NAG 2-189
DRA / AMES

A Simple Algorithm for Computing Canonical Forms

H. Ford, L.R. Hunt, Renjeng Su

Abstract

It is well known that all linear time-invariant controllable systems can be transformed to Brunovsky canonical form by a transformation consisting only of coordinate changes and linear feedback. However, the actual procedures for doing this have tended to be overly complex. The technique introduced here is envisioned as an on-line procedure and is inspired by George Meyer's tangent model for nonlinear systems. The process utilizes Meyer's block triangular form as an intermediate step in going to Brunovsky form. The method also involves orthogonal matrices, thus eliminating the need for the computation of matrix inverses. In addition, the Kronecker indices can be computed as a by-product of this transformation so it is not necessary to know them in advance.

(NASA-CR-177105) A SIMPLE ALGORITHM FOR COMPUTING CANONICAL FORMS (Texas Technological Univ.) 27 p CSCL 09B

N86-29548

Unclas
G3/61 43239

A Simple Algorithm for Computing Canonical Forms

H. Ford*, L.R. Hunt*, and Renjeng Su**

I. Introduction

In his work at NASA Ames Research Center, George Meyer is applying the theory of transformations of nonlinear systems to linear systems in order to design automatic flight controllers for vertical and short take off aircraft [1], [2], [3], [4], [5], [6]. In these articles he introduces a linear system (called the tangent model) which approximates the nonlinear system, and he mentions the importance of taking the tangent model to Brunovsky [7] canonical form. Thus we feel that an on-line procedure for transforming a controllable linear system to Brunovsky form is highly desirable. Additionally in [8] the authors together with George Meyer present a modification of the tangent model in which the procedure introduced here plays a central role in the linearization

*Research supported by NASA Ames Research Center under grant NAG2-189 and the Joint Services Electronics Program under ONR contract N0014-76-C1136.
**Research supported by NASA Ames Research Center under grant NAG2-203 and the Joint Services Electronics Program under ONR contract N0014-76-C1136.

process itself. This process involves a Taylor Series expansion using lie derivatives.

Our computation of the transformation proceeds in two steps:

1) An orthogonal coordinate change is used to move the linear system to Meyer's block triangular form (see [5]).

2) Once we are in block triangular form, the process becomes trivial and formally involves "Lie differentiation" of certain coordinate functions. The on-line procedure for doing this and its application to automatic flight control are given in [5].

In step 1) it is not necessary to calculate a matrix inverse and no systems of linear equations need be solved. As a by-product of step 1) we find the Kronecker indices of the system.

Understanding the fact that the block triangular form is a natural intermediate step in transforming to Brunovsky form is our main contribution in this paper. We have computer programs to carry out the entire process.

In section II, we describe how to transform a single control linear system to a string of integrators. In section III, we show how to generalize the results of section II to a multi-control system. Detailed algorithms for transforming linear systems to Brunovsky form are given in section IV, including pseudocode programs. Results achieved using a three control system and concluding comments are given in section V. We want to emphasize the simplicity and ease of implementation of the algorithm.

Though the explanation of why it works may seem cumbersome, the algorithm is very straightforward.

II. Single Control Case

Definition 2.1. An n dimensional single control system $\dot{x} = f(x,u)$ is called block triangular if $\dot{x}_i$ is a function only of $x_1, x_2, \ldots, x_{i+1}$ for $i = 1, \ldots, n$ where $x_{n+1} = u$ (see [5]).

For a linear single control system $\dot{x} = Ax + bu$, this is equivalent to saying that the square matrix $H = \begin{bmatrix} A & b \\ 0 & 0 \end{bmatrix}$ is a lower Hessenberg matrix [9], that is, all elements above the first super diagonal are zero. Notice that all elements on the first super diagonal must be nonzero if we are to have a controllable system. The outstanding characteristic of the block triangular system is that if we start with x_1, it is necessary to take n derivatives of x_1 with respect to t before reaching the control.

Definition 2.2. An n dimensional single control system $\dot{x} = f(x,u)$ is called a string of integrators if $\dot{x}_i = x_{i+1}$ for $i = 1, \ldots, n$ where $x_{n+1} = u$.

Once we have a block triangular linear system $\dot{x} = Ax + bu$, we can transform it to a string of integrators $\dot{y} = \tilde{A}y + \tilde{b}v$ by simply letting $y_1 = x_1$, $y_2 = \dot{y}_1$, $\ldots, y_n = \dot{y}_{n-1}$ and $v = \dot{y}_n$.

Theorem 2.1. The above transformation from block triangular form to a string of integrators consists of only coordinate changes and linear feedback.

Proof: We know that $y_1 = x_1$ as mentioned above.

