Electronic Transactions on Numerical Analysis.
Volume 65, pp. 271–280, 2026.

ISSN 1068–9613.
DOI: 10.1553/etna_vol65s271

ETNA
Kent State University and
Johann Radon Institute (RICAM)

AGAIN ABOUT THE NUMERICAL SOLUTION OF SYSTEMS OF LINEAR ALGEBRAIC EQUATIONS WITH SYMMETRIC NONSINGULAR NONDEFINITE MATRICES BY THE LANCZOS METHOD*

LEONID KNIZHNERMAN†

Abstract. The Lanczos method for approximating solutions of systems of linear algebraic equations with symmetric nondegenerate (possibly nondefinite) matrices is investigated. We recall our 2002 and 1995 theorems, which contain Cauchy–Hadamard-type estimates in terms of bounded self-adjoint operators in Hilbert spaces and the mth root of the residual or the error norm at step m; those theorems assert that a "reasonably good" bound holds at least at every other step. Using a slight and easy modification of the old proof, we prove a non-asymptotic upper bound for matrices in terms of the individual minimal residuals; this new result retains the every-other-step formulation. A lower residual bound for the case of even discrete spectral measures is also obtained.

Key words. Lanczos method, residual, minimal residual method (MINRES), system of linear algebraic equations with symmetric nondefinite matrix, upper and lower residual bounds

AMS subject classifications. 65F10

1. Introduction: statement of the problem and motivation.

1.1. Description of the problem. Let A be a real symmetric matrix of size $n \times n$ or a bounded self-adjoint operator in a Hilbert space $\mathcal{H}$ and b be a nonzero vector from $\mathbb{R}^n$ or $\mathcal{H}$, respectively. The Lanczos algorithm[1], applied to the pair (A, b) for m steps and using a three-term recurrence, builds[2] a basis $q_1, \dots, q_m$ of the Krylov subspace

$$\mathcal{K}^m(A, b) = \operatorname{span}\{A^0 b, A^1 b, \dots, A^{m-1} b\} \tag{1.1}$$

(see [15, Chapter 13]). After that it is possible to compute the Lanczos approximant

$$u_m = \|b\| Q f(H) e_1 \tag{1.2}$$

to the vector $u = f(A)b$; here f is a function, well defined on A and H, H is the projection of A onto the subspace (1.1) (the $m \times m$-matrix consisting of the recurrence coefficients), $e_1 = [1, 0, \dots, 0]^T$ is the first canonical unit vector from $\mathbb{R}^m$, and Q is the matrix of basis vectors, $Q = [q_1 \cdots q_m]$. See, for example, [5] or [8, Section 13.2].

In this paper we are interested in the "simplest" case $f(A) = A^{-1}$, i.e., in solving systems of linear algebraic equations (i.e., the Full Orthogonalization Method, for short FOM; see [16, Section 6.4]). Of course, it is assumed that the inverse matrix or the bounded inverse operator A^{-1} exists.

If $A > 0$, then (for matrices—in exact arithmetic) the Lanczos method produces the same approximants (see, e.g., [11, Item 3.1]) u_m as the Conjugate Gradient method from [7] (see also [16, Section 6.7]); the theory of the latter has been well developed. However, we shall concentrate on the subcase of a nondefinite matrix or operator A. The Lanczos process will

*Received March 24, 2025. Accepted June 3, 2026. Published online on July 3, 2026. Recommended by Vladimir L. Druskin. The work is supported by the Moscow Center of Fundamental and Applied Mathematics at INM RAS (agreement with the Ministry of Education and Science of the Russian Federation No 075-15-2025-347).

†Corresponding author. Marchuk Institute of Numerical Mathematics of RAS, Russia, 119333, Moscow, Gubkin St., 8 (lknizhnerman@gmail.com).

[1]By saying "the Lanczos method", we mean the Krylov one.

