Electronic Transactions on Numerical Analysis.
Volume 14, pp. 1-19, 2002.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

SUPERLINEAR CG CONVERGENCE FOR SPECIAL RIGHT-HAND SIDES*

BERNHARD BECKERMANN† AND ARNO B. J. KUIJLAARS‡

Abstract. Recently, we gave a theoretical explanation for superlinear convergence behavior observed while solving large symmetric systems of equations using the Conjugate Gradient method. Roughly speaking, one may observe superlinear convergence while solving a sequence of (symmetric positive definite) linear systems if the asymptotic eigenvalue distribution of the sequence of the corresponding matrices of coefficients is far from an equilibrium distribution.

However, it is well known that the convergence of the Conjugate Gradient or other Krylov subspace methods does not only depend on the spectrum but also on the right-hand side of the underlying system and the starting vector. In this paper we present a family of examples based on the discretization via finite differences of the one dimensional Poisson problem where the asymptotic distribution equals an equilibrium distribution but one may as well observe superlinear convergence according to the particular choice of the right-hand sides.

Our findings are related to some recent results concerning asymptotics of discrete orthogonal polynomials. An important tool in our investigations is a constrained energy problem in logarithmic potential theory, where an additional external field is used being related to our particular right-hand sides.

Key words. Superlinear convergence, Conjugate gradients, Krylov subspace methods, Logarithmic potential theory.

AMS subject classifications. 65F10, 65E05, 31A99, 41A10.

1. Introduction.

The Conjugate Gradient (CG) method is widely used for solving systems of linear equations $Ax = b$ with a positive definite symmetric matrix A of order N. The CG method is popular as an iterative method for large systems, stemming e.g. from the discretization of boundary value problems for elliptic PDEs. The rate of convergence of CG depends on the distribution of the eigenvalues of A. A well-known upper bound for the error e_n in the A-norm after n steps is

$$\frac{\|e_n\|_A}{\|e_0\|_A} \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n \tag{1.1}$$

where e_0 is the initial error and the condition number $\kappa = \lambda_{\max}/\lambda_{\min}$ is the ratio of the two extreme eigenvalues of A. In practical situations, this bound is often too pessimistic, and one observes an increase in the convergence rate as n increases. This phenomenon is known as superlinear convergence of the CG method.

Once one knows more about the spectrum of A, one may exploit the link to some polynomial extremal problem (see (2.12) below) and obtain sharper error estimates using techniques from approximation theory; see for instance [13, 16, 17, 35, 36] and the references therein. Indeed, according to a well-known observation in numerical linear algebra, superlinear convergence behavior means that some of the Ritz values approach very well some eigenvalues; see for instance [33, 15, 32, 13]. However, it seems that an analytic description of this superlinear convergence behavior for matrices has only been given recently: in [5] (see also ([6]), the authors take the point of view that there is a sequence of systems $A_N x = b_N$ to be solved, A_N of order N, with the symmetric positive definite matrices A_N having an asymptotic eigenvalue distribution. More precisely, denoting by $\Lambda(A_N)$ the spectrum of A_N, we

*Received November 2, 2000. Accepted for publication March 7, 2001. Communicated by Sven Ehrich.

†Laboratoire d'Analyse Numérique et d'Optimisation, UFR IEEA – M3, UST Lille, F-59655 Villeneuve d'Ascq CEDEX, France, e-mail: bbecker@ano.univ-lille1.fr. Supported in part by INTAS project 2000-272

‡Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200 B, B-3001 Leuven, Belgium, e-mail: arno@wis.kuleuven.ac.be. This author's research is supported in part by FWO research project G.0278.97 and by a research grant of the Fund for Scientific Research – Flanders

suppose that there is a probability measure σ such that, for every function $f \in \mathcal{C}(\mathbb{R})$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\lambda\in\Lambda(A_N)} f(\lambda) = \int f(\lambda)\, d\sigma(\lambda). \tag{1.2}$$

This assumption is natural in a number of applications such as the solution of (block) Toeplitz systems (see, e.g., [10, 8] and [5, §4]) or the discretization of elliptic PDEs via finite difference techniques (see, e.g., [30, 31, 34] and the discussions in [5, §5] or [4]).

Under some additional weak regularity assumptions stated in §2 below, the following estimate for the error in Conjugate Gradients[1] after n iterations for the system $A_N x = b_N$ was established in [5, Theorem 2.1]

$$\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}} \lessapprox \exp\Big(-N \cdot \int_0^{n/N} g_{S(\tau)}(0)\, d\tau\Big), \tag{1.3}$$

where $g_S(z)$ is the Green function with pole at ∞ for the complement of the compact set S. Here $\lessapprox$ means that $\leq$ holds if we take Nth roots on both sides and let n, N tend to infinity in such a way that n/N tends to a limit $t \in (0,1)$. The set $S(t)$ equals the free part of the extremal measure in a constrained energy problem connected to σ. Since this set decreases as t increases, the logarithm of the right-hand side of (1.3) describes some concave curve for fixed N and for increasing n, leading to a CG convergence estimate which corresponds to superlinear convergence.

We should mention that our estimate is only proved to be an upper bound after taking roots and passing to the limit. However, in all examples studied so far (see [5, 4] and below) the right-hand side of (1.3) is actually an upper bound (as long as the effects of finite precision arithmetic can be neglected). Moreover, if we distinguish between the ranges of sublinear, linear and superlinear convergence (see, e.g., Nevanlinna [25]), then one observes quite often (at least for random right-hand sides and random starting vectors) that, in the ranges of linear and superlinear convergence, the CG error curve coincides up to some constant with the curve corresponding to our estimate.

Besides by the distribution of eigenvalues, the CG convergence rate is also governed by the components of the initial residual $r_{0,N} = b_N - A_N x_{0,N}$ in the directions of the eigenvectors of A_N, where $x_{0,N}$ is a starting vector. We choose here $x_{0,N} = 0$, so that $r_{0,N} = b_N$. Then, if the eigenvector components of b_N vary widely in size, the estimate (1.3) may be too pessimistic. It is the aim of this paper to analyze this situation in the spirit of [5] and to improve the estimate (1.3) for certain situations with special right-hand sides. Our assumption is that there is a non-negative function $Q(\lambda)$ such that the component of the right-hand side b_N in the direction of the eigenvector associated with eigenvalue λ is in absolute value less than

$$\exp(-N(Q(\lambda) + o(1))) \qquad \text{as } N \to \infty,$$

with $o(1)$ uniform in λ.

As in [5] our tools come from the asymptotic theory of discrete orthogonal polynomials and from extremal problems in logarithmic potential theory [29]. The function Q plays the role of an external field, whereas the measure σ comes in as a constraint. The interaction of the external field with the constraint is quite delicate, and not yet fully understood. As a

[1] We assume exact arithmetic. Notice that the restriction to Conjugate Gradients is just for the ease of presentation. Similar techniques apply to other iterative Krylov subspace methods, as for instance BiCG and MINRES, but also for GMRES in case of normal matrices.

ETNA
Kent State University
etna@mcs.kent.edu

result, the potential theoretic analysis is more complicated than the one in [5]. We are able to prove the analogue of (1.3) only under some simplifying technical assumptions given in Definition 2.4 and Theorem 2.5 below.

As a motivating model problem, we consider the one dimensional Poisson equation

$$-u''(x) = f(x), \quad x \in [0,1],$$

with homogeneous Dirichlet boundary conditions $u(0) = u(1) = 0$. The usual central finite difference approximation on the uniform grid $j/(N+1)$, $j = 0, 1, \ldots, N+1$, leads to a linear system $A_N x = b_N$ with N equations and unknowns, where

$$A_N = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix}_{N \times N}, \quad b_N = (N+1)^2 \cdot \begin{bmatrix} f(1/(N+1)) \\ f(2/(N+1)) \\ f(3/(N+1)) \\ \vdots \\ f(N/(N+1)) \end{bmatrix}. \tag{1.4}$$

Both the one dimensional Poisson problem and the system $A_N x = b_N$ are easy to solve; however, this toy problem can serve to explain convergence behavior observed also in less trivial situations.

As it is well known, the eigenvalues of A_N are given by $\lambda_{j,N} = 2 - 2\cos(\pi j/(N+1))$, $j = 1, ..., N$, leading to the asymptotic eigenvalue distribution σ, with

$$\frac{d\sigma}{dx}(x) = \frac{1}{\pi\sqrt{x(4-x)}}, \quad x \in [0,4], \tag{1.5}$$

the equilibrium measure of the interval $[0,4]$. In the setting of [5], one checks that $S(t) = [0,4]$ for all $t \in (0,1)$, and thus the right-hand side of (1.3) trivially equals one. Indeed, for general f, we may not observe superlinear convergence, as being confirmed by the first two plots of Figure 1.1.

