SIAM J. MATRIX ANAL. APPL.
Vol. 0, No. 0, pp. 0000–0000

PROPERTIES OF A UNITARY MATRIX OBTAINED FROM A SEQUENCE OF NORMALIZED VECTORS*

CHRISTOPHER C. PAIGE† AND WOLFGANG WÜLLING‡

Abstract. In [*SIAM J. Matrix Anal. Appl.*, 31 (2009), pp. 565–583] it was shown how a special $(n+k) \times (n+k)$ unitary matrix can be defined from any sequence of k vectors in $\mathbb{C}^n$ having unit Euclidean norms. This unitary matrix can be called an augmented orthogonal matrix when applied in the analysis of any algorithm that seeks to compute k orthonormal n-vectors, but where the computed, then theoretically normalized, vectors v_j in $V_k = [v_1, \dots, v_k]$ have a significant loss of orthogonality. These unitary matrices can occur in other situations, being in fact products of k particular Householder matrices (unitary elementary Hermitians), and they have many interesting theoretical properties. Several new results concerning them have been collected here so that they can be easily referenced, our main purpose being to facilitate the rounding error analyses of iterative orthogonalization algorithms which lose significant orthogonality, such as the Lanczos process and its many related procedures. A key component of the analysis is the $k \times k$ strictly upper triangular matrix S_k arising from V_k. The singular value decomposition of S_k reveals the CS decomposition of the $(n+k) \times (n+k)$ unitary matrix, the null space of V_k, and properties of the orthogonality and loss of orthogonality resulting from its columns. Among other things these properties are used to analyze the passage towards a complete set of orthonormal vectors in $\mathbb{C}^n$ and the contribution to orthogonality of any subsequent unit norm vector v_{k+1}.

Key words. orthogonality, augmented orthogonal matrix, singular value decomposition, CS decomposition, rank deficiency, products of projectors

AMS subject classifications. 65F25, 65G50, 15A18, 15A23, 15A57

DOI. 10.1137/120897687

1. Introduction.

The so-called Modified Gram–Schmidt (MGS) method for computing a sequence of orthonormal vectors $v_1, v_2, \dots$ in $\mathbb{C}^n$ from a given sequence in $\mathbb{C}^n$ is well known for its significant loss of orthogonality if used in finite precision computation. Following a comment by Sheffield [17], Björck and Paige [1] proposed a class of matrices for analyzing the MGS method which showed just how this orthogonality was lost, and how it could be regained [1, 2]. In [9] Paige pointed out that the unitary variant (obtained by restriction to unit length vectors v_j) of the more general matrix in [1, Theorem 4.1] could be applied to *any* sequence of unit length vectors v_j, opening up the possibility of more complete analyses of any such orthogonalizing algorithms. The approach was applied in [10] to give a new stability result for the symmetric matrix tridiagonalization process proposed by Lanczos in [8]. A more complete history of the development of these ideas was given in [10, section 2.2].

The above unitary variant can be called an augmented orthogonal matrix when applied in the analysis of any algorithm that seeks to produce k orthonormal n-vectors, but which significantly loses orthogonality. Theorem 3.1 here summarizes the relevant part of [1, Theorem 4.1], while Theorem 2.1 here summarizes the relevant part of [9, Theorem 2.1], which extends the original results in the directions required here. These augmented orthogonal matrices have several interesting properties, and

*Received by the editors November 5, 2012; accepted for publication (in revised form) by J. Liesen January 17, 2014; published electronically DATE.
http://www.siam.org/journals/simax/x-x/89768.html
†School of Computer Science, McGill University, Montreal, Quebec, H3A 0E9, Canada (paige@cs.mcgill.ca). This author's work was supported by NSERC of Canada grant OGP0009236.
‡W2 Actuarial & Math Services Ltd., Blackpool, Lancashire, FY4 5PN, UK (wolfgang.wuelling@web.de).

we collect the new ones we have encountered here so that they can be easily referenced in the analysis of any relevant algorithm. Such results are needed, for example, to complete the analysis of the symmetric Lanczos process started in [10], while the combined approach of the style of analysis in [10] with the results here would lead to a different analysis of MGS-GMRES (the MGS variant of GMRES by Saad and Schultz [14]) than the analysis given by Paige, Rozložník, and Strakoš in [12].

The MGS and MGS-GMRES methods orthogonalize each new vector against *all* the previous supposedly orthogonal vectors, and so are reasonably well behaved. As a result MGS applied to $m \times n$ linear least squares problems and MGS-GMRES applied to $n \times n$ linear systems of equations have been shown to be n-step methods in finite precision [1, 12]. They produce vectors which can quickly lose orthogonality, but the matrix having these vectors as columns can only become rank deficient at a satisfactory numerical solution to these problems. In contrast the symmetric Lanczos process [8], the orthogonal bidiagonalization algorithm of Golub and Kahan [5], and the orthogonal tridiagonalization algorithm of Saunders, Simon, and Yip (see [15, 6, 13]) only orthogonalize against recent vectors, the remaining orthogonalizations being implicit. The supposedly orthonormal matrices can rapidly become numerically rank deficient, and the resulting analyses are much more difficult, encouraging this study of the theoretical properties of the concomitant augmented orthogonal matrices.

These augmented orthogonal matrices are in fact products of Householder matrices (unitary elementary Hermitians) and this work will add to previous results such as those given by Schreiber and Van Loan in [16]. The style of approach in [10] has also been applied by Paige, Panayotov, and Zemke in [11] to the unsymmetric matrix tridiagonalization process proposed by Lanczos in [8]. After k steps applied to an $n \times n$ unsymmetric matrix A this produces two $n \times k$ matrices V_k and $\widehat{V}_k$ which ideally satisfy $\widehat{V}_k^H V_k = I$, but which in practice rapidly lose biorthogonality. The understanding gained here could also ease the completion of that analysis.

We have only applied the results of this paper to general unit Euclidean norm vectors, and to the symmetric Lanczos process [8], and so our comments on numerical behavior will largely be limited to these cases. However, we are confident that many of these comments will apply more widely, especially to the methods mentioned above.

1.1. Notation and terminology. We will use "$\triangleq$" for "is defined to be", and "$\equiv$" for "is equivalent to". Let O_n and I_n denote the $n \times n$ zero and unit matrices, respectively, where e_j will be the jth column of a unit matrix I. We will say $Q_1 \in \mathbb{C}^{n\times k}$ has orthonormal columns if $Q_1^H Q_1 = I$ and write $Q_1 \in \mathcal{U}^{n\times k}$. For singular values we will write $\sigma(\cdot)$, ordered $\sigma_1(B) \geq \sigma_2(B) \geq \cdots \geq 0$. We will denote the Frobenius norm by $\|B\|_F \triangleq \sqrt{\mathrm{trace}(B^H B)}$, the Euclidean norm by $\|v\|_2 \triangleq \sqrt{v^H v}$, and the spectral norm by $\|B\|_2 \triangleq \sigma_{\max}(B)$. We will write $V_{i,j} \triangleq [v_i, v_{i+1}, \ldots, v_j]$ and $V_k \triangleq V_{1,k}$. Here $\Re\{\cdot\}$ means "the real part of", $\mathcal{N}(B)$ denotes the null space of B, and $B^\dagger$ denotes the pseudoinverse of B. For simplicity $p_j \in P_i$ will mean that p_j is a column of the matrix P_i.

We will often index matrices by dimensional subscripts as in V_k when the $(k+1)$st matrix can be obtained from the kth by adding a column, or a column and a row. This holds for $V_k \in \mathbb{C}^{n\times k}$ and $S_k \in \mathbb{C}^{k\times k}$. Otherwise, we will use superscripts, as in $Q^{(k)}$, and then subscripts will denote partitioning, as in $Q^{(k)} \equiv [Q_1^{(k)} | \, Q_2^{(k)}]$. We will often omit the superscript $\cdot^{(k)}$ (but not others, e.g., $\cdot^{(k+1)}$) when the meaning is clear.

We use "sut$(\cdot)$" to give the matrix in parentheses with its lower triangle set to zero; thus $\mathrm{sut}(\alpha) = 0$ for a scalar α. To paraphrase the later result (2.3) where an important matrix S_k will be defined, we will say there is no loss of orthogonality when $\|S_k\|_2 = 0$, some loss of orthogonality when $0 < \|S_k\|_2 < 1$, and complete loss of orthogonality when $\|S_k\|_2 = 1$. See also Remark 5.1 for the definitions of "zero triplets" and "unit triplets", and (6.10) for η_k^2.

2. Obtaining a unitary matrix from unit Euclidean norm n-vectors. The next theorem was given in full in [9]. It allows us to develop a theoretical $(n+k)\times(n+k)$ unitary matrix $Q^{(k)}$ from any $n \times k$ matrix V_k with unit Euclidean norm columns.

THEOREM 2.1 (see [9, Theorem 2.1]). *For integers $n \geq 1$ and $k \geq 1$, and $V_j \triangleq [v_1, \dots, v_j] \in \mathbb{C}^{n\times j}$ with $\|v_j\|_2 = 1$, $j = 1, \dots, k+1$, define the strictly upper triangular matrix S_k as follows:*

$$S_k \triangleq (I_k + U_k)^{-1} U_k = U_k (I_k + U_k)^{-1} \in \mathbb{C}^{k\times k}, \qquad U_k \triangleq \operatorname{sut}(V_k^H V_k) \tag{2.1}$$

(where clearly $I_k \pm S_k$ and $I_k \pm U_k$ are always nonsingular). Then

$$U_k S_k = S_k U_k, \quad U_k = (I_k - S_k)^{-1} S_k = S_k (I_k - S_k)^{-1}, \quad (I_k - S_k)^{-1} = I_k + U_k, \tag{2.2}$$

$$\|S_k\|_2 \leq 1; \quad V_k^H V_k = I \Leftrightarrow \|S_k\|_2 = 0; \quad V_k^H V_k \;\; \textit{singular} \Leftrightarrow \|S_k\|_2 = 1. \tag{2.3}$$

Most importantly, S_k is the unique *strictly upper triangular $k \times k$ matrix such that*

$$Q^{(k)} \equiv \left[\begin{array}{c|c} Q_{11}^{(k)} & Q_{12}^{(k)} \\ \hline Q_{21}^{(k)} & Q_{22}^{(k)} \end{array}\right] \triangleq \left[\begin{array}{c|c} S_k & (I_k - S_k)V_k^H \\ \hline V_k(I_k - S_k) & I_n - V_k(I_k - S_k)V_k^H \end{array}\right] \in \mathcal{U}^{(n+k)\times(n+k)}. \tag{2.4}$$

We also write $Q^{(k)} \equiv [\underbrace{Q_1^{(k)}}_{k} \mid \underbrace{Q_2^{(k)}}_{n}]$. With (2.2) we have $S_{k+1} = \begin{bmatrix} S_k & s_{k+1} \\ 0 & 0 \end{bmatrix}$, where

$$s_{k+1} = (I_k - S_k)V_k^H v_{k+1}, \tag{2.5}$$

$$Q_1^{(k+1)} = \left[\begin{array}{c} S_{k+1} \\ \hline V_{k+1}(I_{k+1} - S_{k+1}) \end{array}\right] = \left[\begin{array}{c|c} S_k & s_{k+1} \\ 0 & 0 \\ \hline V_k(I_k - S_k) & v_{k+1} - V_k s_{k+1} \end{array}\right]. \tag{2.6}$$

One consequence of Theorem 2.1 is that while $\|S_k\|_2 < 1$, $\|S_k\|_2$ gives an indication of the overall loss of orthogonality in the (unit length) columns of V_k; see, e.g., [12, Lemma 5.1] and [9, Corollary 5.2].

3. Some properties of the unitary matrix in Theorem 2.1. In [9] the construction in Theorem 2.1 was called a *unitary or orthonormal augmentation of an array or sequence of unit length vectors* (the "augmentation" from V_k to $Q_1^{(k)}$ in (2.4)), and the full $Q^{(k)}$ will be called an augmented orthogonal matrix. Here $Q^{(k)}$ is the equivalent of the more general matrix in [1, Theorem 4.1] restricted to unit length vectors v_j, giving elementary orthogonal projectors $\mathcal{P}_j$ as follows.

THEOREM 3.1 (see [1, Theorem 4.1]). *Let $V_k = [v_1, \dots, v_k] \in \mathbb{C}^{n\times k}$, and define*

$$\mathcal{P}_j \triangleq I - v_j v_j^H, \quad \tilde{p}_j \triangleq \begin{bmatrix} -e_j \\ v_j \end{bmatrix} \in \mathbb{C}^{n+k}, \quad \widetilde{P}_j \triangleq I - \tilde{p}_j \tilde{p}_j^H, \quad j = 1, \dots, k. \tag{3.1}$$

Then with S_k as in (2.1) and the partitioning we use throughout this paper

$$Q^{(k)} = \widetilde{P}_1\widetilde{P}_2\cdots\widetilde{P}_k = \begin{array}{c} \\ k \\ n \end{array}\left[\begin{array}{c|c} k & n \\ Q_{11}^{(k)} & Q_{12}^{(k)} \\ \hline Q_{21}^{(k)} & Q_{22}^{(k)} \end{array}\right] = \left[\begin{array}{c|c} S_k & (I-S_k)V_k^H \\ \hline V_k(I-S_k) & I - V_k(I-S_k)V_k^H \end{array}\right]$$

$$(3.2) \qquad = \left[\begin{array}{cccccc|c} 0 & v_1^H v_2 & v_1^H\mathcal{P}_2 v_3 & \cdots & v_1^H\mathcal{P}_2\mathcal{P}_3\cdots\mathcal{P}_{k-1}v_k & v_1^H\mathcal{P}_2\mathcal{P}_3\cdots\mathcal{P}_k \\ 0 & 0 & v_2^H v_3 & \cdots & v_2^H\mathcal{P}_3\mathcal{P}_4\cdots\mathcal{P}_{k-1}v_k & v_2^H\mathcal{P}_3\mathcal{P}_4\cdots\mathcal{P}_k \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & v_{k-1}^H v_k & v_{k-1}^H\mathcal{P}_k \\ 0 & 0 & 0 & \cdots & 0 & v_k^H \\ \hline v_1 & \mathcal{P}_1 v_2 & \mathcal{P}_1\mathcal{P}_2 v_3 & \cdots & \mathcal{P}_1\mathcal{P}_2\cdots\mathcal{P}_{k-1}v_k & \mathcal{P}_1\mathcal{P}_2\cdots\mathcal{P}_k \end{array}\right].$$

$Q^{(k)}$ is unitary if and only if $\|v_j\|_2 = 1$ for $j = 1, \ldots, k$ (in which case $Q^{(k)}$ is the matrix in (2.4)); and $S_k = 0$ if and only if $V_k^H V_k$ is diagonal.

