SIAM J. SCI. COMPUT.
Vol. 22, No. 2, pp. 582–603

FOURIER ANALYSIS OF GMRES(m) PRECONDITIONED BY MULTIGRID*

ROMAN WIENANDS†, CORNELIS W. OOSTERLEE†, AND TAKUMI WASHIO‡

Abstract. This paper deals with convergence estimates of GMRES(m) [Saad and Schultz, *SIAM J. Sci. Statist. Comput.*, 7 (1986), pp. 856–869] preconditioned by multigrid [Brandt, *Math. Comp.*, 31 (1977), pp. 333–390], [Hackbusch, *Multi-Grid Methods and Applications*, Springer, Berlin, 1985]. Fourier analysis is a well-known and useful tool in the multigrid community for the prediction of two-grid convergence rates [Brandt, *Math. Comp.*, 31 (1977), pp. 333–390], [Stüben and Trottenberg, in *Multigrid Methods*, Lecture Notes in Math. 960, K. Stüben and U. Trottenberg, eds., Springer, Berlin, pp. 1–176]. This analysis is generalized here to the situation in which multigrid is a preconditioner, since it is possible to obtain the whole spectrum of the two-grid iteration matrix. A preconditioned Krylov subspace acceleration method like GMRES(m) implicitly builds up a minimal residual polynomial. The determination of the polynomial coefficients is easily possible and can be done explicitly since, from Fourier analysis, a simple block-diagonal two-grid iteration matrix results. Based on the GMRES(m) polynomial, sharp theoretical convergence estimates can be obtained which are compared with estimates based on the spectrum of the iteration matrix. Several numerical scalar test problems are computed in order to validate the theoretical predictions.

Key words. Fourier analysis, multigrid, restarted GMRES

AMS subject classifications. 65N55, 65N12, 65F10

PII. S1064827599353014

1. Introduction.

Nowadays, it has become popular to study the convergence of multilevel methods and to use them in combination with a Krylov subspace acceleration method, in other words, to use multilevel methods as a preconditioner. Often the cycling in a hierarchy of grids is simplified compared to standard multigrid cycling; i.e., the multiplicative cycling between the fine and coarse grids is replaced by an additive (simultaneous) cycling. It is then necessary to use this additive method as a preconditioner to obtain a converging method even for simple problems. Recently, standard (multiplicative) multigrid solvers [1], [8] have been used as preconditioners as well. This application of standard multigrid is beneficial in situations where standard multigrid alone does not converge fully satisfactorily, because certain error frequencies are not reduced well enough. This occurs mainly when complicated PDEs are solved. In general, it is difficult to construct *robust* multigrid solvers for large classes of problems. From this point of view, it makes sense to increase the class of problems for which proven efficient multigrid solvers exist by accelerating (standard) multigrid by a Krylov subspace method. Among many other papers, this approach has been presented in [11], where a symmetric multigrid method was accelerated by CG, in [3] for problems with fine level structures, in [13] for linear singularly perturbed problems, and in [14] for nonlinear problems.

Here, we theoretically analyze one combined solution method, restarted GMRES, GMRES(m), preconditioned by multigrid. We derive quantitative convergence results. The basis for the theoretical convergence estimates is Fourier analysis, which

*Received by the editors March 10, 1999; accepted for publication (in revised form) October 25, 1999; published electronically August 3, 2000.

http://www.siam.org/journals/sisc/22-2/35301.html

†GMD, Institute for Algorithms and Scientific Computing, D-53754 Sankt Augustin, Germany (wienands@gmd.de, oosterlee@gmd.de).

‡C&C Media Research Laboratories, NEC Cooperation, 1-1, Miyazaki 4-chome, Miyamae-ku, Kawasaki, Kanagawa 216-8555, Japan (washio@ccm.cl.nec.co.jp).

is a well-known tool for obtaining two-grid convergence factors (see [1], [19]). The typical use of the two-grid Fourier analysis in a multigrid context is that the spectral radius is obtained theoretically. It is the basis for asymptotic multigrid convergence estimates [19]. Two-grid Fourier analysis transforms a two-grid iteration matrix into a block-diagonal matrix from which more information about the spectrum is easily obtained. Fourier analysis uses unitary basis transformations to simplify the representation of the two-grid iteration matrix. It is therefore suitable for the estimate of the multigrid preconditioned GMRES(m) convergence since unitary transformations do not affect the convergence behavior of GMRES [15]. A preconditioned Krylov subspace acceleration method like GMRES(m) implicitly builds up a minimal residual polynomial. The determination of the polynomial coefficients is easily possible and can be done explicitly, since a block-diagonal matrix results with the help of the Fourier analysis.

In this paper, we restrict ourselves to standard problems coming from linear scalar PDEs discretized with finite differences on uniform Cartesian grids and we keep the multigrid method simple. For example, we restrict ourselves to point smoothers combined with standard grid coarsening and analyze the effect of Krylov subspace acceleration for anisotropic diffusion problems and for problems with mixed derivatives, although it is known that other (more expensive) smoothers or other coarsening strategies are appropriate remedies for the poor convergence of such equations. In this way we gain insight in the behavior of the combination of multigrid and GMRES(m).

In section 2 a multigrid preconditioner is presented and the rigorous two-grid Fourier analysis (see [19], [20]) is repeated briefly for 3D problems. The information on the multigrid preconditioner obtained is used in section 3, where different ways to analyze GMRES(m) preconditioned by multigrid are discussed.

There are actually two types of Fourier analysis available in the multigrid community. The first one is the analysis called "rigorous analysis" here, since it explicitly takes boundary conditions into account and discrete eigenfunctions are related to the discrete mesh size. Discrete sine functions are the basis in case Dirichlet boundary conditions are considered [19]. This analysis is, however, restricted to model situations, i.e., to symmetric problems and standard multigrid components. The second possibility is known as local mode analysis [1]. Exponential functions are the basis for the analysis. A discretization at domain boundaries cannot be taken into account. This analysis can be used for a wider range of PDEs and multigrid components. Local mode two-grid analysis and its generalization to multigrid preconditioned GMRES(m) is explained in section 4. Several numerical tests compare the theoretical convergence estimates with the actual numerical convergence.

For symmetric equations it is possible to construct a symmetric multigrid cycle and to apply CG as Krylov subspace acceleration method. We will not do this in section 3 because CG cannot be applied to the unsymmetric multigrid solvers MG-RB (multigrid solver based on pre- and post-smoothing by red-black Gauss–Seidel relaxation) from section 2 and MG-FF (multigrid solver based on pre- and post-smoothing by lexicographical Gauss–Seidel relaxation) from section 4. Furthermore, we are dealing with unsymmetric equations in sections 4.2 and 4.3. However, the analysis as described in section 3 carries over to an acceleration by CG in a straightforward manner. (See also the first remark in section 4.1.)

Furthermore, a relation to other techniques in which (multistage) parameters for smoothing methods, or acceleration parameters for coarse grid corrections [4], are optimized is presented in section 4.3. Such approaches can easily be viewed in the

framework of Krylov subspace acceleration.

The multigrid purist may not like the approach presented here with multigrid viewed as a preconditioner. It is, of course, also possible to try to construct an optimal multigrid method for each separate problem. For the purist, the analysis might give an indication as to how far his present method is from an optimal multigrid solution method. In other words, if an additional acceleration of the convergence is found by the Fourier analysis of multigrid preconditioned GMRES(m), it must also be possible to tune one of the multigrid components, so that multigrid as a stand alone solver is improved.

2. Rigorous Fourier analysis of multigrid. The Poisson equation, a simple test problem, is chosen to explain the possibilities of convergence estimates by Fourier analysis for multigrid. We apply the standard 5-point discretization on a square ($d = 2$) and the 7-point discretization on a cube ($d = 3$) with Dirichlet boundary conditions using a uniform mesh with meshsize $h = 1/n$:

$$\begin{aligned} A_h u_h(\boldsymbol{x}) &= -\Delta_h u_h(\boldsymbol{x}) = b_h(\boldsymbol{x}) \quad \text{on } \Omega_h = (0,1)^d \cap G_h, \\ u_h(\boldsymbol{x}) &= 0 \quad \text{on } \Gamma_h = ([0,1]^d \setminus \Omega_h) \cap G_h, \end{aligned} \tag{1}$$

with $G_h := \{\boldsymbol{x} = h\boldsymbol{j},\ \boldsymbol{j} \in \mathbb{Z}^d\}$. Equation (1), in stencil notation, looks like

$$A_h u_h(\boldsymbol{x}) = \sum_{\boldsymbol{\kappa} \in J} a_{\boldsymbol{\kappa}} u_h(\boldsymbol{x} + \boldsymbol{\kappa} h) = b_h(\boldsymbol{x}) \ \text{ on } \Omega_h, \tag{2}$$

with $J = \{(-1,0,0),(1,0,0),(0,-1,0),(0,1,0),(0,0,-1),(0,0,1),(0,0,0)\}$ ($d = 3$). We combine stencil, operator, and matrix notation. Multigrid is commonly explained with operators (see, for example, [19], [1]), whereas Krylov subspace methods are often described with matrices. An operator is denoted by a subscript "h" (K_h, Δ_h, M_h^{2h}, etc.) and the corresponding Fourier matrix representation is denoted by a tilde ($\widetilde{K}$, $\widetilde{M}$ etc.).

A well-known efficient multigrid method (see [19], [20]) for problem (1) consists of a red-black Gauss–Seidel smoothing method (GS-RB), full weighting of residuals to obtain the right-hand side on the coarse grids, bi- or tri-linear interpolation (in 2D and 3D, respectively) of corrections from coarse to fine grids, and the coarse grid discretization of the PDE on a grid with meshsize $2h$ in each direction. We call this method the red-black multigrid solver or MG-RB.

For the d-dimensional problem (1) and MG-RB, it is possible to apply rigorous two-grid Fourier analysis, as it was presented for 2D problems in [19] and used for 3D (Poisson-type) problems in [20]. We repeat the 3D analysis in some detail here.

The existence of the basis of discrete eigenfunctions

$$\begin{aligned} \varphi_h^{k,\ell,m}(x,y,z) &= \sin k\pi x \sin \ell\pi y \sin m\pi z, \\ \text{with } \ k,\ell,m &= 1,\ldots,n-1 \ \ ((x,y,z) = \boldsymbol{x} \in \Omega_h) \end{aligned} \tag{3}$$

for the operator (1) is crucial for the 3D discrete Fourier analysis. The scaled basis functions $\varphi_h^{k,\ell,m}(\boldsymbol{x})$ generate the space of all grid functions, $F(\Omega_h)$, and are orthogonal with respect to the discrete inner product on Ω_h:

$$(v_h, w_h) := h^3 \sum_{\boldsymbol{x} \in \Omega_h} v_h(\boldsymbol{x})\ w_h(\boldsymbol{x}) \quad \text{with } v_h, w_h \in F(\Omega_h).$$

The discrete solution u_h and the current approximation u_j can be represented by linear combinations of the basis functions $\varphi_h^{k,\ell,m}(\boldsymbol{x})$. The same holds for the error $v_{j-1} = u_{j-1} - u_h$ before and $v_j = u_j - u_h$ after the jth two-grid cycle. The error v_{j-1} is transformed by a two-grid cycle as follows:

$$v_j = M_h^{2h}\, v_{j-1} \quad \text{with} \quad M_h^{2h} = S_h^{\nu_2} K_h^{2h} S_h^{\nu_1} = S_h^{\nu_2}(I_h - I_{2h}^h(\Delta_{2h})^{-1} I_h^{2h}\Delta_h) S_h^{\nu_1}, \tag{4}$$

where S_h is the smoothing operator, ν_1 and ν_2 indicate the number of pre- and post-smoothing iterations, K_h^{2h} is the coarse grid correction operator, Δ_{2h} the discretization of the operator on a $2h$ coarse grid, I_{2h}^h and I_h^{2h} are transfer operators from coarse to fine grids, and vice versa.

