# Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction

BRUCE C. MOORE

*Abstract*—**Kalman's minimal realization theory involves geometric objects (controllable, unobservable subspaces) which are subject to structural instability. Specifically, arbitrarily small perturbations in a model may cause a change in the dimensions of the associated subspaces. This situation is manifested in computational difficulties which arise in attempts to apply textbook algorithms for computing a minimal realization.**

**Structural instability associated with geometric theories is not unique to control; it arises in the theory of linear equations as well. In this setting, the computational problems have been studied for decades and excellent tools have been developed for coping with the situation. One of the main goals of this paper is to call attention to *principal component analysis* (Hotelling, 1933), and an algorithm (Golub and Reinsch, 1970) for computing the *singular value decomposition* of a matrix. Together they form a powerful tool for coping with structural instability in dynamic systems.**

**As developed in this paper, principal component analysis is a technique for analyzing signals. (Singular value decomposition provides the computational machinery.) For this reason, Kalman's minimal realization theory is recast in terms of responses to injected signals. Application of the signal analysis to controllability and observability leads to a coordinate system in which the "internally balanced" model has special properties. For asymptotically stable systems, this yields working approximations of $X_c$, $X_{\bar{o}}$, the controllable and unobservable subspaces. It is proposed that a natural first step in model reduction is to apply the mechanics of minimal realization using these working subspaces.**

## I. Introduction

A COMMON and quite legitimate complaint directed toward multivariable control literature is that the apparent strength of the theory is not accompanied by strong numerical tools. Kalman's minimal realization theory [2], [3], for example, offers a beautifully clear picture of the structure of linear systems. Practically every linear systems text gives a discussion of controllability, observability, and minimal realization. The associated textbook algorithms are far from satisfactory, however, serving mainly to illustrate the theory with textbook examples.

The problem with textbook algorithms for minimal realization theory is that they are based on the literal content of the theory and cannot cope with structural discontinuities (commonly called "structural instabilities") which arise. Uncontrollability corresponds to the situation where a certain subspace (controllable subspace) is proper, but arbitrarily small perturbations in an uncontrollable model may make the subspace technically not proper. Hence, for the perturbed model, the theory, taken literally, says that (assuming observability) there is no lower order model with the same impulse response matrix. There may well exist, however, a lower order model which has effectively the same impulse response matrix. There is a gap between minimal realization theory and the problem of finding a lower order approximation, which we shall refer to as the "model reduction problem."

The purpose of this paper is to show that there are some very useful tools which can be used to cope with these structural instabilities. Specifically, the tools will be applied to the model reduction problem. We shall draw heavily from the work of others in statistics and computer science, where the problem of structural instability associated with geometric theories has been studied intensely. Principal component analysis, introduced in statistics (1933) by Hotelling [4], [5] will be used together with the algorithm by Golub and Reinsch [6] (see [7] for working code) for computing the singular value decomposition of matrix. Dempster [8] gives an excellent geometric treatment of principal component analysis as well as an overview of its history. A thorough discussion of the singular value decomposition and its history is given in a recent paper by Klema and Laub [9]. There are excellent books [10]–[15] within the area of numerical linear algebra which explain how structural instabilities arise and are dealt with in the theory of linear equations.

The material given in Sections II and III of this paper is more general than that appearing in the remaining sections. In Section II minimal realization theory is reviewed from a "signal injection" viewpoint. The main advantage of this viewpoint is that the relevant subspaces are characterized in terms of responses to injected signals rather than in terms of the model parameters $(A, B, C)$. The full power of the ability to inject signals of various types is not fully exploited in this paper. Section III contains very general results which are valid whenever one is trying to find approximate linear relationships that exist among a set of time variables. In no other way is linearity required. (See [16] for ideas about nonlinear applications.)

In Section IV controllability and observability analysis is discussed. Most of the effort is spent coming to grips with the problem of internal coordinate transformations.

Manuscript received July 17, 1978; revised October 25, 1979 and June 5, 1980. This work was supported by the Natural Sciences and Engineering Research Council of Canada under Operating Grant A3925. This work represents a major revision of [1].

The author is with the Department of Electrical Engineering, University of Toronto, Toronto, Ont., Canada.

0018-9286/81/0200-0017$00.75 

This leads to the description of a coordinate system (generically unique within sign changes) in which the model has special properties. This "internally balanced" model is essentially the same as the "principal axis realization" defined in a filtering context by Mullis and Roberts [17].

For asymptotically stable systems, one can view the controllability, observability analysis as a process by which one computes working approximations of $\mathbf{X}_c$, $\mathbf{X}_{\bar{o}}$, the controllable and unobservable subspaces. In Section V we pursue the natural idea of applying the mechanics of minimal realization theory using these working subspaces instead of the exact subspaces, which are subject to structural instability. Examples are given in Section VI.

*Notation*

The symbols $\mathbb{R}$, $PC[t_1, t_2]$ represent the field of real numbers and the ring of piecewise continuous functions on the interval $[t_1, t_2]$, respectively. The corresponding vector spaces $\underbrace{\mathbb{R}\times\mathbb{R}\cdots\times\mathbb{R}}_{m}$ and $\underbrace{PC[t_1,t_2]\times\cdots\times PC[t_1,t_2]}_{m}$ will be denoted by $\mathbb{R}^m$, $PC^m[t_1, t_2]$. For a subspace $\mathbf{S}\in\mathbb{R}^n$, $\mathbf{S}^\perp$ will represent its orthogonal complement in $\mathbb{R}^n$ with respect to a basis defined by the context. A matrix $U$ will be referred to as an *orthonormal basis matrix* for $\mathbf{S}$ if its columns form an orthonormal basis.

For a map $M\colon \mathbb{R}^k\to\mathbb{R}^n$, $\ker(M)$ and $\mathrm{im}(M)$ will represent the kernel and the image of the map. The symbol $M$ will also be used to represent a matrix representation, in which case we shall write $M\in\mathbb{R}^{n\times k}$. For such a matrix, $M^T$ will represent its transpose, and $\|M\|_F$, $\|M\|_2$ will represent, respectively, the Frobenius and spectral norm. For a vector $v\in\mathbb{R}^n$, $\|v\|$ will represent the Euclidean norm $\sqrt{v^Tv}$.

## II. A "Signal Injection" View of Minimal Realization Theory

The two tools, principal component analysis and singular value decomposition, are ideal for analyzing vector time signals. In an effort to make later use of these tools more transparent, minimal realization theory will be reviewed from a "signal injection" point of view. Specifically, the controllable subspace and the unobservable subspace will be characterized in terms of vector time responses of the model to test signals injected at appropriate points.

To avoid confusion, special care must be taken to establish the state space setting. The assumed situation is the following. There is a plant, illustrated in Fig. 1, with $m$ inputs ($\hat{u}$) and $r$ outputs ($\hat{y}$), operating quietly (at rest) at an equilibrium point $(\hat{y}_e, \hat{u}_e)$. Furthermore, there is a corresponding model

$$\dot{x}(t)=Ax(t)+Bu(t)$$

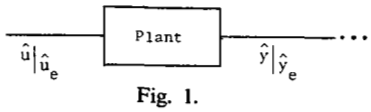

Fig. 1.

$$y(t)=Cx(t) \tag{1}$$

($x(t)\in\mathbb{R}^n$) which, when started at rest ($x(0)=0$), simulates exactly the small signal input–output characteristics of the plant with the coordinate system translated to $\hat{u}_e, \hat{y}_e$.

The model (1) is required to simulate only the input–output characteristics of the plant, and it is often true that little can be said about the relationship of $x(t)$, $A$, $B$, $C$ to the physical system. Furthermore, since (1) is only a *model*, we are free to insert a "test" input vector $d(t)$ as follows:

$$\dot{x}(t)=Ax(t)+Bu(t)+d(t)$$

$$y(t)=Cx(t). \tag{2}$$

The vector $d(t)$ is introduced to aid in the discussion of observability of the model and has no connection with the physical plant.

Minimal realization theory gives a clear geometric answer to the following questions. Under what conditions does there exist a lower order model which, when started at rest, also simulates *exactly* the small signal characteristics of the plant? If there is such a lower order model, in what way does it relate to the original model? The remainder of this section offers a review of minimal realization theory.

*Basic Geometric Review*

The geometric picture which goes along with minimal realization theory is very simple and well known. There are two important subspaces of the state space ($\mathbb{R}^n$): $\mathbf{X}_c$, the controllable subspace, and $\mathbf{X}_{\bar{o}}$, the unobservable subspace. The subspace $\mathbf{X}_c$ is the smallest subspace which contains the state response ($x(0)=0$) to every piecewise continuous vector signal injected at the model input terminals ($u(t)$). The subspace $\mathbf{X}_{\bar{o}}$ is the largest subspace in which arbitrary piecewise continuous signals can be injected [through $d(t)$] with no output response.

With $x(0)=0$ every state response can be decomposed into the sum of two orthogonal vector signals, one in $\mathbf{X}_{co}\triangleq(\mathbf{X}_c\cap\mathbf{X}_{\bar{o}})^\perp\cap\mathbf{X}_c$ and one in $\mathbf{X}_{co}^\perp$. The signals in $\mathbf{X}_{co}$ completely determine the input–output response of the model. If $U$ is an orthonormal basis matrix for $\mathbf{X}_{co}$ then

$$\dot{x}_1(t)=U^TAUx_1(t)+U^TBu(t)$$

$$y(t)=CUx_1(t)$$

is a minimal order model with the same input–output characteristics as (1).

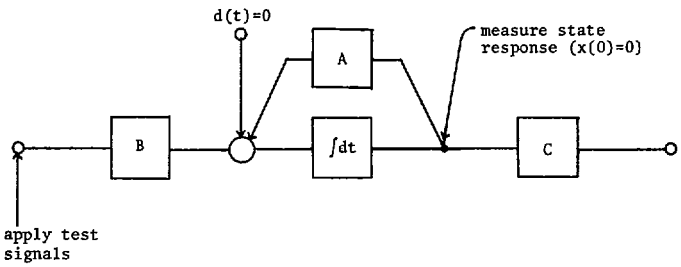


Fig. 2.

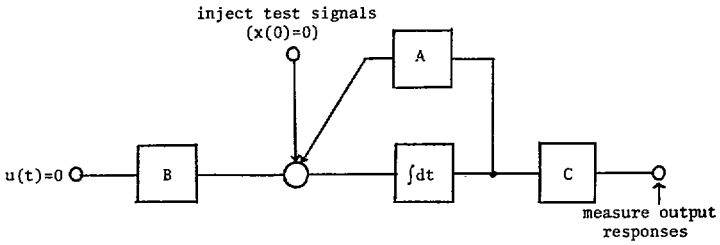


Fig. 3.

*Characterization of* $\mathbf{X}_c$

It is convenient to think of $\mathbf{X}_c$ in terms of state responses to test signals, as indicated in Fig. 2. In this paper we shall consider a sequence of test signals $u^i(t)$, $1 \leqslant i \leqslant m$ given by $u^i(t) = e_i \delta(t)$ where $e_i$ is the $i$th column of the $m \times m$ identity matrix and $\delta(t)$ is the scalar unit impulse function.[1] Let $x^1(t), \cdots, x^m(t)$ be functions corresponding to the state responses of the model to the respective test input signals, and let

$$X(t) \triangleq (x^1(t)\ x^2(t) \cdots x^m(t)).$$

For every instant of time we have $X(t) \in \mathbb{R}^{n \times m}$, and $\mathbf{X}_c$ can be characterized as follows.