[2]This holds in the absence of early termination, which is a pleasant but improbable event. Formally, from the operator viewpoint, the matrix case with $m = n$ can also be treated as an early termination one.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

not yield approximants to $A^{-1}b$ at steps with a zero Ritz value (i.e., with a zero eigenvalue of the matrix H)[3]. Nevertheless, a residual estimate (though not for every step) can be obtained.

It is significant that the existence of Krylov subspace methods that minimize residuals such as the Minimal Residual method (for short MINRES, see [13]) and the Generalized Minimal Residual method (for short GMRES, see [17] or [16, Section 6.5]) does not make an investigation of the Full Orthogonalization Method, for short FOM (see [16, Section 6.4]), senseless.[4] It is said in [3, the first paragraph of Section 3.2] that "in many cases the error norms of FOM are smaller than those of GMRES"; a symmetric numerical example is presented in Figure 3.1 therein.[5] It is also said in [3, the last paragraph of Section 3.3], with a reference to [1], for the symmetric case that the convergence of FOM "tends to follow a general downward trend approximately matching the best possible approximation from the Krylov subspace".

1.2. Notation. Without loss of generality, we assume further on that the starting vector b is normalized: $\|b\| = 1$. We use the following notation and, in describing the terms related to the Lanczos process, we follow [15, Chapter 13]:

$$Q = \begin{bmatrix} q_1 & q_2 & \cdots & q_m \end{bmatrix} \in \mathbb{R}^{n\times m}$$

is the matrix consisting of the first m Lanczos vectors as columns, $Q^TQ = I_m$;

$$H = \begin{bmatrix} \alpha_1 & \beta_1 & & & & \\ \beta_1 & \alpha_2 & \beta_2 & & & \\ & \beta_2 & \alpha_3 & \beta_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_{m-2} & \alpha_{m-1} & \beta_{m-1} \\ & & & & \beta_{m-1} & \alpha_m \end{bmatrix} \in \mathbb{R}^{m\times m}$$

is the symmetric tridiagonal matrix of the coefficients of the Lanczos recurrence; $\beta_m = h_{m+1,m}$ is "the next" off-diagonal entry of the matrix $H = (h_{ij})$;

$$r_m = b - A \cdot QH^{-1}e_1$$

is the method's residual at step m (we assume that $\|r_m\| = +\infty$ if the matrix H is singular). Let

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n, \qquad (z_1, z_2, \ldots, z_n)$$

be the eigenvalues of A (and the corresponding orthonormal system of eigenvectors of A);

$$b = \sum_{j=1}^{n} b_j z_j, \qquad b_j \in \mathbb{R},$$

be the decomposition of the right-hand side b with respect to the eigenvectors;

$$\mu(\lambda) = \sum_{j=1}^{n} b_j^2 \delta(\lambda - \lambda_j)$$

[3]For a matrix H that is close to a degenerate one, a formally existing approximant u_m from (1.2) in some step may be of bad quality.

[4]FOM is the same as Arnoldi's method for SLAEs; the Lanczos method for SLAEs is a particular case of the former.

[5]In the paper [3] the nonsymmetric case has generally been considered: the methods FOM and GMRES are compared there.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

be a spectral measure (δ is the unit measure located at the point 0);

$$(p_0 = 1, p_1, \dots, p_{n-1})$$

be the system of polynomials that are orthonormal with respect to the measure μ with positive leading coefficients;

$$d = \frac{1}{\|A^{-1}\|} = \min_{j=1,\dots,n} |\lambda_j| \tag{1.3}$$

be the distance from zero to the spectrum of A, and let

$$\|r_m^{\mathrm{MR}}\| = \min_{u(X)\in\mathbb{R}[X],\ \deg u \le m-1} \|b - Au(A)b\|$$

be the residual norm of MINRES (see [13]).

