Electronic Transactions on Numerical Analysis.
Volume 23, pp. 5-14, 2006.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

CONDITION NUMBERS OF THE KRYLOV BASES AND SPACES ASSOCIATED WITH THE TRUNCATED QZ ITERATION*

ALEXANDER MALYSHEV† AND MILOUD SADKANE‡

Abstract. We propose exact and computable formulas for computing condition numbers of the Krylov bases and spaces associated with the Hessenberg-Triangular reduction of a regular linear matrix pencil.

Key words. condition number, Krylov spaces, QZ algorithm, generalized Arnoldi algorithm

AMS subject classifications. 65F35, 15A21

1. Introduction. The QZ algorithm [5] is the most popular algorithm for solving the dense generalized eigenvalue problem

$$(\lambda B - A)x = 0, \tag{1.1}$$

where $A, B \in \mathbb{R}^{n\times n}$. It is a generalization of the QR algorithm [2, Sec. 7.7], which solves the standard eigenproblem $Ax = \lambda x$. The first step of the QZ algorithm reduces, via orthogonal matrices V and U, the pair (A, B) to (H, R), where the matrices $H = U^TAV$ and $R = U^TBV$ are respectively upper Hessenberg and upper triangular:

$$H = \begin{pmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,n} \\ h_{2,1} & h_{2,2} & \dots & h_{2,n} \\ & \ddots & \ddots & \vdots \\ & & h_{n,n-1} & h_{n,n} \end{pmatrix}, \quad R = \begin{pmatrix} r_{1,1} & r_{1,2} & \dots & r_{1,n} \\ & r_{2,2} & \dots & r_{2,n} \\ & & \ddots & \vdots \\ & & & r_{n,n} \end{pmatrix}. \tag{1.2}$$

The Hessenberg-Triangular reduction can be recast as

$$\left\{ \begin{array}{l} AV = UH, \\ BV = UR. \end{array} \right. \tag{1.3}$$

It is possible to construct the reduction (1.3) iteratively, starting from a vector v with $\|v\|_2 = 1$. The symbol $\| \ \|_2$ stands for the Euclidean vector norm. The following algorithm, taken from [7], can be used to accomplish the iterative reduction. It is written in MATLAB style.

ALGORITHM 1 (Generalized Arnoldi Reduction [7]).

- **Choose** $v \in \mathbb{R}^n$ with $\|v\|_2 = 1$
- **Set** $V_1 = [v]$; $u = Bv$; $\rho = \|u\|_2$; $R_1 = [\rho]$; $U_1 = [u/\rho]$;
- **Set** $z = Av$; $H_1 = [U_1^T z]$; $f_1 = z - U_1 H_1$;
- **for** $j = 1, 2 \dots$ **do**
  $\gamma = \|f_j\|_2$; $u = f_j/\gamma$;
  $U_{j+1} = [U_j \ , \ u]$; $H_j = [H_j \ ; \ \gamma e_j^T]$;

*Received January 3, 2004. Accepted for publication November 15, 2005. Recommended by D. Sorensen.
†University of Bergen, Department of Mathematics, Johannes Brunsgate 12, N-5008 Bergen, Norway (malyshev@mi.uib.no).
‡Université de Brest, Département de Mathématiques, 6 avenue Victor Le Gorgeu, CS 93837, F-29238 Brest Cedex 3, France (sadkane@univ-brest.fr).

Solve $B\hat{v} = u$;
$z = V_j^T \hat{v}$; $\hat{v} = \hat{v} - V_j z$; $\rho = 1/\|\hat{v}\|_2$;
$v = \rho\hat{v}$; $r = -R_j z \rho$;
$V_j = [V_j \ , \ v]$; $R_{j+1} = [R_j \ , \ r \ ; \ 0 \ , \ \rho]$;
$z = Av$; $h = U_{j+1}^T z$;
$f_{j+1} = z - U_{j+1} h$; $H_{j+1} = [H_j \ , \ h]$;
end for

After $k \le n$ steps of Algorithm 1 we obtain a truncation of (1.3) in the form

$$\text{(1.4)} \qquad \begin{cases} AV_k = U_k H_k + f_k e_k^T, \\ BV_k = U_k R_k, \\ V_k^T V_k = I_k, \ U_k^T U_k = I_k, \ U_k^T f_k = 0, \end{cases}$$

where the matrices $V_k \equiv V(1\!:n, 1\!:k)$ and $U_k \equiv U(1\!:n, 1\!:k)$ represent the first k columns of V and U respectively. The matrices $H_k = H(1\!:k, 1\!:k)$ and $R_k = R(1\!:k, 1\!:k)$ are the leading $k \times k$ submatrices of H and R. The vector e_k represents the k-th canonical vector and I_k is the identity matrix of order k.

