Lyapunov, Lanczos, and Inertia*†

A.C. Antoulas‡ and D.C. Sorensen§

October 19, 2000

Abstract. We present a new proof of the inertia result associated with Lyapunov equations. Furthermore we present a connection between the Lyapunov equation and the Lanczos process which is closely related to the Schwarz form of a matrix. We provide a method for reducing a general matrix to Schwarz form in a finite number of steps ($\mathcal{O}(n^3)$). Hence, we provide a finite method for computing inertia without computing eigenvalues. This scheme is unstable numerically and hence is primarily of theoretical interest.

Keywords: Lanczos method, Lyapunov Equation, Inertia, Stability, Control

AMS subject classifications. Primary 65F15, Secondary 65G05

1 Introduction

The Lyapunov equation

$$\mathbf{AP} + \mathbf{PA}^* = \mathbf{M} \tag{1}$$

is important in linear system and control theory. Its solution has generated a lot of activity, for example [5, 4, 6, 9, 16, 18], to name but a few. There are well known connections between system theory, particularly linear time invariant systems, and the Lanczos method for reducing a general matrix to tridiagonal form [12, 19, 7, 10, 11, 13, 15]; see also [14], [3]. In this paper, we present a way to use the Lanczos method to solve the Lyapunov equation directly using $\mathcal{O}(n^3)$ floating point operations without computing eigenvalues. This scheme is mainly of theoretical interest since it is numerically unstable.

In addition to solving the Lyapunov equation, a slight modification of this scheme has the ability to compute the inertia of almost any real n by n matrix. Here, the inertia triplet is defined to be the number of eigenvalues in the left halfplane, on the imaginary axis, and in the right halfplane.

This solution scheme relies upon a special construction of the starting vectors for the Lanczos process. Our scheme is essentially an alternative to a method proposed by Schwarz [17] and further studied in [1]; see also [4]. The Schwarz scheme provides a means to reduce a real Hessenberg matrix to a special tridiagonal form introduced below in Section 3. From this form, one can determine the inertia of the Hessenberg matrix from the signs of the off diagonal elements of the tridiagonal matrix. This generalizes the Sturm sequence property typically used to determine the inertia of a symmetric tridiagonal matrix.

The Schwarz method was designed to determine when a given matrix is stable. It is mentioned in several sources but is essentially unknown to the numerical linear algebra community. Perhaps this is because the method as proposed by Schwarz ([17]) relies upon elementary upper triangular 2×2 eliminators that are clearly numerically unstable. There is no opportunity for pivoting and this can even cause the transformations to be undefined under certain conditions.

We develop a specialized Lanczos process to reduce a matrix to Schwarz form. The scheme we propose is also numerically unstable and may break down mathematically. Such breakdowns coincide with serious

*This work was supported in part by the NSF Grant DMS-9972591 and by the NSF Grant CCR-9988393.

†Part of the research reported in this paper was carried out while the second author was on sabbatical leave at the Royal Institute of Technology, Stockholm. The hospitality of KTH is greatfully acknowledged.

‡Dept. of Electrical and Computer Engineering, Rice University, Houston, TX 77005-1892 (aca@ece.rice.edu)

§Dept. of Computational and Applied Mathematics, Rice University, Houston, TX 77005-1892 (sorensen@caam.rice.edu)

breakdowns of the non-symmetric Lanczos process. In spite of the numerical instability, this technique is of interest for two reasons. First, it provides a means to pre-determine the number of eigenvalues of a matrix which lie in any specified vertical strip in the complex plane. Second, it provides a direct solution method for the Lyapunov equation. Neither of these requires computation of eigenvalues and both may be accomplished by referencing the given matrix only through matrix-vector products. For sparse matrices, this means there is potentially an $\mathcal{O}(n^2)$ method for finding inertia, and in general there is an $\mathcal{O}(n^3)$ direct method for solving the Lyapunov equation.

This paper has the following organization. We first (Section 2) introduce a generalized notion of inertia and develop a relation between the generalized inertia of the given matrix $\mathbf{A}$ and the standard notion for inertia of a symmetric solution $\mathbf{P}$ to the Lyapunov equation (1). Then, we introduce the notion of Schwarz form in Section 3 and indicate how it may be used to determine inertia. We also derive a special Lanczos method to compute the Schwarz form, and show how to use this form to determine inertia, and to solve the Lyapunov equation. We also indicate the connections between this scheme and the more classic tests for stability used in system theory and control. Finally, we discuss the computational complexity of this scheme in Section 4 and give some concluding remarks.

2 Inertia and the Lyapunov Equation

In this section we present a self-contained proof by induction, of the inertia result associated with the Lyapunov equation. For this purpose we will assume that in (1) $\mathbf{M}$ is semi-definite, namely $\mathbf{M} = -\mathbf{B}\mathbf{B}^*$. Let $\mathbf{P}$ be a symmetric solution of the *Lyapunov equation*

$$\mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^* + \mathbf{B}\mathbf{B}^* = 0, \quad \text{with} \quad \mathbf{P} = \mathbf{P}^* \in \mathbb{R}^{n\times n} \tag{2}$$

where

$$\mathbf{A} \in \mathbb{R}^{n\times n}, \ \mathbf{B} \in \mathbb{R}^{n\times m} \tag{3}$$

Throughout this discussion we will make the assumption

$$(\mathbf{A}, \mathbf{B}) \text{ controllable}, \tag{4}$$

that is, $\operatorname{rank}[\mathbf{B}, \mathbf{AB}, \mathbf{A}^2\mathbf{B}, \cdots, \mathbf{A}^{n-1}\mathbf{B}] = n$.

Definition 2.1 *Given a square matrix* $\mathbf{A}$*, let the number of eigenvalues in the left halfplane, on the imaginary axis, in the right halfplane be denoted by* $in_-(A)$*,* $in_0(A)$*,* $in_+(A)$ *respectively. The triple* $(in_-(A), in_0(A), in_+(A))$ *is called the* **inertia** *of* $\mathbf{A}$ *and is denoted by* $\text{inertia}(\mathbf{A})$*.*

First we will collect some well known consequences of the above assumptions.