Proof. Because it is useful here, we give a quick induction proof of how the form $I - V_k(I-S_k)V_k^H$ of $Q_{22}^{(k)}$ in (3.2) is equivalent to $\mathcal{P}_1\mathcal{P}_2\cdots\mathcal{P}_k$. Since $S_1 = [0]$ we have $Q_{22}^{(1)} = I - v_1 v_1^H = \mathcal{P}_1$. Suppose there is equivalence up to k, then using (2.4)–(2.6),

$$\begin{aligned} Q_{22}^{(k+1)} &= I - \begin{bmatrix} V_k & v_{k+1} \end{bmatrix} \begin{bmatrix} I-S_k & -s_{k+1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_k^H \\ v_{k+1}^H \end{bmatrix} \\ (3.3) \qquad &= I - V_k(I-S_k)V_k^H - (v_{k+1} - V_k s_{k+1})v_{k+1}^H \\ &= Q_{22}^{(k)} - [v_{k+1} - V_k(I-S_k)V_k^H v_{k+1}]v_{k+1}^H = Q_{22}^{(k)}(I - v_{k+1}v_{k+1}^H) = Q_{22}^{(k)}\mathcal{P}_{k+1}, \end{aligned}$$

completing the proof of this part. ∎

Here we only consider unit length vectors v_j, so the $\mathcal{P}_j$ will always be orthogonal projectors. This is important for $Q_{22}^{(k)} = \mathcal{P}_1\cdots\mathcal{P}_k$, but note how these projectors occur throughout (3.2). Note from (2.4)–(2.6) that $Q^{(k+1)}e_{k+1}$ is just $Q_2^{(k)}v_{k+1}$ with a zero inserted after the kth element, since from (2.4) and (2.5) $Q_{12}^{(k)}v_{k+1} = s_{k+1}$ and

$$(3.4) \qquad v_{k+1} - V_k s_{k+1} = [I - V_k(I-S_k)V_k^H]v_{k+1} = Q_{22}^{(k)}v_{k+1} = \mathcal{P}_1\cdots\mathcal{P}_k v_{k+1}.$$

With the norm of the last column in (2.6) being unity, this also shows that

$$(3.5) \qquad 1 = \|s_{k+1}\|_2^2 + \|v_{k+1} - V_k s_{k+1}\|_2^2 = \|s_{k+1}\|_2^2 + \|Q_{22}^{(k)}v_{k+1}\|_2^2.$$

We note that the $\widetilde{P}_j$ in (3.1) are Householder matrices (unitary elementary Hermitians; see, e.g., [7, section 5.1]) when the v_j are unit length vectors. This implies that $Q^{(k)}$ is a product of k Householder matrices of a particular form.

The expression for $Q^{(k)}$ in (2.4) is a special case of the matrix in [1, Theorem 4.1], but the relationships of S_k to U_k were not realized in [1]. The relationships (2.1) and $(I+U_k)(I-S_k) = I$ in (2.2) were first published by Giraud, Gratton, and Langou (see [4, (3.1)]), where [4] extended ideas from [1]. It was also pointed out in [4, section 3.1] that S_k is closely related to the T-factor of the YTY-representation derived by Schreiber and Van Loan in [16] for products of Householder matrices. For $Q^{(k)}$ in (2.4),

$$Y \triangleq \begin{bmatrix} -I_k \\ V_k \end{bmatrix}, \quad T \triangleq S_k - I_k, \quad I + YTY^H = I_{k+n} - \begin{bmatrix} -I_k \\ V_k \end{bmatrix}(I_k - S_k)\begin{bmatrix} -I_k & V_k^H \end{bmatrix} \equiv Q^{(k)}.$$

4. The singular value decomposition (SVD) of S_k. When we first encountered S_k in [1] we did not realize that it was meaningful for *any* sequence of unit length vectors. We did realize that there was a lot of structure in its elements (see (3.2)), but had only a very limited appreciation of what properties this led to. We now document several properties of the SVD of S_k when it arises from a general matrix V_k with unit length columns. This SVD is used throughout this paper, and turns out to be important for further work we intend to publish later.

Let $\check{S}_k$ be S_k less its zero first column and last row, with SVD $\check{S}_k = \check{W}^{(k)}\check{\Sigma}^{(k)}\check{P}^{(k)H}$. This will be used in Remark 5.1 and later, but because we apply matrices of singular vectors to the right of $n\times k$ V_k, we seek final results using the SVD of $k\times k$ S_k rather than that of $\check{S}_k$, while maintaining the relationships (with obvious notation)

$$(4.1)\quad S_k = W^{(k)}\Sigma^{(k)}P^{(k)H} \triangleq \begin{bmatrix}\check{W}^{(k)} & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\check{\Sigma}^{(k)} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0 & \check{P}^{(k)H}\\ 1 & 0\end{bmatrix} = \begin{bmatrix}0 & \check{S}_k\\ 0 & 0\end{bmatrix},$$

$$(4.2)\quad w_j^{(k)} \triangleq \begin{bmatrix}\check{w}_j^{(k)}\\ 0\end{bmatrix},\quad p_j^{(k)} \triangleq \begin{bmatrix}0\\ \check{p}_j^{(k)}\end{bmatrix},\quad j=1,\dots,k-1;\qquad w_k^{(k)} \triangleq e_k,\quad p_k^{(k)} \triangleq e_1,$$

where this $w_k^{(k)}$ and $p_k^{(k)}$ will *always* correspond to the added zero singular value in going from $\check{S}_k$ to S_k. Then we have the *derived vector* updates when $k \to k+1$:

$$(4.3)\qquad \begin{aligned} V_kW^{(k)} &= [V_{k-1}\check{W}^{(k)}, v_k] \to V_{k+1}W^{(k+1)} = [V_k\check{W}^{(k+1)}, v_{k+1}],\\ V_kP^{(k)} &= [V_{2,k}\check{P}^{(k)}, v_1] \;\;\to V_{k+1}P^{(k+1)} = [V_{2,k+1}\check{P}^{(k+1)}, v_1].\end{aligned}$$

We often omit the superscript $\cdot^{(k)}$ for readability, and write, e.g., $S_k = W\Sigma P^H$. We never omit other superscripts such as $\cdot^{(k+1)}$, so, e.g., W will always mean $W^{(k)}$, P_1 will always mean $P_1^{(k)}$, Σ_2 will always mean $\Sigma_2^{(k)}$, etc. From (2.3) we know that $\sigma_{\max}(S_k) \le 1$, and we will see that unit singular values are crucial in the analysis. Also, if $V_k^HV_k = I$, then $S_k = 0$ in (2.1), and it will help to group and label the singular vectors of S_k according to its unit and zero singular values.

Definition 4.1 (partitioned SVD of S_k). *Let the $k\times k$ matrix S_k in Theorem 2.1 have m_k unit and n_k zero singular values. Let*

$$(4.4)\qquad \begin{aligned} S_k &= W\Sigma P^H \equiv W_1P_1^H + W_2\Sigma_2P_2^H,\\ I - S_kS_k^H &= W\Gamma^2W^H \equiv W_2\Gamma_2^2W_2^H + W_3W_3^H,\end{aligned}$$

denote the SVDs of S_k and $I - S_kS_k^H$, where

$$\begin{aligned} W &\equiv [w_1,\dots,w_k] \equiv [W_1, W_2, W_3] \in \mathcal{U}^{k\times k},\ W_1 \in \mathcal{U}^{k\times m_k},\ W_3 \in \mathcal{U}^{k\times n_k},\\ P &\equiv [p_1,\dots,p_k] \equiv [P_1, P_2, P_3] \in \mathcal{U}^{k\times k},\ P_1 \in \mathcal{U}^{k\times m_k},\ P_3 \in \mathcal{U}^{k\times n_k},\\ \Sigma &\equiv \mathrm{diag}(\sigma_1,\dots,\sigma_k) \equiv \mathrm{diag}(I_{m_k}, \Sigma_2, O_{n_k}),\end{aligned}$$

$$(4.5)\quad \Gamma^2 \triangleq I_k - \Sigma^2,\quad \Gamma \equiv \mathrm{diag}(\gamma_1,\dots,\gamma_k) \equiv \mathrm{diag}(O_{m_k}, \Gamma_2, I_{n_k}),\ \Gamma_2 \text{ positive definite.}$$

The singular values σ_j, $1 \le j \le k$, of S_k are arranged as follows:

$$(4.6)\qquad 1 = \sigma_1 = \cdots = \sigma_{m_k} > \sigma_{m_k+1} \ge \cdots \ge \sigma_{k-n_k} > \sigma_{k-n_k+1} = \cdots = \sigma_k = 0.$$

These singular vectors of S_k combine with (2.4) to reveal key properties of V_k:

(4.7)
$$Q^{(k)}\begin{bmatrix}P\\0\end{bmatrix}=\left[\begin{array}{c}S_kP\\\hline V_k(I_k-S_k)P\end{array}\right]=\left[\begin{array}{c|c|c}W_1 & W_2\Sigma_2 & 0\\\hline V_k(P_1-W_1) & V_k(P_2-W_2\Sigma_2) & V_kP_3\end{array}\right]=\left[\begin{array}{c|c|c}W_1 & W_2\Sigma_2 & 0\\\hline 0 & \widetilde{V}_2\Gamma_2 & \widetilde{V}_3\end{array}\right],$$

(4.8)
$$Q^{(k)H}\begin{bmatrix}W\\0\end{bmatrix}=\left[\begin{array}{c}S_k^HW\\\hline V_k(I_k-S_k)^HW\end{array}\right]=\left[\begin{array}{c|c|c}P_1 & P_2\Sigma_2 & 0\\\hline V_k(W_1-P_1) & V_k(W_2-P_2\Sigma_2) & V_kW_3\end{array}\right]=\left[\begin{array}{c|c|c}P_1 & P_2\Sigma_2 & 0\\\hline 0 & \widehat{V}_2\Gamma_2 & \widehat{V}_3\end{array}\right],$$

where $[\widetilde{V}_2,\widetilde{V}_3]$ and $[\widehat{V}_2,\widehat{V}_3]$ are defined by (4.7) and (4.8), and defined again in the following theorem. The first equality in each of (4.7) and (4.8) follows directly from the structure of $Q^{(k)}$, and the second by applying (4.4). But the columns in each expression are orthonormal, giving the structure in the fourth expressions. Then, with (4.5), each of $[\widetilde{V}_2,\widetilde{V}_3]$, $[\widehat{V}_2,\widehat{V}_3]$ has orthonormal columns that span $\mathrm{Range}(V_k)$. This structure is very revealing, and is used to prove the following theorem.

THEOREM 4.2 (range and null space of V_k; SVD of $V_k(I_k-S_k)$ and $V_k(I_k-S_k)^H$; CSD of $Q^{(k)}$ and expression for $Q_{22}^{(k)}$). *With the notation in Theorems* 2.1 and 3.1 *and Definition* 4.1, *define* $\widetilde{V}_2\triangleq V_k(P_2-W_2\Sigma_2)\Gamma_2^{-1}$, $\widetilde{V}_3\triangleq V_kP_3$, $\widehat{V}_2\triangleq V_k(W_2-P_2\Sigma_2)\Gamma_2^{-1}$ *and* $\widehat{V}_3\triangleq V_kW_3$. *Let the columns of* $\widehat{V}_0$ *comprise an orthonormal basis of* $\mathrm{Range}(V_k)^\perp$. *Then*

$$\mathrm{Range}(V_k)=\mathrm{Range}([\widetilde{V}_2,\widetilde{V}_3])=\mathrm{Range}([\widehat{V}_2,\widehat{V}_3])\perp\mathrm{Range}(\widehat{V}_0),\ \ \mathrm{rank}(V_k)=k-m_k, \tag{4.9}$$

$$\mathcal{N}(V_k)=\mathrm{Range}(P_1-W_1),\quad P_1-W_1\in\mathbb{C}^{k\times m_k},\quad \mathrm{rank}(P_1-W_1)=m_k. \tag{4.10}$$

We define the matrices $\widetilde{V}\triangleq[\widehat{V}_0,\widetilde{V}_2,\widetilde{V}_3]\in\mathcal{U}^{n\times n}$ *and* $\widehat{V}\triangleq[\widehat{V}_0,\widehat{V}_2,\widehat{V}_3]\in\mathcal{U}^{n\times n}$.

