Available online at www.sciencedirect.com
ScienceDirect
Applied Numerical Mathematics 58 (2008) 407–421
APPLIED NUMERICAL MATHEMATICS
www.elsevier.com/locate/apnum

Multigrid solution techniques for anisotropic structured linear systems

Rainer Fischer *, Thomas Huckle

Institut für Informatik, Technische Universität München, Boltzmannstr. 3, 85748 Garching b. München, Germany

Available online 24 February 2007

Abstract

Multigrid methods are highly efficient solution techniques for large sparse structured linear systems which are positive definite and ill-conditioned. In a recent paper [R. Fischer, T. Huckle, Multigrid methods for anisotropic BTTB systems, Linear Algebra Appl. (2005), submitted for publication], multigrid methods have been developed which are especially designed for anisotropic matrices belonging to the two-level Toeplitz class. These methods are primarily based on the use of a suitable combination of semicoarsening and full coarsening steps. In this paper the main focus is on the design of efficient smoothing techniques. Moreover, we are not only interested in two-level Toeplitz matrices, but also in matrices of two-level trigonometric matrix algebras. First, we describe methods for systems with anisotropy along coordinate axes. Although some of the ideas are known from the solution of partial differential equations, we present them here in a more formal way using generating functions and their level curves. This allows us not only to obtain theoretical results on convergence and reduction of anisotropy, but also to carry over the results to systems with anisotropy in other directions. We introduce new coordinates in order to describe these more complicated systems in terms of generating functions. This enables us to develop smoothers which are especially suitable for these more complicated systems.

PACS: 65N55; 65F10; 65F22; 65F35

Keywords: Multilevel Toeplitz; Circulant and tau algebra; Multigrid methods; Iterative methods; Preconditioners

1. Introduction

Multigrid methods belong to the fastest iterative methods for the solution of large sparse structured linear systems of equations. Many applications such as discretization of partial differential equations lead to two-level Toeplitz, tau, or circulant systems. If these are positive definite, i.e. corresponding to strictly positive generating functions, they are easily solved with the conjugate gradient algorithm preconditioned by circulant or tau matrices [4]. If the generating function has up to a finite number of zeros, the use of multigrid methods is significantly more efficient, see e.g. [5,3]. If systems are anisotropic, the classical convergence theory for (multilevel) Toeplitz, tau, or circulant matrices [3,11, 12,2] still holds, but standard multigrid methods converge so slowly that they become totally impractical. Therefore, we have devised a multilevel method which is especially designed for application to anisotropic two-level Toeplitz

* Corresponding author.
E-mail addresses: fischerr@in.tum.de (R. Fischer), huckle@in.tum.de (T. Huckle).

0168-9274/$30.00 © 2007 Published by Elsevier B.V. on behalf of IMACS.
doi:10.1016/j.apnum.2007.01.022

problems [6]. This method is entirely based on the use of semicoarsening combined with standard smoothers. The main results are theoretical, concerning the reduction of anisotropy and convergence properties of the two-grid method. In this paper, multilevel methods are developed not only for two-level Toeplitz systems, but also for matrices belonging to an class such as circulant and tau matrices with special attention to the anisotropic case (for the isotropic case see [3, 11,12,2]). Especially if these systems are sparse, multigrid methods are the fastest iterative solvers [5]. Furthermore, we focus on the use of more sophisticated smoothing techniques which can be combined with standard coarsening as well as semicoarsening. Generating functions will be a very helpful tool for the development of our methods. Numerical results will be given to illustrate and compare the different multilevel methods.

The article is organized as follows. In Section 2 we explain fundamental properties of multigrid methods for structured linear systems and describe the problems arising from anisotropic systems. In Section 3 we consider matrices corresponding to generating functions where anisotropy occurs along coordinate axes. After reviewing some results from [6] we focus on the use of suitable smoothers and on the numerical comparison of the different methods. The problems considered in Section 4 are more difficult to solve, because anisotropy occurs in arbitrary directions. We develop multigrid methods which are suitable for this case by carrying over the results from Section 3, again laying special emphasis on the use of smoothers and on numerical tests.

2. Multigrid methods for structured linear systems

Both two-level Toeplitz (BTTB) matrices and matrices forming an algebra such as two-level tau (B$\tau\tau$B) or circulant (BCCB) matrices are closely related to generating functions. These functions will be essential in this paper for the derivation of multigrid methods with certain properties. For the Toeplitz case the correspondence between matrices and generating functions is for example described in [10,7]. Suppose $f(x, y)$ is a real-valued Lebesgue integrable function which is defined on $[-\pi, \pi]^2$, and periodically extended on the whole plane. Then the Fourier coefficients of f are given by

$$a_{k,l} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(x, y)\, \mathrm{e}^{-\mathrm{i}kx - \mathrm{i}ly}\, \mathrm{d}x\, \mathrm{d}y \quad (k, l \in \mathbb{Z}).$$

$A_{mn}[f]$ is the corresponding mn-by-mn BTTB matrix with entries $(A_{mn}[f])_{(j,k)(p,q)} = a_{j-k,p-q}$ $(0 \leqslant j, k < m,$ $0 \leqslant p, q < n)$, where (j, k) indicates the block in $A_{mn}[f]$ and (p, q) the position within the block. If $f_{\min}$ and $f_{\max}$ denote the infimum and supremum values of f, and if $f_{\min} < f_{\max}$, then for all $m, n \geqslant 1$, the eigenvalues of $A_{mn}[f]$ lie in the interval $(f_{\min}, f_{\max})$. For $n, m \to \infty$, the extreme eigenvalues tend to $f_{\min}$ and $f_{\max}$.

Matrices belonging to a two-level trigonometric algebra are diagonalized by a unitary transform Q_{mn}, i.e. they are of the form

$$A_{mn}[f] = Q_{mn}^H \Lambda_{mn}[f] Q_{mn}, \tag{1}$$

where $\Lambda_{mn}[f]$ is the diagonal matrix containing the eigenvalues $\lambda_{k,l}$ of $A_{mn}[f]$. For two-level circulant matrices these are given by $\lambda_{k,l} = f(\frac{2\pi l}{m}, \frac{2\pi k}{n})$ with $0 \leqslant k \leqslant n-1, 0 \leqslant l \leqslant m-1$, and for two-level tau matrices by $\lambda_{k,l} = f(\frac{\pi l}{m+1}, \frac{\pi k}{n+1})$ with $1 \leqslant k \leqslant n, 1 \leqslant l \leqslant m$. This implies that unlike Toeplitz matrices, matrices from trigonometric algebras can become singular if f is zero at one of the grid points. If this happens, $A_{mn}[f]$ is usually replaced by the so called Strang correction [15], which was originally defined for circulant matrices, but which can be used for other matrix algebras as well [2].