Proposition 2.1 (a) (Popov-Belevich-Hautus) *Assume* $\mathbf{A}$*,* $\mathbf{B}$ *are as in (3). Then* $(\mathbf{A}, \mathbf{B})$ *is controllable if and only if no left eigenvector of* $\mathbf{A}$ *is in the left kernel of* $\mathbf{B}$ *(i.e.* $\mathbf{z}^*\mathbf{A} = \lambda\mathbf{z}^*$ *implies* $\mathbf{z}^*\mathbf{B} \neq \mathbf{0}$*).* **(b)** *Assumptions (3), (2), (4), imply that both* $in_0(\mathbf{A}) = 0$ *and* $in_0(\mathbf{P}) = 0$*, i.e.* $\mathbf{A}$ *has no eigenvalues on the imaginary axis and both matrices are non-singular.*

We intend to show that the inertia of $\mathbf{A}$ can be derived from the inertia of $\mathbf{P}$ and that the inertia of the symmetric matrix $\mathbf{P}$ is readily available as a byproduct of solving the Lyapunov equation. The following results are well known in system theory (see e.g. [2]). We believe the following proof is more concise than others we are aware of. Moreover, it is important for the sequel to establish Proposition 2.2 in a form that will allow us to infer the inertia of a matrix in the Schwarz form to be introduced in the next section.

To continue our discussion it will be convenient to assume that $\mathbf{A}$ is in Schur form, i.e. $\mathbf{A}$ is upper triangular. There is no loss in generality with this assumption as it amounts to a transformation of (2) to an equivalent system using the Schur form basis vectors. Once the system is in Schur form, partition $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{P}$ satisfying (2), compatibly as follows:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} \end{pmatrix}, \ \mathbf{B} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix}, \ \mathbf{P} = \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^* & \mathbf{P}_{22} \end{pmatrix} \tag{5}$$

where $\mathbf{A}_{11}$ and $\mathbf{A}_{22}$ are upper triangular.

Proposition 2.2 *Assume* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{P}$ *satisfy (2), (4) and have been partitioned as in (5). The following statements hold:* **(a)** *The pair* $\mathbf{A}_{22}$, $\mathbf{B}_2$, *is controllable.* **(b)** $in_0(\mathbf{P}_{22}) = 0$, *i.e.* $\mathbf{P}_{22}$ *is non-singular.* **(c)** *The pair* $(\mathbf{A}_{11}, \hat{\mathbf{B}}_1)$ *is controllable, where* $\hat{\mathbf{B}}_1 = \mathbf{B}_1 - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{B}_2$.

Proof. **(a)** Let $\mathbf{z}_2$ be any left eigenvector of $\mathbf{A}_{22}$. Then $\mathbf{z} = [0, \mathbf{z}_2^*]^*$ is a left eigenvector of $\mathbf{A}$ and the PBH condition (part **(a)** of Proposition 2.1) implies $0 \neq \mathbf{z}^*\mathbf{B} = \mathbf{z}_2^*\mathbf{B}_2$. Since this is true for any left eigenvector of $\mathbf{A}_{22}$, the PBH condition also implies the controllability of $(\mathbf{A}_{22}, \mathbf{B}_2)$.

(b) Since $\mathbf{A}_{22}\mathbf{P}_{22} + \mathbf{P}_{22}\mathbf{A}_{22}^* + \mathbf{B}_2\mathbf{B}_2^* = 0$, part **(b)** follows from part **(b)** of proposition 2.1, stated earlier.

(c) As a consequence of **(b)**, the Lyapunov equation (2) can be transformed to

$$\hat{\mathbf{A}}\hat{\mathbf{P}} + \hat{\mathbf{P}}\hat{\mathbf{A}}^* + \hat{\mathbf{B}}\hat{\mathbf{B}}^* = 0 \tag{6}$$

where $\hat{\mathbf{A}} = \mathbf{TAT}^{-1}$, $\hat{\mathbf{B}} = \mathbf{TB}$, $\hat{\mathbf{P}} = \mathbf{TPT}^*$, and

$$\mathbf{T} = \begin{pmatrix} \mathbf{I} & -\mathbf{P}_{12}\mathbf{P}_{22}^{-1} \\ 0 & \mathbf{I} \end{pmatrix},\ \hat{\mathbf{A}} = \begin{pmatrix} \mathbf{A}_{11} & \hat{\mathbf{A}}_{12} \\ 0 & \mathbf{A}_{22} \end{pmatrix},\ \hat{\mathbf{B}} = \begin{pmatrix} \hat{\mathbf{B}}_1 \\ \mathbf{B}_2 \end{pmatrix},\ \hat{\mathbf{P}} = \begin{pmatrix} \hat{\mathbf{P}}_{11} & 0 \\ 0 & \mathbf{P}_{22} \end{pmatrix} \tag{7}$$

with $\hat{\mathbf{A}}_{12} = \mathbf{A}_{12} - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{A}_{22} + \mathbf{A}_{11}\mathbf{P}_{12}\mathbf{P}_{22}^{-1}$, $\hat{\mathbf{B}}_1 = \mathbf{B}_1 - \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{B}_2$, and $\hat{\mathbf{P}}_{11} = \mathbf{P}_{11} + \mathbf{P}_{12}\mathbf{P}_{22}^{-1}\mathbf{P}_{12}^*$. From (6) and (7) follow the three equations:

$$\mathbf{A}_{11}\hat{\mathbf{P}}_{11} + \hat{\mathbf{P}}_{11}\mathbf{A}_{11}^* + \hat{\mathbf{B}}_1\hat{\mathbf{B}}_1^* = 0,\ \mathbf{A}_{22}\mathbf{P}_{22} + \mathbf{P}_{22}\mathbf{A}_{22}^* + \mathbf{B}_2\mathbf{B}_2^* = 0,\ \hat{\mathbf{A}}_{12} = \hat{\mathbf{B}}_1\mathbf{B}_2^*\mathbf{P}_{22}^{-1}. \tag{8}$$

Suppose that there is a left eigenvector $\mathbf{z}_1$ of $\mathbf{A}_{11}$ such that $\mathbf{z}_1^*\hat{\mathbf{B}}_1 = 0$. Then $\mathbf{z}_1^*\hat{\mathbf{A}}_{12} = 0$ and it follows that $\mathbf{z} = [\mathbf{z}_1^*, 0]^*$ is a left eigenvector of $\hat{\mathbf{A}}$ such that $\mathbf{z}^*\hat{\mathbf{B}} = \mathbf{z}_1^*\hat{\mathbf{B}}_1 = 0$ in contradiction of the PBH condition. ■

Now, we are ready to prove the main theorem 2.1, of this section. It is based on lemma 2.1.

Definition 2.2 *A diagonal matrix is called a* **signature** *if its diagonal entries consist only of* 1 *or* -1.

Lemma 2.1 *Let* $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{P}$ *satisfy (2), together with assumptions (3) and (4). If* $\mathbf{A}$ *is in Schur form, then* $\mathbf{P}$ *can be expressed in factored form:* $\mathbf{P} = \mathbf{USU}^*$, *where* $\mathbf{U}$ *is upper triangular and* $\mathbf{S}$ *is a signature matrix.*

Proof. The proof will be given by induction on n, the order of $\mathbf{A}$. The required property clearly holds for $n = 1$. Assume that it holds for Lyapunov equations of order $k < n$, where (4) is satisfied. We will show that the same property must also hold for Lyapunov equations of order n, satisfying (4).