Assuming $y_i = \sum_{j=1}^{i} \alpha_j x_j$ for some $i \le n - 1$, we have

$$y_{i+1} = \dot{y}_i = \sum_{j=1}^{i} \alpha_j \dot{x}_j = \sum_{j=1}^{i} \alpha_j \sum_{k=1}^{j+1} a_{j,k} x_k = \sum_{k=1}^{i+1} \beta_k x_k$$

Thus the state variables y are functions only of the state variables x. The new control $v = \dot{y}_n = \sum_{k=1}^{n} \beta_k \dot{x}_k$

$$= \sum_{k=1}^{n-1} \beta_k \dot{x}_k + \beta_n \dot{x}_n = \sum_{k=1}^{n-1} \beta_k \sum_{\alpha=1}^{k+1} a_{k,j} x_j + \beta_n \left(\sum_{j=1}^{n} a_{n,j} x_j + bu \right)$$

$$= \sum_{\ell=1}^{n} \gamma_\ell x_\ell + \gamma_{n+1} u,$$

and our feedback is linear. $\square$

If we start with a system $\dot{x} = Ax + bu$ which is not block triangular, we first form the augmented matrix $H = \begin{bmatrix} A & b \\ 0 & 0 \end{bmatrix}$. From Stewart [9] it is known that there exists an orthogonal transformation $C = \begin{bmatrix} \tilde{C} & 0 \\ 0 & 1 \end{bmatrix}$ such that CHC^T is lower Hessenberg.

$$CHC^T = \begin{bmatrix} \tilde{C} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & b \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{C}^T & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \tilde{C}A\tilde{C}^T & \tilde{C}b \\ 0 & 0 \end{bmatrix}$$

Thus the transformation $y = \tilde{C}x$ yields a block triangular system $\dot{y} = \tilde{C}A\tilde{C}^T y + \tilde{C}bu$.

Notice $\begin{bmatrix} y \\ u \end{bmatrix} = C \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} \tilde{C} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} \tilde{C}x \\ u \end{bmatrix}$ is a transformation involving coordinate changes only. Once we have a block triangular system, we go to a string of integrators as before.

III. Multi-Control Case

For a multi-control system, the situation is slightly more complicated than for the single control case. Below is the usual definition of Kronecker indices [7].

<u>Definition 3.1</u>. Let $\dot{x} = Ax + Bu$ be a time-invariant n dimensional linear control system with m controls.

Let $r_o = \text{rank } B$

$$r_j = \text{rank}\,\{B, AB, \ldots, A^j B\} - \text{rank}\{B, AB, \ldots, A^{j-1}B\}$$

We define the Kronecker indices κ_i as the number of r_j's that are $\geq i$. Notice $\kappa_1 \geq \kappa_2 \geq \cdots \geq \kappa_m$ and $\sum_{j=1}^{m} \kappa_j = n$.

<u>Definition 3.2.</u> By the <u>Brunovsky</u> [7] <u>canonical</u> <u>form</u>, we mean a linear system $\dot{y} = \hat{A}y + \hat{B}v$ such that $\hat{A}$ equals

$$\begin{array}{c} \\ \kappa_1\left\{\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array}\right. \\ \\ \kappa_2\left\{\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array}\right. \\ \\ \\ \kappa_m\left\{\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array}\right. \end{array}
\left[\begin{array}{c|c|c|c}
0\,1\,0 \ldots\ldots 0 & 0\,0 \ldots\ldots 0 & & 0\,0 \ldots\ldots 0 \\
0\,0\,1\,0 \ldots 0 & 0\,0 \ldots\ldots 0 & & \cdot\ \cdot \qquad \cdot \\
\cdot\ \cdot \quad \cdot & \cdot\ \cdot \qquad \cdot & \cdots & \cdot\ \cdot \qquad \cdot \\
\qquad \cdot & & & \\
\qquad\quad \cdot & & & \\
\qquad\qquad 1 & & & \\
0\,0 \ldots\ldots 0 & 0\,0 \ldots\ldots 0 & & 0\,0 \ldots\ldots 0 \\
\hdashline
0\,0 \ldots\ldots 0 & 0\,1\,0 \ldots 0 & & 0\,0 \ldots\ldots 0 \\
\cdot\ \cdot \qquad \cdot & 0\,0\,1\,0 \ldots 0 & & \cdot\ \cdot \qquad \cdot \\
\cdot\ \cdot \qquad \cdot & \cdot \quad \cdot & \cdots & \cdot\ \cdot \qquad \cdot \\
 & \qquad \cdot & & \\
 & \qquad\quad \cdot & & \\
 & \qquad\qquad 1 & & \\
0\,0 \ldots\ldots 0 & 0\,0 \ldots\ldots 0 & & 0\,0 \ldots\ldots 0 \\
\hdashline
\vdots & \vdots & \ddots & \vdots \\
\hdashline
0\,0 \ldots\ldots 0 & 0\,0 \ldots\ldots 0 & & 0\,1\,0 \ldots 0 \\
\cdot\ \cdot \qquad \cdot & \cdot\ \cdot \qquad \cdot & & 0\,0\,1\,0 \ldots 0 \\
\cdot\ \cdot \qquad \cdot & \cdot\ \cdot \qquad \cdot & \cdots & \qquad \cdot \\
 & & & \qquad\quad \cdot \\
 & & & \qquad\qquad \cdot \\
 & & & \qquad\qquad 1 \\
0\,0 \ldots\ldots 0 & 0\,0 \ldots\ldots 0 & & 0\,0 \ldots\ldots 0
\end{array}\right],$$

