Electronic Transactions on Numerical Analysis.
Volume 1, pp. 89-103, December 1993.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

A CHEBYSHEV-LIKE SEMIITERATION FOR INCONSISTENT LINEAR SYSTEMS *

MARTIN HANKE† AND MARLIS HOCHBRUCK‡

Dedicated to Wilhelm Niethammer on the occasion of his 60th birthday.

Abstract. Semiiterative methods are known as a powerful tool for the iterative solution of nonsingular linear systems of equations. For singular but consistent linear systems with coefficient matrix of index one, one can still apply the methods designed for the nonsingular case. However, if the system is inconsistent, the approximations usually fail to converge. Nevertheless, it is still possible to modify classical methods like the Chebyshev semiiterative method in order to fulfill the additional convergence requirements caused by the inconsistency. These modifications may suffer from instabilities since they are based on the computation of the diverging Chebyshev iterates. In this paper we develop an alternative algorithm which allows to construct more stable approximations. This algorithm can be efficiently implemented with short recurrences. There are several reasons indicating that the new algorithm is the most natural generalization of the Chebyshev semiiteration to inconsistent linear systems.

Key words. Semiiterative methods, singular systems, Zolotarev problem, orthogonal polynomials.

AMS subject classifications. 65F10, 65F20.

1. Introduction.

It is quite well known that the discrete modeling of Neumann problems for elliptic partial differential equations (cf. Hackbusch [13]) and problems from statistics such as finding the steady state of a Markov chain (cf. Barker [1]) lead to singular systems of linear equations. In addition, from the early days of computerized tomography until today the solution of singular systems of equations has played a substantial role (cf. Natterer [18]). Finally, overdetermined systems of equations are intimately connected to the concept of frames which presently undergoes a revival due to its impact on wavelet analysis (cf. Daubechies [5]) and irregular sampling algorithms (cf. Feichtinger [9]).

With an appropriate transformation (postmultiplication, preconditioning, etc.), all these applications eventually lead to a square, singular linear system of equations

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

We concentrate on the case where the matrix A is structured in a way which favors the application of iterative methods rather than direct methods for the solution of (1.1). Here, we are mainly concerned with semiiterative methods as introduced by Varga [23].

The Chebyshev method is a semiiterative method which can be applied to consistent singular systems (1.1) when the spectrum of A is real and nonnegative:

$$\sigma(A) \subset \{0\} \cup [c-d, c+d], \qquad 0 < d < c \tag{1.2}$$

(cf. Manteuffel [17], Woźniakowski [24]). A semiiterative method can be described by its associated sequence $\{p_n\}_{n\geq 0}$ of so-called residual polynomials normalized at the

* Received October 5, 1993. Accepted for publication November 30, 1993. Communicated by M. Eiermann.
† Institut für Praktische Mathematik, Universität Karlsruhe, D–76128 Karlsruhe, Germany, (email: hanke@ipmsun1.mathematik.uni-karlsruhe.de).
‡ Institut für Angewandte Mathematik und Statistik, Universität Würzburg, Am Hubland, D–97074 Würzburg, Germany, (email: marlis@mathematik.uni-wuerzburg.de).

ETNA
Kent State University
etna@mcs.kent.edu

origin by $p_n(0) = 1$. For the Chebyshev method the residual polynomials are shifted and translated Chebyshev polynomials of the first kind. Recall the following three fundamental properties of these polynomials:

(i) they are orthogonal with respect to the equilibrium distribution on the interval $[c-d, c+d]$;
(ii) they satisfy a three-term recurrence formula;
(iii) the nth residual polynomial p_n minimizes the L^∞ norm on $[c-d, c+d]$ among all residual polynomials of degree less or equal n.

It is because of the second property that the iterates of the Chebyshev method can be computed efficiently via coupled two-term recursions. The third property guarantees that the Chebyshev method is optimal in the sense that no other semiiteration can converge faster for all problems (1.1) with spectral enclosure (1.2). Note that both of these properties are somehow connected to the first one; in particular, the second property is a well-known consequence of the orthogonality.

For various reasons, e. g. approximation errors, discretization errors, or measuring errors, the final system (1.1) can be inconsistent even if the underlying physical model predicts solvability. In this case, a generalized solution of the discretized equations is sought. However, for inconsistent problems the Chebyshev method will fail to converge [24]. A theory for semiiterative methods for general singular linear systems has been developed by Eiermann, Marek, and Niethammer [6]. In the particular case where the index of A is one, several authors, e. g. in [24, 7, 15], suggested modifications of the Chebyshev algorithm which guarantee convergence of the iterates to the so called group inverse solution (cf. Campbell and Meyer [3]). These modifications maintain Property (ii) but they may fail to be numerically stable as they are based on the original (diverging) Chebyshev method. Further on, the corresponding residual polynomials no longer have an optimality property similar to (iii). On the other hand, one could try to impose some analog of (iii): Eiermann and Starke [8], for example, constructed residual polynomials that are “near optimal” in an L^∞ sense, but cannot be computed by means of short recurrences.

Here we propose a different approach of constructing semiiterative methods for inconsistent systems which is based on orthogonality, i.e., an appropriate analog of Property (i). It turns out, that the resulting residual polynomials are also near optimal in the sense of [8], *and* that the iterates can be computed efficiently via short recurrences. In other words, this new method, especially designed for inconsistent problems, shares all main features (i), (ii), and (iii) of the Chebyshev method.

We would like to stress that conjugate gradient type methods for inconsistent linear systems have been considered in the Hermitian case only, cf. Paige and Saunders [20]. For consistent but non-Hermitian linear systems convergence results for Krylov subspace methods can be found in [11]. However, conjugate gradient type methods require the computation of several inner products per iteration. On modern computer architectures this may be a severe algorithmic disadvantage (see [16] for a discussion of this subject including experimental illustrations). Like the classical Chebyshev method, our new method requires no inner products at all which may favor its application on modern supercomputers.

This paper is organized as follows: after a brief review of semiiterative methods for singular systems in Section 2, we derive a proper analog of the above Property (i) for inconsistent problems in Section 3. In Section 4 we develop simple recursions for the resulting residual polynomials and an efficient implementation of the corresponding semiiterative method. The asymptotic properties of this algorithm are studied in

Section 5. Finally, in Section 6, we present a basic numerical example illustrating the theoretical results.

