STABILITY OF LINEAR GMRES CONVERGENCE WITH RESPECT TO COMPACT PERTURBATIONS[*]

JAN BLECHTA[†]

Abstract. Suppose that a linear bounded operator B on a Hilbert space exhibits at least linear GMRES convergence, i.e., there exists $M_B < 1$ such that the GMRES residuals fulfill $\|r_k\| \leq M_B \|r_{k-1}\|$ for every initial residual r_0 and step $k \in \mathbb{N}$. We prove that GMRES with a compactly perturbed operator $A = B + C$ admits the bound $\|r_k\|/\|r_0\| \leq \prod_{j=1}^{k} \left(M_B + (1 + M_B) \|A^{-1}\| \sigma_j(C) \right)$, i.e., the singular values $\sigma_j(C)$ control the departure from the bound for the unperturbed problem. This result can be seen as an extension of [I. Moret, *SIAM J. Numer. Anal.*, 34 (1997), pp. 513–516], where only the case $B = \lambda I$ is considered. In this special case $M_B = 0$ convergence is superlinear.

Key words. GMRES, linear convergence, compact perturbation

AMS subject classification. 65F10

DOI. 10.1137/20M1340848

1. Introduction. For a bounded linear operator A on a Hilbert space, we consider the equation $Ax = b$ and its GMRES approximations x_k with residuals $r_k^{A,r_0} := b - Ax_k$, i.e., $\|r_k^{A,r_0}\| = \min_{p \in \mathcal{P}_k,\, p(0)=1} \|p(A)r_0\|$. In this paper we are concerned with convergence of the GMRES method for $A = B + C$, where B is such that it exhibits at least linear GMRES convergence and C is compact. The main result, the proof of which we postpone to section 4, is as follows.

THEOREM 1.1. *Let H be a complex separable Hilbert space. Suppose $A = B + C$ is invertible with B an invertible bounded linear operator on H and C a compact linear operator on H with singular values $\|C\| = \sigma_1(C) \geq \sigma_2(C) \geq \cdots \geq 0$. Further assume that B exhibits linear GMRES convergence, i.e, that B's linear reduction factor*

$$(1.1) \qquad M_B := \sup_{r_0 \in H} \sup_{k \in \mathbb{N}} \frac{\|r_k^{B,r_0}\|}{\|r_{k-1}^{B,r_0}\|}$$

satisfies $M_B < 1$. Then for any $r_0 \in H$ the GMRES residuals r_k^{A,r_0} fulfill

$$(1.2) \qquad \frac{\|r_k^{A,r_0}\|}{\|r_0\|} \leq \prod_{j=1}^{k} \left(M_B + (1 + M_B) \|A^{-1}\| \sigma_j(C) \right) \qquad \textit{for all } k \in \mathbb{N}$$

and

$$(1.3) \qquad \limsup_{k \to \infty} \frac{\|r_k^{A,r_0}\|}{\|r_{k-1}^{A,r_0}\|} \leq M_B.$$

Consequently, if there exists $p \geq 1$ such that

$$\|C\|_{\mathcal{S}_p} := \left(\sum_{j=1}^{\infty} \sigma_j(C)^p \right)^{\frac{1}{p}} < \infty,$$

[†]Faculty of Mathematics, Chemnitz University of Technology, Chemnitz, 09126 Germany (jan.blechta@math.tu-chemnitz.de).

then

$$(1.4) \qquad \left(\frac{\|r_k^{A,r_0}\|}{\|r_0\|}\right)^{\frac{1}{k}} \leq M_B + k^{-\frac{1}{p}}\left(1 + M_B\right)\|A^{-1}\|\,\|C\|_{\mathcal{S}_p} \qquad \text{for all } k \in \mathbb{N}.$$

A preliminary version of the result in Theorem 1.1 has been published in the thesis [3, Appendix III.B], where it has been used to study the pressure convection-diffusion preconditioner [7] for the solution of the Navier–Stokes equations. We expect the result to be applicable to analysis of a broad range of problems coming from the numerical solution of partial differential equations, where compact perturbations of "nice" operators often appear.

The linear reduction factor M_B as given by (1.1) is the best constant of the estimate

$$(1.5) \qquad \|r_k^{B,r_0}\| \leq M_B\|r_{k-1}^{B,r_0}\| \qquad \text{for all } k \in \mathbb{N} \text{ and all } r_0 \in H.$$

Obviously it holds $0 \leq M_B \leq 1$ but Theorem 1.1 is nontrivial only when $M_B < 1$. The result can be interpreted as a type of stability[1] of linear GMRES convergence with respect to compact perturbations. The bound (1.2) controls the departure from linear convergence (1.5) in terms of the singular values of C. Recall that $\sigma_j(C) \to 0$ as $j \to \infty$ for C compact. If additionally $\sum_{j=1}^{\infty} \sigma_j(C)^p$ is finite,[2] then, in view of (1.4), the mean departure from the linear convergence is superlinear $O\left(k^{-\frac{1}{p}}\right)$. Furthermore, if C has finite rank L, i.e., $\sigma_L(C) > \sigma_{L+1}(C) = 0$, then (1.2) easily implies

$$(1.6) \qquad \frac{\|r_k^{A,r_0}\|}{\|r_0\|} \leq \eta\, M_B^k \quad \text{for all } k \in \mathbb{N}, \qquad \eta \coloneqq \prod_{j=1}^{L}\left(1 + (1 + M_B^{-1})\|A^{-1}\|\,\sigma_j(C)\right)$$

if $M_B > 0$; the special case $M_B = 0$, which we omit, corresponds to a finite-rank perturbation of the identity.

For the sake of illustration we explicitly give a variant of Theorem 1.1 for the finite-dimensional case.

COROLLARY 1.2. *Let $B \in \mathbb{C}^{N \times N}$ be a nonsingular matrix and $C \in \mathbb{C}^{N \times N}$ be such that $A = B + C$ is nonsingular. Let $M_B \in [0,1]$ be given by (1.1) or, equivalently, let $M_B \in [0,1]$ be the best possible constant from the inequality (1.5). If $M_B < 1$ then (1.2) holds true and (1.4), in particular, implies that*

$$\left(\frac{\|r_k^{A,r_0}\|}{\|r_0\|}\right)^{\frac{1}{k}} \leq M_B + \frac{1}{\sqrt{k}}\left(1 + M_B\right)\|A^{-1}\|\,\|C\|_{\mathcal{S}_2} \qquad \text{for all } k \in \mathbb{N},$$

where $\|C\|_{\mathcal{S}_2}$ is the Frobenius matrix norm of C.

