SIAM J. NUMER. ANAL.
Vol. 39, No. 1, pp. 300–329

SUPERLINEAR CONVERGENCE OF CONJUGATE GRADIENTS*

BERNHARD BECKERMANN† AND ARNO B. J. KUIJLAARS‡

Abstract. We give a theoretical explanation for superlinear convergence behavior observed while solving large symmetric systems of equations using the conjugate gradient method or other Krylov subspace methods. We present a new bound on the relative error after n iterations. This bound is valid in an asymptotic sense when the size N of the system grows together with the number of iterations. The bound depends on the asymptotic eigenvalue distribution and on the ratio n/N. Under appropriate conditions we show that the bound is asymptotically sharp.

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.

The new asymptotic bounds for the rate of convergence are illustrated by discussing Toeplitz systems as well as a model problem stemming from the discretization of the Poisson equation.

Key words. superlinear convergence, conjugate gradients, Krylov subspace methods, Toeplitz systems, logarithmic potential theory

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

PII. S0036142999363188

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. 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. It is the purpose of this paper to give an explanation for this behavior in an asymptotic sense.

The error bounds are derived from the following polynomial minimization problem. For any compact set $S \subset \mathbb{R}$, we define

$$E_n(S) = \min_{p \in P_n} \max_{\lambda \in S} |p(\lambda)| , \tag{1.2}$$

where P_n is the set of polynomials p of degree at most n with $p(0) = 1$. The standard convergence analysis of the CG method leads to

$$\frac{\|e_n\|_A}{\|e_0\|_A} \leq E_n(\Lambda(A)) , \tag{1.3}$$

*Received by the editors October 28, 1999; accepted for publication (in revised form) November 24, 2000; published electronically March 15, 2001.
http://www.siam.org/journals/sinum/39-1/36318.html

†Laboratoire d'Analyse Numérique et d'Optimisation, UFR IEEA – M3, UST Lille, F-59655 Villeneuve d'Ascq CEDEX, France (bbecker@ano.univ-lille1.fr).

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

where $\Lambda(A)$ is the spectrum of A. The usual way to analyze (1.3) is to include the spectrum into a "continuous" compact set S so that

$$\frac{\|e_n\|_A}{\|e_0\|_A} \leq E_n(S). \tag{1.4}$$

The quantity $E_n(S)$ can be estimated using notions from potential theory, since

$$\lim_{n\to\infty} \frac{1}{n} \log E_n(S) = -g_S(0), \tag{1.5}$$

where $g_S(z)$ is the Green function for the complement of S with pole at ∞. Thus one arrives at

$$\frac{\|e_n\|_A}{\|e_0\|_A} \lessapprox \exp(-n g_S(0)) \tag{1.6}$$

as an upper estimate for the error. For example, if one chooses $S = [\lambda_{\min}, \lambda_{\max}]$, then the Green function evaluated at 0 is known to be

$$g_S(0) = \log\left(\frac{\sqrt{\lambda_{\max}} + \sqrt{\lambda_{\min}}}{\sqrt{\lambda_{\max}} - \sqrt{\lambda_{\min}}}\right),$$

which leads to the asymptotic estimate

$$\frac{\|e_n\|_A}{\|e_0\|_A} \lessapprox \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n \tag{1.7}$$

in terms of the condition number $\kappa = \lambda_{\max}/\lambda_{\min}$, which is in agreement with (1.1). We refer the reader to the survey paper [DTT98] of Driscoll, Toh, and Trefethen for an excellent account on the interaction between iterative methods in numerical linear algebra and logarithmic potential theory.

The above analysis of the polynomial approximation problem (1.2)–(1.3) is only the starting point of much more detailed investigations on the convergence behavior of Krylov subspace methods. An extensive literature has emerged over the years; see, e.g., the works [DTT98, Gr97, GrTr94, Tre90, TrBa97], where the approximation point of view is emphasized. Two main ideas should be mentioned in the context of these works: first, provided that the eigenvalues of the matrix A are exactly known, the minmax polynomial of (1.3) can be (at least numerically) determined, leading to error estimates often being much sharper than (1.1). Of course, in practical situations one can only hope to have partial information about the distribution of eigenvalues. Here formulas (1.2)–(1.3) potentially give some intuition of what is a "good" or a "bad" distribution. It is our aim to supplement this second statement with some analytic considerations valid in an asymptotic setting.

The estimate (1.7) is typically accurate at early stages of the iteration. The reason for this is that for small n a polynomial $p \in P_n$ that is small on $\Lambda(A)$ has to be uniformly small on the full interval $[\lambda_{\min}, \lambda_{\max}]$ as well. When n gets larger, however, a better strategy for p is to have some of its zeros very close to some of the eigenvalues of A, thereby annihilating the value of p at those eigenvalues, while being uniformly small on a subcontinuum of S only. Then the right-hand side of (1.7) may become a great overestimation of the error. This effect is the reason for the superlinear convergence behavior of the CG iteration, observed in practical situations.

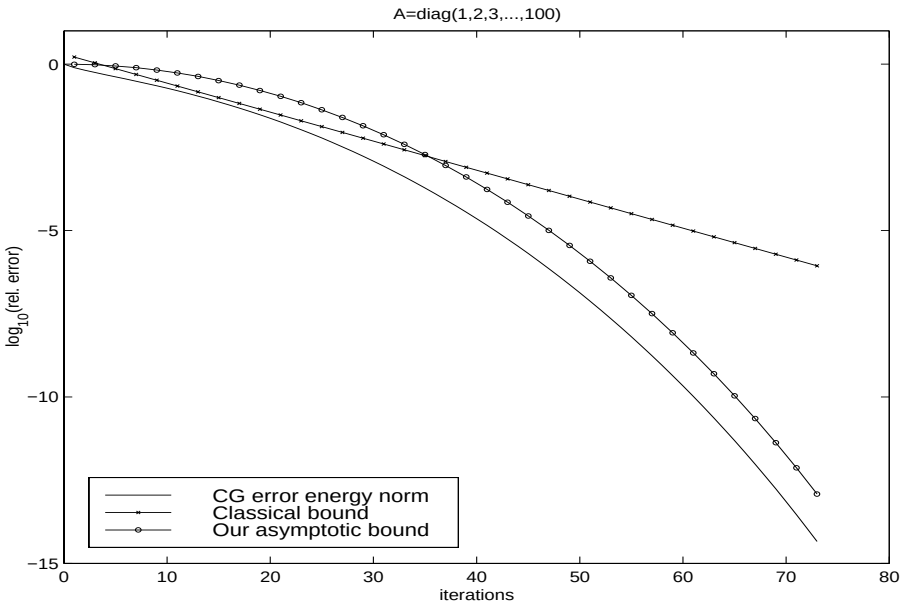

FIG. 1. *The CG error curve versus the two upper bounds for the system* $Ax = b$ *with* $A = \mathrm{diag}(1, 2, \ldots, 100)$*, random solution* x*, and initial residual* $r_0 = (1, \ldots, 1)^T$*. Our new asymptotic bound is given in formula* (3.11).

As an illustration we look at the case of a matrix A with 100 equally spaced eigenvalues $\Lambda(A) = \{1, 2, \ldots, 100\}$. The error curve computed for this example is the solid line in Figure 1. See also [DTT98, p. 560]. The classical error bound given by (1.1) with $\kappa = 100$ is the straight line in Figure 1. For smaller values of n, the classical error bound gives an excellent approximation to the actual error. The other curve (the one with the dots) is the new asymptotic bound for the error that we find in Corollary 3.2 below. This curve follows the actual error especially well in the region of superlinear convergence (for $n \geq 40$).

The phenomenon of superlinear convergence has been understood for compact operators; see [Win80, Mor97, Nev93]. Also, the above heuristic for the convergence behavior of CG for large matrices has been discussed and further analyzed by several authors [VSVV86, Gre79, SlVS96, DTT98]. To our knowledge, a formula for the relative error improving (1.7) and explaining the superlinear convergence is still lacking.

Our goal in this paper is to provide a better understanding of the superlinear convergence of CG iteration, and in particular to explain the form of the error curve as seen in Figure 1. We will argue that for a large $N \times N$ matrix A, the error $E_n(\Lambda(A))$ in the polynomial minimization problem (1.2) is approximately

$$\frac{1}{n} \log E_n(\Lambda(A)) \lessapprox -\frac{1}{t} \int_0^t g_{S(\tau)}(0) d\tau, \tag{1.8}$$

where $t = n/N \in (0, 1)$ and $S(\tau)$, $0 < \tau < 1$, is a decreasing family of sets, depending on the distribution of the eigenvalues of A. The sets $S(\tau)$ have the following interpretation [BeKu00]: $S(\tau)$ is the subcontinuum of $[\lambda_{\min}, \lambda_{\max}]$ where the optimal polynomial of degree $[\tau N]$ is uniformly small.

From (1.3) and (1.8) we find the improved estimate

$$\frac{\|e_n\|_A}{\|e_0\|_A} \lessapprox \rho_t^n \tag{1.9}$$

with

$$\rho_t = \exp\left(-\frac{1}{t}\int_0^t g_{S(\tau)}(0)d\tau\right). \tag{1.10}$$

Note that ρ_t depends on n, since $t = n/N$. As the sets $S(\tau)$ are decreasing as τ increases, their Green functions $g_{S(\tau)}(0)$, evaluated at 0, increase with τ. Hence the numbers ρ_t decrease with increasing n (see also Remark 2.3 below), and this explains the effect of superlinear convergence.

The phenomenon of superlinear convergence may also occur for other Krylov subspace methods applied to a system $Ax = b$, where A is no longer symmetric positive definite. For instance, for symmetric but not positive definite A one often applies the minimal residual method (MINRES) . The GMRES method may be applied in the case of a general matrix A. Supposing that A is diagonalizable, i.e., $A = V \cdot D \cdot V^{-1}$ with D a diagonal matrix containing the (possibly complex) eigenvalues of A, the nth relative residual may be bounded for these methods by

$$\frac{\|r_n\|}{\|r_0\|} \leq \|V\| \cdot \|V^{-1}\| \cdot E_n(\Lambda(A)) \tag{1.11}$$

(see, e.g., [Saa96, Proposition 6.15]). In particular, for symmetric or more generally normal matrices, V is unitary, and thus again we may give bounds for the relative residual by describing the (asymptotic) behavior of $E_n(\Lambda(A))$. Indeed, for the ease of presentation our results are stated for real spectra, but they remain equally valid for complex spectra (see also Remark 2.4 below). Finally, similar techniques may be applied for bounding the error while solving unsymmetric systems using the biconjugate gradients; here, instead of eigenvalues the asymptotic distribution of singular values will intervene.

The paper is organized as follows. In section 2, we describe the (sequence of) matrices A_N under considerations. We explain the potential-theoretic origin of our sets $S(t)$ and establish in Theorem 2.1 the estimate (1.8). Under some stronger assumption concerning the clustering of eigenvalues, we prove in Theorem 2.2 that estimate (1.8) is sharp. Section 3 contains a description of eigenvalue distributions where our sets $S(t)$ are explicit intervals. Subsequently, we give an analysis of the plot of Figure 1. In section 4 it is shown that our assumptions are valid for a large class of symmetric positive definite Toeplitz matrices. Our findings are illustrated by considering a Toeplitz matrix from time series analysis. The discretized two-dimensional Poisson equation on a uniform grid is analyzed in section 5. Finally, a lemma used in the proof of Theorem 2.1 is proved in the appendix.

We should mention that our results concerning the two applications above are more of theoretical nature since in the present paper neither preconditioning nor finite precision arithmetic is considered. The main aim of this paper is to illustrate that some recent results in logarithmic potential theory may help us to understand better a classical phenomenon in numerical linear algebra (see also [BeSa99, Kuij00]).