2. Semiiterative methods. In this section, we briefly review semiiterative methods for inconsistent problems as developed by Eiermann, Marek, and Niethammer [6]. The nth iterate x_n of a semiiterative method lives in the shifted nth Krylov subspace

$$x_0 + \mathcal{K}_n(A, r_0) := x_0 + \operatorname{span}\{r_0, Ar_0, \ldots, A^{n-1}r_0\}.$$

Here, x_0 is any initial guess and $r_0 = b - Ax_0$ denotes the corresponding residual. Expanding x_n with respect to the Krylov basis we can find a polynomial q_{n-1} of degree $n-1$ with

$$x_n = x_0 + q_{n-1}(A)r_0 = p_n(A)x_0 + q_{n-1}(A)b, \tag{2.1}$$

where

$$p_n(\lambda) = 1 - \lambda q_{n-1}(\lambda) \tag{2.2}$$

is the nth residual polynomial.

Let $\mathcal{R}(A)$ and $\mathcal{N}(A)$ denote the range and null space of A, respectively, and let

$$b = b_{\mathcal{R}} + b_{\mathcal{N}}, \qquad b_{\mathcal{R}} \in \mathcal{R}(A), \quad b_{\mathcal{N}} \in \mathcal{N}(A)$$

be the decomposition of b into its “solvable” and “unsolvable” components. Such a decomposition always exists, and it is unique if the index of A equals one, i. e., if $\mathcal{N}(A) = \mathcal{N}(A^2)$, which we will assume from now on. We shall write P for the corresponding projector onto $\mathcal{N}(A)$; hence we have $Pb = b_{\mathcal{N}}$. Finally, we denote by $x = x(x_0)$ the unique solution of $Ax = b_{\mathcal{R}}$ with $x - x_0 \in \mathcal{R}(A)$. The corresponding (linear) map $b \mapsto x(x_0) - Px_0$ is called the group inverse of A; if A is Hermitian then $x(x_0)$ is the least squares solution of (1.1) closest to x_0 in norm (cf. Campbell and Meyer [3]).

For the error $e_n = x - x_n$ we obtain from (2.1):

$$\begin{aligned} e_n &= x - p_n(A)x_0 - q_{n-1}(A)(Ax + b_{\mathcal{N}}) \\ &= p_n(A)(x - x_0) - q_{n-1}(0)b_{\mathcal{N}}. \end{aligned} \tag{2.3}$$

Hence, if $b_{\mathcal{N}} \neq 0$, that is, if the system (1.1) is inconsistent, then we observe that $x_n \to x(x_0)$ as $n \to \infty$ for any initial guess x_0, if and only if

$$q_{n-1}(0) \to 0 \qquad \text{and} \qquad p_n(A)v \to 0 \quad \text{for all } v \in \mathcal{R}(A).$$

Note that, in view of (2.2), $q_{n-1}(0) = -p_n'(0)$. In other words, $x_n \to x(x_0)$ for any x_0 if and only if

$$p_n'(0) \to 0 \qquad \text{and} \qquad p_n(A)v \to 0 \quad \text{for all } v \in \mathcal{R}(A). \tag{2.4}$$

It is especially the first condition that will cause difficulties since $p_n'(0)$ usually diverges to infinity.

Example 1. Given the information (1.2), the residual polynomials t_n of the Chebyshev method are shifted and translated Chebyshev polynomials of the first kind, i. e.

$$t_n(\lambda) = \frac{T_n(z(\lambda))}{T_n(z(0))}, \qquad z(\lambda) = (c - \lambda)/d,$$

where

$$T_n(z) = \begin{cases} \cos(n \arccos(z)), & z \in [-1,1], \\ \cosh(n \mathrm{Arcosh}(z)), & z \notin [-1,1]. \end{cases} \tag{2.5}$$

Denote by κ the root convergence factor associated with the interval $[c-d, c+d]$ (Niethammer and Varga [19]), i.e.,

$$\kappa = e^{-\mathrm{Arcosh}(c/d)} = \frac{c - \sqrt{c^2 - d^2}}{d} < 1. \tag{2.6}$$

It is now easily verified that

$$\tau_n := t_n'(0) = -\frac{1}{\sqrt{c^2 - d^2}}\, n + O(n\kappa^{2n}) \to \infty, \qquad n \to \infty, \tag{2.7}$$

and hence the Chebyshev method fails to converge for inconsistent problems. Woźniakowski [24] was the first to modify the iteration in order to overcome divergence. He proposed the residual polynomials

$$p_n^{\mathrm{I}}(\lambda) = (1 - \tau_{n-1}\lambda) t_{n-1}(\lambda), \quad n \geq 2.$$

These polynomials satisfy $(p_n^{\mathrm{I}})'(0) = 0$, hence, the first condition in (2.4) is trivially fulfilled. Note that the nth iterate of this scheme is easily obtained from the $(n-1)$st Chebyshev iterate x_{n-1} via

$$x_{n-1} - \tau_{n-1} r_{n-1}, \qquad r_{n-1} = b - A x_{n-1}.$$

However, this construction is unstable, since the Chebyshev iterates diverge to infinity in norm.

Example 2. Another modification of the Chebyshev method was suggested in [15]. In this scheme, two subsequent iterates are extrapolated to approximate

$$x \approx x_{n-1} - \frac{\tau_{n-1}}{\tau_n - \tau_{n-1}}(x_n - x_{n-1}).$$

Note that τ_n is strictly decreasing so that no division by zero can occur. The residual polynomials of this modification are

$$p_n^{\mathrm{II}}(\lambda) = -\frac{\tau_{n-1}}{\tau_n - \tau_{n-1}} t_n(\lambda) + \frac{\tau_n}{\tau_n - \tau_{n-1}} t_{n-1}(\lambda).$$

Again, the derivative of p_n^{II} at $\lambda = 0$ vanishes. The drawback of the method is the same as in Example 1: the computation is based on the diverging sequence of Chebyshev iterates. However, the numerical results in [15] indicate a slightly improved stability, see also Section 6.

We emphasize that both modifications of the Chebyshev algorithm use residual polynomials satisfying

$$p_n(0) = 1 \quad \text{and} \quad p_n'(0) = 0, \tag{2.8}$$

and, for the sake of simplicity, we will also restrict our attention to residual polynomials satisfying (2.8). Let Π_n denote the set of all real polynomials of degree at most n, and

$$\Pi_n^0 := \{p \in \Pi_n \mid p(0) = 1,\ p'(0) = 0\}.$$

As analog of Property (iii) from the introduction we now consider the following polynomial minimization problem:

$$\|p_n\|_{[c-d,c+d]} \to \min, \qquad p_n \in \Pi_n^0, \tag{2.9}$$

where $\|\cdot\|_{[c-d,c+d]}$ denotes the L^∞ norm on the interval $[c-d, c+d]$. Note that for any diagonalizable matrix A with spectrum (1.2), and any residual polynomials $p_n \in \Pi_n^0$ we immediately obtain the following bound for the iteration error (2.3):

$$\|e_n\| = \|x - x_n\| \le \|p_n(A)\| \, \|x - x_0\| \le C \|p_n\|_{[c-d,c+d]} \, \|x - x_0\| .$$

Here, C is the condition number of the matrix of eigenvectors of A; in particular, $C = 1$ when A is Hermitian, and it is easy to construct examples where the above upper bound is attained. Thus, problem (2.9) arises quite naturally in this context.

