Multigrid Waveform Relaxation on Spatial Finite Element Meshes: The Discrete-Time Case

Jan Janssen
Stefan Vandewalle

CRPC-TR94450
May 1994

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

This work was supported in part by the NSF.

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

JAN JANSSEN* AND STEFAN VANDEWALLE†

Abstract. The efficiency of numerically solving time-dependent partial differential equations on parallel computers can be greatly improved by computing the solution on many time-levels simultaneously. The theoretical properties of one such method, namely the discrete-time multigrid waveform relaxation method, are investigated for systems of ordinary differential equations obtained by spatial finite element discretisation of linear parabolic initial boundary value problems. The results are compared to the corresponding continuous-time results. The theory is illustrated for a one-dimensional and a two-dimensional model problem and checked against results obtained by numerical experiments.

Key words. parabolic partial differential equations, waveform relaxation, multigrid, linear multistep methods

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

1. Introduction. We consider the numerical solution of a linear parabolic initial boundary value problem, spatially discretised by a conforming Galerkin finite element method. This leads to a linear system of ordinary differential equations (ODEs), see e.g. [5, 14],

$$B\dot{u} + Au = f\ ,\quad u(0) = u_0\ ,\quad t > 0\ , \tag{1.1}$$

with B the symmetric positive definite mass matrix, A the stiffness matrix, and $u(t) = (u_1(t), u_2(t), \ldots, u_d(t))^t$ the unknown solution vector.

In [5] we considered solving (1.1) with the continuous-time multigrid waveform relaxation method. This method is based on waveform relaxation, a highly parallel technique for solving very large systems of ODEs, [7, 10]. It is accelerated by using multigrid, a very efficient method for solving elliptic partial differential equations, see e.g. [2, 17]. The continuous-time waveform relaxation method differs from standard ODE-solvers in that it computes a solution along a continuous time-interval. It requires the analytical solution of certain ODEs and the exact continuous representation of certain functions. The method is therefore mainly of theoretical interest. In an actual implementation of the method, the algorithm is replaced by a discrete-time algorithm. That is, functions are represented discretely as vectors defined on successive time-levels, and the ODEs are solved by using standard time-stepping techniques.

In this paper, we continue our study of the multigrid waveform relaxation method for systems of the form (1.1). In particular, we analyse the effect of time-discretisation when linear multistep formulae are used. The structure of this paper is similar to the structure of [5]. In §2, we analyse the spectral properties of certain operators that arise in the formulation of the waveform relaxation methods. After a brief review of some definitions and properties of linear multistep methods in §3, we investigate the convergence of the discrete-time standard waveform relaxation method (§4) and of its

* Katholieke Universiteit Leuven, Department of Computer Science, Celestijnenlaan 200A, B-3001 Leuven, Belgium. This text presents research results of the Belgian Incentive Program "Information Technology" - Computer Science of the future, initiated by the Belgian State - Prime Minister's Service - Science Policy office. The scientific responsibility is assumed by its authors.

† California Institute of Technology, Applied Mathematics 217-50, Pasadena, CA 91125. This work was supported in part by the NSF under Cooperative Agreement No. CCR-9120008.

two-grid acceleration (§5), both on finite and infinite time-intervals. For systems of the form (1.1) with $B = I$, the discrete-time waveform method and its multigrid variant have been investigated in [9, 11, 12, 15]. Our results are qualitatively very similar, and generalise the ones found in these references. Finally, in §6, extensive numerical results are reported for a one-dimensional and two-dimensional model problem.

2. Spectral properties of a special operator. We will show in §4 and §5 that the discrete-time waveform relaxation method and its two-grid acceleration can be written as successive approximation schemes of the form

$$u_\tau^{(\nu)} = \mathcal{H}_\tau u_\tau^{(\nu-1)} + \varphi_\tau \ . \tag{2.1}$$

We use subscript τ-notation to denote vectors or sequences, e.g. $u_\tau^{(\nu)} = \{u_i^{(\nu)}\}_{i=0}^{N-1}$, where N is the (possibly infinite) number of components. Each component is a d-vector, and will typically approximate the solution of the system of d differential equations (1.1) at a given time-level. Operator $\mathcal{H}_\tau$ is a linear discrete convolution operator with matrix-valued kernel h_τ,

$$(\mathcal{H}_\tau u_\tau)_j = (h_\tau \star u_\tau)_j = \sum_{i=0}^{j} h_{j-i} u_i \ , \quad j = 0, \ldots, N-1 \ .$$

The convergence properties of operator $\mathcal{H}_\tau$ will be analysed in the spaces of $\mathbb{C}^d$-valued p-summable sequences of length N, $l_p(N;\mathbb{C}^d)$, or $l_p(N)$ for short. These are Banach spaces with norms given by

$$||u_\tau||_p = \begin{cases} \sqrt[p]{\sum_{i=0}^{N-1} ||u_i||^p} & 1 \le p < \infty \\ \sup\limits_{0 \le i < N} \{||u_i||\} & p = \infty \end{cases} \ , \tag{2.2}$$

with $||.||$ any usual $\mathbb{C}^d$ vector-norm. Recall that the iterative scheme (2.1) is convergent if and only if the spectral radius of $\mathcal{H}_\tau$, denoted by $\rho(\mathcal{H}_\tau)$, is smaller than one. The spectral radius is defined as the largest value ρ for which $|\lambda| > \rho$ implies that $\lambda - \mathcal{H}_\tau$ has a bounded inverse. When N is finite, it equals the magnitude of the largest eigenvalue of $\mathcal{H}_\tau$.

2.1. Spectral radius on finite time-intervals.

LEMMA 2.1. *Consider $\mathcal{H}_\tau$ as an operator in $l_p(N)$, with $1 \le p \le \infty$ and N finite. Then, $\mathcal{H}_\tau$ is a bounded operator and*

$$\rho(\mathcal{H}_\tau) = \rho(h_0) = \rho(\mathbf{H}_\tau(\infty)) \ , \tag{2.3}$$

with $\mathbf{H}_\tau(z) = \sum_{i=0}^{N-1} h_i z^{-i}$ the discrete Laplace-transform of the kernel of $\mathcal{H}_\tau$.

Proof. Since $\mathcal{H}_\tau$ is a linear operator in a finite-dimensional space, boundedness of $\mathcal{H}_\tau$ follows. The operation $\mathcal{H}_\tau u_\tau$ can be represented in a standard linear algebra notation as a matrix-vector product,

$$\begin{bmatrix} h_0 & & & & & \\ h_1 & h_0 & & & & \\ h_2 & h_1 & h_0 & & & \\ \cdot & \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \\ h_{N-1} & \cdot & \cdot & h_2 & h_1 & h_0 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \cdot \\ \cdot \\ \cdot \\ u_{N-1} \end{bmatrix} . \tag{2.4}$$

The spectral radius of operator $\mathcal{H}_\tau$ equals the spectral radius of the $N \times N$ block lower triangular Toeplitz matrix in (2.4). By consequence, $\rho(\mathcal{H}_\tau) = \rho(h_0)$. The second equality follows immediately. □

2.2. Spectral radius on infinite time-intervals.

LEMMA 2.2. *Suppose $h_\tau \in l_1(\infty)$, and consider $\mathcal{H}_\tau$ as an operator in $l_p(\infty)$, with $1 \leq p \leq \infty$. Then, $\mathcal{H}_\tau$ is bounded and*

$$\rho(\mathcal{H}_\tau) = \sup_{|z|\geq 1} \rho(\mathbf{H}_\tau(z)) \tag{2.5}$$

$$= \sup_{|z|=1} \rho(\mathbf{H}_\tau(z)) \ , \tag{2.6}$$

with $\mathbf{H}_\tau(z) = \sum_{i=0}^{\infty} h_i z^{-i}$ the discrete Laplace-transform of h_τ.

The outline of our proof is very similar to the one given in [11, Th. 3.1], yet it is phrased in terms of general convolution operators. A similar line of arguments is implied in the proof of [9, Prop. 9]. The proof is based on the discrete version of the Paley-Wiener Theorem, [8]. This theorem states that the solution of a discrete Volterra convolution equation $x_\tau + h_\tau * x_\tau = f_\tau$ with $f_\tau \in l_p(\infty)$ and $h_\tau \in l_1(\infty)$ is bounded in $l_p(\infty)$ if and only if $\det(I + \mathbf{H}_\tau(z)) \neq 0$ for $|z| \geq 1$, with $\mathbf{H}_\tau(z)$ the discrete Laplace-transform of h_τ.