*Proposition 1:* The controllable subspace $\mathbf{X}_c$ is the subspace of least dimension which contains $\mathrm{im}(X(t))$ for all $t \in [0, T]$, $T > 0$.

*Proof:* With impulses used as test signals, $X(t) = e^{At}B$ and the proof is straightforward. ■

*Characterization of* $\mathbf{X}_{\bar{o}}$

For this we apply a series of tests as indicated by Fig. 3.

Within this paper the test signals $d^i(t)$ $1 \leqslant i \leqslant n$ are given by $d^i(t) = e_i \delta(t)$ where $e_i$ is the $i$th column of the $n \times n$ identity matrix and $\delta(t)$ is the scalar unit impulse function.[1]

Let $y^1(t), \cdots, y^n(t)$ be the output responses corresponding to the respective test signals, and let

$$Y(t) \triangleq (y^1(t)\ y^2(t) \cdots y^n(t)).$$

For every instant of time we have $Y(t) \in \mathbb{R}^{r \times n}$, and $\mathbf{X}_{\bar{o}}$ is characterized as follows.

*Proposition 2:* The unobservable subspace $\mathbf{X}_{\bar{o}}$ is the subspace of greatest dimension which is contained in $\ker(Y(t))$ for all $t \in [0, T]$, $T > 0$.

*Proof:* With impulses used as test signals, $Y(t) = Ce^{At}$ and the proof is straightforward. ■

*Remark:* Since $\ker(Y(t)) = \mathrm{im}(Y^T(t))^{\perp}$, it is also true that $\mathbf{X}_{\bar{o}}^{\perp}$ *is the subspace of least dimension which contains* $\mathrm{im}(Y^T(t))$ *for all* $t \in [0, T]$, $T > 0$.

[1]The results of this section are valid with $\delta(t)$ selected to be other scalar functions, such as the unit step. In this paper we shall not consider tests other than impulses, but preliminary work indicates that future refinment of the model reduction framework developed here will take advantage of this freedom to select from a large class of test signals.

## III. Principal Component Analysis

The results given in this section support the analysis techniques to be used in later sections. The organization given here is self-contained and, to this author's knowledge, unique. Similar mathematical constructions appear in control literature, but the material is closest in spirit to Hotelling's work in statistics [4], [5].

Let $F$: $\mathbb{R} \to \mathbb{R}^{n \times m}$ be a piecewise continuous map represented in matrix form by $F(t)$. One can think of $F(t)$ as a set of $m$ vector signals involving $n$ variables, i.e., each column represents a single vector signal in $\mathbb{R}^n$. The Gramian[2]

$$W^2 \triangleq \int_{t_1}^{t_2} F(t) F^T(t)\, dt$$

is a positive semidefinite matrix with a set of nonnegative real eigenvalues $\sigma_1^2 \geqslant \sigma_2^2 \geqslant \cdots \geqslant \sigma_n^2 \geqslant 0$ and corresponding mutually orthogonal unit eigenvectors $v_1, v_2, \cdots, v_n$.

The map $F$ may be represented with $v_1, \cdots, v_n$ used as orthonormal basis vectors for $\mathbb{R}^n$, i.e.,

$$F(t) = v_1 f_1^T(t) + v_2 f_2^T(t) + \cdots + v_n f_n^T(t)$$

where $f_i^T(t) \triangleq v_i^T F(t)$ for $1 \leqslant i \leqslant n$. Throughout the paper we shall refer to the "$i$th"

principal component $\triangleq v_i f_i^T(t)$,

component vector $\triangleq v_i$,

component magnitude $\triangleq \sigma_i$, and

component function vector $\triangleq f_i(t)$.

*Proposition 3:* The following relationships hold for $1 \leqslant (i, j) \leqslant n$:

1) $\int_{t_1}^{t_2} f_i^T(t) f_j(t)\, dt = 0$ for $i \neq j$

2) $\int_{t_1}^{t_2} \| f_i(t) \|^2\, dt = \sigma_i^2$

3) $\int_{t_1}^{t_2} \| F(t) \|_F^2 dt = \sum_{i=1}^{n} \sigma_i^2$.

*Proof:* It is a simple matter to show that

$$\int_{t_1}^{t_2} f_i^T(t) f_j(t)\, dt = v_i^T W^2 v_j$$

[2]The Gramian appears in deterministic control literature [18], [19], but its eigenvector structure appears not to have been fully exploited.

from which the first two properties follow easily. The third is a consequence of the fact that

$$\int_{t_1}^{t_2} \| F(t) \|_F^2 dt = \operatorname{tr} \left\{ \int_{t_1}^{t_2} F(t) F^T(t)\, dt \right\}. \quad \blacksquare$$

There is a nice visual image which can be useful in interpreting this result. One can think of $\int_{t_1}^{t_2} \| F(t) \|^2 dt$ as the total energy in the signal set over $[t_1, t_2]$, and the component vectors and magnitudes reflect the spatial distribution of this energy. Specifically, assuming that $\sigma_n > 0$, the set

$$\left\{ v \in \mathbb{R}^n : \int_{t_1}^{t_2} \| v^T F(t) \|^2 dt = 1 \right\} \tag{3}$$

is an ellipsoid with semi-axes $(1/\sigma_i) v_i$, $1 \leqslant i \leqslant n$.

*Linear Dependence and Least Squares Approximations*

It is often important to determine fixed (independent of $t$) linear relationships which exist among the rows of $F(t)$ over $[t_1, t_2]$. From a geometric point of view, this amounts to finding the subspace $\mathbf{S}_F^{\perp}$ where

$$\mathbf{S}_F \triangleq \{ v : v \in \operatorname{im}(F(t)), t \in [t_1, t_2] \}.$$

It is clear that $\mathbf{S}_F$ is spanned by the component vectors corresponding to nonzero magnitudes.

In many cases exact linear dependence is not the case ($\mathbf{S}_F = \mathbb{R}^n$) and one looks instead for "approximate" linear relationships over $[t_1, t_2]$. The following result shows that the component magnitudes and vectors reveal the possibilities with respect to such approximations.

*Proposition 4:* Let $k$ be a fixed integer, $1 \leqslant k \leqslant n$. Over the class of piecewise continuous $F_A(t)$ satisfying $\dim\{\mathbf{S}_{F_A}\} = k$, the residuals

$$JF \triangleq \int_{t_1}^{t_2} \| F(t) - F_A(t) \|_F^2 dt;$$

$$JS \triangleq \max_{\| v \| = 1} \int_{t_1}^{t_2} \| v^T (F(t) - F_A(t)) \|^2 dt$$

are minimized with

$$F_A(t) = F_k(t) \triangleq \sum_{i=1}^{k} v_i f_i^T(t).$$

The error residuals are

$$JF = \sum_{i=k+1}^{n} \sigma_i^2; \qquad JS = \sigma_{k+1}^2.$$

*Proof:* It is easy to verify that $F_k(t)$ gives the stated residuals. For an approximation $F_A$ which minimizes $JF$ or $JS$, it is necessarily true that $\int_{t_1}^{t_2} F_A(t) E_A^T(t)\, dt = 0$, where $E_A(t) = F(t) - F_A(t)$. Hence, for such an approximation

$$W^2 = \int_{t_1}^{t_2} F_A(t) F_A^T(t)\, dt + \int_{t_1}^{t_2} E_A(t) E_A^T(t)\, dt.$$

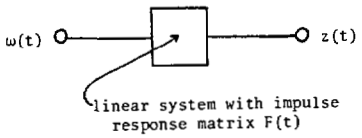


Fig. 4.

If $\int_{t_1}^{t_2} F_A(t) F_A^T(t)\, dt$ has rank $k$, it follows from perturbation properties of singular values that

$$JF = \operatorname{tr} \int_{t_1}^{t_2} E_A(t) E_A^T(t)\, dt \geqslant \sum_{i=k+1}^{n} \sigma_i^2$$

$$JS = \left\| \int_{t_1}^{t_2} E_A(t) E_A^T(t)\, dt \right\|_2 \geqslant \sigma_{k+1}^2. \quad \blacksquare$$

*Principal Components of Impulse Response Matrices*

If $F(t)$ is the impulse response matrix of a linear time-invariant system, the principal components over $[0, T]$ can be given sharp systemic interpretations. Consider the system depicted in Fig. 4 where $\omega(t) \in \mathbb{R}^m$ and $z(t) \in \mathbb{R}^n$, and let $\Omega$ represent the class of all input functions $\omega(\cdot)$ which are piecewise continuous on $[0, T]$ and satisfy the norm bound $(\int_0^T \| \omega(t) \|^2 dt)^{1/2} \leqslant 1$.

It is well known (see [19, p. 75]) that the image of the convolution map

$$\left\{ z \in \mathbb{R}^n : z = \int_0^t F(t-\tau) \omega(\tau)\, d\tau,\ t \leqslant T,\ \omega(\cdot) \in PC^m[0, T] \right\}$$

is precisely the space $\mathbf{S}_F$. The *set* (contained in $\mathbf{S}_F$)

$$S = \left\{ z \in \mathbb{R}^n : z = \int_0^t F(t-\tau) \omega(\tau)\, d\tau,\ t \leqslant T,\ \omega(\cdot) \in \Omega \right\}$$

provides more detailed structural information than the space $\mathbf{S}_F$. In the next few paragraphs we shall show that $S$ is a region in $\mathbb{R}^n$ whose surface is an ellipsoid [closely related to (3)] defined by the component magnitudes and vectors of $F(t)$ over $[0, T]$.

Let $\sigma_i$, $v_i$, $1 \leqslant i \leqslant n$ be the component magnitudes and vectors of $F(t)$ over $[0, T]$, and let $\Sigma$, $V \in \mathbb{R}^{n \times n}$ be defined as follows:

$$\Sigma \triangleq \operatorname{diag}\{\sigma_1, \sigma_2, \cdots, \sigma_n\}$$

$$V \triangleq (v_1\, v_2 \cdots v_n).$$

The following proposition shows that the set $\hat{S} = \{ z : z = V \Sigma p,\ \| p \| = 1 \}$ (an ellipsoid with semi-axes $\sigma_i v_i$, $1 \leqslant i \leqslant n$) corresponds to the surface of $S$.

*Proposition 5:* The set $S$ can be characterized as follows

$$S = \{ z : z = \alpha \hat{z} \text{ with } \hat{z} \in \hat{S} \text{ and } 0 \leqslant \alpha \leqslant 1 \}.$$

*Proof:* For simplicity we shall assume that $\sigma_n > 0$. The general proof is basically the same only more tedious. First it will be shown that $\hat{S}$ is contained in $S$. For every

$\hat{z} \in \hat{S}$ there is a vector $p$, $\|p\|=1$, satisfying $\hat{z}=V\Sigma p$. From the definition of principal components, we know that

$$W^2 = \int_0^T F(t)F^T(t)\,dt = V\Sigma^2 V^T.$$

This means that $q \triangleq V\Sigma^{-1}p$ satisfies $\hat{z}=W^2 q$, which implies (see [19, p. 76]) that the input $\omega(t)=F^T(T-t)q$ drives $z(t)$ to $\hat{z}$ at time $T$. Furthermore, simple manipulation gives $\int_0^T \omega(t)\omega^T(t)\,dt=1$ which means that $\omega(\cdot)\in\Omega$ and $\hat{z}\in S$.

