Electronic Transactions on Numerical Analysis.
Volume 45, pp. 146–159, 2016.

ISSN 1068–9613.

ETNA
Kent State University
http://etna.math.kent.edu

ON AMG METHODS WITH F-SMOOTHING BASED ON CHEBYSHEV POLYNOMIALS AND THEIR RELATION TO AMGr*

FLORIAN GOSSLER† AND REINHARD NABBEN†

Abstract. MacLachlan, Manteuffel, and McCormick [Numer. Linear Algebra Appl., 13 (2006), pp. 599–620] introduced a new algebraic multigrid method, the so-called reduction-based algebraic multigrid method (AMGr). Different from typical multigrid methods, the smoother of the AMGr method is acting only on the fine-grid points. To analyze the AMGr method, different constants and parameters are used. Here, we further analyze the AMGr method. We show that the parameter used by MacLachlan et al. has another important property. We show that it is closely related to the root of a Chebyshev polynomial. This fact also explains the good performance of AMGr. By examining this relation with Chebyshev polynomials, we extend the concept of the AMGr method. We consider algebraic multigrid methods with fine-grid smoothers and AMG methods that are based on polynomial smoothing. We also establish bounds for the error propagation operator. The bound is minimal if Chebyshev polynomials are chosen. If more than one smoothing step is used, the error bound is smaller than the bound given for the AMGr method. For only one smoothing step, the polynomial-based AMG with Chebyshev polynomials coincides with the AMGr method. In this case, our convergence analysis gives some new explanation of the high performance of the AMGr method as well as the parameters used in the AMGr method.

Key words. AMG, AMGr, Chebyshev polynomials

AMS subject classifications. 65F10, 65F20

1. Introduction.

Nowadays multigrid and multilevel methods are the methods of choice to solve linear systems of the form

$$Ax = b \quad \text{with} \quad A \in \mathbb{R}^{n\times n}, \quad x, b \in \mathbb{R}^n,$$

obtained from a discretization of an elliptic partial differential equation (PDE). These methods can either be used as a solver or as a preconditioner for Krylov subspace methods. Multigrid methods have been known for a long time, dating back to the 1930s. Their potential was first exploited by Fedorenko and Bakhalov in the 1960s [2, 6, 7] and later by Brandt [4] and Hackbusch [9].

Following the traditional way of explaining multigrid methods, the success of multigrid depends on two factors. First, there is a basic iterative method, like the Jacobi or the Gauss-Seidel method, that has strong smoothing effects. By such an iterative method, components of errors which are highly oscillatory can be well reduced. This basic iterative method is therefore called the smoother. The second factor is the so-called coarse-grid correction, in which a projected subsystem is solved approximately by recursively applying the smoothing and coarse-grid correction process.

If multigrid is used as a solver, the iteration matrix (or error propagation matrix) for a so-called nonsymmetric multigrid cycle is given as

$$E = (I - CA)(I - SA)^\nu. \tag{1.1}$$

For a symmetric multigrid cycle we have

$$E_S = (I - S^T A)^\nu (I - CA)(I - SA)^\nu.$$

Here S is the smoother and $(I - CA)$ is the coarse-grid correction operator. The integer ν indicates the number of smoothing steps. To measure the speed of convergence, the spectral

*Received March 6, 2014. Accepted January 21, 2016. Published online on May 6, 2016. Recommended by A. Frommer.

†Institut für Mathematik, MA 3-3, Technische Universität Berlin, D-10623 Berlin, Germany, ({gossler, nabben}@math.tu-berlin.de).

ETNA
Kent State University
http://etna.math.kent.edu

radius of E or E_S or a norm of E or E_S is considered. Since 1960, multigrid methods and algebraic multigrid methods have been analyzed in many publications. Here we just refer to the books of Hackbusch [9], Wesseling [23], Trottenberg, Oosterlee and Schüller [18], and Vassilevski [21].

In [11], MacLachlan, Manteuffel, and McCormick introduced a new algebraic multigrid method called the reduction-based algebraic multigrid method (AMGr method). Different from typical multigrid methods, the smoother of the AMGr method is acting only on the fine-grid points. In [11], convergence bounds for the AMGr method applied to symmetric positive definite systems are established. Independently, these methods were also analyzed by Mense and Nabben in [14, 15]. There, the authors gave convergence and comparison results for these methods (called MAMLI and SMAMLI methods) applied to nonsymmetric M-matrices.

The fast convergence of the AMGr method is demonstrated nicely by numerical and theoretical results in [5, 11]. To analyze the AMGr method, different constants and parameters are used in [11]. Specifically, the smoothing operator uses a parameter. This parameter is chosen such that an upper bound for the A-norm of the error propagation matrix of the AMGr method is small. Here, we further analyze the AMGr method. We show, that the parameter of the smoother used in [11] has another important property. It is closely related to the root of a Chebyshev polynomial. This fact also explains the fast convergence of the AMGr method.

The operator $(I - SA)^\nu$ in (1.1) can be seen as ν steps of the Richardson iteration preconditioned with S. Using the polynomial $P(t) = (1-t)^\nu$, this operator is just $P(SA)$. Replacing this trivial polynomial by one with proven best approximation properties leads to the use of Chebyshev polynomials and Chebyshev iterations [8, 16, 20]. Chebyshev acceleration and polynomial-based methods are well-know techniques in numerical analysis. Among the many publications, we only mention two recent ones [3, 10]. Polynomially-enhanced smoothers have been applied for a long time; see, e.g., [1, 19].

Observing the AMGr relation with Chebyshev polynomials, we extend the concept of the AMGr method. We consider algebraic multigrid methods with fine-grid smoothers and introduce new types of AMGr methods that are based on Chebyshev polynomials. If more than one smoothing step is necessary, we show that it is better to use Chebyshev polynomials rather than one smoother several times. Using the roots of the Chebyshev polynomials, the Chebyshev smoothers can be written in a similar way as the original smoother in (1.1). Thus, this modification leads to no extra computations. However, we establish an upper bound for the A-norm of the error propagation operator using Chebyshev polynomials which is below the bound for the AMGr method. This theoretical result suggests the use of Chebyshev polynomials. Numerical results also support this result.

The paper is organized as follows. In Section 2 we describe the AMGr method. The polynomial-based AMGr (AMGrp) methods are introduced and analyzed in Section 3. In Section 4, Chebyshev polynomials are used for the polynomial AMGr method to get optimal error bounds. Section 5 contains some numerical computations.

