SIAM J. MATRIX ANAL. APPL.
Vol. 33, No. 4, pp. 1369–1394

SOLVING ILL-POSED LINEAR SYSTEMS WITH GMRES AND A SINGULAR PRECONDITIONER*

LARS ELDÉN† AND VALERIA SIMONCINI‡

Abstract. Almost singular linear systems arise in discrete ill-posed problems. Either because of the intrinsic structure of the problem or because of preconditioning, the spectrum of the coefficient matrix is often characterized by a sizable gap between a large group of numerically zero eigenvalues and the rest of the spectrum. Correspondingly, the right-hand side has leading eigencomponents associated with the eigenvalues away from zero. In this paper the effect of this setting in the convergence of the generalized minimal residual (GMRES) method is considered. It is shown that in the initial phase of the iterative algorithm, the residual components corresponding to the large eigenvalues are reduced in norm, and these can be monitored without extra computation. The analysis is supported by numerical experiments. In particular, ill-posed Cauchy problems for partial differential equations with variable coefficients are considered, where the preconditioner is a fast, low-rank solver for the corresponding problem with constant coefficients.

Key words. ill-posed, linear system, GMRES, singular preconditioner, nearly singular

AMS subject classifications. 65F10, 65F22, 65F08, 65M30

DOI. 10.1137/110832793

1. Introduction.

Large, sparse nonsymmetric and *singular* linear systems arise when certain partial differential equations (PDEs) are discretized. In [7] conditions are given for the convergence without breakdown of the generalized minimum residual algorithm (GMRES) [38] applied to singular problems. Since the appearance of [7] many papers have been devoted to the analysis and application of GMRES for exactly singular problems; see [25] for a rather extensive account of the relevant literature.

In this paper we are concerned with *almost singular* (or *numerically singular*) linear systems,

$$Ax = b, \tag{1.1}$$

where $A \in \mathbb{C}^{n\times n}$. Such systems occur in connection with ill-posed problems, and for some problems GMRES works well, while for others it performs badly; see, e.g., [29, Examples 5.3 and 5.1], respectively. Recently it has been demonstrated that GMRES gives a good approximate solution in few iterations for certain ill-posed problems for PDEs when a singular preconditioner is used [34, Part III]. However, so far a deeper analysis of the properties of GMRES applied to almost singular systems is lacking.

The purpose of the present paper is to analyze and explain the convergence behavior of GMRES for linear systems that are almost singular, i.e., the way they occur in ill-posed problems $Ax = b$, where the matrix A is a discretization of a compact operator [14]. In this case, A is extremely ill-conditioned, typically with a gradual decay of singular values and a cluster of singular values at zero. Because of this peculiarity, previous attempts toward the understanding of GMRES convergence have often focused on information associated with the singular value decomposition (SVD)

*Received by the editors May 3, 2011; accepted for publication (in revised form) September 12, 2012; published electronically December 13, 2012.
http://www.siam.org/journals/simax/33-4/83279.html
†Department of Mathematics, Linkping University, SE-58183 Linkping, Sweden (lars.elden@liu.se).
‡Dipartimento di Matematica, Universit di Bologna, Piazza di Porta S. Donato 5, I-40127 Bologna, Italy (valeria@dm.unibo.it).

of the matrix; see, e.g., [29, 20, 6]. Instead, in agreement with, e.g., [8, 9], we will rely on spectral information of the problem, with the Schur decomposition of the coefficient matrix as the core theoretical tool. In some cases, especially in connection with singular preconditioners, the matrix has a cluster of eigenvalues of magnitude $O(1)$ that is well separated from another cluster of eigenvalues of small magnitude. Correspondingly, the right-hand side has large and leading components onto the eigendirections associated with the cluster away from the origin. Assuming that the linear system (1.1) is a perturbation of an exactly singular system of rank m, we will show the following:

(i) In the first iterations GMRES mainly reduces the norm of the residual as if solving the unperturbed system.

(ii) After at most m iterations, but often much earlier, the norm of the residual is of the order of magnitude of the perturbation, and if the GMRES procedure is then stopped, it gives a good approximation of the minimum norm solution of the exactly singular system.

Our theoretical findings generalize and are in agreement with the results discussed in [7, 25] for exactly singular systems. In particular, our analysis specifically explores the case when the condition for obtaining a minimum norm solution is not met, which is the setting usually encountered in ill-posed problems.

We will also consider the case when the eigenvalues are not clustered (when the numerical rank is ill-determined, which is often the case in ill-posed problems; see, e.g., the discussion in [2, 9]), and show theoretically and by examples that GMRES will give a good approximate solution if the iterations are stopped when the residual is of the order of the perturbation.

Numerically singular systems with clustered eigenvalues occur when singular preconditioners are applied to discrete ill-posed linear systems $Ax = b$ [34, Part III]. For such a problem, arising from the discretization of a linear equation with a compact operator, the ill-posedness manifests itself in the blow-up of high frequency components in the numerical solution. In order for the problem to be approximately solvable, the solution must be well represented in terms of the low frequency part of the operator. If the preconditioner M gives a good approximation of the low frequency part of the operator but suppresses the high frequency part completely, then the preconditioned problem $AM_m^\dagger y = b$ has the properties above.[1] Thus $AM_m^\dagger$ is numerically singular but with a well-conditioned low rank part. Computing the minimum norm solution of the preconditioned problem will yield a good approximation to the solution of the ill-posed problem. A similar strategy was explored in [5].

It is well known (see, e.g., [17, 31]) that unpreconditioned iterative methods applied to ill-posed problems exhibit *semiconvergence*: initially the approximate solution converges towards the "true solution," then it deteriorates and finally blows up. Such convergence behavior also occurs in the preconditioned case, and we give a theoretical explanation. However, in the case of singular preconditioners semiconvergence does not apply to the final solution approximation but only to an intermediate quantity. A stopping criterion based on the *discrepancy principle* will give a solution that is close to optimal.

The purpose of the paper is twofold:

- to give a theoretical foundation for the use of singular preconditioners for ill-posed problems;

[1] The notation $M_m^\dagger$ is explained at the end of the introduction.

- to demonstrate the effectiveness of singular preconditioners for GMR applied to Cauchy problems for parabolic and elliptic PDEs with variable coefficients.

The outline of the paper is as follows. In section 2 we introduce the GMRES algorithm and its properties for exactly singular systems. The concept of singular preconditioners is motivated in section 3. The Schur decomposition of the matrix is used in section 4 to analyze GMRES for nearly singular systems, and residual estimates are given. In section 5 we derive error estimates, which explain the regularizing properties of the method and the influence of the fact that the iterative solver is not pursued to convergence. Finally, in section 6 we give numerical examples in one to three dimensions.

We will use the following notation. The conjugate transpose of a matrix A is A^*. The Euclidean vector norm is denoted $\|x\| = (x^*x)^{1/2}$, and the induced matrix (operator) norm is $\|A\| = \max_{\|x\|=1} \|Ax\|$. The Frobenius norm is $\|A\|_F = (\sum_{i,j} |a_{ij}|^2)^{1/2}$. The singular values of $B \in \mathbb{C}^{p\times n}$, where $p \le n$, are denoted $\sigma_i,\ i = 1, 2, \ldots, p$, and are ordered as $\sigma_1 \ge \sigma_2 \ge \cdots \sigma_p \ge 0$; if $\sigma_p \ne 0$, its condition number is $\kappa_2(B) = \sigma_1/\sigma_p$. A^+ denotes the Moore-Penrose pseudoinverse of A. For a singular preconditioner M of rank m we will use $M_m^\dagger$ to denote a low-rank approximation of a matrix A^{-1}. Even if $M_m^\dagger$ may be a generalized inverse (not necessarily a Moore-Penrose pseudoinverse) of M, we are not particularly interested in that relation in this paper.

2. The GMRES algorithm for exactly singular systems. In this section we recall some known facts about the iterative solver GMRES and its convergence properties for singular systems that will be our background throughout the manuscript.

We start by defining the subspace under consideration: given a square matrix A and a vector r_0, a Krylov subspace of dimension k is defined as

$$K_k(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}.$$

In the context of solving (1.1), given a starting guess x_0 and the associated residual $r_0 = b - Ax_0$, GMRES determines an approximate solution x_k to (1.1) as $x_k \in x_0 + K_k(A, r_0)$ by requiring that the corresponding residual $r_k = b - Ax_k$ have minimum norm, namely

$$x_k = \arg \min_{x\in x_0+K_k(A,r_0)} \|b - Ax_k\|. \tag{2.1}$$

The algorithm is a popular implementation of a minimal residual method that fully exploits the properties of the approximation space. For a sound implementation of GMRES we refer the reader to [37, Chap. 6.5].

A key feature of the algorithm is the computation of an orthonormal basis $w_1, \ldots, w_k$ by the Arnoldi iterative method. After k iterations, this process can be conveniently summarized by the Arnoldi relation

$$AW_k = W_{k+1}H_k,$$

with $W_k = [w_1, \ldots, w_k]$, $W_k^*W_k = I_k$, and $H_k \in \mathbb{C}^{(k+1)\times k}$ upper Hessenberg.

The problem of solving a singular linear system $Ax = b$ using GMRES is treated in [7, 25], where the following result is proved.

PROPOSITION 2.1. *GMRES determines a least squares solution x_* of a singular system $Ax = b$, for all b and starting approximations x_0, without breakdown, if and only if $\mathcal{N}(\mathcal{A}) = \mathcal{N}(\mathcal{A}^*)$. Furthermore, if the system is consistent and $x_0 \in \mathcal{R}(\mathcal{A})$, then x_* is a minimum norm solution.*

Assume that the rank of A is equal to m. For the analysis it is no restriction to assume that the matrix of the linear system has the structure[2]

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} = \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix}, \qquad A_{11} \in \mathbb{C}^{m\times m}. \tag{2.2}$$

Throughout we will use the notational convention that $c^{(1)}$ is the upper part of the vector c, according to the splitting of the coefficient matrix, and analogously for other involved matrices. It is easy to see (cf. [25]) that the condition $\mathcal{N}(\mathcal{A}) = \mathcal{N}(\mathcal{A}^*)$ is equivalent to $A_{12} = 0$. Similarly, the consistency condition is equivalent to $c^{(2)} = 0$.

Obviously, applying GMRES to the linear system

$$\begin{bmatrix} A_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} = \begin{bmatrix} c^{(1)} \\ 0 \end{bmatrix} \tag{2.3}$$

is mathematically equivalent to applying GMRES to $A_{11}x^{(1)} = c^{(1)}$. Due to the finite termination property of Krylov methods it will never take more than m steps to obtain the solution of this problem (in exact arithmetic). Finally, in this section, the properties of the Krylov subspace ensure that applying GMRES to (2.2) with $c^{(2)} = 0$ and zero starting approximation is also mathematically equivalent to applying GMRES to $A_{11}x^{(1)} = c^{(1)}$. A more common situation occurs when the (2,2) block of (2.2) is *almost* zero, i.e., it has small but nonzero entries, and in addition $c^{(2)} \neq 0$. In this case, the rank of A is of course larger than m, and the role of A_{12} becomes more relevant. We analyze such a setting in section 4 for a general A by first performing a Schur decomposition.

3. Singular preconditioners for ill-posed problems. In this section we motivate the use of singular preconditioners, having in mind large and sparse ill-posed linear systems that occur, e.g., when inverse problems for PDEs are discretized.

Preconditioners are routinely used for solving linear systems $Ax = b$ using Krylov methods. For this discussion we first assume that the matrix A corresponds to a well-posed problem, by which we mean that its condition number is of moderate magnitude. With right preconditioning one derives and computes a nonsingular approximation M of A and then solves the equivalent linear system

$$AM^{-1}y = b, \qquad x = M^{-1}y,$$

using a Krylov subspace method. The reason why we use a right preconditioner is that we will apply the *discrepancy principle* [14, p. 83], [22, p. 179], which means that we are not interested in solving the linear system $Ax = b$ exactly, but only determine an approximation $\hat{x}$ with residual $\|A\hat{x} - b\| \approx \delta$, where δ is prespecified[3] and is a measure of the noise level of the data. In particular, the monotonicity of the (original system) residual norm provides the proper setting for which the discrepancy principle is most meaningful.

