Orthogonal Filters for Model Error Compensation in the Control of Nonrigid Spacecraft

Robert E. Skelton*
Purdue University, West Lafayette, Ind.
and
Peter W. Likins†
Columbia University, New York, N. Y.

This paper seeks to improve the convergence properties of state estimators as used in spacecraft controller designs, when the linearized models upon which the estimators are based are subject to parameter errors, truncated modes, and neglected disturbances. Instead of choosing mode shapes (which are orthogonal in space) multiplied by time varying coefficients (a conventional approach for modeling elastic modes) to represent the truncated modes, the "model error vector" discussed herein is approximated, over short "observation windows," τ units long, by functions which are orthogonal over the time interval τ, where the coefficients of the orthogonal functions are automatically updated via the use of real-time measurements from the system. The device which updates the coefficients of the orthogonal functions is called an orthogonal filter and takes on the form of a state estimator for the synthetic modes of a "model error system" which generates the orthogonal functions. The method is illustrated for a 14th-order model of a flexible spacecraft, resulting in 2nd-, 3rd-, and 4th-order controllers.

I. Introduction

ONE aspect of spacecraft controller design which complicates the design decisions is that the design cannot be fully verified before "flight." This is because most lightweight, highly flexible spacecraft structures cannot be reliably tested in a 1-g environment, and there is therefore some uncertainty concerning the system dynamics which cannot be removed before flight. This places a heavy burden of responsibility upon the analytical methods utilized in the design. In spite of the existence of elaborate computer programs which generate modal data for the structure, there are various assumptions of mode shapes, modal damping, and mode truncation which contribute to errors in the final set of differential equations (which we will call the controller design model) upon which the feedback control system design is based. As a consequence of these circumstances, any spacecraft control system which is worthy of serious consideration must be tolerant of these various types of modeling errors: parameter errors, truncated modes, and the effects of uncertainties in such disturbances as natural environmental disturbances and amplifier drifts and biases. (The accommodation of known disturbances is discussed in Refs. 1-3.) The application of the linear regulator and estimation theory of modern control must be approached with caution since the theory and its promises rely upon the absolute fidelity of the mathematical model of the physical system. For example, if the linearized model of the physical system is completely controllable and observable, the theory promises that the closed-loop system performance can be "arbitrarily good" in the sense that the eigenvalues of the closed-loop system can be arbitrarily placed. There is nothing in the theory which warns of limits of model fidelity, and the difference between the model's predicted performance and the system's actual performance (which comparison brings the realization that the system eigenvalues *cannot* be arbitrarily placed) can be attributed to modeling errors. Some authors approach this deficiency of the theory by seeking adaptive control schemes,[4,5] while others seek better methods of modeling[6] and model reduction[7-9] for use in nonadaptive controllers.

There has been a concerted effort recently[2,3] to accommodate model errors of the type described in the design of linear optimal regulators and state estimators. The concept of the "orthogonal filter" was introduced in Ref. 2, and this paper shows how to design the orthogonal filter and gives some theorems on filter convergence. The orthogonal filter is designed to approximate, in real time, the model error vector and to utilize this approximation to improve the state estimate of the model. These results take the form of a time invariant linear dynamical controller, although an adaptive version of the controller is also possible. Section II reviews the concepts of model error vectors. Section III presents the full-order and the reduced-order orthogonal filters and Sec. IV discusses the performance results of various controller designs for a highly flexible spacecraft using a standard Kalman filter design and an orthogonal filter design. In this introductory paper no general conclusions are yet available concerning the successful compensation of arbitrary modeling errors in closed-loop control systems, but some basic theorems are developed relating to convergence of the state estimator.

II. Characterizations of the Model Error Vectors

When a flexible spacecraft is idealized as a nonrotating rigid body with an elastic appendage, the linearized equations may take the form,

$$J^*\ddot{\theta} - \Delta^T\ddot{q} = T_d + T_c \qquad \theta \epsilon R^3, T_c \epsilon R^3 \tag{1a}$$

$$M\ddot{q} + Kq - \Delta\ddot{\theta} = f \qquad q \epsilon R^{\bar{N}} \text{(generalized coordinate)} \tag{1b}$$

where T_d, T_c represent disturbance and control torques applied to the rigid body with attitude described by θ, and f represents torques and forces applied to the appendage. For

Received July 8, 1977; revision received Sept. 19, 1977.

Index category: Guidance and Control.

*Assistant Professor. Member AIAA.

†Professor and Dean of Engineering. Member AIAA.

more complete information about the inertia J^*, the rigid to elastic appendage coupling Δ, the mass matrix M, and the stiffness matrix K, see Ref. 6. Model (1) is incomplete. The dynamical subsystems which generate $f(t)$, $T_c(t)$ and $T_d(t)$ and the dynamics of sensors must be described before meaningful truncations of model (1) can occur. Otherwise, significant interaction between these subsystems and the vibrations of the appendage (1b) may be overlooked. See Ref. 7 for a discussion of seven different criteria for truncation of Eq. (1). Coordinate truncation of Eq. (1b) is usually performed after a change of coordinates. Define the new coordinates by

$$q=\Phi\eta=\sum_{i=1}^{\nu}\phi^i\eta_i+\sum_{i=\nu+1}^{\bar{N}}\phi^i\eta_i \tag{2}$$

where ϕ^i are vectors which are orthogonal in the weighted sense

$$\phi^{j^T}M\phi^k=\delta_{jk} \qquad (\delta_{jk}=\text{Kronecker delta})$$

$$\phi^{j^T}K\phi^k=\delta_{jk}\sigma_j{}^2 \qquad \text{for all } j,k\in\{1,\bar{N}\}$$

Truncating the $\bar{N}$ second order modes of vibration to ν modes is equivalent to neglecting the model error vector $e(t)$ in the reduced model

$$J^*\ddot{\theta}-\Delta^T\bar{\Phi}\ddot{\bar{\eta}}=T_c+T_d+e(t) \qquad \Phi=[\bar{\Phi},\bar{\bar{\Phi}}] \tag{3a}$$

$$\ddot{\bar{\eta}}+2\bar{\zeta}\bar{\sigma}\dot{\bar{\eta}}+\bar{\sigma}^2\bar{\eta}=\bar{\Phi}^T(f+\Delta\ddot{\theta}), \quad \bar{\eta}\in R^\nu, \eta^T=(\bar{\eta}^T,\bar{\bar{\eta}}^T) \tag{3b}$$

where $e(t)$ evolves from the "model error system"

$$e(t)=\Delta^T\sum_{i=\nu+1}^{\bar{N}}\phi^i\ddot{\eta}_i(t)=\Delta^T\bar{\bar{\Phi}}\ddot{\bar{\bar{\eta}}} \tag{4a}$$

$$\ddot{\bar{\bar{\eta}}}+2\bar{\bar{\zeta}}\bar{\bar{\sigma}}\dot{\bar{\bar{\eta}}}+\bar{\bar{\sigma}}^2\bar{\bar{\eta}}=\bar{\bar{\Phi}}^T(f+\Delta\ddot{\theta}) \tag{4b}$$

The diagonal matrices $\bar{\zeta}\ \bar{\sigma}\ \bar{\bar{\zeta}}\ \bar{\bar{\sigma}}$ in Eqs. (3) and (4) are defined as $\bar{\zeta}\triangleq\text{diag } \zeta_i$, $i=1,2,...\nu$, $\bar{\bar{\zeta}}\triangleq$ diag ζ_i, $i=\nu+1,...\bar{N}$, $\bar{\sigma}\triangleq$ diag $\sigma_i, i=1,...\nu, \bar{\bar{\sigma}}\triangleq$ diag σ_i, $i=\nu+1,...\bar{N}$, where the modal damping ζ_i has been added to the system.