The form (1.4) is a generalization of the classical Arnoldi reduction [6, Chap. VI] and can, for example, be used to compute eigenpairs of the problem (1.1), when the sizes of A and B are large (see [7]). In exact arithmetic, the sets $\mathcal{B}_k = \{v_1, \ldots, v_k\}$ and $\mathcal{C}_k = \{u_1, \ldots, u_k\}$, $k = 1, 2, \ldots, n$, form bases of the Krylov spaces

$$\text{(1.5)} \qquad \mathcal{K}_k \equiv \mathcal{K}_k(B^{-1}A, v_1) = \text{Span}\{v_1, B^{-1}Av_1, \ldots, (B^{-1}A)^{k-1} v_1\}$$

and

$$\text{(1.6)} \qquad \mathcal{L}_k \equiv \mathcal{L}_k(AB^{-1}, u_1) = \text{Span}\{u_1, AB^{-1}u_1, \ldots, (AB^{-1})^{k-1} u_1\}$$

respectively, with $v_1 = V_1 \equiv v$ and $u_1 = Bv_1/\|Bv_1\|_2$.

Unfortunately, the generalized Arnoldi reduction of the matrix pair (A, B) in a finite precision arithmetic will not generally satisfy the relations (1.4) because the computed matrices V, U, H and R are subject to rounding errors. The effects of rounding errors are proportional to condition numbers, which are usually determined in terms of infinitesimal perturbations. The chief advantage of infinitesimal perturbations is that they allow one to neglect second order terms.

Let $\widetilde{\Delta}_1$ and $\widetilde{\Delta}_2$ be infinitesimal perturbations of A and B respectively, and denote by $\widetilde{H} = \widetilde{U}^T(A + \widetilde{\Delta}_2)\widetilde{V}$ and $\widetilde{R} = \widetilde{U}^T(B + \widetilde{\Delta}_1)\widetilde{V}$ the corresponding Hessenberg and triangular forms, where $\widetilde{U} = [\widetilde{u}_1, \widetilde{u}_2, \ldots, \widetilde{u}_n]$ and $\widetilde{V} = [\widetilde{v}_1, \widetilde{v}_2, \ldots, \widetilde{v}_n]$ are the computed orthogonal matrices with $\widetilde{v}_1 = v_1$.

In the present paper we want to analyze the difference between the exact quantities V, U, $\mathcal{K}_k$, $\mathcal{L}_k$ and their computed counterparts $\widetilde{V}$, $\widetilde{U}$, $\widetilde{\mathcal{K}}_k$, $\widetilde{\mathcal{L}}_k$ under variations of the perturbations $\widetilde{\Delta}_1$ and $\widetilde{\Delta}_2$. More precisely, we are interested in the condition numbers of the Krylov spaces $\mathcal{K}_k$ and $\mathcal{L}_k$ and corresponding bases $\mathcal{B}_k$ and $\mathcal{C}_k$.

We use the arguments similar to those from [1, 3] for the Hessenberg reduction by Arnoldi's method and those from [4] for the bidiagonal reduction by the Lanczos method. We assume that the reader is familiar with the arguments and reasoning used in these references.

2. Condition numbers. Following the strategy used in [1, 3, 4], perturbations of $\widetilde{V}$ and $\widetilde{U}$ are sought in the multiplicative form, that is $\widetilde{V} = (I + \widetilde{X})V$ and $\widetilde{U} = (I + \widetilde{Y})U$. The orthogonality of $\widetilde{V}$, V, $\widetilde{U}$ and U implies the orthogonality of $I + \widetilde{X}$ and $I + \widetilde{Y}$. Since the

ETNA
Kent State University
etna@mcs.kent.edu

perturbations $\widetilde{X}$ and $\widetilde{Y}$ are infinitesimal, the matrices $\widetilde{X}$ and $\widetilde{Y}$ are indistinguishable from skew-symmetric matrices. Discarding quadratic terms in the identity $\widetilde{H} = U^T(I+\widetilde{Y}^T)(A+\widetilde{\Delta}_2)(I+\widetilde{X})V$, we obtain the equation

$$H - \widetilde{H} + U^T\widetilde{\Delta}_2 V + (U^T\widetilde{Y}^T U)H + H(V^T\widetilde{X}V) = 0. \tag{2.1}$$

If we set

$$\Delta_2 = U^T\widetilde{\Delta}_2 V,\ X = V^T\widetilde{X}V,\ \text{and}\ Y = U^T\widetilde{Y}U,$$

then X and Y are indistinguishable from skew-symmetric matrices and satisfy, up to the first order, the equation

$$YH - HX = \Delta_2 + H - \widetilde{H}. \tag{2.2}$$

In a similar way, another equation follows from $\widetilde{R} = U^T(I+\widetilde{Y}^T)(B+\widetilde{\Delta}_1)(I+\widetilde{X})V$, which yields

$$YR - RX = \Delta_1 + R - \widetilde{R} \tag{2.3}$$

with $\Delta_1 = U^T\widetilde{\Delta}_1 V$.

