SIAM J. MATRIX ANAL. APPL.
Vol. 35, No. 2, pp. 599–618
© 2014 Society for Industrial and Applied Mathematics

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

ON A PERTURBATION BOUND FOR INVARIANT SUBSPACES OF MATRICES*

MICHAEL KAROW† AND DANIEL KRESSNER‡

Abstract. Given a nonsymmetric matrix A, we investigate the effect of perturbations on an invariant subspace of A. The result derived in this paper differs from Stewart's classical result and sometimes yields tighter bounds. Moreover, we provide norm estimates for the remainder terms in well-known perturbation expansions for invariant subspaces, eigenvectors, and eigenvalues.

Key words. invariant subspaces, perturbation theory, pseudospectra, quadratic matrix equation

AMS subject classifications. Author must provide

DOI. 10.1137/130912372

1. Introduction.

A subspace $\mathcal{X} \subset \mathbb{C}^n$ of a matrix $A \in \mathbb{C}^{n\times n}$ is called *invariant* if it satisfies

$$A\mathcal{X} \subset \mathcal{X}. \tag{1.1}$$

In this paper, we reconsider the classical question of estimating the impact of perturbations in A on $\mathcal{X}$.

Suppose that the columns $X \in \mathbb{C}^{n\times k}$ form an orthonormal basis of $\mathcal{X}$. Then (1.1) implies the existence of $A_{11} \in \mathbb{C}^{k\times k}$ such that $AX = XA_{11}$. The eigenvalues of A_{11} are independent of the choice of basis and constitute the spectrum of the restriction of A to $\mathcal{X}$. Extending X to a unitary matrix $[X, X_\perp]$ leads to the block Schur decomposition

$$A\begin{bmatrix}X, X_\perp\end{bmatrix} = \begin{bmatrix}X, X_\perp\end{bmatrix}\begin{bmatrix}A_{11} & A_{12}\\ 0 & A_{22}\end{bmatrix}. \tag{1.2}$$

Note that this implies $\Lambda(A) = \Lambda(A_{11}) \cup \Lambda(A_{22})$, where $\Lambda(\cdot)$ denotes the spectrum of a matrix. Throughout this paper, we will assume

$$\Lambda(A_{11}) \cap \Lambda(A_{22}) = \emptyset. \tag{1.3}$$

This is a necessary and sufficient condition for the Lipschitz continuity of $\mathcal{X}$ with respect to perturbations in A [9, Thm. 15.5.1]. (Note that continuity requires a substantially weaker condition [9, Thm. 15.2.1].) Hence, if (1.3) holds, adding a small perturbation $A \mapsto A + E$ implies a change in the invariant subspace that is asymptotically proportional to $\|E\|$.

Various bounds on the change of invariant subspaces under perturbations of A have been derived, notably by Davis and Kahan [6], Stewart [18, 19], Demmel [8], and Sun [24]. In the general nonsymmetric case, these bounds are valid only as long as $\|E\|$ remains sufficiently small. A minimal requirement is that the separation condition (1.3) remains valid under perturbations. In the language of pseudospectra, this means that $\|E\|$ should stay below the critical perturbation level ε for which

*Received by the editors March 7, 2013; accepted for publication (in revised form) by B. Sutton March 17, 2014; published electronically May 13, 2014.
http://www.siam.org/journals/simax/35-2/91237.html
†Institut für Mathematik, TU Berlin, 10623 Berlin, Germany (karow@math.tu-berlin.de).
‡MATHICSE, EPF Lausanne, CH-1015 Lausanne, Switzerland (daniel.kressner@epfl.ch).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

the components of the ε-pseudospectrum containing $\Lambda(A_{11})$ and $\Lambda(A_{22})$ first meet each other [1]. It turns out that some existing perturbation results are unnecessarily restrictive and require $\|E\|$ to stay significantly below this critical level. The main contribution of this paper consists of a novel perturbation bound for invariant subspaces; see Theorem 3.1 below. To derive this bound, we employ pseudospectral techniques in the analysis of a quadratic matrix equation.

The rest of this paper is organized as follows. In section 2, we introduce the basic tools for the perturbation analysis of invariant subspaces and recall some existing results. Section 3 contains the statement and proof of our main result, a new perturbation bound for invariant subspaces. In section 3.2, it is shown that this bound is sharp for a 2×2 example. Section 3.3 discusses a variation of the main result based on the block diagonalization of A, while section 3.5 provides a comparison to existing perturbation bounds. In section 4, the former is used to quantify existence conditions and remainder terms for well-known eigenvalue and eigenvector expansions.

2. Preliminaries and existing results. The goal of this section is to summarize some existing perturbation results for invariant subspaces and introduce notation, needed in the rest of the paper, on the way. Let us first recall some basic tools used in the perturbation analysis.

2.1. Separation between matrices. The condition (1.3) can be quantified by the *separation* between A_{11} and A_{22}. Based on Varah's original definition [27], Demmel [8] has proposed

$$\mathrm{sep}_\lambda(A_{11}, A_{22}) := \sup\{\varepsilon > 0 \mid \quad \Lambda(A_{11} + E_{11}) \cap \Lambda(A_{22} + E_{22}) = \emptyset \\ \forall\, E_{11}, E_{22} \text{ with } \max\{\|E_{11}\|_2, \|E_{22}\|_2\} \le \varepsilon\}. \tag{2.1}$$

This definition has an important interpretation in terms of ε-pseudospectra, defined as

$$\Lambda_\varepsilon(M) = \left\{z \in \mathbb{C} \mid z \in \Lambda(M + E) \text{ for some } E \in \mathbb{C}^{n\times n} \text{ with } \|E\|_2 \le \varepsilon\right\}$$

for a matrix $M \in \mathbb{C}^{n\times n}$ [26]. The separation (2.1) is the minimum value of ε such that $\Lambda_\varepsilon(A_{11}) \cap \Lambda_\varepsilon(A_{22})$ is nonempty. This interpretation yields the expression

$$\mathrm{sep}_\lambda(A_{11}, A_{22}) = \inf_{\lambda\in\mathbb{C}} \max\left\{\sigma_{\min}(A_{11} - \lambda I), \sigma_{\min}(A_{22} - \lambda I)\right\},$$

where $\sigma_{\min}(\cdot)$ denotes the smallest singular value of a matrix. Based on this expression, an algorithm for computing sep_λ has been developed by Gu and Overton [11].

Although sep_λ is not an appropriate measure to quantify the perturbation of invariant subspaces, it will still play a role in our derivations. Stewart [19] has introduced a different notion of separation based on the observation that (1.3) is satisfied if and only if the Sylvester operator

$$\mathbb{T} : \mathbb{C}^{(n-k)\times k} \to \mathbb{C}^{(n-k)\times k}, \qquad \mathbb{T} : Z \mapsto ZA_{11} - A_{22}Z$$

is nonsingular. The separation of A_{11} and A_{22} with respect to an arbitrary norm $\|\cdot\|$ is defined as

$$\mathrm{sep}(A_{11}, A_{22}) := \min_{\|Z\|=1} \|\mathbb{T}(Z)\| = \min_{\|Z\|=1} \|ZA_{11} - A_{22}Z\|. \tag{2.2}$$

Then $\mathrm{sep}(A_{11}, A_{22}) \neq 0$ if and only if (1.3) holds. Examples in [8, 27] show that the quantity $\mathrm{sep}(A_{11}, A_{22})$ can be magnitudes smaller than $\mathrm{sep}_\lambda(A_{11}, A_{22})$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

In the following, $\mathrm{sep}_F(A_{11}, A_{22}) = \min_{\|Z\|_F=1} \|ZA_{11} - A_{22}Z\|_F$ denotes the separation with respect to the Frobenius norm. An efficient algorithm for estimating $\mathrm{sep}_F(A_{11}, A_{22})$ can be derived from the inverse power method [4]. Various lower and upper bounds for $\mathrm{sep}_F(A_{11}, A_{22})$ can be found in [22, 23, 28]. Lower bounds for other norms are discussed in [29].

Proposition 2.1 below relates sep with sep_λ and the distance between the spectra of A_{11} and A_{22}. Most of its statements are well known; we have included the proof for the convenience of the reader. Recall that a norm $\|\cdot\|$ is said to be unitarily invariant if $\|UZV\| = \|Z\|$ for all unitary matrices U, V of compatible size. A norm is unitarily invariant if and only if

$$\|XZY\| \le \|X\|_2 \|Z\| \, \|Y\|_2 \tag{2.3}$$

for all matrices of compatible size.

PROPOSITION 2.1. *Let* $A_{11}, E_{11} \in \mathbb{C}^{k\times k}$, $A_{22}, E_{22} \in \mathbb{C}^{(n-k)\times(n-k)}$, *and* $\lambda \in \mathbb{C}$. *Let* $\delta = \min\{\, |\lambda-\mu| \;|\; \lambda \in \Lambda(A_{11}), \mu \in \Lambda(A_{22}) \,\}$ *denote the distance between the spectra* $\Lambda(A_{11})$ *and* $\Lambda(A_{22})$.

(a) *For any norm,* $\mathrm{sep}(A_{11}, A_{22}) \le \delta$.

(b) *For any unitarily invariant norm,*

(i) $\mathrm{sep}(\lambda I,\, A_{22}) = \sigma_{\min}(\lambda I - A_{22})$,

(ii) $\mathrm{sep}(A_{11} + E_{11}, A_{22} + E_{22}) \ge \mathrm{sep}(A_{11}, A_{22}) - \|E_{11}\|_2 - \|E_{22}\|_2$,

(iii) $\mathrm{sep}(A_{11}, A_{22}) \le 2 \cdot \mathrm{sep}_\lambda(A_{11}, A_{22}) \le \delta$.

(c) *If* A_{11} *and* A_{22} *are normal matrices, then* $\mathrm{sep}_F(A_{11}, A_{22}) = 2 \cdot \mathrm{sep}_\lambda(A_{11}, A_{22}) = \delta$.

Proof. Consider vectors x and y^* such that $\|x\|_2 = \|y\|_2 = 1$, $y^*A_{11} = \lambda\, y^*$, and $A_{22}x = \mu\, x$ for $\lambda \in \Lambda(A_{11})$, $\mu \in \Lambda(A_{22})$. Let $Z = xy^*$. Then $ZA_{11} - A_{22}Z = (\lambda-\mu)Z$. This implies (a). To show the statements of (b), let $E_{11} = \frac{1}{2}(\mu - \lambda)yy^*$ and $E_{22} = \frac{1}{2}(\lambda-\mu)xx^*$. Then $y^*(A_{11}+E_{11}) = \frac{1}{2}(\lambda+\mu)y^*$ and $(A_{22}+E_{22})x = \frac{1}{2}(\lambda+\mu)x$. Thus, $\Lambda(A_{11} + E_{11}) \cap \Lambda(A_{22} + E_{22}) \ne \emptyset$ for perturbations satisfying $\|E_{11}\|_2 = \|E_{22}\|_2 = |\lambda-\mu|/2$. This yields the second inequality in (b)(iii). The statement of (b)(ii) follows from the inequality

$$\|Z(A_{11} + E_{11}) - (A_{22} + E_{22})Z\| \ge \|ZA_{11} - A_{22}Z\| - (\|E_{11}\|_2 + \|E_{22}\|_2)\|Z\|.$$

In particular, if $\max\{\|E_{11}\|_2, \|E_{22}\|\} < \mathrm{sep}(A_{11}, A_{22})/2$, then $\mathrm{sep}(A_{11} + E_{11}, A_{22} + E_{22}) > 0$. Thus, $\Lambda(A_{11} + E_{11}) \cap \Lambda(A_{22} + E_{22}) = \emptyset$. This yields the first inequality of (b)(iii).