There may be extraneous information in Eq. (4) which is not essential to include in the control design process. Also, many high frequency modes which are usually poorly modeled are contained in Eq. (4), and the controller order increases by two for every second order mode kept in the reduced design model [Eq. (3)]. [The order of the design model (3) should be kept as low as possible in order to keep the flight controller complexity at a minimum.] On the other hand, complete neglect of the model error vector $e(t)$ may lead to controllers which are highly sensitive to the effects of the modeling errors inherent in the reduced model (3). For the purposes of model error approximation we choose not to represent the model error vector in the manner of Eq. (4) as a collection of *orthogonal "mode shapes"* ϕ^i (which are often assumed) multiplied by time varying coefficients $\ddot{\eta}_i(t)$ [which evolve from a high order and perhaps poorly modeled subsystem (4)]. We will instead represent the model error vector as some arbitrary collection of prescribed functions of time which are *orthogonal over an interval* τ. The coefficients of the orthogonal functions are to be automatically updated by the use of real time measurements from the system in an "adaptive curve fitting" spirit. These suggestions are made more precise in section III. For the present, we wish to generalize the preceding notions of model error vectors to include parameter errors, truncated modes and neglected disturbances. We will return to the flexible spacecraft problem in section IV.

The assumption is now made that the larger model S_1 which is to be used for controller evaluation, and the truncated model S_3 are placed in state variable form,

Evaluation model

$$S_1\begin{cases}\dot{x}^1=A^1x^1+B^1u+\Gamma^1w^1 & x^1\in R^{n_1},u\in R^m\\ y^1=C^1x^1 & y^1\in R^k\\ z^1=M^1x^1+v^1 & z^1\in R^l\end{cases} \tag{5}$$

Reduced model

$$S_3\begin{cases}\dot{x}^3=A^3x^3+B^3u & x^3\in R^{n_3},n_3<n_1\\ y^3=C^3x^3 & y^3\in R^k\\ z^3=M^3x^3 & z^3\in R^l\end{cases} \tag{6}$$

where y^1 is the vector we wish to regulate to zero by some control $u(t)$ and $z^1(t)$ represents the l measurements available for use in a feedback controller. To explain the deficiencies of model S_3 (complete with parameter errors, truncated modes, and neglected disturbances) with respect to S_1 it is convenient to use the model error definitions e_{xy}^{31} and e_z^{31}.

$$\dot{x}^3=A^3x^3+B^3u+e_{xy}^{31}(t) \tag{7}$$

$$\left.\begin{aligned} y^2&\triangleq C^3x^3 &&=y^1(t)\\ z^2&\triangleq M^3x^3+e_z^{31}(t)&&\equiv z^1(t)\end{aligned}\right|\ \text{for some choice of } e_z^{31}(t),e_{xy}^{31}(t)$$

When the evaluation model S_1 quite accurately represents the physical system, then e_z^{31} represents the difference between the measurements $z^1(t)$ and those predicted by $z^3(t)$. The actual model error vector can never be completely known a priori, but various attempts to approximate it by specifying a "model error system" of the form,

$$S_e\begin{cases} e=\begin{pmatrix}e_{xy}^{31}\\ e_z^{31}\end{pmatrix}=\begin{bmatrix}P_x\\ P_z\end{bmatrix}s=Ps \quad s\in R^d,d<n_1-n_3\\ \\ \dot{s}=Ds+A_xx^3+B_su\end{cases} \tag{8}$$

lead to a number of well-known results [by appropriate choices for (P, D, A_x, B_s)] when the model error system in Eq. (8) is augmented to the reduced model S_3 in Eq. (6). The controllers discussed in this paper are all defined by a dynamical system of the form

$$\left.\begin{aligned} u&=-G(S_2)\hat{x}^2 &&\text{control law}\\ \hat{x}^2&=\hat{x}^2(z,S_2) &&\text{state estimator}\end{aligned}\right\}\text{controller} \tag{9}$$

where the controller design model S_2 represents the composite of model S_3, as represented in Eq. (7), augmented by the model error system S_e. The output of the controller is the vector u which represents those m entities which are manipulated for control [for example, if u is defined by $u\triangleq T_c$ where T_c is defined in Eq. (3), then $m=3$]. The vector $z(t)$ represents the actual l measurements, and serves as the input to the controller. The vector $\hat{x}^2$ is the estimate of the state vector of the controller design model S_2. The control law might be selected to minimize

$$\frac{1}{2}\int_0^T(y^{2^T}Qy^2+u^TRu)\mathrm{d}t+\frac{1}{2}y^{2^T}(t)Q_Ty^2(T) \tag{10}$$

subject to model S_2, yielding

$$u = -R^{-1}B^{2^T}K(t,T,A^2,B^2,C^2)\hat{x}^2(z,A^2,B^2,C^2,M^2) \qquad (11a)$$

where

$$\left.\begin{aligned} &\dot{K} = -KA^2 - A^{2^T}K + KB^2R^{-1}B^{2^T}K - C^{2^T}QC^2, \\ &K(T) = C^{2^T}Q_TC^2 \\ &Q>0, R>0, Q_T \geq 0 \end{aligned}\right\} \qquad (11b)$$

although we will usually be interested in the degenerate case where $T \to \infty$.

It was shown in Ref. 2 that special choices for the parameters of the model error system (P, D, A_x, B_s) lead to the singular perturbation controller.[10] This method is useful if the dominant errors in model S_3 are due to truncated modes. The state variables s_i of the model error system in this case represent the truncated coordinates. If the dominant model errors are due to small parameter uncertainties, then the parameters (P, D, A_x, B_s) may be selected to give s_i the meaning $s_i \triangleq \partial x_3/\partial p_i$, where p_i, $i=1 \to d$ represents parameters considered uncertain. In this circumstance $P=0$ and the equations for the model S_3 and the sensitivity subsystem S_e are uncoupled. The control policy which minimizes Eq. (10) [where for minimum sensitivity we may take $y^{2^T} \triangleq \{y^{3^T}, (P_vS)^T\}$ for some P_v, (see Ref. 3)] tends to reduce system sensitivity to parameter errors.[11,12] If the dominant model errors are due to neglected disturbances and differential equations are written to describe the disturbances, then the model error system S_e becomes the disturbance system and the disturbance accommodating controller methods[1,13] may be useful. If the model error vectors e_{xy}^{31} and e_z^{31} are characterized as random processes with only means and covariances available, then the minimization of the expected value of Eq. (10) leads to the Kalman filter.

Beginning with the best controller design model S_2 is obviously important in the controller design problem, and research in model reduction methods is very important. However, this paper is concerned with the opposite point of view. That is, we deal with the task of improving state estimates (and the resulting controller designs) of any given low-order model by partially compensating for the inevitable presence of modeling errors; not by exactly modeling the truncated effects as in Eq. (4), but by the addition of certain synthetic modes designed to approximate only the dominant effects of Eq. (4). Likins[14] has shown that the effects of truncation of any number of modes of Eq. (1) can be exactly compensated *in the steady state* by the construction of a model error system [Eq. (4)] with just four synthetic modes. In this work we will also use synthetic modes (in S_e) to approximate the effects of model errors but attention will not be restricted to steady-state effects. The special case of Eq. (8) which will be of primary interest in this paper leads to the orthogonal filter of the following section.

