A New Model Performance Index for Engineering Design of Flight Control Systems

Herman A. Rediess*
NASA Flight Research Center, Edwards, Calif.
AND
H. Philip Whitaker†
Massachusetts Institute of Technology, Cambridge, Mass.

The theory and application of a new performance index, the Model PI, that brings engineering design specifications into the analytical design process is presented. A parameter optimization design procedure is established that starts with practical engineering specifications and uses the Model PI as a synthesis tool to obtain a satisfactory design. The Model PI represents a new criterion for approximating one dynamical system by another, based on a novel geometrical representation of linear autonomous systems. It is shown to be an effective performance index in designing practical systems and to be substantially more efficient to use than a comparable model-referenced integral squared error performance index. The design procedure is demonstrated by designing a lateral-directional stability augmentation system for the X-15 aircraft.

Nomenclature

$\tilde{\mathbf{a}}$ = system extended coefficient vector
a_i = system characteristic polynomial coefficient of s^i
b_i = system transfer function numerator polynomial coefficient of s^i
l = order of model
m = number of system zeros
n = order of system
$\tilde{\mathbf{Q}}$ = Model PI weighting matrix, $\tilde{\alpha}\tilde{\alpha}'/\|\tilde{\alpha}\|^2$
r = pseudo initial condition weighting factor
u = input variable
$\tilde{\mathbf{W}}$ = extended pseudo initial condition weighting matrix
x = system transient response variable
$\tilde{\mathbf{x}}$ = system extended state vector or trajectory
y = output variable
$\tilde{\alpha}$ = model extended coefficient vector in system's extended state space

Subscripts

m = model
0 = initial time

Special notation

$(\ )^{(n)}$ = nth derivative of the variable ( )
$(\ )'$ = transpose of ( )
$\|\mathbf{v}\|$ = $(\mathbf{v}'\mathbf{v})$ which is the length of v
$\|\mathbf{v}\|_M^2$ = $\mathbf{v}'M\mathbf{v}$

Introduction

THE conventional design process for flight control systems has become complex, lengthy, and laborious. Even with the aid of digital computers for root locus and Bode diagram computations and automatic root locus plotters, these techniques are rather slow when used to synthesize complex multiloop and multiple input/output systems. This is due mainly to their "trial-and-error" nature, with only one of several design parameters treated at a time. The iterations between the effects of several parameters and between inner and outer loop closures can become tedious.

This paper presents a systematic design technique based on a new performance index[1] that allows the designer to introduce practical engineering specifications into the design process, in the convenient form of a model, and yet implement the technique in a general and efficient manner on a digital computer. Models have been included in previous analytical design techniques, notably in the form of model-referenced integral squared error (ISE) performance indices (e.g., Refs. 2–5). This should not be equated to "model-following," which is a specific design configuration that results from using a model-referenced ISE performance index in the optimal regulator problem. In the parameter optimization approach the designer selects the configuration, which need not include the model per se, and the optimization procedure determines the values of the free design parameters that minimize the model-referenced ISE. One problem that has hindered its application to practical system design is that even for moderately high-order systems and models the computational task can become enormous. However, the new "Model Performance Index" (Model PI) presented here has the potential of reducing the computation time by 70–85% of that required for the model-referenced ISE procedure.

The Model PI includes a model to represent the design specifications in an entirely different manner from that used with the familiar model-referenced ISE performance index. It is based on a new geometrical criterion for approximating one dynamical system (the model) by another (the actual system). The basic form of the resulting Model PI is the same as that of quadratic functionals frequently appearing in modern control theory. The important distinction, however, is the ability to interpret the performance index directly in terms of a model of the desired system response, without actually having the model's time response in the performance index. In this respect the Model PI theory can be related to the previous works by Aizerman[6] and Rekasius[7] on generalized performance indices. However, the Model PI does not have the serious limitations that Aizerman's and Rekasius'

Received August 13, 1969; presented as Paper 69-885 at the AIAA Guidance, Control, and Flight Mechanics Conference, Princeton, N.J., August 18–20, 1969; revision received December 22, 1969. This work was performed at the Massachusetts Institute of Technology in partial fulfillment of the requirements of the degree of Doctor of Philosophy (Aeronautical and Astronautical Engineering) and was sponsored by the NASA Electronics Research Center, Cambridge, Mass., through NASA Grant NGR 22-009-229.

* Chief, Flight Control Analysis. Member AIAA.
† Professor, Department of Aeronautics and Astronautics. Member AIAA.

performance indices have, which restrict their application to essentially academic examples.[1] The Model PI also has an interesting relationship to the works of Kalman[8] and Schultz and Melsa[9] in linear optimal control theory, which will be presented in a future paper.

Model Performance Index Theory

There are three major steps to establishing a rigorous basis for the Model PI. First, the input/output transfer characteristics are represented by a linear autonomous system with a specific set of pseudo initial conditions (IC). These pseudo IC's contain the effect of system zeros and are also functions of the characteristic equation coefficients. Second, the linear autonomous system and pseudo IC's are represented geometrically in the system's extended state space by a characteristic plane and pseudo IC vector. Third, by transforming both the system and the model into this geometrical representation, a new criterion for approximating one dynamical system by another is established, which results in the Model PI.

A Geometrical Representation of Linear Systems

Consider a single input/output system that can be described by a closed-loop transfer function of the form

$$\frac{y(s)}{u(s)} = \frac{b_m s^m + \ldots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \ldots + a_2 s^2 + a_1 s + a_0} \quad (1)$$

where y is the output, u is the input, n is the order of the system, and $m \leqq n - 1$. The system's actual initial conditions are assumed to be zero.

