Proceedings of the
15th IFAC Symposium on System Identification
Saint-Malo, France, July 6-8, 2009

Subdiagonal pivot structures and associated canonical forms under state isometries

Bernard Hanzon * **Martine Olivi** ** **Ralf Peeters** ***

* *School of Mathematical Sciences, University College Cork, Cork, Ireland, b.hanzon@ucc.ie*
** *INRIA, Projet APICS, BP 93, 06902 Sophia-Antipolis Cedex, France*
*** *Department of Mathematics, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, Netherlands*

Abstract: Many normalizations in various classes of linear dynamical state-space systems lead to system representations which are determined up to a state isometry. Here we present a new set of techniques to obtain (local) canonical forms under state isometries, using what we call sub-diagonal pivot structures. These techniques lead to a very flexible, straightforward algorithm to put any system into canonical form under state isometries. The parametrization of these canonical forms is discussed for a number of classes, including lossless systems and input-normal stable systems both in discrete time and in continuous time.

1. INTRODUCTION

Consider a linear time-invariant state-space system in discrete time with m inputs and p outputs:

$$\begin{cases} x_{t+1} = Ax_t + Bu_t \\ y_t \ = Cx_t + Du_t \end{cases} \quad t \in \mathbb{Z}$$

or in continuous time

$$\begin{cases} \dot{x}_t = Ax_t + Bu_t \\ y_t = Cx_t + Du_t \end{cases} \quad t \in \mathbb{R}$$

$x_t \in \mathbb{R}^n$ for some nonnegative integer n (the state space dimension), $u_t \in \mathbb{R}^m$ and $y_t \in \mathbb{R}^p$. The matrices A, B, C and D have real-valued entries and are of compatible sizes: $n \times n$, $n \times m$, $p \times n$ and $p \times m$, respectively.

The corresponding transfer matrix of this system is given by $G(z) = D+C(zI_n-A)^{-1}B$, which is a $p\times m$ matrix with rational functions as its entries. From realization theory it follows that, conversely, any $p\times m$ rational matrix function $G(z)$ analytic at infinity can be written in the form

$$G(z) = D + C(zI_n - A)^{-1}B \qquad (1)$$

where (A, B, C, D) is an appropriate quadruple of matrices and n a suitable state space dimension. Such a quadruple with the associated expression (1) is called a *state-space realization* of $G(z)$.

The controllability matrix K and the observability matrix O associated with this system are defined as the block-partitioned matrices

$$K = [B, AB, \ldots, A^{n-1}B], \qquad O = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}. \qquad (2)$$

The system (or its input pair (A, B)) is called controllable if K has full row rank n and the system (or its output pair (C, A)) is called observable if O has full column rank n. Minimality holds iff both controllability and observability hold, which holds iff the McMillan degree of $G(z)$ is equal to n.

Two minimal realizations (A, B, C, D) and $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ of a given function $G(z)$ are always similar: there exists a unique invertible matrix T such that $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}) = (TAT^{-1}, TB, CT^{-1}, D)$. We define a *state isomorphism* as a map :

$$(A, B, C, D) \mapsto (TAT^{-1}, TB, CT^{-1}, D), \quad T \in Gl_n(\mathbb{R})$$

and a *state isometry* as a map

$$(A, B, C, D) \mapsto (QAQ', QB, CQ', D),$$

where Q is an orthogonal matrix $Q'Q = QQ' = I$, and Q' denotes the transpose of Q.

Consider the following well-known characterizations of controllability:

(i) The pair (A, B) is controllable if and only if there exists $T \in Gl_n(\mathbb{R})$ such that $T.K$ contains $e_1, e_2, \ldots, e_n$ as columns, where e_i denotes the ith standard basis vector $e_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0)'$ with the number 1 in the $i-$th position.
(ii) The pair (A, B) is controllable if and only if there exists $T \in Gl_n(\mathbb{R})$ such that $(T.B, T.A.T^{-1})$ contains $e_1, e_2, \ldots, e_n$ as columns in such a way that $T.K$ contains $e_1, e_2, \ldots, e_n$ as columns.

Consider the question whether there are, and if so what are analogous results for *state isometries*? Previously this has been investigated in Hanzon and Ober [1998] and Peeters et al. [2007], where the structure of the controllability matrix plays a central role. Here we reconsider the question in a somewhat different way, where we stress the structure of the pair (A, B). This turns out to lead to much simpler answers.

Reasons for considering state isometries instead of state isomorphisms are:

- Application of orthogonal group elements is in general numerically more stable than application of elements of the general linear group $Gl_n(\mathbb{R})$.
- State isometries appear in the theory of input-normal forms, output-normal forms, "equal Gramians" balanced forms (where the Gramians do *not* have to be diagonal, just equal), balanced forms for lossless systems, and other normal forms for various classes of linear systems (some examples are given in section 3).

