PURDUE UNIVERSITY
SCHOOL OF AERONAUTICS AND ASTRONAUTICS

West Lafayette, Indiana 47907

Progress Report
NAG-1-857 ✓

Sensitivity, Optimal Scaling, and Minimum Roundoff Errors in Flexible Structure Models

Robert E. Skelton
Purdue University

Table of Contents

A Tutorial On Model Error Concepts In Control Design

R. E. Skelton
School of Aeronautics and Astronautics
Purdue University
West Lafayette, IN 47907

Abstract

Traditional modeling notions presume the existence of a "truth" model that relates the input to the output, without advanced knowledge of the input. This has led to the evolution of education and research approaches (including the available control and robustness theories) that treat the modeling and control design as separate problems. This paper explores the subtleties of this presumption that the modeling and control problems are separable. A detailed study of the nature of modeling errors is useful to gain insight into the limitations of traditional control and identification points of view. Modeling errors need not be "small" but simply "appropriate" for control design. Furthermore the modeling and control design processes are inevitably iterative in nature.

Introduction

It is difficult to know *a priori* what type of modeling errors will be significant in the control design problem, and this issue is typically left to *ad hoc* approaches with the hope that "robust" control design techniques will somehow compensate for any error left in the model. However, significant performance improvements are often possible by obtaining more appropriate models at the outset, as opposed to placing all the burdens for the compensation of modeling errors on the control design. Even in identification approaches to modeling, the adequacy of the model for control design is unknown [1]. It is useful to revisit the formal structure of modeling errors to point out some common misunderstandings about their nature.

Perfect models are never required nor are they possible to construct. Some modeling errors are always acceptable. For example, in most aircraft or spacecraft control problems, these errors are probably acceptable without reservation: (i)

microscopic effects (molecular motion in the wings), (ii) relativistic effects, (iii) the higher order effects of earth oblateness on the environmental disturbances.

It is common practice in engineering to develop models of dynamic systems without regard to the impact that input forces have on the validity of the model. Unlike relativistic errors, this is a serious impact that needs further clarification. We shall see that the nature of the input forces dictate the validity of the model. If the input forces are to be controlled by a feedback controller, this means that the controller design dictates the validity of the mathematical model, rather than the other way around. (Control texts like to say that the fidelity of the controller is dictated by the fidelity of the model). An example clarifies the idea. A plate of steel might be characterized as a rigid body. A rigid body model is appropriate if the forces to be applied are relatively small. Large forces will deform the plate and can render a rigid body model useless for predicting the dynamic response. This argument continues indefinitely. The more accuracy required in the predictions of dynamic response, the more careful one has to be about including small effects in the dynamic model (non-homogeneous material properties of the steel, air turbulence, thermal gradients, etc.). This quickly leads to the conclusion that the validity of a model cannot be assessed by its open loop response. The modeling errors that are acceptable depend upon the control forces.

Figure 1 illustrates the steps in the modeling process. The first step is to adopt an "idealization" of the system. Two examples will clarify the meaning of "idealization" 1) If the physical phenomenon is a mechanical structure, we might "idealize" the structure as a rigid body. Another idealization is a flexible plate. The actual structure is neither a rigid body nor a flexible plate, but such a hypothesis (e.g. rigid body) forms our "idealization" of the structure and this occurs before any mathematical models are yet formulated. 2) Another type of idealization occurs in identification experiments. When we try to fit a canonical model (e.g. phase variable form) to the input output data, we say that the canonical structure forms our "idealization" of the plant. (The word *plant* is used

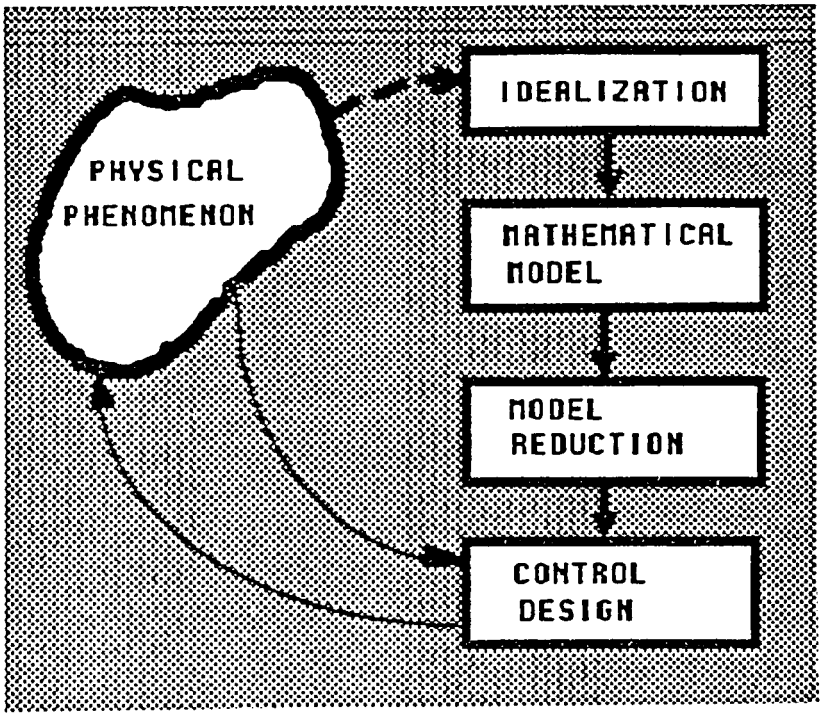


Fig. 1 Steps in the Modeling Process

to include the characterizations of the disturbance environment as well). Hence, idealizations lead to errors in model structure and model order.

The second step in the modeling process is to apply known physical laws to develop a mathematical model of the idealization. This step introduces parameter errors (values of mass, inertia, spring constants, etc.). This step might be omitted if the model is developed from experimental data.

The third step is the simplification of the model. High order models may be reduced to low order models. This step leads to errors in model order and parameters.

The fourth step is controller design based upon the model obtained in step III. The purpose of this paper is to discuss the types of errors that have small and large impacts on the controller design.

The controller design procedures which are tractable are usually of high order. (Both LQG and H^{∞} controllers are of order generally equal to the model order). Hence, a fifth step in the modeling and control design process is controller reduction. (If an

optimization is performed to obtain the controller of specified low order directly, the controller reduction and/or the model reduction step is omitted).

This paper is a tutorial on the concepts of model errors in model based control design. The paper poses some old questions, some new ones, and adds some conjectures. The intent is to clarify some common misunderstandings about the nature of modeling errors and their impact in feedback control.

The paper is organized as follows. Section 2.0 cites simple examples which motivate the need for a more precise way to characterize "acceptable" modeling errors. Section 3.0 characterizes the structure of all modeling errors, and describes a modeling and control inseparability principle. Section 4.0 describes the closed loop impact of modeling errors and presents an uncontrollability, unobservability principle. Section 5.0 offers some conclusions.

2.0 Motivating Examples and Some Modeling Principles

This section illustrates four modeling principles by examples.

MODELING PRINCIPLE I: *arbitrarily small modeling errors can lead to arbitrarily bad closed-loop performance*

For the system described by the transfer function

$$G(s) = \frac{1}{(1+s)(1+\varepsilon s)^{\alpha}} \quad , \tag{1}$$

let $\varepsilon > 0$ be small, possibly representing fact actuator or sensor dynamics. If the fast dynamics are ignored (a common approach in practice) then the control design model becomes

$$G_R(s) = \frac{1}{1+s} \tag{2}$$

where, for $\alpha = 1$, it may be shown that the step input errors between G(s) and $G_2(s)$ are bounded by

$$|y(t) - y_R(t)| = L^{-1}\left\{\frac{1}{s}[G(s) - G_R(s)]\right\} < \frac{2\varepsilon e^{-t}}{1-\varepsilon}, \quad t > 0. \tag{3}$$

Hence the modeling errors as measured by the open-loop response are arbitrarily small if ε is arbitrarily small. Suppose an output feedback controller with $u = -Ky$. If

$$K \ll \frac{1}{4\varepsilon}, \tag{4}$$

then the model $G_R(s)$ is also useful for predicting closed loop behavior, since $y_R(s)\left[\frac{G_R(s)}{1+KG_R(s)}\right]\frac{1}{s} \simeq y(s)$. The interesting observation about (4) is that it tells that the usefulness of the control design model $G_R(s)$ for predicting closed loop behavior *depends* both upon the modeling error (characterized by ε) and the controller gain K. This illustrates the theme of this entire paper; *the modeling and control problems are not independent.*

Now suppose the requirements on modeling error and control design (4) are violated to the extent that

$$K \gg \frac{1}{4\varepsilon} \tag{5}$$

(The reader may verify that $K \simeq \frac{1}{4\varepsilon}$ corresponds to the breakaway point on the root locus). Now the model $G_R(s)$ and controller (5) predict a closed loop system with no overshoot, no oscillations and a small steady state error. However, the actual response will yield severely underdamped oscillations for $\alpha = 1$, and instability for $\alpha = 2$).

The point of this example is that arbitrarily small modeling errors do not lead necessarily to small errors in the closed loop predictions. The following example

illustrates that small modeling errors can lead to large errors in optimality and maximal accuracy predictions.

Consider (1) with $\alpha = 2$ and the same $G_R(s) = 1/s+1$. Let K be the optimal control for the model $G_R(s)$ so as to minimize

$$V = \int_0^\infty [y_R^2(t) + \rho u^2(t)]dt = V_y + \rho V_u \tag{6}$$

Then

$$K = -1 + \sqrt{1+\rho^{-1}} \tag{7}$$

and the closed-loop system performance of (1) with $u = -Ky$ is described in Fig. 2 in terms of V_y versus V_u as ρ varies from $\infty \rightarrow 0$.

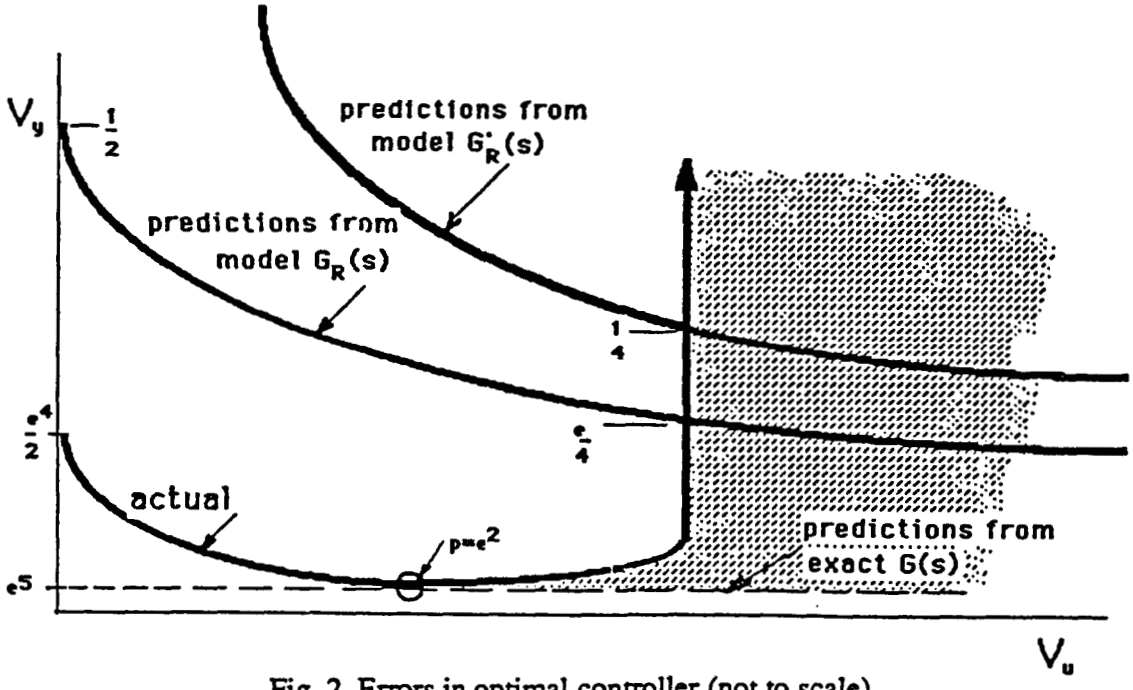


Fig. 2 Errors in optimal controller (not to scale)

For $\rho < \frac{\varepsilon^2}{4}$ the actual closed loop system is unstable, whereas the predicted behavior based on $G_R(s)$ approaches its maximal accuracy, $[V_y \text{ (predicted)}] \rightarrow 0$, as $\rho \rightarrow 0$.

Hence, in the neighborhood of *maximal* accuracy predictions for $G_R(s)$ in Fig. 2, the actual system delivers its worst performance.

Now for the system (1), consider the "absurd" design model

$$G_R'(s) = \frac{1}{(1+\varepsilon s)} \tag{8}$$

in lieu of $G_R(s) = \frac{1}{s+1}$. This model is "absurd" to the extent that the open-loop step response error is not necessarily small. The optimal control for this realization of the plant ($A = -\frac{1}{\varepsilon}$, $B = \frac{1}{\varepsilon}$, $C = 1$) yields the optimal control $u = -Ky$, $K = -1 + \sqrt{1+\rho^{-1}}$, which is precisely the same K as for model $G_R(s) = \frac{1}{s+1}$! Hence the actual performance is the same as that described in Fig. 2, and models $G_R(s)$ and $G_R'(s)$ are *equivalent* for the purposes of control design by (6). There is an infinite number of controllers in this comparison (one for each selected value of ρ). Hence, this equivalence is not just for an isolated control based upon models $G_R'(s)$ and $G_R(s)$.

The performance *predicted* by model $G_R'(s)$ is shown in Fig. 2 to be arbitrarily far from the actual performance for large values of ρ and for small values of ρ, although the predictions are close to those of model $G_R(s)$ for small ρ.

In Fig. 2, the comparison between the "actual" and "predictions from G(s)" indicate that for small gains (small V_u, large ρ) the actual performance agrees with the performance which would be optimal using the exact model G(s). That is, model errors do no damage for small enough control efforts. The actual performance has a "best" performance at a particular value of control effort ($\rho = \varepsilon^2$) and increasing control efforts beyond this point degrades performance leading eventually to instability ($\rho < \varepsilon^2/4$). All controllers will drive a physical system unstable as the control effort is increased enough (see the conjecture below). The shaded area in Fig. 2 describes the difference in performance between that which would be optimal for the exact plant G(s) and that

which results from controllers which are optimal with respect to an erroneous model {$G_R(s)$ or $G_R'(s)$ in our example}. This property is generic, according to the conjecture below.

Let S_c be a linear dynamic controller which is optimal for the linear model S_2. Let the actual linear plant be S_o and its optimal controller be S_{co}(which we cannot construct!). Since mathematical models are always approximations of the physical plant $S_o \neq S_2$. Let V_y denote the output norm when S_c drives S_o. Let V_{yo} denote the output norm when S_{co} drives S_o. V_y takes on its smallest value $\underline{V}_y$ when the control norm has value $\underline{V}_y$. V_{yo} takes on its smallest value $\underline{V}_{yo}$ when V_u is arbitrarily large $\underline{V}_{uo}$, see Fig. 2.

Conjecture:

Controller S_c *(optimal for model* S_2*) driving plant* S_o *always yields an unstable closed-loop system in the neighborhood of maximal accuracy predictions using any* S_2. *Hence,*

$$\lim_{\rho \to 0} V_y(\rho) = \infty , \tag{9a}$$

$$\lim_{\rho \to \infty} [V_y(\rho) - V_{yo}(\rho)] = 0 \tag{9b}$$

$$\underline{V}_y > \underline{V}_{yo} \tag{9c}$$

$$\underline{V}_u < \underline{V}_{uo} \tag{9d}$$

indicating that the actual maximal accuracy $\underline{V}_y \triangleq \{\min V_y(\text{actual})\}$ *occurs always at a lower value of control effort* $\underline{V}_u$ *than the predicted maximal accuracy* $\underline{V}_{yo} \triangleq \{\min V_y(\text{predicted})\}$ *which occurs at a value of* $\underline{V}_{uo}$.

This conjecture asserts that real controllers (always based upon erroneous models) will always be bad (unstable) for large control effort V_u and will always be as good as the optimal controller for arbitrarily small control effort (9b). For quadratic criteria, this

suggests that a meaningful comparison of candidate controllers should be conducted in the neighborhood of the maximal accuracy V_{yA} for each controller. Each controller design, (based perhaps upon different reduced order models), will yield a different V_{yA}. Furthermore one cannot say *a priori* which is the best reduced order model (as the above examples illustrate). Hence, again we see the dependence between the modeling and control problems.

MODELING PRINCIPLE II: *large open-loop modeling errors do not necessarily lead to large closed-loop prediction errors*

Most of the available theories on model reduction [1-14] try to achieve "small" modeling errors, according to some open-loop criterion for minimization. (Example 1 may even provide some motivation for this goal). The appreciated virtue of balancing and Hankel norm methods [14] is the existence of an upper bound on the size of the transfer function error. The purpose of this example is to illustrate that it is not necessary to have small modeling errors if the control scheme has the right characteristics (remember the theme of the paper). Consider a plant described by

$$G(s) = \frac{1}{s+1} \tag{10}$$

and an approximate model

$$G_R(s) = 1/s\,. \tag{11}$$

The controller again is output feedback $u = -Ky$. Note that the actual system (10) is asymptotically stable, whereas the approximation (11) is not. $G_R(s)$ would not be deemed a good approximation of G(s) by any of the model reduction theories available in the literature. It is interesting, however, to ask "in what sense is $G_R(s)$ a good model for predicting closed-loop performance of the plant?"

(i) Stability of $G_R(s)$ under the feedback gain K will also yield stability of G(s) under this K.

(ii) Errors in predicting performance of the step response are given by

$$|y(t) - y_R(t)| < \frac{1}{1+K}, \quad t > 0. \tag{12}$$

Hence the errors in the closed loop response predictions can be made as small as desired by choosing an appropriate K, despite the fact that open-loop errors are large. Thus, closed loop effects of the modeling errors may be smaller than the (open-loop) modeling errors. The conclusion from this example is that the modeling errors should be *appropriate* for the controller design and not necessarily *small*. (We do *not* suggest that high controller gains are always appropriate).

To illustrate that large modeling errors can yield even *zero* closed loop errors, consider the following examples. We have already shown that two models of the same order, yet not related by a coordinate transformation, can yield the same controller (7). This concept is not limited to models of the same order. Consider models for standard LQG design:

$$\dot{x} = Ax + Bu + Dw, \quad z = Cx, \quad z = Mx + v$$

with zero mean uncorrelated white noises w, v, with intensities $W = 4$, $V = 1$. The same controller given by

$$\dot{x}_c = A_c x_c + F_z, \quad u = Gx_c$$

$$A_c = -3, \quad F = 2, \quad G = -2,$$

Optimizes all three of these models:

I)

$$A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{matrix} C_1 = [1\ 0] \\ M_1 = [0\ 1] \end{matrix}$$

II)

$$A_2 = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} .2 \\ 0 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 1.5 \\ .417 \end{bmatrix}, \quad \begin{matrix} C_2 = [\sqrt{3},\ 4.079], \\ M_2 = [.8,\ .5] \end{matrix}$$

(III)

$$A_3 = -1, \quad B_3 = .2, \quad D_3 = 1.5, \quad C_3 = \sqrt{3}, \quad M_3 = .8$$

with respect to the design objective V,

$$V = E_\infty(8y^2 + u^2), \quad E_\infty \triangleq \lim_{t \to \infty} E \quad .$$

That is, the same closed loop value of the performance metric is obtained with *all three* models (I-III). Hence models (I-III) are *equivalent models with respect the control objective* V. The striking observation here is that the open loop responses of the "control equivalent" models (I-IV) are not similar nor "close" by any open loop measure, yet all three models are equally good for the control design. Fig. 3 shows an open loop impulse response of the models. Two of these models were taken from [26]. Other control equivalent models are discussed in [27,28].

Let a given system S_o be controlled by a controller $S_c(S_1)$ based upon a model S_1. Modeling Principles I and II and the above numerical examples lead immediately to the conclusion that making errors "small" is extremely difficult and also unnecessary.

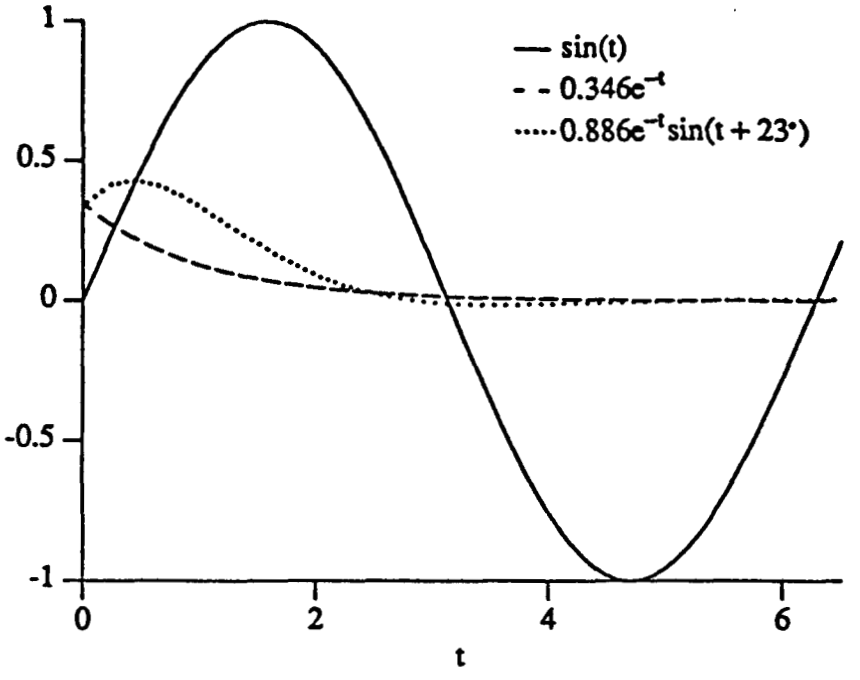


Fig. 3 Impulse response of three "control equivalent" models

Modeling Principle III is a natural consequence of these two facts.

MODELING PRINCIPLE III: *open-loop modeling errors (and hence their bounds) do not generally constitute enough information for successful control design.*

Of course, the previous discussions and examples provide the clue to the missing information. Knowledge of *inputs* are required for any assessment of fidelity of the model. This principle is not limited to quadratic criteria for system evaluation. See for example from the Nyquist plot of Fig. 4 that many different models of a plant may yield controllers with a common gain and phase margin. Hence, models may be "control equivalent" by any control design criterion, classical or modern. Such equivalence seems to have no direct relationship to the open loop modeling errors.