There are many alternate ways to represent mathematically the transfer characteristics of a system. The representation considered here is the transient portion of the time response of the system for a unit step input. Assuming that a finite steady-state value of the output y_{ss} exists for a step input, the transient portion of the response is defined as

$$x(t) = y(t) - y_{ss} \quad (2)$$

The transform of the transient response is easily obtained from Eq. (1) as

$$x(s) = \frac{b_m s^{m-1} + \ldots + b_3 s^2 + b_2 s + b_1}{s^n + a_{n-1}s^{n-1} + \ldots + a_2 s^2 + a_1 s + a_0} - \frac{b_0/a_0(s^{n-1} + a_{n-1}s^{n-2} + \ldots + a_2 s + a_1)}{s^n + a_{n-1}s^{n-1} + \ldots + a_2 s^2 + a_1 s + a_0} \quad (3)$$

The work that follows is based on representing the transfer characteristics of a linear invariant system by a linear autonomous system with a specific set of pseudo IC's. These initial conditions will be referred to as pseudo to emphasize that they are not the actual initial conditions of the original system.

By replacing the step input with a specific set of hypothetical or pseudo IC's for $x(t)$, one can produce a response identical to the unit step response. The reason for doing this is that the effect of the system's closed-loop transfer function zeros can be included in a performance index developed subsequently by these pseudo IC's. The transient response is then described by the homogeneous differential equation

$$x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \ldots + a_1\dot{x}(t) + a_0 x(t) = 0 \quad (4)$$

with pseudo IC's denoted by

$$\left.\begin{aligned} x(0) &= x_0 \\ \dot{x}(0) &= \dot{x}_0 \\ &\cdot \\ &\cdot \\ &\cdot \\ x^{(n-1)}(0) &= x_0^{(n-1)} \end{aligned}\right\} \quad (5)$$

where the values of x_0, $\dot{x}_0$, etc., are yet to be determined to give the desired equivalence. These can be established easily by taking the Laplace transform of Eq. (4) and comparing the result with Eq. (3), as follows:

$$\begin{aligned} (s^n + a_{n-1}s^{n-1} + \ldots + a_2 s^2 + a_1 s + a_0)x(s) = {} & (\dot{x}_0)s^{n-2} + (\ddot{x}_0 + a_{n-1}\dot{x}_0)s^{n-3} + \ldots (x_0^{(n-3)} + \\ & a_{n-1}x_0^{(n-4)} + \ldots + a_4\dot{x}_0)s^2 + (x_0^{(n-2)} + \\ & a_{n-1}x_0^{(n-3)} + \ldots + a_4\ddot{x}_0 + a_3\dot{x}_0)s + (x_0^{(n-1)} + \\ & a_{n-1}x_0^{(n-2)} + \ldots + a_3\ddot{x}_0 + a_2\dot{x}_0) + x_0(s^{n-1} + \\ & a_{n-1}s^{n-2} + \ldots + a_2 s + a_1) \end{aligned} \quad (6)$$

Note that x_0 occurs in Eq. (6) only as a factor of the last term. From the numerator of Eq. (3) it is clear that $x_0 = -b_0/a_0$. Then, equating the coefficients of like powers of s of the remaining terms on the right-hand side of Eq. (6) with those of the numerator of the first term in Eq. (3) results in the relationships

$$\left.\begin{aligned} x_0 &= -b_0/a_0 \\ x_0^{(n-i)} &= \begin{cases} 0 & \text{for } i > m \\ b_i - \sum\limits_{j=n-m}^{n-i-1} a_{j+i}x_0^{(j)} & \text{for } i = 1,2,\ldots,m \end{cases} \end{aligned}\right\} \quad (7)$$

for $i = 1,2,3,\ldots, n - 1$. Thus, if one uses the initial condition given by Eq. (7) with the homogeneous Eq. (4), the time response would be identical to the transient response of the original system equation for a unit step input. An alternate interpretation of the pseudo IC's is that they are the values of the transient response variable, $x(t)$, and its $(n - 1)$th derivatives at an infinitely small time increment after the application of the unit step input.

The linear autonomous system [Eq. (4)] can be written in the following compact form‡:

$$\tilde{\mathbf{x}}'(t)\tilde{\mathbf{a}} = 0 \quad (8)$$

with pseudo IC vector $\tilde{\mathbf{x}}_0$, where $\tilde{\mathbf{x}}(t)$ and $\tilde{\mathbf{a}}$ are $(n + 1) \times 1$ vectors, defined as

$$\tilde{\mathbf{x}}'(t) = [x(t)\dot{x}(t) \ldots x^{(n-1)}(t)x^{(n)}(t)] \quad (9)$$

$$\tilde{\mathbf{a}}' = [a_0 a_1 \ldots a_{n-1} 1] \quad (10)$$

The first n elements of $\tilde{\mathbf{x}}_0$ are determined from Eq. (7), and the $(n + 1)$th element is determined by making $\tilde{\mathbf{x}}_0$ satisfy Eq. (8) at $t = 0$. The $(n + 1)$th-dimensional space is referred to here as the extended state space, and $\tilde{\mathbf{x}}(t)$ and $\tilde{\mathbf{a}}$ are referred to as the extended state and coefficient vectors, respectively. The $(n + 1)$th-dimensional trajectory that $\tilde{\mathbf{x}}(t)$ represents as a function of time is determined by Eq. (8) together with the pseudo IC vector $\tilde{\mathbf{x}}_0$. Equation (8) defines a plane§ in the extended state space that is normal to $\tilde{\mathbf{a}}$, so that $\tilde{\mathbf{x}}(t)$ lies within this plane for all values of time. Any other linear autonomous system producing a trajectory that lies within this plane can differ from the first only by its pseudo IC's. Thus, the plane defined by $\tilde{\mathbf{a}}$ contains the trajectories of all possible systems with the same characteristic equation and is referred to here as the characteristic plane or simply the $\tilde{\mathbf{a}}$ plane. Therefore, a linear, invariant system can be described by its characteristic plane and pseudo IC's.

This method of describing a linear system is illustrated geometrically in Fig. 1 for a second-order system given by

$$y(s)/u(s) = 1/(s^2 + 1.414s + 1) \quad (11)$$

‡ The transpose of a vector or matrix ( ) is denoted by a prime ( )′.

§ In general, this is more correctly referred to as a hyperplane or linear manifold, but, for convenience, it will be referred to here simply as a plane.