2. PIVOT STRUCTURES IN INPUT PAIRS

Definition 2.1. Let n be a positive integer. Consider a vector $v = (v(1), v(2), \ldots, v(n))' \in \mathbf{R}^n$. The vector v is called a *pivot vector* with pivot at position k, $1 \le k \le n$, or pivot-k vector for short, if $v(k) > 0$ and if $v(j)$ with $j > k$ are all zero.

Definition 2.2. An $n \times r$ matrix M, $r \ge n$ is said to have a *full pivot structure* $J = \{j_1, j_2, \ldots, j_n\}$ if for each $k \in \{1, 2, \ldots, n\}$ it holds that column j_k of M is a pivot-k vector.

Example 2.1.

$$M = \begin{bmatrix} * & * & * & * & * & * & + & * \\ + & * & * & * & * & * & 0 & * \\ 0 & * & * & * & + & * & 0 & * \\ 0 & * & + & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 & + & 0 & * \end{bmatrix}$$

where $*$ denotes an arbitrary number and $+$ denotes a (strictly) positive number, has a full pivot structure $J = \{7, 1, 5, 3, 6\}$.

Definition 2.3. Consider the partitioned matrix $[B|A]$ in $\mathbb{R}^{n\times(m+n)}$. We say this has a *sub-diagonal pivot structure* if

(i) $[B|A]$ has a full pivot structure
(ii) the prescribed pivot columns of A have the property that a column with pivot at position k has column number $p_k < k$ in A (hence column number $j_k = m + p_k < m + k$ in $[B|A]$), for each $k \in \{1, 2, \ldots, n\}$.

Example 2.2.

$$[B|A] = \left[\begin{array}{cc|cccc} + & * & * & * & * & * \\ 0 & * & + & * & * & * \\ 0 & + & 0 & * & * & * \\ 0 & 0 & 0 & + & * & * \end{array}\right]$$

has a sub-diagonal pivot structure.

Example 2.3.

$$[B|A] = \left[\begin{array}{cc|cccc} + & * & * & * & * & * \\ 0 & * & + & * & * & * \\ 0 & + & 0 & * & * & * \\ 0 & 0 & 0 & * & + & * \end{array}\right]$$

has a sub-diagonal pivot structure.

Example 2.4.

$$[B|A] = \left[\begin{array}{cc|cccc} + & * & * & * & * & * \\ 0 & * & * & * & + & * \\ 0 & + & * & * & 0 & * \\ 0 & 0 & * & * & 0 & + \end{array}\right]$$

does not have a sub-diagonal pivot structure.

Theorem 1. If $[B|A]$ has a sub-diagonal pivot structure then the pair (A, B) is controllable.

Proof. Recall Hautus (or PBH) criterion (Hautus [1969]): the pair (A, B) is controllable if and only if $\nexists(w, \lambda) \in \mathbb{C}^n \times \mathbb{C}$ such that $w^*B = 0$ and $w^*A = \lambda w^*, w \neq 0$. Now suppose that $w \in \mathbb{C}^n$ is such that $w^*B = 0$ and there exists $\lambda \in \mathbb{C}$ such that $w^*A = \lambda w^*$. Suppose further that $w(i) \neq 0$ and $w(j) = 0$ for all $1 \le j < i$. Then, one of the following situations occurs

(1) B contains the pivot-i vector:
$$w^*B = 0 \Leftrightarrow w(i) = 0,$$
which yields a contradiction.
(2) A contains the pivot-i vector, say in column j, $j < i$:
$$w^*A = \lambda w^* \Leftrightarrow w^*Ae_j = \lambda w^*e_j = 0,$$
and since Ae_j is the pivot-i vector, this again implies $w(i) = 0$ and we get a contradiction.

So by induction, $w = 0$.

Theorem 2. For any given pivot structure which is not sub-diagonal, we can find an example $[B|A]$ for which (A, B) is not controllable.

Proof. We distinguish two cases, depending on whether there is a pivot on the diagonal of the matrix A or not. Firstly, if there is a pivot on the diagonal position (i, i) of the matrix A then consider the case where all pivots are taken to be equal to 1 and all other entries of $[B|A]$ are taken to be zero. Consider $w = e_i$, the i-th standard basis vector. Clearly $w'B = 0$ because B does not contain the i-th pivot vector. Furthermore $w'A = e_i' = w'$ because in its $i-$th row A has only one non-zero entry, namely the value 1 on the diagonal position (i, i) of A. Therefore according to the Hautus test, (A, B) is an uncontrollable pair.