2. The main result. Properly speaking, the concept of superlinear convergence for the CG method applied to a single linear system does not make sense. Indeed, in

the absence of roundoff errors, the iteration will terminate after N steps if N is the size of the system. Also the notion that the eigenvalues are distributed according to some continuous distribution is problematic when considering a single matrix.

Therefore, we are not going to consider a single matrix A, but instead we will consider a sequence $(A_N)_N$ of symmetric positive definite (or more generally invertible symmetric) matrices. The matrix A_N has size $N \times N$, and we are interested in asymptotics for large N. These matrices need to have an *asymptotic eigenvalue distribution*. By this we mean that there exists a positive Borel measure σ with compact support $\mathrm{supp}(\sigma)$ such that the following condition is satisfied.

CONDITION (i). *The spectra $\Lambda(A_N)$ are all contained in a fixed compact set $S \subset \mathbb{R}$, and for every function f continuous on S we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\lambda\in\Lambda(A_N)} f(\lambda) = \int f(\lambda)\, d\sigma(\lambda). \tag{2.1}$$

This condition is equivalent to the weak* convergence of the normalized eigenvalue counting measures ν_N defined by

$$\nu_N = \frac{1}{N} \sum_{\lambda\in\Lambda(A_N)} \delta_\lambda,$$

where δ_λ is the unit point mass at λ, to σ. As all the A_N have spectra contained in S, the measure σ is supported on S. The total mass $\|\sigma\|$ is at most one, and it can be strictly less than one if the matrices A_N have many coinciding eigenvalues. Note that in the sum in (2.1) each λ in $\Lambda(A_N)$ is taken only once, regardless of its multiplicity (see also Remark 2.5 below).

For the use of the potential theory in what follows, we need to impose a condition on σ. The *logarithmic potential* of a Borel measure μ with compact support is the function

$$U^\mu(\lambda) = \int \log \frac{1}{|\lambda - \lambda'|}\, d\mu(\lambda'), \qquad \lambda \in \mathbb{C}.$$

This is a superharmonic function on $\mathbb{C}$ taking values in $(-\infty, \infty]$. In particular it is lower semicontinuous. We refer the reader to [Ran95, SaTo97] for detailed accounts of logarithmic potential theory. Our assumption is the following.

CONDITION (ii). *The logarithmic potential U^σ of the measure σ from Condition* (i) *is a continuous real-valued function on $\mathbb{C}$.*

Condition (ii) is not very restrictive. For example, if σ is absolutely continuous with respect to Lebesgue measure with a bounded density, then Condition (ii) is satisfied. It is also satisfied if the density has only logarithmic-type or power-type singularities at a finite number of points. On the other hand, Condition (ii) is not satisfied if σ has point masses. A consequence of Condition (ii) is that for any measure μ satisfying $\mu \le \sigma$, the potential U^μ is also continuous. Indeed, U^μ is lower semicontinuous, and since $U^\mu = U^\sigma - U^{\sigma-\mu}$ with U^σ continuous and $U^{\sigma-\mu}$ lower semicontinuous, U^μ is also upper semicontinuous; hence U^μ is continuous.

There is a third condition we impose on the sequence $(A_N)_N$.

CONDITION (iii). *The limit* (2.1) *also holds for $f(\lambda) = \log|\lambda|$.*

Notice that Condition (iii) follows from Condition (i) if $0 \notin S$, or even if the (in modulus) small eigenvalues of A_N do not approach zero too fast. If Condition (iii)

would not hold, then the matrices A_N are too ill-conditioned and the estimate (2.9) given below may very well fail.

In many practical applications, the family $(A_N)_N$ of matrices appears as discretizations of a continuous operator, and then Conditions (i)–(iii) are natural conditions; see, for instance, [Ser00, SeTi00, Til98] and the discussion in sections 4 and 5 below.

The sets $S(t)$ that were announced in (1.8) depend only on the asymptotic eigenvalue distribution σ. They are determined by the solution of an energy minimization problem which we describe now.

The *logarithmic energy* of a Borel measure μ with compact real support is the double integral

$$I(\mu) = \int U^{\mu}(\lambda)\, d\mu(\lambda) = \iint \log \frac{1}{|\lambda - \lambda'|}\, d\mu(\lambda)\, d\mu(\lambda'). \tag{2.2}$$

For every $t \in (0, \|\sigma\|)$, we define the class

$$\mathcal{M}(t;\sigma) = \{\mu \text{ is a Borel probability measure on } \mathbb{R} : 0 \le t\mu \le \sigma\},$$

and we let μ_t be the unique measure minimizing the logarithmic energy (2.2) in the class $\mathcal{M}(t;\sigma)$ (compare [Rak96; DrSa97, Theorem 2.1]). Thus

$$I(\mu_t) = \inf\{I(\mu) : \mu \in \mathcal{M}(t;\sigma)\}. \tag{2.3}$$

The minimizer μ_t depends on t and σ. The minimization problem (2.3) is a constrained problem, since measures in $\mathcal{M}(t;\sigma)$ are dominated by the constraint σ/t. It is known that the minimizer μ_t is characterized by the following variational conditions associated with (2.3). There exists a constant F_t such that

$$U^{\mu_t}(\lambda) = F_t \qquad \text{for } \lambda \in \operatorname{supp}(\sigma - t\mu_t), \tag{2.4}$$

$$U^{\mu_t}(\lambda) \le F_t \qquad \text{for } \lambda \in \mathbb{R} \setminus \operatorname{supp}(\sigma - t\mu_t); \tag{2.5}$$

see [Rak96, Theorem 3; DrSa97, Theorem 2.1]. From these variational conditions one obtains

$$\operatorname{supp}(\mu_t) = \operatorname{supp}(\sigma). \tag{2.6}$$

Finally, the sets $S(t)$ which are crucial in our findings are defined by

$$S(t) = \operatorname{supp}(\sigma - t\mu_t), \qquad t \in (0, \|\sigma\|). \tag{2.7}$$

The extremal problem (2.3) has been studied before in connection with the asymptotic behavior of discrete orthogonal polynomials; see, e.g., [Rak96, DrSa97, KuVA99, Beck00, BeSa99, Joh00]. In particular, the monic analogue of (1.2) is covered by these results, that is, the study of

$$E_n^*(S) = \min_{p \in P_n^*} \max_{\lambda \in S} |p(\lambda)|, \tag{2.8}$$

where P_n^* denotes the class of *monic* polynomials of degree n. Notice that if $S \subset [0, \infty)$ (as, for instance, for symmetric positive definite matrices), then $E_n^*(S)$ and $E_n(S)$ are realized (up to scaling) by the same polynomial, namely, the generalized Chebyshev polynomial.

Our main result is the following.

THEOREM 2.1. *Let $(A_N)_N$ be a sequence of symmetric invertible matrices, A_N of size $N \times N$, satisfying Conditions* (i), (ii), *and* (iii) *for some measure σ. Let the measures μ_t, the constants F_t, and the sets $S(t)$ be defined by* (2.3), (2.4)–(2.5), *and* (2.7), *respectively. Then for $t \in (0, \|\sigma\|)$, we have*

$$\limsup_{\substack{n,N\to\infty \\ n/N\to t}} \frac{1}{n} \log E_n(\Lambda(A_N)) \leq -F_t + U^{\mu_t}(0) \tag{2.9}$$

$$= -\frac{1}{t} \int_0^t g_{S(\tau)}(0)\, d\tau. \tag{2.10}$$

Proof. Let $t \in (0, \|\sigma\|)$, and let $n = n(N)$ depend on N in such a way that $n/N \to t$ as $N \to \infty$. Our goal is to construct for every large N a polynomial p_N in P_n which is sufficiently small on $\Lambda(A_N)$ so as to obtain the estimate (2.9).

We fix $\epsilon > 0$ and define

$$K_\epsilon = \{\lambda \in \mathbb{R} : U^{\mu_t}(\lambda) \leq F_t - \epsilon\}. \tag{2.11}$$

Since U^{μ_t} is a continuous function (cf. the discussion following Condition (ii)), the set K_ϵ is closed. It is disjoint from $S(t)$ because of (2.4) and (2.7). Thus $\sigma(K_\epsilon) = t\mu_t(K_\epsilon)$. By choosing a smaller ϵ if necessary, we may assume that $\sigma(\partial K_\epsilon) = 0$. Then it is possible to find for every large N a set Z_N such that

(1) $\#Z_N = n$,
(2) $\Lambda(A_N) \cap K_\epsilon \subset Z_N \subset \Lambda(A_N)$, and
(3) for all continuous functions f

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\lambda\in Z_N} f(\lambda) = t \int f(\lambda)\, d\mu_t(\lambda). \tag{2.12}$$

For the proof that this is indeed possible, we refer to Lemma 5.1 in the appendix.

We write for N large,

$$p_N(\lambda) = \prod_{\lambda'\in Z_N} \left(1 - \frac{\lambda}{\lambda'}\right). \tag{2.13}$$

Then $p_N \in P_n$ by property (1). We want to estimate $\max_{\lambda\in\Lambda(A_N)} |p_N(\lambda)|$. Let $\lambda_N \in \Lambda(A_N)$ be such that

$$|p_N(\lambda_N)| = \max_{\lambda\in\Lambda(A_N)} |p_N(\lambda)|. \tag{2.14}$$

Since p_N vanishes on $\Lambda(A_N) \cap K_\epsilon$ by property (2) and the definition (2.13), we have

$$\lambda_N \in \mathbb{R} \setminus K_\epsilon = \{\lambda \in \mathbb{R} : U^{\mu_t}(\lambda) > F_t - \epsilon\}$$

and the latter is a bounded set. Passing to a subsequence, if necessary, we may assume that the sequence (λ_N) converges as $N \to \infty$ with limit

$$\lambda^* = \lim_{N\to\infty} \lambda_N \in \{\lambda \in \mathbb{R} : U^{\mu_t}(\lambda) \geq F_t - \epsilon\}. \tag{2.15}$$

We have by (2.13)

$$\frac{1}{n}\log|p_N(\lambda_N)| = \frac{1}{n}\sum_{\lambda\in Z_N}\log|\lambda_N-\lambda| - \frac{1}{n}\sum_{\lambda\in Z_N}\log|\lambda|. \tag{2.16}$$

From property (3) we have that the normalized counting measures of Z_N, i.e.,

$$\zeta_N(X) = \frac{\#(Z_N\cap X)}{N} \qquad \text{for } X\subset\mathbb{R},$$

converge in weak* sense to $t\mu_t$. The principle of descent (see [SaTo97, Theorem I.6.8]) and (2.15) then imply that

$$U^{t\mu_t}(\lambda^*) \le \liminf_{N\to\infty} U^{\zeta_N}(\lambda_N).$$

Since $n/N\to t$, this gives

$$\limsup_{N\to\infty}\frac{1}{n}\sum_{\lambda\in Z_N}\log|\lambda_N-\lambda| \le -U^{\mu_t}(\lambda^*),$$

and thus by (2.15)

$$\limsup_{N\to\infty}\frac{1}{n}\sum_{\lambda\in Z_N}\log|\lambda_N-\lambda| \le -F_t+\epsilon. \tag{2.17}$$

The principle of descent also implies

$$U^{t\mu_t}(0) \le \liminf_{N\to\infty} U^{\zeta_N}(0). \tag{2.18}$$

By property (2) we have $\zeta_N\le\nu_N$, where ν_N is the normalized counting measure of $\Lambda(A_N)$. Since $\nu_N\to\sigma$ by Condition (i), we find that $(\nu_N-\zeta_N)_N$ is a sequence of positive measures that converges to $\sigma-t\mu_t$ in weak* sense. Applying the principle of descent once more, we obtain

$$U^{\sigma-t\mu_t}(0) \le \liminf_{N\to\infty} U^{\nu_N-\zeta_N}(0). \tag{2.19}$$

Also, Condition (iii) gives

$$U^{\sigma}(0) = \lim_{N\to\infty} U^{\nu_N}(0). \tag{2.20}$$

The relations (2.18)–(2.20) easily imply that

$$U^{t\mu_t}(0) = \lim_{N\to\infty} U^{\zeta_N}(0),$$

and this is equivalent to

$$\lim_{N\to\infty} -\frac{1}{n}\sum_{\lambda\in Z_N}\log|\lambda| = U^{\mu_t}(0). \tag{2.21}$$

Combining (2.16) with (2.17) and (2.21), we obtain

$$\limsup_{N\to\infty}\frac{1}{n}\log|p_N(\lambda_N)| \le -F_t+U^{\mu_t}(0)+\epsilon.$$

By (2.14) and the definition (1.2) of E_n, we then see

$$\limsup_{N\to\infty} \frac{1}{n} E_n(\Lambda(A_N)) \leq -F_t + U^{\mu_t}(0) + \epsilon.$$

The number $\epsilon > 0$ can be chosen arbitrarily close to 0. Hence (2.9) follows.

