BIT
2000, Vol. 40, No. 2, pp. 395–403

0006-3835/00/4002-0395 $15.00

A NOTE ON SIN Θ THEOREMS FOR SINGULAR SUBSPACE VARIATIONS *

FROILÁN M. DOPICO

Departamento de Matemáticas, Universidad Carlos III, Avda. de la Universidad, 30 28911 Leganés, Madrid, Spain. email: dopico@math.uc3m.es

Abstract.

New perturbation theorems for bases of singular subspaces are proved. These theorems complement the known $\sin\Theta$ theorems for singular subspace perturbations, taking into account a kind of sensitivity of singular vectors discarded by previous theorems. Furthermore these results guarantee that high relative accuracy algorithms for the SVD are able to compute reliably simultaneous bases of left and right singular subspaces.

AMS subject classification: 65F15, 15A18.

Key words: Singular value decomposition, perturbation bounds.

1 Introduction.

Let A and $\widetilde{A}$ be two complex $m \times n$ $(m \geq n)$ matrices with conformally partitioned singular value decompositions (SVD):

$$A = \begin{pmatrix} U_1 & U_2 & U_3 \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1 & V_2 \end{pmatrix}^*, \tag{1.1}$$

$$\widetilde{A} = \begin{pmatrix} \widetilde{U}_1 & \widetilde{U}_2 & \widetilde{U}_3 \end{pmatrix} \begin{pmatrix} \widetilde{\Sigma}_1 & 0 \\ 0 & \widetilde{\Sigma}_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \widetilde{V}_1 & \widetilde{V}_2 \end{pmatrix}^* \tag{1.2}$$

where Σ_1 and Σ_2 have dimensions $k \times k$ and $(n-k) \times (n-k)$ respectively, and B^* is the conjugate transpose of B for any matrix B. No particular order is assumed on the singular values. Let the following residuals be defined:

$$R = A\widetilde{V}_1 - \widetilde{U}_1\widetilde{\Sigma}_1 = (A - \widetilde{A})\widetilde{V}_1 \tag{1.3}$$

and

$$S = A^*\widetilde{U}_1 - \widetilde{V}_1\widetilde{\Sigma}_1 = (A^* - \widetilde{A}^*)\widetilde{U}_1. \tag{1.4}$$

In this paper we will be interested in using the residuals (1.3)–(1.4) to obtain reliable perturbations bounds on singular vectors. One of the most useful ways of doing this is to bound the sines of the canonical angles (see [10] or [8]) between

*Received June 1999. Communicated by Lars Eldén.

the column spaces $\mathcal{R}(U_1)$ of U_1 and $\mathcal{R}(\widetilde{U}_1)$ of $\widetilde{U}_1$, as well as between the column spaces of V_1 and $\widetilde{V}_1$. This approach was taken by Wedin in [9], extending the well-known sinΘ theorems for hermitian matrices of Davis and Kahan [1]. If we denote by Φ the matrix of canonical angles between $\mathcal{R}(U_1)$ and $\mathcal{R}(\widetilde{U}_1)$, and by Θ the matrix of canonical angles between $\mathcal{R}(V_1)$, $\mathcal{R}(\widetilde{V}_1)$, Wedin proved the following classic theorem in the Frobenius norm:

THEOREM 1.1. *Let A and $\widetilde{A}$ be two $m \times n$ $(m \geq n)$ complex matrices with SVDs* (1.1) *and* (1.2). *Define*

$$\delta = \min_{\widetilde{\mu} \in \sigma(\widetilde{\Sigma}_1), \mu \in \sigma_{ext}(\Sigma_2)} |\widetilde{\mu} - \mu| \tag{1.5}$$

where, for any matrix B, $\sigma(B)$ denotes the set of its singular values and $\sigma_{ext}(\Sigma_2) \equiv \sigma(\Sigma_2) \cup \{0\}$ if $m > n$ and $\sigma_{ext}(\Sigma_2) \equiv \sigma(\Sigma_2)$ if $m = n$. If $\delta > 0$ then

$$\sqrt{\| \sin \Phi\|_F^2 + \| \sin \Theta\|_F^2} \leq \frac{\sqrt{\|R\|_F^2 + \|S\|_F^2}}{\delta}. \tag{1.6}$$

By imposing further restrictions on the singular values, Wedin also proved bounds on any unitarily invariant norm, but we will focus on the previous and more general theorem. Recently Li [6] has obtained relative $\sin\Theta$ theorems for singular subspaces in the case of multiplicative perturbations. These theorems lead to sharper relative bounds when the multiplicative perturbations are close to the identity matrix.