The polynomials $p_n^\star$ minimizing (2.9) are rescaled Zolotarev polynomials. Among all polynomials of the form $\lambda^n + \sigma\lambda^{n-1} + \ldots$, the nth Zolotarev polynomial is the polynomial which has minimum L^∞ norm over the interval $[c-d, c+d]$. Like $p_n^\star$, the Zolotarev polynomial is characterized by equioscillating on the given interval, hence there exists $\sigma \in \mathbb{R}$, such that the associated Zolotarev polynomial and $p_n^\star$ only differ by a scaling factor. As the derivative of $p_n^\star$ vanishes at $\lambda = 0$, i. e., outside the interval $[c-d, c+d]$, it follows that this Zolotarev polynomial can be expressed in terms of elliptic functions, cf. Carlson and Todd [4].

As was shown by Bernstein [2], $\{p_n^\star\}$ satisfies

$$\|p_n^\star\|_{[c-d,c+d]} \sim 2(\kappa^{-1} - \kappa) \; n\kappa^n, \qquad n \to \infty, \tag{2.10}$$

where $a_n \sim b_n$ means that a_n/b_n tends to one as n goes to infinity. Opposed to this, we have for the Chebyshev polynomials:

$$\|t_n\|_{[c-d,c+d]} \sim 2\,\kappa^n, \qquad n \to \infty.$$

In other words, the additional interpolation condition $p_n'(0) = 0$ is responsible for the extra factor n in (2.10). We refer to Eiermann and Starke [8] for generalizations of this result. Following [8], we call a sequence of residual polynomials $p_n \in \Pi_n^0$ *near optimal* whenever the strong asymptotics (2.10) hold for $\{p_n\}$.

We would like to stress that no short recursions are known for the optimal polynomials $p_n^\star$, nor for the polynomials considered in [8]. However, short recursions are essential for the construction of efficient semiiterative methods. On the other hand, while the methods outlined in the examples above have short recurrences, they are not near optimal as we will show in Section 5.

3. A new approach based on orthogonal polynomials.

As mentioned as Property (i) in the introduction, the Chebyshev polynomials t_n are orthogonal with respect to the inner product $\langle\cdot,\cdot\rangle$ corresponding to the equilibrium distribution on the interval $[c-d, c+d]$, i. e.,

$$\langle\varphi,\psi\rangle \;:=\; \int_{c-d}^{c+d} \varphi(\lambda)\psi(\lambda) \frac{d\lambda}{\sqrt{(d+c-\lambda)(\lambda-c+d)}}. \tag{3.1}$$

Note that the zeros of orthogonal polynomials with respect to any nonnegative weight function on $[c-d, c+d]$ are located inside this interval and hence the same fact is true for the zeros of their derivatives. Thus, such polynomials do not belong to Π_n^0 and are

ETNA
Kent State University
etna@mcs.kent.edu

therefore not suited as residual polynomials for semiiterative methods for inconsistent systems (compare also Eiermann and Reichel [7, Theorem 3.2]).

On the other hand, it is a well-known consequence of the orthogonality relation (3.1), cf. Stiefel [21], that t_n solves the minimization problem

$$|||p|||^2 := \langle p, \frac{1}{\lambda} p\rangle \ \rightarrow \ \min \tag{3.2}$$

among all polynomials of degree at most n normalized by $p(0) = 1$, and we may ask whether a similar optimality property holds for polynomials in Π_n^0.

THEOREM 3.1. *Problem* (3.2) *has a unique solution* p_n *in* Π_n^0 *which is characterized by*

$$\langle p_n, \lambda^j\rangle = 0, \qquad \textit{for} \quad j = 1, \dots, n-1. \tag{3.3}$$

Proof. Rewriting $p \in \Pi_n^0$ as $p(\lambda) = 1 - \lambda^2 u(\lambda)$, where u is a polynomial of degree $n-2$, we observe that (3.2) is equivalent to searching for the best approximation u of λ^{-2} from the set of polynomials of degree at most $n-2$ in the Hilbert space induced by the inner product $\langle \cdot, \lambda^3 \cdot\rangle$. Since this equivalent approximation problem has a unique solution, there is a unique minimizer of (3.2) in Π_n^0.

Let p_n be this minimizing polynomial, and let $1 \le j \le n-1$ be arbitrarily chosen. Consider $p = p_n + \alpha\lambda^{j+1}$, $\alpha \in \mathbb{R}$, which obviously belongs to Π_n^0. Hence,

$$|||p_n|||^2 \le |||p|||^2 = |||p_n|||^2 + 2\alpha\langle p_n, \frac{1}{\lambda}\lambda^{j+1}\rangle + \alpha^2\,|||\lambda^{j+1}|||^2.$$

Choosing the sign of α appropriately while letting $\alpha \to 0$, we conclude that this inequality holds if and only if

$$\langle p_n, \frac{1}{\lambda}\lambda^{j+1}\rangle = \langle p_n, \lambda^j\rangle = 0.$$

To proof the opposite direction, let p_n satisfy the orthogonality relations (3.3) and let p be an arbitrary polynomial in Π_n^0. Then $p - p_n$ has degree n and a zero of multiplicity two at the origin. Hence,

$$u := (p - p_n)/\lambda \in \operatorname{span}\{\lambda, \lambda^2, \dots, \lambda^{n-1}\}.$$

¿From this we conclude

$$\begin{aligned} |||p|||^2 &= |||p_n + \lambda u|||^2 \\ &= |||p_n|||^2 + 2\langle p_n, u\rangle + |||\lambda u|||^2 \\ &= |||p_n|||^2 + |||\lambda u|||^2 \\ &\ge |||p_n|||^2. \end{aligned}$$

This completes the proof. □

Let us briefly return to the examples of the previous section. For the polynomials p_n^{I} we immediately obtain

$$\langle p_n^{\mathrm{I}}, \lambda^j\rangle = \langle t_{n-1}, (1 - \tau_{n-1}\lambda)\lambda^j\rangle = 0, \qquad j = 0, \dots, n-3,$$

while for the polynomials p_n^{II} we have

$$\langle p_n^{\mathrm{II}}, \lambda^j\rangle = -\frac{\tau_{n-1}}{\tau_n - \tau_{n-1}}\langle t_n, \lambda^j\rangle + \frac{\tau_n}{\tau_n - \tau_{n-1}}\langle t_{n-1}, \lambda^j\rangle = 0, \qquad j = 0, \dots, n-2.$$

It is instructive to compare these orthogonality relations with those of Theorem 3.1. In particular, we note that the polynomials p_n^{I} lose one degree of orthogonality compared to p_n^{II}; this might indicate a certain superiority of p_n^{II}.