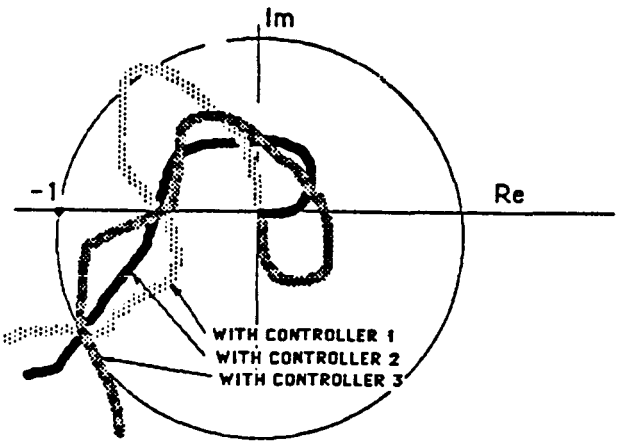


Fig. 4 Three controllers yielding the same gain & phase margins.

3.0 The Structure of Modeling Errors

For the sake of discussion only, imagine a mathematical model so accurate as to mimic the physical phenomena for all practical purposes. Then in the following arguments we refer (by a slight abuse of language in this Section only) to the model

$$\dot{x} = Ax + B\,\mathrm{u} + w + \mathrm{f}(x, \mathrm{u}, \mathrm{t}) \tag{13a}$$

$$z = M\mathrm{x} + v + \mathrm{g}(x, \mathrm{u}, \mathrm{t}) \tag{13b}$$

as the "physical system" with the control inputs $\mathrm{u} \in R^{n_u}$, disturbance inputs $w \in R^{n_w}$ and measurements $z \in R^{n_z}$. The terms f(x, u, t), and g(x, u t) represent nonlinearities and w(t) represents any time varying disturbances (w(t) is not a function of x or u). Of course, to attribute "physical system" status to (13) the dimension of the state x approaches infinity. We shall consider it large but finite. Now with respect to *any* mathematical model of the form

$$\dot{x} = Ax + Bu + w, \quad x \varepsilon R^{n_x} \tag{14a}$$

$$z = Mx + v \quad z \varepsilon R^{n_z} \tag{14b}$$

which might be used to represent the actual system (13), we wish to completely characterize the model errors of (14) with respect to (13). For this purpose rewrite (13) in the partitioned form

$$\begin{bmatrix} \dot{x}_R \\ \dot{x}_T \end{bmatrix} = \begin{bmatrix} A_R & A_{RT} \\ A_{TR} & A_T \end{bmatrix} \begin{bmatrix} x_R \\ x_T \end{bmatrix} + \begin{bmatrix} B_R \\ B_T \end{bmatrix} u + \begin{bmatrix} w_R \\ w_T \end{bmatrix} + \begin{bmatrix} f_R(x, u, t) \\ f_T(x, u, t) \end{bmatrix} \tag{15}$$

$$z = [M_R \, M_T] \begin{bmatrix} x_R \\ x_T \end{bmatrix} + v + g(x,u,t)$$

where $x_R \varepsilon R^{n_x}$ has the dimension of x. Using the following definitions

$$\Delta A \triangleq A_R - A, \quad \Delta B \triangleq B_R - B, \quad \Delta M \triangleq M_R - M, \tag{16}$$

The equation for x_R and z can be written, from (15), (16)

$$\dot{x}_R = (A+\Delta A)x_R + A_{RT}x_T + (B+\Delta B)u + w_R + f_R(x, u, t) \tag{17}$$

$$z = (M+\Delta M)x_R + M_T x_T + v + g(x, u, t)$$

or simply,

$$\dot{x}_R = Ax_R + Bu + w + e_x, \tag{18}$$

$$z = Mx_R + v + e_z$$

where

$$e \triangleq \begin{bmatrix} e_x \\ e_z \end{bmatrix} = \begin{bmatrix} \Delta A \\ \Delta M \end{bmatrix} x_R + \begin{bmatrix} \Delta B \\ 0 \end{bmatrix} u + \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} x_T + \begin{bmatrix} w_R - w \\ v - v \end{bmatrix} + \begin{bmatrix} f_R(x_R, x_T, u, t) \\ g(x_R, x_T, u, t) \end{bmatrix}$$

$$e = e_p + e_o + e_d + e_n \tag{19}$$

We now have the following conclusion from (19).

MODELING PRINCIPLE IV: *For every model of the form* (14) *there exists "model error vectors"* e_x *and* e_z *which represent "corrections" to the state equation* (14a) *and measurement equation* (14b), *respectively, such that z evolving from* (14) *matches the measurements z from the physical system* (13) *if* e_x *is added to* (14a) *and* e_z *is added to* (14b). *Furthermore, the model error vectors can always be decomposed into the sum of four kinds of errors: parameter errors* e_p, *errors in model order* e_o, *neglected disturbances* e_d *and nonlinearities* e_n.

Note that all four types of modeling errors e_p, e_o, e_d, and e_n are *always* present with any mathematical characterization of a physical plant. We wish also to declare at the outset that there exists no control theories which can promise satisfactory control in the simultaneous presence of all four categories of modeling error. Techniques are available which have made progress only in the accommodation of a *subset* of these four types of modeling errors.

Certain other conclusions are also obvious from (16). The partitioning of the state vector (15) was necessary to define "parameter errors" with respect to a specified model of lower order. The matrices A_R, B_R, and M_R result from a partitioning of (13) after (13) is written in a selected coordinate frame. It should be clear from (16) that since the parameters A_R, B_R, and M_R depend upon the initial coordinates in (13), the phrase "parameter errors" in dynamic systems has no precise meaning. Indeed, the definitions (16) are as precise as one can be, yet these definitions are arbitrary to within a coordinate transformation on (15). Note that if one chooses a canonical structure for (A, B, M) in (16), such as phase variable or Hessenberg form, etc. the parameters of (A_R, B_R, M_R)

may not have the same structure. Hence, $(\Delta A,\Delta B,\Delta M)$ does not generally have a canonical structure. This suggests that one should not expect parameter adaptive control schemes to converge if a presumed canonical structure of the parameters has been presumed. Also, note from (19) that a change in coordinates which changes the e_p term will also change the other terms e_o and e_n. Hence, the individual terms e_p, e_o, e_d and e_n in the model error decomposition (19) are not unique.

It does not serve our purpose to characterize the model error vector explicitly in terms of higher order model states x_T. For the reduced model (18) the vector functions e_x and e_z can be considered functions only of x_R and u and t. Eq. (18) shows that e_o is a function of x_T, but x_T is in turn a function of x_R, u, and t, as the solution of the second equation in (15) reveals

$$x_T(x_R,u,t) = \Phi_T(t,0)x_T(0)+\int_0^t \Phi_T(t-\sigma)(A_{TR}x_R+B_T u+w_T+f_T(x_R,x_T,u,\sigma))d\sigma \qquad (20)$$

where Φ_T is the state transition matrix for A_T. Hence, for (18) we may consider e_x to depend only on (x_R, u, t), and when we need to do so we shall write $e_x(x_R,u,t)$. In fact, using (20) the model order error e_o of (19) may be characterized by

$$e_o = \underbrace{\begin{bmatrix} A_{RT} \\ M_T \end{bmatrix}[\Phi_T(t,0)x_T(0)]}_{e_{ot}} + \underbrace{\begin{bmatrix} A_{RT} \\ M_T \end{bmatrix}\left[\int_0^t \Phi_T(t-\sigma)w_T(\sigma)d\sigma\right]}_{e_{od}} + \underbrace{\begin{bmatrix} A_{RT} \\ M_T \end{bmatrix}\left[\int_0^t \Phi_T(t-\sigma)A_{TR}x_R d\sigma\right]}_{e_{ox}}$$

$$+ \underbrace{\begin{bmatrix} A_{RT} \\ M_T \end{bmatrix}\left[\int_0^t \Phi_T(t-\sigma)B_T u d\sigma\right]}_{e_{ou}} + \underbrace{\begin{bmatrix} A_{RT} \\ M_T \end{bmatrix}\left[\int_0^t \Phi_T(t-\sigma)f_T(x_R,x_T,u,\sigma)d\sigma\right]}_{e_{on}} \qquad (21)$$

From (19) and (4) conclude that e may be written

$$e = \{e_{px}+e_{ox}\} + \{e_{pu}+e_{ou}\} + \{e_{ot}+e_{od}+e_d\} + \{e_n+e_{on}\} . \qquad (22)$$

With respect to the reduced model (14) the first bracketed term in (22) denotes errors which depend on the state x_R (e_{px} is a linear function of x_R, while e_{ox} is an integral operator on x_R), the second bracket denotes errors which depend on the control u·e_{pu} is a linear function of x_R, while e_{ou} is an integral operator on u, the third bracket in (22) denotes errors which depend only on time, and finally the last bracket in (22) denotes the errors due to nonlinearities.

It is important to note that the model error vector e_x in (18) depends upon the *integral* of the input u. Hence, one cannot assess the impact of modeling errors e_x without knowledge of the nature of the controls. Even small inputs u can have an arbitrarily large effect in $e_x(t)$. To see this let A_T be a positive scalar and u(t) = ε = constant. Then $e_{ou}(t) = (A_{RT}B_T\varepsilon/A_T)(e^{A_T t}-1)$ gets arbitrarily large in this example. Hence, *the homogeneous part of the system may be modeled arbitrarily closely* (e_p arbitrarily small or zero) *and yet the model may not be acceptable for control design (Modeling Principle I).*

The Model Error System

The model error structure (21) can be further detailed. See that e_{ox} satisfies the differential equations

$$\dot{e}_{ox} = \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} A_{TR}x_R + \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} \int_0^t A_T\Phi_T(t-\sigma)A_{TR}x_R d\sigma \qquad (23)$$

and the definition

$$e_{ix} \triangleq \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} \int_0^t A_T^i\Phi_T(t-\sigma)A_{TR}x_R d\sigma \qquad (24)$$

allows (23) to be written

$$\dot{e}_{ox} = \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} A_{TR} x_R + e_{1x} \tag{25}$$

and likewise for e_{1x} and e_{ix}

$$\dot{e}_{1x} = \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} A_T A_{TR} x_R + e_{2x}$$

$$\dot{e}_{ix} = \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} A_T^i A_{TR} x_R + e_{(i+1)x} . \quad i = 0, 1, ..., . \tag{26}$$

Similarly for e_{ou} in (21)

$$\dot{e}_{ou} = \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} B_T u + e_{1u}$$

$$\dot{e}_{iu} = \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} A_T^i B_T u + e_{(i+1)u} \qquad i = 0, 1, ..., . \tag{27}$$

The deleted state x_T, and hence A_T, is usually quite large so that the range of i is a large finite number.

Equations (26), (27) and (22) are now combined with (17) to give the exact structure of the model error system

$$\dot{x}_R = Ax_R + Bu + w + E_{10}[(e_{ox}+e_{ou})+(e_{px}+e_{pu})+e_t+e_N] \tag{28a}$$

$$\dot{e}_{ix} = P_i x_R + e_{(i+1)x} \qquad i = 0, 1, ..., . \tag{28b}$$

$$\dot{e}_{iu} = Q_i u + e_{(i+1)u} \qquad i = 0, 1, ..., . \tag{28c}$$

$$z = Mx_R + v + E_{01}[(e_{ox}+e_{ou}) + (e_{px}+e_{pu}) + e_t + e_N] \tag{28d}$$

where $E_{10} \triangleq [I\ 0]$, $E_{01} \triangleq [0\ I]$,

$$P_i \triangleq \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} A_T^i A_{TR}, \quad Q_i \triangleq \begin{bmatrix} A_{RT} \\ M_T \end{bmatrix} A_T^i B_T, \tag{29}$$

and

$$e_t \triangleq e_{ot} + e_{od} + e_d, \quad e_N \triangleq e_n + e_{on}, \tag{30}$$

using the definitions of e_{ot}, e_d, e_n, e_{on} in (19), (21).

In state form (28) becomes

$$\begin{bmatrix} \dot{x}_R \\ \dot{e}_{ox} \\ \dot{e}_{ou} \\ \dot{e}_{1x} \\ \dot{e}_{1u} \\ \dot{e}_{2x} \\ \dot{e}_{2u} \\ \cdot \\ \cdot \\ \cdot \\ z \end{bmatrix} = \begin{bmatrix} A+\Delta A & E_{10} & E_{10}0 & 0 & 0\,0 & 0 & 0 & \cdots \\ P_o & 0 & 0\,I & 0 & 0\,0 & 0 & 0 & \cdots \\ 0 & 0 & 0\,0 & I & 0\,0 & 0 & 0 & \cdots \\ P_1 & 0 & 0\,0 & 0 & I\,0 & 0 & 0 & \cdots \\ 0 & 0 & 0\,0 & 0 & 0\,I & 0 & 0 & \cdots \\ P_2 & 0 & 0\,0 & 0 & 0\,0 & I & 0 & \cdots \\ 0 & 0 & 0\,0 & 0 & 0\,0 & 0 & I & \cdots \\ \cdot & \cdot & \cdot\,\cdot & \cdot & \cdot\,\cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot\,\cdot & \cdot & \cdot\,\cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot\,\cdot & \cdot & \cdot\,\cdot & \cdot & \cdot & \\ M+\Delta M & E_{o1} & E_{o1}0 & 0 & 0\,0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_R \\ e_{ox} \\ e_{ou} \\ e_{1x} \\ e_{1u} \\ e_{2x} \\ e_{2u} \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$$

$$+ \begin{bmatrix} B+\Delta B \\ 0 \\ Q_o \\ 0 \\ Q_1 \\ 0 \\ Q_2 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} u + \begin{bmatrix} w+E_{10}(e_t+e_N) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ v+E_{o1}(e_t+e_N) \end{bmatrix} \tag{31}$$

Note that the matrices which are *unknown* are P_i, Q_i, ΔA, ΔB. Knowledge of the model error structure (31) might be very useful in analysis (predictions of performance) and control design. Most adaptive approaches to control ignore the e_{ix}, e_{iu} $i = 0, 1, \cdots$ terms in the model error system (31). It might be useful to research the inclusion of one extra term.

The transfer functions of (28) are developed as follows. In the context of linear systems, we shall ignore e_N. Take the Laplace transform of (28) and see that

$$z(s) = (M+\Delta M+E_{o1}(\sum_{i=0}^{\infty} P_i/s^{i+1}))[(sI-A)-\Delta A-E_{10}(\sum_{i=0}^{\infty} P_i/s^{i+1})]^{-1}[(B+\Delta B+$$

$$E_{10}(\sum_{i=0}^{\infty} Q_i/s^i))u(s)+E_{10}e_t+w] + E_{o1}(\sum_{i=0}^{\infty} Q_i/s^{i+1})u(s)+v(s)+E_{10}e_t(s)$$

To simplify this expression define

$$P_{ix} \triangleq \begin{cases} \Delta A & \text{when } i = 0 \\ E_{10}P_{i-1} & \text{when } i > 0 \end{cases} \qquad P_{iz} = \begin{cases} \Delta M & \text{when } i = 0 \\ E_{o1}P_{i-1} & \text{when } i > 0 \end{cases} \tag{32a}$$

$$Q_{ix} \triangleq \begin{cases} \Delta B & \text{when } i = 0, \\ E_{10}Q_{i-1} & \text{when } i > 0 \end{cases} \qquad Q_{iz} \triangleq E_{o1}Q_{i-1}$$

$$\Phi(s) \triangleq (sI-A)^{-1} \tag{32b}$$

then

$$z(s) = \{[M+\sum_{i=0,} P_{iz}/s^i][\Phi^{-1}(s)-\sum_{i=0,} P_{ix}/s^i]^{-1}[B+\sum_{i=0,} Q_{ix}/s^i]$$

$$+ \sum_{i=1,} Q_{iz}/s^i\} u(s) + e_t(s) \tag{33}$$

where $e_t(t)$ represents the effects of only *time* -dependent terms,

$$e_t(s) = [M+\sum_{i=0} P_{iz}/s^i][\Phi^{-1}(s)-\sum_{i=0} P_{ix}/s^i]^{-1}[w(s)+E_{10}e_t(s)]$$

$$+ E_{o1}e_t(s) + v(s) .$$

Let the inverse of the sum of two matrices be written

$$[\Phi^{-1}+\Psi]^{-1} = \Phi \sum_{j=0,}^{\infty} [(-1)^j(\Psi\Phi)^j] .$$

This can be verified by writing $[\Phi^{-1}+\Psi]x = y$ as $\Phi^{-1}x = (y-\Psi x)$ or

$$\begin{aligned} x &= \Phi(y-\Psi x) \\ &= \Phi y - \Phi\Psi\Phi(y-\Psi x) \\ &= \text{etc.} \end{aligned}$$

Now let

$$\Psi_1 \triangleq -\sum_{i=0,} P_{ix}/s^i,\ \Psi_2 \triangleq \sum_{i=0,} P_{iz}/s^i,\ \Psi_3 \triangleq \sum_{i=0,} Q_{ix}/s^i,\ \Psi_4 \triangleq \sum_{i=0}^{\infty} Q_{iz}/s^i,$$

$$J(s) \triangleq \sum_{j=1,} (-1)^j \Psi_1^j(s)\Phi^j(s)$$

Then

$$z(s) = \{[M+\Psi_2(s)]\Phi(s)[I+J(s)][B+\Psi_3(s)]+\Psi_4(s)\}u(s)+e_t(s)$$

$$= [G(s)+\Delta G(s)]u(s) + e_t(s) \qquad (34)$$

expressed in terms of the (known) transfer function $G(s) = M\Phi(s)B$, where

$$\Delta G(s) \triangleq \Psi_4(s)+M\Phi(s)\Psi_3(s)+[M\Phi(s)J(s)+\Psi_2(s)\Phi(s)+\Psi_2(s)\Phi(s)J(s)][B+\Psi_3(s)]$$

$$e_t(s) = [M+\Psi_2(s)]\Phi(s)[I+J(s)][w(s)+E_{10}e_t(s)]+E_{o1}e_t(s)+v(s) \qquad (35)$$

Both expressions (31) and (35) simplify greatly by special choices of coordinates of (15). Without loss of generality one can take $\Delta M = 0$ and $M_T = 0$ if $\dim z \le \dim x_R$. To see this, note that a similarity transformation on the state $x = T\eta$ can always take the measurement $z = Mx$ to $z = MT\eta = [I\ 0]\eta$ if rank $M = \dim z \le \dim x$. Of course, we cannot construct this T (since we don't know M), but we know that it exists. Therefore by assuming that the states of our model are not fewer than the number of measurements, we can, from (29, 32) without loss of generality set

$$E_{o1}P_i = 0\ ,\quad E_{o1}Q_i = 0\ .$$

Hence, from (32), $Q_{iz} = 0$, and $P_{iz} = 0$ for all i. This simplifies (35) to

$$\Delta G(s) = M\Phi(s)\Psi_3(s) + [M+\Psi_2(s)]\Phi(s)J(s)[B+\Psi_3(s)]\ .$$

$$e_t(s) = M\Phi(s)[I+J(s)][w(s)+E_{10}e_t(s)] + v(s) \qquad (36)$$

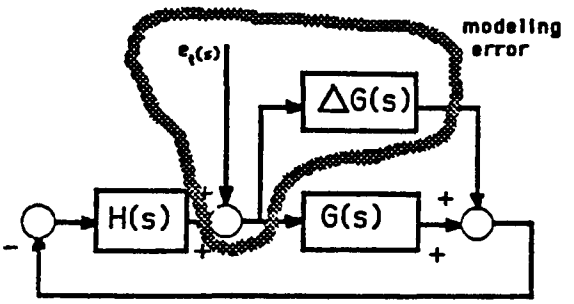
The state equations (31) also simplify in an obvious way when $\Delta M = 0$. The term e_t in (33) and (36) represents unknown time varying excitations arising from a combination of both external disturbances and errors of model order. When $\Delta G(s)$ is assumed zero, some attention has been paid in the literature [16] to the determination of upper bounds on e_t which can be tolerated before losing stability. In these studies, the $\Delta G(s)$ term in (34) and (36) has been *ignored* in the control design. See that u(s) multiplies $\Delta G(s)$. Hence, it should be emphasized that *the effects of model error* e_x *in* (18) *cannot be assessed independently of the control law* u(·), and this is the fundamental pitfall that prevents the modeling problems and the control problems from being separable.

It is also common to analyze the effects of $\Delta G(s)$, but ignore the effects of $e_t(s)$ in (34),(36) [17-18]. However, these terms both *coexist* in the presence of errors in model order. Imagine the physical process $z(s) = [G(s)+\Delta G(s)]u(s) + e_t(s)$ and a model G(s)

which might be used for design of the controller H(s). See Fig. 5a. Control theory has reached a sophisticated level of maturity *assuming* that the model $G(s) + \Delta G(s)$ is specified *a priori*, or, if not given, then the assumption is that the model is something that *exists* in an absolute sense irrespective of the control policy (which is yet to be developed). The thesis of this paper is that the model and the control policy must be developed *together* and that no meaning (in an absolute sense) can be attached to *either* one in isolation. These concepts do not readily lend themselves to established definitions of *robustness*, or to certain *identification* concepts.

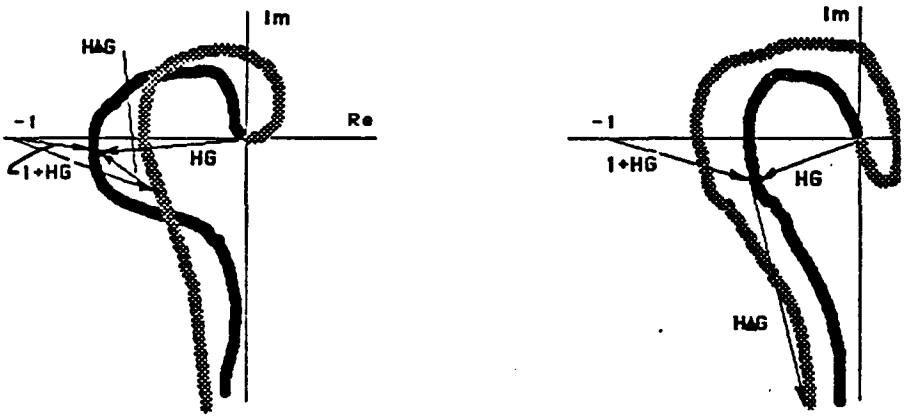
In the time domain characterization of the model errors, the parameters P_i, Q_i in (29)-(31) are key. In the frequency domain characterization of model errors, the spectrum $\Delta G(j\omega)$ and $e_t(j\omega)$ in (36) are key. Which is easier to approximate or parameterize in practice? We cannot say. H^∞ methods rely on a characterization of $\Delta G(j\omega)$, $e_t(j\omega)$, [17, 18, 20, 21, 25]. This is never possible exactly, but neither is an exact determination of P_i, Q_i in (31). One available method does not need either the frequency or time domain characterization of modeling error, but only needs a geometric condition about the space of the model error vector e_x in (18). If e_x lies in the column space of B, (a "matching condition") then only an upperbound on e_x is needed and not its spectral content [22]. The matching conditions are easily violated in problems with errors of model order. Yedavalli [23] shows bounds on each element of the A matrix perturbations which preserve stability.