To complete the proof it is sufficient (because of linearity) to show that the input $\omega(\cdot)$ constructed in the previous paragraph is the minimum norm input which drives the system to $\hat{z}$ at time $T$. Let $\hat{\omega}(t)$ be any input which drives the system to $\hat{z}$ at time $T$. With $\Delta(t)=\hat{\omega}(t)-\omega(t)$, simple manipulation gives $\int_0^T \hat{\omega}^T(t)\Delta(t)=0$ and it follows that

$$\int_0^T \|\hat{\omega}(t)\|^2\,dt = \int_0^T \|\omega(t)\|^2\,dt + \int_0^T \|\Delta(t)\|^2\,dt \geq 1. \quad \blacksquare$$

It is interesting to note that if $\hat{z}=\sigma_i v_i$, then the minimum norm input function is

$$\omega(t)=F^T(T-t)v_i\sigma_i^{-1}=\frac{1}{\sigma_i}f_i(T-t)$$

which is the $i$th component function vector, normalized and reflected in time.

*Computation of Component Magnitudes and Vectors*

First let us deal with the special situation where one has an asymptotically stable model $(A, B, C)$ and wishes to compute the components of $e^{At}B(e^{A^Tt}C^T)$ over $[0,\infty)$. In this case it is often convenient to use the fact that

$$W_c^2 \triangleq \int_0^\infty e^{At}BB^Te^{A^Tt}\,dt$$

$$W_0^2 \triangleq \int_0^\infty e^{A^Tt}C^TCe^{At}\,dt$$

are the unique symmetric positive semidefinite matrices which satisfy

$$AW_c^2 + W_c^2A^T = -BB^T$$

$$A^TW_0^2 + W_0^2A = -C^TC.$$

One can first solve for $W_c^2(W_o^2)$ (see [20] for one algorithm), and then use a specialized routine for computing the eigenvalues and eigenvectors of a symmetric matrix.

For the more general case it is necessary to compute approximate component magnitudes and vectors by sampling $F(t)$. With evenly spaced sample points $\tau_0, \tau_1, \cdots, \tau_N$ ($\tau_0 \triangleq t_1$, $\tau_N \triangleq t_2$), it follows from rectangular approximation that if $N$ is large, then

$$W^2 \approx W_N^2 \triangleq \left(\frac{t_2-t_1}{N}\right)\sum_{i=1}^{N} F(\tau_i)F^T(\tau_i).$$

It is not necessary or desirable, however, to compute such an approximation for $W^2$. Imagine (it need not be actually constructed) a data matrix

$$D_N \triangleq \left(\frac{t_2-t_1}{N}\right)^{1/2}(F(\tau_1)\ F(\tau_2)\cdots F(\tau_N)).$$

Since $W_N^2 = D_N D_N^T$, it follows that the singular values and left singular vectors of $D_N$ approximate the component magnitudes and vectors of $F(t)$ over $[t_1, t_2]$.

It is not necessary to store $D_N$ in memory—this is important because the data matrices may contain many column vectors. Instead, one may preprocess the data by recursively (treating as few as one column at a time) reducing $D_N$ to a unitarily equivalent matrix (see [12, p. 383]):

$$R = D_N Q \qquad (Q \text{ unitary—not stored}).$$

The matrix $R\in\mathbb{R}^{n\times n}$ has singular values and left singular vectors equal (to machine precision) to those of $D_N$. The algorithm (*SVD*) developed by Golub and Reinsch [6] (working code in [7]) can be used to compute the singular values and left singular vectors of $R$.

The main advantage in using *SVD*, instead of computing the eigenvalues and eigenvectors of $W_N^2$, is in reduced resolution requirements. The "squaring process" doubles the demand for resolution in the computations. Specifically, suppose there is 12 bit resolution associated with the samples $F(\tau_i)$. To preserve this same resolution with the computed singular values,

$$\sigma_i - 2^{-12}\sigma_1 \leq \{\text{computed value of } \sigma_i\} \leq \sigma_i + 2^{-12}\sigma_1$$

requires at least 12 bit resolution using *SVD* and at least 24 bit resolution using the squared up version.

*Perturbation Properties of Component Magnitudes and Vectors*

Suppose $F(t)$ is perturbed by $\Delta F(t)$ so that $F_\Delta(t) \triangleq F(t)+\Delta F(t)$ is piecewise continuous. As one might expect from the preceding paragraphs, the perturbation of component magnitudes may be bound in much the same way as singular values of a matrix.

*Proposition 6:* Let $\sigma_i$, $\sigma_i^\Delta$ be the $i$th component magnitudes of $F(t)$, $F_\Delta(t)$, respectively. Then

$$|\sigma_i - \sigma_i^\Delta| \leq \left\| \int_{t_1}^{t_2} \Delta F(t)\Delta F^T(t) \right\|_2^{1/2}.$$

*Proof:* From the discussion of the computation of components, the perturbed Gramian is

$$W_\Delta^2 = \lim_{N\to\infty} (D_N + \Delta D_N)(D_N + \Delta D_N)^T$$

where $\Delta D_N = ((t_2-t_1)/N)^{1/2}(\Delta F(\tau_1)\ \Delta F(\tau_2)\cdots\Delta F(\tau_N))$. Standard perturbation results for singular values implies that each singular value of $D_N + \Delta D_N$ is perturbed from

that of $D_N$ by no more than

$$\|\Delta D_N\|_2 = \left\| \left( \frac{t_2 - t_1}{N} \right) \sum_{i=1}^{N} \Delta F(\tau_i) \Delta F^T(\tau_i) \right\|_2^{1/2}.$$

In the limit the right-hand expression approaches $\| \int_{t_1}^{t_2} \Delta F(t) \Delta F^T(t)\, dt \|_2^{1/2}$. ■

As is the case with singular vectors, the perturbation of component vectors requires some explanation. Let $v_i^\Delta$ represent the $i$th component vector of $F_\Delta(t)$ and visualize the two ellipsoids $E$, $E_\Delta$, the first with semi-axes $\sigma_i v_i$, and the perturbed one with $\sigma_i^\Delta v_i^\Delta$. If $\Delta F(t)$ has very small principal components, then $E_\Delta$ is close to $E$; that is, no point in $E$ is perturbed by more than $\| \int_{t_1}^{t_2} \Delta F(t) \Delta F^T(t)\, dt \|_2^{1/2}$. If, however, $E$ has two or more semi-axes of nearly equal lengths (circular cross section), then these axes may be very sensitive to small perturbations. The space spanned by them is not sensitive, however, and this presents no real difficulty. The algorithm *SVD* gives singular vectors which are orthogonal (to machine precision), even with repeated singular values.

## IV. Controllability and Observability Analysis

We are now prepared to consider the application of principal component analysis to responses of the model

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t)$$

with the idea of combining it with the signal-injection view of minimal realization. The basic idea is close to one advanced by Friedland [21] and, as mentioned in the Introduction, is very closely related to the filtering work of Mullis and Roberts [17].

A central problem to be dealt with is the fact that internal responses $e^{At}B$, $e^{A^Tt}C^T$ depend on the internal coordinate system. This means that unless there is some special significance attached to the internal coordinate system, the existence of "small" components in $e^{At}B$ or $e^{A^Tt}C^T$ implies nothing about their importance with respect to input–output properties of the model. To overcome this problem, we derive a special coordinate system where input–output properties *are* reflected by internal principal components.

In the development we shall find it useful to carry along the discrete time subordinate obtained by sampling (and holding inputs) every $t_s$ seconds.

$$\left.\begin{aligned} x_{k+1}^* &= Fx_k^* + Gu_k^* \\ y_k^* &= Cx_k^* \end{aligned}\right\} \quad F \triangleq e^{At_s}; \quad G \triangleq \int_0^{t_s} e^{A(t_s - \tau)} B\, d\tau.$$

In analyzing responses of $(A, B, C)$ over an interval $[0, T]$, it will be assumed that $N = T/t_s$ is an integer.

*Relationships Between $(F, G, C)$ and Components of $e^{At}B$, $e^{A^Tt}C^T$*

Let $Q_c(t_s)$, $Q_o(t_s)$ represent the extended controllability and observability matrices corresponding to $(F, G, C)$:

$$Q_c(t_s) \triangleq (G \; FG \; \cdots \; F^N G);$$

$$Q_o(t_s) \triangleq \begin{bmatrix} C \\ CF \\ \vdots \\ CF^N \end{bmatrix}. \tag{4}$$

The matrix $Q_c(t_s)$ is a data matrix which is closely related to $e^{At}B$, and $Q_o^T(t_s)$ is one formed by sampling $Ce^{At}$ every $t_s$ seconds over the time interval $[0, T]$. To aid in the discussion, we shall adopt the following notation.

1) $V_c \triangleq (v_{c1} v_{c2} \cdots v_{cn})$; $\Sigma_c \triangleq \operatorname{diag}\{\sigma_{o1}, \cdots, \sigma_{on}\}$ where $v_{ci}$, $\sigma_{ci}$ represent the $i$th component vector and magnitude of $e^{At}B$.

2) $V_o \triangleq (v_{o1} v_{o2} \cdots v_{on})$; $\Sigma_o \triangleq \operatorname{diag}\{\sigma_{o1}, \cdots, \sigma_{on}\}$ where $v_{oi}$, $\sigma_{oi}$ represent the $i$th component vector and magnitude of $e^{A^Tt}C^T$.

3) $V_c^*(t_s)\Sigma_c^*(t_s)U_c^{*T}(t_s) \triangleq$ the singular value decomposition of $Q_c(t_s)$.

4) $V_o^*(t_s)\Sigma_o^*(t_s)U_o^{*T}(t_s) \triangleq$ the singular value decomposition of $Q_o^T(t_s)$.

*Proposition 7:* The singular values satisfy

$$\lim_{t_s \to 0} \frac{1}{\sqrt{t_s}} \Sigma_c^*(t_s) = \Sigma_c; \qquad \lim_{t_s \to 0} \sqrt{t_s}\, \Sigma_o^*(t_s) = \Sigma_o.$$

If the diagonal elements of $\Sigma_c$, $\Sigma_o$ are distinct, then

$$\lim_{t_s \to 0} V_c^*(t_s) = V_c; \qquad \lim_{t_s \to 0} V_o^*(t_s) = V_o.$$

*Proof:* It is easy to see, using rectangular approximation of integration, that

$$\int_0^T e^{A^Tt} C^T C e^{At}\, dt = \lim_{t_s \to 0} \left( \sqrt{t_s}\, Q_0(t_s) \right)^T \left( \sqrt{t_s}\, Q_o(t_s) \right)$$

and we shall prove that

$$\int_0^T e^{At} B B^T e^{A^Tt}\, dt = \lim_{t_s \to 0} \left( \frac{1}{\sqrt{t_s}} Q_c(t_s) \right) \left( \frac{1}{\sqrt{t_s}} Q_c(t_s) \right)^T. \tag{5}$$

These two relationships imply that the claimed limiting relationships hold.