Secondly, suppose there is a super-diagonal pivot. Let the super-diagonal pivot with the smallest row index be in position (i, j) in the matrix A, (hence in position $(i, m + j)$ in the matrix $[B|A]$) where $j > i$. Let the $j-$th pivot be in column $m + k$ in $[B|A]$, where $k \in \{-m + 1, -m + 2, \ldots, n\}$. Let $[\tilde{B}|\tilde{A}]$ denote the matrix that is obtained by taking all pivots equal to 1 and all other entries equal to zero. Now form the matrix

$$[B|A] = [\tilde{B}|\tilde{A}] - E_{i,m+k} + E_{i,m+i}$$

where for any (r, s) with $1 \le r \le n$, $1 \le s \le m + n$, $E_{r,s}$ denotes the $n \times (m + n)$ matrix with all entries zero except for the entry (r, s), which is 1. Note that $[B|A]$ has the same pivot structure as $[\tilde{B}|\tilde{A}]$, because in column $m + k$ one has the $j-$th pivot and $j > i$, so the subtraction of $E_{i,m+k}$ does not change the pivot structure; and in column $m + i$ there is no pivot on a super-diagonal position, because otherwise i would not be the smallest row-number with a super-diagonal pivot. Therefore column i either contains a sub-diagonal pivot or no prescribed pivot at all. In both cases addition of $E_{i,m+i}$ does not alter the pivot structure.

Now note that $w = e_i + e_j \in \mathbb{R}^n$ has the property that $w'[B|A] = [0|w']$. This follows from the facts that $e_i'\tilde{B} = 0$, $e_i'\tilde{A} = e_j'$, so $e_i'[\tilde{B}|\tilde{A}] = e_{m+j}'$ and $e_j'[\tilde{B}|\tilde{A}] = e_{m+k}'$, $e_i'(-E_{i,m+k} + E_{i,m+i}) = -e_{m+k}' + e_{m+i}'$. So $w \neq 0$ is a left eigenvector of A in the left kernel of B, hence (A, B) is uncontrollable according to the Hautus criterion.

3. CANONICAL FORMS UNDER STATE ISOMETRIES

We give some examples in which a normalization of the system by a state isomorphism leaves us the freedom of applying an orthogonal transformation (i.e. a state isometry).

Example 3.1. Let

$$K_n = [B|AB|\dots|A^{n-1}B] \in \mathbb{R}^{n\times nm}$$

denote the finite controllability matrix of (A,B,C,D). Suppose that it has rank n. We can find a nonsingular square matrix T such that $(TK_n)(TK_n)' = I_n$. Then $(\tilde{A},\tilde{B},\tilde{C},\tilde{D}) := (TAT^{-1}, TB, CT^{-1}, D)$ is normalized in the sense that its finite controllability matrix has orthonormal rows. Now note that this property is kept even if we apply a state isometry to the system: $(Q\tilde{A}Q', Q\tilde{B}, \tilde{C}Q', \tilde{D})$ has finite controllability matrix with orthonormal rows as well, for any orthogonal matrix Q.

Example 3.2. One can do the analogous normalization for the finite observability matrix of the system.

Example 3.3. If we have a stable discrete time system, we can define the controllability Gramian of the system in the usual way:

$$W_c = K_\infty K_\infty' = \sum_{k=0}^{\infty} A^k BB'(A')^k.$$

This Gramian solves the Lyapunov-Stein equation

$$W_c - AW_cA' = BB'.$$

We can normalize the system to be input-normal, i.e. such that $W_c = I$. If (A,B,C,D) is input-normal so is (QAQ', QB, CQ', D) for any orthogonal Q.

Example 3.4. One can do the analogous normalization using the observability Gramian W_o. It satisfies

$$W_o - A'W_oA = C'C.$$

The resulting system is output-normal if $W_o = I$. Again there is the freedom of applying a state isometry without losing this property.

Example 3.5. Assume the eigenvalues of A are in the open left half-plane. Then we can define the continuous time controllability Gramian W_c. It satisfies the Lyapunov equation $AW_c + W_cA' = -BB'$. The system is continuous-time input-normal if $W_c = I_n$. Similarly, the system is continuous-time output-normal if $W_o = I_n$ where $A'W_o + W_oA = -C'C$. Again the freedom of applying an orthogonal transformation remains.

Example 3.6. Another well-known normalization is obtained by requiring the Gramians to be equal: $W_c = W_o$ (this can be done both in the discrete time case and in the continuous time case). This can be called Gramian-balanced (or balanced in the sense of Helmke and Moore [1993]). Again this normalization is possible for any stable system and leaves the freedom of applying a state isometry. Effect of a state isometry is the same on both Gramians: $W_c \mapsto QW_cQ'$ and, using $(Q^{-1})' = Q$, $W_o \mapsto QW_oQ'$.