III. Orthogonal Filters

The objective of the orthogonal filter[2] is to approximate the actual unknown model error vector with orthogonal functions in real time and to use this estimate of the model error vector to improve the estimate of the state vector x^3. The resulting estimate of the state vector x^3 is used for state feedback control. It may also happen that the estimate of the model error vector is useful in the control policy. The representation of the model error vector as a collection of orthogonal functions is accomplished by the choice [a special case of Eq. (8)],

$$e = Ps \qquad e \in R^r \qquad (12a)$$

$$\dot{s} = Ds \qquad s \in R^d \qquad (12b)$$

where the choices for P and D depend upon the particular orthogonal functions chosen. Suggestions for choosing P will not come before section IIIC. In this section we will be concerned with the choice of D. In Ref. 2 various D matrices are constructed which generate, as eigenfunctions of Eq. (12), Jacobi polynomials and some of their special cases: Legendre, Chebyshev, and Fourier. For example, if D is chosen as

$$D = \frac{2}{\tau}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 3 & 0 & 6 & 0 & 0 \\ 0 & 8 & 0 & 8 & 0 & \ddots \end{bmatrix}_{d\times d} = \frac{2}{\tau}D_0 \qquad (13)$$

then the particular solution of Eq. (12) which has the initial condition $s_0^T = (1, -1, 1, -1, \ldots)$ yields $e(t) = PT$, where T is a column of Chebyshev polynomials which may be expressed in time, $t \in [0,\tau]$ or normalized time, $\sigma \in [-1, 1]$ by $T_i = \cos[(i-1)\cos^{-1}\sigma]$, $i=1,2,\ldots d$, with $T^T \triangleq (T_1, T_2 \ldots T_d)$, where $\sigma \triangleq 2(t/\tau) - 1$, and T_i is a Chebyshev polynomial of the first kind of degree $i-1$. These functions are orthogonal on the interval $\sigma \in [-1,1]$ with respect to the weight $g(\sigma) = (1-\sigma^2)^{-1/2}$. That is, the matrix

$$\Lambda \triangleq \int_{-1}^{1} T(\sigma)T^T(\sigma)g(\sigma)\mathrm{d}\sigma$$

$$= \int_0^\tau \frac{2}{\tau}T(t)T^T(t)g(t)\mathrm{d}t = \frac{\pi}{2}\left[\begin{array}{c|c} 2 & 0 \\ \hline 0 & I \end{array}\right]_{d\times d} \qquad (14)$$

is diagonal, constant, and nonsingular. The model error system in Eq. (12) is obtained by noting that, from Eq. (13)

$$\frac{\mathrm{d}T}{\mathrm{d}\sigma} = D_0T, \text{ and } \frac{\mathrm{d}T}{\mathrm{d}\sigma} = \frac{\mathrm{d}T}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\sigma} = \frac{\mathrm{d}T}{\mathrm{d}t}\frac{\tau}{2} \qquad (15)$$

together with the observation that $T^T(t=0) = T^T(\sigma=-1) = (1, -1, 1, -1, \ldots)$. A more useful fact than the above result is that *any* solution of Eq. (12) (arbitrary initial condition s_0) yields $e(t)$ as a sum of Chebyshev polynomials, $e = \bar{P}T$, where the matrix $\bar{P}$ is a function of s_0. In fact, a stronger result is stated in theorem 1.

Theorem 1: The Chebyshev error system (12) can generate, for *each* of the r components of $e(t)$, an arbitrary sum of Chebyshev polynomials of degree $\rho - 1$, where ρ is the maximum integer for which the matrix $\mathcal{Q}$ with elements $\mathcal{Q}_{ij} \triangleq P\mathcal{Q}_i\mathcal{Q}_j^TP^T$, $i,j=1,\ldots\rho$, represents a nonsingular matrix, where the $d \times d$ matrices $\mathcal{Q}_i$ are defined by

$$\mathcal{Q}_1 = I + D_0 + \frac{3}{2}\frac{1}{2!}D_0^2 + \frac{5}{2}\frac{1}{3!}D_0^3 + \ldots = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ 13 & 18 & 6 & 1 \\ & & & & \ddots \end{bmatrix} \qquad (16a)$$

$$\mathcal{Q}_2 = D_0 + 2\frac{1}{2!}D_0^2 + \frac{15}{4}\frac{1}{3!}D_0^3 + \ldots = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 18 & 24 & 6 & 0 \\ & & & & \ddots \end{bmatrix} \qquad (16b)$$

$$\mathcal{Q}_3 = \frac{1}{2}\frac{1}{2!}D_0^2 + \frac{3}{2}\frac{1}{3!}D_0^3 + \ldots = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 6 & 6 & 0 & 0 \\ & & & & \ddots \end{bmatrix} \tag{16c}$$

$$\mathcal{Q}_4 = \frac{1}{4}\frac{1}{3!}D_0^3 + \ldots = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ & & & & \ddots \end{bmatrix} \tag{16d}$$

The above matrices $\mathcal{Q}_i$ are applicable for any $d=1, 2, 3, 4$ by construction from the top left proceeding diagonally downward to form the correct $d \times d$ matrices up to $d=4$. Theorem 1 asserts that there exists an initial condition vector s_0 such that $r\rho$ of the coefficients $\bar{P}$ of the Chebyshev polynomial representation of the model error vector, $e=\bar{P}T$, as generated by Eq. (12), can take on arbitrary values. On the strength of this fact we can guarantee that, over an interval τ, any *particular* coefficients which are appropriate for fitting the *actual* model error vector can be automatically determined by employing state estimation techniques with model S_2. Before proving theorem 1, it is helpful to consider the following theorem 2.

Theorem 2: The Chebyshev error system in Eq. (12) is observable if and only if the last column of the matrix P is nonzero.

Discussion of Theorem 2: It may be shown by direct (but long and tedious) construction that the determinant of the "observability matrix,"

$$[P^T, D^T P^T, \ldots D^{d-1^T} P^T] \begin{bmatrix} P \\ PD \\ \vdots \\ PD^{d-1} \end{bmatrix} = \sum_{k=0}^{d-1} D^{k^T} P^T P D^k \tag{17}$$

has the form $P^{d^T} P^d(\cdot)$ where P^d is the dth column of P, and $(\cdot)$ is a positive number with the lower bound $(\cdot) \geq P^{d^T} P^d$. The matrix D is given by Eq. (13).

Proof of Theorem 1: The matrix $\bar{P}$ is constructed as follows. The solution of Eq. (12) is

$$e = Pe^{D_0\hat{\sigma}} s_0 \quad \text{where } \hat{\sigma} = \sigma - \sigma_0 = \sigma + 1 \tag{18}$$

But since, for the D_0 defined by Eq. (13),

$$e^{D_0\hat{\sigma}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \hat{\sigma} & 1 & 0 & 0 \\ 2\hat{\sigma}^2 & 4\hat{\sigma} & 1 & 0 \\ 3\hat{\sigma}+4\hat{\sigma}^3 & 12\hat{\sigma}^2 & 6\hat{\sigma} & 1 \\ & & & & \ddots \end{bmatrix}_{d \times d} \tag{19a}$$

using the definition of T_i allows the expression

$$e^{D_0\hat{\sigma}} = \left[\begin{array}{l|l|l|l} T_0 & 0 & 0 & 0 \\ T_0 + T_1 & T_0 & 0 & 0 \\ 3T_0 + 4T_1 + T_2 & 4T_0 + 4T_1 & T_0 & 0 \\ 13T_0 + 18T_1 + 6T_2 + T_3 & 18T_0 + 24T_1 + 6T_2 & 6T_0 + 6T_1 & T_0 \end{array}\right] \tag{19b}$$

which can be rearranged to write

$$e^{D_0\hat{\sigma}} s_0 = [\mathcal{Q}_1 s_0, \mathcal{Q}_2 s_0, \mathcal{Q}_3 s_0, \mathcal{Q}_4 s_0 \ldots] T \tag{19c}$$

where the $d \times d$ matrices $\mathcal{Q}_i$ are defined in Eq. (16). From Eqs. (18) and (19c),

$$e = \bar{P}T = [P\mathcal{Q}_1 s_0, P\mathcal{Q}_2 s_0, P\mathcal{Q}_3 s_0, P\mathcal{Q}_4 s_0, \ldots] T \tag{20}$$

and it is clear from Eq. (20) that there exists an s_0 which will simultaneously yield ρ arbitrarily specifiable columns $\bar{P}^i$, $i=1, 2, \ldots \rho$, of the matrix $\bar{P}$, if $\{s_0^T [\mathcal{Q}_1^T P^T, \ldots \mathcal{Q}_\rho^T P^T] = s_0^T \bar{\mathcal{Q}}^T = [\bar{P}^1, \bar{P}^2, \ldots \bar{P}^\rho) = \text{arbitrary vector}\}$ has a solution, s_0. A solution is guaranteed if the $r\rho \times d$ matrix $\bar{\mathcal{Q}}$ has maximal rank and ρ is an integer satisfying $r\rho \leq d$. More generally, however, ρ is the largest integer for which $\mathcal{Q}$ is a nonsingular matrix (this follows by forming the inner product of the matrix $\bar{\mathcal{Q}}$ with itself, given that $r\rho \leq d$). Q.E.D.