2. Preliminaries.

2.1. Notations. In this paper we consider symmetric positive definite (spd) matrices and symmetric positive semidefinite (spsd) matrices. If for all $x \in \mathbb{R}^n$ and for two spd matrices $A, B \in \mathbb{R}^{n\times n}$ the condition

$$x^T Ax \geq x^T Bx,$$

holds, we write $A \succeq B$.

Further, we use $\lambda_{\max}(\cdot), \lambda_{\min}(\cdot), \rho(\cdot), \kappa_2(\cdot)$, and $\sigma(\cdot)$ for the largest eigenvalue, smallest eigenvalue, the spectral radius, the condition number, and the spectrum of an spd matrix, respectively.

ETNA
Kent State University
http://etna.math.kent.edu

The space $\mathbb{R}_{\leq\nu}[t]$ denotes the set of all polynomials of degree less than or equal to ν with real coefficients.

2.2. Chebyshev polynomials. In this section, we briefly describe properties of the real Chebyshev polynomials. Here we follow [16]; see also [17] and [22]. The Chebyshev polynomial of the first kind of degree ν is defined as

$$C_\nu(t) = \cos[\nu \cos^{-1}(t)] \qquad \text{for} \quad -1 \leq t \leq 1.$$

It can be shown that this is a polynomial with respect to t. In fact, with $C_o(t) = 1$ and $C_1(t) = t$, the Chebyshev polynomials satisfy the three-term recurrence relation

$$C_{\nu+1} = 2tC_\nu - C_{\nu-1}(t).$$

The normalized Chebyshev polynomials have a well-known minimizing property on a given interval $[a, b]$, $a > 0$ in $\mathbb{R}$.

THEOREM 2.1 ([16, Theorem 6.25]). *The normalized Chebyshev polynomial on* $[a, b]$, $a > 0$ *is given by*

$$T_\nu(t) = \frac{C_\nu\left(\frac{b+a-2t}{b-a}\right)}{C_\nu\left(\frac{b+a}{b-a}\right)}. \tag{2.1}$$

Moreover,

$$\max_{t\in[a,b]} |\, T_\nu(t)\,| = \min_{p\in\mathbb{R}^0_{\leq\nu}[t]} \max_{t\in[a,b]} |\, p(t)\,|,$$

where $\mathbb{R}^0_{\leq\nu}[t] := \{p \in \mathbb{R}_{\leq\nu}[t] \,|\, p(0) = 1\}$.

The maximum of $C_\nu(t)$ in $[-1, 1]$ is 1. With $\kappa = \frac{b}{a}$ we then have

$$\max_{t\in[a,b]} T_\nu(t) = 2\left[\left(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\right)^\nu + \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^\nu\right]^{-1}. \tag{2.2}$$

The roots of the normalized Chebyshev polynomial of degree ν are given by

$$c_i = \tfrac{1}{2}\,(b + a - t_i(b - a)) \qquad \text{for} \quad t_i = \cos\left(\tfrac{\pi}{2} \cdot \tfrac{2i-1}{\nu}\right), \quad i = 1, 2, \ldots, \nu.$$

Note that we have given only the definition and properties of the real Chebyshev polynomials we need in the following sections. For more details, we refer to [17] and [22].

2.3. Polynomial-based AMGs. The classical use of the smoother $(I - SA)^\nu$ in (1.1) involves ν steps of the Richardson iteration (see [20]) preconditioned with S. The Richardson iteration is a special stationary iterative method resulting from a splitting of the matrix A. Let S be an approximation of A^{-1}. Then we obtain the stationary iteration

$$x^{i+1} = (I - SA)x^{i+1} + Sb, \quad i = 1, 2, \ldots$$

with arbitrary $x^0 \in \mathbb{R}^n$.

The error propagation operator is then $I - SA$. To get fast convergence, the spectral radius of $I - SA$ should be small. After ν steps, we obtain $(I - SA)^\nu$ as the error propagation operator. This can be written as $P(SA)$ where $P(t) = (1 - t)^\nu$ is a polynomial of degree ν satisfying $P(0) = 1$. Replacing this trivial polynomial by another polynomial of degree ν leads to the so-called semi-iterative methods. Since the spectral radius of $I - SA$ should be

ETNA
Kent State University
http://etna.math.kent.edu

small, the best polynomial should satisfy some (discrete) minimization property. Here, the Chebyshev polynomials with their (continuous) minimization property are the first choice. The resulting method is called Chebyshev iteration [8, 16, 20].

Similar techniques can be used for the multigrid method. Replacing $P(t) = (1-t)^\nu$ in the smoother of the multigrid iteration operator by another polynomial P_ν of degree ν with $P_\nu(0) = 1$ leads to the polynomial-based multigrid methods. The error propagation operator has the form

$$(I - CA)P_\nu(SA) \tag{2.3}$$

and, in the symmetric cycle,

$$P_\nu(S^T A)(I - CA)P_\nu(SA), \tag{2.4}$$

where $I - CA$ is the coarse-grid correction operator and S is the smoother.

Using the roots of the polynomial P_ν, we obtain the following representation of P_ν, which we will use later on.

PROPOSITION 2.2. *Let P_ν be a real polynomial of degree ν with $P_\nu(0) = 1$. Let $r_1, \ldots, r_\nu$ be the roots of P_ν. Then, P_ν can be factorized as*

$$P_\nu(t) = \left(1 - \frac{1}{r_1}t\right)\left(1 - \frac{1}{r_2}t\right)\ldots\left(1 - \frac{1}{r_\nu}t\right). \tag{2.5}$$

Proof. Obviously the polynomial of the right-hand side in (2.5) is of degree ν, has roots $r_1, \ldots, r_\nu$, and equals one for $t = 0$. Hence it must be P_ν. ∎

With Proposition 2.2 the polynomial-based multigrid error propagation operator in (2.3) can be written as

$$(I - CA)P_\nu(SA) = (I - CA)(I - \omega_1 SA)(I - \omega_2 SA)\ldots(I - \omega_\nu SA),$$

where $\omega_i = \frac{1}{r_i}$, for $i = 1, \ldots, \nu$. Similarly, the symmetric version in (2.4) can be represented. Hence, the difference between the classical smoothing operator $(I - SA)^\nu$ and $P_\nu(SA)$ is only the use of different weights ω_i in the factors. Therefore, there is no extra computational cost for using the polynomial version as long as the roots are known.