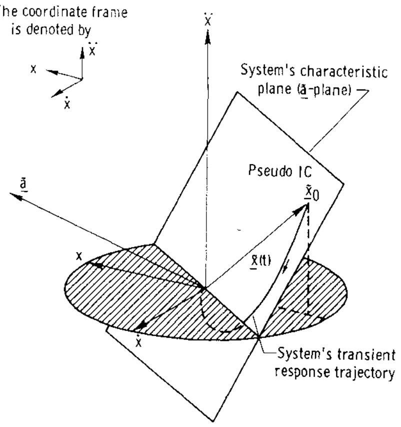

Fig. 1 Geometrical representation of a linear autonomous system; pseudo IC vector $\tilde{\mathbf{x}}_0' = [-1\ 0\ 1]$.

or

$$\tilde{\mathbf{x}}'(t)\tilde{\mathbf{a}} = 0$$

with

$$\begin{aligned} \tilde{\mathbf{a}}' &= [\ \ 1 \quad 1.414 \quad 1] \\ \tilde{\mathbf{x}}'_0 &= [-1 \quad\ \ 0 \quad\ \ 1] \end{aligned} \tag{12}$$

The time response trajectory is seen to start at $\tilde{\mathbf{x}}_0$ and traverse toward the origin, always lying in the $\tilde{\mathbf{a}}$ plane. The trajectory of any other system with the same poles as Eq. (11), but with zeros, would also lie within this plane but would start at a different $\tilde{\mathbf{x}}_0$.

Representing a system by its characteristic plane and pseudo IC's does not set forth any new mathematical information. The statement that the trajectory of any nth-order system lies within a plane in the extended state space of $(n + 1)$th order merely relates the fact that the nth derivative is a linear combination of the state variables, as in Eq. (4). However, it provides a useful way of visualizing the process of approximating one system by another using the performance index developed in the next section.

Model Performance Index (Model PI)

The term "model" as used here relates specifically to dynamic response specifications of a system design problem. Usually it refers to a mathematical model of the desired, or at least satisfactory, response characteristics, so that the design objective would be to approximate the model closely.

As pointed out in the previous section, a linear, invariant system can be represented geometrically by its characteristic plane and pseudo IC. If both the model and the system to be designed are represented in this fashion, one can establish criteria for approximating the model by the system in terms of their characteristic planes and psuedo IC's. The Model PI is one such criterion. The basic form of the Model PI can be thought of as a generalized measure of the distance between the system's time response trajectory and the model's characteristic plane. This concept can be explained most clearly for the situation in which the closed-loop system to be designed has no zeros and the model is of the same order as the system. The general case, for systems with zeros and models of any order, will be discussed subsequently.

Figure 2 illustrates this geometrical interpretation for a second order model and system. The model's extended coefficient vector and pseudo IC vector are denoted $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\mathbf{x}}_{m_0}$, respectively. The time response trajectories for the system and model are shown to lie within their respective characteristic planes, starting at their respective pseudo IC vectors and traversing to the origin.

The system design problem can be thought of as trying to make the system's response trajectory match or closely approximate that of the model by selecting the best $\tilde{\mathbf{a}}$, subject to certain constraints. The coefficient vector is a function of the free design parameters so that varying the design parameters varies $\tilde{\mathbf{a}}$ in a certain constrained manner. Ideally, the designer would like to make these trajectories coincide, which means perfect model matching. That is, the time histories of the system's and model's response to a step input would be identical. It is shown in Ref. 1 that, for systems without zeros, this will occur if and only if the system's trajectory, $\tilde{\mathbf{x}}(t)$, lies within the model's characteristic plane. Furthermore, if the system's trajectory could be made to lie close to the model's characteristic plane, it would be close to the model's trajectory. Therefore, a criterion for approximating the model by the system can be established based on minimizing some generalized measure of the distance between the system's trajectory $\tilde{\mathbf{x}}(t)$ and the $\tilde{\boldsymbol{\alpha}}$ plane.

The instantaneous distance from $\tilde{\mathbf{x}}(t)$ to the $\tilde{\boldsymbol{\alpha}}$ plane is given by¶

$$\|\tilde{\mathbf{x}}'(t)\tilde{\boldsymbol{\alpha}}\|/\|\tilde{\boldsymbol{\alpha}}\| \tag{13}$$

A generalized measure of the distance from $\tilde{\mathbf{x}}(t)$ to the $\tilde{\boldsymbol{\alpha}}$ plane over the time interval 0 to ∞ is the integral square of Eq. (13), which can be written in the compact form

$$\mathrm{PI} = \int_0^\infty \|\tilde{\mathbf{x}}(t)\|^2{}_{\tilde{\mathbf{Q}}}dt \tag{14}$$

where

$$\mathbf{Q} = \tilde{\boldsymbol{\alpha}}\tilde{\boldsymbol{\alpha}}'/\|\tilde{\boldsymbol{\alpha}}\|^2 \tag{15}$$

Equation (14) defines the basic form of the Model PI. It is termed the Model PI because the state vector weighting matrix $\tilde{\mathbf{Q}}$ is used to represent the model, or, more correctly, the model's characteristic plane, in the performance index. It possesses all the desired properties of a performance index, that is, it is a nonnegative number that decreases as the system better approximates the model (in the sense defined earlier), goes to zero if the system matches the model exactly, and is of a convenient mathematical form. The weighting matrix $\tilde{\mathbf{Q}}$ is always a positive semidefinite matrix by its definition and can be written directly from the model's characteristic equation.

The most important property of the Model PI is that it represents the desired results of approximating a model without including the model's time response. This means that minimizing the Model PI numerically involves a computation task set by the dimension of the system only. Any per-

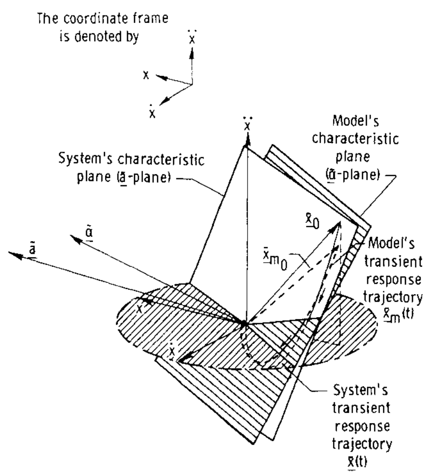

Fig. 2 Geometrical representation of approximating one dynamical system by another.