To obtain (2.10) we need to show that

$$F_t - U^{\mu_t}(0) = \frac{1}{t}\int_0^t g_{S(\tau)}(0)\, d\tau. \tag{2.22}$$

To establish this and the inclusion property

$$S(t) \subset S(\tau), \qquad 0 < \tau < t < \|\sigma\|, \tag{2.23}$$

claimed in the introduction, we recall the connection of the constrained minimization problem (2.3) with the energy problem in the presence of an external field. For a continuous function $Q : \mathbb{R} \to \mathbb{R}$ with sufficient growth at $\pm\infty$, the logarithmic energy of a measure μ in the presence of the external field Q is

$$I_Q(\mu) = \iint \log \frac{1}{|\lambda - \lambda'|}\, d\mu(\lambda)\, d\mu(\lambda') + 2\int Q(\lambda)\, d\mu(\lambda).$$

The minimizer μ_s^Q for the extremal problem

$$\inf\left\{ I_Q(\mu) : \int d\mu = s \right\}$$

exists and is unique if

$$0 < s < \liminf_{\lambda\to\pm\infty} \frac{Q(\lambda)}{\log|\lambda|} \tag{2.24}$$

and it is characterized by the conditions

$$U^{\mu_s^Q}(\lambda) + Q(\lambda) = G_s, \qquad \lambda \in \operatorname{supp}(\mu_s^Q), \tag{2.25}$$

$$U^{\mu_s^Q}(\lambda) + Q(\lambda) \geq G_s, \qquad \lambda \in \mathbb{R} \tag{2.26}$$

for some constant G_s; see, e.g., [SaTo97, Theorem I.1.3]. Buyarov and Rakhmanov [BuRa99, Theorem 2] proved the following formula for μ_s^Q:

$$\mu_s^Q = \int_0^s \omega_{\operatorname{supp}(\mu_\tau^Q)}\, d\tau, \tag{2.27}$$

where ω_S is the equilibrium measure for the set S (see, e.g., [Ran95, section 3.3] for the notion of equilibrium measure of a compact set). In fact, the authors consider external fields where the limit on the right-hand side of (2.24) equals $+\infty$ and $s \in (0, +\infty)$. However, from their proof (see the last paragraph of [BuRa99, section 2]) it becomes clear that (2.27) remains valid as long as (2.24) holds.

Now, in our situation with the constraint σ, we take

$$Q(\lambda) = -U^\sigma(\lambda).$$

By comparing the conditions (2.4)–(2.5) with (2.25)–(2.26) we can easily check that for $s, t > 0$ with $s + t = \|\sigma\|$, we have

$$\mu_s^Q = \sigma - t\mu_t, \qquad \operatorname{supp}(\mu_s^Q) = S(t), \qquad G_s = -tF_t. \tag{2.28}$$

In particular, $(\mu_s^Q)_s$ converges in weak* sense to σ for $s \to \|\sigma\|$. Then the Buyarov–Rakhmanov formula (2.27) gives

$$\begin{aligned} t\mu_t = \sigma - \mu_s^Q &= \int_0^{\|\sigma\|} \omega_{\operatorname{supp}(\mu_\tau^Q)}\, d\tau - \int_0^s \omega_{\operatorname{supp}(\mu_\tau^Q)}\, d\tau \\ &= \int_0^t \omega_{S(\tau)}\, d\tau. \end{aligned} \tag{2.29}$$

From (2.29) we obtain the inequality $\tau\mu_\tau \le t\mu_t$ for $\tau < t$, and thus (2.23) holds. In order to show (2.22), notice that the Green function is connected with the potential of the equilibrium measure by the formula

$$g_{S(\tau)}(\lambda) = -\log \operatorname{cap}(S(\tau)) - U^{\omega_{S(\tau)}}(\lambda),$$

where cap denotes the logarithmic capacity. Combining this with (2.29), we obtain for $\lambda \in \mathbb{C}$

$$\begin{aligned} t(F_t - U^{\mu_t}(\lambda)) &= tF_t - \int_0^t U^{\omega_{S(\tau)}}(\lambda)\, d\tau \\ &= tF_t + \int_0^t g_{S(\tau)}(\lambda)\, d\tau + \int_0^t \log \operatorname{cap}(S(\tau))\, d\tau. \end{aligned} \tag{2.30}$$

For $\lambda \in S(t)$, the left-hand side of (2.30) vanishes according to (2.4). Also, by (2.23), each $\lambda \in S(t)$ belongs to $S(\tau)$ for all $\tau < t$ so that the integral in (2.30) involving the Green functions vanishes for $\lambda \in S(t)$. Consequently,

$$tF_t = -\int_0^t \log \operatorname{cap}(S(\tau))\, d\tau,$$

and (2.22) follows from (2.30). This completes the proof of Theorem 2.1. ∎

Under additional conditions the inequality (2.9) can be improved to give equality

$$\lim_{\substack{n,N\to\infty \\ n/N\to t}} \frac{1}{n} \log E_n(\Lambda(A_N)) = -\frac{1}{t}\int_0^t g_{S(\tau)}(0) d\tau. \tag{2.31}$$

These additional conditions are related to the separation of the eigenvalues. If many eigenvalues are very close to each other, then the inequality (2.9) may be strict. For the extremal problem (2.8), various separation conditions were considered by Rakhmanov [Rak96], Dragnev and Saff [DrSa97], Kuijlaars and Van Assche [KuVA99], and Beckermann [Beck00]; see also [KuRa98] for a survey.

If one of these conditions holds in the present situation, the limit (2.31) can be proved. Indeed, according to Theorem 2.1, we require only a sharp lower bound for $E_n(\Lambda(A_N))$. For sets S with positive capacity, lower bounds for $E_n(S)$ are usually obtained by applying the Bernstein–Walsh inequality. In our discrete setting, some analogue of the Bernstein–Walsh inequality in terms of the extremal measure μ_t exists (see [KuVA99, Lemma 8.1 and Corollary 8.2; Beck00, Theorem 1.4(c)]), implying (2.31).

We will give here a proof using the separation condition of Beckermann [Beck00], which was first conjectured by Rakhmanov [KuRa98]. For a finite subset $Z \subset \mathbb{C}$, we introduce

$$I^*(Z) := \frac{1}{(\#Z)^2} \sum_{\lambda \in Z} \sum_{\substack{\lambda' \in Z \\ \lambda' \neq \lambda}} \log \frac{1}{|\lambda - \lambda'|},$$

which may be thought of as the discrete energy of a system of $\#Z$ particles each having a charge $1/\#Z$. Beckermann's condition is the following.

Condition (iv). *With $I(\sigma)$ as in* (2.2)*, we have*

$$\lim_{N \to \infty} I^*(\Lambda(A_N)) = \frac{1}{\|\sigma\|^2} I(\sigma). \tag{2.32}$$

It can be shown that $\liminf I^*(\Lambda(A_N)) \geq \frac{1}{\|\sigma\|^2} I(\sigma)$ already follows from Condition (i). Condition (iv) is, for example, satisfied if there is a positive constant C independent of N such that

$$\min_{\substack{\lambda, \lambda' \in \Lambda(A_N) \\ \lambda \neq \lambda'}} |\lambda - \lambda'| \geq \frac{C}{N}.$$

This is Rakhmanov's separation condition of [Rak96]. Also, the separation condition of Dragnev and Saff [DrSa97] implies Condition (iv). On the other hand, if

$$\min_{\substack{\lambda, \lambda' \in \Lambda(A_N) \\ \lambda \neq \lambda'}} |\lambda - \lambda'| = O\left(e^{-CN^2}\right)$$

for some $C > 0$, then (2.32) does not hold.

It is not difficult to prove (using, for instance, [Beck00, Lemma 2.2(b)]) that Condition (iv) is equivalent to the fact that

$$I^*(X_N) \to I(\mu) \tag{2.33}$$

whenever (X_N) is a sequence of sets satisfying $X_N \subset \Lambda(A_N)$ for each N, $\lim(\#X_N)/N > 0$, and

$$\frac{1}{\#X_N} \sum_{\lambda \in X_N} \delta_\lambda \to \mu.$$

In this form Condition (iv) will be used.

Theorem 2.2. *Suppose that the assumptions of Theorem* 2.1 *are satisfied and that, in addition, Condition* (iv) *holds. Then for every $t \in (0, \|\sigma\|)$, the limit* (2.31) *holds.*

Proof. Let $t \in (0, \|\sigma\|)$. As in the proof of Theorem 2.1 we assume that $n = n(N)$ depends on N in such a way that $n/N \to t$.

For every $N \in \mathbb{N}$, let Φ_N be a set of $n+1$ Fekete points in $\Lambda(A_N)$. That is, Φ_N has $n+1$ points, denoted by $\lambda_{0,N}, \ldots, \lambda_{n,N}$, and it maximizes the product

$$\prod_{\substack{j,k=0 \\ j \neq k}}^{n} |\lambda_j - \lambda_k|$$

among all $n+1$-point subsets of $\Lambda(A_N)$. Equivalently, Φ_N minimizes the discrete energy $I^*(\Phi_N)$. Since $\Phi_N \subset \Lambda(A_N)$, it is clear that

$$E_n(\Lambda(A_N)) \geq E_n(\Phi_N). \tag{2.34}$$

Our first goal is to show that the normalized counting measures of the Fekete points tend to μ_t, that is,

$$\frac{1}{n+1}\sum_{j=0}^{n} \delta_{\lambda_{j,N}} \to \mu_t \tag{2.35}$$

as $N \to \infty$. Since the sets Φ_N are all contained in the compact S, Helly's theorem asserts that from any subsequence of the sequence of normalized counting measures of the Fekete points, we may extract a further subsequence having a weak* limit μ^* (which clearly is an element of $\mathcal{M}(t;\sigma)$). Our claim (2.35) follows by showing that $\mu^* = \mu_t$. According to (2.33), we find that along an appropriate subsequence we then have

$$I^*(\Phi_N) \to I(\mu^*).$$

Let $(Z_N)_N$ be a sequence of sets satisfying $\#Z_N = n+1$, $Z_N \subset \Lambda(A_N)$, and

$$\frac{1}{n+1}\sum_{\lambda \in Z_N} \delta_\lambda \to \mu_t.$$

It follows from Lemma 5.1 in the appendix that such a sequence exists. Again by (2.33) we find that

$$I^*(Z_N) \to I(\mu_t).$$

Since $I(\mu_t) \leq I(\mu^*)$ by (2.3), and $I^*(Z_N) \geq I^*(\Phi_N)$ by the definition of Fekete points, we may conclude that $I(\mu^*) = I(\mu_t)$, and thus $\mu^* = \mu_t$ by the uniqueness of the minimizer in (2.3). This proves the claim (2.35).