The Krylov spaces $\mathcal{K}_k$ and $\mathcal{L}_k$ are the vector spaces spanned by the columns of V_k and U_k respectively. Their perturbations $\widetilde{\mathcal{K}}_k$ and $\widetilde{\mathcal{L}}_k$ are spanned by the columns of $\widetilde{V}_k = (I+\widetilde{X})V_k$ and $\widetilde{U}_k = (I+\widetilde{Y})U_k$. The Frobenius norm of the difference

$$\left\|V_k - \widetilde{V}_k\right\|_F = \left\|VXV^TV_k\right\|_F = \|X(1{:}n, 1{:}k)\|_F \tag{2.4}$$

can be used to measure the conditioning $\kappa_b \equiv \kappa_b(B^{-1}A, v_1)$ of the basis $\mathcal{B}_k$, which is defined as (see [1, 3])

$$\kappa_b = \inf_{\epsilon>0}\left\{ \sup_{\substack{\|\Delta_1\|_F \le \epsilon \\ \|\Delta_2\|_F \le \epsilon}} \frac{\|V_k - \widetilde{V}_k\|_F}{\sqrt{\frac{\|\Delta_1\|_F^2}{\|B\|_F^2} + \frac{\|\Delta_2\|_F^2}{\|A\|_F^2}}} \right\}.$$

In other words, κ_b is the smallest constant such that

$$\|X(1{:}n, 1{:}k)\|_F \le \kappa_b \sqrt{\frac{\|\Delta_1\|_F^2}{\|B\|_F^2} + \frac{\|\Delta_2\|_F^2}{\|A\|_F^2}}. \tag{2.5}$$

Similarly, the quantity $\|Y(1{:}n, 1{:}k)\|_F$ can be used to measure the conditioning of the basis $\mathcal{C}_k$. Because of similarities between the bases $\mathcal{B}_k$ and $\mathcal{C}_k$ we will only analyze the conditioning of $\mathcal{B}_k$. To this end, we will give a computable estimate of the condition number κ_b.

Let us take the components below the main diagonal in (2.3) and those below the subdiagonal in (2.2). The operation of taking the components below the main diagonal is denoted by $\mathcal{L}^{(1)}$ and that of below the subdiagonal by $\mathcal{L}^{(2)}$. Thus, from equations (2.3), (2.2) we derive the system of linear equations

$$\mathcal{L}^{(1)}(YR - RX) = \mathcal{L}^{(1)}(\Delta_1), \tag{2.6}$$

$$\mathcal{L}^{(2)}(YH - HX) = \mathcal{L}^{(2)}(\Delta_2). \tag{2.7}$$

ETNA
Kent State University
etna@mcs.kent.edu

Note that the diagonal elements of X and Y are equal to zero because the matrices X and Y are real skew-symmetric. Moreover, since $\widetilde{v}_1 = v_1$ and $\tilde{V} = \left(I + \widetilde{X}\right)V$, we have the identity $\widetilde{X}\widetilde{v}_1 = 0$ and, therefore, the first column and row of X are also equal to zero. The structure of X and Y can be described with the help of vectors $x_p \in \mathbb{R}^{n-p-1}$, $p = 1, \dots, n-2$, and $y_q \in \mathbb{R}^{n-q}$, $q = 1, \dots n-1$, such that

$$X = \begin{pmatrix} 0 & 0 & & \dots & 0 & 0 \\ 0 & 0 & & & -x_1^T & \\ \vdots & & \ddots & & -x_2^T & \\ \vdots & & & \ddots & \vdots & \\ 0 & x_1 & x_2 & \dots & 0 & \\ 0 & & & & & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & & & & -y_1^T & \\ & 0 & & & -y_2^T & \\ & & \ddots & & -y_3^T & \\ & & & \ddots & \vdots & \\ y_1 & y_2 & y_3 & \dots & 0 & \\ & & & & & 0 \end{pmatrix}.$$

Similarly, with the help of vectors $f_p \in \mathbb{R}^{n-p}$, $p = 1, \dots, n-1$, and $g_q \in \mathbb{R}^{n-q-1}$, $q = 1, \dots, n-2$, the matrices Δ_1 and Δ_2 are written as

$$\Delta_1 = \begin{pmatrix} \times & \times & & & \times \\ & \times & \ddots & & \\ & & \ddots & \times & \\ f_1 & f_2 & f_3 & \times & \times \\ & & & & \times \end{pmatrix}, \qquad \Delta_2 = \begin{pmatrix} \times & \times & & & \times \\ \times & \times & \ddots & & \\ & \times & \ddots & \times & \\ & & . & . & \times \\ g_1 & g_2 & g_3 & \times & \times \end{pmatrix}.$$

By the aid of the vectors $x_1, \dots, x_{n-2}$ and $y_1, \dots, y_{n-1}$ we determine the lower triangular parts of X and Y:

$$\widehat{X} = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & x_1 & x_2 & 0 & 0 \\ 0 & & & & 0 \end{pmatrix}, \quad \widehat{Y} = \begin{pmatrix} 0 & 0 & & \dots & 0 \\ & 0 & & \dots & 0 \\ & & \ddots & & \vdots \\ y_1 & y_2 & y_3 & 0 & \\ & & & & 0 \end{pmatrix}.$$

Then $X = \widehat{X} - \widehat{X}^T$ and $Y = \widehat{Y} - \widehat{Y}^T$.