In recent years, multigrid methods turned out to be the most efficient solution techniques for ill-conditioned linear systems of Toeplitz, tau, or circulant type whose generating function has isolated zeros [5,11,2,12,13,8]. Alle these methods are based on the algebraic multigrid method (AMG) developed by Ruge and Stüben [9], which as a purely algebraic method does not use real grids. To develop an AMG method we have to define a smoother and a coarse grid correction operator on each level. Usually, a simple smoother such as the damped Richardson or Jacobi method is applied, which is denoted by $S : \mathbb{R}^{mn} \to \mathbb{R}^{mn}$. To compute the coarse grid correction operator we need to define a restriction matrix $P : \mathbb{R}^{mn} \to \mathbb{R}^{m_C n_C}$ with $n_C m_C$ being the dimension of the coarse grid system matrix A_C. The transpose P^{T} is chosen to be the prolongation matrix. The matrix $P = BE$ formally consists of two parts. B is defined to deal with the zero of f, whereas E is the two-dimensional elementary projection matrix, picking every second column and every second block column of a matrix. It is obtained from the one-dimensional projection matrices E_m

and E_n by $E = E_m \otimes E_n$. In the circulant case, we use $E_n(:, 1:2:n)$ with even n, whereas in the Toeplitz and tau case $E_n(:, 2:2:n)$ with odd n. If the Toeplitz structure should be retained on coarser levels, which is for example necessary for some convergence proofs, we have to apply a cutting matrix, which eliminates the first few rows and columns of the coarse grid matrix. Different cutting matrices were suggested in [2]. The matrix A_C is then computed with the Galerkin approach, i.e. as the product

$$A_C = P A_{mn}[f] P^{\mathrm{T}} = E^{\mathrm{T}}\big(B^{\mathrm{T}} A_{mn}[f] B\big) E. \tag{2}$$

The product $\hat{A} = B^{\mathrm{T}} A_{mn}[f] B$ translated into generating functions gives

$$\hat{f}(x, y) = f(x, y) b(x, y)^2, \tag{3}$$

whereas the elementary projection $A_C = E^{\mathrm{T}} \hat{A} E$ becomes

$$f_2(x, y) = \frac{1}{4}\left[\hat{f}\left(\frac{x}{2}, \frac{y}{2}\right) + \hat{f}\left(\frac{x}{2} + \pi, \frac{y}{2}\right) + \hat{f}\left(\frac{x}{2}, \frac{y}{2} + \pi\right) + \hat{f}\left(\frac{x}{2} + \pi, \frac{y}{2} + \pi\right)\right]. \tag{4}$$

In the Toeplitz and tau case $x/2 + \pi$ and $y/2 + \pi$ are replaced by $\pi - x/2$ and $\pi - y/2$. This means f_2 is obtained from the Fourier series of $\hat{f}$ by picking every second coefficient in x and every second coefficient in y. So far, we have not yet chosen a function $b(x, y)$ for prolongation. Fiorentino and Serra [5] suggest to use a nonnegative function which is strictly positive at (x_0, y_0) and zero at the so called mirror points

$$M\big((x_0, y_0)\big) := (x_0 + \pi, y_0), (x_0, y_0 + \pi), \text{ and } (x_0 + \pi, y_0 + \pi). \tag{5}$$

In the Toeplitz and tau case, terms such as $x_0 + \pi$ are replaced by $\pi - x_0$. For this choice they showed that A_C is also positive definite, and its generating function has the single zero $(2x_0, 2y_0)$. The choice

$$b(x, y) = \big(1 + \cos(x - x_0)\big)\big(1 + \cos(y - y_0)\big), \tag{6}$$

satisfies these properties and corresponds to a matrix which is extremely sparse. The coarse grid correction operator can be written as $X = I_{mn} - P^{\mathrm{T}} A_C^{-1} P A_{mn}[f]$, leading to a global iteration matrix of the two-level method $G = S^{\nu_2} X S^{\nu_1}$, where ν_1 denotes the number of presmoothing steps and ν_2 the number of postsmoothing steps. If the zero at (x_0, y_0) is of order higher than 2, $b(x, y)$ from (6) is not enough. In [2] the authors point out that $b(x, y)$ must be chosen as a power of $(1 + \cos(x - x_0))(1 + \cos(y - y_0))$ in this case. A multilevel method is defined by using the two-level method recursively to approximate the inverse of A_C.

All convergence proofs developed so far for our structured matrix classes are based on general convergence results by Ruge and Stüben [9]. In order to state their theorem for the two-level method we must define the following inner products in addition to the Euclidean inner product $\langle u, v\rangle$ for the system matrix A:

$$\begin{aligned} &\langle u, v\rangle_0 = \big\langle \operatorname{diag}(A) u, v\big\rangle, \langle u, v\rangle_1 = \langle Au, v\rangle, \\ &\langle u, v\rangle_2 = \big\langle \operatorname{diag}(A)^{-1} Au, Av\big\rangle. \end{aligned} \tag{7}$$

The respective norms, which are derived from these inner products, are denoted $\|\cdot\|_i, i = 0, 1, 2$.

Theorem 1. *(Ruge and Stüben [9]) Let A be a positive definite mn-by-mn matrix, and let S be a smoother satisfying the presmoothing and the postsmoothing condition, i.e. there exist $\alpha_{\mathrm{pre}}, \alpha_{\mathrm{post}} > 0$ such that*

$$\big\|Se^h\big\|_1^2 \leqslant \big\|e^h\big\|_1^2 - \alpha_{\mathrm{pre}} \big\|Se^h\big\|_2^2, \quad \forall e^h \in \mathbb{R}^{mn}. \tag{8}$$

$$\big\|Se^h\big\|_1^2 \leqslant \big\|e^h\big\|_1^2 - \alpha_{\mathrm{post}} \big\|e^h\big\|_2^2, \quad \forall e^h \in \mathbb{R}^{mn}. \tag{9}$$

Furthermore, suppose that the restriction operator P has full rank and that the correcting condition is satisfied, i.e. there exists a $\beta > 0$ such that

$$\min_{e^H \in \mathbb{R}^{m_C n_C}} \big\|e^h - P^{\mathrm{T}} e^H\big\|_0^2 \leqslant \beta \big\|e^h\big\|_1^2, \quad \forall e^h \in \mathbb{R}^{mn}. \tag{10}$$

Then $\beta > \alpha_{\mathrm{post}}$, and the convergence factor of the two-level method $\|TG\|_1$ is bounded by

$$\|TG\|_1 \leqslant \sqrt{\frac{1 - \alpha_{\mathrm{post}}/\beta}{1 + \alpha_{\mathrm{pre}}/\beta}}.$$

Remark 2. Theorem 1 still holds if $\operatorname{diag}(A)$ in (7) is replaced by any Hermitian positive definite matrix Y (see Remark 2.2 in [2] for the mathematical motivation). We will make use of this degree of freedom in the proof of Theorem 7.

Sun, Jin and Chang [13] use Theorem 1 to prove the optimal convergence rate of the two-grid method and level-independency for BTTB systems where f has a zero of order at most two in the origin, i.e. where f satisfies

$$\min_{(x,y)\in[-\pi,\pi]^2} \frac{f(x,y)}{2-\cos x-\cos y} = C > 0. \tag{11}$$

This is sufficient to obtain W-cycle convergence, but not necessarily V-cycle convergence, see the counterexamples in [2] and, for a more general discussion of V-cycle and W-cycle convergence, the book of Trottenberg et al. [14]. For multilevel tau and circulant matrices optimal V-cycle convergence, also for functions with zeros of higher order, has been proved in [2]. The conditions

$$\limsup_{(x,y)\to(x_0,y_0)} \left|\frac{b_j((x_i,y_i))}{f_j((x,y))}\right| < \infty \quad \text{for } (x_i,y_i)\in M\big((x_0,y_0)\big), \tag{12}$$

$$0 < \sum_{(x_i,y_i)\in M((x_0,y_0))\cup\{(x_0,y_0)\}} b_j^2\big((x_i,y_i)\big) \tag{13}$$

must hold on each grid j of the multigrid method.

Remark 3. Let $f(x,y)$ be a nonnegative generating function which has a zero at $(0,0)$. If f has another zero at one of the mirror points from (5), the multigrid method from above fails completely in all numerical experiments. If f is zero at the origin and at one of the mirror points, conditions (12) and (13) cannot be satisfied for both points. Therefore, the convergence theory developed in [11,2,1,12] does not hold anymore: this observation is motivated theoretically in Remark 5.1 of [2] and, even in more detail, at pages 60 and 70 of [12] even if its origin can be tracked in Section 5.1 of [5]. Furthermore, at p. 60 of [12] there is a short discussion on possible constructive proposals, by R. Chan et. al. and by Huckle, in order to overcome this problem (see Refs. [6] and [12] in the paper [12]). Even if f is close to zero at one of the three points $(0,\pi)$, $(\pi,0)$, (π,π), convergence of the multigrid method is extremely slow.