¶ The notation $\|\mathbf{v}\|_M{}^2$, where M is a square matrix, means $\mathbf{v}'M\mathbf{v}$ and $\|\mathbf{v}\|$ means simply $(\mathbf{v}'\mathbf{v})^{\frac{1}{2}}$, which is the length of $\mathbf{v}$.

formance index that includes the model's time response, such as a model-referenced ISE, involves a computation task set by the dimensions of both the system and the model. The difference in computational efficiency can be substantial and makes a significant impact on the total amount of computer time used in a parameter optimization design method. Using the Model PI method could give up to an 85% savings in computation time over a comparable model-referenced ISE method, as will be shown subsequently.

The development of this point assumed that the model was the same order as the system. The Model PI concept is extended in Ref. 1 to models of lower order than the system. In that case, one obtains the same form for the Model PI as Eq. (14) but with a different, yet compatible, definition for $\tilde{\alpha}$. If l is the order of the model, then the first $(l + 1)$ elements of $\tilde{\alpha}$ are the coefficients of the model's characteristic equation and the remaining elements are zero. Minimizing the Model PI tends to make the system's trajectory match the model's trajectory in the model's extended state space; in other words, the system's output and its first l derivatives tend to match the corresponding model output variables.

Also, the development to this point was for systems without zeros. To extend the Model PI to systems with zeros, it is necessary to apply some constraint to the pseudo IC vector.

The motivation for selecting the Model PI for systems without zeros was to force the system's transient response trajectory to lie close to the model's characteristic plane. In that case the first l pseudo IC's of the system and the model are always identical, so that matching characteristic planes correspond to matching trajectories. This is not necessarily true for systems with zeros. As mentioned previously, the pseudo IC's contain the effect of the system zeros, whereas the characteristic plane is only dependent on the poles. For a given set of poles, changing the zeros of a system corresponds to picking a different starting point for the trajectory in the characteristic plane. A system with zeros and a model with or without zeros would generally not have the same pseudo IC's except in some special cases that are discussed in Ref. 1.

By virtue of the relationship in Eq. (7), the system's pseudo IC's are in general functions of the free design parameters. They are not arbitrary constants as normal initial conditions are usually considered to be. This is a critical point to understand both for extending the Model PI and for deriving an optimization algorithm. The design process of selecting the free parameters changes the pseudo IC's as well as the system's characteristic plane. Therefore, in the general case it is necessary to add a constraint on the relative location of the system's and the model's pseudo IC's. The constraint should be such that if it is satisfied identically, matching the characteristic planes would again correspond to matching the trajectories.

One such constraint that is convenient to use is to place a quadratic penalty on the error between the pseudo IC vectors projected into the model's l-dimensional state space, that is,

$$\|\tilde{\mathbf{x}}_0 - \tilde{\mathbf{x}}_{m_0}\|^2_{\tilde{W}} \tag{16}$$

where

$$\tilde{W} = \left[\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right] \tag{17}$$

and I is the $l \times l$ identity matrix and the 0's are appropriately dimensional null matrices. This places constraints only on the first l pseudo IC's, which is sufficient. If Eq. (16) is identically zero, matching the characteristic planes would correspond to matching the trajectories, at least in the $(l + 1)$-dimensional space. In general it tends to force the first l elements of $\tilde{\mathbf{x}}_0$ to be near the corresponding model pseudo IC's. Selecting the $\tilde{\mathbf{a}}$-plane, via the free parameters, to make the system's trajectory lie close to the $\tilde{\alpha}$ plane then corresponds to approximating the model's trajectory.

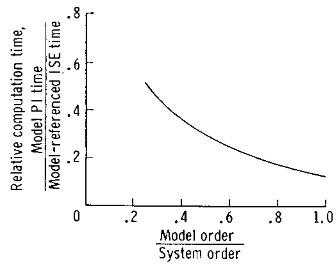

Fig. 3 Approximate relative computation time for the Model PI and model-referenced ISE methods as a function of the ratio of model to system order.

The general form of the Model PI is thus defined as

$$\mathrm{PI} = r\|\tilde{\mathbf{x}}_0 - \tilde{\mathbf{x}}_{m_0}\|^2_{\tilde{W}} + \int_0^\infty \|\tilde{\mathbf{x}}(t)\|^2_{\tilde{Q}} dt \tag{18}$$

The scalar r is selected by the designer to set the relative weighting between matching pseudo IC's and characteristic planes. The form of the Model PI [Eq. (14)] defined for systems without zeros is clearly a special form of Eq. (18).

Once the procedure is established for systems with zeros, it is then possible to treat multivariable systems. Three different methods for treating multivariable systems are presented in Ref. 1.

Design via Parameter Optimization

It is assumed here that the design specifications for the dynamic response of the closed-loop system are given in a graphical form representing an envelope of acceptable design characteristics, either in the time or frequency domain or in the s plane. Starting with these engineering design specifications, the basic Model PI design method proceeds as outlined in the following steps: 1) select a linear model to represent the dynamic response specifications, 2) select a compensation configuration for the control system, 3) form the closed-loop transfer function as a function of the free design parameters, 4) apply the general computer program for control system design using the Model PI, and 5) compare the resulting closed-loop design to the engineering specifications. If they are not satisfied, repeat steps 2–5 with a different compensation configuration. Some variations in these steps, especially step 5, may be necessary for particular design problems.

The state space formulation of analytical design problems makes it possible to establish general digital computer programs for designing linear control systems by parameter optimization. Such a program has been established for the Model PI design procedure,[1] which only requires providing the appropriate input data cards and writing one simple subroutine to change from one design problem to another. The numerical optimization procedure uses a technique, derived in Ref. 1, for evaluating the gradient of the performance index directly rather than using a numerical difference procedure.