We now show that the iterates of the semiiterative method based on the polynomials $\{p_n\}$ of Theorem 3.1 can be computed with short recursions. For this we consider the update from x_n to x_{n+1}: from (2.1) we obtain

$$x_{n+1} - x_n = \big(q_n(A) - q_{n-1}(A)\big)r_0 =: u_n(A)r_0 \tag{3.4}$$

with the so-called update polynomials

$$u_n(\lambda) = q_n(\lambda) - q_{n-1}(\lambda) = \frac{p_n(\lambda) - p_{n+1}(\lambda)}{\lambda}. \tag{3.5}$$

Note that u_n is a polynomial of degree n with $u_n(0) = 0$. Furthermore, if p is any polynomial of degree at most $n-2$, then

$$\langle \frac{u_n}{\lambda}, \lambda^3 p \rangle = \langle \frac{p_n - p_{n+1}}{\lambda^2}, \lambda^3 p \rangle = \langle p_n - p_{n+1}, \lambda p \rangle = 0$$

by virtue of the orthogonality relation (3.3). This means that $\{u_n/\lambda\}_{n\geq 1}$ – polynomials of degree $n-1$, respectively – are classical orthogonal polynomials with respect to the real inner product $\langle \cdot, \lambda^3 \cdot \rangle$. They therefore satisfy a three-term recurrence relation; hence, after multiplication by λ we obtain

$$u_n = \omega_n \lambda u_{n-1} + \mu_n u_{n-1} + \nu_n u_{n-2}, \qquad n \geq 2, \quad \nu_2 = 0, \tag{3.6}$$

with certain uniquely defined coefficients ω_n, μ_n and ν_n, $n \geq 2$. By (3.4), this leads to the following iterative scheme to compute x_n, $n \geq 2$:

$$x_{n+1} = x_n + \omega_n A(x_n - x_{n-1}) + \mu_n(x_n - x_{n-1}) + \nu_n(x_{n-1} - x_{n-2}). \tag{3.7}$$

The coefficients $\{\omega_n, \mu_n, \nu_n\}_{n\geq 2}$ are not known explicitly, but can be obtained from the recursion coefficients of the Chebyshev polynomials $\{t_n\}$ in the course of the iteration, cf. Gautschi [12] or Fischer and Golub [10]. We will derive the corresponding formulas in the following section.

We point out that so far, we have not used the special form of the weight function in (3.1). In fact, all we need is that (3.2) defines a norm.

4. The algorithm. Following Manteuffel [17], the translated Chebyshev polynomials $\{t_n\}_{n\geq -1}$ satisfy the following recurrence relation*:

$$\begin{aligned} &t_{-1} \equiv 0, \qquad t_0 \equiv 1, \\ &t_{n+1} = -\alpha_n \lambda t_n + (1+\beta_n)t_n - \beta_n t_{n-1}, \qquad n \geq 0, \end{aligned} \tag{4.1}$$

with

$$\begin{aligned} &\alpha_0 = 1/c, &&\beta_0 = 0, \\ &\alpha_1 = 2c/(2c^2 - d^2), &&\beta_1 = c\alpha_1 - 1, \\ &\alpha_n = 1/(c - \big(\tfrac{d}{2}\big)^2 \alpha_{n-1}), &&\beta_n = c\alpha_n - 1, \qquad n > 1. \end{aligned}$$

* Note that we use c for center and d for distance, which differs from the notation used in [17].

We will now see how this can be utilized for a computation of p_n:

LEMMA 4.1. *Let* $\tau_n = t_n'(0)$ *and* $\sigma_n = t_n''(0)$. *Then the polynomials* p_n *of Theorem* 3.1 *can be expressed as*

$$p_n = \frac{1}{\lambda}\left(\gamma_n t_{n+1} - (\gamma_n - \delta_n)t_n - \delta_n t_{n-1}\right), \qquad n \geq 0, \tag{4.2}$$

where $\gamma_0 = -c$, $\delta_0 = 0$, *and*

$$\gamma_n = (\sigma_n - \sigma_{n-1})/\rho_n, \quad \delta_n = (\sigma_n - \sigma_{n+1})/\rho_n, \qquad n \geq 1, \tag{4.3}$$

with

$$\rho_n = (\tau_{n+1} - \tau_n)(\sigma_n - \sigma_{n-1}) - (\tau_n - \tau_{n-1})(\sigma_{n+1} - \sigma_n).$$

Proof. We expand λp_n in terms of the Chebyshev polynomials $\{t_n\}$:

$$\lambda p_n = \sum_{j=0}^{n+1} \pi_{j,n} t_j. \tag{4.4}$$

Because of (3.3),

$$\langle \lambda p_n, t_j \rangle = \langle p_n, t_j \lambda \rangle = 0 \qquad \text{for} \quad 0 \leq j \leq n-2.$$

Since the left-hand side is a positive multiple of $\pi_{j,n}$ we conclude that only t_{n-1}, t_n, and t_{n+1} contribute to λp_n in (4.4). With $\gamma_n := \pi_{n+1,n}$ and $\delta_n := -\pi_{n-1,n}$ we find $\pi_{n,n} = \delta_n - \gamma_n$ since λp_n vanishes at the origin; hence we obtain (4.2). The values of γ_n and δ_n can be determined from the first two derivatives of λp_n at $\lambda = 0$:

$$\begin{bmatrix} (\lambda p_n)'(0) \\ (\lambda p_n)''(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \tau_{n+1} - \tau_n & \tau_n - \tau_{n-1} \\ \sigma_{n+1} - \sigma_n & \sigma_n - \sigma_{n-1} \end{bmatrix} \begin{bmatrix} \gamma_n \\ \delta_n \end{bmatrix}.$$