We will devote section 3 to obtaining an intrinsic characterization of M_B which does not depend on GMRES. We will show in Propositions 3.1 and 3.3 that

$$M_B = \min_{\lambda \in \mathbb{C}}\|I - \lambda B\| = \min_{\lambda \in \mathbb{C}}\|I - \lambda B^{-1}\|$$

---

[1]Here the term "stability" is *not* used in the sense of studying the behavior of GMRES in finite arithmetic precision, which itself is a broad research field [19, section 5.10].

[2]Note that $\|\cdot\|_{\mathcal{S}_p}$ in Theorem 1.1 is the norm (quasinorm) on the p-Schatten–von-Neumann ideal $\mathcal{S}_p(H)$ when $p \geq 1$ ($0 < p < 1$). The Hilbert–Schmidt norm $\|\cdot\|_{\mathcal{S}_2}$ takes the form of the Frobenius matrix norm in the finite-dimensional case.

and that the minima are attained on $\overline{\mathrm{Num}}\,B^{-1}$ and $\overline{\mathrm{Num}}\,B$, respectively, where $\overline{\mathrm{Num}}\,T$ stands for the closure of

$$\mathrm{Num}\,T \coloneqq \big\{(Tz, z),\, z \in H,\, \|z\| = 1\big\},$$

the numerical range (field of values) of T. The well-known sufficient condition for $M_B < 1$ that $0 \notin \overline{\mathrm{Num}}\,B$ is due to Elman's bound [6]

$$M_B \le \sqrt{1 - \frac{\nu_B^2}{\|B\|^2}}$$

and the improved bound by Starke [27], Eiermann, and Ernst [5],

$$M_B \le \sqrt{1 - \nu_B \nu_{B^{-1}}} \le \sqrt{1 - \frac{\nu_B^2}{\|B\|^2}},$$

where $\nu_T = \inf_{\lambda \in \mathrm{Num}\,T} |\lambda|$. (Note that $\nu_{T^{-1}} > 0$ if and only if $\nu_T > 0$ for invertible T.) If H is finite dimensional, then the linear convergence (1.5) with some $M_B < 1$ is equivalent to strictly monotone convergence for any initial residual, i.e.,

$$(1.7) \quad \|r_k^{B,r_0}\| < \|r_{k-1}^{B,r_0}\| \quad \text{or} \quad \|r_{k-1}^{B,r_0}\| = 0 \qquad \text{for every } k \in \mathbb{N} \text{ and every } r_0 \in H.$$

We will also show in Proposition 3.1 that (1.5) with $M_B < 1$ holds if and only if $0 \notin \overline{\mathrm{Num}}\,B$, that (1.7) holds if and only if $0 \notin \mathrm{Num}\,B$, and that the two situations are not, in general, equivalent unless H is finite dimensional or B is self-adjoint. In fact, we will give an example in Remark 3.2 of a unitary B with $0 \notin \mathrm{Num}\,B$ and $0 \in \overline{\mathrm{Num}}\,B$ so that B exhibits strictly monotone GMRES convergence (1.7) but does not exhibit linear GMRES convergence of any rate, i.e., the best M_B of (1.5) is $M_B = 1$.

The definition of M_B that we consider corresponds to the optimal constant in bounds of the form $\|r_k^{B,r_0}\| \le M_B^k \|r_0\|$. On the other hand, one might have a bound of the type $\|r_k^{B,r_0}\| \le \eta\, M_{B,\eta}^k \|r_0\|$ with some $\eta > 1$ and $M_{B,\eta} < 1$, which can be a better bound in the sense that $M_{B,\eta} < M_B$. These types of bounds appear frequently in the existing literature; see, e.g., [2, 20]. It is interesting to observe that we have already linked the two types of bounds by (1.5) and (1.6) for the case where $A - B$ is of finite rank. As a consequence, the bound $\|r_k^{A,r_0}\| \le M_A^k \|r_0\|$ with $M_A \le 1$ can be improved to the bound (1.6) with $M_B < M_A \le 1$ and $\eta > 1$ by considering a splitting $A = B + C$ with a finite-rank C. Nevertheless, in the remainder of this paper we focus on the case without $\eta > 1$ and will only shortly revisit the topic of splittings, albeit in a different context, in the closing section 5. Note that various splittings $A = B + C$ with B "good" in some sense and C of low rank were also considered by Huhtanen and Nevanlinna [16].

Theorem 1.1 can be seen as a generalization of Moret's result [22], which only considers $B = \lambda I$, $\lambda \ne 0$. Indeed, $M_{\lambda I} = 0$ and hence (1.3) gives the *Q-superlinear convergence* $\lim_{k \to \infty} \frac{\|r_k\|}{\|r_{k-1}\|} = 0$ of [22, Theorem 1] and (1.4) gives the rate $\|r_k\|^{\frac{1}{k}} = O\big(k^{-\frac{1}{p}}\big)$ of [22, eq. (1.1)]; only the estimate [22, inequality (2.7)] is finer than ours:

$$\frac{\|r_k^{A,r_0}\|}{\|r_0\|} \le \prod_{j=1}^k \sigma_j(A^{-1})\, \sigma_j(C) \le \prod_{j=1}^k \|A^{-1}\|\, \sigma_j(C);$$

the first inequality is due to Moret and the rightmost expression comes from the bound (1.2).