To prove this, we can assume without loss of generality that the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{P}$ (where $\mathbf{A}$ has dimension n) are partitioned as in (5), where the $(1,1)$ block has dimension $k < n$ and the $(2,2)$ block has dimension $n - k < n$. Due to controllability, we may also assume that these matrices are in the form (7) and satisfy the transformed Lyapunov equation (6). By Proposition 2.2, both of the pairs $(\mathbf{A}_{11}, \hat{\mathbf{B}}_1)$ and $(\mathbf{A}_{22}, \mathbf{B}_2)$ are controllable and the induction hypothesis can be applied to each of the two reduced order Lyapunov equations giving: $\hat{\mathbf{P}}_{11} = \mathbf{U}_{11}\mathbf{S}_1\mathbf{U}_{11}^*$ and $\mathbf{P}_{22} = \mathbf{U}_{22}\mathbf{S}_2\mathbf{U}_{22}^*$. Transforming back from equation (6) gives $\mathbf{P} = \mathbf{USU}^*$ with

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ 0 & \mathbf{U}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{P}_{12}\mathbf{P}_{22}^{-1} \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{11} & 0 \\ 0 & \mathbf{U}_{22} \end{bmatrix},$$

and $\mathbf{S} = \begin{bmatrix} \mathbf{S}_1 & 0 \\ 0 & \mathbf{S}_2 \end{bmatrix}$. The induction is thus complete. ■

The next result is a consequence of this lemma.

Theorem 2.1 *Assume that* $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{P}$ *satisfy the Lyapunov equation (2) as well as assumptions (3) and (4). Then*

$$in_-(\mathbf{A}) = in_+(\mathbf{P}) \ \text{ and } \ in_+(\mathbf{A}) = in_-(\mathbf{P}). \tag{9}$$

Proof. Again, the proof will be given by induction on n, the order of $\mathbf{A}$. First, assume that $\mathbf{A}$ is in *Schur form*. Properties (9) clearly hold for $n = 1$. Assume they hold for Lyapunov equations of order $k < n$, satisfying (4). We will show as a consequence that (9) must also hold for Lyapunov equations of order n, satisfying (4).

If we partition the matrices as in the proof of Lemma 2.1, it follows that the Lyapunov equations (8) are satisfied. Each one of these has size less than n and hence the induction hypothesis applies:

$$in_-(\mathbf{A}_{11}) = in_+(\hat{\mathbf{P}}_{11}),\ in_-(\hat{\mathbf{P}}_{11}) = in_+(\mathbf{A}_{11}) \ \text{ and } \ in_-(\mathbf{A}_{22}) = in_+(\mathbf{P}_{22}),\ in_-(\mathbf{P}_{22}) = in_+(\mathbf{A}_{22}).$$

Since $\mathbf{A}$ is in Schur form, there holds $in_-(\mathbf{A}) = in_-(\mathbf{A}_{11}) + in_-(\mathbf{A}_{22})$, $in_+(\mathbf{A}) = in_+(\mathbf{A}_{11}) + in_+(\mathbf{A}_{22})$; due to the structure of $\mathbf{P}$ we have $in_-(\mathbf{P}) = in_-(\hat{\mathbf{P}}_{11}) + in_-(\mathbf{P}_{22})$, $in_+(\mathbf{P}) = in_+(\mathbf{P}_{11}) + in_+(\mathbf{P}_{22})$, completing the induction.

If $\mathbf{A}$ is *not* in Schur form, the upper triangular $\mathbf{U}$ in the considerations above, will be replaced by $\mathbf{QU}$, where $\tilde{\mathbf{A}} = \mathbf{Q}^*\mathbf{AQ}$ is the original matrix (not in Schur form). The solution of the corresponding Lyapunov equation is $\tilde{\mathbf{P}} = (\mathbf{QU})\mathbf{S}(\mathbf{QU})^*$. ■

Remark. The considerations layed out in the proof of the theorem above lead to a UL factorization of the solution $\mathbf{P}$ to the Lyapunov equation. If A is in Schur form, the factorization $\mathbf{P} = \mathbf{USU}^*$ holds, where $\mathbf{U}$ is upper triangular and $\mathbf{S}$ is a signature matrix. The question is, when does the solution $\tilde{\mathbf{P}}$ in the original coordinate system, possess such a factorization?

If the principal minors $\det \tilde{\mathbf{P}}(k:n, k:n)$, $k = 1, \cdots, n$, are different from zero, then the UL factorization of $\tilde{\mathbf{P}}$ exists; let $\tilde{\mathbf{P}} = \bar{\mathbf{U}}\bar{\mathbf{L}}$, where the diagonal entries of $\bar{\mathbf{U}}$ can be chosen to be positive and those of $\bar{\mathbf{L}}$ can be chosen to have the same magnitude as the corresponding entries of $\bar{\mathbf{U}}$. Since $\tilde{\mathbf{P}}$ is symmetric there exists a signature matrix $\tilde{\mathbf{S}}$ such that $(\tilde{\mathbf{S}})^{-1}\bar{\mathbf{L}} = \bar{\mathbf{U}}^*$, and the required factorization follows. It should be noticed that the condition that the minors defined above be different from zero is basis dependent, and cannot always be satisfied. This is the case whenever A has eigenvalues with both positive and negative real parts. Actually it is easy to show in this case, that there always exists a basis such that $\tilde{\mathbf{P}}$ does *not* have an LU factorization. For example if $n = 2$, let the solution $\mathbf{P}_1$ be diagonal; by basis change the transformed solution $\mathbf{P}_2$,

$$\mathbf{P}_1 = \begin{pmatrix} \alpha & 0 \\ 0 & -\beta \end{pmatrix},\ \alpha, \beta > 0 \ \Rightarrow \ \mathbf{P}_2 = \begin{pmatrix} 0 & \sqrt{\alpha\beta} \\ \sqrt{\alpha\beta} & \alpha - \beta \end{pmatrix},$$

does not have an LU factorization. Of course, if $\mathbf{P}$ is positive or negative definite, the result is the Cholesky factorizaton which always exists. ■

The theorem we have just established has some important consequences. We intend to use solutions to specific Lyapunov equations to ascertain the inertia of a matrix $\mathbf{A}$. Theorm 2.1 provides a condition that will determine inertia($\mathbf{A}$) if a symmetric solution to a Lyapunov equation involving $\mathbf{A}$ is available (as it will be in our special case). No apriori assumptions on the spectrum of $\mathbf{A}$ are required.