Next, we define for $N \in \mathbb{N}$ and $k = 0, 1, \ldots, n$, the polynomial

$$P_{k,N}(\lambda) = \prod_{\substack{j=0 \\ j\neq k}}^{n} (\lambda - \lambda_{j,N}). \tag{2.36}$$

Then $P_{k,N}$ has degree n, and any polynomial $p \in P_n$ can be written in the form

$$p(\lambda) = \sum_{k=0}^{n} a_k \frac{P_{k,N}(\lambda)}{P_{k,N}(0)} \tag{2.37}$$

with coefficients a_k satisfying $\sum_{k=0}^{n} a_k = 1$, since $p(0) = 1$. Then

$$p(\lambda_{k,N}) = a_k \frac{P_{k,N}(\lambda_{k,N})}{P_{k,N}(0)}$$

and

$$\begin{aligned} 1 = \sum_{k=0}^{n} a_k \;&\leq\; \sum_{k=0}^{n} |a_k| \\ &= \sum_{k=0}^{n} |p(\lambda_{k,N})| \left| \frac{P_{k,N}(0)}{P_{k,N}(\lambda_{k,N})} \right| \\ &\leq (n+1) \left(\max_k |p(\lambda_{k,N})| \right) \left(\max_k \left| \frac{P_{k,N}(0)}{P_{k,N}(\lambda_{k,N})} \right| \right). \end{aligned} \tag{2.38}$$

Let $k_N \in \{0, 1, \dots, n\}$ be such that it maximizes

$$\left| \frac{P_{k,N}(0)}{P_{k,N}(\lambda_{k,N})} \right|$$

among all $k \in \{0, 1, \dots, n\}$. Then it follows from (2.38) that

$$\max_k |p(\lambda_{k,N})| \geq \frac{1}{n+1} \left| \frac{P_{k_N,N}(\lambda_{k_N,N})}{P_{k_N,N}(0)} \right|.$$

Since this holds for every $p \in P_n$, we find

$$E_n(\Phi_N) \geq \frac{1}{n+1} \left| \frac{P_{k_N,N}(\lambda_{k_N,N})}{P_{k_N,N}(0)} \right|. \tag{2.39}$$

We write shorter

$$\tilde{\Phi}_N = \Phi_N \setminus \{\lambda_{k_N,N}\}$$

with normalized counting measure

$$\phi_N = \frac{1}{n} \sum_{\lambda \in \tilde{\Phi}_N} \delta_\lambda.$$

Because of (2.35) we see that (ϕ_N) has the weak* limit μ_t for $N \to \infty$. From the principle of descent [SaTo97, Theorem I.6.8] it follows that

$$\limsup_{N\to\infty} \frac{1}{n} \log |P_{k_N,N}(0)| = -\liminf_{N\to\infty} U^{\phi_N}(0) \leq -U^{\mu_t}(0). \tag{2.40}$$

Also, we will show below that

$$\liminf_{N\to\infty} \frac{1}{n} \log |P_{k_N,N}(\lambda_{k_N,N})| \geq -F_t \tag{2.41}$$

(compare with [Beck00, Lemma 2.6]). Combining (2.40) and (2.41) with (2.34) and (2.39), we may conclude that

$$\liminf_{N\to\infty} \frac{1}{n} \log E_n(\Lambda(A_N)) \geq \liminf_{N\to\infty} \frac{1}{n} \log \left| \frac{P_{k_N,N}(\lambda_{k_N,N})}{P_{k_N,N}(0)} \right| \geq -F_t + U^{\mu_t}(0),$$

which in view of Theorem 2.1 is the inequality required for the proof of Theorem 2.2.

Finally, in order to establish (2.41), we note that by the definition of Fekete points we have for every $\lambda \in \Lambda(A_N)$,

$$|P_{k_N,N}(\lambda)| \leq |P_{k_N,N}(\lambda_{k_N,N})|.$$

Taking logarithms, and adding the inequalities for $\lambda \in \Lambda(A_N) \setminus \tilde{\Phi}_N$, we obtain

$$(\#\Lambda(A_N) - n) \log |P_{k_N,N}(\lambda_{k_N,N})| \geq \sum_{\lambda \in \Lambda(A_N) \setminus \tilde{\Phi}_N} \log |P_{k_N,N}(\lambda)|,$$

and therefore

$$\frac{\#\Lambda(A_N) - n}{n^2} \log |P_{k_N,N}(\lambda_{k_N,N})| \geq -\frac{1}{n^2} \sum_{\lambda \in \Lambda_N \setminus \tilde{\Phi}_N} \log \frac{1}{|P_{k_N,N}(\lambda)|}. \tag{2.42}$$

One easily verifies that the right-hand side of (2.42) equals

$$-\frac{1}{2}\left[\frac{(\#\Lambda(A_N))^2}{n^2}I^*(\Lambda(A_N)) - I^*(\tilde{\Phi}_N) - \frac{(\#\Lambda(A_N)-n)^2}{n^2}I^*(\Lambda(A_N)\setminus\tilde{\Phi}_N)\right],$$

which according to (2.33) converges to

$$-\frac{1}{2t^2}\left[I(\sigma) - I(t\mu_t) - I(\sigma - t\mu_t)\right] = -\frac{1}{t}\int U^{\mu_t}(\lambda)\, d(\sigma - t\mu_t)(\lambda) = -\frac{\|\sigma\| - t}{t}\cdot F_t,$$

where for the last equality we have used the variational condition (2.4). Since

$$(\#\Lambda(A_N) - n)/n \to (\|\sigma\| - t)/t,$$

assertion (2.41) follows from (2.42), and Theorem 2.2 is proved. ∎

Remark 2.3. We have shown in Theorem 2.1 that, for $n, N \to \infty$, the quantity $\log E_n(\Lambda(A_N))$ is asymptotically bounded above by

$$\begin{aligned}\log \rho_t^n &= N \cdot t \cdot \log \rho_t \\ &= -N\int_0^t g_{S(\tau)}(0)\, d\tau, \qquad t = n/N,\end{aligned} \tag{2.43}$$

and this bound is sharp (under some additional assumptions) according to Theorem 2.2. This confirms our claims (1.9) and (1.10) of the introduction. We note that Theorem 2.1 is also sharp in a different sense as explained in [BeKu00]. The graph of $Nt\log\rho_t$ for fixed N and varying $n = t\cdot N$ is drawn in the plots of Figures 1, 2, and 4. From (2.43) one sees that $Nt\log\rho_t$ is differentiable, up to at most a countable number of points, with derivative

$$\frac{d}{dn}\left(Nt\log\rho_t\right) = -g_{S(n/N)}(0).$$

Thus it follows that (2.43) is decreasing. Also because of (2.23) one sees that $g_{S(n/N)}(0)$ is increasing with n, and therefore the graph of (2.43) is concave.

If S is a compact set containing all the spectra $\Lambda(A_N)$, then $S(t) \subset S$, for every $t \in (0, \|\sigma\|)$, and one easily checks that

$$\rho_t^n \le \exp(-n g_S(0)).$$

In other words, the bound (1.9) is sharper than (1.6). The equality

$$\rho_t^n = \exp(-n g_S(0))$$

holds if and only if $S(\tau) = S$ for $0 < \tau < t = n/N$, which again is true if and only if the equilibrium measure ω_S of S has a density which is less than or equal to σ/t. This may be translated by saying that, roughly, about tN out of the eigenvalues of A_N are asymptotically distributed like the equilibrium distribution of S. In such a situation one does not observe the effect of superlinear convergence.

Recall that the equilibrium measure of S is the unique probability measure on S that minimizes the logarithmic energy (2.2) among all probability measures; see [Ran95, SaTo97].

Remark 2.4. Theorems 2.1 and 2.2 are equally valid for complex discrete sets $\Lambda(A_N)$; here $\mathrm{supp}(\sigma)$ may be a subset of the complex plane. Indeed, the energy

problems with constraint have been studied in a complex setting (see, e.g., [DrSa97]), and it is possible to show that the representation of $F_t - U^{\mu_t}$ in terms of Green functions remains equally true. Furthermore, all other arguments used in the proofs of Theorems 2.1 and 2.2 still apply for complex sets $\Lambda(A_N)$. As a consequence, our theorems also can be used for bounding the relative residual while solving systems of linear equations with normal matrices A_N via methods like MINRES or GMRES.

Remark 2.5. In many applications (as, for instance, for symmetric Toeplitz matrices) it is difficult to know in advance the multiplicities of the eigenvalues, and one only obtains a measure $\tilde{\sigma}$ defined by a modification of Condition (i) where multiple eigenvalues are counted according their multiplicities. We will refer to this modification as Condition (i′). Condition (ii) with this (possibly) new measure $\tilde{\sigma}$ will be called (ii′), and accordingly Condition (iii) becomes (iii′), where again we count multiplicities. Notice that Theorem 2.1 remains valid if assumptions (i), (ii), and (iii) are replaced by (i′), (ii′), and (iii′) (and σ is replaced by the new measure $\tilde{\sigma}$). In case of, e.g., real S, Conditions (i′) and (iii′) have an interesting interpretation in terms of asymptotics of determinants: indeed,

$$\frac{1}{N}\log|\det(\lambda I_N - A_N)| = -\frac{1}{N}\sum_{\lambda'\in\Lambda(A_N)}\log\frac{1}{|\lambda-\lambda'|}$$

(in this formula we count multiplicities), and from logarithmic potential theory we know that relation (2.1) holds for every function f continuous on S if and only if

$$\lim_{N\to\infty}\frac{1}{N}\log|\det(\lambda I_N - A_N)| = -U^{\tilde{\sigma}}(\lambda) \tag{2.44}$$

for all $\lambda \in \mathbb{C}\setminus S$. Furthermore, it is sufficient that (2.44) holds for $\lambda \in \Sigma$, where $\Sigma \subset \mathbb{C}$ has a finite accumulation point outside of S. Notice that Condition (iii′) may be rewritten as (2.44) with $\lambda = 0$. Finally, Condition (i′) is known to hold if and only if

$$\lim_{N\to\infty}\frac{1}{N}\mathrm{trace}\left((\lambda I_N - A_N)^{-1}\right) = \int\frac{d\tilde{\sigma}(\lambda')}{\lambda-\lambda'} \tag{2.45}$$

for all $\lambda \in \Sigma$, $\Sigma \subset \mathbb{C}$ as above. Using (2.45) one can, for instance, easily show that Condition (i′) remains valid (with the same measure) if A_N is perturbed by some matrix B_N with $\sup_N \|B_N\| < \infty$ and $\mathrm{rank}(B_N)/N \to 0$.

