	0.357	0.371

TABLE 5.7
Averaged convergence factors for the two-dimensional heat equation, bilinear basis functions.

h	1/4	1/8	1/16	1/32
Gauss-Seidel	0.356	0.780	0.941	0.985
V-cycle	0.137	0.299	0.353	0.365
W-cycle	0.137	0.294	0.344	0.355

theoretical formulae were found. The multigrid waveform relaxation convergence factors are obviously bounded by a constant, independent of h.

5.2. A two-dimensional problem. Tables 5.6 and 5.7 contain the observed averaged convergence factors, both for Gauss-Seidel waveform relaxation and multigrid waveform relaxation, for the two-dimensional heat equation (5.2). In the latter method the smoother is of four-colour Gauss-Seidel waveform relaxation type. The prolongation operators are, respectively, seven-point prolongation (linear basis functions) and nine-point prolongation (bilinear basis functions) [5, p. 60]. The Gauss-Seidel convergence factors satisfy a relation of the form $1 - O(h^2)$, while the multigrid convergence factors are bounded by a small, h-independent constant.

Appendix. An elementary proof of Lemma 2.1.

Proof. From the proof of Lemma 2.1, we know that $\mathcal{H}$ is a bounded operator, with $||\mathcal{H}|| \leq ||H|| + T\,||h_c||$.

If $H = 0$, we can use a general functional analysis result, which states that the spectrum of a linear Volterra convolution operator with continuous kernel equals the singleton $\{0\}$; see, e.g., [8, p. 33]. Hence, $\rho(\mathcal{H}) = \rho(H) = 0$.

Further on, we assume $H \neq 0$. The n-fold application of $\mathcal{H}$ to u then includes 2^n terms. Each term consists of a combination of matrix multiplication and Volterra convolution operators applied to u. The norm of a term with $n-i$ matrix multiplications and i convolutions can be bounded by

$$||H||^{n-i}||h_c||^i \left(\int_0^t \int_0^{s_1} \int_0^{s_2} \cdots \int_0^{s_{i-1}} ds_i \ldots ds_3 ds_2 ds_1 \right) ||u|| \leq ||H||^{n-i}||h_c||^i \frac{T^i}{i!}||u|| \,.$$

Note that for each i there are $\binom{n}{i}$ terms satisfying the above bound. We get

$$||\mathcal{H}^n u|| \leq ||H||^n \left(\sum_{i=0}^{n} \binom{n}{i} \frac{c^i}{i!} \right) ||u|| \,,$$

with $c = (||h_c||\,T)/||H||$. Using property (2.3) of the spectral radius, we obtain

$$\rho(\mathcal{H}) \leq ||H|| \lim_{n\to\infty} \sqrt[n]{\sum_{i=0}^{n} \binom{n}{i} \frac{c^i}{i!}} \,.$$

To calculate the limit, we observe that [1, Eq. 22.3.9]

$$\sum_{i=0}^{n} \binom{n}{i} \frac{c^i}{i!} = L_n^{(0)}(-c) ,$$

where $L_n^{(0)}(x)$ denotes the nth Laguerre polynomial. Using Perron's asymptotic formula for Laguerre polynomials [17, Thm. 8.22.3], we obtain

$$\sqrt[n]{L_n^{(0)}(-c)} = \sqrt[n]{\frac{1}{2}\pi^{-1/2}e^{-c/2}c^{-1/4}n^{-1/4}e^{2\sqrt{nc}}}\ \sqrt[n]{1 + O(n^{-1/2})} .$$

Taking the limit $n \to \infty$, both factors tend to 1, and, by consequence, $\rho(\mathcal{H}) \leq ||H||$.

To prove that $\rho(\mathcal{H})$ is independent from the choice of the $\mathbb{C}^d$ vector-norm, we use two different $\mathbb{C}^d$ vector-norms $||\,.\,||_1$ and $||\,.\,||_2$, and their associated maximum-norms $||\,.\,||_{T1}$ and $||\,.\,||_{T2}$. Since all $\mathbb{C}^d$ vector-norms are equivalent [6, p. 7, Thm. 2], there exist $m, M > 0$ such that $m\,||u||_{T1} \leq ||u||_{T2} \leq M\,||u||_{T1}$. Hence, $\lambda - \mathcal{H}$ has a bounded inverse with regard to the maximum-norm $||\,.\,||_{T1}$ if and only if $(\lambda - \mathcal{H})^{-1}$ is bounded with regard to $||\,.\,||_{T2}$, and the independency follows by definition of $\rho(\mathcal{H})$. Consequently, $\rho(\mathcal{H}) \leq ||H||$ for every induced matrix-norm, or

$$\rho(\mathcal{H}) \leq \inf_{\{||\,.\,||\}} ||H|| = \rho(H) , \tag{A.1}$$

where the infimum is taken over all matrix-norms induced by a $\mathbb{C}^d$ vector-norm. The equality of (A.1) follows from a well-known characterisation of the spectral radius of a matrix [6, p. 14].