This yields (4.3). Note that ρ_n is the determinant of the matrix on the right-hand side. It must be nonzero, since p_n is uniquely determined by Theorem 3.1. ▯

The coefficients τ_n and σ_n are easily obtained from (4.1), namely we have

$$\begin{aligned} &\tau_0 = 0, \quad \tau_1 = -\alpha_0, \\ &\tau_{n+1} = -\alpha_n + (1+\beta_n)\tau_n - \beta_n \tau_{n-1}, \qquad n \geq 1, \end{aligned}$$

$$\begin{aligned} &\sigma_0 = 0, \quad \sigma_1 = 0, \\ &\sigma_{n+1} = -2\alpha_n \tau_n + (1+\beta_n)\sigma_n - \beta_n \sigma_{n-1}, \qquad n \geq 1. \end{aligned}$$

Now we are in a position to determine the recursion coefficients ω_n, μ_n and ν_n in (3.6). Using (3.5) and Lemma 4.1 we first rewrite the update polynomials in terms of Chebyshev polynomials:

$$u_n = \frac{1}{\lambda^2}\left(-\gamma_{n+1}t_{n+2} + (\gamma_n + \gamma_{n+1} - \delta_{n+1})t_{n+1} + (\delta_{n+1} + \delta_n - \gamma_n)t_n - \delta_n t_{n-1}\right),$$

valid for $n \geq 0$; consequently, when $n \geq 2$, similar expansions hold for u_n, u_{n-1} and u_{n-2}. Then, inserting these expansions into (3.6) and replacing the resulting terms λt_j according to (4.1), namely

$$\lambda t_j = -\frac{1}{\alpha_j}\,t_{j+1} + \frac{1+\beta_j}{\alpha_j}\,t_j - \frac{\beta_j}{\alpha_j}\,t_{j-1}, \qquad n-2 \leq j \leq n+1,$$

this yields an identity

$$\frac{1}{\lambda^2}\sum_{j=n-3}^{n+2}\eta_{j,n}t_j \equiv 0, \qquad n \ge 2;$$

the coefficients $\eta_{j,n}$ in there must equal zero, which gives us a set of linear equations for the unknown recursion coefficients. In particular, setting $\eta_{j,n}=0$ for $j=n+2,n+1$ and $j=n-3$ we eventually obtain

$$\begin{aligned}
\omega_n &= -\alpha_{n+1}\,\frac{\gamma_{n+1}}{\gamma_n}\ ,\\
\mu_n &= \frac{1}{\gamma_n}\left(\delta_{n+1}-\gamma_n+\gamma_{n+1}\big(\beta_{n+1}+\frac{\alpha_{n+1}}{\alpha_n}\big)+(\delta_n-\gamma_{n-1})\frac{\omega_n}{\alpha_n}\right),\\
\nu_n &= \frac{\omega_n}{\alpha_{n-2}}\frac{\delta_{n-1}}{\delta_{n-2}}\beta_{n-2}.
\end{aligned}$$

Since $\sigma_n > \sigma_{n-1}$ for $n \ge 2$, this implies $\gamma_n > 0$ for $n \ge 2$ and $\delta_n < 0$ for $n \ge 1$, cf. (4.3). Further on, since t_{n+1} has exact degree n, we have $\alpha_n \ne 0$ for every $n \ge 0$. It follows that ω_n and μ_n are well defined for $n \ge 2$ and ν_n is well defined for $n \ge 3$; recall that we have set $\nu_2 = 0$ in (3.6).

It remains to determine x_1 and x_2 for a correct initialization of (3.7). Obviously, there is only one polynomial in Π_1^0, namely $p_0 \equiv p_1 \equiv 1$. Hence,

$$x_1 = x_0.$$

Moreover, from $p_2 \in \Pi_2^0$ we conclude

$$p_2(\lambda) = 1 - \varrho\lambda^2,$$

and we may determine ϱ from $\langle p_2, \lambda\rangle = 0$, cf. Theorem 3.1; elementary integration yields

$$\varrho = \frac{\langle 1,\lambda\rangle}{\langle 1,\lambda^3\rangle} = \frac{2}{2c^2+3d^2}.$$

By virtue of (2.1) we therefore find

$$x_2 = x_0 + \varrho A(b - Ax_0).$$

We stress that the computation of x_2 is the only part of the entire algorithm where the right-hand side b of (1.1) is used. The complete algorithm is summarized in Algorithm 4.1: for the actual implementation we updated the differences $\tau_n - \tau_{n-1}$ and $\sigma_n - \sigma_{n-1}$, rather than τ_n and σ_n themselves.

5. Near optimal asymptotic behavior. The aim of the following investigations is to show that (2.10) holds for the polynomials p_n of Theorem 3.1. Recall that

$$t_n(\lambda) = \frac{T_n(z(\lambda))}{T_n(z(0))}, \qquad z(\lambda) = (c-\lambda)/d,$$

where $T_n(z)$ are the usual Chebyshev polynomials of the first kind. Let $\Re$ denote the real part of a complex number and i the imaginary unit. Then, for $\lambda \in [c-d, c+d]$ we have $z(\lambda) \in [-1,1]$ and thus, by (2.5),

$$T_n(z(\lambda)) = \cos(n \arccos z(\lambda)) = \Re w(\lambda)^n \tag{5.1}$$

ETNA
Kent State University
etna@mcs.kent.edu


```
     /* let  sigd(n) = sig(n)-sig(n-1)   taud(n) = tau(n)-tau(n-1)  */
alp(1) = 2c/(2c*c-d*d);              bet(1) = c*alp(1)-1;
alp(2) = 1/(c-d*d*alp(1)/4);         bet(2) = c*alp(2)-1;
alp(3) = 1/(c-d*d*alp(2)/4);         bet(3) = c*alp(3)-1;
tau = -2alp(1);                      taud(2) = tau+1/c;
sigd(2) = 2/c*alp(1);                sigd(3) = -2alp(2)tau + bet(2)sigd(2);
taud(3) = -alp(2) + bet(2)taud(2);   tau = tau + taud(3);
gam(1) = 0;                          del(1) = -c;
rho = taud(3)sigd(2) - taud(2)sigd(3);
gam(2) = sigd(2)/rho;                del(2) = -sigd(3)/rho;
nu = 0;

                     /* let  xd(n) = x(n)-x(n-1)  */
xd(1) = 0;                           xd(2) = 2/(2c*c+3d*d) A*(b-A*x(0));
x  = x0 + xd(2);

for n=3 until ...  do
    sigd(n+1) = -2alp(n)tau + bet(n)sigd(n);
    taud(n+1) = -alp(n) + bet(n)taud(n);
    tau = tau + taud(n+1);
    rho = taud(n+1)sigd(n) - taud(n)sigd(n+1);
    gam(n) = sigd(n)/rho;
    del(n) = -sigd(n+1)/rho;
    om = -alp(n)gam(n)/gam(n-1);
    mu = ( del(n) - gam(n-1) + gam(n)(bet(n)+alp(n)/alp(n-1))
                  + (del(n-1)-gam(n-2))om/alp(n-1) )/gam(n-1);
    if (n > 3)
        nu = om*del(n-2)bet(n-3)/(alp(n-3)del(n-3));
    end if;
    alp(n+1) = 1/(c-d*d*alp(n)/4);
    bet(n+1) = c*alp(n+1)-1;

    xd(n) = om A*xd(n-1) + mu xd(n-1) + nu xd(n-2);
    x = x + xd(n);
end for;
```