In case of standard multigrid coarsening ($H = 2h$) in 3D, it is convenient to divide $F(\Omega_h)$ into a direct sum of (at most) eight-dimensional subspaces, the $2h$-*harmonics* [20]:

$$\begin{aligned} E_h^{k,\ell,m} = \operatorname{span}\Big[&\varphi_h^{k,\ell,m}, \varphi_h^{n-k,n-\ell,n-m}, \varphi_h^{n-k,\ell,m}, \varphi_h^{k,n-\ell,n-m}, \varphi_h^{k,n-\ell,m}, \\ &\varphi_h^{n-k,\ell,n-m}, \varphi_h^{k,\ell,n-m}, \varphi_h^{n-k,n-\ell,m}\Big] \quad \text{for} \quad k,\ell,m = 1,\ldots,\frac{n}{2}. \end{aligned} \tag{5}$$

This distinction is motivated by the observation that the *low-frequency* harmonic $\varphi_h^{k,\ell,m}$ ($1 \le k,\ell,m \le \frac{n}{2}$) is also visible on the coarse grid Ω_{2h}, whereas the (at most) seven corresponding *high-frequency* harmonics coincide (up to their sign) with $\varphi_h^{k,\ell,m}$. If one, two, or three of the indices equal $n/2$, the dimension of $E_h^{k,l,m}$ is four, two, or one, respectively. The coarse grid correction operator K_h^{2h} leaves the (at most) eight-dimensional spaces of $2h$-*harmonics* $E_h^{k,\ell,m}$ with an arbitrary $k,\ell,m = 1,\ldots,\frac{n}{2}$ invariant; see [20]:

$$K_h^{2h} \;:\; E_h^{k,\ell,m} \to E_h^{k,\ell,m} \quad k,\ell,m = 1,\ldots,\frac{n}{2}\,.$$

This is a consequence of the following relations of the transfer and discretization operators:

$$\begin{aligned} \Delta_h \quad &: \quad \operatorname{span}\,[\varphi_h^{k,\ell,m}] \to \operatorname{span}\,[\varphi_h^{k,\ell,m}], \\ I_h^{2h} \quad &: \quad E_h^{k,\ell,m} \to \operatorname{span}\,[\varphi_{2h}^{k,\ell,m}], \\ \Delta_{2h}^{-1} \quad &: \quad \operatorname{span}\,[\varphi_{2h}^{k,\ell,m}] \to \operatorname{span}\,[\varphi_{2h}^{k,\ell,m}], \\ I_{2h}^h \quad &: \quad \operatorname{span}\,[\varphi_{2h}^{k,\ell,m}] \to E_h^{k,\ell,m}. \end{aligned}$$

Because of this invariance property, the corresponding matrix representation of K_h^{2h} with respect to the spaces $E_h^{k,\ell,m}$ leads to a block-diagonal matrix $\widetilde{K}$:

$$\widetilde{K} := \Big[\widehat{K}(k,\ell,m)\Big]_{k,\ell,m=1,\ldots,\frac{n}{2}} \triangleq K_h^{2h} \quad \text{with} \quad \widehat{K}(k,\ell,m) \triangleq K_h^{2h}|_{E_h^{k,\ell,m}}.$$

The same invariance property is true for the GS-RB relaxation, $S_h = S_h^{BLACK} \cdot S_h^{RED}$ [20], where S_h^{RED} and S_h^{BLACK} are the partial step operators, leading to a matrix $\widetilde{S}$:

$$\widetilde{S} := \Big[\widehat{S}(k,\ell,m)\Big]_{k,\ell,m=1,\ldots,\frac{n}{2}} \triangleq S_h \quad \text{with} \quad \widehat{S}(k,\ell,m) \triangleq S_h|_{E_h^{k,\ell,m}}.$$

Hence, M_h^{2h} is orthogonally equivalent to a block matrix $\widetilde{M}$ consisting of (at most) 8×8 blocks (see [19], [20]):

$$\widetilde{M} := \Big[\widehat{M}(k,\ell,m)\Big]_{k,\ell,m=1,\ldots,\frac{n}{2}} \triangleq M_h^{2h} \quad \text{with} \quad \widehat{M}(k,\ell,m) \triangleq M_h^{2h}|_{E_h^{k,\ell,m}}.$$

The tilde $\widetilde{\ }$ indicates the *Fourier matrix representation* of the related operator. By M we denote the matrix representation of the multigrid operator M_h^{2h}, and by A we denote the matrix representation of Δ_h with respect to the Euclidean basis. We have

$$\widetilde{M} = UMU^{-1} \tag{6}$$

with a unitary matrix U. Eight consecutive rows of U are given by the $(n-1)^3$-dimensional orthogonal eigenvectors of A related to the eigenfunctions from (5):

$$\begin{aligned} &\varphi^{k,\ell,m},\ \varphi^{n-k,n-\ell,n-m},\ \varphi^{n-k,\ell,m},\ \varphi^{k,n-\ell,n-m},\ \varphi^{k,n-\ell,m},\ \varphi^{n-k,\ell,n-m},\\ &\varphi^{k,\ell,n-m},\ \varphi^{n-k,n-\ell,m} \quad \text{for} \quad k,\ell,m = 1,\ldots,\frac{n}{2}, \quad \text{with, for example,} \end{aligned} \tag{7}$$

$$\varphi^{k,\ell,m} = (\ \sin k\pi h \sin \ell\pi h \sin m\pi h, \ldots, \sin k\pi(1-h) \sin \ell\pi(1-h) \sin m\pi(1-h)\ )^T.$$

Using discrete Fourier frequencies $\boldsymbol{\theta} = \pi h(k,\ell,m)$ $(k,\ell,m = 1,\ldots,n-1)$, this gives

$$A\,\varphi^{k,\ell,m} = \lambda(\boldsymbol{\theta},h)\,\varphi^{k,\ell,m} \quad \text{with} \quad \lambda(\boldsymbol{\theta},h) = \sum_{\boldsymbol{\kappa}\in J} a_{\boldsymbol{\kappa}} \cos(\boldsymbol{\theta}\boldsymbol{\kappa}). \tag{8}$$

The dimension of $\widetilde{M}$ is given by $(n/2-1)^3 \cdot 8 + 3(n/2-1)^2 \cdot 4 + 3(n/2-1)\cdot 2 + 1\cdot 1 = (n-1)^3$ due to the dimensions of the *2h-harmonics* (5). The representation of $\widehat{M}(k,\ell,m) = \widehat{S}^{\nu_2}(k,\ell,m)\cdot\widehat{K}(k,\ell,m)\cdot\widehat{S}^{\nu_1}(k,\ell,m)$ $(k,\ell,m = 1,\ldots,\frac{n}{2})$ is obtained by straightforward generalization of the 2D case in [19]. From the block matrices $\widehat{M}(k,\ell,m)$, we obtain the two-grid convergence factor by

$$\rho_F := \rho(\widetilde{M}) = \max_{1\le k,\ell,m\le \frac{n}{2}} \rho(\widehat{M}(k,\ell,m)). \tag{9}$$

It is the asymptotic two-grid convergence which is a theoretical estimate of the multigrid convergence. Furthermore, it is possible to determine the whole spectrum and thus the eigenvalue distribution of $\widetilde{M}$ by calculating all eigenvalues of $\widehat{M}(k,\ell,m)$ for $k,\ell,m = 1,\ldots,\frac{n}{2}$. We use this distribution in section 3 for the analysis of the convergence of the multigrid preconditioned GMRES(m) method.

The *smoothing factor* μ_F is based on the ideal coarse grid correction operator Q_h^{2h} from [19] which annihilates the low-frequency error components and leaves the high-frequency components unchanged. Q_h^{2h} is a projection operator onto the space of high frequencies represented by a block-diagonal matrix $\widetilde{Q}$ consisting of (at most) 8×8 diagonal blocks: $\widehat{Q} = \mathrm{diag}(0,1,1,1,1,1,1,1)$. This leads to the following definition of the smoothing factor [19]:

$$\mu_F := \rho(\widetilde{S}^{\nu_2}\widetilde{Q}\widetilde{S}^{\nu_1}) = \rho(\widetilde{Q}\widetilde{S}^{\nu_1+\nu_2}) = \max_{1\le k,\ell,m\le \frac{n}{2}} \rho(\widehat{Q}\widehat{S}(k,\ell,m)^{\nu_1+\nu_2}). \tag{10}$$

The smoothing factor indicates the smoothing effect of the relaxation method under consideration. μ_F is a realistic estimate of the two-grid convergence factor ρ_F as long as the ideal coarse grid correction operator is a good approximation of K_h^{2h}.

2.1. Fourier results for multigrid. In [23] it was shown that an overrelaxation parameter ($\omega > 1$) improves the smoothing properties of GS-RB (leading to ω-GS-RB) and therefore improves the convergence of MG-RB for d-dimensional Poisson-type equations. The extra computational work for performing the overrelaxation in the smoother is worthwhile if $d = 3$. With the 3D Fourier two-grid and smoothing analysis explained above, we are able to estimate this convergence improvement for (1) very accurately; see Table 1. The average numerical convergence rate ρ_h^{mg} from MG-RB (after 100 multigrid iterations) is compared to ρ_F in Table 1. The number of multigrid levels used is 5, 6, and 6, respectively, for the 32^3, 64^3, and 96^3 problems.

TABLE 1
W(1,1)-cycle multigrid convergence ρ_h^{mg}, predicted convergence ρ_F, and smoothing factor μ_F for the 3D Poisson equation.