Notice that the bound (1.6) implies that if the residuals are small, and the gap δ is not too small, the canonical angles are also small. However, this does not mean that the differences $U_1 - \widetilde{U}_1$ and $V_1 - \widetilde{V}_1$ are small. When the angles in Φ are small one can only prove that there is an orthonormal basis of $\mathcal{R}(U_1)$ which is close to the orthonormal basis of $\mathcal{R}(\widetilde{U}_1)$ formed by the columns of $\widetilde{U}_1$. The same can be said for the V's if the angles of Θ are small. This is equivalent to saying that there are two $k \times k$ unitary matrices P and Q such that $\|U_1 P - \widetilde{U}_1\|_F$ and $\|V_1 Q - \widetilde{V}_1\|_F$ are small (see [8, Theorem I.5.2]). Here the drawback of $\sin\Theta$ theorems for singular subspaces becomes apparent: they do not prevent P and Q from being different. However, P and Q should be the same if we want to take the columns of the matrices $\widetilde{U}$ and $\widetilde{V}$ corresponding to nonzero singular values as reliable approximations of a pair of singular vector matrices of the unperturbed matrix A. To be precise, let $\sigma(\Sigma_1)$ and $\sigma(\Sigma_2)$ be disjoint with Σ_1 nonsingular. Then one can easily check that for any other SVD of A of the form

$$A = \begin{pmatrix} \widehat{U}_1 & \widehat{U}_2 & \widehat{U}_3 \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \widehat{V}_1 & \widehat{V}_2 \end{pmatrix}^*$$

there exists a unitary $k \times k$ matrix Q such that

$$\widehat{U}_1 = U_1 Q \quad \text{and} \quad \widehat{V}_1 = V_1 Q. \tag{1.7}$$

In other words any pairs of simultaneous bases of left and right singular subspaces must be related through the same unitary change of basis. (In fact, Q is block diagonal, with each block corresponding to a different singular value.)

A simple example may help to clarify our point.

EXAMPLE 1.1. Let A and $\widetilde{A}$ be

$$A = \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{and} \quad \widetilde{A} = \begin{pmatrix} -\epsilon & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{1.8}$$

where $\epsilon > 0$ is small enough. The left and right singular vectors of A corresponding to the singular value ϵ are (up to a complex factor of modulus one) $u_1 = (1, 0, 0)^T$ and $v_1 = (1, 0, 0)^T$. The corresponding vectors of $\widetilde{A}$ are $\widetilde{u}_1 = (1, 0, 0)^T$ and $\widetilde{v}_1 = (-1, 0, 0)^T$. One can easily check that $\|R\|_F = \|S\|_F = 2\epsilon$ and $\delta = 1 - \epsilon$ in Theorem 1.1 (A and $\widetilde{A}$ are square). Thus the bound on the sines of the angles given by (1.6) is $(2\sqrt{2}\epsilon)/(1 - \epsilon)$. This bound is, of course, valid because the singular subspaces of A and $\widetilde{A}$ coincide. However, it does not give any information about the fact that for any $\epsilon > 0$, any pair of simultaneous singular vectors u_1 and v_1 of A is necessarily far from any pair of simultaneous singular vectors $\widetilde{u}_1$ and $\widetilde{v}_1$ of $\widetilde{A}$.

The purpose of this note is to complement absolute and relative $\sin\Theta$ theorems for singular subspaces with results that give information on the relation of *simultaneous bases* of perturbed left and right singular subspaces with corresponding unperturbed simultaneous bases. This will be done in the second section, leading to important changes in the $\sin\Theta$ bounds only in the absolute case, but not in the relative case. The third section contains the proofs of the theorems in Section 2 and also a proof that the new gap appearing in Theorem 2.1 is necessary.