The SVDs of $V_k(I_k-S_k)$ *and* $V_k(I_k-S_k)^H$ *in* (2.4) *are given by*

$$V_k(I_k-S_k)=\widetilde{V}_2\Gamma_2P_2^H+\widetilde{V}_3P_3^H\quad\textit{and}\quad V_k(I_k-S_k)^H=\widehat{V}_2\Gamma_2W_2^H+\widehat{V}_3W_3^H, \tag{4.11}$$

and the CS-Decomposition (CSD) [3, 18] *of* $Q^{(k)}$ *is given by*

$$\widetilde{Q}\triangleq\left[\begin{array}{c|c}W^H & 0\\\hline 0 & \widetilde{V}^H\end{array}\right]\left[\begin{array}{c|c}S_k & (I_k-S_k)V_k^H\\\hline V_k(I_k-S_k) & I_n-V_k(I_k-S_k)V_k^H\end{array}\right]\left[\begin{array}{c|c}P & 0\\\hline 0 & \widehat{V}\end{array}\right] \tag{4.12}$$
$$=\left[\begin{array}{ccc|ccc}I_{m_k} & 0 & 0 & 0 & 0 & 0\\ 0 & \Sigma_2 & 0 & 0 & \Gamma_2 & 0\\ 0 & 0 & O_{n_k} & 0 & 0 & I_{n_k}\\\hline 0 & 0 & 0 & I_{n-(k-m_k)} & 0 & 0\\ 0 & \Gamma_2 & 0 & 0 & -\Sigma_2 & 0\\ 0 & 0 & I_{n_k} & 0 & 0 & O_{n_k}\end{array}\right]\equiv\left[\begin{array}{c|c}D_{11} & D_{12}\\\hline D_{21} & D_{22}\end{array}\right],$$

where $n-(k-m_k)\geq 0$ *since from* (4.9) $k-m_k=\mathrm{rank}(V_k)\leq n$. *When* $\mathrm{rank}(V_k)=n$ *the* $I_{n-(k-m_k)}$ *term disappears. The trailing* $n\times n$ *block* $Q_{22}^{(k)}$ *can be expressed as*

$$Q_{22}^{(k)}=\mathcal{P}_1\cdots\mathcal{P}_k=\widetilde{V}D_{22}\widehat{V}^H\equiv[\widehat{V}_0,\ \widetilde{V}_2]\,\mathrm{diag}(I_{n-(k-m_k)},-\Sigma_2)[\widehat{V}_0,\ \widehat{V}_2]^H. \tag{4.13}$$

Proof. The results in (4.9) follow from (4.7), (4.8), and the nonsingularity of Γ_2 and $I-S_k$, while $\widetilde{V}$ and $\widehat{V}$ are unitary because of the definition of $\widehat{V}_0$ and the unitary columns of (4.7) and (4.8). Then (4.10) follows since $V_k(P_1 - W_1) = 0$ in (4.7), and from $S_kP_1 = W_1$ we have $\mathrm{rank}(P_1 - W_1) = \mathrm{rank}((I - S_k)P_1) = m_k$. The first SVD in (4.11) follows by multiplying $V_k(I - S_k)P$ in (4.7) on the right by P^H, the second by multiplying $V_k(I - S_k)^H W$ in (4.8) on the right by W^H. For (4.12) D_{21} and D_{12} follow from (4.11), D_{11} is given by $S_k = W\Sigma P^H$ in (4.4), while D_{22} follows since $\widetilde{Q}$ is unitary and $\widehat{V}_0^H[I_n - V_k(I_k - S_k)V_k^H]\widehat{V}_0 = I_{n-(k-m_k)}$, since $\widehat{V}_0^H V_k = 0$ by definition. Then (4.13) is obvious from (3.2) and (4.12). ∎

If $k = 2$, $v_2 \not\perp v_1$, and $v_2 \neq v_1$, then we see from (4.7), (4.8), and (4.3) that $\widetilde{V}_3^{(2)} = [v_1]$ while $\widehat{V}_3^{(2)} = [v_2]$, so, in general, $\mathrm{Range}(\widetilde{V}_3^{(k)}) \neq \mathrm{Range}(\widehat{V}_3^{(k)})$ in (4.9).

We will analyze the "input" vector v_{k+1} in terms of its components in the ranges of $\widehat{V}_0^{(k)}$, $\widehat{V}_2^{(k)}$, and $\widehat{V}_3^{(k)}$, and consider the "outputs", the desired orthonormal sets, as $\widetilde{V}_2^{(k+1)}$ and $\widetilde{V}_3^{(k+1)}$, or as $\widehat{V}_2^{(k+1)}$ and $\widehat{V}_3^{(k+1)}$. So both (4.7) and (4.8) are essential.

An interesting example of the structure in V_k arises from the finite precision Lanczos process [8], where computations show that $k \ll n$ steps can create a large numerical rank deficiency in V_k corresponding to many singular values of S_k being unity to essentially machine precision. Equation (4.9) shows that $k - m_k \leq n$, so that if $k > n$, then S_k has at least $k - n$ singular values which are exactly 1.

For a given v_{k+1} the vector s_{k+1} is immediately available via (2.5), but knowing s_{k+1} only provides part of v_{k+1}. The next corollary reveals the close relationships between certain orthogonal components of v_{k+1} and s_{k+1}.

Corollary 4.3. *With the notation in Theorem* 2.1*, Definition* 4.1*, and Theorem* 4.2*, using unitary matrices* $\widehat{V} = [\widehat{V}_0, \widehat{V}_2, \widehat{V}_3]$ *and* $W = [W_1, W_2, W_3]$ *at step* k*,*

$$(4.14)\quad W_1^H s_{k+1} = 0, \quad s_{k+1} = W_2\Gamma_2\widehat{V}_2^H v_{k+1} + W_3\widehat{V}_3^H v_{k+1},$$

$$(4.15)\quad v_{k+1} = \widehat{V}_0\widehat{V}_0^H v_{k+1} + \widehat{V}_2\widehat{V}_2^H v_{k+1} + \widehat{V}_3\widehat{V}_3^H v_{k+1} = \widehat{V}_0\widehat{V}_0^H v_{k+1} + \widehat{V}_2\Gamma_2^{-1}W_2^H s_{k+1} + \widehat{V}_3 W_3^H s_{k+1},$$

$$(4.16)\quad s_{k+1} \in \mathrm{Range}(W_2) \Leftrightarrow \widehat{V}_3^H v_{k+1} = 0, \quad s_{k+1} \in \mathrm{Range}(W_3) \Leftrightarrow \widehat{V}_2^H v_{k+1} = 0.$$

Proof. With the expression for s_{k+1} in (2.5) and using (4.4) and (4.8),

$$W_1^H s_{k+1} = W_1^H(I - S_k)V_k^H v_{k+1} = (W_1^H - P_1^H)V_k^H v_{k+1} = 0,$$
$$W_2^H s_{k+1} = W_2^H(I - S_k)V_k^H v_{k+1} = (W_2^H - \Sigma_2 P_2^H)V_k^H v_{k+1} = \Gamma_2\widehat{V}_2^H v_{k+1},$$
$$W_3^H s_{k+1} = W_3^H(I - S_k)V_k^H v_{k+1} = W_3^H V_k^H v_{k+1} = \widehat{V}_3^H v_{k+1},$$

giving (4.14), then (4.16) holds since Γ_2 is positive definite, and (4.15) follows. ∎

This corollary, especially (4.14), will help in section 8 to analyze the effect of v_{k+1} on orthogonality in terms of its components in the ranges of $\widehat{V}_0$, $\widehat{V}_2$, and $\widehat{V}_3$.

5. Persistence of unit and zero singular values of S_k. Because S_k is the leading principal submatrix of strictly upper triangular S_ℓ for $k \leq \ell$, its SVD has other useful properties. The next corollary proves the persistence ($m_{k+1} \geq m_k$ and $n_{k+1} \geq n_k$) of unit and zero singular values of S_k as k increases, with similar properties for their singular vectors. Here we need superscripts.

Corollary 5.1. *With the notation in Theorem* 2.1 *and Definition* 4.1*, if* $\ell > k$ *and* $\left[\begin{smallmatrix} S_k & S_{12}^{(k\ell)} \\ 0 & S_{22}^{(k\ell)} \end{smallmatrix}\right] \triangleq S_\ell$*, then* $\left[\begin{smallmatrix} W_1^{(k)} \\ 0 \end{smallmatrix}\right]$ *and* $\left[\begin{smallmatrix} P_1^{(k)} \\ 0 \end{smallmatrix}\right]$ *give left and right singular vectors of* S_ℓ*:*

$$(5.1)\quad S_\ell \begin{bmatrix} P_1^{(k)} \\ 0 \end{bmatrix} = \begin{bmatrix} W_1^{(k)} \\ 0 \end{bmatrix}, \quad \begin{bmatrix} W_1^{(k)H} & 0 \end{bmatrix} S_\ell = \begin{bmatrix} P_1^{(k)H} & 0 \end{bmatrix}, \quad W_1^{(k)H} S_{12}^{(k\ell)} = 0.$$

$\begin{bmatrix} P_3^{(k)} \\ 0 \end{bmatrix}$ *gives right singular vectors of* S_ℓ, *and if* $\begin{bmatrix} W_{13}^{(\ell)} \\ W_{23}^{(\ell)} \end{bmatrix} \triangleq W_3^{(\ell)}$, $W_{13}^{(\ell)} \in \mathbb{C}^{k\times n_\ell}$, *then*

$$S_\ell \begin{bmatrix} P_3^{(k)} \\ 0 \end{bmatrix} = 0, \quad 0 = W_3^{(\ell)H} \begin{bmatrix} S_k \\ 0 \end{bmatrix} = W_{13}^{(\ell)H} S_k, \quad \text{Range}(W_{13}^{(\ell)}) \subseteq \text{Range}(W_3^{(k)}), \tag{5.2}$$

$$e_k^T W^{(k)} = e_k^T, \qquad e_1^T P^{(k)} = e_k^T. \tag{5.3}$$

Proof. To prove (5.1), note from (4.4) that

$$S_\ell \begin{bmatrix} P_1^{(k)} \\ 0 \end{bmatrix} = \begin{bmatrix} S_k P_1^{(k)} \\ 0 \end{bmatrix} = \begin{bmatrix} W_1^{(k)} \\ 0 \end{bmatrix},$$
$$\begin{bmatrix} W_1^{(k)H} & 0 \end{bmatrix} S_\ell = \begin{bmatrix} W_1^{(k)H} S_k & W_1^{(k)H} S_{12}^{(k\ell)} \end{bmatrix} = \begin{bmatrix} P_1^{(k)H} & W_1^{(k)H} S_{12}^{(k\ell)} \end{bmatrix},$$

proving the first part of (5.1), then from (2.3) and e_j^T times this second equation

$$1 \geq \left\| e_j^T \begin{bmatrix} W_1^{(k)H} & 0 \end{bmatrix} S_\ell \right\|_2^2 = \|p_j^{(k)H}\|_2^2 + \|w_j^{(k)H} S_{12}^{(k\ell)}\|_2^2 = 1 + \|w_j^{(k)H} S_{12}^{(k\ell)}\|_2^2,$$

so that $\|w_j^{(k)H} S_{12}^{(k\ell)}\|_2^2 = 0$, $j = 1, \ldots, m_k$, completing the proof of (5.1).

The first two parts of (5.2) follow immediately from the SVD (4.4), and the last part follows since then $\text{Range}(W_{13}^{(\ell)}) \subseteq \mathcal{N}(S_k^H) = \text{Range}(W_3^{(k)})$.

Finally, (5.3) follows directly from the definitions in (4.1). □

Remark 5.1. If we call $\{1, w_j^{(k)}, p_j^{(k)}\}$, $p_j^{(k)} \in P_1^{(k)}$, the unit (SVD) triplets of S_k, and $\{0, w_j^{(k)}, p_j^{(k)}\}$, $p_j^{(k)} \in P_3^{(k)}$, the zero triplets, then we have shown that any unit (or zero) triplet of S_k implies there is a unit (or zero) triplet of all S_ℓ having S_k as leading principal submatrix. Singular vectors are not unique when there are repeated singular values, but we can always obtain an *SVD* of S_ℓ with essentially the same unit triplets as S_k, and essentially the same p_j in each zero triplet as S_k, for $\ell > k$. We show this by advancing one step from the *SVD* of $\check{S}_k$ to that of $\check{S}_{k+1}$ in (4.1). Suppose the SVD of S_k has m_k unit and n_k zero triplets, then from (4.1) and (4.4) define

$$\begin{aligned} &\begin{bmatrix} \check{W}_j^{(k)} \\ 0 \end{bmatrix} \triangleq W_j^{(k)}, \ \begin{bmatrix} 0 \\ \check{P}_j^{(k)} \end{bmatrix} \triangleq P_j^{(k)}, \ j = 1, 2; \ \begin{bmatrix} \check{W}_3^{(k)} & 0 \\ 0 & 1 \end{bmatrix} \triangleq W_3^{(k)}; \ \begin{bmatrix} 0 & 1 \\ \check{P}_3^{(k)} & 0 \end{bmatrix} \triangleq P_3^{(k)}; \\ &\begin{bmatrix} \check{s} \\ \check{\sigma} \end{bmatrix} \triangleq s_{k+1}; \ \begin{bmatrix} \check{S}_k & \check{s} \\ 0 & \check{\sigma} \end{bmatrix} \triangleq \check{S}_{k+1}; \quad W^{(k)H} \equiv \left[\begin{array}{c|c} \check{W}^{(k)H} & 0 \\ \hline 0 & 1 \end{array}\right] \quad \text{in (5.5) below;} \end{aligned} \tag{5.4}$$

so that $\check{W}^{(k)} \equiv [\check{W}_1^{(k)}, \check{W}_2^{(k)}, \check{W}_3^{(k)}]$ and $\check{P}^{(k)} \equiv [\check{P}_1^{(k)}, \check{P}_2^{(k)}, \check{P}_3^{(k)}]$. Note from (4.14) that $0 = W_1^{(k)H} s_{k+1} = \check{W}_1^{(k)H} \check{s}$, so we then have (see (4.1) and Definition 4.1)

$$\left[\begin{array}{c|c} \check{W}^{(k)H} & 0 \\ \hline 0 & 1 \end{array}\right] \left[\begin{array}{c|c} \check{S}_k & \check{s} \\ \hline 0 & \check{\sigma} \end{array}\right] \left[\begin{array}{c|c} \check{P}^{(k)} & 0 \\ \hline 0 & 1 \end{array}\right] = \left[\begin{array}{ccc|c} I_{m_k} & 0 & 0 & 0 \\ 0 & \Sigma_2^{(k)} & 0 & \check{W}_2^{(k)H}\check{s} \\ 0 & 0 & O_{n_k-1} & \check{W}_3^{(k)H}\check{s} \\ \hline 0 & 0 & 0 & \check{\sigma} \end{array}\right] \tag{5.5}$$
$$\to \left[\begin{array}{cc|cc} I_{m_k} & 0 & 0 & 0 \\ 0 & \Sigma_2^{(k)} & 0 & \check{W}_2^{(k)H}\check{s} \\ \hline 0 & 0 & O_{n_k-1} & e_1\rho^{(k)} \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{c|cc|c} I_{m_k} & 0 & 0 & 0 \\ \hline 0 & \Sigma_2^{(k)} & \check{W}_2^{(k)H}\check{s} & 0 \\ 0 & 0 & \rho^{(k)} & 0 \\ \hline 0 & 0 & 0 & O_{n_k-1} \end{array}\right].$$

Here there is a unitary transformation $\mathrm{diag}(I_{k-n_k}, H^{(k)H})$ from the left of the second expression (top right matrix) to transform it to the third, where from (5.4) and (4.14),

$$\check{W}_2^{(k)H}\check{s} = W_2^{(k)H}s_{k+1} = \Gamma_2^{(k)}\widehat{V}_2^{(k)H}v_{k+1}, \qquad \begin{bmatrix}\check{W}_3^{(k)H}\check{s}\\ \check{\sigma}\end{bmatrix} = W_3^{(k)H}s_{k+1} = \widehat{V}_3^{(k)H}v_{k+1}, \tag{5.6}$$

$$H^{(k)H}W_3^{(k)H}s_{k+1} = H^{(k)H}\widehat{V}_3^{(k)H}v_{k+1} = e_1\rho^{(k)}, \qquad \rho^{(k)} \triangleq \|\widehat{V}_3^{(k)H}v_{k+1}\|_2. \tag{5.7}$$

The fourth expression is obtained from the third by shifting the last column to become the $(k-n_k+1)$st. Then the unitary transformations required to give the SVD of the central $((k-m_k)-(n_k-1))$-square submatrix of the fourth expression (and so of $\check{S}_{k+1}$, and thus of S_{k+1} via (4.1)) do not alter $W_1^{(k)}$, $P_1^{(k)}$, or $\check{P}_3^{(k)}$, but $W_3^{(k)} \to W_3^{(k)}H^{(k)}$. We then permute so that any new unit or zero triplet will be added to the end of the previous ones in $\check{S}_k$, ensuring that earlier unit and zero triplets of $\check{S}_k$ will have more trailing zeros. We will use this structure from now on.