Next, we show (c). Suppose that A_{11} and A_{22} are normal, and let U, V be unitary matrices such that $A_{11} = U\mathrm{diag}(\lambda_1, \ldots, \lambda_k)U^*$, $A_{22} = V\mathrm{diag}(\mu_1, \ldots, \mu_{n-k})V^*$, where λ_j and μ_i are the eigenvalues of A_{11} and A_{22}, respectively. Let $W = V^*ZU = [w_{ij}]$. Then

$$\begin{aligned}
\|ZA_{11} - A_{22}Z\|_F^2 &= \|V^*(ZA_{11} - A_{22}Z)U\|_F^2 \\
&= \|W\mathrm{diag}(\lambda_1, \ldots, \lambda_k) - \mathrm{diag}(\mu_1, \ldots, \mu_{n-k})W)\|_F^2 \\
&= \|\, [(\lambda_j - \mu_i)w_{ij}] \,\|_F^2 = \sum_{ij} |\lambda_i - \mu_j|^2 |w_{ij}|^2 \\
&\ge \delta^2 \|W\|_F^2 = \delta^2 \|Z\|_F^2.
\end{aligned}$$

Thus, $\mathrm{sep}_F(A_{11}, A_{22}) \ge \delta$. Combined with (b)(iii), this yields (c).

It remains to prove (b)(i). Let (u, v) be a pair of normalized singular vectors belonging to the smallest singular value $\sigma_{\min}$ of $\lambda I - A_{22}$ such that $(\lambda I - A_{22})v = \sigma_{\min}u$,

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and $\|u\|_2 = \|v\|_2$. Then $\|vu^*\| = \|uu^*\|$, since vu^* and uu^* have the same singular values $1, 0, \ldots, 0$. Setting $Z = vu^*/\|uu^*\|$ we obtain $\|Z\| = 1$ and $\|(\lambda I - A_{22})Z\| = \sigma_{\min}$. Thus $\operatorname{sep}(\lambda I, A_{22}) \leq \sigma_{\min}$. On the other hand,

$$\operatorname{sep}(\lambda I, A_{22}) = \min_{Z \neq 0} \frac{\|(\lambda I - A_{22})Z\|}{\|Z\|} = \min_{W \neq 0} \frac{\|W\|}{\|(\lambda I - A_{22})^{-1} W\|}$$
$$\geq \min_{W \neq 0} \frac{\|W\|}{\|(\lambda I - A_{22})^{-1}\|_2 \, \|W\|} = \sigma_{\min}. \qquad \square$$

2.2. Invariant subspaces and a quadratic matrix equation. Let us consider a general matrix $A \in \mathbb{C}^{n \times n}$ and partition

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad A_{11} \in \mathbb{C}^{k \times k}, \qquad A_{22} \in \mathbb{C}^{(n-k) \times (n-k)}.$$

For given $Z \in \mathbb{C}^{(n-k) \times k}$, consider the similarity transformation

$$\begin{bmatrix} I & 0 \\ -Z & I \end{bmatrix} A \begin{bmatrix} I & 0 \\ Z & I \end{bmatrix} = \begin{bmatrix} A_{11} + A_{12}Z & A_{12} \\ A_{21} + A_{22}Z - ZA_{11} - ZA_{12}Z & A_{22} - ZA_{12} \end{bmatrix} \tag{2.4}$$

which becomes block upper triangular if and only if the quadratic matrix equation

$$0 = f(A, Z) := A_{21} + A_{22}Z - ZA_{11} - ZA_{12}Z \tag{2.5}$$

is satisfied. This implies Lemma 2.2 below. Note that a subspace $\mathcal{Y}$ of row vectors is called a *left invariant subspace* if $\mathcal{Y}A \subset \mathcal{Y}$.

LEMMA 2.2. *Using the notation introduced above, the following statements are equivalent:*

(i) *The columns of* $[I \;\; Z^\top]^\top$ *span a right invariant subspace of* A *such that*

$$A \begin{bmatrix} I \\ Z \end{bmatrix} = \begin{bmatrix} I \\ Z \end{bmatrix} (A_{11} + A_{12}Z).$$

(ii) *The rows of* $[-Z \;\; I]$ *span a left invariant subspace of* A *such that*

$$\begin{bmatrix} -Z & I \end{bmatrix} A = (A_{22} - ZA_{12}) \begin{bmatrix} -Z & I \end{bmatrix}.$$

(iii) *The quadratic matrix equation* (2.5) *is satisfied.*

2.3. An asymptotic result. Perturbation bounds that are asymptotically valid as $\|E\| \to 0$ can be obtained in a relatively straightforward way from truncating perturbation expansions. For invariant subspaces, such expansions have been discussed in [5, 14, 25]. In the following, we will illustrate this approach.

LEMMA 2.3. *Given* $A \in \mathbb{C}^{n \times n}$, *suppose that there exists* $Z \in \mathbb{C}^{(n-k) \times k}$ *such that* $f(A, Z) = 0$, *with* f *defined as in* (2.5). *If* $\Lambda(A_{11} + A_{12}Z) \cap \Lambda(A_{22} - ZA_{12}) = \emptyset$, *then there exist an open neighborhood* $\mathcal{E} \subset \mathbb{C}^{n \times n}$ *of* 0 *and an open neighborhood* $\mathcal{Z} \subset \mathbb{C}^{(n-k) \times k}$ *of* Z *such that for each* $E \in \mathcal{E}$ *the equation* $f(A + E, Z_E) = 0$ *has a unique solution* $Z_E \in \mathcal{Z}$. *Moreover,* Z_E *depends holomorphically on* E *and admits the first-order expansion*

$$Z_E = Z + \mathbb{T}_Z^{-1}(E_{21}) + O(\|E\|^2)$$

with the Sylvester operator $\mathbb{T}_Z : \Delta Z \mapsto \Delta Z(A_{11} + A_{12}Z) - (A_{22} - ZA_{12})\Delta Z$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. Clearly, f is a holomorphic function in the entries of A and Z. The derivative of f with respect to the variable Z equals the Sylvester operator $-\mathbb{T}_Z$. Since $A_{22} - ZA_{12}$ and $A_{11} + A_{12}Z$ have disjoint spectra, the operator $\mathbb{T}_Z$ is invertible. Thus, the lemma follows from the implicit function theorem [15]. ∎

The way Lemma 2.3 is stated will be convenient for later purposes. However, for the sake of an asymptotic result, we may assume without loss of generality that the unperturbed matrix A is already in block triangular form,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \qquad A_{11} \in \mathbb{C}^{k\times k}, \qquad A_{22} \in \mathbb{C}^{(n-k)\times(n-k)}. \tag{2.6}$$

By Lemma 2.2, this is equivalent to requiring that $X = \left[\begin{smallmatrix} I_k \\ 0 \end{smallmatrix}\right]$ spans an invariant subspace of A. Combining the statement of Lemma 2.3 with Lemma 2.2 then yields the following result.

COROLLARY 2.4. *Let A be in block triangular form* (2.6) *and assume $\Lambda(A_{11}) \cap \Lambda(A_{22}) = \emptyset$. Then, for every E with $\|E\|$ sufficiently small, there exists an invariant subspace $\mathcal{X}_E = \operatorname{span}\left[\begin{smallmatrix} I \\ Z_E \end{smallmatrix}\right]$ of $A + E$ such that Z_E admits the first-order expansion*

$$Z_E = \mathbb{T}^{-1}\big(E_{21}\big) + O(\|E\|^2), \qquad \mathbb{T} : Z \mapsto ZA_{11} - A_{22}Z. \tag{2.7}$$

Recently, Stewart [20] derived bounds on $\|E\|$ for which Z_E, as a function of E, is Fréchet differentiable.

Once we have obtained Z_E, there are different ways of comparing the two invariant subspaces

$$\mathcal{X} = \operatorname{span}\begin{bmatrix} I \\ 0 \end{bmatrix}, \qquad \mathcal{X}_E = \operatorname{span}\begin{bmatrix} I \\ Z_E \end{bmatrix}$$

of the matrices A and $A + E$, respectively. If $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k$ denote the singular values of Z_E, then the ith canonical angle between $\mathcal{X}$ and $\mathcal{X}_E$ is given by $\theta_i(\mathcal{X}, \mathcal{X}_E) = \arctan\sigma_i$. Defining $\Theta(\mathcal{X}, \mathcal{X}_E) := \operatorname{diag}(\theta_1, \ldots, \theta_k)$, it is well known [21, sect. II.4] that $\|\sin(\Theta(\mathcal{X}, \mathcal{X}_E))\|$ generates a metric on the k-dimensional subspaces of $\mathbb{C}^n$ for any unitarily invariant matrix norm $\|\cdot\|$. However, it is sometimes more convenient to simply use

$$\|Z_E\| = \|\tan(\Theta(\mathcal{X}, \mathcal{X}_E))\| \tag{2.8}$$

for measuring the distance, which remains close to $\|\sin(\Theta(\mathcal{X}, \mathcal{X}_E))\|$ as long as $\|Z_E\|$ is small. The first-order result

$$\|\tan(\Theta(\mathcal{X}, \mathcal{X}_E))\| = \|Z_E\| = \frac{\|E_{21}\|}{\operatorname{sep}(A_{11}, A_{22})} + O(\|E\|^2) \tag{2.9}$$

is now readily obtained from Corollary 2.4. This also confirms that $\operatorname{sep}(A_{11}, A_{22})^{-1}$ is the condition number of $\mathcal{X}$ [2].

2.4. Nonasymptotic results. The derivation of nonasymptotic results requires a more careful study of the quadratic matrix equation $f(A + E, Z_E) = 0$ with f as in (2.5) and

$$A + E = \begin{bmatrix} A_{11} + E_{11} & A_{12} + E_{12} \\ E_{21} & A_{22} + E_{22} \end{bmatrix}. \tag{2.10}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

In particular, A is assumed to be block upper triangular. Then Stewart's result [18, Thm. 4.1] see also [21, Thm. 2.7]) reads as follows.