2 Results and comments.

Our first theorem uses the same notation of Theorem 1.1:

THEOREM 2.1. *Let A and $\widetilde{A}$ be two $m \times n$ ($m \geq n$) complex matrices with SVDs* (1.1) *and* (1.2). *Define*

$$\delta_b = \min\left\{ \left(\min_{\widetilde{\mu} \in \sigma(\widetilde{\Sigma}_1), \mu \in \sigma_{ext}(\Sigma_2)} |\widetilde{\mu} - \mu| \right),\ \sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1) \right\} \tag{2.1}$$

where $\sigma_{\min}(\Sigma_1)$ and $\sigma_{\min}(\widetilde{\Sigma}_1)$ denote the minimum of the singular values of Σ_1 and $\widetilde{\Sigma}_1$ respectively. If $\delta_b > 0$ then

$$\min_{W \text{unitary}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2} \leq \sqrt{2} \frac{\sqrt{\|R\|_F^2 + \|S\|_F^2}}{\delta_b}. \tag{2.2}$$

Moreover, the left hand side of (2.2) *is minimized for $W = YZ^*$, where YSZ^* is any SVD of $U_1^* \widetilde{U}_1 + V_1^* \widetilde{V}_1$, and the equality can be attained.*

Remark 2.1. The left hand side of (2.2) is an extension to pairs of subspaces of the unitarily invariant metric on the set of subspaces introduced by Paige [7] in a different context (see also [8, Section II.4.2]).

Apart from the numerical factor $\sqrt{2}$ the bound in Theorem 2.1 differs from that in Wedin's Theorem 1.1 in the definition of the gap δ_b, which is the minimum of Wedin's gap δ and the new term $\sigma_{\min}(\widetilde{\Sigma}_1)+\sigma_{\min}(\Sigma_1)$. This term is only relevant in the case $m = n$, because $\delta \leq (\sigma_{\min}(\widetilde{\Sigma}_1)+\sigma_{\min}(\Sigma_1))$ when $m > n$ and therefore $\delta_b = \delta$. This case $m > n$ is easily understood: if $m > n$ it is well known that the subspaces are ill-determined whenever small singular values are present in $\widetilde{\Sigma}_1$. This is due to the existence of $m - n$ "ghost" singular values equal to zero and their associated singular subspace $\mathcal{R}(U_3)$, which perturbs $\mathcal{R}(U_1)$[1](see [8, p. 262]). Thus there is no hope for the bases to be well-determined in this case. However in the square case $m = n$ the new term $\sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1)$ can play an important role in δ_b and its presence, which reflects the different sensitivities of simultaneous bases and subspaces, has to be explained. Although in Section 3 we will prove that $\sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1)$ is necessary to bound the left hand side of (2.2), in the next paragraph some arguments are provided for the presence of this term in the definition of δ_b.

In the first place, when Theorem 2.1 is applied to Example 1.1 both the left and right hand side of (2.2) are equal to 2, because $\delta_b = \sigma_{\min}(\widetilde{\Sigma}_1)+\sigma_{\min}(\Sigma_1) = 2\epsilon \ll \delta$. This proves that bound (2.2) is sharp in this case due precisely to the presence of the new term. In the second place, analogous terms have already appeared in the different but related problem of obtaining perturbation bounds for the unitary polar factor (see [4] and references therein). The relationship between this problem and ours is that the unitary polar factor depends on matrices of *simultaneous* left and right singular vectors. In fact, it is easy to obtain from (2.2) a perturbation bound for the unitary polar factor which, although weaker than that provided in [4], reflects the same sensitivity. This is again possible due to the presence of $\sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1)$. Finally, it should be noticed that (1.7) is no longer valid when Σ_1 is singular. Thus in this case it is hopeless to look for reliable bounds on the left hand side of (2.2). Therefore it is natural that the gap δ_b depends on how far is Σ_1 ($\widetilde{\Sigma}_1$) from being singular.