2.4. Reduction-based AMG (AMGr). Here we consider a symmetric positive definite $n \times n$ matrix A partitioned into a 2×2 block form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}. \tag{2.6}$$

We assume that the partition of $\mathbb{R}^n$ in two sets, fine- and coarse-grid points, is given so that $A_{11} \in \mathbb{R}^{n_1 \times n_1}$, $A_{22} \in \mathbb{R}^{n_2 \times n_2}$ and $n_1 + n_2 = n$. The AMGr method presented by MacLachlan, Manteuffel, and McCormick in [11] is motivated from a reduction point of view. Using only an F-smoother (a smoother that operates only on $\mathbb{R}^{n_1}$), it can be seen that $P_{\text{opt}} := \begin{bmatrix} -A_{11}^{-1}A_{12} \\ I \end{bmatrix}$ is the optimal interpolation matrix; see [5, 11, 13].

In general, computing A_{11}^{-1} is expensive, and often A_{11}^{-1} is not sparse, so an easy invertible and sparse approximation $D \in \mathbb{R}^{n_1 \times n_1}$ of A_{11} is used instead. The resulting interpolation is then defined by

$$\mathcal{P} = \begin{bmatrix} -D^{-1}A_{12} \\ I \end{bmatrix},$$

ETNA
Kent State University
http://etna.math.kent.edu

which results in the AMGr method with error propagation matrix

$$E_{\mathrm{amg}_r} := I - B^{-1}_{\mathrm{amg}_r} A := (I - CA)\left(I - \sigma \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} A\right), \tag{2.7}$$

where

$$C := \mathcal{P}(\mathcal{P}^T A \mathcal{P})^{-1} \mathcal{P}^T.$$

In [11], D has to satisfy two conditions:

1. $\sigma(D^{-1}A_{11}) \subseteq [1, 1+\varepsilon]$, and
2. $\begin{bmatrix} D & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ has to be symmetric positive semidefinite,

where $\varepsilon > 0$ is a given parameter. The constant σ is determined by $\sigma = \frac{2}{2+\varepsilon}$, motivated by the following inequality:

$$\left\| \left[I - \sigma D^{-1} A_{11} \right] \right\|_{A_{11}} = \rho\left(I - \sigma A_{11}^{\frac{1}{2}} D^{-1} A_{11}^{\frac{1}{2}} \right) \leq \max\left(|\, 1 - \sigma \,|, |\, (1+\varepsilon)\sigma - 1 \,| \right).$$

The last term is minimal if

$$|\, 1 - \sigma \,| = 1 - \sigma = (1+\varepsilon)\sigma - 1 = |\, (1+\varepsilon)\sigma - 1 \,|,$$

which holds for $\sigma = \frac{2}{2+\varepsilon}$. The following convergence result is proven in [11].

THEOREM 2.3. *Let A be spd and partitioned as in (2.6). Further, let D be an approximation of A_{11} with $\frac{1}{1+\varepsilon} A_{11} \preceq D \preceq A_{11}$ such that $\begin{bmatrix} D & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ is spsd. Then, the error propagation matrix $E_{\mathrm{amg}_r} := I - B^{-1}_{\mathrm{amg}_r} A$ in (2.7) with $\sigma = \frac{2}{2+\epsilon}$ satisfies*

$$\left\| I - B^{-1}_{\mathrm{amg}_r} A \right\|_A \leq \left(\frac{\varepsilon}{1+\varepsilon} \left(1 + \frac{\varepsilon}{(2+\varepsilon)^2} \right) \right)^{\frac{1}{2}} < 1.$$

The above result can be generalized to an AMGr method that uses more than one smoothing step. The following bound is given in [11].

THEOREM 2.4. *Let A be spd and partitioned as in (2.6). Further, let D be an approximation of A_{11} with $\frac{1}{1+\varepsilon} A_{11} \preceq D \preceq A_{11}$ so that $\begin{bmatrix} D & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ is spsd. Then,*

$$\left\| I - B^{-1}_{\mathrm{amg}_{r,\nu}} A \right\|_A \leq \left(\frac{\varepsilon}{1+\varepsilon} \left(1 + \left(\frac{\varepsilon}{2+\varepsilon} \right)^{2(\nu-1)} \left(\frac{\varepsilon}{(2+\varepsilon)^2} \right) \right) \right)^{\frac{1}{2}} < 1,$$

where $B_{\mathrm{amg}_{r,\nu}}$ is defined by

$$I - B^{-1}_{\mathrm{amg}_{r,\nu}} A = (I - CA) \cdot \left(I - \sigma \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} A \right)^{\nu}.$$

One of the aims of this paper is to understand the choice of the weight σ in the AMGr method. For one smoothing step, we will prove that the value $\sigma = \frac{2}{2+\epsilon}$ in Theorem 2.3 is optimal. However, if more than one smoothing step is used, this optimality is lost.

Finally, we remark that the use of two different approximations is possible; see [5]. Here, we use $D_p \approx A_{11}$ in $\mathcal{P}$ and $M_s \approx A_{11}$ in the smoother. We shall see that the choice of σ depends only on the approximation in the smoother as well as on the number of relaxation (or smoothing) steps.

ETNA
Kent State University
http://etna.math.kent.edu

3. Polynomial-based AMGr (AMG$_{\mathrm{rp}}$). In this section we introduce a class of AMGr methods using a polynomial-based relaxation. We are able to prove an error bound for these methods. Later on, we will establish a relation between the polynomial-based AMG and the reduction-based AMG introduced in [11].

3.1. Definition and first characterizations. In the last section we described the AMGr method which uses a special weight σ for acceleration. In this section we consider a similar AMG method using more than one smoothing step. In contrast to employing the same smoother several times, we use different smoothers of the same type. Our smoothers differ only in the weights that are used. We will establish the best choice for every weight. The weights have a strong connection to a polynomial. Thus, we define the resulting AMGr as *polynomial-based AMGr*, (*AMGrp*) and the symmetrized one as *symmetrized polynomial-based AMGr*, (*AMGrps*). For a polynomial P_ν of degree ν with $P_\nu(0) = 1$, we define

$$I - B^{-1}_{\mathrm{amg}_{rp}} A := (I - CA) \cdot P_\nu(MA), \tag{3.1}$$

$$I - B^{-1}_{\mathrm{amg}_{rps}} A := P_\nu(M^T A) \cdot (I - CA) \cdot P_\nu(MA), \tag{3.2}$$

with

$$M = \begin{bmatrix} M_s^{-1} & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathcal{P} = \begin{bmatrix} -D_p^{-1} A_{12} \\ I \end{bmatrix}, \quad C = \mathcal{P}(\mathcal{P}^T A \mathcal{P})^{-1} \mathcal{P}^T, \tag{3.3}$$

where D_p and M_s are nonsingular approximations of A_{11}.