Example 3.7. (A,B,C,D) is called Frobenius-balanced (or Euclidean balanced) if

$$\min_{T\in Gl_n(\mathbb{R})} \left(\|\tilde{A}\|_F^2 + \|\tilde{B}\|_F^2 + \|\tilde{C}\|_F^2 + \|\tilde{D}\|_F^2 \right)$$

is attained at $T = I$ where

$$\tilde{A} = TAT^{-1}, \tilde{B} = TB, \tilde{C} = CT^{-1}, \tilde{D} = D$$

and $\|M\|_F^2 = \mathrm{Tr}(MM')$. This normalization is again well-defined up to a state isometry (cf the book of Helmke and Moore [1993]).

Example 3.8. The so-called LQG-input normal, bounded-real input normal and positive-real input normal forms described in Hanzon and Ober [1998] again leave the freedom of applying a state isometry.

In all these cases it makes sense to look for canonical forms under state isometry. We now want to show that the sub-diagonal pivot structures correspond to local canonical forms under state isometry, and these local canonical forms are covering all cases.

Theorem 3. Suppose that $[B|A]$ and $[QB|QAQ']$ have the same sub-diagonal pivot structure $J = \{j_1, j_2, \dots, j_n\}$ and that Q is orthogonal. Then $Q = I$.

Proof. First note that the pivot-1 column must be a column of B (it cannot lie in A) and of QB, so $j_1 \le m$. Therefore the first column of Q is of the form λe_1, $\lambda > 0$, hence must be e_1; therefore Q can be partitioned as

$$Q = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q_{n-1} & \\ 0 & & & \end{pmatrix}.$$

Now assume (induction hypothesis) we have shown that Q is of the form

$$Q = \begin{pmatrix} I_i & 0 \\ 0 & Q_{n-i} \end{pmatrix}, \ i \in \{1, \dots, n-1\}$$

If the pivot-$(i+1)$ vector is in B we can use the same argument to conclude that the induction hypothesis holds for $i+1$. If the pivot-$(i+1)$ vector is in A then: partition A as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \ A_{11} : i\times i; A_{21} : (n-i)\times i.$$

We get

$$QAQ' = \begin{pmatrix} I_i & 0 \\ 0 & Q_{n-i} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I_i & 0 \\ 0 & Q_{n-i}' \end{pmatrix} =$$

$$\begin{pmatrix} A_{11} & A_{12}Q_{n-i}' \\ Q_{n-i}A_{21} & Q_{n-i}A_{22}Q_{n-i}' \end{pmatrix}$$

The $(i+1)$st pivot lies in the first row of A_{21} and in the first row of $Q_{n-i}A_{21}$ hence the first column of Q_{n-i} is e_1. So

$$Q = \begin{pmatrix} I_{i+1} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q_{n-i-1} & \\ 0 & & & \end{pmatrix}$$

The theorem now follows by induction.

To show that every controllable pair (A,B) can be mapped to a sub-diagonal pivot structure by an appropriate state isometry we will present a recursive algorithm "which cannot get stuck, if controllability holds" and which is rather insensitive to small perturbations. We will use the well-known fact that for any vector $x \ne 0$ in $\mathbb{R}^{n-i}$ there exists an orthogonal matrix Q_{n-i} such that

$$Q_{n-i}\, x = \begin{pmatrix} \|x\| \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

First Step: (A, B) is controllable so $B \neq 0$. It follows that B has at least one non-zero column. Choose one such column. Heuristically one can take the *largest* column. Find $Q_n,\ Q_n Q_n' = I_n$ which maps this column to a vector of the form $\lambda e_1,\ \lambda > 0$.
Recursion Step: Suppose the pivot$-k$ column j_k has been determined for $k = 1, 2, \dots, i$, in a sub-diagonal manner ($j_k < m + k$). Partition $B = \begin{pmatrix} \hat{B}(i) \\ B(n-i) \end{pmatrix}$ where $\hat{B}(i)$ is $i \times m$; and $A = \begin{pmatrix} \hat{A}(i,i) & * \\ A(n-i,i) & * \end{pmatrix}$ where $\hat{A}(i,i)$ is $i \times i$.
Note that