We have shown that each singular vector of a unit triplet of $\check{S}_k$ will simply have zeros appended to it to become a singular vector of later $\check{S}_\ell$, $\ell > k$, and similarly for the columns of $\check{P}_3^{(k)}$, with equivalent results for S_k. This extends (5.3) as follows.

COROLLARY 5.2. *With the notation in Definition* 4.1 *and the construction in Remark* 5.1*, for unit and zero triplets the singular vectors of* S_{k+1} *in* (4.1) *satisfy*

$$\begin{array}{lll}
m_{k+1} = m_k & \Rightarrow e_k^T W_1^{(k+1)} = 0, & e_{k+1}^T P_1^{(k+1)} = 0.\\
m_{k+1} = m_k+1 & \Rightarrow e_k^T W_1^{(k+1)} = (e_k^T w_{m_{k+1}}^{(k+1)})e_{m_{k+1}}^T, & e_{k+1}^T P_1^{(k+1)} = (e_{k+1}^T p_{m_{k+1}}^{(k+1)})e_{m_{k+1}}^T.\\
n_{k+1} = n_k & \Rightarrow & e_{k+1}^T P_3^{(k+1)} = 0.\\
n_{k+1} = n_k+1 & \Rightarrow e_k^T W_3^{(k+1)} = e_{n_k}^T, & e_{k+1}^T P_3^{(k+1)} = (e_{k+1}^T p_k^{(k+1)})e_{n_k}^T.
\end{array}$$

Proof. From (5.4) $e_k^T W_1^{(k+1)} = e_k^T \check{W}_1^{(k+1)}$, $e_k^T W_3^{(k+1)} = e_k^T[\check{W}_3^{(k+1)}, 0]$, $e_{k+1}^T P_1^{(k+1)} = e_k^T \check{P}_1^{(k+1)}$, $e_{k+1}^T P_3^{(k+1)} = e_k^T[\check{P}_3^{(k+1)}, 0]$, and the results follow from the transformations in (5.5) as follows. If $m_{k+1} = m_k$, then $\check{W}_1^{(k+1)} = \left[\begin{smallmatrix}\check{W}_1^{(k)}\\ 0\end{smallmatrix}\right]$ and $\check{P}_1^{(k+1)} = \left[\begin{smallmatrix}\check{P}_1^{(k)}\\ 0\end{smallmatrix}\right]$, proving line 1. Then line 2 follows from

$$m_{k+1} = m_k+1 \Rightarrow \check{W}_1^{(k+1)} = \left[\begin{array}{c|c}\check{W}_1^{(k)} & \check{w}_{m_{k+1}}^{(k+1)}\\ 0 & \end{array}\right], \quad \check{P}_1^{(k+1)} = \left[\begin{array}{c|c}\check{P}_1^{(k)} & \check{p}_{m_{k+1}}^{(k+1)}\\ 0 & \end{array}\right]. \tag{5.8}$$

If $n_{k+1} = n_k$, then $\check{P}_3^{(k+1)} = \left[\begin{smallmatrix}\check{P}_3^{(k)}\\ 0\end{smallmatrix}\right]$, proving line 3. From (5.5)–(5.6) $n_{k+1} = n_k+1$ is only possible if $W_3^{(k)H}s_{k+1} = 0$, so that $\rho^{(k)} = 0$, and line 4 follows from (4.2) and (5.9)

$$n_{k+1} = n_k+1 \Rightarrow \check{W}_3^{(k+1)} = \begin{bmatrix}\check{W}_3^{(k)} & 0\\ 0 & 1\end{bmatrix}, \quad \check{P}_3^{(k+1)} = \left[\begin{array}{c|c}\check{P}_3^{(k)} & \check{p}_k^{(k+1)}\\ 0 & \end{array}\right], \quad p_k^{(k+1)} = \begin{bmatrix}0\\ \check{p}_k^{(k+1)}\end{bmatrix}. \qquad \square$$

EXAMPLE 5.1. *It is possible that* $m_{k+1} = m_k+1$ *and* $n_{k+1} = n_k+1$ *simultaneously:*

$$V_3 = \begin{bmatrix}\frac{1}{\sqrt{2}} & 1 & 0\\ \frac{1}{\sqrt{2}} & 0 & 1\\ 0 & 0 & 0\end{bmatrix}, \quad U_3 = S_3 = \begin{bmatrix}0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}, \quad \begin{array}{ll} n_2 = 1, & n_3 = 2,\\ m_2 = 0, & m_3 = 1.\end{array}$$

It can be shown here that $v_3 = e_2 \in \mathrm{Range}(\widehat{V}_2^{(2)})$*, see item* 3 *in Table* 1.

6. Properties of $Q_{22}^{(k)}$ in (2.4). We repeat the useful (4.13) here for ease:

$$Q_{22}^{(k)} = \widetilde{V}\,\mathrm{diag}(I_{n-k+m_k}, -\Sigma_2, O_{n_k})\widehat{V}^H = \widehat{V}_0\widehat{V}_0^H - \widetilde{V}_2\Sigma_2\widehat{V}_2^H. \tag{6.1}$$

If V_k has all orthonormal columns, then from (2.1) and (2.4) $S_k = 0$ and $Q_{22}^{(k)} = I - V_kV_k^H$. In this ideal case $Q_{22}^{(k)}$ has k zero singular values and $n-k$ unit singular values, so that the squared Frobenius norm $\|Q_{22}^{(j)}\|_F^2$, $j = 1, \ldots, k$, decreases by one each step, and we will have $Q_{22}^{(n)} = 0$. In general, if $Q_{22}^{(k)} = 0$, then from (6.1) $n = k - m_k$ while $\Sigma_2^{(k)}$, $\widehat{V}_0^{(k)}$, and $\widehat{V}_2^{(k)}$ are nonexistent, so from (4.9) $\mathrm{Range}(V_k) = \mathrm{Range}(\widehat{V}_3^{(k)})$ gives a complete set of orthonormal vectors in $\widehat{V}_3^{(k)} \in \mathbb{C}^{n\times n}$. For example, with the finite precision Lanczos process, when $\|Q_{22}^{(k)}\|_F^2 \approx 0$ we can show that the Ritz approximations are very accurate. It also follows from the augmented stability result in [10] that *all* the eigenvalues of A whose eigenvectors have nonnegligible components in v_1 have been well approximated, so we need to understand how $\|Q_{22}^{(j)}\|_F^2$ decreases.

Now $[\widehat{V}_0, \widehat{V}_2, \widehat{V}_3] \in \mathcal{U}^{n\times n}$, and since $Q_{22}^{(k)}\widehat{V}_0 = \widehat{V}_0$ and $Q_{22}^{(k)}\widehat{V}_3 = 0$, both $Q_{22}^{(k)}$ and $I - Q_{22}^{(k)}$ behave like orthogonal projectors on $\mathrm{Range}([\widehat{V}_0, \widehat{V}_3])$, but not on $\mathrm{Range}(\widehat{V}_2)$. Here $\widehat{V}_0^{(k)}\widehat{V}_0^{(k)H}$ is an orthogonal projector which is sporadically decreasing in rank.

LEMMA 6.1. *For the pseudoinverse $Q_{22}^{(k)\dagger}$ of $Q_{22}^{(k)}$ in* (6.1)

$$Q_{22}^{(k)\dagger} = \widehat{V}_0\widehat{V}_0^H - \widehat{V}_2\Sigma_2^{-1}\widetilde{V}_2^H, \qquad (Q_{22}^{(k)H})^\dagger = (Q_{22}^{(k)\dagger})^H, \tag{6.2}$$

$$Q_{22}^{(k)}Q_{22}^{(k)\dagger} = Q_{22}^{(k)H\dagger}Q_{22}^{(k)H} = \widehat{V}_0\widehat{V}_0^H + \widetilde{V}_2\widetilde{V}_2^H, \quad Q_{22}^{(k)\dagger}Q_{22}^{(k)} = \widehat{V}_0\widehat{V}_0^H + \widehat{V}_2\widehat{V}_2^H. \tag{6.3}$$

Also with Definition 4.1 *and the notation in Theorem* 4.2*,*

$$\mathcal{N}(Q_{22}^{(k)}) = \mathrm{Range}(\widehat{V}_3), \qquad \mathcal{N}(Q_{22}^{(k)H}) = \mathrm{Range}(\widetilde{V}_3), \tag{6.4}$$

$$(I - Q_{22}^{(k)})\widehat{V}_2 = V_kW_2\Gamma_2, \qquad (I - Q_{22}^{(k)H})\widetilde{V}_2 = V_kP_2\Gamma_2. \tag{6.5}$$

Proof. Equations (6.2)–(6.4) follow directly from (6.1) and the properties of the pseudoinverse. Then (6.5) follows from (6.1), (4.8), and (4.7). For example,

$$(I - Q_{22}^{(k)})\widehat{V}_2\Gamma_2 = (\widehat{V}_2 + \widetilde{V}_2\Sigma_2)\Gamma_2 = V_k(W_2 - P_2\Sigma_2 + P_2\Sigma_2 - W_2\Sigma_2^2) = V_kW_2\Gamma_2^2. \qquad \square$$

Since it can be important that $Q_{22}^{(k)} \to 0$ rapidly, we now show that the singular values of $Q_{22}^{(k)}$, and so $\|Q_{22}^{(k)}\|_F$, are always nonincreasing with k. We then show in (6.10) how the decrease in $\|Q_{22}^{(k)}\|_F$ relates to $\|s_{k+1}\|_2$.

COROLLARY 6.2. *For the lagging principal $n\times n$ block $Q_{22}^{(k)}$ of $Q^{(k)}$ in Theorem* 2.1 *consider the eigenvalues $\lambda_i^{(k)} \triangleq \lambda_i(Q_{22}^{(k)}Q_{22}^{(k)H})$, so $\lambda_i^{(k)} = 1$, $i = 1, \ldots, n-(k-m_k)$, from* (4.12), (4.6)*, and $\lambda_i^{(k)} = (\sigma_{i+k-n}^{(k)})^2$, $i = n-(k-m_k)+1, \ldots, n$. Then for $k = 1, 2, 3, \ldots,$*

$$\lambda_i^{(k+1)} \le \lambda_i^{(k)}, \quad i = 1, \ldots, n, \quad \|Q_{22}^{(k+1)}\|_{2,F} \le \|Q_{22}^{(k)}\|_{2,F}, \tag{6.6}$$

$$Q_{22}^{(k)}v_{k+1} \ne 0 \Leftrightarrow \|Q_{22}^{(k+1)}\|_F < \|Q_{22}^{(k)}\|_F \Leftrightarrow \lambda_i^{(k+1)} < \lambda_i^{(k)} \text{ for at least one } i, \tag{6.7}$$

$$\{\|s_{k+1}\|_2 = 1\} \Leftrightarrow \{Q_{22}^{(k)}v_{k+1} = 0\} \Leftrightarrow \{Q_{22}^{(k+1)} = Q_{22}^{(k)}\}, \tag{6.8}$$

$$\mathrm{rank}(V_k) < n \Rightarrow \|Q_{22}^{(k)}\|_2 = 1, \quad Q_{22}^{(k)} = 0 \Rightarrow \mathrm{rank}(V_k) = n. \tag{6.9}$$

In all cases the squared Frobenius norm of $Q_{22}^{(k)}$ *decreases each step as follows:*

$$\eta_{k+1}^2 \triangleq \|Q_{22}^{(k)}\|_F^2 - \|Q_{22}^{(k+1)}\|_F^2 = \|Q_{22}^{(k)} v_{k+1}\|_2^2 = 1 - \|s_{k+1}\|_2^2 \tag{6.10}$$
$$= \|\widehat{V}_0^{(k)H} v_{k+1}\|_2^2 + \|\Sigma_2^{(k)} \widehat{V}_2^{(k)H} v_{k+1}\|_2^2,$$

$$\|Q_{22}^{(k)}\|_F^2 = n - (k - \|S_k\|_F^2) = n - \eta_1^2 - \eta_2^2 - \cdots - \eta_k^2, \quad \eta_1^2 \triangleq 1. \tag{6.11}$$

Proof. Equation (3.3) shows that

$$Q_{22}^{(k+1)} Q_{22}^{(k+1)H} = Q_{22}^{(k)} \mathcal{P}_{k+1} Q_{22}^{(k)H} = Q_{22}^{(k)} Q_{22}^{(k)H} - Q_{22}^{(k)} v_{k+1} v_{k+1}^H Q_{22}^{(k)H}, \tag{6.12}$$

which together with Weyl's inequalities (see, e.g., [19, pp. 102 and 104]), yields

$$\lambda_i(Q_{22}^{(k)} Q_{22}^{(k)H}) - \|Q_{22}^{(k)} v_{k+1}\|_2^2 \le \lambda_i(Q_{22}^{(k+1)} Q_{22}^{(k+1)H}) \le \lambda_i(Q_{22}^{(k)} Q_{22}^{(k)H}),$$

proving (6.6). Next, (6.12) leads to

$$\sum_{i=1}^{n} \lambda_i^{(k+1)} = \|Q_{22}^{(k+1)}\|_F^2 = \text{trace}(Q_{22}^{(k+1)} Q_{22}^{(k+1)H}) = \|Q_{22}^{(k)}\|_F^2 - \|Q_{22}^{(k)} v_{k+1}\|_2^2 \tag{6.13}$$
$$= \sum_{i=1}^{n} \lambda_i^{(k)} - \|Q_{22}^{(k)} v_{k+1}\|_2^2,$$

proving (6.7). The first equality in (6.10) follows from (6.13), the second from (3.5), and the third from (6.1), then (6.11) follows with (4.12). The first implication in (6.8) follows from (3.5), the second from (3.3). If $\text{rank}(V_k) < n$, there exists nonzero v such that $V_k^H v = 0$, so that $Q_{22}^{(k)} v = v$ (see (2.4)) proving the first part of (6.9). The second part also follows from the form of $Q^{(k)}$ in (2.4). ∎

7. Effectiveness of v_{k+1}. We will consider two obvious ways of measuring the effectiveness of v_{k+1} in contributing to orthogonality, and show their relationship.