1.3. Some existing results in the nonsymmetric case. Let $\|r_j^{\mathrm{F}}\|$ be the residual of FOM at a step j, and let $\|r_j^{\mathrm{G}}\|$ be the residual of GMRES at a step j. These residuals are related by the equality

$$\|r_m^{\mathrm{F}}\| = \frac{\|r_m^{\mathrm{G}}\|}{\sqrt{1 - \left(\|r_m^{\mathrm{G}}\| / \|r_{m-1}^{\mathrm{G}}\|\right)^2}} \tag{1.4}$$

(see [2] or [4, Formula (3) in Theorem 1]).[6]

The estimate

$$\min_{1\le j\le m} \|r_j^{\mathrm{F}}\| \le \sqrt{m}\cdot \|r_m^{\mathrm{G}}\| \tag{1.5}$$

is derived from (1.4) in [16, Proposition 6.15]. An estimate close to (1.5), as well as its unimprovability, was established in [3] (see Theorems 2.1 and 2.3 therein).

Evidently, estimate (1.5) invites the user to choose the best residual between the current step and all the previous Arnoldi steps. Of course, it is impossible to always take the current step since the corresponding approximant may not exist. Our estimate for the symmetric case, shown below, provides the user with the simplest possible rule: to test the current Lanczos step and the previous one.

1.4. Earlier results on the adaptation to the spectrum in terms of bounded operators in Hilbert spaces. Theorem 2 from the 2002 paper [10] is formulated as follows:

THEOREM 1.1. *Let A be a bounded self-adjoint operator with an infinite spectrum S in a Hilbert space, and let g be the generalized Green function for S.*[7] *If $0 \notin S$, then the inequality*

$$\varlimsup_{m\to\infty} \min\left(\|r_{m-1}\|, \|r_m\|\right)^{1/m} \le e^{-g(0)} \tag{1.6}$$

holds, where r_k is the residual of the approximant (1.2) for $f(A) = A^{-1}$ at a Lanczos step k.

[6] In the case of stagnation, $\|r_m^{\mathrm{G}}\| = \|r_{m-1}^{\mathrm{G}}\|$, the denominator in (1.4) is zero, and one sets $\|r_m^{\mathrm{F}}\| = +\infty$.
[7] Spectra and Green functions are described, e.g., in [12, Sections 2.5 and 5.5].

Earlier, in 1995 in the preprint [9], an analogous theorem (Theorem 1) was stated in terms of the error.[8]

The key point in (1.6) is that a natural residual bound of the Lanczos method holds at least *at every other step*. On the right-hand side in (1.6) one finds an asymptotic (in the Cauchy–Hadamard sense) worst-case convergence factor of MINRES.

However, many matrix algebraists prefer estimates in terms of the residual norm of MINRES $\|r_j^{\mathrm{MR}}\|$, and, moreover, in the direct, not asymptotic, sense. A request of this kind—an estimate of $\|r_m\|$ in terms of $\|r_j^{\mathrm{MR}}\|$ themselves— was initiated implicitly in the symmetric case in [3, the last paragraph of Section 3.1].

We shall modify in Section 2 our proof from [10] so that the result is stated in terms that are common for matrix algebraists while the quality does not decrease. To this end it is sufficient to apply moderate changes. In Section 3 a lower bound is given.

2. An upper residual bound.

In this section we formulate and prove an upper bound for the minimum of the Lanczos residual norms at two consecutive Lanczos steps.

THEOREM 2.1. *For $2 \le m \le n-1$, the inequality*[9]

$$\min(\|r_{m-1}\|, \|r_m\|) \le \frac{\sqrt{2}\,\beta_m}{d}\|r_{m-1}^{\mathit{MR}}\| \tag{2.1}$$

is valid.

Proof. The Lanczos residuals are related to the underlying orthonormal polynomials. Namely, [14, Formula (3.8)] can be rewritten in the form

$$r_j = \frac{p_j(A)b}{p_j(0)}.$$

Whence, thanks to orthonormality, for the residual norm, one has

$$\|r_j\| = \frac{1}{|p_j(0)|}. \tag{2.2}$$

Introducing the vector

$$\nu = \begin{bmatrix} 1 & p_1(0) & \dots & p_{m-1}(0) \end{bmatrix}^T \in \mathbb{R}^m \tag{2.3}$$