ALGORITHM 4.1. *Chebyshev-like algorithm for inconsistent problems*

with

$$w(\lambda) = e^{i \arccos z(\lambda)}, \qquad |w(\lambda)| = 1.$$

Further on, when $\lambda = 0$ then $z(0) = c/d > 1$; from (2.5), with the root convergence factor κ defined in (2.6), we thus obtain

$$T_n(z(0)) = \frac{1}{2}\left(e^{n\mathrm{Arcosh}(c/d)} + e^{-n\mathrm{Arcosh}(c/d)}\right) \sim \frac{1}{2}\,\kappa^{-n}, \qquad n \to \infty. \tag{5.2}$$

Consequently, (5.1) and (5.2) together yield the following asymptotics for the residual polynomials t_n of the Chebyshev method:

$$t_n(\lambda) \sim 2\kappa^n\,\Re w(\lambda)^n, \qquad n \to \infty. \tag{5.3}$$

For later use, we also mention two useful identities which readily follow from (2.6):

$$\kappa^{-1} + \kappa = 2\,\frac{c}{d}, \qquad \kappa^{-1} - \kappa = 2\frac{\sqrt{c^2 - d^2}}{d}. \tag{5.4}$$

We now turn to an analysis of $\{p_n\}$ based on the representation (4.2). First we state the asymptotic behavior of the corresponding coefficients γ_n and δ_n.

LEMMA 5.1. *Let γ_n and δ_n be defined as in Lemma* 4.1. *Then we have*

$$\begin{aligned} \gamma_n &= n\sqrt{c^2-d^2}+O(1), \\ \delta_n &= -n\sqrt{c^2-d^2}+O(1), \end{aligned} \qquad n\to\infty.$$

Proof. Here we only sketch the main steps of the proof because of its many tedious calculations. Using the explicit representation (2.5) of the Chebyshev polynomials (and (5.2)), one obtains the asymptotics (2.7) for τ_n, and similarly,

$$\sigma_n = t_n''(0) = \frac{1}{c^2-d^2}\,n^2 \;-\; \frac{c}{(c^2-d^2)^{3/2}}\,n \;+\; O(n\kappa^{2n}), \qquad n\to\infty.$$

Now we can evaluate ρ_n defined in Lemma 4.1:

$$\rho_n = \frac{2}{(c^2-d^2)^{3/2}} + O(n^2\kappa^{2n}), \qquad n\to\infty.$$

Inserting these asymptotics into (4.3) completes the proof. ▯ Combining Lemma 5.1, Lemma 4.1, and (5.3) yields the dominating term in the asymptotic expansion of p_n for $\lambda \in [c-d, c+d]$:

$$p_n(\lambda) \sim \frac{2}{\lambda}\sqrt{c^2-d^2}\;\Re\big(\frac{w^{n-1}}{\kappa}(\kappa w-1)^2\big)\;n\kappa^n, \qquad n\to\infty. \tag{5.5}$$

Here, as throughout the following manipulations, $w = w(\lambda)$, and we have

$$\lambda = c - dz(\lambda) = c - d\,\Re w.$$

Hence, using (5.4) and keeping in mind that $|w(\lambda)| = 1$ we obtain

$$\lambda = \frac{d}{2}\,(\kappa^{-1}+\kappa-w-\overline{w}) = \frac{d}{2}\,\kappa^{-1}\big(\kappa w-1\big)\big(\kappa\overline{w}-1\big).$$

Inserting this into (5.5) and using (5.4) we conclude, as $n\to\infty$,

$$\begin{aligned} p_n(\lambda) &\sim 4\,\frac{\sqrt{c^2-d^2}}{d}\;\Re\left(w^{n-1}\frac{(\kappa w-1)^2}{(\kappa w-1)(\kappa\overline{w}-1)}\right)\,n\kappa^n \\ &= 2(\kappa^{-1}-\kappa)\;\Re\left(w^{n-1}\frac{\kappa w-1}{\kappa\overline{w}-1}\right)\,n\kappa^n. \end{aligned}$$

Note that (5.5) holds uniformly for $\lambda\in[c-d,c+d]$ and hence

$$\lim_{n\to\infty}\frac{1}{n}\,\kappa^{-n}p_n(\lambda) = 2(\kappa^{-1}-\kappa)\;\Re\left(w^{n-1}\frac{\kappa w-1}{\kappa\overline{w}-1}\right),$$

uniformly for $\lambda\in[c-d,c+d]$. Taking absolute values therefore yields

$$\frac{1}{n}\,\kappa^{-n}\,\|p_n\|_{[c-d,c+d]} \le 2(\kappa^{-1}-\kappa)+o(1), \qquad n\to\infty.$$

Since Bernstein's result (2.10) constitutes a lower bound for $\|p_n\|_{[c-d,c+d]}$ we have actually shown

ETNA
Kent State University
etna@mcs.kent.edu

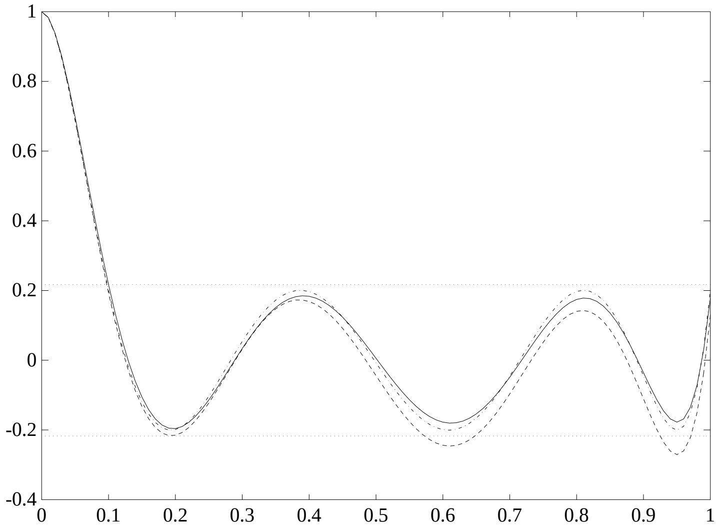

FIG. 5.1. *Optimal and near optimal polynomials of degree* 6

THEOREM 5.2. *The polynomials p_n of Theorem 3.1 are near-optimal, i.e.,*

$$\|p_n\|_{[c-d,c+d]} \sim 2(\kappa^{-1} - \kappa)\, n\kappa^n, \qquad n \to \infty.$$