It may also be noted that the $d \times dr$ matrix $\bar{\bar{\mathcal{Q}}}^T \triangleq [\mathcal{Q}_1^T \; P^T, \mathcal{Q}_2^T P^T, \ldots \mathcal{Q}_d^T P^T]$ has rank d if and only if the Chebyshev error system, Eq. (12) or Eq. (15), is observable. To see this, expand $\bar{\mathcal{Q}}$ in terms of the columns of P using Eq. (16) to get

$$\bar{\bar{\mathcal{Q}}} = \begin{bmatrix} P^1 + P^2 + 3P^3 + 13p^4 + .. & P^2 + 4P^3 + 18P^4 + .. & P^3 + 6P^4 + .. & P^4 + .. \\ P^2 + 4P^3 + 18P^4 + .. & 4P^3 + 24P^4 + .. & 6P^4 + .. & 0 \\ P^3 + 6P^4 + .. & 6P^4 + .. & 0 & 0 \\ P^4 + .. & 0 & 0 & 0 \\ & & & & \ddots \end{bmatrix}_{rd \times d} \tag{21}$$

where the expansion has been continued out to the fourth column (as for a case of $d=4$). Thus, $\bar{Q}$ has maximal rank if and only if the last column of P (in this case P^4) is nonzero, which is the same as the condition for observability (from theorem 2). It should also be mentioned that observability of the pair (D,P) is neither necessary nor sufficient to establish a Chebyshev error system with a specified number ρ of arbitrary columns of the $\bar{P}$ matrix. In the next section a state estimator is employed to claim the benefits of theorem 1.

A. Full-Order Orthogonal Filters

Theorem 1 allows the conclusion that if s_0 could somehow be updated in response to the *actual* measurements, then the actual model error vector can be approximated as a sum of Chebyshev polynomials [evolving from Eq. (12)] whose coefficients can change from time to time in an arbitrary way. Since any square-integrable function of time $(\int f^T f dt < \infty)$ can be fitted in the mean-square sense with orthogonal functions (Chebyshev in the present case) of some degree, we have only to examine the questions a) Over what interval τ should the functions be orthogonal? and b) What degree of

the polynomials is appropriate for the modeling of $e(t)$? Questions a and b are obviously related. [The smaller τ allows fitting of $e(t)$, $t\in[0,\tau]$ with polynomials of lower degree (smaller d). Also, for pointwise fitting of $e(t)$ the assumption of continuous $e(t)$ is necessary.] We continue our collection of facts with the following statement based upon a given τ.

Theorem 3: For any fixed observation window $\tau>0$, there exists an integer d [the order of the Chebyshev model error system in Eq. (12)] such that, in the presence of any model errors associated with the model S_3 (parameter errors, truncated modes, and neglected disturbances), the state estimate $\hat{x}^3(t)$ converges to $x^3(t)$ if the estimate $\hat{x}^3(t)$ is given by the full-order Chebyshev filter

$$\hat{x}^3=[I,0]\hat{x}^2 \qquad z\in R^l, \qquad \hat{x}^3\in R^{n_3} \tag{22a}$$

$$\dot{\hat{x}}^2=[A^2-\hat{G}M^2]\hat{x}^2+B^2u+\hat{G}z, \quad \hat{x}^2\in R^{n_3+d}, u\in R^m \tag{22b}$$

where $\hat{G}$ stabilizes $[A^2-\hat{G}M^2]$ such that the estimator poles satisfy, for $i=1,2,\ldots n_3+d$,

$$-Re\lambda_i[A^2-\hat{G}M^2]\gg 1/\tau \tag{23a}$$

and (A^2, B^2, M^2) are given by

$$\left\{ A^2=\begin{bmatrix} A^3 & P_x \\ 0 & D\end{bmatrix}, \quad \begin{matrix} B^{2^T}=[B^{3^T},0] \\ M^2=[M^3,P_z]\end{matrix} \right\} \tag{23b}$$

where it is assumed that (A^2, M^2) and (D, P) are both observable pairs, and D is given by Eq. (13).

Proof of Theorem 3: The generality of the claim of theorem 3 requires us to consider in the proof the model error vector $e(t)$ to be an arbitrary square-integrable single valued vector function of time. We must show that this model error vector can be determined from the measurements and that the state estimate for x^3 takes into account the presence of e according to Eq. (7). Let time be divided into equal intervals τ units long. First, consider the actual model error vector $e(t)$ over the interval $t\in[0,\tau]$. To prove theorem 3, we must show 1) that the arbitrary function of time $e(t)$, $t\in[0,\tau]$ can be expressed as a solution of Eq. (12) for some s_0 and some d; 2) that the particular solution of Eq. (12) which actually yields $e(t)$, $t\in[0,\tau]$, can be determined prior to $t=\tau$, using the real-time measurements $z(t)$; 3) that once $e(t)$, $t\in[0,\tau]$ is known, $x^3(t)$ can be uniquely determined prior to $t=\tau$; and 4) that once $x^3(t)$ is determined on the interval $t\in[0,\tau]$ the preceding events repeat on the interval, $t\in[\tau,2\tau]$, etc.

Part 1: Since any square-integrable function can be fitted with orthogonal functions (in the present case Chebyshev polynomials) of some degree, the proof of part 1 follows immediately from the statement of theorem 1, together with the requirement that (D, P) be an observable pair.

Parts 2 and 3: Observability of the pair (A^2, M^2) guarantees that there exists a $\hat{G}$ such that $\hat{x}^2\rightarrow x^2$ arbitrarily fast if Eq. (24) defines x^2, Eq. (25) defines $\hat{x}^2$,

$$\dot{x}=A^2x^2+B^2u \tag{24}$$

$$\dot{\hat{x}}=[A^2-\hat{G}M^2]\hat{x}^2+B^2u+\hat{G}z \tag{25}$$

and if $z=M^2x^2$ defines the actual measurements. For this last statement, we rely on the accurate modeling of e (large enough d) as discussed in part 1, so that Eq. (7) holds. The estimator is "fast enough" to accomplish convergence in less than τ seconds if the estimator eigenvalues satisfy Eq. (23a).

Part 4: Since convergence of $\hat{x}^2$ to x^2 occurs before $t=\tau$, by virtue of Eq. (23a), then the onset of a new "observation window" $t\in[\tau,2\tau]$, brings a new function $e(t)$, $t\in[\tau,2\tau]$, and a new set of coefficients of the Chebyshev polynomials which must be determined, and the entire sequence of events is the same as in the interval $t\in[0,\tau]$. Q.E.D.