Here we will be interested in what happens for the CG starting vector 0 and particularly smooth functions f, namely

$$f(x) = \sum_{j=1}^{\infty} f_j \sin(\pi j x), \quad x \in [0,1],$$

where we suppose that the Fourier coefficients satisfy

$$r := \limsup_{j \to \infty} |f_j|^{1/j} \in (0,1). \tag{1.6}$$

As we will see below in Lemma 3.1, here the eigenvector components of b_N decrease exponentially for increasing j. Thus, roughly speaking, eigenvalues close to zero will be earlier matched by Ritz values, leading to superlinear convergence. For this model problem, we will show that an improved estimate of the form (1.3) holds with sets $S(t) = [\alpha(t), \beta(t)] \subset (0,4)$ explicitly given in terms of Jacobi elliptic functions, see Lemma 3.2 below. The set $S(t)$ corresponds to the part of the support of an extremal measure in a constrained equilibrium problem with external field where the constraint is not met.

We already noted that the additional external field does not occur (or, what amounts to the same, equals zero) in the analysis of [3, 4, 5, 6, 19]. Indeed, the analysis presented in

ETNA
Kent State University
etna@mcs.kent.edu

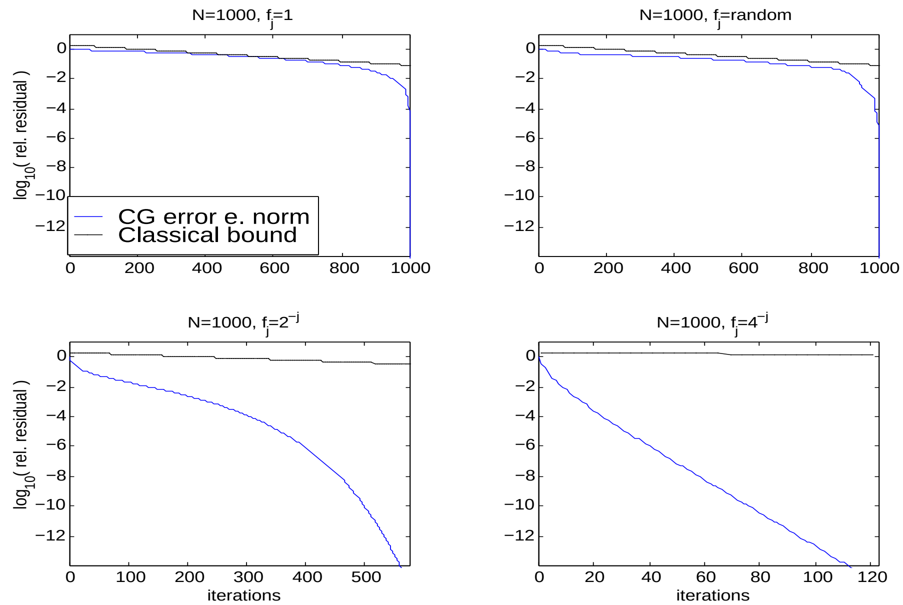

FIG. 1.1. *The CG relative error curve (energy norm, blue) for the one dimensional Poisson problem discretized on a uniform grid ($N = 1000$, $x_{0,N} = 0$) for four choices of $f(x) = \sum_{j=1}^{N} f_j \sin(j\pi x)$: there is hardly any superlinear convergence for the upper two examples ($f_j = 1$, f_j random) and one has to reach approximately the dimension of the system in order to achieve full precision. In contrast, for the lower two examples ($f_j = 2^{-j}$, $f_j = 4^{-j}$) we observe superlinear convergence. For the sake of completeness, we have also drawn in black the error bound (1.1) for the four examples. This bound is only significant for the first two examples.*

these papers is based only on the asymptotic eigenvalue distribution and allows for any starting residual; therefore one may understand these results somehow as worst case estimates. In contrast, here we discuss the situation where the eigenvector components of the starting residual strongly vary in size: we obtain superlinear convergence for CG even for an a priori "bad" eigenvalue distribution. We have the impression that this phenomenon also occurs in less trivial applications in the discretization of PDEs for smooth right-hand sides since quite often high frequencies have a much smaller amplitude than small frequencies.

In order to conclude this introduction, let us return to the correspondence between superlinear convergence and the convergence of Ritz values. Under some stronger assumptions — being for instance true for many right-hand sides of our model problem — one may show generalizing ideas of [19, 3] that the extremal measure mentioned above just gives the asymptotic distribution of Ritz values as $n, N \to \infty$, $n/N \to t$. More precisely, one may show that the nth Ritz values match all eigenvalues within $[0, \alpha(n/N)]$ with an exponential rate, and that there are essentially no nth Ritz values in $[\beta(n/N), 4]$. A precise formulation of the assumptions and a proof of this statement will be given in a future publication. Here we only show some numerical results which fully confirm our claim; see Figure 1.2.

The remainder of the paper is organized as follows: in §2 we recall the well-known connection between CG convergence and discrete orthogonal polynomials. We describe and discuss our assumptions for the sequence of systems under consideration, and briefly recall some extremal problem in logarithmic potential theory. Subsequently, we show in Theorem 2.2 that the solution of this extremal problem leads to a new asymptotic estimate taking into account the asymptotic distribution of the spectrum, but also the characteristics of particular starting residuals. In Theorem 2.5 we show that a formula like (1.3) remains valid for

ETNA
Kent State University
etna@mcs.kent.edu

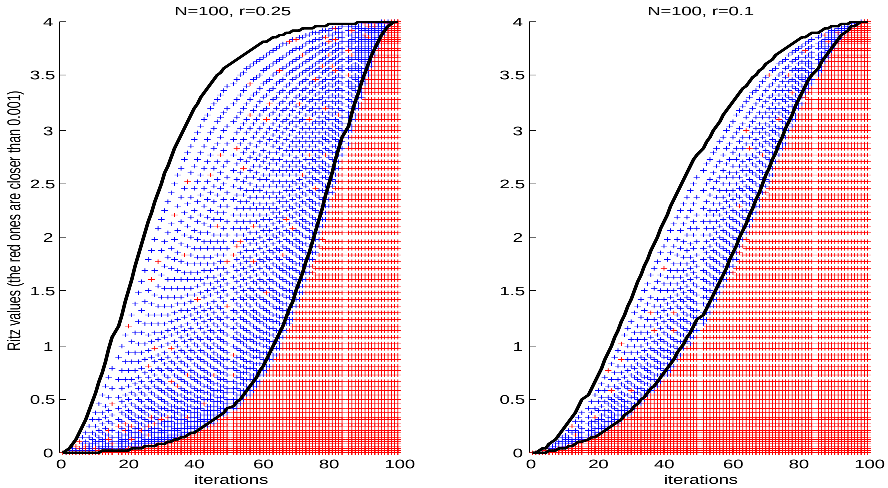

FIG. 1.2. *Convergence of Ritz values for our model problem with* $N = 100$, $x_{0,N} = 0$, *and* $f(x) = \sum_{j=1}^{N} r^j \sin(j\pi x)$, $r \in \{1/4, 1/10\}$. *The two black curves indicate the graphs of* α, β. *We draw in the* n*th column,* $1 \leq n \leq N$, *the position of the* n*th Ritz values (red or blue) within the interval* $[0, 4]$. *A red cross is used if the Ritz value is closer than* 0.001 *to some eigenvalue of* A_N. *Notice that nearly all Ritz values are red in the range* $[0, \alpha(t)]$, $t = n/N$, *that there are no Ritz values in* $[\beta(t), 4]$, *and that in the range* $[\alpha(t), \beta(t)]$ *hardly any Ritz value converged (up to some exceptions by "accident").*

our more general setting. We conclude §2 by recalling in Theorem 2.7 and Theorem 2.8 some known results for computing $S(t)$, especially for a particular interval case, often referred to as the "left ansatz". In §3 we give a detailed analysis for our model problem, and finally present some numerical results.

2. The link with an extremal problem in potential theory. In what follows we denote by $e_{n,N} = A_N^{-1} b_N - x_{n,N}$ the error in Conjugate Gradients after n iterations for the system $A_N x = b_N$ using the starting vector $x_{0,N}$, and by $r_{n,N} = A e_{n,N} = b_N - A_N x_{n,N}$ the corresponding residual.