THEOREM 2.5. *Let $\|\cdot\|$ denote a consistent family of norms (i.e. $\|XY\| \le \|X\|\,\|Y\|$ whenever the matrix product XY is defined). Let $A + E \in \mathbb{C}^{n\times n}$ be partitioned as in* (2.10) *and set*

$$s_E := \operatorname{sep}(A_{11}, A_{22}) - \|E_{11}\| - \|E_{22}\|.$$

If $s_E > 0$ and

$$\|E_{21}\|(\|A_{12}\| + \|E_{12}\|) < s_E^2/4, \tag{2.11}$$

then there exists a unique solution Z_E of $f(A + E, Z_E) = 0$ satisfying

$$\|Z_E\| \le \frac{2\|E_{21}\|}{s_E + \sqrt{s_E^2 - 4\|E_{21}\|(\|A_{12}\| + \|E_{12}\|)}} < 2\frac{\|E_{21}\|}{s_E}. \tag{2.12}$$

Interestingly, Theorem 2.5 can be derived directly from the Newton–Kantorovich theorem; see Appendix A.

Remark 2.6. For a unitarily invariant family of norms $\|\cdot\|$, the result of Theorem 2.5 still holds if s_E is replaced with $\tilde{s}_E = \operatorname{sep}(A_{11}, A_{22}) - \|E_{11}\|_2 - \|E_{22}\|_2$, $\|E_{12}\|$ is replaced with $\|E_{12}\|_2$, and $\|A_{12}\|$ is replaced with $\|A_{12}\|_2$. See Appendix A.

A different analysis for the case of the Frobenius norm has been given by Demmel [8]; see also [16].

THEOREM 2.7 (see [16, Thm. 1.15]). *Using the notation of Theorem* 2.5*, assume that*

$$\|E\|_F < \frac{\operatorname{sep}_F(A_{11}, A_{22})}{4\|P\|_2}, \tag{2.13}$$

where P denotes the spectral projector of A belonging to $\Lambda(A_{11})$. Then there exists a unique solution Z_E of $f(A + E, Z_E) = 0$ satisfying

$$\|Z_E\|_F < \frac{4\|E\|_F}{\operatorname{sep}_F(A_{11}, A_{22}) - 4\|P\|_2\|E\|_F}.$$

A comparison of Theorems 2.5 and 2.7 with our results is given in section 3.5.

3. Main result. Our main result admits the use of different norms for measuring the perturbation E and the solution Z. More specifically, we consider two norms $\|\cdot\| : \mathbb{C}^{(n-k)\times k} \to \mathbb{R}$ and $|\!|\!|\cdot|\!|\!| : \mathbb{C}^{n\times n} \to \mathbb{R}$ that satisfy the inequalities

$$\begin{aligned} &\text{(N1)} \quad \|Z\|_2 \le \|Z\|, && \text{(N2)} \quad \|UZV\| \le \|U\|_2\,\|Z\|\,\|V\|_2, \\ &\text{(N3)} \quad \|E\|_2 \le |\!|\!|E|\!|\!|, && \text{(N4)} \quad \|XEY\| \le \|X\|_2\,|\!|\!|E|\!|\!|\,\|Y\|_2 \end{aligned}$$

for all $E \in \mathbb{C}^{n\times n}$, $Z \in \mathbb{C}^{(n-k)\times k}$, $X \in \mathbb{C}^{k\times n}$, $Y \in \mathbb{C}^{n\times k}$, $U \in \mathbb{C}^{(n-k)\times(n-k)}$, and $V \in \mathbb{C}^{k\times k}$. Condition (N2) is equivalent to requiring that $\|\cdot\|$ is unitarily invariant. In particular, there is a symmetric gauge function Φ such that

$$\|Z\| = \Phi(\sigma_1(Z), \ldots, \sigma_{\min\{k,n-k\}}(Z)),$$

where $\sigma_1(Z) \ge \cdots \ge \sigma_{\min\{k,n-k\}}(Z)$ are the singular values of Z in nonincreasing order; see [3, Thm. IV.2.1]. Condition (N1) is equivalent to requiring that

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$\Phi(1,0,\dots,0) \geq 1$. If we now define $|\!|E|\!|$ for $E \in \mathbb{C}^{n\times n}$ by the same formula as $\|Z\|$, that is

$$|\!|E|\!| = \Phi(\sigma_1(E),\dots,\sigma_{\min\{k,n-k\}}(E)),$$

then conditions (N1)–(N4) are all satisfied. It is important to note that only $\min\{k, n-k\}$ of the n singular values of E are involved in the definition of $|\!|E|\!|$.

We remark that $|\!|\cdot|\!|$ need not be unitarily invariant. For instance, the conditions (N1)–(N4) are also valid if $\|\cdot\| = \|\cdot\|_F$ is the Frobenius norm and $|\!|E|\!| = \sum_{i,j} |e_{ij}|$.

THEOREM 3.1. *Consider the block triangular matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \qquad A_{11} \in \mathbb{C}^{k\times k}, \qquad A_{22} \in \mathbb{C}^{(n-k)\times(n-k)},$$

and assume that $\Lambda(A_{11}) \cap \Lambda(A_{22}) = \emptyset$. *Let* $\|\cdot\|$ *and* $|\!|\cdot|\!|$ *be norms on* $\mathbb{C}^{(n-k)\times k}$ *and* $\mathbb{C}^{n\times n}$, *respectively, that satisfy conditions* (N1)–(N4). *Let*

$$s := \operatorname{sep}(A_{11}, A_{22}) = \min_{\|Z\|=1} \|ZA_{11} - A_{22}Z\|.$$

For $\varepsilon \geq 0$ *define* $g(\varepsilon) = \sqrt{\varepsilon(\varepsilon + \|A_{12}\|_2)}$ *and let* $\rho \geq 0$ *be such that* $g(\rho) = \frac{s}{2}$, *i.e.,* $\rho = \frac{1}{2}(\sqrt{s^2 + \|A_{12}\|_2^2} - \|A_{12}\|_2)$. *Finally, let* $\mathcal{B}_\rho := \{\, E \in \mathbb{C}^{n\times n} \mid\ |\!|E|\!| < \rho\}$. *Then the following statements hold:*

(a) *For all* $\varepsilon \geq 0$, $\Lambda_\varepsilon(A) \subseteq \Lambda_{g(\varepsilon)}(A_{11}) \cup \Lambda_{g(\varepsilon)}(A_{22})$.
(b) *If* $\varepsilon < \rho$, *then* $\Lambda_{g(\varepsilon)}(A_{11}) \cap \Lambda_{g(\varepsilon)}(A_{22}) = \emptyset$.
(c) *There exists a unique holomorphic function*

$$\mathcal{B}_\rho \ni E \longmapsto Z_E \in \mathbb{C}^{(n-k)\times k}$$

with the following properties:
(i) *The columns of* $[I \;\; Z_E^\top]^\top$ *span a right invariant subspace* $\mathcal{X}_E$ *of* $A + E$.
(ii) *The rows of* $[-Z_E \;\; I]$ *span a left invariant subspace* $\mathcal{Y}_E$ *of* $A + E$.
(iii) *The spectrum of the restriction of* $A + E$ *to* $\mathcal{X}_E$ *is contained in the pseudospectrum* $\Lambda_{g(\|E\|_2)}(A_{11})$. *The spectrum of the restriction of* $A+E$ *to* $\mathcal{Y}_E$ *is contained in the pseudospectrum* $\Lambda_{g(\|E\|_2)}(A_{22})$.
(iv) *The matrix* Z_E *satisfies*

$$\|Z_E\| \leq \frac{2|\!|E|\!|}{s + \sqrt{s^2 - 4|\!|E|\!|(|\!|E|\!| + \|A_{12}\|_2)}} \leq \frac{2}{s}\,|\!|E|\!| \tag{3.1}$$

as well as

$$\|Z_E - \mathbb{T}^{-1}(E_{21})\| \leq \frac{6}{s^2}\|E\|_2\,|\!|E|\!|, \tag{3.2}$$

where $\mathbb{T} : Z \mapsto ZA_{11} - A_{22}Z$.

Remark 3.2. Inequality (3.2) gives a bound for the remainder $O(\|E\|^2)$ in the first-order expansion (2.7).

Remark 3.3. The first bound in (3.1) looks quite complicated. A slightly weaker but more appealing estimate for $\|Z_E\|$ is

$$\|Z_E\| \leq \frac{|\!|E|\!|}{s}\left(1 + \frac{|\!|E|\!|}{\rho}\right), \tag{3.3}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

which can be derived as follows. Setting $\psi(\epsilon) = 2/(s + \sqrt{s^2 - 4\epsilon(\epsilon + \|A_{12}\|)})$ the first bound in (3.1) can be written as $\|Z_E\| \leq \psi(|E|)\,|E|$. A direct computation yields $\psi''(\epsilon) \geq 0$. Thus, ψ is a convex function. It follows that $\psi(\epsilon) \leq \psi(0) + (\epsilon/\rho)(\psi(\rho) - \psi(0)) = s^{-1}(1 + (\epsilon/\rho))$. This shows (3.3).

3.1. Proof of Theorem 3.1. Statement (a) of Theorem 3.1 has been shown by Grammont and Largillier [10, Proposition 3.1]; see also [13]. Statement (b) is a consequence of Proposition 2.1(b)(iii). It remains to prove statement (c), in particular the upper bound (3.1). For this purpose, we will first derive an auxiliary result showing that this upper bound and a certain lower bound are mutually exclusive. The main part of the proof then consists of ruling out the lower bound by a continuity argument. In the following, we will write Z instead of Z_E for notational convenience.

LEMMA 3.4. *With the notation and assumptions stated in Theorem* 3.1, *define*

$$r_+(\varepsilon) = 2\varepsilon/\big(s + \sqrt{s^2 - 4\varepsilon(\varepsilon + \|A_{12}\|_2)}\big), \quad r_-(\varepsilon) = 2\varepsilon/\big(s - \sqrt{s^2 - 4\varepsilon(\varepsilon + \|A_{12}\|_2)}\big).$$

If $|E| \leq \rho$ *and* $f(A + E, Z) = 0$, *then* $\|Z\| \leq r_+(|E|)$ *or* $\|Z\| \geq r_-(|E|)$.

Proof. By direct computation, it can be verified that $r_-(\varepsilon)$ and $r_+(\varepsilon)$ are the zeros of the quadratic polynomial $p_\varepsilon(r) := (\varepsilon + \|A_{12}\|_2)r^2 - s\,r + \varepsilon$. Since its leading coefficient is positive, we have

$$p_\varepsilon(r) < 0 \quad \Leftrightarrow \quad r_+(\varepsilon) < r < r_-(\varepsilon). \tag{3.4}$$

The assumption $\Lambda(A_{11}) \cap \Lambda(A_{22}) = \emptyset$ of Theorem 3.1 implies that the Sylvester operator $\mathbb{T} : Z \mapsto ZA_{11} - A_{22}Z$ is nonsingular, and hence $s = \|\mathbb{T}^{-1}\|^{-1}$. The equation $f(A + E, Z) = 0$ can be written as

$$\mathbb{T}(Z) = [-Z \;\; I]\,E\,\begin{bmatrix} I \\ Z \end{bmatrix} - ZA_{12}Z$$

or, equivalently,

$$Z = \mathbb{T}^{-1}\left([-Z \;\; I]\,E\,\begin{bmatrix} I \\ Z \end{bmatrix} - ZA_{12}Z\right).$$

Using (2.3) and the fact that $\|[-Z \;\; I]\|_2 = \|\,[I \;\; Z^\top]^\top\,\|_2 = \sqrt{1 + \|Z\|_2^2} \leq \sqrt{1 + \|Z\|^2}$, we conclude that $\|Z\| \leq \big(|E|(1 + \|Z\|^2) + \|A_{12}\|_2\,\|Z\|^2\big)/s$, which is equivalent to $0 \leq p_{|E|}(\|Z\|)$. Thus, the claim follows from (3.4). ∎ Note that the upper bound $\|Z\| \leq r_+(|E|)$ in Lemma 3.4 is equivalent to the first inequality in (3.1).

