MATHEMATICS OF COMPUTATION
VOLUME 60, NUMBER 202
APRIL 1993, PAGES 473–493

FOURIER ANALYSIS OF MULTIGRID METHODS FOR GENERAL SYSTEMS OF PDES

PER LÖTSTEDT AND BERTIL GUSTAFSSON

ABSTRACT. Most iteration methods for solving boundary value problems can be viewed as approximations of a time-dependent differential equation. In this paper we show that the multigrid method has the effect of increasing the time-step for the smooth part of the solution leading back to an increase of the convergence rate. For the nonsmooth part the convergence is an effect of damping. Fourier analysis is used to find the relation between the convergence rate for multigrid methods and singlegrid methods. The analysis is performed for general partial differential equations and an arbitrary number of grids. The difference in the behavior of the iterations between first- and second-order equations is discussed. The theoretical results are confirmed in simple numerical experiments.

1. INTRODUCTION

The convergence analysis of multigrid methods for solving numerical approximations to partial differential equations is usually based on the assumption that the problem is solved exactly on the coarsest grid. In this way high convergence rates are often predicted, at least for elliptic model problems. The situation is different for large-scale real-life problems, where the geometry and the structure of the grid is such that a grid coarse enough to permit exact solutions is never reached. Instead, the smoothing operator (for example Jacobi, Gauss-Seidel, conjugate gradient, Runge-Kutta iteration) applied on the finer grids is also used on the coarsest grid, and the number of grids is usually low, typically two, three or four. The convergence rates observed for this kind of computations are often lower than the ones predicted by too simplified model problems.

The traditional way of performing convergence analysis is to estimate the magnitude of the eigenvalues of the iteration matrix M. For equations of simple structure, boundary conditions can sometimes be included, since the set of eigenvectors can be derived, which in turn permits the calculation of eigenvalues. However, the analysis is in most cases based on Fourier modes (see, for example, [1], [16], [18]), which means that the solutions are assumed to be periodic in space or that the domain is unbounded. If ξ is the wave number, h is the fine-grid stepsize, the differential equation has no lower-order terms, and $\widehat{M}(h\xi)$ is the symbol of M, there is always one eigenvalue $\lambda(h\xi)$ of $\widehat{M}(h\xi)$ with $\lambda(0) = 1$. However, under the assumption that an exact solution can be

Received by the editor February 19, 1991 and, in revised form, March 16, 1992.
1991 *Mathematics Subject Classification.* Primary 65N22, 65F10.

obtained on the coarsest grid, it is often possible to show that the magnitude of the eigenvalues are uniformly bounded by $1-\delta$ for $h\xi > 0$, where $\delta > 0$ is independent of $h\xi$. If one assumes that the constant mode corresponding $\xi = 0$ is not present, the number $1-\delta$ is then given as the convergence factor.

If the fine-grid smoother is used also on the coarse grids, the eigenvalues of the modified symbol $\widehat{M}$ are continuous functions of $h\xi$. This means that the convergence rate is arbitrarily low for small values of $h\xi$. The interpretation is that the iteration procedure is closely connected to a time-dependent differential equation. This was illustrated in [5] for a simple model problem. It was shown that for low wave numbers the two-grid procedure used plays the role of scaling up the time-variable, compared to what it would have been for the single-grid method. For first-order equations this means that the long waves move faster, but there is very little damping. On a finite domain fast convergence can still be obtained because the long waves move quickly out of the domain. The important fact is that both the wave propagation properties and the damping properties must be taken into account when constructing multigrid methods for first-order systems.

In this paper we first prove that under very natural conditions multigrid methods are consistent with a time-dependent differential equation where the time variable is scaled up compared to the corresponding singlegrid method. This was conjectured by Jesperson [11], see also [10]. However, the properties of the time-dependent differential equation obtained in this way can be used only for the smooth part of the solution. The remaining part of the discrete solution is completely independent of the differential equation. In our analysis we take into account the interaction between the two parts. We use Fourier analysis to derive precise results concerning the behavior of the low and the high wave number parts of the solution. The analysis is carried out for discretizations of general constant-coefficient differential equations of arbitrary order for a V-cycle on an arbitrary number of grids in two space dimensions.

Based on the results of this paper, grid-independent convergence is proved for systems of first-order PDEs in [14]. The usefulness of Fourier analysis in predicting convergence rates for such problems has been demonstrated recently in [3]. An early and short version of this paper is [6].

Our approach of introducing a time-variable is only for the analysis, and it has been used for other iteration methods, see e.g. [4]. We emphasize that our aim is not to solve time-dependent problems. Attempts have been made to use the multigrid technique to speed up the calculation also for that kind of problem, see e.g. [11]. However, in that case the solution must be accurate also on the coarse grids. Since the fine structure in the solution can be represented only on the fine grids, the true time-dependent behavior can never be obtained on the coarser ones where larger time-steps are used. Therefore, the multigrid technique is useful only if the finest grid is unnecessarily fine for some reason.

2. Consistency with a time-dependent system

In this section we shall prove that a full multigrid iteration is consistent with a time-dependent differential equation where the time-variable is scaled compared to the equation which corresponds to a singlegrid iteration. We begin by giving the notation.

We shall use $L+1$ grids $\{G_l\}_{l=0}^{L}$, where G_L is the finest one. For convenience it is assumed that the stepsize h_l on grid G_l is equal in all directions. Q_l is an ordinary finite difference or finite volume approximation of the linear differential operator

$$P = \sum_{\nu} A_\nu(x) \left(\frac{\partial}{\partial x^1}\right)^{\nu_1} \cdots \left(\frac{\partial}{\partial x^d}\right)^{\nu_d}, \tag{2.1}$$

$$A_\nu \in \mathbf{R}^s \times \mathbf{R}^s, \quad \nu = (\nu_1, \ldots, \nu_d), \quad x = (x^1, \ldots, x^d)^T,$$

on the grid G_l, $l = 0, 1, \ldots, L$. The matrices A_ν are assumed to be sufficiently smooth. We seek the solution to

$$Q_L u = f, \tag{2.2}$$

where u and f are vector functions with s components. The restriction operator from G_l to G_{l-1} is r_l, and the prolongation operator from G_{l-1} to G_l is p_l, $l = 1, 2, \ldots, L$. On each grid G_l there is an iterative method

$$R_l(u, f) \to u,$$

which is applied p times before and q times after the coarse-grid corrections. On the coarsest grid we use $p+q$ iterations. The complete multigrid V-cycle is defined by (see [7, §4.1])

```
procedure MG(l, u, f)
if l = 0 then for j := 1 step 1 until p + q do u := R_0(u, f)
else
begin
        for j := 1 step 1 until p do u := R_l(u, f);
            d := r(Q_l u - f);
            v := 0;
            MG(l - 1, v, d);
            u := u - pv;
        for j := 1 step 1 until q do u := R_l(u, f);
end;
    u := u^n;
    MG(L, u, f);
    u^{n+1} := u;
```

We write the iteration operator as

$$R_l(u, f) = S_l u + T_l f, \tag{2.3}$$

where consistency requires

$$T_l Q_l = I - S_l. \tag{2.4}$$

When analyzing the error and its convergence to zero, it is sufficient to consider the case $f = 0$. Let n be the iteration index. Then the multigrid V-cycle can be written as, cf. [7, Lemma 7.1.4],

$$\begin{aligned} M_{-1} &= I, \\ M_l &= S_l^q (I - p_l (I - M_{l-1}) Q_{l-1}^{-1} r_l Q_l) S_l^p, \qquad l = 0, 1, \ldots, L, \\ u^{n+1} &= M_L u^n. \end{aligned} \tag{2.5}$$

(If the exact solution is computed on G_0, then $M_0 = 0$ in the recursion (2.5) for $l \geq 1$.)

From now on we use the notation $u^{(l)}$ for a grid function defined on G_l. Consider the time-dependent problems

$$\frac{\partial u^{(l)}}{\partial t} + Q_l u^{(l)} = f^{(l)}, \qquad l = 0, 1, \ldots, L, \tag{2.6}$$

and introduce the time steps Δt_l, $l = 0, 1, \ldots, L$. We also use the notation

$$\Delta t = \Delta t_L, \qquad h = h_L, \qquad \alpha_l = \frac{\Delta t_l}{\Delta t}, \quad l = 0, 1, \ldots, L.$$

The idea is to consider the iteration $R_l(u^{(l)}, f^{(l)})$ as one time step Δt_l in a solution procedure of (2.6). If the whole multigrid cycle (2.5) is considered as one time step Δt, we want to relate it to the time-dependent problem (2.6) with $l = L$.

When calling a function u smooth on a certain grid G_l, we mean that the divided differences $D_+^\nu u_j$ are bounded on that grid.

In all of the following assumptions, u denotes a smooth function, $u \in S$, and g denotes a bounded function, $g \in B$.

Assumption 2.1 (assumptions on Q_l).