The matrix M may represent, e.g., the discretization of a related but simplified operator, or a structure-capturing matrix with an inexpensive-to-apply inverse. Now assume that the linear system of equations $Ax = b$ represents a discrete, ill-posed problem. The problem with a preconditioner as described above is that if M is a good approximation of A, then also M is ill-conditioned, with M^{-1} very large in norm. For concreteness, let M be a circulant matrix [20], written as

[2] In [25] a transformation of the system is done by decomposing the space $\mathbb{C}^n$ into $\mathcal{R}(\mathcal{A})$ and $\mathcal{R}(\mathcal{A})^\perp$.

[3] In some cases δ can be estimated from the data; see [27].

$$M = F\Lambda F^*, \tag{3.1}$$

where F is the Fourier matrix and Λ is a diagonal matrix of eigenvalues. In order to "regularize" the preconditioner, the small eigenvalues, corresponding to high frequencies, are replaced by ones; i.e., the preconditioner is chosen as

$$M_I = F \begin{bmatrix} \Lambda_1 & 0 \\ 0 & I \end{bmatrix} F^*, \tag{3.2}$$

which has an inverse with a norm that is not too large. This approach is investigated in several papers [20, 18, 19, 31, 32, 24]. In the current paper, motivated by the application to Cauchy problems for elliptic and parabolic PDEs in two space dimensions (see [34, Part III] and sections 6.3 and 6.4), we instead choose to use another type of regularized, *singular preconditioner*, defined using a low-rank approximation of the solution operator. If we were to use the analogue of this idea in the case of a circulant preconditioner, we would take

$$M_m^\dagger = F \begin{bmatrix} \Lambda_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} F^*. \tag{3.3}$$

Thus we solve the singular linear system $(AM_m^\dagger)y = b$, with the GMRES method, and then compute $x = M_m^\dagger y$. A somehow related approach was proposed in [2], where, however, the singular preconditioner was generated by means of a projection argument instead of a generalized inverse strategy.

We show in section 4 that the distribution of eigenvalues of $AM_m^\dagger$ determines the rate of convergence and the quality of the GMRES solution. In fact, the regularized singular preconditioner also induces regularization on the solution; see section 5.

4. The GMRES algorithm for nearly singular systems. Consider a preconditioned least squares problem

$$\min_y \|(AM_m^\dagger)y - b\|, \tag{4.1}$$

where, in exact arithmetic, $\operatorname{rank}(M_m^\dagger) = m$. For the purpose of analysis we will use the Schur decomposition $AM_m^\dagger = UBU^*$ [15, p. 313], where B is upper triangular with diagonal elements ordered by decreasing magnitude. By a change of variables we get the equivalent linear least squares problem $\min_d \|Bd - c\|$, with $c = U^*b$, which we partition as

$$\min_d \left\| \begin{bmatrix} L_1 & G \\ 0 & L_2 \end{bmatrix} \begin{bmatrix} d^{(1)} \\ d^{(2)} \end{bmatrix} - \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix} \right\|, \tag{4.2}$$

where $L_1 \in \mathbb{C}^{m\times m}$ is nonsingular. We emphasize that the use of the Schur decomposition in this context is due only to numerical convenience and to consistency with later computational experiments. Any decomposition that provides a 2×2 *block* upper triangular form by unitary transformation with the same spectral properties would yield the same setting. In particular, we shall not use the fact that both L_1, L_2 are upper triangular.

Since in many cases neither A nor $M_m^\dagger$ will be explicitly available, but only as operators acting on vectors, we cannot presuppose in our analysis that $L_2 = 0$. Instead we assume[4] that

[4]The meaning of the "much larger than" symbol will depend on the context: in the case of singular preconditioners it can be several orders of magnitude, while in the case when GMRES is applied directly to an ill-posed problem, it may be only two orders of magnitude; see the numerical examples.

$$|\lambda_{\min}(L_1)| \gg |\lambda_{\max}(L_2)|, \qquad \|c^{(1)}\| \gg \|c^{(2)}\| = \delta. \tag{4.3}$$

By $\lambda_{\min}(L_1)$ we mean the eigenvalue of smallest modulus. We also assume that L_1 is well conditioned; i.e., $\|L_1^{-1}\|$ is not large. The eigenvalue condition in (4.3) is related to the assumption that B is almost singular. Thus L_2 can be considered as a perturbation of zero, corresponding to either floating point round-off or some other type of "noise," and the same applies to $c^{(2)}$. We shall also assume that $\|G\|$ has a small or moderate value, excluding the occurrence of nonnormality influencing the two diagonal blocks.[5] The assumptions in (4.3) also exclude the case, for instance, where the given problem is a perturbation of a nonsymmetric matrix with all zero eigenvalues and a single eigenvector; cf., e.g., [29, sect. 5.1, Example]. The eigenvalues of such a perturbed matrix will tend to distribute in a small disk around the origin. This last assumption is not restrictive, since it is already known that GMRES will perform very badly in this setting; see, e.g., [30, Example R, p. 787].

Now, since (4.2) can be seen as a perturbation of

$$\min_d \left\| \begin{bmatrix} L_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} d^{(1)} \\ d^{(2)} \end{bmatrix} - \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix} \right\|, \tag{4.4}$$

we may ask whether it is possible to "solve" (4.2) as efficiently as we would do with (4.4) (cf. the discussion around (2.3)). We will show in section 4.1 that in the first few (fewer than m) steps of GMRES, a sufficiently small residual will be obtained, whose size depends on $\|L_2\|$, $\|G\|$, and $\|c^{(2)}\|$, as expected. Other quantities also enter the picture. In section 4.2 we will use the approximation properties of Krylov subspaces to derive more accurate bounds for the residual norm, which involve the spectral distance between L_1 and L_2. We also remark that the model derived by splitting the spectral domain into a "good" part and a "bad" part has been used; see, e.g., [3] and [2]. In both cited cases, however, the aim is to computationally exploit an approximate decomposition so as to accelerate the convergence of the employed method. Here the *exact* splitting is a theoretical device used to explain the practical behavior of GMRES in certain circumstances.

Example 4.1. In section 6.2 we consider an ill-posed "model problem": the Cauchy problem for a parabolic equation in one space dimension (referred to as Cauchy-1D), which is solved using right-preconditioned GMRES. The preconditioner is singular, which leads to an almost singular linear system, whose Schur decomposition has the structure (4.2)–(4.3). The relevant quantities are

$$|\lambda_{\min}(L_1)| = 0.6768, \quad |\lambda_{\max}(L_2)| = 2.4 \cdot 10^{-16},$$
$$\|G\| = 0.0962, \quad \|c^{(1)}\| = 0.6753, \quad \|c^{(2)}\| = 0.006573.$$

Clearly, the assumptions of Proposition 2.1 are not satisfied. Figure 4.1 shows that GMRES quickly reduces the relative residual norm to 10^{-2} and then stagnates. We will see later that the approximate solution after 4 steps is acceptable (for an ill-posed problem). In sections 4.1 and 4.2 we will show that in the first few steps of the Arnoldi recursion the L_1 block dominates, and, essentially, the well-conditioned system $L_1 d^{(1)} = c^{(1)}$ is solved, before the small L_2 block comes into play.

To proceed we need to introduce some notation and definitions. Under the eigenvalue assumption in (4.3) we can write

$$B = \begin{bmatrix} L_1 & G \\ 0 & L_2 \end{bmatrix} = X B_0 X^{-1} = [X_1, X_2] \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix} \begin{bmatrix} Y_1^* \\ Y_2^* \end{bmatrix}, \tag{4.5}$$

[5] More precisely, we assume that G is small in the sense that the matrix P in (4.7) is not large.

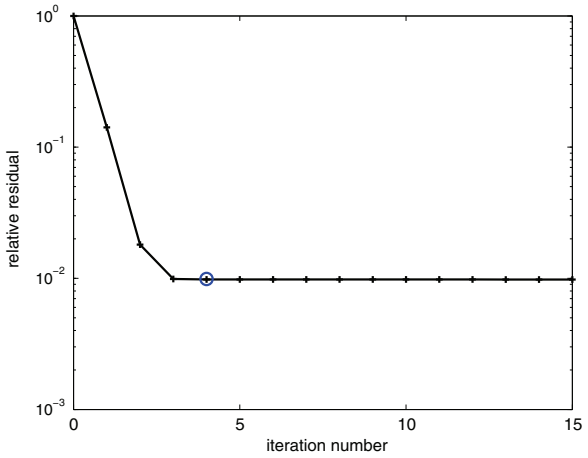

FIG. 4.1. *Example* 4.1 *(Cauchy-1D). Relative residual as a function of iteration index.*

where $[Y_1, Y_2]^* = [X_1, X_2]^{-1}$,

$$[X_1, X_2] = \begin{bmatrix} I & P \\ 0 & I \end{bmatrix}, \qquad [Y_1, Y_2] = \begin{bmatrix} I & 0 \\ -P^* & I \end{bmatrix}, \tag{4.6}$$

and P is the unique solution of the Sylvester equation $L_1 P - P L_2 = -G$. Note that

$$\|X_2\| \le 1 + \|P\|, \quad \|Y_1\| \le 1 + \|P\|, \quad \text{where} \quad \|P\| \le \frac{\|G\|}{\operatorname{sep}(L_1, L_2)}, \tag{4.7}$$

and $\operatorname{sep}(L_1, L_2)$ is the separation function.[6] It is known (cf., e.g., [39, Thm. V.2.3]) that $\operatorname{sep}(L_1, L_2) \le \min_{i,j} |\lambda_i(L_1) - \lambda_j(L_2)|$, where $\lambda_i(X)$ denotes the ith eigenvalue of X. It is also easy to show, using the definition of the matrix norm, that $\|X\| \le 1 + \|P\|$.

DEFINITION 4.2 (see [42, p. 36]). *The grade of a matrix L with respect to a vector v is the degree of the lowest degree monic polynomial p such that $p(L)v = 0$.*

The polynomial giving the grade is unique and is referred to in the literature as the minimum polynomial; see, e.g., [16, 28]. In this paper we shall adopt the term *grade polynomial* to avoid confusion with the minimum residual GMRES polynomial.

4.1. Estimating the residual. We start by establishing a relation between Arnoldi recursions for the block-triangular system $Bd = c$ and the block-diagonal system $B_0 d_0 = u$, where $u = X^{-1}c$. Assume that for any $k \ge 1$ (with, of course, $k < n$) we have generated a Krylov decomposition of B,

$$BW_k = W_{k+1} H_k, \qquad w_1 = \frac{\hat{w}}{\|\hat{w}\|}, \quad \hat{w} = \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix}, \tag{4.8}$$

with $H_k \in \mathbb{C}^{(k+1)\times k}$ upper Hessenberg. Using the relation $B = X B_0 X^{-1}$, we get $X B_0 X^{-1} W_k = W_{k+1} H_k$. Using the thin QR decompositions

$$X^{-1} W_i = V_i S_i, \quad i = k, k+1, \tag{4.9}$$

we obtain the Krylov decomposition of B_0,

[6] The sep function is defined as $\operatorname{sep}(L_1, L_2) = \inf_{\|P\|=1} \|T(P)\|$, where $T : P \mapsto L_1 P - P L_2$ (cf., e.g., [39, sect. V.2.1]).

$$B_0 V_k = V_{k+1}\left(S_{k+1} H_k S_k^{-1}\right), \qquad v_1 = \frac{u}{\|u\|}, \quad u = \begin{bmatrix} u^{(1)} \\ u^{(2)} \end{bmatrix} = \begin{bmatrix} c^{(1)} - Pc^{(2)} \\ c^{(2)} \end{bmatrix}, \tag{4.10}$$

where $S_{k+1} H_k S_k^{-1}$ is upper Hessenberg. Thus, the Arnoldi method applied to B with starting vector c uniquely defines another sequence of vectors, which can be generated by the Arnoldi method applied to B_0 with starting vector v_1 defined by (4.10).