We cannot associate gain and phase margins with tolerance of any one of the four categories of model error in (19), although errors in model order ($x_T \neq 0$) will certainly modify the phase of the system. It can be said (without any sort of precision), that larger phase margins may allow the design to be less sensitive to errors in model order. Gain margins on the other hand do not necessarily provide tolerance to either parameter error or model order error. Refer to Fig. 5 and $\Delta G(s)$ to see that the Nyquist test for stability is satisfied if over all frequencies ω, the full length of the vector $1+HG(j\omega)$ is larger than



Block Diagram of Perturbed System

Fig. 5a



Nyquist Plot of Perturbed System

Fig. 5b

the length of the vector $H\Delta G(j\omega)$, since the -1 point cannot be encircled in this case. Hence, a sufficient condition for stability is

$$1+HG(j\omega) > H\Delta G(j\omega) \quad \text{for all } \omega$$

or

$$1 > [1+HG(j\omega)]^{-1}[H\Delta G(j\omega)] , \quad \text{for all } \omega \tag{47a}$$

This condition may be extended to the matrix case as follows, [11, 14, 15, 18],

$$1 > \max_{\omega} \lambda \{[I+HG(j\omega)]^{-1}H\Delta G(j\omega)\} . \tag{47b}$$

where $\lambda\{\cdot\}$ denotes the square root of the largest eigenvalue of matrix $\{\cdot\}^*\{\cdot\}$. Note, however, that this can be an extremely conservative condition, since it is possible for the vector $H\Delta G(j\omega)$ to be much longer than $1+HG(j\omega)$ without causing an encirclement. See Fig. 5b for a stable situation which violates (47). However, a more fundamental limitation of these results is due to the fact that Fig. 5a does *not* describe most physical situations if the $e_t(s)$ term in (35), (36) has been ignored. The significant term e_t is composed of the same source of errors which make up $\Delta G(s)$, (See that J(s) is a model order error term which appears in both eqns. of (35). Hence, the above stability results are extremely conservative and they ignore $e_t(s)$. Therefore they do not readily extend to include performance guarantees. Stability is the most studied subject in control, but stability is usually not sufficient for successful operation.

CRITICISM OF THE "MODELING PROBLEM"

The traditional *idea* of the "modeling problem" is as follows:

The Modeling Problem:

Find a set (or a class) of differential equations S_1 *describing the dynamical relationships between the response* z(t) *and the (unspecified) inputs {controls* u(t), *disturbances* w(t), *and initial conditions}.*

The flaw in this task statement is the presumption that *there exists* a set of differential equations which relate $z(t)$ and u(t), irrespective of u(t). We argue that *any* set of differential equations S_1 is only an approximation of the physical phenomenon S_o, and it was shown in (28) that the errors associated with this approximation cannot be assessed, qualitatively or quantitatively, independently of u(t). In other words, knowledge of the *control* inputs u(t) are required in any assessment of *model* fidelity. In modeling and identification literature it is common to talk about model errors with respect to a *truth* model, S. There is no truth model ($S_1 \neq S_o$). The model and its controller should be discussed as a *pair*. They have no significance separately.

CRITICISM OF THE "CONTROL PROBLEM"

The traditional *idea* of the "control problem" is as follows:

The Control Problem:

Given the set of models $\sum$ *which describe the dynamical process, find an appropriate control* u(t) *or controller* u(z(t), t) *to meet a specified set of control objectives.*

The flaw in this task statement is the presumption that the class of models $\sum$ which appropriately describe the process exists independently of knowledge of u(t), or consequently, of knowledge of the controller generating u(t). Now the control law cannot logically be specified prior to model development. Thus, if one wishes to squeeze the best possible performance from the controller design, then one cannot ignore the

MODELING PRINCIPLE V:

The modeling and control problems are not separable and are necessarily iterative.

This means that for any given $G(s) + \Delta G(s)$, in Fig. 5 the development of a G(s) and H(s) is an iterative process.

Several implications of the *modeling and control inseparability principle* are:

(a) The phrase "model of the plant" is a misnomer. We must refer to a model as being appropriate *under the influence of a particular controller*. Hence, we must refer to a (model, controller) *pair* as appropriate or inappropriate for each other, with respect to a given plant.

(b) Only "local" properties can be stated concerning the model and the controller. This means that the interpretation of both classical and modern control theory must be tempered with this knowledge, since parameters of *neither* plant model nor controller can be taken to infinity (or wide ranges). Three examples follow.

(b1) The Root Locus theory presumes a fixed plant while the controller gain goes to infinity. But the fidelity of the plant model *depends* upon the control gain. Hence, *the same model of the plant is not appropriate at both the vicinity of the open loop poles and the open loop zeros,* and a given root locus plot is never reliable in the vicinity of both the open loop poles and the open loop zeros.

(b2) The Nyquist plots are reliable only over a limited frequency range and certainly not reliable in the vicinity of the origin, where $\omega \rightarrow \infty$. If this region of uncertainty extends to a *unit* radius around the origin, then even the stability results of the Nyquist plot are suspect. See Fig. 6.

(b3) In LQG theory it is presumed that the model is fixed and that the weights in the performance index may be varied over wide ranges. This generates the theoretical

predictions of maximal accuracy in Fig. 7 (solid curve). However, the actual performance follows the dotted curve, due to modeling errors (note the decreasing effect of modeling errors as the control effort decreases). Thus, the deviation of the actual performance from the theoretical is *greatest* where maximal accuracy predictions are made. Hence there is often a large discrepancy between achieved and predicted maximal accuracy, and large errors in the value of the control effort at which maximal accuracy occurs. These inequalities always hold, from (9c, 9d).

$$u_a^2 < u_t^2 , \qquad y_a^2 > y_t^2 .$$

The class of all stabilizing controllers *for a given plant model* is described in [25] and its references. The advantage of this knowledge is that the total design freedom is characterized for accomplishing performance beyond stability. However, the given model might not accurately describe the plant over "all stabilizing controllers (for the *given* model)," due to MODELING PRINCIPLE V. Hence, the "robust controller" based upon the *given model* might actually destabilize the actual system due to the fact that when the observer based controller is far away from its nominal design (for the *given model),* the model that accurately portrays the plant is not close to the *given* plant model. Much more work is required to capture the class of observer based controllers which allow appropriate changes in the model as a function of the controller.

A "three-model" control theory would serve to keep control designs honest: Model 1: an *Evaluation Model* used to simulate the real system (until prototype testing is available); Model 2: a high order *model for anlytical predictions of controlled performance;* and Model 3: a low order *model for control design.*

"Honesty" is maintained by maintaining distinction between the three models. Model 2 can never be equal to model 1 due to Modeling Principle IV. Model 3 must change as a function of controller design due to Modeling Principle

V. Dangerous (faulty) prections occur whenever *any two* of these models are coincident, for in such cases control design mathematics can easily use high gain controllers to achieve a high level of performance without paying the (inevitable) penalty for modeling errors, see conjecture (9). See the "Controller data base" S_o, S_1, S_2, in Fig 6a.

4.0 The Structure of Errors in The Closed-Loop System

Suppose our hypothetical "physical system" (1) is now driven by the linear dynamical controller with transfer matrix H(s)

$$u(s) = H(s)z(s) \tag{37}$$

Without loss of generality we may associate a state space realization with (37) and write

$$H(s) = G(sI - A_c)^{-1}F \tag{38}$$

or in state form

$$\begin{aligned} \dot{x}_c &= A_c x_c + Fz, \quad x_c \in R^{n_z} \\ u &= Gx_c \quad , \quad u \in R^{n_u} \end{aligned} \tag{39}$$

In this Section we ask how the poles of the closed loop system behave as either G or F approaches zero. Next arbitrary gains G and F are considered but with a restrictive assumption about parameter errors ($e_p = 0$). It will prove convenient to write results in terms of a set of nonzero auxillary matrices (A, B, M) which can *always* be found satisfying

$$A + BG - FM \triangleq A_c \tag{40}$$

for any given A_c, G, F.

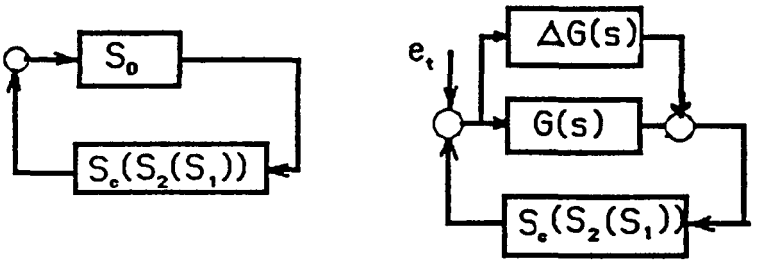


Fig. 6a Controller Data Base

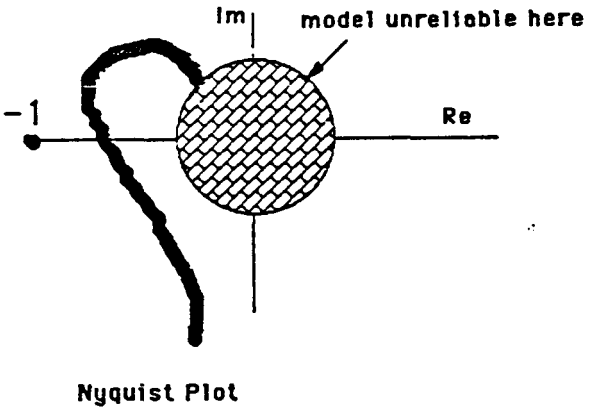


Fig. 6b Nyquist Diagram

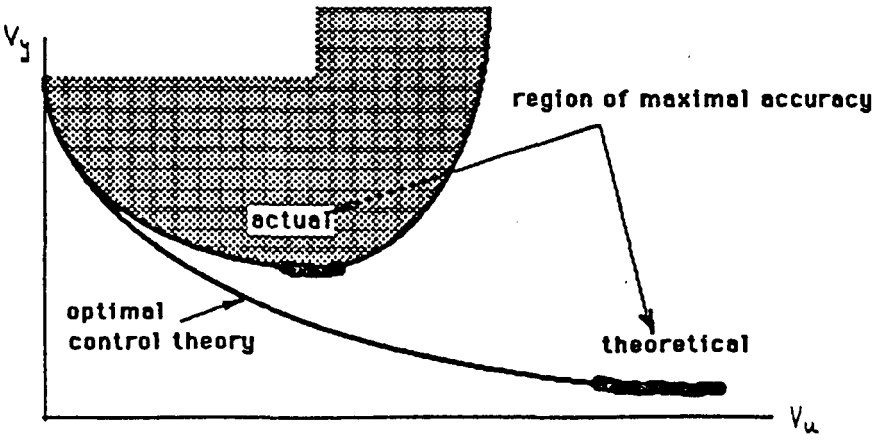


Fig. 7 Performance Evaluation (with "Truth" model)

The system (13) is now driven by controller (21). Subtract (39) from (17). This gives a differential equation for $\tilde{x} \triangleq x_R - x_C$. Now writing (13), (39) in terms of the states x and $\tilde{x}$ yields the homogeneous part of the closed loop system

$$\begin{bmatrix} \dot{x} \\ \dot{\tilde{x}} \end{bmatrix} = \left[\begin{array}{c|c} A + BGE_{10} & -BG \\ \hline (A_{RT} - FM_T)E_{01} + (\Delta A - F\Delta M + \Delta BG)E_{10} & A - FM - \Delta BG \end{array}\right] \begin{bmatrix} x \\ \tilde{x} \end{bmatrix} \tag{41}$$

$$\begin{bmatrix} y \\ u \end{bmatrix} = \left[\begin{array}{c|c} C & 0 \\ \hline GE_{10} & -G \end{array}\right] \begin{bmatrix} x \\ \tilde{x} \end{bmatrix}$$

where $E_{10} \triangleq [I_{n_x}\ 0]$ and $E_{01} \triangleq [0_{n_x}\ I]$.

In the limit as $G \to 0$, the eigenvalues of (41) become those of the block diagonal matrices A and A - FM. This conclusion is summarized as follows.

Theorem 1: *In the limit as* $G \to 0$, *any linear system* (13) *driven by any controller of the form* (38), *has the eigenvalues of A and* (A − FM).

Theorem 1 suggests that the low gain controller (characterized by small G) is stable if the open-loop system is stable and if the matrix (A - FM) (which is entirely under the design of the analyst) is stable.

Now multiply (39) by E_{10}^T, then subtract this equation from (13). This defines the vector $\bar{x} \triangleq x - E_{10}^T x_c$. The homogeneous part of the closed loop system (13), (39) can now be described in terms of states $\bar{x}$ and x_c, yielding

$$\begin{bmatrix} \dot{\bar{x}} \\ \dot{x}_c \end{bmatrix} = \left[\begin{array}{c|c} A - E_{10}^T FM & E_{10}^T(\Delta A + \Delta BG - F\Delta M) + E_{01}^T(A_{TR} + B_T G) \\ \hline FM & A + BG + F\Delta M \end{array}\right] \begin{bmatrix} \bar{x} \\ x_c \end{bmatrix} \tag{42}$$

$$\begin{bmatrix} y \\ u \end{bmatrix} = \left[\begin{array}{c|c} C & CE_{10}^T \\ \hline 0 & G \end{array}\right] \begin{bmatrix} \bar{x} \\ x_c \end{bmatrix}$$

In the limit $F \rightarrow 0$ the eigenvalues of (42) become those of the block diagonal matrices A and $A + BG$. This conclusion is summarized as follows.

Theorem 2: *In the limit as* $F \rightarrow 0$, *any linear system* (13) *driven by any linear controller of the form* (39) *has the eigenvalues of A and* $(A + BG)$.

Again, the reader is reminded that $A + BG$ is under the design of the analyst.

For further insight into the effects of modeling errors suppose that the gains G and F are not small, but the parameter errors are zero ($\Delta A = 0$, $\Delta B = 0$, $\Delta M = 0$). Then the following is true.

Theorem 3: In the absence of parameter errors ($\Delta A = 0, \Delta B = 0, \Delta M = 0$), *the closed-loop system eigenvalues are those of* $(A - FM)$, A_T, *and* $(A + BG)$, *if* x_T *is either uncontrollable or unmeasurable.*

Proof: Letting ($\Delta A = 0$, $\Delta B = 0$, $\Delta M = 0$) in (42) yields

$$\begin{bmatrix} \dot{x} \\ \dot{x}_c \end{bmatrix} = \left[\begin{array}{c|c} A - E_{10}^T FM & E_{01}^T (A_{TR} + B_T G) \\ \hline FM & A + BG \end{array}\right] \begin{bmatrix} x \\ x_c \end{bmatrix}$$

$$\begin{bmatrix} y \\ u \end{bmatrix} = \left[\begin{array}{c|c} C & C E_{10}^T \\ \hline 0 & G \end{array}\right] \begin{bmatrix} x \\ x_c \end{bmatrix} \tag{43}$$

The block diagram of the homogeneous part of the closed-loop system using the notation of (15) and (39) is given in Fig. 8. Using the definitions of E_{10} and E_{01}, (43) is further expanded as follows.

$$\begin{bmatrix} \dot{\hat{x}}_R \\ \dot{\hat{x}}_T \\ \dot{x}_c \end{bmatrix} = \left[\begin{array}{c|c:c} A_R - FM_R & A_{RT} - FM_T & 0 \\ \hline A_{TR} & A_T & A_{TR} + B_T G \\ \hdashline FM_R & FM_T & A + BG \end{array}\right] \begin{bmatrix} \hat{x}_R \\ \hat{x}_T \\ x_c \end{bmatrix} \tag{44}$$

$$\begin{bmatrix} y \\ u \end{bmatrix} = \left[\begin{array}{c|c:c} C_R & C_T & C_R \\ 0 & 0 & G \end{array}\right] \begin{bmatrix} \hat{x}_R \\ \hat{x}_T \\ x_c \end{bmatrix}$$

We shall first let x_T in (15) be uncontrollable. This is equivalent to the statement that ($A_{TR} = 0, B_T = 0$) in (15) and in Fig. 8. (Recall the controllable canonical form). This makes the upper right matrix of (44) zero (using the *dotted* line partitions). In this event the eigenvalues of (44) become those of (A+BG) and those of the upper left block matrix, (using again the *dotted* line partitions). But since $A_{TR} = 0$ this upper left matrix is now also block diagonal, and therefore has the eigenvalues of (A_R–FM_R) and A_T. In the absence of parameter error $A_R = A$, $M_R = M$ and the first part of the theorem is proved.

For the second part of the proof let x_T in (15) be unmeasurable (unobservable from the measurement z). This is equivalent to the statement that ($A_{RT} = 0, M_T = 0$) in (15) and in Fig. 8. (Recall the observable canonical form). This makes the upper right matrix of (44) zero using the *solid* line partitions. In this event the eigenvalues of (44) become those of (A_R–FM_R), and those of the lower right partition, again using the *solid* line partitions. But since $M_T = 0$ this lower right matrix is now also block diagonal, and therefore has the eigenvalues of A_T and (A+BG). Now in the absence of parameter errors $A_R = A_R$, $M_R = M$. This gives the same set of eigenvalues (of (A-FM), A_T, and (A+BG)) as in the first part of the proof. Hence, the theorem is proved. #

A clear discussion of the concepts in Theorem 3 first appeared in [29] using modal coordinates and the phrases "observation spillover" and "control spillover", (signals "a" and "b" in Fig. 8 respectively). While such phrases are descriptive, they disguise the fact that the "controllability" and "observability" of the states they x_T are the important

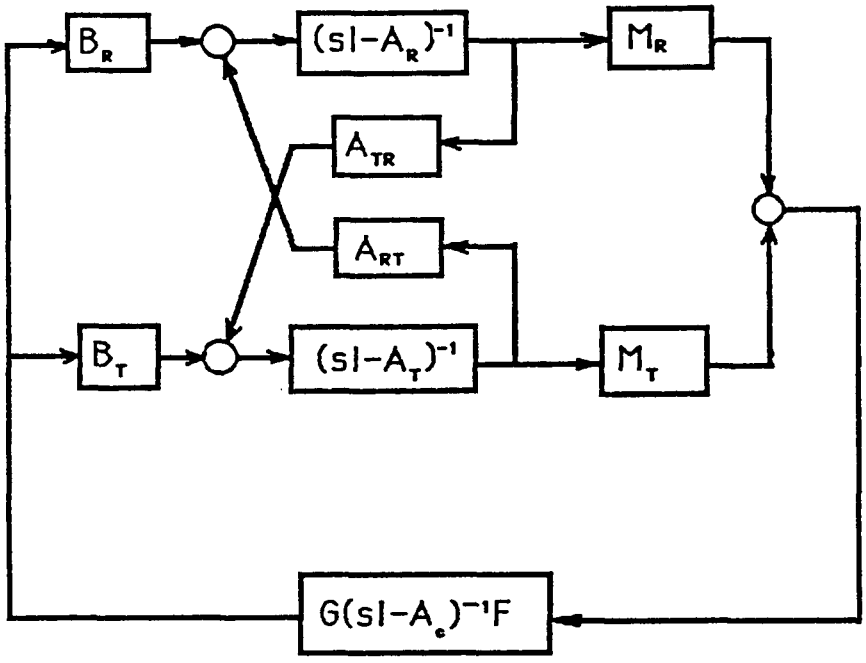


Figure 8 Structure of Linear System Control

concepts. Hence, we prefer the standard "controllability, observability" description of the phenomenon to the phrase "spillover".

Theorems 1 - 3 suggest that it is important that the matrices {(A-FM), (A+BG), A_T} be asymptotically stable, and more significantly, that this conclusion remain *independent* of the methods by which the controller parameters (A_c, F, G) were derived! (Notice that the optimal LQG controller stabilizes (A-FM), (A+BG)). Suppose the controller (39) was designed based upon the assumed model (A_R, M_R, B_R) so that this model had certain desired behavior in the closed loop. Fig. 8 makes it clear that, in the absence of parameter errors, the fundamental cause of deviation in system behavior from predicted behavior is the *relative degree of controllability and observability of the states* x_T. While controllability and observability prove to be of great benefit for control in the *absence* of model errors, complete controllability and observabilty would be a serious handicap in the real world. Control designers are indeed *fortunate* that most of the "deleted states" associated with "real world x_T" are uncontrollable and unobservable. Otherwise there would be even fewer successful control designs to celebrate in practice.

Modeling Principle VI: (Uncontrollability and Unobservability of Dynamic Systems)

Using any number of sensors and actuators, the physical plant will not be completely observable nor controllable.

This proposition requires some explanation since observability and controllability are mathematical properties associated with a mathematical model, whereas the "physical plant" defies exact description by any mathematical model. Suppose one improves a given mathematical representation of the physical plant by adding additional dynamics which were originally ignored in the model. As one continues this process, adding more and more details so that the new model more accurately models the physical plant, the mathematical model eventually becomes both uncontrollable and unobservable. In other words, an uncontrollable, unobservable model can always be constructed to provide a

closer representation of the physical plant than any controllable, observable model. It takes little convincing to see that this argument is correct. Taken to an extreme it is perhaps obvious that the molecular motions in an aircraft wing or the seat cushion vibrations in the aircraft cannot be controlled by aileron actions nor observed by rate gyros. However, one need not resort to such extreme examples, using infinitesimal effects. Note that the example in Fig. 9 describes an uncontrollable system, and the uncontrollable part is a nontrivial part of the system dynamics. The control is the torque applied to the reaction wheel.

Recall that if (A,B) is controllable [or (A,C) is observable] then A is asymptotically stable if and only if there exists a positive definite solution to $0 = XA^{*}+AX+BB^{*}$ [or to $0 = KA+A^{*}K+C^{*}C$]. The important impact of the lack of observability or controllability is that the stability proofs which rely on observability or controllability cannot be used to assure that the physical plant will be stable. Indeed, stability is a mathematical concept relating to a mathematical model, and hence the physical plant can never be *proved* stable by mathematics. (Due to the fact that the physical plant defies exact mathematical description, and the amount of modeling precision required to predict stability is not known a priori). We can only say that the *model* is stable subject to a given range of parameter values, or a given magnitude of the model error vector e_x, etc.