Using Proposition 2.2, this polynomial AMGr can be written in the following form

$$I - B^{-1}_{\mathrm{amg}_{rp}} A = (I - CA) \cdot \left(I - \omega_1 \begin{bmatrix} M_s^{-1} & 0 \\ 0 & 0 \end{bmatrix} A \right) \ldots \left(I - \omega_\nu \begin{bmatrix} M_s^{-1} & 0 \\ 0 & 0 \end{bmatrix} A \right),$$

and the weights ω_i are the reciprocals of the roots of the polynomial P.

These polynomial-based AMGs use ν presmoothing steps and one coarse-grid correction, or in the symmetric cycle, ν presmoothing steps, one coarse-grid correction, and ν postsmoothing steps. Note that the factors in the smoother commute. For the special smoother M in (3.3) one can obtain a simpler expression of the polynomial $P_\nu(MA)$ in (3.1).

THEOREM 3.1. *Let A be spd and partitioned as in (2.6) and let $P_\nu \in \mathbb{R}_{\leq \nu}[t]$ be a polynomial of degree ν with $P_\nu(0) = 1$. Moreover, let $M_s \in \mathbb{R}^{n_1 \times n_1}$ be spd. Let M be defined as in (3.3). Then, $P_\nu(MA)$ and $P_\nu(M^T A)$ satisfy*

$$P_\nu(MA) = I - \begin{bmatrix} \left[I - P_\nu(M_s^{-1} A_{11}) \right] A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} A, \tag{3.4}$$

$$P_\nu(M^T A) = I - \begin{bmatrix} \left[I - P_\nu(M_s^{-T} A_{11}) \right] A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} A. \tag{3.5}$$

If in addition $| P_\nu(t) | < 1$ for $t \in \sigma(M_s^{-1} A_{11})$, then $\left[I - P_\nu(M_s^{-1} A_{11}) \right] A_{11}^{-1}$ is spd.

Proof. Let Q be a polynomial with $Q(0) = 0$. Then, we can prove easily by induction for $j \geq 1$ that

$$\left(\begin{bmatrix} M_s^{-1} & 0 \\ 0 & 0 \end{bmatrix} A \right)^j = \begin{bmatrix} (M_s^{-1} A_{11})^j A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} A.$$

Hence, we obtain

$$Q\left(\begin{bmatrix} M_s^{-1} & 0 \\ 0 & 0 \end{bmatrix} A \right) = \begin{bmatrix} Q(M_s^{-1} A_{11}) A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} A.$$

ETNA
Kent State University
http://etna.math.kent.edu

Since $P_\nu(0) = 1$, we get

$$P_\nu\left(\begin{bmatrix} M_s^{-1} & 0 \\ 0 & 0 \end{bmatrix} A\right) - I = \begin{bmatrix} (P_\nu(M_s^{-1}A_{11}) - I)A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} A,$$

which gives the desired representation in (3.4) and (3.5).

Next we prove the last statement, i.e., $\left[I - P_\nu(M_s^{-1}A_{11})\right] A_{11}^{-1}$ is spd if $|\, P_\nu(t)\,| < 1$ for $t \in \sigma(M_s^{-1}A_{11})$. We consider

$$\left(I - P_\nu(M_s^{-1}A_{11})\right) A_{11}^{-1} = A_{11}^{-\frac{1}{2}} \left(I - P_\nu\left(A_{11}^{\frac{1}{2}} M_s^{-1} A_{11}^{\frac{1}{2}}\right)\right) A_{11}^{-\frac{1}{2}}.$$

Obviously, this matrix is symmetric. Since

$$|\, P_\nu(t)\,| < 1 \quad \text{on} \quad \sigma(M_s^{-1}A_{11}) = \sigma\left(A_{11}^{\frac{1}{2}} M_s^{-1} A_{11}^{\frac{1}{2}}\right),$$

the positive definiteness follows. □

3.2. Convergence of the polynomial-based AMGr. The convergence analysis of the polynomial-based AMGrp defined in (3.1) is given in the next theorem.

THEOREM 3.2. *Let A be spd and partitioned as in (2.6), let $D_p \in \mathbb{R}^{n_1 \times n_1}$ be an spd approximation of A_{11} with*

$$\alpha A_{11} \preceq D_p \preceq A_{11}$$

for some $\alpha \leq 1$ so that $\begin{bmatrix} D_p & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ is spsd. Further let $M_s \in \mathbb{R}^{n_1 \times n_1}$ be another approximation of A_{11}. Let $P_\nu \in \mathbb{R}_{\leq \nu}[t]$ be a polynomial of degree ν satisfying $P_\nu(0) = 1$ and $|\, P_\nu(t)\,| < 1$ for $t \in \sigma(M_s^{-1}A_{11})$. Assume that the largest singular value $\sigma_{\max}$ of $P_\nu(A_{11}^{\frac{1}{2}} M_s^{-1} A_{11}^{\frac{1}{2}})$ satisfies $\sigma_{\max} < 1$. Consider the polynomial-based AMGr defined in (3.1). Then,

$$\left\| I - B_{\mathrm{amg}_{rp}}^{-1} A \right\|_A \leq \left(1 - \alpha(1 - \sigma_{\max}^2)\right)^{\frac{1}{2}} < 1.$$

Proof. First, note that every $\overline{e} \in \mathbb{R}^n$ can be written as the A-orthogonal sum

$$\overline{e} = \eta \begin{bmatrix} -A_{11}^{-1}A_{12} \\ I \end{bmatrix} v + \mu \begin{bmatrix} I \\ 0 \end{bmatrix} w, \tag{3.6}$$

where $\|v\|_S = \|w\|_{A_{11}} = 1$. Here, $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$.