The last step in the Model PI design procedure is to determine if the closed-loop system design resulting from the optimization process meets the engineering specifications. If the system response lies within the tolerance envelope, the design problem is complete. At this point, it is immaterial whether or not the system response closely matches the model response as long as it satisfies the engineering specifications. If they are not satisfied, the design process is repeated with some alteration that is likely to result in a better design.

This is essentially the same procedure that would be used for a model-referenced ISE performance index. The major difference is in the computational task of the digital computer program. It is possible to establish a rough estimate of the relative computational effort for these two methods in terms

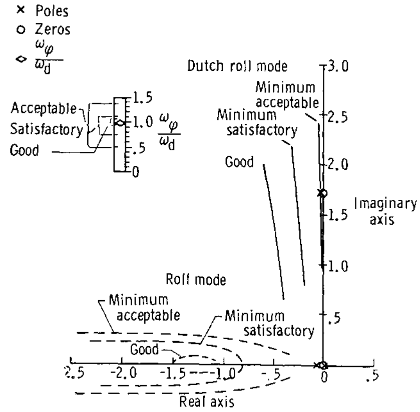

Fig. 4 Lateral-directional handling qualities of the X-15 aircraft (SAS off) at Mach 5.5 and an altitude of 147,000 ft; roll control to roll rate transfer function of X-15.

of the orders of the system and model n and l, respectively. By far the most time consuming portion of the numerical optimization algorithm is the evaluation of the performance index and its gradient. In the program written for this study, the computation time is approximately proportional to the cube of the dimension of the state vector appearing in the performance index. Only the system state vector appears in the Model PI, so in that case this dimension is n. But both the system and model state vectors appear in any model-referenced ISE performance index, so that the dimension is $(n + l)$.** A rough estimate of the relative computation times is then given by

$$\left(\frac{\text{Model PI compute time}}{\text{Model-referenced ISE compute time}}\right) \approx \frac{n^3}{(n+l)^3} \qquad (19)$$

which is plotted as a function of l/n in Fig. 3. If the system and model are of the same order, one can see from the figure that the Model PI method would require only about $\frac{1}{8}$ as much computation time as the model-referenced ISE would require, or about an 85% savings in computation time. If the model is of lower order than the system, the savings would not be as great. It appears unlikely that the model would be less than half the order of the system so that the savings indicated in Fig. 3 would still be about 70%.

A Flight Control System Design Example

Design of a multivariable flight control system by the Model PI method is illustrated here by an example that was previously treated in Ref. 11 by linear optimal control theory and in Ref. 5 by the model-referenced ISE method of parameter optimization. The treatment presented differs from these previous works in respect to the design object. In Refs. 5 and 11, the design objectives were to match the transient responses of a linear model to an initial condition in sideslip angle and to an initial condition in roll rate. The model used corresponds to the dynamic characteristics of a T-33 trainer modified to have generally good handling qualities. Although one might state qualitatively that producing a design with transient responses similar to those of the model for the same initial conditions is generally desirable, one must still check the actual characteristics of the resulting design against the appropriate handling qualities criteria. In this example, the design specifications are in the form of lateral-directional handling qualities criteria. Starting at this point, models are selected to represent the criteria. Then, a stability augmentation system (SAS) design is synthesized using the Model PI. Finally, the design is judged by the actual handling qualities criteria.

** This is based on forming an augmented state vector of $(n + l)$ dimension that includes both the system and model state vectors and then reformulating the state equations and performance index in terms of the augmented state vector, which is a fairly standard procedure.[10] One could reduce the computation time for the model-referenced ISE method by partitioning certain matrix equations, but it could still be two to three times slower than the comparable Model PI method.

Problem Formulation

Generally, an aircraft's lateral-directional motion for small perturbations can be described approximately by

$$\begin{aligned} \dot{p} &= L_p p + L_r r + L_\beta \beta + L_{\delta_a}\delta_a + L_{\delta_v}\delta_v \\ \dot{r} &= N_p p + N_r r + N_\beta \beta + N_{\delta_a}\delta_a + N_{\delta_v}\delta_v \\ \dot{\beta} &= (g/V)\varphi - r + Y_\beta \beta + Y_{\delta_v}\delta_v \end{aligned} \qquad (20)$$

where p is the roll rate, r is the yaw rate, β is the angle of sideslip, φ is the bank angle, δ_a is the roll control surface deflection, and δ_v is the vertical stabilizer deflection, L_p, L_r, L_β, L_{δ_v}, N_p, N_r, N_β, N_{δ_a}, N_{δ_v}, Y_β, and Y_{δ_v} are dimensional stability and control derivatives.

The numerical values used for this example in Refs. 5 and 11 are the characteristics of the X-15 aircraft at a Mach number of 5.5 and at an altitude of 147,000 ft (flight configuration and weight were not specified).

Using the numerical values in Eq. (20), one can obtain the open-loop X-15 transfer function of roll control surface deflection to roll rates as

$$\frac{p(s)}{\delta_a(s)} = \frac{23{,}700s\{1 + [2(0.007)/(1.71)]s + (s/1.71)^2\}}{\underbrace{(1 - s/0.0033)}_{\text{Spiral}}\underbrace{(1 + s/0.0645)}_{\text{Roll subsidence}}\underbrace{\{1 + [2(0.0078)/(1.72)]s + (s/1.72)^2\}}_{\text{Dutch roll}}} \qquad (21)$$

The specific lateral-directional handling qualities criteria for this example are shown in Fig. 4. (See Ref. 1 for a discussion of these criteria.) These criteria consist of regions in the s plane for the roll mode and Dutch roll mode poles and the numerator to denominator frequency ratio of the Dutch roll mode corresponding to good, satisfactory, and acceptable handling qualities. The poles and zeros of Eq. (21) superimposed on Fig. 4 show that the X-15 at this flight condition without a SAS would be predicted to have unacceptable roll and Dutch roll handling qualities based on these criteria. It is apparent that a SAS is needed to increase the roll damping and Dutch roll mode damping in order to obtain good or at least satisfactory handling qualities.