and using the equalities ($e_m = [0, \dots, 0, 1]^T$ is the mth canonical unit vector from $\mathbb{R}^m$)

$$H\nu = -\beta_m p_m(0) e_m \tag{2.4}$$

([15, Exercise 7.10.3]) and

$$AQ - QH = \beta_m q_{m+1} e_m^T$$

([15, Formula (13.1.1)], the Lanczos recurrence in matrix form), we obtain[10]

$$\begin{aligned} \|AQ\nu\| &= \|(QH + \beta_m q_{m+1} e_m^T)\nu\| = \|QH\nu + \beta_m q_{m+1}(e_m^T \nu)\| \\ &= \| - \beta_m p_m(0) q_m + \beta_m p_{m-1}(0) q_{m+1}\| = \beta_m \sqrt{p_{m-1}(0)^2 + p_m(0)^2}. \end{aligned} \tag{2.5}$$

[8] The term "adaptation to the spectrum" is caused by the fact that the right-hand side of the estimate (1.6) depends on the spectrum S.

[9] The condition number of A explicitly appears if one accounts for the fact that $\beta_m \le \|A\|$.

[10] The original versions of (2.3) and (2.4) are written in terms of the monic, not μ-normalized, polynomials. One should account for the fact that the leading coefficient of p_k is $1/(\beta_1 \cdots \beta_k)$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

In view of (1.3), we have

$$\|Q\nu\| = \|A^{-1}AQ\nu\| \le \|A^{-1}\| \cdot \|AQ\nu\| = \frac{1}{d}\|AQ\nu\|$$

or

$$\|AQ\nu\| \ge d\|Q\nu\|. \tag{2.6}$$

Combining equality (2.5) and inequality (2.6), we get

$$\beta_m\sqrt{p_{m-1}(0)^2 + p_m(0)^2} \ge d\|Q\nu\| = d\sqrt{\sum_{k=0}^{m-1} p_k(0)^2},$$

whence

$$\max(|p_{m-1}(0)|, |p_m(0)|) \ge \frac{d}{\sqrt{2}\,\beta_m}\sqrt{\sum_{k=0}^{m-1} p_k(0)^2}.$$

•[11] From this and from (2.2) we derive

$$\min(\|r_{m-1}\|, \|r_m\|) = \min\left(\frac{1}{|p_{m-1}(0)|}, \frac{1}{|p_m(0)|}\right) \le \frac{\sqrt{2}\,\beta_m}{d} \cdot \frac{1}{\sqrt{\sum_{k=0}^{m-1} p_k(0)^2}}. \tag{2.7}$$

It remains to apply [12, Chapter 2, Proposition 7.2]:

$$\begin{aligned}\left[\sum_{k=0}^{m-1} p_k(0)^2\right]^{-1} &\overset{[12]}{=} \min_{t(X)\in\mathbb{R}[X],\ \deg t\le m-1,\ t(0)=1} \int_{-\infty}^{+\infty} t(\lambda)^2\, d\mu(\lambda) \\ &= \min_{u(X)\in\mathbb{R}[X],\ \deg u\le m-2} \int_{-\infty}^{+\infty} (1-\lambda u(\lambda))^2\, d\mu(\lambda) = \|r_{m-1}^{\mathrm{MR}}\|^2.\end{aligned} \tag{2.8}$$

Equality (2.8) in conjunction with inequality (2.7) gives the estimate (2.1). ☐

It is evident that the estimate (2.1) of Theorem 2.1 corresponds to the promise made by us in Section 1.4: for any Lanczos step m, a "good" residual norm of the approximant u_m or the previous approximant u_{m-1} is guaranteed, and the estimate is not asymptotic.

REMARK 2.2. We reckon that ideologically the "every other step" statement is closely related to the fact that the Lanczos recurrence (and, equivalently, the recurrence for the associated orthonormal polynomials) is a three-term one. A technical explanation follows. Imagine for a minute that we allowed the matrix H to be upper Hessenberg, which is characteristic of the more general Arnoldi method. Then we would find that, for the derivation of formula (2.4) to hold, H must be tridiagonal symmetric, which implies the recurrences being three-term.