Under the conditions of Theorem 1.1, Hansmann's theorem [14, Theorem 2.1] provides a bound for the accumulation rate of discrete eigenvalues of A at the numerical range of B. A straightforward application of the theorem gives the estimate

$$\sum_{\lambda \in \sigma_{\mathrm{disc}}(A)} \mathrm{dist}\big(\lambda, \overline{\mathrm{Num}}\, B\big)^p \le \|C\|_{\mathcal{S}_p}^p \qquad \text{for any } p > 1,$$

where $\sigma_{\mathrm{disc}}(A)$ is the set of eigenvalues of A with finite algebraic multiplicity and each eigenvalue is counted in the sum according to its algebraic multiplicity. This estimate can be understood as the counterpart to the perturbation result of Theorem 1.1 concerning the perturbation of spectra.[3]

Nevanlinna [24, 25, 23], Malinen [21], and Hyvönen and Nevanlinna [17] deal with GMRES for compactly perturbed operators using techniques of complex analysis, which we completely avoid in this paper. On the other hand, Huhtanen and Nevanlinna [16] avoid complex analysis, consider splittings $A = B + C$ with B normal and C compact (or low rank), and obtain lower bounds on GMRES convergence speed.

The outline of the paper is as follows. In section 2 we fix the notation and gather some basic facts. Section 3 contains a characterization of linear convergence and the reduction factor M_B. In section 4, building on Moret's result [22], we prove Theorem 1.1. We briefly conclude and mention interesting questions this research gave rise to in section 5.

2. Preliminaries. Throughout the paper we assume that H is a complex Hilbert space with an inner product $(\cdot, \cdot)$ which is linear and antilinear in the first and second argument, respectively, and the symbol $\|\cdot\|$ stands for either the norm on H induced by $(\cdot, \cdot)$ or the induced operator norm on $\mathcal{L}(H)$, the space of bounded linear operators on H. For operator $A \in \mathcal{L}(H)$, initial residual $r_0 \in H$, and integer $k = 0, 1, 2, \ldots$ we define the Krylov subspace $\mathcal{K}_k(A, r_0) := \mathrm{span}\{r_0, Ar_0, A^2 r_0, \ldots, A^{k-1} r_0\} \subset H$. For a right-hand side $b \in H$ and an initial guess $x_0 \in H$, the initial residual is given by $r_0 = b - Ax_0$. Then the GMRES algorithm constructs a sequence $\{x_k\}_{k=1}^{\infty} \subset H$ given by minimizing the norm of residuals $r_k^{A,r_0} = b - Ax_k$ over $x_k \in x_0 + \mathcal{K}_k(A, r_0)$, i.e.,

$$(2.1) \qquad x_k = \underset{\substack{x_k \in H \\ x_k - x_0 \in \mathcal{K}_k(A, r_0)}}{\arg\min} \|b - Ax_k\|.$$

We will write just r_k instead of r_k^{A,r_0} if there is no risk of confusion.

Assume that $t_1, t_2, \ldots, t_k$ is the ascending orthonormal basis of the spaces $\mathcal{K}_k(A, r_0)$, $k = 1, 2, \ldots$, and $z_1, z_2, \ldots, z_k$ is the ascending orthornormal basis of the spaces $A\mathcal{K}_k(A, r_0)$, $k = 1, 2, \ldots$. This is well-defined if

$$\mathcal{K}_{k+1}(A, r_0) \supsetneq \mathcal{K}_k(A, r_0) \qquad \text{for all } k = 1, 2, \ldots.$$

It is well known that in the converse case, when $\mathcal{K}_{m+1}(A, r_0) = \mathcal{K}_m(A, r_0) \supsetneq \mathcal{K}_{m-1}(A, r_0) \supsetneq \cdots$ for certain m, the solution has been reached, i.e., $Ax_m = b$,

---

[3]Here, recall that concerning the finite-dimensional case the spectrum itself is not sufficient for obtaining any useful information about the behavior of GMRES [11, 10]. On the other hand, in the infinite-dimensional case the spectrum determines, in a certain sense, the behavior of GMRES at high iteration counts $k \to \infty$; see section 5 and [23]. The relationship between the finite-dimensional and infinite-dimensional cases is not, to our best knowledge, well established.

provided that A is invertible. To see this, observe that $A\mathcal{K}_m(A, r_0) \subset \mathcal{K}_{m+1}(A, r_0)$ but at the same time $\dim A\mathcal{K}_m(A, r_0) = \dim \mathcal{K}_m(A, r_0) = \dim \mathcal{K}_{m+1}(A, r_0)$ by the invertibility of A; hence $A\mathcal{K}_m(A, r_0) = \mathcal{K}_{m+1}(A, r_0) \ni r_0$, i.e., $r_0 = \hat{p}(A)Ar_0$ with some $\hat{p} \in \mathcal{P}_{m-1}$ which means that $r_m = r_0 - \hat{p}(A)Ar_0 = 0$; here and throughout the paper $\mathcal{P}_n$ stands for the space of polynomials of degree at most n.

Hence, if A is invertible, we can assume that $\{t_j\}_{j=1}^\infty$ and $\{z_j\}_{j=1}^\infty$ are extended orthonormal systems, i.e., either (i) $\{t_j\}_{j=1}^\infty$ and $\{z_j\}_{j=1}^\infty$ are orthonormal systems or (ii) $\{t_j\}_{j=1}^\infty$ and $\{z_j\}_{j=1}^\infty$ are such that $\{t_j\}_{j=1}^m$ and $\{z_j\}_{j=1}^m$ are orthonormal systems and $\{t_j\}_{j=m+1}^\infty$ and $\{z_j\}_{j=m+1}^\infty$ are zero sequences. In both cases $t_1, t_2, \ldots, t_k$ and $z_1, z_2, \ldots, z_k$ are orthonormal bases of $\mathcal{K}_k(A, r_0)$ and $A\mathcal{K}_k(A, r_0)$, respectively, for each $k \in \mathbb{N}$.

Recall that the numerical range of a bounded linear operator A on H is a subset of $\mathbb{C}$ given by

$$\operatorname{Num} A := \big\{ (Az, z),\, z \in H,\, \|z\| = 1 \big\}.$$

The distance of $\operatorname{Num} A$ to the origin is denoted

$$\nu_A := \inf_{\lambda \in \operatorname{Num} A} |\lambda|.$$

The closure of the numerical range, $\overline{\operatorname{Num} A}$, contains the spectrum of A. The Toeplitz–Hausdorff theorem [4, Theorem 9.3.1], [13] says that $\operatorname{Num} A$ is a convex set. The Lax–Milgram theorem guarantees that A is invertible provided $0 \notin \overline{\operatorname{Num} A}$ and in such a situation it holds that $\|A^{-1}\| \leq \nu_A^{-1}$.