However, the condition $(\mathbf{A}, \mathbf{B})$ controllable is not sufficient for the existence of a symmetric solution to the Lyapunov equation as the example $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ clearly shows. It is easily seen that if a solution $\mathbf{P}$ to the Lyapunov equation (2) is unique (for any $\mathbf{B}$), then it must be symmetric. Of course, there is a well known condition, not involving controllablility, that implies uniqueness. For the sake of completeness, we shall end this section with a statement of that well known result.

Theorem 2.2 *Assume* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *with* $\sigma(\mathbf{A}) \cap \sigma(-\mathbf{A}^*) = \emptyset$. *Then the Lyapunov equation (2) has a unique solution* $\mathbf{P} = \mathbf{P}^*$ *for any* $\mathbf{B} \in \mathbb{R}^{n\times m}$. *Moreover, in an appropriate basis, this solution can be written in the form* $\mathbf{P} = \mathbf{USU}^*$ *where* $\mathbf{U}$ *is upper triangular and* $\mathbf{S}$ *is a signature matrix.*

The proof of this theorem may be obtained through the construction that leads to the Bartels-Stewart algorithm [6], as modified for the Lyapunov case by Hammarling.

3 Schwarz Form and the Lanczos Process

In this section we introduce the Schwarz form mentioned in Section 1 and relate it to the solution of a special Lyapunov equation that leads to the determination of inertia. We rely upon the main result, Theorem 2.1, of the previous section restricted to the special case where $m = 1$, i.e. $\mathbf{B} = \mathbf{b}$ is a column vector. Our discussion will require the following definitions:

Definition 3.1 (a) *A matrix* $\mathbf{A}$ *is called* **sign-symmetric** *if* $\mathbf{AS} = \mathbf{SA}^*$ *for some signature matrix* $\mathbf{S}$ *and* **skew-symmetric** *if* $\mathbf{A} = -\mathbf{A}^*$. **(b)** *A tridiagonal matrix*

$$\mathbf{T} = \begin{bmatrix} \alpha & \gamma_1 & & & \\ \beta_1 & 0 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 0 & \gamma_{n-1} \\ & & & \beta_{n-1} & 0 \end{bmatrix} \tag{10}$$

is said to be in **pre-Schwarz** *form if* $\alpha \leq 0$, $\beta_j > 0$, $|\gamma_j| = \beta_j$, $j = 1, 2, \cdots, n-1$. $\mathbf{T}$ *is said to be in* **Schwarz form** *if in addition* $\gamma_j = -\beta_j$, $j = 1, 2, \cdots, n-1$.

It is easily seen that if $\mathbf{T}$ is in pre-Schwarz form, one can construct signatures $\mathbf{S}_o, \mathbf{S}_1$ such that $\mathbf{S}_0\mathbf{T}\mathbf{S}_1$ is in Schwarz form. The following lemma provides a useful consequence of this.

Lemma 3.1 *Let* $\mathbf{T}$ *be tridiagonal. If* $\mathbf{S}_0\mathbf{T}\mathbf{S}_1$ *is in Schwarz form for two signatures* $\mathbf{S}_0, \mathbf{S}_1$, *let* $\mathbf{S} = \mathbf{S}_0\mathbf{S}_1$. *Then* $\text{inertia}(\mathbf{S}) = (n_L, n_I, n_R)$ *implies* $\text{inertia}(\mathbf{T}) = (n_R, n_I, n_L)$.

Proof. Suppse $\mathbf{S}_0\mathbf{T}\mathbf{S}_1$ is in Schwarz form as defined above. Then $\hat{\mathbf{T}} = (\mathbf{S}_0\mathbf{T}\mathbf{S}_0)\mathbf{S} + |\alpha|\mathbf{e}_1\mathbf{e}_1^*$, is skew-symmetric so that $\hat{\mathbf{T}} + \hat{\mathbf{T}}^* = \mathbf{0}$. Therefore, $(\mathbf{S}_0\mathbf{T}\mathbf{S}_0)\mathbf{S} + \mathbf{S}(\mathbf{S}_0\mathbf{T}\mathbf{S}_0)^* + 2|\alpha|\mathbf{e}_1\mathbf{e}_1^* = \mathbf{0}$. It is easily seen that the pair $(\mathbf{T}, \mathbf{e}_1)$ is controllable. Hence, the result follows from Theorem 2.1. ∎

We shall utilize this result to determine inertia through the reduction of a general matrix to Schwarz form. Our scheme is to first reduce a given matrix $\mathbf{A}$ to Hessenberg form $\mathbf{H}$ with an Arnoldi process, and then further reduce the Hessenberg matrix to pre-Schwarz form with a nonsymmetric Lanczos process.

The Arnoldi Process: Given a vector $\mathbf{b}$ of unit length, this process produces a sequence of partial factorizations of the form

$$\mathbf{AV}_k = \mathbf{V}_k\mathbf{H}_k + \mathbf{f}_k\mathbf{e}_k^T, \quad k = 1, 2, \ldots, n,$$

with $\mathbf{V}_k\mathbf{e}_1 = \mathbf{b}$, $\mathbf{V}_k^*\mathbf{V}_k = \mathbf{I}_k$, $\mathbf{V}_k^*\mathbf{f}_k = 0$, and $\mathbf{H}_k$ a $k \times k$ upper Hessenberg matrix with non-negative subdiagonal elements. It is well known that $\mathbf{H}_k$ has positive subdiagonal elements at each step $k = 1, 2, \ldots, n$ if and only if the pair $(\mathbf{A}, \mathbf{b})$ is controllable. Moreover, it is easily seen from the conditions on $\mathbf{V}_k$ and $\mathbf{f}_k$ that $\mathbf{f}_n = \mathbf{0}$ and when the pair $(\mathbf{A}, \mathbf{b})$ is controllable this amounts to a unitary similarity transformation of $\mathbf{A}$ to Hessenberg form $\mathbf{H}_n$ which is uniquely determined by the given vector $\mathbf{b}$.