Consider next an A-normalized $e \in \mathbb{R}^n$ ($\|e\|_A = 1$) which implies that $\eta^2 + \mu^2 = 1$. With Theorem 3.1, we have

$$P_\nu(MA) = I - \begin{bmatrix} \left[I - P_\nu(M_s^{-1}A_{11})\right] A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} A.$$

Thus,

$$\begin{aligned} P_\nu(MA)e &= e - \mu \begin{bmatrix} \left[I - P_\nu(M_s^{-1}A_{11})\right] w \\ 0 \end{bmatrix} \\ &= \eta \begin{bmatrix} -A_{11}^{-1}A_{12} \\ I \end{bmatrix} v + \mu \begin{bmatrix} P_\nu(M_s^{-1}A_{11}) \\ 0 \end{bmatrix} w. \end{aligned}$$

ETNA
Kent State University
http://etna.math.kent.edu

We follow the ideas of [11] and estimate the second term with respect to the A-norm,

$$\left\| \begin{bmatrix} P_\nu(M_s^{-1}A_{11}) \\ 0 \end{bmatrix} w \right\|_A^2 = \left\| P_\nu(M_s^{-1}A_{11})w \right\|_{A_{11}}^2$$
$$= \max_{x \neq 0} \frac{x^T P_\nu \left(A_{11}^{\frac{1}{2}} M_s^{-T} A_{11}^{\frac{1}{2}} \right) P_\nu \left(A_{11}^{\frac{1}{2}} M_s^{-1} A_{11}^{\frac{1}{2}} \right) x}{x^T x}$$
$$= \sigma_{\max} \left(P_\nu \left(A_{11}^{\frac{1}{2}} M_s^{-1} A_{11}^{\frac{1}{2}} \right) \right)^2 = \sigma_{\max}^2.$$

Therefore, for any e of the form (3.6), there exists a $\widehat{w}$ with $\|\widehat{w}\|_{A_{11}} = 1$ and

$$P_\nu(MA)e = \eta \begin{bmatrix} -A_{11}^{-1}A_{12} \\ I \end{bmatrix} v + \widehat{\mu} \begin{bmatrix} I \\ 0 \end{bmatrix} \widehat{w}$$

with $|\,\widehat{\mu}\,| \leq |\,\mu\sigma_{\max}\,|$.

Next, let $\widehat{e} = P_\nu e$. We obtain

$$\|\mathrm{C}\,\widehat{e}\|_A = \min_u \left\| \widehat{e} - \begin{bmatrix} -D_p^{-1}A_{12} \\ I \end{bmatrix} u \right\|_A.$$

With $u = \eta\theta v$, minimizing over θ, and using that

$$\begin{bmatrix} -D_p^{-1}A_{12} \\ I \end{bmatrix} v = \begin{bmatrix} -A_{11}^{-1}A_{12} \\ I \end{bmatrix} v + \begin{bmatrix} -(D_p^{-1} - A_{11}^{-1})A_{12} \\ 0 \end{bmatrix} v,$$

we get

$$\|\mathrm{C}\,\widehat{e}\|_A^2 \leq \min_\theta \left\{ \eta^2(1-\theta)^2 + \left(\widehat{\mu} + \eta\theta \left\| (D_p^{-1} - A_{11}^{-1})A_{12}v \right\|_{A_{11}} \right)^2 \right\}.$$

Since $\begin{bmatrix} D_p & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ is spsd, we then have $A_{22} \succeq A_{21}D_p^{-1}A_{12}$. Furthermore, we assume that $\lambda_{\max}(D_p^{-1}A_{11}) \leq \frac{1}{\alpha}$. For convenience, we set $1 + \varepsilon = \frac{1}{\alpha}$ for $\varepsilon > 0$, which implies

$$\left\| (D_p^{-1} - A_{11}^{-1})A_{12}) \right\|_{A_{11}}^2 \leq \varepsilon.$$

This yields

$$\left\| (I - B_{amg_{rp}}^{-1}A)e \right\|_A^2 \leq \min_\theta \left\{ \eta^2(1-\theta)^2 + \left(\widehat{\mu} + \eta\theta\sqrt{\varepsilon} \right)^2 \right\}.$$

The zero of the derivative of this function is given by

$$\theta = \frac{\eta - \sqrt{\varepsilon}\widehat{\mu}}{(1+\varepsilon)\eta}.$$

ETNA
Kent State University
http://etna.math.kent.edu

Combining these two terms gives

$$\begin{aligned}\left\|(I - B_{amg_{rp}}^{-1}A)e\right\|_A^2 &\leq \eta^2 \left(\frac{(1+\varepsilon)\eta - \eta + \sqrt{\varepsilon}\widehat{\mu}}{(1+\varepsilon)\eta}\right)^2 + \left(\widehat{\mu} + \frac{\eta - \sqrt{\varepsilon}\widehat{\mu}}{1+\varepsilon}\sqrt{\varepsilon}\right)^2 \\ &= \frac{\varepsilon}{(1+\varepsilon)^2}\left(\left(\sqrt{\varepsilon}\eta + \widehat{\mu}\right)^2 + \left(\frac{\widehat{\mu}}{\sqrt{\varepsilon}} + \eta\right)^2\right) \\ &= \frac{\varepsilon}{(1+\varepsilon)^2}\left((1+\varepsilon)\eta^2 + (1+\varepsilon)\left(\frac{\widehat{\mu}}{\sqrt{\varepsilon}}\right)^2 + (1+\varepsilon)2\eta\frac{\widehat{\mu}}{\sqrt{\varepsilon}}\right) \\ &= \frac{\varepsilon}{1+\varepsilon}\left(\eta + \frac{\widehat{\mu}}{\sqrt{\varepsilon}}\right)^2.\end{aligned}$$

Next, we maximize the last term with respect to $\mu = \sqrt{1-\eta^2}$ and $\eta \in [0,1]$.

We obtain the maximum for $\eta = \sqrt{\frac{\varepsilon}{\sigma_{\max}^2+\varepsilon}}$ and $\mu = \sqrt{\frac{\sigma_{\max}^2}{\sigma_{\max}^2+\varepsilon}}$, remembering that $|\,\widehat{\mu}\,| \leq |\,\mu\sigma_{\max}\,|$. Thus,

$$\begin{aligned}\left\|(I - B_{amg_{rp}}^{-1}A)e\right\|_A^2 &\leq \frac{\varepsilon}{1+\varepsilon}\left(\sqrt{\frac{\varepsilon}{\sigma_{\max}^2+\varepsilon}} + \sqrt{\frac{\sigma_{\max}^2}{\varepsilon}\cdot\frac{\sigma_{\max}^2}{\sigma_{\max}^2+\varepsilon}}\right)^2 \\ &= \frac{\varepsilon}{1+\varepsilon}\left(\frac{\varepsilon + \sigma_{\max}^2}{\sqrt{\sigma_{\max}^2+\varepsilon}\cdot\sqrt{\varepsilon}}\right)^2 = \frac{1}{1+\varepsilon}\cdot\frac{1}{\sigma_{\max}^2+\varepsilon}\left(\varepsilon + \sigma_{\max}^2\right)^2 \\ &= \frac{\varepsilon}{1+\varepsilon}\left(1 + \frac{\sigma_{\max}^2}{\varepsilon}\right).\end{aligned}$$

Finally, we substitute $\varepsilon = \frac{1}{\alpha} - 1$ and obtain

$$\begin{aligned}\left\|(I - B_{amg_{rp}}^{-1}A)e\right\|_A^2 &\leq \frac{\frac{1}{\alpha}-1}{\frac{1}{\alpha}}\left(1 + \frac{\sigma_{\max}^2}{\frac{1}{\alpha}-1}\right) = (1-\alpha)\left(1 + \frac{\sigma_{\max}^2\alpha}{1-\alpha}\right) \\ &= 1 - \alpha + \sigma_{\max}^2\alpha = 1 - \alpha(1-\sigma_{\max}^2).\end{aligned}$$

This completes the proof. ∎

With Theorem 3.2 we obtain the convergence of the symmetrized AMGrps easily.