We denote by $(\lambda_{j,N}, v_{j,N})$, $j = 1, 2, ..., N$, $0 < \lambda_{1,N} \leq \lambda_{2,N} \leq ... \leq \lambda_{N,N}$, $||v_{j,N}|| = 1$, the eigensystem of the matrix A_N and use the spectral decomposition of the first residual

$$r_{0,N} = \sum_{j=1}^{N} w_{j,N} v_{j,N}, \quad w_{j,N} = (r_{0,N}, v_{j,N}) \in \mathbb{R}. \tag{2.1}$$

It will be useful to adopt polynomial language in order to describe the CG error [14]. We consider the *discrete scalar product with varying weights*

$$(p, q)_N = \sum_{j=1}^{N} w_{j,N}^2 p(\lambda_{j,N}) q(\lambda_{j,N}) \tag{2.2}$$

with corresponding norm $||p||_N = \sqrt{(p,p)_N}$, and denote the corresponding orthonormal polynomials by $p_{0,N}, p_{1,N}, ...$ As it is well known [14, 28, 36], these discrete orthonormal polynomials may be used to describe the error in CG as follows:

$$e_{n,N} = \frac{1}{p_{n,N}(0)} p_{n,N}(A) e_{0,N} = \frac{1}{p_{n,N}(0)} A^{-1} p_{n,N}(A) r_{0,N}, \tag{2.3}$$

ETNA
Kent State University
etna@mcs.kent.edu

and thus

$$\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}} = \frac{\|r_{n,N}\|_{A_N^{-1}}}{\|r_{0,N}\|_{A_N^{-1}}} = \frac{1}{|p_{n,N}(0)|} \frac{\|p_{n,N}(\lambda)/\sqrt{\lambda}\|_N}{\|1/\sqrt{\lambda}\|_N}. \tag{2.4}$$

In numerical linear algebra terminology, the zeros of $p_{n,N}$ are referred to as nth Ritz values of A_N. Let us notice that these Ritz values are used to approximate (parts of) the spectrum of A_N. In order to exploit (2.4), we need to know the asymptotic behavior of discrete orthogonal polynomials and in particular their zero distribution.

Starting with a paper of Rakhmanov [27], a number of people [12, 23, 2, 7, 18] contributed to the theory of asymptotics of discrete orthogonal polynomials, see also [22] for a survey. This study is related to an extremal problem in logarithmic potential theory. Before going into details, let us first state the precise assumptions in terms of A_N, b_N.

ASSUMPTION 2.1.

(a) *The eigenvalues of the matrices A_N are uniformly bounded, so that they are all in a fixed interval $[0, R]$ for some $R > 0$.*

(b) *The matrices A_N have the asymptotic eigenvalue distribution σ defined by (1.2).*

(c) *The logarithmic potential U^σ of σ, given by*

$$U^\sigma : \mathbb{C} \to (-\infty, +\infty] : \lambda \mapsto \int \log \frac{1}{|\lambda - \lambda'|} \, d\sigma(\lambda'),$$

is a continuous, real-valued function on $\mathbb{C}$.

(d) *The limit (1.2) also holds for the function $f(\lambda) = \log \lambda$.*

(e) *There exists a nonnegative continuous function $Q : S \to [0, +\infty)$ such that*

$$\limsup_{N\to\infty} \max_j \exp(Q(\lambda_{j,N})) \cdot \left(\frac{|(r_{0,N}, v_{j,N})|}{||r_{0,N}||} \right)^{1/N} \le 1. \tag{2.5}$$

It follows from Assumption 2.1(a),(b) that σ has a compact support in $[0, R]$. Assumption 2.1(b) is natural in a number of applications like for instance in the finite difference discretization of elliptic PDEs. Notice that the measure σ in (1.2) remains invariant if one slightly perturbs the matrix A_N, see, e.g., [37]. For example, σ is preserved in the context of elliptic PDEs if we change the boundary conditions. Here we agree to count eigenvalues in (1.2) according to their multiplicities. Indeed, all following results remain valid (and may become sharper) if we count an eigenvalue only once disregarding multiplicities as in [5], but in applications it is in general difficult to know in advance the multiplicity of eigenvalues.

With regard to Assumption 2.1(c), recall that the logarithmic potential U^σ is always superharmonic on $\mathbb{C}$ (in particular lower semi-continuous), and harmonic on $\mathbb{C} \setminus \mathrm{supp}(\sigma)$, see [29, 26]. The regularity assumption (c) does not allow σ to have point masses. In applications, σ will typically have a density with respect to Lebesgue measure. Assumption 2.1(c) is satisfied for example if the density is continuous.

Assumption 2.1(d) prevents eigenvalues from approaching 0 too fast as $N \to \infty$ (see [5] for an interpretation in terms $\det(A_N)$). Finally, notice that Assumption 2.1(e) trivially holds with $Q \equiv 0$. Indeed, the function $Q \equiv 0$ has been chosen in [3, 4, 5, 6, 19]. As a consequence, these results describe the CG convergence behavior for a "favorable" eigenvalue distribution and for the "worst case" choice of the starting residual. In contrast, if the eigenvector components of the starting residual corresponding to eigenvalues in some part of S decrease exponentially (as it is the case for our model problem) then we may choose Q to be strictly positive on this part of S. This will enable us to prove superlinear convergence for particularly smooth starting residuals even in the case of a "worst case" eigenvalue distribution.

ETNA
Kent State University
etna@mcs.kent.edu

The sets $S(t)$ appearing in (1.3) for some fixed $t \in (0,1)$ are characterized using a *constrained energy problem with external field* where σ acts as the constraint and Q as an external field. This extremal problem is to minimize the weighted logarithmic energy

$$\iint \left(\log \frac{1}{|\lambda - \lambda'|} + 2Q(\lambda) \right) d\nu(\lambda) d\nu(\lambda') \tag{2.6}$$

among all Borel measures on $[0, R]$ that satisfy

$$0 \le \nu \le \sigma, \qquad \int d\nu =: ||\nu|| = t. \tag{2.7}$$

The condition $\nu \le \sigma$ means that $\nu(B) \le \sigma(B)$ for every Borel set B.

Rakhmanov [27, Theorem 3] and Dragnev and Saff [12, Theorem 2.1] showed that there is a unique solution ν_t of this extremal problem,[2] and that, among all measures satisfying (2.7), this solution is uniquely characterized by equilibrium conditions: there exists a constant $F_t \in \mathbb{R}$ such that

$$U^{\nu_t}(\lambda) + Q(\lambda) \left\{ \begin{array}{ll} \le F_t, & \lambda \in \mathrm{supp}(\nu_t), \\ \ge F_t, & \lambda \in \mathrm{supp}(\sigma - \nu_t). \end{array} \right. \tag{2.8}$$

Now the sets $S(t)$ are defined as

$$S(t) = \mathrm{supp}(\sigma - \nu_t) \cap \mathrm{supp}(\nu_t). \tag{2.9}$$

Note that the minimizer ν_t and the set $S(t)$ also depend on σ and Q. However, we do not indicate this in the notation. For the ease of presentation we will suppose in the sequel that $\mathrm{supp}(\sigma)$ is an interval, implying that $S(t) \neq \emptyset$ and thus the extremal constant F_t is unique. Let us mention that in the case of a trivial external field $Q \equiv 0$ one can establish the relationship $\mathrm{supp}(\nu_t) = \mathrm{supp}(\sigma)$ (and therefore $S(t) = \mathrm{supp}(\sigma - \nu_t)$), and the inclusion $S(t) \subset S(\tau)$ for $0 < \tau < t < 1$. These latter properties may get lost for $Q \not\equiv 0$; see our example in §3.

For the particular case of a trivial external field $Q \equiv 0$, the asymptotic bound (1.3) was established in [5, Theorem 2.1] under Assumptions 2.1(a)–(d). As main guides in the proof we were inspired by a result of Rakhmanov [27, Theorem 1] who showed (under some additional assumptions concerning the separation of eigenvalues) that the Ritz values (zeros of $p_{n,N}$) are asymptotically distributed according to the measure ν_t/t as $n, N \to \infty$ with $n/N \to t$. From this it follows that there are roughly speaking as many eigenvalues of A_N outside $S(t)$ as there are Ritz values. See [19, 2] for estimates on the distance between Ritz values and eigenvalues in this setting. As a second important result, we showed using fundamental ideas of Buyarov and Rakhmanov [9] that in the case $Q \equiv 0$ there holds

$$\int_0^t g_{S(\tau)}(0)\, d\tau = F_t - U^{\nu_t}(0). \tag{2.10}$$

The theory of asymptotics of discrete orthogonal polynomials has been developed also for the general case of a nontrivial external field. The aim of the following considerations is to show that we may generalize [5, Theorem 2.1], leading to some counterpart of estimate (1.3) for particularly smooth starting residuals. We will proceed in two steps. Firstly, we

[2] In the presence of an external field we found it convenient to use a normalization different from, e.g., [5, 6]: if $Q \equiv 0$, the extremal measure μ_t of [5, 6] satisfies $\mu_t = \nu_t/t$.

ETNA
Kent State University
etna@mcs.kent.edu

establish in Theorem 2.2 a bound in terms of the extremal quantities ν_t and F_t in a quite general setting, and discuss the sharpness of this new asymptotic bound in Remark 2.3. As second step, we express in Theorem 2.5 our new bound in terms of Green functions for the special case of an interval $S(t)$. We believe that such a formula may be established in a more general setting following ideas of Buyarov and Rakhmanov [9]. However, in order to avoid technical difficulties, we will restrict ourselves to this simpler case which enables us already to fully analyze our model problem.