There are several subtleties in theorem 3 which should be mentioned so that we do not read too much into the theorem. Nothing has been said so far that promises stability of the closed-loop system. We have only discussed performance of the estimator and we have not invoked any "separation theorem" to allow feedback of the state estimate $\hat{x}^2$ independently of considerations of the estimator design. Also of concern in theorem 3 is that the order d of the model error system which satisfies theorem 3 might be quite large. If we wish to design an on-line controller, we are usually interested in finding the *minimal*-order controller which will cause the physical system to meet its control objectives. It is not yet clear, however, whether such a "minimal model" is obtained with straightforward model reduction procedures with no pretense to compensate for the resulting model errors, or whether the "minimal model" (upon which the controller design is based and which yields satisfactory performance) is obtained by first reducing the model severely and then augmenting to it a low-order "model error system" which gives to the resulting controller a certain quality of model error tolerance.

Since in physical systems the "actuator" device which achieves control of the system has a finite "bandwidth" (sometimes called the "spectrum of control authority") it is pointless to pass through the estimator signals to which the actuator cannot respond. In this case it might be advisable to *avoid* the attempt to model the error vector exactly, as was the spirit of theorem 3, and instead, to attempt to model a certain "average" of the model error effects over the interval τ.‡ This can drastically reduce the order of the model error system without sacrificing information to which the actuator can respond. In the interest of minimizing the final controller order, it will be our standard procedure to determine the order of the model error system (which is to be augmented to any given model S_3) by iterating with designs beginning with $d=0$, 1, 2, etc. The details of such an approach are discussed elsewhere.[15] The method of evaluating the design may be analytical (i.e. evaluate an index of performance) or by simulation. Concern for minimal order of the controller design leads us to the considerations of the next section.

Minimal-Order Orthogonal Filter

The following result follows from the arguments of theorem 3 with the exception of estimator synthesis.

Theorem 4: For any fixed observation window $\tau>0$, there exists an integer d [the order of the Chebyshev model error system in Eq. (12)] such that, in the presence of any model errors associated with the model S_3 (parameter errors, truncated modes, and neglected disturbances), the state estimate $\hat{x}^3(t)$ converges to $x^3(t)$ if the estimate $\hat{x}^3(t)$ is given by the minimal-order Chebyshev filter,

$$\hat{x}=[I,0]\hat{x}^2 \qquad \hat{x}^3\in R^{n_3}, \hat{x}^2\in R^{n_3+d} \tag{26a}$$

$$\hat{x}^2=S_1z+S_2\xi \qquad z\in R^l, \xi\in R^q \tag{26b}$$

$$\dot{\xi}=\Gamma A^2\hat{x}^2+\Gamma B^2u=\Gamma[A^2-B^2G]\hat{x}^2 \tag{26c}$$

‡A property[16] of the Chebyshev polynomials T_i in the approximation of $e(t)$ is that, if $e(t)$ is continuous, then

$$\max_{t\in[0,\tau]} |e(t)-\bar{P}T(t)|$$

is minimized, lending a certain quality of smoothness to the approximation of the model error vector. If $e(t)$ is not continuous then the fitting is still accomplished in the mean-square sense, although exact pointwise fitting is possible only if $e(t)$ is continuous.

where $q \triangleq n_3 + d - l$, and,

$$S_2 \triangleq \begin{bmatrix} -M_l^{-1}M_2 \\ I \end{bmatrix} \qquad S_l \triangleq \begin{bmatrix} M_l^{-1}[I - M_2 S_{2l}] \\ S_{2l} \end{bmatrix}$$

$$M^2 \triangleq [M_l, M_2], \; |M_l| \neq 0, \; \Gamma \triangleq [0_{q\times l}, I_{q\times q}] - [S_{2l} M^2]$$

where the $q \times l$ matrix S_{2l} is selected to stabilize the matrix $\Gamma A^2 S_2$ such that

$$-Re\lambda_i[\Gamma A^2 S_2] \geqslant l/\tau \tag{27}$$

and where A^2, M^2 are defined by Eq. (23b). In writing Eq. (26c), we have presumed a control policy of the form $u = -G\hat{x}^2$.

Following the results of theorem 3, the only new point to prove is that $\hat{x}^2(t)$ converges to $x^2(t)$ quickly compared to the observation window τ. The minimal-order estimator for x^2 which completes the proof of theorem 4 is constructed in the Appendix.

C. Chebyshev-Fourier Coefficients

Thus far, the D matrix in the Chebyshev error system [Eq. (12)] has been fixed, but aside from the rather indirect information concerning P in theorems 1 and 2, no specific guidelines have been offered for the choice of the matrix P in Eq. (12). Three alternate methods for selecting P are suggested in the following. The procedures are described for the case of e_z approximation. (When z instead of z' is used in the definitions in Eq. (7), then the notation e_x and e_z will be used instead of e_x^{3l}, e_z^{3l}.)

Method 1: Adaptive P Based upon $e_z(t)$ Data over the Previous Observation Window τ

By observation of the "measurement residual" $z - \hat{z}^3$, [where $\hat{z}^3 = M^3\hat{x}^3$, and $\hat{x}^3$ is the output of an orthogonal filter from Eqs. (22) or (26)] an instantaneous estimate of e_z is available as $\hat{e}_z(t) = z(t) - \hat{z}^3(t)$. Integrating this estimate of the model error vector over the interval τ, in the manner

$$P = \int_{-l}^{l} \hat{e}_z(\sigma) T^T(\sigma) g(\sigma) \mathrm{d}\sigma \Lambda^{-l}, \; \sigma \triangleq \frac{2t}{\tau} - l \tag{28}$$

where σ, $T(\sigma)$, Λ, and $g(\sigma)$ are defined in Eq. (14), yields the P which gives a least-squares fit to the model error vector in the sense that

$$J(P) = \int_{-l}^{l} [\hat{e}_z - PT]^T g[\hat{e}_z - PT] \mathrm{d}\sigma \tag{29}$$

is minimized by the choice of Eq. (28), using Chebyshev polynomials, T for the approximation. The P given in Eq. (28) is called the Chebyshev-Fourier coefficient matrix.[16] A slight variation of Eq. (28) may also prove useful in adaptive orthogonal filters. If we make the assumption that $\hat{s} = T$ we can change Eq. (28) to read, for the update of P at time $t = (k+1)\tau$, $k = 0, 1, 2, \ldots$

$$P(k+l) = \int_{k\tau}^{(k+l)\tau} (z - M_x \hat{x}^2)\hat{x}^{2^T} M_s^T g(t) \mathrm{d}t \Lambda^{-l} \tag{30}$$

where $P(0)$ is taken as any from the observable pairs (D,P) and

$$g(t) = (l - \sigma^2)^{-1/2}, \quad \sigma \triangleq 2(t - k\tau)/\tau - l$$

$$M_x = [M^3, 0], \qquad M_s = [0, I_{d\times d}]$$

where $\hat{x}^2$ is provided by the orthogonal filter. The matrix P, according to Eqs. (28) or (30) is discretely updated every τ seconds, creating an adaptive loop whose stability properties must be investigated in future work.

Method 2: Constant P Based upon a "Characteristic Waveform" in $e_z(t)$

If certain dominant waveforms are known to be present in $e_z(t)$ [i.e., lightly damped truncated modes will contribute lightly damped oscillations in $e_z(t)$] over brief intervals τ units long, this characteristic waveform $e_z(t')$, $t' \in [0, \tau]$ can be utilized to compute P according to Eq. (28). The only difference is that now P is constant for all time and is not updated as in method 1.

Method 3: Constant P Satisfying Observability of the Pair (D, P)

The easiest decision for P is to choose any maximally ranked P which yields an observable Chebyshev error system [Eq. (12)]. In this case the degree of the Chebyshev polynomials whose coefficients can take on arbitrary values is, from theorem 1, the maximum integer ρ for which $\rho \leq d/r$. This easiest choice for P is used in the example of Sec. IV.