3. Characterization of linear GMRES convergence.

PROPOSITION 3.1. *Let B be a bounded linear operator on a complex Hilbert space H. Further suppose that B is invertible. The linear reduction factor* (1.1) *fulfills*

$$(3.1) \qquad M_B = \inf_{\lambda \in \mathbb{C}} \|I - \lambda B\| = \inf_{\lambda \in \mathbb{C}} \|I - \lambda B^{-1}\| \leq \sqrt{1 - \nu_B \nu_{B^{-1}}} \leq \sqrt{1 - \frac{\nu_B^2}{\|B\|^2}}.$$

Furthermore, the following assertions are equivalent:
  (L1) *the operator B exhibits linear GMRES convergence* (1.5) *with some $M_B < 1$;*
  (L2) $\nu_B > 0$;
  (L3) $\nu_{B^{-1}} > 0$.
The following assertions are equivalent as well:
  (M1) *the operator B exhibits strictly monotone GMRES convergence* (1.7);
  (M2) $0 \notin \operatorname{Num} B$;
  (M3) $0 \notin \operatorname{Num} B^{-1}$.
If H is finite dimensional or B is self-adjoint then (L1), (L2), (L3), (M1), (M2), *and* (M3) *are equivalent.*

Proof. From (2.1) we have, for all $\lambda \in \mathbb{C}$, $r_0 \in H$, and $k \in \mathbb{N}$,

$$\|r_k\| = \min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(B)r_0\| \leq \|I - \lambda B\| \min_{\substack{p \in \mathcal{P}_{k-1} \\ p(0)=1}} \|p(B)r_0\| = \|I - \lambda B\|\, \|r_{k-1}\|.$$

Dividing by $\|r_{k-1}\|$, taking the infimum over $\lambda \in \mathbb{C}$ and the supremum over $r_0 \in H$ and $k \in \mathbb{N}$, we immediately obtain $M_B \leq \inf_{\lambda \in \mathbb{C}} \|I - \lambda B\|$. We continue by showing

the opposite inequality. The minimax theorem of Asplund and Pták [1] says that[4]

$$
(3.2) \qquad \inf_{\lambda \in \mathbb{C}} \|I - \lambda B\| = \inf_{\lambda \in \mathbb{C}} \sup_{\substack{z \in H \\ \|z\|=1}} \|z - \lambda Bz\| = \sup_{\substack{z \in H \\ \|z\|=1}} \inf_{\lambda \in \mathbb{C}} \|z - \lambda Bz\|.
$$

The minimization on the right-hand side is nothing other than the first GMRES step so that $\inf_{\lambda \in \mathbb{C}} \|I - \lambda B\| = \sup_{r_0 \in H} \frac{\|r_1\|}{\|r_0\|} \leq M_B$ and the first equality in (3.1) is thus proved. Furthermore, the minimization on the right-hand side of (3.2) can be carried out explicitly: for any $z \in H$, $z \neq 0$ it holds that

$$
(3.3) \qquad \min_{\lambda \in \mathbb{C}} \frac{\|z - \lambda Bz\|^2}{\|z\|^2} = 1 - \frac{|(Bz, z)|^2}{\|Bz\|^2 \|z\|^2}
$$

and the minimum is attained for $\lambda = \frac{(z, Bz)}{\|Bz\|^2}$. Equations (3.2) and (3.3) imply

$$
(3.4a) \qquad \inf_{\lambda \in \mathbb{C}} \|I - \lambda B\| = \sup_{z \in H} \sqrt{1 - \frac{|(Bz, z)|^2}{\|Bz\|^2 \|z\|^2}},
$$

$$
(3.4b) \qquad \inf_{\lambda \in \mathbb{C}} \|I - \lambda B^{-1}\| = \sup_{z \in H} \sqrt{1 - \frac{|(B^{-1}z, z)|^2}{\|B^{-1}z\|^2 \|z\|^2}},
$$

but by virtue of the invertibility of B, the right-hand sides in (3.4) are the same, which shows the second equality in (3.1). Using the inequalities

$$
(3.5a) \qquad \inf_{z \in H} \frac{|(Bz, z)|^2}{\|Bz\|^2 \|z\|^2} \geq \inf_{z \in H} \frac{|(Bz, z)|}{\|z\|^2} \inf_{z \in H} \frac{|(Bz, z)|}{\|Bz\|^2} = \nu_B \nu_{B^{-1}},
$$

$$
(3.5b) \qquad \nu_{B^{-1}} = \inf_{z \in H} \frac{|(B^{-1}Bz, Bz)|}{\|Bz\|^2} \geq \inf_{z \in H} \frac{|(z, Bz)|}{\|B\|^2 \|z\|^2} = \frac{\nu_B}{\|B\|^2},
$$

$$
(3.5c) \qquad \nu_B = \inf_{z \in H} \frac{|(BB^{-1}z, B^{-1}z)|}{\|B^{-1}z\|^2} \geq \inf_{z \in H} \frac{|(z, B^{-1}z)|}{\|B^{-1}\|^2 \|z\|^2} = \frac{\nu_{B^{-1}}}{\|B^{-1}\|^2},
$$

together with (3.4) immediately yields both inequalities in (3.1). Thus (3.1) is proved.

The inequalities (3.5b), (3.5c) establish the equivalence of (L2) and (L3). The assertion (L2) or (L3) implies (L1) by (3.1). We proceed by showing that (L1) implies (L2). Assume that (L2) is violated, i.e., that $\nu_B = 0$. Hence there exists $\{z_j\}_{j=1}^{\infty} \subset H$ such that $\|z_j\| = 1$ and $(Bz_j, z_j) \to 0$ as $j \to \infty$. Therefore $\frac{|(Bz_j, z_j)|}{\|Bz_j\|} \leq \|B^{-1}\| \, |(Bz_j, z_j)| \to 0$. Thus by (3.3) we obtain

$$
\inf_{\lambda \in \mathbb{C}} \|z_j - \lambda Bz_j\|^2 = 1 - \frac{|(Bz_j, z_j)|^2}{\|Bz_j\|^2} \to 1.
$$

Thus

$$
\sup_{\substack{r_0 \in H \\ \|r_0\|=1}} \inf_{\lambda \in \mathbb{C}} \|r_0 - \lambda Br_0\| = 1
$$

---

[4]The equality (3.2) expresses the "equivalence of true GMRES and ideal GMRES in the first step." This topic has received considerable attention in the GMRES literature; see, e.g., [9, 12, 18]. For our purpose, in the context of operators on a Hilbert space, we refer to [1], where (3.2) has been demonstrated to hold true if and only if H is an inner-product space, possibly infinite dimensional.

so that $M_B = 1$ and implication (L1)$\Rightarrow$(L2) is proved.