We will now analyze GMRES for $Bd = c$ in terms of the equivalent recursion for $B_0 d_0 = u$. Denote the grade of L_1 with respect to $u^{(1)}$ by m_*. We will first show that the upper block of V_{m_*}, denoted by $V_{m_*}^{(1)} \in \mathbb{C}^{m \times m_*}$, has full column rank. Due to the structure of B_0, the Arnoldi method applied to the linear system $B_0 d = u$ generates a basis for the Krylov subspace

$$\mathcal{K}_{m_*}(B_0, u) = \operatorname{span}\left\{ \begin{bmatrix} u^{(1)} \\ u^{(2)} \end{bmatrix}, \begin{bmatrix} L_1 u^{(1)} \\ L_2 u^{(2)} \end{bmatrix}, \ldots, \begin{bmatrix} L_1^{m_*-1} u^{(1)} \\ L_2^{m_*-1} u^{(2)} \end{bmatrix} \right\}. \tag{4.11}$$

LEMMA 4.3. *Assume that the orthonormal columns of the matrix*

$$V_{m_*} = \begin{bmatrix} V_{m_*}^{(1)} \\ V_{m_*}^{(2)} \end{bmatrix} \in \mathbb{C}^{n \times m_*}$$

span the Krylov subspace (4.11). *Then the upper $m \times m_*$ block $V_{m_*}^{(1)}$ has full column rank. In addition, the overdetermined linear system $L_1 V_{m_*}^{(1)} z = u^{(1)}$ is consistent.*

Proof. Let $K^{(i)} = [u^{(i)}, L_i u^{(i)}, \ldots, L_i^{m_*-1} u^{(i)}]$, $i = 1, 2$, and

$$K = \begin{bmatrix} K^{(1)} \\ K^{(2)} \end{bmatrix}.$$

The columns of $K^{(1)}$ are linearly independent; otherwise the zero linear combination would imply the existence of a polynomial p of degree strictly less than m_* such that $p(L^{(1)})u^{(1)} = 0$, which is a contradiction to the definition of grade. Therefore, the matrix $K^*K = (K^{(1)})^* K^{(1)} + (K^{(2)})^* K^{(2)}$ is nonsingular, and the columns of $Q_{m_*} = K(K^*K)^{-\frac{1}{2}}$ are orthonormal with first block $Q_{m_*}^{(1)}$ having full column rank. Any other orthonormal basis spanning range(Q_{m_*}) differs from Q_{m_*} in a right multiplication by a unitary matrix, leaving the full rank property of the first block unchanged. The consistency follows from the definition of grade. ∎

The lemma shows that, since $V_k^{(1)}$ has full rank for $k \le m_*$, GMRES "works on reducing the significant part" of the residual of the linear system until m_* steps have been performed.

THEOREM 4.4. *Assume that m_* is the grade of L_1 with respect to $u^{(1)} = c^{(1)} - Pc^{(2)}$. If the projection matrix W_{m_*} is constructed using the Arnoldi method applied to the system $Bd = c$, with starting vector $w_1 = c/\|c\|$, then*

$$\begin{aligned} \|r_{m_*}\| &= \min_f \|BW_{m_*} f - c\| \\ &\le (1 + \|P\|)\left(\|L_2 V_{m_*}^{(2)}\| \, \|(L_1 V_{m_*}^{(1)})^+\| (1 + \|P\|)\|c^{(1)}\| + \|c^{(2)}\|\right), \end{aligned} \tag{4.12}$$

where W_{m_*} and V_{m_*} are related by (4.9), *and P is defined in* (4.6).

Proof. Using $B = XB_0X^{-1}$ and (4.9) we have, for any y,

$$\begin{aligned} \|BW_{m_*} f - c\|^2 &= \|XB_0X^{-1}W_{m_*} f - c\|^2 \le \|X\|^2 \, \|B_0 V_{m_*} S_{m_*} f - u\|^2 \\ &= \|X\|^2 \left(\|L_1 V_{m_*}^{(1)} z - u^{(1)}\|^2 + \|L_2 V_{m_*}^{(2)} z - c^{(2)}\|^2\right), \end{aligned} \tag{4.13}$$

where $z = S_{m_*} f$ and $u^{(2)} = c^{(2)}$. Since, by Lemma 4.3, the equation $L_1 V_{m_*}^{(1)} z = u^{(1)}$ is consistent, we can make the first term in (4.13) equal to zero by choosing $z = (L_1 V_{m_*}^{(1)})^+ u^{(1)}$. Thus, we have

$$\|r_{m_*}\| \le \|X\| \, \|L_2 V_{m_*}^{(2)} z - c^{(2)}\|.$$

The result now follows by using the triangle inequality, $\|X\| \le 1 + \|P\|$, and $\|u^{(1)}\| \le (1 + \|P\|)\|c^{(1)}\|$. ∎

The aim of Theorem 4.4 is to describe why in the first few steps of GMRES mainly the residual of the "significant part" of the linear system is reduced in norm; a few comments are in order. For $k \le m_*$, assume that $L_1 V_k^{(1)}$ is well-conditioned. Then, as in the above proof,

$$\begin{aligned} \|r_k\|^2 &= \min_y \|BW_k \, y - c\|^2 \\ &\le \|X\|^2 \left(\|L_1 V_k^{(1)} z - u^{(1)}\|^2 + \|L_2 V_k^{(2)} z - c^{(2)}\|^2 \right). \end{aligned} \tag{4.14}$$

Due to the assumption (4.3), the second term will be small for z of moderate norm. Therefore, since almost nothing can be done in reducing the second term, GMRES will give a solution that almost optimizes the first term; i.e., it will give a solution $\hat{y} \approx S_k^{-1} \hat{z}$, with $\hat{z} = (L_1 V_k^{(1)})^+ u^{(1)}$. After m_* steps the first part of the residual can be made equal to zero, and the norm of the overall residual is small due to the assumption (4.3). Furthermore, since the Arnoldi recursion for B_0 can be seen as a perturbation of that for

$$\begin{bmatrix} L_1 & 0 \\ 0 & 0 \end{bmatrix},$$

in the first steps of the recursion the matrix $V_k^{(1)}$ is close to orthogonal, and therefore the assumption that $L_1 V_k^{(1)}$ is well-conditioned is justified for small values of k.

We expect that for general problems the grade m_* will be close to m. However, in the case of preconditioning of ill-posed equations, L_1 may have extremely close or even multiple eigenvalues (depending on the quality of the preconditioner), so that the method will reduce $\|X\| \, \|L_1 V_k^{(1)} \hat{z} - u^{(1)}\|$ to a level below $\|c^{(2)}\|$ after only a few steps. This is illustrated in the following example.

Example 4.5. The example presented in section 6.2 has numerical rank 20 ($m = 20$), and we solve it using GMRES. For this example $\|X\| \le 1.08$, $\|L_1^{-1}\| \approx 1.58$, $\|L_2\| \approx 5.8 \cdot 10^{-16}$. The convergence history is illustrated in the plots of Figure 4.2. We see that $\|X\| \, \|L_1 V_k^{(1)} \hat{z} - u^{(1)}\|$ is reduced under the level $\|c^{(2)}\|$ already after three steps. Furthermore, the grade m_* is equal to 20 since the residual $\|L_1 V_k^{(1)} \hat{z} - u^{(1)}\|$ is zero after 20 steps. On the other hand, since in this example it is not necessary to reduce the residual much below the level $\|c^{(2)}\|$, the method does not need to reach the number of iterations corresponding to the grade.

We see that the residual estimate from Theorem 4.4 is realistic in Example 4.5, but in many cases it may be a gross overestimate of the actual convergence. Indeed, the result only exploits the Krylov decomposition (4.8), and therefore any approximation space whose basis satisfies this type of equation for $k = m_*$ could be used to obtain the bound in Theorem 4.4. A fundamental property of Krylov subspaces, which has not been employed so far, is that there is an underlying polynomial approximation taking place; this will be explored in the next section.

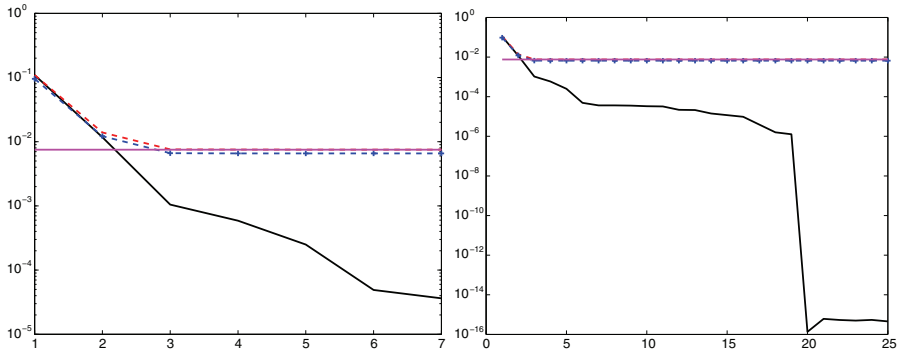

FIG. 4.2. *Example* 4.5, *residuals. First* 7 *steps (left),* 25 *steps (right).* $\|X\|\,\|L_1V_k^{(1)}\hat{z}-u^{(1)}\|$ *(solid),* $\|X\|\,(\|L_1V_k^{(1)}\hat{z}-u^{(1)}\|^2+\|c^{(2)}\|^2)^{1/2}$ *(dashed), true residual (dashed with +), and the estimate* (4.12) *(solid straight line).*

4.2. An improved residual estimate. For any polynomial p_m of degree not greater than m, we can write

$$p_m(B)c=[X_1,X_2]\begin{bmatrix}p_m(L_1)Y_1^*c\\ p_m(L_2)Y_2^*c\end{bmatrix}=X_1p_m(L_1)Y_1^*c+X_2p_m(L_2)Y_2^*c,$$

where $[Y_1,Y_2]^*=[X_1,X_2]^{-1}$ (cf. (4.6)). Therefore, using $X_1^*X_1=I$ and $Y_2^*c=c^{(2)}$,

$$\|p_m(B)c\|\le\|p_m(L_1)Y_1^*c\|+\|X_2p_m(L_2)c^{(2)}\|. \tag{4.15}$$

We denote by $\mathbb{P}_k$ the set of polynomials p of degree not greater than k and such that $p(0)=1$. We also recall that k iterations of GMRES generate an approximate solution d_k for $Bd=c$ with $d_k\in K_k(B,c)$ (for a zero initial guess) by minimizing the residual $r_k=c-Bd_k$ [37]. In terms of polynomials, this implies that $r_k=p_k(B)c$ where $p_k=\arg\min_{p\in\mathbb{P}_k}\|p(B)c\|$; p_k is called the GMRES residual polynomial.

The following theorem provides a bound of the GMRES residual when the spectra of L_1 and L_2 are well separated and the magnitude of $c^{(2)}$ is small compared with that of the whole vector c, as is the case in our setting. The proof is in the spirit of that in [10].

THEOREM 4.6. *Let* m_* *be the grade of* L_1 *with respect to* Y_1^*c. *Assume* k *iterations of GMRES have been performed on* $Bd=c$, *and let* r_k *be the corresponding residual. Let* Δ_2 *be a circle centered at the origin and having radius* ρ, *enclosing all eigenvalues of* L_2.

(i) *If* $k<m_*$, *let* $s_k^{(1)}=\phi_k(L_1)Y_1^*c$ *be the GMRES residual associated with* $L_1z=Y_1^*c$, *where* $\phi_k\in\mathbb{P}_k$. *Then*

$$\|r_k\|\le\|s_k^{(1)}\|+\|X_2\|\gamma_k\tau,\qquad\tau=\rho\max_{z\in\Delta_2}\|(zI-L_2)^{-1}c^{(2)}\|, \tag{4.16}$$

where $\gamma_k=\max_{z\in\Delta_2}\prod_{i=1}^k|\theta_i-z|/|\theta_i|$ *and* θ_i *are the roots of* ϕ_k.

(ii) *If* $k=m_*+j$, $j\ge0$, *let* $s_j^{(2)}=\varphi_j(L_2)c^{(2)}$ *be the GMRES residual associated with* $L_2z=c^{(2)}$ *after* j *iterations, where* $\varphi_j\in\mathbb{P}_j$, *so that* $\|s_j^{(2)}\|\le\|c^{(2)}\|$. *Then*

$$\|r_k\|\le\rho\gamma_{k_*}\|s_j^{(2)}\|\|X_2\|\max_{z\in\Delta_2}\|(zI-L_2)^{-1}\|, \tag{4.17}$$

where $\gamma_{m_*}=\max_{z\in\Delta_2}\prod_{i=1}^{m_*}|\theta_i-z|/|\theta_i|$ *and* θ_i *are the roots of the grade polynomial of* L_1.

Proof. Let us write $r_k = p_k(B)c$, where p_k is the GMRES residual polynomial.