Notice that Wedin's Theorem 1.1 and Theorem 2.1 are independent. It is easy to prove that $\|\sin\Phi\| \leq \|U_1 W - \widetilde{U}_1\|$ and $\|\sin\Theta\| \leq \|V_1 W - \widetilde{V}_1\|$ for any unitarily invariant norm $\|\cdot\|$ and any unitary matrix W. Then the bound (2.2) implies a bound for the sines of the canonical angles but weaker than that of Wedin (1.6) (specially when $m = n$). On the other hand a bound on the sines of canonical angles cannot imply a simultaneous bound for left and right singular bases, as Example 1.1 shows. In this sense Theorem 2.1 complements the classical Wedin's theorem. When Wedin's bound is satisfactory we can ask:

[1]The right subspace $\mathcal{R}(V_1)$ is not affected by "ghost" singular values, although this cannot be seen in a joint bound like Wedin's (1.6). A finer analysis can circumvent this difficulty [8, p. 262] for subspaces. However this problem cannot be solved for bounds on simultaneous bases, because both the bases of $\mathcal{R}(U_1)$ and $\mathcal{R}(V_1)$ are ill-determined if small singular values are present in Σ_1 ($\widetilde{\Sigma}_1$).

is also the bound (2.2) satisfactory? The answer is "yes" if the singular values of Σ_1 ($\widetilde{\Sigma}_1$) are far from zero.

Now we will consider multiplicative perturbations and relative perturbation bounds. These bounds play an important role in devising high relative accuracy algorithms to compute singular value decomposition [2], and have been developed by several authors. One of the most general and systematic treatments is that of Li [6]. We will state a relative version of Theorem 2.1 for multiplicative perturbations which complements the $\sin\Theta$ relative Theorem 4.1 in [6]. Notation and definitions of previous theorems are used:

THEOREM 2.2. *Let A and $\widetilde{A} = D_1^* A D_2$ be two $m \times n$ complex matrices, where $m \geq n$ and D_1, D_2 are nonsingular matrices. Let A and $\widetilde{A}$ have SVDs* (1.1) *and* (1.2)*, with Σ_1 and $\widetilde{\Sigma}_1$ non-singular. Define*

$$\rho_2 = \min_{\mu \in \sigma(\Sigma_1), \widetilde{\mu} \in \sigma_{ext}(\widetilde{\Sigma}_2)} \frac{|\mu - \widetilde{\mu}|}{\sqrt{\mu^2 + \widetilde{\mu}^2}}, \tag{2.3}$$

$$\rho_b = \min\left\{ \left(\min_{\mu \in \sigma(\Sigma_1), \widetilde{\mu} \in \sigma_{ext}(\widetilde{\Sigma}_2)} \frac{|\mu - \widetilde{\mu}|}{\widetilde{\mu}} \right), \left(\min_{\mu \in \sigma(\Sigma_1), \widetilde{\mu} \in \sigma(\widetilde{\Sigma}_1)} \frac{\mu + \widetilde{\mu}}{\widetilde{\mu}} \right) \right\}. \tag{2.4}$$

If $\rho_2, \rho_b > 0$ then

$$\min_{W \text{unitary}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2} \tag{2.5}$$

$$\leq \frac{\sqrt{2}}{\rho_2} \sqrt{\|(I - D_1^*)U_1\|_F^2 + \|(I - D_2^*)V_1\|_F^2 + \|(I - D_1^{-1})U_1\|_F^2 + \|(I - D_2^{-1})V_1\|_F^2}$$

and

$$\min_{W \text{unitary}} \sqrt{\|U_1 W - \widetilde{U}_1\|_F^2 + \|V_1 W - \widetilde{V}_1\|_F^2} \tag{2.6}$$

$$\leq \sqrt{2}\Bigg(\sqrt{\|(I - D_1^*)U_1\|_F^2 + \|(I - D_2^*)V_1\|_F^2} + \frac{1}{\rho_b} \sqrt{\|(D_1^* - D_1^{-1})U_1\|_F^2 + \|(D_2^* - D_2^{-1})V_1\|_F^2} \Bigg).$$

Moreover the left hand sides of (2.5) *and* (2.6) *are minimized for $W = YZ^*$, where YSZ^* is any SVD of $U_1^* \widetilde{U}_1 + V_1^* \widetilde{V}_1$.*

REMARK 2.2. In the definitions of ρ_2 and ρ_b any quotient $0/0$ must be taken as 0. The assumption about the non-singularity of Σ_1 and $\widetilde{\Sigma}_1$ is done only for simplifying the expression for ρ_2, otherwise ρ_2 is given by (3.8). Moreover, when Σ_1 or $\widetilde{\Sigma}_1$ is singular it makes no sense to look for relations between the simultaneous bases of the corresponding singular subspaces.