(i) Q_l is consistent with P, $0 \leq l \leq L$, i.e.,

$$Q_l u = Pu + h_l g.$$

(ii) $\Delta t_l Q_l$ is a bounded operator.

If Q_l is an ordinary difference operator, then the condition 2.1(ii) implies

$$\Delta t_l \leq \text{const.}\, h_l^{\max(\nu_1+\nu_2+\cdots+\nu_d)}. \tag{2.7}$$

(For first-order systems: $\Delta t_l \leq \text{const.}\, h_l$.) Also by (2.1), Pu is a smooth function.

Assumption 2.2 (assumptions on S_l).

(i) $S_l u = (I - \Delta t_l Q_l)u + \Delta t_l h_l g$, $0 \leq l \leq L$.

(ii) S_l is a bounded operator.

Assumption 2.3.

(i) $Q_{l-1}^{-1} r_l Q_l S_l u = v + h_l \Delta t_l g$, $1 \leq l \leq L$, where $v \in S$.

(ii) $Q_{l-1}^{-1} r_l Q_l$ is a bounded operator.

Assumption 2.4 (assumptions on p_l, r_l).

(i) $p_l r_l u = u + h_l g$, $1 \leq l \leq L$.

(ii) p_l, r_l are bounded operators.

Theorem 2.1. *If Assumptions* 2.1–2.4 *hold, and if the multigrid iteration* (2.5) *is considered as one time step* Δt *in a time-dependent procedure, then it is consistent with*

$$\frac{\partial u}{\partial t} + (p+q)\sum_{l=0}^{L} \alpha_l P u = 0. \tag{2.8}$$

For $p = 1$, $q = 0$, $\alpha_l = 2^{L-l}$, $l = 0, 1, \ldots, L$, *it is consistent with*

$$\frac{\partial u}{\partial t} + (2^{L+1} - 1)Pu = 0.$$

Proof. The major part of the proof is found in Lemma A.2 of Appendix A in the Supplement section, where an expression for $M_L u$ is derived. In general, $M_L u$ cannot be expected to be smooth, since p_l may be such that it returns a nonsmooth function, even if it is applied to a smooth one. However, the nonsmooth part is of $O(h\Delta t)$.

It follows from (2.5) and Lemma A.2 that with u^n smooth,

$$u^{n+1} = M_L u^n = \left(I - (p+q)\Delta t \sum_{l=0}^{L} \alpha_l Q_L \right) u^n + h\Delta t g, \tag{2.9}$$

where g is bounded.

Consider now a smooth solution $u(x, t)$ of the differential equation (2.8) substituted into the iteration formula (2.5). The truncation error τ is defined by

$$\Delta t \tau(x, t) = u(x, t + \Delta t) - M_L u(x, t),$$

and consistency requires that $\tau(x, t) \to 0$ as $\Delta t \to 0$, $h \to 0$. We have by (2.9) and Assumption 2.1(i),

$$\begin{aligned} \tau(x, t) &= \frac{u(x, t+\Delta t) - u(x, t)}{\Delta t} + (p+q) \sum_{l=0}^{L} \alpha_l Q_L u(x, t) - hg(x, t) \\ &= \frac{\partial u}{\partial t}(x, t) + O(\Delta t) + (p+q) \sum_{l=0}^{L} \alpha_l P u(x, t) + O(h) - hg(x, t) \\ &= O(\Delta t + h), \end{aligned}$$

which proves the theorem. □

The theorem shows that on a fixed number of grids the iteration formula converges to the modified time-dependent equation (2.8) as $\Delta t \to 0$. In practice this means that for first-order systems we can expect the waves corresponding to low wave numbers to move $(\sum_{l=0}^{L} \alpha_l)$ times faster by using the multigrid procedure instead of a singlegrid solver. Alternatively, we can of course consider the procedure as an increase in the time step for the original system (2.6), and this interpretation applies to all types of operators P. Note that we have assumed consistency also on the coarsest grid G_0. The practical implication of this is that G_0 must be fine enough such that the low-frequency part of the solution can be represented. If there are only two points, say, in each direction of G_0, the theorem has no meaning. This does not mean that one should avoid very coarse grids if the geometry of the computational domain permits it. On the contrary, it may accelerate the convergence as a result of stronger damping.

Instead of (2.6) we could of course consider the more general systems

$$\frac{\partial u^{(l)}}{\partial t} + D_l Q_l u^{(l)} = D_l f^{(l)},$$

where D_l are nonsingular operators. But this is just a preconditioning of the original system, e.g., with "local time-stepping" or residual smoothing [10], [2]. Let the preconditioner D_l be included in Q_l everywhere, also when the residual is determined. Then, with minor modifications of Assumption 2.1, a theorem

similar to Theorem 2.1 can be proved when Q_l and P are not necessarily difference and differential operators.

3. Fourier analysis of the multigrid cycle

In the previous section we have analyzed the multigrid method in a general case with smooth solutions by considering the multigrid procedure as an iteration forward in time (or pseudotime). In this section we present a complete convergence analysis, where we take also the nonsmooth part of the solution into account. We restrict the analysis to the constant-coefficient case in order to be able to use Fourier analysis as our main tool of investigation. The results are less general here, but more details about the convergence are revealed. For the sake of notational simplicity the analysis is carried out for two space dimensions, but the results generalize to any finite number d of dimensions. All variables are associated with level l except when the level is explicitly written as a subscript or superscript on the variable. The norm in what follows is the Euclidean vector norm and the subordinate spectral matrix norm.

The main result of the analysis is that two effects are responsible for the convergence: the amplification of the time scale for low wave number modes and damping of intermediate and high wave number modes.

3.1. **Fourier representation.** In the analysis we need a Fourier representation of the solution

$$u(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\xi\cdot x}\hat{u}(\xi)\,d\xi_1\,d\xi_2\,, \qquad x = (x^1, x^2)^T,\ \xi = (\xi_1, \xi_2)^T.$$

We are interested in the solution of the Cauchy problem at discrete points $x_{\mu\nu}$ on level l,

$$x_{\mu\nu} = x_0 + h\begin{pmatrix}\mu\\ \nu\end{pmatrix}, \qquad (\mu, \nu) \in \mathbf{Z}\times\mathbf{Z},\ \mathbf{Z} = \{\text{the integer numbers}\}. \tag{3.1}$$

Then $u^{(l)}(x_{\mu\nu})$ can be written

$$\begin{aligned} u_{\mu\nu} &= u^{(l)}(x_{\mu\nu}) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(i(\xi_1(x_0^1 + \mu h) + \xi_2(x_0^2 + \nu h)))\hat{u}(\xi)\,d\xi_1\,d\xi_2\,. \end{aligned} \tag{3.2}$$

Replace the integral over ξ_1 in (3.2) by

$$\int_{-\infty}^{\infty} I(\xi_1)\,d\xi_1 = \sum_{j\in\mathbf{Z}}\int_{j\cdot 2\pi/h}^{(j+1)\cdot 2\pi/h} I(\xi_1)\,d\xi_1 = \sum_{j\in\mathbf{Z}}\int_0^{2\pi/h} I(\xi_1' + j\cdot 2\pi/h)\,d\xi_1'\,,$$

where

$$\xi_1 = \xi_1' + j\cdot 2\pi/h\,, \quad j\in\mathbf{Z}, \quad \xi_1'\in[0, 2\pi/h].$$

The integral in (3.2) over ξ_2 is rewritten in the same manner and ξ_2 is substituted by

$$\xi_2 = \xi_2' + k\cdot 2\pi/h\,, \quad k\in\mathbf{Z}, \quad \xi_2'\in[0, 2\pi/h].$$

After simplification and introduction of

$$\hat{u}'(\xi') = \sum_{k,j\in\mathbf{Z}} \exp(i\xi\cdot x_0)\hat{u}(\xi_1' + j\cdot 2\pi/h, \xi_2' + k\cdot 2\pi/h),$$

the expression for $u_{\mu\nu}$ is

$$u_{\mu\nu} = \int_{C_l} \exp(i(\xi_1'\mu + \xi_2'\nu)h)\hat{u}'(\xi')\,d\xi', \tag{3.3}$$

$$d\xi' = d\xi_1'\,d\xi_2', \quad C_l = [0, 2\pi/h_l] \times [0, 2\pi/h_l].$$

The integral (3.3) is split once more into a sum of integrals,

$$\int_{C_l} I(\xi_1')\,d\xi_1' = \sum_{j=0}^{m-1} \int_{j\cdot 2\pi/h_0}^{(j+1)\cdot 2\pi/h_0} I(\xi_1')\,d\xi_1' = \sum_{j=0}^{m-1} \int_0^{2\pi/h_0} I(\xi_1'' + j\cdot 2\pi/h_0)\,d\xi_1'',$$

where

$$\xi_1' = \xi_1'' + j\cdot 2\pi/h_0, \quad m = 2^l, \quad h_l = h_0/m.$$

With a similar treatment of the ξ_2'-variable the integral in (3.3) takes the form

(3.4)

$$u_{\mu\nu} = \int_{C_0} \sum_{k=0}^{m-1}\sum_{j=0}^{m-1} \exp(i(\xi_1''\mu h + j\cdot 2\pi\mu/m))$$
$$\cdot \exp(i(\xi_2''\nu h + k\cdot 2\pi\nu/m))\hat{u}'(\xi_1'' + j\cdot 2\pi/h_0, \xi_2'' + k\cdot 2\pi/h_0)\,d\xi''.$$

Henceforth, we drop the primes on ξ and $\hat{u}$.