The equivalence of (M2) and (M3) follows directly from the definition of numerical range owing to the invertibility of B. We proceed by showing that (M2) implies (M1). We fix $k \in \mathbb{N}$. If $r_{k-1} = 0$ we are done so let us assume the converse. We have

$$\frac{\|r_k\|}{\|r_{k-1}\|} = \min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \frac{\|p(B)r_0\|}{\|r_{k-1}\|} \le \frac{\|(I - \lambda B)r_{k-1}\|}{\|r_{k-1}\|}$$

for any $\lambda \in \mathbb{C}$. Hence with λ minimizing the right-hand side we obtain, by (3.3) and (M2),

$$\frac{\|r_k\|}{\|r_{k-1}\|} \le \sqrt{1 - \frac{|(Br_{k-1}, r_{k-1})|^2}{\|Br_{k-1}\|^2 \|r_{k-1}\|^2}} < 1$$

so that (M1) follows. On the other hand, if (M2) is violated then there exists $r_0 \in H$ with $r_0 \ne 0$ and $(Br_0, r_0) = 0$ so that (3.3) immediately implies $\|r_1\| = \|r_0\|$, which contradicts (M1). Hence the equivalence of (M1) and (M2) is now clear.

If H is finite dimensional then $\operatorname{Num} B$ is closed and thus (L2) and (M2) are equivalent. Now assume B is self-adjoint. Clearly (L2) implies (M2). To prove the opposite, assume that $\nu_B = 0$ but $0 \notin \operatorname{Num} B$. As $\operatorname{Num} B$ is real and convex, it must be of the form $[a, 0)$, $(a, 0)$, $(0, b)$, or $(0, b]$ with some $a < 0$ or $b > 0$. On the other hand, the endpoints of $\operatorname{Num} B$ belong to the spectrum of B [15, Satz 1], which is always closed. Hence the conclusion is that 0 is in the spectrum, which contradicts the invertibility of B. $\qquad\square$

Remark 3.2. We give an example of a unitary operator fulfilling (M1) but contradicting (L1), thus showing that (L1) and (M1) are not in general equivalent. Let B be a diagonal operator on ℓ^2 given by

$$B = \sum_{j \in \mathbb{Z}} \lambda_j e_j(\cdot, e_j), \qquad \lambda_j = e^{i \arctan j} \text{ for } j \in \mathbb{Z},$$

where $\{e_j\}_{j \in \mathbb{Z}}$ is the canonical basis in ℓ^2, which is orthonormal with respect to $(\cdot, \cdot)$. The eigenvalues $\{\lambda_j\}_{j \in \mathbb{Z}}$ fulfill $|\lambda_j| = 1$ and $\operatorname{Re} \lambda_j > 0$ and, indeed, B is unitary and hence invertible. In fact, the spectrum of B is $\{\lambda_j\}_{j \in \mathbb{Z}} \cup \{i, -i\}$. Indeed, $\|(B \mp iI)e_k\| = |\lambda_k \mp i| \to 0$ as $k \to \pm\infty$, i.e., $\pm i$ is in the approximate point spectrum of B. For $f_k := e_k + e_{-k}$ it holds that $\frac{(Bf_k, f_k)}{\|f_k\|^2} = \operatorname{Re} \lambda_k \to 0$ as $k \to \pm\infty$. Hence $\nu_B = 0$, and thus (L2), and in turn (L1), do not hold true. On the other hand, $(Bz, z) = \sum_{j \in \mathbb{Z}} \lambda_j |(z, e_j)|^2$ which is nonzero whenever $z \ne 0$, because $\operatorname{Re} \lambda_j > 0$ for all $j \in \mathbb{Z}$. Thus $0 \notin \operatorname{Num} B$, which is (M2) so that (M1) holds true.

The following result characterizes λ that minimize (3.1).

PROPOSITION 3.3. *Let H be a complex Hilbert space and $B \in \mathcal{L}(H)$ be invertible. There exist $\lambda_B \in \mathbb{C}$ and $\lambda_{B^{-1}} \in \mathbb{C}$ such that the infima on the left-hand sides of (3.4a) and (3.4b), respectively, are attained. Any such λ_B and $\lambda_{B^{-1}}$ fulfill*

$$\lambda_B \in \overline{\operatorname{Num}} B^{-1}, \qquad\qquad \|\lambda_B B\| \le 1 + M_B,$$
$$\lambda_{B^{-1}} \in \overline{\operatorname{Num}} B, \qquad\qquad \|\lambda_{B^{-1}} B^{-1}\| \le 1 + M_B.$$

Proof. We will first prove that any minimizer of (3.4a), if it exists, fulfills $\lambda_B \in \overline{\operatorname{Num}} B^{-1}$. For the sake of contradiction assume that $\lambda_B \in \mathbb{C} \setminus \overline{\operatorname{Num}} B^{-1}$. Let $z \in H$

with $\|z\| = 1$ be arbitrary. The function $\lambda \mapsto \|z - \lambda Bz\|^2$ is minimized by $\lambda_z = \frac{(z, Bz)}{\|Bz\|^2} \in \operatorname{Num} B^{-1}$; see (3.3). Hence $|\lambda_B - \lambda_z| \geq C > 0$, where $C = \operatorname{dist}(\lambda_B, \operatorname{Num} B^{-1})$ is independent of z. Now consider that

$$\|z - \lambda_B Bz\|^2 = 1 - 2 \operatorname{Re}\big(\lambda_B (Bz, z)\big) + |\lambda_B|^2 \|Bz\|^2,$$
$$\|z - \lambda_z Bz\|^2 = 1 - \frac{|(Bz, z)|^2}{\|Bz\|^2},$$

where the second equality has already been shown in (3.3). Combining the last two equalities we obtain

$$\|z - \lambda_B Bz\|^2 - \|z - \lambda_z Bz\|^2 = \frac{|(Bz, z)|^2}{\|Bz\|^2} - 2 \operatorname{Re}\big(\lambda_B (Bz, z)\big) + |\lambda_B|^2 \|Bz\|^2$$
$$= \|Bz\|^2 |\lambda_B - \lambda_z|^2 \geq \|B^{-1}\|^{-2} C^2.$$