Let $\theta \triangleq \theta(v_{k+1}, V_k)$ be the acute angle between v_{k+1} and $\text{Range}(V_k)$, then an indicator of the lack of orthogonality in v_{k+1} is $\cos^2\theta$. We can, therefore, consider $\sin^2\theta$ as an indicator of the effectiveness of v_{k+1} in contributing to orthogonality. From (4.9) let $P_{V_k} \triangleq [\widehat{V}_2, \widehat{V}_3][\widehat{V}_2, \widehat{V}_3]^H$ be the orthogonal projector onto $\text{Range}(V_k)$ ($[\widehat{V}_2, \widehat{V}_3] \equiv [\widehat{V}_2^{(k)}, \widehat{V}_3^{(k)}]$), then $\cos\theta = \|P_{V_k} v_{k+1}\|_2$, and from (4.15) with $v_{k+1}^H v_{k+1} = 1$,

$$\cos^2\theta = v_{k+1}^H P_{V_k} v_{k+1} = v_{k+1}^H \widehat{V}_2 \widehat{V}_2^H v_{k+1} + v_{k+1}^H \widehat{V}_3 \widehat{V}_3^H v_{k+1},$$
$$\sin^2\theta = 1 - \cos^2\theta = v_{k+1}^H \widehat{V}_0 \widehat{V}_0^H v_{k+1} = \|\widehat{V}_0^H v_{k+1}\|_2^2, \quad \theta \triangleq \theta(v_{k+1}, V_k). \tag{7.1}$$

However, we saw in section 6 that it can be meaningful for $\|Q_{22}\|_F^2$ to decrease as fast as possible, and then η_{k+1}^2 in (6.10) is the desired measure of effectiveness, where

$$\eta_{k+1}^2 = \|Q_{22} v_{k+1}\|_2^2 = 1 - \|s_{k+1}\|_2^2 = \sin^2\theta(v_{k+1}, V_k) + \|\Sigma_2 \widehat{V}_2^H v_{k+1}\|_2^2. \tag{7.2}$$

Thus η_{k+1}^2 is an upper bound on the orthogonality indicator $\sin^2\theta$, while $\|s_{k+1}\|_2^2$ is a lower bound on the loss of orthogonality indicator $\cos^2\theta$. Here we choose η_{k+1}^2 as our indicator of effectiveness of v_{k+1}, first because it measures the decrease in $\|Q_{22}\|_F^2$, and second because, as we will show in the next section, the component of v_{k+1} in $\text{Range}(\widehat{V}_2)$ contributes to the improved orthonormality in $[\widetilde{V}_2 \Gamma_2, \widetilde{V}_3]$; see (4.7).

8. Orthogonality obtained from v_{k+1}. Ideally $v_{k+1} \perp \text{Range}(V_k)$, but when it is not, we need to examine the effect of the different possible unit norm v_{k+1}. We see from (6.10) how $\|Q_{22}\|_F^2$ is reduced by the components of v_{k+1}. We have also seen the effect on s_{k+1} in Corollary 4.3. We now show the possible resulting m_{k+1} and n_{k+1} in (4.6), before looking at the contributions of v_{k+1} to the orthonormal matrices $\widehat{V}_2^{(k+1)}$, $\widehat{V}_3^{(k+1)}$, $\widetilde{V}_2^{(k+1)}$, and $\widetilde{V}_3^{(k+1)}$.

From the definitions of m_k and n_k in Definition 4.1, (4.9) shows that the rank deficiency $m_k = k - \text{rank}([\widehat{V}_2, \widehat{V}_3])$ can increase only if $v_{k+1} \in \text{Range}([\widehat{V}_2, \widehat{V}_3])$, so

$$\widehat{V}_0^H v_{k+1} \neq 0 \Rightarrow m_{k+1} = m_k, \qquad v_{k+1} \perp \text{Range}(\widehat{V}_0) \Rightarrow m_{k+1} = m_k+1, \tag{8.1}$$

while from (5.5)–(5.7) $n_{k+1} = n_k+1 \Leftrightarrow \rho^{(k)} \triangleq \|\widehat{V}_3^H v_{k+1}\|_2 = 0$, so

$$v_{k+1} \perp \text{Range}(\widehat{V}_3) \Leftrightarrow n_{k+1} = n_k+1. \tag{8.2}$$

These implications lead to Table 1, which gives the values of m_{k+1} and n_{k+1} resulting from the possible nonzero projections of v_{k+1} on the ranges of $\widehat{V}_0$, $\widehat{V}_2$, and $\widehat{V}_3$, where $[\widehat{V}_0, \widehat{V}_2, \widehat{V}_3] \equiv [\widehat{V}_0^{(k)}, \widehat{V}_2^{(k)}, \widehat{V}_3^{(k)}]$ is unitary; see (4.9). For example, item 1 is equivalent to $v_{k+1} \in \text{Range}(\widehat{V}_0) \Rightarrow \{m_{k+1} = m_k \ \& \ n_{k+1} = n_k+1\}$ in (8.20). Table 1 can also be derived from observations on (5.5)–(5.7) with (4.1).

Table 1
m_{k+1} and n_{k+1} resulting from possible components of v_{k+1}.

Item	v_{k+1} has nonzero components only in:	Implication	$m_{k+1} =$	$n_{k+1} =$
1	$\text{Range}(\widehat{V}_0)$	$\Rightarrow$	m_k	$n_k + 1$
2	$\text{Range}(\widehat{V}_0)$ & $\text{Range}(\widehat{V}_2)$	$\Rightarrow$	m_k	$n_k + 1$
3	$\text{Range}(\widehat{V}_2)$	$\Leftrightarrow$	$m_k + 1$	$n_k + 1$
4	$\text{Range}(\widehat{V}_3)$	$\Rightarrow$	$m_k + 1$	n_k
5	$\text{Range}(\widehat{V}_2)$ & $\text{Range}(\widehat{V}_3)$	$\Rightarrow$	$m_k + 1$	n_k
6	$\text{Range}(\widehat{V}_0)$ & $\text{Range}(\widehat{V}_3)$	$\Rightarrow$	m_k	n_k
7	$\text{Range}(\widehat{V}_0)$ & $\text{Range}(\widehat{V}_2)$ & $\text{Range}(\widehat{V}_3)$	$\Rightarrow$	m_k	n_k

We will use this table to divide the analysis into four cases: items 1, 2, and 3, items 3, 4, and 5, item 6, and finally item 7. Item 3 appears twice, once for $n_{k+1} = n_k+1$ and once for $m_{k+1} = m_k + 1$, while for items 6 and 7 only the dimensions of $\Sigma_2^{(k)}$ increase. Where possible, the following case descriptions will use the central block in the last matrix of (5.5) to find the contributions of v_{k+1} to the orthonormal matrices $\widehat{V}_2^{(k+1)}$, $\widehat{V}_3^{(k+1)}$, $\widetilde{V}_2^{(k+1)}$, and $\widetilde{V}_3^{(k+1)}$. All the results here assume (4.1) and the construction in Remark 5.1, and we will take the decrease η_{k+1}^2 in $\|Q_{22}^{(k)}\|_F^2$ as a simple measure of the effectiveness of v_{k+1} (see (6.10)), where ideally $\eta_{k+1}^2 = 1$.

8.1. Case 1. Items 1, 2, and 3: $v_{k+1} \perp \text{Range}(\widehat{V}_3)$, so $n_{k+1} = n_k+1$; see (8.2).

Here S_{k+1} has one more zero singular value than S_k, and the next theorem derives the resulting new vectors leading to orthonormal $\widehat{V}_3^{(k+1)}$ and $\widetilde{V}_3^{(k+1)}$.

Theorem 8.1. *With Definition* 4.1 *and the notation in Theorem* 4.2*, if*

$$v_{k+1} = \widehat{V}_0 \widehat{V}_0^H v_{k+1} + \widehat{V}_2 \widehat{V}_2^H v_{k+1}, \tag{8.3}$$

then $\widehat{V}_3$ and $\widetilde{V}_3$ are each updated with one new column as follows:

$$\widehat{V}_3^{(k+1)} = V_{k+1} W_3^{(k+1)} = [V_k \check{W}_3^{(k+1)}, v_{k+1}] = [V_{k-1} \check{W}_3^{(k)}, v_k, v_{k+1}] = [\widehat{V}_3^{(k)}, v_{k+1}], \tag{8.4}$$

$$\widetilde{V}_3^{(k+1)} = [V_{2,k+1} \check{P}_3^{(k+1)}, v_1] = [V_{2,k} \check{P}_3^{(k)}, V_{k+1} p_k^{(k+1)}, v_1], \quad \widetilde{V}_3^{(k)} = [V_{2,k} \check{P}_3^{(k)}, v_1], \tag{8.5}$$

where

$$(8.6)\qquad V_{k+1}p_k^{(k+1)} = (\widehat{V}_0\widehat{V}_0^H - \widetilde{V}_2\Sigma_2^{-1}\widehat{V}_2^H)v_{k+1}\|Q_{22}^{(k)H\dagger}v_{k+1}\|_2^{-1} \perp \mathrm{Range}(\widetilde{V}_3^{(k)})$$

is the desired new vector inserted into $\widetilde{V}_3$ *to produce orthonormal* $\widetilde{V}_3^{(k+1)}$*, increasing the dimension of the null space of* D_{22} *and* Q_{22} *in* (4.12)*.*

Proof. From (8.3) and (5.7) we see that $\rho^{(k)} = 0$ in (5.5), so there is a new zero singular value and $\widehat{V}_3^{(k+1)}$ has one more column than $\widehat{V}_3$. With (4.8), (4.3) shows that $\widehat{V}_3^{(k+1)} = V_{k+1}W_3^{(k+1)} = [V_k\check{W}_3^{(k+1)}, v_{k+1}]$, and since $n_{k+1} = n_k + 1$, (5.9) proves (8.4), so that $\widehat{V}_3$ gives the first n_k columns of $\widehat{V}_3^{(k+1)}$. Similarly (4.7), (4.3), and (5.9) prove (8.5), where $\widetilde{V}_3 = V_kP_3$ gives n_k columns of $\widetilde{V}_3^{(k+1)}$. To find $V_{2,k+1}\check{p}_k^{(k+1)} = V_{k+1}p_k^{(k+1)} = \widetilde{V}_3^{(k+1)}e_{n_k}$ in (8.5) and (8.6), note from (6.4) and (3.3) that

$$0 = Q_{22}^{(k+1)H}\widetilde{V}_3^{(k+1)} = (I - v_{k+1}v_{k+1}^H)Q_{22}^{(k)H}\widetilde{V}_3^{(k+1)},$$

meaning $Q_{22}^{(k)H}V_{k+1}p_k^{(k+1)}$ is proportional to v_{k+1}. But $\widetilde{V}_3^{(k+1)}$ has orthonormal columns, so that $\widetilde{V}_3^{(k)H}V_{k+1}p_k^{(k+1)} = 0$, giving $V_{k+1}p_k^{(k+1)} \in \mathrm{Range}([\widehat{V}_0, \widetilde{V}_2])$; see Theorem 4.2 and $\|\widetilde{V}_3^{(k+1)}e_{n_k}\|_2 = 1$. Thus with (6.3) and (6.2) we obtain (8.6) via

$$\begin{aligned} Q_{22}^{(k)H\dagger}Q_{22}^{(k)H}V_{k+1}p_k^{(k+1)} &= V_{k+1}p_k^{(k+1)} = Q_{22}^{(k)H\dagger}v_{k+1}\|Q_{22}^{(k)H\dagger}v_{k+1}\|_2^{-1} \\ &= (\widehat{V}_0\widehat{V}_0^H - \widetilde{V}_2\Sigma_2^{-1}\widehat{V}_2^H)v_{k+1}\|Q_{22}^{(k)H\dagger}v_{k+1}\|_2^{-1} \perp \mathrm{Range}(\widetilde{V}_3). \qquad \square \end{aligned}$$

It might seem strange that (8.4) implies $v_{k+1} \perp v_k$, but from (4.3) $\widehat{V}_3e_{n_k} = v_k$, and the criterion for case 1 is $v_{k+1} \perp \mathrm{Range}(\widehat{V}_3)$, so it is necessarily true for case 1.