1. $[\hat{B}(i)|\hat{A}(i,i)]$ has a full pivot structure.
2. $[B(n-i)|A(n-i,i)]$ has at least i zero columns.
3. $[B(n-i)|A(n-i,i)] \neq 0$; this can be seen as follows: if $[B(n-i)|A(n-i,i)] = 0$ then
$$B = \begin{pmatrix} \hat{B}(i) \\ 0 \end{pmatrix}, A = \begin{pmatrix} \hat{A}(i,i) & * \\ 0 & * \end{pmatrix}$$
hence the controllability matrix can be partitioned as
$$K = \begin{pmatrix} \hat{B}(i) & \hat{A}(i,i)\hat{B}(i) & \hat{A}(i,i)^2\hat{B}(i) & \dots \\ 0 & 0 & 0 & \dots \end{pmatrix}$$
which has rank $i < n$. This contradicts the assumption that (A, B) is controllable, q.e.d.
4. Application of a state isometry of the form
$$Q = \begin{pmatrix} I_i & 0 \\ 0 & Q_{n-i} \end{pmatrix}$$
transforms $[B(n-i)|A(n-i,i)]$ into
$$[Q_{n-i}B(n-i)|Q_{n-i}A(n-i,i)]$$

Choose a non-zero column from $[B(n-i)|A(n-i,i)]$. Heuristically one can take the *largest* column. Find $Q_{n-i},\ Q_{n-i}Q_{n-i}' = I_{n-i}$ which maps this column to a vector of the form $\lambda e_1 \in \mathbb{R}^{n-i},\ \lambda > 0$. Then the corresponding vector (in $\mathbb{R}^n$) in $[B|A]$ is a pivot-$(i+1)$ vector.

Applying the recursion step for each $i = 1, 2, \dots, n-1$, we obtain a sub-diagonal pivot structure for $[B|A]$. So we can construct a state isometry that brings $[B|A]$ into a sub-diagonal pivot structure. It is clear that the same sub-diagonal pivot structure can still be used under sufficiently small perturbations (namely if the length of the chosen (non-zero) vectors in each step does not reduce to zero).

Remarks

- This also gives a *controllability test*: If (A, B) is controllable a sub-diagonal pivot structure is obtained. If (A, B) is *not* controllable an orthogonal Q is found such that
$$(QAQ', QB) = \left(\begin{pmatrix} \hat{A}(i,i) & * \\ 0 & * \end{pmatrix}, \begin{pmatrix} \hat{B}(i) \\ 0 \end{pmatrix} \right)$$
where $[\hat{B}(i)|\hat{A}(i,i)]$ has a sub-diagonal pivot structure. When applied to input-normal discrete-time stable systems, $[B|A]$ has orthonormal rows. In that case the size of the pivots is less than or equal to 1, so one may have a good idea whether a pivot is "numerically zero."
- *Model reduction by truncation.* If $[B|A]$ is in sub-diagonal pivot form and we truncate the last $n-i$ rows and columns, we again have a controllable pair in sub-diagonal pivot form.
- The staircase pivot structures (see Peeters et al. [2007]) form a sub-class of sub-diagonal pivot structures. A staircase pivot structure has the nice property that it corresponds to a pivot structure in the controllability matrix. However the recursive algorithm is *not* guaranteed to work for the staircase pivot structures. See section 6.

4. SUBDIAGONAL PIVOT STRUCTURE AND SCHUR PARAMETRIZATION OF DISCRETE-TIME LOSSLESS SYSTEMS

A system is called discrete-time *lossless* if it is stable and its $m \times m$ transfer matrix $G(z)$ is unitary for all complex z with $|z| = 1$. It is well-known (cf., e.g., Proposition 3.2 in Hanzon et al. [2006] and the references given there) that the *realization matrix*, i.e. the block-partitioned matrix

$$R = \begin{bmatrix} D & C \\ B & A \end{bmatrix} \tag{3}$$

associated with (1), is a real balanced minimal realization matrix of a lossless system if and only if R is an orthogonal matrix and A is asymptotically stable. It then holds that $W_c = W_o = I_n$. For further background on lossless systems, see e.g. Genin et al. [1983]. Note that there is a one-to-one correspondence between McMillan degree n lossless functions $G(z)$, up to a left orthogonal factor, and controllable input-normal pairs $[B|A]$ up to a state isometry.

In Hanzon et al. [2006] an atlas of overlapping balanced (local) canonical forms for lossless discrete-time systems of order n is presented: balanced state space realizations are constructed recursively, in line with the tangential Schur algorithm. Each such balanced canonical form is characterized by a fixed sequence of n numbers $w_k, |w_k| < 1, k = 1, \dots, n$, called the *interpolation points*, and a fixed sequence of n unit vectors $u_k \in \mathbb{R}^m, \|u_k\| = 1, k = 1, \dots, n$, called the *direction vectors* (which are not to be confused with the input signal applied to a system). Here we will consider the case $w_k = 0, k = 1, \dots, n$ hence each balanced canonical form that we consider is determined by the choice of direction vectors. Each such balanced canonical form is then parameterized by an $m \times m$ orthogonal matrix D_0 and a sequence of n vectors $v_k, \|v_k\| < 1, k = 1, \dots, n$ which are called the *Schur parameter vectors*.