THEOREM 2.2. *Let* (A_N) *be a sequence of symmetric positive definite matrices such that* Assumptions 2.1(a)-(e) *are satisfied, with* $\mathrm{supp}(\sigma)$ *being an interval. Then, for every* $t \in (0,1)$,

$$\limsup_{\substack{n,N\to\infty \\ n/N\to t}} \frac{1}{N} \log\left(\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}}\right) \le U^{\nu_t}(0) - F_t, \tag{2.11}$$

where ν_t *and* F_t *are from* (2.8).

Proof. As our proof follows closely the lines of the proof of the corresponding part in [5, Theorem 2.1] for $Q \equiv 0$, we only indicate the necessary changes (compare also [12, Lemma 5.3] and [23, Lemma 7.1]). Let $t \in (0,1)$ and $n = n_N$ be fixed such that $n_N/N \to t$ for $N \to \infty$. By definition of CG, for any polynomial p of degree $\le n$ there holds

$$\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}} \le \frac{1}{|p(0)|} \frac{\|p(\lambda)/\sqrt{\lambda}\|_N}{\|1/\sqrt{\lambda}\|_N}, \tag{2.12}$$

where we recall that, according to (2.4), there is equality for $p = p_{n,N}$. Since all eigenvalues are bounded below by $\lambda_{1,N}$ and above by R, we have the rough estimate

$$\begin{aligned} &\frac{\|p/\sqrt{\lambda}\|_N}{\|1/\sqrt{\lambda}\|_N} \le \frac{R}{\lambda_{1,N}} \cdot \frac{\|p\|_N}{||r_{0,N}||} \\ &\le \frac{R}{\lambda_{1,N}} \sqrt{N} \cdot \left(\max_k e^{-NQ(\lambda_{k,N})} |p(\lambda_{k,N})|\right) \cdot \left(\max_j Q(\lambda_{j,N})) \frac{|(r_{0,N}, v_{j,N})|}{||r_{0,N}||}\right). \end{aligned}$$

It follows from Assumption 2.1(a),(d) that

$$\lim_{N\to\infty} |\lambda_{1,N}|^{1/N} = 1.$$

Combining these findings and using Assumption 2.1(e) we may conclude that, for any sequence of polynomials $(p_N)_N$, p_N of degree at most $n = n_N$, there holds

$$\begin{aligned} &\limsup_{N\to\infty} \frac{1}{N} \log\left(\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}}\right) \\ &\le \limsup_{N\to\infty} \frac{1}{N} \log\left(\max_k \frac{|e^{-NQ(\lambda_{k,N})} p_N(\lambda_{k,N})|}{|p_N(0)|}\right). \end{aligned} \tag{2.13}$$

We now consider for $\epsilon > 0$ the set

$$K_\epsilon := \{\lambda \in \mathbb{R} : U^{\nu_t}(\lambda) + Q(\lambda) \le F_t - \epsilon\}$$

and notice that, by (2.8), the two measures ν_t and σ coincide on K_ϵ. Following the construction in [5, Proof of Theorem 2.1] based on a discretization of ν_t and making use of the

ETNA
Kent State University
etna@mcs.kent.edu

Assumption 2.1(a)-(d), we find monic polynomials p_N of degree $n = n_N$ such that

$$\begin{aligned}
&\text{all zeros of } p_N \text{ lie in } \Lambda(A_N),\\
&p_N(\lambda_{k,N}) = 0 \ \text{ for } \lambda_{k,N} \in K_\epsilon,\\
&\lim_{N\to\infty} \frac{1}{N} \log |p_N(0)| = -U^{\nu_t}(0), \quad \text{and}\\
&\limsup_{N\to\infty} \frac{1}{N} \log \left(\max_k e^{NU^{\nu_t}(\lambda_{k,N})} |p_N(\lambda_{k,N})| \right) \le 0.
\end{aligned}$$

Injecting these properties in (2.13), we obtain

$$\begin{aligned}
&\limsup_{\substack{n,N\to\infty\\ n/N\to t}} \frac{1}{N} \log \left(\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}} \right)\\
&\le U^{\nu_t}(0) + \limsup_{N\to\infty} \max_{\lambda_{k,N}\notin K_\epsilon} \Big(-Q(\lambda_{k,N}) - U^{\nu_t}(\lambda_{k,N}) \Big)\\
&\le U^{\nu_t}(0) - F_t + \epsilon.
\end{aligned}$$

Since ϵ can be chosen arbitrarily close to zero, assertion (2.11) follows. □

REMARK 2.3. We claim that estimate (2.11) cannot be improved. In order to see this let us briefly describe a subclass of problems for which equality is attained. We suppose instead of (2.5) that

$$\begin{aligned}
&\lim_{N\to\infty} \max_j \exp(Q(\lambda_{j,N})) \cdot \left(\frac{|(r_{0,N}, v_{j,N})|}{||r_{0,N}||} \right)^{1/N}\\
&= \lim_{N\to\infty} \min_j \exp(Q(\lambda_{j,N})) \cdot \left(\frac{|(r_{0,N}, v_{j,N})|}{||r_{0,N}||} \right)^{1/N} = 1,
\end{aligned} \tag{2.14}$$

that is, the Nth root of the normalized eigenvector component of the starting residual behaves like $\exp(-Q(\lambda_{j,N}))$ uniformly in j. Furthermore, we require that the eigenvalues in $\Lambda(A_N)$ are sufficiently separated. More precisely, we suppose that one of the various separation conditions considered by Rakhmanov [27], Dragnev–Saff [12], Kuijlaars–Van Assche [23], or Beckermann [2] holds; see also [22] for a survey. Then for the corresponding orthonormal polynomials we have nth root asymptotics (see, e.g., [2]), namely

$$\lim_{\substack{n,N\to\infty\\ n/N\to t}} \frac{1}{N} \log |p_{n,N}(z)| = F_t - U^{\nu_t}(z) \tag{2.15}$$

for all $z \in \mathbb{C} \setminus S$. Moreover, following the techniques of [19] one shows using Assumption 2.1(d) that (2.15) remains valid for $z = 0$. Injecting this information in (2.4) enables us to determine the exact rate of convergence:

$$\lim_{\substack{n,N\to\infty\\ n/N\to t}} \frac{1}{N} \log \left(\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}} \right) = U^{\nu_t}(0) - F_t.$$

DEFINITION 2.4. (see [11, Chapter 11] and [21]) *We say that the "left ansatz" holds if* $\mathrm{supp}(\sigma) = [A, B]$, *and if there are continuous and increasing functions* $\alpha, \beta : (0, 1) \to (A, B)$ *such that, for all* $t \in (0, 1)$,

$$\mathrm{supp}(\nu_t) = [A, \beta(t)], \quad \mathrm{supp}(\sigma - \nu_t) = [\alpha(t), B], \quad \alpha(t) < \beta(t). \tag{2.16}$$

ETNA
Kent State University
etna@mcs.kent.edu

The left ansatz plays a prominent role in the study of the continuum limit of the Toda lattice, see, e.g., [11, 20, 21]. Notice that here $S(t) = [\alpha(t), \beta(t)]$ by (2.9) and (2.16). For this interval case we recover the particular form of our estimate (1.3).

THEOREM 2.5. *If the left ansatz holds, then*

$$F_t - U^{\nu_t}(0) = \int_0^t g_{S(\tau)}(0)\, d\tau + \min_{\lambda \in [A,B]} Q(\lambda). \tag{2.17}$$

In particular, if A_N, b_N, σ are as in Theorem 2.2 *and if the left ansatz holds, then*

$$\limsup_{\substack{n,N\to\infty \\ n/N\to t}} \frac{1}{N} \log \left(\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}} \right) \le - \int_0^t g_{S(\tau)}(0)\, d\tau. \tag{2.18}$$

Proof. We only need to show formula (2.17) since (2.18) is a consequence of (2.11), (2.17), and the fact that Q is nonnegative on $[A, B]$ by Assumption 2.1(e). Consider for $0 < t < t' < 1$ the sets $S_{t,t'} \subset S'_{t,t'}$ being defined by

$$S_{t,t'} := S(t) \cap S(t') = [\alpha(t'), \beta(t)], \quad S'_{t,t'} := S(t) \cup S(t') = [\alpha(t), \beta(t')],$$

together with

$$\nu_{t,t'} := \frac{\nu_{t'} - \nu_t}{t' - t}, \quad F_{t,t'} := \frac{F_{t'} - F_t}{t' - t},$$

where $t' - t$ is sufficiently small such that $S_{t,t'} \neq \emptyset$. According to [20, Proposition 4.1], $\nu_{t,t'}$ is a probability measure which by (2.9) is supported on $S'_{t,t'}$. From (2.8) and (2.16) we know that $U^{\nu_{t,t'}}$ equals $F_{t,t'}$ on $S_{t,t'}$, and is less than or equal to this constant on $S'_{t,t'}$. Consequently,

$$U^{\nu_{t,t'}}(\lambda) + g_{S_{t,t'}}(\lambda) \ge F_{t,t'}, \quad \lambda \in S_{t,t'}, \tag{2.19}$$

$$U^{\nu_{t,t'}}(\lambda) + g_{S'_{t,t'}}(\lambda) \le F_{t,t'}, \quad \lambda \in S'_{t,t'}. \tag{2.20}$$