In Figure 5.1 we show the polynomial p_6 (solid line), when $[c-d, c+d] = [0.1, 1]$. We compare p_6 with the near optimal polynomial (dashed line) constructed by Eiermann and Starke [8], and with the optimal polynomial $p_6^\star$ (dashdotted line) which solves (2.9) and which we computed with a weighted Remez algorithm. The horizontal dotted lines indicate the L^∞ norm of p_6 over $[0.1, 1]$ which is attained when $\lambda = 0.1$. It can be seen that the polynomials are close together which means that the asymptotics describe the behavior of the polynomials reasonably well already for small n.

In Figure 5.2 we compare p_6 (solid line) with the polynomials p_6^{I} (dashed line) and p_6^{II} (dashdotted line) from the examples in Section 2. Here the differences are significant. This can also be established theoretically: both polynomials, p_n^{I} and p_n^{II}, for every $n \in \mathbb{N}$, attain their maximum absolute value over $[c-d, c+d]$ at $\lambda = c+d$. This yields

$$\|p_n^{\mathrm{I}}\|_{[c-d,c+d]} = \big(1 - \tau_{n-1}(c+d)\big)\, \|t_{n-1}\|_{[c-d,c+d]} \sim \frac{2}{\kappa}\left(\frac{c+d}{c-d}\right)^{1/2} n\kappa^n$$

and

$$\|p_n^{\mathrm{II}}\|_{[c-d,c+d]} \sim n(\|t_n\|_{[c-d,c+d]} + \|t_{n-1}\|_{[c-d,c+d]}) \sim 2(1+\kappa^{-1})\, n\kappa^n.$$

In both cases, the factors in front of $n\kappa^n$ are bigger than $2\kappa^{-1}$ which in turn is larger than the corresponding factor in Theorem 5.2. Moreover, when d approaches c, i. e., when the problem gets more ill conditioned, then the factor in Theorem 5.2 tends to

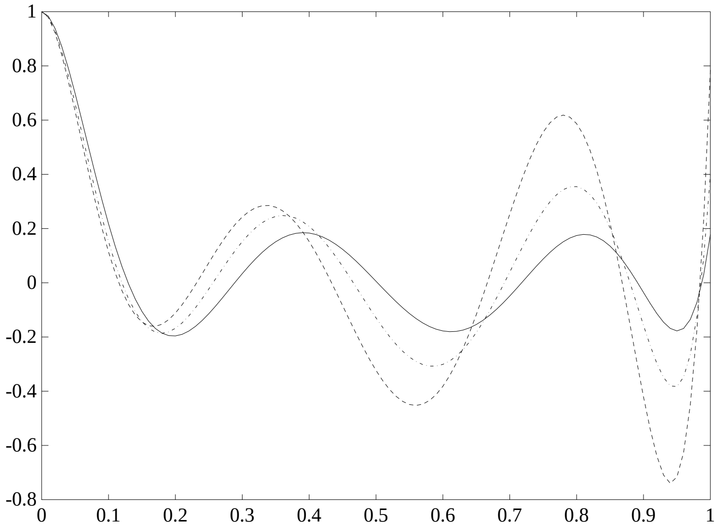

FIG. 5.2. *Polynomials of degree* 6 *from the examples in Section* 2 *and from the new approach*

zero, while the factor corresponding to p_n^{II} tends to four and the one corresponding to p_n^{I} goes to infinity.

6. A numerical example. We tested the numerical properties of the new method for a simple model problem taken from [15]. In this example we consider the solution of the Poisson equation with Neumann boundary conditions on the unit square. On an equidistant grid with mesh size h we discretize the Laplace operator and the boundary conditions with central differences, cf. Hackbusch [13, Chapter 4.7.2]. The grid points are arranged in the red-black ordering. In this way, we end up with a non-Hermitian matrix M with 'Property A'. We have chosen this (somewhat academic) discretization, because we can easily compute the (real) eigenvalues of the associated matrices. Note that M is singular with a one dimensional null space spanned by the vector $e = [1 \ \cdots \ 1]^T$. Even if the continuous problem has a solution, the discretized problem need not be consistent, cf. [13, Remark 4.7.10]. Recall that the semiiterative methods described in this paper do not require the matrix to be Hermitian.

¿From Young's SOR theory (see for example [23], and Hadjidimos [14] for an analysis of the singular case) we can determine the optimal SOR parameter but it is known that the Gauss-Seidel method with appropriate semiiterative acceleration yields the same convergence rate, cf. Varga [22]. In fact, it can be shown that the Gauss-Seidel preconditioned coefficient matrix $I-\mathcal{L}_1$ has a nonnegative, real spectrum contained in $\{0\} \cup [\gamma^2, 1]$, where $\gamma = (1-\cos(\pi h))/2$ and $1-\gamma$ is the subdominant eigenvalue of the corresponding Jacobi operator $\mathcal{J} = I - M/4$. As was shown in [14], the matrix $I - \mathcal{L}_1$ has index 1; therefore, the Gauss-Seidel preconditioned problem satisfies the requirements imposed in Section 1. For our computations we choose $h = 1/63$ which yields a coefficient matrix of order 4096. Moreover, we have $\gamma = 6.2 \cdot 10^{-4}$ and a convergence factor $\kappa = 0.9319$ of the semiiterative methods.