	$\omega = 1$			$\omega = 1.1$			$\omega = 1.15$ [23]		
Grid	ρ_h^{mg}	ρ_F	μ_F	ρ_h^{mg}	ρ_F	μ_F	ρ_h^{mg}	ρ_F	μ_F
32^3	0.192	0.194	0.194	0.089	0.091	0.090	0.070	0.072	0.088
64^3	0.196	0.197	0.197	0.091	0.092	0.092	0.074	0.074	0.088
96^3	0.196	0.198	0.197	0.091	0.093	0.093	0.074	0.075	0.088

In [23] it was shown furthermore that ω-GS-RB also improves the convergence factor of the 2D anisotropic diffusion equation,

$$\begin{aligned} A_h u_h(\boldsymbol{x}) &= \frac{1}{h^2}\begin{bmatrix} & -1 & \\ -\varepsilon & 2\varepsilon+2 & -\varepsilon \\ & -1 & \end{bmatrix} u_h(\boldsymbol{x}) = b_h(\boldsymbol{x}) \ \text{ on } \Omega_h, \\ u_h(\boldsymbol{x}) &= 0 \quad \text{on } \Gamma_h. \end{aligned} \tag{11}$$

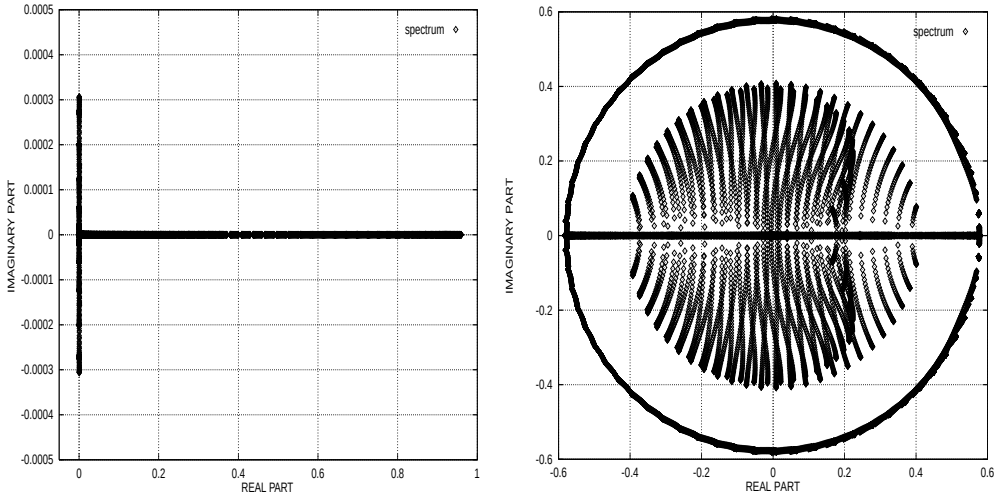
For moderate values of ε, like $100 \geq \varepsilon \geq 0.01$, the extra computational work pays off. Anisotropic problems can be solved with line smoothers in the multigrid process [19], by which much better convergence rates are obtained. However, line smoothers cannot be applied to unstructured grids in a straightforward way and their parallelization can be expensive. Here, we investigate the possibility to improve a multigrid method with a point smoother for moderate anisotropies with GMRES(m) acceleration. (Results are presented in the next sections.)

The rigorous Fourier analysis also applies for (11). In 2D we have to deal with (at most) four-dimensional $2h$-harmonics (see (5)) and therefore $\widetilde{M}$ consists of (at most) 4×4 blocks [19]. We perform numerical tests with $\varepsilon = 0.1$ and $\varepsilon = 0.01$. In both cases a large overrelaxation parameter (ω_{opt}=1.41, 1.76) has to be selected (for these values, see [23]) to obtain an impressive multigrid convergence improvement. The numerical multigrid convergence rates ρ_h^{mg} and the Fourier values μ_F and ρ_F are shown in Table 2. The W-cycle with $\nu_1 = \nu_2 = 1$ (W(1,1)) is chosen with seven multigrid levels for problem (11) discretized on a 128×128 grid.

TABLE 2
W(1,1)-cycle multigrid convergence ρ_h^{mg}, predicted convergence ρ_F, and smoothing factor μ_F for the 2D anisotropic diffusion equation, $h = 1/128$.

	$\omega = 1.0$			$\omega = \omega_{opt}$		
ε	ρ_h^{mg}	ρ_F	μ_F	ρ_h^{mg}	ρ_F	μ_F
$\varepsilon = 0.1$	0.679	0.682	0.682	0.193	0.210	0.219
$\varepsilon = 0.01$	0.957	0.960	0.960	0.566	0.583	0.590

In Figure 1, we show the effect of an overrelaxation parameter $\omega = 1.76$ on the eigenvalue distribution of the two-grid iteration matrix for the anisotropic ($\varepsilon = 0.01$) diffusion problem. In contrast to the almost real-valued spectrum for $\omega = 1.0$, which is uniformly distributed on the interval $[0.0, 0.960]$, we find complex eigenvalues clustered around the origin and uniformly distributed on a circle by the overrelaxation.



1 (a). $\omega = 1.0$ 1 (b). $\omega = 1.76$

FIG. 1. *Eigenvalue spectra of the red-black two-grid 2D anisotropic diffusion solver from rigorous Fourier analysis;* $\nu_1 = \nu_2 = 1, \varepsilon = 0.01, h = 1/128$. (*Note the different scaling of the axes in Figure* 1 (a).)

From the comparison of the predicted convergence rates ρ_F and the numerically obtained ρ_h^{mg}, it is clear that the two-grid Fourier analysis provides excellent quantitative estimates for the actual multigrid convergence of MG-RB for Poisson-type equations. These results encourage the following considerations concerning Fourier analysis of the multigrid preconditioned GMRES(m).

3. Fourier analysis of GMRES(m) preconditioned by multigrid. In this section we discuss two approaches for the analysis of GMRES(m) preconditioned by multigrid. For both approaches it is crucial that the discrete Fourier analysis uses unitary basis transformations (6) which simplify the representation of the multigrid iteration matrix.

We consider the linear system related to (2),

$$Au = b. \tag{12}$$

A two-grid (or more generally, a multigrid) cycle for solving (12) is described by the following matrix splitting:

$$Cu_j + (A - C)u_{j-1} = b, \tag{13}$$

where u_j represents a new and u_{j-1} a previous approximation. This formulation is equivalent to

$$u_j = u_{j-1} + C^{-1}(b - Au_{j-1}), \qquad r_j = (I - AC^{-1})r_{j-1}, \tag{14}$$

with r_j, r_{j-1} the residual vectors and $I - AC^{-1}$ represents the two-grid (multi-grid) iteration matrix.

The multigrid method (13) considered was described in the previous section. It is used as a *right preconditioner* for GMRES(m). The parameter m represents the number of search vectors stored after which the GMRES(m) algorithm will restart. GMRES(m) searches for a solution u_j in the following subspace:

$$\begin{aligned} C(u_j - u_{j-m}) &\in \text{span}[r_{j-m}, (AC^{-1})r_{j-m}, \ldots, (AC^{-1})^{m-1}r_{j-m}] \\ &=: K^m(AC^{-1}, r_{j-m}), \end{aligned} \tag{15}$$

where $K^m(AC^{-1}, r_{j-m})$ is the Krylov subspace. It selects a new approximation u_j by minimizing the corresponding residual r_j in the L_2-norm

$$\min_{C(u_j - u_{j-m}) \in K^m(AC^{-1}, r_{j-m})} ||b - Au_j||_2. \tag{16}$$

Any element w in the affine subspace $u_{j-m} + C^{-1}K^m(AC^{-1}, r_{j-m})$ (see (14), (15)) can be represented by

$$w = u_{j-m} + C^{-1}(\alpha_1 r_{j-m} + \alpha_2 AC^{-1}r_{j-m} + \cdots + \alpha_m (AC^{-1})^{m-1} r_{j-m}). \tag{17}$$

By substituting (17) into the residual equation, we obtain

$$\begin{aligned} r_w &= b - Aw \\ &= r_{j-m} - \alpha_1 AC^{-1}r_{j-m} - \alpha_2 (AC^{-1})^2 r_{j-m} - \cdots - \alpha_m (AC^{-1})^m r_{j-m} \\ &= P_m(AC^{-1}) r_{j-m}, \end{aligned} \tag{18}$$

where P_m is an mth order polynomial defined by $P_m(\lambda) = 1 - \sum_{k=1}^{m} \alpha_k \lambda^k$. Comparing (16) and (18), we see that the L_2-norm of r_j satisfies the following minimization property:

$$||r_j||_2 = \min\{||P_m(AC^{-1})r_{j-m}||_2 \ \mid \ P_m \in \mathbb{P}_m\}, \tag{19}$$

where $\mathbb{P}_m$ denotes the set of all polynomials of degree at most m with $P_m(0) = 1$.

Since we are interested in the convergence of multigrid preconditioned GMRES with a restart after m iterations, we consider the residuals $r_m, r_{2m}, \ldots, r_{i \cdot m}$. This leads to the following definition of the reduction factors ρ_i for a *complete* ith iteration, consisting of m multigrid preconditioned GMRES(m) steps:

$$\rho_i := \left[\left(\frac{||r_{i \cdot m}||_2}{||r_0||_2} \right)^{1/m} \right]^{1/i}.$$

In subsection 3.1, we review estimates for ρ_i which are based on an analysis of the spectrum σ of AC^{-1} (see [5], [6], [15], [16], [17]). In general, these estimates exhibit a qualitative character, and concrete values are rarely found in the literature because the spectrum is usually not available. However, with the Fourier analysis these estimates can be calculated explicitly since the spectrum is easily obtained if multigrid is used as a preconditioner. In subsection 3.2 we introduce a sharper estimate, which is based on an analysis of the GMRES(m) polynomial.

3.1. Analysis with the spectrum of the iteration matrix. The usual way to analyze the convergence of GMRES is to exploit information about the spectrum of the iteration matrix AC^{-1}. If we compare the error transformation (4) and the residual transformation (14) by a two-grid cycle, the Fourier matrix $\widetilde{M}$ (6) is related to the two-grid iteration matrix $I - AC^{-1}$ via the identity

$$\widetilde{M} = UA^{-1}(I - AC^{-1})AU^{-1} \Longleftrightarrow \widetilde{M} = I - UC^{-1}AU^{-1}, \tag{20}$$

with the unitary matrix U from (6). With (20) we have

$$UAC^{-1}U^{-1} = I - UAU^{-1}\widetilde{M}(UAU^{-1})^{-1} = I - \widetilde{A}\widetilde{M}\widetilde{A}^{-1}. \tag{21}$$

$\widetilde{A}$ and $\widetilde{A}^{-1}$ are diagonal matrices since the rows of U are eigenvectors of A (see (7) and section 2). This means that $UAC^{-1}U^{-1}$ is a block-diagonal matrix which consists of (at most) 8×8 blocks in the 3D case and of (at most) 4×4 blocks in the 2D case. The eigenvalues of these block matrices are easily calculated.

We are dealing with positive real iteration matrices AC^{-1} (due to the multigrid preconditioner), so the symmetric part of AC^{-1} is positive definite. It is shown in [6], [5] that GMRES(m) always converges for this type of matrices. More generally, it is sufficient that the *field of values* of the iteration matrix $\{\bar{x}^T AC^{-1}x/\bar{x}^T x : x \in \mathbb{C}^{(n-1)^d}\}$ lies in an open half-plane $\{z : \mathrm{Re}(e^{-i\theta}z) > 0\}$. This is called the *half-plane condition* [12]. In particular the following error bound can be established for any restart parameter m and iteration index i [6]:

$$||r_{i\cdot m}||_2 \leq (1 - \alpha/\beta)^{m/2}||r_{(i-1)m}||_2, \tag{22}$$

with $\alpha = (\,\lambda_{min}(\frac{1}{2}(AC^{-1} + (AC^{-1})^T))\,)^2$ and $\beta = \lambda_{max}(\,(AC^{-1})^T AC^{-1}\,)$. Using (22), we obtain

$$||r_{i\cdot m}||_2 \leq (1 - \alpha/\beta)^{i\cdot m/2}||r_0||_2 \quad \Longrightarrow \quad \rho_i \leq (1 - \alpha/\beta)^{1/2} =: \rho_{hpc} < 1. \tag{23}$$

In general, the m- and i-independent value ρ_{hpc} is not a realistic quantitative estimate of ρ_i, but the residual bound from (23) is an important qualitative result because it ensures the convergence of GMRES(m) for all m.