The Lanczos Process: Given vectors $\mathbf{b}$ and $\mathbf{c}$ such that $\mathbf{c}^*\mathbf{b} = 1$, this process produces a sequence of partial factorizations of the form

$$\mathbf{AV}_k = \mathbf{V}_k\mathbf{T}_k + \mathbf{f}_k\mathbf{e}_k^T, \tag{11}$$

$$\mathbf{A}^*\mathbf{W}_k = \mathbf{W}_k\mathbf{T}_k^* + \mathbf{g}_k\mathbf{e}_k^T, \quad k = 1, 2, \ldots, n, \tag{12}$$

with $\mathbf{V}_k\mathbf{e}_1 = \mathbf{b}$, $\mathbf{W}_k\mathbf{e}_1 = \mathbf{c}$, $\mathbf{W}_k^*\mathbf{V}_k = \mathbf{I}_k$, $\mathbf{W}_k^*\mathbf{f}_k = 0$, $\mathbf{V}_k^*\mathbf{g}_k = 0$, and $\mathbf{T}_k$ a $k \times k$ tridiagonal matrix with non-negative subdiagonal elements. We shall specify the normalization $|\mathbf{T}_k| = |\mathbf{T}_k|^T$. If this process completes to n steps, then we have a similarity transformation of $\mathbf{A}$ to tridiagonal form $\mathbf{T}_n$:

$$\mathbf{T}_n = \begin{bmatrix} \tau_1 & \gamma_1 & & & \\ \beta_1 & \tau_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \tau_{n-1} & \gamma_{n-1} \\ & & & \beta_{n-1} & \tau_n \end{bmatrix}. \tag{13}$$

Observe that this matrix will be in pre-Schwarz form if we can construct starting vectors $\mathbf{b}$ and $\mathbf{c}$ that cause $|\tau_j| = 0$, $j = 2, 3, \ldots, n$. It turns out that we can accomplish this when $\mathbf{A}$ is upper Hessenberg and thus we begin with a unitary reduction of $\mathbf{A}$ to Hessenberg form $\mathbf{H}$ and then apply the Lanczos process to reduce $\mathbf{H}$ to pre-Schwarz form using starting vectors $\mathbf{b} = \mathbf{e}_1$ together with a special choice of $\mathbf{c}$.

Consider the application of Lanczos with $\mathbf{H}$ in place of $\mathbf{A}$ and with $\mathbf{V}_k\mathbf{e}_1 = \mathbf{e}_1$ in (11). It is easily shown that $\mathbf{v}_j = \mathbf{V}_k e_j = p_{j-1}(\mathbf{H})\mathbf{e}_1$ where p_{j-1} is a polynomial of degree $j-1$ (in fact p_{j-1} is a multiple of the characteristic polynomial of $\mathbf{T}_{j-1}$). Since $\mathbf{H}$ is upper Hessenberg, this implies that $\mathbf{e}_i^T\mathbf{v}_j = 0$ for $i > j$. Hence,

$$\mathbf{U} \equiv \mathbf{V}_n \text{ is upper triangular and } \mathbf{L} \equiv \mathbf{W}_n \text{ is lower triangular}$$

since $\mathbf{I}_n = \mathbf{W}_n^*\mathbf{V}_n = \mathbf{W}_n^*\mathbf{U}$ implies $\mathbf{W}^* = \mathbf{U}^{-1}$ is upper triangular. We now must find a choice of $\ell_1 = \mathbf{L}\mathbf{e}_1 = \mathbf{c}$ such that $\mathbf{T}$ is in pre-Schwarz form. The following theorem gives the proper choice.

Theorem 3.1 *Let the characteristic polynomial of* $\mathbf{A}$ *be* $\chi_{\mathbf{A}}(\lambda) = \det(\lambda\mathbf{I} - \mathbf{H}) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \alpha_{n-2}\lambda^{n-2} + \alpha_{n-3}\lambda^{n-3} + \cdots + \alpha_1\lambda + \alpha_0$. *Assume* $\alpha_{n-1} \neq 0$. *Then* $\mathbf{T}$ *is in pre-Schwarz form if and only if* ℓ_1 *is of the form*

$$\ell_1^* = \alpha_{n-1}\mathbf{e}_n^*\mathbf{H}^{n-1} + \alpha_{n-3}\mathbf{e}_n^*\mathbf{H}^{n-3} + \alpha_{n-5}\mathbf{e}_n^*\mathbf{H}^{n-5} + \cdots \tag{14}$$

The proof of this theorem will be accomplished in two steps. First, we show that if the Lanczos process results in a tridiagonal $\mathbf{T}$ in pre-Schwarz form, then ℓ_1 must be (up to a scale factor) of the form (14). Then, we verify that if the Lanczos process is applied to $\mathbf{H}$ with this starting vector, it must produce a $\mathbf{T}$ of the required form.

As we have shown, given any starting vector ℓ_1, barring serious breakdown of the process, the Lanczos scheme with $\mathbf{u}_1 = \mathbf{e}_1$ will produce

$$\mathbf{H}\mathbf{U} = \mathbf{U}\mathbf{T}, \quad \mathbf{L}^*\mathbf{H} = \mathbf{T}\mathbf{L}^*, \quad \mathbf{L}^*\mathbf{U} = \mathbf{I}, \tag{15}$$

with $\mathbf{U}$ upper triangular, $\mathbf{L}$ lower triangular, and $\mathbf{T}$ tridiagonal.

Proof of necessity of (14). To see that ℓ_1 must be of the form (14), consider the upper triangular matrix $\mathbf{K}$ and the upper Hessenberg companion matrix $\mathbf{G}$ defined by

$$\mathbf{K} = \begin{bmatrix} \mathbf{e}_n^*\mathbf{H}^{n-1} \\ \mathbf{e}_n^*\mathbf{H}^{n-2} \\ \vdots \\ \mathbf{e}_n^*\mathbf{H} \\ \mathbf{e}_n^* \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} -\alpha_{n-1} & -\alpha_{n-2} & \cdots & -\alpha_1 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & 0 \\ & 1 & \cdots & 0 & 0 \\ & & \ddots & \vdots & \vdots \\ & & & 0 & 0 \\ & & & 1 & 0 \end{bmatrix} \tag{16}$$

The Cayley-Hamilton theorem may be used to validate the relation $\mathbf{K}\mathbf{H} = \mathbf{G}\mathbf{K}$. We note that $\mathbf{K}$ is upper triangular and nonsingular since $\mathbf{H}$ is unreduced upper Hessenberg (i.e. the pair $(\mathbf{H}^*, \mathbf{e}_n)$ is contollable). Thus, the matrix $\mathbf{K}\mathbf{U}$ is nonsingular and upper triangular. Moreover,

$$\mathbf{K}\mathbf{U}\mathbf{T} = \mathbf{K}\mathbf{H}\mathbf{U} = \mathbf{G}\mathbf{K}\mathbf{U}.$$

If we define $\mathbf{R} \equiv (\mathbf{K}\mathbf{U})^{-1} = \mathbf{L}^*\mathbf{K}^{-1}$, then

$$\mathbf{R}\mathbf{G} = \mathbf{T}\mathbf{R}. \tag{17}$$

Moreover,

$$\mathbf{R}\mathbf{K} = \mathbf{L}^* \;\Rightarrow\; \mathbf{r}_1^*\mathbf{K} = \mathbf{e}_1^*\mathbf{L}^* = \ell_1^*,$$

where $\mathbf{r}_1^* = \mathbf{e}_1^*\mathbf{R}$, so that

$$\ell_1^* = [\rho_{1,1}\mathbf{e}_n^*\mathbf{H}^{n-1} + \rho_{1,2}\mathbf{e}_n^*\mathbf{H}^{n-2} + \rho_{1,3}\mathbf{e}_n^*\mathbf{H}^{n-3} + \cdots \rho_{1,n-1}\mathbf{e}_n^*\mathbf{H} + \rho_{1,n}\mathbf{e}_n^*].$$

To complete the proof of necessity, we shall determine the structure of $\mathbf{R}$ that must be imposed if $\mathbf{T}$ is in pre-Schwarz form in (17).