In the analysis of the multigrid iteration in one space dimension, two wave numbers $\xi + j\cdot 2\pi/h_0$ and $\xi + (j + m/2)\cdot 2\pi/h_0$, $j = 0, 1, \ldots, m/2 - 1$, on a grid l correspond to one wave number $\xi + j\cdot 2\pi/h_0$ on the next coarser grid $l-1$ [5]. In d space dimensions, 2^d wave numbers on mesh l are reduced to one wave number on the next coarser grid $l-1$ by "aliasing" in the restriction process. In two space dimensions it is natural to treat the wave numbers in groups of four. In Figure 3.1 the wave numbers in C_l are mapped on C_{l-1} when we restrict a solution on grid G_l to grid G_{l-1}. In the prolongation process from level $l-1$ to l the wave number domain C_{l-1} is expanded to C_l.

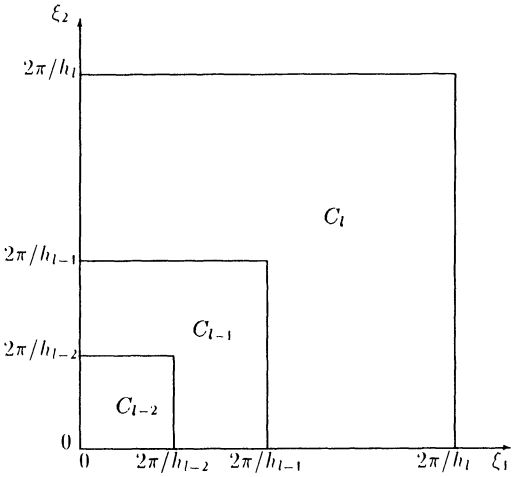

FIGURE 3.1. The definitions of the wave number domains C_l, C_{l-1}, and C_{l-2}

The notation is simplified if at $x_{\mu\nu}$ and level l we let

$$e_j^1 = \exp(i(\xi_1 + j \cdot 2\pi/h_0)x_{\mu\nu}^1),$$
$$e_k^2 = \exp(i(\xi_2 + k \cdot 2\pi/h_0)x_{\mu\nu}^2),$$
$$\hat{u}_{jk} = \hat{u}(\xi_1 + j \cdot 2\pi/h_0, \xi_2 + k \cdot 2\pi/h_0).$$

For wave numbers ξ_1 such that $0 \le \xi_1 \le \xi_*$, where ξ_* is small and $x_\mu^1 = \mu h$,

$$e_0^1 = \exp(i\xi_1 \mu h)$$

has a slow variation with μ. This is also true for e_{m-1}^1 with $\xi_1 = 2\pi/h_0 - \xi_1'$, $0 \le \xi_1' \le \xi_*$,

$$e_{m-1}^1 = \exp(i(\xi_1 + (m-1) \cdot 2\pi/h_0)\mu h) = \exp(-i\xi_1' \mu h). \tag{3.5}$$

The function $e_{m/2}^1$ with ξ_1 small is highly oscillatory, since

$$e_{m/2}^1 = \exp(i(\xi_1 + m\pi/h_0)\mu h) = \exp(i\xi_1 \mu h)(-1)^\mu.$$

The basis function e_k^2 has the same properties. Let

$$\begin{aligned} E_{jk} &= (e_j^1, e_{j+m/2}^1) \otimes (e_k^2, e_{k+m/2}^2) \otimes I_s \in \mathbf{C}^s \times \mathbf{C}^{4s}, \\ U_{jk} &= (\hat{u}_{jk}^T, \hat{u}_{j,k+m/2}^T, \hat{u}_{j+m/2,k}^T, \hat{u}_{j+m/2,k+m/2}^T)^T \in \mathbf{C}^{4s}. \end{aligned} \tag{3.6}$$

The Kronecker product is denoted by $\otimes$ and is defined in [13], and I_s is the identity matrix of dimension s. The number of unknown variables at each grid point is s as defined in §2. The array E_{jk} in (3.6) consists of the four Fourier basis functions that coalesce into the basis function $e_j^1 e_k^2$ on level $l-1$. The Fourier coefficients associated with the wave numbers in E_{jk} are stored in U_{jk}.

Finally, define

$$\mathbf{E}_{\mu\nu} = \mathbf{E}(x_{\mu\nu}) = (E_{00}, \ldots, E_{jk}, \ldots), \qquad \tilde{u} = ((U_{00})^T, \ldots, (U_{jk})^T, \ldots)^T.$$

The exact order of the components in $\mathbf{E}$ and $\tilde{u}$ will be determined later. Then (3.4) can be rewritten

$$u_{\mu\nu} = \int_{C_0} \mathbf{E}_{\mu\nu} \tilde{u}\, d\xi. \tag{3.7}$$

The difference operator Q is described in §2. Here we assume that A_ν in (2.1) is constant. The symbol of Q is denoted by $\hat{Q}$, and

$$\hat{Q}_{jk}(\xi_1 h, \xi_2 h, h) = \hat{Q}((\xi_1 + j \cdot 2\pi/h_0)h, (\xi_2 + k \cdot 2\pi/h_0)h, h) \in \mathbf{C}^s \times \mathbf{C}^s. \tag{3.8}$$

Four of the matrices $\hat{Q}_{jk}$ are collected in a block diagonal matrix $\tilde{Q}_{jk}$ such that the blocks on the diagonal correspond to the wave numbers in E_{jk}. Then create a matrix $\tilde{Q}$ of dimension sm^2,

$$\tilde{Q} = \operatorname{diag}(\tilde{Q}_{jk}), \quad \tilde{Q}_{jk} \in \mathbf{C}^{4s} \times \mathbf{C}^{4s}, \quad j = 0, 1, \ldots, m/2-1, \quad k = 0, 1, \ldots, m/2-1.$$

When Q operates on $u_{\mu\nu}$ in (3.7), we arrive at

$$Qu_{\mu\nu} = \int_{C_0} Q\mathbf{E}_{\mu\nu} \tilde{u}\, d\xi = \int_{C_0} \mathbf{E}_{\mu\nu} \tilde{\mathbf{Q}} \tilde{u}\, d\xi_1. \tag{3.9}$$

We now turn to the matrix structure of the symbol of the restriction operator r. Assume that r can be written as

$$r = r_1 \otimes r_2 \otimes I_s . \tag{3.10}$$

The restriction in the x^i-direction is denoted by r_i in (3.10). As an example, take $s = 1$ and

$$r_i = \tfrac{1}{4}E_i^{-1} + \tfrac{1}{2} + \tfrac{1}{4}E_i , \qquad i = 1, 2,$$

where E_i is the shift operator in the x^i-direction. Then the stencil of r in (3.10) is

$$\left(\tfrac{1}{4} \quad \tfrac{1}{2} \quad \tfrac{1}{4}\right) \otimes \begin{pmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{pmatrix} \otimes 1 = \begin{pmatrix} \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \end{pmatrix} .$$

The restriction of E_{jk} in (3.6) is

$$rE_{jk} = (r_1 e_j^1 , r_1 e_{j+m/2}^1) \otimes (r_2 e_k^2 , r_2 e_{k+m/2}^2) \otimes I_s = e_j^1 e_k^2 \tilde{R}_{jk} , \tag{3.11}$$

where

$$\tilde{R}_{jk} = (\hat{r}_{1j} , g_1 \hat{r}_{1, j+m/2}) \otimes (\hat{r}_{2k} , g_2 \hat{r}_{2, k+m/2}) \otimes I_s \in \mathbf{C}^s \times \mathbf{C}^{4s} ,$$

and

$$\hat{r}_{ij} = \hat{r}_i((\xi_i + j \cdot 2\pi/h_0)h) , \qquad i = 1, 2,$$

is the symbol of the one-dimensional restriction r_i, and g_i is a constant such that $g_i e_j^i = e_{j+m/2}^i$. In the above example,

$$\begin{aligned} \hat{r}_{ij} &= \cos^2((\xi_i + j \cdot 2\pi/h_0)h/2) , \\ \hat{r}_{i, j+m/2} &= \sin^2((\xi_i + j \cdot 2\pi/h_0)h/2) , \\ g_i &= 1 \quad \text{on a mesh } x_{\mu\nu} = (\mu h , \nu h) . \end{aligned} \tag{3.12}$$