3. Equidistant eigenvalues. In order to apply Theorems 2.1 and 2.2 we have to calculate the sets $S(t)$ from the eigenvalue distribution σ. This is a problem in itself. In general, the sets $S(t)$ can have a complicated form. They may consist of several intervals or even have a Cantor-like structure. The easiest case would be if all $S(t)$ are single intervals. This would also be the most convenient case for the computation of the Green function at 0, since for an interval $[a,b]$ we have

$$g_{[a,b]}(0) = \log\left(\frac{\sqrt{b}+\sqrt{a}}{\sqrt{b}-\sqrt{a}}\right). \tag{3.1}$$

LEMMA 3.1. *Suppose that σ is supported on the interval $[a,b]$ and has a density $w(\lambda)$ with respect to Lebesgue measure. We write*

$$\tilde{w}(\lambda) := \pi\sqrt{(\lambda-a)(b-\lambda)}\,w(\lambda).$$

(a) *Suppose $\tilde{w}$ is strictly increasing on (a,b). Then $S(t)$ is an interval containing b for every $t \in (0, \|\sigma\|)$. More precisely, we have*

$$S(t) = [a,b] \qquad \text{if } t \le \tilde{w}(a+)$$

and

$$S(t) = [r(t), b] \qquad \text{if } t \in (\tilde{w}(a+), \|\sigma\|),$$

where $r = r(t)$ is the unique solution in (a,b) of the equation

$$t = \int_a^r \sqrt{\frac{b-\lambda}{r-\lambda}} w(\lambda)\, d\lambda. \tag{3.2}$$

(b) *Suppose $\tilde{w}$ is strictly decreasing on (a,b). Then $S(t)$ is an interval containing a for every $t \in (0, \|\sigma\|)$. More precisely, we have*

$$S(t) = [a,b] \qquad \text{if } t \le \tilde{w}(b-)$$

and

$$S(t) = [a, r(t)] \qquad \text{if } t \in (\tilde{w}(b-), \|\sigma\|),$$

where $r = r(t)$ is the unique solution in (a,b) of the equation

$$t = \int_r^b \sqrt{\frac{\lambda-a}{\lambda-r}} w(\lambda)\, d\lambda.$$

(c) *Suppose $\tilde{w}$ is symmetric with respect to the midpoint $m := (a+b)/2$ and strictly decreasing on (m,b). Let $t \in (0, \|\sigma\|)$. Then*

$$S(t) = [a,b] \qquad \text{if } t \le \tilde{w}(b-)$$

and

$$S(t) = [m - r(t), m + r(t)] \qquad \text{if } t \in (\tilde{w}(b-), \|\sigma\|),$$

where $r = r(t)$ is the unique solution in $(0, (b-a)/2)$ of the equation

$$t = 2\int_a^{m-r} \frac{m-\lambda}{\sqrt{(m-\lambda)^2 - r^2}} w(\lambda) d\lambda.$$

Proof. (a) We consider as in the proof of Theorem 2.1 the external field

$$Q(\lambda) = -U^\sigma(\lambda) = \int_a^b \log|\lambda - \lambda'| w(\lambda')\, d\lambda'. \tag{3.3}$$

Let μ_s^Q be the extremal measure with external field Q and normalization $s \in (0, \|\sigma\|)$ (cf. the paragraph preceding formula (2.26)). In [KuDr99, Theorem 2] it was proved that the support of μ_s^Q is an interval of the form $[r,b]$ if Q and w are related as in (3.3) and if $\tilde{w}(\lambda)$ increases on $[a,b]$. In [KuDr99] this is stated under the assumption that Q is differentiable with a Hölder continuous derivative. An inspection of the proof, however, shows that this assumption is not necessary. It was also assumed that $s = 1$.

This is also not essential. Since $S(t) = \operatorname{supp}(\mu_s^Q)$ for $s + t = \|\sigma\|$ by (2.28), it thus follows that $S(t)$ is an interval containing b for every t.

We show that $S(t) = [a, b]$ if and only if $t \le \tilde{w}(a+)$. For $t \le \tilde{w}(a+)$, we have from the fact that $\tilde{w}$ is strictly increasing

$$\frac{t}{\pi\sqrt{(b-\lambda)(\lambda-a)}} < w(\lambda), \qquad \lambda \in (a, b). \tag{3.4}$$

Thus the equilibrium measure $\omega_{[a,b]}$ of $[a, b]$, i.e.,

$$d\omega_{[a,b]}(\lambda) = \frac{1}{\pi\sqrt{(b-\lambda)(\lambda-a)}}\, d\lambda,$$

belongs to the class $\mathcal{M}(\sigma; t)$. Since $\omega_{[a,b]}$ minimizes the energy (2.2) among all probability measures on $[a, b]$, it is then also the minimizer over $\mathcal{M}(\sigma; t)$. Thus $\mu_t = \omega_{[a,b]}$, and it follows from (3.4) that

$$S(t) = \operatorname{supp}(\sigma - t\mu_t) = [a, b].$$

Conversely, if $S(t) = [a, b]$, then μ_t is a probability measure on $[a, b]$ whose potential is constant on $[a, b]$ by (2.4). This implies that $\mu_t = \omega_{[a,b]}$. Hence $t\omega_{[a,b]} \le \sigma$ and (3.4) holds. Thus $t \le \tilde{w}(a+)$.

For the rest of the proof we assume that $t \in (\tilde{w}(a+), \|\sigma\|)$. Then $S(t) = [r(t), b]$ with $a < r(t) < b$. From [DrSa97, Corollary 2.15] we know that $t\mu_t = \sigma - \hat{\sigma} + t\omega_{[r,b]}$, where $\hat{\sigma}$ is the balayage (see [SaTo97, section II.4]) of σ onto the interval $[r, b]$. Consequently, according to [SaTo97, (II.4.47)], $t\mu_t$ has the density

$$v(\lambda) = \begin{cases} w(\lambda) & \text{if } \lambda \in (a, r), \\ \dfrac{t}{\pi\sqrt{(\lambda-r)(b-\lambda)}} - \dfrac{1}{\pi}\displaystyle\int_a^r \sqrt{\frac{(\lambda'-r)(\lambda'-b)}{(\lambda-r)(b-\lambda)}}\,\frac{w(\lambda')\,d\lambda'}{\lambda-\lambda'} & \text{if } \lambda \in (r, b). \end{cases}$$

For $\lambda \in (r, b)$, we rewrite the density as

$$v(\lambda) = \frac{1}{\pi\sqrt{(\lambda-r)(b-\lambda)}}\left[t - \int_a^r \frac{\sqrt{(b-\lambda')(r-\lambda')}}{\lambda-\lambda'} w(\lambda')\,d\lambda'\right]. \tag{3.5}$$

Since $a < r < b$ and $0 \le v(\lambda) \le w(\lambda) < \infty$ for $\lambda \in (a, b)$, we must have

$$\lim_{\lambda\to r+} \left(\pi\sqrt{(\lambda-r)(b-\lambda)}\right) v(\lambda) = 0.$$

In view of (3.5), the relation (3.2) follows.

To show that there is only one r satisfying (3.2), we rewrite the right-hand side of (3.2) as

$$\begin{aligned} \int_a^r \sqrt{\frac{b-\lambda}{r-\lambda}} w(\lambda) d\lambda &= \frac{1}{\pi}\int_a^r \tilde{w}(\lambda) \frac{d\lambda}{\sqrt{(r-\lambda)(\lambda-a)}} \\ &= \frac{2}{\pi}\int_0^{\pi/2} \tilde{w}(a + (r-a)\sin^2\theta)\, d\theta, \end{aligned} \tag{3.6}$$

where for the second equality, we used the change of variables $\lambda = a + (r-a)\sin^2\theta$. Since $\tilde{w}$ is strictly increasing, it is then clear that (3.6) strictly increases for $r \in (a, b)$. This completes the proof of part (a).

(b) The proof of part (b) is similar.

(c) Part (c) follows using a quadratic transformation; compare this with [Kuij00, Proof of Theorem 5.1]. ∎

Lemma 3.1 allows us to determine the sets $S(t)$ in a number of situations. We consider here the case of equidistant eigenvalues.

Suppose A_N has N equidistant eigenvalues $1, 2, \dots, N$. Multiplying the matrix by a positive constant does not change the numbers $E_n(\Lambda(A_N))$. We multiply A_N by $1/N$ and so we consider instead matrices with spectrum

$$\Lambda\left(\frac{1}{N}A_N\right) = \left\{\frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}, 1\right\}.$$

These matrices have an asymptotic eigenvalue distribution

$$d\sigma(\lambda) = d\lambda, \qquad \lambda \in [0,1], \tag{3.7}$$

and Conditions (i)–(iv) are satisfied.

The explicit solution of the energy minimization problem (2.3) with σ given by (3.7) is due to Rakhmanov [Rak96]. We show how the sets $S(t)$ can be determined from Lemma 3.1. The assumptions of Lemma 3.1(c) are clearly satisfied with $a = 0$, $b = 1$, and $m = 1/2$. Therefore, we have for $0 < t < 1$,

$$\begin{aligned} t &= 2\int_0^{1/2-r} \frac{1/2-\lambda}{\sqrt{(1/2-\lambda)^{1/2} - r^2}}\, d\lambda \\ &= -2\sqrt{(1/2-\lambda)^2 - r^2}\Big|_{\lambda=0}^{\lambda=1/2-r} \\ &= \sqrt{1-4r^2}. \end{aligned}$$

Thus

$$r = r(t) = \sqrt{1-t^2}/2 \tag{3.8}$$

and

$$S(t) = [m-r, m+r] = \left[\frac{1}{2} - \frac{1}{2}\sqrt{1-t^2}, \frac{1}{2} + \frac{1}{2}\sqrt{1-t^2}\right].$$

Using (3.1) and (3.8) we find after a brief calculation

$$g_{S(t)}(0) = \log\left(\frac{\sqrt{m+r}+\sqrt{m-r}}{\sqrt{m+r}-\sqrt{m-r}}\right) = \frac{1}{2}\log\left(\frac{1+t}{1-t}\right). \tag{3.9}$$

Hence

$$\begin{aligned} \frac{1}{t}\int_0^t g_{S(\tau)}(0)\, d\tau &= \frac{1}{2t}\int_0^t \log\left(\frac{1+\tau}{1-\tau}\right) d\tau \\ &= \frac{(1+t)\log(1+t) + (1-t)\log(1-t)}{2t}. \end{aligned} \tag{3.10}$$

Theorem 2.2 and (3.10) then give the following result.

COROLLARY 3.2. *For every $t \in (0,1)$ we have*

$$\lim_{\substack{n,N\to\infty \\ n/N\to t}} \frac{1}{n}\log E_n(\{1,2,\dots,N\}) = -\frac{(1+t)\log(1+t) + (1-t)\log(1-t)}{2t}. \tag{3.11}$$

Corollary 3.2 gives the theoretical justification for our CG bound in the case of equidistant eigenvalues as given in Figure 1. Notice that, already for $N = 100$, the approximation for $\log E_n(\{1, 2, \dots, N\})$ obtained by multiplying the right-hand side of (3.11) by n is quite accurate.

To conclude this section, we note that Lemma 3.1(c) also applies to the case of ultraspherical eigenvalue distributions. The corresponding sets $S(t)$ were determined in [Kuij00].

4. Applications to Toeplitz matrices. Toeplitz matrices provide interesting examples for our results. Toeplitz systems arise in a variety of applications, such as signal processing and time series analysis; see [ChNg96] and the references cited therein.

Let $\phi : [-\pi, \pi] \to [0, \infty)$ be an integrable function with Fourier coefficients

$$\phi_k := \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(\theta) e^{-ik\theta}\, d\theta, \quad k = 0, \pm 1, \pm 2, \dots .$$

We assume ϕ is bounded and not equal to a constant. The Nth Toeplitz matrix with symbol ϕ is given by

$$T_N(\phi) = \begin{bmatrix} \phi_0 & \phi_{-1} & \cdots & \phi_{1-N} \\ \phi_1 & \phi_0 & \cdots & \phi_{2-N} \\ \vdots & \vdots & & \vdots \\ \phi_{N-1} & \phi_{N-2} & \cdots & \phi_0 \end{bmatrix}.$$

Then $T_N(\phi)$ is a Hermitian matrix and it is well known that $(\phi_{\inf}, \phi_{\sup})$ is the smallest interval containing the spectrum of $T_N(\phi)$ for every N, where $\phi_{\inf}$ and $\phi_{\sup}$ denote the essential infimum and essential supremum of ϕ, respectively. Thus, since ϕ is nonnegative, all eigenvalues $\lambda_{1,N} \le \lambda_{2,N} \le \cdots \le \lambda_{N,N}$ of $T_N(\phi)$ are strictly positive, and the matrix $T_N(\phi)$ is positive definite.