Modeling Principle VI seems to be at odds with the notion of generic controllability, observability [24]. This result [24] states correctly that adding arbitrarily small numbers to every element of (A,B,C) will make the matrix triple controllable and observable. Hence, the notion of generic controllability, observability suggests that physical systems (such as Fig. 9) are arbitrarily "close" to observability controllability. This is consistent with the view of at least half a dozen control texts which state emphatically that all real world plants are controllable and observable. But the addition of arbitrarily small numbers in every element of (A, B, C) is a mathematical exercise that

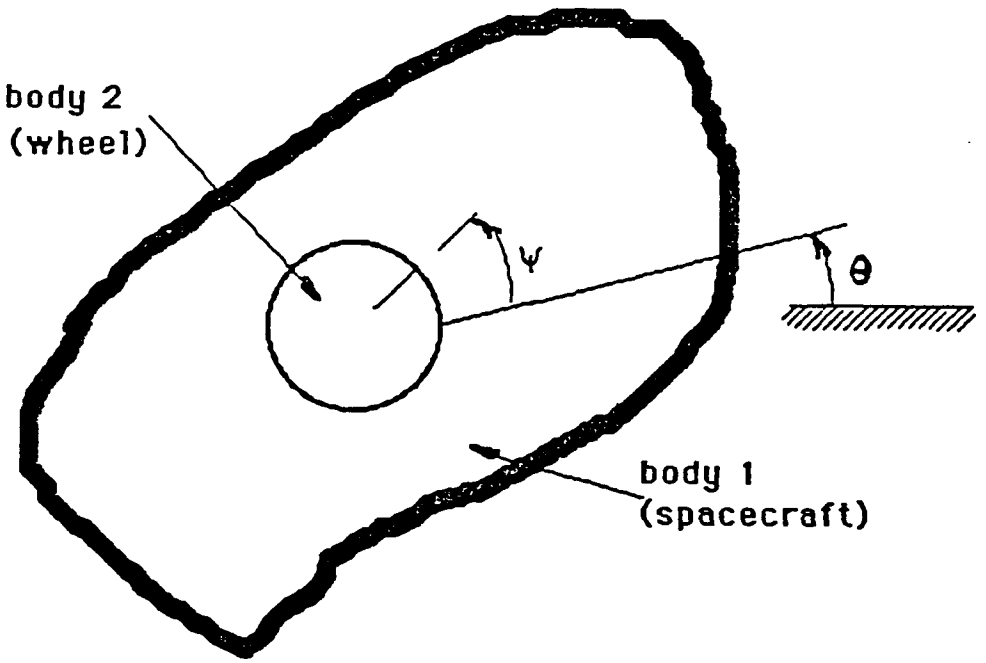


Fig. 9 An Uncontrollable System

has no basis in the physics of the problem and cannot describe physical behavior. For the physical system in Fig. 9, a control which can simultaneously regulate both $\theta(t)$ and $\psi(t)$ to arbitrary values would contradict the conservation of angular momentum. Hence, we can safely claim that this system is arbitrarily "far" away from controllability rather than arbitrarily "close" as the notion of generic controllability suggests. Rather, we should interpret the generic controllability result as good reason not to believe *computer calculations* of controllability, observability, since roundoff errors are the equivalent of adding small errors to every element of (A, B, C), as in the thesis of generic controllability, observability. A common reaction of those that insist that physical systems are state controllable is "Who would design an uncontrollable system?" Actually, it is impossible to do otherwise. The physical system will always be uncontrollable (and that is good news not bad news), but we should not discard the uncontrollable part in the control design process. This is verified as follows.

It can be shown that the stable, uncontrollable, observable system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u$$

$$y = [C_1\ 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

minimizes

$$V = \int_0^\infty (y^T Q y + u^T R u) dt$$

with the control law

$$u = G_1 x_1 + G_2 x_2$$

where G_1 is the standard state feedback gain for optimizing the plant (A_{11}, B_1, C_1) and

G_2 satisfying $0 = KA_{22} + [A_{11} + B_1G_1]^TK + K_{11}A_{12}$, $G_2 = -R^{-1}B^TK$ depends upon (A_{22}, A_{11}, B_1, G_1) and is *not* zero unless x_2 is unobservable in y. Hence, the often heard argument "the uncontrollable part of the system should be deleted since we cannot control it" is faulty logic. The optimal controller (or any other reasonable controller) will alter its control of x_1 (the controllable part) with the knowledge of the dynamical interactions between x_1 and x_2, even though x_2 itself cannot be controlled. For example, the minimal energy optimal control that pushes x_1 toward a certain desired value $\bar{x}_1$ might be zero if the dynamics between x_1 and x_2 are such that x_1 is *naturally* (without control action) driven toward $\bar{x}_1$. If x_2 were deleted *a priori* the optimal control would be greater than zero, hence not optimal for the complete system.

To illustrate the relationship between controllability observability and stability consider the second order system in modal coordinates.

$$\begin{bmatrix} \dot{x}_R \\ \dot{x}_T \end{bmatrix} = \begin{bmatrix} A_R & 0 \\ 0 & A_T \end{bmatrix} \begin{bmatrix} x_R \\ x_T \end{bmatrix} + \begin{bmatrix} B_R \\ B_T \end{bmatrix} u, \quad z = (M_R, M_T)x$$

Note that mode x_T is observable (controllable) if and only if $M_T \neq (B_T \neq 0)$. Now consider any first order controller

$$u(s) = H(s)z(s)$$

where

$$H(s) = G(sI - A_c)^{-1}F$$

and where we have chosen to describe A_c in the form

$$A_c = A_R + B_RG - FM_R$$

where G is chosen so that

$$A_R + B_R G < 0$$

and F is chosen so that

$$A_R - F M_R < 0 .$$

The necessary and sufficient closed loop stability requires the product of controllability and observability of x_T to be limited by

$$|M_T B_T| < |A_R - F M_R|\,|A_T|\,|A_R + B_R G|\,\frac{1}{|G A_R F|} \qquad (46)$$

This simple third order example provides important insight into the necessary and sufficient condition for stability. The closed loop will be stable if the neglected part is unobservable ($M_T = 0$) or uncontrollable ($B_T = 0$), as promised by the sufficient condition, Theorem 3. But (46) also shows the upper bound on $|M_T B_T|$ which will allow stability even if M_T, B_T are not zero. Hence, the right hand side of (46) is a measure of the conservatism in Theorem 3, and this measure is a function of *both* the modeling errors (A_T) and the control gains (G, F).

Control Design Considerations: (Trading Stability and Performance)

Practical control designs always require some iteration or some fine tuning during experimentation with the real hardware. There are two fundamentally different strategies for these two phases of design. These strategies differ by the manner in which they handle stability and performance concerns.

The Stability Design Scenario:

Design phase A: Design for stability

Design phase B: Tune for performance.

The Performance Design Scenario:

Design Phase A: Design for performance

Design Phase B: Tune for stability

"Stability Design" includes (i) guaranteeing specified gain and phase margins by root locus, Bode, Nyquist, (ii) designing by pole or pole region assignment, (iii) designing by Liapunov techniques, (iv) designing by H^{∞} techniques, etc. "Performance Design" includes satisfying input or output variance (or L_2) bounds, perhaps by optimal control.

Phase B is intended to be a simple gain change. In the "Stability Design" scenario the gain is usually tuned *up* a bit to improve performance without destroying stability. In the "Performance Design" scenario the gain is usually turned *down* a bit to improve stability margins while maintaining acceptable performance.

Of course, the design tools have been oversimplified in these scenarios, in order to draw distinctions among points of view. There are circumstances where each scenario is preferred. In some situations in the Stability Design scenario, adequate performance may be impossible to achieve by a simple gain change in Phase B. In this case the Performance Design scenario is preferred. In some situations with the Performance Design scenario, adequate stability margins may be impossible to achieve by a simple gain change in Phase B. In this case the Stability Design scenario is preferred. An overwhelming proportion of control literature has focused on *stability* to support the Stability Design scenario. However, as society demands more *performance* in modern systems, the development of practical and theoretical tools which focus more on performance will strengthen the Performance Design scenario.

Stability and performance are competing partners in the design process. They usually do not naturally cooperate. Improving one usually (but not always) degrades the other, and good tradeoff methods are needed from the research community. In fact, there are many examples where stability, performance, and sensitivity are *three* competitors in

the design process, where design changes that improve one of the three necesarily degrades the other two. Modeling and control design theory has not yet provided a convenient and practical tradeoff among these three design goals. Part of the problem is lack of an agreement about what *kind* of sensitivity, stability, and performance measures to use. For example, in [33] we have the following conclusion

Theorem

Let λ_i, i=1,2,..,m, *denote the distinct eigenvalues of* A. *Define a measure of root sensitivity by*

$$S \triangleq \sum_{i=1}^{n} \left\| \frac{\partial \lambda_i}{\partial A} \right\| \quad , \quad \|[\cdot]\|^2 = \mathrm{tr}\,[\cdot]^*[\cdot] \ .$$

The lowest bound on root sensitivity is $S \geq n$ *and* $S = n$ *if and only if* A *is normal,* $(AA^T = A^TA)$.

Similar results are available for $\left\| \frac{\partial \,\mathrm{Re}\,\lambda_i}{\partial A} \right\|$, $\left\| \frac{\partial \,\mathrm{Im}\,\lambda_i}{\partial A} \right\|$ types of sensitivity [34]. For a simple pitch control problem for an aircraft it was shown in [33] that minimal root sensitivity in the closed-loop system is achieved (that is, normality is achieved $(A_cA_c^T = A_c^TA_c,\ A_c \triangleq A + BG)$ only at values of G which were destabilizing. Furthermore, as a design objective a "nearly normal" objective can be added subject to performance or stability constraints. In this event the design can depart arbitrarily far from minimal root sensitivity (even though the design is "nearly" normal). This is due to the fact that the "abnormality" measure $\hat{a} \triangleq \|A_cA_c^T - A_c^TA_c\|$ can be a convex function of the parameters in A_c even when the sensitivity measure S is not. This means that "nearly normal" *does not* mean "nearly minimally sensitive". The "nearly normal" design objective has been widely used in both frequency and time domain designs. It has popular appeal due to the fact that symmetric matrices are normal, easy to work with, and have orthogonal eigenvectors. However, such examples in [33] point out that while

normal matrices (or orthogonal eigenvectors, or symmetric matrices) coorespond to global minimal root sensitivity, a "nearly normal" design goal is suspect since "nearly normal" does not coorespond to "nearly insensitive".

Since root sensitivity is not easily incorporated in designs that minimize other types of performance, some authors seek to incorporate different kinds of *response* sensitivity. The max entropy approach of [30] minimizes an L_2 norm of the inputs and outputs while modeling the uncertain parameters as zero mean white noise. This leads to two Riccati-like plus two Liapunov-like equations to solve by iteration. A case study of this approach is presented in [32], by comparing it with a different optimal controller, one that minimizes the L_2 norm of nominal performance plus a weighted L_2 norm of input and output sensitivity $\|\frac{\partial u}{\partial p}\|^2$, $\|\frac{\partial y}{\partial p}\|^2$ where p is the vector of uncertain parameters (not random). These conclusions emerged from these case studies (comparing methods [30], [31], and [32]):

(i) Presently method [31] cannot treat uncertain parameters appearing in the measurement matrix, method [30] cannot treat parameters p_i appearing both in B and C, while method [32] can do both.

(ii) Presently method [30] cannot treat parameters p_i appearing nonlinearly in (A, B, C), while methods [31] and [32] require only differentiability of (A(p), B(p), C(p)) with respect to p.

(iii) When comparing L_2 performance of inputs, outputs with stability margins, the case studies favored the results of method [32] over methods [30] and [31] in most cases studied, but ranked the results of method [32] equivalent to the results of method [30] in two of the cases studied.

Case studies should be encouraged in the areas of robustness and model error compensation since applicable theory is quite limited, and case studies can point the way

VI) The physical plant is always unobservable and uncontrollable (fortunately).

The impact of each of these principles is discussed to explore limitations of available theory. An understanding of these principles can aid in the search for successful control designs and for improved "robustness" definitions.

References

[1] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification.* MIT Press, Cambridge, MA, 1987.

[2] M. F. Hutton and G. B. Friedland, "Routh Approximations for Reducing Order of Linear, Time-Invariant Systems," *IEEE Trans. Autom. Control,* AC-20(3), 329-337, 1975.

[3] W. Eckhaus, "Formal Approximations and Singular Perturbation Techniques," *SIAM Rev.,* 19, 593, 1977.

[4] E. I. Verriest, "Low Sensitivity Design and Optimal Order Reduction for the LQG Problem," *Proc. 24th Symp. Circuits Systems,* Chicago, June 1981, pp. 365-369.

[5] E. C. Y. Tse, J. V. Medanic, and W. R. Perkins, "Generalized Hessenberg Transformations for Reduced-Order Modeling of Large Scale Systems," *Int. J. Control,* 27, 493-512, 1978.

[6] R. N. Mishra and D. A. Wilson, "A New Algorithm for Optimal Reduction of Multivariable Systems," *Int. J. Control,* 31(3), 443-466, 1980.

[7] D. C. Hyland and D. S. Bernstein, "The Optimal Projection Approach to Model Reduction and the Relationship Between the Methods of Wilson and Moore," *Proc. 23rd IEEE CDC,* Las Vegas, December 1984.

[8] B. C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability and Model Reduction," *IEEE Trans. Autom. Control,* AC-26(1), 17-32, 1981.

[9] L. Parnebo and L. M. Silverman, "Model Reduction Via Balanced State Space Representation," *IEEE Trans. Autom. Control,* AC-27, 1982.

[10] J. Rissanen and T. Kailath, "Partial Realization of Random Systems," *Automatica*, 8, 389-396, 1972.

[11] K. Glover, "All Optimal Hankel-Norm Approximations of Linear Multivariable Systems and Their L^∞-Error Bounds, *Int. J. Control*, 39(6), 1115-1193, 1984.

[12] C. D. Villemagne and R. E. Skelton, "Model Reductions Using a Projection Formulation", *Int. J. Control*, 1987, Vol. 46, No. 6, 2141-2169.

[13] P. V. Kokotovic, R. E. O'Malley, Jr., and P. Sannuti, "Singular Perturbations and Order Reduction in Control Theory - An Overview," *Automatica*, 12, 123-132, 1976.

[14] B. D. O. Anderson and Yi Liu, "Controller Reduction: Concepts and Approaches," *ACC*, 1987 Plenary Paper, Minneapolis.

[15] D. Enns, "Model Reduction for Control Design," Ph.D. Thesis, 1984, EE Dept., Stanford University, Stanford, California.

[16] D. D. Siljak, *Large-Scale Dynamic Systems, Stability and Structure*. North-Holland, 1978.

[17] J. C. Doyle, "Quantitative Feedback Theory and Robust Control," *Proceedings ACC*, Seattle, 664-665 (1986).

[18] Yi Liu and B. D. O. Anderson, "Controller Reduction via Stable Factorization and Balancing," *Int. J. Control*, 1986, 44, pp. 507-531.

[19] J. C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," *IEEE Trans. Auto. Control*, AC-26, No. 1, 1981, pp. 4-16.

[20] S. Boyd, V. Balakrishnan, and P. Kabamba, "On Computing the H_∞ Norm of a Transfer Matrix," to appear.

[21] M. J. Grimble, "H_∞ and LQG Robust Design Methods for Uncertain Linear Systems," pp. 91-96, in *Model Error Concepts and Compensation,* Pergamon Press 1986, New York.

[22] B. R. Barmish, M. Corless, G. Leitmann, "A New Class of Stabilizing Controllers for Uncertain Dynamical Systems," *SIAM J. Control,* 21(2), 246-255 (1983).

[23] R. K. Yedavalli, "Perturbation Bounds for Robust Stability in Linear State Space Models," *Int. J. Control,* Vol. 42, pp. 1507-1517, Dec. 1985.

[24] W. M. Wonham, *Linear Multivariable Control,* Springer-Verlag, New York, 1979, p. 43-44.

[25] S. P. Boyd, V. Balakrishnan, C. H. Barratt, N. M. Khraishai, X. Li, D. G. Meyer, and S. A. Norman, "A N ew CAD Method and Associated Architectures for Linear Controllers."

[26] A. Yousuff, "Model Errors: Should they Always be Minimized?" *IEEE Midwest Symposium on Circuits and Systems,* 1985.

[27] N. H. McClamroch and J. K. Aggarwal, "On Equivalent Systems in Optimal Control and Stability Theory," *IEEE Trans. Auto. Control,* 1967.

[28] R. Liu and R. J. Leake, "Exhaustive Equivalence Classes of Optimal Systems with Separable Controls," *J. SIAM on Control,* vol. 4, pp. 678-685, November 1966.

[29] M. J. Balas,"Feedback Control of Flexible Systems" *IEEE Trans. Auto. Control* Vol. AC-23, Aug. 1978.

[2]

A Sensitivity Controller For Uncertain Systems

Kenji Okada and Robert E. Skelton

Abstract

In this paper a new controller design, which we shall call the "Trajectory Sensitivity Optimization" method is presented to improve the robustness for parameter variations. The method uses the sensitivity trajectory to model the parameter uncertainty and introduces a special quadratic cost function involving an input and output sensitivity term.

Necessary conditions are derived to obtain the dynamic controller. The necessary conditions consist of two Lyapunov equations and two controller gain equations which have no closed form solution. Therefore a special iterative algorithm was developed to obtain the numerical solution.

The method can deal with a wider class of parameter uncertainty than existing methods. Numerical examples show that the method is effective in improving the robustness to parameter variations.

1. Introduction

The LQG theory is well established as a multivariable control design synthesis, but it suffers from a poor sensitivity to certain classes of plant parameter uncertainty [1]. This sensitivity problem for parameter uncertainty becomes extremely important in flexible structure control where there is large parameter uncertainty. To cope with the problem various design syntheses have been proposed.

We are motivated by the trajectory sensitivity approach of Yedavalli and Skelton [13] where the necessary conditions are described to solve our problem. By restricting our attention to controllers of order n (equal to plant order) we will be able to make further progress toward solutions.

The sensitivity controller proposed by Wagie and Skelton [2] uses a trajectory sensitivity model to include the effects of parameter uncertainty and a special cost function involving both an output and input sensitivity term. This paper shows how to reduce the sensitivity model to tractable order, while preserving the correlations between outputs and all their sensitivities. The main drawback of this sensitivity controller design method is that the method does not deal with parameter uncertainty in the measurement matrix.

The Maximum Entropy Method has been applied by Hyland and Bernstein to the flexible structure problems [3]. This method uses a stochastic modeling for the parameter uncertainty in order to improve the robustness for the parameter variation. The design synthesis provides a direct method to the design of robust, reduced order controllers in which robust controller design and controller order reduction are performed simultaneously. The necessary conditions obtained by this method consist of two modified Riccati equations and two modified Lyapunov equations coupled by stochastic effects. Two restrictions of the method relate to the structure of the parameter uncertainties permitted. The uncertain parameters must appear linearly in the plant, input and output matrices. It also requires that the control-and-measurement-dependent uncertain parameters are uncorrelated. Because of this requirement the method cannot be applied directly to the problems in which there exists parameter uncertainty that affects the control matrix and the measurement matrix simultaneously. The method also cannot deal with parameter uncertainties in the disturbance matrix and in the output matrix. This may cause the unnecessary degradation of the closed loop system performance.

The approach developed by Tahk and Speyer [4] is called asymptotic LQG design synthesis. This method uses the internal feedback loop to model the parameter variations and serves to improve the stability robustness and reduce the sensitivity to parameter variation. This approach is a generalization of the LQG/LTR technique introduced by Doyle and Stein [5]. The approach has difficulties when there exist parameter variations

in the input matrix B or in the measurement matrix M. In this case the method requires augmentation of the state space so that ΔB and ΔM are embedded in the state matrix of the augmented system. This augmentation of the state space eventually leads to the increase of controller order.

As explained so far, the existing robust controller design methods for parameter uncertainty have some restrictions on the structure of parameter uncertainty. Hence, the main purpose of this paper is to propose a new robust controller design synthesis which can deal with wider classes of parameter uncertainty. The proposed method uses trajectory sensitivity to model the parameter uncertainty and introduces the special cost function which includes the output and input sensitivity terms in addition to the nominal input and output cost. The controller parameters are determined such that the given cost function is minimized. Through this minimization procedure, the controller obtains a robustness property with respect to parameter variation. The fundamental idea of this method is the same as the Wagie, Skelton Method, although the approach to obtaining the controller is different.

This paper is organized as follows. Section 2 discusses the modeling of parameter uncertainty using a trajectory sensitivity model. Section 3 introduces the newly developed "Trajectory Sensitivity Optimization" method and provides the necessary conditions for the sensitivity reducing controller and the algorithm to obtain the solution. Section 4 deals with the numerical examples to demonstrate the effectiveness of the proposed method and provides performance comparisons with other design methods. Finally Section 5 contains conclusions.

2. Modeling of Parameter Uncertainty

2.1 Trajectory Sensitivity Model

In this section we derive the trajectory sensitivity model for a simply supported beam example which will be used in the numerical example in section 4.