This remark will be crucial for anisotropic systems, which arise for example from discretization of partial differential equations. One important model problem for our experiments is obtained from the following equation, which is closely related to the Poisson equation $-\epsilon u_{xx} - u_{yy} = f$. Finite difference discretization of this equation with a five point stencil on a uniform mesh leads to the following linear system.

Example 4. Let $A_{mn}[f]$ be the matrix corresponding to the generating function

$$f(x,y) = \alpha\big(1-\cos(x)\big) + \big(1-\cos(y)\big). \tag{14}$$

If $\alpha = 1$, we get one of the isotropic standard model problems, the discrete Poisson equation. For $\alpha \ll 1$, the problem becomes strongly anisotropic.

The function

$$f(x,y) = \alpha x^2 + y^2 \quad (\alpha \ll 1) \tag{15}$$

corresponds to linear systems which have a similar type of anisotropy, but which are not sparse.

With these functions we can illustrate why the standard multigrid methods from Section 2 should not be used for the solution of anisotropic systems. Again, generating functions turn out to be a helpful tool for the analysis of multigrid methods.

- If the anisotropy is strong, i.e. if $\alpha \ll 1$, the function f becomes close to zero on the whole x-axis, and especially at $(\pi,0)$. From Remark 3 we know that convergence is extremely slow in this case.

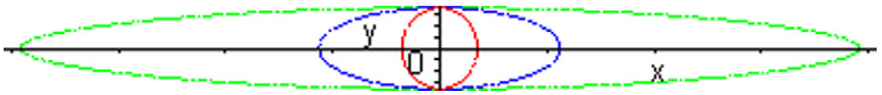

Fig. 1. Curves $f(x, y) = 0.01$ for the function from Example 4 with $\alpha = 1, 0.1, 0.01$.

- Even for moderately anisotropic systems we obtain a weak connection in one direction due to small coefficients in front of the x-terms in f. This can be described very well if we look at level curves of the generating function f from Example 4. Fig. 1 depicts the curve $f(x, y) = 0.01$ for three different values of α, i.e. for three different degrees of anisotropy. The solid curve for $\alpha = 1$ is totally isotropic, i.e. almost like a circle. The dashed curve with $\alpha = 0.1$ is only moderately anisotropic, and the dotted curve for $\alpha = 0.01$ is significantly more anisotropic, which means that the curve is very flat. If anisotropy is even stronger, the value of the function hardly depends on x, making coarsening in x-direction rather useless.

In the following we wish to develop multigrid methods for anisotropic systems with special emphasis on the development of suitable smoothers. Therefore, we divide these systems into two classes. The first class contains matrices where anisotropy occurs along coordinate axes such as the matrices from Example 4. The methods will be described in a formal way using the notation of generating functions, giving us a different view on this type of anisotropic problems which are known from the solution of partial differential equations. This enables us to carry over the results to the second class of anisotropic problems, where anisotropy occurs in other directions.

3. Anisotropy along coordinate axes

Two-level structured matrices where anisotropy occurs along coordinate axes were introduced in Example 4. Multigrid methods which do not suffer from the problems described in the previous section can be constructed in two different ways, either using semicoarsening or more sophisticated smoothers. First, we review some results from [6], where all multigrid methods are constructed with semicoarsening. Then, we develop methods using line smoothers and standard coarsening, and finally, both methods are tested and compared numerically.

3.1. Theoretical results on semicoarsening

One possible way to get rid of the two problems described in the previous section is to use semicoarsening in the direction perpendicular to the anisotropy. The smoother is then chosen to be a pointwise one such as the damped Jacobi method. For example, if coarsening is done in y-direction only, the function corresponding to the matrix B in $P = BE$ is

$$b(x, y) = 1 + \cos(y - y_0), \tag{16}$$

instead of (6). If the zero is of higher order, we chose $(1 + \cos(y - y_0))^k$. The product $\hat{A} = BA_{mn}[f]B$, i.e. $\hat{f}(x, y) = f(x, y)b(x, y)^2$ is computed as in the isotropic case. The elementary projection matrix E is chosen to be $E_{mn} = I_m \otimes E_n$ with the one-dimensional projection matrix E_n from Section 2. Translated to generating functions this becomes

$$f_2(x, y) = \frac{1}{2}\left(\hat{f}\left(x, \frac{y}{2}\right) + \hat{f}\left(x, \frac{y}{2} + \pi\right)\right). \tag{17}$$

We obtain a coarse grid matrix A_C with half as many blocks as $A_{mn}[f]$, but with the same block size. In the following, let us assume without loss of generality that the zero (x_0, y_0) of $f(x, y)$ is located at the origin.

In [6] we obtained two important theoretical results. The first of them concerns the reduction of anisotropy, which is measured with the ratio $r_F = x_F/y_F$ of the points $(x_F, 0)$ and $(0, y_F)$ where curves $f(x, y) = c$ intersect x- and y-axis for very small $c > 0$. Since it is stated purely in terms of generating functions, it holds for circulant and tau matrices as well.

Theorem 5. *Let f be a nonnegative generating function with a zero of order 2 at the origin which is of the form*

$$f(x, y) = \big[\lambda_1\big(1 - \cos(x)\big) + \lambda_2\big(1 - \cos(y)\big)\big]h(x, y) \tag{18}$$

with $h(x, y) > 0$ and $\lambda_1, \lambda_2 > 0$. Let f_2 be the function obtained by one semicoarsening step with b from (16)*. Then the degree of anisotropy, measured with r_F, is reduced by a factor* 2*. Since f_2 is also of the form* (18)*, each semicoarsening step reduces the anisotropy by a factor* 2*.*

This theorem yields a straightforward heuristic for the development of a multigrid method: *Apply semicoarsening until the system is not anisotropic anymore, i.e. until r_F is reduced to almost* 1*, then switch to full coarsening.*

The other theoretical result from [6] concerns two-grid convergence and level independency. The proofs of [3] and [13] are carried over to anisotropic two-level Toeplitz systems. If anisotropy occurs along the x-axis, i.e. if $f(x, y)$ is small for $y = 0$ and all $x \in [0, 2\pi]$, coarsening is done only in y. In this case, f is allowed to be zero on the whole line $y = 0$. Thus, (11) is replaced by

$$\min_{(x,y)\in[-\pi,\pi]^2} \frac{f(x, y)}{1 - \cos y} = C > 0. \tag{19}$$

The following theorem proves convergence of the two-level method.

Theorem 6. *([6]) Let $A_{mn}[f]$ be a positive definite BTTB matrix whose generating function f is real-valued even and satisfies* (19)*. Moreover, let the prolongation matrix be given by* (16)*, and let the smoother be the damped Richardson or Jacobi method.*

Then, the convergence factor of the two-level method is uniformly bounded below 1 *independent of m and n. The following estimate for the convergence factor holds*:

$$\|TG\|_1 \leqslant \sqrt{1 - C/\big(2\rho\big(A_{mn}[f]\big)\big)}.$$

It is also shown that if (19) holds on some level, it also holds on the next coarser level after one semicoarsening step. If q levels of semicoarsening are used, the following estimate for the convergence factor holds:

$$\big\|TG^q\big\|_1 \leqslant \sqrt{1 - \frac{\alpha^q}{\beta^q}} = \sqrt{1 - \frac{\alpha^h}{4^{q-1}\beta^h}}. \tag{20}$$

A similar estimate holds if we mix semicoarsening and full coarsening steps.