Proof. The boundedness of $\mathcal{H}_\tau$ follows from the fact that $l_1 * l_p \subset l_p$. Indeed, applying Young's inequality for discrete convolution products, [4, p. 198], yields

$$||\mathcal{H}_\tau u_\tau||_p \leq ||h_\tau||_1 \ ||u_\tau||_p \ .$$

By definition, the spectral radius of $\mathcal{H}_\tau$ is the smallest value of ρ for which $|\lambda| > \rho$ implies that $\lambda - \mathcal{H}_\tau$ has a bounded inverse in $l_p(\infty)$. Consider

$$\lambda u_\tau - \mathcal{H}_\tau u_\tau = \lambda u_\tau - h_\tau * u_\tau = f_\tau \ ,$$

with $f_\tau \in l_p(\infty)$. Suppose $\lambda \neq 0$, then this can be rewritten as a convolution equation

$$u_\tau - \frac{1}{\lambda} h_\tau * u_\tau = \frac{1}{\lambda} f_\tau \ .$$

By the Paley-Wiener Theorem, it follows that u_τ is bounded if and only if

$$\det\left(I - \frac{1}{\lambda}\mathbf{H}_\tau(z)\right) \neq 0 \quad \text{for} \quad |z| \geq 1 \ ,$$

or, equivalently,

$$\rho(\mathcal{H}_\tau) = \sup_{|z|\geq 1} \rho(\mathbf{H}_\tau(z)) \ .$$

Note that $\mathbf{H}_\tau(z)$ is analytic for $|z| > 1$, including $z = \infty$, and, since $h_\tau \in l_1(\infty)$ it is continuous for $|z| \geq 1$. As the spectral radius satisfies the maximum principle, we obtain equality (2.6). □

REMARK 2.1. In the case of $d = 1$, this lemma corresponds to a well-known spectral property of semi-infinite Toeplitz operators, [13, Th. 2.1].

In $l_2(\infty)$, an analogous result holds for the norm.

LEMMA 2.3. *Suppose $h_\tau \in l_1(\infty)$, and consider $\mathcal{H}_\tau$ as an operator in $l_2(\infty)$. Denote by $||.||_2$ the l_2-norm and by $||.||$ the standard Euclidean vector norm. Then,*

$$||\mathcal{H}_\tau||_2 = \sup_{|z|\geq 1} ||\mathbf{H}_\tau(z)|| \tag{2.7}$$

$$= \sup_{|z|=1} ||\mathbf{H}_\tau(z)|| \ , \tag{2.8}$$

with $\mathbf{H}_\tau(z)$ the discrete Laplace-transform of h_τ.

Proof. The proof is based on Parseval's relation for l_2-sequences, [18, p. 42],

$$||u_\tau||_2 := ||\{u_i\}_{i=0}^{\infty}||_2 = ||\sum_{i=0}^{\infty} u_i z^{-i}||_{H_2} \ ,$$

where $||\cdot||_{H_2}$ is the norm in the Hardy-Lebesgue space of square integrable functions analytic outside the unit disk,

$$||f(z)||_{H_2} = \sup_{r>1} \left(\frac{1}{2\pi} \int_0^{2\pi} ||f(re^{i\theta})||^2 d\theta \right)^{1/2} \ .$$

By definition of operator norm and by Parseval's relation, we have

$$||\mathcal{H}_\tau||_2 = \sup \frac{||\mathcal{H}_\tau u_\tau||_2}{||u_\tau||_2} = \sup \frac{||\mathbf{H}_\tau(z)\tilde{u}_\tau(z)||_{H_2}}{||\tilde{u}_\tau(z)||_{H_2}} \ ,$$

with $\tilde{u}_\tau(z)$ the discrete Laplace-transform of u_τ. The latter can be seen to be equal to $\sup_{|z|\geq 1} ||H(z)||$. (For the technical details of this last step, we refer to the proof of a very similar theorem, [1, Th. 2.2], which deals with operator-norms of Fourier multipliers.) □

REMARK 2.2. From (2.3) and (2.5), it follows that the spectral radius of $\mathcal{H}_\tau$ on finite time-intervals is smaller than the spectral radius of $\mathcal{H}_\tau$ on infinite time-intervals.

3. Some linear multistep formulae. For the reader's convenience, we recall the general linear multistep formula for calculating the solution to the ODE $\dot{y} = f(t, y)$ with $y(0) = y_0$, see e.g. [6, p. 11],

$$\frac{1}{\tau} \sum_{j=0}^{k} \alpha_j y_{n+j} = \sum_{j=0}^{k} \beta_j f_{n+j} \ . \tag{3.1}$$

In this formula, α_j and β_j are real constants, and τ denotes a constant step-size. We shall assume that k starting values $y_0, y_1, \ldots, y_{k-1}$ are given.

DEFINITION 3.1. *The characteristic polynomials of the linear multistep method are given by*

$$a(z) = \sum_{j=0}^{k} \alpha_j z^j \quad \text{and} \quad b(z) = \sum_{j=0}^{k} \beta_j z^j \ .$$

Throughout this paper we adhere to some common assumptions. The linear multistep method is irreducible: $a(z)$ and $b(z)$ have no common roots; the linear multistep method is consistent: $a(1) = 0$ and $a'(1) = b(1)$; the linear multistep method is zero-stable: all roots of $a(z)$ are inside the closed unit disk and every root with modulus

one is simple. For future reference, we also define the stability region of a linear multistep method, and the related notion of $A(\alpha)$-stability, see e.g. [3, 6].

DEFINITION 3.2. *The stability region S consists of those $\mu \in \bar{\mathbb{C}}$ for which the polynomial $a(z) - \mu b(z)$ (around $\mu = \infty$: $\mu^{-1}a(z) - b(z)$) satisfies the root condition: all roots satisfy $|z_j| \leq 1$ and those of modulus 1 are simple.*

DEFINITION 3.3. *A multistep method is called*
i) $A(\alpha)$-stable, $0 < \alpha < \frac{\pi}{2}$, if $S \supset \Sigma_\alpha = \{z : |Arg(-z)| < \alpha,\ z \neq 0\}$
ii) A-stable if S contains the left half complex plane.

4. The waveform relaxation method. The continuous-time waveform relaxation method for solving initial value problem (1.1) is defined by the splittings $B = M_B - N_B$, $A = M_A - N_A$, and the iteration scheme

$$M_B \dot{u}^{(\nu)} + M_A u^{(\nu)} = N_B \dot{u}^{(\nu-1)} + N_A u^{(\nu-1)} + f \ , \tag{4.1}$$

with $u^{(\nu)}(0) = u_0$. We assume the splitting is such that M_B is invertible. This iterative scheme can be written in explicit form as $u^{(\nu)} = \mathcal{K}u^{(\nu-1)} + \varphi$. The convergence properties of iteration operator $\mathcal{K}$, the *continuous-time waveform relaxation operator*, have been studied in [5]. They are expressed in terms of the waveform relaxation matrix

$$\mathbf{K}(z) = (zM_B + M_A)^{-1}(zN_B + N_A) \ . \tag{4.2}$$

It was shown that, respectively on finite and infinite time-intervals, and with $i = \sqrt{-1}$,

$$\rho(\mathcal{K}) = \rho(\mathbf{K}(\infty)) \quad \text{and} \quad \rho(\mathcal{K}) = \sup_{\mathrm{Re}(z)\geq 0} \rho(\mathbf{K}(z)) = \sup_{\xi \in \mathbb{R}} \rho(\mathbf{K}(i\xi)) \ . \tag{4.3}$$

4.1. The discrete-time waveform relaxation operator. Application of linear multistep formula (3.1) to the continuous-time iteration scheme (4.1) leads to

$$\begin{aligned} &\frac{1}{\tau}\sum_{j=0}^{k} \alpha_j M_B u_{n+j}^{(\nu)} + \sum_{j=0}^{k} \beta_j M_A u_{n+j}^{(\nu)} = \\ &\qquad \frac{1}{\tau}\sum_{j=0}^{k} \alpha_j N_B u_{n+j}^{(\nu-1)} + \sum_{j=0}^{k} \beta_j N_A u_{n+j}^{(\nu-1)} + \sum_{j=0}^{k} \beta_j f_{n+j} \ , \quad n \geq 0 \ . \end{aligned} \tag{4.4}$$

We do not iterate on the k starting values, i.e., $u_j^{(\nu)} = u_j^{(\nu-1)} = u_j$, for $j < k$. In the remainder of the text we shall concentrate on the use of implicit methods, i.e., $\beta_k \neq 0$. Equation (4.4) can then be solved uniquely for every n if and only if the following condition is satisfied:

$$\frac{\alpha_k}{\beta_k} \notin \sigma\left(-\tau M_B^{-1} M_A\right) \ , \tag{4.5}$$

where $\sigma(\cdot)$ denotes the spectrum. Further on we shall refer to this condition as the discrete solvability condition.