Hence

$$\|I - \lambda_B B\|^2 = \sup_{\|z\|=1} \|z - \lambda_B Bz\|^2 \geq \sup_{\|z\|=1} \|z - \lambda_z Bz\|^2 + \|B^{-1}\|^{-2} C^2$$
$$> \sup_{\|z\|=1} \|z - \lambda_z Bz\|^2 = \inf_{\lambda \in \mathbb{C}} \|I - \lambda B\|^2,$$

where we used (3.2) in the last equality, and this gives the desired contradiction. Hence any minimizers of (3.4a) belong to the compact set $\overline{\operatorname{Num}} B^{-1}$, where the function to be minimized is continuous, so the existence is also proved. By the triangle inequality we have $\|\lambda_B B\| \leq \|I\| + \|I - \lambda_B B\| = 1 + M_B$. The proof for $\lambda_{B^{-1}}$ is analogous. □

4. Proof of the main result. Moret [22, Lemma 6], as well as Eiermann and Ernst [5, Lemma 3.3], proves the following auxiliary result, which will be crucial for us.

LEMMA 4.1 (Moret's formula). *Let H be a separable complex Hilbert space, $A \in \mathcal{L}(H)$ be invertible, and $r_0 \in H$ be given. For every $k \in \mathbb{N}$ and $\lambda \in \mathbb{C}$ the GMRES residuals $r_k := r_k^{A, r_0}$ fulfill*

$$(4.1) \qquad \|r_k\| = |(t_{k+1}, z_k)| \|r_{k-1}\| = |(t_{k+1}, (I - \lambda A^{-1}) z_k)| \|r_{k-1}\|,$$

where $\{t_j\}_{j=1}^{\infty}$ and $\{z_j\}_{j=1}^{\infty}$ are the ascending extended orthonormal bases of $\{\mathcal{K}_1(A, r_0), \mathcal{K}_2(A, r_0), \ldots\}$ and $\{A\mathcal{K}_1(A, r_0), A\mathcal{K}_2(A, r_0), \ldots\}$, respectively.

Note that the second equality in (4.1) follows trivially from the definition of t_{k+1} and z_k; indeed $A^{-1} z_k \in \mathcal{K}_k(A, r_0)$, which is a space spanned by $\{t_1, \ldots, t_k\}$, and $(t_{k+1}, t_j) = 0$ for all $j \leq k$.

The approximation numbers of a bounded linear operator T on a Hilbert space H are defined by

$$(4.2) \qquad \sigma_j(T) = \inf_{\substack{M \in \mathcal{L}(H) \\ \operatorname{rank}(M) < j}} \|T - M\|, \qquad j = 1, 2, \ldots.$$

The approximation numbers satisfy

$$(4.3a) \qquad \|T\| = \sigma_1(T) \geq \sigma_2(T) \geq \cdots \geq 0,$$
$$(4.3b) \qquad \sigma_{j+k-1}(S + T) \leq \sigma_j(S) + \sigma_k(T), \qquad j, k = 1, 2, \ldots,$$
$$(4.3c) \qquad \sigma_j(WTU) \leq \|W\| \sigma_j(T) \|U\|, \qquad j = 1, 2, \ldots,$$

with any $S, T, U, W \in \mathcal{L}(H)$; see [26, paragraph 2.2.1, p. 79; Theorem 2.3.3, p. 83]. If T is compact, the numbers $\sigma_j(T)$ are the singular values of T.

LEMMA 4.2 (Pietsch [26, Lemma 2.11.13, p. 125]). *Let T be a bounded linear operator on a complex Hilbert space H. Let $\sigma_1(T) \geq \sigma_2(T) \geq \sigma_3(T) \geq \cdots \geq 0$ denote the approximation numbers of T as defined by* (4.2). *Then for any pair of orthonormal families $\{f_1, f_2, \ldots, f_k\}$, $\{g_1, g_2, \ldots, g_k\} \subset H$ it holds that*

$$(4.4) \qquad \det\{(Tf_i, g_j)\}_{i,j=1}^{k} \leq \prod_{j=1}^{k} \sigma_j(T).$$

Moret [22, inequality (2.7)] proves that if $A - \lambda I$ is compact for some $\lambda \in \mathbb{C}$, then

$$\frac{\|r_k\|}{\|r_0\|} \leq \prod_{j=1}^{k} \sigma_j(A - \lambda I)\,\sigma_j(A^{-1}).$$

We will need a modification, which follows.

LEMMA 4.3. *Suppose that the assumptions of Lemma* 4.1 *are fullfilled. Then for every $k \in \mathbb{N}$ and $\lambda \in \mathbb{C}$ the GMRES residuals $r_k := r_k^{A,r_0}$ fulfill*

$$(4.5) \qquad \frac{\|r_k\|}{\|r_0\|} \leq \prod_{j=1}^{k} \sigma_j(I - \lambda A^{-1}),$$

where $\sigma_1(I - \lambda A^{-1}) \geq \sigma_2(I - \lambda A^{-1}) \geq \cdots \geq 0$ are the approximation numbers of $I - \lambda A^{-1}$ as defined by (4.2).