LEMMA 3.5. *For* $E \in \mathcal{B}_\rho$ *let* $\mathbb{I}_E$ *denote the set of* $t \in [0, 1]$ *such that there exists a matrix* Z *with the following properties:*

(α) $f(A + tE, Z) = 0$;
(β) $\Lambda_1(t) \subseteq \Lambda_{g(\|E\|_2)}(A_{11})$, *where* $\Lambda_1(t) := \Lambda(A_{11} + t(E_{11} + E_{12}Z))$;
(γ) $\Lambda_2(t) \subseteq \Lambda_{g(\|E\|_2)}(A_{22})$, *where* $\Lambda_2(t) := \Lambda(A_{22} + t(E_{22} - ZE_{21}))$;
(δ) $\|Z\| \leq r_+(|E|)$.

Then $1 \in \mathbb{I}_E$.

Proof. The proof of the statement proceeds via analytic continuation in three steps.

Step 1. Since the conditions (α)–(δ) hold for $t = 0$ and $Z = 0$, it follows that $0 \in \mathbb{I}_E$.

Step 2. We now claim that there exists $\varepsilon > 0$ such that $[\hat{t}, \hat{t} + \varepsilon) \subset \mathbb{I}_E$ for any $\hat{t} \in \mathbb{I}_E$ with $\hat{t} < 1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Let $\hat{Z}$ be such that $f(A+\hat{t}E,\hat{Z})=0$. The pseudospectra $\Lambda_{g(\|E\|_2)}(A_{11})$ and $\Lambda_{g(\|E\|_2)}(A_{22})$ are disjoint by Theorem 3.1(b). Thus $\Lambda_1(\hat{t})$ and $\Lambda_2(\hat{t})$ are disjoint, too. Hence, Lemma 2.3 applied to $A+\hat{t}E$ implies that there exist $\varepsilon>0$ and a holomorphic function

$$[\hat{t},\hat{t}+\varepsilon)\ni t\longmapsto Z_t\in\mathbb{C}^{m\times l}$$

such that $f(A+tE,Z_t)=0$ and $Z_{\hat{t}}=\hat{Z}$. We may assume that $\hat{t}+\varepsilon<1$. The set $\Lambda_1(t)$ is the spectrum of $A+tE$ restricted to the right invariant subspace $[I\ \ Z_t^\top]^\top$. Thus,

$$\Lambda_1(t)\subset\Lambda(A+tE)\subset\Lambda_{g(|E|)}(A_{11})\cup\Lambda_{g(|E|)}(A_{22}).$$

However, since the latter pseudospectra are disjoint closed sets it follows from $\Lambda_1(\hat{t})\subseteq\Lambda_{g(|E|)}(A_{11})$ and the continuity of eigenvalues that $\Lambda_1(t)\subseteq\Lambda_{g(|E|)}(A_{11})$ for all $t\in[\hat{t},\hat{t}+\varepsilon)$. Analogously, we conclude that $\Lambda_2(t)\subseteq\Lambda_{g(|E|)}(A_{22})$ for all $t\in[\hat{t},\hat{t}+\varepsilon)$. It remains to verify property (δ). By Lemma 3.4, we have for every t that $\|Z_t\|\le r_+(t|E|)\le r_+(|E|)$ or $\|Z_t\|\ge r_-(t|E|)\ge r_-(|E|)$. Since the former inequality holds for $t=\hat{t}$ and $r_+(|E|)<r_-(|E|)$, the continuity of $t\mapsto Z_t$ implies that $\|Z_t\|\le r_+(|E|)$ for all $t\in[\hat{t},\hat{t}+\varepsilon)$. This establishes the claim.

Step 3. We now claim that the set $\mathbb{I}_E$ is closed.

Let (t_j) be a sequence in $\mathbb{I}_E$ with limit t_*. Then there exists a sequence (Z_j) such that the pairs (t_j,Z_j) satisfy (α)–(δ). In particular $\|Z_j\|\le r_+(|E|)$ for all j. By compactness the sequence (Z_j) has a convergent subsequence. Let Z_* denote its limit. Then (t_*,Z_*) satisfies (α)–(δ). In particular, the properties (β) and (γ) for (t_*,Z_*) are consequences of the continuity of eigenvalues and the closedness of the pseudospectra $\Lambda_{g(\|E\|_2)}(A_{11}),\Lambda_{g(\|E\|_2)}(A_{22})$. This establishes the claim.

From Step 3, together with Steps 1 and 2, it follows that $1=\sup\mathbb{I}_E\in\mathbb{I}_E$. ∎

Proof of Theorem 3.1(c). By applying Lemma 2.2 to $A+E$, each of the conditions (c)(i) and (c)(ii) of Theorem 3.1 is equivalent to property (α) of Lemma 3.5 for $t=1$. Again by Lemma 2.2, condition (c)(iii) is equivalent to the properties (β) and (γ) for $t=1$. Thus, Lemma 3.5 establishes the existence of Z satisfying (c)(i)–(c)(iii) as well as the first inequality in (3.1).

Next we prove uniqueness of any such matrix Z. For this purpose, suppose that $f(A+E,Z)=f(A+E,\tilde{Z})=0$. Then

$$\begin{aligned}0&=f(A+E,Z)-f(A+E,\tilde{Z})\\&=E_{21}+(A_{22}+E_{22})Z-Z(A_{11}+E_{11})-Z(A_{12}+E_{12})Z\\&\quad-\big[E_{21}+(A_{22}+E_{22})\tilde{Z}-\tilde{Z}(A_{11}+E_{11})-\tilde{Z}(A_{12}+E_{12})\tilde{Z}\big]\\&=\underbrace{(A_{22}+E_{22}-\tilde{Z}(A_{12}+E_{12}))}_{=:\tilde{A}_{22}}(Z-\tilde{Z})-(Z-\tilde{Z})\underbrace{(A_{11}+E_{11}+(A_{12}+E_{12})Z)}_{=:\tilde{A}_{11}}.\end{aligned}$$

Since both Z and $\tilde{Z}$ satisfy properties (β) and (γ) for $t=1$, the spectra $\Lambda(\tilde{A}_{22})$ and $\Lambda(\tilde{A}_{11})$ are disjoint. Thus, $Z-\tilde{Z}=0$.

In summary, we have shown the existence and uniqueness of a function $E\mapsto Z_E$ satisfying conditions (i)–(iii) in Theorem 3.1. Lemma 2.3 implies that this function is holomorphic. It remains to prove the inequality (3.2). The relation $f(A+E,Z_E)=0$ is equivalent to

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$Z_E = \mathbb{T}^{-1}(E_{21} + E_{22}Z_E - Z_E E_{11} - Z_E E_{12} Z_E).$$

Furthermore, by (3.1) we have $\|Z_E\| \le 2|E|/s < 2\rho/s \le 1$. Thus,

$$\begin{aligned}\|Z_E - \mathbb{T}^{-1}(E_{21})\| &= \|\mathbb{T}^{-1}(E_{22}Z_E - Z_E E_{11} - Z_E E_{12} Z_E)\| \\ &\le \frac{1}{s}(2\|E\|_2 \|Z_E\| + \|E\|_2 \|Z_E\|^2) \\ &\le \frac{3}{s}\|E\|_2 \|Z_E\| \le \frac{6}{s^2}\|E\|_2 |E|.\end{aligned}$$

This concludes the proof of Theorem 3.1. ∎

3.2. The 2×2 case. In this section we show for a 2×2 example that the bounds in Theorem 3.1 are sharp. Let

$$A = \begin{bmatrix} -s/2 & c \\ 0 & s/2 \end{bmatrix}, \qquad E = \begin{bmatrix} 0 & \varepsilon \\ -\varepsilon & 0 \end{bmatrix}, \qquad s > 0, \quad c, \varepsilon \ge 0.$$

Then s is the separation of the diagonal elements of A and $\|E\|_2 = \varepsilon$. Furthermore, let $\rho = \frac{1}{2}(\sqrt{s^2 + c^2} - c)$. Then $\rho(\rho + c) = s^2/4$. The ε-pseudospectrum of A can be calculated as

$$\begin{aligned}\Lambda_\varepsilon(A) &= \{ z \in \mathbb{C} \mid \sigma_{\min}(zI - A) \le \varepsilon \} \\ &= \left\{ z \in \mathbb{C} \mid \frac{1}{2}\left(\sqrt{\left(|z - \tfrac{s}{2}| + |z + \tfrac{s}{2}|\right)^2 + c^2} - \sqrt{\left(|z - \tfrac{s}{2}| - |z + \tfrac{s}{2}|\right)^2 + c^2} \right) \le \varepsilon \right\};\end{aligned}$$

see also [13]. By Theorem 3.1(a),

$$\Lambda_\varepsilon(A) \subseteq \mathcal{D}_{\sqrt{\varepsilon(\varepsilon+c)}}(-s/2) \cup \mathcal{D}_{\sqrt{\varepsilon(\varepsilon+c)}}(s/2), \tag{3.5}$$

where $\mathcal{D}_r(z) \subset \mathbb{C}$ denotes the closed disk of radius $r \ge 0$ around $z \in \mathbb{C}$. If $\varepsilon < \rho$, then $\sqrt{\varepsilon(\varepsilon + c)} < s/2$ and the disks in (3.5) are disjoint. Hence, the pseudospectrum $\Lambda_\varepsilon(A)$ has two connected components. For $\varepsilon = \rho$ the disks in (3.5) touch each other at $0 \in \mathbb{C}$. From (3.5) it follows that also $0 \in \sigma_\rho(A)$. Hence, $\sigma_\varepsilon(A)$ has only one connected component for $\varepsilon \ge \rho$. The eigenvalues of $A + E$ are

$$\lambda_\pm(\varepsilon) = \pm\frac{1}{2}\sqrt{s^2 - 4\varepsilon(\varepsilon + c)}.$$