It is important now to observe that

$$\mathcal{L}^{(1)}\left(\widehat{Y}^T R\right) = 0, \quad \mathcal{L}^{(1)}\left(R\widehat{X}^T\right) = 0, \quad \mathcal{L}^{(2)}\left(\widehat{Y}^T H\right) = 0, \quad \mathcal{L}^{(2)}\left(H\widehat{X}^T\right) = 0.$$

As a result,

$$\mathcal{L}^{(1)}\left(\widehat{Y}R - R\widehat{X}\right) = \mathcal{L}^{(1)}\left(\Delta_1\right), \tag{2.8}$$

$$\mathcal{L}^{(2)}\left(\widehat{Y}H - H\widehat{X}\right) = \mathcal{L}^{(2)}\left(\Delta_2\right). \tag{2.9}$$

Equation (2.8) is equivalent to the system

$$\begin{cases} f_1 = r_{1,1}y_1, \\ f_i = \sum_{j=1}^{i} r_{j,i}\mathcal{I}_{n-i}^{n-j}y_j - \check{R}_i x_{i-1}, \quad i = 2, \dots, n-1, \end{cases} \tag{2.10}$$

ETNA
Kent State University
etna@mcs.kent.edu

where $\check{R}_i = R(i+1\!:\!n, i+1\!:\!n) \in \mathbb{R}^{(n-i)\times(n-i)}$ and for $i \le j$, $\mathcal{I}_i^j = \left(0_{i\times(j-i)}\ ,\ I_i\right) \in \mathbb{R}^{i\times j}$ is the matrix whose first $j-i$ columns are equal to zero and the last i columns form the identity matrix of order i. Similarly, equation (2.9) is equivalent to

$$\begin{cases} g_1 = h_{1,1}\mathcal{I}_{n-2}^{n-1}y_1 + h_{2,1}y_2, \\ g_i = \sum_{j=1}^{i+1} h_{j,i}\mathcal{I}_{n-i-1}^{n-j}y_j - \check{H}_i x_{i-1}, \quad i = 2,\dots,n-2, \end{cases} \tag{2.11}$$

where $\check{H}_i = H(i+2\!:\!n, i+1\!:\!n) \in \mathbb{R}^{(n-i-1)\times(n-i)}$.

Let us summarize the above derivations. For $k = 1, 2, \dots, n-1$, we have a series of systems of linear equations

$$\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_k \end{pmatrix} = \begin{pmatrix} r_{1,1}I_{n-1} & & & \\ r_{1,2}\mathcal{I}_{n-2}^{n-1} & r_{2,2}I_{n-2} & & \\ \vdots & \ddots & \ddots & \\ r_{1,k}\mathcal{I}_{n-k}^{n-1} & & r_{k-1,k}\mathcal{I}_{n-k}^{n-k+1} & r_{k,k}I_{n-k} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix} - \begin{pmatrix} 0 & & & \\ \check{R}_2 & 0 & & \\ & \ddots & \ddots & \\ & & \ddots & 0 \\ & & & \check{R}_k \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \end{pmatrix},$$

$$\begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_{k-1} \end{pmatrix} = \begin{pmatrix} h_{1,1}\mathcal{I}_{n-2}^{n-1} & h_{2,1}I_{n-2} & & \\ h_{1,2}\mathcal{I}_{n-3}^{n-1} & h_{2,2}\mathcal{I}_{n-3}^{n-2} & h_{3,2}I_{n-3} & \\ \vdots & & & \\ h_{1,k-1}\mathcal{I}_{n-k}^{n-1} & h_{2,k-1}\mathcal{I}_{n-k}^{n-2} & \dots & h_{k,k-1}I_{n-k} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix} - \begin{pmatrix} 0 & & & \\ \check{H}_2 & 0 & & \\ & \ddots & \ddots & \\ & & \check{H}_{k-1} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \end{pmatrix},$$

which we can write in the compact form

$$\begin{cases} f = M_{11}y - M_{12}x, \\ g = M_{21}y - M_{22}x. \end{cases} \tag{2.12}$$

Since Algorithm 1 proceeds if and only if $r_{k,k} \neq 0$ and $h_{k,k-1} \neq 0$, the matrix M_{11} is nonsingular. It follows that the system (2.12) has the solution

$$x = M_x \begin{pmatrix} f \\ g \end{pmatrix} \quad \text{and} \quad y = M_y \begin{pmatrix} f \\ g \end{pmatrix}, \tag{2.13}$$

where

$$\begin{cases} M_x = (M_{22} - M_{21}M_{11}^{-1}M_{12})^{-1}\left(M_{21}M_{11}^{-1}\ ,\ -I_{N_k}\right) \\ M_y = M_{11}^{-1}M_{12}M_x + \left(M_{11}^{-1}\ ,\ 0\right). \end{cases} \tag{2.14}$$

Here I_{N_k} (I_{M_k}) is the identity matrix of order N_k (M_k) with $N_k = n(k-1) - \frac{k(k+1)}{2} + 1$ and $M_k = nk - \frac{k(k+1)}{2}$.

