COMPACT FOURIER ANALYSIS FOR MULTIGRID METHODS
BASED ON BLOCK SYMBOLS

THOMAS K. HUCKLE* AND CHRISTOS KRAVVARITIS†

Abstract. The notion of Compact Fourier Analysis (CFA) is discussed. The CFA allows
description of multigrid (MG) in a nutshell and offers a clear overview on the MG components. The
principal idea of CFA is to model the MG mechanisms by means of scalar generating functions and
matrix functions (block symbols). The formalism of the CFA approach is presented by describing
the symbols of the fine and coarse grid problems, the prolongation and restriction, the smoother, and
the coarse grid correction, resp. smoothing corrections. A first application is to utilize the CFA for
deriving MG as a direct solver, i.e. an MG cycle that will converge in just one iteration step. Necessary
and sufficient conditions that have to be fulfilled by the MG components are given for obtaining
MG functioning as a direct solver. Furthermore, new general and practical smoothers and transfer
operators that lead to efficient MG methods are introduced. Approximate inverse smoothers and
submatrix smoothers are derived and analyzed. In addition, we study sparse approximations of the
dense Galerkin coarse grid operator yielding efficient and practicable MG algorithms (approximately
direct solvers). Numerical experiments demonstrate the theoretical results.

Key words. multigrid, Fourier analysis, generating function, block symbol, Toeplitz matrices

AMS subject classifications. 65N55, 65F10, 65F15, 65N12, 15A12

1. Introduction. A crucial point for the efficiency of a multigrid (MG) method
is the appropriate choice of its components, which allows for an efficient interplay
between smoother and coarse grid correction. In many cases, this coordination can
be made by use of Local Fourier Analysis (LFA), which is an important quantitative
tool for the development of efficient MG methods [2, 24, 23, 13, 3]. This approach
has been generalized by employing generating functions expressing, e.g. the symbol of
the smoother or of the standard discrete Laplace operator in terms of a trigonometric
polynomial. It is based on the connection between Toeplitz, resp. circulant matrices,
and trigonometric functions. It was analyzed by S. Serra Capizzano, R. Chan, T.
Huckle, and coauthors in [9, 10, 5, 14, 16, 17, 15, 19, 20, 21]. In this paper we complete
this formal approach in order to represent a full two-grid (TG) step in terms of the
block symbol, called also block generating (matrix) function.

Here we consider the notion of Compact Fourier Analysis (CFA), which can be
seen as a reformulation of LFA. Instead of discrete operators on a grid we consider an-
alytic matrix functions (block symbols) that capture the behavior of the full matrices
[11, 16, 17]. This allows the use of matrix features, such as product and eigendecom-
position, to describe the smoothing factor and the convergence of a two-grid method.
CFA offers a more elegant and easier to handle description than the LFA and a clear
overview on the MG components. The principal idea of CFA is to fully model the MG
mechanisms by means of generating functions and block symbols. The block symbol
is simple to treat, since it is a matrix function of small order. The main properties of
a structured matrix can be reduced to a simple function and to its range [11, 15, 19].
We use CFA to
- derive necessary conditions for smoothing property and convergence expressed
  as characteristics of the underlying block symbol,
- determine all projections and smoothers that lead to MG as a direct solver,
- to design practicable MG components and improved MG algorithms, e.g. such that MG acts nearly as a direct solver.

The paper is organized as follows. In sections 1.1 and 1.2 we describe the formalism of CFA with a simple 1D example and explain the connection between discretization stencil of the original PDE and block symbol. Section 1.3 discusses how CFA can be used for determining smoothing factors. Section 2 illustrates conditions for the MG convergence in terms of the block symbols. In section 3 we derive necessary and sufficient conditions for MG as a direct solver. In section 4 we develop a strategy for deriving sparse approximate inverse smoothers with CFA and scalar generating functions. Section 5 is devoted to sparse approximations of the dense Galerkin coarse grid operator, which lead with special combinations of smoother and projection to MG as a nearly (approximately) direct solver. Section 6 is concerned with numerical experiments demonstrating the theoretical results. Section 7 summarizes the conclusions of this study and points out related future research possibilities. The appendices at the end aim at offering more insight and understanding on the subblock smoother and full projection.

Notations and preliminary remarks. Throughout this paper $B_R$ and $B_P$ denote the block symbols of the restriction and prolongation, respectively. The first columns of the corresponding block symbols will be denoted by $b_{1,R}$ and $b_{1,P}$. The term projection is used for describing both transfer operators if distinction between restriction and prolongation is not necessary or when they are applied symmetrically for defining the coarse grid matrix according to the Galerkin principle. The corresponding scalar generating function and block symbol are denoted by $b$ and $B$ (first column $b_1$). $M_R$ and $M_L$ denote the presmoother and postsmoother, respectively. When a distinction between them is unnecessary or when we refer generally to a smoother, we denote it by $M$. $S_R = I - M_R^{-1} A$ and $S_L = I - M_L^{-1} A$ are the block symbols for the post- and (pre-)smoothing defined by $M_R$ and $M_L$, respectively. The smoother given by $M$ defines the smoothing iteration $x_{k+1} = x_k + M^{-1}(b - Ax_k)$ for the iterative solution of a linear system $Ax = b$. Similar is the use of the denotation $N$ for approximate inverse smoother. The smoothers may be given in terms of an approximation $M$ of a matrix or of an approximation $N$ of its inverse.

$F^+$ stands for the adjoint matrix of $F$, so $F^{-1} = (\det F)^{-1} \cdot F^+$. If the entries of $F$ are trigonometric polynomials, the same is true for $F^+$. $F^+$ is then also related to a sparse matrix. Throughout the paper $\alpha = 1 + \exp(ix)$ and $\beta = 1 + \exp(iy)$. Square brackets $[ \ ]$ will be used for stencils and parentheses $()$ for matrices. Whenever distinction between block symbol and full matrix is necessary we add on the dependency on the variable(s) for the block symbols and the size of the full matrices when dealing with full matrices. So, for example, $TGS(x, y)$ denotes the block symbol of the matrix describing the full error reduction by a two-grid step (TGS), while $TGS_{n^2}$ denotes the corresponding full matrix of order $n^2$. Otherwise, when no confusion is possible we simply write the name of the matrix, e.g. $TGS$. $CGC$ is the coarse grid correction.

Definition 1.1. A two-grid method is considered to be a direct (exact) solver if the total error is removed after one iteration.

In other words, it must hold $TGS = S_L \cdot CGC \cdot S_R = 0$. Evidently it is sufficient that $S_L \cdot CGC = 0$ or $CGC \cdot S_R = 0$. In this case, the smoother and the projection interact in such a manner that the range of the one matrix is in the nullspace of the other and the actually iterative MG solver degenerates to a non-iterative, direct method. If the coarse system is the same as the original problem up to a constant factor, then the same projection and smoother again make the two-grid method a
direct solver on each level, hence the whole MG method can be considered as a direct solver. More informations about MG as a direct solver can be found in [23, 14, 3].

**Theorem 1.2.** [7, p. 72] A circulant matrix $C_n$ has the decomposition $C_n = F_n^H \Lambda_n F_n$, where $\Lambda_n$ is the diagonal matrix containing the eigenvalues of $C_n$ and $F_n$ is the Fourier matrix, which is unitary.

**Definition 1.3.** Let $\omega = \exp(i\theta)$ with $\theta \in [-\pi, \pi]$. An $n \times n$ matrix $W_n$ is called $\omega$-circulant if it has the spectral decomposition $W_n = \Omega_n F_n^H \Lambda_n F_n \Omega_n^H = \Omega_n C_n \Omega_n^H$. $F_n$ is the Fourier matrix, $\Lambda_n$ is diagonal containing the eigenvalues of $W_n$, $\Omega_n = \text{diag}(1, \omega^{1/n}, ..., \omega^{(n-1)/n})$ and $C_n$ denotes the circulant matrix from Theorem 1.2.

The following theorem, which computes, given the inverse of a matrix $A$, the inverse of a rank-one perturbation of this matrix in a convenient (i.e. numerically cheap) way, will be used repeatedly in section 3.