(i) For $k < m_*$, we have $\|r_k\| = \min_{p\in\mathbb{P}_k} \|p(B)c\| \le \|\phi_k(B)c\|$, where ϕ_k is the GMRES residual polynomial associated with L_1 and Y_1^*c. Using (4.15), we have

$$\|\phi_k(B)c\| \le \|\phi_k(L_1)Y_1^*c\| + \|X_2\phi_k(L_2)c^{(2)}\| \le \|s_k^{(1)}\| + \|X_2\|\,\|\phi_k(L_2)c^{(2)}\|.$$

To evaluate the last term we use the Cauchy integral representation. From $\phi_k(L_2)c^{(2)} = \frac{1}{2\pi i}\int_{\Delta_2} \phi_k(z)(zI - L_2)^{-1}c^{(2)}dz$, we obtain

$$\|\phi_k(L_2)c^{(2)}\| \le \rho \max_{z\in\Delta_2} |\phi_k(z)| \max_{z\in\Delta_2} \|(zI - L_2)^{-1}c^{(2)}\|.$$

Using $\phi_k(z) = \prod_{i=1}^k (1 - z/\theta_i)$, the first result follows.

For $k \ge m_*$, we select the polynomial $p_k(z) = q_{m_*}(z)\varphi_j(z)$, where q_{m_*} is the grade polynomial, namely it satisfies $q_{m_*}(L_1)Y_1^*c = 0$, so that $p_k(L_1)Y_1^*c = 0$; moreover, $\varphi_j(z)$ is the GMRES residual polynomial after j iterations on $L_2z = c^{(2)}$. Then

$$\begin{aligned}\|r_k\| &\le \|p_k(B)c\| \le \|p_k(L_1)Y_1^*c\| + \|X_2p_k(L_2)c^{(2)}\| \\ &\le \|X_2\|\,\|p_k(L_2)c^{(2)}\| \le \|X_2\|\,\|q_{m_*}(L_2)\|\,\|\varphi_j(L_2)c^{(2)}\|.\end{aligned}$$

Once again, using the Cauchy integral representation,

$$\|q_{m_*}(L_2)\| \le \rho \max_{z\in\Delta_2} |q_{m_*}(z)| \max_{z\in\Delta_2} \|(zI - L_2)^{-1}\|.$$

Since $q_{m_*}(z) = \prod_{i=1}^{m_*}(1 - z/\theta_i)$, the result follows. ∎

A few comments are in order before we proceed with some examples. Assuming that $m_* \ll n$, Theorem 4.6(i) shows that the behavior of the first few iterations of GMRES is driven by the convergence of the reduced system $L_1d^{(1)} = Y_1^*c$ through the quantity $\|s_k^{(1)}\|$. During these iterations, the noise-related part of the problem may affect the bound on $\|r_k\|$ with the quantities $\|X_2\|$ and τ if B is nonnormal; otherwise the first term $\|s_k^{(1)}\|$ dominates. Such nonnormality reveals itself in two different ways: (a) the quantity τ may be large if the second diagonal block L_2 is very nonnormal, so that its resolvent norm may be large even for z not too close to the spectrum; (b) due to (4.7), $\|P\|$ and thus $\|X_2\|$ may be large if L_1 and L_2 are not well separated in terms of sep function, while the norm of the "coupling" matrix G is sizable. If $G = 0$, then X_2 has orthonormal columns and only the nonnormality of L_2 plays a role in the balance between the two terms in (4.16).

For k sufficiently large, we expect that $\|s_k^{(1)}\|$ will become smaller than the second term in (4.16), so that the second term $\|X_2\|\gamma_k\tau$ will start to dominate. For $k > m_*$ (item (ii) in Theorem 4.6), the first term is zero, so that a bound based on the system in L_2 may be obtained, as in (4.17). We also remark that this second bound differs considerably from that obtained in Theorem 4.4, which was stated for $k = m_*$.

We also need to comment on the expected size of τ and γ_k. The quantity τ collects information on the nonnormality of L_2 and on the size of the data perturbation. We already mentioned the role of the transfer function norm, which appears as $\|(zI - L_2)^{-1}c^{(2)}\| \le \|(zI - L_2)^{-1}\|\,\|c^{(2)}\|$. Therefore, the size of the noise-related data, $\|c^{(2)}\|$, may be amplified significantly on a nonnormal problem. On the other hand, the radius ρ also plays a role. We recall that $\|(zI - L_2)^{-1}\| \le \operatorname{dist}(z, \mathcal{F}(L_2))^{-1}$, where $\mathcal{F}(L_2)$ is the field of values[7] of L_2. Therefore, the circle Δ_2 may be set to be

[7]The field of values of an $n \times n$ matrix L is defined as $\mathcal{F}(L) = \{z^*Lz : z \in \mathbb{C}^n, \|z\| = 1\}$.

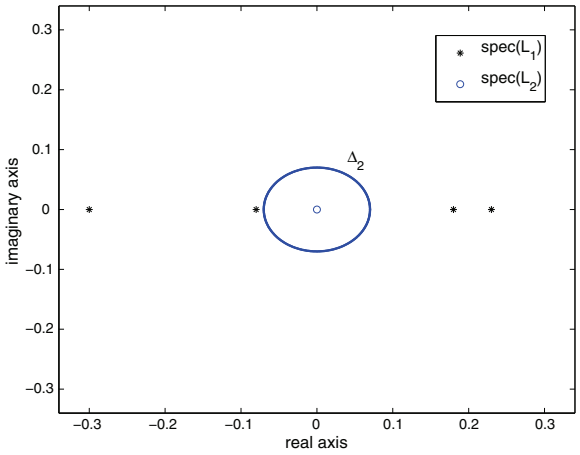

FIG. 4.3. *Location of the spectra of L_1 and L_2, and choice of the circle Δ_2 in Theorem* 4.6.

sufficiently far from $\mathcal{F}(L_2)$ (see Figure 4.3), so that $\|(zI - L_2)^{-1}\|$ is of moderate size, while maintaining ρ not too large, so as not to influence γ_k (see below). In that case, $\rho\|(zI - L_2)^{-1}\| \ll 1$, implying $\tau \approx \|c^{(2)}\|$. Similar considerations hold for the bound (4.17). The quantity γ_k is the maximum value of the GMRES residual polynomial on the circle Δ_2. If the circle tightly surrounds zero, then γ_k is very close to one since the residual polynomial ϕ_k satisfies $\phi_k(0) = 1$. Circles of larger radius may cause γ_k to assume significantly larger values, depending on the location of the polynomial roots θ's. We found that values of the radius ρ within $\|L_1\|$ provided good bounds; in general, however, we tried to select significantly smaller ρ's; see the examples below.

Theorem 4.4 gives good estimates only when $\|L_2\|$ is very small and L_1 is well-conditioned, which is the case when a good, singular preconditioner is used for an ill-posed problem. On the other hand, the improved result in Theorem 4.6 can be applied to estimate the behavior of GMRES applied directly (i.e., without preconditioner) to an ill-posed problem.

Example 4.7. We consider the `wing` example from the MATLAB Regularization Toolbox [21, 23] of dimension $n = 100$, and the largest few eigenvalues of A in absolute value are

$$3.7471 \cdot 10^{-1},\ -2.5553 \cdot 10^{-2},\ 7.6533 \cdot 10^{-4},\ -1.4851 \cdot 10^{-5},$$
$$2.1395 \cdot 10^{-7},\ -2.4529 \cdot 10^{-9},\ 2.3352 \cdot 10^{-11},\ -1.8998 \cdot 10^{-13},\ 1.3260 \cdot 10^{-15}.$$

We perturb the exact right-hand side b_e as $b = b_e + \varepsilon p$, with p having normally distributed random entries and $\|p\| = 1$. With the explicit Schur decomposition of the matrix, we take as L_1 the portion of B corresponding to the largest six eigenvalues in absolute value (that is $m_* = 6$), down to $\lambda_6 = -2.4529 \cdot 10^{-9}$; for this choice we have $\|G\| = 2.29 \cdot 10^{-5}$ and $\|P\| = 10.02$. This choice of L_1 was used to ensure that there is a sufficiently large gap between L_1 and L_2, while still being able to assume that $\|L_2\|$ is mainly noise. Note that since all relevant eigenvalues are simple, $m_* = m$ for this example. We then take a circle of radius $\rho = 2 \cdot 10^{-9} < \mathrm{dist}(\mathrm{spec}(L_1), 0)$. We compute the invariant subspace basis $[X_1, X_2]$ as in (4.6), where P was obtained by solving the associated Sylvester equation. We note that for $\varepsilon = 10^{-7}$ we have $\|Y_1^* c\| = 1$ and $\|Y_2^* c\| = 6.7 \cdot 10^{-7}$, while for $\varepsilon = 10^{-5}$ we obtain $\|Y_2^* c\| = 6.49 \cdot 10^{-5}$; all these are consistent with the used perturbation ε.

TABLE 4.1

Example 4.7. `wing` *data. Key quantities of Theorem* 4.6. L_1 *of size* 6×6 *(*$m_* = 6$*), so that* $\|G\| = 2.29 \cdot 10^{-5}$ *and* $\|P\| = 10.02$. *Circle of radius* $\rho = 2 \cdot 10^{-9}$.

				Bound	
ε	k	$\|s_k^{(1)}\|$	$\|X_2\|\gamma_k\tau$	(4.16) or (4.17)	$\|r_k\|$
10^{-7}	2	1.640e-03	6.770e-06	1.647e-03	1.640e-03
	3	3.594e-05	6.770e-06	4.271e-05	3.573e-05
	10			6.712e-06	6.311e-07
10^{-5}	2	1.621e-03	6.770e-04	2.298e-03	1.640e-03
	3	6.568e-05	6.770e-04	7.427e-04	7.568e-05
	10			6.442e-04	6.308e-05

TABLE 4.2

Example 4.8. `baart` *data. Key quantities of Theorem* 4.6. L_1 *of size* 7×7 *(*$m_* = 7$*), so that* $\|G\| = 6.4357 \cdot 10^{-3}$ *and* $\|P\| = 1.48$. *Circle of radius* $\rho = 2 \cdot 10^{-7}$.

				Bound	
ε	k	$\|s_k^{(1)}\|$	$\|X_2\|\gamma_k\tau$	(4.16) or (4.17)	$\|r_k\|$
10^{-7}	2	1.590e-02	5.851e-08	1.590e-02	1.590e-02
	3	5.105e-06	5.851e-08	5.165e-06	5.105e-06
	10			1.062e-07	3.188e-08
10^{-5}	2	1.590e-02	5.851e-06	1.590e-02	1.590e-02
	3	5.404e-06	5.851e-06	1.125e-05	6.110e-06
	10			1.062e-05	3.188e-06

Table 4.1 reports some key quantities in the bound of Theorem 4.6 for a few values of ε at different stages of the GMRES convergence. For $k < m_* = 6$ we see that the two addends of the bound in (4.16) perform as expected: $\|s_k^{(1)}\|$ dominates for the first few iterations, after which the second term leads the bound, providing a quite good estimate of the true residual norm, $\|r_k\|$. A larger perturbation ε makes this dominance effect more visible at an earlier stage.

Example 4.8. We consider the `baart` example from the same toolbox as in the previous example. This example will be considered again in later sections. The leading eigenvalues for the 100×100 matrix are

$$2.5490 \cdot 10^{0},\ -7.2651 \cdot 10^{-1},\ 6.9414 \cdot 10^{-2},\ -4.3562 \cdot 10^{-3},$$
$$2.0292 \cdot 10^{-4},\ -7.5219 \cdot 10^{-6},\ 2.3168 \cdot 10^{-7},\ -6.1058 \cdot 10^{-9},$$
$$1.4064 \cdot 10^{-10},\ -2.8770 \cdot 10^{-12},\ 5.2962 \cdot 10^{-14}.$$

We consider $m_* = 7$, giving $\|G\| = 6.4357 \cdot 10^{-3}$ and $\|P\| = 1.48$, and we chose $\rho = 2 \cdot 10^{-7}$. Also in this case, $m_* = m$ as all involved eigenvalues are simple. For $\varepsilon = 10^{-7}$ we have $\|Y_1^* c\| = 1$ and $\|Y_2^* c\| = 3.26 \cdot 10^{-8}$, while for $\varepsilon = 10^{-5}$ we obtain $\|Y_2^* c\| = 3.26 \cdot 10^{-6}$.

Table 4.2 reports some key quantities in the bound of Theorem 4.6 for a few values of ε at different stages of the GMRES convergence.

The digits in the table fully confirm what we found in the previous example, although here the addend carrying the perturbation is less dominant in the early phase of the convergence history.