Although the handling qualities criteria used here are described entirely by terms appearing in the $p(s)/\delta_a(s)$ transfer function, they cannot be adequately represented in an analytical design method by just a model of the roll command to roll rate transfer characteristics. Part of the criteria re-

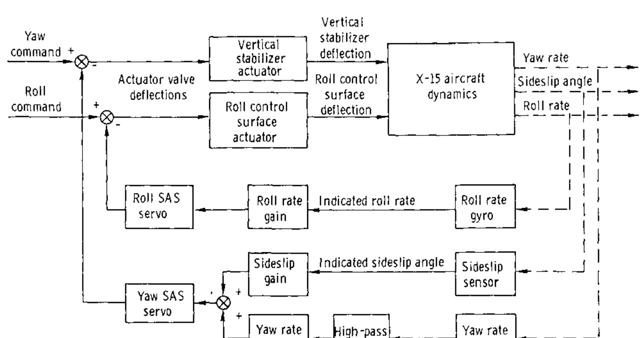

Fig. 5 Functional block diagram of a lateral-directional SAS configuration.

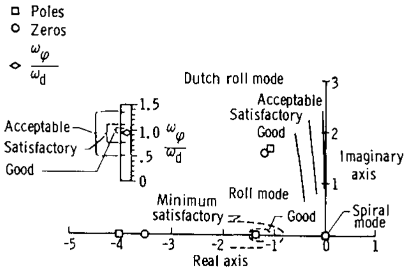

Fig. 6 Estimated handling qualities of the roll control model; roll control to roll rate transfer function [Eq. (23)].

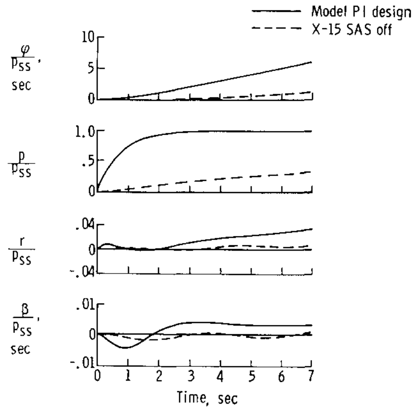

Fig. 7a) Time history comparison of Model PI designed SAS and SAS-off responses of the X-15; step roll command input.

quires the complex zeros to cancel or nearly cancel the Dutch roll mode poles. A model with this characteristic would have such a small Dutch roll mode residue that it would essentially neglect that mode. In addition to a model of the roll command to roll rate transfer characteristics, it is necessary to have a model of some transfer characteristics in which the Dutch roll mode is dominant. A suitable one is the yaw command to sidelip angle transfer function. Before selecting the specific models to use in this case, it is helpful first to form the corresponding system closed-loop transfer functions.

A possible SAS configuration†† is shown in Fig. 5. This configuration is the same as that considered in Ref. 5, which pointed out the need for a high-pass (washout) filter to eliminate the yaw rate feedback in a steady turn. In this example, the dynamics of the motion sensors, the SAS servos, and the control surface actuators are neglected. The washout filter shown, with a time constant of 1 sec, is the same as that used in Ref. 5. The roll rate, sideslip, and yaw rate feedback gains, denoted by k_p, k_β, and k_r, respectively, are the free design parameters. The control surface deflections $\delta_a(s)$ and $\delta_v(s)$ are related to the feedback variables and the commands by

$$\left.\begin{aligned}\delta_a(s) &= \delta_{a_p}(s) - k_p p(s)\\ \delta_v(s) &= \delta_{v_p}(s) - k_r[s/(1+s)]r(s) - k_\beta \beta(s)\end{aligned}\right\} \quad (22)$$

where δ_{a_p} is the pilot roll command and δ_{v_p} is the pilot yaw command. The desired closed-loop transfer functions for pilot roll command to roll rate and pilot yaw command to sideslip angle as functions of k_p, k_β, and k_r are obtained from the Laplace transform of Eqs. (20) and (22).

Two models can now be established to represent the design specifications for these two transfer relationships. In this case, it is relatively simple to select models with the same structure as the closed-loop system transfer functions by referring directly to the handling qualities criteria (Fig. 4).

Four poles of the models can be selected in regions corresponding to good handling qualities in Fig. 4, and the complex zeros can be selected sufficiently close to the Dutch roll poles. The washout filter causes an extra pole and zero not covered explicitly by the criteria in Fig. 4. However, the residue of this extra mode must be kept small in order for it to affect the handling qualities, so it is wise to select the zero

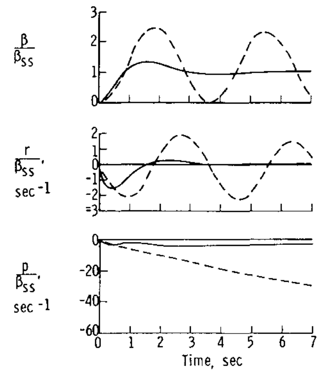

Fig. 7b) Time history comparison of Model PI designed SAS and SAS-off responses of the X-15; step yaw command input.

relatively close to the pole. Following these guidelines, the model for roll command to roll rate is selected to be as shown in Eq. (23), where the sensitivity selected is such that the model would produce the same rolling moment due to roll control surface deflection as the X-15. The model's poles and zeros are superimposed on the handling qualities criteria in Fig. 6. An aircraft with these characteristics would be predicted to have good handling qualities.

$$\frac{p_m(s)}{\delta_{a_p}(s)} = \frac{-624s(1+s/3.5)\{1+[2(0.63)/1.98]s+(s/1.98)^2\}}{\underbrace{(1+s/4.0)}_{\text{filter}}\underbrace{(1+s/1.4)}_{\text{roll subsidence}}\underbrace{(1+s/0.005)}_{\text{spiral}}\underbrace{\{1+[2(0.6)/2.0]s+(s/2.0)^2\}}_{\text{Dutch roll}}} \quad (23)$$

It is only necessary to select zeros to establish the yaw command to sideslip angle model. The handling qualities criteria do not place any requirements on these zeros, but the Dutch roll mode should be dominant. One zero is essentially at the origin, and one is typically near the roll mode pole. The washout filter pole becomes a closed-loop zero in this transfer function. The remaining zero is typically far into the right-half plane. On the basis of these guidelines, the model for yaw command to sideslip angle is selected to be as shown in Eq. (24). The sensitivity used is such that the model would produce the same side force due to the vertical stabilizer deflection as the X-15.