Iteration (4.4) can be rewritten as $u_\tau^{(\nu)} = \mathcal{K}_\tau u_\tau^{(\nu-1)} + \varphi_\tau$. Because we do not iterate on the starting values, we use a slightly different subscript τ-notation here than the one in (2.1); that is,

$$u_\tau = \{u_{k+i}\}_{i=0}^{N-1} \ . \tag{4.6}$$

(Alternatively, we could have used negative indices to denote the time-levels associated with the k starting values, as is done in [8, 9]. This, however, would require some shifting in the indices of formulae (3.1) and (4.4).) The precise expression for φ_τ can be calculated following the lines of [12, p. 536-537]. It depends on the values of $f_n, n \geq 0$ and on the starting values u_n, $n < k$. In order to determine the nature of $\mathcal{K}_\tau$, the *discrete-time waveform relaxation operator*, we rewrite (4.4) using $e_n^{(\nu)} = u_n^{(\nu)} - u_n$. Here, u_n is the exact solution of ODE (1.1) when discretised using the linear multistep method. This gives

$$\frac{1}{\tau}\sum_{j=0}^{k}\alpha_j M_B e_{n+j}^{(\nu)} + \sum_{j=0}^{k}\beta_j M_A e_{n+j}^{(\nu)} = \frac{1}{\tau}\sum_{j=0}^{k}\alpha_j N_B e_{n+j}^{(\nu-1)} + \sum_{j=0}^{k}\beta_j N_A e_{n+j}^{(\nu-1)} \ , \quad n \geq 0 \ .$$

With $C_j = \frac{1}{\tau}\alpha_j M_B + \beta_j M_A$, and $D_j = \frac{1}{\tau}\alpha_j N_B + \beta_j N_A$, this becomes

$$\sum_{j=0}^{k} C_j e_{n+j}^{(\nu)} = \sum_{j=0}^{k} D_j e_{n+j}^{(\nu-1)} \ , \quad n \geq 0 \ . \tag{4.7}$$

Note that $e_j^{(\nu)} = e_j^{(\nu-1)} = 0$, $j < k$. When we combine the first N equations, i.e., the equations for the unknowns on time-steps $k, \ldots, N+k-1$, and after introducing vector $E^{(\nu)} = [e_k^{(\nu)} e_{k+1}^{(\nu)} \ldots e_{N+k-1}^{(\nu)}]^t$, we get

$$E^{(\nu)} = C^{-1} D \, E^{(\nu-1)} \ . \tag{4.8}$$

Matrices C and D are $N \times N$ block lower triangular matrices with $k+1$ constant diagonals. The blocks on the j-th diagonal are given respectively by C_{k-j} and D_{k-j}. It follows immediately that matrix $C^{-1}D$ is a $N \times N$ block lower triangular Toeplitz matrix. Hence, $\mathcal{K}_\tau$ is a discrete linear convolution operator on the l_p-space of vectors or sequences of length N. The j-th component of the matrix-valued discrete convolution kernel k_τ equals the (constant) submatrix on the j-th lower block diagonal of $C^{-1}D$.

In the theory we shall need the discrete Laplace-transform of the convolution kernel. It can be found by discrete Laplace-transforming equation (4.7). If $\tilde{e}_\tau^{(\nu)}(z)$ denotes the transform of $e_\tau^{(\nu)}$, we obtain

$$\tilde{e}_\tau^{(\nu)}(z) = \mathbf{K}_\tau(z)\tilde{e}_\tau^{(\nu-1)}(z) \ ,$$

with the discrete-time waveform relaxation matrix given by

$$\mathbf{K}_\tau(z) = (a(z)M_B + \tau b(z)M_A)^{-1}(a(z)N_B + \tau b(z)N_A) \ . \tag{4.9}$$

By comparison to (4.2) the following relation results,

$$\mathbf{K}_\tau(z) = \mathbf{K}(\frac{1}{\tau}\frac{a}{b}(z)) \ . \tag{4.10}$$

Note that (4.10) still holds when $\frac{a}{b}(z)$ is set to ∞ in the case of $b(z) = 0$. (In this case $a(z) \neq 0$, since the characteristic polynomials have no common roots.)

4.2. Convergence analysis.

4.2.1. Convergence on finite time-intervals.

THEOREM 4.1. *Assume that condition (4.5) is satisfied, and consider $\mathcal{K}_\tau$ as an operator in $l_p(N)$, with $1 \leq p \leq \infty$ and N finite. Then, $\mathcal{K}_\tau$ is bounded and*

$$\rho(\mathcal{K}_\tau) = \rho\left(\mathbf{K}\left(\frac{1}{\tau}\frac{\alpha_k}{\beta_k}\right)\right) . \tag{4.11}$$

Proof. The theorem follows from Lemma 2.1 and the observation that

$$\lim_{z\to\infty} \mathbf{K}_\tau(z) = \lim_{z\to\infty} \mathbf{K}\left(\frac{1}{\tau}\frac{a}{b}(z)\right) = \mathbf{K}\left(\frac{1}{\tau}\frac{\alpha_k}{\beta_k}\right) .$$

□

4.2.2. Convergence on infinite time-intervals. The following lemma deals with the boundedness of the discrete-time waveform relaxation operator $\mathcal{K}_\tau$. It is proved using a matrix-valued version of Wiener's inversion Theorem, [8, p. 446] and [11, p. 577], which is stated here for the reader's convenience.

THEOREM 4.2 (WIENER'S INVERSION THEOREM). *Given a matrix valued sequence A_τ such that $A_\tau \in l_1(\infty)$, and assume that*

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

for $|z| \geq 1$. Setting $\sum_{i=0}^{\infty} B_i z^{-i} = \left(\sum_{i=0}^{\infty} A_i z^{-i}\right)^{-1}$, we have $B_\tau \in l_1(\infty)$.

LEMMA 4.3. *If $\sigma(-\tau M_B^{-1} M_A) \subset int\, S$, then $\mathcal{K}_\tau$ is bounded in $l_p(\infty)$.*

Proof. It is sufficient to prove that the kernel k_τ of the discrete convolution operator $\mathcal{K}_\tau$ is an l_1-sequence. To this end, consider first the l_1-sequence

$$\alpha_k M_B + \tau\beta_k M_A,\ \alpha_{k-1} M_B + \tau\beta_{k-1} M_A,\ \ldots,\ \alpha_0 M_B + \tau\beta_0 M_A,\ 0,\ 0, \ldots$$

Its discrete Laplace-transform equals the matrix function $z^{-k}\left(a(z)M_B + \tau b(z)M_A\right)$. By Wiener's Theorem, we have that the inverse, $\left(a(z)M_B + \tau b(z)M_A\right)^{-1} z^k$, is the transform of another l_1-sequence, say r_τ, if

$$\det\left(a(z)M_B + \tau b(z)M_A\right) \neq 0 \quad \text{for} \quad |z| \geq 1 . \tag{4.12}$$

Next, consider the l_1-sequence

$$s_\tau = \alpha_k N_B + \tau\beta_k N_A,\ \alpha_{k-1} N_B + \tau\beta_{k-1} N_A,\ \ldots,\ \alpha_0 N_B + \tau\beta_0 N_A,\ 0,\ 0,\ \ldots ,$$

the discrete Laplace-transform of which is given by $z^{-k}(a(z)N_B + \tau b(z)N_A)$. The convolution of r_τ and s_τ is another l_1-sequence, which can be seen to be equal to the kernel k_τ. Indeed, the discrete Laplace-transform of $r_\tau \star s_\tau$ is identical to $\mathbf{K}_\tau(z)$. As a result, it follows that $\mathcal{K}_\tau$ is bounded if (4.12) is satisfied.

Suppose there is a z with $|z| \geq 1$ such that

$$\det\left(a(z)M_B + \tau b(z)M_A\right) = 0 . \tag{4.13}$$

Then necessarily $b(z) \neq 0$. (If $b(z) = 0$ then $a(z) \neq 0$, because $a(z)$ and $b(z)$ have no common roots. Since M_B is assumed to be invertible, equality (4.13) can not hold.) Hence, we obtain

$$\det\left(\frac{a}{b}(z)M_B + \tau M_A\right) = 0 ,$$

and therefore $\frac{a}{b}(z) \in \sigma(-\tau M_B^{-1} M_A)$. Since $|z| \geq 1$, it follows that $-\tau M_B^{-1} M_A$ has an eigenvalue which is not an interior point of S. This contradicts the assumption of the lemma. Hence, (4.12) is satisfied. □

REMARK 4.1. Condition $\sigma(-\tau M_B^{-1} M_A) \subset \operatorname{int} S$ implies the discrete solvability condition (4.5). Indeed, since $\frac{\alpha_k}{\beta_k} = \frac{a}{b}(\infty)$, it follows that $\frac{\alpha_k}{\beta_k} \notin \operatorname{int} S$, and, therefore, $\frac{\alpha_k}{\beta_k} \notin \sigma(-\tau M_B^{-1} M_A)$.

REMARK 4.2. Condition $\sigma(-\tau M_B^{-1} M_A) \subset \operatorname{int} S$ implies that all poles of $\mathbf{K}(z)$ are in the interior of the scaled stability region $\frac{1}{\tau} S$.

THEOREM 4.4. *Assume $\sigma(-\tau M_B^{-1} M_A) \subset int S$, and consider $\mathcal{K}_\tau$ as an operator in $l_p(\infty)$, with $1 \leq p \leq \infty$. Then,*

$$\rho(\mathcal{K}_\tau) = \sup\{\rho(\mathbf{K}(z)) | \tau z \in \bar{\mathbb{C}} \setminus int S\} \tag{4.14}$$

$$= \sup_{\tau z \in \partial S} \rho(\mathbf{K}(z)) \ . \tag{4.15}$$

Proof. As $\sigma(-\tau M_B^{-1} M_A) \subset \operatorname{int} S$, it follows that $k_\tau \in l_1(\infty)$. Lemma 2.2 yields

$$\rho(\mathcal{K}_\tau) = \sup_{|z| \geq 1} \rho(\mathbf{K}_\tau(z)) = \sup_{|z| \geq 1} \rho\left(\mathbf{K}\left(\frac{1}{\tau}\frac{a}{b}(z)\right)\right) \ .$$

By definition of the stability region,

$$\bar{\mathbb{C}} \setminus \operatorname{int} S = \left\{\frac{a}{b}(z) : |z| \geq 1\right\} \ ,$$

and thereby (4.14) follows. Equality (4.15) is obtained by the maximum principle. □

In $l_2(\infty)$, a similar result holds for the norm by application of Lemma 2.3.