The result below was first given in [8], but without a full proof (only up to $n = 4$); in this same reference the entries of the triangular $\mathbf{T}$ are obtained from the Routh table. For the sake of completeness, we shall prove this result.

Proposition 3.1 *Suppose* $\mathbf{T}$ *is in pre-Schwarz form and that* $\mathbf{R}$ *and* $\mathbf{T}$ *satisfy (17). Then* $\mathbf{R}$ *has a checkerboard pattern with* $\rho_{j-(2t+1),j} = 0$*, for* $t = 0, 1, \cdots$ *and* $j = 1, 2, ..., n$*. Moreover,*

$$[\ \rho_{1,1},\ \rho_{1,2},\ \rho_{1,3},\ \rho_{1,4},\ \rho_{1,5},\ \cdots] = \frac{1}{\alpha_{n-1}}[\ \alpha_{n-1},\ 0,\ \alpha_{n-3},\ 0,\ \alpha_{n-5},\ 0,\ \cdots].$$

Proof. Consider (17) and equate the last column on both sides: $\mathbf{RGe}_n = \mathbf{TRe}_n$. Rearranging terms provides

$$\begin{array}{rcl} -\alpha_0\rho_{1,1} & = & -\alpha_{n-1}\rho_{1,n} + \gamma_1\rho_{2,n} \\ \beta_1\rho_{1,n} & = & -\gamma_2\rho_{3,n} \\ \beta_2\rho_{2,n} & = & -\gamma_3\rho_{4,n} \\ & \vdots & \\ \beta_{n-4}\rho_{n-4,n} & = & -\gamma_{n-3}\rho_{n-2,n} \\ \beta_{n-3}\rho_{n-3,n} & = & -\gamma_{n-2}\rho_{n-1,n} \\ \beta_{n-2}\rho_{n-2,n} & = & -\gamma_{n-1}\rho_{n,n} \\ \beta_{n-1}\rho_{n-1,n} & = & 0 \end{array} \tag{18}$$

The controllability assumption implies that none of the β_j or γ_j are zero. Thus, beginning with $\beta_{n-1}\rho_{n-1,n} = 0$ and following the consequences of the recurrence $\beta_{n-j}\rho_{n-j,n} = -\gamma_{n-j+1}\rho_{n-j-2,n}$ backwards for j odd, implies that $\rho_{n-(2t+1),n} = 0$ for $t = 0, 1, \cdots$. Furthermore, either $\rho_{1,n} = 0$ (n even) or $\rho_{2,n} = 0$ and $\alpha_{n-1}\rho_{1,n} = \alpha_0\rho_{1,1}$ (n odd).

Now, for $j = n-1, n-2, \ldots, 1$ we equate the j-th columns on both sides: $\mathbf{RGe}_j = \mathbf{TRe}_j$ to see

$$\begin{array}{rcl} -\alpha_{n-1}\rho_{1,j} & = & \rho_{1,j+1} - \gamma_1\rho_{2,j} - \alpha_{n-j}\rho_{1,1} \\ \beta_1\rho_{1,j} & = & \rho_{2,j+1} - \gamma_2\rho_{3,j} \\ \beta_2\rho_{2,j} & = & \rho_{3,j+1} - \gamma_3\rho_{4,j} \\ & \vdots & \\ \beta_{j-3}\rho_{j-3,j} & = & \rho_{j-2,j+1} - \gamma_{j-2}\rho_{j-1,j} \\ \beta_{j-2}\rho_{j-2,j} & = & \rho_{j-1,j+1} - \gamma_{j-1}\rho_{j,j} \\ \beta_{j-1}\rho_{j-1,j} & = & \rho_{j,j+1} \\ \beta_j\rho_{j,j} & = & \rho_{j+1,j+1} \end{array} \tag{19}$$

From this we conclude that $\beta_{j-(2t+1)}\rho_{j-(2t+1),j} = \rho_{j+1-(2t+1),j+1} - \gamma_{j-2t}\rho_{j-(2t-1),j}$. Hence $\rho_{j-(2t+1),j} = 0$, $t = 0, 1, 2, \cdots$. Finally, either $\rho_{1,j} = 0$ (j even) or $\rho_{2,j} = 0$ and $\alpha_{n-1}\rho_{1,j} = \alpha_{n-j}\rho_{1,1}$ (j odd).

We have established that $\mathbf{R}$ has the claimed checker board pattern and also that

$$\alpha_{n-1}[\ \rho_{1,1},\ \rho_{1,2},\ \rho_{1,3},\ \rho_{1,4},\ \rho_{1,5},\ \cdots] = [\ \alpha_{n-1},\ 0,\ \alpha_{n-3},\ 0,\ \alpha_{n-5},\ 0,\ \cdots]\rho_{1,1}.$$

Note that $\rho_{1,1} = 1$ may be specified without loss of generality in (17) and the assumption that $\alpha_{n-1} \neq 0$ provides the desired result. ■

Proof of sufficiency of (14). Now that we have established the necessity for ℓ_1 to have the form of (14), let us consider the consequences of applying the Lanczos process with this starting vector for the ℓ-sequence. We wish to show that the diagonal elements of $\mathbf{T}$ in positions $2, 3, \cdots, n$ are all zero.