When comparing the relative bounds (2.5) and (2.6) with those of [6, Theorem 4.1] the situation is different to that in the absolute case: here the main difference is in the factor $\sqrt{2}$ and not in the gaps.[2] In fact ρ_2 is the same quantity as in [6], while ρ_b differs from the quantity in [6] only in the presence of the second argument of (2.4). However, the consequences of this change are not so drastic as in the absolute case because

$$\min_{\mu\in\sigma(\Sigma_1),\widetilde{\mu}\in\sigma(\widetilde{\Sigma}_1)} \frac{\widetilde{\mu}+\mu}{\widetilde{\mu}} \geq 1$$

Thus if ρ_b is smaller than one, this must be due to the gap defined in [6, Theorem 4.1].

These remarks are interesting when computing the SVD of a matrix A with high relative accuracy [2], because when this is possible the algorithms essentially give the SVD of a small multiplicative perturbation of A. In this case Theorem 2.2 guarantees that for well separated clusters of nonzero singular values (regardless of their magnitudes) the simultaneous bases of left and right singular subspaces are reliably computed. To our knowledge this has not been previously observed. As Theorem 2.1 and Example 1.1 show, the same cannot be stated for standard SVD algorithms [3, p. 261].

To finish we would like to point out that Theorem 2.2 and Li's Theorem 4.1 [6] are independent. However, as in the absolute case, Theorem 2.2 implies a bound on the sines of the canonical angles, which now is essentially equivalent to Theorem 4.1 of [6] up to a factor $\sqrt{2}$.

3 Proofs of Theorems 2.1 and 2.2.

Proof of Theorem 2.1. Let us define the following $(m+n)\times k$ matrices with orthonormal columns:

$$X_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} U_1 \\ V_1 \end{pmatrix} \quad \text{and} \quad \widetilde{X}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} \widetilde{U}_1 \\ \widetilde{V}_1 \end{pmatrix},$$

and notice that

$$\min_{W\text{unitary}} \sqrt{\|U_1W-\widetilde{U}_1\|_F^2 + \|V_1W-\widetilde{V}_1\|_F^2} = \sqrt{2} \min_{W\text{unitary}} \|X_1W-\widetilde{X}_1\|_F. \tag{3.1}$$

Now we use [8, Theorem II.4.11] to prove that

$$\min_{W\text{unitary}} \|X_1W-\widetilde{X}_1\|_F$$

$$= \sqrt{\|I-\cos\Theta(\mathcal{R}(X_1),\mathcal{R}(\widetilde{X}_1))\|_F^2 + \|\sin\Theta(\mathcal{R}(X_1),\mathcal{R}(\widetilde{X}_1))\|_F^2} \tag{3.2}$$

$$\leq \sqrt{2}\|\sin\Theta(\mathcal{R}(X_1),\mathcal{R}(\widetilde{X}_1))\|_F.$$

[2] In the relative gaps defined in [6, Theorem 4.1], if Σ_1 is non-singular it is not necessary to distinguish between $\sigma_{ext}(\widetilde{\Sigma}_2)$ $(m>n)$ and $\sigma(\widetilde{\Sigma}_2)$ $(m=n)$ when the minimums are taken. The same is true for Theorem 2.2. However we consider it clearer to state Theorem 2.2 in the same way as Theorem 4.1 [6].