These eigenvalues lie close to the boundary of $\Lambda_\varepsilon(A)$. The situation is illustrated in Figure 1, where the shaded regions represent pseudospectra, the circles represent the boundaries of the disks $\mathcal{D}_{\sqrt{\varepsilon(\varepsilon+c)}}(\pm s/2)$, and the dots mark the eigenvalues $\lambda_\pm(\varepsilon)$. A right eigenvector to the eigenvalue $\lambda_-(\varepsilon)$ is given by $[1 \;\; z_\varepsilon]^\top$, where $z_\varepsilon = 2\varepsilon/(s + \sqrt{s^2 - 4\varepsilon(\varepsilon + c)})$. If $\varepsilon < \rho$, then the eigenvalues are real and distinct. If $\varepsilon = \rho$, then $A + E$ is similar to a Jordan block. More specifically, in this case we have

$$A + E = \begin{bmatrix} 1 & -2/s \\ 2\varepsilon/s & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2/s \\ 2\varepsilon/s & 0 \end{bmatrix}^{-1}.$$

If $\varepsilon > \rho$, then the eigenvalues $\lambda_\pm(\varepsilon)$ are purely imaginary. Note that the function $\varepsilon \mapsto z_\varepsilon$ is not differentiable at $\varepsilon = \rho$.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

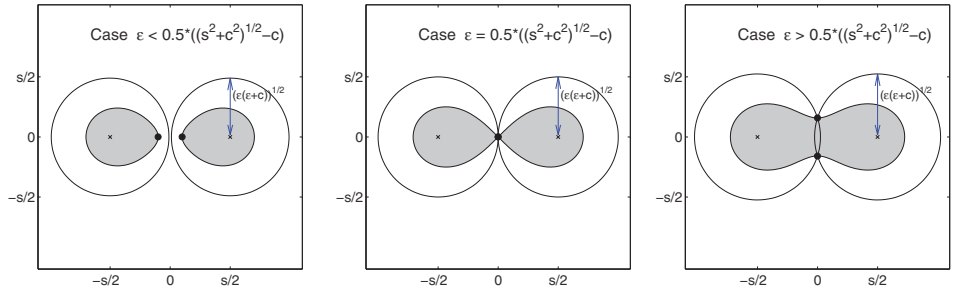


FIG. 1. *ε-pseudospectra of a* 2×2 *matrix for three values of* ε.

3.3. A bound in terms of the spectral decomposition. The purpose of this section is to derive a bound based on the block diagonalization of A. We assume the setting of Theorem 3.1. In particular, A is block triangular and $\Lambda(A_{11}) \cap \Lambda(A_{22}) = \emptyset$. Then there exists a unique $R \in \mathbb{C}^{k\times(n-k)}$ such that $RA_{22} - A_{11}R = A_{12}$. We have

$$A \begin{bmatrix} R \\ I \end{bmatrix} = \begin{bmatrix} R \\ I \end{bmatrix} A_{22}.$$

Hence, the columns of $[R^\top \quad I]^\top$ span a right invariant subspace $\mathcal{X}^c$ of A which is complementary to $\mathcal{X} = \text{range}([I \;\; 0]^\top)$. The projector onto $\mathcal{X}$ along $\mathcal{X}^c$ is the spectral projector $P = \begin{bmatrix} I & -R \\ 0 & 0 \end{bmatrix}$. Let

$$p = \sqrt{1 + \|R\|_2^2}, \qquad \kappa = p + \|R\|_2 = p + \sqrt{p^2 - 1}, \qquad G = \begin{bmatrix} I & R/p \\ 0 & I/p \end{bmatrix}. \tag{3.6}$$

Then $p = \|P\|_2$, $\kappa = \|G\|_2\|G^{-1}\|_2$ is the condition number of G [8] and

$$A = G \,\text{diag}(A_{11}, A_{22})\, G^{-1}, \qquad G^{-1} = \begin{bmatrix} I & -R \\ 0 & p\, I \end{bmatrix}.$$

Furthermore, we have

$$\|R\|_2 = (\tan\varphi)^{-1}, \quad p = (\sin\varphi)^{-1}, \quad \kappa = (\tan\frac{\varphi}{2})^{-1},$$

where φ is the smallest angle between the subspaces $\mathcal{X}$ and $\mathcal{X}^c$; see [7].

With these preparations we are in a position to state and prove the following theorem.

THEOREM 3.6. *Let* A *and* s *be defined as in Theorem* 3.1*, and let* $\|\cdot\|$ *and* $|\!|\!|\cdot|\!|\!|$ *be unitarily invariant norms on* $\mathbb{C}^{(n-k)\times k}$ *and* $\mathbb{C}^{n\times n}$*, respectively, satisfying the conditions* (N1)–(N4)*. Let* R *and* κ *be defined as above. If* $|\!|\!|E|\!|\!| < s/(2\kappa)$*, then there exists a unique* $W_E \in \mathbb{C}^{(n-k)\times k}$*, depending holomorphically on* E *with the following properties:*

(i) *The columns of* $\left[\begin{smallmatrix} I & R \\ 0 & I \end{smallmatrix}\right]\left[\begin{smallmatrix} I \\ W_E \end{smallmatrix}\right]$ *span a right invariant subspace* $\mathcal{X}_E$ *of* $A + E$.
(ii) *The rows of* $\left[\begin{smallmatrix} -W_E & I \end{smallmatrix}\right]\left[\begin{smallmatrix} I & -R \\ 0 & I \end{smallmatrix}\right]$ *span a left invariant subspace* $\mathcal{Y}_E$ *of* $A + E$.
(iii) *The spectrum of the restriction of* $A + E$ *to* $\mathcal{X}_E$ *is contained in the pseudospectrum* $\Lambda_{\kappa\,|\!|\!|E|\!|\!|_2}(A_{11})$*. The spectrum of the restriction of* $A + E$ *to* $\mathcal{Y}_E$ *is contained in the pseudospectrum* $\Lambda_{\kappa\,|\!|\!|E|\!|\!|_2}(A_{22})$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(iv) *The matrix W_E satisfies*

$$\|W_E\| \leq \frac{2\kappa}{s\,p}\,|E| \tag{3.7}$$

and

$$\|W_E - \mathbb{T}^{-1}(E_{21})\| \leq \frac{6\,\kappa^2}{p\,s^2}\|E\|_2\,|E|, \tag{3.8}$$

where $\mathbb{T} : Z \mapsto ZA_{11} - A_{22}Z$.

(v) *Let $Z_E = W_E(I + RW_E)^{-1} = (I + W_E R)^{-1}W_E$. Then the columns of $[I \quad Z_E^\top]^\top$ span $\mathcal{X}_E$, the rows of $[-Z_E \quad I]$ span the subspace $\mathcal{Y}_E$, and*

$$\|Z_E\| \leq \frac{\frac{2\kappa}{p}|E|}{s - \frac{2\kappa}{p}\|R\|_2|E|}. \tag{3.9}$$

Proof. Let $\hat{A} = \operatorname{diag}(A_{11}, A_{22})$ and $\hat{E} = G^{-1}EG$. Then $A + E = G(\hat{A} + \hat{E})G^{-1}$. Furthermore, the equivalences

$$\begin{aligned} (\hat{A} + \hat{E})U = UL \quad &\Leftrightarrow \quad (A + E)(GU) = (GU)L, \\ V(\hat{A} + \hat{E}) = MV \quad &\Leftrightarrow \quad (VG^{-1})(A + E) = M(VG^{-1}) \end{aligned} \tag{3.10}$$

hold for any $U \in \mathbb{C}^{n\times k}$, $L \in \mathbb{C}^{k\times k}$, $V \in \mathbb{C}^{(n-k)\times n}$, $M \in \mathbb{C}^{(n-k)\times(n-k)}$. Thus, the columns of GU span a right invariant subspace of $A + E$ if and only if the columns of U span a right invariant subspace $\hat{A} + \hat{E}$. Furthermore, the rows of VG^{-1} span a left invariant subspace of $A + E$ if and only if the rows of V span a left invariant subspace $\hat{A} + \hat{E}$. Suppose $|E| < s/(2\kappa)$. Then, since $|\cdot|$ is unitarily invariant,

$$|\hat{E}| \leq \kappa|E| < s/2. \tag{3.11}$$

Hence, according to Theorem 3.1 there exists a unique $Z_{\hat{E}} \in \mathbb{C}^{(n-k)\times k}$ depending holomorphically on $\hat{E}$ with the following properties:

(i′) The columns of $[I \;\; Z_{\hat{E}}^\top]^\top$ span a right invariant subspace $\mathcal{X}_{\hat{E}}$ of $\hat{A} + \hat{E}$.

(ii′) The rows of $[-Z_{\hat{E}} \;\; I]$ span a left invariant subspace $\mathcal{Y}_{\hat{E}}$ of $\hat{A} + \hat{E}$.

(iii′) The spectrum of the restriction of $\hat{A} + \hat{E}$ to $\mathcal{X}_{\hat{E}}$ is contained in the pseudospectrum $\Lambda_{\|\hat{E}\|_2}(A_{11})$. The spectrum of the restriction of $A + E$ to $\mathcal{Y}_{\hat{E}}$ is contained in the pseudospectrum $\Lambda_{\|\hat{E}\|_2}(A_{22})$.

(iv′) The matrix $Z_{\hat{E}}$ satisfies

$$\|Z_{\hat{E}}\| \leq \frac{2}{s}\,|\hat{E}| \qquad \text{as well as} \qquad \|Z_{\hat{E}} - \mathbb{T}^{-1}(\hat{E}_{21})\| \leq \frac{6}{s^2}\|\hat{E}\|_2|\hat{E}|.$$

Let $U = [I \;\; Z_{\hat{E}}^\top]^\top$, $V = [-Z_{\hat{E}} \;\; I]$, and $W_E = Z_{\hat{E}}/p$. Then

$$\begin{bmatrix} I & R \\ 0 & I \end{bmatrix}\begin{bmatrix} I \\ W_E \end{bmatrix} = GU, \qquad \begin{bmatrix} -W_E & I \end{bmatrix}\begin{bmatrix} I & -R \\ 0 & I \end{bmatrix} = VG^{-1}.$$

Hence, the claims (i)–(iv) follow from (i′)–(iv′), the equivalences (3.10), the inequality (3.11), and the fact that $\hat{E}_{21} = p\,E_{21}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Now, redefine Z_E as $Z_E := W_E(I + RW_E)^{-1} = (I + RW_E)^{-1}W_E$. The first statements of claim (v) follow from the identities

$$\begin{bmatrix} I & R \\ 0 & I \end{bmatrix} \begin{bmatrix} I \\ W_E \end{bmatrix} = \begin{bmatrix} I \\ Z_E \end{bmatrix} (I+RW_E), \qquad \begin{bmatrix} -W_E & I \end{bmatrix} \begin{bmatrix} I & -R \\ 0 & I \end{bmatrix} = (I+W_ER) \begin{bmatrix} -Z_E & I \end{bmatrix}.$$

Since $\|Z_E\| = \|W_E(I + RW_E)^{-1}\| \le \|W_E\|/(1 - \|R\|_2\|W_E\|)$ the inequality (3.9) is a consequence of (3.7). ∎

Remark 3.7. If E and the underlying norms fulfill the assumptions of both Theorem 3.1 and Theorem 3.6, then the solution matrices Z_E established in these theorems are identical. This follows from the uniqueness statement in Theorem 3.1. Note that the bound (3.1) is tighter than the bound (3.9).