Proposition 3.2 *Let* τ_i *be the* i*-th diagonal element of* $\mathbf{T}$ *produced by the Lanczos process in (15), where we assume that the starting vector* ℓ_1 *has been chosen to be as specified in (14) and that* $\mathbf{u}_1 = \mathbf{e}_1$*. Then* $\tau_1 = -\alpha_{n-1}$ *and* $\tau_i = 0$ *for* $i = 2, 3, \cdots, n$*.*

Proof. Since $\mathbf{KH} = \mathbf{GK}$, it is sufficient to consider the equivalent Lanczos process applied to the companion matrix $\mathbf{G}$. This will produce $\mathbf{GU} = \mathbf{UT}$ and $\mathbf{RG} = \mathbf{TR}$. Let $\mathbf{r}^* = \mathbf{e}_j^*\mathbf{R}$ and $\mathbf{u}_j = \mathbf{U}\mathbf{e}_j$; then $\mathbf{u}_1 = \mathbf{e}_1$ and

$$\mathbf{r}_1^* = \frac{1}{\alpha_{n-1}}[\ \alpha_{n-1},\ 0,\ \alpha_{n-3},\ 0,\ \alpha_{n-5},\ 0,\ \cdots].$$

It is now straightforward to see that the second row of $\mathbf{R}$ is

$$\mathbf{r}_2^* = \mathbf{r}_1^*(\mathbf{G} + \alpha_{n-1}\mathbf{I}) = [\ 0,\ \hat{\alpha}_{n-2},\ 0,\ \hat{\alpha}_{n-4},\ 0,\ \hat{\alpha}_{n-6},\ 0,\ \cdots],$$

where $\hat{\alpha}_{n-2t} = \frac{\alpha_{n-2t-1}}{\alpha_{n-1}} - \alpha_{n-2t}$ for $t = 1, 2, \cdots$. We also see that $\mathbf{u}_2 = (\mathbf{G} + \alpha_{n-1}\mathbf{I})\mathbf{u}_1 = \mathbf{e}_2$. Now, it is easy to check that $\mathbf{r}_2^*\mathbf{G}\mathbf{u}_2 = 0$ and this will establish $\tau_2 = 0$. These initial vectors set the pattern for the remaining sequence of vectors produced by the process. Namely, if $\tau_k = \tau_{k-1} = \cdots = \tau_3 = \tau_2 = 0$, then

$$\mathbf{r}_{j+1}^* = \mathbf{r}_j^*\mathbf{G} - \gamma_{j-1}\mathbf{r}_{j-1}^* \quad \text{and} \quad \gamma_j\mathbf{u}_{j+1} = \mathbf{G}\mathbf{u}_j - \mathbf{u}_{j-1} \quad \text{for} \quad j = 2, 3, \cdots, k$$

A slightly tedious but straightfoward calculation will show that

$$\mathbf{r}_{j+1}^* = [\ 0,\ \cdots,\ \rho_{j+1,j+1},\ 0,\ \rho_{j+1,j+3},\ 0, \cdots]$$

and that $\mathbf{u}_{j+1} = [\ 0,\ \cdots,\ 0,\ \mu_{j-1,j+1},\ 0,\ \mu_{j+1,j+1}, \cdots,\ 0]^*$. From this we may conclude that

$$\mathbf{r}_{j+1}^*\mathbf{G}\mathbf{u}_{j+1} = [\rho_{j+1,j+1},\ 0,\ \rho_{j+1,j+3},\ 0] \begin{bmatrix} 0 \\ \mu_{j-1,j+1} \\ 0 \\ \mu_{j+1,j+1} \end{bmatrix} = 0$$

and hence that $\tau_{j+1} = 0$. Inductively, we see that this pattern continues. The proof is thus complete. ■

From this discussion, which established the *necessity* and *sufficiency* of (14), we conclude that the Lanczos process with the starting vectors specified as above will result in a special Lanczos process where the diagonal elements are set to zero. Hence the special tridiagonal form is achieved with this scheme, which is summarized in Figure 1.

Computation of (14). There are several options for efficiently producing the coefficients α_j and the special starting vector ℓ_1. One possibility is based upon the classic moment equation

$$\mathcal{T}\mathbf{a} = -\mathbf{t} \tag{20}$$

where $\mathcal{T}_{ij} = \mathbf{e}_n^*\mathbf{H}^{(n+i-j-1)}\mathbf{e}_1$ and $\mathbf{t}(i) = \mathbf{e}_n^*\mathbf{H}^{(i+n-1)}\mathbf{e}_1$. To validate this recursion, we again use the Cayley-Hamilton theorem to see that

$$\begin{aligned} 0 &= \mathbf{e}_n^*\mathbf{H}^{i-1}(\mathbf{H}^n + \alpha_{n-1}\mathbf{H}^{n-1} + \alpha_{n-2}\mathbf{H}^{n-2} + \cdots + \alpha_1\mathbf{H} + \alpha_0\mathbf{I})\mathbf{e}_1 \\ &= \mathbf{e}_n^*\mathbf{H}^{n+i-1}\mathbf{e}_1 + \alpha_{n-1}\mathbf{e}_n^*\mathbf{H}^{n+i-2}\mathbf{e}_1 + + \cdots + \alpha_1\mathbf{e}_n^*\mathbf{H}^i\mathbf{e}_1 + \alpha_0\mathbf{e}_n^*\mathbf{H}^{i-1}\mathbf{e}_1 \\ &= \mathbf{e}_n^*\mathbf{H}^{n+i-1}\mathbf{e}_1 + \alpha_{n-1}\mathbf{e}_n^*\mathbf{H}^{n+i-2}\mathbf{e}_1 + \alpha_{n-2}\mathbf{e}_n^*\mathbf{H}^{n+i-3}\mathbf{e}_1 + \cdots + \alpha_{n-i}\mathbf{e}_n^*\mathbf{H}^{n-1}\mathbf{e}_1 \\ &= \mathbf{e}_n^*\mathbf{H}^{n+i-1}\mathbf{e}_1 + \Sigma_{j=1}^i\alpha_{n-j}\mathbf{e}_n^*\mathbf{H}^{n+i-j-1}\mathbf{e}_1 \end{aligned}$$

where the third equality is derived from the fact that $\mathbf{e}_n^*\mathbf{H}^m\mathbf{e}_1 = 0$ for $m < n-1$. The coefficients of the characteristic polynomial are then given by $\alpha_{n-j} = \mathbf{a}(j)$, for $j = 1, 2, \cdots, n$, where $\mathbf{a}$ is the solution to the lower triangular system (20). This formula provides a recursion for computing these coefficients and constructing the vector ℓ_1. This recursion does not require storage of $\mathcal{T}$. It amounts to nothing more than the standard forward substitution scheme applied to (20) while utilizing the Toeplitz structure of $\mathcal{T}$ in (20).