THEOREM 3.3. *Let A be spd and partitioned as in (2.6), and let $D_p \in \mathbb{R}^{n_1\times n_1}$ be an spd approximation of A_{11} with*

$$\alpha A_{11} \preceq D_p \preceq A_{11}$$

for $\alpha \leq 1$ so that $\begin{bmatrix} D_p & A_{12} \\ A_{21} & A_{22}\end{bmatrix}$ is spsd. Furthermore, let $M_s \in \mathbb{R}^{n_1\times n_1}$ be another approximation of A_{11}. Let $P_\nu \in \mathbb{R}_{\leq\nu}[t]$ be a polynomial of degree ν with ν satisfying $P_\nu(0) = 1$ and $|\,P_\nu(t)\,| < 1$ for $t \in \sigma(M_s^{-1}A_{11})$. Assume that the largest singular value $\sigma_{\max}$ of $P_\nu(A_{11}^{\frac{1}{2}}M_s^{-1}A_{11}^{\frac{1}{2}})$ satisfies $\sigma_{\max} < 1$. Consider the symmetrized polynomial-based AMGr defined in (3.2). Then,

$$\left\|I - B_{\mathrm{amg}_{rps}}^{-1}A\right\|_A \leq 1 - \alpha(1-\sigma_{\max}^2) < 1.$$

ETNA
Kent State University
http://etna.math.kent.edu

Proof. To prove the result we use two properties of the method. First, the coarse-grid correction $I-CA$ is a projection, thus $(I-CA)^2 = I-CA$. Second, the postsmoother $P_\nu(M^TA)$ is the adjoint of $P_\nu(MA)$ in the inner product $\langle x,y\rangle_A := x^TAy$, thus $P_\nu(M^TA)^* = P_\nu(M^TA)$. Therefore, we obtain

$$\left\|I - B^{-1}_{\mathrm{amg}_{rps}}A\right\|_A = \left\|I - B^{-1}_{\mathrm{amg}_{rp}}A\right\|_A^2,$$

and with Theorem 3.2, the upper bound follows. ∎

The resulting convergence bounds of Theorems 3.2 and 3.3 depend on the choice of the polynomial. The next section deals with the minimization of these upper bounds using polynomials which have a minimizing property.

4. Chebyshev polynomials and the polynomial-based AMGr. In Theorem 3.2, we did not choose a special polynomial. If the matrix M_s used in the smoother is spd, we have $\sigma(M_s^{-1}A_{11}) \subseteq [a,b]$. Thus, we obtain for the largest singular value

$$\begin{aligned}\sigma_{\max} &= \sigma_{\max}\left(P_\nu\left(A^{\frac12}M_s^{-1}A^{\frac12}\right)\right)\\ &= \max_k\left|\lambda_k\left(P_\nu\left(A^{\frac12}M_s^{-1}A^{\frac12}\right)\right)\right| \le \max_{t\in[a,b]} P_\nu(t).\end{aligned} \tag{4.1}$$

To minimize the right-hand side, we have to choose the polynomial with its smallest maximum value in the interval $[a,b]$, and this leads us to the Chebyshev polynomials with their minimization property. This choice is not only the optimal one but also helps to establish upper bounds that are independent of the singular values.

THEOREM 4.1. *Let A be spd and partitioned as in (2.6). Let $D_p \in \mathbb{R}^{n_1\times n_1}$ be spd such that $\alpha A_{11} \preceq D_p \preceq A_{11}$ and $\begin{bmatrix} D_p & A_{12}\\ A_{21} & A_{22}\end{bmatrix}$ is spsd. Further let $M_s \in \mathbb{R}^{n_1\times n_1}$ be spd with $\sigma(M_s^{-1}A_{11}) \subseteq [a,b]$ and $\kappa := \frac{b}{a}$. Let T_ν be given by (2.1). Consider the error propagation matrix of the polynomial-based AMG given by (3.1). Then,*

$$\left\|I - B^{-1}_{amg_{rp}}A\right\|_A \le \left(1-\alpha\left(\frac{1-\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2\nu}}{1+\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2\nu}}\right)^2\right)^{\frac12} < 1.$$

Proof. In Theorem 3.2, the upper bound depends on the maximum of the singular values of $A_{11}^{\frac12}M_s^{-1}A_{11}^{\frac12}$. For an spd matrix M_s with $\sigma(M_s^{-1}A_{11}) \subseteq [a,b]$, we obtain with (2.2)

$$\begin{aligned}\sigma_{\max} &= \sigma_{\max}\left(T_\nu\left(A^{\frac12}M_s^{-1}A^{\frac12}\right)\right) = \max_k\left|\lambda_k\left(T_\nu\left(A^{\frac12}M_s^{-1}A^{\frac12}\right)\right)\right|\\ &\le \max_{t\in[a,b]} T_\nu(t) = 2\left(\left(\tfrac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\right)^\nu + \left(\tfrac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^\nu\right)^{-1}.\end{aligned}$$

Hence

$$1-\sigma_{\max}^2 \ge \left(\frac{1-\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2\nu}}{1+\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^{2\nu}}\right)^2,$$

which leads to the desired result. ∎

ETNA
Kent State University
http://etna.math.kent.edu

Next, we compare our result with the upper bound of the AMGr method given by Theorem 2.3. First, we note that for $\nu = 1$, $a = 1$, and $b = 1 + \varepsilon$ (this is the setting for the AMGr method) the weight used in the smoothing operator using Chebyshev polynomials is given by

$$\omega_1 = \left(\tfrac{1}{2}\left(1 + \varepsilon + 1 - \cos(\tfrac{\pi}{2} \cdot \tfrac{2-1}{1})(1 + \varepsilon - 1)\right)\right)^{-1} = \frac{2}{2+\varepsilon}.$$