Assume there are h uncertain parameters $p_1, p_2, \cdots, p_h$ and a space-state model of the system is given by

$$\dot{x} = A(p)x + B(p)u + D(p)w \tag{2.1}$$

$$y = C(p)x \tag{2.2}$$

$$z = M(p)x + v \tag{2.3}$$

where x, y, z, u, w, and v are respectively, state vector of dimension n, output vector of dimension k, measurement vector of dimension l, input vector of dimension m, zero-mean white noise of dimension d with intensity W, and zero-mean white noise of dimension l with intensity V, and p is given by $p = [p_1, p_2, \cdots, p_h]^T$

Then the resulting sensitivity system can be expressed as follows:

$$\frac{d}{dt}x_s = A_s x_s + B_s u_s + D_s w_s \tag{2.4}$$

$$y_s = C_s x_s \tag{2.5}$$

$$z_s = M_s x_s + v_s \tag{2.6}$$

where:

$$x_s = \begin{Bmatrix} x \\ x_p \end{Bmatrix}, \; y_s = \begin{Bmatrix} y \\ y_p \end{Bmatrix}, \; z_s = \begin{Bmatrix} z \\ z_p \end{Bmatrix}, \; u_s = \begin{Bmatrix} u \\ u_p \end{Bmatrix}, \quad w_s = w, \quad v_s = \begin{Bmatrix} v \\ 0 \end{Bmatrix}$$

$$A_s = \begin{bmatrix} A & 0 \\ A_p & \tilde{A} \end{bmatrix} \quad B_s = \begin{bmatrix} B & 0 \\ B_p & \tilde{B} \end{bmatrix} \quad C_s = \begin{bmatrix} C & 0 \\ C_p & \tilde{C} \end{bmatrix} \quad D_s = \begin{bmatrix} D \\ D_p \end{bmatrix} \quad M_s = \begin{bmatrix} M & 0 \\ M_p & \tilde{M} \end{bmatrix}$$

$[\tilde{\cdot}] \triangleq$ block diag $\{ [\cdot], \cdots, [\cdot] \}$

$$[\cdot]_p \triangleq \left[\frac{\partial [\cdot]^T}{\partial p_1} \cdots \frac{\partial [\cdot]^T}{\partial p_h} \right]^T$$

where the matrices A_s, B_s, C_s, D_s, and M_s are evaluated at $p = \bar{p}$ (nominal vector value of

p). The basic idea to improve the robustness for parameter variations is to use a cost function V_D given by

$$V_D = E_{\_} [y^T Q y + u^T R u + \sum_{i=1}^{h} (y_{pi}^T Q_i y_{pi} + u_{pi}^T R_i u_{pi})] \tag{2.7}$$

where $y^T Q y + u^T R u$ is the part of the cost function for the standard LQG design, and $\sum_{i=1}^{h}(y_{pi}^T Q_i y_{pi} + u_{pi}^T R_i u_{pi})$ are added sensitivity terms. We seek a controller which minimizes the cost function V_D. Then the sensitivity of the controller to the parameter uncertainty p_i is reduced in increasing the norm of weighting matrices Q_i and R_i.

The nominal LQG controller is obtained by setting Q_i, R_i to zero, and a controller which minimizes sensitivity only (neglecting nominal performance requirements) is obtained by setting Q, R to zero.

2.2 Trajectory Sensitivity Matrices for An Example

In order to construct the trajectory sensitivity matrices A_p, B_p, C_p, D_p, and M_p for the physical system, we deal with an Euler Bernoulli simply supported beam shown in Fig 1. This example, is used later for sensitivity reducing controller design. In this example we take the following three quantities as the uncertain parameters.

(1) ρ : Mass Density of Beam (per length)

(2) EI : Flexual Rigidity of Beam

(3) K_a : Actuator Gain

It is well known that the natural frequency ω_i and mode shape $\Psi_i(r)$ for i-th mode of a simply supported beam are given by

$$\omega_i = \sqrt{\frac{EI}{\rho}}(\frac{i\pi}{L})^2, \quad \Psi_i(r) = \sqrt{\frac{2}{\rho L}} \sin(\frac{i\pi}{L} r)$$

(L: Length of Beam)

If we assume a torquer at $r = r_c$, a linear displacement measurement sensor at $r = r_m$, and a linear displacement at $r = r_o$ for the output, then we obtain the following equations of motion:

$$\ddot{q}_i + 2\zeta_i\omega_i\dot{q}_i + \omega_i^2 q_i = b_i(u + w), \quad (i = 1, \ldots, n)$$

$$y = \mu(r_o, t) = \sum_{i=1}^{n} \Psi_i(r_o)\, q_i$$

$$z = \mu(r_m, t) + v = \sum_{i=1}^{n} \Psi_i(r_m)\, q_i + v$$

where: $b_i = \Phi_i(r_c) = \frac{\partial}{\partial r}\Psi_i(r)\,|_{r=r_c} = \frac{i\pi}{L}\sqrt{\frac{2}{\rho L}} \cos(i\pi/L) r_c$

If we choose the state variable x by $x = [q^T, \dot{q}^T]^T$, then the above equations can be transformed into the following state-space expression.

$$\dot{x} = A x + B u + D w, \quad y = C x, \quad z = M x + v$$

where

$$A = \begin{bmatrix} 0 & I \\ -\Omega^2 & -2\zeta\Omega \end{bmatrix} \quad \Omega = \text{diag}\,[\omega_1, \cdots, \omega_n] \quad \zeta\text{: scaler}$$

$$B = [0 \cdots 0\ \Phi_1(r_c) \cdots \Phi_n(r_c)]^T \triangleq [0\ \ B_u^T]^T$$

$$C = [\Psi_1(r_o) \cdots 0 \cdots 0\,] \triangleq [C_y\ \ 0]$$

$$D = B$$

$$M = [\Psi_1(r_m) \cdots \Psi_n(r_m)\ 0 \cdots 0\,] \triangleq [M_z\ \ 0]$$

Using the above plant model, we obtain the following trajectory sensitivity matrices A_{pi}, B_{pi}, C_{pi}, D_{pi}, and M_{pi} for the three uncertain physical parameters.

• Uncertain Parameter ρ : $p_1 = \dfrac{\rho}{(\rho)_{NOM}}$ $\bar{p}_1 = 1$

$$A_{p_1} = \begin{bmatrix} 0 & 0 \\ \Omega^2_{NOM} & \zeta\Omega_{NOM} \end{bmatrix}, \quad B_{p_1} = \begin{bmatrix} 0 \\ -\frac{1}{2}(B_u)_{NOM} \end{bmatrix}, \quad C_{p_1} = \begin{bmatrix} -\frac{1}{2}(C_y)_{NOM} & 0 \end{bmatrix}$$

$$D_{p_1} = B_{p_1}, \quad M_{p_1} = \begin{bmatrix} -\frac{1}{2}(M_y)_{NOM} & 0 \end{bmatrix}$$

• Uncertain Parameter EI : $p_2 = \dfrac{EI}{(EI)_{NOM}}$, $\bar{p}_2 = 1$

$$A_{p2} = \begin{bmatrix} 0 & 0 \\ -\Omega^2_{NOM} & -\zeta\Omega_{NOM} \end{bmatrix}$$

$$B_{p_2} = 0, \quad C_{p_2} = 0, \quad D_{p_2} = 0 \quad M_{p_2} = 0$$

• Uncertain Parameter K_a : $p_3 = \dfrac{K_a}{(K_a)_{NOM}}$, $\bar{p}_3 = 1$

$$A_{p_3} = 0 \quad B_{p_3} = B \quad C_{p_3} = 0 \quad D_{p_3} = 0 \quad M_{p_3} = 0$$

3. Trajectory Sensitivity Optimization Method

In this section we introduce our controller design synthesis which reduces the sensitivity to parameter variations of the plant, using the trajectory sensitivity derived in section 2. The basic idea of this method is similar to that of the Wagie and Skelton method [2], but the advantages of the new method are: 1.) The order of the controller (the number of states) obtained by this method is smaller than that of the Wagie and Skelton method. 2.)The method can deal with a wider class of parameter uncertainty than the Wagie and Skelton method (The Wagie and Skelton method cannot deal with the parameter uncertainty related to measurement matrix. i.e., M_p term)

The method has these advantages over the Maximum Entropy Method: It can deal with (a) parameter uncertainty in a nonlinear manner, (b) parameter uncertainty in the disturbance matrix D and the output matrix C, and (c) parameter uncertainty appeared both in control matrix B and measurement matrix M at the same time. The advantage of the Maximum Entropy Method is fewer equations to solve. Some discussion of the convergence of the method appears in [12].

3.1 Problem Statement

We consider the following problem.

For the system

- n-th order Plant:

$$\dot{x} = A(p)x + B(p)\,u + D(p)\,w \tag{3.1}$$

$$y = C(p)\,x \tag{3.2}$$

$$z = M(p)\,x + v \tag{3.3}$$

where,

$$E\{w(t)w(\tau)^T\} = W\delta(t-\tau) \quad E\{v(t)v(\tau)^T\} = V\delta(t-\tau)$$

$$E\{w(t)\} = 0 \quad E\{v(t)\} = 0$$

$p = (p_1, \cdots, p_h)$ Uncertain Parameters

The n-th order Controller is:

$$u = G\,x_c \tag{3.4}$$

$$\dot{x}_c = A^c x_c + B^c u + F\,(z - M^c x_c)$$

$$= (A^c + B^c G - F M^c)x_c + F\,z \tag{3.5}$$

where,

$$A^c = A(\bar{p}) \quad B^c = B(\bar{p}) \quad M^c = M(\bar{p})$$

$\bar{p}$: Nominal Value of p

$$\frac{\partial F}{\partial p_i} = 0 \quad \frac{\partial G}{\partial p_i} = 0 \quad (i=1,\ldots,h)$$

Find F and G such that the cost function V_D is minimized.

$$V_D = E_w\{y^TQy + u^TRu + \sum_{i=1}^{h}\beta_i\left[\left[\frac{\partial y}{\partial p_i}\right]^T Q\left[\frac{\partial y}{\partial p_i}\right] + \left[\frac{\partial u}{\partial p_i}\right]^T R\left[\frac{\partial u}{\partial p_i}\right]\right]\} \qquad (3.6)$$

If we set $\beta_i = \beta\sigma_i$, where $\sigma_i = |\Delta p_i|$ is the magnitude of the expected variations in p_i, then we need only to determine β, Q, and R as design parameters. The weight β is usually determined through tradeoff between the robustness to parameter variation and the nominal performances of input and output cost. the weights Q and R may be determined under nominal conditions ($\sigma_i = 0$) to satisfy $E_w y_i^2 \le \sigma_i^2$, for a specified σ_i, i=1, 2, ..., n_y while minimizing u^TRu. The algorithm for such weights is given in chapter 8 of [10].

3.2 Derivation of Necessary Conditions

Let

$$\tilde{x} = x - x_c \qquad (3.7)$$

then the equations for the closed loop system are given by

$$\dot{\tilde{x}} = (A-FM)\tilde{x} + \{(A+BG-FM) - (A^c + B^cG-FM^c)\}\, x_c + Dw - Fv \qquad (3.8)$$

$$\dot{x}_c = FM\tilde{x} + \{A^c + B^cG + F(M - M^c)\}\, x_c + Fv \qquad (3.9)$$

$$y = C\tilde{x} + Cx_c \qquad (3.10)$$

$$u = Gx_c \qquad (3.11)$$

The above equations are transformed into the following matrix forms:

$$\dot{x}_a = A_a x_a + D_a w_a \tag{3.12}$$

$$y = C_a x_a \tag{3.13}$$

$$u = G_a x_a \tag{3.14}$$

where :

$$x_a = \begin{Bmatrix} \tilde{x} \\ x_c \end{Bmatrix} \quad w_a = \begin{Bmatrix} w \\ v \end{Bmatrix}$$

$$A_a = \begin{bmatrix} A-FM & (A+BG-FM)-(A^c+B^cG-FM^c) \\ FM & A^c+B^cG+F(M-M^c) \end{bmatrix}$$

$$D_a = \begin{bmatrix} D & -F \\ 0 & F \end{bmatrix} \quad C_a = [C \; C] \quad G_a = [0 \; G]$$

Let

$$x_s = [\tilde{x}^T \; x_c^T \; \tilde{x}_{p1}^T \; x_{c_{p1}}, \ldots, \tilde{x}_{ph}^T \; x_{c_{ph}}^T]^T \tag{3.15}$$

$$y_s = [y^T \; u^T \; y_{p1}^T \; u_{p1}^T, \ldots, y_{ph}^T \; u_{ph}^T]^T \tag{3.16}$$

$$w_s = [w^T \; v^T]^T \tag{3.17}$$

then the closed loop sensitivity system is expressed by

$$\dot{x}_s = A_s x_s + D_s w_s \tag{3.18}$$

$$y_s = C_s x_s \tag{3.19}$$

where

$$A_s = \begin{bmatrix} A_0 & 0 & 0 & \cdots & 0 \\ A_1 & A_0 & 0 & \cdots & 0 \\ \vdots & 0 & A_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ A_h & 0 & 0 & \cdots & A_0 \end{bmatrix} \tag{3.20}$$

$$A_0 = A_s(\bar{p}) = \begin{bmatrix} A(\bar{p})-FM(\bar{p}) & 0 \\ FM(\bar{p}) & A(\bar{p})+B(\bar{p})G \end{bmatrix} \tag{3.21}$$

$$A_i = \frac{\partial(A_s)}{\partial p_i}\Big|_{p=\bar{p}} = \begin{bmatrix} A_{pi}(\bar{p})-FM_{pi}(\bar{p}) & A_{pi}(\bar{p})+B_{pi}(\bar{p})G-FM_{pi}(\bar{p}) \\ FM_{pi}(\bar{p}) & FM_{pi}(\bar{p}) \end{bmatrix} \tag{3.22}$$

$$(i=1, \ldots, h)$$

$$D_s = \begin{bmatrix} D & -F \\ 0 & F \\ D_{p1} & 0 \\ 0 & 0 \\ \vdots & \vdots \\ D_{ph} & 0 \\ 0 & 0 \end{bmatrix}, \quad C_s = \begin{bmatrix} C & C & 0 & 0 & \ldots & 0 & 0 \\ 0 & G & 0 & 0 & \ldots & 0 & 0 \\ C_{p1} & C_{p1} & C & C & \ldots & 0 & 0 \\ 0 & 0 & 0 & G & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ C_{ph} & C_{ph} & 0 & 0 & \ldots & C & C \\ 0 & 0 & 0 & 0 & \ldots & 0 & G \end{bmatrix} \tag{3.23),(3.24}$$

From the triangular structure of A_s in Eq.(3.20) note that the poles of the closed loop sensitivity system given by Eq.(3.18) are equal to those of the closed loop system without sensitivity states repeated (h+1) times. Therefore if the closed loop system is stable at $p = \bar{p}$, then the closed loop sensitivity system given by Eq. (3.18) is always stable.

The steady state covariance is defined by

$$X_s = E_{ss}\{x_s x_s^T\} = E_{ss}\begin{bmatrix} \tilde{x}\tilde{x}^T & \tilde{x}x_c^T & \tilde{x}\tilde{x}_{p1}^T & \tilde{x}x_{c_{p1}}^T & \dots & \tilde{x}\tilde{x}_{ph}^T & \tilde{x}_{ph}x_{c_{ph}}^T \\ x_c\tilde{x}^T & x_c x_c^T & x_c\tilde{x}_{p1}^T & x_c x_{c_{p1}}^T & \dots & x_c\tilde{x}_{ph}^T & x_c x_{c_{ph}}^T \\ & & & & & \vdots & \vdots \\ & & & & & \tilde{x}_{ph}\tilde{x}_{ph}^T & \tilde{x}_{ph}x_{c_{ph}}^T \\ S. & Y. & M. & . & & x_{c_{ph}}\tilde{x}_{ph}^T & x_{c_{ph}}x_{c_{ph}}^T \end{bmatrix} \tag{3.25}$$

then X_s is obtained as the solution of the following Lyapunov equation:

$$A_s X_s + X_s A_s^T + D_s W_s D_s^T = 0$$

where :

$$W_s = \begin{bmatrix} W & 0 \\ 0 & V \end{bmatrix} \tag{3.27}$$

using the above X_s, we can express the cost function V_D by

$$V_D = tr\,[X_s C_s^T Q_s C_s] \tag{3.28}$$

where : $Q_s = \text{block diag}(Q, R, \beta_1 Q, \beta_1 R, \ldots, \beta_h Q, \beta_h R)$ (3.29)

By augmenting the constraints (3.26) to the objective functions (3.28) by use of Lagrange multipliers, we introduce H given by

$$H = tr\,[X_s C_s^T Q_s C_s] + tr\,[K_s(A_s X_s + X_s A_s^T + D_s W_s D_s^T)] \tag{3.30}$$

Then the solution to the problem satisfies the following conditions:

$$\frac{\partial H}{\partial X_s} = 0 \quad \frac{\partial H}{\partial K_s} = 0 \quad \frac{\partial H}{\partial G} = 0 \quad \frac{\partial H}{\partial F} = 0 \tag{3.31}$$

Relying on standard matrix calculus, the following necessary conditions are derived:

$$\bullet \ \frac{\partial H}{\partial X_s} = 0: \quad K_s A_s + A_s^T K_s + C_s^T Q_s C_s = 0 \tag{3.32}$$

$$\bullet \ \frac{\partial H}{\partial K_s} = 0: \quad A_s X_s + X_s A_s^T + D_s W_s D_s^T = 0 \tag{3.33}$$

$$\bullet \ \frac{\partial H}{\partial G} = 0: \quad G = -R^{-1}\{B^T \sum_{k=1}^{h+1} K_{2k} X_{2k} + \sum_{k=1}^{h} B_{pk}^T (K_{2k+1} X_2)\}$$
$$\times [X_{22} + \sum_{k=1}^{h} \beta_k X_{2k+1,2k+1}]^{-1} \tag{3.34}$$

$$\bullet \ \frac{\partial H}{\partial F} = 0: \quad F = -(K_{11} - K_{12} - K_{12}^T + K_{22})^{-1} [\{\sum_{k=1}^{h+1} (K_{2k} - K_{2k-1}) X_{2k-1}\} M^T$$
$$+ \sum_{k=1}^{h} (K_{2k+2} - K_{2k+1})(X_1 + X_2) M_{pk}^T] V^{-1} \tag{3.35}$$

where $K_1, \ldots, K_{2h+2}$ and $X_1, \ldots, X_{2h+2}$ are defined by

$$K_s = [K_1^T \ K_2^T \ \cdots \ K_{2h+2}^T]^T \quad K_i : n \times 2n(h+1) \tag{3.36}$$

$$X_s = [X_1 \ X_2 \ \cdots \ X_{2h+2}] \quad X_i : 2n(h+1) \times n \tag{3.37}$$

X_{ij} and K_{ij} are the (i,j) block of X_s and K_s.

$$X_s = \begin{bmatrix} X_{1,1} & \cdots & X_{1,2h+2} \\ \vdots & \cdots & \vdots \\ X_{2h+2,1} & \cdots & X_{2h+2,2h+2} \end{bmatrix} \quad X_{ij} : n \times n \tag{3.38}$$

$$K_s = \begin{bmatrix} K_{1,1} & \cdots & K_{1,2h+2} \\ \vdots & \cdots & \vdots \\ K_{2h+2,1} & \cdots & K_{2h+2,2h+2} \end{bmatrix} \quad K_{ij} : n \times n \tag{3.39}$$

As we can see in the definition of A_s, A_s contains F and G. Therefore the equations Eq.(3.32) - Eq.(3.35) are coupled. However these equations can be solved numerically by using the iterative method presented later in this section. Before proceeding to the iterative method, we investigate the Lyapunov equations given by Eq.(3.32) and Eq.(3.33).

The size of the Lyapunov equations is $2n(1+h)\times 2n(1+h)$. Therefore if the number of uncertain parameters is big, the equations may be too large to solve directly. The structure of the equations, however, allows them to be solved in partioned forms.

• Partitioning of the Lyapunov equation for X_s

Let Z_{ij} represent the $2n\times 2n$ matrix

$$Z_{ij} \triangleq \begin{bmatrix} X_{2i-1,2j-1} & X_{2i-1,2j} \\ X_{2i,2j-1} & X_{2i,2j} \end{bmatrix} \tag{3.40}$$

and D_{ij} be the $2n\times 2n$ matrices given by

$$D_{11} \triangleq \begin{bmatrix} DWD^T+FVF^T & -FVF^T \\ -FVF^T & FVF^T \end{bmatrix} \tag{3.41}$$

$$D_{ij} \triangleq \begin{bmatrix} D_{p_{i-1}} W D_{p_{j-1}}^T & 0 \\ 0 & 0 \end{bmatrix} \tag{3.42}$$

$$(i \geq 1 \quad j \geq 2 \text{ and } j \geq i, \; D_{p_0} \triangleq D)$$

then Z_{ij} is the $2n\times 2n$ block of X_s in the ij position, and D_{ij} is the $2n\times 2n$ block of $D_s W_s D_s^T$ in the ij position. The partitioned matrices Z_{ij} $(j \geq i \geq 1)$ are obtained by solving the following equations:

$$(1,1)\text{ block:} \quad A_0 Z_{11} + Z_{11} A_0^T + D_{11} = 0 \tag{3.43a}$$

$$(1,j)\text{ block:} \quad A_0 Z_{1j} + Z_{1j} A_0^T + Z_{11} A_{j-1}^T + D_{1j} = 0 \tag{3.43b}$$

$$(j = 2,\ldots, h+1)$$

$$(i,j)\text{ block:} \quad A_0 Z_{ij} + Z_{ij} A_0^T + (A_{i-1} Z_{1j} + Z_{1j}^T A_{j-1}^T) + D_{ij} = 0$$

$$(2 \leq i \leq j \leq h+1)$$

Since the (1,1) block element given by Eq.(3.43a) is the standard Lyapunov equation, it may be solved directly for Z_{11}. Substitution of Z_{11} obtained in Eq.(3.43a) into the (1,j) block elements reduces Eq.(3.43b) to standard Lyapunov equations. In a similar way the (i,j) block ($2 \le i \le j \le h+1$) elements can be solved by substituting Z_{1i} and Z_{1j} obtained in the previous calculations.

Therefore the Lyapunov equation for X_s of order 2n(1+h) can be reduced to the (h+1)(h+2)/2 Lyapunov equations of order 2n. It can also be shown easily that each partitioned Lyapunov equation obtained above can be partitioned further into four sub-block elements of order n which can be solved separately in the sequence of (1,1), (1,2), (2,1), (2,2) sub-block. Hence the total number of Lyapunov equations of order n becomes 2(h+1)(h+2). However, from the symmetric property of X_s the diagonal block elements Z_{ii} (i=1,...,h+1) are reduced to three sub-block elements of order n instead of four sub-block elements. Therefore the total number of Lyapunov equations to be solved is (h+1)(2h+3).