Proof. From (4.1) we have

$$\frac{\|r_k\|}{\|r_0\|} = \prod_{j=1}^{k} |(t_{j+1}, (I - \lambda A^{-1})z_j)|.$$

The matrix

$$\left\{|(t_{i+1}, (I - \lambda A^{-1})z_j)|\right\}_{i,j=1}^{k}$$

is upper triangular because by construction,

$$0 = (t_{j+2}, z_j) = (t_{j+3}, z_j) = \cdots,$$
$$0 = (t_{j+1}, A^{-1}z_j) = (t_{j+2}, A^{-1}z_j) = \cdots.$$

This implies that $\prod_{j=1}^{k} |(t_{j+1}, (I - \lambda A^{-1})z_j)| = \det\{|(t_{i+1}, (I - \lambda A^{-1})z_j)|\}_{i,j=1}^{k}$ which is bounded by $\prod_{j=1}^{k} \sigma_j(I - \lambda A^{-1})$ due to (4.4). □

Proof of Theorem 1.1. By virtue of the invertibility of $A = B + C$ and the invertibility of B we have $A^{-1} = B^{-1} - B^{-1}CA^{-1}$. Hence by Lemma 4.1 we have, for any $\lambda \in \mathbb{C}$,

$$\frac{\|r_k^{A,r_0}\|}{\|r_{k-1}^{A,r_0}\|} \leq |(t_{k+1}, (I - \lambda B^{-1})z_k)| + |(t_{k+1}, \lambda B^{-1}CA^{-1}z_k)|$$

$$\leq \|I - \lambda B^{-1}\| + \|\lambda B^{-1}CA^{-1}z_k\|.$$

The sequence $\{z_k\}_{k=1}^{\infty}$ is an extended orthonormal system, so that Bessel's inequality

$$\sum_{k=1}^{\infty} |(y, z_k)|^2 \leq \|y\|^2 \qquad \text{for all } y \in H$$

immediately implies that $(y, z_k) \to 0$ as $k \to \infty$ for every $y \in H$, i.e., $\{z_k\}_{k=1}^{\infty}$ is weakly null. The compactness of C implies the strong convergence $\|CA^{-1}z_k\| \to 0$ and consequently

$$\limsup_{k\to\infty} \frac{\|r_k^{A,r_0}\|}{\|r_{k-1}^{A,r_0}\|} \leq \|I - \lambda B^{-1}\| \qquad \text{for any } \lambda \in \mathbb{C}.$$

By minimizing the right-hand side in λ and using the equality in (3.1) we obtain (1.3).

Using (4.5) and (4.3) we obtain, for any $\lambda \in \mathbb{C}$,

$$\frac{\|r_k^{A,r_0}\|}{\|r_0\|} \leq \prod_{j=1}^{k} \sigma_j(I - \lambda A^{-1}) \leq \prod_{j=1}^{k} \Big(\sigma_1(I - \lambda B^{-1}) + \sigma_j(\lambda B^{-1}CA^{-1}) \Big)$$

$$\leq \prod_{j=1}^{k} \Big(\|I - \lambda B^{-1}\| + \|\lambda B^{-1}\| \, \|A^{-1}\| \, \sigma_j(C) \Big).$$

Taking $\lambda \in \mathbb{C}$ minimizing $\|I - \lambda B^{-1}\|$, we obtain (1.2) using Proposition 3.3.

From (1.2) we obtain, using the inequality of arithmetic and geometric means and Hölder's inequality,

$$\left(\frac{\|r_k^{A,r_0}\|}{\|r_0\|} \right)^{\frac{1}{k}} \leq \tfrac{1}{k} \sum_{j=1}^{k} \Big(M_B + (1 + M_B) \|A^{-1}\| \, \sigma_j(C) \Big)$$

$$= M_B + \tfrac{1}{k} (1 + M_B) \|A^{-1}\| \sum_{j=1}^{k} \sigma_j(C)$$

$$\leq M_B + \tfrac{1}{k} (1 + M_B) \|A^{-1}\| \left(\sum_{j=1}^{k} \sigma_j(C)^p \right)^{\frac{1}{p}} k^{\frac{p-1}{p}},$$

which immediately yields (1.4). $\qquad\square$

5. Conclusion and outlook. We generalized Moret's result [22] in Theorem 1.1 and thus obtained a certain stability of linear GMRES convergence under compact perturbations. Indeed, the superlinear convergence of [22] follows as a corollary. In section 3 we obtained a useful characterization of linear convergence, which sheds some light on the limits of the applicability of Theorem 1.1.

Bound (1.3) has an interesting consequence:

$$(5.1) \qquad \limsup_{k\to\infty} \frac{\|r_k^{A,r_0}\|}{\|r_{k-1}^{A,r_0}\|} \leq \inf_{K \in \mathcal{C}(H)} M_{A+K} =: M_{\pi(A)},$$

where the infimum is taken over $\mathcal{C}(H)$, the ideal of all compact operators in $\mathcal{L}(H)$. Here π is the quotient map $\pi : \mathcal{L}(H) \to \mathcal{L}(H)/\mathcal{C}(H)$ with the quotient space known as the Calkin algebra and the quotient norm $\|\pi(T)\| = \inf_{K \in \mathcal{C}(H)} \|T + K\|$. In particular,

every finite-dimensional eigenspace can be "removed" from A when computing the Q-rate (5.1). Interestingly, for the R-rate, one has

$$\limsup_{k\to\infty}\left(\frac{\|r_k^{A,r_0}\|}{\|r_0\|}\right)^{\frac{1}{k}} \leq \lim_{k\to\infty}\inf_{\substack{p\in\mathcal{P}_k\\p(0)=1}}\|p(A)\|^{\frac{1}{k}} = \inf_{\substack{k\in\mathbb{N}\\p\in\mathcal{P}_k\\p(0)=1}}\max_{z\in\sigma_0(A)}|p(z)|^{\frac{1}{k}},$$

where the equality is due to Nevanlinna [23, Theorem 3.3.4] and $\sigma_0(A)$ consists of the spectrum of A except for its isolated points. By this analogy, one could conjecture that $M_{\pi(A)}$ also "cannot see" any eigenspaces, including those of infinite dimension. One might wonder whether there is a useful characterization of $M_{\pi(A)}$.

Another interesting question concerns s-step linear convergence $\|r_k\| \leq M\|r_{k-s}\|$ with some $s\in\mathbb{N}$ and $M < 1$. The present result assumes linear reduction in every step, i.e., $s = 1$. On the other hand, an important class of problems, e.g., self-adjoint operators, exhibit a reduction in every second step, i.e., $s = 2$ and $M < 1$. To this end, we cannot see how the technique of section 4 generalizes to $s \geq 2$. Note that [16] is concerned with related issues but does not directly provide such stability as we obtained for the case $s = 1$. Also note that pseudospectral techniques have been used to obtain bounds for the case $s > 1$; see, e.g., [8, 28].