Now to see that (5) holds, note that

$$G = \int_0^{t_s} e^{A(t_s - \tau)} B\, d\tau = t_s B + \frac{t_s^2}{2} AB + \cdots$$

and

$$GG^T = t_s^2 BB^T + \frac{t_s^3}{2} (ABB^T + BB^T A^T) + \cdots.$$

Since $e^{At}$ is bounded on $[0, T]$, there exists finite $K$, $\delta > 0$

such that for $t_s < \delta$

$$\| F^i\left(t_s^2 BB^T - GG^T\right)F^{iT} \| < Kt_s^3$$

for all $i \geq 0$. Hence,

$$\| t_s F^i BB^T F^{iT} - \frac{1}{t_s} F^i GG^T F^{iT} \| < Kt_s^2,$$

and for $t_s < \delta$,

$$\| \sum_{k=0}^{N} t_s e^{At_s k} BB^T e^{A^T t_s k} - \frac{1}{t_s} Q_c(t_s) Q_c^T(t_s) \| < KNt_s^2 = KTt_s.$$

Since $T$, $K$ are fixed constants, the result follows. ■

*Principal Components of $e^{At}B$, $e^{A^Tt}C^T$*

The component vectors corresponding to the nonzero principal components of $e^{At}B$ span $\mathsf{X}_c$, the controllable subspace. For simplicity we shall assume that there are no components which are identically zero. For this author, the mental image developed in the previous section is helpful. The ellipsoid with semi-axes $\sigma_{c1}v_{c1}$, $\sigma_{c2}v_{c2}, \cdots, \sigma_{cn}v_{cn}$ is the surface of the region in the state space corresponding to points which can be reached from the origin with input vectors satisfying $\int_0^T \| \omega(t) \|^2 dt \leq 1$. In some respects the ratio $\mu_c = \sigma_{c1}/\sigma_{cn}$ serves as a condition number with respect to pointwise state control.[3]

The component vectors corresponding to the nonzero principal components of $e^{A^Tt}C^T$ span $\mathsf{X}_o^{\perp}$. Again we shall rule out the trivial case where there are components which are identically zero. Here one can imagine an ellipsoid in $\mathbb{R}^n$ with semi-axes (descending order according to length)

$$\sigma_{on}^{-1}v_{on}, \sigma_{on-1}^{-1}v_{on-1}, \cdots, \sigma_{o1}^{-1}v_{o1}$$

which corresponds to the set of all initial conditions which satisfy $\int_0^T \| Ce^{At}x_0 \|^2 dt = 1$. The ratio $\mu_o = \sigma_{o1}/\sigma_{on}$ acts in some respects as a condition number with respect to zero-input state observation.

It may be tempting to treat very small components of $e^{At}B$ or $e^{A^Tt}C^T$ as though they were identically zero (i.e., approximate $\mathsf{X}_c$, $\mathsf{X}_{\bar{o}}$ using least squares). A little reflection reveals that this is not generally appropriate. Consider the model

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 10^{-6} \\ 10^{6} \end{pmatrix} u(t)$$

$$y(t) = (10^6 \quad 10^{-6})\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

It is easy to see that $e^{At}B$, $e^{A^Tt}C^T$ both have one very small principal component with $\mu_o \approx \mu_c \approx 10^{12}$. In this case the highly distorted ellipsoids simply reflect an internal scaling imbalance. Rescaling $\hat{x}_1(t) = 10^6 x_1(t)$; $\hat{x}_2(t) = 10^{-6}x_2(t)$ gives

$$\begin{pmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u(t)$$

$$y(t) = (1 \quad 1)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

The preceding example illustrates the fact that $e^{At}B$, $Ce^{At}$ depend upon the internal coordinate system. A coordinate transformation $x(t) = P\hat{x}(t)$ gives the model

$$\dot{\hat{x}} = \hat{A}\hat{x}(t) + \hat{B}u(t)$$
$$y(t) = \hat{C}x(t)$$

where $\hat{A} = P^{-1}AP$, $\hat{B} = P^{-1}B$, $\hat{C} = CP$. It is important to observe that

$$e^{\hat{A}t}\hat{B} = P^{-1}e^{At}B; \qquad \hat{C}e^{\hat{A}t} = Ce^{At}P.$$

In discussing coordinate transformations, the following notation will be adopted:

$$W_c^2(P) \triangleq P^{-1}\left(\int_0^T e^{At}BB^Te^{A^Tt}\,dt\right)P^{-1^T}$$
$$= \int_0^T e^{\hat{A}t}\hat{B}\hat{B}^Te^{\hat{A}^Tt}\,dt$$
$$W_0^2(P) \triangleq P^T\left(\int_0^T e^{A^Tt}C^TCe^{At}\,dt\right)P$$
$$= \int_0^T e^{\hat{A}^Tt}\hat{C}^T\hat{C}e^{\hat{A}t}\,dt.$$

For the case where $P = I$ (original coordinate system), we shall simply write $W_c^2$, $W_o^2$.

*Coordinate Invariant Values—Second-Order Modes*

Looking again at the example used to illustrate the dependence on internal coordinates, one can see that small components of $e^{A^Tt}C^T$ may be offset by large components of $e^{At}B$. Intuitively, it seems this problem could be alleviated by transforming to a coordinate system where the components of $e^{At}B$ all have unit magnitude, i.e., select $P$ so that $W_c^2(P) = I$. Since $W_c^2 = V_c\Sigma_c^2V_c^T$ it is clear that this can be achieved by setting $P = V_c\Sigma_c$. A little algebra gives $W_o^2(P) = H^TH$ where

$$H \triangleq \Sigma_o V_o^T V_c \Sigma_c.$$

Note that the singular values of $H$ are the component magnitudes of $e^{\hat{A}^Tt}\hat{C}^T$. Although the matrix $H$ depends on the initial coordinate system, one can easily verify that its singular values do not. These values, which will be shown to reflect input–output properties of the system, play a

[3] This ratio depends on the selected coordinate system—a poor condition number may simply reflect a poor internal coordinate system. Discussion of this will appear later in this section.

central role in the remainder of this paper. For this reason we shall attach special terminology.[4]

*Definition:* The singular values of $H$ will be represented by $\sigma_1^2 \geqslant \sigma_2^2 \geqslant \cdots \geqslant \sigma_n^2 > 0$ and will be referred to as *second-order modes* of the system.

There is an interesting connection between the second-order modes and discrete-time Hankel matrices corresponding to $(F, G, C)$. Recall that for a given interval $[0, T]$ and sampling time $t_s$, we have the extended controllability and observability matrices $Q_c(t_s)$, $Q_o(t_s)$ defined by (4). With $T$ fixed, the corresponding Hankel matrix $M_H(t_s) = Q_o(t_s)Q_c(t_s)$ grows in size (number of rows and columns) without bound as $t_s \to 0$, but $M_H$ has at most $n$ nonzero singular values.

*Proposition 8:* Let $\sigma_i^*(t_s)$, $1 \leqslant i \leqslant n$ be the ordered singular values of $M_H(t_s)$. Then for $1 \leqslant i \leqslant n$,

$$\lim_{t_s \to 0} \sigma_i^*(t_s) = \sigma_i^2$$

where $\sigma_i^2$ is the $i$th second-order mode.

*Proof:* Proposition 7 implies that

$$\frac{1}{\sqrt{t_s}} Q_c(t_s) = (V_c \Sigma_c + R_1(t_s)) U_c^{*T}(t_s)$$

$$\sqrt{t_s}\, Q_o(t_s) = U_o^*(t_s)(\Sigma_o V_o^T + R_2(t_s))$$

where $R_1(t_s)$, $R_2(t_s)$ are $n \times n$ matrices satisfying

$$\lim_{t_s \to 0} \| R_1(t_s) \| = \lim_{t_s \to 0} \| R_2(t_s) \| = 0.$$

It follows, therefore, that

$$\begin{aligned} M_H(t_s) &= Q_o(t_s) Q_c(t_s) \\ &= U_o^*(t_s)(\Sigma_o V_o^T V_c \Sigma_c) U_c^{*T}(t_s) \\ &\quad + U_o^*(t_s) E(t_s) U_c^{*T}(t_s) \end{aligned}$$

where $E(t_s)$ is $n \times n$ and satisfies $\lim_{t_s \to 0} \| E(t_s) \| = 0$. Since $U_o^*(t_s)$, $U_c^*(t_s)$ each has orthonormal column vectors, and the second-order modes are the singular values of $H \triangleq \Sigma_o V_o^T V_c \Sigma_c$, the result follows. ■

*Models with Normalized and Balanced Internal Dynamics*

It is evident that the condition $W_c^2(P) = I$ does not define a unique coordinate system. We shall now define and discuss three closely related, essentially unique coordinate systems[5] in which the structure associated with the second-order modes is displayed clearly. Let $\Sigma^2 = \operatorname{diag}\{\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2\}$.

*Definitions:* The model $(\hat{A}, \hat{B}, \hat{C})$ is *input-normal* on $[0, T]$ if $W_c^2(P) = I$; $W_o^2(P) = \Sigma^4$; *output-normal* on $[0, T]$ if $W_c^2(P) = \Sigma^4$; $W_o^2(P) = I$; and *internally balanced* on $[0, T]$ if $W_c^2(P) = W_o^2(P) = \Sigma^2$.

Let $P_{ib}$ represent a transformation which gives an internally balanced model. It is easy to verify that $P_{on} \triangleq P_{ib}\Sigma^{-1}$, $P_{in} \triangleq P_{ib}\Sigma$ give, respectively, output-normal and input-normal models. The three coordinate systems are, therefore, related by simple scale factors. The internally balanced model will receive the most attention in later sections, largely because of the property given in the next paragraph.

Let $\mu_c^2(P)$, $\mu_o^2(P)$ represent the condition numbers (with respect to inversion) of $W_c^2(P)$, $W_o^2(P)$. Then, as discussed in previous paragraphs, $\mu_c(P)$, $\mu_o(P)$ act as condition numbers of the model $(\hat{A}, \hat{B}, \hat{C})$ with respect to pointwise state control and zero input state observation, respectively. The following result shows that, in one sense, the internally balanced model gives the best compromise between the two condition numbers.

*Proposition 9:* The quantity $\max(\mu_c(P), \mu_o(P))$ achieves its minimum with $P = P_{ib}$.

*Proof:* Let $\mu(M)$ represent the condition number of the matrix $M$ with respect to inversion. It is easy to show that

$$\mu_c(P) = \mu(P^{-1} V_c \Sigma_c); \qquad \mu_o(P) = \mu(\Sigma_o V_o^T P).$$

Furthermore, it is a simple matter[6] to show that

$$\mu(M_1)\mu(M_2) \geqslant \mu(M_1 M_2)$$

for square nonsingular matrices $M_1$, $M_2$. It follows that

$$\begin{aligned} \mu_o(P)\mu_c(P) &= \mu(\Sigma_o V_o^T P)\mu(P^{-1} V_c \Sigma_c) \\ &\geqslant \mu(H) = \left(\frac{\sigma_1}{\sigma_n}\right)^2 \end{aligned}$$

and that

$$\max(\mu_c(P), \mu_o(P)) \geqslant \left(\frac{\sigma_1}{\sigma_n}\right).$$

Since $\mu_c(P_{ib}) = \mu_o(P_{ib}) = (\sigma_1/\sigma_n)$, this completes the proof. ■

The following result shows that if the second-order modes are distinct, then the internally balanced model is essentially unique.