A classical result of Szegő (see, e.g., [GrSz84, pp. 63–65; BoSi99, Theorem 5.10 and Corollary 5.11]) says that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} f(\lambda_{j,N}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\phi(\theta))\, d\theta \tag{4.1}$$

for every continuous function f on $[\phi_{\min}, \phi_{\max}]$. It follows that the sequence $(T_N(\phi))$ satisfies Condition (i′); see Remark 2.5 with the measure σ given by

$$\int f(\lambda) d\sigma(\lambda) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\phi(\theta))\, d\theta. \tag{4.2}$$

Then σ is a probability measure and its support is equal to the essential range of ϕ.

Another result of Szegő (see [Sz67, (12.3.3); GrSz84, pp. 44, 66]) is that

$$\lim_{N\to\infty} \frac{\det(T_N(\phi))}{\det(T_{N-1}(\phi))} = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log \phi(\theta)\, d\theta \right) \tag{4.3}$$

provided that ϕ satisfies the Szegő condition

$$\int_{-\pi}^{\pi} \log \phi(\theta)\, d\theta > -\infty.$$

Notice that this condition can be rewritten as $U^\sigma(0) < +\infty$. It follows from (4.2), (4.3) that

$$\lim_{N\to\infty} \frac{\log\det T_N(\phi)}{N} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\phi(\theta)\,d\theta = \int \log\lambda\,d\sigma(\lambda) \in \mathbf{R},$$

and Condition (iii′) is satisfied.

Consequently, for Toeplitz matrices $T_N(\phi)$ with nonnegative, integrable, and bounded symbol ϕ and continuous real-valued potential U^σ, Conditions (i′)–(iii′) are satisfied, and we may apply Theorem 2.1.

We will discuss an example of Kac, Murdock, and Szegő [KaMuSz53, p. 783]

$$\phi(\theta) = \frac{1-\gamma^2}{1-2\gamma\cos\theta+\gamma^2}$$

with $\gamma \in (-1,1)$. Toeplitz matrices with this symbol (or with a multiple of this symbol) arise as covariance matrices of first-order autoregressive processes [ChNg96, section 4.6.1]. The corresponding Fourier coefficients are given by

$$\phi_k = \gamma^{|k|}, \qquad k = 0, \pm1, \pm2, \ldots .$$

Suppose without loss of generality that $\gamma > 0$. Then the measure σ from (4.2) has support $[a,b]$, where

$$a = \frac{1-\gamma}{1+\gamma}, \qquad b = \frac{1+\gamma}{1-\gamma}.$$

Since ϕ is even we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\phi(\theta))\,d\theta = \frac{1}{\pi}\int_{0}^{\pi} f(\phi(\theta))\,d\theta.$$

Making the substitution $\lambda = \phi(\theta)$, we obtain after some calculations

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\phi(\theta))\,d\theta = \frac{1}{\pi}\int_{a}^{b} f(\lambda)\frac{d\lambda}{\lambda\sqrt{(\lambda-a)(b-\lambda)}}.$$

Thus the measure σ has density

$$w(\lambda) = \frac{1}{\pi\lambda\sqrt{(\lambda-a)(b-\lambda)}}, \qquad a < \lambda < b, \tag{4.4}$$

with respect to Lebesgue measure. From (4.4) it is easy to show that U^σ is continuous, so that Theorem 2.1 applies in this case.

Now we apply Lemma 3.1(b) in order to compute $S(t)$. Notice that for $r \in [a,b)$,

$$\int_r^b \sqrt{\frac{\lambda-a}{\lambda-r}}\,w(\lambda)\,d\lambda = \frac{1}{\pi}\int_r^b \frac{d\lambda}{\lambda\sqrt{(\lambda-r)(b-\lambda)}} = \sqrt{\frac{a}{r}}.$$

Consequently, by Lemma 3.1(b), we have $S(t) = [a, r(t)]$ with

$$r(t) = \begin{cases} b & \text{if } 0 < t \le \sqrt{\dfrac{a}{b}} = a, \\ \dfrac{a}{t^2} & \text{if } a < t < 1. \end{cases}$$

In particular, we get from (1.10) and (3.1) the convergence rate

$$\log \rho_t = -g_{[a,b]}(0) = \log(\gamma) \qquad \text{if } 0 < t \le a,$$

whereas for $a < t < 1$, we have

$$\begin{aligned} t \log \rho_t &= a \log(\gamma) - \int_a^t \log \frac{\sqrt{r(\tau)} + \sqrt{a}}{\sqrt{r(\tau)} - \sqrt{a}}\, d\tau \\ &= a \log(\gamma) - \int_a^t \log \left(\frac{1+\tau}{1-\tau} \right) d\tau \\ &= a \log(\gamma) - \left[(1+\tau)\log(1+\tau) + (1-\tau)\log(1-\tau) \right]_a^t \\ &= \log \left(\frac{4\gamma}{(1+\gamma)^2} \right) - (1+t)\log(1+t) - (1-t)\log(1-t). \end{aligned}$$

It is quite interesting that, in the superlinear range, we obtain (up to some linear transformation) the same function as for equidistant nodes.

Numerical experiments for the symmetric positive definite Toeplitz matrix T_{200} of order 200 of Kac, Murdock, and Szegő are given in Figure 2. The four different plots correspond to the choices $\gamma \in \{1/2, 2/3, 5/6, 19/20\}$ of the parameter. Notice that the CG error curve (solid line) of the last two plots is clearly affected by rounding errors leading to loss of orthogonality, whereas the GMRES relative residual curves (dotted line) behave essentially as predicted by our theory.[1] In particular, the classical bound (1.1), (1.11) (crosses) no longer describes correctly the size of the relative residual of GMRES for $n \ge 20$ and $\gamma \in \{5/6, 19/20\}$. Experimentally we observe that the range of superlinear convergence starts in the different examples approximately at the iteration indices ≥ 50, 30, 20, and 10, respectively. This has to be compared with the predicted quantity $N \cdot a$ which for the different choices of γ approximately takes the values 66, 40, 29, and 5, respectively. Though these numbers differ slightly, we observe that the new bound (1.9), (1.10) reflects quite precisely the shape of the relative residual curve, and in particular allows us to detect the ranges of linear and of superlinear convergence.

As our second illustrating example let us mention the Toeplitz matrices occurring in the context of the first-order moving average process [ChNg96, section 4.6.1], where the symbol is given by

$$\phi(\theta) = \eta^2(1 + 2\gamma \cos\theta + \gamma^2).$$

Here the eigenvalues are asymptotically distributed like the equilibrium distribution on $[\eta^2(1-\gamma)^2, \eta^2(1+\gamma)^2]$, and therefore there will be no superlinear convergence in this case.

5. The model problem. Consider the two-dimensional Poisson equation

$$-\frac{\partial^2 u(x,y)}{\partial x^2} - \frac{\partial^2 u(x,y)}{\partial y^2} = f(x,y)$$

[1] In the context of finite precision arithmetic we should report about some recent work [Stra00] showing that the convergence of CG is delayed by one iteration step per each loss of linear dependence among the CG residual vectors. In contrast, for GMRES there is no delay since one may establish some correspondence between loss of orthogonality and convergence of the method.

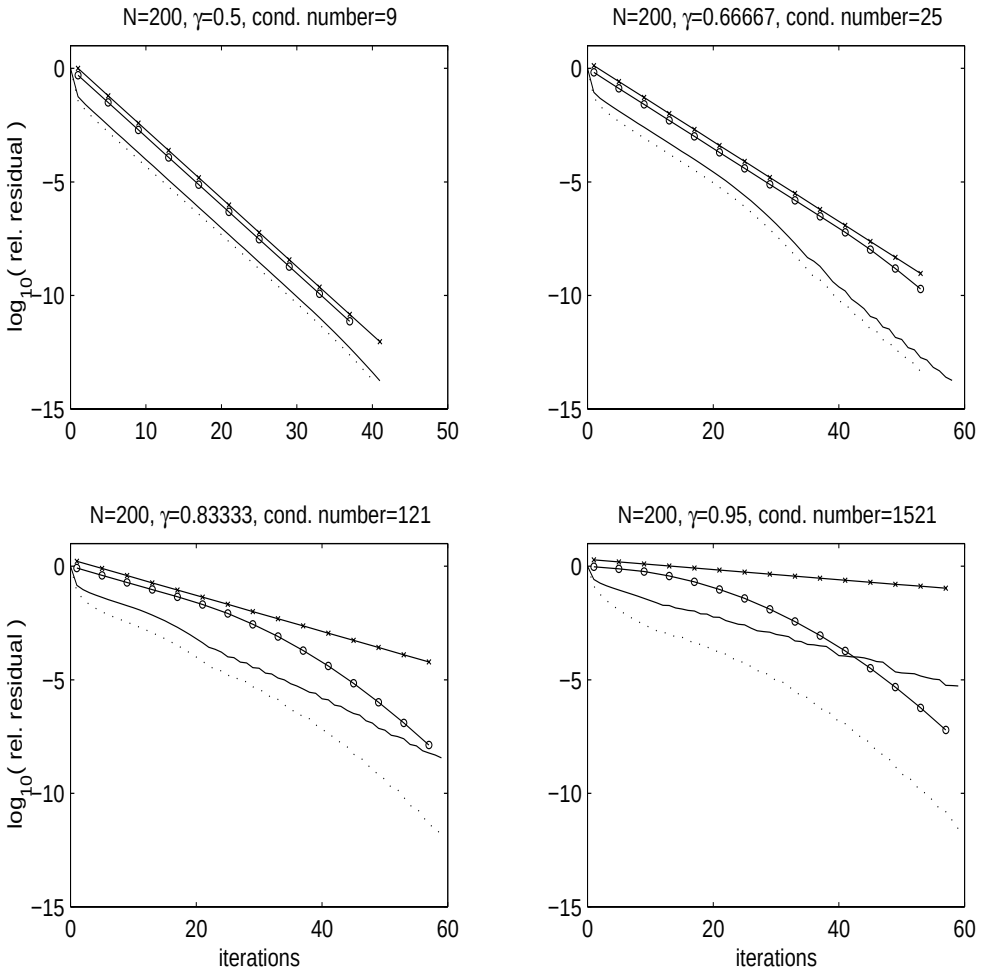

FIG. 2. *The error curve of CG (solid line) and GMRES (dotted line) versus the classical upper bound (crosses) and our asymptotic upper bound (circles) for the system* $T_{200}x = b$, *with random solution* x *and initial residual* $r_0 = (1, \ldots, 1)^T$. *Here* T_N *is the Kac, Murdock, and Szegő matrix of section* 4 *with parameter* $\gamma \in \{1/2, 2/3, 5/6, 19/20\}$.

for (x, y) in the unit square $0 < x, y < 1$ with Dirichlet boundary conditions on the boundary of the square. The usual five-point finite difference approximation on the uniform grid

$$(j/(m+1), k/(m+1)), \qquad j, k = 0, 1, \ldots, m+1,$$

leads to a linear system of size $N \times N$, where $N = m^2$. After rescaling, the coefficient matrix of the system may be written as a sum of Kronecker products

$$A_N = B_m \otimes I_m + I_m \otimes B_m, \tag{5.1}$$

where

$$B_m = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix}_{m\times m} \tag{5.2}$$

and I_m is the identity matrix of order m. It is well known and easy to verify that the eigenvalues of B_m are

$$\mu_k = 2 - 2\cos\frac{\pi k}{m+1}, \qquad k = 1, \dots, m,$$

and that the eigenvalues $\lambda_{j,k}$ of A_N are connected with the eigenvalues of B_m via

$$\lambda_{j,k} = \mu_j + \mu_k = 4 - 2\left(\cos\frac{\pi j}{m+1} + \cos\frac{\pi k}{m+1}\right), \qquad j,k = 1, \dots, m. \tag{5.3}$$

Since $\lambda_{j,k} = \lambda_{k,j}$, most of the eigenvalues have multiplicity at least 2. Also, $\lambda_{j,m+1-j} = 4$ for all $j = 1, \dots, m$, and the eigenvalue 4 has multiplicity m. We suspect that $N/2 + o(N)$ eigenvalues have multiplicity 2, which is confirmed by our numerical experiments presented below.