3.2. The use of line smoothers

Anisotropic systems can be solved with a different multigrid strategy where standard coarsening from Section 2 can still be used. However, this requires application of an adequate smoother on each level. Line smoothers such as the damped block Jacobi method smooth along a whole line of grid points, which corresponds to a block of unknowns in the solution vector. These unknowns, and therefore also the rows and columns of the matrix, must be permuted and then grouped into blocks. Fig. 2 shows how the blocks of the matrix are built, i.e. how the rows and columns of

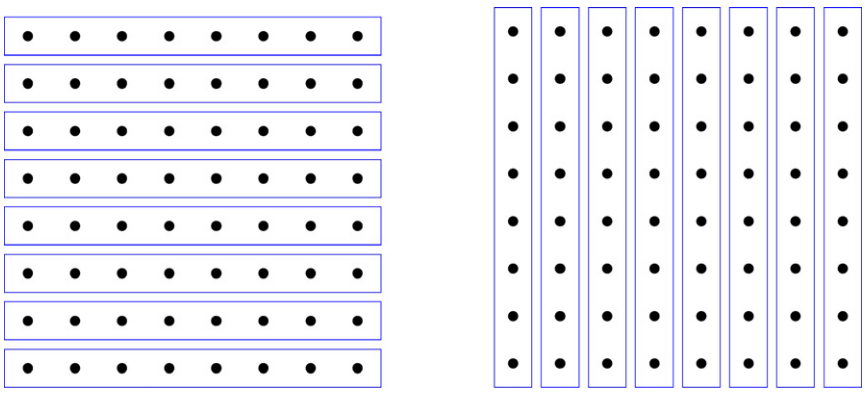

Fig. 2. Partitioning of the original matrix into blocks.

$A_{mn}[f]$ must be permuted for smoothing. Each point in the pictures corresponds to one unknown in the solution and to one line or column of the matrix, similar to the discretization of PDEs. If anisotropy occurs along the y-axis, the blocks must be constructed as it is shown in the left picture. No permutation is necessary, and all blocks have size n. If anisotropy occurs along the x-axis, the right picture applies. The rows and the columns are permuted with the vector

$$\begin{aligned}&\big(1, n+1, 2n+1, \ldots, (m-1)n+1, 2, n+2, 2n+2, \ldots, (m-1)n+2, \ldots,\\ &\quad n-1, n+n-1, 2n+n-1, \ldots, (m-1)n+n-1\big),\end{aligned}$$

which can also be interpreted as a change of the variables, i.e. as a new coordinate system, where x and y are interchanged. This idea will become more important in Section 4. The blocks in this case are of size m. The following theorem shows that the block Jacobi method indeed satisfies the smoothing conditions (8) and (9) of the Ruge–Stüben-theorem if anisotropy occurs along the y-axis. The condition for anisotropy in x-direction follows immediately.

Theorem 7. *Let f be a nonnegative generating function with a zero of order 2 at the origin which is of the form*

$$f(x,y) = \big[\big(1-\cos(x)\big) + \alpha\big(1-\cos(y)\big)\big]h(x,y) \tag{21}$$

with $0 < \alpha \ll 1$ and the trigonometric polynomial h satisfying $0 < h_{\min} \leqslant h(x,y) \leqslant h_{\max} < \infty$. Let $A = A_{mn}[f]$ be the corresponding two-level tau or circulant matrix, and D the block diagonal matrix with the same diagonal blocks as A, corresponding to the generating function

$$g(x,y) = \big(1+\alpha-\cos(x)\big)\tilde{h}(x). \tag{22}$$

$\tilde{h}$ is obtained by eliminating all terms in h containing y. Let $M_f, M_g, M_{f/g}$ denote the maximum values of $f, g, f/g$.

If $\tilde{h}(x) > 0$, then the block Jacobi method $x^{(k+1)} = x^{(k)} + \omega D^{-1}(b - Ax^{(k)})$, i.e. $S = I - \omega D^{-1}A$, satisfies the smoothing conditions (8) and (9). *More precisely, for $0 \leqslant \omega \leqslant 2/M_{f/g}$ there exist nonnegative $\alpha_{\text{pre}}, \alpha_{\text{post}}$ with*

$$\begin{aligned}&\alpha_{\text{pre}} \leqslant \min\left\{2\omega, \frac{\omega(2-\omega M_{f/g})}{(1-\omega M_{f/g})^2}\right\},\\ &\alpha_{\text{post}} \leqslant \omega(2-\omega M_{f/g}).\end{aligned} \tag{23}$$

Proof. First of all, we show that the functions f, g, and f/g are bounded by M_f, M_g, and $M_{f/g}$.

$$\begin{aligned}M_f &= \max_{(x,y)\in[-\pi,\pi]^2} f(x,y) = (2+2\alpha)h_{\max},\\ M_g &= \max_{(x,y)\in[-\pi,\pi]^2} g(x,y) = (2+\alpha)\tilde{h}_{\max},\\ \frac{f(x,y)}{g(x,y)} &\leqslant \frac{1-\cos(x)+\alpha(1-\cos(y))}{1-\cos(x)+\alpha}\frac{h_{\max}}{\tilde{h}_{\min}} \leqslant \left(1+\frac{-\alpha\cos(y)}{1-\cos(x)+\alpha}\right)\frac{h_{\max}}{\tilde{h}_{\min}}\\ &\leqslant \left(1+\frac{-\alpha\cdot(-1)}{\alpha}\right)\frac{h_{\max}}{\tilde{h}_{\min}} \leqslant 2\frac{h_{\max}}{\tilde{h}_{\min}}.\end{aligned}$$

This implies that $M_{f/g} = \max_{(x,y)\in[-\pi,\pi]^2} f(x,y)/g(x,y) = 2h_{\max}/\tilde{h}_{\min}$. Moreover, f/g has the minimum value $m_{f/g} = 0$.

In the following we prove the smoothing conditions with $Y = D^{-1}$ (see Remark 2). It should be noted that the proof technique of translating the conditions of Theorem 1 into function inequalities was introduced in [11]. In [1] the authors use a similar proof technique with $Y = I$ for the damped Richardson method. With our choice of Y the presmoothing condition (8) can be written $SAS \leqslant A - \alpha_{\text{pre}} SA^2D^{-1}S$, which is equivalent to

$$\big(I-\omega D^{-1}A\big)A\big(I-\omega D^{-1}A\big) \leqslant A - \alpha_{\text{pre}}\big(I-\omega D^{-1}A\big)D^{-1}A^2\big(I-\omega D^{-1}A\big). \tag{24}$$

This is implied by the function inequality

$$\left(1-\omega\frac{f}{g}\right)f\left(1-\omega\frac{f}{g}\right) \leqslant f - \alpha_{\text{pre}}\left(1-\omega\frac{f}{g}\right)\frac{f^2}{g}\left(1-\omega\frac{f}{g}\right), \tag{25}$$

and because of $f/g \geqslant 0$ by

$$1 + \alpha_{\text{pre}} \frac{f}{g} \leqslant \frac{1}{(1-\omega f/g)^2}. \tag{26}$$

There exists a nonnegative α_{pre} in (26) only for $0 \leqslant \omega \leqslant \frac{2}{M_{f/g}}$. Since f/g can take values between 0 and $M_{f/g}$, (26) holds if

$$1 + \alpha_{\text{pre}} t \leqslant \frac{1}{(1-\omega t)^2} \tag{27}$$

is true for all $0 < t \leqslant M_{f/g}$. As in [1] we can deduce that this holds if $\alpha_{\text{pre}} \leqslant 2\omega$ for $0 \leqslant \omega \leqslant \frac{2}{M_{f/g}}$, and in addition $1 + \alpha_{\text{pre}} M_{f/g} \leqslant \frac{1}{(1-\omega M_{f/g})^2}$ for $\frac{1}{M_{f/g}} < \omega \leqslant \frac{2}{M_{f/g}}$.