*Proposition 10:* If the second-order modes are distinct, then the basis vectors defining the internally balanced model are unique within a change of sign.

*Proof:* Suppose $A$, $B$, $C$ is internally balanced so that $W_c^2 = W_o^2 = \Sigma^2$ and assume that $P$ satisfies $W_c^2(P) = W_o^2(P) = \Sigma^2$. This implies that

$$P^{-1}\Sigma^2 P^{-1^T} = P^T \Sigma^2 P = \Sigma^2$$

which further implies that $P^{-1}\Sigma^4 P = \Sigma^4$. These equations constrain $P$ to be a diagonal matrix with $\pm 1$ in each diagonal entry. ■

[4] This terminology is borrowed from Mullis and Roberts [17].

[5] These correspond to "principal axis realizations" introduced by Mullis and Roberts [17].

[6] It is not a simple matter, as one reviewer observed, to show that this result holds in general for nonsquare matrices, but the general result is not needed here.

The transformation $P_{ib}$ can be executed with the following algorithm.

*Step 1)* Compute $V_c$, $\Sigma_c$ and apply the transformation $P_1 = V_c\Sigma_c$ to give $W_c^2(P_1) = I$. Let $\tilde{A} \triangleq P_1^{-1}AP_1$; $\tilde{B} \triangleq P_1^{-1}B$; $\tilde{C} \triangleq CP_1$.

*Step 2)* Compute $\tilde{V}_o$, $\tilde{\Sigma}_o$ corresponding to the component vectors and magnitudes of $e^{\tilde{A}^T t}\tilde{C}^T$. Apply to $(\tilde{A}, \tilde{B}, \tilde{C})$ the transformation $P_2 = \tilde{V}_o\tilde{\Sigma}_o^{-1/2}$ to give the internally balanced model

$$\hat{A} = P_2^{-1}P_1^{-1}AP_1P_2;$$
$$\hat{B} = P_2^{-1}p_1^{-1}B;$$
$$\hat{C} = CP_1P_2.$$

*Properties of Asymptotically Stable, Internally Balanced Models*

Let us now limit the discussion to asymptotically stable systems and assume that $(A, B, C)$ is internally balanced over $[0, \infty)$, i.e.,

$$\int_0^\infty e^{At}BB^Te^{A^Tt}\,dt = \int_0^\infty e^{A^Tt}C^TCe^{At}\,dt = \Sigma^2$$

or, equivalently,

$$A\Sigma^2 + \Sigma^2A^T = -BB^T \tag{6}$$
$$A^T\Sigma^2 + \Sigma^2A = -C^TC. \tag{7}$$

Under these conditions the model $(A, B, C)$ has a rather extraordinary property involving the stability of subsystems. Consider arbitrary reorganization of the internally balanced system into two interconnected subsystems, i.e., let $(\hat{A}, \hat{B}, \hat{C})$ be obtained by reordering the state variables and partition $(\hat{A}, \hat{B}, \hat{C})$ as follows:

$$\begin{bmatrix}\dot{\hat{x}}_1\\ \dot{\hat{x}}_2\end{bmatrix} = \begin{bmatrix}\hat{A}_{11} & \hat{A}_{12}\\ \hat{A}_{21} & \hat{A}_{22}\end{bmatrix}\begin{bmatrix}\hat{x}_1\\ \hat{x}_2\end{bmatrix} + \begin{bmatrix}\hat{B}_1\\ \hat{B}_2\end{bmatrix}u$$
$$y = (\hat{C}_1 \quad \hat{C}_2)\begin{bmatrix}\hat{x}_1\\ \hat{x}_2\end{bmatrix}.$$

*Lemma 4.1:*[7] The matrices $\hat{A}_{11}$, $\hat{A}_{22}$ satisfy $\lim_{t\to\infty} e^{\hat{A}_{11}t} = K_1$, $\lim_{t\to\infty} e^{\hat{A}_{22}t} = K_2$ where generically $K_1 = 0$, $K_2 = 0$.

*Proof:* It is a simple matter to show that $\hat{A}\hat{\Sigma}^2 + \hat{\Sigma}^2\hat{A}^T = -\hat{B}\hat{B}^T$ where $\hat{\Sigma}$ is a nonsingular diagonal matrix (reordered version of $\Sigma$). This implies that

$$\hat{A}_{11}\hat{\Sigma}_1^2 + \hat{\Sigma}_1^2\hat{A}_{11}^T = -\hat{B}_1\hat{B}_1^T,$$

and it follows (integration by parts, or see [22, p. 299]) that for $t > 0$

$$\hat{\Sigma}_1^2 = e^{\hat{A}_{11}t}\hat{\Sigma}_1^2e^{\hat{A}_{11}^Tt} + \int_0^t e^{\hat{A}_{11}\tau}\hat{B}_1\hat{B}_1^Te^{\hat{A}_{11}^T\tau}\,d\tau.$$

[7]During the process of revising this paper, Parnebo and Silverman proved a stronger stability result [23].

The symmetric, positive semidefinite matrix $\int_0^t e^{\hat{A}_{11}\tau}\hat{B}_1\hat{B}_1^T \cdot e^{\hat{A}_{11}^T\tau}\,d\tau$ is, therefore, norm bounded and it follows that $\lim_{t\to\infty} e^{\hat{A}_{11}t}\hat{B}_1 = 0$ and that $\lim_{t\to\infty} e^{\hat{A}_{11}t}\hat{\Sigma}_1$ exists. Since $\hat{\Sigma}_1$ is nonsingular, $\lim_{t\to\infty} e^{\hat{A}_{11}t}$ must exist. A straightforward argument (omitted) shows that controllability of $\hat{A}_{11}$, $\hat{B}_1$ is a generic property, and this implies further that asymptotic stability of $\hat{A}_{11}$ is a generic property. Similar arguments apply for $\hat{A}_{22}$. ■

The previous stability result is really more general than we need here. Consider the case where the state variables of the internally balanced model $(A, B, C)$ are simply partitioned and not reordered. Specifically, suppose the matrix $\Sigma^2$ of second-order modes is partitioned where $\Sigma_1 = \text{diag}\{\sigma_1, \cdots, \sigma_k\}$, $\Sigma_2 = \text{diag}\{\sigma_{k+1}, \cdots, \sigma_n\}$. This imposes a corresponding partition of the state variables of the internally balanced model

$$\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix} = \begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} + \begin{pmatrix}B_1\\ B_2\end{pmatrix}u(t) + \begin{pmatrix}d_1(t)\\ d_2(t)\end{pmatrix}$$
$$y(t) = (C_1 \quad C_2)\begin{pmatrix}x_1\\ x_2\end{pmatrix}$$

where we have again inserted the test injection input $d(t)$. We may view this as two interconnected subsystems: $(A_{11}, B_1, C_1)$ and $(A_{22}, B_2, C_2)$. Using Lemma 7, it is not difficult to prove the following.

*Proposition 11:* The two subsystems $(A_{11}, B_1, C_1)$, $(A_{22}, B_2, C_2)$ are generically asymptotically stable and internally balanced with

$$\int_0^\infty e^{A_{11}t}B_1B_1^Te^{A_{11}^Tt}\,dt = \int_0^\infty e^{A_{11}^Tt}C_1^TC_1e^{A_{11}t}\,dt = \Sigma_1^2$$
$$\int_0^\infty e^{A_{22}t}B_2B_2^Te^{A_{22}^Tt}\,dt = \int_0^\infty e^{A_{22}^Tt}C_2^TC_2e^{A_{22}t} = \Sigma_2^2.$$

*Proof:* Stability was proved in Lemma 7; the second part follows from simple inspection of (6) and (7) in partitioned form. ■

*Prelude to Model Reduction*

The basic model reduction idea to be pursued in the next section follows intuitively from Proposition 11. Consider the case where $\sigma_k^2 \gg \sigma_{k+1}^2$. Then the subspace

$$\mathsf{X}_1 \triangleq \text{im}\begin{pmatrix}I_k\\ 0\end{pmatrix}$$

(where $I_k$ is the $k \times k$ identity matrix) in some sense "acts like" both $\mathsf{X}_c$ and $\mathsf{X}_{\bar{o}}^{\perp}$. If the mechanics of Kalman's minimal realization theory are applied to the internally balanced model with $\mathsf{X}_1$ used as a working approximation of $\mathsf{X}_{co}$, the resulting lower order model is $(A_{11}, B_1, C_1)$. This model is generically asymptotically stable and internally balanced.

This idea is given strength by Proposition 10 which shows that the second-order modes are the singular values

of the (infinite) Hankel matrix of the discrete-time subordinate with $t_s \to 0$. It is well known that the number of nonzero singular values of this matrix determines the order of the model. Hence, if $\sigma_k^2 \gg \sigma_{k+1}^2$, then the Hankel matrix is nearly singular. In some model reduction work [24]–[27], model reduction is based on near singularity of the Hankel matrix.

## V. Tools for Model Reduction

Model reduction involves a tradeoff between model order and the degree to which the characteristics of the plant are reflected by the model. Because the relative importance of various plant characteristics is highly dependent upon the application, there can be no universal model reduction *algorithm*. The best one can hope for is a good set of tools and some reliable guidelines for using them.

In this section we shall illustrate one way that principal component analysis can be applied to the model reduction problem for the case where the full model is asymptotically stable. The intention is to convey a way of thinking about model reduction which lends itself well to the signal injection view of minimal realization theory. The results are promising but incomplete; future research will hopefully refine the viewpoint and sharpen the tools.

In this section a reduced-order model $(A_R, B_R, C_R)$ will be judged by its impulse response matrix. The error impulse response matrix

$$H_e(t) \triangleq Ce^{At}B - C_R e^{A_R t}B_R$$

characterizes the error. We shall say that a reduced order model is "good" if the largest principal component of $H_e(t)$ over $[0, \infty)$ is "small" compared to the smallest principal component of $Ce^{At}B$, i.e., if

$$\left\| \int_0^\infty H_e(t) H_e^T(t)\, dt \right\|_2^{1/2} \ll \min_{\|v\|=1} \left\{ v^T \left( \int_0^\infty Ce^{At}BB^T e^{A^T t} C^T dt \right) v \right\}^{1/2}. \quad (8)$$

This viewpoint forces an assumption: the map $Ce^{At}B$ may have pathologically small or zero components (e.g., two rows of $C$ are identical), and we must assume that this situation has been corrected by a projection onto an output space of appropriate dimension. To make the condition (8) simpler, it will be further assumed that an output coordinate transformation has been applied so that

$$\int_0^\infty Ce^{At}BB^T e^{A^T t} C^T = I$$

in which case (8) may be replaced by

$$\left\| \int_0^\infty H_e(t) H_e^T(t)\, dt \right\|_2^{1/2} \ll 1. \quad (9)$$

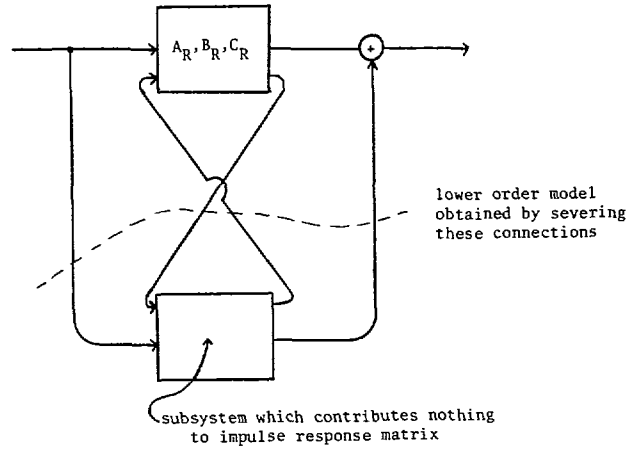


Fig. 5.