In fact the orthogonal realization matrix can be written as a *product of matrices* of size $(m+n) \times (m+n)$:

$$R = \Gamma_n \Gamma_{n-1} \cdots \Gamma_1 \Gamma_0 \Delta_1^T \Delta_2^T \cdots \Delta_n^T, \tag{4}$$

where for $k = 1, \dots, n$:

$$\Gamma_k = \begin{bmatrix} I_{n-k} & 0 & 0 \\ 0 & V_k & 0 \\ 0 & 0 & I_{k-1} \end{bmatrix},$$

$$\Delta_k = \begin{bmatrix} I_{n-k} & 0 & 0 \\ 0 & U_k & 0 \\ 0 & 0 & I_{k-1} \end{bmatrix}$$

with an $(m+1)\times(m+1)$ orthogonal matrix block V_k given by

$$V_k = \begin{bmatrix} v_k & I_m - (1-\sqrt{1-\|v_k\|^2}\frac{v_k v_k^T}{\|v_k\|^2} \\ \sqrt{1-\|v_k\|^2} & -v_k^T \end{bmatrix},$$

and an $(m+1)\times(m+1)$ orthogonal matrix block U_k given by

$$U_k = \begin{bmatrix} u_k & I_m - u_k u_k^T \\ 0 & u_k^T \end{bmatrix}$$

and furthermore

$$\Gamma_0 = \begin{bmatrix} I_n & 0 \\ 0 & D_0 \end{bmatrix}.$$

Note that here we consider the *real* case with *real* direction vectors and *real* Schur parameter vectors. Note further that $\Gamma_0,\ldots,\Gamma_n$ and $\Delta_1,\ldots,\Delta_n$ are all orthogonal matrices. It is important to note and not too difficult to see that the orthogonal matrix product

$$\Gamma = \Gamma_n\Gamma_{n-1}\cdots\Gamma_1\Gamma_0 \tag{5}$$

in fact forms a *positive m-upper Hessenberg matrix*, i.e. an $(m+n)\times(m+n)$ matrix of which the m-th sub-diagonal only has positive entries and of which the last $n-1$ sub-diagonals are all zero. It also follows almost directly that if the direction vectors $u_1,\ldots,u_n$ are taken to be standard basis vectors, then the matrix product

$$\Delta^T = \Delta_1^T\Delta_2^T\cdots\Delta_n^T \tag{6}$$

yields a permutation matrix. Hence in that case the balanced realization matrix R is obtained as a column permutation of an orthogonal positive m-upper Hessenberg matrix. More precisely, we have:

Theorem 4. Let R be given by (4) and let the direction vectors be standard basis vectors:
for $k=1,\ldots,n$, $u_{n-k+1}=e_{i(k)}$, $i(k)\in\{1,\ldots,m\}$.
Then the sub-matrix $[B|A]$ has a sub-diagonal pivot structure. This pivot structure is completely determined by the sequence $i(1),i(2),\ldots,i(n)$ as follows: Let $k\in\{1,2,\ldots,n\}$.

(i) If for all $j<k$, $i(j)\neq i(k)$, then $j_k=i(k)$ and pivot-k is in B;
(ii) otherwise let $l=\sup\{j|\ j<k,\ i(j)=i(k)\}$, then $j_k=l+m$ and pivot-k is in the first $k-1$ columns of A.

Proof. Indeed, $R=\Gamma\Delta^T$ in which Γ, given by (5), is a positive m-upper Hessenberg matrix and Δ^T the permutation matrix (6). The sub-matrix $[B|A]$ is obtained from $H=[0\ I_n]\,\Gamma$, which possesses the full pivot structure $J=\{1,2,\ldots,n\}$, by applying the permutation matrix Δ^T. For any $j\in\{1,2,\ldots,n\}$, post-multiplication by Δ^T_{n-j+1} only acts on the columns j to $j+m$ as follows (where (i),(ii),(iii) apply simultaneously):
(i) column j is moved into column position $j+i(j)-1$
(ii) column $j+i(j)$ is moved into column position $j+m$
(iii) columns $j+1,j+2,\ldots j+i(j)-1$ and $j+i(j)+1,j+i(j)+2,j+m$ are moved one position to the left, into column positions $j,j+1,\ldots j+i(j)-2$ and $j+i(j),j+i(j)+1,j+m-1$.