Since $\|L_1^{-1}\|\,\|L_2\| \approx 0.061$, we see that Theorem 4.4 would give a much worse residual estimate for this example, where the eigenvalues are not as well separated as in Example 4.5.

5. Estimating the error. In this section we derive two error estimates. The first one is a standard estimate for ill-posed problems that is used in the literature to demonstrate continuous dependence on the data for a regularization method (especially when the problem is formulated in function spaces). In the second one we estimate the error due to the approximate solution of the least squares problem (4.1).

5.1. Error estimate for the singularly preconditioned problem. Assume that $A \in \mathbb{C}^{n\times n}$ is a matrix corresponding to a compact operator, i.e., an ill-conditioned matrix obtained by discretizing an ill-posed problem. Consider the linear system of equations $Ax = b$, where $b = b_e + \eta$, where b_e is an *exact* right-hand side, and η is a noise vector, which is assumed to be small in norm. For simplicity we assume that the smallest singular value of A is nonzero, but may be very small. The exact linear system $Ax = b_e$ has the solution[8] $x_e = A^{-1}b_e$. Let $M_m^\dagger \in \mathbb{C}^{n\times n}$ be a rank-m approximation[9] of A^{-1}. Then, in the preconditioned context, we have a rank-deficient least squares problem of type (4.1), with least norm solution $y_m = (AM_m^\dagger)^+ b$. The corresponding approximate solution of $Ax = b$ is $x_m = M_m^\dagger (AM_m^\dagger)^+ b$. To estimate $\|x_e - x_m\|$ we first introduce the generalized SVD (GSVD) [41, 33] of A^{-1} and $M_m^\dagger$,

$$A^{-1} = Z\Omega^{-1}P^*, \qquad M_m^\dagger = Z\Lambda^+ Q^*, \tag{5.1}$$

where $\Omega = \mathrm{diag}(\omega_1, \ldots, \omega_n)$ with $\omega_1 \geq \omega_2 \geq \cdots \geq \omega_n > 0$, and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, with $\lambda_1, \ldots, \lambda_m > 0$ and $\lambda_{m+1} = \cdots = \lambda_n = 0$. The matrices P and Q are unitary, while Z is only nonsingular.

PROPOSITION 5.1. *With the notation defined above, we can estimate*

$$\|x_e - x_m\| \leq \|Sx_e\| + \|M_m^\dagger (AM_m^\dagger)^+ (b_e - b)\|, \tag{5.2}$$

where

$$S = Z \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} Z^{-1}. \tag{5.3}$$

If $\hat{y}$ denotes any least squares solution of (4.1)*, and $\hat{x} = M_m^\dagger \hat{y}$, then $\hat{x} = x_m$.*

Proof. Let $\Lambda_m = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$. It is straightforward to show that

$$AM_m^\dagger = P \begin{bmatrix} \Omega_m \Lambda_m^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q^*, \qquad M_m^\dagger (AM_m^\dagger)^+ = Z \begin{bmatrix} \Omega_m^{-1} & 0 \\ 0 & 0 \end{bmatrix} P^*, \tag{5.4}$$

where $\Omega_m = \mathrm{diag}(\omega_1, \omega_2, \ldots, \omega_m)$. It follows immediately that

$$M_m^\dagger (AM_m^\dagger)^+ A = Z \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} Z^{-1} = I - S, \tag{5.5}$$

where S is defined by (5.3). We can now estimate

$$\|x_e - x_m\| = \|x_e - M_m^\dagger (AM_m^\dagger)^+ b\| \leq \|x_e - M_m^\dagger (AM_m^\dagger)^+ b_e\| + \|M_m^\dagger (AM_m^\dagger)^+ (b_e - b)\|.$$

For the first term we use $b_e = Ax_e$ and (5.5) and get $\|x_e - M_m^\dagger (AM_m^\dagger)^+ b_e\| = \|Sx_e\|$, from which the error bound follows.

[8] Note that in the context of ill-posed problem the inverse A^{-1} makes sense only in connection with exact data; cf. [14, Chapter 3].

[9] Recall from section 3 that $M_m^\dagger$ approximates the low frequency part of A^{-1}.

For the second part of the proposition, partition $Q = (Q_1\ Q_2)$, where $Q_1 \in \mathbb{C}^{n\times m}$. Then from (5.4) we see that the columns of Q_2 form a unitary basis for the nullspace of $AM_m^\dagger$. Then assume that our solver does not give the exact minimum norm least squares solution of (4.1) but also has a component in the nullspace; i.e., we have $\hat{y} = \left(AM_m^\dagger\right)^+ b + Q_2 w$ for some w. Since the nullspaces of $M_m^\dagger$ and $AM_m^\dagger$ coincide, multiplication by $M_m^\dagger$ annihilates $Q_2 w$, and $\hat{x} = M_m^\dagger \hat{y} = x_m$. ∎

For the discussion, let the SVD of A be $A = U\Sigma V^*$. The ideal rank-m preconditioner M is the best rank-m approximation (in Frobenius or operator norm) of A, $M = U_m \Sigma_m V_m^*$, with U_m, V_m collecting the first m columns of U and V, respectively, and Σ_m being the leading $m \times m$ portion of Σ. Then $M_m^\dagger = M^+ = V_m \Sigma_m^{-1} U_m^*$. In the nonideal case, the better the rank-m preconditioner $M = Q\Lambda Z^{-1}$ approximates A in some sense, the closer the decomposition of A corresponding to (5.1) is to the SVD, and the closer Z is to the matrix V in the SVD of A. Therefore, *for a good preconditioner the matrix S in* (5.3) *is a projection onto the high frequency part of the solution.* Thus the estimate (5.2), and its worst case version,

$$\|x_e - x_m\| \le \|Sx_e\| + \|M_m^\dagger (AM_m^\dagger)^+\|\, \|b_e - b\|,$$

are analogous to those in the proofs of Proposition 3.7 and Theorem 3.26 in [14], where with an assumption about the smoothness of the exact solution x_e and with a suitable regularization parameter choice rule (e.g., the discrepancy principle), continuous dependence on the data ($\|b - b_e\|$) is proved.

5.2. The GMRES approximation error. The GMRES algorithm delivers a monotonically nonincreasing residual norm $\|r\| = \|b - Ax_m\| = \|c - Bd\|$, and we have shown that under certain spectral hypotheses on B, this norm can be sufficiently small. We will now consider (4.1) and discuss the error in the solution approximation that arises due to the fact that we do not solve that least squares problem exactly. We will assume[10] that $\operatorname{rank}(AM_m^\dagger) = m$, which implies that the Schur decomposition is

$$U^*(AM_m^\dagger)U = \begin{bmatrix} L_1 & G \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} B_1 \\ 0 \end{bmatrix} = B. \tag{5.6}$$

The least squares problem (4.2) can then be written as

$$\min_d \{\|B_1 d - c^{(1)}\|^2 + \|c^{(2)}\|^2\}, \qquad c = U^* b = \begin{bmatrix} c^{(1)} \\ c^{(2)} \end{bmatrix}, \tag{5.7}$$

which, due to the nonsingularity of L_1, is equivalent to the underdetermined system

$$B_1 d = c^{(1)}. \tag{5.8}$$

THEOREM 5.2. *Assume that* $\operatorname{rank}(AM_m^\dagger) = m$ *with Schur decomposition* (5.6). *Let* $x_m = M_m^\dagger y_m$, *where* $y_m = \left(AM_m^\dagger\right)^+ b$ *is the minimum norm least squares solution of* (4.1), *and let* $\hat{x}_m = M_m^\dagger \hat{y}_m$, *where* $\hat{y}_m$ *is an approximate solution of* (4.1). *Then*

$$\|x_m - \hat{x}_m\| \le 2\kappa(AM_m^\dagger)\, \|M_m^\dagger\|\, \frac{\|r^{(1)}\|}{\|AM_m^\dagger\|\, \|\hat{y}\| + \|c^{(1)}\|} + O(\|r^{(1)}\|^2),$$

where $\|r^{(1)}\| = \|B_1 \hat{d} - c^{(1)}\|$, $\hat{d} = U^* \hat{y}_m$, *and* $c^{(1)}$ *is defined by* (5.7).

[10] When in section 4 we analyzed the behavior of GMRES, it was necessary to take into account a nonzero block L_2, due to round-off. Here, since we are using a preconditioner of rank m, it makes sense to compare the approximate solution with the one that would be obtained for $L_2 = 0$.

Proof. From the second part of Proposition 5.1 we see that we do not need to take into account any component of $\hat{y}$ (or, equivalently, $\hat{d}$) in the nullspace of $AM_m^\dagger$, since that part will be annihilated in the multiplication by $M_m^\dagger$. Therefore, the sensitivity to perturbations of the underdetermined problem (5.8) is equivalent to that of a corresponding square problem. Using standard results for linear systems [26, section 7.1], we get

$$\|x_m - \hat{x}_m\| \le 2\kappa(AM_m^\dagger)\,\|M_m^\dagger\|\,\frac{\|r^{(1)}\|}{\|B_1\|\,\|\hat{d}\| + \|c^{(1)}\|} + O(\|r^{(1)}\|^2),$$

from which the result follows. ∎

In our numerical experiments we have observed that GMRES applied to (4.2) can produce approximate solutions y such that $\|r^{(1)}\| = \|B_1 y - c^{(1)}\| \ll \|By - c\| = \|r\|$. In actual large-scale computations we do not have access to the Schur decomposition,[11] so we cannot obtain $r^{(1)}$. However, consider the quantity

$$B^* r = B^* \begin{bmatrix} r^{(1)} \\ r^{(2)} \end{bmatrix} = \begin{bmatrix} L_1^* r^{(1)} \\ G^* r^{(1)} + L_2^* r^{(2)} \end{bmatrix}.$$

Since we have assumed that $\|L_2\| \ll \|L_1\|$, we see that the occurrence that $\|B^* r\| \ll \|r\|$ gives an indication that $\|r^{(1)}\|$ is considerably smaller than $\|r\|$. Indeed, for $\sigma_{\min}(L_1) \gg 0$, the condition $\|B^* r\| \ll \|r\|$ corresponds to $\|L_1^* r^{(1)}\|^2 + \|G^* r^{(1)} + L_2^* r^{(2)}\|^2 \ll \|r\|^2$ with $\|L_1^* r^{(1)}\| \ge \sigma_{\min}(L_1)\|r^{(1)}\|$, from which the assertion follows.

The same is true if $\|A^* s\| \ll \|s\|$, where $s = b - Ax$, since $\|A^* s\| = \|B^* r\|$. Residuals and this estimate are illustrated in Figure 6.8. In light of these considerations, (and in cases when the computation of $A^* s$ is not prohibitively expensive), we would like to encourage monitoring $\|A^* s\|$ during the GMRES iterations as a companion of a stopping criterion based on the discrepancy principle.

By combining the estimates in Proposition 5.1 and Theorem 5.2, we get an estimate for the total error $\|x_e - \hat{x}_m\|$. Assuming that M is a good low-rank approximation of A, the pseudoinverse of the preconditioned matrix, $(AM_m^\dagger)^+$, is small in norm. Furthermore, since $M_m^\dagger$ is an approximate solution operator for the ill-posed problem, $\|M_m^\dagger\|$ is only as large as is needed for obtaining a reasonable regularized solution.

Normally when an iterative solver is used for an ill-posed problem, it is the number of iterations that acts as regularization parameter. However, here the error estimates show that the regularization is mainly due to the preconditioner.

6. Numerical examples. In this section we solve numerically four ill-posed problems. Perturbations to the data were added to illustrate the sensitivity of the solution to noise (cf. Proposition 5.1), and to also simulate measurement errors that occur in real applications. The first two examples are small and are chosen to illustrate different aspects of the theory. The last two are problems where the use of singular preconditioners is particularly useful: to our knowledge there are no papers in the literature describing the solution of ill-posed problems with variable coefficient PDEs in two or three space dimensions.

6.1. An ill-posed problem. Our first example is a discretization $Kf = g$ of an integral equation of the first kind [1] (test problem `baart` in [21, 23]),

[11] This is because it is either too expensive to compute the Schur decomposition or the matrix A is not available explicitly. See sections 3 and 6.4.

$$\int_0^\pi \exp(s\,\cos t) f(t)dt = 2\sinh(s)/s, \qquad 0 \le s \le \pi/2,$$

with solution $f(t) = \sin t$. The results are typical for unpreconditioned GMRES applied to an ill-posed problem and clearly show the phenomenon of semiconvergence.