In (8.3) v_{k+1} with the component $\widehat{V}_2\widehat{V}_2^Hv_{k+1} \in \mathrm{Range}(\widehat{V}_2)$ is inserted in $\widehat{V}_3$ to give $\widehat{V}_3^{(k+1)}$ (see (8.4)), while the vector $V_{k+1}p_k^{(k+1)}$ with a relatively larger component of $\widetilde{V}_2\Sigma_2^{-1}\widehat{V}_2^Hv_{k+1} \in \mathrm{Range}(\widetilde{V}_2)$ is inserted in $\widetilde{V}_3$ to give $\widetilde{V}_3^{(k+1)}$ (see (8.5)–(8.6)), where $[\widetilde{V}_2, \widetilde{V}_3]$ is an already developing orthonormal set. This suggests that $[\widetilde{V}_2, \widetilde{V}_3]$ might be the more useful set of "output" vectors in applications of this theory than $[\widehat{V}_2, \widehat{V}_3]$, which is so useful for examining the components of the "input" vector v_{k+1}.

Note that in case 1, both $\widehat{V}_3$ and $\widetilde{V}_3$ just have one column added, while $\widehat{V}_2$ and $\widetilde{V}_2$ are altered without an increase in their number of columns (this number will decrease if $v_{k+1} \in \mathrm{Range}(\widehat{V}_2)$ as in item 3 of Table 1, see section 8.2 below).

8.2. Case 2. Items 3, 4, and 5: $v_{k+1} \perp \mathrm{Range}(\widehat{V}_0)$, so $m_{k+1} = m_k + 1$; see (8.1).

Since S_{k+1} has one more unit singular value than S_k, we derive the corresponding right and left unit singular vectors and the resulting (redundant) derived vectors.

THEOREM 8.2. *With Definition* 4.1*, the notation in Theorem* 4.2*,* $f \triangleq \widehat{V}_2^Hv_{k+1}$*,* $\rho \triangleq \|\widehat{V}_3^Hv_{k+1}\|_2$*,* $\mu \triangleq \sqrt{\|\Gamma_2^{-1}f\|_2^2 + \rho^2}$*, and* $B \triangleq \left[\begin{smallmatrix}\Sigma_2 & \Gamma_2 f\\ 0 & \rho\end{smallmatrix}\right]$*, if* $\widehat{V}_0^Hv_{k+1} = 0$*, then*

$$(8.7)\quad Bu = z, \quad B^Hz = u, \quad u \triangleq \begin{bmatrix}\Gamma_2^{-1}\Sigma_2 f\\ 1\end{bmatrix}\mu^{-1}, \quad z \triangleq \begin{bmatrix}\Gamma_2^{-1}f\\ \rho\end{bmatrix}\mu^{-1}, \quad u^Hu = z^Hz = 1.$$

There is a new unit triplet $\{1, \check{w}^{(k+1)}_{m_{k+1}}, \check{p}^{(k+1)}_{m_{k+1}}\}$ *of* $\check{S}_{k+1}$ *(and* S_{k+1} *see* (4.2)*), and the resulting new columns* $V_{k+1}p^{(k+1)}_{m_{k+1}}$ *of* $V_{k+1}P^{(k+1)}_1$ *and* $V_{k+1}w^{(k+1)}_{m_{k+1}}$ *of* $V_{k+1}W^{(k+1)}_1$ *are*

$$\check{p}^{(k+1)}_{m_{k+1}} = \begin{bmatrix} \check{P}^{(k)}_2\Gamma_2^{-1}\Sigma_2 f \\ 1 \end{bmatrix}\mu^{-1}, \quad V_{k+1}p^{(k+1)}_{m_{k+1}} = (V_kP_2\Gamma_2^{-1}\Sigma_2\widehat{V}_2^H + I)v_{k+1}\mu^{-1}, \tag{8.8}$$

$$\check{w}^{(k+1)}_{m_{k+1}} = (W_2\Gamma_2^{-1}\widehat{V}_2^H + W_3\widehat{V}_3^H)v_{k+1}\mu^{-1}, \tag{8.9}$$

$$V_{k+1}w^{(k+1)}_{m_{k+1}} = (V_kW_2\Gamma_2^{-1}\widehat{V}_2^H + \widehat{V}_3\widehat{V}_3^H)v_{k+1}\mu^{-1}. \tag{8.10}$$

Proof. First, $\widehat{V}_0^H v_{k+1} = 0$ implies $f^H f + \rho^2 = 1$; see (4.15). Note with (5.6) that B is just the center block of the last expression in (5.5). Thus evaluating Bu, $B^H z$, $u^H u$, and $z^H z$ and using the stated properties proves (8.7), and so reveals the new unit triplet. Therefore, applying the parts of u in (8.7) appropriately to the right of the P matrix in the first expression in (5.5) gives $\check{p}^{(k+1)}_{m_{k+1}}$ in (8.8). Then $V_{k+1}p^{(k+1)}_{m_{k+1}}$ follows from (4.2). Similarly applying z^H from (8.7) appropriately to the left of the third expression in (5.5), then using $H^{(k)}e_1\rho = W_3^H s_{k+1} = \widehat{V}_3^H v_{k+1}$ from (5.7), gives

$$\mu\check{w}^{(k+1)H}_{m_{k+1}} = [0^H, f^H\Gamma_2^{-1}, \rho e_1^T]\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & H^{(k)H}\end{bmatrix}\begin{bmatrix} W_1^H \\ W_2^H \\ W_3^H \end{bmatrix} = f^H\Gamma_2^{-1}W_2^H + v_{k+1}^H\widehat{V}_3W_3^H,$$

proving (8.9). Then since $\widehat{V}_3 = V_kW_3$ in (4.8), (8.10) follows from (4.2). ∎

Note from (8.8)–(8.10) that for $V_{k+1}(W_1^{(k+1)} - P_1^{(k+1)})e_{m_{k+1}}$, as in (4.8) and (4.10),

$$\begin{aligned}(V_{k+1}w^{(k+1)}_{m_{k+1}} - V_{k+1}p^{(k+1)}_{m_{k+1}})\mu &= [V_k(W_2 - P_2\Sigma_2)\Gamma_2^{-1}\widehat{V}_2^H + \widehat{V}_3\widehat{V}_3^H - I]v_{k+1}, \\ &= (\widehat{V}_2\widehat{V}_2^H + \widehat{V}_3\widehat{V}_3^H - I)v_{k+1} = -\widehat{V}_0\widehat{V}_0^H v_{k+1} = 0.\end{aligned} \tag{8.11}$$

Since $\widehat{V}_0^H v_{k+1} = 0$ here, we want $\|\Sigma_2\widehat{V}_2^H v_{k+1}\|_2$ to be as large as possible in (6.10).

If $v_{k+1} \in \text{Range}(\widehat{V}_3)$, which is item 4 in Table 1, we see that $\mu = \rho = 1$ in Theorem 8.2, and $\eta^2_{k+1} = 0$ in (6.10), so $V_{k+1}p^{(k+1)}_{m_{k+1}} = V_{k+1}w^{(k+1)}_{m_{k+1}} = v_{k+1}$ makes no contribution to the orthogonal vectors. If $v_{k+1} \in \text{Range}(\widehat{V}_2)$, which is item 3 and is thus also in case 1, this creates a new column of $\widetilde{V}_3^{(k+1)}$ solely from Range$(\widetilde{V}_2)$ via (8.6), as well as producing the redundant vector in (8.8) and (8.10); see (8.11).

8.3. Case 3. Item 6: $v_{k+1} = \widehat{V}_0\widehat{V}_0^H v_{k+1} + \widehat{V}_3\widehat{V}_3^H v_{k+1}$, where both components are nonzero.

Here S_{k+1} has one new singular value σ added to those of $\Sigma_2^{(k)}$. For simplicity we take $\Sigma_2^{(k+1)} = \text{diag}(\Sigma_2^{(k)}, \sigma)$ before possible reordering, where $\sigma^2 + \gamma^2 = 1$ from (4.5).

Theorem 8.3. *With Definition* 4.1 *and the notation in Theorem* 4.2*, if*

$$v_{k+1} = \widehat{V}_0\widehat{V}_0^H v_{k+1} + \widehat{V}_3\widehat{V}_3^H v_{k+1}, \quad \gamma \triangleq \|\widehat{V}_0^H v_{k+1}\|_2, \quad \sigma \triangleq \|\widehat{V}_3^H v_{k+1}\|_2, \quad \gamma\sigma > 0, \tag{8.12}$$

and $\ell \triangleq k - n_k + 1$*, then* $\gamma = \gamma_\ell^{(k+1)} = \sqrt{1-\sigma^2}$*, where* $\sigma = \sigma_\ell^{(k+1)}$ *is the end new singular value in (unordered)* $\Sigma_2^{(k+1)}$ *with the singular vectors* $\check{p}_\ell^{(k+1)}$ *and* $\check{w}_\ell^{(k+1)}$ *of* $\check{S}_{k+1}$*. The resulting new columns* $V_{k+1}p_\ell^{(k+1)}$ *of* $V_{k+1}P_2^{(k+1)}$ *and* $V_{k+1}w_\ell^{(k+1)}$ *of* $V_{k+1}W_2^{(k+1)}$*,* *and* $\tilde{v}\gamma \triangleq \widetilde{V}_2^{(k+1)}\Gamma_2^{(k+1)}e_{\ell-m_k}$ *and* $\hat{v}\gamma \triangleq \widehat{V}_2^{(k+1)}\Gamma_2^{(k+1)}e_{\ell-m_k}$ *in* (4.7) *and* (4.8), *are*

$$\check{p}_\ell^{(k+1)} = e_k, \qquad V_{k+1}p_\ell^{(k+1)} = v_{k+1}, \tag{8.13}$$

$$V_{k+1}w_\ell^{(k+1)} = \widehat{V}_3\widehat{V}_3^H v_{k+1}\sigma^{-1}, \tag{8.14}$$

$$\tilde{v}\gamma = \widehat{V}_0\widehat{V}_0^H v_{k+1} \perp \mathrm{Range}([\widetilde{V}_2, \widetilde{V}_3]), \tag{8.15}$$

$$\hat{v}\gamma = (\widehat{V}_3\widehat{V}_3^H v_{k+1}\gamma^2 - \widehat{V}_0\widehat{V}_0^H v_{k+1}\sigma^2)\sigma^{-1} \perp \mathrm{Range}([\widehat{V}_2^{(k)}, \widehat{V}_3^{(k+1)}]). \tag{8.16}$$

Proof. In (5.5) $\check{W}_2^H\check{s} = 0$ (see (5.6)) and so $\sigma_\ell^{(k+1)} = \rho^{(k)} = \|W_3^H s_{k+1}\|_2 = \|\widehat{V}_3^H v_{k+1}\|_2$ (see also (5.7)), then (8.13) follows from (5.5) and (4.2), while $\gamma_\ell^{(k+1)} = \sqrt{1-(\sigma_\ell^{(k+1)})^2} = \|\widehat{V}_0^H v_{k+1}\|_2$; see (8.12). Now $[\begin{smallmatrix}\check{W}_3^{(k)H} & 0\\ 0 & 1\end{smallmatrix}] = W_3^{(k)H}$ in (5.4) is transformed from the left by $H^{(k)H}$ in (5.5) (see (5.6)–(5.7)) to give the last n_k+1 rows of $W^{(k+1)H}$, which are

$$\begin{bmatrix} w_\ell^{(k+1)H} \\ W_3^{(k+1)H}\end{bmatrix} = \begin{bmatrix} H^{(k)H}W_3^H & 0 \\ 0 & 1\end{bmatrix}, \quad \begin{bmatrix} W_3H^{(k)} & 0 \\ 0 & 1\end{bmatrix} = \left[\; w_\ell^{(k+1)} \;\middle|\; W_3^{(k+1)} \;\right]. \tag{8.17}$$

Thus with $V_kW_3 = \widehat{V}_3$ in (4.8), $H^{(k)}e_1\rho^{(k)} = \widehat{V}_3^H v_{k+1}$ in (5.7), and $\sigma = \sigma_\ell^{(k+1)} = \rho^{(k)}$,

$$V_{k+1}w_\ell^{(k+1)} = V_kW_3H^{(k)}e_1 = \widehat{V}_3\widehat{V}_3^H v_{k+1}\sigma^{-1},$$

proving (8.14). From these and (4.7), for (8.15),

$$\tilde{v}\gamma \triangleq \widetilde{V}_2^{(k+1)}\Gamma_2^{(k+1)}e_{\ell-m_k} = V_{k+1}(p_\ell^{(k+1)} - w_\ell^{(k+1)}\sigma) = v_{k+1} - \widehat{V}_3\widehat{V}_3^H v_{k+1} = \widehat{V}_0\widehat{V}_0^H v_{k+1},$$

so with (4.9), (8.15) holds. Similarly from (4.8), (8.13), and (8.14) with (8.12),

$$\begin{aligned}\hat{v}\gamma &\triangleq \widehat{V}_2^{(k+1)}\Gamma_2^{(k+1)}e_{\ell-m_k} = V_{k+1}(w_\ell^{(k+1)} - p_\ell^{(k+1)}\sigma) = \widehat{V}_3\widehat{V}_3^H v_{k+1}\sigma^{-1} - v_{k+1}\sigma \\ &= \widehat{V}_3\widehat{V}_3^H v_{k+1}\sigma^{-1} - (\widehat{V}_0\widehat{V}_0^H v_{k+1} + \widehat{V}_3\widehat{V}_3^H v_{k+1})\sigma,\end{aligned}$$

which with $1-\sigma^2 = \gamma^2$ shows $\hat{v}\gamma$ is correct in (8.16), so $\hat{v}\gamma \perp \mathrm{Range}(\widehat{V}_2)$. Multiplying the last equation in (8.17) on the left by V_{k+1} and using (4.8) gives

$$\left[V_kW_3H^{(k)}, v_{k+1}\right] = \left[\widehat{V}_3H^{(k)}, v_{k+1}\right] = \left[V_{k+1}w_\ell^{(k+1)}, \widehat{V}_3^{(k+1)}\right]. \tag{8.18}$$

We will prove that $\hat{v}^H\widehat{V}_3^{(k+1)} = 0$ in (8.16) by showing that the last n_{k+1} columns of $\hat{v}^H \times$(8.18) are zero. From (8.12)

$$\gamma\sigma\hat{v}^H v_{k+1} = \|\widehat{V}_3^H v_{k+1}\|_2^2\gamma^2 - \|\widehat{V}_0^H v_{k+1}\|_2^2\sigma^2 = \sigma^2\gamma^2 - \gamma^2\sigma^2 = 0,$$

and now we show $\sigma\hat{v}^H\widehat{V}_3H^{(k)} = \gamma\rho^{(k)}e_1^T$ to complete the proof. From (8.16) and (5.7)

$$H^{(k)H}\widehat{V}_3^H\hat{v}\sigma = H^{(k)H}\widehat{V}_3^H v_{k+1}\gamma = e_1\rho^{(k)}\gamma. \qquad \square$$

The orthogonality result in (8.16) holds from the general form in (4.8), but the proof was given here to gain an understanding of how this occurs in a specific case.