3. A lower residual bound. A natural question arises: is it possible to make our upper bound smaller than it is now? To what extent? For example, this could be related to investigating the precise exponent of d in the denominators, and we shall deal just with this parameter d in this section.

[11] This point is the end of the old part of the proof. We have marked the point to confirm that our present modification is slight.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

The next lemma yields "building material" for our "bad" example.

LEMMA 3.1. *Let us fix a number γ, $0 < \gamma < 1$. For the spectra $\{\lambda_1, \ldots, \lambda_n\}$ of the symmetric tridiagonal matrices of even size n,*

$$A_n = \begin{bmatrix} 0 & \frac{1+\gamma}{2} & & & \\ \frac{1+\gamma}{2} & 0 & \frac{1-\gamma}{2} & & \\ & \frac{1-\gamma}{2} & 0 & \frac{1+\gamma}{2} & \\ & & \ddots & \ddots & \ddots \\ & & & \frac{1+\gamma}{2} & 0 \end{bmatrix}, \tag{3.1}$$

the relations

$$\lim_{n\to\infty,\ n \text{ even}} \lambda_{n/2} = -\gamma, \qquad \lim_{n\to\infty,\ n \text{ even}} \lambda_{n/2+1} = \gamma$$

hold. In particular,

$$d = d_n \to \gamma, \qquad \textit{as } n \to \infty,\ n \textit{ even}. \tag{3.2}$$

Proof. According to [15, Formula (12.6.6)], the orthonormal polynomials p_j, induced by the Lanczos process with the matrix (3.1) of infinite size ($n = \infty$) and the initial vector $e_1 = [1, 0, 0, \ldots]^T$ in the space l_2, satisfy the recurrence relation

$$\begin{gathered} p_0(\lambda) = 1, \qquad \beta_0 \equiv 0, \\ \lambda p_{k-1}(\lambda) = \beta_{k-1} p_{k-2}(\lambda) + \alpha_k p_{k-1}(\lambda) + \beta_k p_k(\lambda), \quad k \geq 1, \end{gathered} \tag{3.3}$$

where the coefficients of the Lanczos recurrence (in the given case) are expressed by the equalities

$$\alpha_k = 0 \quad (k \geq 1), \qquad \beta_k = \begin{cases} \frac{1+\gamma}{2} & k \text{ odd}, \\ \frac{1-\gamma}{2} & k \text{ even}. \end{cases} \tag{3.4}$$

From this we derive, for $k \geq 1$,

$$\begin{aligned} \lambda^2 p_{2k}(\lambda) &= \beta_{2k} \cdot \lambda p_{2k-1}(\lambda) + \beta_{2k+1} \cdot \lambda p_{2k+1}(\lambda) \\ &= \beta_{2k}[\beta_{2k-1} p_{2k-2}(\lambda) + \beta_{2k} p_{2k}(\lambda)] \\ &\qquad + \beta_{2k+1}[\beta_{2k+1} p_{2k}(\lambda) + \beta_{2k+2} p_{2k+2}(\lambda)] \\ &= \beta_{2k-1}\beta_{2k} p_{2k-2}(\lambda) + (\beta_{2k}^2 + \beta_{2k+1}^2) p_{2k}(\lambda) + \beta_{2k+1}\beta_{2k+2} p_{2k+2}(\lambda) \\ &= \frac{1-\gamma^2}{4} \cdot p_{2k-2}(\lambda) + \frac{1+\gamma^2}{2} \cdot p_{2k}(\lambda) + \frac{1-\gamma^2}{4} \cdot p_{2k+2}(\lambda) \end{aligned} \tag{3.5}$$

and also

$$\lambda^2 p_0(\lambda) = \left(\frac{1+\gamma}{2}\right)^2 p_0(\lambda) + \frac{1-\gamma^2}{4} \cdot p_2(\lambda). \tag{3.6}$$