ETNA
Kent State University
etna@mcs.kent.edu

Using the identity $\|x\|_2 = \left(\|x_1\|_2^2 + \ldots + \|x_k\|_2^2\right)^{\frac{1}{2}} = \|X(1\colon n, 1\colon k)\|_F$ we arrive at the formula

$$\|X(1\colon n,1\colon k)\|_F = \left\| M_x \begin{pmatrix} f \\ g \end{pmatrix} \right\|_2$$

$$= \left\| M_x \begin{pmatrix} \|R\|_F I_{M_k} & 0 \\ 0 & -\|H\|_F I_{N_k} \end{pmatrix} \begin{pmatrix} f/\|R\|_F \\ -g/\|H\|_F \end{pmatrix} \right\|_2$$

$$\leq \left\| M_x \begin{pmatrix} \|R\|_F I_{M_k} & 0 \\ 0 & -\|H\|_F I_{N_k} \end{pmatrix} \right\|_2 \left\| \begin{pmatrix} f/\|R\|_F \\ -g/\|H\|_F \end{pmatrix} \right\|_2 \tag{2.15}$$

$$\leq \left\| M_x \begin{pmatrix} \|R\|_F I_{M_k} & 0 \\ 0 & -\|H\|_F I_{N_k} \end{pmatrix} \right\|_2 \sqrt{\frac{\|\Delta_1\|_F^2}{\|B\|_F^2} + \frac{\|\Delta_2\|_F^2}{\|A\|_F^2}}. \tag{2.16}$$

If we choose f and g such that

$$\left\| M_x \begin{pmatrix} f \\ g \end{pmatrix} \right\|_2 = \left\| M_x \begin{pmatrix} \|R\|_F I_{M_k} & 0 \\ 0 & -\|H\|_F I_{N_k} \end{pmatrix} \right\|_2 \left\| \begin{pmatrix} f/\|R\|_F \\ -g/\|H\|_F \end{pmatrix} \right\|_2$$

and

$$\left\| \begin{pmatrix} f/\|R\|_F \\ -g/\|H\|_F \end{pmatrix} \right\|_2 = \sqrt{\frac{\|\Delta_1\|_F^2}{\|B\|_F^2} + \frac{\|\Delta_2\|_F^2}{\|A\|_F^2}},$$

then inequalities (2.15) and (2.16) become equalities (see [4] for a similar reasoning).

We thus obtain a series of condition numbers of the orthonormal bases (not spaces!) for $k = 1, 2, \ldots, n-1$

$$\kappa_{k,b}(B^{-1}A, v_1) = \left\| M_x \begin{pmatrix} \|R\|_F I_{M_k} & 0 \\ 0 & -\|H\|_F I_{N_k} \end{pmatrix} \right\|_2$$

$$= \left\| \left(M_{22} - M_{21}M_{11}^{-1}M_{12}\right)^{-1} \left(\|R\|_F M_{21}M_{11}^{-1}\ ,\ \|H\|_F I_{N_k}\right) \right\|_2. \tag{2.17}$$

In order to derive condition numbers of the corresponding Krylov spaces $\mathcal{K}_k$ we use the same arguments as in [4]. The main idea is to compare $\widetilde{V}_k = (I + \widetilde{X})V_k$ with $V_k Q_k$ instead of V_k for all orthogonal matrices Q_k of order k. The minimum

$$\min_{\substack{Q \in \mathbb{R}^{k\times k} \\ Q^T Q = I_k}} \left\| \widetilde{V}_k - V_k Q \right\|_F$$

is attained at $Q_k = W_l W_r^T$, where $W_l^T \left(V_k^T \widetilde{V}_k\right) W_r = \Sigma$ is the singular value decomposition of $V_k^T \widetilde{V}_k$ (see, e.g., [2, p.582]). Since

$$V_k^T \widetilde{V}_k = I_k + V_k^T \widetilde{X} V_k = I_k + X(1:k, 1:k) \tag{2.18}$$

is indistinguishable from an orthogonal matrix, we can take $Q_k = W_l = I_k + X(1\colon k, 1\colon k)$ and $W_r = \Sigma = I_k$. This leads to the following interpretation. In fact, the leading $k \times k$ submatrix of X is cut off. The rest of X represents the distance between the Krylov spaces, i.e., only last $n - k$ components of each vector x_i contribute to the difference between the

Krylov spaces. Let us introduce the matrix Π_k, whose columns are the coordinate vectors e_i corresponding to the contributing components. Then the condition numbers of the Krylov spaces are determined for $k = 1, 2, \ldots, n-1$ as

$$\kappa_k(B^{-1}A, v_1) = \left\| \Pi_k (M_{22} - M_{21}M_{11}^{-1}M_{12})^{-1} \left(\|R\|_F M_{21}M_{11}^{-1}, \|H\|_F I_{N_k} \right) \right\|_2 . \tag{2.19}$$