and $\hat{B}$ equals

$$
\begin{array}{c}
\kappa_1 \left\{ \begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array} \right. \\
\kappa_2 \left\{ \begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array} \right. \\
\\
\kappa_m \left\{ \begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array} \right.
\end{array}
\left[\begin{array}{cccccc}
0 & 0 & \cdot & \cdot & \cdot & 0 \\
0 & 0 & \cdot & \cdot & \cdot & 0 \\
\vdots & \vdots & & & & \vdots \\
1 & 0 & \cdot & \cdot & \cdot & 0 \\
\hdashline
0 & 0 & \cdot & \cdot & \cdot & 0 \\
0 & 0 & \cdot & \cdot & \cdot & 0 \\
\vdots & \vdots & & & & \vdots \\
0 & 1 & 0 & \cdot & \cdot & 0 \\
\hdashline
 & & & \vdots & & \\
\hdashline
0 & 0 & \cdot & \cdot & \cdot & 0 \\
0 & 0 & & & & 0 \\
\vdots & \vdots & & & & \\
0 & 0 & \cdot & \cdot & \cdot & 1
\end{array}\right] .
$$

That is, the Brunovsky form consists of m strings of integrators whose respective lengths are the Kronecker indices. For instance a ten dimensional system with three controls and Kronecker indices 5,3, and 2 will have Brunovsky form

$$(3.1)\quad \begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \\ \dot{y}_4 \\ \dot{y}_5 \\ \dot{y}_6 \\ \dot{y}_7 \\ \dot{y}_8 \\ \dot{y}_9 \\ \dot{y}_{10} \end{bmatrix} = \left[\begin{array}{ccccc|ccc|cc} 0&1&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \\ \hline 0&0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0&0&0 \\ \hline 0&0&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&0&0&0&0 \end{array}\right] \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \end{bmatrix} + \begin{bmatrix} 0&0&0 \\ 0&0&0 \\ 0&0&0 \\ 0&0&0 \\ 1&0&0 \\ 0&0&0 \\ 0&0&0 \\ 0&1&0 \\ 0&0&0 \\ 0&0&1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Notice that for a system with only one control, the Brunovsky form is a string of integrators.

Though not necessary for the implementation of our algorithm, we shall for theoretical reasons want to be able to identify individual Kronecker indices with individual controls. For that reason we shall introduce the following alternate definition of Kronecker indices.

<u>Definition 3.3.</u> Let $\dot{x} = Ax + Bu$ be an n dimensional linear system with m controls and with $B = [b_1,\ldots,b_m]$. Consider the vectors $b_m, b_{m-1},\ldots,b_1, Ab_m,\ldots,Ab_1, A^2b_m,\ldots,A^2b_1,\ldots$ until we come to a vector dependent on the previous vectors, call it $A^{\kappa_i}b_i$. Then κ_i is the

Kronecker index associated with the control u_i. If we continue in this manner we will get m Kronecker indices whose sum is n. This definition will give us the same Kronecker indices as before. However, the subscripts will now associate each Kronecker index with a control rather than ordering the Kronecker indices (we can obviously renumber our controls so that $\kappa_1 \geq \kappa_2 \geq \kappa_3 \geq \ldots \geq \kappa_m$). This definition will be very useful in showing that the block triangular form of a system contains all information on the Kronecker indices.

What do we mean by the block triangular form of a linear system with m controls ? The most obvious choice would be a system consisting of m block triangular systems, each in one control. For instance, a ten dimensional block triangular system with three controls and Kronecker indices 5,3, and 2 might be of the form

$$(3.2)\quad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \\ \dot{x}_8 \\ \dot{x}_9 \\ \dot{x}_{10} \end{bmatrix} = \begin{bmatrix} X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & X & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & X & X & 0 & 0 & 0 & 0 & 0 \\ X & X & X & X & X & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & X & X & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & X & X & X & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & X & X & X & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X & X \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X & X \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ X & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & 0 \\ 0 & 0 & X \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

Here each of the 0's represents zero. Each of the rightmost X's represents a nonzero element. The other X's may or may not be zero. We emphasize that 5 derivatives of x_1 must be taken to reach a control, 3 derivatives of x_6, and 2 derivatives of x_9, precisely the Kronecker indices.

We found it convenient to rename the state variables, ordering them in terms of their distance from the controls. By distance, we mean the number of derivatives of a variable we must take to reach the controls. Renaming $x_1 \to x_1$, $x_2 \to x_2$, $x_3 \to x_3$, $x_6 \to x_4$, $x_4 \to x_5$, $x_7 \to x_6$, $x_9 \to x_7$, $x_5 \to x_8$, $x_8 \to x_9$, and $x_{10} \to x_{10}$. We then get

$$(3.3)\qquad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \\ \dot{x}_8 \\ \dot{x}_9 \\ \dot{x}_{10} \end{bmatrix} = \begin{bmatrix} X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & 0 & X & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & X & 0 & X & 0 & 0 & 0 & 0 \\ X & X & X & 0 & X & 0 & 0 & X & 0 & 0 \\ 0 & 0 & 0 & X & 0 & X & 0 & 0 & X & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & X & 0 & 0 & X \\ X & X & X & 0 & X & 0 & 0 & X & 0 & 0 \\ 0 & 0 & 0 & X & 0 & X & 0 & 0 & X & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & X & 0 & 0 & X \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ X & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & X \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