Collect the submatrices $\tilde{R}_{jk}$ along the diagonal of $\tilde{R} \in \mathbf{C}^{sm^2/4} \times \mathbf{C}^{sm^2}$. If the wave numbers are ordered properly, then the conclusion from (3.11) is that

$$\begin{aligned} &r(E_{jk} , E_{j, k+m/4} , E_{j+m/4, k} , E_{j+m/4, k+m/4}) \\ &\quad = (e_j^1 e_k^2 \tilde{R}_{jk} , e_j^1 e_{k+m/4}^2 \tilde{R}_{j, k+m/4} , e_{j+m/4}^1 e_k^2 \tilde{R}_{j+m/4, k} , \\ &\qquad\qquad e_{j+m/4}^1 e_{k+m/4}^2 \tilde{R}_{j+m/4, k+m/4}) \\ &\quad = E_{jk}^{l-1} \begin{pmatrix} \tilde{R}_{jk} & & & 0 \\ & \tilde{R}_{j, k+m/4} & & \\ & & \tilde{R}_{j+m/4, k} & \\ 0 & & & \tilde{R}_{j+m/4, k+m/4} \end{pmatrix} , \\ &\qquad j = 0, 1, \ldots, m/4 - 1 , \quad k = 0, 1, \ldots, m/4 - 1 . \end{aligned} \tag{3.13}$$

The last matrix in (3.13) is a $4s \times 16s$ submatrix on the diagonal of $\tilde{R}$. The restriction of $u_{\mu\nu}$ in (3.7) is now

$$ru_{\mu\nu} = \int_{C_0} r\mathbf{E}_{\mu\nu} \tilde{u} \, d\xi = \int_{C_0} \mathbf{E}_{\mu\nu}^{l-1} \tilde{R} \tilde{u} \, d\xi . \tag{3.14}$$

The simple structure of $\tilde{R}$ relies on the order of the wave numbers in $\mathbf{E}_{\mu\nu}$. Suppose that the order is the suitable one on level $l - 1$ in $\mathbf{E}_{\mu\nu}^{l-1}$. Then it

follows from (3.13) how the wave numbers shall be grouped on level l in $\mathbf{E}_{\mu\nu}$. At the coarsest level $l = 0$, $e_0^1 e_0^2$ is the only component in $\mathbf{E}_{\mu\nu}^0$ and therefore always in the first position. In this way the order of the indices of E_{jk}^l in $\mathbf{E}_{\mu\nu}$ is determined recursively beginning at $l = 0$. The result is that the first matrix E_{jk}^l in $\mathbf{E}_{\mu\nu}$ at every level $l > 0$ has index $j = k = 0$.

The structure of the symbol $\tilde{P}$ of the prolongation operator p from level $l-1$ to level l is determined by an analysis similar to the analysis of the restriction. The matrices $\tilde{P}$ and $\tilde{R}^T$ possess the same zero pattern outside the diagonal blocks.

In the sequel we assume that the symbol matrix $\tilde{S}$ of the smoothing operator S has the same block diagonal structure as the $\tilde{Q}$-matrix. This is, e.g., the case if we use Runge-Kutta time-stepping in the smoothing iterations as in [9]. A reader familiar with numerical linear algebra would call this smoothing scheme Richardson iteration [19].

The definitions and discussion concerning the two-dimensional problem in this subsection are easily reduced to one space dimension or extended to the three-dimensional case.

3.2. **Analysis of the iteration matrix.** In this subsection the properties of the Fourier transform $\widehat{M}$ of the multigrid iteration matrix M in (2.5) are derived. The results hold for differential equations of arbitrary order. It follows from (3.9) and (3.14) that

$$rQu_{\mu\nu} = \int_{C_0} \mathbf{E}_{\mu\nu}^{l-1} \tilde{R}\tilde{Q}\tilde{u}\, d\xi .$$

The successive application of operators to $u_{\mu\nu}$ corresponds to multiplication of the symbol matrices inside the integral. Hence, for the multigrid operator M performing one V-cycle,

$$Mu_{\mu\nu} = \int_{C_0} \mathbf{E}_{\mu\nu} \widetilde{M}\tilde{u}\, d\xi , \tag{3.15}$$

where

$$\widetilde{M} = \tilde{S}^q (I - \tilde{P}(I - \widetilde{M}_{l-1})(\tilde{Q}_{l-1})^{-1}\tilde{R}\tilde{Q})\tilde{S}^p , \tag{3.16}$$

cf. (2.5). The transform matrices $\tilde{R}$, $\tilde{Q}$, $\tilde{P}$, and $\tilde{S}$ are defined and discussed in §3.1.

Let the superscript i on $\tilde{u}^i$ denote the number of the iteration as in (2.5). By (3.15) we find that

$$\tilde{u}^n = \widetilde{M}^n \tilde{u}^0 . \tag{3.17}$$

The convergence of the iteration depends critically on the behavior of $\widetilde{M}^n$. We analyze this matrix in the theorem of this subsection.

First we need some additional notation and state a few assumptions. The symbol of the difference operator $\tilde{Q}_{jk}$ is defined in (3.8). The restriction and prolongation symbols are functions of

$$\eta_l = \begin{pmatrix} \eta_{1l} \\ \eta_{2l} \end{pmatrix} = h_l \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} .$$

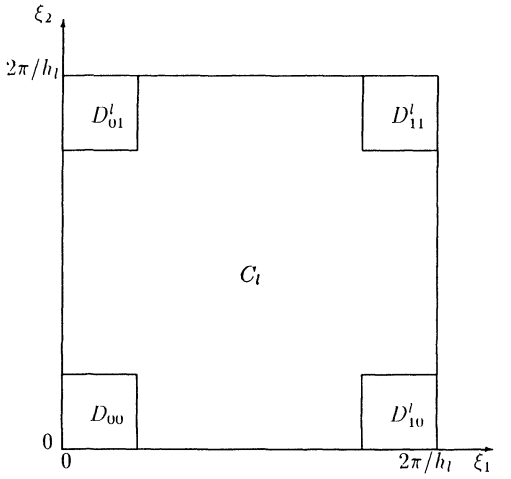

FIGURE 3.2. The definition of $C_l^* = D_{00} \cup D_{01}^l \cup D_{10}^l \cup D_{11}^l$ as a subset of C_l

We are interested in the properties of certain submatrices of $\widetilde{M}$. Of special interest are the wave numbers of modes with slow variation and the corresponding submatrix of $\widetilde{M}$. On each grid these wave numbers are contained in C_l^* defined as follows:

$$
\begin{aligned}
&D_0 = [0, \xi_*], \; D_1^l = [2\pi/h_l - \xi_*, 2\pi/h_l], \\
&D_{00} = D_0 \times D_0, \qquad D_{01}^l = D_0 \times D_1^l, \\
&D_{10}^l = D_1^l \times D_0, \qquad D_{11}^l = D_1^l \times D_1^l, \\
&C_l^* = D_{00} \cup D_{01}^l \cup D_{10}^l \cup D_{11}^l .
\end{aligned} \tag{3.18}
$$

In order to have good resolution on the coarsest grid, we take $\xi_* h_0/2\pi \ll 1$. In Figure 3.2 the definitions in (3.18) are illustrated.

If $\xi_1 \in D_1^l$, then with $\xi_1 = 2\pi/h_l - \xi_1'$, we obtain

$$\exp(i(2\pi/h_l - \xi_1')\mu h_l) = \exp(-i\xi_1'\mu h_l)$$

as in (3.5). Hence, for a wave number ξ in C_l^* the spatial variation is slow. Since $\xi_* < \pi/h_0$, there is no overlap on the coarsest mesh

$$C_0^* \subset C_0 .$$

Moreover, if $\xi_1 \in D_1^l$, then $\tilde{\xi}_1 \in D_1^0$, where

$$\xi_1 = \tilde{\xi}_1 + (m-1) \cdot 2\pi/h_0 .$$

With

$$\xi_1 = 2\pi/h_l - \xi_1' = 2\pi/h_0 - \xi_1' + (m-1) \cdot 2\pi/h_0, \qquad \xi_1' \in D_0,$$

it follows for a difference operator, e.g., Q with its symbol $\hat{Q}$ defined in (3.8), that

$$
\begin{aligned}
\hat{Q}_{m-1,k}((2\pi/h_0 - \xi_1')h, \xi_2 h, h) &= \hat{Q}(\xi_1 h, (\xi_2 + k \cdot 2\pi/h_0)h, h) \\
&= \hat{Q}(-\xi_1' h, (\xi_2 + k \cdot 2\pi/h_0)h, h) = \hat{Q}_{0k}(-\xi_1' h, \xi_2 h, h) .
\end{aligned} \tag{3.19}
$$

The restriction and prolongation symbols have the same property. For completeness we formulate all the assumptions we need in the Fourier analysis, even if some of them follow from the assumptions in §2. Note that

$$(\xi_i + (m/2)\cdot 2\pi/h_0)h_l = \eta_{il} + \pi\,, \qquad i = 1\,,\ 2\,.$$

Assumption 3.1.

$$\begin{aligned}
&\hat{r}_{i0} = \hat{r}_{i0}(\eta_{il}) = 1 + O(\|\eta_l\|) \quad \text{for small } \|\eta_l\|\,,\\
&\hat{r}_{i\,,m/2} = \hat{r}_{i0}(\eta_{il} + \pi) = O(\|\eta_l\|)\,, \qquad i = 1\,,\ 2\,,\\
&\hat{p}_{ij} \text{ has the same properties,} \qquad 0 \le l \le L\,.
\end{aligned}$$

This assumption concerns the behavior of the restriction and prolongation operators at low wave numbers with $\|\eta_l\|$ small (cf. Definition 2.1) and at oscillatory wave numbers $\eta_{il} + \pi$ with $\|\eta_{il}\|$ small. This assumption is very natural and is satisfied in (3.12). Its counterpart for the physical variables is Assumption 2.4(i). The requirements on r and p with respect to the order of Q are derived by Fourier analysis in [8]. The low- and high-frequency order in the terminology of [8, equations (16), (17)] is at least 1 in the assumption.