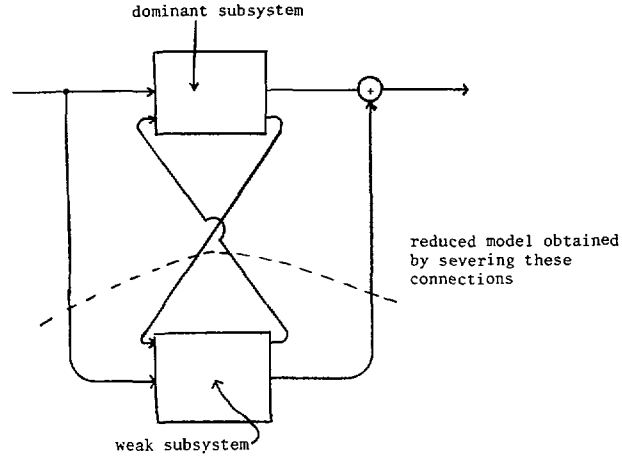


Fig. 6.

### *Reduction by Subsystem Elimination*

Minimal realization theory says that there is an exact lower order model if and only if in some coordinate system the full model can be organized as follows.

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} A_R & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} B_R \\ B_2 \end{pmatrix} \omega(t)$$

$$y(t) = (C_R \quad C_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

where the subsystem $(A_R, B_R, C_R)$ has the same impulse response matrix as the full model. This is illustrated by Fig. 5.

The main idea underlying the model reduction work here is to eliminate any *weak* subsystem which contributes little to the impulse response matrix. In other words, we shall try to reorganize the full model with an internal coordinate transformation as illustrated by Fig. 6. This implicitly defines the meaning of a "dominant" subsystem: it is one whose impulse response matrix is close (in the sense discussed in the beginning of this paragraph) to that of the full model.

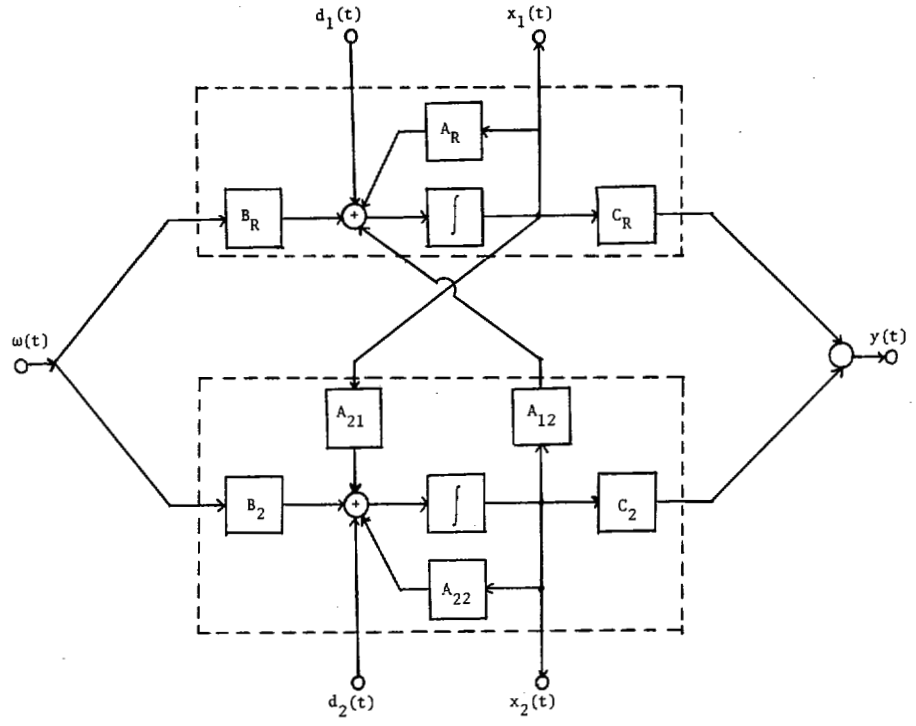


Fig. 7.

It is here that we must face up to a theoretical gap in the paper. We are not able to use this definition of dominance directly, and we introduce instead the concept of *internal dominance* which is natural within the framework developed in Sections II–IV. Toward the end of this section, a conjecture will be advanced concerning the relationship between this concept and actual dominance.

*Internal Dominance*

Consider now the organization of $(A, B, C)$ shown in Fig. 7. Loosely speaking, internal dominance of $(A_R, B_R, C_R)$ means that signal injection tests involving $d_1$, $x_1$ give much stronger signal components than corresponding tests at the second set of terminals $d_2$, $x_2$.

Care must be taken here for the same reasons which called for balancing in Section IV. Clearly, if internal dominance is to relate to actual dominance, it should be invariant under coordinate transformations of the type $x_1(t)=T_1\hat{x}_1(t)$, $x_2(t)=T_2\hat{x}_2(t)$, but the responses to tests are *not* invariant under such transformations. Again the idea of balancing may be used to resolve this problem.

In matrix form, we may express the model in Fig. 1 with arbitrary transformation $x_1(t)=T_1\hat{x}_1(t)$, $x_2(t)=T_2\hat{x}_2(t)$ as follows.

$$\begin{bmatrix}\dot{\hat{x}}_1\\ \dot{\hat{x}}_2\end{bmatrix}=\begin{bmatrix}T_1^{-1}A_RT_1 & T_1^{-1}A_{12}T_2\\ T_2^{-1}A_{21}T_1 & T_2^{-1}A_{22}T_2\end{bmatrix}\begin{bmatrix}\hat{x}_1\\ \hat{x}_2\end{bmatrix}+\begin{bmatrix}T^{-1}B_R\\ T_2^{-1}B_2\end{bmatrix}\omega(t)+\begin{bmatrix}I_k\\ 0\end{bmatrix}\hat{d}_1(t)+\begin{bmatrix}0\\ I_{n-k}\end{bmatrix}\hat{d}_2(t)$$

$$y(t)=(C_RT_1 \quad C_2T_2)\begin{bmatrix}\hat{x}_1\\ \hat{x}_2\end{bmatrix} \tag{10}$$

where $I_k$, $I_{n-k}$ represent $k\times k$ and $(n-k)\times(n-k)$ identity matrices, respectively. With

$$V_1=\begin{pmatrix}I_k\\ 0\end{pmatrix};\qquad V_2=\begin{pmatrix}0\\ I_{n-k}\end{pmatrix}$$

injection of impulses at $\omega(t)$, $d_1(t)$, respectively, gives

$$\hat{X}_1(t)\triangleq T_1^{-1}V_1^Te^{At}B,\qquad \hat{Y}_1(t)\triangleq Ce^{At}V_1T_1,$$

and it follows that

$$\int_0^\infty \hat{X}_1(t)\hat{X}_1^T(t)\,dt=T_1^{-1}(V_1^TW_c^2V_1)T_1^{-1^T}$$

$$\int_0^\infty \hat{Y}_1^T(t)\hat{Y}(t)\,dt=T_1^T(V_1^TW_o^2V_1)T_1.$$

Similar expressions hold for $\hat{X}_2(t)$, $\hat{Y}_2(t)$. ($V_2, T_2$ replace $V_1$, $T_1$.)

These expressions indicate that we may select $T_1$ so that the principal components of $\hat{X}_1(t)$, $\hat{Y}^T(t)$ are aligned with equal magnitudes, i.e.,

$$\int_0^\infty \hat{X}_1(t)\hat{X}_1^T(t)\,dt=\int_0^\infty \hat{Y}_1^T(t)\hat{Y}_1(t)\,dt=\hat{\Sigma}_1^2. \tag{11}$$

Similarly, $T_2$ may be chosen so that

$$\int_0^\infty \hat{X}_2(t)\hat{X}_2^T(t)\,dt=\int_0^\infty \hat{Y}_2^T(t)\hat{Y}_2(t)\,dt=\hat{\Sigma}_2^2. \tag{12}$$

*Definition:* The model (10) is said to be "balanced with respect to $X_1$" iff (11) holds and "balanced with respect to $X_2$" iff (12) holds.

If the model $(A, B, C)$ in Fig. 1 is transformed to (10), which is balanced[8] with respect to both $X_1$, $X_2$, then we shall say that $(A, B, C)$ in Fig. 1 "has been balanced with respect to $X_1$, $X_2$." We are now prepared to give a definition of internal dominance.

*Definition:* The system $(A_R, B_R, C_R)$ is an *internally dominant subsystem* iff in some coordinate system the full model $(A, B, C)$ can be organized as in Fig. 1 such that, when balanced with respect to $X_1$, $X_2$,

$$\|\hat{\Sigma}_1^2\|_F \gg \|\hat{\Sigma}_2^2\|_F.$$

*Internal Dominance and Second-Order Modes*

The following result shows that the second-order modes reflect the existence of internally dominant subsystems.

*Proposition 12:* There exists an internally dominant subsystem of order $k$ if and only if

$$\left(\sum_{i=1}^{k} \sigma_i^4\right)^{1/2} \gg \left(\sum_{i=k+1}^{n} \sigma_i^4\right)^{1/2} \tag{13}$$

where $\sigma_i^2$, $1 \leqslant i \leqslant n$ are the second-order modes.

*Proof:* First sufficiency. Suppose $(A, B, C)$ is internally balanced and let $\Sigma_1 \triangleq \operatorname{diag}\{\sigma_1, \sigma_2, \cdots, \sigma_k\}$, $\Sigma_2 \triangleq \operatorname{diag}\{\sigma_{k+1}, \cdots, \sigma_n\}$. It can be easily verified that $(A, B, C)$ is balanced with respect to $X_1$, $X_2$ with $\hat{\Sigma}_1 = \Sigma_1$, $\hat{\Sigma}_2 = \Sigma_2$. Hence, if (11) holds, then

$$\|\hat{\Sigma}_1^2\|_F = \left(\sum_{i=1}^{k} \sigma_i^4\right)^{1/2} \gg \left(\sum_{i=k+1}^{n} \sigma_i^4\right)^{1/2} = \|\hat{\Sigma}_2^2\|_F.$$

To establish necessity, assume that $(\hat{A}, \hat{B}, \hat{C})$ is balanced with respect to $X_1$, $X_2$ and satisfies $\|\hat{\Sigma}_1^2\|_F \gg \|\hat{\Sigma}_2^2\|_F$. This model is related to the internally balanced model $(A, B, C)$ by a coordinate transformation

$$x(t) = (P_1 \quad P_2)\begin{pmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \end{pmatrix};$$

$$\begin{pmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \end{pmatrix} = \begin{pmatrix} Q_1^T \\ Q_2^T \end{pmatrix} x(t) \qquad \left(Q^T \triangleq P^{-1}\right),$$

and since $PQ^T = I$, we may write

$$\Sigma^2 = \Sigma P Q^T \Sigma = \Sigma P_1 Q_1^T \Sigma + \Sigma P_2 Q_2^T \Sigma.$$

Since $\Sigma P_1 Q_1^T \Sigma$ has rank $k$, it follows from perturbation properties of singular values that

$$\|\Sigma P_2 Q_2^T \Sigma\|_F \geqslant \left(\sum_{i=k+1}^{n} \sigma_i^4\right)^{1/2}.$$

[8]Note that this does *not* imply that the model (8) is internally balanced or that the elements of $\hat{\Sigma}_1^2$, $\hat{\Sigma}_2^2$ are second-order modes.