Next, we discuss considerations on the convergence of GMRES(m) from [15], [16], [17], which are also based on the spectrum of the iteration matrix. From (19) it follows that

$$||r_{i\cdot m}||_2 \leq ||P_m(AC^{-1})||_2\,||r_{(i-1)m}||_2 \quad \text{for all} \quad P_m \in \mathbb{P}_m. \tag{24}$$

Suppose that all eigenvalues of AC^{-1} are located in an ellipse $E(c, d, a)$ which excludes the origin. c denotes the center, d the focal distance, and a the major semi-axis. Then it is known [17] that the absolute value of the polynomial

$$t_m(z) := T_m\Big(\frac{c}{d} - \frac{1}{d}\,z\Big)/T_m\Big(\frac{c}{d}\Big) = T_m(\widehat{z})/T_m\Big(\frac{c}{d}\Big) \quad \text{with} \quad z,\, \widehat{z} := \Big(\frac{c}{d} - \frac{1}{d}\,z\Big) \in \mathbb{C}$$

is small on the spectrum of AC^{-1}. Here, T_m represents the Chebychev polynomial of degree m of the first kind which is given by the three-term recurrence relation

$$T_0(\widehat{z}) = I\,, \quad T_1(\widehat{z}) = \widehat{z}\,, \quad T_{m+1}(\widehat{z}) = 2T_m(\widehat{z}) - T_{m-1}(\widehat{z}) \quad \text{for} \quad m \geq 1.$$

If AC^{-1} is diagonalizable, $AC^{-1} = XDX^{-1}$; then (24) yields

$$||r_{i\cdot m}||_2 \leq ||t_m(AC^{-1})||_2 \, ||r_{(i-1)m}||_2 \leq \left[||t_m(AC^{-1})||_2\right]^i \, ||r_0||_2 \tag{25}$$

$$\leq \left[||X||_2 \, ||X^{-1}||_2 \max_{\lambda_j \in \sigma} |t_m(\lambda_j)|\right]^i \, ||r_0||_2 \tag{26}$$

$$\leq \left[\kappa_2(X)\, T_m\left(\frac{a}{d}\right) / T_m\left(\frac{c}{d}\right)\right]^i \, ||r_0||_2, \tag{27}$$

where $\kappa_2(X)$ denotes the spectral condition number of the transformation matrix X (see [16], [17]). Inequality (27) uses the fact that the maximum modulus of a complex analytical function is attained on the boundary of the domain, which is due to Liouville's theorem [16]. Heuristically, we expect this inequality to be sharp if the interior of the ellipse $E(c, d, a)$ is well covered by the spectrum σ. Using (25) and (27), we obtain the following well-known estimates [16], [17] for ρ_i:

$$\rho_i \leq N_m^E \leq (\kappa_2(X))^{1/m} \, T_m^E, \tag{28}$$

$$\text{with} \quad N_m^E := (||t_m(AC^{-1})||_2)^{1/m} \quad \text{and} \quad T_m^E := \left(T_m\left(\frac{a}{d}\right) / T_m\left(\frac{c}{d}\right)\right)^{1/m}. \tag{29}$$

T_m^E can be further approximated by

$$T_m^E = \left(\frac{\left(\frac{a}{d} + \sqrt{(\frac{a}{d})^2 - 1}\right)^m + \left(\frac{a}{d} + \sqrt{(\frac{a}{d})^2 - 1}\right)^{-m}}{\left(\frac{c}{d} + \sqrt{(\frac{c}{d})^2 - 1}\right)^m + \left(\frac{c}{d} + \sqrt{(\frac{c}{d})^2 - 1}\right)^{-m}} \right)^{1/m} \approx \frac{a + \sqrt{a^2 - d^2}}{c + \sqrt{c^2 - d^2}}. \tag{30}$$

Approximation (30) is sharp if $a \gg d$ and $c \gg d$ or $m \gg 1$ holds. The first inequality is fulfilled if the ellipse $E(c, d, a)$ tends to a circle which means that the focal distance d becomes very small compared to a (in our tests it is usually sufficient if approximately $a \geq 2d$). In all cases considered, the multigrid preconditioner ensures that the spectrum σ is clustered around 1, which means that the second inequality is satisfied. If (30) is sharp due to a circular shape of σ or due to a large m, T_m^E is independent of the restart parameter m. In this case, it is difficult to further accelerate the multigrid method. A large Krylov subspace (15) does not lead to additional acceleration.

The main drawback of the estimates (28) for the reduction factors ρ_i is their i-independence. The relative amount of residual decrease from early GMRES cycles compared to later ones depends on both the matrix and the initial residual. The reduction by early cycles may be very different from the reduction by later ones. More precisely, the estimates from (28) are based on the worst previous residual that is possible due to the definition of the matrix norm $||A||_2 := \sup\{||Ar||_2 / ||r||_2 : r \in \mathbb{C}^{(n-1)^d}\}$. Thus, in general they cannot be expected to be sharp approximations for large i because $r_{(i-1)m}$ can be completely different from the worst residual since many iterations have already been performed and the worst search directions have already been removed from the current approximation $u_{(i-1)m}$ by GMRES(m). Furthermore, $\kappa_2(X)$ can be very large if the iteration matrix AC^{-1} is far from normal. But it is found in [12] that the condition number $\kappa_2(X)$ is not always relevant for the convergence of GMRES. In [17] it is stated that (28) is an asymptotic result and that the actual residual norm should behave like T_m^E without $\kappa_2(X)$. Therefore, the explicit values for T_m^E are also given in subsection 3.3 and compared with the asymptotic numerical convergence results. However, T_m^E is a heuristic estimate for $\rho_{i \gg 1}$ and not an upper bound like N_m^E.

3.2. Analysis with the GMRES-polynomial. To obtain a quantitative result for the convergence of GMRES(m), we present a second approach to estimate ρ_i, which explicitly depends on the iteration index i. In this way, it is possible to take the benefits of the previous iterations into account, in contrast to the above estimates. An initial residual r_0 is prescribed, and the GMRES(m)-polynomials P_m^i are explicitly determined for every i. The choice of r_0 does not influence the average reduction factor for $i \gg 1$ in which we are interested.

For any unitary matrix U, and in particular for U from (6), we have $||r_{i\cdot m}||_2 = ||Ur_{i\cdot m}||_2$. In order to find the coefficients α_k^i $(k = 1, \ldots, m)$ of the GMRES(m)-polynomials P_m^i, the following function f is minimized:

$$\begin{aligned} f(\alpha_1^i, \ldots, \alpha_m^i) :=& \Big(UP_m^i(AC^{-1})r_{(i-1)m}, UP_m^i(AC^{-1})r_{(i-1)m} \Big) \\ =& \Big(P_m^i(UAC^{-1}U^{-1})Ur_{(i-1)m}, P_m^i(UAC^{-1}U^{-1})Ur_{(i-1)m} \Big) \\ =& \sum_{l,k=1}^{m} \alpha_l^i \alpha_k^i \Big((UAC^{-1}U^{-1})^l Ur_{(i-1)m}, (UAC^{-1}U^{-1})^k Ur_{(i-1)m} \Big) \\ &+ 2\sum_{l=1}^{m} \alpha_l^i \Big((UAC^{-1}U^{-1})^l Ur_{(i-1)m}, Ur_{(i-1)m} \Big). \end{aligned}$$

The α_k^i are obtained by solving the following linear system of equations:

$$\begin{aligned} \frac{\partial f}{\partial \alpha_l^i} =& 2\sum_{k=1}^{m} \alpha_k^i \Big((UAC^{-1}U^{-1})^l Ur_{(i-1)m}, (UAC^{-1}U^{-1})^k Ur_{(i-1)m} \Big) \\ &+ 2\Big((UAC^{-1}U^{-1})^l Ur_{(i-1)m}, Ur_{(i-1)m} \Big) = 0 \quad \text{for } l = 1, \ldots, m. \end{aligned} \tag{31}$$

In (21) it is seen that the argument $UAC^{-1}U^{-1} = I - \widetilde{A}\widetilde{M}\widetilde{A}^{-1}$ of the GMRES(m)-polynomials P_m^i has a simple block structure. Therefore, due to the sparse structure of $(UAC^{-1}U^{-1})^l$ $(l = 1, \ldots, m)$, the solution of the linear system (31) is easily computed for every complete iteration i, as soon as the previous transformed residual $Ur_{(i-1)m}$ is given. We prescribe right-hand side $b = 0$ and a randomly chosen initial approximation Uu_0, which yield a transformed initial residual $Ur_0 = -\widetilde{A}Uu_0$. More precisely, we select an initial solution $u_0 = \sum_{k,\ell,m=1}^{n-1} \beta(k,\ell,m)\varphi^{k,\ell,m}$, where the amplitudes $\beta(k,\ell,m)$ of the eigenvectors are randomly chosen. Hence, we get

$$Uu_0 = \Big(\beta(1,1,1), \beta(n-1,n-1,n-1), \ldots, \beta\Big(\frac{n}{2}, \frac{n}{2}, \frac{n}{2}\Big) \Big)^T;$$

see (7). From Ur_0 it is possible to evaluate the α_k^1 by solving (31) for $i = 1$. This gives

$$Ur_m = P_m^1(UAC^{-1}U^{-1})Ur_0.$$

It follows that $Ur_{i\cdot m}$ is easily calculated for every i. This leads to the definition of an average reduction factor ρ_i^U obtained by the Fourier analysis:

$$\rho_i^U := \left[\left(\frac{||Ur_{i\cdot m}||_2}{||Ur_0||_2} \right)^{1/m} \right]^{1/i}. \tag{32}$$

A straightforward way to compute ρ_i^U is to run GMRES(m) directly for the prescribed residual r_0 and the transformed matrix $UACU^{-1}$. However, it turns out (see subsection 4.3) that an explicit calculation of the coefficients α_k^i $(k = 1, \ldots, m)$ of the GMRES(m)-polynomials gives an interesting insight into the behavior of GMRES(m).

In most of the tests, presented in the next subsection, ρ_i^U becomes constant for $i \geq 20$. Thus, ρ_{20}^U is expected to match numerical reference values very well and, in particular, to be sharper than the estimates from subsection 3.1.

3.3. Fourier results of multigrid preconditioned GMRES(m). Here, we investigate the quality of the theoretical considerations by comparing them to the numerical convergence of the multigrid preconditioned GMRES(m) method for the test problems from subsection 2.1. We compare the average accelerated multigrid convergence rate, ρ_h^{acc}, from numerical experiments with the theoretical prediction $\rho_{i=20}^U$, the heuristic estimate T_m^E, and the upper bound N_m^E. Insight into the behavior of the acceleration method is gained by considering different combinations of the relaxation parameter ω and the GMRES(m) restart parameter m.

In order to make sure that the asymptotic convergence behavior is observed, we present ρ_h^{acc} after 40 or after 100 multigrid preconditioned GMRES(m) iterations. These numbers of iterations match $\rho_{i=20}^U$ for $m = 2$ and for $m = 5$, respectively. However, the numerical convergence stabilizes much earlier and often the asymptotic values are obtained after only a few complete iterations.