Noticing with the help of the recurrence (3.5)–(3.6) that the polynomials of even degrees p_{2k} are even and defining

$$\kappa = \lambda^2 \quad \text{and} \quad p_{2k}(\lambda) = s_k(\lambda^2), \qquad \deg s_k = k,$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

we derive the recurrence relation for the polynomials s_k

$$\kappa s_k(\kappa) = \begin{cases} \left(\frac{1+\gamma}{2}\right)^2 s_0(\kappa) + \frac{1-\gamma^2}{4}\cdot s_1(\kappa) & k=0, \\ \frac{1-\gamma^2}{4}\cdot s_{k-1}(\kappa) + \frac{1+\gamma^2}{2}\cdot s_k(\kappa) + \frac{1-\gamma^2}{4}\cdot s_{k+1}(\kappa) & k\geq 1. \end{cases}$$

The roots of the polynomial s_n are the eigenvalues of the symmetric tridiagonal $n\times n$-matrix

$$\begin{bmatrix} \left(\frac{1+\gamma}{2}\right)^2 & \frac{1-\gamma^2}{4} & & & \\ \frac{1-\gamma^2}{4} & \frac{1+\gamma^2}{2} & \frac{1-\gamma^2}{4} & & \\ & \frac{1-\gamma^2}{4} & \frac{1+\gamma^2}{2} & \frac{1-\gamma^2}{4} & \\ & & \ddots & \ddots & \ddots \end{bmatrix} = \frac{1-\gamma^2}{4}\begin{bmatrix} \frac{1+3\gamma}{1+\gamma} & 1 & & & \\ 1 & 2 & 1 & & \\ & 1 & 2 & 1 & \\ & & \ddots & \ddots & \ddots \end{bmatrix} + \gamma^2 I_n, \tag{3.7}$$

or

$$B_n = \frac{1-\gamma^2}{4}\cdot C_n + \gamma^2 I_n,$$

where B_n and C_n are the symmetric tridiagonal $n\times n$-matrices in (3.7) on the left-hand side and the first one on the right-hand side, respectively.

It is sufficient to demonstrate that (relative spectral problems are considered in [18, Section 19])

$$\lim_{n\to\infty}\lambda_{\min}(C_n) = 0. \tag{3.8}$$

Indeed, if (3.8) holds, then we have

$$\lim_{n\to\infty}\lambda_{\min}(B_n) = \gamma^2,$$

that is, the smallest root of the polynomial s_k tends to γ^2 as $k\to\infty$, whereas the roots of least modulus of p_{2k} tend to $\pm\gamma$. This implies (3.2).

The matrix C_n is nonnegative definite:

$$\begin{aligned} x^T C_n x &= \frac{1+3\gamma}{1+\gamma}x_1^2 + 2\sum_{j=2}^{n} x_j^2 + 2\sum_{j=1}^{n-1} x_j x_{j+1} \\ &= \frac{2\gamma}{1+\gamma}x_1^2 + x_n^2 + \sum_{j=1}^{n-1}(x_j+x_{j+1})^2 \geq 0, \qquad x=(x_1,\dots,x_n)^T. \end{aligned} \tag{3.9}$$

Using the simplest residual theorem (see [15, Equation (4.5.1)]) with the trial vector

$$x = \begin{bmatrix} 1 & -1 & \cdots & 1 & -1 \end{bmatrix}^T,$$

we derive

$$\min_{1\leq j\leq n}|\lambda_j(C_n)| \leq \frac{\|C_n x\|}{\|x\|} = \frac{\frac{2\gamma}{1+\gamma}}{\sqrt{n}} = o(1), \qquad n\to\infty.$$

This in conjunction with (3.9) leads to (3.8). $\square$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

THEOREM 3.2. *For symmetric (even) measures μ, there does not exist an estimate of the type*

$$\min(\|r_{m-1}\|, \|r_m\|) \le \frac{c\beta_m}{d^{\epsilon}} \|r_{m-1}^{MR}\|, \tag{3.10}$$

where

$$c > 0 \quad \text{and} \quad \epsilon < \frac{1}{2} \quad \text{are constants.} \tag{3.11}$$

Proof. Suppose that there exists a pair of parameters from (3.11) for which the estimate (3.10) is valid. As a counterexample we take the example from Lemma 3.1. Let us choose and fix γ, $0 < \gamma < 1$, such that

$$c\gamma^{1/2-\epsilon} < 1. \tag{3.12}$$

It is possible to set $m = 2k$, $n = 2m = 4k$.