IV. Model Error Compensation in Flexible Spacecraft

The previous sections have been concerned with a particular approach to model error compensation in the design of state estimators. In this section we will apply two different model error compensation schemes to the design of a feedback controller for attitude control of a certain solar electric spacecraft, described more completely in Ref. 17. Table 1 summarizes a 14th-order evaluation model of the pitch axis of the spacecraft, where $\delta^T = (172.16, -4.95, -41.33, 2.48, -1.63, 12.22)$, $\sigma^T = (0.07, 0.106, 0.156, 0.228, 0.236, 0.238)2\pi$, $\sigma_D \triangleq \mathrm{diag}\ \sigma_i$, $i = 1, 2, \ldots 6$, and the Gaussian white noises representing the uncertainties in position measurements and control accelerations are denoted by $v \sim N(0.3, 2.1 \times 10^{-8})$, $w \sim N(0.06, 10^{-6})$, respectively, with means 0.3 and 0.06 and variances 2.1×10^{-8} and 10^{-6}. The output vector y^l denotes those variables (pitch and pitch rate) that must be regulated to zero. In all of the designs which follow, the model S_3 is taken as the rigid body model and all flexible appendage modes have been truncated. This is an example and is not a restriction of the methods. Generally, certain flexible modes are retained in S_3 if they are known to present serious control problems and are known to be reliably modeled. All other modes are truncated from S_3 and assigned to the responsibility of the model error system for approximate compensation.

A. Orthogonal Filter Designs

This section does not explore the adaptive orthogonal filter, nor even all types of time-invariant orthogonal filters which have been discussed in Sec. III. It only illustrates the Chebyshev minimal-order filter for the case of e_x approximation. Whenever there is substantial measurement

Table 1 Spacecraft model S_l

A_1	0	I	0
	$J^{-1}\Omega_1$	$J^{-1}\Omega_2$	$J^{-1}b_r$
C_1	1,0	0	
	0	1,0	
M_1	300,0	0	

$$B_1 = \Gamma_1, \quad J \triangleq \begin{bmatrix} J_{xx} & -\delta^T \\ -\delta & I \end{bmatrix}$$

$$b_r = \begin{pmatrix} -1.12 \\ 0 \end{pmatrix} \qquad \Omega_1 \triangleq \begin{bmatrix} 0 & 0 \\ 0 & -2\zeta\sigma_D \end{bmatrix} \qquad \Omega_2 \triangleq \begin{bmatrix} 0 & 0 \\ 0 & -\sigma^2_D \end{bmatrix}$$

$\zeta = .005$, $Q = 10^4 I_{2\times 2}$

$J_{xx} = 33{,}353$ slug ft^2, $R = 1$

noise, the full-order filter [Eq. (22)] should be employed to allow the opportunity to attenuate the noise while simultaneously estimating the state. Otherwise, the feed-forward structure of the minimal-order filter [Eq. (26)] may yield controllers which are too noisy. Table 2 shows the cases of the Chebyshev filter which have been examined where the rigid body model S_3 is given by

$$C^3 = 10^2 I_{2\times 2},\quad A^3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\quad B^3 = \begin{pmatrix} 0 \\ -\mu \end{pmatrix},\quad \mu = \frac{b}{J_{xx}}$$
$$M^3 = (300,0) \tag{31}$$

We will discuss the cases $d=0$, 1, 2 in that order.

Case $d=0$

In the case of $d=0$ the design model is $S_3 = S_2$. The full-order estimator [Eq. (22)] is applied in this case where $\hat{G}$ in Eq. (22b) may be taken as any $\hat{G}$ which satisfies

$$\hat{G} = \hat{K}M^{3^T}\hat{R}^{-1},\quad 0 = A^3\hat{K} + \hat{K}A^{3^T} - \hat{G}\hat{R}\hat{G}^T + B^3\hat{Q}B^{3^T} \tag{32}$$

for any positive definite $\hat{R}$ and $\hat{Q}$. The particular choice $\hat{R} = 2.1\times 10^{-8}$, $\hat{Q} = 10^{-6}$ yields the Kalman filter. In this case,

$$\hat{G} = \begin{pmatrix} [2\mu/K_{sx}\sqrt{\hat{Q}/\hat{R}}]^{1/2} \\ \mu\sqrt{\hat{Q}/\hat{R}} \end{pmatrix} = \begin{pmatrix} 1.24\times 10^{-3} \\ 2.31\times 10^{-4} \end{pmatrix} \tag{33}$$

The eigenvalues of this state estimator (Kalman filter) are

$$\lambda[A^3 - \hat{G}M^3] = -(\hat{Q}/\hat{R})^{1/4}\sqrt{1/2 K_{sx}\mu}\,(1\pm j1)$$
$$= -0.186 \pm j0.186 \tag{34}$$

The control policy and the gains which minimize Eq. (10), $T\to\infty$, are [in this model from Eq. (31), $x^3 \triangleq x^2$]

$$u = -G_x x^3,\qquad G_x = (-100, -2440) \tag{35}$$

The eigenvalues of the closed-loop design model are

$$\lambda[A^3 - B^3 G] = -0.041 \pm j0.041 \tag{36}$$

and the response of S_1 for an initial condition of 4.5 mrad in position is shown by curve KF = CFO in Fig. 1. Since the fidelity of the rigid body model is questionable beyond the spectrum of truncated modes, a "model credibility spectrum" $c(S_2)$ may be constructed in Fig. 2 by the rule $0 < |\lambda| < |\lambda_t|$, where λ_t are truncated modes. That is, when the theory places the eigenvalues outside of this region $c(S_2)$ it is less likely that the system will behave as predicted by the model. Consider the constant damping line given by Eq. (34) along which move the Kalman filter poles as the ratio $\hat{Q}/\hat{R}$ varies. Note that the estimator poles leave the model credibility spectrum $c(S_2)$ as $\hat{Q}/\hat{R}\to 0$ and again for large values of $\hat{Q}/\hat{R}$. Under such conditions the Kalman filter may diverge by making "predictions" (integration of the model state equations) over time intervals (too long in the case of $\hat{Q}/\hat{R}\to 0$, and too short in the case of large $\hat{Q}/\hat{R}$) for which the model is not reliable.

Table 2 Chebyshev filter cases studied

d	τ	(Filter)	S_3	Fig. 1 curve no.
0	–	$-0.186\pm j0.186$	rigid body	KF = CFO
1	–	$-0.1, -0.1$	rigid body	CF1
2	24	$0.09\pm j0.04, -0.1$	rigid body	CF2

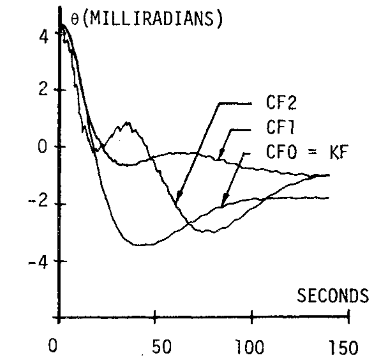

Fig. 1 Performance comparisons.

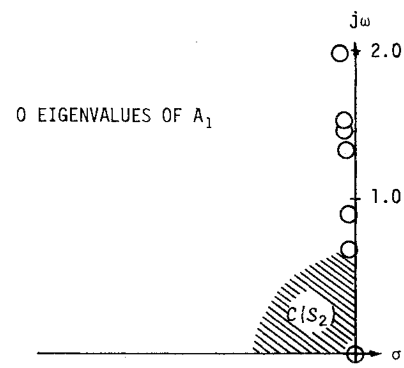

Fig. 2 Model credibility spectrum.