THEOREM 4.5. *Assume $\sigma(-\tau M_B^{-1} M_A) \subset int S$, and consider $\mathcal{K}_\tau$ as an operator in $l_2(\infty)$. Denote by $||.||_2$ the l_2-norm and by $||.||$ the standard Euclidean vector-norm. Then,*

$$||\mathcal{K}_\tau||_2 = \sup\{||\mathbf{K}(z)|| : \tau z \in \bar{\mathbb{C}} \setminus int S\} \tag{4.16}$$

$$= \sup_{\tau z \in \partial S} ||\mathbf{K}(z)|| \ . \tag{4.17}$$

In analogy to the discussion in [12, Th. 4.2] we can make the following note.

REMARK 4.3. When the assumption in the above theorems is violated, a weaker condition may be satisfied: $\sigma(-\tau M_B^{-1} M_A) \subset \operatorname{int} S_{\gamma\tau}$, where $S_{\gamma\tau}$ consists of all μ for which $a(e^{-\gamma\tau} z) - \mu b(e^{-\gamma\tau} z)$ (around $\mu = \infty$: $\mu^{-1} a(e^{-\gamma\tau} z) - b(e^{-\gamma\tau} z)$) satisfies the root condition. The analysis then can be redone using an exponentially scaled norm,

$$||u_\tau||_\gamma = ||\{u_i\}||_\gamma = ||\{e^{-\gamma\tau i} u_i\}|| \ . \tag{4.18}$$

The norm in the right-hand side is a standard p-norm (2.2). With this change of norm, the suprema in Theorems 4.4 and 4.5 have to be taken over all τz in $\bar{\mathbb{C}} \setminus \operatorname{int} S_{\gamma\tau}$, or, after application of the maximum principle, over $\partial S_{\gamma\tau}$.

4.3. Discrete-time versus continuous-time results. The continuous-time results (4.3) are regained when we let $\tau \to 0$ in the convergence formulae for operator $\mathcal{K}_\tau$. For finite time-intervals, we have

$$\lim_{\tau \to 0} \rho(\mathcal{K}_\tau) = \lim_{\tau \to 0} \rho\left(\mathbf{K}\left(\frac{1}{\tau}\frac{\alpha_k}{\beta_k}\right)\right) = \rho(\mathbf{K}(\infty)) = \rho(\mathcal{K}) \ .$$

A similar result is found for infinite time-intervals. Note that the tangent to ∂S in the origin of the complex plane is the imaginary axis, for any consistent linear multistep method. As such, the boundary of the scaled stability region $\partial(\frac{1}{\tau}S)$ tends to the imaginary axis when $\tau \to 0$. Consequently,

$$\lim_{\tau\to 0} \rho(\mathcal{K}_\tau) = \lim_{\tau\to 0} \sup_{\tau z\in\partial S} \rho(\mathbf{K}(z)) = \sup_{\xi\in\mathbb{R}} \rho(\mathbf{K}(i\xi)) = \rho(\mathcal{K}) \ .$$

Furthermore, for a fixed time-step τ, we can prove the following theorem for $A(\alpha)$-stable linear multistep methods (see Definition 3.3). The theorem is closely related to [9, Prop. 9], where multigrid waveform relaxation on finite difference grids is analysed. We reformulate the proof, using our notations, for completeness.

THEOREM 4.6. *Assume $\sigma(-\tau M_B^{-1} M_A) \subset \Sigma_\alpha$. Consider $\mathcal{K}_\tau$ as an operator in $l_p(\infty)$ and $\mathcal{K}$ as an operator in $L_p(0,\infty)$, with $1 \le p \le \infty$. Then,*
i) if the linear multistep method is A-stable, then $\rho(\mathcal{K}_\tau) \le \rho(\mathcal{K})$.
ii) if the linear multistep method is $A(\alpha)$-stable, then

$$\rho(\mathcal{K}_\tau) \le \sup_{z\in\Sigma_\alpha^c} \rho(\mathbf{K}(z)) = \sup_{z\in\partial\Sigma_\alpha^c} \rho(\mathbf{K}(z)) \ , \tag{4.19}$$

with $\Sigma_\alpha^c = \bar{\mathbb{C}} \setminus \Sigma_\alpha = \{z : |Arg(z)| \le \pi - \alpha\}$.

Proof. Part i) is a special case of ii) with $\alpha = \pi/2$, combined with the second equality of (4.3). For part ii), we notice that we may apply Theorem 4.4 since $\sigma(-\tau M_B^{-1} M_A) \subset \Sigma_\alpha \subset \operatorname{int} S$. Therefore,

$$\rho(\mathcal{K}_\tau) = \sup_{|z|\ge 1} \rho\left(\mathbf{K}\left(\frac{1}{\tau}\frac{a}{b}(z)\right)\right) \ . \tag{4.20}$$

If the multistep method is $A(\alpha)$-stable, then $\frac{a}{b}(z) \in \Sigma_\alpha^c$ for $|z| \ge 1$. Combining the latter with (4.20) yields the inequality of (4.19). The equality is obtained by the maximum principle. □

5. The multigrid waveform relaxation method. The splittings of matrices B and A used in actual computations typically correspond to Gauss-Seidel or weighted Jacobi splittings. Each iteration defined by (4.1) can then be computed as the solution of d ordinary differential equations, each in a single unknown. The resulting iteration can be accelerated by using the multigrid principle, in a very similar way as the standard point-wise relaxation methods are accelerated when solving elliptic partial differential equations.

The continuous-time two-grid waveform relaxation scheme is sketched below. We refer to [5] for a more elaborate description. The algorithm uses two nested grids, a coarse grid Ω_H and a fine grid Ω_h. Grid functions are mapped from the one grid to the other by a prolongation (or interpolation) operator ($p : \Omega_H \to \Omega_h$) and a restriction operator ($r : \Omega_h \to \Omega_H$). The discretisation on the fine grid is defined by the matrices B_h and A_h, the discretisation on the coarse grid by B_H and A_H. One iteration transforms iterate $u^{(\nu-1)}$ into $u^{(\nu)}$ in three steps.

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

$$M_{B_h}\dot{x}^{(\nu)} + M_{A_h}x^{(\nu)} = N_{B_h}\dot{x}^{(\nu-1)} + N_{A_h}x^{(\nu-1)} + f_h \ , \text{ with } x^{(\nu)}(0) = u_0 \ . \tag{5.1}$$

(ii) Coarse grid correction. Calculate the defect

$$d_h = B\dot{x}_h^{(\nu_1)} + A_h x_h^{(\nu_1)} - f_h \ .$$

Solve the coarse-grid defect equation

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

and correct,

$$\bar{x} = x^{(\nu_1)} - p v_H .$$

(iii) Post-smoothing. Set $x^{(0)} = \bar{x}$ and perform ν_2 fine-grid waveform relaxation steps (5.1). Set $u^{(\nu)} = x^{(\nu_2)}$.

This two-grid cycle can be written as $u^{(\nu)} = \mathcal{M} u^{(\nu-1)} + \varphi$, where $\mathcal{M}$ is called the *continuous-time two-grid waveform relaxation operator.* The convergence formulae of $\mathcal{M}$ as an iteration operator resemble those of the standard waveform relaxation method. More precisely, in [5] we find respectively for the finite and for the infinite time-interval case

$$\rho(\mathcal{M}) = \rho(\mathbf{M}(\infty)) \quad \text{and} \quad \rho(\mathcal{M}) = \sup_{\mathrm{Re}(z)\geq 0} \rho(\mathbf{M}(z)) = \sup_{\xi \in \mathbb{R}} \rho(\mathbf{M}(i\xi)) . \tag{5.2}$$

$\mathbf{M}(z)$, the continuous-time two-grid waveform relaxation matrix, is given by

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

with $\mathbf{K}(z)$ the fine-grid matrix given in (4.2).

5.1. The discrete-time two-grid waveform relaxation operator. We discretise the equations of the continuous-time two-grid cycle using a linear multistep method with a fixed time-step τ. As before, we assume that we do not iterate on k given starting values. The discrete-time two-grid cycle defines a linear operator $\mathcal{M}_\tau$, which satisfies

$$u_\tau^{(\nu)} = \mathcal{M}_\tau u_\tau^{(\nu-1)} + \varphi_\tau \quad \text{and} \quad e_\tau^{(\nu)} = \mathcal{M}_\tau e_\tau^{(\nu-1)} , \tag{5.3}$$

where $e_\tau^{(\nu)}$ is the error of the ν-th iterate. Our notation is again similar to (4.6). $\mathcal{M}_\tau$ is called the *discrete-time two-grid waveform relaxation operator.*

The second equation of (5.3) can be reformulated in a similar way as we arrived at (4.8),

$$E^{(\nu)} = (C_h^{-1} D_h)^{\nu_2} (I - P F_H^{-1} R F_h)(C_h^{-1} D_h)^{\nu_1} E^{(\nu-1)} . \tag{5.4}$$

Here, $E^{(\nu)} = [e_k^{(\nu)} e_{k+1}^{(\nu)} \ldots e_{N+k-1}^{(\nu)}]^t$. Matrices C_h, D_h, F_H and F_h are $N \times N$ block lower triangular matrices with $k+1$ constant diagonals. The blocks of the j-th diagonal equal respectively $(C_h)_{k-j}$, $(D_h)_{k-j}$, $(F_H)_{k-j}$ and $(F_h)_{k-j}$, with

$$(C_h)_j = \frac{1}{\tau}\alpha_j M_{B_h} + \beta_j M_{A_h} , \quad (D_h)_j = \frac{1}{\tau}\alpha_j N_{B_h} + \beta_j N_{A_h} ,$$

and

$$(F_H)_j = \frac{1}{\tau}\alpha_j B_H + \beta_j A_H , \quad (F_h)_j = \frac{1}{\tau}\alpha_j B_h + \beta_j A_h .$$

Matrices P and R are block diagonal with constant diagonal blocks respectively equal to matrices p and r. I is the identity matrix of dimension $d \times N$. The resulting

discrete-time two-grid cycle is well-defined, if and only if the following conditions hold,

$$\frac{\alpha_k}{\beta_k} \notin \sigma(-\tau M_{B_h}^{-1} M_{A_h}) \quad \text{and} \quad \frac{\alpha_k}{\beta_k} \notin \sigma(-\tau B_H^{-1} A_H) \ . \tag{5.5}$$

We shall refer to (5.5) as the discrete solvability conditions for the two-grid algorithm.