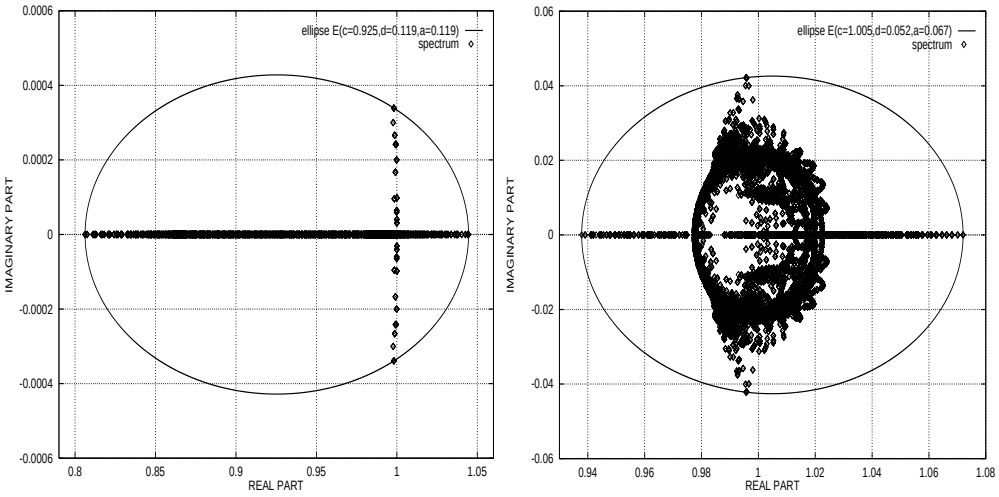
The results for the 3D Poisson equation with right-hand side $b_h(\boldsymbol{x}) = 0$ and initial guess $u_0 = (1, \ldots, 1)^T$ are presented in Table 3. We observe that $\rho_{i=20}^U$ and T_m^E provide accurate predictions of ρ_h^{acc} for different overrelaxation parameters ω. The upper bound N_m^E becomes sharper with increasing m as T_m^E and N_m^E coincide for large m, because the condition number $\kappa_2(X)$ cancels out (see (28)). Comparing Tables 1 and 3, we observe a satisfactory convergence improvement for $\omega = 1$. If overrelaxation is selected, the benefits of the GMRES(m) acceleration are small and the convergence cannot be improved significantly by using a larger Krylov space (15). This behavior is expected from the spectral picture in Figure 2. The ellipses become more orbital (note the different scaling of the axes) with the overrelaxation so that approximation (30) gets sharper.

TABLE 3

W(1,1)*-cycle multigrid preconditioned GMRES(m) convergence* ρ_h^{acc} *and convergence estimates* $\rho_{i=20}^U$, T_m^E, N_m^E *for the 3D Poisson equation,* $h = 1/32$.

Relaxation parameter	Acc. MG-RB, $m = 2$				Acc. MG-RB, $m = 5$			
	ρ_h^{acc}	$\rho_{i=20}^U$	T_m^E	N_m^E	ρ_h^{acc}	$\rho_{i=20}^U$	T_m^E	N_m^E
$\omega = 1.0$	0.085	0.090	0.091	0.186	0.070	0.072	0.074	0.115
$\omega = 1.1$	0.045	0.054	0.055	0.133	0.042	0.045	0.047	0.082
$\omega = 1.15$ [23]	0.050	0.054	0.056	0.129	0.045	0.049	0.055	0.077

Similar statements hold for the 2D anisotropic diffusion equation, where right-hand side $b_h(\boldsymbol{x}) = 0$ and initial guess $u_0 = (1, \ldots, 1)^T$ are chosen, as in the 3D case. Theoretical predictions $\rho_{i=20}^U$ match very well with ρ_h^{acc}, as can be seen from Table 4. In most cases the same observation can be made for the values T_m^E and ρ_h^{acc}. N_m^E is not a good approximation for ρ_h^{acc} but it improves with increasing m. If we choose an optimal overrelaxation parameter for $\varepsilon = 0.1$ and $\varepsilon = 0.01$, it is not really possible to further accelerate the ω-MG-RB method due to the circular shape of the corresponding spectrum σ shown in Figure 3. For the same reason further



2 (a). $\omega = 1.0$ 2 (b). $\omega = 1.15$

FIG. 2. *Eigenvalue spectra of the red-black two-grid 3D Poisson solver from Fourier analysis;* $\nu_1 = \nu_2 = 1$, $h = 1/32$. *(Note the different scaling of the axes.)*

TABLE 4

W(1,1)*-cycle multigrid preconditioned GMRES(m) convergence* ρ_h^{acc} *and convergence estimates* $\rho_{i=20}^U$, T_m^E, N_m^E *for the 2D anisotropic diffusion equation,* $h = 1/128$.

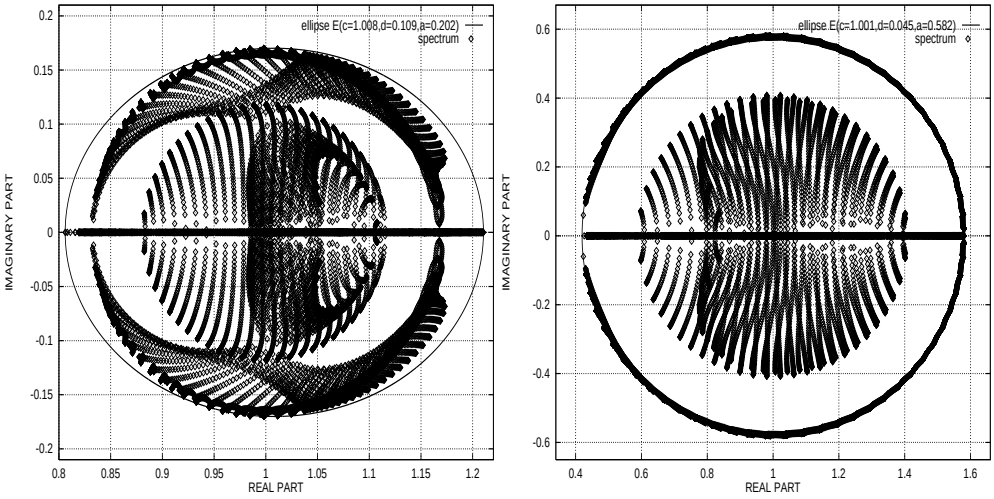
Acc.	$\varepsilon = 0.1$, $\omega = 1.0$				$\varepsilon = 0.1$, $\omega = 1.41$			
MG-RB	ρ_h^{acc}	$\rho_{i=20}^U$	T_m^E	N_m^E	ρ_h^{acc}	$\rho_{i=20}^U$	T_m^E	N_m^E
$m = 2$	0.350	0.392	0.405	0.734	0.173	0.184	0.186	0.724
$m = 5$	0.300	0.327	0.333	0.409	0.169	0.180	0.185	0.354

Acc.	$\varepsilon = 0.01$, $\omega = 1.0$				$\varepsilon = 0.01$, $\omega = 1.76$			
MG-RB	ρ_h^{acc}	$\rho_{i=20}^U$	T_m^E	N_m^E	ρ_h^{acc}	$\rho_{i=20}^U$	T_m^E	N_m^E
$m = 2$	0.795	0.800	0.861	1.460	0.568	0.579	0.580	2.385
$m = 5$	0.723	0.740	0.763	1.054	0.566	0.579	0.580	1.182

improvement cannot be achieved by a larger Krylov space. For $\omega = 1.0$ a satisfactory convergence improvement can be observed comparing Table 2 and Table 4.

As mentioned before, the Fourier analysis for multigrid is based on $\widetilde{M} = I - UC^{-1}AU^{-1}$ (see (20)), the *error* reduction in a two-grid algorithm, whereas Krylov methods minimize the *residual* norm and are based on AC^{-1}. Thus, we have for the corresponding spectra: $\sigma(AC^{-1}) = 1 - \sigma(\widetilde{M})$ (compare Figure 1 (b) and Figure 3 (b)). As a consequence the iteration matrices AC^{-1} are always positive real here because the spectrum of the multigrid iteration matrix $\sigma(\widetilde{M})$ is mainly clustered around the origin. In particular $\rho(\widetilde{M}) = \rho_F < 1$ holds because the multigrid preconditioner is already a convergent method. Thus, any eigenvalue of AC^{-1} lies in the right half-plane and we have $\rho_{hpc} < 1$.

Remark. In this connection it is interesting to note that the *residual* reduction norm $||\widetilde{A}\widetilde{M}\widetilde{A}^{-1}||_2$ decreases with an increasing number of post-relaxations ν_2 in contrast to the *error* reduction norm $||\widetilde{M}||_2$ which decreases with an increasing ν_1 [19]. According to this observation, one should choose similar values for ν_1 and ν_2 and prefer $\nu_1 \geq \nu_2$ to $\nu_1 \leq \nu_2$ when multigrid is selected as a solver [19]. The reversed



3 (a). $\varepsilon = 0.1$, $\omega = 1.41$ 3 (b). $\varepsilon = 0.01$, $\omega = 1.76$

FIG. 3. *Eigenvalue spectra of the red-black two-grid* $2D$ *anisotropic diffusion solver from Fourier analysis;* $\nu_1 = \nu_2 = 1$, $h = 1/128$.

statement (prefer $\nu_1 \leq \nu_2$ to $\nu_1 \geq \nu_2$) holds if we apply multigrid as a preconditioner for a Krylov acceleration technique.

From Tables 3 and 4, it follows that a significant acceleration of the ω-MG-RB method with the help of GMRES(m) is possible if $\omega = 1.0$. This acceleration is observed if we compare the numerical results ρ_h^{mg} and ρ_h^{acc} from subsections 2.1 and 3.3 and if we compare the Fourier estimates ρ_F and $\rho_{i=20}^U$. If, in some sense, an *optimal multigrid method* is selected by using the optimal overrelaxation parameter, a further improvement is almost impossible.

$\rho_{i=20}^U$ can be considered as a quantitative value which is found to be close to the numerical convergence rate ρ_h^{acc}. The heuristic estimate T_m^E gives good insight into the asymptotic convergence behavior, whereas N_m^E is too pessimistic an approximation in general. The qualitative value ρ_{hpc} is important for theoretical purposes, but it doesn't help in choosing between solution methods. In most cases ρ_{hpc} is close to 1 and far away from ρ_h^{acc}.

4. Local mode analysis of multigrid preconditioned GMRES(m). The Fourier analysis as presented in section 2 is restricted to model operators A_h, for which the discrete functions (3) form a basis of eigenfunctions. Furthermore, it is not possible to apply this analysis for multigrid methods based on Gauss–Seidel relaxations with a lexicographical ordering of the grid points (GS-LEX). The eigenfunctions (3) are no longer invariant under the GS-LEX operators. Instead they are intermixed, which leads to a full matrix $\widetilde{S}$. For a general linear operator A_h with constant or frozen coefficients it is, however, possible to use the *local mode Fourier analysis* [1] for analyzing multigrid methods based on GS-LEX as the smoother. The theoretical foundations (for example, presented in [18], [2]) for the local mode analysis are not as straightforward as they are for the analysis from section 2. Basically the local mode analysis is valid if the influence of a domain boundary is negligible [2]. Often, this requirement can be satisfied by performing extra local boundary relaxation. In practice, the local mode (or infinite grid) two-grid and smoothing analysis gives satisfactory

sharp estimates of the actual multigrid convergence.