**Theorem 1.4** (The Sherman-Morrison formula [6, p. 137]). Let $A$ be a nonsingular matrix and $u, v$ two $n$-vectors with $v^T A^{-1} u \neq 1$. Then

$$
(A - uv^T)^{-1} = A^{-1} + \frac{A^{-1}uv^T A^{-1}}{1 - v^T A^{-1} u}.
$$

**Model problem.** The model problem for the introduction of the block symbol formalism and for the numerical results is the two dimensional Poisson equation

$$
u_{xx} + \nu_{yy} = g(x, y) \quad \text{(1.1)}
$$

with Dirichlet boundary conditions. Equation (1.1) is replaced by second order finite difference approximations leading to the linear system $Ax = b$, where the matrix $A$ is described in compact form by the 5-point stencil

$$A_5 = \begin{bmatrix}
-1 & -1 & 4 & -1 & -1 \\
1 & 0 & 0 & 0 & 1 \\
-1 & 4 & -1 & -1 & -1
\end{bmatrix}.
$$

**1.1. Multigrid and the block symbol formalism: A first example.** We want to describe the generating function and the block symbol formalism for the simple 1D Laplacian. The boundary value problem $\nu_{xx} = -g(x)$, $x \in [a, b]$, $u(a) = u(b) = 0$ in standard discretization leads to the stencil $[-1 2 -1]$ and the matrix $A = \text{tridiag}(-1, 2, -1)$. The full weighting restriction operator is related to the stencil $[1 2 1]$. The scalar generating functions (called also scalar symbols) to these two stencils are given by $f(x) = -\exp(ix) + 2 - \exp(-ix) = 2(1 - \cos(x))$ and $b(x) = 2(1 + \cos(x))$, respectively [14]. The corresponding block symbols can be derived by considering the Toeplitz tridiagonal matrices as block Toeplitz matrices with $2 \times 2$ blocks and related block stencil, e.g. $A = \text{tridiag}(-1, 2, -1)$ can be described in block stencil notation as

$$
\begin{bmatrix}
0 & -1 & 0 & 0 \\
0 & 0 & -1 & 2 \\
0 & 0 & -1 & 0 \\
-1 & 2 & -1 & 0
\end{bmatrix},
$$

which leads to the generating block function

$$
F(x) = e^{ix} \begin{bmatrix}
0 & -1 \\
0 & 0
\end{bmatrix} + \begin{bmatrix}
2 & -1 \\
-1 & 2
\end{bmatrix} e^{-ix} \begin{bmatrix}
0 & 0 \\
0 & 0
\end{bmatrix} = \begin{bmatrix}
2 & -\alpha \\
-\alpha & 2
\end{bmatrix}.
$$
The block symbol related to \( b(x) \) is

\[
B(x) = \begin{pmatrix} 2 & \alpha \\ \bar{\alpha} & 2 \end{pmatrix}.
\]

From the block symbol \( B \) only the first column \( b_1 \) is used to define the projection from fine to coarse grid.

The connection between block symbol and full matrix can be studied with the example \( A = \text{tridiag}(-1,2,-1) \). Let \( P \) be the odd/even permutation, which brings the even indices forwards. Then \( B = PAP^T \) has the block structure

\[
B = \begin{pmatrix} 2I & C_1 \\ D_1 & 2I \end{pmatrix},
\]

where \( C_1 = \text{tridiag}(-1,-1,0) \) and \( D_1 = \text{tridiag}(0,-1,-1) \). Application of the block sinus transformation \( S = \text{diag}(S_{n/2},S_{n/2}) \) on \( B \) from both sides diagonalizes \( \text{tridiag}(-1,-1,0) \) and we have

\[
SBS = \begin{pmatrix} 2I & C_2 \\ D_2 & 2I \end{pmatrix} =: A_1,
\]

where \( C_2 = \text{diag}(-\exp(ix_j) - 1) \) and \( D_2 = \text{diag}(-1 - \exp(-ix_j)) \) for equidistant points \( x_j \). One more application of the odd/even permutation yields

\[
P^T A_1 P = \text{blockdiag} \begin{pmatrix} 2 \\ -1 - \exp(-ix_j) \\ -1 - \exp(ix_j) \end{pmatrix} = \text{blockdiag}(F(x_j)).
\]

The last expression indicates that the block symbol is

\[
F(x) = \begin{pmatrix} 2 \\ -1 - \exp(-ix) \\ -1 - \exp(ix) \end{pmatrix}.
\]

Similarly can be handled the case of circulant matrices by making use of the Fourier transformation.

Also the smoother can be described by a block symbol. For the Jacobi smoother the block symbol is given by \( \text{diag}(2,2) \) and the Gauss-Seidel smoother is related to the block symbol

\[
\begin{pmatrix} 2 & -e^{ix} \\ -1 & 2 \end{pmatrix} = e^{ix} \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix},
\]

which corresponds to the lower triangular matrix \( \text{tridiag}(-1,2,0) \). Furthermore, since coarsening is equivalent to picking only the first column/row in the block symbol, the scalar symbol for the coarse grid problem according to the Galerkin approach is derived as

\[
f_c(x) = \begin{pmatrix} 1 & 0 \end{pmatrix} B(x) F(x) B(x) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = b_1(x)^H F(x) b_1(x) = 4(1 - \cos(x)) = 2f(x).
\]

The error after one TG step can be analyzed by the error reduction induced by the smoothing operator \( S = I - M^{-1} A \) and by the coarse grid correction \( \text{CGC} = I - RA^{-1} R^T \cdot A \) (\( R \) is the restriction) in the form \( S_L^{n+1} \cdot \text{CGC} \cdot S_R^n \) with postsmoother
\( S_L \) and presmoother \( S_R \) applied \( \nu_l \), resp. \( \nu_r \) times. Then the error transformation by the coarse grid correction is described by
\[
CGC = I - b_{1,R} b_{1,R}^H F / f_c,
\]
where the scalar symbol for the coarse problem is given according to the Galerkin principle as \( f_c = b_{1,R}^H F b_{1,R} \). The restriction and prolongation can be applied also symmetrically, leading to
\[
CGC = I - b_{1} b_{1}^H F / f_c.
\]

The error reduction of a two-grid cycle with coarse grid correction and one pre- and postsmoothing step is described by
\[
S_L \cdot CGC \cdot S_R.
\]
It is easy to see that \( b_{1,R} \) belongs to the nullspace of \( CGC \), while \( b_{1,R}^H F \) belongs to the left nullspace of \( CGC \).

\[
CGC \cdot b_{1,R} = (I - b_{1,R} b_{1,R}^H F / f_c) \cdot b_{1,R} = b_{1,R} - b_{1,R} b_{1,R}^H F / b_{1,R}^H F b_{1,R} = b_{1,R} - b_{1,R} = 0,
\]
\[
b_{1,R}^H F \cdot CGC = b_{1,R}^H F \cdot (I - b_{1,R} b_{1,R}^H F / f_c) = b_{1,R}^H F - b_{1,R}^H F / b_{1,R}^H F b_{1,R} = 0.
\]

1.2. Block symbols and stencils. First we investigate the relationship between PDE, stencil, matrix and block symbol for the general 2D case. The stencil determines the series coefficients of the multilevel Toeplitz, resp. circulant matrix, related to the discretization of the given constant coefficient PDE. Hence, related to the two dimensional stencil
\[
\begin{bmatrix}
  
  \vdots & \vdots & \vdots & 
  \vdots & a_{1,1} & a_{0,1} & a_{-1,1} \\
  \vdots & a_{1,0} & a_{0,0} & a_{-1,0} \\
  \vdots & a_{1,-1} & a_{0,-1} & a_{-1,-1} \\
  \vdots & \vdots & \vdots & 
  
\end{bmatrix}
\] (1.6)
is the scalar 2D symbol
\[
f(x, y) = \cdots a_{1,1} e^{ix+iy} + a_{0,1} e^{iy} + a_{-1,1} e^{-ix+iy} + \cdots \\
\cdots a_{1,0} e^{ix} + a_{0,0} + a_{-1,0} e^{-ix} + \cdots \\
\cdots a_{1,-1} e^{ix-iy} + a_{0,-1} e^{-iy} + a_{-1,-1} e^{-ix-iy} + \cdots .
\] (1.7)

For the block symbol we distinguish between odd and even numbered grid points in each dimension. There are different ways to describe the connection between stencil, matrix, and block symbol. First, like in 1D (cf. [14]), we can overlay a twolevel block structure over the discretization matrix of the 2D problem, which is \( \hat{A} = A \otimes I + I \otimes A \),
with $A = \text{tridiag}(-1, 2 - 1)$. The $4 \times 4$ blocks are built out of $2 \times 2$ blocks in each dimension. Based on these blocks $F_{j,k}$ we describe the block symbol for the twoflevel Block Toeplitz matrix by a series in the form

$$F(x, y) = \sum_{j,k} F_{j,k} \exp(ijx + iky).$$

A second approach can be applied only if the given matrix is a multilevel Toeplitz matrix with known scalar symbol and we want to redefine the matrix as a multilevel Block Toeplitz matrix with block symbol. From the scalar generating function $f$ we extract odd/even scalar coefficients directly to derive the scalar generating functions $F_{jk}(x, y)$, which are the entries in $F$. These methods are described in [14]. For 2D with the 5-point stencil $A_5$ we get

$$F(x, y) = \begin{pmatrix} 4 & -\alpha & -\beta & 0 \\ -\bar{\alpha} & 4 & 0 & -\beta \\ -\beta & 0 & 4 & -\alpha \\ 0 & -\beta & -\bar{\alpha} & 4 \end{pmatrix}.$$