Now consider $k\in\{1,2,\ldots,n\}$. The $k-$th column of H with a pivot at position k is only affected by the last k permutation matrices Δ^T_{n-k+1}, $\Delta^T_{n-k+2},\cdots,\Delta^T_n$. In particular, by considering the special case $k=1$, it follows that only the matrix Δ_n^T acts on the first column of H (with pivot-1) which is moved into column $i(1)$ of B. More generally, column k of H ends up in the first $m+k-1$ columns of $[B|A]$, which ensures that $[B|A]$ does have a sub-diagonal pivot structure.

The progression of column k of H with pivot at position k can be monitored as follows:
For $j=k$: Post-multiplication by Δ^T_{n-k+1} moves column k into column $k-1+i(k)$.
For each $j=k-1,k-2,\ldots$ as long as $i(j)\neq i(k)$, post-multiplication by Δ^T_{n-j+1} moves the pivot-k vector one position to the left from column position $j+i(k)$ into column position $j-1+i(k)$. Now, one of the following possibilities occurs:
(i) for all $1\le j<k$, $i(j)\neq i(k)$. Then column k of H is ultimately moved into column $i(k)$ of B which inherits a pivot in position k.
(ii) there exists l, $l=\sup\{j|k>j\ge 1; i(j)=i(k)\}$, then eventually post-multiplication by $\Delta^T_{n-(l+1)+1}$ moves the pivot-k vector from column position $l+i(l)$ into column position $l+m$ where it remains and the pivot-k vector ends up in column l of A.

Conversely, we have

Theorem 5. Let $J=(j_1,j_2,\ldots,j_n)$ be a given sub-diagonal pivot structure for $n\times(m+n)$ matrices. For each $k=1,2,\ldots,n$, choose the direction vector $u_{n+1-k}=e_{i(k)}$ in the Schur algorithm by induction : $i(1):=j_1$, and for $k=2,\ldots,n$

(i) if $j_k\le m$, then $i(k):=j_k$,
(ii) if $m<j_k\le m+n$ define $p_k:=j_k-m<k$ (sub-diagonal structure); then $i(k):=i(p_k)$.

For any choice of the Schur parameter vectors $v_1,v_2,\ldots,v_n$ (all of length <1) and for any choice of the orthogonal matrix D_0, consider the $(m+n)\times(m+n)$ orthogonal realization matrix R given by (4). Then, the sub-matrix $[B|A]$ possesses the sub-diagonal pivot structure J.

Remark. If a given lossless system allows a sub-diagonal pivot structure J then the previous result shows how to choose the corresponding direction vectors. Once these are known one can construct the corresponding Schur parameter vectors of the lossless system (cf Hanzon et al. [2006]). This implies that we have an explicit parametrization of the local canonical form of lossless systems that allow a given sub-diagonal pivot structure.

5. SUB-DIAGONAL PIVOT STRUCTURE AND CONTINUOUS-TIME LOSSLESS SYSTEMS

In the remark at the end of section 4 in Hanzon and Ober [1998] on stable all-pass systems (i.e. on lossless systems) it was concluded that one obtains an atlas of local balanced canonical forms for lossless systems using the particular pivot structures presented in that paper. These pivot structures are staircase pivot structures hence are sub-diagonal. The atlas as well as the parametrizations presented there for each of the charts can be generalized straightforwardly to the case where one allows all possible sub-diagonal pivot structures. The parametrization of a local balanced canonical form associated to a sub-diagonal pivot structure $J = (j_1, j_2, \ldots, j_n)$ is obtained as follows: Firstly let $[B|\tilde{A}]$ be an $n \times (m+n)$ (real) matrix with sub-diagonal pivot structure J and with skew-symmetric matrix $\tilde{A}$, i.e. $\tilde{A} + \tilde{A}^T = 0$. The set of all such matrices can be parameterized by $\mathbf{R}_+^n \times \mathbf{R}^{(m-1)n}$ where the positive numbers are describing the pivots and the other numbers are describing the remaining free entries in B and in the strictly lower triangular part of the skew symmetric matrix $\tilde{A}$. Secondly let V be the unique upper triangular matrix satisfying $V + V^T = -BB^T$. Thirdly choose D to be an arbitrary $m \times m$ orthogonal matrix and let $C = -DB^T$. Finally let $A = \tilde{A} + V$. Then (A, B, C, D) is the balanced realization with the sub-diagonal pivot structure J, of a lossless system. Note that $A + A^T = -BB^T = -C^T C$ showing that the state-space system is balanced. In this way one obtains a chart for each sub-diagonal pivot structure J; varying over all sub-diagonal pivot structures one obtains an atlas covering the manifold of McMillan degree n continuous-time lossless systems.