Thus, the AMGrp and AMGr methods use the same weight. Hence, they are identical. Next, we show that the upper bounds of the error propagation operator with $\alpha = \frac{1}{1+\varepsilon}$ are also the same. We use $\kappa = \frac{b}{a} = 1 + \varepsilon$

$$\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^2 = \frac{2+\varepsilon-2\sqrt{1+\varepsilon}}{2+\varepsilon+2\sqrt{1+\varepsilon}}.$$

Thus,

$$1 - \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^2 = \frac{4\sqrt{1+\varepsilon}}{2+\varepsilon+2\sqrt{1+\varepsilon}},$$

and

$$1 + \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^2 = \frac{4+2\varepsilon}{2+\varepsilon+2\sqrt{1+\varepsilon}}.$$

With these two equations, we obtain

$$\begin{aligned} 1 - \alpha\left(\frac{1 - \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^2}{1 + \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^2}\right)^2 &= 1 - \frac{1}{1+\varepsilon} \cdot \frac{4(1+\varepsilon)}{(2+\varepsilon)^2} = \frac{\varepsilon}{(2+\varepsilon)^2} \cdot (\varepsilon + 4) \\ &= \frac{\varepsilon}{(2+\varepsilon)^2} \cdot \frac{\varepsilon^2 + 5\varepsilon + 4}{1+\varepsilon} = \frac{\varepsilon}{(2+\varepsilon)^2} \cdot \frac{(2+\varepsilon)^2 + \varepsilon}{1+\varepsilon} \\ &= \frac{\varepsilon}{1+\varepsilon}\left(1 + \frac{\varepsilon}{(2+\varepsilon)^2}\right), \end{aligned}$$

which gives the upper bound of the AMGr method proved in [11].

5. Numerical comparisons. Here we illustrate the theoretical results of this paper with some numerical computations. Similarly, as done in [11], we consider Poisson's equation in two dimensions, $-\Delta u = f$. The performance of the AMGr method is already demonstrated in [11].

Here, we consider the resulting matrix $A \in \mathbb{R}^{n \times n}$ using homogeneous Dirichlet boundary conditions and linear finite elements on a regular mesh in $[0,1]^2$. To get a partition into fine- (F) and coarse-grid (C) points, we use the Greedy coarsening algorithm for symmetric problems [12]. The algorithm can be found in Appendix A. This coarser does not use weak and strong connections but a special diagonal dominance of A depending on a parameter $\theta \in (\frac{1}{2}, 1)$. By choosing $\theta \approx \frac{1}{2}$, we get fewer coarse-grid points at the cost of an increased condition number of $M_s^{-1}A_{11}$: if we use $M_s = \operatorname{diag}(A_{11})$, it can be shown that $\sigma(M_s^{-1}A_{11}) \subseteq [2 - \frac{1}{\theta}, \frac{1}{\theta}]$; see [12]. In our numerical experiments we choose $M_s = D_p = [d_{ij}]$ with $d_{ij} = 0$ for $i \neq j$ and $d_{ii} = \sum_{i \in F} a_{ij}$ for $i \in F$. Thus, $\begin{bmatrix} D_p & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ is spsd and $(2\theta - 1)A_{11} \preceq D_p \preceq A_{11}$; see [12, Corollary 2 and Theorem 5]. Notice that we consider only two-level methods and the coarse-grid system is solved exactly.

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.1
Comparison of AMGr and AMGp based on Chebyshev polynomials for the test problem $-\Delta u = f$.

grid points	θ		AMGr $\nu=1$	AMGr $\nu=2$	AMGr $\nu=4$	AMGr $\nu=6$	AMGrp $\nu=1$	AMGrp $\nu=2$	AMGrp $\nu=4$	AMGrp $\nu=6$
16×16	.55	$\rho(t)$	.82	.67	.55	.54	.82	.50	.58	.52
		$\rho(e)$	.66	.71	.60	.56	.66	.63	.54	.53
	.60	$\rho(t)$	.67	.44	.20	.09	.67	.29	.10	.09
		$\rho(e)$	.40	.24	.12	.10	.40	.17	.10	.09
	.65	$\rho(t)$	.54	.29	.09	.09	.54	.17	.11	.09
		$\rho(e)$	.40	.24	.12	.10	.40	.17	.10	.09
132×32	.55	$\rho(t)$	.82	.67	.56	.54	.82	.50	.58	.52
		$\rho(e)$	.66	.71	.60	.56	.66	.63	.54	.53
	.60	$\rho(t)$	.67	.44	.20	.09	.67	.29	.10	.09
		$\rho(e)$	.40	.24	.12	.10	.40	.17	.10	.09
	.65	$\rho(t)$	.54	.29	.09	.09	.54	.17	.10	.09
		$\rho(e)$	.40	.24	.12	.10	.40	.17	.10	.09
164×64	.55	$\rho(t)$	.82	.67	.56	.54	.82	.50	.59	.52
		$\rho(e)$	.66	.71	.60	.56	.66	.63	.54	.53
	.60	$\rho(t)$	.67	.44	.20	.09	.67	.29	.10	.09
		$\rho(e)$	.40	.24	.12	.10	.40	.17	.10	.09
	.65	$\rho(t)$	.54	.29	.09	.09	.54	.17	.10	.09
		$\rho(e)$	.40	.24	.12	.10	.40	.17	.10	.09

In Table 5.1 we consider the spectral radii of the AMGrp operators based on Chebyshev polynomials compared to the classical AMGr method. We consider two cases. First, denoted by $\rho(t)$, we list the spectral radii using the parameter as described above, i.e., we use

$$[a,b] = \left[2 - \frac{1}{\theta}, \frac{1}{\theta}\right], \quad \alpha = 2\theta - 1, \quad \varepsilon = \frac{1}{\theta}. \tag{5.1}$$

Second, denoted by $\rho(e)$, we use the exact interval $[a,b]$, thus a and b are the extreme eigenvalues of $M_s^{-1}A_{11}$. Moreover, ν is the number of smoothing steps in one AMG cycle.

For $\nu = 1$ we obtain that both methods are identical, thus the spectral radii are the same. For $\nu > 1$ and the estimates of a and b as in (5.1), it happens only in very few cases that the spectral radius of the AMGr operator is below the spectral radius of the AMGrp operator. This is no contradiction to the optimality of the Chebyshev polynomials. In this case we have a strict inequality in (4.1), namely

$$\sigma_{\max} = \sigma_{\max}\big(P_\nu(A^{\frac{1}{2}}M_s^{-1}A^{\frac{1}{2}})\big) = \max_k \left|\, \lambda_k(P_\nu(A^{\frac{1}{2}}M_s^{-1}A^{\frac{1}{2}}) \,\right| < \max_{t\in[a,b]} P_\nu(t).$$

For the exact values of a and b, the spectral radius of the AMGrp operator is always below the one of the AMGr operator.