REFERENCES

[1] E. ASPLUND AND V. PTÁK, *A minimax inequality for operators and a related numerical range*, Acta Math., 126 (1971), pp. 53–62, https://doi.org/10.1007/BF02392025.

[2] B. BECKERMANN, S. A. GOREINOV, AND E. E. TYRTYSHNIKOV, *Some remarks on the Elman estimate for GMRES*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 772–778, https://doi.org/10.1137/040618849.

[3] J. BLECHTA, *Towards Efficient Numerical Computation of Flows of Non-Newtonian Fluids*, Ph.D. thesis, Faculty of Mathematics and Physics, Charles University, Prague, Czechia, 2019, https://hdl.handle.net/20.500.11956/108384.

[4] E. B. DAVIES, *Linear Operators and their Spectra*, Cambridge Stud. Adv. Math., Cambridge University Press, Cambridge, 2007, https://doi.org/10.1017/CBO9780511618864.

[5] M. EIERMANN AND O. G. ERNST, *Geometric aspects of the theory of Krylov subspace methods*, Acta Numer., 10 (2001), pp. 251–312, https://doi.org/10.1017/S0962492901000046.

[6] H. ELMAN, *Iterative Methods for Large, Sparse, Nonsymmetric Systems of Linear Equations*, Ph.D. thesis, Yale University, New Haven, CT, 1982, http://ftp.cs.yale.edu/publications/techreports/tr229.pdf,

[7] H. C. ELMAN, D. J. SILVESTER, AND A. J. WATHEN, *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, 2nd ed., Oxford University Press, Oxford, 2014, https://doi.org/10.1093/acprof:oso/9780199678792.001.0001.

[8] M. EMBREE, *How descriptive are GMRES convergence bounds?*, Technical report NA-99-08, Oxford University Computing Laboratory, Oxford University, Oxford, 1999, https://ora.ox.ac.uk/objects/uuid:8ca2d383-4d7d-4e21-805c-98e16537d3d3.

[9] A. GREENBAUM AND L. GURVITS, *Max-min properties of matrix factor norms*, SIAM J. Sci. Comput., 15 (1994), pp. 348–358, https://doi.org/10.1137/0915024,

[10] A. GREENBAUM, V. PTÁK, AND Z. STRAKOŠ, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 465–469, https://doi.org/10.1137/S0895479894275030.

[11] A. GREENBAUM AND Z. STRAKOŠ, *Matrices that generate the same Krylov residual spaces*, in Recent Advances in Iterative Methods, IMA Vol. Math. Appl. 60, Springer, New York, 1994, pp. 95–118, https://doi.org/10.1007/978-1-4613-9353-5_7.

[12] A. GREENBAUM AND L. N. TREFETHEN, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368, https://doi.org/10.1137/0915025.

[13] K. GUSTAFSON, *The Toeplitz-Hausdorff theorem for linear operators*, Proc. Amer. Math. Soc., 25 (1970), pp. 203–204, https://doi.org/10.2307/2036559.

[14] M. HANSMANN, *An eigenvalue estimate and its application to non-selfadjoint Jacobi and Schrödinger operators*, Lett. Math. Phys., 98 (2011), pp. 79–95, https://doi.org/10.1007/s11005-011-0494-9.

[15] S. HILDEBRANDT, *Über den numerischen Wertebereich eines Operators*, Math. Ann., 163 (1966), pp. 230–247, https://doi.org/10.1007/BF02052287.

[16] M. HUHTANEN AND O. NEVANLINNA, *Minimal decompositions and iterative methods*, Numer. Math., 86 (2000), pp. 257–281, https://doi.org/10.1007/PL00005406.

[17] S. HYVÖNEN AND O. NEVANLINNA, *Robust bounds for Krylov methods*, BIT, 40 (2000), pp. 267–290, https://doi.org/10.1023/A:1022390923774.

[18] W. JOUBERT, *A robust GMRES-based adaptive polynomial preconditioning algorithm for nonsymmetric linear systems*, SIAM J. Sci. Comput., 15 (1994), pp. 427–439, https://doi.org/10.1137/0915029,

[19] J. LIESEN AND Z. STRAKOŠ, *Krylov subspace methods: Principles and analysis*, Numer. Math. Sci. Comput., Oxford University Press, Oxford, 2013, https://doi.org/10.1093/acprof:oso/9780199655410.001.0001.

[20] J. LIESEN AND P. TICHÝ, *The Field of Values Bounds on Ideal GMRES*, 2020, preprint, https://arxiv.org/abs/1211.5969v3.

[21] J. MALINEN, *On the Properties for Iteration of a Compact Operator with Unstructured Perturbation*, Technical report A360, Helsinki University of Technology, Institute of Mathematics, Espoo, Finland, 1996, https://math.aalto.fi/~jmalinen/MyPSFilesInWeb/SmallCompact.pdf.

[22] I. MORET, *A note on the superlinear convergence of GMRES*, SIAM J. Numer. Anal., 34 (1997), pp. 513–516, https://doi.org/10.1137/S0036142993259792.

[23] O. NEVANLINNA, *Convergence of Iterations for Linear Equations*, Lectures in Math. ETH Zürich, Birkhäuser, Basel, Switzerland, 1993, https://doi.org/10.1007/978-3-0348-8547-8.

[24] O. NEVANLINNA, *Convergence of Krylov methods for sums of two operators*, BIT, 36 (1996), pp. 775–785, https://doi.org/10.1007/BF01733791.

[25] O. NEVANLINNA, *Meromorphic functions and linear algebra*, Fields Inst. Monogr. 18, American Mathematical Society, Providence, RI, 2003, https://doi.org/10.1090/fim/018.

[26] A. PIETSCH, *Eigenvalues and s-Numbers*, Cambridge Stud. Adv. Math. 13, Cambridge University Press, Cambridge, 1987.

[27] G. STARKE, *Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems*, Numer. Math., 78 (1997), pp. 103–117, https://doi.org/10.1007/s002110050306.

[28] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra. The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, NJ, 2005.