By the principle of domination for logarithmic potentials [29, Theorem II.3.2], the inequalities (2.19) and (2.20) remain valid for all $\lambda \in \mathbb{C}$. Comparing the values at infinity we obtain

$$-\log(\text{cap}\,(S'_{t,t'})) \le F_{t,t'} \le -\log(\text{cap}\,(S_{t,t'})), \tag{2.21}$$

where cap $(.)$ denotes the logarithmic capacity, e.g., cap $([a, b]) = (b - a)/4$. Hence for all $\lambda \in \mathbb{C}$,

$$\log(\text{cap}\,(S'_{t,t'})) + g_{S_{t,t'}}(\lambda) \ge -U^{\nu_{t,t'}}(\lambda) \ge \log(\text{cap}\,(S_{t,t'})) + g_{S'_{t,t'}}(\lambda). \tag{2.22}$$

Passing to the limit $t' \to t+$ and taking into account the continuity of α and β, we get from (2.22) the relation

$$-\frac{d}{dt} U^{\nu_t}(\lambda) = \log(\text{cap}\,(S(t))) + g_{S(t)}(\lambda), \quad t \in (0, 1), \quad \lambda \in \mathbb{C}.$$

Notice also that, by the monotone convergence theorem,

$$\lim_{t\to 0+} U^{\nu_t}(\lambda) = 0, \quad \lambda \in \mathbb{C}.$$

Integrating the above identity with respect to t gives us for every $t \in (0,1)$ and every $\lambda \in \mathbb{C}$

$$-U^{\nu_t}(\lambda) = G_t + \int_0^t g_{S(\tau)}(\lambda)\, d\tau,$$

where

$$G_t = \int_0^t \log(\text{cap}\,(S(\tau)))\, d\tau.$$

Similarly, (2.21) leads to $dF_t/dt = -\log(\text{cap}\,(S(t)))$, and hence for all $t, t' \in (0,1)$

$$G_t - G_{t'} = -F_t + F_{t'}.$$

Therefore

$$C := U^{\nu_t}(\lambda) - F_t + \int_0^t g_{S(\tau)}(\lambda)\, d\tau = -F_t - G_t \tag{2.23}$$

is independent of λ and t. In particular, we may choose $\lambda = \beta(t)$ which by (2.16) is an element of $\text{supp}(\nu_t) \cap \text{supp}(\sigma - \nu_t)$. Applying (2.8) we obtain

$$C := -Q(\beta(t)) + \int_0^t g_{S(\tau)}(\beta(t))\, d\tau, \quad 0 < t < 1,$$

leading to $C = -Q(\beta(0+))$ by continuity of β and Q. Finally, we consider

$$M := \bigcap_{0<t<1} \text{supp}(\nu_t).$$

It is shown in [21, Lemma 5.2] that Q attains its minimum on $[A, B]$ exactly in the set M. In our setting there holds $M = [A, \beta(0+)]$ by (2.16). Hence the minimum value of Q in $[A, B]$ is $Q(\beta(0+)) = -C$, and assertion (2.17) follows from (2.23). □

REMARK 2.6. We recall from the proof of Theorem 2.5 that a combination of (2.11) and (2.17) leads to the more precise assertion

$$\limsup_{\substack{n,N\to\infty \\ n/N\to t}} \frac{1}{N} \log\left(\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}}\right) \leq -\int_0^t g_{S(\tau)}(0)\, d\tau - Q(A)$$

which may be a strict improvement of (2.18). Indeed, $Q(A) \geq 0$ is the minimum of Q on $\text{supp}(\sigma)$, which can be strictly larger than zero.[3] In terms of linear algebra, this happens if there is some "outlier" eigenvalue $\lambda \in \Lambda(A_N) \backslash \text{supp}(\sigma)$ for all N such that the corresponding eigenvector component of $r_{0,N}$ strongly dominates the other eigenvector components.

We still need a simple criterion to decide whether the left ansatz is true, and in this case how to compute $S(t) = [\alpha(t), \beta(t)]$. Here we may apply the following two known results.

THEOREM 2.7. *Let* $\text{supp}(\sigma) = [A, B]$, *and suppose that the functions* Q *and* $\widetilde{Q}$ *defined by*

$$\widetilde{Q}(x) = -Q(x) - U^{\sigma}(x)$$

[3]It can be shown using Assumption 2.1(e) that the minimum of Q on S has to be equal to 0.

are continuous in $[A, B]$ *and have a continuous derivative in* (A, B)*. Suppose in addition that the functions* $x \mapsto (x - A)Q'(x)$ *and* $x \mapsto (B - x)\widetilde{Q}'(x)$ *are increasing functions on* $[A, B]$*. Then the left ansatz holds.*

Proof. See [21, Lemma 3.1 and Theorem 3.3] and [20, Proposition 4.1]. Parts of this assertion have been already shown in [12, Theorem 2.16(b)]. □

THEOREM 2.8. *If the left ansatz holds then, for any* $t \in (0, 1)$*, the quantities* $\alpha(t)$ *and* $\beta(t)$ *satisfy*

$$0 = \frac{1}{\pi}\int_{\alpha(t)}^{\beta(t)} \frac{Q'(\lambda)\,d\lambda}{\sqrt{(\beta(t)-\lambda)(\lambda-\alpha(t))}} - \int_A^{\alpha(t)} \frac{d\sigma(\lambda)}{\sqrt{(\beta(t)-\lambda)(\alpha(t)-\lambda)}}, \tag{2.24}$$

$$t = \frac{1}{\pi}\int_{\alpha(t)}^{\beta(t)} \frac{\lambda Q'(\lambda)\,d\lambda}{\sqrt{(\beta(t)-\lambda)(\lambda-\alpha(t))}} - \int_A^{\alpha(t)} \frac{\lambda\, d\sigma(\lambda)}{\sqrt{(\beta(t)-\lambda)(\alpha(t)-\lambda)}}. \tag{2.25}$$

Proof. See [11, Chapter 4] and [21, Proof of Lemma 6.2]. □

The two assertions above do apply to our model problem, but also to several other cases.

In order to conclude this section, let us give another interpretation of estimate (1.3) (or (2.18)) in terms of some "effective" condition number, compare, e.g., the discussion in the introduction of [33].

REMARK 2.9. If $S(t)$ equals some interval $[\alpha(t), \beta(t)]$ for all $t \in (0, 1)$, then we can use the fact that

$$g_{[\alpha,\beta]}(0) = \log\Big(\frac{\sqrt{\beta}+\sqrt{\alpha}}{\sqrt{\beta}-\sqrt{\alpha}}\Big), \quad 0 \le \alpha < \beta \tag{2.26}$$

in order to rewrite (1.3) (or (2.18)). Indeed, replacing the integral in (1.3) by some rectangle quadrature formula and the Green function by (2.26), we obtain the estimate

$$\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}} \lessapprox \prod_{j=0}^{n-1} \frac{\sqrt{\beta(j/N)/\alpha(j/N)}-1}{\sqrt{\beta(j/N)/\alpha(j/N)}+1}. \tag{2.27}$$

Comparing with (1.1) we see that the ratio $\beta(j/N)/\alpha(j/N)$ has an interpretation as "effective" condition number after j iterations of CG for the system $A_N x = b_N$. Moreover, the product of these first n effective condition numbers (instead of the condition number of A_N) leads to error estimates reflecting superlinear convergence.

3. The analysis of the model problem and examples. In this section we want to show how our model problem (1.4) with (1.6) enters in the general frame described in Section §2. Our analysis consists of two steps: in Lemma 3.1 we show that Assumption 2.1(e) holds with the function

$$Q(\lambda) = \frac{\log(1/r)}{\pi} \arccos\Big(\frac{2-\lambda}{2}\Big), \quad \lambda \in [0, 4]. \tag{3.1}$$

Recall that Assumptions 2.1(a)-(b) have been already verified in §1: $S = [0, 4]$, and σ is the equilibrium measure of S. Assumption 2.1(c) follows from $U^\sigma + g_S \equiv 0$, and a verification of Assumption 2.1(d) is straightforward; we omit details. Our second step consists then in determining, in Lemma 3.2, the set $S(t)$ in terms of Jacobi elliptic functions as a function of r and t. At the end of this section we present some numerical experiments.

LEMMA 3.1. *Let* A_N, b_N *be as in (1.4), and take as starting vector for CG the trivial choice* $x_{0,N} = 0$*.*

ETNA
Kent State University
etna@mcs.kent.edu

If f is as in (1.6) then Assumption 2.1(e) *holds with the function Q from (3.1). Moreover, for the particular function*

$$f(x) = \sum_{j=1}^{\infty} r^j \sin(\pi j x) = \frac{r \sin(\pi x)}{r^2 + 1 - 2R\cos(\pi x)}$$

of (1.6) we have the stronger condition (2.14).