The postsmoothing condition is equivalent to

$$\big(I - \omega D^{-1} A\big) A \big(I - \omega D^{-1} A\big) \leqslant A - \alpha_{\text{post}} D^{-1} A^2. \tag{28}$$

This is translated to generating functions and because of $f/g \geqslant 0$ simplified to

$$\left(1 - \omega \frac{f}{g}\right)^2 \leqslant 1 - \alpha_{\text{post}} \frac{f}{g}, \tag{29}$$

which leads to nonnegative α_{post} only for $0 \leqslant \omega \leqslant \frac{2}{M_{f/g}}$. (29) is satisfied if

$$(1-\omega t)^2 \leqslant 1 - \alpha_{\text{post}} t \tag{30}$$

holds for $0 < t \leqslant M_{f/g}$. Similar to [1] this is shown to be true if $1 - \alpha_{\text{post}} M_{f/g} \geqslant (1 - \omega M_{f/g})^2$. □

In the following corollary we deduce optimal values for $\alpha_{\text{pre}}, \alpha_{\text{post}}, \omega$ from (23).

Corollary 8. *Under the same assumptions as in Theorem 7 we obtain the following optimal values for the parameters* $\alpha_{\text{pre}}, \alpha_{\text{post}}, \omega$:

(1) *If one presmoothing step and no postsmoothing is performed, then*

$$\omega_{\text{best}} = \frac{3}{2M_{f/g}}, \qquad \alpha_{\text{pre,best}} = \frac{3}{M_{f/g}}.$$

(2) *If one postsmoothing step and no presmoothing is performed, then*

$$\omega_{\text{best}} = \frac{1}{M_{f/g}}, \qquad \alpha_{\text{post,best}} = \frac{1}{M_{f/g}}.$$

(3) *If only one step of smoothing shall be performed, then the optimal rate of convergence is obtained with*
- *one presmoothing step with* $\omega = 3/(2M_{f/g})$ *if* $\alpha_{\text{post}}/\beta \in [0, \frac{2}{3}]$,
- *one postsmoothing step with* $\omega = 1/M_{f/g}$ *if* $\alpha_{\text{post}}/\beta \in [\frac{2}{3}, 1]$.

Proof. The proof uses the same technique as the one in [12]. The first two parts are proved using the estimates for α_{pre} and α_{post} from Theorem 7. The first inequality in (23) implies that

$$\alpha_{\text{pre,best}} = \max_{\omega \in (0, 2/M_{f/g})} \begin{Bmatrix} 2\omega & \text{if } \omega \leqslant 3/(2M_{f/g}) \\ (\omega(2-\omega M_{f/g}))/(1-\omega M_{f/g})^2 & \text{if } \omega > 3/(2M_{f/g}) \end{Bmatrix},$$

which is obtained for $\omega = 3/(2M_{f/g})$. The second inequality in (23) leads to

$$\alpha_{\text{pre,best}} = \max_{\omega \in (0, 2/M_{f/g})} \omega(2 - \omega M_{f/g}).$$

The third part is proved by comparing the convergence factors from Theorem 1. One presmoothing step without postsmoothing leads to $\sqrt{1/(1+\alpha_{\text{pre,best}}/\beta)} = \sqrt{1/(1+3\alpha_{\text{post,best}}/\beta)}$, whereas one postsmoothing step without presmoothing leads to $\sqrt{1-\alpha_{\text{post,best}}/\beta}$. □

Table 1
Iteration numbers for $T_{nn}[f]$ with f from Example 4 with $\alpha = 0.001$

Coarsening	$n = 2^6 - 1$	$n = 2^7 - 1$	$n = 2^8 - 1$
y,xy,xy,xy,xy	170	> 200	>200
y,y,y,xy,xy	12	19	23
y,y,y,y,y	7	7	7

Table 2
Iteration numbers for $T_{nn}[f]$ with f from (15) with $\alpha = 0.001$

Coarsening	$n = 2^5 - 1$	$n = 2^6 - 1$	$n = 2^7 - 1$
y,xy,xy,xy,xy	76	182	>200
y,y,y,xy,xy	9	11	15
y,y,y,y,y	9	9	9

Remark 9. An extension of Theorem 7 and Corollary 8 to the BTTB case seems rather difficult. For example, the implication from (24) to (25) does not hold in the Toeplitz case. We have information on the localization of the spectrum of A, A^2, $D^{-1}A$ (see [10]), but not on their linear combinations. The nontrivial structure of D poses a serious problem. Nonetheless, block Jacobi smoothing leads to fast multigrid convergence for BTTB systems.

3.3. Numerical results

In the previous two sections we have theoretically analyzed two different multigrid methods for anisotropic problems. Now these methods shall be tested and compared numerically. The first type of multilevel method consists of a certain number of semicoarsening steps followed by full coarsening steps. The damped Jacobi method or the Gauss–Seidel method is used as a smoother. Theorem 5 states that the ratio r_F is reduced by a factor 2 with each semicoarsening step. Therefore, we apply semicoarsening until this ratio is close to one, i.e. until the system is not anisotropic anymore, and then proceed with standard coarsening. For the function $f(x, y)$ from Example 4 with $\alpha = 0.001$ the ratio r_F is 31.62, which means that on the five finest levels semicoarsening should be performed. In our numerical experiments we test different coarsening strategies on the matrix $T_{nn}[f]$, which belongs both to the two-level Toeplitz and the two-level tau class. The first strategy (denoted y,xy,xy,xy,xy) consists of one semicoarsening step, followed by four full coarsening steps, the second (y,y,y,xy,xy) of three semicoarsening steps and two full coarsening steps, and the third (y,y,y,y,y) of five semicoarsening steps. Table 1 shows the number of V-cycle iterations our method requires until the residual is smaller than 10^{-6}. These results are obtained with one presmoothing step and one postsmoothing step on each level with the symmetric Gauss–Seidel method. If the damped Jacobi method is used, the difference between the coarsening strategies is even more striking. Since we are also interested in multilevel Toeplitz matrices which are not necessarily connected with PDEs, we take a look at the BTTB matrices with f from (15), which are not sparse. Again, we choose $\alpha = 0.001$ and use one iteration of the Gauss–Seidel method as pre- and postsmoother. Table 2 shows a similar behavior of the multigrid method as we have observed in Example 4. Again, the damped Jacobi smoother leads to similar results. It is cheaper concerning computational cost, but the V-cycle requires a few more iterations. Multigrid methods for circulant matrices are usually only efficient if the matrix is sparse, because for dense matrices the inverse can be directly computed in $O(n \log(n))$ with the FFT. The circulant matrix $C_{mn}[f]$ with f from Example 4, however, is singular. Therefore, we add $\frac{1}{(mn)^2} I_{mn}$ to $C_{mn}[f]$, which corresponds to a shift of the grid points and results in the ill-conditioned matrix $\tilde{C}_{mn}[f]$. In our numerical experiments we use a V-cycle with six levels and the damped Jacobi method as a smoother. For $\alpha = 0.02$, the value r_F suggest that we use three steps of semicoarsening followed by full coarsening. For the stronger anisotropy and $\alpha = 0.001$ five semicoarsening steps are supposed to yield optimal results. The results of the numerical calculations in Table 3 confirm that these suggestions lead to the fastest convergence in both cases.