Hence, we need only show that internal dominance implies that

$$\|\Sigma P_1 Q_1^T \Sigma\|_F \gg \|\Sigma P_2 Q_2^T \Sigma\|_F.$$

To verify this relationship, observe that

$$\int_0^\infty e^{\hat{A}t} \hat{B}\hat{B}^T e^{\hat{A}^T t}\, dt = \begin{pmatrix} P_1^T \\ P_2^T \end{pmatrix} \Sigma^2 (P_1 \quad P_2)$$

$$\int_0^\infty e^{\hat{A}^T t} \hat{C}^T \hat{C} e^{\hat{A}t}\, dt = \begin{pmatrix} Q_1^T \\ Q_2^T \end{pmatrix} \Sigma^2 (Q_1 \quad Q_2)$$

and since $(\hat{A}, \hat{B}, \hat{C})$ is balanced with respect to $X_1$, $X_2$,

$$\hat{\Sigma}_1^2 = P_1^T \Sigma^2 P_1 = Q_1^T \Sigma^2 Q_1$$

$$\hat{\Sigma}_2^2 = P_2^T \Sigma^2 P_2 = Q_2^T \Sigma^2 Q_2.$$

This means that $\Sigma P_1 = U_1 \hat{\Sigma}_1$, $Q_1^T \Sigma = \hat{\Sigma}_1 W_1$, $\Sigma P_2 = U_2 \hat{\Sigma}_2$, $Q_2^T \Sigma = \hat{\Sigma}_2 W_2$ for some orthogonal matrices $U_1$, $W_1$, $U_2$, $W_2$. It follows, therefore, that

$$\Sigma P_1 Q_1^T \Sigma = U_1 \hat{\Sigma}_1^2 W_1; \qquad \Sigma P_2 Q_2^T \Sigma = U_2 \hat{\Sigma}_2^2 W_2.$$

Thus $\|\Sigma P_1 Q_1^T \Sigma\|_F = \|\hat{\Sigma}_1^2\|_F \gg \|\hat{\Sigma}_2^2\|_F = \|\Sigma P_2 Q_2^T \Sigma\|_F$ and the proof is complete. ■

This result, Proposition 12, suggests a natural first step in model reduction: compute the internally balanced model and inspect the second-order modes. If condition (11) is not satisfied, there is no internally dominant subsystem. If (11) is satisfied, then the subsystem corresponding to the first $k$ state variables of the model is internally dominant and, generically, internally balanced and asymptotically stable. This subsystem is precisely that obtained by applying the mechanics of minimal realization (Section II) using $\begin{pmatrix} I_k \\ 0 \end{pmatrix}$ as a working basis for $X_{co}$.

Now let us return to the issue of dominance versus internal dominance. An internally dominant subsystem may be *tested* for dominance by applying principal component analysis to the impulse response error matrix $H_e(t)$. The interesting question is whether it is possible for a dominant subsystem of order $k$ to exist with no corresponding internally dominant subsystem of order $k$.

*Conjecture:* Every dominant subsystem is internally dominant.

*On the Question of Optimality*

Wilson [28], [29] and Applevich [30] have given necessary conditions for optimality of the reduced order model, i.e., necessary conditions for $\int_0^\infty \|H_e(t)\|_F^2 dt$ to be a minimum. It is natural to question the relationship between their work and the model reduction results given here.

The first point to be made is that Wilson and Applevich consider the entire class of $k$th-order $r$-output, $m$-input linear systems as possible reduced-order models of order

$k$. This is a *larger*[9] class than that considered here: those which can be obtained by subsystem elimination. It is not surprising, then, that the reduced model of order $k$ defined by the internally balanced model *is not optimal*. Computational experience of this author suggests, however, that if it corresponds to an internally dominant subsystem, it is *near* optimal.[10]

In the situation where one is trying to find an optimal $k$th-order model (from the large class) where there is no internally dominant subsystem of order $k$, it seems reasonable to use the $k$th-order subsystem defined by the internally balanced model as a *starting point* in the model reduction computations.

## VI. Illustrative Examples

To illustrate internal balancing and the proposed model reduction framework, we shall consider two single input, single output examples. The examples are not particularly important other than for the purpose of illustration.

As stated in Section V, we shall judge potential reduced models on the basis of the relative error in the impulse response

$$\frac{\left(\int_0^\infty H_e(t)^2\,dt\right)^{1/2}}{\left(\int_0^\infty H(t)^2\,dt\right)^{1/2}}.$$

*Example 1:* The first example illustrates that phase canonical models, often chosen for convenience, may be extremely poorly balanced.

$$A=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -50 & -79 & -33 & -5 \end{bmatrix}; \qquad B=\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$C=(50 \quad 15 \quad 1 \quad 0)$$

$$C(sI-A)^{-1}B=\frac{(s+10)(s+5)}{(s+1+j4.9)(s+1-j4.9)(s+1)(s+2)}.$$

The principal component vectors and magnitudes of $e^{At}B$ are given by

$$V_c=\begin{bmatrix} 0 & -0.0643 & 0 & \underline{0.998} \\ -0.0335 & 0 & \underline{0.999} & 0 \\ 0 & \underline{0.998} & 0 & 0.0643 \\ \underline{0.999} & 0 & 0.0335 & 0 \end{bmatrix};$$

$$\Sigma_c=\begin{bmatrix} 0.461 & & & \\ & 0.04183 & & \\ & & 0.01456 & \\ & & & 0.0112 \end{bmatrix}$$

and those of $e^{A^Tt}C^T$ are

$$V_o=\begin{bmatrix} \underline{0.899} & -0.426 & 0.0991 & 0.0225 \\ 0.431 & \underline{0.824} & -0.361 & -0.0738 \\ 0.0721 & 0.368 & \underline{0.927} & -0.00891 \\ 0.0122 & 0.0738 & -0.0206 & \underline{0.997} \end{bmatrix}$$

$$\Sigma_o=\begin{bmatrix} 47.95 & & & \\ & 2.96 & & \\ & & 0.667 & \\ & & & 0.244 \end{bmatrix}.$$

The underlined entries in $V_c$ show that it is essentially antidiagonal (the higher derivatives of the output are most excited by an impulse). For $V_o$ it is the opposite; $V_o$ is essentially diagonal (initial conditions involving lower derivatives cause the greatest output response). The second-order modes are

$$\{\sigma_1^2,\sigma_2^2,\sigma_3^2,\sigma_4^2\}=\{0.576,0.147,0.0904,0.0192\}.$$

The internally balanced model[11] for this system is given by

$$\hat{A}=\begin{bmatrix} -0.5183 & 1.45 & -0.3911 & 0.3501 \\ -1.45 & -2.195 & 4.753 & -1.218 \\ -0.3911 & -4.753 & -0.6297 & 1.196 \\ -0.3501 & -1.218 & -1.196 & -1.657 \end{bmatrix}$$

$$\hat{B}=\begin{bmatrix} 0.7729 \\ 0.8047 \\ 0.3373 \\ 0.2523 \end{bmatrix}$$

$$\hat{C}=(0.7729 \quad -0.8047 \quad 0.3373 \quad -0.2523).$$

Table I summarizes the respective reduced models found by subsystem elimination.

*Example 2:* The following example was considered by Wilson [28].

$$A=\begin{bmatrix} 0 & 0 & 0 & -150 \\ 1 & 0 & 0 & -245 \\ 0 & 1 & 0 & -113 \\ 0 & 0 & 1 & -19 \end{bmatrix}; \qquad B=\begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$C=(0 \quad 0 \quad 0 \quad 1)$$

$$C(sI-A)^{-1}B=\frac{s+4}{(s+1)(s+3)(s+5)(s+10)}.$$

For this system the component vectors and magnitudes of $e^{At}B$ are

[9] It is intuitively clear that the class of models of order $k$ which can be achieved by subsystem elimination *does not* contain all models of order $k$. This fact is shown by example in [31].

[10] Indeed, Example 2 was considered by Wilson who reported a *larger* residual. This undoubtedly arises from the nature of his search algorithm.

[11] It is interesting to observe the "absolute value" symmetry of the model. This property has appeared in every single input, single output example tested by this author.

TABLE I
REDUCED ORDER MODEL PROPERTIES—EXAMPLE 1

| order (k) | $\left(\sum_{i=k+1}^{n}\sigma_i^4\right)^{1/2} \Big/ \left(\sum_{i=1}^{k}\sigma_i^4\right)^{1/2}$ | actual relative error | transfer function |
|---|---|---|---|
| 3 | .03193 | .05988 | $\frac{.0637(s-2.14+j17.38)(s-2.14-j17.38)}{(s+.738)(s+1.3+j4.88)(s+1.3-j4.88)}$ |
| 2 | .1553 | .3332 | $\frac{-.05014(s-54.4)}{(s+1.357+j1.183)(s+1.357-j1.183)}$ |
| 1 | .302 | .4849 | $\frac{.5974}{s+.5183}$ |

$$V_c=\begin{bmatrix} -0.8897 & -0.4495 & 0.0769 & 0.002 \\ 0.4510 & 0.8385 & -0.2951 & -0.008 \\ 0.0706 & 0.3073 & 0.8981 & 0.3065 \\ 0.359\times10^{-2} & 0.01903 & 0.3169 & 0.9483 \end{bmatrix};$$

$$\Sigma_c=\begin{bmatrix} 4.13 & & & \\ & 0.367 & & \\ & & 2.35\times10^{-3} & \\ & & & 2.5\times10^{-4} \end{bmatrix}$$

$$V_o=\begin{bmatrix} 0 & -0.0809 & 0 & +0.9967 \\ -0.987\times10^{-3} & 0 & 0.999 & 0 \\ 0 & 0.9967 & 0 & +0.0809 \\ +0.9999 & 0 & 0.987\times10^{-3} & 0 \end{bmatrix};$$

$$\Sigma_o=\begin{bmatrix} 0.174 & & & \\ & 0.0173 & & \\ & & 4.48\times10^{-3} & \\ & & & 3.68\times10^{-3} \end{bmatrix}.$$

The second-order modes are

$$\{\sigma_1^2,\sigma_2^2,\sigma_3^2,\sigma_4^2\}$$
$$=\{0.0159,0.272\times10^{-2},0.127\times10^{-3},0.8\times10^{-5}\}$$

and the internally balanced model is

$$\hat{A}=\begin{bmatrix} -0.4378 & -1.168 & -0.4143 & 0.05098 \\ 1.168 & -3.135 & -2.835 & 0.3288 \\ -0.4143 & 2.835 & -12.48 & 3.249 \\ -0.05098 & 0.3288 & -3.249 & -2.952 \end{bmatrix}$$

$$\hat{B}=\begin{bmatrix} -0.1181 \\ 0.1307 \\ -0.05634 \\ -0.006875 \end{bmatrix}$$

$$\hat{C}=(-0.1181 \quad -0.1307 \quad -0.05634 \quad 0.006875).$$

Table II summarizes properties of lower order models. It is interesting to note that the relative error of Wilson's second-order model is 0.04097. The fact that his "optimal" model gives higher relative error than the one obtained here must be due to the details of his search algorithm.