Additionally, we want to describe the direct connection between block symbol and stencil. In 1D, the stencil is considered as blocks of pairs in the form

$$[ \cdots | 0 \, -1 \, | \, 2 \, -1 \, | \, 0 \, 0 \, | \, \cdots ].$$

This stencil is related to the block symbol

$$e^{ix} \begin{pmatrix} 0 & -1 \end{pmatrix} + e^{-ix} \begin{pmatrix} 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -\alpha \end{pmatrix}.$$

In the same way we can write the shifted stencil (applied on the neighboring grid point)

$$[ \cdots | 0 \, 0 \, | \, -1 \, 2 \, | \, -1 \, 0 \, | \, \cdots ].$$

with block symbol

$$e^{ix} \begin{pmatrix} 0 & 0 \end{pmatrix} + e^{-ix} \begin{pmatrix} -1 & 2 \end{pmatrix} = \begin{pmatrix} -\bar{\alpha} & 2 \end{pmatrix}.$$

Combining these two block symbols in a $2 \times 2$ matrix yields the block symbol for $A = \text{tridiag}(-1, 2, -1)$. In general the rows of the block symbol $F$ define two - possibly different - stencils, that are related to two typical neighboring rows of the matrix $A$ and to the $2 \times 2$ block structure of $A$.

The same approach can be applied in 2D. The stencil describes a typical row in the matrix. In 2D with $2 \times 2$ block size we could have 4 different stencils, resp. four typical rows in $A$ or 4 different rows in the block symbol. To derive the entries in the block symbol directly, we can use the stencil. But in contrast to the relation of stencil and scalar symbol, cf. (1.6), (1.7), here we choose a center point in the stencil (there are four independent choices); around this center point we delete every second entry in each dimension. The resulting stencil can be transformed into a scalar function like before, and this scalar function is one entry in the block symbol.

For the 2D five-point stencil we choose e.g. as center point $-1$ and define a scalar function by building a Fourier series collecting every second entry as displayed in Figure 1.1. This gives the scalar function $-1 - \exp(ix) = -\alpha$. In the same way we get all scalar functions of the block symbol $F$. 
1.3. Using block symbols for determining smoothing factors. In the scalar case the smoothing property is analyzed by considering the scalar symbol of the smoothing correction $s = 1 - f/m$ for the high frequency components (e.g. in 1D $\pi/2 \leq x \leq \pi$). To discuss the smoothing property for the block symbol we consider $S = I - M^{-1}F$ on the subspace related to the nonsingular eigenvectors [14]. Let $u_j, j = 0, ..., k - 1$ be the eigenvectors of the $k \times k$ matrix function $F$, related to eigenvalues $\lambda_j, j = 0, ..., k - 1$. $\lambda_0(0) = 0$ represents the singularity of $F$ while $\lambda_j > 0$ and bounded away from 0 for $j = 1, ..., k - 1$. The order $k$ of the block symbol $F$ is defined in dependence on the dimension and the coarsening strategy as $k = k_d \cdot k_t$, where $k_d$ is the dimension of the problem and $k_t$ is the ratio of coarse to fine grid size.

Note that $\lambda_0$, the scalar symbol $f$ and $\det F$ have the same zero of the same order at the origin. Then, the smoothing property can be determined by considering the projection of $S$ on the subspace spanned by $u_1, ..., u_{k-1}$, namely the spectral radius of $U^H S U_1$ for $x \in [0, 2\pi]$, where $U_1 = (u_1 \ldots u_{k-1})$. Examples: $b_R, b_P$ in 1D as $(1, 0), (1, \exp(ix)), (2, \alpha)$, or in 2D as trivial, standard, or with mixed zeros in the mirror points $(0, \pi), (\pi, 0), \text{and} (\pi, \pi)$.

In the general case we can analyse the $k \times k$ block symbol $F$ more precisely. As in 1D, cf. (1.2), and as seen in the recursive derivation of the block symbol described in [14], $F$ is up to a unitary diagonal similarity transform - a multilevel circulant matrix (number of levels depending on the dimension, block size depending on the projection: every second, third, ... entry in each dimension). Hence, we can diagonalize $F$ by a unitary diagonal matrix and a multilevel Fourier transform.

In 1D the block symbol $F$ can be transformed via $D_\alpha := \text{diag}(1, \alpha/|\alpha|)$ into the real circulant matrix

$$
\begin{pmatrix}
2 & -|\alpha| \\
-|\alpha| & 2
\end{pmatrix} := C.
$$

So, $F$ is an 1-level $\omega$-circulant matrix with $\omega = \alpha/|\alpha|$. Hence, $C$ can be diagonalized via the Fourier matrix $F_2$ of order 2. Then

$$
F = D_\alpha F_2^H \Lambda F_2 D_\alpha^H,
$$

where $\Lambda$ is a diagonal matrix containing the eigenvalues $\lambda_0 = f(x/2)$ and $\lambda_1 = f(x/2 + \pi)$ of $F$ and $C$.

In 2D we can apply the diagonal transformation

$$
D = D_\alpha \otimes D_\beta = \text{diag}(1, \alpha/|\alpha|, \beta/|\beta|, \alpha \beta/|\alpha \beta|)
$$
Table 2.1

<table>
<thead>
<tr>
<th>restriction</th>
<th>projection</th>
<th>norm</th>
<th>spectral radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>constant</td>
<td>0.6979</td>
<td>0.4934</td>
</tr>
<tr>
<td>standard</td>
<td>trivial</td>
<td>0.6993</td>
<td>0.4943</td>
</tr>
<tr>
<td>trivial</td>
<td>standard</td>
<td>0.0063</td>
<td>0.0043</td>
</tr>
</tbody>
</table>

for making \( F(x, y) \) real. Indeed,

\[
DFD^H = \begin{pmatrix}
4 & -|\alpha| & -|\beta| & 0 \\
-|\alpha| & 4 & 0 & -|\beta| \\
-|\beta| & 0 & 4 & -|\alpha| \\
0 & -|\beta| & -|\alpha| & 4
\end{pmatrix}.
\]

This yields a block circulant matrix with circulant blocks (BCCB), or a 2-level circulant matrix \( C \), with eigenvectors given by \( U = F_2 \otimes F_2 \). The eigenvalues of \( F \) and \( C \) are then \( \lambda_{00} = f(x/2, y/2) \), and generally \( \lambda_{kj} = f(x/2 + j\pi, y/2 + k\pi) \), \( j, k = 0, 1 \). Because \( F \) is singular at the origin, at least one of the eigenvalues, say \( \lambda_0 \), has to have the same singularity.

The diagonal matrix containing the eigenvalues is in the 2h harmonic LFA the symbol of the coarse grid system [23, 24]. Therefore, the block symbol of the CFA, which is described by generating functions, can be considered actually as an intermediate step in LFA. For instance, in order to determine the smoothing factor in the block symbol representation, the corresponding block symbols must be diagonalized employing \( D \) and \( U \). The spectral radius of the matrix function projected on the eigenspace corresponding to the eigenvalues different from \( \lambda_{00} \) gives the smoothing factor

\[
\max_{x \in \text{span}\{u_{10}, u_{01}, u_{11}\}} x^H (I - M^{-1} F) x.
\]

2. Necessary Conditions for the projections for feasible MG in terms of the block symbol. In the literature related to algebraic MG [18] and MG for Toeplitz matrices there are described necessary conditions that have to be satisfied by the restriction and prolongation operators, by the smoother, and by the coarse grid problem in order to derive convergence. The following analysis uses the eigen-decomposition of the \( k \times k \) matrix function \( F \) and the different character of eigenvalues: eigenvalue \( \lambda_0 \equiv f(x/2) \) in 1D) with the same singularity at the origin as \( f \) and \( \det F \). CGC has to reduce the error relative to the singularity, and therefore \( \|\text{CGC} \cdot u_0\| < \|u_0\| \) must hold, or \((F u_0)/f_c = \lambda_0 u_0/f_c\) has to be bounded and different from zero. Hence, \( f_c, \lambda_0, f \), and \( \det F \) have to have the same zeros (note that \( \lambda_0, f \) and \( \det F \) are equal up to a positive trigonometric function).

The CGC \( k \times k \) block symbol has eigenvalue 1 of order \( k - 1 \) and a single eigenvalue \( 1 - b^H F b_1, F b_1, /f_c \). \( F \) has one eigenvalue \( \lambda_0 \equiv f(x/2) \) in 1D) with the same singularity at the origin as \( f \) and \( \det F \). CGC has to reduce the error relative to the singularity, and therefore \( \|\text{CGC} \cdot u_0\| < \|u_0\| \) must hold, or \((F u_0)/f_c = \lambda_0 u_0/f_c\) has to be bounded and different from zero. Hence, \( f_c, \lambda_0, f \), and \( \det F \) have to have the same zeros (note that \( \lambda_0, f \) and \( \det F \) are equal up to a positive trigonometric function).