Assumption 3.2.

$$(\widetilde{Q}_{l-1})_{00}^{-1}(\widetilde{Q}_l)_{00} = I + O(\|\eta_{l-1}\|)\,, \qquad 1 \le l \le L\,.$$

Suppose that the numerical approximation of Q is at least first-order accurate and that

$$(\widetilde{Q}_l)_{00} = \widehat{Q}(I + O(\|\eta_l\|))\,,$$

where $\widehat{Q}$ is the symbol of Q, see (3.8). With such a $\widehat{Q}_{00}$ the assumption is fulfilled. The assumption corresponds to [8, equation (31)].

Assumption 3.3. $\widetilde{S}_l$ is a block diagonal matrix as specified in §3.1 and bounded uniformly when $\xi \in C_0^*$. The upper left block is

$$(\widehat{S}_l)_{00} = I - \Delta t_l \widehat{H}_l + O(\Delta t_l(\Delta t_l + \|\eta_l\|))\,,$$

where $\widehat{H}_l$ is the leading term in ξ in the smoothing operator for small ξ, and Δt_l is a small parameter of $O(h_l^\nu)$, $\nu \ge 1$, $0 \le l \le L$.

The assumption is related to Assumption 2.2(i). In case we choose Runge-Kutta time-stepping in the smoothing iterations, then

$$\widehat{Q}_{00}(\xi h_l\,,\ h_l) = \widehat{H}_l(\xi) + O(\|\eta_l\|)\,,$$

i.e., $\widehat{H}_l$ is the symbol $\widehat{Q}$. It is remarked in [14] that Chebyshev iteration [15] has the same property. In numerical experiments in [14], GMRES [17] also appears to have a similar property but with Δt_l dependent on the iteration number n.

The time step Δt_l can be regarded as a smoothly varying scaling parameter of the equations. With a Runge-Kutta scheme, the remainder term in $(\widehat{S}_l)_{00}$ is Δt_l times the sum of the truncation error due to the discretization in space of $O(\|\eta_l\|)$ at least and $\beta\Delta t_l\widehat{H}_l(\widehat{H}_l + O(\|\eta_l\|))$, where β is a constant depending on the scheme.

Assumption 3.4. Define

$$f^l_{j+\alpha m/2,k+\beta m/2}(\eta_l) = \hat{r}_{1,j+\alpha m/2}\hat{r}_{2,k+\beta m/2} g_1^\alpha g_2^\beta \overline{Q}_{jk\alpha\beta},$$

$$\overline{Q}_{jk\alpha\beta} = (\hat{Q}_{l-1})^{-1}_{jk}(\hat{Q}_l)_{j+\alpha m/2,k+\beta m/2},$$

$$j,k = 0,1,2,\dots,m/2-1,\quad 1 \le l \le L,\quad \alpha,\beta = 0,1,$$

where g_i, $i = 1,2$, are defined in (3.11). For all l, j, k, α, and β,

$$\|f^l_{j+\alpha m/2,k+\beta m/2}\|$$

is bounded for η_l such that $\|\eta_l\| < \delta$.

This rather technical assumption is a sufficient condition to obtain bounded elements in $\widehat{M}$. Note that $\overline{Q}_{jk\alpha\beta}$ may be unbounded, but f^l is not. This is the case in the simple one-dimensional example in [5, §5.3]. When comparing with the assumptions in §2, we note that the bound in Assumption 2.3(ii) includes also oscillatory modes. The main result in [8] provides necessary conditions for the assumption to hold in a two-grid iteration. It follows already from Assumptions 3.1 and 3.2 that $\|f^l_{00}\|$ is of $O(1)$.

Assumption 3.5. The wave numbers are ordered such that the first submatrix in $\mathbf{E}_{\mu\nu}$ for each l is

$$\begin{aligned} &e^1_0 e^2_0 I_s, && \xi \in D_{00},\\ &e^1_0 e^2_{m-1} I_s, && \xi \in D^0_{01},\\ &e^1_{m-1} e^2_0 I_s, && \xi \in D^0_{10},\\ &e^1_{m-1} e^2_{m-1} I_s, && \xi \in D^0_{11}, \end{aligned}$$

where the D's are defined in (3.18) and ξ_* is sufficiently small.

The ordering of the wave numbers is chosen as above so that the mode of slowest variation is located in the first position of $\mathbf{E}_{\mu\nu}$. How to do this for $\xi \in D_{00}$ is indicated in §3.1, but the desired ordering can also be obtained for the other corners in C_0 by a minor redefinition of E_{jk} in (3.6).

Partition $\widetilde{M}$ in (3.16) for $\xi \in D_{00}$, $\|\eta_l\|$ sufficiently small and $1 \le l \le L$ as follows:

$$\widetilde{M} = \begin{pmatrix} \widetilde{M}_{00} & \widetilde{M}_{0I} \\ \widetilde{M}_{I0} & \widetilde{M}_{II} \end{pmatrix}, \qquad \widetilde{M}_{00} \in \mathbf{C}^s \times \mathbf{C}^s,\ \widetilde{M}_{II} \in \mathbf{C}^{s(m^2-1)} \times \mathbf{C}^{s(m^2-1)}. \tag{3.20}$$

Assumption 3.6. $\Delta t_0 < \delta$, $\|\eta_0\| < \delta$, for some $\delta > 0$, when $\xi \in C_0^*$ defined in (3.18).

The assumption restricts the size of Δt_0 and $\eta_0 = h_0\xi$ on the coarsest grid.

Assumption 3.7. $\widehat{H}_l$ has simple eigenvalues for $\|\eta_0\| < \delta$, and $\widehat{H}_l$ is independent of l, $\widehat{H}_l = \widehat{H}(\xi)$.

The first part of the assumption is trivially true for scalar equations. If the same type of smoothing iteration is chosen at each level, then $\widehat{H}_l$ remains the same for each l.

Assumption 3.8. $\widetilde{M}_{II}$ has simple eigenvalues $\mu_k(\widetilde{M}_{II})$, $k = 1,2,\dots,s(m^2-1)$, under Assumption 3.6 and if $\delta < \delta_0$, then there is an ε_0, $\varepsilon_0 > 0$, such that

$$|\mu_k| \le \theta < 1 - \varepsilon_0.$$

Two auxiliary lemmas, see Appendix B, are needed in the proof of the theorem in this section. The proofs of the lemmas are simplified considerably if Assumptions 3.7 and 3.8 are satisfied. With the help of the theory in [12] and more detailed properties of $\widetilde{S}$, less restrictive conditions can be derived.

We are now ready to state the theorem giving the properties of $\widetilde{M}^n$. The parameters α_j are defined in §2. The eigenvector matrices of $\widehat{H}$ and $\widetilde{M}_{\mathrm{II}}$ are denoted by T_0 and V_0, respectively.