## VII. CONCLUDING REMARKS

There are several points in this paper which, in this author's opinion, are natural points of departure for further study. With respect to model reduction, the signal injection results given in Section II provide one such point. The results in Sections IV and V were developed for impulse injection, but this may not be appropriate if, say, one is modeling in an environment where computer control is to be used to implement low gain feedback for disturbance rejection. Preliminary results show that one can "tune" the model reduction tools to certain classes of inputs.

Another point of departure that is probably more important than the one discussed in the preceding paragraph is Section III, which describes very general tools for detecting near linear dependence. The focus of this paper, model reduction as it relates to minimal realization theory, is *very narrow* compared to the domain where the tools are applicable.

One point in this paper might better be described as a "loose end" than a "point of departure." The relationship

TABLE II
REDUCED ORDER MODEL PROPERTIES—EXAMPLE 2

| order (k) | $\left(\sum_{i=k+1}^{n}\sigma_i^4\right)^{1/2} \Big/ \left(\sum_{i=1}^{k}\sigma_i^4\right)^{1/2}$ | actual relative error | transfer function |
|---|---|---|---|
| 3 | .000495 | .001311 | $\frac{4.72\times10^{-5}(s-27.2+j147.6)(s-27.2-j147.6)}{(s+.99)(s+3.477)(s+11.58)}$ |
| 2 | .007882 | .03938 | $\frac{-.003127(s-23.14)}{(s+1.113)(s+2.46)}$ |
| 1 | .1711 | .4321 | $\frac{.01395}{s+.4378}$ |

between subsystem dominance and internal dominance is not clear; the latter concept was clearly created for pragmatic reasons. The relationship between general model reduction and reduction by subsystem elimination is not well understood, either. (A discussion of this is given in [31].)

As a final comment, this work has left the author with a strong bias toward operating directly on *signals* whenever possible. This means, of course, a bias against working with secondary objects such as model parameters. There are at least two reasons which support this bias. First, limitations of physical hardware (measurement accuracies, regions of linearity, etc.) can usually be stated directly in terms of signals. Second, tools for coping with multiple signals (principal component analysis + singular value decomposition) are available.

## ACKNOWLEDGMENT

For discussions during the writing of the original draft of this paper, the author is indebted to A. J. Laub and W. M. Wonham. Of considerable help in the revision process were constructive comments by L. M. Silverman, M. G. Safonov, S. Y. Kung, and the reviewers of the first draft [1].

## REFERENCES

[1] B. C. Moore, "Singular value analysis of linear systems, parts I, II," Dep. Elec. Eng., Univ. Toronto, Toronto, Ont., Syst. Contr. Rep. 7801 and 7802, July 1978; also in *Proc. IEEE Conf. Dec. Contr.*, pp. 66–73.
[2] R. E. Kalman, "Irreducible realizations and the degree of a rational matrix," *SIAM J. Appl. Math.*, vol. 13, no. 2, pp. 520–544, 1965.
[3] ——, "Algebraic structure of finite dimensional dynamical systems," *Proc. Nat. Acad. Sci., U.S.A.*, vol. 54, pp. 1503–1508, 1965.
[4] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *J. Educ. Psych.*, vol. 24, pp. 417–441, 498–520, 1933.
[5] ——, "Simplified calculation of principal components," *Psychometrica*, vol. 1, pp. 27–35, 1935.
[6] G. H. Golub and C. Reinsch, "Singular value decomposition and least squares solutions," *Numer. Math.*, vol. 14, pp. 403–420, 1970.
[7] G. B. Garbow *et al.*, *EISPACK Guide*, *Lecture Notes in Computer Science*, 2nd ed., vol. 6. Berlin: Springer-Verlag, 1976.
[8] A. P. Dempster, *Elements of Continuous Multivariate Analysis*. Reading, MA: Addison-Wesley, 1969.
[9] V. C. Klema and A. J. Laub, "The singular value decomposition: Its computation and some applications," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 164–176, Apr. 1980.
[10] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. New York: Oxford Univ. Press, 1965.
[11] G. E. Forsythe and C. B. Moler, *Computer Solution of Linear Algebraic Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1967.
[12] G. W. Stewart, *Introduction to Matrix Computations*. New York: Academic, 1973.
[13] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Englewood Cliffs, NJ: Prentice-Hall, 1974.
[14] G. Strang, *Linear Algebra and its Applications*. New York: Academic, 1976.
[15] G. Forsythe, M. A. Malcolm, and C. B. Moler, *Computer Methods for Mathematical Computations*, Englewood Cliffs, NJ: Prentice-Hall, 1977.
[16] B. C. Moore, "Principal component analysis in nonlinear systems: Preliminary results," in *Proc. 1979 IEEE Conf. Decision Contr.*, Fort Lauderdale, FL., Dec. 1979, pp. 1057–1060.
[17] C. T. Mullis and R. A. Roberts, "Synthesis of minimum roundoff noise fixed point digital filters," *IEEE Trans. Circuits Syst.*, vol. CAS-23, pp. 551–562, Sept. 1976.
[18] E. G. Lee and L. Marcus, *Foundations of Optimal Control Theory*. New York: Wiley, 1967.
[19] R. W. Brockett, *Finite Dimensional Linear Systems*. New York: Wiley, 1970.
[20] R. H. Bartels and G. W. Stewart, "Solution of the matrix equation $AX+XB=C$," *Commun. Ass. Comput. Mach.*, vol. 15, pp. 820–826, 1972.
[21] B. Friedland, "Controllability index based on conditioning number," *Trans. ASME, J. Dynamic Syst., Measurement, Contr.*, pp. 444–445, Dec. 1975.
[22] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*, 2nd ed. New York: Springer-Verlag, 1979.
[23] L. Pernebo and L. M. Silverman, "Balanced systems and model reduction," in *Proc. 18th IEEE Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 12–14, 1979.
[24] J. Rissanen, "Recursive identification of linear systems," *SIAM J. Contr.*, vol. 9, pp. 420–430, 1971.
[25] H. P. Zeiger and J. McEwen, "Approximate linear realizations of given dimension via Ho's algorithm," *IEEE Trans. Automat. Contr.*, vol. AC-19, p. 153, 1974.
[26] L. S. DeJong, "Numerical aspects of realization algorithms in linear systems theory," Ph.D. dissertation, Eindhoven Univ. Technol., Eindhoven, The Netherlands, 1975.
[27] S. Kung, "A new identification and model reduction algorithm via singular value decompositions," in *Proc. 12th Ann. Asilomar Conf. Circuits, Syst. Comput.*, Nov. 1978.
[28] D. A. Wilson, "Optimum solution of model-reduction problem," *Proc. IEEE*, vol. 117, pp. 1161–1165, 1970.
[29] ——, "Model reduction for multivariable systems," *Int. J. Contr.*, vol. 20, pp. 57–64, 1974.
[30] J. D. Applevich, "Approximation of discrete linear systems," *Int.*

*J. Contr.*, vol. 17, pp. 565–575, 1973.
[31] P. J. M. Pinquet, "State space formulation of a class of model reduction methods," M. Phil. thesis (supervised by K. Glover), Dep. Eng., Univ. Cambridge, Cambridge, England, 1978.

**Bruce Moore** was born in Russellville, AR, on January 8, 1947. He received the B.S. degree in electrical engineering from Louisiana State University, Baton Rouge, in 1969, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 1970 and 1973, respectively.

From 1972 to 1975 he was on the faculty of the Department of Computer Science, Louisiana State University, Baton Rouge. Since 1975 he has been with the Department of Electrical Engineering, University of Toronto, Toronto, Canada, currently as an Associate Professor. His interests include multivariable signal analysis, computer process control, and numerical linear systems theory.

# Principal Gains and Principal Phases in the Analysis of Linear Multivariable Feedback Systems

IAN POSTLETHWAITE, MEMBER, IEEE, JOHN M. EDMUNDS, AND ALISTAIR G. J. MACFARLANE, SENIOR MEMBER, IEEE

***Abstract*—The concepts of *principal gain* and *principal phase* are introduced for linear multivariable systems, and their use in the analysis of feedback behavior is demonstrated. A sufficient Nyquist-type stability criterion is presented in terms of these quantities and is used to characterize the robustness of the closed-loop stability property when the system model is subjected to a linear perturbation (either multiplicative or additive) at any point in the feedback configuration. The results presented are less conservative than those obtained via the small gain theorem.**

## I. Introduction

In recent years there has been a revival of interest in the development and application of frequency-response techniques to the design and analysis of linear multivariable feedback control systems (see, for example, [1]–[3]). One of the major reasons for this has been the availability of increasingly inexpensive computers and the consequent increase in all branches of engineering of interactive computing facilities to assist in design and analysis. The frequency-response approach is particularly attractive in this context since, having a strong complex-variable content, it lends itself well to graphical representations which are a way of presenting results that engineers have found helpful in the past, as shown by the success of the classic Nyquist and Bode plots. Technically, frequency-response methods have the advantage of being largely insensitive to small errors in a system model. Should the actual system suffer from large parameter variations, however, or should the model be very inaccurate because of various approximations and uncertainties, then the control system should naturally be designed to have a *large* degree of stability. The mere presence of feedback is not sufficient to guarantee the robustness of the stability property, and so techniques for assessing the relative stability of a multivariable design are required. This problem has been studied by Doyle [4] who characterizes the robustness of the closed-loop stability property in terms of the spectral norm of an appropriate frequency-response matrix. In this paper we give a less conservative characterization of the robustness of the closed-loop stability property in terms of a Nyquist-type plot by introducing phase information.

The phase information used terms from the polar decomposition of a complex matrix [5] which is defined as follows. Analogous to the polar form of a complex number, a complex matrix $T$ can be represented in the forms

$$T = UH_R \tag{1.1}$$

$$T = H_L U \tag{1.2}$$

Manuscript received March 11, 1980; revised September 22, 1980. This work was supported by the Science Research Council.

I. Postlethwaite was with the Research and Development Center, General Electric Company, Schenectady, NY 12301. He is now with the Department of Engineering, Control and Management Systems Division, Cambridge University, Cambridge, England.

J. M. Edmunds and A. G. J. MacFarlane are with the Department of Engineering, Control and Management Systems Division, Cambridge University, Cambridge, England.

0018-9286/81/0200-0032$00.75 