This leads us to consider the perturbation problem for the Jordan–Wielandt matrices

$$C = \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \quad \text{and} \quad \widetilde{C} = \begin{pmatrix} 0 & \widetilde{A} \\ \widetilde{A}^* & 0 \end{pmatrix}$$

because X_1 and $\widetilde{X}_1$ are the matrices of orthonormal eigenvectors of C and $\widetilde{C}$ corresponding to the eigenvalues of Σ_1 and $\widetilde{\Sigma}_1$. If we set

$$T_1 = C\widetilde{X}_1 - \widetilde{X}_1\widetilde{\Sigma}_1,$$

then

$$\|\sin\Theta(\mathcal{R}(X_1), \mathcal{R}(\widetilde{X}_1))\|_F \le \frac{\|T_1\|_F}{\delta_b} \tag{3.3}$$

using Davis and Kahan's $\sin\Theta$ theorem for the Hermitian matrices C and $\widetilde{C}$ [1]. Furthermore it is easy to prove that

$$\|T_1\|_F^2 = \frac{1}{2}\left(\|R\|_F^2 + \|S\|_F^2\right). \tag{3.4}$$

The expression (2.2) follows from combining (3.1)–(3.4). The solution of the orthogonal Procrustes problem for X_1 and $\widetilde{X}_1$ implies that the unitary matrix $W = YZ^*$ minimizes the left hand side of (2.2) [3, Sec. 12.4.1]. Finally, when (2.2) is applied to Example 1.1 equality is attained. □

REMARK 3.1. The necessity of $\sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1)$:
The previous proof is rather short because it relies on known results, but does not show clearly why the term $\sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1)$ appears in the gap δ_b. This can be seen as follows. Notice that (3.1) and (3.2) prove that to bound the left hand side of (2.2) we have to bound the sines of the canonical angles between $\mathcal{R}(X_1)$ and $\mathcal{R}(\widetilde{X}_1)$. We will do this by a direct calculation, assuming for simplicity that $m = n$. It is known that

$$\begin{aligned} \left\|\sin\Theta(\mathcal{R}(X_1), \mathcal{R}(\widetilde{X}_1))\right\|_F^2 &= \left\| \frac{1}{\sqrt{2}} \begin{pmatrix} U_1 \\ V_1 \end{pmatrix}^* \frac{1}{\sqrt{2}} \begin{pmatrix} \widetilde{U}_2 & -\widetilde{U}_1 & -\widetilde{U}_2 \\ \widetilde{V}_2 & \widetilde{V}_1 & \widetilde{V}_2 \end{pmatrix} \right\|_F^2 \\ &= \frac{1}{4}\left(2\|U_1^*\widetilde{U}_2\|_F^2 + 2\|V_1^*\widetilde{V}_2\|_F^2 + \|V_1^*\widetilde{V}_1 - U_1^*\widetilde{U}_1\|_F^2\right). \end{aligned} \tag{3.5}$$

Notice that

$$\|U_1^*\widetilde{U}_2\|_F^2 + \|V_1^*\widetilde{V}_2\|_F^2 = \|\sin\Phi\|_F^2 + \|\sin\Theta\|_F^2 \tag{3.6}$$

can be bounded using δ due to Theorem 1.1. Therefore $\sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1)$ must come from the contribution of $\|V_1^*\widetilde{V}_1 - U_1^*\widetilde{U}_1\|_F^2$, and indeed this is the case. As in [4], we pose the Sylvester equation

$$U_1^*R - V_1^*S = \Sigma_1(V_1^*\widetilde{V}_1 - U_1^*\widetilde{U}_1) + (V_1^*\widetilde{V}_1 - U_1^*\widetilde{U}_1)\widetilde{\Sigma}_1.$$

Use a special case of a theorem of Davis–Kahan [1] (see also [4, Lemma 1]) to get

$$\|V_1^*\widetilde{V}_1 - U_1^*\widetilde{U}_1\|_F \le \frac{\|U_1^*R - V_1^*S\|_F}{\sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1)}. \tag{3.7}$$

The next bound follows from combining (3.5)–(3.7) and Wedin's bound (1.6):

$$\left\| \sin\Theta(\mathcal{R}(X_1), \mathcal{R}(\widetilde{X}_1)) \right\|_F^2 \leq \frac{1}{2} \frac{\|R\|_F^2 + \|S\|_F^2}{\delta^2} + \frac{1}{4} \frac{\|U_1^* R - V_1^* S\|_F^2}{(\sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1))^2}.$$

This can be used in (3.1) and (3.2) to get a bound for the left hand side of (2.2). The bound so obtained is different and more complicated than that of (2.2). For instance, it is not sharp for Example 1.1 and can be easily shown to be weaker than (2.2) if $\delta = \delta_b$.