Summarizing. the polynomial-based AMGrp is a good alternative to the AMGr method. The computational costs of the methods are the same.

6. Conclusion. We presented a further analysis of the AMGr method introduced in [11]. Our analysis lead to polynomial-based AMGr methods. We established an error bound for these methods. If more than one smoothing step is used, our bound for the error propagation

ETNA
Kent State University
http://etna.math.kent.edu

operator of the polynomial-based AMGr method is below those for the AMGr method. The polynomial-based AMGr methods can be implemented with the same computational cost as the AMGr method.

Acknowledgment. We would like to thank Scott MacLachlan and the referee for many helpful comments. Moreover, we thank Scott for sharing his Greedy Coarser with us.

Appendix A. Greedy coarsening algorithm.

Algorithm 1: Greedy coarsening.

Input: $A, \theta \in (\frac{1}{2}, 1)$
Output: $F, C, \theta_i, i = 1, 2 \ldots, n$
Set $U = \{1, 2, \ldots, n\}$, $F = C = \emptyset$
Compute $\theta_i = \frac{|\, a_{ii} \,|}{\sum_{j \in F} |\, a_{ij} \,|}$
for $i = 1, \ldots, n$ **do**
  if $\theta_i \geq \theta$ **then**
    $F = F \cup \{i\}, U = U \backslash \{i\}$
while $U \neq \emptyset$ **do**
  Find $j = \underset{i \in U}{\operatorname{argmin}} \{\theta_i\}$
  $C = C \cup \{j\}, U = U \backslash \{j\}$
  for $i \in U \cap \mathrm{Adj}(j)$ *with* $\mathrm{Adj}(j) := \{k : a_{jk} \neq 0\}$ **do**
    Update $\theta_i = \frac{|\, a_{ii} \,|}{\sum_{j \in F} |\, a_{ij} \,|}$ **if** $\theta_i \geq \theta$ **then**
      $F = F \cup \{i\}, U = U \backslash \{i\}$

REFERENCES

[1] M. ADAMS, M. BREZINA, J. HU, AND R. TUMINARO, *Parallel multigrid smoothing: polynomial versus Gauss-Seidel*, J. Comput. Phys., 188 (2003), pp. 593–610.
[2] N. S. BAKHVALOV, *On the convergence of a relaxation method under natural contraints on an elliptic operator*, Ž. Vyčisl. Mat. i Mat. Fiz., 6 (1966), pp. 861–883.
[3] M. BENZI AND V. KUHLEMANN, *Chebyshev acceleration of the GeneRank algorithm*, Electron. Trans. Numer. Anal., 40 (2013), pp. 311–320. http://etna.ricam.oeaw.ac.at/vol.40.2013/pp311-320.dir/pp311-320.pdf
[4] A. BRANDT, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.
[5] J. BRANNICK, A. FROMMER, K. KAHL, S. MACLACHLAN, AND L. ZIKATANOV, *Adaptive reduction-based multigrid for nearly singular and highly disordered physical systems*, Electron. Trans. Numer. Anal., 37 (2010), pp. 276–295. http://etna.ricam.oeaw.ac.at/vol.37.2010/pp276-295.dir/pp276-295.pdf
[6] R. P. FEDORENKO, *A relaxation method for solving elliptic difference equations*, USSR Comp. Math. Math. Phys., 1 (1962), pp. 1092–1096.
[7] ———, *On the speed of convergence of an iteration process*, Ž. Vyčisl. Mat. i Mat. Fiz., 4 (1964), pp. 559–564.
[8] G. GOLUB AND C. VAN LOAN, *Matrix Computations*, 4th ed., Johns Hopkins University Press, Baltimore, 2013.
[9] W. HACKBUSCH, *Multigrid Methods and Applications*, Springer, Berlin, 1985.
[10] J. KRAUS, P. VASSILEVSKI, AND L. ZIKATANOV, *Polynomial of best uniform approximation to $1/x$ and smoothing in two-level methods*, Comput. Methods Appl. Math., 12 (2012), pp. 448–468.
[11] S. MACLACHLAN, T. MANTEUFFEL, AND S. MCCORMICK, *Adaptive reduction-based AMG*, Numer. Linear Algebra Appl., 13 (2006), pp. 599–620.
[12] S. MACLACHLAN AND Y. SAAD, *A greedy strategy for coarse-grid selection*, SIAM J. Sci. Comput., 29 (2007), pp. 1825–1853.

ETNA
Kent State University
http://etna.math.kent.edu

[13] S. P. MACLACHLAN, *Improving Robustness in Multiscale Methods*, Ph.D. Thesis, Dept. of Applied Math, University of Colorado, Boulder, 2004.
[14] C. MENSE AND R. NABBEN, *On algebraic multi-level methods for non-symmetric systems—comparison results*, Linear Algebra Appl., 429 (2008), pp. 2567–2588.
[15] C. MENSE AND R. NABBEN, *On algebraic multilevel methods for non-symmetric systems—convergence results*, Electron. Trans. Numer. Anal., 30 (2008), pp. 323–345.
http://etna.ricam.oeaw.ac.at/vol.30.2008/pp323-345.dir/pp323-345.pdf
[16] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
[17] V. I. SMIRNOV AND N. A. LEBEDEV, *Functions of a Complex Variable: Constructive Theory*, MIT Press, Cambridge, 1968.
[18] U. TROTTENBERG, C. W. OOSTERLEE, AND A. SCHÜLLER, *Multigrid*, Academic Press, San Diego, 2001.
[19] P. VANĚK AND M. BREZINA, *Nearly optimal convergence result for multigrid with aggressive coarsening and polynomial smoothing*, Appl. Math., 58 (2013), pp. 369–388.
[20] R. S. VARGA, *Matrix Iterative Analysis*, Springer, Berlin, 2000.
[21] P. S. VASSILEVSKI, *Multilevel Block Factorization Preconditioners*, Springer, New York, 2008.
[22] J. L. WALSH, *Interpolation and Approximation by Rational Functions in the Complex Domain*, 4th ed., AMS, Providence, 1965.
[23] P. WESSELING, *An Introduction to Multigrid Methods*, Wiley, Chichester, 1992.