It can be seen that the matrix pre-multiplying $E^{(\nu-1)}$ in (5.4) is block lower triangular of dimension N. This implies that $\mathcal{M}_\tau$ is a discrete linear convolution operator. The discrete Laplace-transform of its matrix-valued kernel can be found by transforming the equations of the discrete-time two-grid cycle. It is denoted by $\mathbf{M}_\tau(z)$, the discrete-time two-grid waveform relaxation matrix, and equals

$$\mathbf{M}_\tau(z) = \mathbf{K}_\tau{}^{\nu_2}(z)\, \mathbf{C}_\tau(z)\, \mathbf{K}_\tau{}^{\nu_1}(z) \ ,$$

with $\mathbf{K}_\tau(z)$ given by (4.9) and $\mathbf{C}_\tau(z)$ given by

$$I - p\,(a(z)B_H + \tau b(z)A_H)^{-1}\, r\,(a(z)B_h + \tau b(z)A_h) \ .$$

Matrix $\mathbf{M}_\tau(z)$ satisfies as similar relation as $\mathbf{K}_\tau(z)$ does in (4.10),

$$\mathbf{M}_\tau(z) = \mathbf{M}\left(\frac{1}{\tau}\frac{a}{b}(z)\right) \ . \tag{5.6}$$

5.2. Convergence analysis. The convergence analysis of operator $\mathcal{M}_\tau$ is very similar to the convergence analysis of the standard waveform relaxation operator $\mathcal{K}_\tau$.

5.2.1. Convergence on finite time-intervals.

THEOREM 5.1. *Assume that conditions (5.5) are satisfied, and consider $\mathcal{M}_\tau$ as an operator in $l_p(N)$, with $1 \leq p \leq \infty$ and N finite. Then, $\mathcal{M}_\tau$ is bounded and*

$$\rho(\mathcal{M}_\tau) = \rho\left(\mathbf{M}\left(\frac{1}{\tau}\frac{\alpha_k}{\beta_k}\right)\right) \ . \tag{5.7}$$

Proof. The theorem follows from Lemma 2.1 and (5.6),

$$\rho(\mathcal{M}_\tau) = \rho(\mathbf{M}_\tau(\infty)) = \rho\left(\mathbf{M}\left(\frac{1}{\tau}\frac{a}{b}(\infty)\right)\right) = \rho\left(\mathbf{M}\left(\frac{1}{\tau}\frac{\alpha_k}{\beta_k}\right)\right) \ .$$

□

5.2.2. Convergence on infinite time-intervals. We first prove the boundedness of $\mathcal{M}_\tau$, i.e, we prove the two-grid equivalent of Lemma 4.3.

LEMMA 5.2. *Assume $\sigma(-\tau M_{B_h}^{-1} M_{A_h}) \cup \sigma(-\tau B_H^{-1} A_H) \subset int S$. Then, $\mathcal{M}_\tau$ is bounded in $l_p(\infty)$.*

Proof. It is sufficient to prove that the kernel of $\mathcal{M}_\tau$ belongs to $l_1(\infty)$. We shall analyse each of the factors in the formula for $\mathbf{M}_\tau(z)$ separately.

We have, from the proof of Lemma 4.3, that $\mathbf{K}_\tau(z)$ is the discrete Laplace-transform of an l_1-sequence, say q_τ, if

$$\det\,(a(z)M_{B_h} + \tau b(z)M_{A_h}) \neq 0 \ , \quad |z| \geq 1 \ . \tag{5.8}$$

Consider the l_1-sequence

$$\alpha_k B_H + \tau\beta_k A_H,\ \alpha_{k-1}B_H + \tau\beta_{k-1}A_H, \ldots, \alpha_0 B_H + \tau\beta_0 A_H,\ 0,\ 0, \ldots$$

Its transform is given by $z^{-k}(a(z)B_H + \tau b(z)A_H)$. By Wiener's inversion Theorem, $(a(z)B_H + \tau b(z)A_H)^{-1}z^k$ is the transform of an l_1-sequence, say w_τ, if

$$\det(a(z)B_H + \tau b(z)A_H) \neq 0 \ , \quad |z| \geq 1 \ . \tag{5.9}$$

Next, consider the l_1-sequences

$$\begin{aligned} i_\tau &= I,\ 0,\ \ldots,\ 0,\ 0,\ 0,\ \ldots \\ v_\tau &= \alpha_k B_h + \tau\beta_k A_h,\ \alpha_{k-1}B_h + \tau\beta_{k-1}A_h\ ,\ldots,\ \alpha_0 B_h + \tau\beta_0 A_h,\ 0,\ 0,\ldots \end{aligned}$$

I is the $d \times d$ identity matrix. Their transforms are given respectively by

$$I \qquad \text{and} \qquad z^{-k}\,(a(z)B_h + \tau b(z)A_h) \ .$$

Now, consider the sequence

$$\underbrace{q_\tau \star q_\tau \star \ldots \star q_\tau}_{\nu_2 \text{ times}} \star (\ i_\tau - p\, w_\tau \ \star\ r\, q_\tau\ ) \ \star \underbrace{q_\tau \star q_\tau \star \ldots \star q_\tau}_{\nu_1 \text{ times}} \ . \tag{5.10}$$

If conditions (5.8) and (5.9) are satisfied, it follows that this sequence is in l_1. (l_1 is closed under convolution and addition. The multiplication of an l_1-sequence by a matrix is an l_1-sequence.) The discrete Laplace-transform of sequence (5.10) equals $\mathbf{M}_\tau(z)$, hence the sequence equals the kernel of $\mathcal{M}_\tau$. To conclude, $\mathcal{M}_\tau$ is bounded under conditions (5.8) and (5.9).

Suppose one of these conditions is violated. That is to say, there is a z with $|z| \geq 1$ such that $\det(a(z)M_{B_h} + \tau b(z)N_{B_h}) = 0$ or $\det(a(z)B_H + \tau b(z)A_H) = 0$. That would mean that $\frac{a}{b}(z) \in \sigma(-\tau M_{B_h}^{-1}M_{A_h}) \cup \sigma(-\tau B_H^{-1}A_H)$. Since $|z| \geq 1$ this violates the assumption of the lemma. □

REMARK 5.1. The assumption of Lemma 5.2 implies the two-grid discrete solvability conditions (5.5).

REMARK 5.2. The assumption of Lemma 5.2 implies that all poles of $\mathbf{M}(z)$ are inside the scaled stability region $\frac{1}{\tau}S$.

THEOREM 5.3. *Assume* $\sigma(-\tau M_{B_h}^{-1}M_{A_h}) \cup \sigma(-\tau B_H^{-1}A_H) \subset int\,S$, *and consider* $\mathcal{M}_\tau$ *as an operator in* $l_p(\infty)$, *with* $1 \leq p \leq \infty$. *Then,*

$$\rho(\mathcal{M}_\tau) = \sup\{\rho(\mathbf{M}(z)) | \tau z \in \bar{\mathbb{C}} \setminus int\,S\} \tag{5.11}$$

$$= \sup_{\tau z \in \partial S} \rho(\mathbf{M}(z)) \ . \tag{5.12}$$

Proof. The proof is a direct consequence of Lemma 2.2, and is similar to the proof of Theorem 4.4. □

Application of Lemma 2.3 yields the following result for the l_2-norm of $\mathcal{M}_\tau$.

THEOREM 5.4. *Assume* $\sigma(-\tau M_{B_h}^{-1}M_{A_h}) \cup \sigma(-\tau B_H^{-1}A_H) \subset int\,S$, *and consider* $\mathcal{M}_\tau$ *as an operator in* $l_2(\infty)$. *Denote by* $||.||_2$ *the* l_2*-norm and by* $||.||$ *the standard Euclidean vector-norm. Then,*

$$||\mathcal{M}_\tau||_2 = \sup\{||\mathbf{M}(z)|| : \tau z \in \bar{\mathbb{C}} \setminus int\,S\} \tag{5.13}$$

$$= \sup_{\tau z \in \partial S} ||\mathbf{M}(z)|| \ . \tag{5.14}$$

REMARK 5.3. If the assumption of the former theorems is violated, but the weaker condition $\sigma(-\tau M_{B_h}^{-1}M_{A_h}) \cup \sigma(-\tau B_H^{-1}A_H) \subset \text{int}\,S_{\gamma\tau}$ holds, then we can formulate an analogous remark to Remark 4.3.