For $\hat{Q}/\hat{R}$ too small, the time constant of the estimator $\tau_e = 1/(Re\lambda)\to\infty$ and $\hat{G}\to 0$. Since the theory [represented by the Kalman filter design Eqs. (32)] relies upon the absolute fidelity of the model, there is nothing in the theory to warn against violating the right boundary of $c(S_2)$ and causing the state estimator to disconnect itself from the measurements (by allowing $\hat{G}\to 0$). Such a result would guarantee filter divergence since no model is sufficiently reliable to rely on its own uncorrected predictions forever. When the estimator places its poles to the left of $c(S_2)$, the estimator may again diverge due to predictions over intervals inappropriate for the model S_2 (this time the intervals are too short $\tau_e = 1/\lambda_e$, $\lambda_e > \lambda_t$). In view of these problems we might note that, in situations in which the Kalman filter diverges, the following statements are generally *false*:

1) The estimator can be stabilized by better statistical data. Even if $\hat{Q}$ and $\hat{R}$ *actually* represent the properties of the physical system, pole placement according to Eq. (34) may still fall outside $c(S_2)$ because information concerning the boundaries of $c(S_2)$ is *not* contained in Eq. (34).

2) The estimator can be stabilized by better plant parameters. Just as with the statistical parameters discussed earlier, there may not exist *any* values for the plant parameters (A^2, B^2, M^2) which will stabilize the divergent Kalman filter (see Ref. 18 for further examples of this circumstance).

3) The estimator can be stabilized by an adaptive Kalman filter. Most adaptive filtering schemes increase $\hat{Q}$ in some way (so as to add artificial noise to "cover" modeling errors). Note from Eq. (34) that increasing $\hat{Q}$ may help if the estimator poles are initially near the origin. However, if the nomial $\hat{Q}$, $\hat{R}$ correspond to poles near the left boundary of $c(S_2)$, the adaptive Kalman filter will actually cause the state estimates to get worse.

The conclusion drawn from this digression is that poor performance from a Kalman filter does not necessarily imply a poor choice of parameters for either the noise covariances or the plant. The more likely case is that the model error vector is not appropriately modeled as uncorrelated white noise. In this event the model error systems of Sec. III may be helpful. Also to be concluded from this example is the fact that similar

difficulties are faced with a deterministic state estimator. Thus, the often printed "rule of thumb" for designing Luenberger observers – "make the observer *fast* compared to the plant" – may be poor advice considering the existence of a left boundary of $c(S_2)$. The eigenvalues of the Kalman filter in Eq. (34) seem compatible with the model credibility spectrum $c(S_2)$ of Fig. 2 and yield the performance illustrated in Fig. 1. The design is stable but the steady-state offset in position is about 1.9 mrad due to the uncompensated bias in w^1.

Case d = 1

The resulting control law [from minimization of Eq. (10)]

$$u = -G\hat{x}^2,\ G = [G_x G_s] = (-100, -2440, -29780) \tag{37}$$

requires feedback of both the estimate of the state of the model $\hat{x}^3$ and the state of the model error system $\hat{s}$. It can be shown that G_s may be obtained from

$$G_s = R^{-1} B^{3^T} L \tag{38a}$$

using the recursive formula,

$$L^i = -\{\tilde{A}^{3^T}\}^{-1}\left\{KP^i + \sum_{j=i}^{d-1} D_{(j+1),i} L^{j+1}\right\} \tag{38b}$$

where $i = d,\ d-1,\ d-2,\ldots 1$, and $\tilde{A}^3 \triangleq A^3 - B^3\ G_x$, $G_x = R^{-1}B^{3^T}K$, and K is the Riccati matrix which results from the optimal control problem for model S_3 (or, equivalently, from the control problem for S_2 with $P = 0$). While it cannot be argued that the control [Eq. (37)] is optimal [because (A^2, B^2) is uncontrollable] it was shown in Ref. 2 that G_2 as given by Eq. (38) does exist for Chebyshev error systems augmented to models S_3 which are controllable, observable. See the system S_1 response in Fig. 1, curve CF1, for the minimum-order Chebyshev filter design [Eq. (26)].

Case d = 2

The minimum-order Chebyshev filter design [Eq. (26)] follows the preceding procedure with the resulting control gains: $G = [G_x G_s] = (-100,\ -2440,\ 0,\ -29780)$, and the resulting performance of curve CF2 in Fig. 1. Note that the transient properties have not improved although the steady-state value is the same as for $d = 1$. The observation window $\tau = 24$ for CF2 was selected from the dominant time constant of the rigid body design S_3 with the motivation that the model predictions might not be reliable much above or below the characteristic time constant of the model.

B. Fourier Error Systems

As a special case of the Chebyshev filter, the Fourier filter can be obtained. The model error system of Eq. (12) generates a Fourier series if D in Eq. (12) is multiplied by $\frac{1}{2}\tau\omega_0 \sin\omega_0 t$ where ω_0 may be selected as the fundamental frequency expected in the model error vector.[2] For example, when a certain spacecraft flexible mode is particularly troublesome in the control problem and the frequency of that mode is known fairly accurately, but the mode shape and damping are poorly known, then it may be helpful to choose the model error system [Eq. (12)], where

$$D = \omega_0 \sin\omega_0 t \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{39}$$

generates a Fourier series truncated to two terms. The controller design based upon the orthogonal filter then yields a state space version of the classical "notch filter" coupled with an "integral controller" design. To eliminate the time dependence of D in Eq. (39) the time invariant form

$$D_F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ \omega_0^2 & -\omega_0^2 & 0 \end{bmatrix} \tag{40}$$

may be used to obtain, using either the full-order [Eq. (22)] or reduced-order [Eq. (26)] versions of the orthogonal filter, a time invariant linear dynamical controller.

The performance of the time invariant Chebyshev filter may eventually be limited for large d by closed-loop stability considerations, because the Chebyshev model error system adds to the model S_3, d uncontrollable, unstable modes (the eigenvalues of D_0 are d repeated zeros). This is a manifestation of the fact that model errors may be successfully accounted for by adding integrators in the state estimator, but model error compensation may not always yield a *stable* system. (In Ref. 19 a similar point is made by showing, in the servomechanism problem, that tracking of input signals may be accomplished by adding integrators, but the tracking cannot always be accomplished with a stable system.) Hence, the time invariant Chebyshev filter seems most appropriate for investigating minimal bandwidth designs, where the dominant model error effects (including disturbances) can be approximated with low order polynomials. The detailed analysis of the characteristics of the adaptive orthogonal filter must await further research, although Ref. 15 illustrates an adaptive Chebyshev filter of degree 1 for this same spacecraft.

V. Conclusions

Specific guidelines are given herein for specifying the parameters which cause the model error system to generate orthogonal functions (and in the special case of this paper, Chebyshev polynomials). The orthogonal functions can be used in a real time algorithm for model error compensation by combining notions of state estimation. The special case of the Chebyshev filter is the focus of this paper in an application to a 14th-order model of a flexible spacecraft. Using a rigid body model of the spacecraft as a design model, the Chebyshev filter leads to a controller which compensates for certain of the modeling errors which are not compensated in the standard Kalman filter design. One can also view the orthogonal filter work as a particular attempt to compensate for "time-correlated" disturbances.