A completely different strategy consists of using standard coarsening in combination with a line smoother such as the damped block Jacobi method. If the anisotropy is very strong, this method converges extremely fast, because the block diagonal matrix of the smoother is a very good approximation of $A_{mn}[f]$. If the problem is only mildly

Table 3
Iteration numbers for $\tilde{C}_{mn}[f]$ with f from Example 4 with $\alpha = 0.02$ and $\alpha = 0.001$

α	Coarsening	$n = 2^6$	$n = 2^7$	$n = 2^8$
0.02	y,xy,xy,xy,xy	24	23	23
0.02	y,y,y,xy,xy	5	5	5
0.02	y,y,y,y,y	8	8	8
0.001	y,xy,xy,xy,xy	>200	>200	>200
0.001	y,y,y,xy,xy	27	27	26
0.001	y,y,y,y,y	5	5	5

Table 4
Iteration numbers for $T_{nn}[f]$ with f from Example 4 with different α

α	$n = 2^6 - 1$	$n = 2^7 - 1$	$n = 2^8 - 1$
0.1	8	8	8
0.001	3	4	6
0.00001	2	2	3

Table 5
Iteration numbers for $\tilde{C}_{nn}[f]$ with f from Example 4 with different α

α	$n = 2^6 - 1$	$n = 2^7 - 1$	$n = 2^8 - 1$
0.1	5	5	5
0.001	5	5	5
0.00001	5	5	5

anisotropic, we observe the typical multigrid convergence behavior which is fairly fast and independent of the matrix size. In this case, standard coarsening still works, because the anisotropy is not so strong, and the block Jacobi methods has good smoothing properties. The most difficult case for our method are problems which are quite anisotropic, e.g. when α is between 0.01 and 0.001 in our examples. However, even in this case we obtain fast convergence if we apply two block Jacobi iterations as pre- and postsmoother, although the number of V-cycle iterations increases slightly with the matrix size. Table 4 summarizes the results for Example 4 and different degrees of anisotropy. In each case, we use a V-cycle with five levels and standard coarsening. Since the diagonal blocks have to be inverted, the block Jacobi method is too expensive for dense structured matrices. For sparse matrices, however, their results are comparable with those obtained by semicoarsening. Furthermore, it is possible to start with some semicoarsening steps, and then on the coarser levels use the block Jacobi method. Such a mixture of both types of algorithms is less expensive and further improves convergence, especially for those α where standard coarsening combined with a line smoother has difficulties.

The circulant matrix corresponding to the function $f(x, y)$ from Example 4 is also solved quite efficiently with standard coarsening and the block Jacobi smoother. Table 5 shows the number of V-cycle iterations. For BCCB matrices, the block Jacobi method can be applied even if the diagonal blocks are not sparse, because inversion of a circulant block has a generic case and worst case complexity of $O(n \log(n))$.

4. Anisotropy in other directions

The more general case of anisotropic linear systems, where anisotropy occurs in other directions, is more difficult to handle. However, we will carry over some results on generating functions from the previous section in order to solve these systems with similar techniques. The following two functions illustrate what problems arise from such matrices.

Example 10. Let $A_{mn}[f]$ and $A_{mn}[g]$ be the two-level Toeplitz, tau, or circulant matrices corresponding to one of the functions

$$f(x, y) = \alpha\big(1 - \cos(x + y)\big) + \big(1 - \cos(x - y)\big),$$

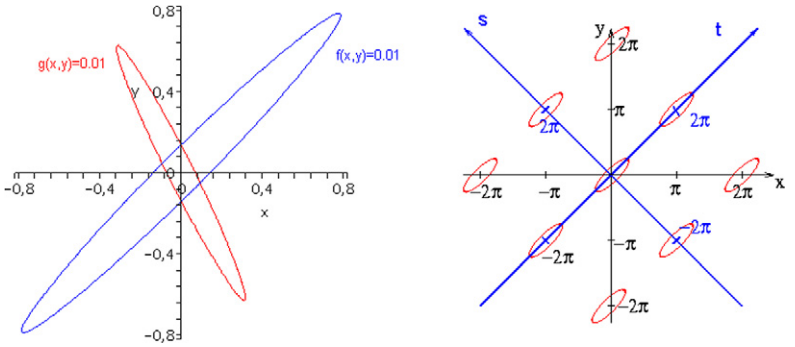

Fig. 3. Level curves $f(x, y) = 0.01$ and $g(x, y) = 0.01$ for the functions from Example 10 with $\alpha = 0.01$ (left) and introduction of new coordinates for $f(x, y)$ (right).

$$g(x, y) = \big(1 - \cos(2x + y)\big) + \alpha\big(1 - \cos(x - 2y)\big). \tag{31}$$

The left picture in Fig. 3 shows how the two functions behave in the neighborhood of their zero at the origin, i.e. with what kind of anisotropy we have to deal with. f is anisotropic along the line $y = x$, which means it is rotated by an angle of 45 degrees from the x-axis. The anisotropy of g occurs along $y = -2x$, which corresponds to an angle of 30 degrees from the y-axis.

If the anisotropy is strong, neither standard multigrid nor the methods from Section 3 work properly, and for $\alpha \to 0$ they fail completely. Semicoarsening along an axis does not help to treat anisotropy in other directions well. Furthermore, $f(x, y)$ has another zero at (π, π), which is a another obstacle to a convergent multigrid method, see Remark 3.

We will modify the ideas from Section 3 to make them suitable for these problems. In the first part, we focus on the coarsening strategy and use a simple smoother. First, we review results from [6] on BTTB systems. Then we define a similar semicoarsening method for BCCB matrices. The second part of this section defines a multilevel method which uses a line smoother and a block version of standard coarsening. Numerical tests will conclude this section. We will mainly use the 45° case to explain the basic features of the algorithms and only briefly mention generalizations to more general angles $\frac{k}{k+l}90°$.

4.1. Theoretical results on semicoarsening

Again, we wish to describe a multigrid method which uses a combination of semicoarsening and full coarsening steps and a standard smoother. The following strategy was suggested in [6].

- Before we start with the computation of prolongation and coarse grid functions, we define a new coordinate system (s, t) with

$$s := x - y \quad \text{and} \quad t := x + y \tag{32}$$

such that anisotropy occurs along one of the axes. Then, the function f from Example 10 becomes

$$\tilde{f}(s, t) = \alpha \cdot \big(1 - \cos(s)\big) + \big(1 - \cos(t)\big). \tag{33}$$

The right picture of Fig. 3 illustrates the consequences of this transformation. For $(s, t) \in [-\pi, \pi[^2$, the function $\tilde{f}$ has only one zero, and anisotropy occurs along the s-axis. This translates into matrices as a permutation of rows and columns and a partitioning of the resulting matrix into blocks. This is done as shown in the left picture of Fig. 4. Each block of the matrix corresponds to one diagonal in the picture. This means permutation must be done by the permutation vector

$$\begin{aligned}&\big(1, 2, n + 1, 3, n + 2, 2n + 3, \ldots, n, n + n - 1, 2n + n - 2, \ldots,\\ &\quad (m - 1)n + 1, \ldots, (m - 1)n, mn - 1, 2n + n - 1, mn\big),\end{aligned} \tag{34}$$

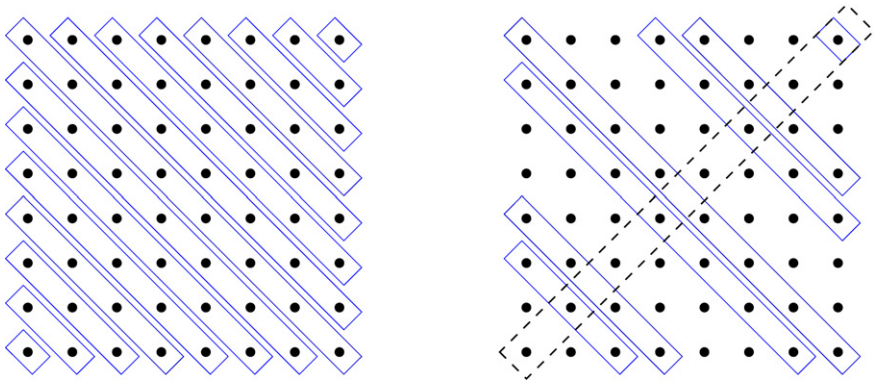

Fig. 4. Partitioning of the original matrix into blocks for semicoarsening (left) and full coarsening (right).

and the blocks are of size $1, 2, 3, \ldots, n-1, n, n-1, \ldots, 2, 1$. The resulting matrix is denoted $\tilde{T}_{mn}$.