The recurrence (3.3) with (3.4) implies, by induction, the equalities

$$|p_{2k}(0)| = \left(\frac{1+\gamma}{1-\gamma}\right)^k, \qquad p_{2k+1}(0) = 0, \qquad k \ge 0,$$

which in view of (2.2) gives

$$\min\{\|r_{m-1}\|, \|r_m\|\} = \min\left\{+\infty, \frac{1}{|p_{2k}(0)|}\right\} = \left(\frac{1-\gamma}{1+\gamma}\right)^k. \tag{3.13}$$

As $k \to \infty$,

$$\begin{aligned}
\sum_{l=0}^{2k-1} p_l(0)^2 = \sum_{l=0}^{k-1} p_{2l}(0)^2 = \sum_{l=0}^{k-1} \left(\frac{1+\gamma}{1-\gamma}\right)^{2l} &= \frac{\left(\frac{1+\gamma}{1-\gamma}\right)^{2k} - 1}{\left(\frac{1+\gamma}{1-\gamma}\right)^2 - 1} \\
&= \left(\frac{1+\gamma}{1-\gamma}\right)^{2k} (1 + o(1)) \cdot \frac{(1-\gamma)^2}{1 + 2\gamma + \gamma^2 - 1 + 2\gamma - \gamma^2} \\
&= \left(\frac{1+\gamma}{1-\gamma}\right)^{2k} (1 + o(1)) \cdot \frac{(1-\gamma)^2}{4\gamma},
\end{aligned}$$

whence, by virtue of (2.8) and (3.4),

$$\begin{aligned}
\beta_m \|r_{m-1}^{\mathrm{MR}}\| &= \beta_{2k} \left[\left(\frac{1+\gamma}{1-\gamma}\right)^{2k} \cdot \frac{(1-\gamma)^2}{4\gamma} \cdot (1 + o(1))\right]^{-1/2} \\
&= \frac{1-\gamma}{2} \cdot \left(\frac{1-\gamma}{1+\gamma}\right)^k \frac{2\sqrt{\gamma}}{1-\gamma} \cdot (1 + o(1)) \\
&= \left(\frac{1-\gamma}{1+\gamma}\right)^k \sqrt{\gamma} \cdot (1 + o(1)).
\end{aligned} \tag{3.14}$$

Suppose that the estimate (3.10) holds. In our concrete case, compare the actual residual and the corresponding hypothetical bound with (3.13) and (3.14):

$$\left(\frac{1-\gamma}{1+\gamma}\right)^k \le cd_n^{-\epsilon} \cdot \left(\frac{1-\gamma}{1+\gamma}\right)^k \sqrt{\gamma}(1 + o(1)),$$

or

$$c\sqrt{\gamma}d_n^{-\epsilon}(1+o(1)) \geq 1.$$

Letting k tend to infinity, in view of (3.2), we obtain

$$c\gamma^{1/2-\epsilon} \geq 1.$$

This contradicts inequality (3.12), which proves the theorem. □

4. Concluding remarks. Finally, we would like to comment on a few aspects of the subject.

REMARK 4.1. Theorem 2.1 extends to bounded self-adjoint operators. Indeed, we have exploited only the properties of orthonormal polynomials induced by the spectral measure, and nothing prevents the sequence of orthonormal polynomials from being infinite. In this sense, the matrix case is nothing but a "degenerate case" of the general theory. For this reason we worked with operators from the beginning, and we thus consider this topic to be covered.