• Partitioning of the Lyapunov equation for K_s

Let Y_{ij} represent the 2n×2n matrix

$$Y_{ij} \triangleq \begin{bmatrix} K_{2i-1,2j-1} & K_{2i-1,2j} \\ K_{2i,2j-1} & K_{2i,2j} \end{bmatrix} \tag{3.44}$$

and C_{ij} be the 2n×2n matrices given by

$$C_{11} \triangleq \begin{bmatrix} C^TQC + \sum_{k=1}^{h} C_{p_k}^T Q_k C_{p_k} & C^TQC + \sum_{k=1}^{h} C_{p_k}^T Q_k C_{p_k} \\ C^TQC + \sum_{k=1}^{h} C_{p_k}^T Q_k C_{p_k} & C^TQC + \sum_{k=1}^{h} C_{p_k}^T Q_k C_{p_k} + G^TRG \end{bmatrix} \tag{3.45}$$

$$C_{1j} \triangleq \begin{bmatrix} C_{p_{j-1}}^T Q_{j-1} C & C_{p_{j-1}}^T Q_{j-1} C \\ C_{p_{j-1}}^T Q_{j-1} C & C_{p_{j-1}}^T Q_{j-1} C \end{bmatrix} \quad (2 \le j \le h+1) \tag{3.46}$$

$$C_{ij} \triangleq \begin{bmatrix} C^TQ_{i-1}C & C^TQ_{i-1}C \\ C^TQ_{i-1}C & C^TQ_{i-1}C + G^TR_{i-1}G \end{bmatrix} \quad (2 \le i \le h+1) \tag{3.47}$$

$$C_{ij} \triangleq 0 \quad (2 \le i < j) \tag{3.48}$$

where $Q_R = \beta_k Q$, $R_k = \beta_k R$, then Y_{ij} is the 2n×2n block of K_s in the ij position, and C_{ij} is the 2n×2n block of $C_s^T Q_s C_s$ in the ij position. The partitioned matrices Y_{ij} $(j \ge i \ge 1)$ are obtained by solving the following equations:

$$(j,j)\text{ block}: \quad Y_{jj}A_0 + A_0^T Y_{jj} + C_{jj} = 0 \ , \ (1 < j \le h+1) \tag{3.49a}$$

$$(i,j)\text{ block}: \quad Y_{ij}A_0 + A_0^T Y_{ij} + C_{ij} = 0 \ , \ (1 < i < j) \tag{3.49b}$$

$$(1,j)\text{ block}: \quad Y_{1j}A_0 + A_0^T Y_{1j} + A_{j-1}^T Y_{jj} + C_{1j} = 0 \ , \ (2 < j) \tag{3.49c}$$

$$(1,1)\text{ block}: \quad Y_{11}A_0 + A_0^T Y_{11} + (Y_{12}A_1 + Y_{13}A_2 + \cdots + Y_{1,h+1}A_h) + (Y_{12}A_1 + Y_{13}A_2 + \cdots + Y_{1,h+1}A_h)^T + C_{11} = 0 \ . \tag{3.49d}$$

From Eq.(3.49) and $C_{ij} = 0$ $(2 \le i < j)$, we obtain

$$Y_{ij} = 0 \quad \text{for} \quad 1 < i < j \tag{3.50}$$

Therefore the structure of K_s is given by

$$K_s = \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} & \cdots & Y_{1,h+1} \\ & \beta_1 Y_0 & 0 & \cdots & 0 \\ & & \beta_2 Y_0 & \cdots & \vdots \\ & & & \cdots & 0 \\ S. & Y. & M. & & \beta_h Y_0 \end{bmatrix} \tag{3.51}$$

where Y_0 is the solution for the following Lyapunov equations:

$$Y_0 A_0 + A_0^T Y_0 + C_q = 0 \tag{3.52}$$

$$C_q = \begin{bmatrix} C^T Q C & C^T Q C \\ C^T Q C & C^T Q C + G^T R G \end{bmatrix} \tag{3.53}$$

The number of Lyapunov equations to be solved becomes h+2 Lyapunov equations of order 2n or 4h+6 Lyapunov equations of order n. When we choose $\beta_1 = \beta_2 = \cdots = \beta_h = 0$, only the (1,1) block K_s becomes non-zero. In this case it can be easily shown that the equations for G and F reduce to the standard LQG equations.

3.3 Algorithm to obtain the Numerical Solution

Since the equations obtained as the necessary conditions are coupled, a special numerical algorithm is required. The approach taken is similar to that of the Wagie and Skelton method. The algorithm is summarized as followed:

- Algorithm to obtain F and G

STEP 1 Choose initial F and G

$$F = F_0 \qquad G = G_0$$

(e.g. use the solution for the standard LQG problem)

STEP 2 For the given F and G, solve X_s and K_s from Eq.(3.32) and Eq.(3.33) using the partitioned Lyapunov equations.

STEP 3 Using X_s and K_s obtained in STEP 2, Calculate F and G. Set $F_{NEW} = F$ $G_{NEW} = G$

STEP 4 If $| G_{NEW} - G_{OLD} | < \varepsilon_g$ and $| F_{NEW} - F_{OLD} | < \varepsilon_f$
then Stop : Solution Completed. Otherwise set
$F = F_{NEW}\alpha + F_{OLD}(1 - \alpha)$,
$G = G_{NEW}\alpha + G_{OLD}(1 - \alpha)$,
Return to STEP 2.

where α is the coefficient which dictates the convergence of the solution. Usually 0.2 - 0.5 is used as the value of α. But when the weight β_i to the sensitivity part of the cost is large, a smaller value may be necessary.

4. Numerical Example and PerformanceComparison with other Controller Design Methods

In order to investigate the effectiveness of the proposed controller design synthesis and compare the performances with other methods, we take the following three examples.

4.1 Simply Supported Beam

We have already derived the sensitivity trajectory model for a simply supported beam in section 2. Here we consider the same example. As numerical values we use the following values:

- Beam Parameter: $L = \pi \quad \rho = \frac{2}{L} \quad EI = \rho$

- Sensor, Actuator, Output Position : $r_m = 0.30\,L \quad r_c = 0 \quad r_0 = 0.45\,L$

- Noises : $V = 1 \quad W = 1$

Then ω_i, $\Psi_i(r)$, $\Phi_i(r)$ are given by

$$\omega_i = i^2 \quad \Psi_i(r) = \sin\left(i\frac{\pi}{L}r\right) \quad \Phi_i(r) = i\cos\left(i\frac{\pi}{L}r\right)$$

$$(i = 1, \cdots, n)$$

If we choose the first 4 modes as our design and evaluation model, then we obtain

$$\Omega = \mathrm{diag}\,(1^2, 2^2, 3^2, 4^2) \qquad B_u = [1\ 2\ 3\ 4\ ]^T$$

$$C_y = [\sin(0.45\,\pi)\ \sin(2\times0.45\,\pi)\ \ \sin(3\times0.45\pi)\ \ \sin(4\times0.45\,\pi)]$$

$$M_z = [\sin(0.30\pi)\ \sin(2\times0.30\pi)\ \ \sin(3\times0.30\pi)\ \ \sin(4\times0.30\pi)]$$

Substituting these matrices into A, B, C, D, M and A_{p_i}, B_{p_i}, C_{p_i}, D_{p_i}, M_{p_i}, we obtain the required data for the Trajectory Sensitivity Optimization (TSO) method design synthesis.

If we choose Q = 1, R = 1 for the weights of the cost function and apply the standard LQG method, then we obtain the controller whose stability range for parameter variation and input & output cost are summarized in Table 1. As we can see from Table 1, the standard LQG controller is sensitive to the parameter variation ρ and EI.

We apply the TSO to the same system to reduce the sensitivity. First we investigate one uncertain parameter case in which only one uncertain parameter is considered for the design of sensitivity reducing controller. Next we deal with two uncertain parameter case in which two uncertain parameters are considered at the same time for the controller design.

- Sensivity Reducing Controllers for One Uncertain Parameter Case

  Controller Type A

  In this case the sensitivity part with respect to mass density variation is weighted for the TSO cost. Table 2 shows the stability range for parameter variation and input & output cost of Type A controllers for different weights β. When we compare Table 1 and Table 2, we see that the stability range for ρ variation increases as the weight β increases. In this case the sensitivity to EI variation is also reduced.

  The gain margin, however, decreases as the weight β increases. Fig 2 shows output cost performance change due to ρ variation for different controllers. As we see in Fig. 2 and Table 2, the output cost increase rate is maintained relatively small while the input cost increases pretty rapidly as β increases. Therefore tradeoff between the robustness to parameter variation and the input & output cost should be made to determine the appropriate weight β.

- Sensivity Reducing Controllers for Two Uncertain Parameter Cased

  Controller Type B

  In this case two sensitivity terms (i.e, sensitivity terms with respect to ρ and K_a) are weighted at the same time for the TSO cost. Table 3 shows the performance of Type B controllers Comparing the results with the standard LQG controller performance, we notice that controller B-2 can achieve better robustness for parameter variations ρ, EI and K_a also.

4.2 Cart with an Inverted Pendulum

Next we consider the cart with an inverted pendulum shown in Fig 3. The linearized equations expressed in state-variables form are given by

$$\frac{d}{dt}\begin{Bmatrix} z \\ \dot{z} \\ \theta \\ \dot{\theta} \end{Bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{m}{M}g & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{M+m}{Ml}g & 0 \end{bmatrix} \begin{Bmatrix} z \\ \dot{z} \\ \theta \\ \dot{\theta} \end{Bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{M} \\ 0 \\ -\frac{1}{Ml} \end{bmatrix} u$$

or

$$\dot{x} = Ax + B(u+w)$$

$$y = z = [\,1\,0\,0\,0\,]\,x = C\,x$$

$$z_m : z_m = \begin{Bmatrix} z \\ \theta \end{Bmatrix} + \begin{Bmatrix} v_1 \\ v_2 \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x + v = M\,x + v$$

We consider the following two uncertain parameters:

a) Actuator Gain (K_a) :

We obtain the following sensitivity data.

Trajectory Sensitivity Data : $p = \dfrac{K_a}{(K_a)_{NOM}}$

$A_p = 0 \quad B_p = B \quad C_p = 0 \quad D_p = 0 \quad M_p = 0$

Maximum Entropy Design Data :

$A_1 = 0 \quad B_1 = 0 \quad M_1 = 0$

b) Angle Sensor Gain (K_s) :

We obtain the following sensitivity data.

Trajectory Sensitivity Data : $p = \dfrac{K_s}{(K_s)_{NOM}}$

$$A_p = 0 \quad B_p = 0 \quad C_p = 0 \quad D_p = 0 \quad M_p = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Maximum Entropy Design Data :

$$A_1 = 0 \quad B_1 = 0 \quad M_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

- Numerical Values: $M = 1\,kg \quad m = 0.1\,kg \quad L = 1.0\,m \quad g = 9.8 m/sec^2$

$$W = 1.0\times10^{-6} \quad V = 1.0\times10^{-6}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad R = 1 \quad Q = 1$$

- Sensitivity Reducing Controllers for K_a Variation

Three different sensitivity reducing controller design synthesis, 1) TSO 2) Wagie and Skelton method 3) Maximum Entropy method, were applied to this problem to compare the performance. The performance curves (input & output cost versus stability range of K_a variation) for different controller designs are shown in Fig 4. As we can see in the figure, the Wagie and Skelton method cannot improve the robustness for K_a variation even if a large β is chosen. The TSO achieves smaller input & output cost than the Maximum Entropy method for the same stability margin. Therefore the TSO is best for this problem in terms of performance cost and robustness to parameter variation.

- Sensitivity Reducing Controllers for K_s Variation

The Wagie and Skelton method cannot deal with the problem in which measurement sensitivity matrix M_p is non-zero. Therefore the TSO and the Maximum Entropy methods were applied to this problem. The results are shown in Fig 5. In this case the TSO achieves smaller output cost than the Maximum Entropy method for the same stability margin, whereas it requires larger input cost than the Maximum Entropy method. Fig 5 suggests that in this case there is no big difference between two methods with respect to performance cost and stability margin for K_a variation.

4.3 Doyle's Example

The problem considered here was first given by Doyle [1] and investigated further by Bernstein in his Maximum Entropy method [7]. The required data for the problem are given by Plant Matrices :

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ b \end{bmatrix} \quad C = [1\ 1] \quad D = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad M = [1\ 0]$$

Uncertain parameter : b in matrix B $(b)_{NOM} = 1$

Like the previous example, three different controller design syntheses were applied to this problem. The performance curves for different controller designs are shown in Fig 6. In this case the controllers obtained by the Wagie and Skelton method show worse robustness for parameter variation b than the standard LQG controller. Therefore the method is unacceptable for this problem. The TSO and the Maximum Entropy method (For the Maximum Entropy method, we used the results presented in [7]) show the similar results to those obtained in the previous example (Uncertain Parameter : K_s Fig 5).

5. Conclusions

In this paper a new controller design synthesis is presented to improve robustness to parameter uncertainty. The proposed method uses the trajectory sensitivity to model the parameter uncertainty and introduces a special cost function to reduce the parameter sensitivity at both the input and output to the plant. The order of the controller is equal to that of the nominal plant. The necessary conditions for the optimization consist of two Lyapunov equations and two gain matrix equations. An iterative algorithm was developed to obtain the solution to these coupled equations. The large sizer of the Lyapunov equations of order 2n(h+1)

is reduced to several smaller equations, using partitioned forms improving numerical efficiency.

The new method can deal with a wider class of parameter uncertainty than either the Wagie and Skelton method or the Maximum Entropy method. The new method can deal with parameters appearing nonlinearly in any place; the plant matrix, the input matrix, the disturbance matrix, the output matrix, and the measurement matrix.

Numerical examples show that the method is effective in improving robustness to parameter variations. For examples with a simple parameter structure other methods can be applied (Maximum Entropy and Wagie, Skelton). In some of these examples the new method performed as well as these existing methods. In other examples the new method performed better. The disadvantage of the method is the lack of a closed form solution. Iterative algorithms are required whose convergence remains an open question for future research.

REFERENCES

[1] J.C. Doyle, "Guaranteed Margins for LQG Regulators," IEEE Trans. Automat. Contr. Vol. AC-23, pp.756-757, 1978

[2] D.A. Wagie and R.E. Skelton, "Model Reduction and Controller Synthesis in the Presence of Parameter Uncertainty," Automatica, 1986, vol.22, No.3 pp.295-308.

[3] D.S. Bernstein and D.C. Hyland, "The Optimal Projection/Maximum Entropy Approach to Designing Low-Order, Robust Controllers for Flexible Structures," Proc. 24th IEEE CDC. Dec. 1985

[4] M. Tahk and J. Speyer, "Modeling of Parameter Variations and Asymptotic LQG synthesis," IEEE Trans. Automat. Contr. Vol. AC-32, no.9 pp. 793-801, Sept. 1987

[5] J.C. Doyle and G. Stein, "Robustness with observers," IEEE Trans. Automat. Contr. Vol. AC-24, Aug. 1979.

[6] R.E. Skelton and D.H. Owens, editors, "Model Error Concepts and Compensation," Pergamon Press, New York, 1986

[7] D.S. Bernstein and S.W. Greeley, "Robust Controller Synthesis Using the Maximum Entropy Design Equations," IEEE Trans. Automat. Contr. Vol. AC-31, No.4, Apr. 1986

[8] D.S. Bernstein, L.D. Davis, S.W. Greeley and D.C. Hyland, "Numerical Solution of the Optimal Projection/Maximum Entropy Design Equations for Low-Order, Robust Controller Design," Proc. 24th IEEE CDC. Dec. 1985.

[9] H.Kwakernaak and R. Sivan, Linear Optimal Control Systems, Wiley, New York, 1972

[10] R.E. Skelton, Dynamic Systems Contro, John Wiley & Sons, New York, Feb. 1988

[11] T.E. Fortmann and K.L. Hitz, An Introduction Linear Control Systems, Marcel Dekker, New York, 1977.

[12] G.N. Mil'stein "Design of Stabilizing Controller with Incomplete State Data for Linear Stochastic Systems with Multiplicative Noise", Auto. Remote Control, vol. 43, 1982, pp. 653-659

[13] R.K. Yedavalli and R.E. Skelton, "Controller Design for Parameter Sensitivity Reduction in Linear Regulators", Optimal Control Applications & Methods, vol. 3, 1982, pp. 221-240.

Table 1 Performances of the Standard LQG Controller

Standard LQG Controller		Performances
Gain Margin	K_{max}	4.087
Stability Range of ρ	P_{max}	1.173
$P = \rho/(\rho)_{NOM}$	P_{min}	0.677
Stability Range of EI	P_{max}	3.226
$P = EI/(EI)_{NOM}$	P_{min}	0.863
Input & Output Cost	V_y	2.760
	V_u	0.488

Table 2 Performances of Controllers Type A

Sensitivity Reducing Controller for ρ Type A		Performances		
		Type A-1 $\beta = 0.01$	Type A-2 $\beta = 0.1$	Type A-3 $\beta = 0.5$
Gain Margin	K_{max}	3.247	2.300	1.706
Stability Range ρ	P_{max}	1.415	1.687	1.802
$P = \rho/(\rho)_{NOM}$	P_{min}	0.681	0.596	0.435
Stability Range of EI	P_{max}	2.774	2.704	2.734
$P = EI/(EI)_{NOM}$	P_{min}	0.733	0.631	0.607
Input & Output Cost	V_y	2.856	3.147	3.821
	V_u	1.300	3.709	11.243

Table 3 Performances of Controllers Type B

Sensitivity Reducing Controllers for ρ and K_a Type B		Performances		
		Type A-2 $\beta_1 = 0.1$ $\beta_2 = 0$	Type B-1 $\beta_1 = 0.1$ $\beta_2 = 1$	Type B-2 $\beta_1 = 0.1$ $\beta_2 = 10$
Gain Margin	K_{max}	2.300	3.054	4.843
Stability Range ρ	P_{max}	1.687	1.567	1.360
$P = \rho/(\rho)_{NOM}$	P_{min}	0.596	0.661	0.687
Stability Range of EI	P_{max}	2.704	1.618	1.526
$P = EI/(EI)_{NOM}$	P_{min}	0.631	0.669	0.755
Input & Output Cost	V_y	3.147	2.913	2.723
	V_u	3.709	3.692	3.865

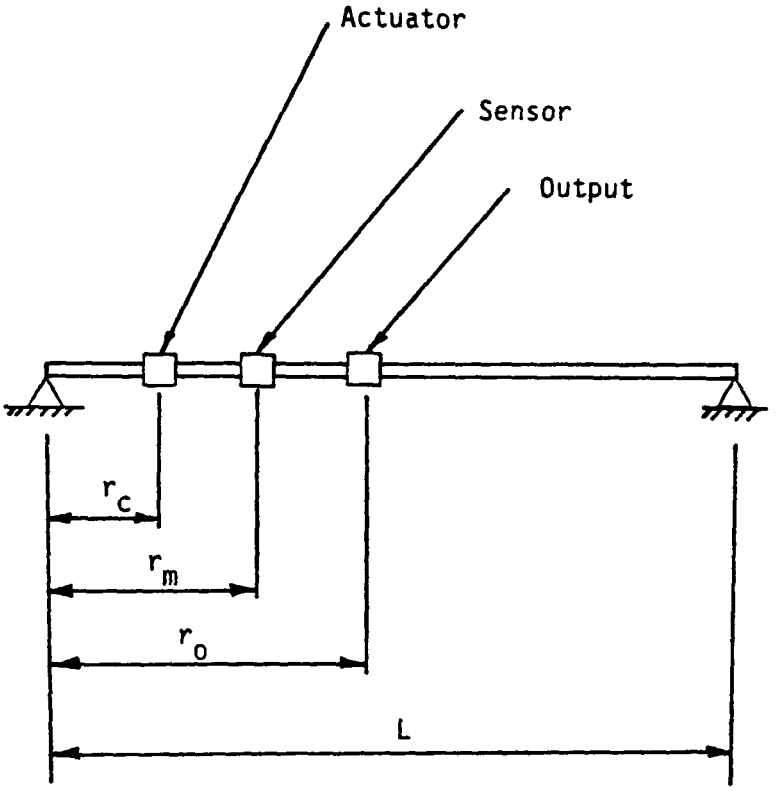


Figure 1 Simply Supported Beam

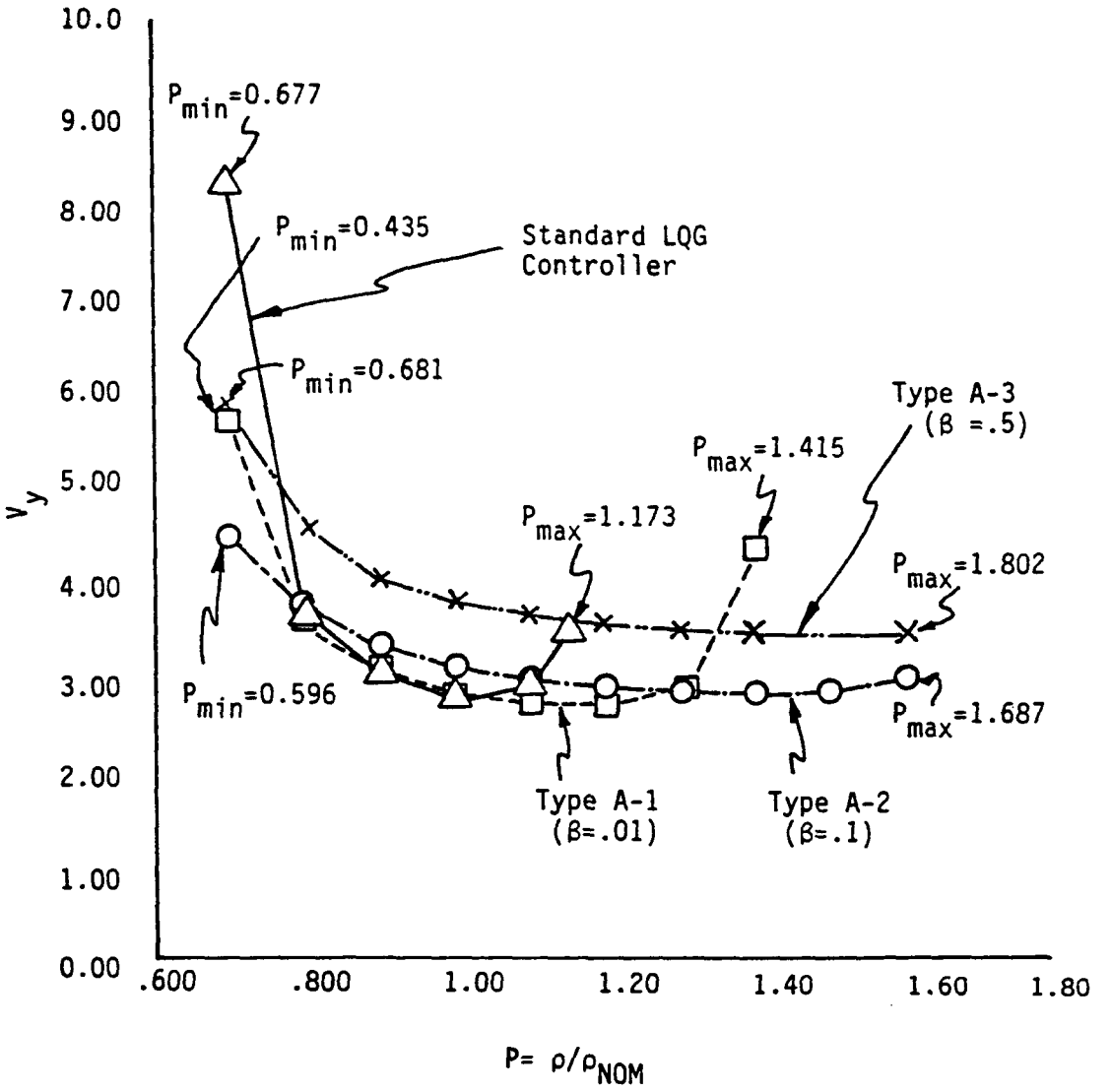


Figure 2 Output Cost Performance Change due to ρ Variation for Different Controllers