In the case of Galerkin coarse grid approach it must hold

\[
f_c = b^H_1 F b_1 / f_c = f \cdot t
\]

for a trigonometric function \( t > 0 \). Note, that this criterion follows from the condition that the scalar symbol \( b_p b_R \) must have zeros at the so called mirror points at least of
the order of the zero $0$ of $f$, and $b_pb_R$ is not zero at the origin. More information about the mirror points and their importance for appropriate choices of transfer operators for MG can be found in [10, 8].

Furthermore, it is advisable to derive a CGC block symbol that is well defined and bounded without poles: in the matrix $b_{1,P}b_{1,R}^H F/f_c$ all coefficients have to be bounded. It is sufficient to assure that $b_{1,R}^H F/f_c$ is bounded. Let $b_{1,R}^H F/f_c =: c^H$, where the vector $\hat{c}$ has trigonometric polynomials as entries that are nonzero at the origin. Taking into account that $f_c$ equals $\det F$ up to constant factor yields the condition

$$b_{1,R}^H = c^H F^+$.$$  

Therefore, the block symbol for CGC is bounded for all chosen $b_{1,P}$. Similarly, for

$$b_{1,P} = F^+ c$$

we get that $F \cdot CGC \cdot F^{-1}$ is bounded for each $b_{1,R}$. In Table 2.1 we see that the choices of restriction and prolongation given in the first two rows do not yield satisfactory convergence properties because CGC is not bounded. On the contrary, in the third row it is clear that a choice of transfer operators leading to bounded CGC yields more effective two-grid corrections.

Neglecting the influence of the smoothing error correction, an optimal pair of $b_P$ and $b_R$ should lead to rank $k - 1$ CGC symbol and minimize $\|CGC\|$. So, we could consider the following problem: Given $f$ and $b_R$ satisfying (2.1) and (2.2) (e.g. $b_R$ with zeros of the right order at the mirror points); find optimal $b_P$ that minimizes $\|CGC\|$.

The requests (2.1), (2.2) and (2.3) are satisfied if $b_R$ and $b_P$ have zeros of the same order as $f$ at $x = 0$ at the mirror points and are nonzero at the origin. One should avoid a splitting of the zeros between $b_R$ and $b_P$ because then only property (2.1) is fulfilled. The scalar symbols $b_P$ and $b_R$ are allowed to have additional zeros at the origin. But then $b_P b_R$ has to match the order of the zero of $f$ at the origin at least in one of the mirror points. So, e.g. in 1D we can choose $b_R = f(2x)$ and $b_P$ as the standard prolongation. In [10] the necessary conditions for the prolongation $b$ are described as:

1. $b(0,0) \neq 0, b(0,\pi) = b(\pi,0) = b(\pi,\pi) = 0$, and
2. $b^2(x,y) + b^2(x,y+\pi) + b^2(x+\pi,y) + b^2(x+\pi,y+\pi) > 0$ for all $(x,y)$.

Considering the block symbol $B(x,y)$ related to $b(x,y)$ if the second condition is violated for a pair $(x_1,y_1)$, then it would hold $B(2x_1,2y_1)$ is the null matrix.

3. Deriving multigrid as a direct solver. We aim at deriving a full classification of all MG designs based on trigonometric polynomials that are related to MG as a direct solver. In the theoretical proofs of this section we assume that conditions (2.1), (2.2) and (2.3) are satisfied, so that a functioning MG algorithm is feasible. Additionally, for the sake of practicability, the restriction and prolongation matrices must be sparse, i.e. the vectors $b_{1,R}$ and $b_{1,P}$ must have trigonometric polynomials as entries.

3.1. The case of smoother $M$. In this subsection we analyze under what conditions it is possible to derive MG as a direct solver, when an explicit smoother $M$ is used. Due to the recursive nature of MG methods, this specific two-grid algorithm can be extended straightforwardly to MG as a direct solver.
Theorem 3.1. Let $b_{1,P}$ be a given prolongation. Multigrid is a direct solver, when a presmoother $M_R$ is used, if and only if
\[ M_R = F + (Fb_{1,P})d^H, \]  
(3.1)
where $d$ is an arbitrary vector.

Proof. Suppose that multigrid is functioning as a direct solver. Since $CGC$ has an one-dimensional nullspace, $I - M_R^{-1}F$ has to eliminate all the remaining components, hence it should be of rank 1. Therefore, we can suppose that $M_R - F = cd^H$ for arbitrary vectors $c$ and $d$ of appropriate dimensions. According to Theorem 1.4 we have
\[ M_R^{-1} = F^{-1} - \frac{F^{-1}cd^H F^{-1}}{\gamma}, \]
where $\gamma = 1 + d^H F^{-1} c$. Multiplication by $F$ from the right yields
\[ M_R^{-1} F = I - \frac{F^{-1}cd^H F^{-1}}{\gamma} \Leftrightarrow I - M_R^{-1} F = (F^{-1} c) \frac{d^H}{\gamma}. \]
Since we presume a direct solver it holds
\[ CGC \cdot (I - M_R^{-1} F) = 0 \Leftrightarrow CGC \cdot (F^{-1} c) \frac{d^H}{\gamma} = 0. \]
Taking into account that $CGC \cdot b_{1,P} = 0$ always applies, cf. (1.4), a straightforward choice for prolongation is $b_{1,P} = F^{-1} c$, which implies $c = Fb_{1,P}$ and the presmoother is defined as $M = F + (Fb_{1,P})d^H$.

Now suppose that $M_R = F + (Fb_{1,P})d^H$ holds. Like before, we see that
\[ M_R^{-1} = F^{-1} - \frac{F^{-1} F bd^H F^{-1}}{1 + d^H F^{-1} F b} = F^{-1} - \frac{bd^H F^{-1}}{1 + d^H b}. \]
It holds
\[ CGC \cdot (I - M_R^{-1} F) = (I - bb^H F/f_c)(I - (F^{-1} - \frac{bd^H F^{-1}}{1 + d^H b})F) = 0, \]
so the presmoother $M_R$ given by (3.1) leads to MG as a direct solver \( \square \)

Theorem 3.2. Let $b_{1,R}$ be a given restriction. Multigrid is a direct solver, when a postsmoother $M_L$ is used, if and only if
\[ M_L = F + cb_{1,R}^H F, \]  
(3.2)
where $c$ is an arbitrary vector.

Proof. Let multigrid be a direct solver, so we have $(I - M_L^{-1} F) CGC = 0$, or equivalently $\frac{1}{2}(F^{-1} c)d^H \cdot CGC = 0$, like in the proof of Theorem 3.1. Since $b_{1,R} F \cdot CGC = 0$ always holds, an obvious choice for $d$ is $b_{1,R} F = d^H$ or $d = Fb_{1,R}$, leading to $M_L = F + c \cdot b_{1,R}^H F$.

If $M_L = F + c \cdot b_{1,R}^H F$, then
\[ M_L^{-1} = F^{-1} - \frac{F^{-1} cb_{1,R}^H FF^{-1}}{1 + b_{1,R} F F^{-1} c} \Rightarrow M_L^{-1} = F^{-1} - \frac{F^{-1} cb_{1,R}^H}{1 + b_{1,R} c}. \]
and
\[ I - M^{-1}_L F = \frac{F^{-1} c b^H F}{1 + b^H c}. \]

Hence
\[ (I - M^{-1}_L F)CGC = \frac{F^{-1} c b^H F}{1 + b^H c} (I - b b^H F) = \frac{F^{-1} c b^H F}{1 + b^H c} (1 + b^H c) b^H F b = 0. \]

However, the expressions (3.1) and (3.2) for \( M \) have the drawback that \( M \) is equally ill-conditioned as \( F \) and \( M^{-1} \) is as complicated to compute as \( F^{-1} \). A possibility to overcome this peculiarity is to introduce the additional constant \( \det F \) and to incorporate it appropriately in the above expressions for \( M \).

**Corollary 3.3.** Let \( b_{1,R}^H = e^H F^+ \) and let \( c = e \cdot (\det F)^{-1} \) be the vector of Theorem 3.2, where \( e \) is a vector with trigonometric polynomial entries, not parallel to \( u_0 \). Then \( M_L = F + ee^H \) leads to MG as a direct solver, \( M_L > 0 \) and \( M_L \) is well conditioned.

**Proof.** From Theorem 3.2 we have that the postsmoother has to be of the form (3.2) for a given restriction \( b_{1,R} \) and an arbitrary vector \( c \), in order to derive MG as a direct solver. Hence
\[ M_L = F + c b_{1,R}^H F = F + \frac{e}{\det F} e^H F^+ F = F + \frac{e}{\det F} e^H (\det(F) \cdot F^{-1}) F = F + ee^H > 0. \]

The next corollary can be proved in a similar fashion.