We must take 5 derivatives of x_1 to reach the controls, 3 derivatives of x_4, and 2 derivatives of x_7. Remarkably, the only thing which affects the number of derivatives necessary to reach the controls is the rightmost nonzero element in each row.
For instance, $\dot{x}_4$ is a function of x_6 but not a function of x_j for $j > 6$ nor u_i for $1 \leq i \leq 3$. $\dot{x}_6$ is a function of x_9 but not a function of x_{10} nor u_i for $1 \leq i \leq 3$. This implies $\ddot{x}_4$ is a function of x_9

but not a function of x_{10} nor u_i for $1 \leq i \leq 3$. $\dot{x}_9$ is a function of u_2 but not u_3. This implies $\dddot{x}_4$ is a function of u_2 but not u_3. Notice in determining that it takes three derivatives of x_4 before we reach the controls, all we used was knowledge of the rightmost nonzero elements. Thus a block triangular system with ten dimensions, three controls, and Kronecker indices 5,3, and 2 could have the form

$$(3.4.)\quad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \\ \dot{x}_8 \\ \dot{x}_9 \\ \dot{x}_{10} \end{bmatrix} = \begin{bmatrix} X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & X & X & 0 & 0 & 0 & 0 & 0 \\ X & X & X & X & X & X & 0 & 0 & 0 & 0 \\ X & X & X & X & X & X & X & X & 0 & 0 \\ X & X & X & X & X & X & X & X & X & 0 \\ X & X & X & X & X & X & X & X & X & X \\ X & X & X & X & X & X & X & X & X & X \\ X & X & X & X & X & X & X & X & X & X \\ X & X & X & X & X & X & X & X & X & X \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ x & 0 & 0 \\ X & X & 0 \\ X & X & X \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

In each row, the X furthest to the right is nonzero. The other X's may or may not be zero. Notice it still takes 5 derivatives of x_1 to reach the controls, 3 derivatives of x_4, and 2 derivatives of x_7.

The above discussion motivates our definition of block triangular for a linear system with several controls.

<u>Definition 3.4.</u> Let $\dot{x} = Ax + Bu$ be an n dimensional system with m controls. We say the system is block triangular if the square matrix $H = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}$ is a generalization of a lower Hessenberg matrix that

we simply call generalized lower Hessenberg and which we now define. First, all elements of H above the mth super diagonal are zero. Second, if H(I,J), with I < n, is a zero element with no nonzero element to its right in row I, then all elements above H(I,J) in the same super diagonal must also be zero. For instance,

$$(3.5.)\quad H = \begin{bmatrix} X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & X & X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ X & X & X & X & X & X & X & X & 0 & 0 & 0 & 0 & 0 \\ X & X & X & X & X & X & X & X & X & 0 & 0 & 0 & 0 \\ X & X & X & X & X & X & X & X & X & X & 0 & 0 & 0 \\ X & X & X & X & X & X & X & X & X & X & X & 0 & 0 \\ X & X & X & X & X & X & X & X & X & X & X & X & 0 \\ X & X & X & X & X & X & X & X & X & X & X & X & X \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where the rightmost X's are nonzero, represents a ten dimensional block triangular system with three controls and Kronecker indices 5,3, and 2. Notice that since H(4,7) = 0, we have H(3,6) = H(2,5) = H(1,4) = 0. Similarly, H(2,4) = 0 implies H(1,3) = 0.

If $\dot{x} = Ax + Bu$ is an n dimensional system with m controls which is not in block triangular form, we can easily make it block triangular in the following way. Form the augmented matrix $H = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}$, and find an orthogonal matrix $C = \begin{bmatrix} \tilde{C} & 0 \\ 0 & I \end{bmatrix}$ (where I is the m dimensional identity) such that CHC^T is generalized lower Hessenberg. We will explain in the next section precisely how to find C. This matrix C satisfies $CHC^T = \begin{bmatrix} \tilde{C} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{C}^T & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} \tilde{C}A\tilde{C}^T & CB \\ 0 & 0 \end{bmatrix}$

Thus the transformation $y = \tilde{C}x$ yields a block triangular system $\dot{y} = \tilde{C}A\tilde{C}^T y + \tilde{C}Bu$. As in the single control case, this transformation is just a change of coordinates on our state space.

Definition 3.5. By zero pattern we shall mean the pattern of the rightmost nonzero elements of the matrix $H = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}$ of a system $\dot{x} = Ax + Bu$ in block triangular form. That is, for each block triangular form there is an n tuple of integers $(\ell_1, \ell_2, \ldots, \ell_n)$ so that ℓ_i is the column number of rightmost nonzero element in the i^{th} row.

Theorem 3.1. Let $\dot{x} = Ax + Bu$ be an n dimensional system with m controls. There is a one to one correspondence between the possible ordered sets of Kronecker indices and the possible zero patterns of the block triangular form. That is, a given set of Kronecker indices with given associations with the controls (see def. 3.3) results in a distinct and unique zero pattern.

Comment: This implies we can retrieve the Kronecker indices of a system knowing only the zero pattern of the block triangular form.

Proof: Consider the original system $\dot{x} = Ax + Bu$.