Theorem 3.1. *Let Assumptions* 3.1–3.8 *be satisfied. If* δ *is sufficiently small, then the transformed multigrid iteration matrix for a* V*-cycle at level* l *after* n *iterations for* $\xi \in C_0^*$ *is*

$$\widetilde{M}^n = \begin{pmatrix} (I - \delta t \widehat{H})^n + O(\Delta t_0 + \|\eta_0\|) & O(\Delta t_0) \\ O(\Delta t_0 \|\eta_0\|) & \widetilde{M}_{\mathrm{II}}^n + O(\Delta t_0 \|\eta_0\|) \end{pmatrix},$$

where

$$\delta t = (p+q) \sum_{j=0}^{l} \Delta t_j = (p+q) \Delta t \sum_{j=0}^{l} \alpha_j .$$

Alternatively, the upper left corner of $\widetilde{M}^n$ *can be written*

$$T_0 \mathrm{diag}(\{\psi_k^n\}_{k=1}^s) T_0^{-1} + O(\Delta t_0 + \|\eta_0\|),$$

where

$$\psi_k^n = (1 - \delta t \lambda_k(\widehat{H}))^n = |\psi_k|^n \exp(-in\delta t \operatorname{Im} \lambda_k(\widehat{H})) + O(\Delta t_0).$$

In the lower right corner of $\widetilde{M}^n$ *we have*

$$\|\widetilde{M}_{\mathrm{II}}^n\| \le \|V_0\| \cdot \|V_0^{-1}\| \theta^n, \qquad \theta < 1 .$$

The proof of the theorem is found in Appendix B. In Theorem 3.1 we have investigated the behavior of $\widetilde{M}(\xi)$ when $\xi \in C_0^*$. For the remaining part of the wave number domain C_0 we merely make

Assumption 3.9.

$$\|\widetilde{M}(\xi)\| \le \theta < 1, \qquad \xi \in C_0 \backslash C_0^*, \ 0 \le l \le L .$$

The consequences of the assumptions and the analysis in this section will be discussed in the next section.

4. Examples and discussion

We discuss the results of the previous section and present two simple numerical illustrations in this section.

Several conclusions can be drawn from Theorem 3.1. Partition the Fourier coefficients $\tilde{u}$ in the same manner as $\widetilde{M}$ in (3.20) for $\xi \in D_{00}$,

$$\tilde{u} = (\hat{u}_{00}^T, \tilde{u}_{\mathrm{I}}^T)^T . \tag{4.1}$$

The low wave number part is

$$\hat{u}_{00} = \hat{u}(\xi_1, \xi_2) \in \mathbf{C}^s ,$$

and the remaining wave number components are collected in $\tilde{u}_{\mathrm{I}}$. Suppose that the conditions in the theorem are satisfied. Then by (3.17) and the theorem, after the nth iteration on the finest grid L,

$$\hat{u}_{00}^{n} = (I - \delta t_L \widehat{H})^n \hat{u}_{00}^{0} + O(\Delta t_0 + \|\eta_0\|), \tag{4.2}$$

$$\tilde{u}_{\mathrm{I}}^{n} = (\widetilde{M}_L)_{\mathrm{II}}^{n} \tilde{u}_{\mathrm{I}}^{0} + O(\Delta t_0 \|\eta_0\|). \tag{4.3}$$

The intermediate and high wave numbers in $\tilde{u}_{\mathrm{I}}^{0}$ are damped by a factor θ in each iteration of the multigrid procedure.

The low wave number part in (4.2) behaves as if we had applied the smoothing operator S to the low wave numbers only, but with "time-step" $(p+q)\Delta t_L \sum_{j=0}^{L} \alpha_j$ instead of Δt_L (see Assumption 3.3 and Theorem 2.1). At the other end of the spectrum, $\xi \in D_{01}^{0} \cup D_{10}^{0} \cup D_{11}^{0}$, the wave numbers with slow spatial variation behave as $\hat{u}_{00}^{n}$ in (4.2), whereas the remaining Fourier components are damped as in (4.3). If Assumption 3.9 is fulfilled, then the description of the evolution of the Fourier components for all wave numbers is complete in C_0, and therefore also in C_L.

We summarize in a theorem the preceding discussion with a comparison between the smoothing iteration and the multigrid iteration.

The Fourier coefficients associated with slowly varying modes, $\xi \in C_L^*$, after n iterations on grid L with only the smoothing operator S are denoted by

$$\hat{u}_S^n(\xi) = \widehat{U}_L(n\Delta t_L, \xi), \qquad \xi \in C_L^*. \tag{4.4}$$

The corresponding coefficients for the multigrid iterations is $\hat{u}_{\mathrm{MG}}^{n}(\xi)$.

Theorem 4.1. *Let the sufficient conditions in Theorem* 3.1 *and Assumption* 3.9 *be fulfilled. Assume that the Fourier symbol* $\widehat{S}$ *of the smoothing operator satisfies*

$$\|\widehat{S}(\xi)\| \le \theta < 1, \qquad \xi \in C_L \backslash C_L^*,$$

and that $\|\hat{u}^0(\xi)\|$ *is bounded for* $\xi \in C_L$. *Then for* $\xi \in C_L^*$,

$$\hat{u}_{\mathrm{MG}}^{n}(\xi) = \widehat{U}_L(n\delta t_L, \xi) + O(\Delta t_0 + \|\eta_0\|),$$

where $\widehat{U}_L$ *and* η_0 *are defined in* (4.4) *and Theorem* 3.1, *respectively. For* $\xi \in C_L \backslash C_L^*$,

$$\|\hat{u}_S^n(\xi)\| \le \theta^n \|\hat{u}^0(\xi)\|,$$
$$\|\hat{u}_{\mathrm{MG}}^{n}(\xi)\| \le c_L(\xi)\theta^n + O(\Delta t_0^{1+1/\nu}),$$

where c_L *depends on* $\hat{u}^0$, *and* ν *is defined in Assumption* 3.3.

Proof. Consider the oscillatory modes, and suppose that $\xi \in C_L \backslash C_L^*$, that there is a ξ' in C_0^* and

$$(\iota, \kappa) \in I_{mm} = I_m \times I_m, \qquad I_m = \{0, 1, \ldots, m-1\},$$
$$\xi = \xi' + 2\pi/h_0(\iota, \kappa)^T. \tag{4.5}$$

Then by Theorem 3.1, when $\xi' \in D_{00}$,

$$\|\hat{u}_{\mathrm{MG}}^{n}(\xi)\|^2 = \|\hat{u}_{\iota\kappa}^{n}(\xi')\|^2 \le \sum_{(j,k) \in I_{mm} \backslash (0,0)} \|\tilde{u}_{jk}^{n}(\xi')\|^2$$
$$= \|\tilde{u}_{\mathrm{I}}^{n}(\xi')\|^2 \le c_V^2 \theta^{2n} \|\tilde{u}_{\mathrm{I}}^{0}(\xi')\|^2 + O(\Delta t_0 \|\eta_0\|).$$

The same bound is valid in $D_{01}^0 \cup D_{10}^0 \cup D_{11}^0$. Since $\|\tilde{u}^0(\xi)\|$ is bounded for $\xi \in C_L$, there is a $c_L(\xi)$ such that

$$\|\hat{u}_{\mathrm{MG}}^n(\xi)\| \le c_L(\xi)\theta^n + O(\Delta t_0 \|\eta_0\|). \tag{4.6}$$

In C_L^* there is a $\hat{c}$ such that

$$\|\eta_0\| \le h_0 \|\xi_*\| \le \hat{c}\Delta t_0^{1/\nu}.$$

When ξ' in (4.5) fulfills $\xi' \in C_0 \backslash C_0^*$, then by Assumption 3.9,

$$\|\hat{u}_{\mathrm{MG}}^n(\xi)\|^2 = \|\hat{u}_{i\kappa}^n(\xi')\|^2 \le \sum_{(j,k)\in I_{mm}} \|\hat{u}_{jk}^n(\xi')\|^2 \le \theta^{2n}\|\tilde{u}^0(\xi')\|^2.$$

With only smoothing iterations, we have

$$\|\hat{u}_S^n(\xi)\| = \|\hat{S}^n(\xi)\hat{u}^0(\xi)\| \le \theta^n \|\hat{u}^0(\xi)\|.$$

The results in the theorem for $\xi \in C_L \backslash C_L^*$ are proved.

If $\xi \in C_L^*$, then $\xi' \in C_0^*$, and it follows from Assumption 3.3 that

$$\hat{u}_S^n(\xi) = (\hat{S}_L)_{00}^n \hat{u}^0(\xi) = (I - \Delta t_L \hat{H})^n \hat{u}^0(\xi) + O(\Delta t_L + \|\eta_L\|) \equiv \hat{U}_L(n\Delta t_L, \xi).$$

By Theorem 3.1 we obtain

$$\begin{aligned} \hat{u}_{\mathrm{MG}}^n(\xi) &= (I - \delta t_L \hat{H})^n \hat{u}^0(\xi) + O(\Delta t_0 + \|\eta_0\|) \\ &= \hat{U}_L(n\delta t_L, \xi) + O(\Delta t_0 + \|\eta_0\|). \quad \square \end{aligned}$$

If we choose Runge-Kutta time-stepping, then

$$\hat{H} = \hat{Q}_{00}, \tag{4.7}$$

and the slow Fourier modes are integrated at least first-order accurately in time, cf. Theorem 2.1, but the time step taken per multigrid iteration is so much larger than it is with only Runge-Kutta iteration. An interpretation of the theorem in this case is that time (or pseudotime) proceeds faster with the multigrid method. This effect is always achieved with multigrid iterations satisfying the assumptions.