Proof. We first determine the spectral decomposition of $r_{0,N} = b_N$. Recall that the (normalized) eigenvector to the eigenvalue $\lambda_{j,N} = 2 - 2\cos(\pi j/(N+1))$ of A_N is given by

$$v_{j,N} = \sqrt{\frac{2}{N+1}} \cdot \Big(\sin(\frac{\pi j k}{N+1})\Big)_{k=1,\dots,N}, \quad j = 1, ..., N.$$

Hence

$$\begin{aligned} w_{j,N} = (v_{j,N}, b_N) &= \sqrt{\frac{2}{N+1}} \cdot \sum_{k=1}^{N} \sin(\frac{\pi j k}{N+1}) f(\frac{k}{N+1}) \\ &= \sqrt{\frac{2}{N+1}} \cdot \sum_{k=1}^{N} \sum_{\ell=1}^{\infty} f_\ell \sin(\frac{\pi j k}{N+1}) \sin(\frac{\pi \ell k}{N+1}) \\ &= \sqrt{\frac{2}{N+1}} \cdot \sum_{k=1}^{N} \sum_{m=1}^{N} \sum_{\substack{\ell=1 \\ \ell \equiv m(2N+2)}}^{\infty} [f_\ell - f_{2N+2-\ell}] \sin(\frac{\pi j k}{N+1}) \sin(\frac{\pi m k}{N+1}) \\ &= \sqrt{\frac{N+1}{2}} \cdot \sum_{\substack{\ell=1 \\ \ell \equiv j(2N+2)}}^{\infty} [f_\ell - f_{2N+2-\ell}] \\ &= \sqrt{\frac{N+1}{2}} \cdot \sum_{\ell=0}^{\infty} [f_{(2N+2)\ell+j} - f_{(2N+2)(\ell+1)-j}]. \end{aligned}$$

Notice that Q in (3.1) is nonnegative on S, and

$$\exp(-Q(\lambda_{j,N})) = r^{j/(N+1)}, \quad j = 1, ..., N.$$

In the particular case $f_j = r^j$, the sums for $w_{j,N}$ can be evaluated, leading to

$$\begin{aligned} e^{Q(\lambda_{j,N})} |w_{j,N}|^{1/(N+1)} &= \frac{((N+1)/2)^{1/(2N+2)}}{r^{j/(N+1)}} \cdot [\frac{r^j - r^{2N+2-j}}{1 - r^{2N+2}}]^{1/(N+1)} \\ &= ((N+1)/2)^{1/(2N+2)} \cdot [\frac{1 - r^{2N+2-2j}}{1 - r^{2N+2}}]^{1/(N+1)} \end{aligned}$$

which uniformly for $1 \leq j \leq N$ tends to 1 for $N \to \infty$. In particular,

$$\lim_{N\to\infty} ||b_N||^{1/N} = \lim_{N\to\infty} \max_j |w_{j,N}|^{1/N} = 1,$$

and condition (2.15) follows.

We now turn to more general coefficients f_j as in (1.6) where similar techniques may be applied. Let $\rho \in (r, 1)$. By assumption, there exists a J such that $|f_j| \leq \rho^j$ for $j \geq J$. Let $N > J$. Then for $1 \leq j < J$ we have

$$(f_j - \frac{2\rho^{2N+2-j}}{1 - \rho^{2N+2}})^{1/(N+1)} \leq \frac{|w_{j,N}|^{1/(N+1)}}{((N+1)/2)^{1/(2N+2)}} \leq (f_j + \frac{2\rho^{2N+2-j}}{1 - \rho^{2N+2}})^{1/(N+1)}$$

and the two bounds tend to 1 uniformly for $1 \le j < J$ for $N \to \infty$. In particular it follows that

$$\liminf_{N\to\infty} ||b_N||^{1/N} \ge \liminf_{N\to\infty} |w_{1,N}|^{1/N} = 1.$$

In the case $j \ge J$ we have

$$\frac{|w_{j,N}|^{1/(N+1)}}{\rho^{j/(N+1)}} \le \frac{((N+1)/2)^{1/(2N+2)}}{\rho^{j/(N+1)}} \cdot [\frac{\rho^j + \rho^{2N+2-j}}{1-\rho^{2N+2}}]^{1/(N+1)}$$

and the right-hand side tends to 1 uniformly for $J \le j \le N$ for $N \to \infty$. Combining these findings, we may conclude that

$$\limsup_{N\to\infty} \max_{1\le j\le N} \rho^{-j/(N+1)} \cdot \Big(\frac{|w_{j,N}|}{||r_{0,N}||}\Big)^{1/N} \le 1$$

for all $\rho > r$, implying (2.5). □

Suppose that $f(x) = \sin(\pi x)\cdot g(\cos(\pi x))$. Then it is not difficult to verify that the series in (1.6) corresponds to an expansion in terms of Chebyshev polynomials of the second kind of g. In particular, denoting by $\mathcal{E}_R$, $R > 1$, the interior of the ellipse with foci at ± 1 and semiaxes $(R \pm 1/R)/2$, assumption (1.6) will be true iff g is analytic in $\mathcal{E}_{1/r}$ but in no larger $\mathcal{E}_R$. Moreover, it is not difficult to verify that the stronger condition (2.14) remains true if g is meromorphic in some larger $\mathcal{E}_R$, with exactly one pole on the boundary of $\mathcal{E}_{1/r}$.

In the second part of this section we explicitly solve the constrained equilibrium problem with external field for the data of our example in terms of Jacobi elliptic functions and complete elliptic integrals of the first kind

$$K(k) = \frac{1}{2}\int_0^1 \frac{dt}{\sqrt{t(1-t)(1-k^2t)}}, \quad k \in [0,1).$$

For some detailed account on these functions see, e.g., [1].

LEMMA 3.2. *Let σ be the equilibrium measure of $S = [0,4]$, and Q as in (3.1). Furthermore, let $k = k(r) \in (0,1)$ be defined by*

$$\frac{\log(1/r)}{\pi} K(k) = K(\sqrt{1-k^2}). \tag{3.2}$$

Then, for $0 < t < 1$, the left ansatz holds, and $S(t) = [\alpha(t), \beta(t)]$, where

$$\alpha(t) = 4\mathrm{cn}^2((1-t)K(k);k), \tag{3.3}$$
$$\beta(t) = \alpha(t)/\mathrm{dn}^2((1-t)K(k);k). \tag{3.4}$$

In addition, we have

$$g_{S(t)}(0) = \log\Big(\frac{1+\mathrm{dn}((1-t)K(k);k)}{1-\mathrm{dn}((1-t)K(k),k)}\Big). \tag{3.5}$$

Proof. Let us show that the assumptions of Theorem 2.7 are satisfied. Indeed, for our particular data σ, Q we have

$$[A,B] = [0,4], \quad \frac{d\sigma}{dx}(x) = \frac{1}{\pi}\frac{1}{\sqrt{x(4-x)}}, \quad Q'(x) = \frac{\log(1/r)}{\pi}\frac{1}{\sqrt{x(4-x)}}.$$

Notice that U^σ is constant on $[0,4]$, and that both functions

$$x \mapsto xQ'(x) = \frac{\log(1/r)}{\pi}\frac{\sqrt{x}}{\sqrt{4-x}} \quad \text{and} \quad x \mapsto (4-x)\widetilde{Q}'(x) = -\frac{\log(1/r)}{\pi}\frac{\sqrt{4-x}}{\sqrt{x}}$$

are increasing on $[0,4]$. Hence Theorem 2.7 may be applied, showing that the left ansatz holds. By combining (2.9) and (2.16) we obtain the particular form of $S(t)$.

It remains to establish the explicit formulas (3.3)-(3.5). In our setting, formula (2.24) takes the form

$$\begin{aligned} &\frac{\log(1/r)}{\pi^2}\int_{\alpha(t)}^{\beta(t)} \frac{dx}{\sqrt{x(\beta(t)-x)(x-\alpha(t))(4-x)}} \\ &= \frac{1}{\pi}\int_0^{\alpha(t)} \frac{dx}{\sqrt{x(\beta(t)-x)(\alpha(t)-x)(4-x)}}. \end{aligned} \tag{3.6}$$

Notice that both integrals are complete elliptic integrals of the first kind, evaluated at complementary modules. In order to see this, recall that, for $x_1 < x_2 < x_3 < x_4$ or for $x_2 < x_3 < x_4 < x_1$,

$$\int_{x_2}^{x_3} \frac{dx}{\sqrt{|(x-x_1)(x-x_2)(x_3-x)(x_4-x)|}} = \frac{2K(k)}{\sqrt{|(x_3-x_1)(x_4-x_2)|}},$$

where

$$k^2 = \frac{(x_3-x_2)(x_4-x_1)}{(x_3-x_1)(x_4-x_2)}.$$

After identification of the involved quantities and some elementary computations we deduce from (3.6) that

$$k^2 = \frac{4(\beta(t)-\alpha(t))}{\beta(t)(4-\alpha(t))} \in (0,1), \tag{3.7}$$

where k is as in (3.2). Notice that (3.2) fixes a unique k as a function of r since $k \mapsto K(k)/K(\sqrt{1-k^2})$ is strictly increasing in $(0,1)$, takes the value 0 at 0 and $+\infty$ at 1.