Let $\tilde{H} = \begin{bmatrix} \tilde{C} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{C}^T & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} \tilde{A} & \tilde{B} \\ 0 & 0 \end{bmatrix}$ be the block triangular representation. Notice that $\dot{y} = \tilde{A}y + \tilde{B}u$ has the same Kronecker indices as the original system $\dot{x} = Ax + Bu$. To see this consider that

$$[\tilde{b}_m,\ldots,\tilde{b}_1,\tilde{A}\tilde{b}_m,\ldots,\tilde{A}\tilde{b}_1,\ldots]=[\tilde{C}b_m,\ldots,\tilde{C}b_1,\tilde{C}A\tilde{C}^T\tilde{C}b_m,\ldots,\tilde{C}A\tilde{C}^T\tilde{C}b_1,\ldots]$$
$$=[\tilde{C}b_m,\ldots,\tilde{C}b_1,\tilde{C}Ab_m,\ldots,\tilde{C}Ab_1,\ldots] = \tilde{C}[b_m,\ldots,b_1,Ab_m,\ldots,Ab_1,\ldots].$$

Since $\tilde{C}$ is nonsingular, the rank of any selection of columns from $[\tilde{b}_m,\ldots,\tilde{b}_1,\tilde{A}\tilde{b}_m,\ldots,\tilde{A}\tilde{b}_1,\ldots]$ is the same as the rank of the corresponding selection of columns from $[b_m,\ldots,b_1,Ab_m,\ldots,Ab_1,\ldots]$. Therefore, we can assume without loss of generality that the original system is already in block triangular form.

The rest of the proof depends on two basic principles from linear algebra:

1. When we multiply a matrix times a vector, the product is a linear combination of the columns of the matrix.
2. If we have a collection of n nonzero vectors $v_1,\ldots,v_n$, then the linear combination $\sum_{i=1}^{n}\alpha_i v_i$, with $\alpha_n \neq 0$, is linearly independent of $v_1,\ldots,v_{n-1}$ if and only if v_n is linearly independent of $v_1,\ldots,v_{n-1}$.

Using the first principle, we see that $Ab_m = a_m b_{n,m}$ where a_n is the n^{th} column of A, and $b_{n,m}$ is the n^{th} entry in b_m. For an integer i_0, $0 \leq i_0 \leq m-1$, $Ab_{m-i_0} = \sum_{i_1=0}^{i_0} a_{n-i_1}\, b_{n-i_1,m-i_0}$

$$A^k b_{m-i_0} = \sum_{i_1=0}^{i_0} a_{n-i_1}\, b_{n-i_1,\, m-i_0}$$

$$A^k b_{m-i_0} = \sum_{i_1=0}^{i_0}\cdots\sum_{i_2=0}^{m+i_1}\cdots\sum_{i_k=0}^{m+i_{k-1}} a_{n-i_k} a_{n-i_k,n-i_{k-1}}\cdots a_{n-i_2,n-i_1} b_{n-i_1,m-i_0}$$

Thus $A^k b_{m-i_0}$ is a linear combination of columns of A.

Using the second principle, Ab_m is linearly independent of $b_m,\ldots,b_1$ if and only if a_n is linearly independent of $b_m,\ldots,b_1$.

Thus for purposes of checking independence Ab_m can be represented by a_n. Similarly $A^k b_{m-i_0}$ can be represented by the leftmost column of A in the linear combination which has a nonzero coefficient. Zero coefficients are caused by previous vectors being dependent. Thus $A^k b_{m-i_0}$ can be represented by column $a_{n-[(k-1)m + i_0]+h}$ where h is the number of vectors already found to be dependent on previous vectors. To see this, consider $A^k b_{m-i_0} = A(A^{k-1} b_{m-i_0})$. The index of the column representing $A^k b_{m-i_0}$ will be the first (from the top) nonzero element of $A^{k-1} b_{m-i_0}$. Since we move up a row for every independent vector, the first nonzero element of $A^{k-1} b_{m-i_0}$ has index $n-[(k-1)m + i_0] + h$.

Thus $A^k b_{m-i_0}$ can be dependent on the previous vectors in two ways. One way is that $A^{k-1} b_{m-i_0}$ was dependent on vectors previous to it. The second way is that the column $a_{n-[(k-1)m + i_0] + h}$ is dependent on columns to the right of it. In this second way $A^k b_{m-i_0}$ is the lowest power of A times b_{m-i_0} which is dependent on previous vectors. Thus there is a one to one correspondence between the ordered sets of Kronecker indices (ordered by def 3.3) and the zero patterns of the block triangular form. □

Once we have the block triangular matrix H, how do we retrieve the lead variables (to be defined) for the Brunovsky form? To see how this is done and to better understand block triangular systems, it is useful to think in terms of derivative levels.

<u>Definition 3.6.</u> Let $\dot{x} = Ax + Bu$ be a control system in block

triangular form. A state variable x_i is said to be on the j^{th} derivative level if it takes j derivatives of x_i to reach the controls. For instance in the block triangular system illustrated in equations (3.4) and (3.5), x_8, x_9, and x_{10} are on the first derivative level. x_5, x_6, and x_7 are on the second derivative level, x_3 and x_4 are on the third level, and x_2 is on the fourth level. Lastly the variable x_1 is on the fifth derivative level.

Definition 3.7. Let $\dot{x} = Ax + Bu$ be a control system in block triangular form. Let x_i be a state variable on the j^{th} derivative level. Then $\dot{x}_i$ is said to be a lead variable if it cannot be reached by taking the derivative of a state variable on the (j+1)th level.