To calculate the asymptotic distribution of the eigenvalues $\lambda_{j,k}$ as $m \to \infty$, we first note that the eigenvalues μ_k of B_m are in $[0,4]$ and have the asymptotic density

$$v(\mu) = \frac{1}{\pi\sqrt{\mu(4-\mu)}}, \qquad 0 < \mu < 4, \tag{5.4}$$

as $m \to \infty$. The asymptotic density of the $\lambda_{j,k} = \mu_j + \mu_k$ is then given by the convolution of v with itself, i.e.,

$$w(\lambda) = \frac{1}{2}\int v(\lambda-\mu)v(\mu)\,d\mu, \qquad 0 < \lambda < 8, \tag{5.5}$$

where the factor $1/2$ is added in accordance with the multiplicities of the eigenvalues of A_N. The density w is symmetric around 4. For $\lambda \in (0,4)$, we have from (5.4) and (5.5)

$$w(\lambda) = \frac{1}{2\pi^2}\int_0^\lambda \frac{1}{\sqrt{\mu(\lambda-\mu)(4-\mu)(4-\lambda-\mu)}}\,d\mu. \tag{5.6}$$

In (5.6) we put $\lambda = 4 - 4x$ with $0 < x < 1$ and make the change of variables

$$\mu = \frac{\lambda x t}{1-(1-x)t}$$

to obtain

$$w(4-4x) = \frac{1}{8\pi^2}\int_0^1 \frac{1}{\sqrt{t(1-t)(1-(1-x^2)t)}}\,dt, \qquad 0 < x < 1. \tag{5.7}$$

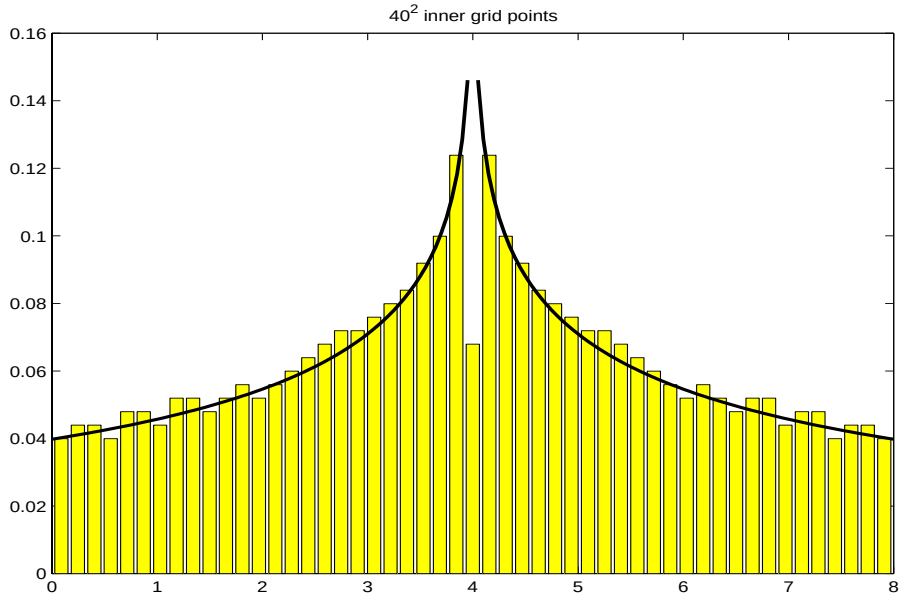

FIG. 3. *Bar chart of the eigenvalue distribution of the matrix A_{1600} (without counting multiplicities) resulting from discretizing the two-dimensional Poisson equation on a uniform grid with $N = 40^2$ inner points. The solid line corresponds to the asymptotic density function.*

By the Euler integral representation for hypergeometric functions, (5.7) is

$$w(4-4x) = \frac{1}{8\pi} F(1/2, 1/2; 1; 1-x^2) = \frac{1}{8\pi} \sum_{k=0}^{\infty} \frac{(1/2)_k (1/2)_k}{k!k!} (1-x^2)^k, \tag{5.8}$$

$0 < x < 1$, where $(1/2)_k$ is a Pochhammer symbol. It is interesting to observe that $4\pi^2 w(4-4x)$ equals the complete elliptic integral of the first kind, evaluated at $\sqrt{1-x^2}$; see [PrBrMa90, section 7.3.2., (75)]. Since w is symmetric around 4 and the right-hand side of (5.8) is even in x, the formula (5.8) holds for $-1 < x < 0$ as well.

In the series in the right-hand side of (5.8) each term is clearly decreasing as $x \in (0,1)$ increases. Thus $w(\lambda)$ is increasing for $\lambda \in (0,4)$, which is also clear from Figure 3. Then $\sqrt{\lambda(8-\lambda)}w(\lambda)$ also increases for $\lambda \in (0,4)$, and therefore the assumptions of Lemma 3.1(c) are satisfied. From Lemma 3.1(c) we then conclude that for every $t \in (0, \|\sigma\|) = (0, 1/2)$, the set $S(t)$ associated with $d\sigma(\lambda) = w(\lambda)d\lambda$ is

$$S(t) = [4-r, 4+r]$$

with $r \in (0,4)$ satisfying

$$t = 2\int_0^{4-r} \frac{4-\lambda}{\sqrt{(4-\lambda)^2 - r^2}} w(\lambda)\, d\lambda. \tag{5.9}$$

Putting $\lambda = 4 - 4x$ in (5.9) we have

$$t = 8\int_{r/4}^{1} \frac{x}{\sqrt{x^2 - (r/4)^2}} w(4-4x)\, dx. \tag{5.10}$$

Inserting the series (5.8) for $w(4-4x)$ and interchanging integration and summation, we find

$$t = \frac{1}{\pi} \sum_{k=0}^{\infty} \frac{(1/2)_k (1/2)_k}{k!k!} \int_{r/4}^{1} \frac{x}{\sqrt{x^2 - (r/4)^2}} (1-x^2)^k \, dx. \tag{5.11}$$

For each k, the integral is easily transformed to a beta-integral, and it follows that

$$\begin{aligned} \int_{r/4}^{1} \frac{x}{\sqrt{x^2 - (r/4)^2}} (1-x^2)^k \, dx &= \frac{k!\sqrt{\pi}}{\Gamma(k+3/2)} (1-(r/4)^2)^{k+1/2} \\ &= \frac{k!}{(3/2)_k} (1-(r/4)^2)^{k+1/2}; \end{aligned}$$

see also [Kuij00]. Inserting this in (5.11), we obtain

$$\begin{aligned} t &= \frac{1}{\pi} \sum_{k=0}^{\infty} \frac{(1/2)_k (1/2)_k}{(3/2)_k k!} (1-(r/4)^2)^{k+1/2} \\ &= \frac{1}{\pi} \sqrt{1-(r/4)^2} F(1/2, 1/2; 3/2; 1-(r/4)^2). \end{aligned} \tag{5.12}$$

This is a known series expansion for the arccos function

$$t = \frac{1}{\pi} \arccos(r/4);$$

see [PrBrMa90, section 7.3.2, (76)]. Inverting this we obtain the remarkably simple formula

$$r = 4\cos(\pi t), \tag{5.13}$$

and so

$$S(t) = [4 - 4\cos(\pi t), 4 + 4\cos(\pi t)]. \tag{5.14}$$

Finally, after a small calculation using (1.10) and (3.1) we obtain the convergence factor

$$\log \rho_t = -\frac{1}{t} \int_0^t \log\left(\tan\left(\frac{\pi}{4}(1+2\tau)\right)\right) d\tau, \qquad 0 < t < 1/2. \tag{5.15}$$

Notice that, for small t, the set $S(t)$ of (5.14) approximately equals the set obtained for equidistant eigenvalues on $[0, 8]$; compare this with section 3. This observation is in accordance with the behavior of the eigenvalues of A_N at the endpoints of $\operatorname{supp}(\sigma) = [0, 8]$.

Numerical results for the discretized two-dimensional Poisson equation are given in Figure 4. One might be curious about what CG error curve is obtained if other boundary conditions are imposed. In this case, we need to modify $\mathcal{O}(m)$ rows of A_N, and such "small rank" perturbations have been covered in Remark 2.5. However, since multiplicities are, in general, not preserved by such modifications, we need to have a closer look in order to obtain sharp error bounds.

In our case we can be more precise since again the eigenvalues can be computed explicitly for a number of configurations (see, e.g., [ChEl89]). For instance, in the

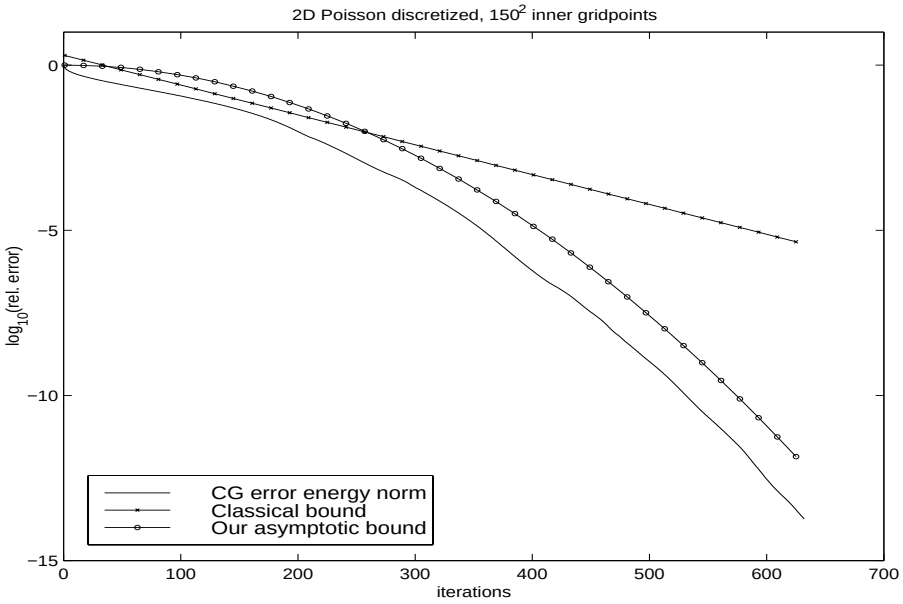

FIG. 4. *The CG error curve versus the two upper bounds for the system $A_N x = b$ resulting from discretizing the two-dimensional Poisson equation on a uniform grid with $N = 150^2$ inner points. We have chosen a random solution x and initial residual $r_0 = (1, \ldots, 1)^T$. As predicted by* (5.15)*, we obtain superlinear convergence from the beginning. Notice that the classical upper bound for CG is far too pessimistic for larger iteration indices. Similar plots are obtained for other mesh sizes.*

case of periodic boundary conditions, most of the eigenvalues are of multiplicity 8. Thus, in accordance with [ChEl89], the convergence behavior for Dirichlet boundary conditions with mesh size h is similar to the one obtained for periodic boundary conditions with mesh size $h/2$. In case of "no-flow" Neumann boundary conditions on the vertical boundaries $x = 0, x = 1$ discretized by a first-order scheme, the corresponding eigenvalues are given by (3.4) plus the m eigenvalues of B_m. Here we may expect the same convergence behavior as for Dirichlet boundary conditions.