Finally, suppose $\lambda \neq 0$, and

$$\lambda u - \mathcal{H}u = f .$$

Evaluating the above equation for $t = 0$ gives

$$(\lambda I - H)u(0) = f(0) . \tag{A.2}$$

If $\det(\lambda I - H) = 0$, then (A.2) has either no solutions or an infinite number of solutions. The eigenvalues of H are therefore not regular values of $\mathcal{H}$ and, consequently, $\rho(\mathcal{H}) \geq \rho(H)$, which completes the proof. □

Acknowledgments. The authors would like to thank Prof. Walter Van Assche for his valuable comments on the use of Laguerre polynomials in the proof of Lemma 2.1 (Appendix) and an anonymous referee for suggesting the proof of Lemma 2.1 based on perturbation theory.

REFERENCES

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York, 1970.

[2] P. Brenner, V. Thomée, and L. B. Wahlbin, *Besov Spaces and Applications to Difference Methods for Initial Value Problems*, Lecture Notes in Mathematics, Vol. 434, Springer-Verlag, Berlin, 1975.

[3] P. G. Ciarlet, *Basic error estimates for elliptic problems*, in Handbook of Numerical Analysis, Vol. 2, part 1, P. G. Ciarlet and J. L. Lions, eds., Elsevier-North Holland, Amsterdam, 1991, pp. 18-351.

[4] G. Gripenberg, S. O. Londen, and O. Staffans, *Volterra Integral and Functional Equations*, Encyclopedia of Mathematics and Its Applications, Vol. 34, Cambridge University Press, Cambridge, UK, 1990.

[5] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer Series in Computational Mathematics, Vol. 4, Springer-Verlag, Berlin, 1985.

[6] E. ISAACSON AND H. B. KELLER, *Analysis of Numerical Methods*, John Wiley and Sons, New York, 1966.
[7] T. KATO, *Perturbation Theory for Linear Operators*, 2nd ed., Springer-Verlag, Berlin, 1984.
[8] R. KRESS, *Linear Integral Equations*, Applied Mathematical Sciences, Vol. 82, Springer-Verlag, New York, 1989.
[9] E. LELARASMEE, A. E. RUEHLI, AND A. L. SANGIOVANNI-VINCENTELLI, *The waveform relaxation method for time-domain analysis of large scale integrated circuits*, IEEE Trans. CAD IC Sys., 1 (1982), pp. 131-145.
[10] C. LUBICH AND A. OSTERMANN, *Multi-grid dynamic iteration for parabolic equations*, BIT, 27 (1987), pp. 216-234.
[11] U. MIEKKALA, *Dynamic iteration methods applied to linear DAE systems*, J. Comput. Appl. Math., 25 (1989), pp. 133-151.
[12] U. MIEKKALA AND O. NEVANLINNA, *Convergence of dynamic iteration methods for initial value problems*, SIAM J. Sci. Statist. Comput., 8 (1987), pp. 459-482.
[13] O. NEVANLINNA, *Remarks on Picard-Lindelöf iteration, Part* I, BIT, 29 (1989), pp. 328-346.
[14] A. R. NEWTON AND A. L. SANGIOVANNI-VINCENTELLI, *Relaxation-based electrical simulation*, SIAM J. Sci. Statist. Comput., 4 (1983), pp. 485-524.
[15] R. E. A. C. PALEY AND N. WIENER, *Fourier Transforms in the Complex Domain*, American Mathematical Society, Providence, RI, 1934.
[16] S. REED AND B. SIMON, *Fourier Analysis, Self-Adjointness*, Methods of Modern Mathematical Physics, Vol. 2, Academic Press, New York, 1975.
[17] G. SZEGÖ, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, Vol. 23, 3rd ed., American Mathematical Society, Providence, RI, 1967.
[18] V. THOMÉE, *Galerkin Finite Element Methods for Parabolic Problems*, Lecture Notes in Mathematics, Vol. 1054, Springer-Verlag, Berlin, 1984.
[19] S. VANDEWALLE, *Parallel Multigrid Waveform Relaxation for Parabolic Problems*, B.G. Teubner, Stuttgart, 1993.
[20] S. VANDEWALLE AND R. PIESSENS, *Efficient parallel algorithms for solving initial-boundary value and time-periodic parabolic partial differential equations*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 1330-1346.
[21] ———, *On dynamic iteration methods for solving time-periodic differential equations*, SIAM J. Numer. Anal., 30 (1993), pp. 286-303.
[22] P. WESSELING, *An Introduction to Multigrid Methods*, John Wiley and Sons, Chicester, UK, 1992.
[23] D. M. YOUNG, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.