5.3. Discrete-time versus continuous-time results. The relation between the two-grid operators $\mathcal{M}_\tau$ and $\mathcal{M}$ is similar to the relation between $\mathcal{K}_\tau$ and $\mathcal{K}$. More precisely, both for finite and infinite intervals:

$$\lim_{\tau \to 0} \rho(\mathcal{M}_\tau) = \rho(\mathcal{M}) \ .$$

We also state the two-grid equivalent of Theorem 4.6, without proof.

THEOREM 5.5. *Assume* $\sigma(-\tau M_{B_h}^{-1} M_{A_h}) \cup \sigma(-\tau B_H^{-1} A_H) \subset \Sigma_\alpha$. *Consider* $\mathcal{M}_\tau$ *as an operator in* $l_p(\infty)$ *and* $\mathcal{M}$ *as an operator in* $L_p(0,\infty)$, *with* $1 \leq p \leq \infty$. *Then,*
i) if the linear multistep method is A-stable, then $\rho(\mathcal{M}_\tau) \leq \rho(\mathcal{M})$.
ii) if the linear multistep method is $A(\alpha)$*-stable, then*

$$\rho(\mathcal{M}_\tau) \leq \sup_{z \in \Sigma_\alpha^c} \rho(\mathbf{M}(z)) = \sup_{z \in \partial\Sigma_\alpha^c} \rho(\mathbf{M}(z)) \ , \tag{5.15}$$

with $\Sigma_\alpha^c = \bar{\mathbb{C}} \setminus \Sigma_\alpha = \{z : |Arg(z)| \leq \pi - \alpha\}$.

6. Model problem analysis and numerical results.

6.1. Analysis of a model problem. In order to clarify the convergence behaviour of the waveform relaxation methods, we shall start with a very simple and small model problem, the one-dimensional heat equation on the unit interval,

$$\frac{\partial \mathbf{u}}{\partial t} - \Delta_1 \mathbf{u} = 0 \ , \quad x \in [0,1] \ , \quad t \in [0,1] \ . \tag{6.1}$$

We impose Dirichlet boundary conditions and an initial condition such that the analytical solution becomes $\mathbf{u}(x,t) = \sin(\pi x)\exp(-\pi^2 t)$. The problem is discretised using linear finite elements on a mesh Ω_h with mesh-size $h = 1/16$. We shall study the convergence of the waveform algorithms and the dependence of the convergence on the nature of the time-discretisation method.

6.1.1. Dependence on the time-discretisation method. We consider the Gauss-Seidel waveform relaxation algorithm and the two-level method, with one red/black Gauss-Seidel pre-smoothing step, one similar post-smoothing step, standard coarsening ($H = 2h$) and linear interpolation. The restriction is defined in the standard way for finite element multigrid methods, i.e., $r = p^t$. For both waveform algorithms, we analyse the use of different time-discretisation formulae, with a constant time-step $\tau = 1/100$. In particular, we consider the trapezoidal rule or Crank-Nicolson (CN) method, and the backward differentiation formulae (BDF) of order 1 up to 5, [3, 6]. The spectral radii of the finite and infinite time-interval operators for the standard and for the two-level algorithm are reported respectively in Table 6.1 and 6.2. The results were computed by numerical evaluation of formulae (4.11) and (4.15), and (5.7) and (5.12).

These results can be understood by looking at the *spectral picture*, [15, p. 107], which enables a graphical inspection of convergence. In the spectral picture a set of contour lines of the function $\rho(\mathbf{K}(z))$ or $\rho(\mathbf{M}(z))$ is plotted for z in a region of the complex plane close to the complex origin. On top of this picture, the scaled stability boundary of the linear multistep methods can be plotted. Figures 6.1 and 6.2 display contour lines of $\rho(\mathbf{K}(z))$ and $\rho(\mathbf{M}(z))$ (respectively for values 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0 and for values 0.1, 0.3, 0.5, 0.7, 0.9, 1.1) for the model problem, together with the scaled stability region boundaries of the CN and BDF methods.

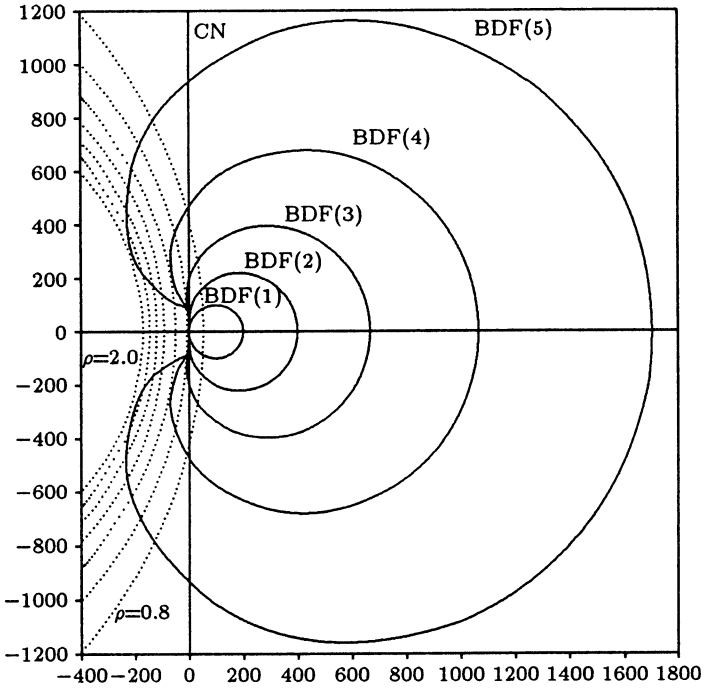

FIG. 6.1. *Graphical convergence test of $\mathcal{K}_\tau$ for model problem (6.1).*

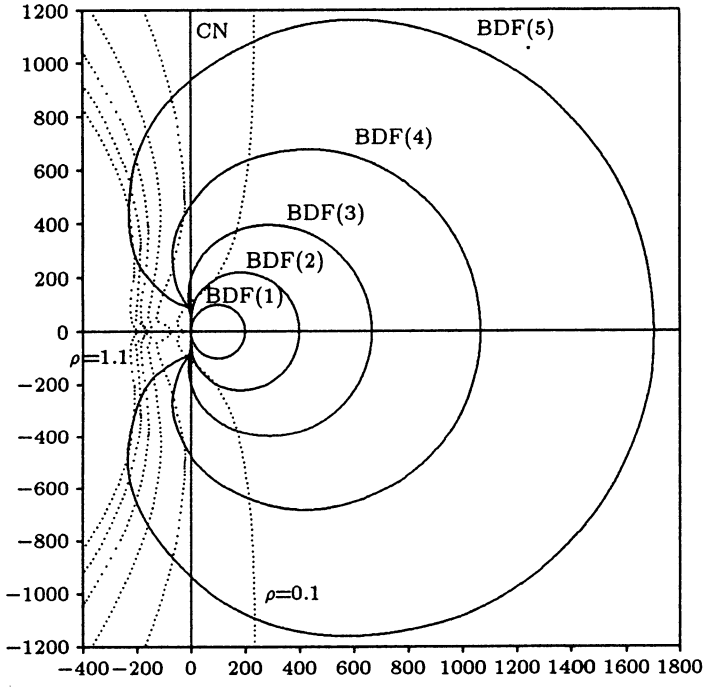

FIG. 6.2. *Graphical convergence test of $\mathcal{M}_\tau$ for model problem (6.1).*

TABLE 6.1
Values of $\rho(\mathcal{K}_\tau)$ for model problem (6.1).

	CN	BDF(1)	BDF(2)	BDF(3)	BDF(4)	BDF(5)
finite length	0.458	0.658	0.548	0.486	0.445	0.414
infinite length	0.962	0.962	0.962	0.976	1.149	1.865

TABLE 6.2
Values of $\rho(\mathcal{M}_\tau)$ for model problem (6.1).

	CN	BDF(1)	BDF(2)	BDF(3)	BDF(4)	BDF(5)
finite length	0.050	0.050	0.052	0.051	0.049	0.047
infinite length	0.264	0.069	0.106	0.170	0.343	1.184

The values of the finite interval spectral radii can be estimated by checking the values of the functions at the points on the real axis given by $\frac{1}{\tau}\frac{\alpha_k}{\beta_k}$ (which are not shown in the picture). With increasing order of the BDF methods, these points move to the right. Indeed, we have that $\frac{\alpha_k}{\beta_k}$ equals 1 (BDF(1)), 3/2 (BDF(2)), 11/6 (BDF(3)), 25/12 (BDF(4)), and 137/60 (BDF(5)). A value of 2 is found for the Crank-Nicolson method.