Instead of considering the finite domain with discrete eigenfunctions (3) of A_h an infinite domain $\overline{\Omega}_h := \{\boldsymbol{x} = (j_x h, j_y h, j_z h) : j_x, j_y, j_z \in \mathbb{Z}\}$ is considered with continuous eigenfunctions

$$\phi_h(\boldsymbol{\theta}, \boldsymbol{x}) = e^{i\boldsymbol{x}\boldsymbol{\theta}/h} = e^{i\boldsymbol{j}\boldsymbol{\theta}} = e^{i(j_x\theta_x + j_y\theta_y + j_z\theta_z)}, \tag{33}$$

where the Fourier frequencies $\boldsymbol{\theta} = (\theta_x, \theta_y, \theta_z)$ vary continuously in $\mathbb{R}^3$. The corresponding eigenvalues are given by $\lambda(\boldsymbol{\theta}, h) = \sum_{\boldsymbol{\kappa}\in J} a_{\boldsymbol{\kappa}} e^{i\boldsymbol{\theta}\boldsymbol{\kappa}}$ (see (8)). Fourier components with $|\boldsymbol{\theta}| := \max\{|\theta_x|, |\theta_y|, |\theta_z|\} \geq \pi$ are not visible on $\overline{\Omega}_h$, since they coincide with components $e^{i\boldsymbol{j}\hat{\boldsymbol{\theta}}}$ where $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} (\mathrm{mod}\, \pi)$. Therefore, the Fourier space

$$\boldsymbol{\varepsilon}^h = \mathrm{span}\{e^{i\boldsymbol{j}\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta = (-\pi, \pi]^3\}$$

contains any infinite grid function on $\overline{\Omega}_h$ [19]. As in the discrete Fourier analysis, we divide the Fourier space $\boldsymbol{\varepsilon}^h$ into eight-dimensional subspaces $\boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}}$, which consist of one low-frequency harmonic ($\boldsymbol{\theta} \in \Theta^l := (-\pi/2, \pi/2]^3$) and seven corresponding high-frequency harmonics ($\boldsymbol{\theta} \in \Theta^h := \Theta \backslash \Theta^l$) (5):

$$\boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}} = \mathrm{span}[\phi(\boldsymbol{\theta}^{\alpha_x\alpha_y\alpha_z}, \boldsymbol{x}) = e^{i\boldsymbol{j}\boldsymbol{\theta}^{\alpha_x\alpha_y\alpha_z}} : \alpha_x, \alpha_y, \alpha_z \in \{0, 1\}], \quad \text{with } \boldsymbol{x} \in \overline{\Omega}_h,$$
$$\boldsymbol{\theta}^{000} \in \Theta^l, \text{ and } \boldsymbol{\theta}^{\alpha_x\alpha_y\alpha_z} = (\theta_x - \alpha_x \mathrm{sign}(\theta_x)\pi, \theta_y - \alpha_y \mathrm{sign}(\theta_y)\pi, \theta_z - \alpha_z \mathrm{sign}(\theta_z)\pi).$$

These spaces $\boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}}$ of $2h$-harmonics are invariant under the coarse grid correction operator K_h^{2h} (4) and also under a GS-LEX relaxation operator. In order to obtain a well-defined two-grid operator we replace the Fourier space $\boldsymbol{\varepsilon}^h$ by a slightly shrunk subspace $\tilde{\boldsymbol{\varepsilon}}^h$ such that $(A_h)^{-1}$ exists and also $(A_{2h})^{-1}$ can be reasonably defined on the coarser grid, as in [19]:

$$\tilde{\boldsymbol{\varepsilon}}^h = \boldsymbol{\varepsilon}^h \backslash \bigcup_{\boldsymbol{\theta}\in\Psi} \boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}} \quad \text{with} \quad \Psi := \{\boldsymbol{\theta} : \lambda(\boldsymbol{\theta}, h) = 0 \text{ or } \lambda(2\boldsymbol{\theta}, 2h) = 0\}.$$

Hence, we have a well-defined operator M_h^{2h} (4) on $\tilde{\boldsymbol{\varepsilon}}^h$ with the invariance property

$$M_h^{2h} : \boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}} \longrightarrow \boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}} \quad \text{with} \quad \boldsymbol{\theta} \in \tilde{\Theta}^l := \Theta^l \backslash \Psi,$$

equivalent to a block-diagonal matrix $\widetilde{M}$. The spectral radius $\rho(\widetilde{M})$ of the corresponding Fourier matrix is determined similarly as in the previous Fourier analysis (9):

$$\rho_F := \rho(\widetilde{M}) = \sup_{\boldsymbol{\theta}\in\tilde{\Theta}^l} \rho(\widehat{M}(\boldsymbol{\theta})) \quad \text{with} \quad \widehat{M}(\boldsymbol{\theta}) \triangleq M_h^{2h}|_{\boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}}}. \tag{34}$$

We consider operators with $\Psi = \{(0, 0, 0)\}$, which is an ellipticity requirement [2]. Then, it is possible to keep the suprenum (34) finite by an appropriate choice of the transfer operators I_h^{2h} and I_{2h}^h [9].

The whole spectrum is obtained similarly as in section 2. We, however, always observe a spectrum from a discrete mesh, whereas we have a continuously defined set of eigenfunctions (33) in order to get h-independent upper bounds for the convergence rates.

The definition of the smoothing factor (10) carries over from the rigorous analysis. The coarse grid correction operator is replaced by an ideal operator Q_h^{2h} which is orthogonally equivalent to a block matrix $\widehat{Q}$ with blocks $\widehat{Q} = \mathrm{diag}(0, 1, 1, 1, 1, 1, 1, 1)$.

$$\mu_F := \rho(\widetilde{S}_2^{\nu_2}\widetilde{Q}\widetilde{S}_1^{\nu_1}) = \rho(\widetilde{Q}\widetilde{S}_1^{\nu_1}\widetilde{S}_2^{\nu_2}) = \sup_{\boldsymbol{\theta}\in\tilde{\Theta}^l} \rho(\widehat{Q}\widehat{S}_1(\boldsymbol{\theta})^{\nu_1}\widehat{S}_2(\boldsymbol{\theta})^{\nu_2}) \quad \text{with} \quad \widehat{S}_*(\boldsymbol{\theta}) \triangleq S_h^*|_{\boldsymbol{\varepsilon}^h_{\boldsymbol{\theta}}}.$$

TABLE 5

Multigrid convergence ρ_h^{mg}, predicted convergence ρ_F, and smoothing factor μ_F for the 2D anisotropic diffusion equation; $h = 1/128$.

Solution method	$\omega = 1.0$			$\omega = \omega_{opt}$		
ε	ρ_h^{mg}	ρ_F	μ_F	ρ_h^{mg}	ρ_F	μ_F
MG-FF, $\varepsilon = 0.1$	0.693	0.696	0.697	0.410	0.433	0.492
MG-FB, $\varepsilon = 0.1$	0.693	0.697	0.697	0.437	0.440	0.492
MG-FF, $\varepsilon = 0.01$	0.957	0.961	0.961	0.755	0.758	0.769
MG-FB, $\varepsilon = 0.01$	0.957	0.962	0.961	0.750	0.759	0.769

TABLE 6

Multigrid preconditioned GMRES(m) convergence ρ_h^{acc} and convergence estimates $\rho_{i=20}^U$, T_m^E, N_m^E for the 2D anisotropic diffusion equation; $h = 1/128$.

Acc.	$\varepsilon = 0.1$, $\omega = 1.0$				$\varepsilon = 0.1$, $\omega = 1.40$			
MG	ρ_h^{acc}	$\rho_{i=20}^U$	T_m^E	N_m^E	ρ_h^{acc}	$\rho_{i=20}^U$	T_m^E	N_m^E
FF, $m = 2$	0.475	0.515	0.529	0.911	0.410	0.420	0.430	0.524
FF, $m = 5$	0.450	0.469	0.519	0.907	0.410	0.420	0.430	0.441
FB, $m = 2$	0.370	0.402	0.408	0.718	0.185	0.185	0.203	0.445
FB, $m = 5$	0.305	0.331	0.331	0.487	0.158	0.160	0.165	0.202

Acc.	$\varepsilon = 0.01$, $\omega = 1.0$				$\varepsilon = 0.01$, $\omega = 1.75$			
MG	ρ_h^{acc}	$\rho_{i=20}^U$	T_m^E	N_m^E	ρ_h^{acc}	$\rho_{i=20}^U$	T_m^E	N_m^E
FF, $m = 2$	0.850	0.895	0.907	0.952	0.753	0.750	0.757	0.923
FF, $m = 5$	0.800	0.837	0.896	0.897	0.753	0.750	0.757	0.813
FB, $m = 2$	0.784	0.850	0.865	1.403	0.397	0.380	0.479	0.717
FB, $m = 5$	0.723	0.755	0.768	1.102	0.362	0.370	0.392	0.457

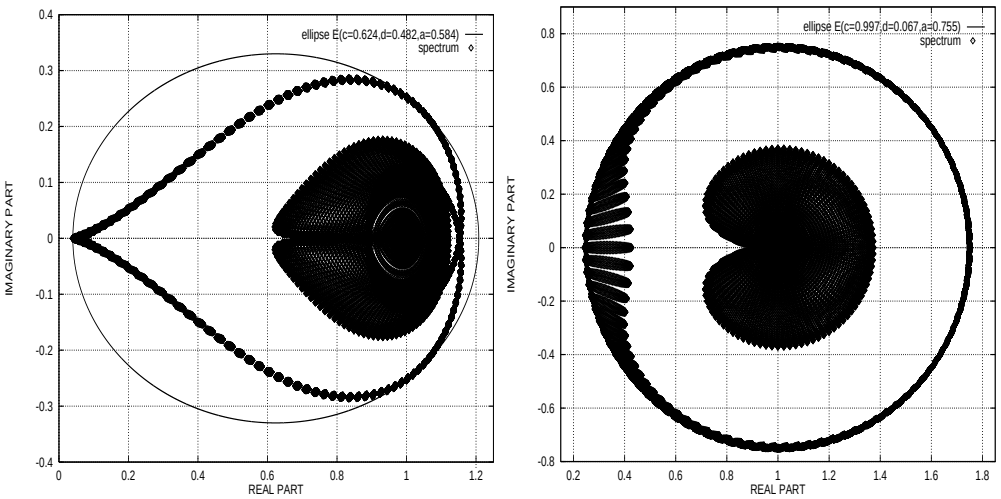
From the local mode analysis, the analytical values $\rho_{i=20}^U$, T_m^E, N_m^E, and ρ_{hpc} for the multigrid preconditioned GMRES(m) method are obtained by straightforward modifications of the definitions (32), (29), and (23) taken from the rigorous case.

4.1. Results for Poisson-type equations. The first example of the local-mode Fourier analysis of multigrid preconditioned GMRES(m) deals again with the anisotropic diffusion equation (2). We analyze the difference in convergence between performing two smoothing iterations ($\nu_1 = \nu_2 = 1$) of the forward lexicographical point Gauss–Seidel (GS-fLEX) method, and one iteration of GS-fLEX and one iteration of backward lexicographical point Gauss–Seidel (GS-bLEX). The other multigrid components are identical to the previously introduced multigrid method MG-RB. In the following, we call these multigrid methods MG-FF and MG-FB, which stand for multigrid forward forward and multigrid forward backward, respectively.

Local mode analysis gives sharp quantitative predictions for the multigrid convergence of MG-FF and MG-FB for the 2D anisotropic diffusion equation with $\varepsilon = 0.1$ and $\varepsilon = 0.01$, as can be seen in Table 5. In the anisotropic case, it is also beneficial to introduce an overrelaxation parameter for these lexicographical smoothers leading to ω-MG-FF and ω-MG-FB. The improvement is less impressive than for ω-MG-RB from [23], even if the optimal overrelaxation parameter ($\omega_{opt} = 1.40$ for $\varepsilon = 0.1, \omega_{opt} = 1.75$, or $\varepsilon = 0.01$) is selected (see Table 5).