A second relation enabling us to prove (3.3)-(3.5) is found as follows: we define $\gamma(t) \in (\alpha(t),\beta(t))$ by

$$\int_{\alpha(t)}^{\beta(t)} \frac{(x-\gamma(t))\,dx}{\sqrt{x(4-x)(\beta(t)-x)(x-\alpha(t))}} = 0.$$

Subtracting $\gamma(t)$ times (2.24) from (2.25), we obtain the relation

$$t = \frac{1}{\pi}\int_0^{\alpha(t)} \frac{(\gamma(t)-x)\,dx}{\sqrt{x(4-x)(\beta(t)-x)(x-\alpha(t))}}. \tag{3.8}$$

Here the right-hand side equals $\omega(\infty)$, where ω is the harmonic measure being 1 on $[0,\alpha(t)]$, 0 on $[\beta(t),4]$, continuous in the extended complex plane $\overline{\mathbb{C}}$ and harmonic in $\Omega := \overline{\mathbb{C}} \setminus \big([0,\alpha(t)] \cup [\beta(t),4]\big)$; see, e.g., [38, §4 and §14]. Thus (3.8) gives us the second relation

$$t = \omega(\infty). \tag{3.9}$$

In order to find the harmonic measure (and especially its value at infinity), it will be useful to map the doubly connected domain Ω to some ring domain using the elliptic sine,

$$\mathrm{sn}(u;k)=x, \quad i.e., \quad u=\int_0^x \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}} \quad (\sqrt{1}=1),$$

being meromorphic and doubly periodic, with periodicities $4K(k)$ and $2iK(k')$, $k'=\sqrt{1-k^2}$, see, e.g., [1, §24, §25, and Example 3 in §49]. Let

$$z=\phi(u)=\frac{4\alpha(t)\mathrm{sn}^2(u;k)}{4\mathrm{sn}^2(u;k)-(4-\alpha(t))}.$$

Then

$$\begin{aligned}&\phi(0)=0, \quad \phi(iK(k'))=\alpha(t),\\ &\phi(-K(k))=\frac{4\alpha(t)}{4-(4-\alpha(t))}=4, \quad \phi(iK(k')-K(k))=\frac{4\alpha(t)/k^2}{4/k^2-(4-\alpha(t))}=\beta(t).\end{aligned}$$

One checks (compare with [1, Example 3 in §49]) that ϕ maps conformally the rectangle $\{u:-K(k)<\Re(u)<0,-K(k')<\Im(u)<K(k')\}$ on $\overline{\mathbb{C}}\setminus[0,4]$. Let

$$u=\psi(w)=\frac{K(k')}{\pi}\log(w) \quad (\log(1)=0),$$

then the set $\{\rho e^{is}:R<\rho<1,-\pi<s<\pi\}$ is mapped via ψ on the above rectangle, where

$$\log(R)=-\pi\frac{K(k)}{K(k')}=\frac{\pi^2}{\log(r)}$$

Pasting at the segment $[-1,-R]$, we have shown that $f:=\phi\circ\psi$ maps the ring domain $\mathcal{R}:=\{w:R<|w|<1\}$ conformally on Ω. Moreover, $f(w)$ approaches $[0,\alpha(t)]$ (and $[\beta(t),4]$, repectively) if w approaches $|w|=1$ (and $|w|=R$, repectively).

Hence for the harmonic measure ω we have the explicit formula

$$\omega(f(w))=\frac{\log(R)-\log(|w|)}{\log(R)}, \quad w\in\mathcal{R},$$

the harmonic measure for the ring domain $\mathcal{R}$. Since

$$f^{-1}(\infty)=\psi^{-1}(u_0), \quad \mathrm{sn}(u_0;k)=-\sqrt{1-\alpha(t)/4},$$

we may conclude using (3.9) that

$$1-t=1-\omega(\infty)=\frac{\log(|\psi^{-1}(u_0)|)}{\log(R)}=u_0\cdot\frac{\pi}{K(k')\log(R)}=-\frac{u_0}{K(k)},$$

and hence

$$\sqrt{1-\alpha(t)/4}=\mathrm{sn}((1-t)K(k);k).$$

Using the relation $\mathrm{cn}^2(u;k)=1-\mathrm{sn}^2(u;k)$ we finally arrive at (3.3). Formula (3.4) is obtained by inserting (3.3) in (3.7), and by using the relation $\mathrm{dn}^2(u;k)=1-k^2\mathrm{sn}^2(u;k)$. Finally, in order to show (3.5) we use (2.25) together with (3.3) and (3.4). $\square$

ETNA
Kent State University
etna@mcs.kent.edu

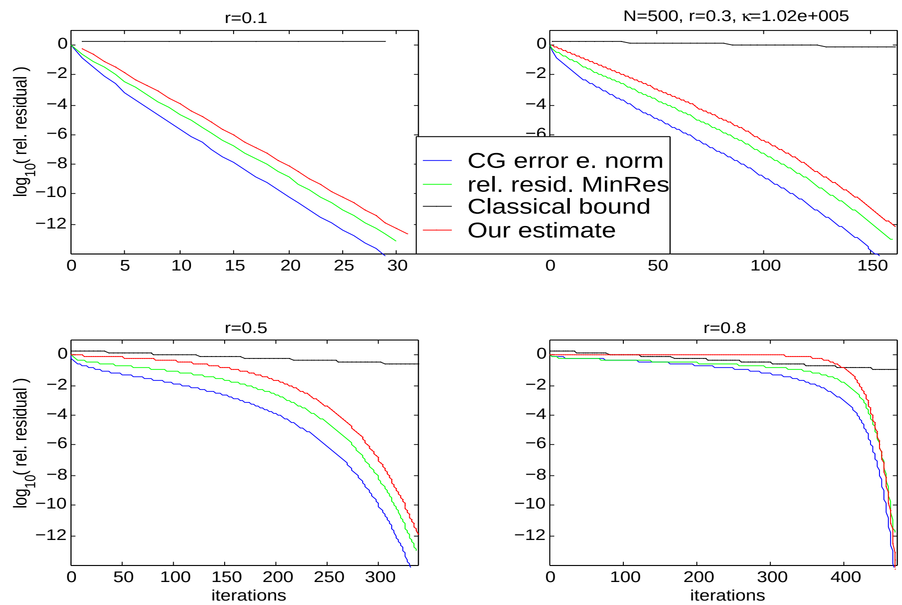

FIG. 3.1. *The one dimensional Poisson problem discretized on a uniform grid ($N = 500$) for $f(x) = \sum_{j=1}^{N} r^j \sin(j\pi x)$, $r = 0.1, 0.3, 0.5, 0.8$. We find the CG (relative) error curve (energy norm) in blue, the classical bound (1.1) in black, and our new asymptotic bound in red. For comparison we give the MINRES relative residual curve in green.*

To summarize, we have shown that the assumptions of Theorem 2.7 and Theorem 2.5 hold. Using the explicit formulas of Lemma 3.2 and Remark 2.9, we obtain for our model problem the error estimate

$$\frac{\|e_{n,N}\|_{A_N}}{\|e_{0,N}\|_{A_N}} \lessapprox \prod_{j=0}^{n-1} \frac{1 - \mathrm{dn}((1-j/N)K(k);k)}{1 + \mathrm{dn}((1-j/N)K(k);k)}, \tag{3.10}$$

where $k = k(r)$ is defined in (3.2).

For the function

$$f(x) = \sum_{j=1}^{N} r^j \sin(j\pi x)$$

and four different choices of r, numerical results are given in Figure 3.1. Notice that our estimator describes quite precisely the convergence rate of CG, and even better the relative residual curve for MINRES, applied to our model problem. In contrast, the classical estimate (1.1), based only on the condition number, is not significant for the convergence process, perhaps up to the case $r = 8/10$ where the superlinear convergence is less pronounced. Again it is striking to observe that our estimate overestimates the actual CG error curve, but that the ratio of estimate and error is constant (about 100) in the superlinear range.