- Semicoarsening must be performed in s- or t-direction. A semicoarsening step in t is defined by replacing x and y by s and t in (16) and (17). $b(s,t)$ corresponds to the block diagonal matrix

$$B_S = \operatorname{diag}(B_1, B_2, \ldots, B_n, \ldots, B_2, B_1), \tag{35}$$

  where $B_1 = 1$ and all other blocks are $B_k = A_k[1 + \cos(x)]$ of size k.
- In the Toeplitz and tau case, the blocks of $\tilde{T}$ and B_S are no longer of equal size, but we can nevertheless compute the coarse grid matrix A_C by applying elementary projection to the matrix $\hat{T} = B_S \tilde{T}_{nn} B_S$. This means we leave the number of blocks unchanged, and within each block we pick every second row and every second column.

The following two theoretical results are proved in [6], carrying over the results from Section 3.1:

(1) Each semicoarsening step reduces the degree of anisotropy by a factor 2;
(2) Convergence of the two-grid method for BTTB matrices is proved in a similar way as in Theorem 6.

These results not only hold for anisotropy occurring in an angle of 45°, but also for arbitrary rational angles $\frac{k}{k+l}90°$. In this more general case, new coordinates are defined with

$$s := kx + ly \quad \text{and} \quad t := lx - ky. \tag{36}$$

A multilevel method which combines semicoarsening and full coarsening similar to the one from Section 3 is defined by carrying out all coarsening steps, including the full coarsening steps, in the rotated coordinates. In each full coarsening step, the matrix B_F corresponds to the prolongation function $b(s,t)$. The construction of the elementary projection matrix is illustrated in the right part of Fig. 4. Within each block we pick every second row and every second column. On the block level, we pick two rows, eliminate the next two, pick another two rows and so on. The solid lines mark the blocks which are retained on the coarser levels, whereas the other blocks are eliminated. The dashed line explains why precisely these blocks have to chosen. Elimination on the block level must be done such that within the dashed line every second element is retained and the other elements are eliminated. This is described in more detail in [6].

For two-level circulant matrices $C_{mn}[f]$, a similar coarsening strategy can be applied which preserves the BCCB structure. Furthermore, if $m = n$, the blocks are all of size n. Again, the left part of Fig. 4 illustrates the construction of the permutation vector

$$\begin{aligned}&\big(1, 2n, 3n-1, \ldots, n^2-(n-2), 2, n+1, 3n, 4n-1, \ldots, n^2-(n-3),\\ &\quad \ldots, n, 2n-1, \ldots, n^2-(n-1)\big).\end{aligned} \tag{37}$$

The first block is obtained from the first and the $(n+1)$th diagonal, the second block from the second and the $(n+2)$th diagonal, and so on. The last block is equal to the nth diagonal. Since the permuted matrix has BCCB structure with blocks of equal size, we can almost directly apply the semicoarsening method from Section 3. Whereas for BTTB

matrices the exact structure is lost on coarser levels and the methods may lose efficiency, for BCCB matrices the methods are as efficient as in Section 3.

4.2. The use of line smoothers

For problems with anisotropy along coordinate axes we have used standard coarsening in combination with a line smoother to define a different multigrid algorithm. The same can be done with the systems from Example 10 if we take into account that the functions $f(x, y)$ and $g(x, y)$ have multiple zeros in $[0, 2\pi[^2$. Again, we have to carry out the permutation of rows and columns only for the smoothing process. For prolongation and restriction as well as for the computation of the system matrices on coarser grids we use the original coordinates x and y. Since we wish to apply a line smoother, the rows and columns of $A_{nn}[f]$ are permuted with the vector from (34) if $A_{nn}[f]$ is Toeplitz ot tau. Circulant matrices are permuted with a circulant version of (34). Then we partition the permuted matrix into blocks in the same way as it was done for the semicoarsening method. Each line in the right picture of Fig. 4 corresponds to one block of the matrix. The following theorem is a generalization of Theorem 7 which covers anisotropies occurring in arbitrary rational angles.

Theorem 11. *Let $f(x, y)$ be a nonnegative generating function with a zero of order 2 at the origin which in the rotated coordinates is of the form*

$$\tilde{f}(s,t) = \big[\big(1-\cos(s)\big) + \alpha\big(1-\cos(t)\big)\big]h(s,t) \tag{38}$$

with the trigonometric polynomial $h(s, t) > 0$ and $0 < \alpha \ll 1$. Let $A = A_{nn}[f]$ be the two-level tau or circulant matrix corresponding to $f(x, y)$, and $\tilde{A}$ its permuted version corresponding to $\tilde{f}(s, t)$. Let D be the matrix corresponding to $g(x, y)$, which in the rotated coordinates is of the form

$$\tilde{g}(s,t) = \big(1+\alpha-\cos(s)\big)\tilde{h}(s), \tag{39}$$

where $\tilde{h}$ is obtained by eliminating all terms in h containing t. $\tilde{D}$, the permuted version of D, is a block diagonal matrix with the same diagonal blocks as $\tilde{A}$. If $\tilde{h}(s) > 0$, then the block Jacobi method satisfies the smoothing conditions (8) *and* (9)*.*

Proof. After replacing the functions $f(x, y)$ and $g(x, y)$ by $\tilde{f}(s, t)$ and $\tilde{g}(s, t)$ the calculations are the same as in the proof of Theorem 7. As in the proof of Theorem 7, the matrix Y is chosen to be D^{-1}. □

Again, the proof does not include BTTB matrices for the technical problems mentioned in Remark 9. However, the numerical results of the multigrid solution of anisotropic BTTB systems with the block Jacobi smoother are quite promising (see Section 4.3).

Computation of the coarse grid matrix A_C cannot be done in such a straightforward way as in Section 3.2, because f has an additional zero at (π, π). Therefore we carry over an idea of Huckle and Staudacher [8] to the two-dimensional case and to the tau and circulant algebras. They interpret a Toeplitz matrix as a block Toeplitz matrix with blocks of size 2, which are themselves not necessarily Toeplitz. The BTTB matrix $T_{nn}[f]$ of size n^2-by-n^2 is considered to be a block BTTB matrix with blocks of size 4. Thus, the generating function becomes a 4-by-4 matrix $F(x, y)$, whose entries are functions in x and y. The eigenvalues of $F(x, y)$ in Example 10 only become zero at $(0, 0)$. Then $B(x, y)$ is for example chosen to be the 4-by-4 diagonal matrix with $b(x, y) = (1 + \cos(x))(1 + \cos(y))$ in each position of the diagonal, taking care of the zero in $F(x, y)$. The coarse grid matrix is computed by picking every second 2-by-2 block on both levels.