Let us consider two simple examples. The smoothing operator is Runge-Kutta time-stepping (or repeated Richardson iteration with a fixed number of steps) and the problem to be solved is scalar, $s = 1$. In the first example, the differential operator P is of first order,

$$Pu = \frac{\partial u}{\partial x^1} + \frac{\partial u}{\partial x^2}.$$

It is approximated by a first-order accurate difference expression such that

$$\hat{Q}_{00} = i(\xi_1 + \xi_2) + O(h_L)$$

for small ξ_j, $j = 1, 2$. According to Theorem 3.1, (4.2) and (4.7), we obtain on the finest grid L

$$\hat{u}_{00}^n = \exp(-i(\xi_1 + \xi_2)n\delta t_L)\hat{u}_{00}^0 + O(h_L). \tag{4.8}$$

For numerical stability, Δt_L is of $O(h_L)$. One part of $u_{\mu\nu}^n$ corresponding to low wave numbers is

$$\hat{u}_{00}^n e_0^1 e_0^2 = \exp(i(\xi_1(x_{\mu\nu}^1 - n\delta t_L) + \xi_2(x_{\mu\nu}^2 - n\delta t_L)))\hat{u}_{00}^0 + O(h_L), \quad \xi \in D_{00}. \tag{4.9}$$

The result is similar in the other three corners of C_L^*. The conclusion from (4.9) is that the smooth part of $u_{\mu\nu}^n$ is propagated a distance $n\delta t_L$ in both spatial directions in n multigrid iterations. If only smoothing iterations are employed at level L, then the corresponding distance is $n\Delta t_L$, cf. Theorem 4.1. The improvement with multigrid iteration with the same number of smoothing iterations on level L is

$$(n\delta t_L(p+q))/n\Delta t_L = \sum_{j=0}^{l} \alpha_j , \tag{4.10}$$

in agreement with Theorem 2.1. In [5] this factor is interpreted as an increase of the group speed of a plane wave.

In the second example the differential operator is of second order,

$$Pu = -\left(\frac{\partial^2 u}{(\partial x^1)^2} + \frac{\partial^2 u}{(\partial x^2)^2}\right).$$

Choose the numerical approximation to be at least first-order accurate,

$$\hat{Q}_{00} = \xi_1^2 + \xi_2^2 + O(h_L)$$

for small ξ_j, $j = 1, 2$. Analogously to (4.8), we derive

$$\hat{u}_{00}^n = \exp(-(\xi_1^2 + \xi_2^2)n\delta t_L)\hat{u}_{00}^0 + O(h_L).$$

Stability requirements force Δt_L to be of $O(h_L^2)$. Here, we have only damping of the Fourier coefficient. Suppose that

$$\Delta t_l = \beta h_l^2 ,$$

where β is a constant. Then

$$\alpha_j = \Delta t_j/\Delta t_L = (h_j/h_L)^2 = 2^{2(L-j)}$$

and

$$\delta t_L = \sum_{j=0}^{L} \alpha_j \Delta t_L = \Delta t_0 \sum_{j=0}^{L} \alpha_j/\alpha_0 = \Delta t_0 \cdot \frac{4}{3}(1 - 0.25^{L+1}).$$

Therefore, the damping of the mode is governed by

$$\exp(-(\xi_1^2 + \xi_2^2)4n\beta h_0^2/3),$$

essentially achieved by the coarse grid. The improvement of multigrid iterations over only Runge-Kutta time-stepping is also as in the first case given by Theorem 4.1, but the convergence is becoming progressively slower as $\|\xi\| \to 0$.

As an illustration of the theoretical results in this paper, the eigenvalues of $\widehat{M}$ in Theorem 3.1 are plotted for a one-dimensional example in Figure 4.1 (next page) and three two-dimensional examples in Figure 4.2 (see p. 491). In Figure 4.1 the model equation to be solved is

$$u_x = f. \tag{4.11}$$

The equation is discretized by a cell-centered finite volume scheme with additional 4th-order artificial viscosity, see [10]. The smoothing iteration is a five-stage Runge-Kutta method and $\Delta t_l = h_l$. The number of presmoothing steps is 1, and there are no postsmoothing iterations.

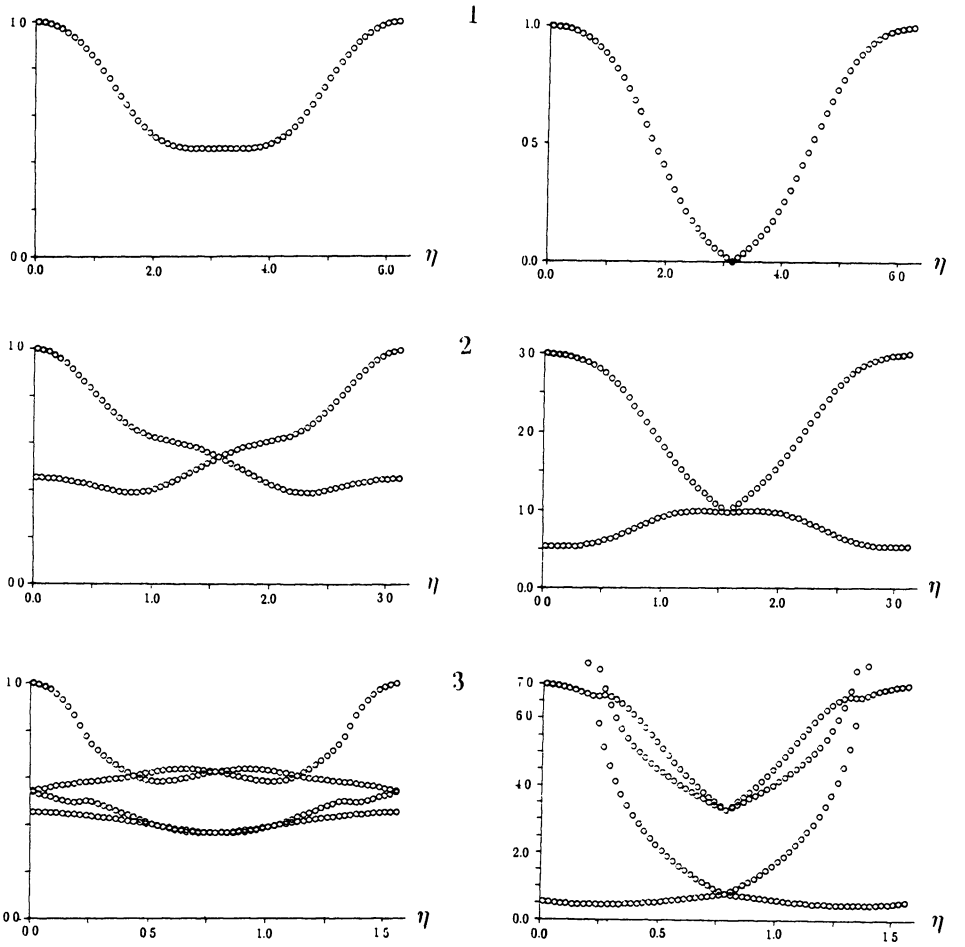

FIGURE 4.1. The eigenvalues $\lambda_j(\eta)$ of the Fourier transform of the multigrid matrix $\widetilde{M}$ are displayed for the model equation $u_x = f$ and a five-stage Runge-Kutta scheme on 1, 2, and 3 grids. On the left, $|\lambda_j|$ is plotted, and on the right, $\operatorname{Im} \log \lambda_j / \zeta \Delta t_l$ shows the speed-up factor for $|\lambda_j|$ close to 1

The symbol of the restriction operator is

$$\hat{r} = \cos(\xi h_l / 2),$$

and $\hat{p} = \hat{r}$.

In the left column, $|\lambda_j(\widetilde{M})|$, $j = 1, 2, \ldots, m = 2^L$, is plotted as a function of $\eta_L = \xi h_L$, where $\xi \in C_0$. Thus,

$$\eta_L \in [0, 2\pi/m].$$

In the right column of Figure 4.1 the factor

$$\operatorname{Im} \log \lambda_j(\widetilde{M}) / \zeta \Delta t_L,$$

$$\zeta = \begin{cases} \xi, & \xi \in [0, \pi/h_0], \\ 2\pi/h_0 - \xi, & \xi \in (\pi/h_0, 2\pi/h_0], \end{cases}$$