The condition numbers of the Krylov bases $\mathcal{C}_k$ and corresponding Krylov spaces $\mathcal{L}_k$, $k = 1, 2, \ldots, n-1$, can be derived analogously by the aid of the expression for y in (2.13):

$$\kappa_{k,b}(AB^{-1}, u_1) = \left\| M_y \begin{pmatrix} \|R\|_F I_{M_k} & 0 \\ 0 & -\|H\|_F I_{N_k} \end{pmatrix} \right\|_2 , \tag{2.20}$$

$$\kappa_k(AB^{-1}, u_1) = \left\| \Pi_k M_y \begin{pmatrix} \|R\|_F I_{M_k} & 0 \\ 0 & -\|H\|_F I_{N_k} \end{pmatrix} \right\|_2 . \tag{2.21}$$

3. Numerical examples. In this section we illustrate behavior of the condition numbers of Krylov bases and subspaces given in (2.17) and (2.19) on some examples.

EXAMPLE 3.1. A is the identity matrix, and B is the upper-Hessenberg Toeplitz matrix

$$B_\alpha = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \alpha & 1 & \ldots & 1 \\ & \ddots & \ddots & \vdots \\ & & \alpha & 1 \end{pmatrix} .$$

The matrix B_α is singular if and only if $\alpha = 1$. Our aim here is to show the effect of conditioning of B_α on the computed basis and subspace condition numbers, when the parameter α varies. The Hessenberg and triangular forms obtained by Algorithm 1 are of order $k = 15$. The starting vector v has all its components equal to 1 before normalization. Table 3.1 summarizes the information obtained on the computed matrices V_k, U_k, R_k and H_k for different parameters α. The standard condition number $\chi(B_\alpha) = \|B_\alpha\|_2 \|(B_\alpha)^{-1}\|_2$ is also given.

TABLE 3.1
Numerical results from Algorithm 1 (Example 3.1)

α	8	4	2	1.3	0
$\\|I_k - V_k^T V_k\\|_2$	$4.55\ 10^{-16}$	$4.83\ 10^{-16}$	$5.42\ 10^{-16}$	$4.59\ 10^{-16}$	$4.52\ 10^{-16}$
$\\|I_k - U_k^T U_k\\|_2$	$2.58\ 10^{-15}$	$8.28\ 10^{-15}$	$3.22\ 10^{-15}$	$3.04\ 10^{-15}$	$1.52\ 10^{-15}$
$\\|H_k - U_k^T A V_k\\|_2$	$2.56\ 10^{-15}$	$8.14\ 10^{-15}$	$2.77\ 10^{-15}$	$2.98\ 10^{-15}$	$1.47\ 10^{-15}$
$\\|R_k - U_k^T B_\alpha V_k\\|_2$	$5.19\ 10^{-14}$	$1.30\ 10^{-13}$	$1.00\ 10^{-11}$	$1.53\ 10^{-6}$	$1.21\ 10^{-14}$
$\\|B_\alpha V_k - U_k R_k\\|_2$	$3.43\ 10^{-14}$	$1.12\ 10^{-13}$	$1.00\ 10^{-11}$	$1.54\ 10^{-6}$	$1.89\ 10^{-15}$
$\\|AV_k - U_k H_k - f_k e_k^T\\|_2$	$2.12\ 10^{-16}$	$6.01\ 10^{-16}$	$5.00\ 10^{-16}$	$4.66\ 10^{-16}$	$2.41\ 10^{-16}$
$\chi(B_\alpha)$	$2.58\ 10^{2}$	$4.28\ 10^{3}$	$1.00\ 10^{7}$	$2.83\ 10^{13}$	$2.60\ 10^{1}$
$\chi(R_k)$	$2.40\ 10^{2}$	$2.63\ 10^{3}$	$9.44\ 10^{5}$	$1.32\ 10^{11}$	$2.59\ 10^{1}$
$\chi(W_\alpha)$	$9.78\ 10^{1}$	$4.91\ 10^{3}$	$4.01\ 10^{7}$	$8.21\ 10^{11}$	∞

The results show that the orthogonality of V_k and U_k is well maintained along with the first equality of (1.4), i.e., $AV_k = U_k H_k + f_k e_k^T$. However, the relations $B_\alpha V_k = U_k R_k$ and $U_k^T B_\alpha V_k = R_k$ deteriorate, when B_α gets ill-conditioned. Figure 3.1 shows the behavior of the condition numbers of the Krylov bases $\kappa_b \left(\equiv \kappa_{b,k}\left(B^{-1}A, v_1\right)\right)$ and Krylov spaces $\kappa \left(\equiv \kappa_k \left(B^{-1}A, v_1\right)\right)$. We observe that the condition numbers of the bases increase with the dimension of the bases and that the condition numbers of the corresponding spaces increase and decrease but are always smaller than those of the bases.