Similar phenomena occur for different choices of N and r. To see this we present in Figure 3.2 a different numerical example based on the model function

$$f(x) = \sum_{j=1}^{\infty} r^j \sin(\pi j x) = \frac{r \sin(\pi x)}{r^2 + 1 - 2r\cos(\pi x)}$$

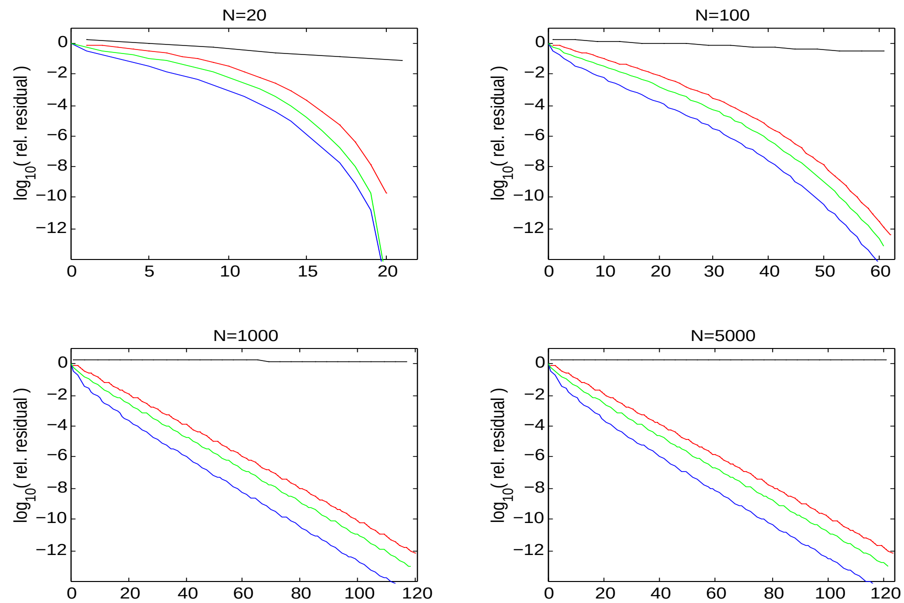

FIG. 3.2. *The one dimensional Poisson problem discretized on a uniform grid for* $f(x) = \sin(\pi x)/[r^2 + 1 - 2r\cos(\pi x)]$*,* $r = 1/4$*, and* $N = 20, 100, 1000, 5000$*. We find the CG (relative) error curve (energy norm) in blue, the MINRES relative residual curve in green, the classical bound (1.1) in black, and our new asymptotic bound in red.*

of Lemma 3.1 with $r = 1/4$ and $N \in \{20, 100, 1000, 5000\}$. It is interesting to observe that the gap between error curve and estimator is about the same for any choice of N; hence it does not depend on the condition number of the underlying system. We also observe that, for sufficiently large N, the number of iterations required to achieve full precision seems to depend no longer on N nor on the condition number of A_N. Such a phenomenon occurred also in [24] where the authors discussed the same Poisson model problem (1.4) with the right-hand side ($\beta > 0$ being some parameter)

$$f(x) = \sum_{j=1}^{N} \frac{\sin^2(\frac{\pi j}{2N+2})}{\cos^\beta(\frac{\pi j}{2N+2})} \cdot \sin(\pi j x).$$

REFERENCES

[1] N.I. AKHIESER, *Elements of the Theory of Elliptic Functions*, Transl. of Math. Monographs **79**, AMS, Providence RI (1990).

[2] B. BECKERMANN, *On a conjecture of E.A. Rakhmanov*, Constr. Approx. 16 (2000), no. 3, 427–448.

[3] B. BECKERMANN, *A note on the convergence of Ritz values for sequences of matrices*, Publication ANO 408, Université de Lille (2000).

[4] B. BECKERMANN, *Superlinear convergence for finite difference approximation of elliptic PDEs*, in preparation (2001).

[5] B. BECKERMANN AND A.B.J. KUIJLAARS, *Superlinear Convergence of Conjugate Gradients*, SIAM J. Num. Anal. **39** (2001) 300-329.

[6] B. BECKERMANN, A.B.J. KUIJLAARS, *On the sharpness of an asymptotic error estimate for Conjugate Gradients*, BIT **41** (2001), 856–867.

[7] B. BECKERMANN AND E.B. SAFF, *The sensitivity of least squares polynomial approximation*, in Int. Ser. of Num. Math. 131, Birkhäuser, Basel (1999), pp. 1–19.

[8] A. BÖTTCHER AND B. SILBERMANN, *Introduction to large truncated Toeplitz matrices*, Universitext, Springer Verlag, New York, 1999.

ETNA
Kent State University
etna@mcs.kent.edu

[9] V.S. BUYAROV AND E.A. RAKHMANOV, *Families of equilibrium measures with external field on the real axis*, Sb. Math. 190 (1999), pp. 791–802.
[10] R.H. CHAN AND M.K. NG, *Conjugate Gradient Methods for Toeplitz systems*, SIAM Rev. 38 (1996), pp. 427–482.
[11] P. DEIFT & K.T-R. MCLAUGHLIN, *A continuum limit of the Toda lattice*, Mem. Amer. Math. Soc. 131/624 (1998).
[12] P.D. DRAGNEV AND E.B. SAFF, *Constrained energy problems with applications to orthogonal polynomials of a discrete variable*, J. d'Analyse Math. 72 (1997), pp. 223–259.
[13] T.A. DRISCOLL, K.-C. TOH AND L.N. TREFETHEN, *From potential theory to matrix iteration in six steps*, SIAM Rev. 40 (1998), pp. 547–578.
[14] B. FISCHER, *Polynomial Based Iteration Methods for Symmetric Linear Systems*, Wiley-Teubner, Sec. Adv. Numer. Math. (1996).
[15] A. GREENBAUM, *Comparisions of splittings used with the conjugate gradient algorithm*, Numer. Math. 33 (1979), pp. 181–194.
[16] A. GREENBAUM, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia PA, 1997.
[17] A. GREENBAUM AND L.N. TREFETHEN, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput. 15 (1994), 359–368.
[18] K. JOHANSSON, *Shape fluctuations and random matrices*, Comm. Math. Phys. 209 (2000), 437–476.
[19] A.B.J. KUIJLAARS, *Which eigenvalues are found by the Lanczos method?* SIAM J. Matrix Anal. Appl. 22 (2000), 306–321.
[20] A.B.J. KUIJLAARS, *On the finite gap ansatz in the continuum limit of Toda lattice.* Duke Math. J. 104 (2000), 433–462.
[21] A.B.J. KUIJLAARS & K.T-R. MCLAUGHLIN, *Long time behavior of the continuum limit of the Toda lattice, and the generation of infinitely many gaps from $\mathcal{C}^\infty$ initial data*, Manuscript (2000).
[22] A.B.J. KUIJLAARS AND E.A. RAKHMANOV, *Zero distributions for discrete orthogonal polynomials*, J. Comput. Appl. Math. 99 (1998), pp. 255–274.
[23] A.B.J. KUIJLAARS AND W. VAN ASSCHE, *Extremal polynomials on discrete sets*, Proc. London Math. Soc. 79 (1999), pp. 191–221.
[24] A.E. NAIMAN, S. ENGELBERG, *A note on conjugate gradient convergence - Part II*, Numer. Math. 85 (2000), pp. 665-683.
[25] O. NEVANLINNA, *Convergence of Iterations for Linear Equations*, Birkhäuser, Basel, 1993.
[26] T. RANSFORD, *Potential theory in the complex plane*, Cambridge University Press, Cambridge, 1995.
[27] E.A. RAKHMANOV, *Equilibrium measure and the distribution of zeros of the extremal polynomials of a discrete variable*, Sb. Math. 187 (1996), pp. 1213–1228.
[28] Y. SAAD, *Iterative methods for sparse linear systems*, PWS Publishing, Boston, MA, 1996.
[29] E.B. SAFF AND V. TOTIK, *Logarithmic potentials with external fields*, Springer, Berlin, 1997.
[30] S. SERRA CAPIZZANO, *A note on the asymptotic spectra of finite difference discretizations of second order elliptic partial differential equations*, Asian J. Math. 4 (2000), 499-514.
[31] S. SERRA CAPIZZANO AND P. TILLI, *From partial differential equations to generalised locally Toeplitz sequences*, manuscript, 2000.
[32] G.L.G. SLEIJPEN AND A. VAN DER SLUIS, *Further results on the convergence behavior of conjugate-gradients and Ritz values*, Linear Algebra Appl. 246 (1996), 233-278.
[33] A. VAN DER SLUIS AND H.A. VAN DER VORST, *The rate of convergence of conjugate gradients*, Numer. Math. 48 (1986), pp. 543–560.
[34] P. TILLI, *Locally Toeplitz sequences: spectral properties and applications*, Linear Algebra Appl. 278 (1998), 91–120.
[35] L.N. TREFETHEN, *Approximation theory and numerical linear algebra*, in: Algorithms for Approximation II, J.C. Mason and M.G. Cox, eds, Chapman & Hall, London, 1990.
[36] L.N. TREFETHEN AND D. BAU III, *Numerical Linear Algebra*, SIAM, Philadelphia PA, 1997.
[37] E.E. TYRTYSHNIKOV, *A uniform approach to some old and new theorems on distribution and clustering*, Linear Algebra Appl. 232 (1996), 1-43.
[38] H. WIDOM, *Extremal Polynomials Associated with a System of Curves and Arcs in the Complex Plane*, Adv. Math., 3 (1969), pp.127-232.