**Corollary 3.4.** Let \( b_{1,P}^H = F^+ e \) and let \( d = e \cdot (\det F)^{-1} \) be the vector of Theorem 3.1, where \( e \) is a vector with trigonometric polynomial entries, not parallel to \( u_0 \). Then \( M_R = F + ee^H \) leads to MG as a direct solver, \( M_R > 0 \) and \( M_R \) is well conditioned.

**Proof.** \( M_R = F + (Fb_{1,P})d^H = F + F F^+ e \frac{e^H}{\det F} = F + ee^H > 0. \)

In the theoretical considerations and numerical experiments in the rest of the paper we use frequently the subblock smoother \( SF \) (cf. [14]), which for the model problem (1.1) with discretization stencil \( A_5 \), has block symbol
\[
SF(x,y) = \begin{pmatrix}
4 & 0 & 0 & 0 \\
-\alpha & 4 & 0 & -\beta \\
-\beta & 0 & 4 & -\alpha \\
0 & -\beta & -\alpha & 4
\end{pmatrix}.
\]

More informations on how linear systems in the SF smoother can be solved effectively are given in Appendix A. The subblock smoother combined appropriately with the full projection, which was introduced in [14] and is further discussed in Appendix B, leads to a practical direct solver. The smoothing process with \( SF \) is closely related to the F-smoothing discussed in [23, Paragraph A.5.1.2].

**Remark 3.5.** The results of Corollaries 3.3 and 3.4 hold also for non symmetric construction of the smoother, i.e. \( M = F + uv^H \). For example, the subblock smoother is constructed as \( SF = F + uv^H \), where \( u = ( -1 \ 0 \ 0 \ 0 )^T \) and \( v = ( 0 \ -\alpha \ -\beta \ 0 ) \).
3.2. The case of approximate inverse smoother \( N \). In this paragraph we demonstrate conditions for deriving multigrid as a direct solver, when an approximate inverse smoother \( N \) is used. We consider only the case of postsmoother application, since the case of presmoother can be handled absolutely similarly.

**Lemma 3.6.** Let \( b_{1,R} \) be a given restriction. Multigrid with approximate inverse postsmoother \( S_L = I - N_L F \) is a direct solver if and only if
\[
S_L = a \cdot b_{1,R}^H F,
\]for some vector \( a \) of appropriate dimension. In this case it holds
\[
F^+ = d \cdot N_L + \tilde{a} \cdot b_{1,R}^H,
\]where \( d = \det F \) and \( \tilde{a} = d \cdot a \).

**Proof.** Suppose that (3.3) holds. For a direct solver with approximate inverse postsmoother it is required \( S_L \cdot CGC = 0 \), where \( CGC \) is given by (1.3). Indeed, the choice \( S_L = ab_{1,R}^H F \) leads to \( S_L \cdot CGC = 0 \) for any arbitrary vector \( a \) of appropriate dimension:
\[
S_L \cdot CGC = ab_{1,R}^H F \left( I - \frac{b_{1,R} b_{1,R}^H}{b_{1,R}^H b_{1,R}} F \right) = ab_{1,R}^H F - \frac{ab_{1,R}^H F b_{1,R} b_{1,R}^H F}{b_{1,R}^H b_{1,R}} F = 0.
\]

If MG is a direct solver, then \( S_L \) must have rank one, as explained in the proof of Theorem 3.1. Hence we can suppose that \( S_L = a \cdot c^H \) for arbitrary vectors \( a \) and \( c \). Considering (1.5) we can select \( c = F b_{1,R} \) and the result follows straightforwardly.

For the second part of the enunciation we have \( S_L = I - N_L F = ab_{1,R}^H F \Leftrightarrow N_L = F^{-1} - ab_{1,R}^H F \Leftrightarrow F^+ - ab_{1,R}^H F = dN_L + \tilde{a} b_{1,R}^H \), where \( \tilde{a} = d \cdot a \).

**Theorem 3.7.** An approximate inverse postsmoother leading to multigrid as a practicable direct solver is given by
\[
\begin{pmatrix}
F_1^{-1} & 0_{3 \times 1} \\
0_{1 \times 3} & 0
\end{pmatrix},
\]where \( F_1 \) is the upper left 3 \( \times \) 3 block of the block symbol \( F \). In this case, the decomposition (3.4) for the adjoint matrix holds, where \( \tilde{a} = (-U_1^{-1}c, 1)^T \), \( U_1 \) is the upper left 3 \( \times \) 3 block of the upper triangular matrix \( U \) of the LU factorization of \( F \), \( c \) is the upper right 3 \( \times \) 1 column vector of \( U \), and \( b_{1,R}^H = F^+ e_4 \), with \( e_4 = (0, 0, 0, 1)^T \).

**Proof.** Let \( d := \det F \). We consider the LU factorization \( F = L \cdot U \), where \( L \) is lower triangular with ones on the diagonal and \( U \) is upper triangular. It holds
\[
FF^+ = LUF^+ = dI \Rightarrow UF^+ = \begin{pmatrix} U_1 & c \\ 0_{1 \times 3} & \gamma \end{pmatrix} F^+ = dL^{-1},
\]where \( \gamma \) is a scalar and \( U_1 \) is a 3 \( \times \) 3 upper triangular matrix. Multiplication of the last equation from the left by \( \begin{pmatrix} U_1^{-1} & 0_{3 \times 1} \\ 0_{1 \times 3} & 0 \end{pmatrix} \) yields
\[
\begin{pmatrix} I & U_1^{-1}c \\ 0_{1 \times 3} & 0 \end{pmatrix} F^+ = d \begin{pmatrix} U_1^{-1} & 0_{3 \times 1} \\ 0_{1 \times 3} & 0 \end{pmatrix} L^{-1}.
\]Taking into account that \( F^+ \) should be of the form (3.4), equation (3.6) can be written as
\[
\left\{ \begin{pmatrix} I & 0_{3 \times 1} \\ 0_{1 \times 3} & 1 \end{pmatrix} + \begin{pmatrix} 0_{3 \times 3} & U_1^{-1}c \\ 0_{1 \times 3} & -1 \end{pmatrix} \right\} F^+ = d \begin{pmatrix} U_1^{-1} & 0_{3 \times 1} \\ 0_{1 \times 3} & 0 \end{pmatrix} L^{-1}.
\]
\[ \nonumber \leftrightarrow F^+ + \begin{pmatrix} 0_{3\times 1} & U_1^{-1}c \\ 0_{1\times 3} & -1 \end{pmatrix} F^+ = d \begin{pmatrix} U_1^{-1} & 0_{3\times 1} \\ 0_{1\times 3} & 0 \end{pmatrix} L^{-1} \]

\[ \nonumber \leftrightarrow F^+ = d \begin{pmatrix} U_1^{-1} & 0_{3\times 1} \\ 0_{1\times 3} & 0 \end{pmatrix} L^{-1} + \begin{pmatrix} -U_1^{-1}c \\ 1 \end{pmatrix} (F^+_{41} F^+_{42} F^+_{43} F^+_{44}). \]

Hence

\[ F^+ = d \begin{pmatrix} U_1^{-1} & 0_{3\times 1} \\ 0_{1\times 3} & 0 \end{pmatrix} L^{-1} + \begin{pmatrix} -U_1^{-1}c \\ 1 \end{pmatrix} (F^+_{41} F^+_{42} F^+_{43} F^+_{44}). \]

From Lemma 3.6 it is easy to see that

\[ N_L = \begin{pmatrix} U_1^{-1} & 0_{3\times 1} \\ 0_{1\times 3} & 0 \end{pmatrix} L^{-1} = \begin{pmatrix} F_1^{-1} & 0_{3\times 1} \\ 0_{1\times 3} & 0 \end{pmatrix} \text{ and } b_{1,R} = F^+ e_4 \]

lead to a practicable direct solver. \qed

**Remark 3.8.** The approximate inverse smoother of Theorem 3.7 is similar to the subblock smoother, both leading to MG as a direct solver. The former is easier to solve but doesn’t exhibit so good smoothing properties as the latter.

The corresponding results for the presmoother can be proved analogously.

**Lemma 3.9.** Let \( b_{1,P} \) be a given prolongation. Multigrid with approximate inverse presmoother \( S_R = I - N_R F \) is a direct solver if and only if

\[ S_R = b_{1,P} \cdot a, \]

for some vector \( a \) of appropriate dimension. In this case it holds

\[ (I - b_{1,P} \cdot a) F^+ = d \cdot N_R, \quad (3.7) \]

where \( d = \det F \).