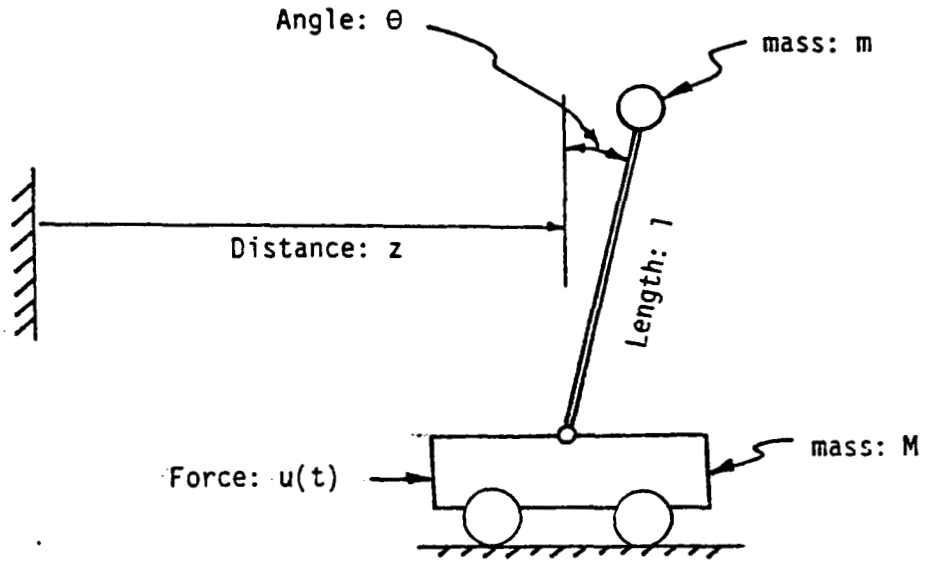


Figure 3 Cart with an Inverted Pendullum

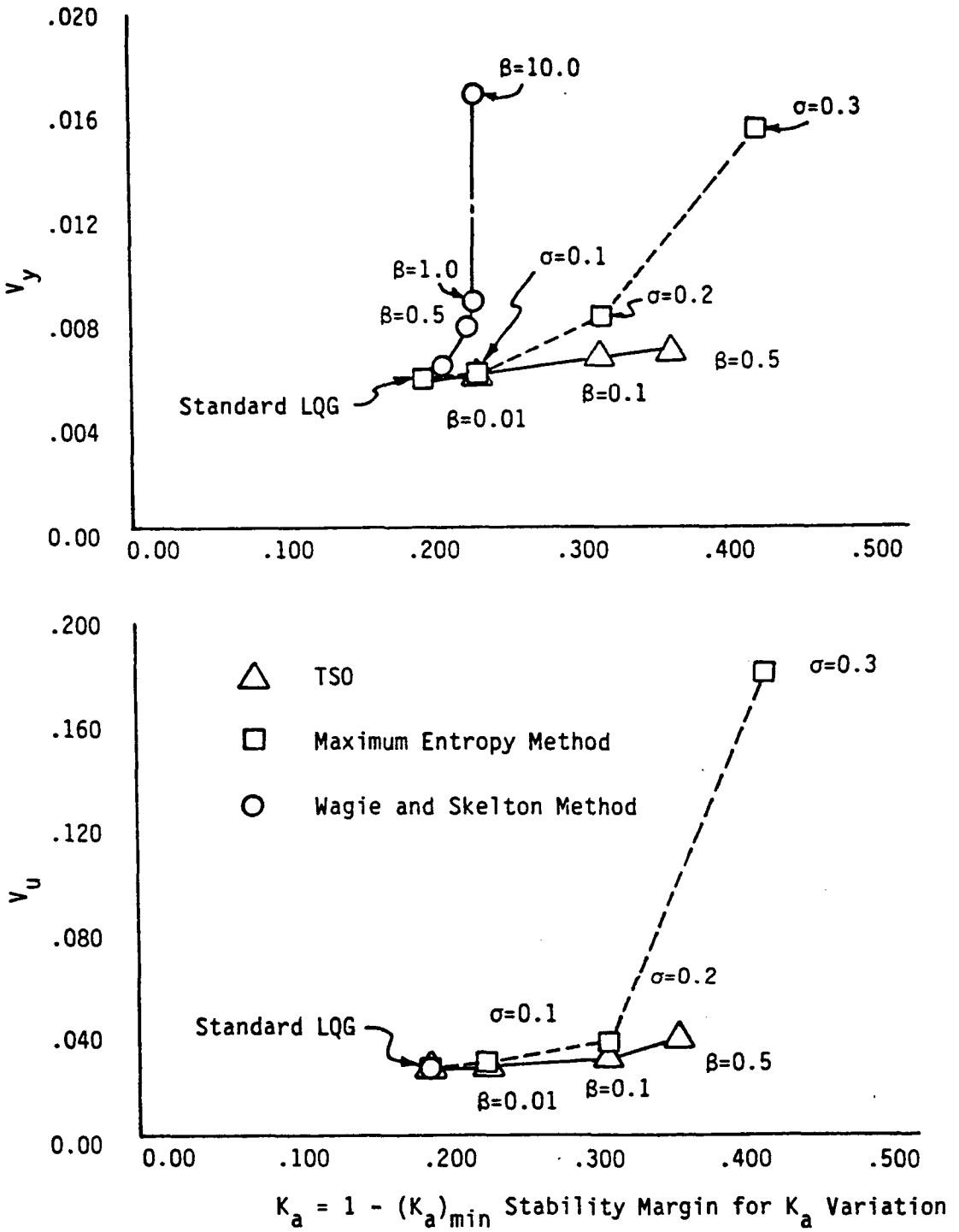


Figure 4 Performance Comparison of Different Controller Designs (Uncertain Parameter: K_a)

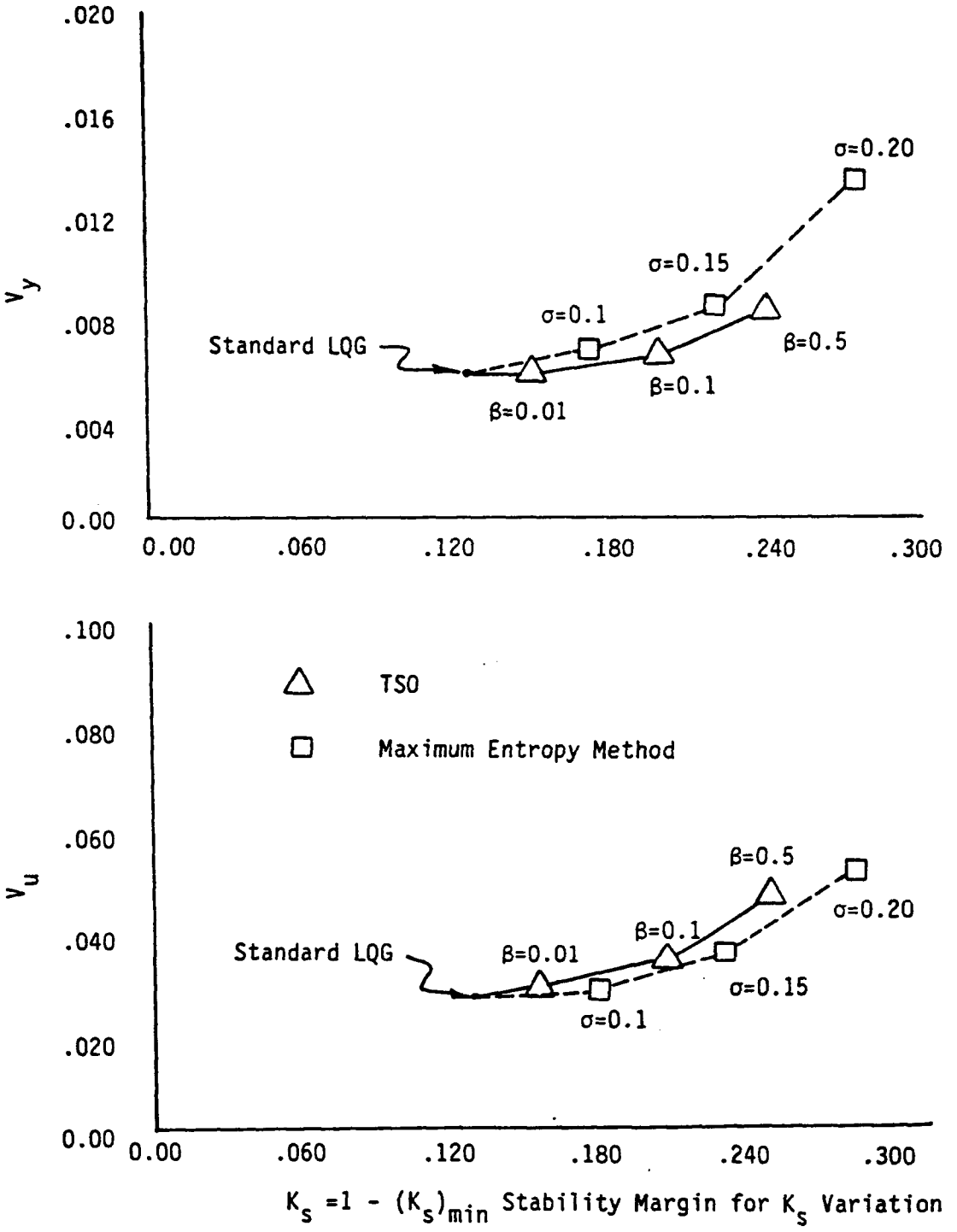


Figure 5 Performance Comparison of Different Controller Designs (Uncertain Parameter: K_s)

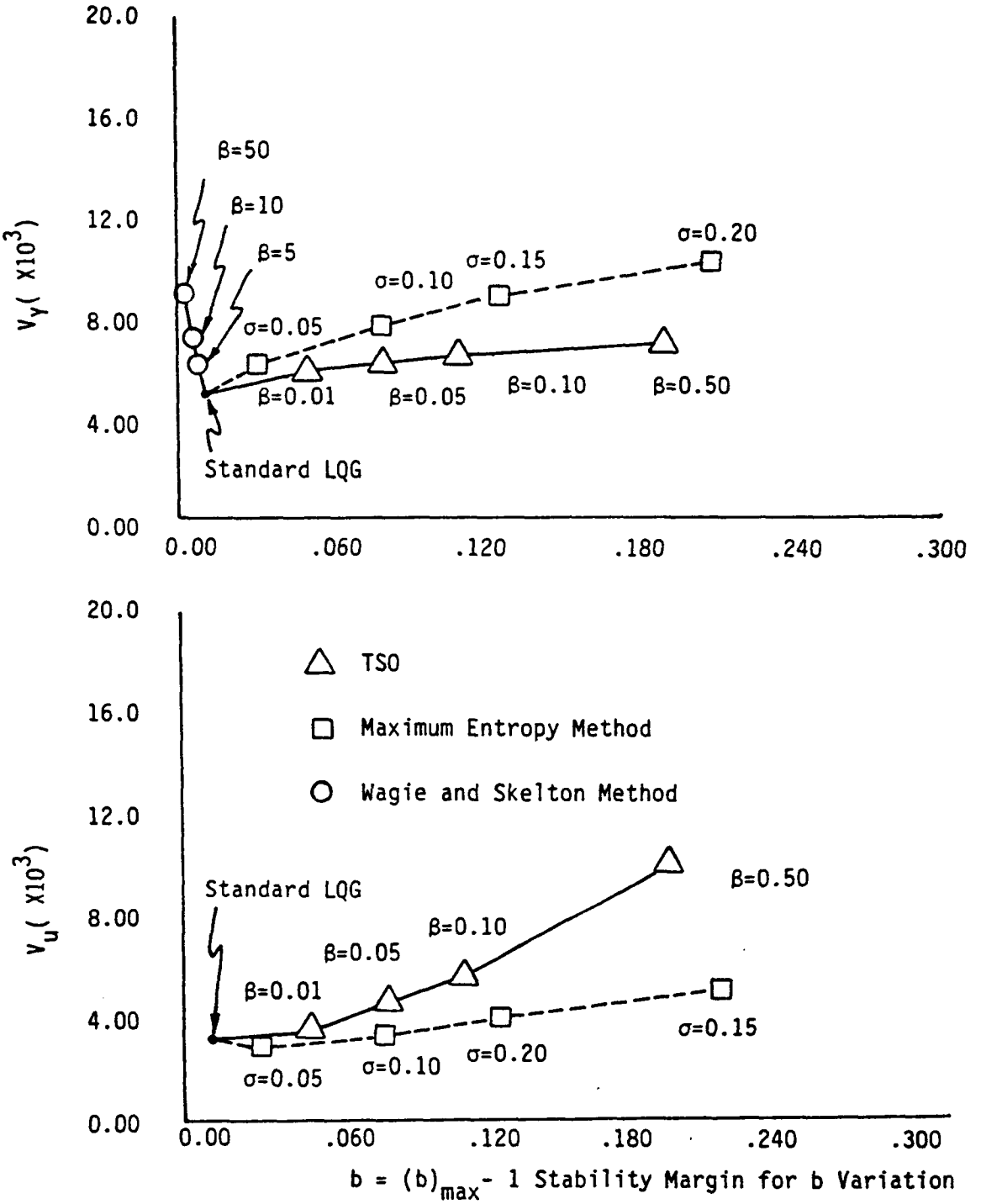


Figure 6 Performance Comparison of Different Controller Designs (Doyle's Example)

Optimal q-Markov Cover for Finite Precision Implementation

Darrell Williamson* and Robert E. Skelton**

Abstract

The existing q-Markov COVER realization theory does not take into account the problems of arithmetic errors due to both the quantization of states and coefficients of the reduced order model. All q-Markov COVERs allow some freedom in the choice of parameters. In this paper we exploit this freedom in the existing theory to optimize the models with respect to these finite wordlength effects.

*Dept. of Systems Engineering, Research School of Physical Sciences, Australian National University Canberra, ACT 2601, Australia

**School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN 47907, U.S.A.

Introduction

An asymptotically stable system can be characterized in terms of its impulse response sequence (Markov parameters) and its output covariance sequence (covariance parameters) due to a zero mean white noise input process. A general approach has been developed [3] for realizing a system which matches q Markov parameters and q covariance parameters. Such a system is referred to as a q-Markov COVER, and q-Markov COVERs may be generated from output data [3,4] or from higher order models [5,6]. The Markov and covariance parameters are not independent and consequently the q-Markov COVER is not unique. In particular, all q-Markov COVERs are not related by state space similarity transformations [4]. In this paper we shall exploit the remaining degrees of freedom to optimize the q-Markov COVER realization with respect to an aspect of its finite wordlength realization.

Specifically, when digital controllers are to be implemented, both the controller coefficients and the controller states must be represented in finite wordlength precision. This finite wordlength (FWL) representation (or quantization) causes inaccuracies in the response when compared to the ideal (i.e. infinite precision) behaviour. Effects of quantization on the controller are increased noise at the output due to internal state quantization, and errors in time and frequency response characteristics due to coefficient errors.

In digital filter design, the FWL effects are known to be most significant when the poles of the filter are very close to the unit circle [12]. In particular, narrow band filters have all these poles near $z = 1 \pm j\omega$. For digital control, the zero-order-hold equivalent of a continuous time model (or controller) with a pole at λ will have a discrete pole at $\exp(\lambda T)$. Hence for fast sampling and/or low damping of the continuous models, the discrete model will behave like a narrow band filter. The synthesis of optimal digital controllers with respect to arithemetic quantization noise is an important consideration in design especially for continuous time systems operating under a fast sampling rate [9,10]. The effects of quantization depend highly on the structure of the controller. This paper seeks to reduce these errors in the synthesis of q-Markov COVERs.

1. Discrete q-Markov COVER

Consider the asymptotically stable nominal discrete system

$$x(k+1) = Ax(k) + Bu(k)\ ;\ \ x(k)\epsilon R^{n_x},\ u(k)\epsilon R^{n_u}$$
$$y(k) = Cx(k)\qquad\ ;\ \ y(k)\epsilon R^{n_y} \tag{1.1}$$

where $\{u(k)\}$ is a zero mean process with unit intensity $E\{u(k)u^*(j)\} = I\delta_{ij}$ and $E\{x(k)u^*(j)\} = 0$ for $k \geq j$. The Markov parameters M_i and covariance parameters R_j of (1.1) are defined by

$$M_i \triangleq CA^iB\ ;\ \ R_j = CA^jXC^*,\ \ j \geq 0,\ R_j = CXA^{*j}C^*,\ \ j \leq 0 \tag{1.2}$$

where the state covariance matrix X satisfies the Lyapunov Equation

$$X = AXA^* + BB^*\ . \tag{1.3}$$

These parameters M_i and R_j appear as coefficients in the expansion of the transfer function $H(z)$ and power spectral density $H(z)H^*(z^{-1})$; that is

$$H(z) = C(zI-A)^{-1}B = \sum_{i=0}^{\infty} M_i z^{-(i+1)};\ \ H(z)H^*(z^{-1}) = \sum_{j=-\infty}^{\infty} R_j z^{-j}$$

We suppose that as data we are given the first q-Markov and first q-covariance parameters $\{M_i, R_i;\ i = 0, 1, ..., q-1\}$ of an asymptotically stable system from which we construct the two data matrices

$$D_q \triangleq R_q - M_q M_q^* \ \epsilon\ R^{n_y q \times n_u q}$$
$$\overline{D}_q \triangleq R_q - \overline{M}_q \overline{M}_q^* \ \epsilon\ R^{n_y q \times n_u q} \tag{1.4a}$$

where R_q, M_q and $\overline{M}_q$ are the Toeplitz matrices of the data as defined by

$$R_q = \begin{bmatrix} R_0 & R_1^* & \dots & R_{q-1}^* \\ R_1 & R_0 & \dots & R_{q-2}^* \\ \vdots & \vdots & & \vdots \\ R_{q-2} & \dots & \dots & \\ R_{q-1} & R_{q-2} & \dots & R_0 \end{bmatrix} \tag{1.4b}$$

$$M_q = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ M_0 & 0 & \dots & 0 & 0 \\ M_1 & M_0 & \dots & . & . \\ \vdots & \vdots & & \vdots & \dots \\ M_{q-2} & M_{q-3} & \dots & M_o & 0 \end{bmatrix},\ \overline{M}_q = \begin{bmatrix} M_0 & 0 & \dots & 0 \\ M_1 & M_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ M_{q-2} & & & 0 \\ M_{q-1} & M_{q-2} & \dots & M_0 \end{bmatrix}$$

The first data matrix D_q in (1.4a) is Hermitian and it is shown in [3-4] to be

positive semidefinite. Hence we can obtain a (nonunique) full rank factorization

$$D_q = P_q P_q^*; \; P_q \varepsilon R^{n_y q x r_q}, \tag{1.5a}$$

where

$$r_q \triangleq \text{rank}\,(D_q) = \text{rank}(P_q) \le n_y q \tag{1.5b}$$

If we partition P_q according to

$$P_q^* = [E_q^* \; F_q^*]; \; E_q \varepsilon R^{n_y x r_q}, \; F_q \varepsilon R^{(q-1)n_y x r_q} \tag{1.6}$$

then it follows that the second data matrix $\bar{D}_q$ can be factored as

$$\bar{D}_q = \bar{P}_q \bar{P}_q^*; \; \bar{P}_q \varepsilon R^{n_y q x r_q} \tag{1.7}$$

where

$$\bar{P}_q^* = [F_q^* \; G_q^*]; \; G_q \varepsilon R^{n_y x r_q} \tag{1.8}$$

for some G_q (to be determined). The following result has been established.