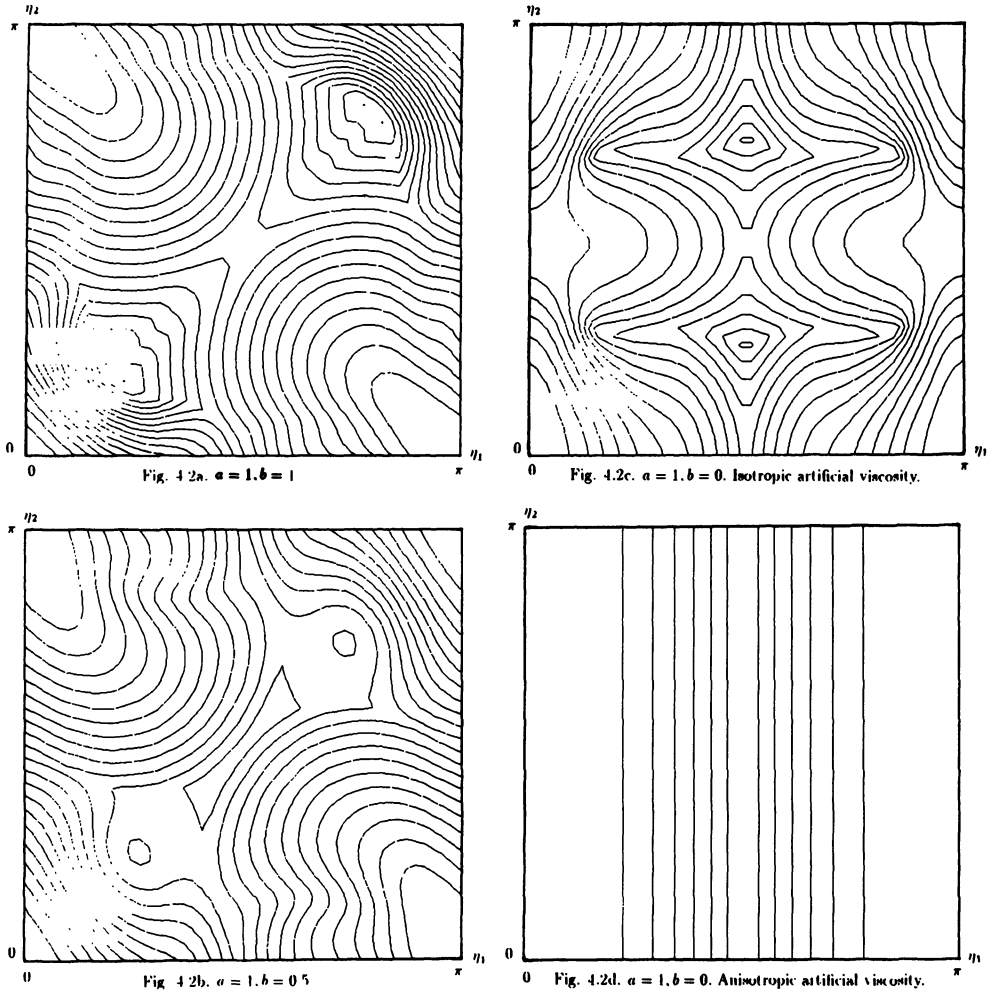

FIGURE 4.2. Isolines of the maximum of $|\lambda_j(\eta)|$ for each η of the Fourier transform of the multigrid matrix $\widehat{M}$ are displayed for the two-dimensional equation (4.12) and a three-stage Runge-Kutta scheme on 2 grids. The maximum is 1 in the corners, and less than 1 in the interior of the η-domain

is displayed as a function of η_L. It follows from Theorem 3.1 that this is the speed-up factor $\gamma = \sum_{j=0}^{L} \alpha_j$ of the slowly varying Fourier mode with $\lambda_1 \approx 1$ for small ξ and ζ close to $2\pi/h_0$. The results are in accordance with the theory in Theorem 3.1. There is one eigenvalue λ_1 in each of the three cases, $L = 0, 1, 2$, for which $|\lambda_1| \approx 1$ when

$$\xi \in [0, \xi_*] \cup [2\pi/h_0 - \xi_*, 2\pi/h_0],$$

with, e.g., $\xi_* < 1/2h_0$, and when ξ is in the neighborhood of 0 or $2\pi/h_0$, we have

$$\gamma = 2^{L+1} - 1.$$

For the other eigenvalues,

$$|\lambda_j(\widetilde{M})| \leq \theta < 1, \qquad j = 2, \ldots, m.$$

The two-dimensional model example is

$$a\frac{\partial u}{\partial x^1} + b\frac{\partial u}{\partial x^2} = f. \tag{4.12}$$

The equation is discretized by the same method as (4.11). The smoothing procedure is a five-stage Runge-Kutta scheme, the multigrid strategy is the same as above and there are two grids, $L = 1$.

Isolines of

$$\max_{j=1,2,3,4} |\lambda_j(\widehat{M}(\xi))|$$

are plotted in Figure 4.2. The distance between two lines is 0.02. In Figure 4.2(a) and Figure 4.2(b), the coefficients in (4.12) are $a = b = 1$ and $a = 1$, $b = 0.5$, respectively. When $\xi \in C_L^*$ (the squares in the corners of the wave number domain), the maximum of $|\lambda_j|$ is close to 1 and exactly 1 at the corners. In the interior of C_L we have

$$|\lambda_j| \leq \theta < 1, \qquad j = 1, 2.$$

The coefficients in (4.12) and Figure 4.2(c) are $a = 1$ and $b = 0$. We still have good damping properties in the interior, and $\max|\lambda_j|$ is close to 1 in the corners. There are no "grid alignment" effects with no damping at all in large parts of the wave number domain. This is sometimes a problem with upwind discretizations of (4.12) [16]. The reason why the centered difference scheme is successful is that the artificial viscosity is isotropic and independent of a and b in (4.12). Similar results are obtained with a three-stage Runge-Kutta scheme in [6].

On the other hand, if we scale the artificial viscosity term in the x^1-direction by a and in the x^2-direction by b, then we have an obvious "grid alignment" problem as in Figure 4.2(d). The multigrid method does not reduce the amplitude of the modes with, e.g., η_1 small and $\eta_2 \approx \pi/2$, simply because there is no artificial viscosity in the x^2-direction.

In [14] we use Parseval's relation for the Fourier representation in §3.1 and Theorem 3.1 to derive the properties of u^n in the multigrid iterations.

Bibliography

1. A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp. **31** (1977), 333–390.
2. R. Enander, *Grid patching and residual smoothing for computations of steady state solutions of first order hyperbolic systems*, Ph.D. thesis, Dept. of Scientific Computing, Uppsala Univeristy, Uppsala, 1991.
3. L. Ferm, *The choice of the number of iterations in every cycle for the two-grid method*, Report 139, Dept. of Scientific Computing, Uppsala University, Uppsala, 1991.
4. P. R. Garabedian, *Estimation of the relaxation factor for small mesh size*, MTAC **10** (1956), 183–185.
5. B. Gustafsson and P. Lötstedt, *Analysis of the multigrid method applied to first order systems*, Proc. Fourth Copper Mountain Conf. on Multigrid Methods (J. Mandel et al., eds.), SIAM, Philadelphia, PA, 1989, pp. 181–233.

6. ——, *Analysis of multigrid methods for general systems of PDE*, Multigrid Methods III (W. Hackbusch and U. Trottenberg, eds.), Birkhäuser Verlag, Basel, 1991, pp. 223–234.
7. W. Hackbusch, *Multigrid methods and applications*, Springer-Verlag, Berlin, Heidelberg, 1985.
8. P. W. Hemker, *On the order of prolongations and restrictions in multigrid procedures*, J. Comput. Appl. Math. **32** (1990), 423–429.
9. A. Jameson, *Solution of the Euler equations for two-dimensional transonic flow by a multigrid method*, Appl. Math. Comput. **13** (1983), 327–355.
10. ——, *Computational transonics*, Comm. Pure Appl. Math. **41** (1988), 507–549.
11. D. C. Jespersen, *A time-accurate multiple-grid algorithm*, AIAA paper 85-1493-CP, (1985).
12. T. Kato, *Perturbation theory for linear operators*, Springer-Verlag, New York, 1966.
13. P. Lancaster, *Theory of matrices*, Academic Press, New York, 1969.
14. P. Lötstedt, *Grid independent convergence of the multigrid method for first-order equations*, SIAM J. Numer. Anal. **29** (1992), 1370–1394.
15. T. A. Manteuffel, *The Tchebychev iteration for nonsymmetric linear systems*, Numer. Math. **28** (1977), 307–327.
16. W. A. Mulder, *A new multigrid approach to convection problems*, J. Comput. Phys. **83** (1989), 303–323.
17. Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput. **7** (1986), 856–869.
18. K. Stüben and U. Trottenberg, *Multigrid methods: Fundamental algorithms, model problem analysis and applications*, Multigrid Methods (Proc., Köln-Porz, 1981), (W. Hackbusch and U. Trottenberg, eds.), Lecture Notes in Math., vol. 960, Springer-Verlag, Berlin, pp. 1–176.
19. R. S. Varga, *Matrix iterative analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.

DEPARTMENT OF SCIENTIFIC COMPUTING, UPPSALA UNIVERSITY, UPPSALA, SWEDEN, AND SAAB-SCANIA, LINKÖPING, SWEDEN
E-mail address: perlo@weald.air.saab.se

DEPARTMENT OF SCIENTIFIC COMPUTING, UPPSALA UNIVERSITY, UPPSALA, SWEDEN
E-mail address: bertil@tdb.uu.se