3.4. The case of the Frobenius norm. We now specialize our bounds to the case that the underlying norm is the Frobenius norm $\|\cdot\|_F$ on $\mathbb{C}^{(n-k)\times k}$. To this end, we define a unitarily invariant norm on $\mathbb{C}^{n\times n}$ by

$$\|E\|_{F,k} = \left(\sum_{j=1}^{\min\{k,n-k\}} \sigma_j(E)^2 \right)^{1/2}, \tag{3.12}$$

where $\sigma_1(E) \ge \sigma_2(E) \ge \cdots \ge \sigma_n(E)$ denote the singular values of $E \in \mathbb{C}^{n\times n}$ in nonincreasing order. Note that $\|E\|_{F,k} \le \|E\|_F$ with equality if and only if $\operatorname{rank} E \le k$. Since conditions (N1)–(N4) are satisfied for $\|\cdot\| = \|\cdot\|_F$ and $|\cdot| = \|\cdot\|_{F,k}$, we have the following corollary to Theorems 3.1 and 3.6.

COROLLARY 3.8. *Let $A \in \mathbb{C}^{n\times n}$ be partitioned as in Theorem* 3.1. *Let R, κ, and p be defined as above, and let $s = \operatorname{sep}_F(A_{11}, A_{22})$. Furthermore, let $\rho = \frac{1}{2}(\sqrt{s^2 + \|A_{12}\|_2^2} - \|A_{12}\|_2)$.*

(i) *If $\|E\|_{F,k} < \rho$, then then the matrix Z_E in Theorem* 3.1 *satisfies*

$$\|Z_E\|_F \le \frac{s}{2}\|E\|_{F,k}. \tag{3.13}$$

(ii) *If $\|E\|_{F,k} < \frac{s}{2\kappa}$, then the matrix Z_E in Theorem* 3.6 *satisfies*

$$\|Z_E\|_F \le \frac{\frac{2\kappa}{p}\|E\|_{F,k}}{s - \frac{2\kappa}{p}\|R\|_2\|E\|_{F,k}}. \tag{3.14}$$

If E satisfies both conditions, $\|E\|_{F,k} < \rho$ and $\|E\|_{F,k} < \frac{s}{2\kappa}$, then the matrices Z_E in (i) *and* (ii) *are identical.*

3.5. Comparison with existing results.

Comparison with Theorem 2.7. Since always $\|E\|_{F,k} \le \|E\|_F$, $\|R\|_2 \le p$, and $\kappa \le 2p$, the bound (3.14) is an improvement of Demmel's bound in Theorem 2.7.

For the particular case that A is block diagonal (i.e., $A_{12} = 0$) we have that $\kappa = p = 1$, $R = 0$, and (3.14) as well as (3.13) state

$$\|Z_E\|_F \le \frac{2\|E\|_{F,k}}{s} \quad \text{if} \quad \|E\|_{F,k} \le \frac{s}{2},$$

where $s = \operatorname{sep}_F(A_{11}, A_{22})$. In contrast, Theorem 2.7 yields

$$\|Z_E\|_F \le \frac{4\|E\|_F}{s} \quad \text{if} \quad \|E\|_F \le \frac{s}{4}.$$

On the other hand, if $\|R\|_2$ is large, then the constants in the bound (3.14) and in Demmel's bound are nearly identical since then $\|R\|_2 \approx p$ and $2\kappa/p \approx 4$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Comparison with Theorem 2.5. It is not that easy to compare our bound (3.1) with Stewart's bound (2.12) from Theorem 2.5. First, the bound (3.1) holds under the condition $|||E|(|E|+\|A_{12}\|)<s^2/4$, while Stewart's bound requires $\|E_{21}\|(\|E_{12}\|+\|A_{12}\|)<s_E^2/4$, where $s=\operatorname{sep}(A_{11},A_{22})$ and $s_E=s-\|E_{11}\|-\|E_{22}\|$. Hence, these bounds have a different range of applicability. The advantage of the bound in Theorem 3.1 over Stewart's result is that it has the separation s in the denominator instead of the smaller number s_E.

On the other hand, (3.1) works with the norm of the whole matrix E and Stewart's bound involves the norms of the blocks E_{ij}, which are smaller than $|||E|||$. In particular, Stewart's bound properly predicts $Z_E=0$ when $E_{21}=0$. Our bound (3.1) does not reflect this fact.

Following advice by Stewart, we performed several numerical experiments with random matrices and perturbations. Not surprisingly, in these experiments Stewart's bound is observed to be tighter than ours in most cases when $|||E|||$ is small compared to $s/2$. The same observation is made when $\|A_{12}\|$ is not significantly smaller than s.

It turns out that our new bound becomes advantageous when $A_{12}=0$, the perturbations are measured in the spectral norm, and $|||E|||$ is not small compared with $s/2$. To demonstrate this, let us consider an $n\times n$ block diagonal matrix A with $k\times k$ and $(n-k)\times(n-k)$ diagonal blocks. With $A_{12}=0$, the perturbation bounds depend on A only via the separation s and they scale nearly proportionally with $1/s$. It therefore suffices to consider one value for s, say, $s=1$. We have then chosen $\varepsilon\in(0,s/2]$ and generated a large set of random perturbations (10,000 for $n=7$ and 1000 for $n=50$) having normally distributed entries and norm ε. The plots in Figure 2 show the percentage of cases our bound is smaller than Stewart's bound. Note that "strong" refers to the stronger bounds (i.e., the first inequalities) while "weak" refers to the weaker but simpler bounds (i.e., the second inequalities) in (2.12) and (3.1). We have considered the spectral norm as well as the Frobenius norm for measuring perturbations. In the latter case, we have used the tighter bound (A.4) instead of (2.12) and set $|||E|||:=\|E\|_{F,k}=\sqrt{\sigma_1(E)^2+\cdots+\sigma^2_{\min\{k,n-k\}}(E)}$. For the spectral norm, Figures 2(a) and (c) show that our new bound performs better for $|||E|||$ close to $s/2$, especially when k is larger. For the Frobenius norm, the new bound seems to be better on average only for $k=1$ and when $|||E|||$ is close to $s/2$.

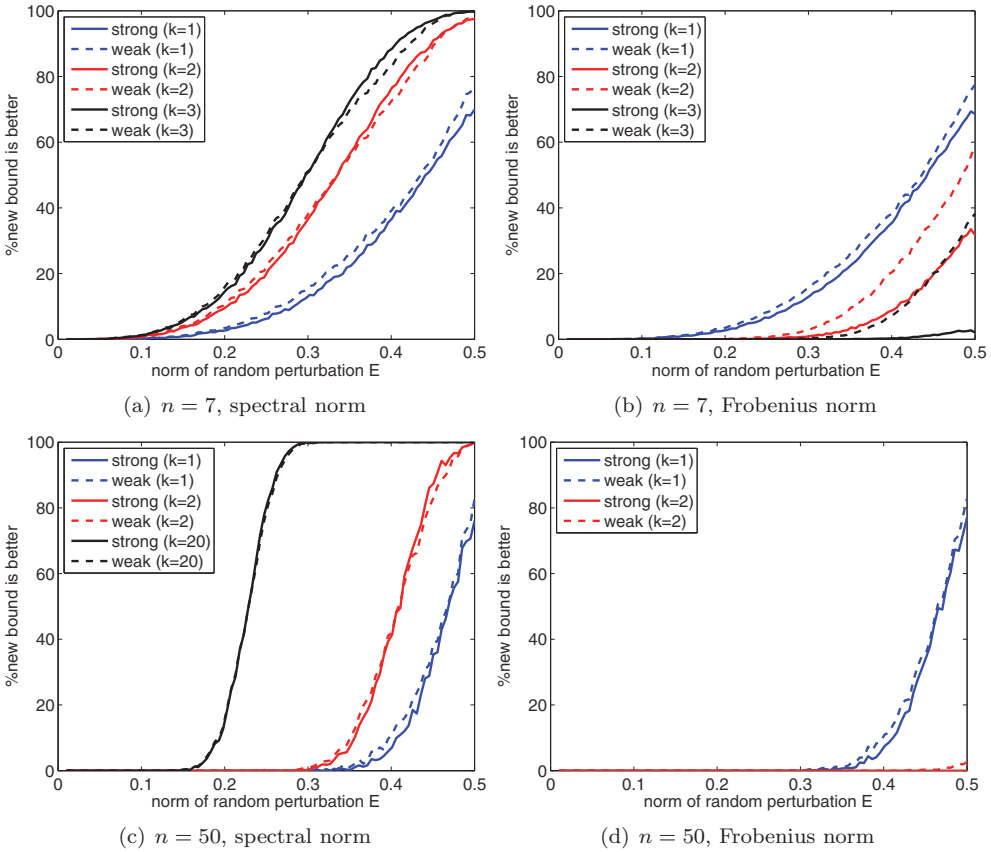
The following example illustrates that Stewart's bound can become very conservative for $|||E|||\approx s/2$.

Example 3.9. For $A_{11},A_{22}\in\mathbb{C}^{2\times 2}$ with disjoint spectra consider

$$A=\left[\begin{array}{c|c}A_{11} & 0\\ \hline 0 & A_{22}\end{array}\right],\quad E_t=\frac{st}{2}\left[\begin{array}{cc|cc}1&0&0&0\\0&0&0&0\\ \hline 0&1&0&0\\0&0&0&1\end{array}\right]=\left[\begin{array}{c|c}E_{11}^t & E_{12}^t\\ \hline E_{21}^t & E_{22}^t\end{array}\right],$$

where $0\le t<1$ and $s=\operatorname{sep}_2(A_{11},A_{22})$ is the separation with respect to the spectral norm. Then $\|E_t\|_2=\|E_{11}^t\|_2=\|E_{22}^t\|_2=\|E_{21}^t\|_2=st/2$, $\|E_{12}^t\|_2=0$, and $s_E=s-\|E_{11}^t\|_2-\|E_{22}^t\|_2=s(1-t)$. If $\|\cdot\|=|||\cdot|||=\|\cdot\|_2$, then our bound (3.1) and Stewart's bound in Theorem 2.5 are both applicable for $t<1$. As t tends to 1, Stewart's bound tends to infinity, while our bound (3.1) tends to 1. If $s=\operatorname{sep}_F(A_{11},A_{22})$ denotes the separation with respect to the Frobenius norm and $|||\cdot|||=\|\cdot\|_{F,2}$, then our bound (3.1) and Stewart's bound in Theorem 2.5 are both applicable for $t<1/\sqrt{2}$. As t tends to $1/\sqrt{2}$, Stewart's bound tends to $1/(2^{3/2}-2)\approx 1.207$, while our bound (3.1) tends to 1.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) $n = 7$, spectral norm

(b) $n = 7$, Frobenius norm

(c) $n = 50$, spectral norm

(d) $n = 50$, Frobenius norm

FIG. 2. *Performance of new perturbation bound* (3.1) *compared to Stewart's bounds* (2.12) *(spectral norm) or* (A.4) *(Frobenius norm) for block diagonal matrices and random perturbations.*

4. The case of a simple eigenvalue. The theorem below gives the second-order expansion of a simple eigenvalue as well as the first-order expansion of the associated eigenvector. These expansions are well known and can be found, e.g., in [17]. Our novel contributions consist of the existence regions and the bounds for the remainders in the expansion. Below, $M^\dagger$ and $M^\sharp$ denote the Moore–Penrose inverse and the Drazin inverse of $M \in \mathbb{C}^{n\times n}$, respectively. Furthermore, $\theta(x, y) = \arccos(|x^*y|/(\|x\|_2\,\|y\|_2))$ denotes the angle between the one-dimensional subspaces $\mathbb{C}x$ and $\mathbb{C}y$.