The singular values and the eigenvalues of the matrix K of dimension $n = 200$ are illustrated in Figure 6.1. Clearly K is numerically singular. However, it is not easy to decide about its numerical rank. No matter what value, between 2 and 11, of the dimension of L_1 in the ordered Schur decomposition we choose, the smallest singular value of L_1 is much smaller than the norm of G.

We added a normally distributed perturbation to the right-hand side, and performed 10 GMRES steps. In Figures 6.3 and 6.4 we illustrate the approximate solution at iterations 2–5. For comparison we also show the solution using Tikhonov regularization, $\min_f\{\|Kf - g_{pert}\|^2 + \mu^2\|Lf\|^2\}$, where L was a discrete first derivative. The value of the regularization parameter was chosen according to the discrepancy principle: it was successively halved until the least squares residual was smaller than a tolerance; see below.

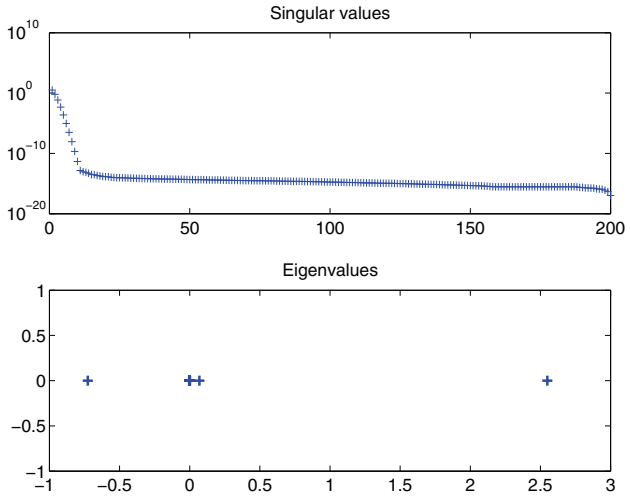

FIG. 6.1. *Singular values and eigenvalues of the matrix K for the* `baart` *problem. Note that all eigenvalues except the three of largest magnitude belong to a cluster at the origin.*

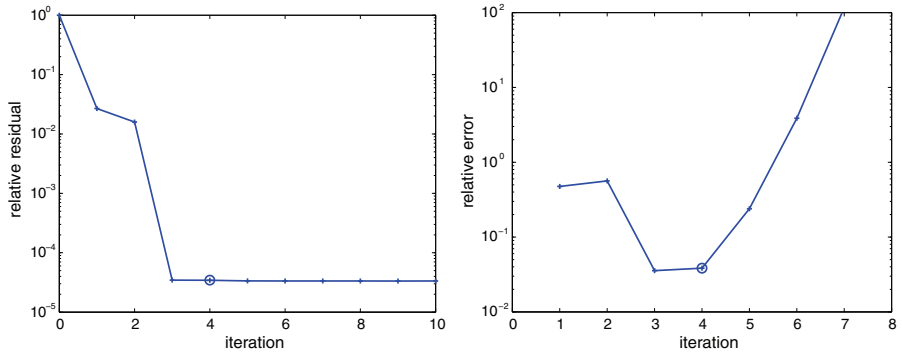

FIG. 6.2. `baart` *example: relative residual (left) and relative error (right) as functions of the GMRES step number. The circle marks when the stopping criterion was first satisfied.*

In Figure 6.2 we give the relative residual and the relative error for the GMRES iterations. Clearly the residual stagnates after 3 steps, and the solution starts to diverge after 4. This is also seen in Figures 6.3–6.4. The discrepancy principle is used as stopping criterion. The data error is $\|g - g_{pert}\|/\|g\| \approx 3.5 \cdot 10^{-5}$. If we choose $m = 4$, then $\|c^{(2)}\| \approx 10^{-4}$. The iterations are stopped when the relative norm of the residual is smaller than $7 \cdot 10^{-5}$. In Figure 6.2 we mark when the stopping criterion was satisfied. The results agree with those in [29, Example 5.3] and are explained by our theoretical analysis in section 4.2.

6.2. A preconditioned ill-posed problem. In this example we solve numerically a Cauchy problem for a parabolic PDE in the unit square (we will refer to it as Cauchy-1D). The purpose is not to propose a method for solving an ill-posed problem in one space dimension (because there are other, simpler methods for that) but to analyze numerically and illustrate why the preconditioned GMRES method works for the corresponding problem in two space dimensions. We also report comparisons with the circulant preconditioners mentioned in section 3.

The Cauchy problem is

$$(\alpha(x)u_x)_x = u_t, \qquad 0 \le x \le 1, \quad 0 \le t \le 1, \tag{6.1}$$

$$u(x,0) = 0, \qquad 0 \le x \le 1, \tag{6.2}$$

$$u_x(1,t) = 0, \qquad 0 \le t \le 1, \tag{6.3}$$

$$u(1,t) = g(t), \quad 0 \le t \le 1, \tag{6.4}$$

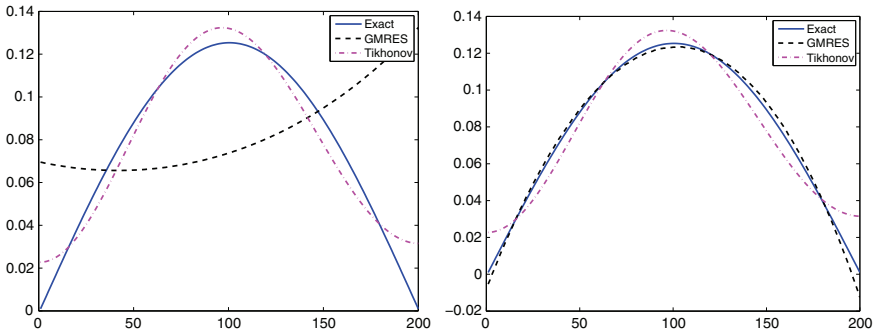

FIG. 6.3. `baart` *example: exact solution (solid), GMRES solution (dashed), and Tikhonov solution for* $\mu = 0.03125$ *(dashed-dotted). Left: after* 2 *GMRES iterations; right: after* 3.

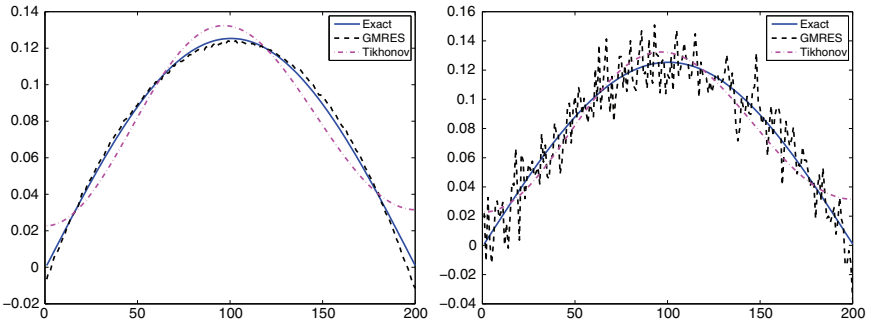

FIG. 6.4. `baart` *example: exact solution (solid), GMRES solution (dashed), and Tikhonov solution (dashed-dotted). Left: after* 4 *GMRES iterations; right: after* 5.

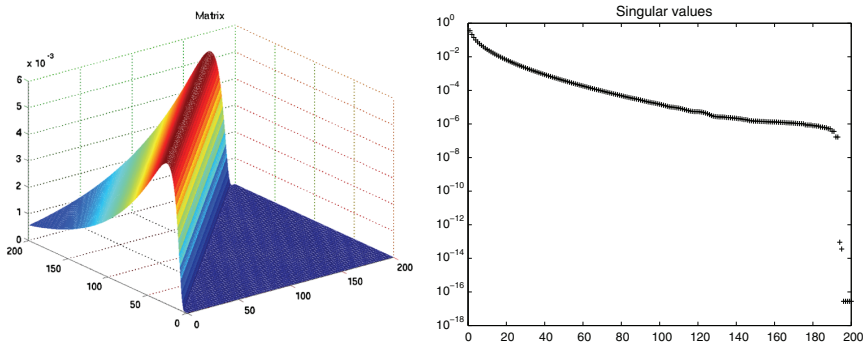

FIG. 6.5. *Cauchy-1D example. Matrix and singular values.*

where the parabolic equation has a variable coefficient

$$\alpha(x) = \begin{cases} 1, & 0 \le x \le 0.5, \\ 2, & 0.5 \le x \le 1. \end{cases}$$

The solution $f(t) = u(0,t)$ is sought. This problem, which we call the *sideways heat equation*, is severely ill-posed; see, e.g., [4, 11, 12]. It can be written as a Volterra integral equation of the first kind,

$$\int_0^t k(t-\tau) f(\tau) d\tau = g(t), \qquad 0 \le t \le 1. \tag{6.5}$$

The kernel $k(t)$ is not known explicitly in the case of a variable coefficient $\alpha(x)$. We compute it by solving (using the MATLAB stiff solver `ode23s`) a well-posed problem (6.1)–(6.3) and as boundary values at $x = 0$ an approximate Dirac delta function at $t = 0$. The integral equation (6.5) is then discretized giving a linear system of equations $Kf = g$ of dimension $n = 200$, where K is a lower triangular Toeplitz matrix, illustrated in Figure 6.5. To construct the data we selected a solution f, solved (6.1)–(6.3) with boundary values $u(0,t) = f(t)$ using the MATLAB `ode23s`. The data vector g was then obtained by evaluating the solution at $x = 1$. To simulate measurement errors we added a normally distributed perturbation such that $\|g_{pert} - g\|/\|g\| = 10^{-2}$.

As the diagonal of K is equal to zero, this is an eigenvalue of multiplicity 200, and the assumptions of section 4 are not satisfied. Therefore it is not surprising that such a linear system cannot be solved by GMRES; see [29, Example 5.1] and [9, Example 4.1], where a closely related sideways heat equation is studied.

On the other hand, for this problem the initial decay rate of the singular values is relatively slow (see Figure 6.5), and therefore it should be possible to solve approximately a regularized version of the system $Kf = g$. To this end we precondition the linear system by a problem with a constant coefficient $\alpha_0 = 1.5$. The kernel functions are given in Figure 6.6.

For the discretized problem with constant coefficient with matrix K_0, we compute the SVD, $K_0 = U\Sigma V^T$, and define the preconditioner as a truncation to rank $m = 20$ of the pseudoinverse, $M_m^\dagger = V_m \Sigma_m^{-1} U_m^T$. The eigenvalues of the preconditioned matrix $KM_m^\dagger$ are illustrated in Figure 6.6. Clearly, the numerical rank of $KM_m^\dagger$ is equal to m. We also computed the ordered Schur decomposition (4.2) of $KM_m^\dagger$. The matrix L_1 had condition number $\kappa_2(L_1) = \sigma_1(L_1)/\sigma_m(L_1) = 1.43$, $\|G\| = 0.0962$, and $\|c^{(2)}\| \approx 0.0066$. Thus, in this example the data perturbation is larger than $\|c^{(2)}\|$.