The ill-conditioning of B_α clearly influences the computed condition numbers (see the case $\alpha = 1.3$ in Table 3.1). It can be more practical to monitor the condition number $\chi(R_k) = \|R_k\|_2 \|R_k^{-1}\|_2$ instead of $\chi(B_\alpha)$. This quantity is always available. Since $\chi(M_{11}) \geq \chi(R_k)$, large $\chi(R_k)$ means essentially that the matrix M_{11} in (2.12) is ill-conditioned. The computed condition numbers might therefore be large or even inaccurate.

One may wonder whether the "non-normality" influences, in a way, the computed condition numbers. Let W_α denote a matrix of eigenvectors of B_α and define

$$\chi(W_\alpha) = \begin{cases} \|W_\alpha\|_2 \|(W_\alpha)^{-1}\|_2 & \text{if } B_\alpha \text{ is diagonalizable,} \\ \infty & \text{otherwise.} \end{cases}$$

The factor $\chi(W_\alpha)$ can be used to quantify the departure from normality of B_α. The smaller $\chi(W_\alpha)$, the closer B_α to a normal matrix. Table 3.1 and Figure 3.1 show that the computed factors $\chi(W_\alpha)$, κ_b and κ do not seem to be related.

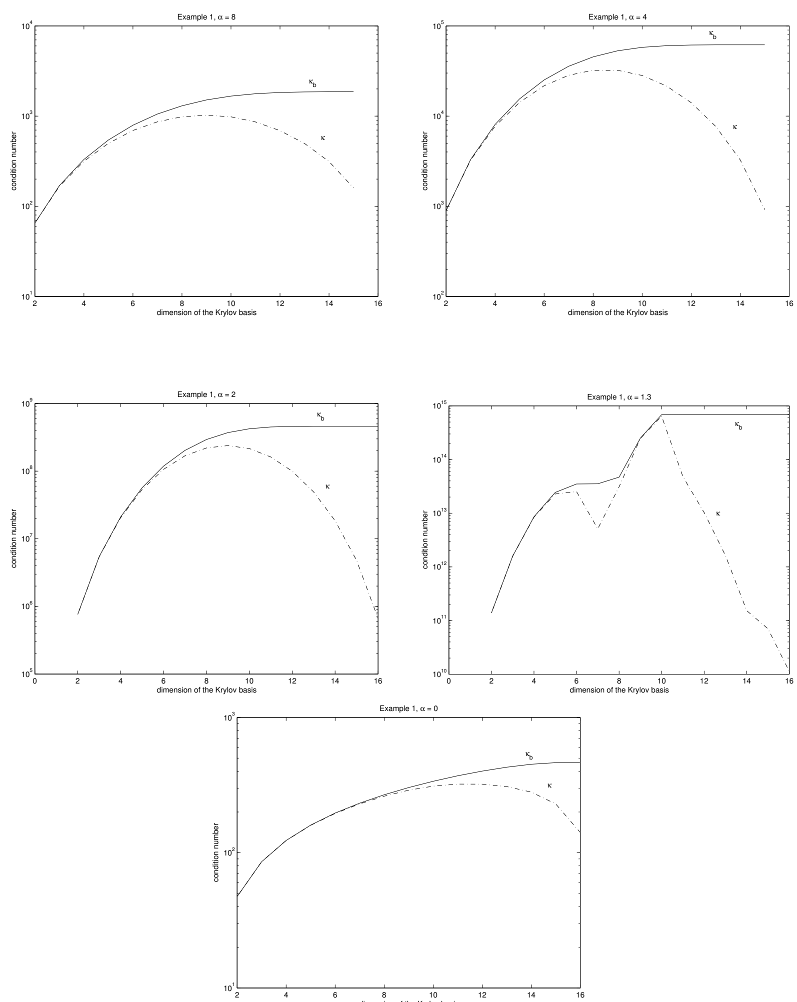

FIG. 3.1. *Condition numbers of the Krylov bases κ_b and spaces κ (Example 3.1).*

TABLE 3.2
Numerical results from Algorithm 1 (Example 3.2)

matrix pair	(A, B)	(A_0, B)	(A, B_0)
$\|I_k - V_k^T V_k\|_2$	$3.01\ 10^{-15}$	$3.02\ 10^{-15}$	$3.00\ 10^{-15}$
$\|I_k - U_k^T U_k\|_2$	$2.08\ 10^{-11}$	$3.33\ 10^{-15}$	$1.32\ 10^{-12}$
$\|H_k - U_k^T A V_k\|_2$	$2.17\ 10^{-9}$	$6.02\ 10^{-12}$	$1.33\ 10^{-10}$
$\|R_k - U_k^T B V_k\|_2$	$4.42\ 10^{-12}$	$4.53\ 10^{-12}$	$3.98\ 10^{-12}$
$\frac{\|BV_k - U_k R_k\|_2}{\|B\|_2}$	$2.01\ 10^{-12}$	$2.07\ 10^{-12}$	$6.55\ 10^{-16}$
$\frac{\|AV_k - U_k H_k - f_k e_k^T\|_2}{\|A\|_2}$	$2.99\ 10^{-13}$	$5.74\ 10^{-16}$	$2.57\ 10^{-14}$
$\chi(R_k)$	$6.10\ 10^{6}$	$6.71\ 10^{8}$	1.95