THEOREM 4.1. *Let $x_0 \in \mathbb{C}^n$ be a normalized right eigenvector of $A \in \mathbb{C}^{n\times n}$ belonging to a simple eigenvalue $\lambda_0 \in \mathbb{C}$ (i.e., $Ax_0 = \lambda_0 x_0$, $\|x_0\|_2 = 1$). Let $y_0 \in \mathbb{C}$ be a left eigenvector such that $y_0^*A = \lambda_0 y_0^*$ and $y_0^*x_0 = 1$. Moreover, let*

$$\begin{aligned} c &= \|x_0^*(\lambda_0 I - A)\|_2, \\ P_0 &= I - x_0x_0^*, \\ p_0 &= \|y_0\|_2, \\ \kappa_0 &= \|y_0\|_2 + \sqrt{\|y_0\|_2^2 - 1}, \\ s_0 &= \sigma_{n-1}(P_0(\lambda_0 I - A)), \end{aligned}$$

where $\sigma_{n-1}(\cdot)$ denotes the second smallest singular value.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(i) *If* $\|E\|_2 < \frac{1}{2}(\sqrt{s_0^2 + c^2} - c)$, *then there exists an eigenvector* x_E *of* $A + E$ *depending holomorphically on* E *such that* $x_0^* x_E = 1$ *and*

$$x_E = (I + [P_0(\lambda I - A)]^\dagger E)x_0 + \xi_E$$

for some $\xi_E \in \mathbb{C}^n$ *with*

$$\|\xi_E\|_2 \leq \frac{6}{s_0^2}\|E\|_2^2.$$

Furthermore,

$$\|x_E - x_0\|_2 = \tan(\theta(x_0, x_E)) \leq 2\,\|E\|_2/s_0.$$

(ii) *If* $\|E\|_2 < s_0/(2\kappa_0)$, *then there exists an eigenvector* $\tilde{x}_E$ *of* $A + E$ *depending holomorphically on* E *such that* $y_0^* \tilde{x}_E = 1$ *and*

$$\tilde{x}_E = x_0 + (\lambda_0 I - A)^\sharp E x_0 + \tilde{\xi}_E$$

for some $\tilde{\xi}_E \in \mathbb{C}^n$ *with*

$$\|\tilde{\xi}_E\|_2 \leq \frac{6\kappa_0^2}{s_0^2}\|E\|_2^2. \tag{4.1}$$

Furthermore,

(4.2)

$$\|\tilde{x}_E - x_0\|_2 \leq \frac{2\kappa_0}{s_0}\|E\|_2 \qquad \textit{and} \qquad \tan(\theta(x_0, \tilde{x}_E)) \leq \frac{\frac{2\kappa_0}{p_0}\|E\|_2}{s_0 - \frac{2\kappa_0}{p_0}\sqrt{p_0^2 - 1}\,\|E\|_2}.$$

The corresponding eigenvalue of $A + E$ *satisfies*

$$\begin{aligned} \lambda_E &= \lambda_0 + y_0^* E \tilde{x}_E \\ &= \lambda_0 + y_0^* E x_0 + \ell_E \\ &= \lambda_0 + y_0^* E x_0 + y_0^* E (\lambda_0 I - A)^\sharp E x_0 + \tilde{\ell}_E \end{aligned} \tag{4.3}$$

with $\ell_E = y_0^*(\tilde{x}_E - x_0)$ *and* $\tilde{\ell}_E = y_0^* E \tilde{\xi}_E$. *We have*

$$|\ell_E| \leq \frac{2\,p_0\,\kappa_0}{s_0}\|E\|_2^2 \qquad \textit{and} \qquad |\tilde{\ell}_E| \leq \frac{6\,p_0\,\kappa_0^2}{s_0^2}\|E\|_2^3. \tag{4.4}$$

Proof. After a unitary similarity transformation we may assume that

$$A = \begin{bmatrix} \lambda_0 & A_{12} \\ 0 & A_{22} \end{bmatrix}, \qquad A_{22} - \lambda_0\, I \text{ nonsingular}, \qquad x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad y_0^* = [1 \;\; r] \in \mathbb{C}^{1\times n},$$

where $r = A_{12}(\lambda_0\, I - A_{22})^{-1}$. Furthermore,

$$(\lambda_0\, I - A)^\sharp = \begin{bmatrix} 0 & A_{12}(\lambda_0\, I - A_{22})^{-2} \\ 0 & (\lambda_0\, I - A_{22})^{-1} \end{bmatrix} = \begin{bmatrix} 0 & r(\lambda_0\, I - A_{22})^{-1} \\ 0 & (\lambda_0\, I - A_{22})^{-1} \end{bmatrix},$$

$$\big[P_0(\lambda_0\, I - A_{22})\big]^\dagger = \begin{bmatrix} 0 & 0 \\ 0 & \lambda_0\, I - A_{22} \end{bmatrix}^\dagger = \begin{bmatrix} 0 & 0 \\ 0 & (\lambda_0\, I - A_{22})^{-1} \end{bmatrix}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The statements of the theorem are obtained by specializing Theorems 3.1 and 3.6 to the case $A_{11} = \lambda_0$, $\|\cdot\| = |\cdot| = \|\cdot\|_2$. In this case we have the following identities:

$$\begin{aligned}
\mathbb{T}(z) &= (\lambda I - A_{22})z,\\
\mathbb{T}^{-1}(E_{21}) &= (\lambda I - A_{22})^{-1}E_{21},\\
\|A_{12}\|_2 &= \|x_0^*(\lambda_0 I - A)\|_2 = c,\\
p &= \sqrt{1+\|r\|_2^2} = \|y_0\|_2 = p_0,\\
\kappa &= \sqrt{1+\|r\|_2^2} + \|r\|_2 = \|y_0\|_2 + \sqrt{\|y_0\|_2^2 - 1} = \kappa_0,\\
\operatorname{sep}_2(\lambda_0 I, A_{22}) &= \sigma_{\min}(\lambda_0 I - A_{22}) = \sigma_{n-1}(P_0(\lambda_0 I - A_{22})) = s_0.
\end{aligned}$$

Let $E \in \mathbb{C}^{n\times n}$ with $\|E\|_2 < \frac{1}{2}(\sqrt{s_0^2+c^2} - c)$. According to Theorem 3.1 there exists a vector $z_E \in \mathbb{C}^{n-1}$ depending holomorphically on E such that $x_E = [1 \;\; z_E^\top]^\top$ is an eigenvector of $A+E$, $\|z_E\|_2 \le \frac{2}{s_0}\|E\|_2$ and $\|z_E - \mathbb{T}^{-1}(E_{21})\|_2 \le 6\|E\|_2^2/s_0^2$. Clearly, $x_0^* x_E = 1$ and $\|z_E\|_2 = \tan(\theta(x_0, x_E))$. It is straightforward to verify that

$$\|\xi_E\|_2 = \|x_E - (I + [P_0(\lambda I - A)]^\dagger E)x_0\|_2 = \|z_E - \mathbb{T}^{-1}(E_{21})\|_2.$$

This concludes the proof of (i).

To show (ii), suppose that $\|E\|_2 < \kappa_0/(2s_0)$. According to Theorem 3.6 there exists a vector $w_E \in \mathbb{C}^{n-1}$ depending holomorphically on E such that $\tilde{x}_E = \begin{bmatrix} 1 & r \\ 0 & I \end{bmatrix}\begin{bmatrix} 1 \\ w_E \end{bmatrix}$ is an eigenvalue of $A+E$ and

$$\|w_E\|_2 \le \frac{2\kappa_0}{p_0\, s_0}\|E\|_2 \qquad \text{as well as} \qquad \|w_E - \mathbb{T}^{-1}(E_{21})\|_2 \le \frac{6\kappa_0^2}{p_0\, s_0^2}\|E\|_2^2.$$

It is easily verified that

$$\begin{aligned}
\tilde{x}_E - x_0 &= \begin{bmatrix} r \\ I \end{bmatrix} w_E \qquad \text{and}\\
\tilde{\xi}_E = \tilde{x}_E - (x_0 + (\lambda_0 I - A)^\sharp E x_0) &= \begin{bmatrix} r \\ I \end{bmatrix} (w_E - \mathbb{T}^{-1}(E_{21})).
\end{aligned}$$

This yields (4.1) and the first inequality in (4.2), since $\|[r^\top \;\; I]^\top\|_2 = p_0$. We have $\tilde{x}_E = (1+rw_E)\begin{bmatrix} 1 & (1+rw_E)^{-1}w_E^\top \end{bmatrix}^\top$. Thus, $\tan(\theta(x_0, \tilde{x}_E)) = \|(1+rw_E)^{-1}w_E\|_2 \le \|w_E\|_2/(1-\|r\|_2\|w_E\|_2)$. This implies the second inequality in (4.2). Equation (4.3) follows by multiplying the identity $(A+E)\tilde{x}_E = \lambda_E \tilde{x}_E$ with y_0^*. The estimates in (4.4) are then obvious. ∎

Let $A \in \mathbb{C}^{n\times n}$ be a normal matrix. Then we can take $y_0 = x_0$ in Theorem 4.1. Moreover, the separation $s_0 = \sigma_{n-1}(P_0(\lambda_0 I - A)) = \sigma_{n-1}(\lambda_0 I - A)$ equals the distance of λ_0 to the set $\Lambda(A) \setminus \{\lambda_0\}$, and the identities $a = 0$, $p_0 = \kappa_0 = 1$, $[P_0(\lambda_0 I - A)]^\dagger = (\lambda_0 I - A)^\dagger = (\lambda_0 I - A)^\sharp$ hold. We thus have the following corollary to Theorem 4.1.