For all block triangular systems, x_1 is a lead variable. For the particular system in equations 3.4, x_4 and x_7 are also lead variables. Notice that x_4 is on the third level; x_2 is the only variable on the fourth level but its derivative is not a function of x_4. Also x_7 is on the second level; x_3 and x_4 are on the third level, but their derivatives do not involve x_7.

Theorem 3.2. Let $\dot{x} = Ax + Bu$ be a control system in block triangular form. The derivative levels of the lead variables are precisely the Kronecker indices.

Proof: This theorem is really just a restatement of Theorem 3.1. Notice that a lead variable occurs because of the inability to reach the variable from a higher level which in turn is caused by a column dependency in A. □

Once we have the lead variables in the block triangular form, we simply let them be the lead variables in the Brunovsky form.

For instance for the block triangular system $\dot{x} = Ax + Bu$ of equation 3.4, we let

$$
\begin{array}{lll}
y_1 = x_1 & y_6 = x_4 & y_9 = x_7 \\
y_2 = \dot{y}_1 & y_7 = \dot{y}_6 & y_{10} = \dot{y}_9 \\
y_3 = \dot{y}_2 & y_8 = \dot{y}_7 & v_3 = \dot{y}_{10} \\
y_4 = \dot{y}_3 & v_2 = \dot{y}_8 & \\
y_5 = \dot{y}_4 & & \\
v_1 = \dot{y}_5 & &
\end{array}
$$

We know the respective lengths of the integral strings because we know the derivative levels of x_1, x_4, and x_7.

Again, the essential characteristic of the block triangular system is that the controls do not appear "too soon". Thus the state variables in Brunovsky form are functions only of the state variables in the original system (that is, not functions of the controls).

Theorem 3.3. The above transformation from block triangular form to Brunovsky canonical form consists of only coordinate changes and linear feedback.

Proof: As in the single control case, the derivative of x_i is a linear combination of x_j's except when x_i is on the first derivative level. In that case $\dot{x}_i$ is a linear combination of the x_j's and u_ℓ's. Subsequent derivatives of the state variables are also linear combinations of the appropriate variables. Let x_i be a lead variable and let k be the derivative level of x_i. For $j < k$, the j^{th} derivative of x_i with respect to $t, x_i^{(j)}$, is a linear combination of x_ℓ's and $x_i^{(k)}$ is a linear combination of the x_ℓ's and u_ℓ's. Since the lengths of the strings in the Brunovsky form are determined by the derivative levels of the lead variables, our proof is complete. □

IV. Algorithms

The actual algorithm for finding a matrix C to transform a matrix H to block triangular form is a generalization of the method found in Stewart [9] for transforming a matrix to Hessenberg form.

We do this by placing the appropriate zeros in H one column at a time starting with the rightmost column. For each column, we multiply on the left and right by what is called an elementary reflector or Householder tranformation u_k (see Stewart [9]). Each of the u_k is of the form $u_k = I - \frac{2vv^T}{v^Tv}$ where I is the identity matrix and v is a vector. To illustrate this process, suppose we want to zero out the first k-1 elements in some column and at the same time insure that the k^{th} element is nonzero, suppose $\begin{bmatrix} v_1 \\ \vdots \\ v_k \\ ? \end{bmatrix}$ is the column in question then we multiply

$$\left\{ I - \frac{1}{\|v\|(v_k + \|v\|)} \begin{bmatrix} v_1 \\ \vdots \\ v_{k-1} \\ v_k + \|v\| \\ 0 \end{bmatrix} [v_1, \ldots, v_{k-1}, v_k + \|v\|, 0] \right\} \begin{bmatrix} v \\ ? \end{bmatrix} =$$

$$= \begin{bmatrix} v \\ ? \end{bmatrix} - \begin{bmatrix} v_1 \\ \vdots \\ v_{k-1} \\ v_{k} + \|v\| \\ 0 \end{bmatrix} \frac{\|v\|^2 + v_k \|v\|}{\|v\|(v_k + \|v\|)} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -\|v\| \\ ? \end{bmatrix}$$

where $\|v\|$ is the Euclidean norm of $\begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}$.

Starting with an index = m and k = n+m we zero out the first k-index -1 in elements column k. Then reduce k by 1 (or move one column to the left) and continue. If we find that the first k-index elements of column k are already zero, we reduce the index by 1, reduce k by 1 and continue. All of the work is done by multiplication on the left, but it is easy to see that multiplication on the right does not undo the work. For instance u_k, the matrix which produces zeros in the k^{th} column is of the form $u_k = \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix}$ where C is at most $(k-1)x(k-1)$. Thus multiplication of H by u_k on the right cannot affect columns k thru n+m, precisely the columns that have already been transformed.