4.3. Numerical results

Again, we wish to test both types of multilevel methods numerically. Let us start with a multigrid method which is constructed as a suitable combination of semicoarsening steps followed by some full coarsening steps. The prolongation/restriction matrices and the elementary projection matrices are defined as described above, the change of coordinates, of course, has to be done only before the first step. Again, we use the same heuristic as in Section 3.

Table 6
Iteration numbers for $T_{mn}[f]$ with f from Example 10 with $\alpha = 0.002$ and $\alpha = 0.0001$

α	Coarsening	$n = 2^6 - 1$	$n = 2^7 - 1$	$n = 2^8 - 1$
0.002	t,t,st,st	10	12	13
0.002	t,t,t,t	6	6	6
0.0001	t,t,st,st	24	51	77
0.0001	t,t,t,t	6	6	6

Table 7
Iteration numbers for the block Jacobi smoother and $T_{nn}[f]$ with f from Example 10

α	$n = 2(2^5 - 1)$	$n = 2(2^6 - 1)$	$n = 2(2^7 - 1)$
0.1	9	8	8
0.001	3	6	11
0.00001	2	2	2

Table 8
Iteration numbers for $\tilde{C}_{mn}[f]$ with f from Example 10 with $\alpha = 0.002$ and $\alpha = 0.0001$

α	Coarsening	$n = 2^6 - 1$	$n = 2^7 - 1$	$n = 2^8 - 1$
0.02	t,st,st,st,st	23	24	24
0.02	t,t,t,st,st	5	5	5
0.02	t,t,t,t,t	8	7	7
0.001	t,st,st,st,st	>200	>200	>200
0.001	t,t,t,st,st	27	26	26
0.001	t,t,t,t,t	5	5	5

Our theoretical results state that the ratio r_F is reduced by a factor 2 in each semicoarsening step. Therefore, semicoarsening steps are applied until level curves are close to circles, i.e. until r_F is almost 1. Then we continue with full coarsening. We wish to test our multilevel method with the function $f(x, y)$ from Example 10, where α takes the values 0.002. The corresponding matrices $T_{nn}[f]$ belong both to the two-level Toeplitz class and to the two-level tau algebra. We use a five-grid method, where one step of symmetric Gauss–Seidel is used as pre- and postsmoother. Our theory suggests to use four semicoarsening steps, because $r_F = 22.36$. Table 6 shows the number of V-cycle iterations is significantly lower if four semicoarsening steps are used instead of only two. For a strongly anisotropic problem such as the same function f with $\alpha = 0.0001$ and $r_F = 100$, two semicoarsening steps followed by full coarsening do not lead to satisfactory convergence at all. If only semicoarsening is used, we observe the same rapid convergence.

The second type of multigrid method uses line smoothers for relaxation and the block strategy mentioned above for full coarsening. Again, this is too expensive if the diagonal blocks are full. However, if the matrix is sparse, this type of method is a good alternative to the multigrid algorithm with semicoarsening. As we have observed for the systems in Section 3, this method obtains its best results if the anisotropy is either moderate or very strong. If α is somewhere between 0.05 and 0.005, the method based on semicoarsening is preferable. The following iteration numbers we obtained with a three-level method, where two steps of block Jacobi were used as pre- and postsmoother.

Let us finally consider a two-level circulant example. The matrix $\tilde{C}_{mn}[f]$ with f from Example 10 is obtained from $C_{mn}[f]$ by adding $1/(mn)^2$. As we have seen in Section 4.1, these BCCB examples can be treated almost like the BCCB systems where anisotropy occurs along coordinate axes. The semicoarsening steps are exactly the same as in Section 3. For full coarsening we have to use the block interpretation of [8], which is equally valid for BCCB matrices. Table 8 shows the number of V-cycle iterations for different numbers of semicoarsening steps. As we expect from our theory, best results for $\alpha = 0.02$ are obtained with three semicoarsening steps and for $\alpha = 0.001$ with five semicoarsening steps.

After applying the permutation vector (37) to $\tilde{C}_{mn}[f]$ we can also define a multilevel method using standard prolongation and a line smoother. The convergence behavior is the same as for the circulant example, where the results were shown in Table 5. However for circulant matrices, inversion of the diagonal blocks is to expensive in most cases,

since the whole system can be solved in $O(n \log n)$ with the FFT. Therefore we suggest to use the multilevel method which is based on semicoarsening for BCCB matrices.

5. Conclusions

This article was devoted to the analysis of anisotropic problems in the context of structured linear systems, generating functions, and their level curves. This point of view allows the development of multilevel methods also for systems where anisotropy occurs in arbitrary directions and not only along coordinate axes. These more complicated systems are solved with the same efficiency.

Several applications where this type of anisotropic systems needs to be solved will be subject of future research. The most interesting of them seems to be the solution of systems corresponding to functions with a whole zero curve such as

$$f(x, y) = \big(\rho - \cos(x) - \cos(y)\big)^2 \quad (\rho < 2),$$

which arise when Helmholtz equations are solved. The aim is to construct a multilevel preconditioner where the building blocks are anisotropic systems, which approximate the zero curve at several of its points.

Moreover, a proof of V-cycle optimality for a multigrid method with semicoarsening is an interesting topic for future work. At least for anisotropy along coordinate axes an extension of the results in [2,1] to the anisotropic case should be possible.

References

[1] A. Arico, M. Donatelli, A V-cycle multigrid for multilevel matrix algebras: proof of optimality, Numer. Math., in press.
[2] A. Arico, M. Donatelli, S. Serra-Capizzano, V-cycle optimal convergence for certain (multilevel) structured linear systems, SIAM J. Matrix Anal. Appl. 26 (2004) 186–214.
[3] R. Chan, Q. Chang, H. Sun, Multigrid methods for ill-conditioned Toeplitz systems, SIAM J. Sci. Comput. 19 (1998) 516–529.
[4] R. Chan, M. Ng, Conjugate gradient methods for Toeplitz systems, SIAM Rev. 38 (1996) 427–482.
[5] G. Fiorentino, S. Serra, Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions, SIAM J. Sci. Comput. 17 (1996) 1068–1081.
[6] R. Fischer, T. Huckle, Multigrid methods for anisotropic BTTB systems, Linear Algebra Appl. 417 (2–3) (2006) 314–334.
[7] U. Grenander, G. Szegö, Toeplitz Forms and Their Applications, second ed., Chelsea Publishing, New York, 1984.
[8] T. Huckle, J. Staudacher, Multigrid preconditioning and Toeplitz matrices, ETNA 13 (2002) 81–105.
[9] J.W. Ruge, K. Stüben, Algebraic multigrid, in: S. McCormick (Ed.), Frontiers in Applied Mathematics: Multigrid Methods, SIAM, Philadelphia, 1987, pp. 73–130.
[10] S. Serra, Preconditioning strategies for asymptotically ill-conditioned block Toeplitz systems, BIT 34 (1994) 579–594.
[11] S. Serra-Capizzano, Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences, Numer. Math. 92 (2002) 433–465.
[12] S. Serra-Capizzano, C. Tablino-Possio, Multigrid methods for multilevel circulant matrices, SIAM J. Sci. Comput. 26 (2004) 55–85.
[13] H. Sun, X. Jin, Q. Chang, Convergence of the multigrid method for ill-conditioned block Toeplitz systems, BIT 41 (2001) 179–190.
[14] U. Trottenberg, C. Oosterlee, K. Schüller, Multigrid, Academic Press, London, 2001.
[15] E. Tyrtyshnikov, Circulant preconditioners with unbounded inverses, Linear Algebra Appl. 216 (1995) 1–23.