Since $\widehat{V}_0\widehat{V}_0^H v_{k+1}$ gives the total contribution to $\tilde{v}\gamma$ in (8.15), but contributes a smaller amount to $\hat{v}\gamma$ in (8.16), this again suggests that $[\widetilde{V}_2, \widetilde{V}_3]$ might be the more useful set of "output" vectors than $[\widehat{V}_2, \widehat{V}_3]$ in applications of this theory.

8.4. Case 4. Item 7: $v_{k+1} \not\perp \mathrm{Range}(\widehat{V}_j)$, $j = 0, 2, 3$.

We were able to obtain singular vectors in cases 1, 2, and 3 because we knew the new singular value 0, 1, or $\sigma = \|\widehat{V}_3^{(k)H} v_{k+1}\|_2$, but this is not possible here. This case is similar to case 3 in that the dimensions of only Σ_2 change, but different in that usually no elements of $\Sigma_2^{(k)}$ and $\Sigma_2^{(k+1)}$ are the same. Because of this difficulty we will not pursue the singular value and vector effects, apart from noting that in (6.10)

$$1 > 1 - \|\widehat{V}_3^H v_{k+1}\|_2^2 > \eta_{k+1}^2 = \|\widehat{V}_0^H v_{k+1}\|_2^2 + \|\Sigma_2 \widehat{V}_2^H v_{k+1}\|_2^2 > 0. \tag{8.19}$$

8.5. Special cases. If $v_{k+1} \perp \mathrm{Range}(\widehat{V}_2)$, then from (5.5) and (5.6) $\check{W}_2^{(k)H} \check{s} = 0$, and the SVD of S_k can easily be updated without an iterative process.

Now look at items 1 and 4 in Table 1 more closely. These give the two extreme cases $\|s_{k+1}\|_2 = 0$ (the ideal case) and 1 (no improvement; see, e.g., (6.8)).

THEOREM 8.4. *With the notation in Definition* 4.1 *and Theorem* 4.2 *we consider two extreme cases. First, we give results for the case when* $v_{k+1} \in \mathrm{Range}(\widehat{V}_0)$*:*

$$v_{k+1} \in \mathrm{Range}(\widehat{V}_0) \Leftrightarrow s_{k+1} = 0 \Leftrightarrow V_k^H v_{k+1} = 0 \Rightarrow \{m_{k+1} = m_k \ \&\ n_{k+1} = n_k + 1\}. \tag{8.20}$$

If $v_{k+1} \in \mathrm{Range}(\widehat{V}_0)$*, the new zero triplet of* S_{k+1} *is* $\{0, w_k^{(k+1)}, p_k^{(k+1)}\} = \{0, e_k, e_{k+1}\}$*,*

$$V_{k+1} W^{(k+1)} = [V_{k-1} \check{W}^{(k)}, v_k, v_{k+1}], \quad V_{k+1} P^{(k+1)} = [V_{2,k} \check{P}^{(k)}, v_{k+1}, v_1], \tag{8.21}$$

$$p_k^{(k+1)H} w_k^{(k+1)} = 0, \qquad V_{k+1} p_k^{(k+1)} = v_{k+1} \perp V_{k+1} w_k^{(k+1)} = v_k, \tag{8.22}$$

and the singular vector component $w_k^{(k+1)} p_k^{(k+1)H}$ *of* S_{k+1} *is strictly upper triangular.*

Now for the case when $v_{k+1} \in \mathrm{Range}(\widehat{V}_3)$*,*

$$\begin{aligned} v_{k+1} \in \mathrm{Range}(\widehat{V}_3) &\Leftrightarrow \|s_{k+1}\|_2 = 1 \Leftrightarrow \{s_{k+1} = W_3 \widehat{V}_3^H v_{k+1} \ \&\ \|\widehat{V}_3^H v_{k+1}\|_2 = 1\} \\ &\Leftrightarrow v_{k+1} = V_k s_{k+1} \Rightarrow \{S_k^H s_{k+1} = 0 \ \&\ m_{k+1} = m_k + 1 \ \&\ n_{k+1} = n_k\}. \end{aligned} \tag{8.23}$$

If $v_{k+1} \in \mathrm{Range}(\widehat{V}_3)$*, the new unit triplet of* S_{k+1} *is*

$$\begin{aligned} &\{1, w_{m_k+1}^{(k+1)}, p_{m_k+1}^{(k+1)}\} = \{1, \left[\begin{smallmatrix} s_{k+1} \\ 0 \end{smallmatrix}\right], e_{k+1}\}, \\ &V_{k+1} w_{m_k+1}^{(k+1)} = V_k s_{k+1} = v_{k+1} = V_{k+1} p_{m_k+1}^{(k+1)}, \qquad p_{m_k+1}^{(k+1)H} w_{m_k+1}^{(k+1)} = 0, \end{aligned} \tag{8.24}$$

and the new singular vector component $w_{m_k+1}^{(k+1)} p_{m_k+1}^{(k+1)H}$ *is strictly upper triangular.*

Proof. First, assume that $v_{k+1} \in \mathrm{Range}(\widehat{V}_0)$. The sequence (8.20) follows from (4.14), (4.9), and Table 1. Since $\left[\begin{smallmatrix} \check{s} \\ \check{\sigma} \end{smallmatrix}\right] \triangleq s_{k+1} = 0$, the first line of (5.5) clearly provides the SVD of $\check{S}_{k+1}$ which then has the new triplet $\{0, \check{w}_k^{(k+1)}, \check{p}_k^{(k+1)}\} = \{0, e_k, e_k\}$. The corresponding new triplet of S_{k+1} follows from this via (4.2), while (8.21) follows from (8.4) and (8.5). Then (8.22) follows immediately using $V_k^H v_{k+1} = 0$.

Now assume that $v_{k+1} \in \mathrm{Range}(\widehat{V}_3)$. The top line of (8.23) follows from (4.14), the next two-way implication and $S_k^H s_{k+1} = 0$ from the orthonormality of the columns in (2.6), while m_{k+1} and n_{k+1} follow from Table 1. Then the unit triplet of S_{k+1}, and so (8.24), follows from (8.8)–(8.10) by using $v_{k+1} \in \mathrm{Range}(\widehat{V}_3)$ and (8.23). ∎

EXAMPLE 8.1. *Here* $m_{k+1} = m_k + 1$ *with* $\|s_{k+1}\|_2 < 1$*. If* $v_3 = v_1$ *and* $v_2^T v_1 = 1/2$*,*

$$U_3 = \begin{bmatrix} 0 & 1/2 & 1 \\ & 0 & 1/2 \\ & & 0 \end{bmatrix}, \quad (I + U_3)^{-1} = \begin{bmatrix} 1 & -1/2 & -3/4 \\ & 1 & -1/2 \\ & & 1 \end{bmatrix} = I - S_3, \quad S_3 = \begin{bmatrix} 0 & 1/2 & 3/4 \\ & 0 & 1/2 \\ & & 0 \end{bmatrix},$$

so that S_2 *has singular values* $1/2,\ 0,$ *and* S_3 *has singular values* $1,\ 1/4,\ 0,$ *where* $\|s_{k+1}\|_2^2 = \|s_3\|_2^2 = 13/16 < 1$*. Here the unit singular vectors of* S_3 *are* $w_1^{(3)T} = \frac{1}{\sqrt{5}}[2,1,0]$ *and* $p_1^{(3)T} = \frac{1}{\sqrt{5}}[0,1,2]$*, so that in contrast to the case of* $m_{k+1} = m_k + 1$ *in Theorem* 8.4 *(see* (8.24)*)* $w_1^{(3)} p_1^{(3)T}$ *is not strictly upper triangular.*

Remark 8.1. We want to maximize $\text{rank}(V_k) = \text{rank}([\widehat{V}_2, \widehat{V}_3]) = \text{rank}([\widetilde{V}_2, \widetilde{V}_3])$ while hopefully maximizing the orthonormality $\text{rank}(\widehat{V}_3) = \text{rank}(\widetilde{V}_3)$; see (4.7)–(4.8). The cases we have treated in this section help us to understand how any given v_{k+1} will contribute to this. We can see from Theorem 8.4 that when $s_{k+1} = 0$, the ideal case, the new vector $v_{k+1} \perp V_k$, becomes a new unit norm column of both $\widehat{V}_3^{(k+1)} = V_{k+1} W_3^{(k+1)}$ and $\widetilde{V}_3^{(k+1)} = V_{k+1} P_3^{(k+1)}$ (see (8.21), (4.7) and (4.8)), and so increases the orthonormality n_k. Next, if $v_{k+1} \in \text{Range}(\widehat{V}_3)$ so that $\|s_{k+1}\|_2 = 1$, the worst case, then from (8.24) $V_{k+1} w_{m_k+1}^{(k+1)} = V_{k+1} p_{m_k+1}^{(k+1)} = v_{k+1}$ is the new unit norm vector added to both $V_k W_1$ and $V_k P_1$. It follows from (8.20)–(8.23) that every other case is intermediate in that $0 < \|s_{k+1}\|_2 < 1$. Clearly $V_k W_1$ and $V_k P_1$ are repositories for vectors which are redundant because they lie in the range of already obtained orthogonal vectors.

9. Lengths of derived vectors arising from the SVD of S_k. In the computational Lanczos process some accuracy measures are inversely proportional to the lengths of certain derived vectors. Thus it is important to study the lengths of the derived vectors $V_k p_j$ and $V_k w_j$ in (4.7) and (4.8).

COROLLARY 9.1. *With Definition* 4.1*, and* S_k *and* U_k *as in* (2.1)*,*

$$\{p_j^H U_k p_j = w_j^H U_k w_j \ \text{and}\ \|V_k p_j\|_2 = \|V_k w_j\|_2\}, \quad j = 1, \ldots, k, \tag{9.1}$$

$$V_k w_j = V_k p_j \ \forall p_j \in P_1; \qquad \|V_k w_j\|_2 = \|V_k p_j\|_2 = 1 \ \forall p_j \in P_3, \tag{9.2}$$

$$\frac{1-\sigma_j}{1+\sigma_j} \le \|V_k p_j\|_2^2 \le \frac{1+\sigma_j}{1-\sigma_j} \ \text{ for all columns } p_j \text{ of } [P_2, P_3], \tag{9.3}$$

$$\|(I - S_k) p_j\|_2 = \|p_j - w_j \sigma_j\|_2 \ge k^{-1} \quad \forall p_j \in P; \quad \text{rank}(P - W\Sigma) = k. \tag{9.4}$$

Proof. The facts that $V_k^H V_k = I + U_k + U_k^H$ and $\|p_j\|_2 = \|w_j\|_2 = 1$ lead to

$$\|V_k p_j\|_2^2 = 1 + 2\Re(p_j^H U_k p_j), \quad \|V_k w_j\|_2^2 = 1 + 2\Re(w_j^H U_k w_j), \quad j = 1, \ldots, k. \tag{9.5}$$

Combining (2.1) with the SVD of S_k in (4.4) gives

$$U_k P = (I + U_k) W \Sigma, \quad W^H U_k = \Sigma P^H (I + U_k), \quad \sigma_j = 0 \Leftrightarrow U_k p_j = 0 \Leftrightarrow w_j^H U_k = 0, \tag{9.6}$$

so if $\sigma_j = 0$, (9.5) shows that (9.1) holds. If $\sigma_j \neq 0$, then from (9.6) for $j = 1, \ldots, k$,

$$U_k p_j = w_j \sigma_j + U_k w_j \sigma_j, \qquad w_j^H U_k = \sigma_j p_j^H + \sigma_j p_j^H U_k,$$
$$w_j^H U_k w_j = (w_j^H U_k p_j - \sigma_j)/\sigma_j, \qquad p_j^H U_k p_j = (w_j^H U_k p_j - \sigma_j)/\sigma_j = w_j^H U_k w_j,$$

which with (9.5) completes (9.1). The first part of (9.2) follows from (4.7), the second part from (9.5) with (9.6). From (4.7), (4.8), and $\Gamma_2^2 = I - \Sigma_2^2$ in (4.5),

$$\widetilde{V}_2 \Gamma_2 = V_k (P_2 - W_2 \Sigma_2), \qquad \widehat{V}_2 \Gamma_2 = V_k (W_2 - P_2 \Sigma_2),$$
$$(\widetilde{V}_2 + \widehat{V}_2 \Sigma_2) = V_k (P_2 - W_2 \Sigma_2 + W_2 \Sigma_2 - P_2 \Sigma_2^2) \Gamma_2^{-1} = V_k P_2 \Gamma_2. \tag{9.7}$$

Thus when p_j is a column of P_2 we have

$$\tilde{v}_j + \hat{v}_j \sigma_j = V_k p_j \gamma_j, \qquad 1 + \sigma_j^2 + 2\Re(\tilde{v}_j^H \hat{v}_j)\sigma_j = \|V_k p_j\|_2^2 (1 - \sigma_j^2),$$

which leads to (9.3) for $p_j \in P_2$. If $\sigma_j = 0$, then $p_j \in P_3$, and (9.3) follows from (9.2). For any column p_j of P, $p_j = (I - S_k)^{-1}(I - S_k)p_j$, so

$$1 \le \|I + S_k + \cdots + S_k^{k-1}\|_2 \|(I - S_k)p_j\|_2 \le k\|(I - S_k)p_j\|_2 = k\|p_j - w_j\sigma_j\|_2,$$

and $x \neq 0 \Rightarrow (P - W\Sigma)x = (I - S_k)Px \neq 0$, proving (9.4). ∎

Remark 9.1. Note the structure that follows from the *SVD* of S_k. For all $j = 1, \ldots, k$ we have $\|V_k w_j\|_2 = \|V_k p_j\|_2$, but with m_k unit and n_k zero singular values there is the additional information that $V_k w_j = V_k p_j$ for $j = 1, \ldots, m_k$; $(1-\sigma_j)/(1+\sigma_j) \le \|V_k w_j\|_2 = \|V_k p_j\|_2 \le (1+\sigma_j)/(1-\sigma_j)$ for $j = m_k+1, \ldots, k-n_k$; and $\|V_k w_j\|_2 = \|V_k p_j\|_2 = 1$ for $j = k-n_k+1, \ldots, k$, where these are exact since S_k is a theoretical construct. We see from (9.2) that $\sigma_j = 1 \Rightarrow V_k w_j = V_k p_j$, but (9.3) gives no useful bounds on $\|V_k p_j\|_2$ in this case. This is understandable in general because of the indeterminacy of the SVD for repeated singular values. When $\sigma_j = 1$, corresponding to rank deficiency of V_k, the special case of Theorem 8.4 (see (8.24)) gives the only useful bounds so far for the column lengths of $V_k P_1$ and $V_k W_1$. It might not be possible, but if $\sigma_j < 1$ and $V_k w_j = V_k p_j$, then since each column of (4.7) has unit length, $1 = \sigma_j^2 + \|V_k p_j(1-\sigma_j)\|_2^2$, and $\|V_k p_j\|_2^2$ attains the upper bound in (9.3).