**Theorem 3.10.** An approximate inverse presmoother leading to multigrid as a practicable direct solver is given by

\[ \begin{pmatrix} F_1^{-1} & 0_{3\times 1} \\ 0_{1\times 3} & 0 \end{pmatrix}, \]

where \( F_1 \) is the upper left 3 \( \times \) 3 block of the block symbol \( F \). In this case, (3.7) holds for the adjoint matrix, where \( b_{1,P} = (-U_1^{-1}c, 1)^T \), \( U_1 \) is the upper left 3 \( \times \) 3 block of the upper triangular matrix \( U \) of the LU factorization of \( F \), \( c \) is the upper right 3 \( \times \) 1 block of \( U \), \( U \) is the upper triangular matrix of the LU factorization of \( F \).

**4. A technique for deriving sparse approximate inverse smoothers.** We are interested in deriving new sparse approximate inverse (SAI) smoothers with good smoothing properties. The advantages when formulating a stationary smoothing iteration with approximate inverses are that singularity is allowed in the smoother, i.e. improved robustness, that the algorithm is inherently more suitable for implementation in a parallel environment and the local adaptivity feature. The benefits of the use of SAI smoothers have been discussed, among others, in [4, 12, 22, 23].
We utilize the formalism of scalar generating functions. The task can be formulated as a double optimization problem as

$$\min_{a,b,c} \max_{x,y} |n(x,y) \cdot f(x,y) - 1|,$$

where $n(x,y) = a + b(\cos x + \cos y + \cos(x-y) + \cos(x+y))$ and $f(x,y)$ are the scalar generating functions of the approximate inverse smoother $N$ and of the discretization matrix, respectively. The full matrix $N$ corresponding to the scalar symbol $n(x,y)$ is a sparse matrix, since $n(x,y)$ is a trigonometric polynomial.

Let $\sigma(x,y) := n(x,y) \cdot f(x,y) - 1$. The maximum of $\sigma(x,y)$ is sought for $(x,y)$ belonging to the interior of the region defined by the solid line in Figure 4.1. This region is part of the subspace of high frequency error components which is important for smoothing. Because of symmetry, the other parts of this subspace are not examined. Actually, it is sufficient to restrict ourselves to the values $0 \leq x \leq \pi/2$ and $0 \leq y \leq \pi$, cf. Figure 4.1.

**Example.** We consider the discretization of model problem (1.1) leading to the 9-point stencil

$$\begin{bmatrix}
-1 & -1 & -1 \\
-1 & 8 & -1 \\
-1 & -1 & -1
\end{bmatrix}$$

with scalar symbol $f(x,y) = 8 - 2 \cos(x) - 2 \cos(y) - 2 \cos(x+y) - 2 \cos(x-y)$ [14]. For the sake of simplicity and without loss of generality we divide it by 8, so that the corresponding full matrix is normalized. Let $h(x,y) := \cos x + \cos y + \cos(x-y) + \cos(x+y)$. Then

$$\sigma(x,y) = (a + bh(x,y)) \left(1 - \frac{h(x,y)}{4}\right) - 1.$$ 

For $0 \leq x \leq \pi/2$ and $0 \leq y \leq \pi$ it holds $h \in [-2,1]$. For $h = -2$, $\sigma(x,y) = (a - 2b)\frac{3}{4} - 1$ and for $h = 1$, $\sigma(x,y) = (a + b)\frac{3}{4} - 1$. These two quantities must be equal, hence $a = 5b$.

Furthermore, let $\psi(h) := (a + bh)(1 - \frac{h}{4}) - 1 = a + bh - \frac{ah}{4} - \frac{bh^2}{4} - 1$. Since we are looking for the maximum of $\psi$ it must hold $\psi' = 0$. Hence, $b - \frac{a}{4} - \frac{bh}{2} = 0 \Rightarrow h = \ldots$
Table 4.1
Smoothing factors for various relaxations. The model problem is discretized with the 5-point stencil \(A_5\).

<table>
<thead>
<tr>
<th>relaxation</th>
<th>smoothing factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\omega)-JAC, (\omega = 0.8)</td>
<td>0.6</td>
</tr>
<tr>
<td>GS-LEX</td>
<td>0.5</td>
</tr>
<tr>
<td>GS-RB</td>
<td>0.25</td>
</tr>
<tr>
<td>SAI</td>
<td>0.2195</td>
</tr>
<tr>
<td>subblock</td>
<td>0.0732</td>
</tr>
</tbody>
</table>

\[2 - \frac{a}{\pi^2} = 2 - \frac{56}{\pi^2} = -\frac{1}{2}\]. For \(h = -\frac{1}{2}\), \(\sigma(x, y) = (a - \frac{1}{2})(1 + \frac{1}{2}) - 1 = \frac{9}{2}b^2 - 1 = \frac{81}{16}b - 1\).

The value of \(\sigma(x, y)\), e.g. for \(h = -2\), is \(\frac{9}{2}b - 1\), hence it must hold

\[\frac{9}{2}b - 1 = 1 - \frac{81}{16}b \Rightarrow b = \frac{32}{153}, a = \frac{160}{153}\].

The smoothing factor is calculated e.g. from the value for \(h = -2\) as \(\frac{9}{2} \cdot \frac{32}{153} - 1 = \frac{1}{17} \approx 0.0588\), leading to a very effective smoothing property.

In a similar manner SAI smoothers can be derived also for other problems and various discretizations. It is important to stress that the SAI must have the same sparsity pattern as the discretization matrix so that the proposed strategy can be applied. For instance, for the 5-point stencil with scalar generating function \(4 - 2 \cos x - 2 \cos y\) [14, 23] one can assume a SAI smoother of the form \(a + b(\cos x + \cos y)\). If \(h(x, y) := \cos x + \cos y\) then \(\sigma(x, y) := (a + bh(x, y))(4 - 2(\cos x + \cos y)) - 1\) can be defined like before. The calculations lead to smoothing factor \(\frac{9}{2} \frac{32}{153} \approx 0.2195\), which indicates very good smoothing behavior and is comparable to the Gauss-Seidel smoother with red-black ordering (GS-RB). Thus it becomes apparent that the choice of discretization for a given problem plays a crucial role for the effectiveness of a SAI smoother.

Table 4.1 shows the smoothing factors of the most frequently used smoothers and of the smoothers that are discussed in this work. The smoothing factor of the subblock smoother is calculated as described in Section 1.3. We observe that the subblock smoother exhibits a remarkable smoothing effect.

5. Sparse approximations of the Galerkin coarse grid operator. As already demonstrated, cf. e.g. [3, 13, 23], the MG algorithms with coarsening based on the Galerkin principle may lead to efficient solvers for various choices of MG components. However, this coarsening strategy has the disadvantage that it leads to coarse grid matrices that become thicker in every next grid, without preserving the sparsity of the initial problem. A possible remedy for this issue was proposed in [1], where a stencil collapsing technique was developed for deriving sparser stencils that are spectrally equivalent to the Galerkin coarse grid operator.

Hence it is interesting to specify sparse approximations of the thick Galerkin coarse grid operator \(f_c = b^H Fb\) leading to a reduction of the computational complexity. It is important that the matrices \(A\) and \(A_c\) are the same up to a constant factor (yet they have different dimension). This is the case when the quotient of \(f_c\) is a constant number, where \(f\) and \(f_c\) are the scalar generating functions of the original matrix \(A\) and of the Galerkin coarse grid matrix \(A_c\), respectively. But in order to be able to generalize the results derived for two-grid to a multigrid cycle it is necessary that this property is satisfied. Therefore it is interesting to study sparse approximations of \(f_c\) of the form \(\frac{g}{q}\), where \(g\) is the scalar generating function corresponding to
a sparse matrix \( G \), i.e. \( g \) is a trigonometric polynomial.

The benefit of this approach is that the coarse grid system is similar to the fine and this results in a practicable algorithm. For instance, for the discretization with the 5-point stencil on the fine grid it is known that all coarse grid matrices are given by the thicker 9-point stencil, and this is a drawback in view of practicability and effectiveness of the method.

In contrast, according to the proposed approach, \( A_c \) is easily calculated as \( A_c = G^{-1}A \). Then on the coarser grid the residual equation must be solved

\[
A_c e = r \Leftrightarrow G^{-1}A e = r \Leftrightarrow A e = G r.
\]

Hence, in this manner we need to solve on every grid a linear system with the original matrix. The right hand side is formed by a cheap matrix-vector multiplication.