PROOF of Theorem 2.2. The same notation as in the proof of Theorem 2.1 will be used.

Consider again the Jordan–Wielandt matrices C and $\widetilde{C}$, and notice that

$$\begin{pmatrix} 0 & \widetilde{A} \\ \widetilde{A}^* & 0 \end{pmatrix} = \begin{pmatrix} D_1^* & 0 \\ 0 & D_2^* \end{pmatrix} \begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}.$$

Theorem 3.1 of [6] applied to C, $\widetilde{C}$ and the subspaces $\mathcal{R}(X_1)$, $\mathcal{R}(\widetilde{X}_1)$, combined with (3.1) and (3.2), yields Theorem 2.2 after some direct manipulations. The only point to stress is that when applying Theorem 3.1 of [6] we obtain

$$(3.8)\ \rho_2 = \min\left\{ \min_{\mu\in\sigma(\Sigma_1),\widetilde{\mu}\in\sigma_{\text{ext}}(\widetilde{\Sigma}_2)} \frac{|\mu - \widetilde{\mu}|}{\sqrt{\mu^2 + \widetilde{\mu}^2}},\ \min_{\mu\in\sigma(\Sigma_1),\widetilde{\mu}\in\sigma(\widetilde{\Sigma}_1)} \frac{|\mu + \widetilde{\mu}|}{\sqrt{\mu^2 + \widetilde{\mu}^2}} \right\}$$

instead of (2.3). Both expressions are equal if $\mu + \widetilde{\mu} > 0$, because

$$\min_{\mu\in\sigma(\Sigma_1),\widetilde{\mu}\in\sigma(\widetilde{\Sigma}_1)} \frac{\mu + \widetilde{\mu}}{\sqrt{\mu^2 + \widetilde{\mu}^2}} \geq 1$$

whereas the first argument of (3.8) is easily proved to be smaller than or equal to 1 [5]. □

Acknowledgments.

I thank Julio Moro for his careful reading of this note and for his constructive comments. I also thank Juan M. Molera for helpful discussions. I am indebted to an anonymous referee for the arguments about the necessity of the term $\sigma_{\min}(\widetilde{\Sigma}_1) + \sigma_{\min}(\Sigma_1)$ in Section 3 and for bringing to my attention reference [4].

REFERENCES

1. C. Davis and W. Kahan, *The rotation of eigenvectors by a perturbation. III*, SIAM J. Numer. Anal., 7 (1970), pp. 1–46.
2. J. Demmel, M. Gu, S. Eisenstat, I. Slapničar, K. Veselić, and Z. Drmač, *Computing the singular value decomposition with high relative accuracy*, Tech. Report CSD-97-934, Computer Science Division, University of California, Berkeley, February 1997; LAPACK Working Note 119.
3. G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd. ed., The Johns Hopkins University Press, Baltimore, MD, 1996.

4. R.-C. Li, *New perturbation bounds for the unitary polar factor*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 327–332.
5. R.-C. Li, *Relative perturbation theory: (I) Eigenvalue and singular value variations*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 956–982.
6. R.-C. Li, *Relative perturbation theory: (II) Eigenspace and singular subspace variations*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 471–492.
7. C. C. Paige, *A note on a result of Sun Ji-Guang: Sensitivity of the CS and GSV decompositions*, SIAM J. Numer. Anal., 21 (1984), pp. 186–191.
8. G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, CA, 1990.
9. P.-Å. Wedin, *Perturbation bounds in connection with singular value decomposition*, BIT, 12 (1972), pp. 99–111.
10. P.-Å. Wedin, *On angles between subspaces.*, in Matrix Pencils, B. Kågström and A. Ruhe, eds., Springer-Verlag, New York, 1983, pp. 263–285.