After k steps of the finite precision Lanczos process on a large problem, $\mathrm{trace}(\Sigma_2^{(k)})$ is quite regularly seen to be very small, so we derive some bounds in terms of this. $\Sigma_2^{(k)}$ is nonexistent if and only if $m_k + n_k = k$ (see (4.12)), and we denote this by $\Sigma_2^{(k)} = 0$.

Corollary 9.2. *With the notation in Definition* 4.1 *and Theorem* 4.2 *let* S_k *have* m_k *unit and* n_k *zero singular values. Then*

$$|\mathrm{trace}(P_1^H W_1)| \le \mathrm{trace}(\Sigma_2), \quad |\|P_1 - W_1\|_F^2 - 2m_k| \le 2\,\mathrm{trace}(\Sigma_2), \tag{9.8}$$

$$\Sigma_2 = 0 \Rightarrow \{\mathrm{trace}(P_1^H W_1) = 0 \;\&\; \|P_1 - W_1\|_F^2 = 2m_k \;\&\; \|V_k P_1\|_F^2 = m_k\}, \tag{9.9}$$

$$\Sigma_2 \neq 0 \Rightarrow \sum_{j=m_k+1}^{k-n_k} \frac{-2\sigma_j}{1-\sigma_j} \le \|V_k P_1\|_F^2 - m_k \le \sum_{j=m_k+1}^{k-n_k} \frac{2\sigma_j}{1+\sigma_j} \le 2\,\mathrm{trace}(\Sigma_2). \tag{9.10}$$

Proof. From (4.4) we have $P^H S_k P = P^H W \Sigma$, so

$$0 = \mathrm{trace}(S_k) = \mathrm{trace}(P^H S_k P) = \mathrm{trace}(P_1^H W_1) + \mathrm{trace}(P_2^H W_2 \Sigma_2),$$
$$|\mathrm{trace}(P_1^H W_1)| = |\mathrm{trace}(P_2^H W_2 \Sigma_2)| \le \mathrm{trace}(\Sigma_2)$$

since the diagonal elements of $|P_2^H W_2|$ are bounded above by unity, while

$$\|P_1 - W_1\|_F^2 = \mathrm{trace}((P_1 - W_1)^H (P_1 - W_1)) = 2m_k - 2\Re\{\mathrm{trace}(P_1^H W_1)\},$$
$$|\|P_1 - W_1\|_F^2 - 2m_k| \le 2|\mathrm{trace}(P_2^H W_2 \Sigma_2)| \le 2\mathrm{trace}(\Sigma_2),$$

proving (9.8). Next (9.9) follows from (9.8) and (9.2) since then $k = m_k + n_k = \|V_k P_1\|_F^2 + \|V_k P_3\|_F^2 = \|V_k P_1\|_F^2 + n_k$. If $\Sigma_2 \neq 0$, 1 minus each term in (9.3) gives

$$\frac{-2\sigma_j}{1-\sigma_j} \le 1 - \|V_k p_j\|_2^2 = 1 - \|V_k w_j\|_2^2 \le \frac{2\sigma_j}{1+\sigma_j}, \; j = m_k + 1, \ldots, k - n_k,$$

and we sum and use $\|V_k P_1\|_F^2 = k - n_k - \|V_k P_2\|_F^2$ to prove (9.10). ∎

10. The cases of rank$(V_k) = n$ and $Q_{22}^{(k)} = 0$. Although $Q_{22}^{(k)}=0 \Rightarrow \mathrm{rank}(V_k)= n$ (see (6.9)) it is not true that $\mathrm{rank}(V_k)=n \Rightarrow Q_{22}^{(k)}=0$, since with $0<\cos\theta \triangleq |v_1^H v_2|< 1$,

$$\mathcal{P}_1\mathcal{P}_2 v_1 = (I-v_1v_1^H)(I-v_2v_2^H)v_1 = -\mathcal{P}_1 v_2(v_2^H v_1), \quad \|\mathcal{P}_1\mathcal{P}_2 v_1\|_2 = \sin\theta\cdot\cos\theta,$$

which is not zero, showing it is not even true for the $n = 2$ case.

It can be meaningful to take more than n steps, e.g., when the finite precision Lanczos process is used to find a large number of the eigenvalues of a matrix. So one may be interested in understanding what happens in such cases.

From (4.13) we see that if $Q_{22}^{(k)} = 0$, then $\widehat{V}_0^{(k)}$, $\widehat{V}_2^{(k)}$, and $\widetilde{V}_2^{(k)}$ no longer exist and $k = n + m_k$, so that in (4.9) $\widehat{V}_3^{(k)}$, $\widetilde{V}_3^{(k)} \in \mathcal{U}^{n\times n}$. Also from (6.10) $\|s_j\|_2 = 1$ and $\eta_j^2 = 0$ for each $j > k$. Then (8.23) and (8.24) hold in Theorem 8.4, and all future vectors v_j are redundant:

(10.1)

$$Q_{22}^{(k)}=0 \;\&\; j>k \Rightarrow p_{m_j}^{(j)H} w_{m_j}^{(j)} = 0 \;\&\; \|p_{m_j}^{(j)} - w_{m_j}^{(j)}\|_2^2 = 2 \;\&\; V_j w_{m_j}^{(j)} = V_j p_{m_j}^{(j)} = v_j.$$

In general, once $\mathrm{rank}(V_k) = n$, we have $k \geq n$ and $[\widehat{V}_2^{(j)}, \widehat{V}_3^{(j)}]$, $[\widetilde{V}_2^{(j)}, \widetilde{V}_3^{(j)}] \in \mathcal{U}^{n\times n}$ for $j \geq k$. Since $\widehat{V}_0^{(j)}$ does not exist, we see that only item 3 in case 1, and items 3, 4, and 5 in case 2 apply. From case 2 there must always be a new unit singular value of S_{j+1} and the results of Theorem 8.2 hold, the only addition being that $v_{j+1} \in [\widehat{V}_2^{(j)}, \widehat{V}_3^{(j)}]$. Then, in theory, only in item 3 where $v_{j+1} \in \mathrm{Range}(\widehat{V}_2^{(j)})$ do we get $\widehat{V}_2^{(j)}$ having one less column, and $\widetilde{V}_3^{(j+1)}$ one more; see (8.4)–(8.6) without $\widehat{V}_0$. But $\|Q_{22}^{(j)}\|_F^2 = \|\Sigma_2^{(j)}\|_F^2$, and from (6.10) this decreases by $\eta_{j+1}^2 = \|\Sigma_2^{(j)}\widehat{V}_2^{(j)H} v_{j+1}\|_2^2$ with no singular value increasing, so that even if $\widehat{V}_3^{(j)H} v_{j+1} \neq 0$, in practice an algorithm which produces the v_j in an effective way will still lead to numerically singular, and eventually numerically nonexistent, $\Sigma_2^{(j)}$. Each new numerically zero singular value of $\Sigma_2^{(j)}$ will effectively result in an element of $\Gamma_2^{(j)}$ being unity, and a new column of $\widetilde{V}_3^{(j)}$; see the last matrix in (4.7).

Finally when $Q_{22}^{(k)}$ is very small but not zero, case 2 must hold, and then for $j > k$ very small $\Sigma_2^{(j-1)}$ in Theorem 8.2 gives $\mu \approx 1$ and in (8.8) $\check{p}_{m_j}^{(j)} \approx e_{j-1}$, $p_{m_j}^{(j)H} w_{m_j}^{(j)} \approx 0$, $\|p_{m_j}^{(j)} - w_{m_j}^{(j)}\|_2^2 \approx 2$, and $V_j p_{m_j}^{(j)} = V_j w_{m_j}^{(j)} \approx v_j$, so that (10.1) holds approximately. More precise results could be found from Corollary 9.2.

11. Summary. In sections 2 and 3 we summarized earlier results from [1, 9] on producing an $(n + k) \times (n + k)$ unitary matrix $Q^{(k)}$ from k unit Euclidean norm n-vectors, the columns of V_k. A key component of this was the $k \times k$ strictly upper triangular matrix S_k in (2.1).

In section 4 we described the special structure of the SVD $S_k = W^{(k)}\Sigma^{(k)}P^{(k)H}$ and its connection to the CSD of the augmented orthogonal matrix $Q^{(k)}$. In section 5 we used this structure to show that when the SVD is updated from $k \to k{+}1$ the unit singular values persist with their vectors essentially unchanged, while zero singular values persist with their right vectors essentially unchanged. This leads to a simple updating procedure for the SVD of S_k.

Based on the SVD of S_k and the CSD of $Q^{(k)}$ we derived several results on $Q_{22}^{(k)}$, S_k, V_k, and v_{k+1}. Section 6 was devoted to the study of $n\times n$ $Q_{22}^{(k)}$. In many situations it can be observed that $Q_{22}^{(k)}$ tends to zero numerically. It is shown how the process $Q_{22}^{(k)} \to 0$ is related to the columns of S_k. In section 7 we proposed two measures of the effectiveness of v_{k+1} in contributing to orthogonality, while in section 8 we investigated what this v_{k+1} actually does. We showed how it affected the rank deficiency m_{k+1} of V_{k+1} and the number n_{k+1} of zero singular values of S_{k+1} (see Table 1) and then looked at the orthogonal and redundant vectors resulting from v_{k+1}.

We continued in section 9 by obtaining some lower and upper bounds on the length of vectors $V_k w_j$ and $V_k p_j$. Since these are closely related questions, we showed how various bounds can be obtained via the nonunit singular values of S_k. Finally, in section 10 we investigated what happens when $\text{rank}(V_k) = n$ and when $Q_{22}^{(k)} = 0$.

Acknowledgments. We would especially like to thank Jens-Peter Zemke for pointing out some errors and many lacks of clarity, and helping us to produce a tighter, clearer exposition in a greatly improved paper. Also Jörg Liesen gave invaluable help and suggestions.

REFERENCES

[1] Å. Björck and C. C. Paige, *Loss and recapture of orthogonality in the modified Gram–Schmidt algorithm*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 176–190.
[2] Å. Björck and C. C. Paige, *Solution of augmented linear systems using orthogonal factorizations*, BIT, 34 (1994), pp. 1–24.
[3] C. Davis and W. M. Kahan, *Some new bounds on perturbations of subspaces*, Bull. Amer. Math. Soc., 75 (1969), pp. 863–868.
[4] L. Giraud, S. Gratton, and J. Langou, *A rank-k update procedure for reorthogonalizing the orthogonal factor from modified Gram–Schmidt*, SIAM J. Matrix Anal. Appl., 25 (2004), pp. 1163–1177.
[5] G. H. Golub and W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, SIAM J. Numer. Anal., 2 (1965), pp. 205–224.
[6] G. H. Golub, M. Stoll, and A. Wathen, *Approximation of the scattering amplitude and linear systems*, Electron. Trans. Numer. Anal., 31 (2008), pp. 178–203.
[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, 1996.
[8] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Research Nat. Bur. Standards, 45 (1950), pp. 255–282.
[9] C. C. Paige, *A useful form of unitary matrix obtained from any sequence of unit 2-norm n-vectors*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 565–583.
[10] C. C. Paige, *An augmented stability result for the Lanczos Hermitian matrix tridiagonalization process*, SIAM. J. Matrix Anal. Appl., 31 (2010), pp. 2347–2359.
[11] C. C. Paige, I. Panayotov, and J.-P. M. Zemke, *An augmented analysis of the perturbed two-sided Lanczos tridiagonalization process*, Linear Algebra Appl., in press.
[12] C. C. Paige, M. Rozložník, and Z. Strakoš, *Modified Gram–Schmidt (MGS), least squares, and backward stability of MGS-GMRES*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 264–284.
[13] L. Reichel and Q. Ye, *A generalized LSQR algorithm*, Numer. Linear Algebra Appl., 15 (2008), pp. 643–660.
[14] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[15] M. A. Saunders, H. D. Simon, and E. L. Yip, *Two conjugate-gradient-type methods for unsymmetric linear equations*, SIAM J. Numer. Anal., 25 (1988), pp. 927–940.
[16] R. Schreiber and C. F. Van Loan, *A storage-efficient WY representation for products of Householder transformations*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 53–57.
[17] C. Sheffield, *comment to Gene Golub*, circa 1967.
[18] G. W. Stewart, *On the perturbation of pseudo-inverses, projections, and linear least squares problems*, SIAM Rev., 19 (1977), pp. 634–662.
[19] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.