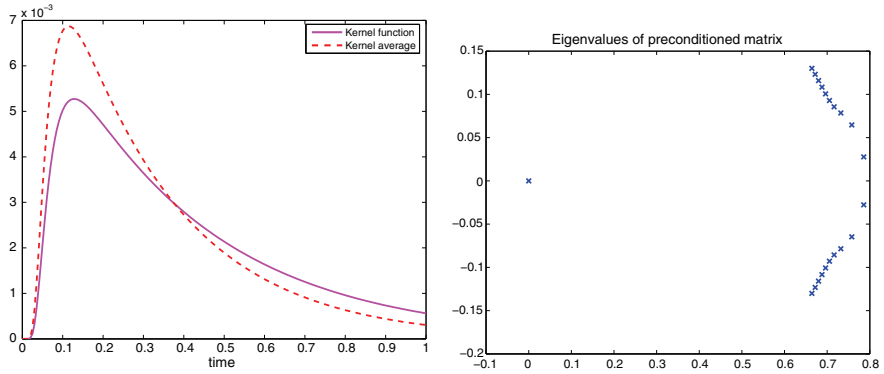

FIG. 6.6. *Cauchy-1D example. Left: kernel function $k(t)$ for the operator with variable coefficients (solid) and for the constant coefficient (dashed). Right: eigenvalues of the preconditioned matrix $KM_m^\dagger$.*

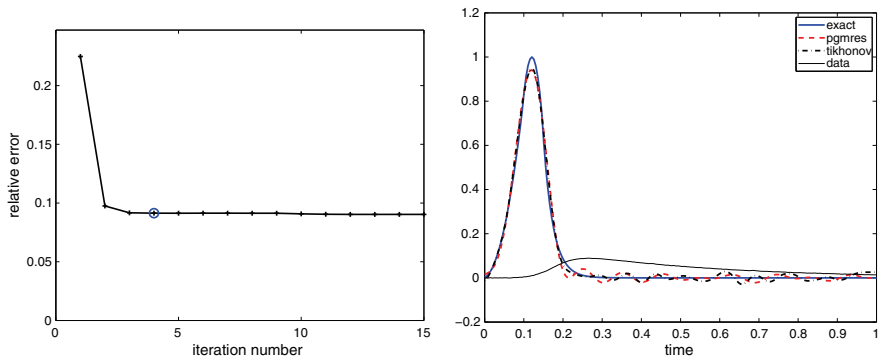

FIG. 6.7. *Cauchy-1D example. Left: relative error as a function of iteration index. The circle marks when the stopping criterion was first satisfied. Right: exact solution (solid), approximate solution after 4 iterations of preconditioned GMRES (dashed), and Tikhonov solution with $\mu = 0.015625$. The lower solid curve is the right-hand side.*

We applied 15 GMRES iterations to the preconditioned system. The relative error is given in Figure 6.7(left) (for the relative residual; cf. Figure 4.2). The numerical solution after 4 steps is illustrated in Figure 6.7(right), where, for comparison, we also show the solution using Tikhonov regularization, implemented as in the previous example. It is seen that the two approximate solutions have comparable accuracy.

The stopping criterion (with a fudge factor of 1.1) was satisfied after 4 GMRES steps. From the left plot of Figure 6.7 we see that the solution accuracy does not deteriorate as the iterations proceed; cf. the last paragraph of section 5.2.

In Figure 6.8 we demonstrate that $\|r^{(1)}\|$ is well approximated by $\|B^*r\|$, and that this part of the residual is much smaller than the overall residual $\|r\|$. Here we illustrate 25 GMRES steps to show that after 20 steps the residual for the first part of the system is of the order of machine precision.

Due to the shift-invariance of the kernel in the integral equation (6.5), the coefficient matrix has Toeplitz structure and can be preconditioned by a circulant matrix. Therefore, in addition to the preconditioner described in this section, we also made some experiments with the Strang preconditioner [40], as in the discussion in section 3. We computed the eigenvalue decomposition (3.1) of the circulant matrix (by

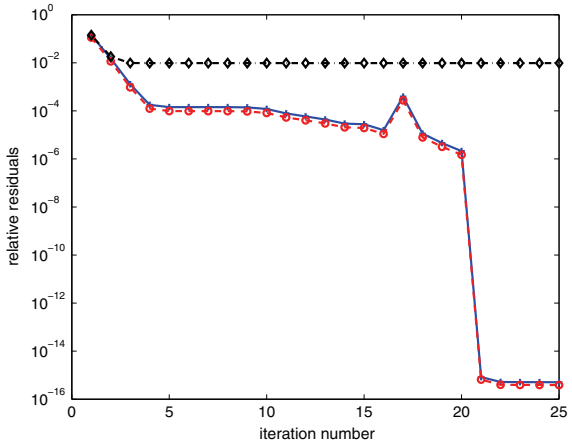

FIG. 6.8. *Cauchy-1D example. Relative residual norm* $\|r\|$ *(diamonds),* $\|r^{(1)}\|$ *(+), and* $\|B^*r\|$ *(o), as functions of iteration index.*

FFT) and retained only the 20 largest eigenvalues, thereby obtaining a singular preconditioner of rank 20 (cf. (3.3)). As a comparison we also used the nonsingular preconditioner (3.2). After 5 GMRES steps the results for both preconditioners were virtually indistinguishable from those reported earlier in this section.

6.3. A preconditioned two-dimensional (2D) ill-posed elliptic problem. It is in the numerical solution of Cauchy problems for PDEs with variable coefficients in two or more space dimensions that the application of a singular preconditioner is particularly interesting. The following elliptic Cauchy problem is severely ill-posed:

$$
\begin{aligned}
&(\beta(y)u_y)_y + (\alpha(x)u_x)_x + \gamma u_x = 0, \quad 0 < x < 1, \quad 0 < y < 1,\\
&u(y,0) = u(y,1) = 0, \qquad\qquad 0 \le y \le 1,\\
&u(x,0) = g(x), \qquad\qquad 0 \le x \le 1,\\
&u_y(x,0) = 0, \qquad\qquad 0 \le x \le 1,
\end{aligned}
\tag{6.6}
$$

where $u(x,1) = f(x)$ is sought from the Cauchy data at the boundary $y = 0$. The coefficients are $\alpha(x) = 1$, $\gamma = 2$, and

$$
\beta(y) = \begin{cases} 50, & 0 \le y \le 0.5, \\ 8 & 0.5 < y \le 1. \end{cases}
$$

We generated a solution $f(x)$ and computed the corresponding data function $g(x)$ by solving the well-posed elliptic equation with boundary data $u(x,1) = f(x)$ and $u_y(x,0) = 0$. Due to the relatively sharp gradients at the ends of the interval and the constant behavior at the middle, the Cauchy problem becomes difficult in the sense that the solution cannot be well represented by a low-rank approximation.

We added zero-mean normally distributed noise such that the data perturbation was $\|g - g_{pert}\|/\|g\| \approx 1.8 \cdot 10^{-3}$. We discretized the problem using finite differences, with 100 unknowns in each dimension.

Had the coefficient $\beta(y)$ been constant, we could have solved (6.6) approximately using an obvious extension of the Krylov-based method in [13]. In that method applied to the Cauchy problem with $\beta(y) = \beta_0$, a low-rank approximation is computed using a basis of a Krylov space for the operator L^{-1}, where $L = (\alpha(x)u_x)_x + \gamma u_x$,

and approximate evaluation of the solution as $f_0 = \cosh((1/\beta_0 L_m)^{1/2})$, with β_0 equal to the mean value of $\beta(y)$ over the interval, and where $\cosh((1/\beta_0 L_m)^{1/2})$ denotes a rank-m approximation of $\cosh((1/\beta_0 L_m)^{1/2})$. Here we used that approximate method as preconditioner, where the rank was determined as large as possible without obtaining an unstable solution (i.e., a solution with large oscillations). The rank was chosen equal to 9. Note that to compute the action of the preconditioning operator to a vector, it is only required to solve a number of well-posed non-self-adjoint one-dimensional (1D) elliptic problems (in this case 15). For a more detailed description of the preconditioner, see [13].

The preconditioned problem that we solved by GMRES was

$$\min_y \|(AM_m^\dagger)y - g\|, \qquad M_m^\dagger = \cosh((1/\beta_0 L_m)^{1/2}),$$

where the action of A to a vector v is equivalent to solving the well-posed problem (6.6) of dimension 10000, with $u(1,x) = v(x)$ replacing $u(x,0) = g(x)$. We performed a small number of iterations (our theory in the preceding section indicates that at most 9 iterations are needed). As stopping criterion we used the discrepancy principle. In Figure 6.9 we plot the relative residual. The stopping criterion with a fudge factor of 1.2 was first satisfied after 3 iterations.

The approximate solution after 3 steps is illustrated in the left plot of Figure 6.10. The "visual quality" of the solution was almost the same with 3–6 steps. In Figure 6.10(left) we also give the approximate solution produced using only the preconditioner of rank 9 as solution operator.

Unpreconditioned GMRES exhibited the typical semiconvergence behavior of an iterative method applied to an ill-posed problem. The smallest error was obtained after 5 steps, with approximate solution illustrated in Figure 6.10(right).

In this problem the linear operator is given only implicitly; hence it is not straightforward to give a measure of the non-self-adjointness. On the other hand the non-self-adjointness is reflected in the Hessenberg matrix H_k occurring in GMRES. Thus for the unpreconditioned iteration, we define $\tilde{H}_{10}$ as the 10×10 leading submatrix of H_k for $k \geq 10$. Then we have $\|\tilde{H}_{10} - \tilde{H}_{10}^T\|/\|\tilde{H}_{10}\| \approx 0.42$.

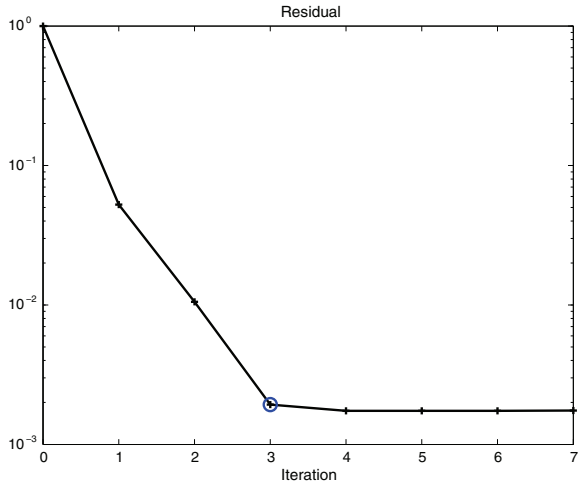

FIG. 6.9. *Elliptic-2D example. Relative residual as function of iteration index.*

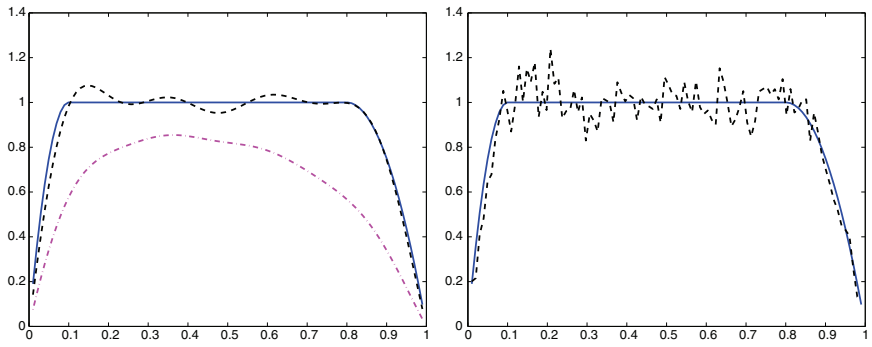

FIG. 6.10. *Elliptic-2D example. Exact solution (solid line). Left: preconditioned GMRES solution after 3 steps (dashed), solution with preconditioner only (dashed-dotted). Right: GMRES solution (no preconditioning) with smallest error, obtained after 5 steps (dashed).*

The approach of this example can be employed for more general operators. Assuming that $\mathcal{L}$ is a 2D elliptic operator (self-adjoint or non-self-adjoint), our methodology can be used to solve three-dimensional elliptic Cauchy problems for equations of the type

$$(d(z)u_z)_z + \mathcal{L}u = 0,$$

with variable coefficient $d(z)$ and cylindrical geometry with respect to z.

6.4. A preconditioned 2D ill-posed parabolic problem. Here we consider the problem

$$\begin{array}{lll} u_t = (\alpha(x)u_x)_x + (\beta(y)u_y)_y, & 0 < x < 1, \quad 0 < y < 1, & 0 \le t \le 1, \\ u(x,y,0) = 0, & 0 \le x \le 1, \quad 0 \le y \le 1, & \\ u(x,0,t) = u(x,1,t) = 0, & 0 \le x \le 1, \quad 0 \le t \le 1, & \\ u(1,y,t) = g(y,t), & 0 \le y \le 1, \quad 0 \le t \le 1, & \\ u_x(1,y,t) = 0, & 0 \le y \le 1, \quad 0 \le t \le 1, & \end{array} \tag{6.7}$$

where $u(0,y,t) = f(y,t)$ is sought from the Cauchy data at the boundary $x = 1$, and

$$\alpha(x) = \begin{cases} 2.5, & 0 \le x \le 0.5, \\ 1.5, & 0.5 < x \le 1, \end{cases} \qquad \beta(y) = \begin{cases} 0.75, & 0 \le y \le 0.5, \\ 1.25 & 0.5 < y \le 1. \end{cases}$$

The solution is taken to be

$$f(y,t) = \exp\left(4 - \frac{1}{y(1-y)}\right) \exp\left(4 - \frac{1}{t(1-t)}\right).$$

An approximate data function g was computed by replacing the condition $u(1,y,t) = g(y,t)$ in (6.7) by $u(0,y,t) = f(y,t)$, which gives a well-posed problem. After finite difference discretization with respect to x and y and 50 unknowns in each dimension, this problem can be considered as a stiff system of ordinary differential equations of dimension 2500, and is solved using the MATLAB `ode23s`. The Cauchy data are then obtained by evaluating the solution at $x = 1$.