In order to compute the relative errors of the approximations, we first construct

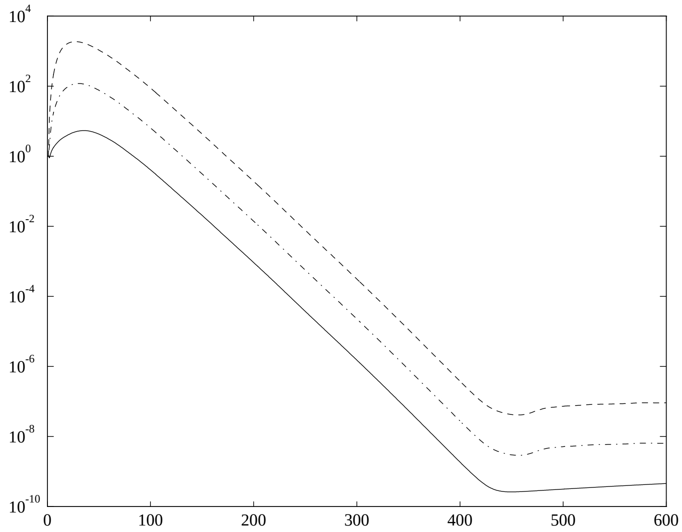

FIG. 6.1. *Relative errors for the two methods from Section* 2 *and for the new algorithm*

a consistent problem with known solution $x \in \mathcal{R}(I - \mathcal{L}_1)$, namely $x = (I - \mathcal{L}_1)y$, where y is a normally distributed random vector. Then we perturb the right-hand side (of the preconditioned problem) with a constant multiple of the null space vector e. In this way we end up with an inconsistent problem with group inverse solution x. For this particular example our perturbation amounts to one percent in norm, i.e., $\|b_{\mathcal{N}}\| / \|b_{\mathcal{R}}\| = 0.01$. The initial vector is always the zero vector. All our computations have been performed in MATLAB 4.0.

Figure 6.1 shows the relative iteration errors of the new method (solid line) and the two methods introduced in the examples in Section 2, namely the dashed line corresponds to Example 1 and the dashdotted line corresponds to Example 2. As expected from the asymptotic analysis and from the graphs of the residual polynomials in Figure 5.2, the new method performs best, followed by the method from Example 2. All curves show exactly the same slope which means that only the different factors in the asymptotics eventually determine the superiority of the new method. The stagnation of the error after about 430 iterations is due to accumulated round-off components of the iterates in the null space of M (see [15] for a further discussion of this topic). We have run the iteration up to this point to demonstrate the stability of the three algorithms. As can be seen from the graphs in Figure 6.1, the new method is not only faster but also achieves higher accuracy. In this application this is not important in view of the discretization error, but it is definitely another advantage of the new method that may pay off in other applications.

REFERENCES

[1] V. A. BARKER, *Numerical solution of sparse singular systems of equations arising from ergodic Markov chains*, Comm. Statist. Stochastic Models 5 (1989), pp. 335–381.
[2] S. BERNSTEIN, *Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle*, Gauthier-Villars, Paris, 1926. (Reprint: Chelsea, New

York, 1970.)

[3] S. L. Campbell and C. D. Meyer, *Generalized Inverses of Linear Transformations*, Pitman, London, San Francisco, Melbourne, 1979.

[4] B. C. Carlson and J. Todd, *Zolotarev's first problem – the best approximation by polynomials of degree* $\leq n-2$ *to* $x^n - n\sigma x^{n-1}$ *in* $[-1,1]$, Aequat. Math. 26 (1983), 1–33.

[5] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.

[6] M. Eiermann, I. Marek, and W. Niethammer, *On the solution of singular linear systems of algebraic equations by semiiterative methods*, Numer. Math. 53 (1988), 265–283.

[7] M. Eiermann and L. Reichel, *On the application of orthogonal polynomials to the iterative solution of singular linear systems of equations*, in J. Dongarra, I. Duff, P. Gaffney, and S. McKee, Editors: Vector and Parallel Computing, Prentice-Hall, Englewood Cliffs, NJ, 1989, pp. 285–297.

[8] M. Eiermann and G. Starke, *The near-best solution of a polynomial minimization problem by the Carathéodory-Fejér method*, Constr. Approx. 6 (1990), pp. 303–319.

[9] H. G. Feichtinger, *Pseudo-inverse matrix methods for signal reconstruction from partial data*, SPIE-Conf. Visual Comm. and Image Processing, Int. Soc. Opt. Eng., Boston, 1991, pp. 766-772.

[10] B. Fischer and G. H. Golub, *How to generate unknown orthogonal polynomials out of known orthogonal polynomials*, J. Comp. Appl. Math. 43 (1992), pp. 99–115.

[11] R. W. Freund and M. Hochbruck, *On the use of two QMR algorithms for solving singular systems and applications in Markov chain modeling*, RIACS Technical Report 91.25, December 1991. To appear in: Journal of Numerical Linear Algebra with Applications.

[12] W. Gautschi, *An algorithmic implementation of the generalized Christoffel theorem*, in Numerical Intergration, G. Hämmerlin, Editor, Birkhäuser, Basel, 1982, pp. 89–106.

[13] W. Hackbusch, *Elliptic Differential Equations, Theory and Numerical Treatment*, Springer–Verlag, Berlin, Heidelberg, New York, 1992.

[14] A. Hadjidimos, *On the optimization of the classical iterative schemes for the solution of complex singular linear systems*, SIAM J. Alg. Disc. Meth. 6 (1985), 555–566.

[15] M. Hanke, *Iterative Lösung gewichteter linearer Ausgleichsprobleme*, Dissertation, Universität Karlsruhe, 1989.

[16] M. Hanke, M. Hochbruck, and W. Niethammer, *Experiments with Krylov subspace methods on a massively parallel computer*, IPS Research Report No. 92-16, IPS, ETH Zürich (1992). To appear in: Appl. Math.

[17] T. Manteuffel, *The Tchebychev iteration for nonsymmetric linear systems*, Numer. Math. 28 (1977), pp. 307–327.

[18] F. Natterer, *The Mathematics of Computerized Tomography*, John Wiley & Sons, Chichester, New York, 1986.

[19] W. Niethammer and R. S. Varga, *The analysis of* k*-step iterative methods for linear systems from summability theory*, Numer. Math. 41 (1983), pp. 177–206.

[20] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal. 12 (1975), pp. 617–629.

[21] E. Stiefel, *Relaxationsmethoden bester Strategie zur Lösung linearer Gleichungssyteme*, Comm. Math. Helv. 29 (1955), pp. 157–179.

[22] R. S. Varga, *A comparison of the succesive overrelaxation method and semi-iterative methods using Chebyshev polynomials*, J. Soc. Indust. Appl. Math. 5 (1957), pp. 39–46.

[23] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.

[24] H. Woźniakowski, *Numerical stability of the Chebyshev method for the solution of large linear systems*, Numer. Math. 28 (1977), pp. 191–209.