EXAMPLE 3.2. The matrix A is $MHD416A$ and the matrix B is $MHD416B$ from the MHD set[1]. These matrices are of order $n = 416$ and arise in the modal analysis of dissipative magnetohydrodynamics. The matrix A is unsymmetric, $\|A\|_2 = 2.52 \cdot 10^3$, $\chi(A) = 9.75 \cdot 10^{23}$. The matrix B is symmetric, $\|B\|_2 = 2.19$, $\chi(B) = 3.99 \cdot 10^9$. The Hessenberg and triangular forms obtained from Algorithm 1 are of order $k = 50$. As in the previous example, the starting vector v has all its components equal to 1 before normalization. Table 3.2 summarizes the information obtained on the computed matrices V_k, U_k, R_k and H_k. The condition number $\chi(R_k)$ is also given. For more comparisons, the table also shows the results obtained with the well-conditioned matrices $A_0 = A + \|A\|_2 I$ and $B_0 = B + \|B\|_2 I$. $\|A_0\|_2 = 4.11 \cdot 10^3$, $\chi(A_0) = 2.61$, $\|B_0\|_2 = 4.39$, $\chi(B_0) = 2$.

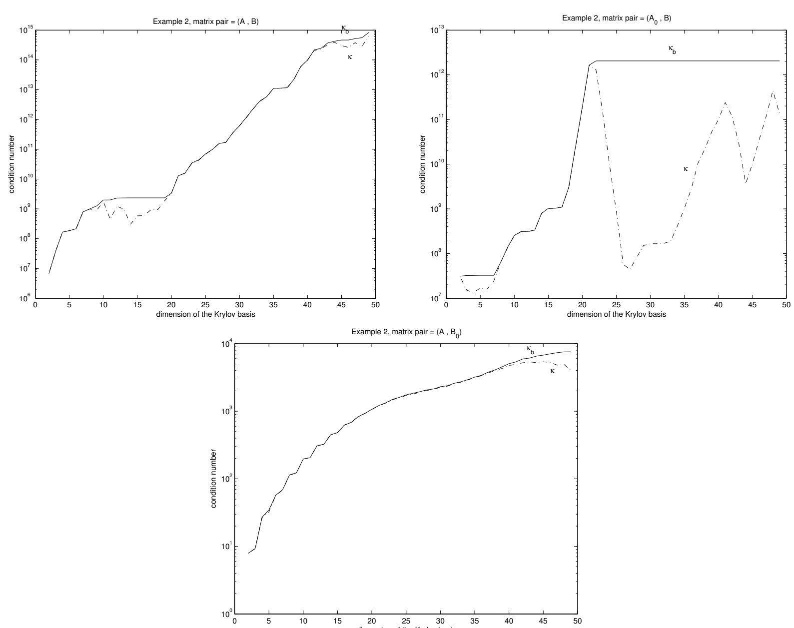

FIG. 3.2. *Condition numbers of the Krylov bases κ_b and spaces κ (Example 3.2)*

From Table 3.2 and Figure 3.2, it is clear that the ill-conditioning of B, also shown in $\chi(R_k)$, is responsible for the large condition numbers. We also see that

[1] see http://math.nist.gov/MatrixMarket/

ETNA
Kent State University
etna@mcs.kent.edu

- the use of A leads to less accuracy in the approximations $U_k^T U_k \approx I_k$, $H_k \approx U_k^T A V_k$,
- the use of B leads to less accuracy in the approximation $BV_k \approx U_k R_k$,
- the less accuracy is more pronounced when both A and B are used.

In conclusion, the numerical experiments indicate that when the truncated reduction is good, i.e., when the relations (1.4) are accurately satisfied, the ill-conditioning of B is responsible for large basis and space condition numbers.

REFERENCES

[1] J.-F. CARPRAUX, S. K. GODUNOV, AND S. V. KUZNETSOV, *Condition number of the Krylov bases and subspaces*, Linear Algebra Appl., 248 (1996), pp. 136–160.

[2] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations, Third editon*, The Johns Hopkins Press, Baltimore, MD., 1996.

[3] S .V. KUZNETSOV, *Perturbation bounds of the Krylov bases and associated Hessenberg forms*, Linear Algebra Appl., 265 (1997), pp. 1–28.

[4] A. MALYSHEV AND M. SADKANE, *Condition numbers for Lanczos bidiagonalization with complete reorthogonalization*, Linear Algebra Appl., 371 (2003), pp. 315–331.

[5] C. B. MOLER AND G. W. STEWART, *An algorithm for generalized matrix eigenvalue problems*, SIAM J. Numer. Anal., 10 (1973), pp. 241–256.

[6] Y. SAAD, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, Manchester, 1992.

[7] D. C. SORENSEN, *Truncated QZ methods for large scale generalized eigenvalue problems*, Electron. Trans. Numer. Anal., 7 (1998), pp. 141–162,
http://etna.math.kent.edu/vol.7.1998/pp141-162.dir/pp141-162.html.