Due to the unimodal nature of the exact solution (cf. Figure 6.12), the problem might seem easy to solve. However, the fact that the solution is close to zero in a relatively large region along the border of the unit square makes it difficult to expand it using a small number of sine functions (as is used in the preconditioner).

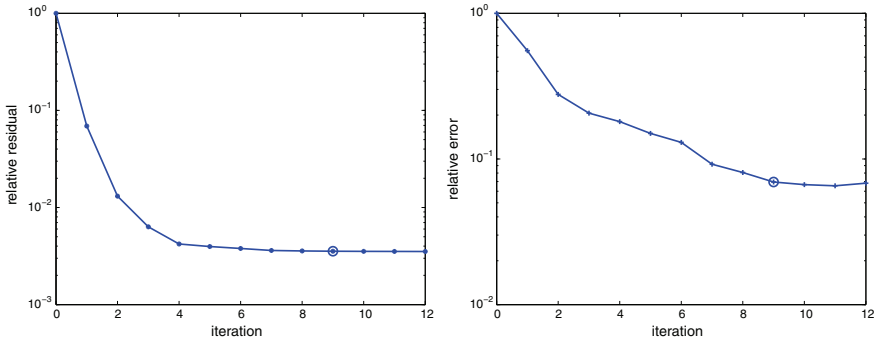

FIG. 6.11. *Parabolic-2D example. Relative residual and error as function of the number of iterations. The stopping criterion was satisfied after* 9 *steps.*

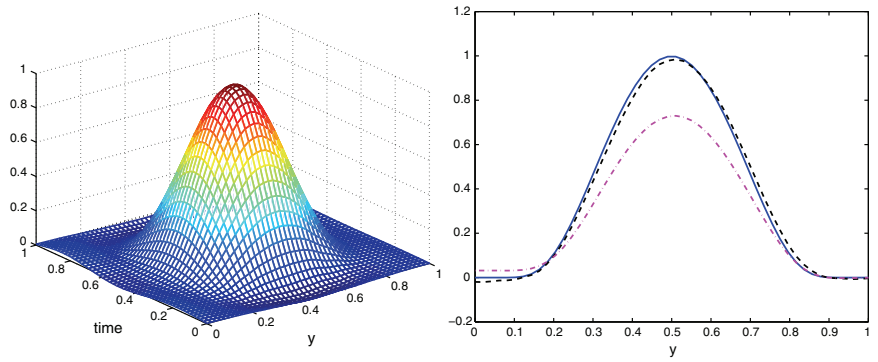

FIG. 6.12. *Parabolic-2D example. The solution after* 9 *iterations (left). Right: the exact solution (solid), the approximate solution (dashed), and the solution with preconditioner only (dashed-dotted) at* $t = 0.5$.

A discretization of the problem would give a linear system $Kf = g$. Since we discretize with $n = 50$ equidistant points in both the y and t directions, that matrix would have dimension 2500. However, due to the variable coefficients, we cannot compute the matrix; instead, when in GMRES we multiply a vector by K, we solve a parabolic equation in a way similar to how we computed the data g, but here we used the Crank–Nicholson method with step size 1/50.

The preconditioner is based on the approximation of the differential operator by a corresponding one with constant coefficients (average values). Then, since the geometry is rectangular, separation of variables can be applied, and a semianalytic solution formula can be applied (see [36]) involving an expansion in Fourier (sine) series. It is the truncation of this series that leads to a singular preconditioner $M_m^\dagger$, whose rank is equal to nq, where q is the number of terms in the series. Each term in the series involves, in addition, the solution of a 1D ill-posed Cauchy problem using Tikhonov regularization. The preconditioner is discussed in detail in [35]. In our numerical experiment the data perturbation was $\|g - g_{pert}\|/\|g\| = 3.6 \cdot 10^{-3}$, the preconditioner regularization parameter was 0.06, and $q = 6$.

In Figure 6.11 we plot the relative residual and the relative error. Note that, as in the previous examples, the solution accuracy is not sensitive to the exact choice of the stopping criterion. The approximate solution after the 9th iteration, when the relative residual was first smaller than $3.6 \cdot 10^{-3}$, is shown in Figure 6.12.

7. Conclusions. The main contributions of the present paper are the following. We give an eigenvalue-based analysis of the use of GMRES for almost singular linear systems of equations, where the eigenvalues are well separated and clustered. This gives a theoretical and algorithmic basis for the use of singular preconditioners for non-self-adjoint ill-posed problems. The GMRES method is used here and in [35] to solve Cauchy problems for parabolic and elliptic equations with variable coefficients, with a singular (low-rank) preconditioner based on a corresponding problem with constant coefficients.

The case of "ill-determined numerical rank" (where there is no distinct eigenvalue gap) is also treated. It is shown that in both cases a stopping criterion based on the discrepancy principle will give a numerical solution that is as good an approximation as is admissible, given the problem properties and the noise level.

The fact that GMRES with a singular preconditioner can be efficiently applied opens new possibilities in the numerical solution of ill-posed problems in two and three space dimensions, self-adjoint or non-self-adjoint, linear or nonlinear. As soon as a nearby linear ill-posed problem has a fast solver that can be regularized by cutting off high frequencies,[12] that solver can be used as preconditioner. Thus, in each step a well-posed problem with variable coefficients is solved, and a fast, regularized solver is applied. With a good preconditioner only a small number of Krylov steps will be required.

Acknowledgments. This work was started when the first author was a visiting fellow at the Institute of Advanced Study, University of Bologna. The authors would like to thank Dianne O'Leary and two anonymous referees for their many helpful comments.

REFERENCES

[1] M. L. Baart, *The use of auto-correlation for pseudo-rank determination in noisy ill-conditioned linear least-squares problems*, IMA J. Numer. Anal., 2 (1982), pp. 241–247.
[2] J. Baglama and L. Reichel, *Decomposition methods for large linear discrete ill-posed problems*, J. Comput. Appl. Math., 198 (2007), pp. 332–343.
[3] C. A. Beattie, M. Embree, and D. C. Sorensen, *Convergence of polynomial restart Krylov methods for eigenvalue computations*, SIAM Rev., 47 (2005), pp. 492–515.
[4] J. V. Beck, B. Blackwell, and S. R. Clair, *Inverse Heat Conduction. Ill-Posed Problems*, Wiley, New York, 1985.
[5] P. Brianzi, F. Di Benedetto, and C. Estatico, *Improvement of space-invariant image deblurring by preconditioned Landweber iterations*, SIAM J. Sci. Comput., 30 (2008), pp. 1430–1458.
[6] P. Brianzi, P. Favati, O. Menchi, and F. Romani, *A framework for studying the regularizing properties of Krylov subspace methods*, Inverse Problems, 22 (2006), pp. 1007–1021.
[7] P. N. Brown and H. F. Walker, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 37–51.
[8] D. Calvetti, B. Lewis, and L. Reichel, *GMRES, L-curves, and discrete ill-posed problems*, BIT, 42 (2002), pp. 44–65.
[9] D. Calvetti, B. Lewis, and L. Reichel, *On the regularizing properties of the GMRES method*, Numer. Math., 91 (2002), pp. 605–625.
[10] S. L. Campbell, I. C. F. Ipsen, C. T. Kelley, and C. D. Meyer, *GMRES and the minimal polynomial*, BIT, 36 (1996), pp. 664–675.
[11] A. Carasso, *Determining surface temperatures from interior observations*, SIAM J. Appl. Math., 42 (1982), pp. 558–574.
[12] L. Eldén, F. Berntsson, and T. Regińska, *Wavelet and Fourier methods for solving the sideways heat equation*, SIAM J. Sci. Comput., 21 (2000), pp. 2187–2205.

[12] This is true, e.g., for an FFT-based fast Poisson solver for elliptic equations.

[13] L. Eldén and V. Simoncini, *A numerical solution of a Cauchy problem for an elliptic equation by Krylov subspaces*, Inverse Problems, 25 (2009), 065002.
[14] H. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.
[15] G. H. Golub and C. F. Van Loan, *Matrix Computations,* 3rd ed., Johns Hopkins Press, Baltimore, MD, 1996.
[16] M. H. Gutknecht and T. Schmelzer, *The block grade of a block Krylov space*, Linear Algebra Appl., 430 (2009), pp. 174–185.
[17] M. Hanke, *On Lanczos based methods for the regularization of discrete ill-posed problems*, BIT, 41 (2001), pp. 1008–1018.
[18] M. Hanke and J. Nagy, *Restoration of atmospherically blurred images by symmetric indefinite conjugate gradient techniques*, Inverse Problems, 12 (1996), pp. 157–173.
[19] M. Hanke and J. G. Nagy, *Inverse Toeplitz preconditioners for ill-posed problems*, Linear Algebra Appl., 284 (1998), pp. 137–156.
[20] M. Hanke, J. Nagy, and R. Plemmons, *Preconditioned iterative regularization for ill-posed problems*, in Numerical Linear Algebra, L. Reichel, A. Ruttan, and R. S. Varga eds., de Gruyter, Berlin, 1993, pp. 141–163.
[21] P. C. Hansen, *Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems*, Numer. Algorithms, 6 (1994), pp. 1–35.
[22] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM, Philadelphia, 1998.
[23] P. C. Hansen, Regularization Tools version 4.0 for Matlab 7.3, Numer. Algorithms, 46 (2007), pp. 189–194.
[24] P. C. Hansen and T. K. Jensen, *Smoothing-norm preconditioning for regularizing minimum-residual methods*, SIAM J. Matrix Anal. and Appl., 29 (2006), pp. 1–14.
[25] K. Hayami and M. Sugihara, *A geometric view of Krylov subspace methods on singular systems*, Numer Linear Algebra Appl., 18 (2011), pp. 449–469.
[26] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd., SIAM, Philadelphia, 2002.
[27] I. Hnětynková, M. Plešinger, and Z. Strakoš, *The regularizing effect of the Golub-Kahan iterative bidiagonalization and revealing the noise level in the data*, BIT, 49 (2009), pp. 669–696.
[28] M. Ilić and I. W. Turner, *Krylov subspaces and the analytic grade*, Numer Linear Algebra Appl., 12 (2005), pp. 55–76.
[29] T. K. Jensen and P. C. Hansen, *Iterative regularization with minimum-residual methods*, BIT, 47 (2007), pp. 103–120.
[30] N. M. Nachtigal, S. C. Reddy, and L. N. Trefethen, *How fast are nonsymmetric matrix iterations?*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 778–795.
[31] J. G. Nagy and K. M. Palmer, *Steepest descent, CG, and iterative regularization of ill-posed problems*, BIT, 43 (2003), pp. 1003–1017.
[32] J. G. Nagy, R. J. Plemmons, and T. C. Torgersen, *Iterative image restoration using approximate inverse preconditioning*, IEEE Trans. Image Process., 5 (1996), pp. 1151–1162.
[33] C. C. Paige and M. A. Saunders, *Towards a generalized singular value decomposition*, SIAM J. Numer. Anal., 18 (1981), pp. 398–405.
[34] Z. Ranjbar, *Numerical Solution of Ill-posed Cauchy Problems for Parabolic Equations*, Linköping studies in Science and Technology. Dissertations. No. 1300, Linköping University, Linköping, Sweden, 2010.
[35] Z. Ranjbar and L. Eldén, *A Preconditioned GMRES Method for Solving a Sideways Parabolic Equation in Two Space Dimensions*, Technical report LiTH-MAT-R-3, Department of Mathematics, Linkping University, Linkping, Sweden, 2010.
[36] H.-J. Reinhardt, *A numerical method for the solution of two-dimensional inverse heat conduction problems*, Internat. J. Numer. Methods Engrg., 32 (1991), pp. 363–383.
[37] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
[38] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[39] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
[40] G. Strang, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.
[41] C. F. Van Loan, *Generalizing the singular value decomposition*, SIAM J. Numer. Anal., 13 (1976), pp. 76–83.
[42] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, UK, 1965.