Design by the Model PI Method

The geometrical representations of the models $p_m(s)/\delta_{a_p}(s)$

$$\frac{\beta_m(s)}{\delta_{v_p}(s)} = \frac{12.06s(1+s)(1+s/1.35)(1-s/400)}{(1+s/4.0)(1+s/1.4)(1+s/0.005)\{1+[2(0.6)/2.0]s+(s/2.0)^2\}} \quad (24)$$

†† This is not the actual X-15 SAS configuration,[12] which has a yaw rate feedback to the roll SAS servo and does not have a sideslip angle feedback to the yaw SAS servo.

and $\beta_m(s)/\delta_{v_p}(s)$ are

$$\tilde{\mathbf{x}}'_{1m}(t)\tilde{\boldsymbol{\alpha}} = 0$$
$$\tilde{\mathbf{x}}'_{2m}(t)\tilde{\boldsymbol{\alpha}} = 0 \tag{25}$$

$$\frac{p(s)}{\delta_{a_p}(s)} = \frac{18{,}153s\left(1+\frac{s}{3.01}\right)\left[1+\frac{2(0.582)}{1.70}s+\left(\frac{s}{1.70}\right)^2\right]}{\left(1+\frac{s}{3.68}\right)\left(1+\frac{s}{1.31}\right)\left(1-\frac{s}{0.00019}\right)\left[1+\frac{2(0.575)}{1.64}s+\left(\frac{s}{1.64}\right)^2\right]} \tag{29}$$

and

$$\frac{\beta(s)}{\delta_{v_p}(s)} = \frac{0.165\left(1-\frac{s}{0.00022}\right)(1+s)\left(1+\frac{s}{1.319}\right)\left(1-\frac{s}{573}\right)}{\left(1+\frac{s}{3.68}\right)\left(1+\frac{s}{1.31}\right)\left(1-\frac{s}{0.00019}\right)\left[1+\frac{2(0.575)}{1.64}s+\left(\frac{s}{1.64}\right)^2\right]} \tag{30}$$

with pseudo IC vectors $\tilde{\mathbf{x}}_{1m_0}$ and $\tilde{\mathbf{x}}_{2m_0}$, respectively, where

$$\boldsymbol{\alpha}' = [0.112\ 22.575\ 35.143\ 22.609\ 7.805\ 1] \tag{26}$$

and

$$\tilde{\mathbf{x}}'_{1m_0} = [\mathbf{x}'_{1m_0} \,\vdots\, -\mathbf{x}'_{1m_0}\boldsymbol{\alpha}]$$
$$\tilde{\mathbf{x}}'_{1m_0} = [0\ -5.116\ 9.23\ -20.73\ 64.7] \tag{27}$$

$$\tilde{\mathbf{x}}'_{2m_0} = [\mathbf{x}'_{2m_0} \,\vdots\, -\mathbf{x}'_{2m_0}\boldsymbol{\alpha}]$$
$$\tilde{\mathbf{x}}'_{2m_0} = [0\ -0.0025\ 1.014\ -5.508\ 21.52] \tag{28}$$

The elements of $\mathbf{x}_{1m_0}$ and $\mathbf{x}_{2m_0}$ were computed from Eq. (7) for Eqs. (23) and (24), respectively.

Only one of the three methods for multivariable systems presented in Ref. 1 will be illustrated here. It is to apply the Model PI optimization procedure for a single input/output system alternately to the design of the closed-loop "systems" represented by the roll command to roll rate and yaw command to sideslip angle transfer functions using the appropriate model. In alternate applications the number of free design parameters (feedback gains, in this case) is reduced by retaining only the most effective ones and fixing the others at the best value from the previous minimization.

The design example is in the proper form to apply the numerical optimization program described in Ref. 1 for minimizing a functional of the form of Eq. (18) with respect to the free parameters. All that is necessary is to write a simple subroutine that relates the free parameters, k_p, k_β, and k_r, to the coefficients of the closed-loop system transfer functions and provide the input data cards for the appropriate model. A pseudo IC weighting factor of 4.0×10^{-6} is used in the Model PI [Eq. (18)] throughout.

First, consider selecting the free parameter values k_p, k_β, and k_r that provide the best Model PI match of the system's p/δ_{a_p} transfer characteristics to the model's p_m/δ_{a_p} transfer characteristics. The Model PI solution for this step is $k_p = -0.270\delta_a$ per p, $k_\beta = 4.206\delta_v$ per β, $k_r = -2.974\delta_v$ per r.

The next step is to hold k_p fixed at the value -0.270 and then determine the Model PI solution for matching the system's β/δ_{v_p} transfer characteristics to the model's β_m/δ_{v_p} characteristics with just k_β and k_r as free parameters. The roll rate feedback gain is held constant because it is known to be the least effective of the three parameters in changing the sideslip response to a yaw command. Starting with an initial choice of k_β and k_r given by the final values of the above step resulted in final values for this step of $k_\beta = 4.562\ \delta_v$ per β, $k_r = -3.084\ \delta_v$ per r.

These values together with the previous value for k_p should be a good compromise between matching the two models. To check this, the first step is repeated with k_p free and k_β and k_r fixed at these latter values. The procedure for doing this should be clear from the previous steps. The resulting Model PI design did not change k_p to within three significant figures. Therefore, the final solution is the following: $k_p = -0.270\ \delta_a$ per p, $k_\beta = 4.562\ \delta_v$ per β, $k_r = -3.084\ \delta_v$ per r. The closed-loop transfer functions corresponding to these feedback gains are given in Eqs. (29) and (30).

This design would be predicted to have good handling qualities characteristics based on Fig. 4.