The values of the infinite interval spectral radii can be estimated by taking the maximum of $\rho(\mathbf{K}(z))$ or $\rho(\mathbf{M}(z))$ over the plotted scaled stability region boundaries. The infinite length discrete-time waveform methods are convergent for the CN method and the low order BDF methods. Divergence is observed for some high order methods. In general, the spectral radius increases with increasing order of the BDF method. This was to be expected from Theorems 4.6 and 5.5, and the knowledge that the BDF methods are $A(\alpha)$-stable with $\alpha = 90°$ (BDF(1), BDF(2)), $\alpha = 88°$ (BDF(3)), $\alpha = 73°$ (BDF(4)) and $\alpha = 51°$ (BDF(5)). Note also that the maximum of $\rho(\mathbf{K}(z))$ over $\frac{1}{\tau}\partial S$ is found at the origin for CN, BDF(1) and BDF(2). Hence, the equality of the corresponding values in Table 6.1.

6.1.2. On the relation between finite time-interval and infinite time-interval spectral radii. We would like to compare some observed experimental convergence factors to the theoretical factors computed in the previous section. More importantly, we want to clarify the relation between the finite time-interval and infinite time-interval spectral radii. To this end, we solve (6.1) using the Gauss-Seidel waveform relaxation method with BDF(2) and BDF(5) time-discretisation and with constant time-step $\tau = 1/100$. (Note that a similar analysis could be done with the two-level method. It would lead to similar conclusions and insights.)

Let $d_\tau^{(\nu)}$ denote the discrete defect or residual in the ν-th iteration. The convergence factor of the ν-th iteration is then defined by

$$\rho^{(\nu)} = ||d_\tau^{(\nu)}||_2 / ||d_\tau^{(\nu-1)}||_2 \,. \tag{6.2}$$

In Figure 6.3 successive convergence factors are plotted for the first 400 waveform Gauss-Seidel iterations, when BDF(2) discretisation is used. These factors appear to remain more or less constant for a large number of iterations. The height of the plateau matches the value obtained in Table 6.1 for infinite time-intervals, i.e., 0.962. (Note that the time-interval in the computation is, of course, finite.) Eventually, the constant plateau in Figure 6.3 is left, and the factors start to decrease. Ultimately, they start to rise again and reach the value 1. This is for purely technical reasons,

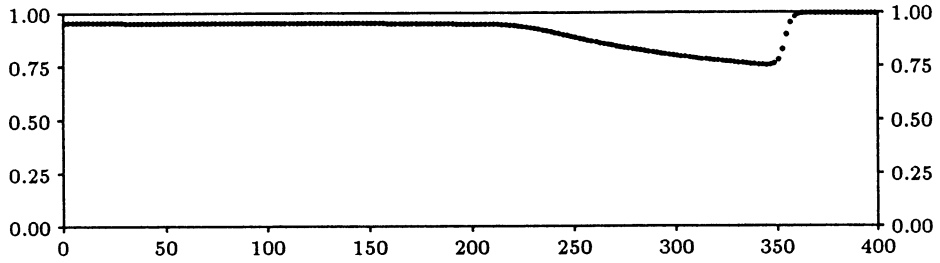

FIG. 6.3. *Convergence factor $\rho^{(\nu)}$ as a function of ν (BDF(2) method).*

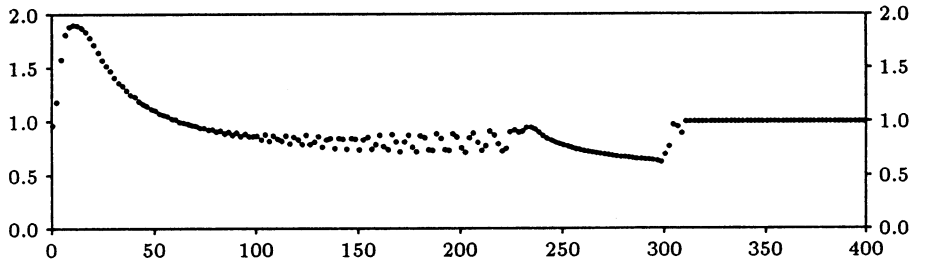

FIG. 6.4. *Convergence factor $\rho^{(\nu)}$ as a function of ν (BDF(5) method).*

because at that time the solution has converged within the finite-precision arithmetic of the implementation.

A similar plot is given in Figure 6.4 for the BDF(5) discretisation. Here, the evolution is much more erratic. The results clearly indicate divergence for a large number of iterations. After sufficient number of iterations, the convergence factors decrease below 1, and the iteration starts to converge rapidly.

This behaviour can be explained by examining the *time-level convergence factors.* These factors are similar to the standard convergence factors (6.2), but are evaluated for each time-level separately,

$$\rho_k^{(\nu)} = ||d_k^{(\nu)}||_2 / ||d_k^{(\nu-1)}||_2 \ .$$

In Figure 6.5, we plotted such time-level convergence factors for the BDF(2) method (for $\nu = 10$, $\nu = 100$, $\nu = 200$ and $\nu = 300$). The factor measured at the first time-level equals 0.548, exactly equal to the value predicted by the finite time-interval analysis in Table 6.1. The convergence factors at the next time-levels increase, and eventually become constant. The height of the plateau matches again the spectral radius value for infinite time-intervals. When more and more iterations are applied, the plateau is forced out of the time-window and the corresponding convergence factors decrease. (Later on, they increase again towards the value 1 because of convergence in finite precision arithmetic.)

In Figure 6.6, we have plotted time-level convergence factors for the BDF(5) method (for $\nu = 1$, $\nu = 5$, $\nu = 50$ and $\nu = 100$). Again, we observe that the factor at the first time-level corresponds to the value predicted by the finite time-interval analysis (0.445). The pictures illustrate the onset of oscillations which rapidly explode. As more iterations are applied, the region of divergent behaviour moves to the right, and is forced out of the time-window. From then on, the iteration converges rapidly, as we observed in Figure 6.4.

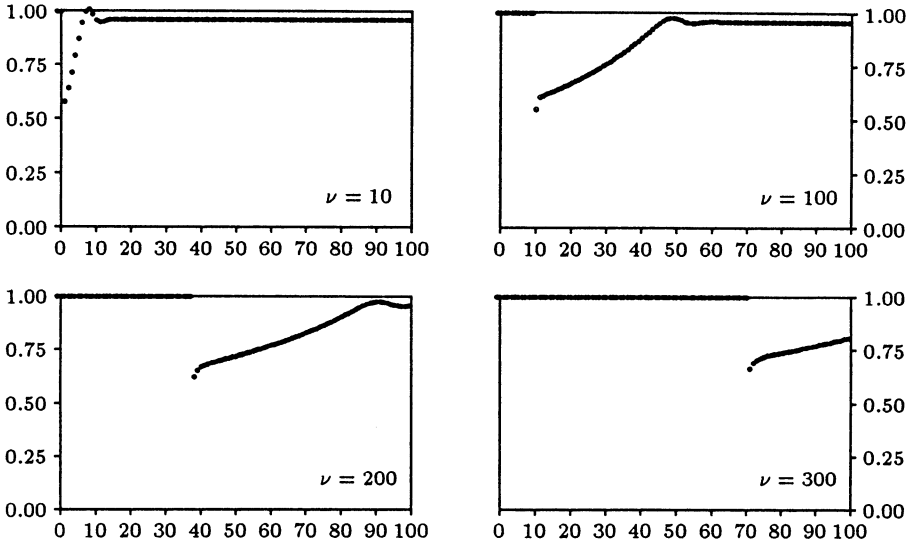

FIG. 6.5. *Time-level convergence factors* $\rho_k^{(\nu)}$ *as a function of k (BDF(2) method)*

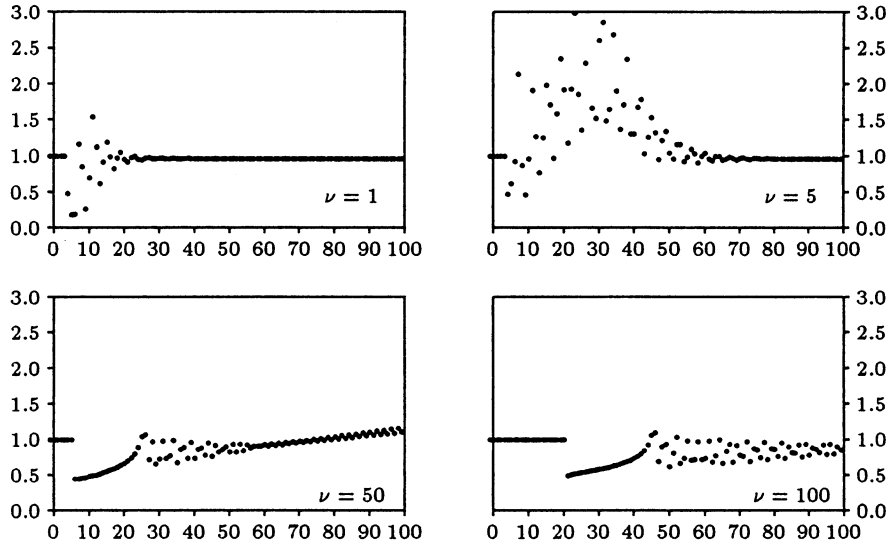

FIG. 6.6. *Time-level convergence factors* $\rho_k^{(\nu)}$ *as a function of k (BDF(5) method)*

6.2. Numerical results. In the current section we shall present numerical results for a model problem that was also considered in [5], i.e., the two-dimensional heat equation,

$$\frac{\partial \mathbf{u}}{\partial t} - \Delta_2 \mathbf{u} = 0 \ , \quad (x,y) \in [0,1] \times [0,1] \ , \quad t \in [0,1] \ , \tag{6.3}$$

with Dirichlet boundary conditions and an initial condition such that the analytical solution equals $\mathbf{u}(x,y,t) = 1 + \sin(\pi x/2)\sin(\pi y/2)\exp(-\pi^2 t/2)$.