Appendix: Construction of the Minimal Order State Estimator of Theorem 4

Minimal-order state estimators, or "observers," for linear dynamical systems have been derived in a number of places. The development here was motivated by the suggestion of Ref. 20 which simplifies the mechanics of the estimator design by first placing the model in a set of coordinates in which the $l \times l$ matrix M_1 is nonsingular, where $M^2 = [M_1\ M_2]$, (matrix M^2 is $l \cdot (n_3 + d)$). A simple rearrangement of the state vector x^2 will accomplish this, assuming an independent set of l measurements. We will write the estimator design model S_2: $\{x^2 = A^2 x^2 + B^2 u,\ z^2 = M^2 x^2\}$ in the form without superscripts, assuming the preceding coordinates so that, $x \in R^{n_3+d}, M = [M_1, M_2]$ where $|M_1| \neq 0$. Define a coordinate transformation $x = S\alpha$ such that the first l components of α are the l measurements z. That is, $\alpha^T \triangleq (z^T, \zeta^T)$. The fundamental principle upon which all reduced estimators rely is that, while the matrix $S^{-1}AS$ in $\dot{\alpha} = S^{-1}AS\alpha + S^{-1}Bu$ has exactly the same set of eigenvalues as A, the eigenvalues of the $(n_3 + d - l) \times (n_3 + d - l)$ submatrix $a_{22} \triangleq [S^{-1}AS]_{22}$ may be arbitrarily placed by some choice of S if the matrix pair (A, M) is observable. Thus, the system

$$\dot{\zeta} = a_{22}\zeta + a_{21}z + \Gamma Bu \qquad a_{21} \triangleq [S^{-1}AS]_{21} \tag{A1a}$$

$$\hat{x} = S_1 z + S_2 \zeta, \qquad [S_1 S_2] \triangleq S \tag{A1b}$$

is a state estimator in the sense that the vector $\bar{\zeta}(t)-\zeta(t)=\Gamma[x(t)-\hat{x}(t)]$ converges to zero as fast as desired by appropriate choice of the eigenvalues of a_{22}, if the model is perfectly accurate so that $z-z^2 \triangleq e_z \equiv 0$. This result follows from the fact that $(\dot{\bar{\zeta}}-\dot{\zeta})=a_{22}(\bar{\zeta}-\zeta)-a_{21}e_z$.

To obtain the form of the estimator which is illustrated in theorem 4 we now organize our search for the S which satisfies the foregoing requirements. From the nonsingularity of S,

$$S^{-1}S=\begin{bmatrix} M_1 & M_2 \\ \Gamma_1 & \Gamma_2 \end{bmatrix}\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}=SS^{-1}$$

$$=\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}\begin{bmatrix} M_1 & M_2 \\ \Gamma_1 & \Gamma_2 \end{bmatrix}=I \tag{A2}$$

we can solve for S_{11}, S_{12}, Γ_1, Γ_2, using the fact that M_1 is nonsingular, and write $S_{11}=M_1^{-1}[I-M_2S_{21}]$, $S_{12}=-M_1^{-1}M_2S_{22}$, $\Gamma=[\Gamma_1,\Gamma_2]=S_{22}^{-1}[-S_{21}M_1, I-S_{21}M_2]$, and $a_{21}\triangleq[S^{-1}AS]_{21}=\Gamma AS_1$. Finally, it can be shown that

$$a_{22}=[S^{-1}AS]_{22}=\Gamma AS_2=S_{22}^{-1}[F_2+S_{21}G_2]S_{22}$$

where

$$F_2\triangleq[0_{q\times l},I_{q\times q}]A\begin{bmatrix} -M_1^{-1}M_2 \\ I \end{bmatrix}, G_2\triangleq -MA\begin{bmatrix} -M_1^{-1}M_2 \\ I \end{bmatrix} \tag{A3}$$

It is obvious that S_{22} does not affect the eigenvalues of a_{22} and we might choose $S_{22}=I$. The estimator design is therefore accomplished with the choice of the $(n_3+d-l)\times l$ matrix S_{21} so that ΓAS_2 is stable and $|\lambda[\Gamma AS_2]|_{\min}\geqslant 1/\tau$ for convergence with τ. Substituting the expressions for a_{22}, a_{21} [Eq. (A1)] yields the form of Eq. (26).

References

[1] Johnson, C. D. and Skelton, R. E., "Optimal Desaturation of Momentum Exchange Control Systems," *Proceedings of Joint Automatic Control Conference*, Atlanta, Ga., June 1970, pp. 638-694; also see *AIAA Journal*, Vol. 9, Jan. 1971, pp. 12-22.

[2] Skelton, R. E. and Likins, P. W., "On the Use of Model Error Systems in the Control of Large Scale Linearized Systems," *Proceedings of the IFAC Symposium on Large Scale Systems, Theory and Application*, Udine, Italy, June 1976.

[3] Skelton, R. E. and Likins, P. W., "Techniques of Modeling and Model Error Compensation in Linear Regulator Problems," *Advances in Control and Dynamic Systems*, Vol. XIV, edited by C. T. Leondes, Academic Press, New York, 1978, Chap. 1 (to be published).

[4] Luders, G. and Narendra, K. S., "A New Canonical Form for an Adaptive Observer," *IEEE Transactions on Automatic Control*, April 1974, pp. 117-119.

[5] Quigley, A.L.C., "An Approach to the Control of Divergence in Kalman Filter Algorithms," *International Journal on Control*, Vol. 17, No. 4, 1973, pp. 741-746.

[6] Likins, P. W., "Analytical Dynamics and Nonrigid Spacecraft Simulation," Jet Propulsion Laboratory, NASA TR32-1953, July 15, 1974.

[7] Likins, P. W. and Ohkami, Y., "Appendage Modal Coordinate Truncation Criteria in Hybrid Coordinate Dynamic Analysis," *Journal of Spacecraft and Rockets*, Vol. 13, Oct. 1976, pp. 611-617.

[8] Poelaert, D.H.L., "A Guideline for the Analysis and Synthesis of a Nonrigid-Spacecraft Control System," *ESA/ASE Scientific & Technical Review*, No. 1, 1975, pp. 203-218.

[9] Davison, E. J., "A Method for Simplifying Linear Dynamic Systems," *IEEE Transactions on Automatic Control*, Vol. AC-11, Jan. 1966, pp. 93-101.

[10] Kokotovic, P. V. and Yackel, R. A., "Singular Perturbation of Linear Regulators: Basic Theorems," *IEEE Transactions on Automatic Control*, Vol. AC-17, Feb. 1972, pp. 29-37.

[11] Kriendler, E., "Formulation of the Minimum Trajectory Sensitivity Problem," *IEE Transactions on Automatic Control*, April 1969, pp. 206-207.

[12] Byrne, P.C. and Burke, M., "Optimization with Trajectory Sensitivity Considerations, *IEEE Transactions on Automatic Control,* Vol. Ac-21, April 1976, pp. 282-283.

[13] Johnson, C. D., "Theory of Disturbance Accommodating Controllers," *Advances in Control and Dynamic Systems: Theory and Applications*, Vol. 12, edited by C. T. Londes, Academic Press, New York, 1975.

[14] Likins, P. W., "Dynamics and Control of Flexible Space Vehicles," Jet Propulsion Laboratory, NASA TR32-1329, Jan. 1970, p. 57.

[15] Skelton, R.E., "An Algorithm for an Approximation of the Minimum Controller Problem," *Journal of Guidance and Control,* Vol. 1, Jan.-Feb. 1978, pp. 90-93.

[16] Ghizzetti, A. and Ossicini, A., *Quadrature Formulae*, Academic Press, New York, 1970, pp. 124-125.

[17] Larson, V. and Likins, P. W., "An Application of Modern Control Theory to an Elastic Spacecraft," *ESA Symposium on Dynamics and Control of Nonrigid Space-Vehicles*, Frascati, Italy, May 1976.

[18] Fitzgerald, R. J., "Divergence of the Kalman Filter," *IEEE Transactions on Automatic Control*, Vol. AC-16, Dec. 1971, pp. 736-747.

[19] Bhattacharyya, S. P. and Pearson, J. B., "On Error Systems and The Servomechanism Problem," *International Journal of Control*, Vol. 15, No. 6, 1972, pp. 1041-1062.

[20] Yuksel, Y. O. and Bongiorno, J. J. Jr., "Observers for Linear Multivariable Systems with Applications," *IEEE Transactions on Automatic Control*, Vol. AC-16, Dec. 1971, pp. 603-613.