Discussion of Results

The first and most important result is that the Model PI produced a conceptual design for a stability augmentation system that would provide satisfactory to good lateral-directional handling qualities for the X-15 at the flight condition considered. That was the objective of the example. The improvement in the dynamic response of the X-15 with a stability augmentation system is illustrated in Fig. 7 for the design obtained.

It is not possible to compare this design directly to the actual X-15 SAS because the feedback configurations are different. But the roll and yaw rate feedback gain ranges used in the actual X-15 SAS should indicate the general gain levels that would be reasonable to implement. The roll rate feedback gain on the X-15 SAS[12] ranges from 0 to $-0.5\ \delta_a$ per p, and the yaw rate feedback gain ranges from 0 to $-0.3\ \delta_v$ per r. The Model PI design has a roll rate gain of $-0.27\ \delta_a$ per p, which is within the X-15 SAS range and thus is reasonable to implement. The yaw rate feedback gain obtained here, about $-3.1\ \delta_v$ per r, is an order of magnitude greater than the corresponding X-15 SAS range, which indicates that it is probably too high or at least higher than necessary. This is not a fault of the design procedure but rather of the model selected. By referring to Fig. 6, one can see that the Dutch roll mode damping selected for the model was very conservative. A much lower damping ratio, as low as 0.2, could be used for the model and still have satisfactory Dutch roll handling qualities characteristics, providing the other conditions are met. Since the yaw rate gain affects the Dutch roll damping primarily, a lower value would result from the Model PI design if a lower damping ratio were used in the model. Repeating the Model PI design procedure with a less conservative model would be a better approach for obtaining a realistic value for the yaw rate gain than placing a constraint on the magnitude of the yaw rate gain. By using the first design as a guide, one could select a new model with a lower Dutch roll damping that could quite likely be matched very closely by the Model PI design. This would give the designer more control over the resulting locations of the poles and zeros than merely placing a constraint on a free parameter.

The sideslip angle feedback gain resulting from the Model PI approach for the model chosen, which is about $4\ \delta_v$ per β, is probably too high. The X-15 SAS yaw servo is authority limited to $\pm 7.5°$ of vertical stabilizer deflection so that less than 2° of sideslip would saturate the yaw servo with this gain. Using a less conservative model as discussed previously would also result in a lower sideslip gain.

This discussion of the practical limitations that might arise in implementing the conceptual design is included to emphasize the synthesis nature of the Model PI method. In complex problems, one should plan on several design iterations to obtain an acceptable, practical design. This is also true with other analytical design methods. For example, the designs in Refs. 5 and 11 for the same problem considered here also have feedback gains that would be too high for practical implementation. To obtain practical designs one would have to repeat the procedures used in the particular reference with an appropriate modification. In the linear optimal control approach used in Ref. 11, the resulting feedback design is much more complicated than necessary, and the designer would want to simplify it, which involves additional synthesis. Although these analytical techniques do not provide an automatic design, they are much simpler to use than conventional techniques on complex design problems such as this example. Anyone familiar with conventional techniques can appreciate the tedious effort that would be involved in designing this multivariable system by root locus, Nyquist, or Bode techniques.

Conclusions

This paper presented a new performance index, the Model PI, that brings engineering design specifications into the analytical design process. Application to flight control systems was emphasized, although the techniques apply to linear, time invariant, deterministic systems in general. The basic form of the Model PI is the same as that of quadratic functionals frequently appearing in modern control theory. The important difference is the ability to interpret the weighting matrix of the Model PI directly in terms of a model that relates to engineering specifications. A parameter optimization design procedure was established that starts with practical engineering specifications and uses the Model PI as a synthesis tool to obtain a satisfactory design.

The Model PI is different from the familiar model-referenced integral squared error (ISE) performance index. It can be used effectively in designing practical control systems, and it is substantially more efficient to use than a comparable model-referenced ISE performance index.

References

[1] Rediess, H. A., "A New Model Performance Index for the Engineering Design of Control Systems," Ph.D. thesis, MIT, Dec. 9, 1968.

[2] Newton, G. C., Jr., Gould, I. A., and Kaiser, J. F., *Analytical Design of Linear Feedback Control*, Wiley, New York, 1957, Chap. 2.

[3] Roberts, J. D., "A Method of Optimizing Adjustable Parameters in a Control System," *Proceedings of the Institution of Electrical Engineers*, Paper 4000M, Nov. 1962, pp. 519–528.

[4] Binulac, S. and Koktovic, P., "Automatic Optimization of Linear Feedback Control Systems on an Analog Computer," *Annales de l'Association Internationale pour le Calcul Analogique*, No. 1, Jan. 1965, pp. 12–17.

[5] Whitaker, H. P. and Potter, J. E., "Optimization of the Use of Automatic Flight Control Systems for Manned Aircraft," Rept. R-558, Aug. 1966, MIT Instrumentation Lab., Cambridge, Mass.

[6] Aizerman, M. A., *Lectures on the Theory of Automatic Control*, 2nd ed., Gostekizdat, Moscow, USSR, 1958, pp. 302–320.

[7] Rekasius, Z. U., "A General Performance Index for Analytical Design of Control Systems," *IRE Transactions on Automatic Control*, Vol. AC-6, No. 2, May 1961, pp. 217–222.

[8] Kalman, R. E., "When is a Linear System Optimal?" *Transactions of the ASME, Ser. D: Journal of Basic Engineering*, Vol. 86, March 1964, pp. 51–60.

[9] Schultz, D. G. and Melsa, J. L., *State Functions and Linear Control Systems*, McGraw-Hill, New York, 1967, Chap. 8.

[10] Kalman, R. E. and Koepcke, R. W., "Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indices," *Transactions of the ASME*, Vol. 80, Nov. 1958, pp. 1820–1826.

[11] Rynaski, E. G., Reynolds, P. A., and Shed, W. H., "Design of Linear Flight Control Systems Using Optimal Control Theory," ASD-TDR-63-376, April 1964, Wright-Patterson Air Force Base, Dayton, Ohio.

[12] Tremant, R. A., "Operational Experiences and Characteristics of the X-15 Flight Control System," TN D-1402, Dec. 1962, NASA.