Problem (6.3) is discretised using linear basis functions (triangular finite elements) or bilinear basis functions (rectangular finite elements) on a discrete mesh Ω_h

TABLE 6.3
Averaged convergence factors for problem (6.3), linear basis functions, $h = 1/32$, $\tau = 1/200$.

	CN	BDF(1)	BDF(2)	BDF(3)	BDF(4)	BDF(5)
Gauss-Seidel	0.990	0.990	0.990	0.998	-	-
V-cycle	0.443	0.186	0.188	0.194	0.374	-
W-cycle	0.307	0.127	0.125	0.126	0.246	-

TABLE 6.4
Averaged convergence factors for problem (6.3), bilinear basis functions, $h = 1/32$, $\tau = 1/200$.

	CN	BDF(1)	BDF(2)	BDF(3)	BDF(4)	BDF(5)
Gauss-Seidel	0.985	0.985	0.985	0.996	-	-
V-cycle	0.442	0.043	0.133	0.332	-	-
W-cycle	0.295	0.038	0.038	0.040	0.558	-

with equal mesh-size in x-direction and y-direction. The resulting system of ODEs of the form (1.1) is solved using the Gauss-Seidel and the multigrid waveform relaxation methods. In the latter we applied standard V-cycles and W-cycles, with one pre-smoothing and one post-smoothing step of four-colour Gauss-Seidel waveform relaxation type. We use standard coarsening down to a mesh with size $h = 1/2$, seven-point prolongation (linear basis functions) and nine-point prolongation (bilinear basis functions). The restriction operator is defined as $r = p^t$.

In Tables 6.3 and 6.4 we report *averaged convergence factors*. These are defined as the average of $\rho^{(\nu)}$ over the region of nearly constant behaviour (see §6.1.2). The dashes ("-") in the tables indicate that the corresponding method showed divergence over a large number of iterations. Both tables illustrate the dependence of the convergence on the nature of the time-discretisation method. These numerical experiments basically confirm the results of the analysis in §6.1.1.

In Tables 6.5 to 6.10, we report averaged convergence factors of the multigrid W-cycle waveform relaxation method for different values of the mesh-size parameters, and for different discretisation schemes. We observe a dependence of the actual convergence factors on h and τ. Yet, for the Crank-Nicolson and BDF(2) methods, these factors appear to be bounded by a constant, smaller than one, independent of the mesh-size.

For a constant value of h, we expect the averaged convergence factors to converge to the continuous-time results when τ decreases, see §4.3 and §5.3. (In [5], the continuous-time convergence factors are approximated by the results obtained with the CN method with $\tau = 1/1000$). This behaviour is recognised clearly for the CN method, in Tables 6.5 and 6.8. Due to the shape of the stability regions of the BDF(2) and BDF(4) methods, it takes a much smaller value of τ before the discrete-time convergence factors tend to the continuous-time ones, see Tables 6.6, 6.7, 6.9 and 6.10.

For a constant value of τ, we observe an initial increase of the convergence factor when h decreases. For sufficiently small h the convergence factor starts to decrease again. This behaviour is similar to what is observed when the multigrid waveform relaxation method is used to solved the ODEs obtained by spatial finite difference discretisation of a parabolic problem. We refer to [15, §3.5] for an intuitive explanation, and to [16] for a discussion based on an exponential Fourier mode analysis.

TABLE 6.5
Averaged convergence factors for problem (6.3), linear basis functions (CN method).

h, τ	0.04	0.02	0.01	0.005	0.0025	0.001
1/4	0.103	0.135	0.133	0.135	0.134	0.134
1/8	0.121	0.234	0.304	0.303	0.303	0.302
1/16	0.105	0.128	0.282	0.357	0.357	0.355
1/32	0.104	0.105	0.135	0.307	0.370	0.371

TABLE 6.6
Averaged convergence factors for problem (6.3), linear basis functions, (BDF(2) method).

h, τ	0.04	0.02	0.01	0.005	0.0025	0.001
1/4	0.057	0.056	0.056	0.055	0.056	0.056
1/8	0.094	0.098	0.094	0.093	0.093	0.092
1/16	0.123	0.124	0.125	0.126	0.122	0.122
1/32	0.127	0.127	0.127	0.125	0.125	0.126

TABLE 6.7
Averaged convergence factors for problem (6.3), linear basis functions, (BDF(4) method).

h, τ	0.04	0.02	0.01	0.005	0.0025	0.001
1/4	0.054	0.072	0.060	0.057	0.056	0.057
1/8	0.146	0.170	0.338	0.444	0.374	0.269
1/16	0.182	0.236	0.261	0.257	0.475	0.586
1/32	0.127	0.126	0.127	0.246	0.242	0.315

TABLE 6.8
Averaged convergence factors for problem (6.3), bilinear basis functions, (CN method).

h, τ	0.04	0.02	0.01	0.005	0.0025	0.001
1/4	0.074	0.133	0.134	0.137	0.137	0.137
1/8	0.136	0.179	0.283	0.293	0.294	0.294
1/16	0.080	0.179	0.273	0.330	0.341	0.343
1/32	0.038	0.078	0.202	0.295	0.343	0.355

TABLE 6.9
Averaged convergence factors for problem (6.3), bilinear basis functions, (BDF(2) method).

h, τ	0.04	0.02	0.01	0.005	0.0025	0.001
1/4	0.052	0.071	0.084	0.088	0.107	0.124
1/8	0.063	0.090	0.130	0.178	0.222	0.240
1/16	0.046	0.048	0.050	0.104	0.167	0.247
1/32	0.038	0.038	0.038	0.038	0.038	0.128

TABLE 6.10
Averaged convergence factors for problem (6.3), bilinear basis functions, (BDF(4) method).

h, τ	0.04	0.02	0.01	0.005	0.0025	0.001
1/4	0.034	0.072	0.161	0.147	0.139	0.135
1/8	0.091	0.324	0.602	0.659	0.403	0.328
1/16	0.069	0.150	0.383	0.737	0.890	0.769
1/32	0.038	0.038	0.112	0.558	0.645	0.935

REFERENCES

[1] P. BRENNER, V. THOMÉE, AND L. B. WAHLBIN, *Besov Spaces and Applications to Difference Methods for Initial Value Problems*, vol. 434 of Springer Lecture Notes in Mathematics, Springer Verlag, Berlin, 1975.

[2] W. HACKBUSCH, *Multi-Grid Methods and Applications*, vol. 4 of Springer Series in Computational Mathematics, Springer Verlag, Berlin, 1985.

[3] E. HAIRER AND G. WANNER, *Solving Ordinary Differential Equations II*, vol. 14 of Springer Series in Computational Mathematics, Springer Verlag, Berlin, 1991.

[4] G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, *Inequalities*, University Press, Cambridge, 2nd ed., 1978.

[5] J. JANSSEN AND S. VANDEWALLE, *Multigrid waveform relaxation on spatial finite element meshes: the continuous-time case*, Tech. Report TW 201, Department of Computer Science, Katholieke Universiteit Leuven, November 1993. (accepted for publication in SIAM J. Numer. Anal.)

[6] J. D. LAMBERT, *Computational Methods in Ordinary Differential Equations*, John Wiley and Sons, Chicester, 1973.

[7] 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, 1 (1982), pp. 131–145.

[8] C. LUBICH, *On the stability of linear multistep methods for Volterra convolution equations*, IMA J. Numer. Anal., 3 (1983), pp. 439–465.

[9] C. LUBICH AND A. OSTERMANN, *Multi-grid dynamic iteration for parabolic equations*, BIT, 27 (1987), pp. 216–234.

[10] U. MIEKKALA AND O. NEVANLINNA, *Convergence of dynamic iteration methods for initial value problems*, SIAM J. Sci. Stat. Comput., 8 (1987), pp. 459–482.

[11] ———, *Sets of convergence and stability regions*, BIT, 27 (1987), pp. 554–584.

[12] O. NEVANLINNA, *Remarks on Picard-Lindelöf iteration, PART II*, BIT, 29 (1989), pp. 535–562.

[13] L. REICHEL AND L. N. TREFETHEN, *Eigenvalues and pseudo-eigenvalues of Toeplitz matrices*, Linear Algebra Appl., 162–164 (1992), pp. 153–185.

[14] V. THOMÉE, *Galerkin Finite Element Methods for Parabolic Problems*, vol. 1054 of Springer Lecture Notes in Mathematics, Springer Verlag, Berlin, 1984.

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

[16] S. VANDEWALLE AND G. HORTON, *Fourier mode analysis of the multigrid waveform relaxation and time-parallel multigrid methods*, Tech. Report IMMD 3, 7/93, Universität Erlangen-Nürnberg, Martensstrasse 3, D-91058 Erlangen, Germany, July 1993.

[17] P. WESSELING, *An Introduction to Multigrid Methods*, John Wiley and Sons, Chicester, 1992.

[18] K. YOSIDA, *Functional Analysis*, Springer-Verlag, New York, 1980.