6. ON THE RELATION OF SUB-DIAGONAL PIVOT STRUCTURES WITH STAIRCASE FORMS

As mentioned previously, the atlas studied in Peeters et al. [2007] provides overlapping balanced canonical forms with a particular sub-diagonal pivot structure : the so-called staircase forms. In these forms, the matrix $[B|A]$ has a particular sub-diagonal pivot structure: if the columns of B contain p_B pivots, the remaining $p_A = n - p_B$ pivots have to be located in the first p_A columns of A with increasing pivot positions. This distribution of pivots implies that the associated controllability matrix contains a pivot at position k for $k = 1, 2, \ldots, n$ so the controllability matrix has a full pivot structure.

The following representation of a sub-diagonal pivot structure $J = (j_1, j_2, \ldots, j_n)$ allows to clarify the connection. An $m \times n$ matrix $Y = (y_{i,j})$ can be associated with the sub-diagonal pivot structure. Its entries are all zeros except n entries corresponding to the pivot positions $1, 2, \ldots, n$. This matrix is constructed by induction :
Pivot-1 is in the j_1-th column of B and we put $Y_{j_1,1} = 1$
Assume that the pivot positions $1, 2, \ldots, k-1$ have been displayed, then
(i) either pivot-k is in the j_k-th column of B, and we put $Y_{j_k,1} = k$
(ii) or pivot-k is in the p_k-th column of A. Since $[B|A]$ has a sub-diagonal structure, then $p_k \le k-1$ has already been displayed: $Y_{i,j} = p_k$ for some i and j. Then we put $Y_{i,j+1} = k$.
In the case of a staircase form, this matrix has a special structure and specifies the pivot structure of the controllability matrix. In view of its relationship with nice selections, it is called a numbered Young diagram (see [Peeters et al., 2007, section 4]).

The recursive algorithm presented in the previous section provides a chart selection algorithm, that was lacking for the staircase forms. To be more precise: it is known that, if (A, B) is controllable, then, if one knows in advance which choices of pivots in B, can be made, one can, in order to obtain a staircase form, again apply the recursive algorithm; in that case one will have to choose the *left-most* possible pivot in A whenever a pivot is not chosen from B at any step in the recursive algorithm. The difficulty here is of course that usually one will *not* know in advance which choices of pivots from B lead to a staircase form. If one makes a wrong choice for the pivots of B in the recursive algorithm, then at some stage in the recursive algorithm in which a pivot, say pivot i, from A needs to be chosen in column j of A, the relevant column vector $(a_{i,j}, a_{i+1,j}, \ldots, a_{n,j})'$ will be equal to the zero vector in $\mathbb{R}^{n-i+1}$. In that case one can still continue the recursive algorithm to obtain a sub-diagonal pivot structure.

If at each stage one takes either a pivot in B or one takes the left-most possible pivot in A, then the corresponding controllability matrix will have a pivot structure! Using that one can use the numbered Young diagram methods of Peeters et al. [2007] to determine a choice of pivots in B that correspond with a staircase form. This can then be used in a second round to obtain a staircase form using the recursive algorithm. This should still work under effects of (sufficiently small) round-off error or more general, when one decides to ignore (sufficiently) small perturbations.

7. CONCLUSION

Usage of sub-diagonal pivot structures and associated algorithms can lead to new flexible canonical forms, for which parametrizations are available in several cases, including discrete-time and continuous-time lossless systems. The results can also be used to construct controllability tests and procedures for model reduction by truncation. As only orthogonal transformations are used, procedures will be numerically stable. Generalizations to the complex-valued case are straightforward.

REFERENCES

Y. Genin, P. Van Dooren, T. Kailath, J.-M. Delosme, and M. Morf. On Σ-lossless transfer functions and related questions. *Linear Algebra and its Applications*, 50:251–275, 1983.

B. Hanzon and R.J. Ober. Overlapping block-balanced canonical forms for various classes of linear systems. *Linear Algebra and its Applications*, 281:171–225, 1998.

B. Hanzon, M. Olivi, and R.L.M. Peeters. Balanced realizations of discrete-time stable all-pass systems and the tangential Schur algorithm. *Linear Algebra and its Applications*, 418:793–820, 2006. A preliminary version is available as Technical Report RR-5111, (2004), INRIA, Sophia-Antipolis, France.

M.L.J. Hautus, Controllability and observability conditions of linear autonomous systems. *Indagationes Mathematicae*, 31:443–448, 1969.

U. Helmke and J. Moore. *Optimization and Dynamical Systems*. Springer-Verlag, 1993.

R. Peeters, B. Hanzon, and M. Olivi. Canonical lossless state-space systems: staircase forms and the schur algorithm. *Linear Algebra and its Applications*, 425:404–433, 2007.