We give the essential part of the algorithm below. In so doing we use the pseudo code INFL of Stewart [4]. Here N is the dimension of the system and M is the number of controls.

```
1.)  INDEX = M

2.)  For K = N+M,...,1

     1.)  If H(K+1 - INDEX , K +1) = 0
          1) INDEX = INDEX -1

     2.)  If K-Index ≤ 1 Exit Loop

     3.)  ETA = max { |H(I,K)|: I = 1,...,K- INDEX}
```

4.) If ETA = 0 Step K

5.) V(I,K) = H(I,K)/ETA I = 1,...,K -INDEX

6.) SIGMA = sign (V(K - INDEX, K)) $\sqrt{V^2(1,K)+\ldots+V^2(K\text{-INDEX},K)}$

7.) V(K-INDEX, K) = V(K-INDEX, K) + SIGMA

8.) PI(K) = SIGMA * V(K-INDEX, K)

After finding the vector V and the scalor π, the elementary reflector U_k consists of $U_k = I - \frac{VV^T}{\pi}$ and the transformation matrix C is $C = U_1U_2\ldots U_{n+m}$. Of course some of the U_k may be identities.

A very similar procedure was used by Minimis and Paige [5], as a first step, for the purpose of placing eigenvalues. They use a generalized upper Hessenberg matrix rather than a generalized lower Hessenberg matrix. With renaming of coordinates they would have obtained the block triangular form. Once in block triangular form, as we have shown in this paper, the transformation to Brunovsky canonical form and hence the placing of eigen values is extremely easy. In Brunovsky form, the system is decoupled into several single control systems, each represented by a string of integrators.

Once we have a block triangular system, we can utilize the zero pattern of H to select the lead variables for the Brunovsky form. We do this by working our way up the super diagonals of H until encountering zeros. The column number of the first zero in each of the first m superdiagonal working upwards tells us which variables have been "skipped over" and must be lead variables. Recall also that x_1 is always a lead variable. The following algorithm shows precisely how this is done.

1.) K = M

2.) I = N

3.) For L = 1, N +M

  1.) If $|H(I,I+K)| \neq 0$

    1.) If I = 1 Exit Loop

    2.) I = I-1

    3.) Step L

  2.) IOTA (K) = I+K

  3.) $Z(K) = \sum_{J=1}^{N} C(I+K,J)\ X(J)$

  4.) K = K-1

  5.) Step L

4.) If $|H(1,J+K)| \neq 0$

  1.) For L = 1,K

    1.) IOTA(L) = L

    2.) $Z(L)\ \sum_{J=1}^{N} C(L,J)\ X(J)$

Notice the lead variables Z of the Brunovsky form are computed in terms of the original x, not in terms of the intermediate block triangular system. The variable $Z(K) = Y(I+K) = \sum_{J=1}^{N} C(I+K,J)\ X(J)$ would be the I+Kth variable in the block triangular system.

Once we know the lead variables in the Brunovsky form, all we need to know are the Kronecker indices. These can easily be found by counting the number of derivatives necessary to go from the lead variables in block triangular form to the controls. The following algorithm does this, using the value of IOTA from the previous algorithm.

1.) For K=1, M

    1.) I = IOTA(K)

    2.) KAPPA(K) = 1

    3.) For J = N+M,...,1

        1.) If $|H(I,J)| \neq 0$

            1.) If J > N Step K

            2.) I = J

            3.) KAPPA(K) = KAPPA(K) + 1

            4.) Begin J Loop again

        2.) Step J

    4.) Step K

V. RESULTS

We apply our theory and algorithms to a linear system on $\mathbb{R}^7$ with 3 controls.

Example: The following results were achieved using single precision Fortran on a VAX 11/780 machine. For purposes of distinguishing zeros, numbers having absolute value less than 0.00001 were considered zero.

Original System

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 5 \\ 3 & 2 & 0 & 0 & 0 & 1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 3 & 0 \\ 2 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

Transformation to Brunovsky form

$Z_1 = 0.16273 x_2 + 0.16781\ x_3 + 0.75770\ x_4 + 0.50853\ x_5 - 0.33563\ x_6$

$\kappa_1 = 3$

$Z_4 = -\ 0.85120\ x_2 - 0.09259\ x_3 - 0.03736\ x_4 + 0.48083\ x_5 + 0.18518\ x_6$

$\kappa_2 = 3$

$Z_7 = -\ x_1$

$\kappa_3 = 1$

Here κ_1, κ_2, and κ_3 are in order as in definition 3.1 and are not necessarily the Kronecker indices associated with the original controls u_1, u_2, and u_3 respectively. That is, the subscript 1 does not identify κ_1 with u_1. In a Brunovsky canonical form, the controls appear only at the end of integral strings; that is, after three derivatives of Z_1, three derivatives of Z_4, and one derivative of Z_7. All state variables Z should be functions only of the state variables x.

To get some idea of the accuracy of our numerical method we will look at the actual results in the first string.

We have

$$Z_2 = \dot{Z}_1 = 0.16273\ x_3 + 0.49327\ x_4\ 0.75770\ x_5 - 0.32546\ x_6$$
$$-\ 0.00002\ x_7 - 0.00003\ u_2$$

Notice that the control u_2 does appear in Z_2, but only negligibly. Computing, we find

$$Z_3 = \dot{Z}_2 = \ddot{Z}_1 = -\ 0.00060\ x_1 - 0.00004\ x_2 + 0.16273\ x_4 + 0.49327\ x_5$$
$$+\ 0.01354\ x_6 - 0.64098\ x_7 - 0.00004\ u_1 - 0.00002\ u_2$$
$$-\ 0.00003\ \dot{u}_2$$

We again have in Z_3 that the appearance of the controls u_1 and u_2 and the derivative u_2 is negligible.