REMARK 4.2. A matrix problem that coincides with the one considered here but is treated in finite-precision computer arithmetic was investigated in the paper [6]. The "every-other-step assertion" is present there, but the adaptation to the spectrum is unfortunately absent: a substantiation of the adaptation requires orthogonality, but orthogonality of the Lanczos vectors in finite-precision computer arithmetic is in general lost.

REMARK 4.3. In Section 3 we considered the case of even measures in order to exclude examples with positive definite matrices (or for matrices close to such in the sense of $\mu(-\infty, 0)$ being small), so as not to transfer the results here to the Conjugate Gradient method. Besides that, the restriction to a measure makes the lower bound stronger.

Acknowledgments. The author thanks Valeria Simoncini for a number of valuable comments. The author thanks the participants of a seminar at the Marchuk Institute of Numerical Mathematics of RAS for useful remarks and discussions.

REFERENCES

[1] N. AMSEL, T. CHEN, A. GREENBAUM, CAM. MUSCO AND CH. MUSCO, *Near-optimal approximation of matrix functions by the Lanczos method*, in Advances in Neural Information Processing Systems (NeurIPS 2024), A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, eds., Curran Associates, Boston, 2024, pp. 139823–139853.
Also: Preprint on arXiv, 2024. https://arxiv.org/abs/2303.03358

[2] P. N. BROWN, *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 58–78.

[3] T. CHEN AND G. MEURANT, *Near-optimal convergence of the full orthogonalization method*, Electron. Trans. Numer. Anal., 60 (2024), pp. 421–427.
https://etna.ricam.oeaw.ac.at/vol.60.2024/pp421-427.dir/pp421-427.pdf

[4] J. CULLUM AND A. GREENBAUM, *Relations between Galerkin and norm-minimizing iterative methods for solving linear systems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 223–247.

[5] V. L. DRUSKIN AND L. A. KNIZHNERMAN, *Two polynomial methods for the computation of functions of symmetric matrices*, U.S.S.R. Comput. Math. and Math. Phys., 29 (1989), pp. 112–121.

[6] A. GREENBAUM, V. L. DRUSKIN, AND L. A. KNIZHNERMAN, *On solving indefinite symmetric linear systems by means of the Lanczos method*, Zh. Vychisl. Mat. Mat. Fiz., 39 (1999), pp. 371–377.

[7] M. R. HESTENES AND E. STIEFEL, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards, 49 (1952), pp. 409–436.

[8] N. J. HIGHAM, *Functions of Matrices*, SIAM, Philadelphia, 2008.

[9] L. A. KNIZHNERMAN, *On adaptation of the Lanczos method to the spectrum*, Tech. Rep., Schlumberger-Doll Research, Res. Note EMG-001-95-12, MIT, Cambridge, 1995.

[10] ———, *Adaptation of the Lanczos and Arnoldi methods to the spectrum, or why the two Krylov subspace methods are powerful*, Chebyshevskiĭ Sb., 3 (2002), pp. 141–164.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[11] G. MEURANT AND Z. STRAKOŠ, *The Lanczos and conjugate gradient algorithms in finite precision arithmetic*, Acta Numer., 15 (2006), pp. 471–542.

[12] E. M. NIKISHIN AND V. N. SOROKIN, *Rational Approximations and Orthogonality*, Transl. Math. Monogr., 92, AMS, Providence, 1991.

[13] C. C. PAIGE AND M. A. SAUNDERS, *Solutions of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[14] C. C. PAIGE, B. N. PARLETT, AND H. A. VAN DER VORST, *Approximate solutions and eigenvalue bounds from Krylov subspaces*, Numer. Linear Algebra Appl., 2 (1995), pp. 115–133.

[15] B. N. PARLETT, *The Symmetric Eigenvalue Problem*, SIAM, Philadelphia, 1998.

[16] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.

[17] Y. SAAD AND M. H. SCHULTZ, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.

[18] V. V. VOEVODIN AND YU. A. KUZNETSOV, *Matrices and Computations*, Nauka, Moscow, 1984 (in Russian).