The purpose is to calculate \( \min_y \| g - \frac{f}{f_c} \| \). As an example we consider \( g_1(x, y) = a_0 + a_1(\cos x + \cos y) \) and the discretization matrix \( A \) described in compact form by the 5-point stencil on the fine grid. Let \( h(x, y) := \cos x + \cos y \). Then

\[
\min_{a_0, a_1} \| g_1 - \frac{f}{f_c} \|^2 = \min_{a_0, a_1} \int_0^\pi \int_0^\pi \left( a_0 + a_1 h - \frac{f}{f_c} \right)^2 dxdy
\]

\[
= \min_{a_0, a_1} \int_0^\pi \int_0^\pi \left( a_0^2 + a_1^2 h^2 + \frac{f^2}{f_c^2} + 2a_0a_1 h - 2a_0 \frac{f}{f_c} - 2a_1 h \frac{f}{f_c} \right) dxdy
\]

\[
= \min_{a_0, a_1} (a_0^2 I_0 + a_1^2 I_1 + I_2 + 2a_0a_1 I_3 - 2a_0 I_4 - 2a_1 I_5) \equiv \min\phi(a_0, a_1),
\]

where

\[
I_0 = \int_0^\pi \int_0^\pi 1 dxdy, \quad I_1 = \int_0^\pi \int_0^\pi h^2 dxdy, \quad I_2 = \int_0^\pi \int_0^\pi \frac{f^2}{f_c^2} dxdy,
\]

\[
I_3 = \int_0^\pi \int_0^\pi h dxdy, \quad I_4 = \int_0^\pi \int_0^\pi \frac{f}{f_c} dxdy, \quad I_5 = \int_0^\pi \int_0^\pi h \frac{f}{f_c} dxdy.
\]

It must hold:

\[
\begin{align*}
\frac{\partial \phi}{\partial a_0} &= 0 : \quad 2a_0 I_0 + 2a_1 I_3 - 2I_4 = 0 \\
\frac{\partial \phi}{\partial a_1} &= 0 : \quad 2a_1 I_1 + 2a_0 I_3 - 2I_5 = 0 \quad \Leftrightarrow \quad \begin{pmatrix} I_0 & I_3 \\ I_3 & I_1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} I_4 \\ I_5 \end{pmatrix}.
\end{align*}
\]

The integrals are evaluated numerically with Maple and then the solution of the minimization problem gives the required \( a_0 \) and \( a_1 \). A similar approach can be applied as well for approximations of the Galerkin operator with trigonometric polynomials of higher order. We consider also

\[
g_2(x, y) = g_1(x, y) + a_2(\cos x - y) + \cos(x + y),
g_3(x, y) = g_2(x, y) + a_3(2\cos 2x + \cos 2y),
g_4(x, y) = g_3(x, y) + a_4(2\cos 2x - y) + \cos(x - 2y) + \cos(2x + y) + \cos(x + 2y)
\]

and the constant case \( g_0(x, y) = a_0 \). We use the following denotation: trigonometric polynomials of degrees 1,2,3,4,5 are called the trigonometric polynomials that have the forms \( g_0, g_1, g_2, g_3, g_4 \), respectively. The results of the above described strategy are given in Table 5.1 for full prolongation and various restrictions.
5.1. MG as approximately direct solver. The goal of this approach is furthermore to derive efficient MG methods which work as approximately direct solvers. As approximately direct solver we declare MG methods that lead to satisfactory enough errors, i.e. of order $10^{-6}$ or $10^{-8}$, after application of 2 or 3 V-cycles.

The appropriate computer program was developed to this end in a hybrid (i.e. symbolic and numerical) environment. For degrees $d = 1, \ldots, 5$ of the approximating trigonometric polynomial and for $k = 1, 2, 3$ applications of the smoothers there has been an exhaustive search over all possible MG components (pre- and (post-)smoother, restriction, prolongation). The corresponding methods have been tested with respect to the norm and spectral radius of $TGC(x,y)$. For pre- and (post-)smoothers there have been examined: subblock smoother, Red-Black Gauss-Seidel (GS-RB), GS with lexicographic ordering (GS-LEX) and $\omega$-Jacobi ($\omega$-JAC) with the optimal relaxation parameter $\omega = 0.8$. Possible restrictions were: full, standard (full weighting), trivial and constant.

Tables 5.2 and 5.3 show the best results that appeared with respect to norm and spectral radius of $TGC(x,y)$, respectively, and also for which restriction. The best results are always obtained with the full prolongation and with the use of the subblock smoother as (pre-) and postsmoother. We see that the approximating trigonometric polynomials of degrees $1, \ldots, 4$ yield approximately direct solvers, while the approximation of degree 5 gives significantly better results. The results can always be further improved by multiple applications of the smoothers.

It is also interesting to mention that approximations of the Galerkin coarse grid operator with higher order trigonometric polynomials don’t always improve the overall behavior of MG. For instance, in Tables 5.2 and 5.3 we observe that the increase of the degree from $d = 1$ to $d = 2$ doesn’t lead to an improvement of the results. Similarly, from $d = 3$ to $d = 4$. The explanation is that with the proposed approach $\min_g \| g - f_c \|$ is determined, which finds an optimal approximation of $f/f_c$, but does not necessarily guarantee that $\min \| S_t \cdot CGC \cdot S_r \|$ is attained.

Hence, we can conclude that it is advisable to use sparse approximations of the Galerkin operators of orders $d = 1, 3, 5$. For $d = 1$ we obtain satisfactory results while preserving at the same time the sparsity of the original problem. For $d = 5$ we have the best results, which lead to an approximately direct solver. However, in this case there appear sparsity properties similar to the 9-point stencil. An average situation between these two extreme cases is for $d = 3$.

6. Experimental Results. Table 6.1 presents the norms and spectral radii of $TGS(x,y)$ for various MG methods, when the computations are based on the block symbol. Such computations reflect also the behavior of the full matrices, which is sufficiently well captured by the block symbol [14]. The presmoother is always the same with the postsmoother and the same also for restriction and prolongation. The first line shows the results for the standard choice of optimal MG components, which
Table 5.2
Optimal norms of $TGC(x,y)$ for various restrictions, full prolongation and subblock (pre-) and postsmoother for approximation of the Galerkin coarse grid scalar symbol with trigonometric polynomials of degree $d$, $n=50$ grid points, $\nu$ is the number of (pre-) and postsmoother applications.

<table>
<thead>
<tr>
<th>$d$</th>
<th>restriction</th>
<th>$\nu=1$</th>
<th>$\nu=2$</th>
<th>$\nu=3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>trivial</td>
<td>0.0447</td>
<td>0.0446</td>
<td>0.0444</td>
</tr>
<tr>
<td>2</td>
<td>standard</td>
<td>0.06</td>
<td>0.0598</td>
<td>0.0596</td>
</tr>
<tr>
<td>3</td>
<td>constant</td>
<td>0.0180</td>
<td>0.0179</td>
<td>0.0178</td>
</tr>
<tr>
<td>4</td>
<td>standard</td>
<td>0.0219</td>
<td>0.0218</td>
<td>0.0217</td>
</tr>
<tr>
<td>5</td>
<td>trivial</td>
<td>0.0098</td>
<td>0.0062</td>
<td>0.0045</td>
</tr>
</tbody>
</table>

Table 5.3
Optimal spectral radii of $TGC(x,y)$ for various restrictions, full prolongation and subblock (pre-) and postsmoother for approximation of the Galerkin coarse grid scalar symbol with trigonometric polynomials of degree $d$, $n=50$ grid points, $\nu$ is the number of (pre-) and postsmoother applications.

<table>
<thead>
<tr>
<th>$d$</th>
<th>restriction</th>
<th>$\nu=1$</th>
<th>$\nu=2$</th>
<th>$\nu=3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>trivial</td>
<td>0.0316</td>
<td>0.0315</td>
<td>0.0314</td>
</tr>
<tr>
<td>2</td>
<td>standard</td>
<td>0.0425</td>
<td>0.0423</td>
<td>0.0421</td>
</tr>
<tr>
<td>3</td>
<td>trivial</td>
<td>0.0127</td>
<td>0.0127</td>
<td>0.0126</td>
</tr>
<tr>
<td>4</td>
<td>standard</td>
<td>0.0155</td>
<td>0.0154</td>
<td>0.0154</td>
</tr>
<tr>
<td>5</td>
<td>trivial</td>
<td>0.0066</td>
<td>0.0042</td>
<td>0.0031</td>
</tr>
</tbody>
</table>

is the use of the Red-Black Gauss-Seidel smoother and of the full weighting projection. In the second line we see that the use of the subblock smoother leads to slightly better results. And in the third line it is evident that the combined use of the subblock smoother together with the full projection leads to a direct solver.

In Table 6.2 we investigate the behavior of the MG algorithm for various components with respect to effectiveness and practicability. A comparison between well known combinations of MG components and the approach proposed in this work demonstrates the superiority of the latter over the classical strategies and also the theoretical results given in previous sections. It is clear that more applications of the smoother lead to slightly better results. However, there is a significant improvement of the method if the approximation of the Galerkin coarse grid operator is done by a trigonometric polynomial of higher order, e.g. 5, eventually combined with multiple applications of the smoother. The drawback of these higher order approximations is that they lead to coarse grid matrices, which are thicker than the original matrix, yet still sparse.