COROLLARY 4.2. *Let $A \in \mathbb{C}^{n\times n}$ be a normal matrix and consider a normalized eigenvector $x_0 \in \mathbb{C}^n$ belonging to a simple eigenvalue λ_0 of A. Let s_0 denote the distance of λ_0 to the rest of the spectrum of A_0, that is, $s_0 = \min\{|\lambda_0 - \nu| : \nu \in \Lambda(A), \nu \ne \lambda_0\}$. Let $E \in \mathbb{C}^{n\times n}$ with $\|E\|_2 < s_0/2$. Then there exists an eigenvector x_E of $A+E$ depending holomorphically on E such that $x_0^* x_E = 1$ and*

$$x_E = (I + (\lambda I - A)^\dagger E)x_0 + \xi_E$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

for some $\xi_E \in \mathbb{C}^n$ *with* $\|\xi_E\|_2 \leq 6\|E\|_2^2/s_0^2$. *Furthermore,* $\|x_E - x_0\|_2 = \tan(\theta(x_0, x_E)) \leq 2\,\|E\|_2/s_0$. *The associated eigenvalue satisfies*

$$\lambda_E = \lambda_0 + x_0^* E x_E = \lambda_0 + x_0^* E x_0 + \ell_E = \lambda_0 + y_0^* E x_0 + x_0^* E (\lambda_0 I - A)^\dagger E x_0 + \tilde{\ell}_E$$

with $|\ell_E| \leq 2\|E\|_2^2/s_0$ *and* $|\tilde{\ell}_E| \leq 6\|E\|_2^3/s_0^2$.

5. Conclusions. By establishing a link to the coalescence of pseudospectral components, we have derived a new perturbation bound for invariant subspaces. As the bound turns out to be sharp for a 2×2 example, no further obvious improvement of the bound seems to be possible. Moreover, we establish a novel bound for the remainder term of a well-known perturbation expansion. Even the (modified) specialization of this remainder bound to the case of individual eigenvectors and eigenvalues appears to be new. We believe that such bounds for remainder terms are important; e.g., they can be used for quantifying the validity for condition numbers frequently used in practice, e.g., in MATLAB and LAPACK [2].

As a side result, we have shown that Stewart's classical result on the perturbation of invariant subspaces is a direct consequence of the Newton–Kantorovich theorem. We believe that there is some interest in this observation, as it may more easily allow for extensions of Stewart's result to different settings.

Appendix A. Stewart's result via the Newton–Kantorovich theorem. In this section, we show that Theorem 2.5 is a special case of the Newton–Kantorovich theorem formulated in [12, p. 536].

THEOREM A.1. *Let* $\mathcal{E}, \mathcal{Z}$ *be Banach spaces and let* $f : \mathcal{Z} \to \mathcal{E}$ *be twice continuously differentiable in a sufficiently large neighborhood* Ω *of* $Z \in \mathcal{Z}$. *Suppose that there exists a linear operator* $\mathbb{T} : \mathcal{Z} \to \mathcal{E}$ *having a continuous inverse* $\mathbb{T}^{-1}$ *and satisfying the following conditions:*

$$\|\mathbb{T}^{-1}(F(Z))\| \leq \eta, \tag{A.1}$$

$$\|\mathbb{T}^{-1} \circ F'(Z) - I\| \leq \delta, \tag{A.2}$$

$$\|\mathbb{T}^{-1} \circ F''(\tilde{Z})\| \leq K \qquad \forall \tilde{Z} \in \Omega. \tag{A.3}$$

If $\delta < 1$ *and* $h := \frac{\eta K}{(1-\delta)^2} < \frac{1}{2}$, *then there exists a solution* Z_E *of* $F(Z_E) = 0$ *such that*

$$\|Z_E - Z\| \leq r_0 \quad \textit{with} \quad r_0 := \frac{2\eta}{(1-\delta)(1+\sqrt{1-2h})}.$$

We apply Theorem A.1 to the setting of section 2.4:

$$0 = F(Z) := -f(A+E, Z) = -E_{21} + Z(A_{11} + E_{11}) - (A_{22} + E_{22})Z + Z(A_{12} + E_{12})Z$$

with $\mathcal{E} = \mathcal{Z} = \mathbb{C}^{(n-k)\times k}$, $Z = 0$, and $\mathbb{T} : Z \mapsto ZA_{11} - A_{22}Z$. In the following, $\|\cdot\|$ denotes a consistent family of matrix norms.

Condition (A.1). We have

$$\|\mathbb{T}^{-1}(F(0))\| = \|\mathbb{T}^{-1}(E_{21})\| \leq \frac{\|E_{21}\|}{s} =: \eta,$$

where $s = \operatorname{sep}(A_{11}, A_{22})$.

Condition (A.2). From

$$F'(0) : \triangle Z \mapsto (A_{22} + E_{22})\triangle Z - \triangle Z(A_{11} + E_{11}) = \mathbb{T}(\triangle Z) + \triangle\mathbb{T}(\triangle Z)$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with $\triangle\mathbb{T}(\triangle Z) := E_{22} \cdot \triangle Z - \triangle Z \cdot E_{11}$, it follows that

$$\|\mathbb{T}^{-1} \circ F'(0) - I\| = \|\mathbb{T}^{-1} \circ \triangle\mathbb{T}\| \leq \frac{\|E_{11}\| + \|E_{22}\|}{s} =: \delta.$$

Condition (A.3). Since the second derivative of f is constant, it immediately follows that

$$\|\mathbb{T}^{-1} \circ F''(\tilde{Z})\| \leq 2\frac{\|A_{12} + E_{12}\|}{s} \leq 2\frac{\|A_{12}\| + \|E_{12}\|}{s} =: K.$$

Summary. Setting $s_E = s - \|E_{11}\| - \|E_{22}\|$, we finally obtain

$$h = \frac{\eta K}{(1-\delta)^2} = 2\frac{\|E_{21}\|(\|A_{12}\| + \|E_{12}\|)}{s_E^2}$$

$$r_0 = \frac{2\eta}{(1-\delta)(1+\sqrt{1-2h})} = \frac{2\|E_{21}\|}{s_E + \sqrt{s_E^2 - 4\|E_{21}\|(\|A_{12}\| + \|E_{12}\|)}}.$$

Theorem A.1 now states the existence of a solution Z_E to $F(Z) = -f(A+E, Z_E) = 0$ with $\|Z_E\| \leq r_0$ if $\delta < 1$ and $h < \frac{1}{2}$. This coincides precisely with the statement of Theorem 2.5.

Extension. The statement of Theorem 2.5 can be improved when we assume that $\|\cdot\|$ is unitarily invariant. In this case, the quantities δ, K in the derivation above can be replaced by the potentially smaller quantities

$$\delta = \frac{\|E_{11}\|_2 + \|E_{22}\|_2}{s}, \quad K = 2\frac{\|A_{12}\|_2 + \|E_{12}\|_2}{s}.$$

Consequently, the bound of Theorem 2.5 becomes

$$\|Z_E\| \leq \frac{2\|E_{21}\|}{s_E + \sqrt{s_E^2 - 4\|E_{21}\|(\|A_{12}\|_2 + \|E_{12}\|_2)}} < 2\frac{\|E_{21}\|}{s_E} \tag{A.4}$$

under the condition $\|E_{21}\|(\|A_{12}\|_2 + \|E_{12}\|_2) < s_E^2/4$ with $s_E = s - \|E_{11}\|_2 - \|E_{22}\|_2$.

Acknowledgments. The authors thank Pete Stewart for very helpful discussions. Parts of this work were prepared while the first author was visiting FIM (Institute for Mathematical Research) at ETH Zurich. The generous hospitality of FIM is gratefully acknowledged.

REFERENCES

[1] R. Alam and S. Bora, *On stable eigendecompositions of matrices*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 830–848.
[2] Z. Bai, J. W. Demmel, and A. McKenney, *On computing condition numbers for the nonsymmetric eigenproblem*, ACM Trans. Math. Software, 19 (1993), pp. 202–223.
[3] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997.
[4] R. Byers, *A LINPACK-style condition estimator for the equation $AX - XB^T = C$*, IEEE Trans. Automat. Control, 29 (1984), pp. 926–928.
[5] F. Chatelin, *Spectral Approximation of Linear Operators*, Academic Press, New York, 1983.
[6] C. Davis and W. M. Kahan, *The rotation of eigenvectors by a perturbation.* III, SIAM J. Numer. Anal., 7 (1970), pp. 1–46.
[7] J. W. Demmel, *The condition number of equivalence transformations that block diagonalize matrix pencils*, SIAM J. Numer. Anal., 20 (1983), pp. 599–610.
[8] J. W. Demmel, *Computing stable eigendecompositions of matrices*, Linear Algebra Appl., 79 (1986), pp. 163–193.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[9] I. Gohberg, P. Lancaster, and L. Rodman, *Invariant Subspaces of Matrices with Applications*, Classics in Appl. Math. 51, SIAM, Philadelphia, 2006.
[10] L. Grammont and A. Largillier, *On ϵ-spectra and stability radii*, J. Comput. Appl. Math., 147 (2002), pp. 453–469.
[11] M. Gu and M. L. Overton, *An algorithm to compute* Sep_λ, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 348–359.
[12] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, 2nd ed., Pergamon Press, Oxford, UK, 1982.
[13] M. Karow, *Inclusion theorems for pseudospectra of block triangular matrices*, in preparation.
[14] T. Kato, *Perturbation Theory for Linear Operators*, Classics in Math., Springer-Verlag, Berlin, 1995.
[15] S. G. Krantz and H. R. Parks, *The Implicit Function Theorem*, Birkhäuser, Boston, 2003.
[16] D. Kressner, *Numerical Methods for General and Structured Eigenvalue Problems*, of Lect. Notes Comput. Sci. Eng. 16, Springer-Verlag, Berlin, 2005.
[17] C. D. Meyer and G. W. Stewart, *Derivatives and perturbations of eigenvectors*, SIAM J. Numer. Anal., 25 (1988), pp. 679–691.
[18] G. W. Stewart, *Error bounds for approximate invariant subspaces of closed linear operators*, SIAM J. Numer. Anal., 8 (1971), pp. 796–808.
[19] G. W. Stewart, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Rev., 15 (1973), pp. 727–764.
[20] G. W. Stewart, *Smooth local bases for perturbed eigenspaces*, UMIACS-TR-2012-08, CS-TR-4010, May 2012, Department of Computer Science, University of Maryland.
[21] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.
[22] J.-G. Sun, *Estimation of the separation of two matrices*, J. Comput. Math., 2 (1984), pp. 189–200.
[23] J.-G. Sun, *Estimation of the separation of two matrices.* II, J. Comput. Math., 3 (1985), pp. 19–26.
[24] J.-G. Sun, *Perturbation expansions for invariant subspaces*, Linear Algebra Appl., 153 (1991), pp. 85–97.
[25] J.-G. Sun, *Stability and Accuracy: Perturbation Analysis of Algebraic Eigenproblems*, Technical report UMINF 98-07, Department of Computing Science, University of Umeå, Umeå, Sweden, 1998.
[26] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra. The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, NJ, 2005.
[27] J. M. Varah, *On the separation of two matrices*, SIAM J. Numer. Anal., 16 (1979), pp. 216–222.
[28] H. Xu, *Bounds on the separation of two matrices*, J. Fudan Univ. Nat. Sci., 33 (1994), pp. 413–420.
[29] S.-F. Xu, *Lower bound estimation for the separation of two matrices*, Linear Algebra Appl., 262 (1997), pp. 67–82.

Downloaded 12/14/17 to 130.149.176.172. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.