In contrast to ω_{opt}-MG-FF (and ω_{opt}-MG-RB), it is possible to further accelerate ω_{opt}-MG-FB significantly by GMRES(m). This can be observed by comparing Table 5 and Table 6 and can be explained as follows. If A is a symmetric matrix, then MG-FB results in an *A-symmetric iteration matrix* due to the construction of the coarse grid correction (FW as the restriction and bilinear interpolation as the prolongation

operator) and the two relaxations which are adjoint to each other [8]. This means that the multigrid iteration matrix M is symmetric with respect to the A-induced inner product $(\,.\,,\,.\,)_A := (A\,.\,,\,.\,)$ (see, for example, [10]). The spectrum of the iteration matrix is real-valued, and uniformly distributed on the interval $[\lambda_{min}, \lambda_{max}]$. This feature is maintained if ω_{opt} is introduced, whereas for ω_{opt}-MG-FF an acceleration by GMRES(m) is not possible because the spectrum becomes orbital. This behavior of ω-MG-FF can be observed in Figure 4, where the spectra of the iteration matrices for $\omega = 1.0$ and $\omega = 1.75$ are shown for $\varepsilon = 0.01$. For other applications it is found in [10] that a symmetrization is not beneficial if Bi-CGstab [21] is used as the Krylov subspace acceleration method.



4 (a). $\omega = 1.0$ 4 (b). $\omega = 1.75$

FIG. 4. *Eigenvalue spectra of the MG-FF two-grid 2D anisotropic diffusion solver from local mode analysis;* $\varepsilon = 0.01, h = 1/128$. (*Note the different scaling of the axes in Figure* 4(a).)

Remark. For A-symmetric matrices it is possible to use the CG iteration as the acceleration method, in which the residual is minimized in the A-norm instead of the 2-norm (see (16)). The asymptotic convergence rate for the preconditioned CG iteration is given by the well-known formula [16]

$$\rho_{cg} := \Big(\min_{P\in\mathbb{P}_m} \max_{\lambda_j\in[\lambda_{min},\lambda_{max}]} |P(\lambda_j)| \Big)^{1/m} = \left(\max_{\lambda_j\in[\lambda_{min},\lambda_{max}]} \frac{T_m(1+2\frac{\lambda_{min}-\lambda_j}{\lambda_{max}-\lambda_{min}})}{T_m(1+2\frac{\lambda_{min}}{\lambda_{max}-\lambda_{min}})} \right)^{1/m}$$
$$= \left(2/T_m\left(\frac{\lambda_{max}+\lambda_{min}}{\lambda_{max}-\lambda_{min}}\right) \right)^{1/m} \le 2\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \qquad \text{with} \quad \kappa := \kappa_A(AC^{-1}) = \frac{\lambda_{max}}{\lambda_{min}}.$$

If we consider the definition T_m^E (29) for a real-valued spectrum $\sigma \subset [\lambda_{min}, \lambda_{max}]$, then the ellipse $E(c, d, a)$ deteriorates to a line. In particular we have $a = d = (\lambda_{max} - \lambda_{min})/2$ and $c = \lambda_{min} + a$, which leads to $T_m^E = (T_m(\frac{a}{a})/T_m(\frac{c}{a}))^{1/m} = \rho_{cg}$ (see (29), (30)). As the iteration matrix is A-symmetric, we find $\kappa_A(X) = 1$, where $\kappa_A(X)$ denotes the condition number of the transformation matrix X (26) with respect to the A-norm. In general, this does not imply that $\kappa_2(X) = 1$, but $\kappa_2(X)$ cancels out of (28) for large m. It can be expected that GMRES(m) has a similar asymptotic convergence as CG.

TABLE 7
Multigrid and multigrid preconditioned GMRES(m) convergence (different m) for the 3D Poisson equation and forward and backward lexicographical Gauss–Seidel smoothers.

Solution method	Grid 32^3	Grid 64^3	Grid 96^3	Fourier values
MG-FF	0.25	0.26	0.26	$\rho_F = 0.266$
Acc. MG-FF, $m = 2$	0.22	0.22	0.22	$\rho^U_{i=20} = 0.233$
Acc. MG-FF, $m = 5$	0.22	0.21	0.21	$\rho^U_{i=20} = 0.233$
MG-FB	0.29	0.29	0.29	$\rho_F = 0.294$
Acc. MG-FB, $m = 2$	0.098	0.10	0.10	$\rho^U_{i=20} = 0.121$
Acc. MG-FB, $m = 5$	0.086	0.085	0.086	$\rho^U_{i=20} = 0.099$

The analytical prediction $\rho^U_{i=20}$ shows a sharp quantitative character again, and the estimate T^E_m gives a reliable indication of the average accelerated convergence rate presented in Table 6. N^E_m is an interesting quantity for large m. For the 3D Poisson equation also, the multigrid convergence of MG-FF and MG-FB is similar. The convergence estimate ρ_F is 0.266 for MG-FF and 0.294 for MG-FB, whereas the spectra are completely different (as in 2D). It can be seen from Table 7 that the convergence rates with and without GMRES(m) are h-independent. Table 7 shows that very satisfactory convergence rates are obtained by the Krylov subspace acceleration of MG-FB. Moreover, due to the symmetry, we can predict the convergence rate of the accelerated MG-FB accurately from ρ_{cg} for different m.

Remark. With this knowledge, a symmetrization of the red-black multigrid Poisson solver is performed by doing the post-smoothing in black-red order. Indeed, the iteration matrix is A-symmetric, but the asymptotic convergence factor ρ_F increases to 0.440, which is also the value for MG-RB with only one smoothing step (see, for example, [7]). In this case the nonsymmetric multigrid iteration matrix leads to much better convergence factors with and without Krylov acceleration.

4.2. Results for a problem with mixed derivative. Next, we discuss an unsymmetric equation where reasonable smoothing properties can be achieved with point smoothers, but the coarse grid correction turns out to be problematic in combination with standard coarsening [19]. This equation is tackled in order to demonstrate that GMRES can also be applied to overcome coarse grid difficulties. We investigate the 2D differential equation with a mixed derivative $-\Delta u - \tau u_{xy} = b$ discretized by the $O(h^2)$ 9-point operator

$$A_h u_h(\boldsymbol{x}) = \frac{1}{h^2} \begin{bmatrix} \frac{\tau}{4} & -1 & -\frac{\tau}{4} \\ -1 & 4 & -1 \\ -\frac{\tau}{4} & -1 & \frac{\tau}{4} \end{bmatrix} u_h(\boldsymbol{x}) = b_h(\boldsymbol{x}) \quad \text{on} \quad \Omega_h.$$

It is no longer elliptic if $|\tau| = 2$ and the efficiency of all previously considered multigrid methods (MG-RB, MG-FF, MG-FB) deteriorates for $|\tau| \to 2$ (see, for example, Table 8 and [19]). This behavior can be explained if we perform a simplified two-grid analysis where only the very low Fourier frequencies along the characteristic direction of the differential operator (first differential approximation (FDA) [22]) are considered. The transfer operators act almost as identities for very low-frequency harmonics and no reduction of amplitudes can be achieved by the Gauss–Seidel relaxation. Thus, we obtain the following two-grid amplification factor (see (4)):

$$\rho^{FDA}(\boldsymbol{\theta}) := 1 - \frac{\lambda(\boldsymbol{\theta}, h)}{\lambda(2\boldsymbol{\theta}, 2h)} \quad \text{for } |\boldsymbol{\theta}| := \max\{|\theta_1|, |\theta_2|\} \longrightarrow 0. \tag{35}$$

TABLE 8
Multigrid and two-grid convergences ρ_h^{mg} and ρ_h^{tg}, predicted convergence ρ_F, and smoothing factor μ_F for the τ-problem, MG-FB, $h = 1/64$.

$\tau = -1.90$				$\tau = -1.99$			
ρ_h^{mg}	ρ_h^{tg}	ρ_F	μ_F	ρ_h^{mg}	ρ_h^{tg}	ρ_F	μ_F
0.747	0.654	0.662	0.499	0.839	0.730	0.745	0.518

TABLE 9
Multi- and two-grid preconditioned GMRES(m) convergence factors ρ_h^{acc} and $\rho_h^{acc}(tg)$ and estimates $\rho_{i=20}^U$, T_m^E for the τ-problem, $h = 1/64$.

Acc.	$\tau = -1.90$				$\tau = -1.99$			
MG-FB	ρ_h^{acc}	$\rho_h^{acc}(tg)$	$\rho_{i=20}^U$	T_m^E	ρ_h^{acc}	$\rho_h^{acc}(tg)$	$\rho_{i=20}^U$	T_m^E
$m = 2$	0.420	0.333	0.372	0.373	0.533	0.404	0.415	0.463
$m = 5$	0.370	0.287	0.302	0.304	0.472	0.348	0.354	0.378

The Fourier symbol $\lambda(\boldsymbol{\theta}, h)$ of the differential operator reads as follows:

$$\begin{aligned}\lambda(\boldsymbol{\theta}, h) &= \frac{1}{h^2}\Big(4 - 2\cos(\theta_1) - 2\cos(\theta_2) + \frac{\tau}{2}\cos(\theta_1 - \theta_2) - \frac{\tau}{2}\cos(\theta_1 + \theta_2)\Big) \\ &= \frac{1}{h^2}\Big(\theta_1^2 + \theta_2^2 + \frac{\tau}{4}(\,(\theta_1 + \theta_2)^2 - (\theta_1 - \theta_2)^2\,) - \frac{\theta_1^4 + \theta_2^4}{12} \\ &\qquad + \frac{\tau}{48}(\,(\theta_1 - \theta_2)^4 - (\theta_1 + \theta_2)^4\,) + \mathrm{O}(|\boldsymbol{\theta}|^6)\Big),\end{aligned}$$

where Taylor's expansion is used. If $\tau \longrightarrow \pm 2$ the characteristic frequencies are given by $\theta_2 = \mp\theta_1$. This gives

$$\begin{aligned}&\lambda(\boldsymbol{\theta}, h) = \frac{1}{h^2}\Big((2 \mp \tau)\theta_1^2 - \Big(\frac{1}{6} \mp \frac{\tau}{3}\Big)\theta_1^4 + \mathrm{O}(\theta_1^6)\Big) \\ \Longrightarrow\quad &\rho^{FDA}(\boldsymbol{\theta}) = 1 - \frac{4}{16} = 0.75 \quad \text{for } |\boldsymbol{\theta}| \longrightarrow 0 \text{ and } \tau \longrightarrow \pm 2.\end{aligned}$$

For $\tau = -1.99$ and MG-FB, this limiting value of 0.75 is almost obtained by the predicted two-grid factor ρ_F and also by the numerically obtained ρ_h^{tg}; see Table 8. The multigrid convergence ρ_h^{mg} increases further since the coarse grid problem occurs on all coarser grids (actually six grids are used).

One way to handle the coarse grid problem is to change the smoother to the more expensive ILU-type relaxation [19], which also takes care of the problematic low-frequency error components. A second way is to replace the grid coarsening by a nonstandard coarsening technique. We keep the point Gauss–Seidel smoother with standard grid coarsening and solve this problem by the Krylov subspace acceleration of MG-FB. In Table 9, a significant improvement of the multigrid convergence due to the acceleration is observed even with a small Krylov subspace. The two-grid convergence $\rho_h^{acc}(tg)$ is very well predicted by $\rho_{i=20}^U$ and T_m^E.