The numerical results in Tables 6.3 and 6.4 verify the above mentioned comments and the theoretical results and they establish a comparison between the new components analyzed in this work and the classical choices for MG ingredients, which are commonly used. In the first two rows of Table 6.4 we see that the appropriate combination of the subblock smoother and the full prolongation with the standard restriction gives a highly efficient solver. The results with the GS-RB smoother are of comparable quality.

7. Conclusions. A compact Fourier analysis was devised for studying MG methods and for designing new efficient MG algorithms. The CFA is principally based on the block symbol. The connection between LFA and CFA is established, however the latter is easier to handle, offers a better overview and is suitable for analyzing all MG components. The block symbol is an intermediate step between the full matrix and the $2h$-harmonics LFA matrix.
We present the CFA framework and give the necessary conditions in terms of the block symbol that have to be satisfied by the projections for deriving functioning MG methods. We characterize generally the various smoothers that lead to MG as a direct solver in terms of given transfer operators. The results on direct solver hold for every stencil, type of discretization and also for anisotropic and indefinite problems (e.g. Helmholtz equation). The paper develops templates, like the full projection and the subblock smoother, and shows how to incorporate them efficiently in MG.

By taking advantage of scalar generating functions that are trigonometric polynomials, i.e. corresponding to sparse full matrices, we derive a new approximate inverse smoother with good smoothing properties. A technique for approximating the dense Galerkin coarse grid operator with sparse matrices was developed, leading to practicable approximately direct solvers. The various numerical experiments confirm the theoretical results and the superiority of the use of the full projection and of the subblock smoother.

A future research direction is the analytical treatment of anisotropic and indefinite problems by making use of the CFA. The corresponding components have to be modeled appropriately by block symbols to this end. Furthermore, it is interesting to study a three-grid Fourier analysis by means of the block symbol and also the case of varying coefficients.

**Appendix A. The subblock smoother.**

**Lemma A.1.** The lower right $3 \times 3$ block $F_3$ of $SF$ is positive definite.

**Proof.** For the subblock $F_3$ holds $F_3 \geq 0$, since $F \geq 0$. The proof will be done by contradiction. Suppose the $F_3$ is singular. This is possible only for the root of the symbol $(x,y) = (0,0)$. In that case there exists an eigenvector $a$ for which $F_3a = 0$ holds. Then $a$ is part of the eigenvector $b$ of $F$ for which $Fb = 0$. This eigenvector $b$ corresponds to the eigenvalue $\lambda_0 = f\left(\frac{x}{2}, \frac{y}{2}\right)$, or that is to say, $b$ is the vector $(1,1,1)^T$, up to an $\omega$-circulant diagonal transformation. But $F_3(1,1,1)^T$ is not 0, and this is a contradiction. 

In order to efficiently use the subblock smoother, we have to solve a linear system...
Table 6.3

Norms and spectral radii of \( TGC(x, y) \) for various multigrid variants with approximation of Galerkin coarsening of degree \( d \), smoother= subblock, restriction and prolongation are full in all cases, \( n=40 \) points, \( \nu \) is the number of smoother applications.

<table>
<thead>
<tr>
<th>( d )</th>
<th>( \nu )</th>
<th>norm</th>
<th>spectral radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
<td>0.0443</td>
<td>0.0313</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0.0438</td>
<td>0.0310</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.0119</td>
<td>0.0080</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>0.0072</td>
<td>0.0051</td>
</tr>
</tbody>
</table>

Table 6.4

Norms and spectral radii of \( TGC(x, y) \) for various multigrid variants with approximate Galerkin coarsening of degree 5, presmoother = subblock in all cases, restriction=standard and prolongation=full, \( n=40 \) points, \( \nu \) is the number of smoother applications.

<table>
<thead>
<tr>
<th>postsmoother</th>
<th>( \nu )</th>
<th>norm</th>
<th>spectral radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>subblock</td>
<td>1</td>
<td>0.0097</td>
<td>0.0065</td>
</tr>
<tr>
<td>subblock</td>
<td>4</td>
<td>0.0036</td>
<td>0.0024</td>
</tr>
<tr>
<td>GS-LEX</td>
<td>1</td>
<td>0.0114</td>
<td>0.0082</td>
</tr>
<tr>
<td>GS-LEX</td>
<td>4</td>
<td>0.0045</td>
<td>0.0032</td>
</tr>
<tr>
<td>GS-RB</td>
<td>1</td>
<td>0.0112</td>
<td>0.0075</td>
</tr>
<tr>
<td>GS-RB</td>
<td>4</td>
<td>0.0044</td>
<td>0.0031</td>
</tr>
<tr>
<td>Jac</td>
<td>1</td>
<td>0.1950</td>
<td>0.1199</td>
</tr>
<tr>
<td>Jac</td>
<td>4</td>
<td>0.0063</td>
<td>0.0044</td>
</tr>
</tbody>
</table>

with matrix

\[
\begin{pmatrix}
F_{1,1} & 0_{1 \times 3} \\
F_{2:3,1} & F_3
\end{pmatrix},
\]

where \( F_3(x, y) := F(2 : 4, 2 : 4)(x, y) \). As submatrix of \( F \), the subblock matrix \( F_3 \) is also positive semidefinite. Furthermore, \( F_3 \) is positive definite and is the block symbol of a symmetric positive definite well conditioned matrix. To solve a linear system in the subblock smoother we mainly have to solve a linear system in the well conditioned \( F_3 \). For the 5-point stencil it holds \( \text{cond}(F_3) = 3 + 2\sqrt{2} \). The condition number of the symmetric Gauss-Seidel preconditioned \( F_3 \) matrix is 2.

For the solution of linear systems with the subblock smoother we can reduce the matrix \( F_3 \) by the Schur complement to

\[
\begin{pmatrix}
4 & 0 & -\beta \\
0 & 4 & -\alpha \\
-\beta & -\alpha & 4
\end{pmatrix} =
\begin{pmatrix}
4 & 0 & 0 \\
0 & 4 & 0 \\
-\beta & -\alpha & s
\end{pmatrix}
\cdot
\begin{pmatrix}
1 & 0 & -\beta/4 \\
0 & 1 & -\alpha/4 \\
0 & 0 & 1
\end{pmatrix}
\]

with Schur complement

\[
s = 1 - |\alpha|^2/16 + |\beta|^2/16 = (16 - |\alpha|^2 - |\beta|^2)/16 = (6 - \cos(x) - \cos(y))/8.
\]

Therefore, for solving \( F_3 \) we mainly have to solve \( s \) with condition number bounded by 2. With symmetric Gauss-Seidel the condition number for \( s \) can further be reduced to 1.5. Hence, the solution of a linear system in the subblock smoother can be reduced to a linear system in \( s \); solving a linear system in \( s \) with symmetric Gauss-Seidel preconditioner and the preconditioned conjugate gradient method in view of condition number 1.5. Hence, the cost for smoothing with the subblock matrix is not essentially higher than for other standard smoothers.
Appendix B. The full projection.

In this paper the full projection \( b_{\text{full}} \), which was introduced in [14], is used for deriving new efficient MG methods. It is defined as

\[
b_{\text{full}}(x, y) := f(x, y + \pi)f(x + \pi, y)f(x + \pi, y + \pi),
\]

where \( f \) is the scalar generating function corresponding to the block symbol \( F \). The block symbol for the full projection can be described according to Lemma B.1.

**Lemma B.1.** The block symbol of the full projection is \( F^+ \), where \( F \) is the block symbol of the original problem.

**Proof.** For the scalar symbol \( f(x, y) = f_{0,0}(x, y) \) in 2D with projection on every second entry in each dimension, we define the three functions \( f_{0,1}(x, y) = f(x, y + \pi) \), \( f_{1,0}(x, y) = f(x + \pi, y) \) and \( f_{1,1}(x, y) = f(x + \pi, y + \pi) \). Sums and differences of these four scalar functions build the entries in the multilevel \( \omega \)-circulant block symbol \( F(x, y) \) as described in [14]. Furthermore, the \( \omega \) multilevel circulant structure yields

\[
det(F(x, y)) = f_{0,0}(x, y/2) \cdot f_{1,0}(x/2, y) \cdot f_{1,1}(x/2, y/2) \cdot f_{0,1}(x/2, y/2).
\]

Let \( b_{\text{full}}(x, y) := f_{0,0}(x, y) \cdot f_{1,0}(x, y) \cdot f_{1,1}(x, y) \). Then the block symbol for \( f_{0,0} \cdot b_{\text{full}} = det(F(x/2, y)) \) is equal to \( det(F(x, y) \cdot I) \), because for functions in \( 2x \) and \( 2y \) - in view of [14] - the nondiagonal entries in the block symbol are 0, and the main diagonal entries reproduce the scalar function up to a scaling factor 2 in \( x \) and \( y \). Therefore it holds \( F(x, y) \cdot B_{\text{full}}(x, y) = det(F(x, y)) \), and hence \( B_{\text{full}}(x, y) = F^+(x, y) \). □

REFERENCES