Appendix. In the appendix we state and prove a lemma that is used in the proof of Theorem 2.1.

LEMMA 5.1. *Let σ be a finite Borel measure on $\mathbb{R}$ with compact support. Suppose $(\Lambda_N)_N$ is a sequence of sets, all contained in a fixed compact set, such that*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{\lambda\in\Lambda_N} f(\lambda) = \int f(\lambda) d\sigma(\lambda)$$

for all continuous functions f on $\mathbb{R}$.

Let $t \in (0, \|\sigma\|)$ and let μ be a Borel probability measure such that $t\mu \le \sigma$. Let $n = n_N \le \#\Lambda_N$ such that $n/N \to t$. Then there exists a sequence of sets $(Z_N)_N$ such that

(a) $\#Z_N = n$,
(b) $Z_N \subset \Lambda_N$, *and*
(c) *for all continuous functions f*

$$\lim_{N\to\infty} \frac{1}{n} \sum_{\lambda\in Z_N} f(\lambda) = \int f(\lambda) d\mu(\lambda).$$

Furthermore, if K is a closed set such that $\sigma(\partial K) = 0$ and $\sigma(K) = t\mu(K)$, then the sets Z_N can be chosen such that in addition to (a), (b), *and* (c), *we also have for N large enough*

(d) $\Lambda_N \cap K \subset Z_N$.

Proof. We have to prove that for some sets Z_N satisfying (a) and (b) the normalized counting measures

$$\zeta_N = \frac{1}{N} \sum_{\lambda \in Z_N} \delta_\lambda$$

converge in weak* sense to $t\mu$. To show this, we proceed in three steps.

Step 1. Suppose we have a finite partition of $\mathbb{R}$ consisting of measurable sets U_j, $j = 1, \dots, k$, such that $\sigma(\partial U_j) = 0$ for every j. Since the normalized counting measures of the sets Λ_N tend to σ, we then have for every j

$$\lim_{N\to\infty} \frac{\#(\Lambda_N \cap U_j)}{N} = \sigma(U_j).$$

Since $t\mu(U_j) \le \sigma(U_j)$, it is then possible to choose, for every j and N, a subset $Z_{N,j} \subset \Lambda_N \cap U_j$ such that

$$\lim_{N\to\infty} \frac{\#Z_{N,j}}{N} = t\mu(U_j).$$

The sets $Z_{N,j}$ are disjoint and for their union

$$Z_N^* := \bigcup_{j=1}^{k} Z_{N,j}$$

we have $Z_N^* \cap U_j = Z_{N,j}$. Hence

$$\lim_{N\to\infty} \frac{\#(Z_N^* \cap U_j)}{N} = t\mu(U_j).$$

Then also

$$\lim_{N\to\infty} \frac{\#Z_N^*}{N} = \sum_{j=1}^{k} \lim_{N\to\infty} \frac{\#(Z_N^* \cap U_j)}{N} = t$$

so that $\#Z_N^* = n + o(N)$ as $N \to \infty$. The set Z_N^* may not have exactly n elements. By adding or deleting $o(N)$ elements, we obtain from Z_N^* a set Z_N with exactly n elements. If we add elements, we choose them from Λ_N. Then $Z_N \subset \Lambda_N$ and the limits

$$\lim_{N\to\infty} \frac{\#(Z_N \cap U_j)}{N} = t\mu(U_j), \qquad j = 1, \dots, k, \tag{5.16}$$

hold.

Step 2. Now assume we have a finite collection U_j, $j = 1, \dots, k$, of measurable sets such that $\sigma(\partial U_j) = 0$ for all j. The sets U_j are not necessarily disjoint. For each $I \subset \{1, \dots, k\}$, we put

$$V_I = \left(\bigcap_{j\in I} U_j \right) \cap \left(\bigcap_{j\notin I} (\mathbb{R} \setminus U_j) \right).$$

The sets V_I with I ranging over all subsets of $\{1,\dots,k\}$ form a partition of $\mathbb{R}$. By Step 1 (see (5.16)) there exist sets Z_N such that $\#Z_N = n$, $Z_N \subset \Lambda_N$, and

$$\lim_{N\to\infty} \frac{\#(Z_N \cap V_I)}{N} = t\mu(V_I) \qquad \text{for all } I.$$

Since every U_j is a finite disjoint union of some of the V_I, it also follows that

$$\lim_{N\to\infty} \frac{\#(Z_N \cap U_j)}{N} = t\mu(U_j), \qquad j = 1,\dots,k. \tag{5.17}$$

Step 3. Now let U_j, $j = 1,2,\dots$, be a basis for the topology of $\mathbb{R}$, chosen such that $\sigma(\partial U_j) = 0$ for every j. From Step 2 we get for each k a sequence of sets $(Z_N^{(k)})_N$ such that $\#Z_N^{(k)} = n$, $Z_N^{(k)} \subset \Lambda_N$, and

$$\lim_{N\to\infty} \frac{\#(Z_N^{(k)} \cap U_j)}{N} = t\mu(U_j), \qquad j = 1,\dots,k;$$

see (5.17). Then by a diagonal argument, it is possible to find a sequence (k_N) tending to infinity, such that the sets Z_N defined by

$$Z_N := Z_N^{(k_N)}$$

satisfy

$$\lim_{N\to\infty} \frac{\#(Z_N \cap U_j)}{N} = t\mu(U_j), \qquad j = 1,2,\dots. \tag{5.18}$$

We also have

$$\#Z_N = n, \qquad Z_N \subset \Lambda_N \tag{5.19}$$

so that (a) and (b) hold.

Now, if ζ_N is the normalized counting measure of Z_N, then by (5.18) we have for every j,

$$\lim_{N\to\infty} \nu_N(U_j) = t\mu(U_j).$$

Since the U_j form a basis for the open sets, it follows that the measures ν_N tend in weak* sense to $t\mu$. Thus (c) holds.

Next, assume that K is a closed set such that $\sigma(\partial K) = 0$ and $\sigma(K) = t\mu(K)$. Then

$$\lim_{N\to\infty} \frac{\#(\Lambda_N \cap K)}{N} = \sigma(K)$$

and since also $\mu(\partial K) = 0$, we have because of (c)

$$\lim_{N\to\infty} \frac{\#(Z_N \cap K)}{N} = t\mu(K).$$

Since $\sigma(K) = t\mu(K)$, we then have

$$\#((\Lambda_N \setminus Z_N) \cap K) = o(N) \text{ as } N \to \infty.$$

Then we modify Z_N by adding the elements of $(\Lambda_N \setminus Z_N) \cap K$ to Z_N and removing $o(N)$ arbitrary elements from $Z_N \setminus K$. This is always possible for N large enough. Then clearly (d) is satisfied, while (a), (b), and (c) continue to hold.

This completes the proof of the lemma. ∎

REFERENCES

[Beck00] B. Beckermann, *On a conjecture of E.A. Rakhmanov*, Constr. Approx., 16 (2000), pp. 427–448.

[BeKu00] B. Beckermann and A.B.J. Kuijlaars, *On the Sharpness of an Asymptotic Error Estimate for Conjugate Gradients*, manuscript, 2000.

[BeSa99] B. Beckermann and E.B. Saff, *The sensitivity of least squares polynomial approximation*, in Internat. Ser. Numer. Math. 131, Birkhäuser, Basel, 1999, pp. 1–19.

[BoSi99] A. Böttcher and B. Silbermann, *Introduction to large truncated Toeplitz matrices*, Universitext, Springer-Verlag, New York, 1999.

[BuRa99] 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.

[ChEl89] T.F. Chan and H.C. Elman, *Fourier analysis of iterative methods for elliptic problems*, SIAM Rev., 31 (1989), pp. 20–49.

[ChNg96] R.H. Chan and M.K. Ng, *Conjugate gradient methods for Toeplitz systems*, SIAM Rev., 38 (1996), pp. 427–482.

[DrSa97] P.D. Dragnev and E.B. Saff, *Constrained energy problems with applications to orthogonal polynomials of a discrete variable*, J. Anal. Math., 72 (1997), pp. 223–259.

[DTT98] T.A. Driscoll, K.-C. Toh, and L.N. Trefethen, *From potential theory to matrix iterations in six steps*, SIAM Rev., 40 (1998), pp. 547–578.

[Gre79] A. Greenbaum, *Comparisions of splittings used with the conjugate gradient algorithm*, Numer. Math., 33 (1979), pp. 181–194.

[Gr97] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Frontiers Appl. Math. 17, SIAM, Philadelphia, 1997.

[GrTr94] A. Greenbaum and L.N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368.

[GrSz84] U. Grenander and G. Szegő, *Toeplitz Forms and Their Applications*, 2nd ed., Chelsea, New York, 1984.

[Joh00] K. Johansson, *Shape fluctuations and random matrices*, Comm. Math. Phys., 209 (2000), pp. 437–476.

[KaMuSz53] M. Kac, W. Murdock, and G. Szegő, *On the eigenvalues of certain Hermitian forms*, J. Rational Mech. Anal., 2 (1953), pp. 767–800.

[Kuij00] A.B.J. Kuijlaars, *Which eigenvalues are found by the Lanczos method?*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 306–321.

[KuDr99] A.B.J. Kuijlaars and P.D. Dragnev, *Equilibrium problems associated with fast decreasing polynomials*, Proc. Amer. Math. Soc., 127 (1999), pp. 1065–1074.

[KuRa98] A.B.J. Kuijlaars and E.A. Rakhmanov, *Zero distributions for discrete orthogonal polynomials*, J. Comput. Appl. Math., 99 (1998), pp. 255–274.

[KuVA99] A.B.J. Kuijlaars and W. Van Assche, *Extremal polynomials on discrete sets*, Proc. London Math. Soc. (3), 79 (1999), pp. 191–221.

[Mor97] I. Moret, *A note on the superlinear convergence of GMRES*, SIAM J. Numer. Anal., 34 (1997), pp. 513–516.

[Nev93] O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Birkhäuser, Basel, 1993.

[PrBrMa90] A.P. Prudnikov, Yu.A. Brychkov, and O.I. Marichev, *Integrals and Series*, Vol. 3, Gordon and Breach, New York, 1990.

[Ran95] T. Ransford, *Potential Theory in the Complex Plane*, Cambridge University Press, Cambridge, 1995.

[Rak96] 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.

[Saa96] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing, Boston, MA, 1996.

[SaTo97] E.B. Saff and V. Totik, *Logarithmic Potentials with External Fields*, Springer, Berlin, 1997.

[Ser00] 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), pp. 499-514.

[SeTi00] S. Serra Capizzano and P. Tilli, *From Partial Differential Equations to Generalised Locally Toeplitz Sequences*, manuscript, 2000.

[SlVS96] 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),

pp. 233–278.

[Stra00] Z. Strakos, *personal communication*, 2000.

[Sz67] G. Szegő, *Orthogonal Polynomials*, 3rd ed., Amer. Math. Soc. Colloq. Publ. 23, AMS, Providence, RI, 1967.

[Til98] P. Tilli, *Locally Toeplitz sequences: Spectral properties and applications*, Linear Algebra Appl., 278 (1998), pp. 91–120.

[Tre90] L.N. Trefethen, *Approximation theory and numerical linear algebra*, in Algorithms for Approximation II, J.C. Mason and M.G. Cox, eds, Chapman and Hall, London, 1990, pp. 336–360.

[TrBa97] L.N. Trefethen and D. Bau III, *Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

[VSVV86] A. van der Sluis and H.A. van der Vorst, *The rate of convergence of conjugate gradients*, Numer. Math., 48 (1986), pp. 543–560.

[Win80] R. Winther, *Some superlinear convergence results for the conjugate gradient method*, SIAM J. Numer. Anal., 17 (1980), pp. 14–17.