Comparing the accelerated two-grid and multigrid convergences from Table 9, it is expected that it is possible to further improve ρ_h^{acc} by incorporating the Krylov acceleration into the multigrid cycle and to apply it also on the coarse grids as in [14].

4.3. Further results and extensions. Here, we present a relation between Krylov subspace acceleration and a coarse grid correction acceleration as presented for the convection-diffusion equation in [4] and a relation between Krylov subspace acceleration and optimal (multistage) overrelaxation in a smoothing method.

A standard upwind discretization of a convection-dominated ($0 < \epsilon \ll 1$) convection-diffusion equation

$$-\epsilon\Delta u(\boldsymbol{x}) + (\, a_1(\boldsymbol{x})u(\boldsymbol{x})\,)_x + (\, a_2(\boldsymbol{x})u(\boldsymbol{x})\,)_y = b(\boldsymbol{x}), \tag{36}$$

with a rotating convection term ($a_1(\boldsymbol{x}) = -\sin(\pi x)\cos(\pi y)$, $a_2(\boldsymbol{x}) = \sin(\pi y)\cos(\pi x)$), is another well-known equation with a problematic coarse grid correction (see [4], [8]). A similar FDA analysis, as in the previous section, shows that characteristic low frequency error components, that are constant along the characteristics of the advection operator, are not reduced efficiently on coarse grids [4]. The different scaling of convection ($a_1(\boldsymbol{x})/h, a_2(\boldsymbol{x})/h$) and diffusion ϵ/h^2 is not approximated properly on the $2h$-grid which leads to a limiting two-grid convergence factor of 0.5 [4]. It is shown in [14] that the problems with rotating convection direction can be solved efficiently by a Krylov subspace acceleration of multigrid (also on the coarser grids of the multigrid preconditioner). In [4] an improved multigrid method is designed for this problem where the residuals in the kth two-grid cycle are overweighted by optimized acceleration parameters η_k. This means that the FDA error amplification factor after the mth cycle is given by the following polynomial (see (35) and [4]):

$$\rho_m^{FDA}(\boldsymbol{\theta}) = \prod_{k=1}^{m}(\, 1 - \eta_k\zeta(\boldsymbol{\theta})\,) \quad \text{with} \quad \zeta(\boldsymbol{\theta}) = \frac{\lambda(\boldsymbol{\theta}, h)}{\lambda(2\boldsymbol{\theta}, 2h)}.$$

The two methods are related. The FDA analysis can also be performed for the argument of the GMRES(m)-polynomials (21). It yields that the block matrix $UAC^{-1}U$ is dominated by $\zeta(\boldsymbol{\theta})$ for very low characteristic frequencies $\boldsymbol{\theta}$. We can construct an initial solution u_0 (necessary to calculate the GMRES(m)-polynomials explicitly; see subsection 3.2), which consists only of characteristic error components and freezes the discretization operator at a fixed (but arbitrary) point $\boldsymbol{x}$. All resulting GMRES(m)-polynomials P_m^i are almost identical for $i \geq 20$. Therefore, always the same coefficients α_k^i ($k = 1, \ldots, m$) are found. Thus, it is possible to identify ρ_m^{FDA} with P_m^i for $i \geq 20$ and we can express the coefficients α_k^i ($i \geq 20$) by the η_k, for example,

$$\alpha_1^i = -\eta_1 \quad \text{for} \quad m = 1\,, \quad \alpha_1^i = -(\eta_1 + \eta_2)\,,\ \alpha_2^i = \eta_1\eta_2 \quad \text{for} \quad m = 2. \tag{37}$$

The above considerations are confirmed by test calculations where such relations as from (37) are established for $i \geq 20$ and also for larger m. For problem (36), for example, we find

$$\alpha_1^i = -1.33 \quad \text{for} \quad m = 1\,, \quad \alpha_1^i = -2.82\,,\ \alpha_2^i = 1.88 \quad \text{for} \quad m = 2,$$

which matches exactly with the reference values η_k; see [4]. We see from this heuristic consideration that an optimal tuning of the coarse grid correction is implicitly done by the Krylov subspace acceleration.

Similarly it is possible to find optimal overrelaxation parameters for smoothing methods, like Jacobi or GS-RB, if the coarse grid correction acts as the optimal operator Q_h^{2h} from sections 2 and 4. For example, consider the multistage Jacobi relaxation for the 2D Poisson equation for which the smoothing operator is given by the polynomial

$$S_h = \prod_{k=1}^{m}(\, I_h - \omega_k\Delta_h/4\,). \tag{38}$$

The corresponding smoothing factor for the multistage relaxation S_h from (38) can be minimized analytically, yielding the optimal ω_k, which are knots of Chebychev polynomials [19]. If we select an initial approximation which consists only of high frequency error components, it is possible to identify S_h with the GMRES(m)-polynomials P_m^i (for $i \geq 20$) as the coarse grid correction leaves the high frequency components almost unchanged. A similar comparison of coefficients ω_k and α_k can be performed as described above for η_k and α_k (37). The related α_k^i, calculated by the analysis from subsection 3.2, match with the optimal ω_k for $i \geq 20$ very well.

For more complicated equations (or smoothing procedures) it is, in general, not possible to determine optimal overrelaxation parameters analytically, but the analysis is easily applicable and the α_k^i can be determined. As an example we consider the discrete biharmonic equation $\Delta_h \Delta_h u_h = b_h$ which is sometimes transformed into a system of two Poisson-type equations in order to achieve better smoothing properties [1]. Here we keep the scalar forth order problem and apply a multistage version of a GS-RB smoothing method

$$S_h^{msRB} = \prod_{k=1}^{m} (I_h - \omega_k (I_h - S_h^{BLACK} \cdot S_h^{RED})),$$

where the multistage parameters ω_k are obtained by an evaluation of the related α_k^i ($i \geq 20$). It can be seen in Table 10 that it is possible to obtain satisfactory smoothing rates also for the scalar biharmonic equation.

TABLE 10

Multistage red-black Gauss–Seidel relaxation parameters and smoothing rate for the biharmonic equation.

	α_1	α_2	α_3	μ_F	$(\mu_F)^{1/m}$
$m = 1$	-1.39			0.501	0.501
$m = 2$	-3.17	2.20		0.152	0.390
$m = 3$	-4.81	7.13	-3.30	0.050	0.369

Finally, it can be concluded that the Krylov subspace acceleration implicitly improves the coarse grid correction or the relaxation procedure, if one of these multigrid components clearly hampers the overall multigrid convergence. Furthermore, the analysis from subsection 3.2 is an easy tool to obtain good overrelaxation parameters for different (and also difficult) problems.

5. Conclusions. In this paper we have presented a way to obtain sharp quantitative convergence estimates for GMRES(m) preconditioned by multigrid on the basis of Fourier analysis. For all the cases considered the estimates are accurate compared to measured numerical convergence of the multigrid preconditioned GMRES(m) method.

It has been shown that it is not easily possible to further accelerate multigrid methods which are optimally tuned with, for example, overrelaxation parameters. In other situations, however, very satisfactory convergence improvement is achieved with the Krylov subspace acceleration. The possibilities for the subspace acceleration of multigrid are not only available in the case of isolated eigenvalues from the multigrid preconditioner, but they depend also on the shape of the spectrum. A spectrum resulting from the preconditioner with a circular shape is not appropriate for an acceleration with the method presented. For a fair comparison one should take the additional work for the Krylov subspace acceleration into account. This, however,

is not the main issue in this paper. The acceleration is, in general, cheap compared to the multigrid method. The proposed acceleration is, of course, also applicable to situations in which it is not easy to tune a multigrid method.

REFERENCES

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

[2] A. Brandt, *Rigorous quantitive analysis of multigrid,* I*: Constant coefficients two-level cycle with L_2-norm*, SIAM J. Numer. Anal., 31 (1994), pp. 1695–1730.

[3] A. Brandt and V. Mikulinsky, *On recombining iterants in multigrid algorithms and problems with small islands*, SIAM J. Sci. Comput., 16 (1995), pp. 20–28.

[4] A. Brandt and I. Yavneh, *Accelerated multigrid convergence and high-Reynolds recirculating flows*, SIAM J. Sci. Comput., 14 (1993), pp. 607–626.

[5] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.

[6] H. C. Elman, *Iterative Methods for Large Sparse Nonsymmetric Systems of Linear Equations*, Ph.D. thesis, Computer Science Department, Yale University, New Haven, CT, 1982.

[7] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations,* translated and revised from the 1991 German original, Appl. Math. Sci. 95, Springer-Verlag, New York, 1994.

[8] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer, Berlin, Germany, 1985.

[9] P. W. Hemker, *On the order of prolongations and restrictions in multigrid procedures*, J. Comput. Appl. Math., 32 (1990), pp. 423–429.

[10] M. Holst and S. Vandewalle, *Schwarz methods: To symmetrize or not to symmetrize*, SIAM J. Numer. Anal., 34 (1997), pp. 699–722.

[11] R. Kettler, *Analysis and comparison of relaxation schemes in robust multigrid and preconditioned conjugate gradient methods*, in Multigrid Methods, Lecture Notes in Math. 960, W. Hackbusch and U. Trottenberg, eds., Springer, Berlin, Germany, 1982, pp. 502–534.

[12] N. M. Nachtigal, S. C. Reddy, and L. N. Trefethen, *How fast are nonsymmetric matrix iterations?*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 778–795.

[13] C. W. Oosterlee and T. Washio, *An evaluation of parallel multigrid as a solver and a preconditioner for singularly perturbed problems*, SIAM J. Sci. Comput., 19 (1998), pp. 87–110.

[14] C. W. Oosterlee and T. Washio, *Krylov subspace acceleration of nonlinear multigrid with application to recirculating flows*, SIAM J. Sci. Comput., 21 (2000), pp. 1670–1690.

[15] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[16] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing, Boston, 1996.

[17] Y. Saad, *Further Analysis of Minimal Residual Iterations*, Tech. report, available online at ftp://ftp.cs.umn.edu/dept/users/saad/reports/FILES/umsi-97-14.ps.gz (1997).

[18] R. Stevenson, *On the Validity of Local Mode Analysis of Multi-Grid Methods*, Ph.D. thesis, Rijks University Utrecht, The Netherlands, 1990.

[19] K. Stüben and U. Trottenberg, *Multigrid methods: Fundamental algorithms, model problem analysis and applications*, in Multigrid Methods, Lecture Notes in Math. 960, W. Hackbusch and U. Trottenberg, eds., Springer, Berlin, Germany, 1982, pp. 1–176.

[20] C. A. Thole and U. Trottenberg, *Basic smoothing procedures for the multigrid treatment of elliptic* 3*-D operators*, Appl. Math. Comput., 19 (1986), pp. 333–345.

[21] H. A. van der Vorst, *BI-CGSTAB: A fast and smoothly converging variant of BI-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Comput., 13 (1992), pp. 631–644.

[22] N. N. Yanenko and Y. I. Shokin, *On the correctness of first differential approximation of difference schemes*, Dokl. Akad. Nauk. SSSR, 182 (1968), pp. 776–778.

[23] I. Yavneh, *On red-black SOR smoothing in multigrid*, SIAM J. Sci. Comput., 17 (1996), pp. 180–192.