4 Concluding Remarks

In this paper we presented a new proof of the inertia result associated with the Lyapunov equation and have related it to the non-symmetric Lanczos process. As it turns out, the Schwarz form of a matrix plays

Let $\det(\lambda \mathbf{I} - \mathbf{H}) =$

$$\lambda^n + \alpha_{n-1}\lambda^{n-1} + \alpha_{n-2}\lambda^{n-2} + \alpha_{n-3}\lambda^{n-3} + \cdots + \alpha_1\lambda + \alpha_0;$$

Put $\hat{\ell}_1^* = \alpha_{n-1}\mathbf{e}_n^*\mathbf{H}^{n-1} + \alpha_{n-3}\mathbf{e}_n^*\mathbf{H}^{n-3} + \alpha_{n-5}\mathbf{e}_n^*\mathbf{H}^{n-5} + \cdots;$

Put $\mathbf{u}_1 = \mathbf{e}_1\frac{1}{\sqrt{|\alpha_{n-1}|}};\;\; \gamma_0 = \sqrt{|\hat{\ell}_1^*\mathbf{u}_1|};\;\; \ell_1^* = \frac{1}{\gamma_0}\hat{\ell}_1;$

$$\mathbf{T} = [\alpha_1];\;\; \mathbf{U} = [\mathbf{u}_1];\;\; \mathbf{L} = [\ell_1];$$

Put $\hat{\mathbf{u}}_2 = (\mathbf{H} + \alpha_{n-1}\mathbf{I})\mathbf{u}_1;$

$$\hat{\ell}_2^* = \ell_1^*(\mathbf{H} + \alpha_{n-1}\mathbf{I});$$

$$\gamma = \hat{\ell}_2^*\hat{\mathbf{u}}_2;\;\; \beta_1 = \sqrt{|\gamma|};\;\; \gamma_1 = \beta_1 \cdot sign(\gamma);$$

$$\mathbf{u}_2 \leftarrow \hat{\mathbf{u}}_2/\beta_1;\;\; \ell_2 \leftarrow \hat{\ell}_2/\gamma_1;$$

$$\mathbf{T} \leftarrow \begin{bmatrix} \mathbf{T} & \gamma_1 \\ \beta_1 & 0 \end{bmatrix};\;\; \mathbf{U} \leftarrow [\mathbf{U}, \mathbf{u}_2];\;\; \mathbf{L} \leftarrow [\mathbf{L}, \ell_2];$$

For $j = 2 : n-1$,

$$\hat{\mathbf{u}}_{j+1} = \mathbf{H}\mathbf{u}_j - \gamma_{j-1}\mathbf{u}_{j-1};$$

$$\hat{\ell}_{j+1}^* = \ell_j^*\mathbf{H} - \beta_{j-1}\ell_{j-1}^*;$$

$$\gamma = \hat{\ell}_j^*\hat{\mathbf{u}}_j;\;\; \beta_j = \sqrt{|\gamma|};\;\; \gamma_j = \beta_j \cdot sign(\gamma);$$

$$\mathbf{u}_{j+1} \leftarrow \hat{\mathbf{u}}_{j+1}/\beta_j;$$

$$\ell_{j+1} \leftarrow \hat{\ell}_{j+1}/\gamma_j;$$

$$\mathbf{T} \leftarrow \begin{bmatrix} \mathbf{T} & \mathbf{e}_j\gamma_j \\ \beta_j\mathbf{e}_j^* & 0 \end{bmatrix};\;\; \mathbf{U} \leftarrow [\mathbf{U}, \mathbf{u}_{j+1}];\;\; \mathbf{L} \leftarrow [\mathbf{L}, \ell_{j+1}];$$

End

Figure 1: Special Lanczos Method

a central role in these considerations. This approach provides a method for solving the Lyapunov equation and hence computing the inertia of a matrix, in $\mathcal{O}(n^3)$ operations without computing eigenvalues.

We have not discussed the possibility of serious breakdown of the Lanczos process, nor have we investigated potential causes for such breakdown. The possibility of near breakdowns indicate potential numerical instabilities. Moreover, the moment equations are notoriously ill-conditioned. We have indeed coded the method presented here and tested it. While it does produce the desired result in many cases, the instabilities do indeed manifest themselves in practice.

References

[1] B.D.O. Anderson, E.I. Jury, and M. Mansour, *Schwarz matrix properties for continuous and discrete time systems*, Int. J. Control, **23**: 1-16 (1976).

[2] A.C. Antoulas, *Lectures on the approximation of large-scale dynamical systems*, Dept. of Electrical and Computer Engineering, Rice University (2000).

[3] A.C. Antoulas and R.H. Bishop, *Continued fraction decomposition of linear systems in the space state*, *Systems and Control Letters,* **9**, pp. 43-53 (1987).

[4] S. Barnett, *Polynomials and linear control systems*, Marcel Dekker, New York (1983).

[5] S. Barnett and C. Storey, *The Liapunov matrix equation and Schwarz's form*, IEEE Trans. Automatic Control, **AC-12**: 117-118 (1967).

[6] R.H. Bartels and G.W. Stewart, *Solution of the matrix equation AX+XB=C*, Communications of the ACM, **15**: 820-826 (1972).

[7] D.L. Boley, *Krylov space methods on state-space control models*, Circuits, Systems, and Signal Processing, **13**: 733-758 (1994).

[8] C.F. Chen and H. Chu, *A matrix for evaluating Schwarz's form*, IEEE Trans. Automatic Control, **AC-11**: 303-305 (1966).

[9] E. de Souza and S.P. Bhattacharyya, *Controllability, observability and the solution of* $AX - XB = C$, Linear Alg. Applic., **39**: 167-188 (1981).

[10] R. Freund, *Computing minimal partial realizations via a Lanczos-type algorithm for multiple starting vectors*, Proc. 36th IEEE CDC, pages 4394-4399 (1997).

[11] W.B. Gragg and A. Lindquist, *On the partial realization problem*, Linear Algebra and Its Applications, Special Issue on Linear Systems and Control, **50**: 277-319 (1983).

[12] E.J. Grimme, D.C. Sorensen, and P. Van Dooren, *Model reduction of state space systems via an implicitly restarted Lanczos method*, Numerical Algorithms, **12**: 1-31 (1995).

[13] M.H. Gutknecht, *The Lanczos process and Padé approximation*, Proc. Cornelius Lanczos Intl. Centenary Conference, edited by J.D. Brown et al., pages 61-75, SIAM, Philadelphia (1994).

[14] R.E. Kalman, *On partial realizations, transfer functions, and canonical forms*, Acta Polyt. Scand. Ma., **31**: 9-32 (1979).

[15] B.N. Parlett, *Reduction to tridiagonal form and minimal realizations*, SIAM J. Matrix Anal. Appl., **13**: 567-593 (1992).

[16] Y. Saad, *Numerical solution of large Lyapunov equations*, in Signal Processing, Scattering and Operator Theory, and Numerical Methods, Proc. MTNS-89, **3**: 503-511, Birkhäuser (1990).

[17] H.-R. Schwarz, *Ein Verfahren zur Stabilitätsfrage bei Matrizen-Eigenwertproblemen*, Zeitschrift für Angewandte Mathematik und Physik, **ZAMP VII**: 473-500 (1956).

[18] V. Simoncini, *On the numerical solution of* $AX - XB = C$, BIT, **36**: 814-830 (1996).

[19] P. Van Dooren, *The Lanczos algorithm and Padé approximations*, Short Course, Benelux Meeting on Systems and Control, (1995).