Theorem 1.1 [3]

Given the q Markov parameters $\{M_i; i = 0,1, ..., q-1\}$ and the q covariance parameters $\{R_i; i = 0,1, ..., q-1\}$ and a matrix G in (1.8) such that (1.7) is satisfied, then the realization $\{A_q, B_q, C_q\}$ of order r_q defined by

$$A_q = P_q^+ \bar{P}_q; \; B_q = P_q^+ [M_0^* \cdots M_{q-1}^*]^*; \; C_q = E_q \tag{1.9}$$

where P_q^+ denotes the Moore-Penrose inverse of P is a q-Markov COVER. The corresponding controllability grammian X_q is given by

$$X_q = I \tag{1.10}$$

Furthermore

$$P_q = [C_q^* \; A_q^* C_q^* \cdots (A_q^{q-1})^* C_q^*]^* \tag{1.11}$$

□□□

This theorem describes a large but *not* complete class C_q of q-Markov COVERs parameterized by $\{G_q\}$ such that for some E_q,F_q the data matrices D_q, $\bar{D}_q$ satisfy (1.5)-(1.8). Each matrix G_q will (generally) result in a q-Markov COVER having a different transfer function. In order to compute the set of all such G_q, observe in (1.5)-(1.8) that

$$D_q = \begin{bmatrix} E_q \\ F_q \end{bmatrix} (E_q^* \; F_q^*) . \tag{1.12a}$$

Then

$$\bar{D}_q = \begin{bmatrix} \bar{D}_{q-1} & \bar{d}_q \\ \bar{d}_q^* & \bar{d}_{qq} \end{bmatrix} = \begin{bmatrix} F_q \\ G_q \end{bmatrix} [F_q^* \; G_q^*] \tag{1.12b}$$

$$\bar{d}_{qq} \in R^{n_y x n_y}$$

implies

$$E_q E_q^* = R_o , \;\; F_q F_q^* = \bar{D}_{q-1}, \;\; F_q G_q^* = \bar{d}_q, \;\; G_q G_q^* = \bar{d}_{qq} \tag{1.13}$$

Now expand D_q in terms of its singular value decomposition

$$D_q = (U_1 \; U_2) \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix} \; ; \;\; \Sigma_1 \in R^{r_q x r_q} . \tag{1.14}$$

Then from (1.12a)

$$(E_q^* \; F_q^*) = \Sigma_1^{1/2} U_1^* \tag{1.15}$$

so that $E_q = C_q$ is defined by the first n_y rows and F_q by the last $(q-r)n_y$ rows of $U_1 \Sigma_1^{1/2}$. Define

$$\rho_q \triangleq \text{rank } (F_q) . \tag{1.16a}$$

Then from (1.15)

$$\rho_q \leq \min (r_q, (q-1)n_y) . \tag{1.16b}$$

Next, expand F_q in (1.13) in terms of its singular value decomposition. If strict inequality occurs in (1.16b) we have

$$F_q = [U_\alpha \; U_\beta] \begin{bmatrix} \Sigma_q & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_\alpha^* \\ V_\beta^* \end{bmatrix} \; ; \;\; \Sigma_q \in R^{\rho_q x \rho_q} \tag{1.17}$$

The Moore-Penrose inverse F_q^+ of F_q is then given by

$$F_q^+ = V_\alpha \Sigma_q^{-1} U_\alpha^* \tag{1.18}$$

Corollary 1.1

Define

$$\text{(i)}\ G_{q1} \triangleq (F_q^+ \bar{d}_q)^* \in R^{n_y x r_q} \tag{1.19}$$

$$\text{(ii)}\ G_{q2} \in R^{n_y x s_q} \text{ such that } G_{q2}G_{q2}^* = \bar{d}_{qq} - \bar{d}_q^* \bar{D}_{q-1}^+ \bar{d}_q$$

where

$$s_q \triangleq \text{rank}\,[\bar{d}_{qq} - \bar{d}_q^* \bar{D}_{q-1}^+ \bar{d}_q] \tag{1.20}$$

and

$$\text{(iii)}\ G_{q3} \triangleq V_\beta^* \in R^{(r_q - \rho_q) x r_q}. \tag{1.21}$$

Then if strict inequality occurs in (1.16b) the set of all G_q which satisfy (1.13) are defined by

$$G_q = G_{q1} + G_{q2} U_q G_{q3} \tag{1.22a}$$

where

$$U_q \in R^{s_q x (r_q - \rho_q)};\ s_q \le r_q - \rho_q \le n_y \tag{1.22b}$$

is an arbitrary row unitary matrix (i.e. $U_q U_q^* = I$). Furthermore, if the Moore-Penrose P_q^+ of

$$P_q = [E_q^*\ F_q^*]^* \tag{1.23}$$

is expressed as

$$P_q^+ = [\tilde{L}_{11}\ L_{12}];\ \tilde{L}_{11} \in R^{r_q x (q-1)n_y},\ L_{12} \in R^{r_q x n_y} \tag{1.24}$$

then the corresponding state space representation $\{A_q, B_q, C_q\}$ of the q-Markov COVER is given by

$$A_q = L_{11} + L_{12} G_q;\ L_{11} = \tilde{L}_{11} F_q \in R^{r_q x r_q}$$

$$B_q = P_q^+ [M_0^* M_1^* \cdots M_{q-1}^*]^*;\ C_q = E_q. \tag{1.25}$$

If $r_q = \rho_q$, then $G_q = G_{q1}$ is unique.

Proof: The expression for $F_q G_q^*$ in (1.13) implies G_q^* is of the form

$$G_q^* = F_q^+ \bar{d}_q + G_{q3}^* M^*;\ M \in R^{n_y x (r_q - \rho_q)}$$

for some M. Then expanding $G_q G_q^*$ using (1.13) we have

$$\bar{d}_{qq} = \bar{d}_q^*(F_q^+)^* F_q^+ \bar{d}_q + \bar{d}_q^*(F_q^+)^* G_{q3}^* M^* + M G_{q3} F_q^+ \bar{d}_q + M G_{q3} G_{q3}^* M$$

Also from (1.13) and (1.21)

$$(F_q^+)^* F_q^+ = \bar{D}_{q-1}^+, \quad G_{q3} G_{q3}^* = I; \quad (F_q^+)^* G_{q3}^* = 0 \tag{1.26}$$

so that

$$MM^* = \bar{d}_{qq} - \bar{d}_q^*(F_q^+)^* F_q^+ \bar{d}_q$$

Since MM^* has rank s_q,

$$s_q = \text{rank}(G_{q2} G_{q2}^*) \le r_q - \rho_q$$

2. Optimal Finite Wordlength q-Markov COVER

A fixed point finite wordlength realization of the ideal (i.e. infinite precision) q-Markov COVER (1.1) shall be referred to as a q-FWL Markov COVER and is described by

$$\hat{x}(k+1) = \hat{A} Q[\hat{x}(k)] + \hat{B}\hat{u}(k)$$

$$\hat{y}(k) = \hat{C} Q[\hat{x}(k)] \tag{2.1}$$

$$Q[\hat{x}(k)] = \hat{x}(k) - e(k)$$

The components of the matrices $\hat{A}$, $\hat{B}$, $\hat{C}$ are assumed to have a W_o bit fractional representation obtained by quantization of the components of A, B, C in (1.1). The components of $\hat{x}(k)$ have a $W+W_0$ bit fractional part while components of $Q[\hat{x}(k)]$ and $\hat{u}(k)$ all have a W bit fractional part. The components of the state residue vector e(k) has a $W+W_o$ bit fractional representation in which the most significant W bits are zero. The LHS and RHS of (2.1) are therefore consistent with respect to their fractional wordlength representation. The number of bits required to represent the integer parts of $\hat{A}$, $\hat{B}$ and $\hat{C}$ depend on the dynamic range of the coefficients. State space structures in which all coefficients are less than unity are therefore advantageous in this regard. The required integer representation of $Q[\hat{x}(k)]$ will depend on the dynamic range of the input signal $\hat{u}(k)$. Inadequate dynamic range will result in arithmetic overflow. The accuracy in the computation of $\hat{x}(k)$ is determined by its fractional wordlength W.

Define the state error vector $\varepsilon_x(k)$ and output error vector $\varepsilon_y(k)$ by

$$\varepsilon_x(k) = \hat{x}(k) - x(k); \quad \varepsilon_y(k) = \hat{y}(k) - y(k) \tag{2.2}$$

Then from (1.1), (2.1) and (2.2)

$$\varepsilon_x(k+1) = A\varepsilon_x(k) + Ae(k) + \Delta AQ[\hat{x}(k)] + \Delta Bu(k) + B\Delta u(k) \tag{2.3}$$

$$\varepsilon_y(k) = C\varepsilon_x(k) + Ce(k) + \Delta CQ[\hat{x}(k)]$$

where

$$\Delta A = \hat{A} - A;\ \Delta B = \hat{B} - B;\ \Delta C = \hat{C} - C$$

$$\Delta u(k) = \hat{u}(k) - u(k)$$

There are five terms which contribute to the output error (i) internal arithmetic errors e(k) due to state quantization (ii) coefficient errors due to errors ΔA in A (iii) ΔB in B (iv) ΔC in C, and (v) input quantization errors $\Delta u(k)$. Under weak 'sufficiently exciting' conditions on the input $\{u(k)\}$ it can be shown [6] that if $Q[\cdot]$ in (2.1) denotes 'roundoff' quantization, then $\{e(k)\}$ is a zero mean uniform white process with covariance

$$E\{e(k)e^*(k)\} = \gamma^2 I;\ \gamma^2 = \frac{1}{12}2^{-2W}. \tag{2.4}$$

Similarly $\{\Delta u(k)\}$ is assumed to be a zero mean white uniform process with

$$E\{\Delta u(k)\Delta^* u(k)\} = \gamma^2 I \tag{2.5}$$

We assume that the quantized coefficients $\hat{A}$, $\hat{B}$, $\hat{C}$ are obtained by rounding A, B, C to W_o bit fractions. Consequently, all components Δp of the error matrices ΔA, ΔB, ΔC satisfy

$$|\Delta p| < \gamma_o;\ \gamma_o = \frac{1}{2}2^{-W_o}. \tag{2.6}$$

For simplicity we normalize the error matrices and define δA, δB, δC by

$$\delta A = \frac{1}{\gamma_o}\Delta A;\ \delta B = \frac{1}{\gamma_o}\Delta B;\ \delta C = \frac{1}{\gamma_o}\Delta C \tag{2.7}$$

so that all components δp of δA, δB and δC satisfy

$$|\delta p| < 1. \tag{2.8}$$

The steady state output error covariance Y of $\{E_y(k)\}$ is then given by

$$Y = CPC^* + \gamma^2 CC^* + \gamma_o^2(\delta C)(\hat{X}+\gamma^2 I)(\delta C)^* - \gamma_o\gamma^2[C(\delta C)^* + (\delta C)C^*], \tag{2.9}$$

where

$$P = E\{\varepsilon_x(k)\varepsilon_x^*(k)\}$$

$$= APA^* + \gamma^2 AA^* + \gamma_o^2(\delta A)(\hat{X}+\gamma^2 I)(\delta A)^* + \gamma_o^2(\delta B)(\delta B)^* + \gamma^2 BB^*$$

and

$$\hat{X} = E\{\hat{x}(k)\hat{x}^*(k)\} = \hat{A}\hat{X}(\hat{A})^* + \gamma^2\hat{A}(\hat{A})^* + (1+\gamma^2)\hat{B}\hat{B}^*$$

For the remainder of this section we assume no coefficient errors (i.e. $\gamma_o = 0$ in (2.9)) and consider only the effects due to *finite state wordlength* (FSWL). The issue of coefficient error shall be resumed in Section 5.

Theorem 2.1

Define the output noise measure

$$J \triangleq \mathrm{tr}[Y] .$$

Then for $\gamma_o = 0$

$$J = \gamma^2\{\mathrm{tr}[K] + \mathrm{tr}[B^*KB]\} \tag{2.10}$$

where

$$K = A^*KA + C^*C . \tag{2.11}$$

Proof: From (2.9)

$$Y = C\bar{P}C^*; \quad \bar{P} = A\bar{P}A^* + \gamma^2 Z = P + \gamma^2 I$$

where

$$Z = I + BB^*;$$

Now

$$\bar{P} = \gamma^2 \sum_{k=0}^{\infty} A^k Z (A^k)^*$$

and

$$K = \sum_{k=0}^{\infty} (A^k)^* C^* C A^k$$

so that

$$\mathrm{tr}[C\bar{P}C^*] = \gamma^2 \mathrm{tr}(ZK) .$$

□□□

A *fixed point q-FSWL Markov COVER* corresponding to the (ideal) q-Markov COVER (1.1) is therefore described by

$$\hat{x}(k+1) = AQ[\hat{x}(k)] + B\hat{u}(k)$$

$$\hat{y}(k) = CQ[\hat{x}(k)] \tag{2.12}$$

$$Q[\hat{x}(k)] = \hat{x}(k) - e(k)$$

The *output noise gain* (η_x) *due to state quantization* and the *output noise gain* (η_u) *due to input quantization* are defined by

$$\eta_x \triangleq \mathrm{tr}[K];\ \eta_u \triangleq \mathrm{tr}[B^*KB] \tag{2.13}$$

The noise gain η_x generally varies with state space representation whereas η_u is independent of the coordinate basis. Specifically, consider the q-FSWL Markov COVER

$$\hat{z}(k+1) = A\,Q[\hat{z}(k)] + B\,\hat{u}(k)$$

$$y(k) = C\,Q[\hat{z}(k)] \tag{2.14a}$$

$$Q[\hat{z}(k)] = z(k) - f(k)$$

where

$$A = T^{-1}AT,\ B = T^{-1}B,\ C = CT \tag{2.14b}$$

and $Q[\hat{z}(k)]$ has a W bit fractional representation. Assuming 'sufficient excitation' by $\hat{u}(k)$, the state residue sequence $\{f(k)\}$ in (2.14a) due to roundoff quantization will again be a zero mean white uniform process with covariance $\gamma^2 I$ as in (2.5). The corresponding output quantization noise gains η_z and $\tilde{\eta}_u$ due respectively to state and input quantization are given by

$$\eta_z = \mathrm{tr}[K_z];\ \tilde{\eta}_u = \mathrm{tr}[B^*K_zB] \tag{2.15}$$

where B is given by (2.14b) and

$$K_z = A\,K_zA^* + C^*C\ . \tag{2.16}$$

But from (2.11), $K_z = T^*KT$, so that

$$\eta_z = \mathrm{tr}[T^*KT];\ \tilde{\eta}_u = \mathrm{tr}[B^*KB] \tag{2.17}$$

Notice from (2.13) that the noise gain η_u due to input quantization errors is *unaffected* by a similarity transformation. Conversely the noise gain η_x due to state quantization generally changes with co-ordinate bases. There is no change if T is

unitary. The q-FSWL Markov COVER (2.14) is superior to the q-FSWL Markov COVER (2.12) if

$$\eta_z < \eta_x . \tag{2.18}$$

However the comparison in (2.18) must be made under the assumption of *identical scaling* of the states $\hat{x}(k)$ and $\hat{z}(k)$. Specifically, equal l_2-scaling of gain α from a zero mean unit intensity white noise input $\hat{u}(k)$ to the state components $\hat{x}_j(k)$ of $\hat{x}(k)$ requires

$$X_{jj} = \alpha \text{ for all } j \tag{2.19}$$

where X_{jj} denotes the jth diagonal component of the state covariance matrix X given by (1.3). Equal l_2-scaling of gain α of components of $\hat{z}(k)$ in (2.14) requires

$$Z_{jj} = \alpha;\ Z = AZA^* + BB^* \tag{2.20}$$

Equality in l_2-scaling of representations (2.12) and (2.14) is equivalent to equality in the state dynamic range (i.e. number of bits in the integer representation of states) for a given probability of overflow. We now state a result which is important for establishing l_2-scaling.

Lemma 2.1 [8,9] Suppose $M = M^* > 0$ is an nxn matrix. Then a necessary and sufficient condition for the existence of a unitary matrix V such that

$$(VMV^*)_{jj} = \alpha \text{ for all } j$$

is

$$tr[M] = n\alpha$$

□□□

We have shown in Lemma 1.1 that different similarity transformations of an ideal q-Markov COVER corresponds to different factorization of the first data matrix D_q in (1.5a). Our aim is to optimize this factorization.

Definition 2.1

The *Optimal q-FSWL Markov COVER* minimizes the output quantization noise gain η due to state quantization errors; that is

$$\eta_{opt} = \min_{T,G_q} \mathrm{tr}[T^* K_q T];\ T^*T = \Lambda^{-1} \tag{2.21}$$

subject to the l_2-scaling constraint:

$$\Lambda_{jj} = \alpha \text{ for all } j \tag{2.22}$$

where the observability grammian K_q satisfies

$$K_q = A_q^* K_q A_q + C_q^* C_q \tag{2.23}$$

with $\{A_q, B_q, C_q\}$ defined by (1.22)-(1.25).

□□□

In corollary 1.1 we have shown that all the degrees of freedom available to select G_q are confined to an arbitrary row unitary matrix U_q. We now show how to optimize U_q.

Theorem 2.1

a. The optimal q-FSWL Markov COVER (1.21a), (1.24) is defined by

$$\eta_{opt} = r_q^{-1} \min_{U_q} (\mathrm{tr}[K_q^{1/2}])^2 \tag{2.24}$$

where $U_q \in R^{s_q \times (r_q - \rho_q)}$ is an arbitrary row unitary matrix and K_q satisfies (2.23).

b. The transfer function of the optimal q-FSWL Markov COVER has Hankel singular values given by the eigenvalues of K_q defined by the minimizing U_q.

c. Suppose $U_q = U_{qo}$ is the minimizing solution corresponding to the optimal $G_q = G_{qo}$ in (1.21). Let $\{A_{qo}, B_{qo}, C_{qo}\}$ be the corresponding state space realization in (1.24). Then the optimal q-FSWL Markov COVER has a (nonunique) state space representation $\{T_o^{-1} A_{qo} T_o, T_o^{-1} B_{qo}, C_{qo} T_o\}$ where

$$T_o = U_o \pi_o V_o^* \tag{2.25}$$

such that

(i) the unitary matrix U_o is defined by

$$U_o^* K_{qo} U_o = \Sigma_o^2 \tag{2.26a}$$

where

$$K_{qo} = A_{qo}K_{qo}A_{qo}^* + C_{qo}^*C_{qo}\ ;\ \sum_o^2 = \text{diag}\{\sigma_{10}^2, \sigma_{20}^2, ..., \sigma_{r_q0}^2\} \qquad (2.26b)$$

in which $\{\sigma_{jo}^2\}$ are the optimal Hankel singular values (eigenvalues of K_{qo}).
(ii)

$$\pi_o^2 = \frac{1}{\alpha^2 r_q}\left(\sum_{k=1}^{r_q}\sigma_{ko}\right)\sum_o^{-1} \qquad (2.27)$$

and (iii) V_o is unitary such that

$$(V_o\sum_o V_o^*)_{jj} = \frac{\sum_{k=1}^{r_q}\sigma_{ko}}{r_q} \text{ for all } j \qquad (2.28)$$

$$\eta_{opt} \triangleq \eta_q \text{ (optimal)} = \frac{1}{\alpha^2 r_q}\left(\sum_{k=1}^{r_q}\sigma_{ko}\right)^2 \qquad (2.29)$$

Proof: By corollary 1.1 we have for G_q defined by (1.22) for any row unitary matrix U_q (of appropriately specified dimensions) that G_q defines a q-Markov COVER. The corresponding realization $\{A_q, B_q, C_q\}$ for each such U_q has identity controllability grammian and observability grammian K_q defined by (2.23). Now given a particular U_q, apply a similarity transformation

$$T = U_o\pi_o V_o^*$$

to the given q-Markov COVER. Then

$$\text{tr}(T^*K_qT) = \text{tr}(\pi_o^2 U_o^* K_q U_o)$$

and

$$(T^*T)^{-1} = V_o\pi_o^{-2}V_o^*$$

By lemma 2.1, the l_2-scaling constant can be satisfied for some V_o provided $\text{tr}(\pi_o^{-2}) = n\alpha$. Following Williamson [1, Theorem 4.1] (with a minor modification of the l_2-scaling constraint), the optimal performance is given by

$$\eta_{qopt} = \frac{\left(\sum_{\alpha=1}^{r_q}\sigma_j\right)^2}{\alpha^2 r_q}$$

where $\{\sigma_j^2\}$ are the eigenvalues of K_q. That is,

$$\mathrm{tr}(K_q^{1/2}) = \sum_{j=1}^{r_q} \sigma_j$$

The *optimal* q-FSWL Markov COVER therefore achieves the minimum in (2.24). The structure of U_o, π_o, V_o in (2.25)-(2.29) follow directly from Williamson [1] (see proof of Theorem 4.1 with U = I).

3. Computation of the Optimal FSWL Markov COVER

Necessary conditions for the optimal solution in Theorem 2.1 can be obtained using the method of Lagrange multipliers. Specifically, let

$$J = (\mathrm{tr}[K_q^{1/2}])^2 + \mathrm{tr}[\Lambda(-K_q + A_q^* K_q A_q + C_q^* C_q)] + \mathrm{tr}[\Omega(I - U_q U_q^*)] \qquad (3.1a)$$

where

$$\Lambda = \Lambda^* \in R^{r_q x r_q}; \ \Omega = \Omega^* \in R^{s_q x s_q} \qquad (3.1b)$$

are symmetric Lagrange multipliers. After taking derivatives of J using (1.22) and (1.25)

$$\frac{\partial J}{\partial \Lambda} = K_q - A_q^* K_q A_q - C_q^* C_q$$

$$\frac{\partial J}{\partial \Omega} = I - U_q U_q^*$$

$$\frac{\partial J}{\partial K_q^{1/2}} = 2I + 2\Lambda K_q^{1/2} + 2A_q^* \Lambda A_q K_q^{1/2} \qquad (3.2)$$

$$\frac{\partial J}{\partial U_q} = 2G_{q2}^* L_{12}^* K_q A_q \Lambda G_{q3}^* - 2\Omega U_q$$

By setting these derivatives to zero we obtain the following result.

Lemma 3.1 Necessary conditions for the derivation of the optimal q-FSWL Markov COVER are given by

$$K_q = A_q^* K_q A_q + C_q^* C_q$$

$$-\Lambda = A_q^* \Lambda A_q + K_q^{-1/2}; \quad \Lambda = \Lambda^* \in R^{r_q x r_q}$$

$$U_q U_q^* = I \qquad ; \ U_q \in R^{s_q x (r_q - \rho_q)} \qquad (3.3)$$

$$\Omega U_q - P_q U_q Q_q = R_q \quad ; \; \Omega = \Omega^* \in R^{s_q x s_q}$$

where

$$P_q = P_q^* = G_{q2}^* L_{12}^* K_1 L_{12} G_{q2} \in R^{s_q x s_q}$$

$$Q_q = Q_q^* = G_{q3} \Lambda G_{q3}^* \in R^{(r_q-\rho_q)x(r_q-\rho_q)} \tag{3.4}$$

$$R_q = G_{q2}^* L_{12}^* K_q (L_{11}+L_{12}G_{q1}) \Lambda G_{q3}^* \in R^{s_q x (r_q-\rho_q)}$$

and A_q, G_{qj}, L_j are defined by (1.20)-(1.24)

□□□

These necessary conditions cannot be solved explicitly for the optimal row unitary matrix U_q and so an iterative procedure is required. One possible algorithm is now described.

Recursive Algorithm for Optimal q-FSWL Markov COVER:

(0) Set j = 0 and choose any row unitary $U_q(0)$ in (1.21a)

(1) Form $A_q(j)$ from

$$A_q(j) = (L_{11}+L_{12}G_{q1}) + L_{12}G_{q2}U_q(j)G_{q3} \tag{3.5a}$$

(2) Compute $K_q(j)$: $K_1(j) = A_q^*(j)K_q(j)A_q(j) + C_q^* C_q$ (3.5b)

(3) Compute $\Lambda(j)$: $-\Lambda(j) = A_q^*(j)\Lambda(j)A_q(j) + K_q^{-½}(j)$; $\Lambda(j) = \Lambda^*(j)$ (3.5c)

(4) Compute $P_q(j)$, $Q_q(j)$ $R_q(j)$:

$$P_q(j) = G_{q2}^* L_{12}^* K_q(j) L_{12} G_{q2}; \; Q_q(j) = G_{q3}\Lambda(j)G_{q3}^*;$$

$$R_q(j) = G_{q2}^* L_{12}^* K_q(j)(L_{11}+L_{12}G_{q1})\Lambda(j)G_{q3}^* \tag{3.5d}$$

(5) Update $U_q(j)$ by solving the nonlinear algebra problem:

$$\Omega(j)U_q(j+1) - P_q(j)U_q(j+1)Q_q(j) = R_q(j); \quad \Omega(j) = \Omega^*(j) \tag{3.5e}$$

$$U_q(j+1)U_q^*(j+1) = I$$

The most difficult step at each stage of the algorithm is to solve (3.5e) for a row unitary $U_q(j+1)$ and symmetric $\Omega(j)$. There is generally no explicit solution