The new control $v_1 = \dot{Z}_3 = \ddot{Z}_2 = \dddot{Z}_1 =$

$$- 1.92294\ x_1 - 1.28202\ x_2 - 0.00004\ x_3 - 0.00014\ x_4$$
$$+ 0.16273\ x_5 + 0.39972\ x_6 - 3.45276\ x_7$$
$$- 1.28196\ u_1 - 0.60036\ u_2 - 0.00024\ u_3 - 0.00004\ \dot{u}_1$$
$$- 0.00002\ \dot{u}_2 - 0.00003\ \ddot{u}_2$$

In the new control $\dot{v}_1$, the original controls appear substantially but derivatives of controls appear only negligibly. These results can be improved by using more precise arithmetic, but we want to show how well it works even in single precision.

The weakest link in the algorithm is of course the recognition of zeros. Because the computer uses discrete arithmetic, there are no absolute zeros and nonzeros. We must have some criteria for deciding which numbers are zero. We chose an arbitrary cutoff point of 0.00001 in our example. Perhaps a more objective approach would be to compare an element with other elements in the matrix by using a matrix norm. We expect however that in an actual application, there will always be some "tailoring" of this cutoff point to suit the particular situation.

As mentioned previously, George Meyer is using transformations to Brunovsky canonical forms to design automatic flight controllers for vertical and short take off aircraft. In his scheme, the weak dependencies of the new state variables on the controls in our example are ignored, and errors are treated by a regulator which resides in the aircraft's on board computer.

In practice, we may be receiving noisy measurements for state and control variables. For our process, once we reach block triangular form, time derivatives are required to complete the transformations to Brunovsky form. Meyer [2] has a beautiful technique for moving from block triangular to Brunovsky form which involves "smoothing integration" using the inflight computer before the differentation process takes place.

The theory of transformations of nonlinear systems to linear systems is developed in [11],[12],[13],[14],[15], [16]. Recent applications of this theory to automatic flight control are found in [17],[18],[19]. The techniques of this paper are employed in [8] to build approximate transformations of nonlinear systems to Brunovsky form.

REFERENCES

[1] Meyer, G. and Cicolani, L., 1975, A formal structure for advanced automatic flight control systems, NASA TN D-7940.

[2] Meyer, G. and Cicolani, L., 1980, Applications of nonlinear system inverses to automatic flight control design-system concepts and flight evaluations. AGARDograph 251 on Theory and Applications of Optimal Control in Aerospace Systems, P. Kant, ed., reprinted by NATO.

[3] Smith, G. A. and Meyer, G., Total aircraft flight control system balanced open- and closed-loop with dynamic trimmaps, 3rd Avionic conference Dallas, 1979.

[4] Smith, G.A. and Meyer, G., Applications of the concept of dynamic trim control to automatic landing of carrier aircraft, NASA TP-1512, 1980.

[5] Meyer, G., The design of exact nonlinear model followers, Proceedings of 1981 Joint Automatic Control Conference, FA3A.

[6] Wehrend, W. R., Jr. and Meyer, G., Flight tests of the total automatic flight control system (TAFCOS) concept on a DHC-6 Twin Otter aircraft., NASA TP-1513, 1980.

[7] Brunovsky, P., A classification of linear controllable systems, Kibernetika (Praha) 6 (1970), pp. 173-188.

[8] Ford, H., Hunt, L. R., Meyer, G. and Su, R., The modified tangent model, submitted.

[9] Stewart, G. W., Introduction to matrix computations (New York: Academic Press), 1973.

[10] Minimis, G. S. and Paige, C.C., An algorithm for pole placement of time invariant multi-input linear systems, 21st IEEE Conference on Decision and Control, Orlando, FL, 1982, 62-67.

[11] Krener, A.J., On the equivalence of control systems and the linearization of nonlinear systems, SIAM J. Control 11, 1973 pp. 670-676.

[12] Brockett, R. W., Feedback invariants for nonlinear systems, IFAC Congress, Helsinki, 1978.

[13] Jakubcyzk, B. and Respondek, W. On linearization of control systems., Bull. Acad. Polon. Sci., Ser. Sci. Math. Astronom. Phys., 28 (1980), pp. 517-522.

[14] Su, R., On the linear equivalents of nonlinear systems, Systems and Control Letters 2, No. 1 (1982), pp. 48-52.

[15] Hunt, L. R., Su, R., and Meyer, G., Design for multi-input systems, Differential Geometric Control Theory Conference, Birkhauser, Boston, R. W. Brockett, R. S. Millman, and H. Sussman, Eds., 27 (1983), 268-298.

[16] Hunt, L. R. and Su, R., Control of nonlinear time-varying systems, 20th IEEE Conference on Decision and Control, San Diego, CA, 1981, pp. 558-563.

[17] Meyer, G., Su, Renjeng and Hunt, L. R., Applications to aeronautics of the theory of transformation of nonlinear systems, 1983, CNRS Conference, pp. 153-162.

[18] Meyer, G., Hunt, L. R., and Su, R., Design of a helicopter autopilot by means of linearizing transformations, NASA Technical Memorandum 84295.

[19] Meyer, G., Su, R., and Hunt, L. R., Application of nonlinear transformations to automatic flight control, submitted.