(NASA-CR-143844) DEVELOPMENT OF A PARAMETER OPTIMIZATION TECHNIQUE FOR THE DESIGN OF AUTOMATIC CONTROL SYSTEMS Final Report (Massachusetts Inst. of Tech.) 125 p HC A06/MF A01 CSCL 01C G3/08

N77-27126

Unclas 35555

NASA CR-143844

DEVELOPMENT OF A PARAMETER OPTIMIZATION TECHNIQUE FOR THE DESIGN OF AUTOMATIC CONTROL SYSTEMS

H. Philip Whitaker

Measurement Systems Laboratory
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

May 1977

Prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Dryden Flight Research Center
Edwards, California 93523

1. Report No. NASA CR-143844	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DEVELOPMENT OF A PARAMETER OPTIMIZATION TECHNIQUE FOR THE DESIGN OF AUTOMATIC CONTROL SYSTEMS		5. Report Date May 1977
		6. Performing Organization Code
7. Author(s) H. Philip Whitaker		8. Performing Organization Report No. H-978
		10. Work Unit No. 505-06-91
9. Performing Organization Name and Address Measurement Systems Laboratory Department of Aeronautics and Astronautics Massachusetts Institute of Technology Cambridge, Mass. 02139		11. Contract or Grant No. NGL-22-009-548
		13. Type of Report and Period Covered Contractor Report - Final
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code

15. Supplementary Notes

NASA Technical Monitor: Kenneth J. Szalai, Dryden Flight Research Center

16. Abstract

This report describes parameter optimization techniques for the design of linear automatic control systems that are applicable to both continuous and digital systems. The model performance index is used as the optimization criterion because of the physical insight that can be attached to it. It is shown that satisfactory design of representative flight vehicle control systems can be achieved by using these techniques. These include aircraft longitudinal and lateral automatic flight control systems and a first stage launch booster attitude control system with bending mode coupling.

The design emphasis is to start with the simplest system configuration that experience indicates would be practical. Design parameters are specified, and a digital computer program is used to select that set of parameter values which minimizes the performance index. The resulting design is examined, and complexity, through the use of more complex information processing or more feedback paths, is added only if performance fails to meet operational specifications. System performance specifications are assumed to be such that the desired step function time response of the system can be inferred.

In essence, the manual effort associated with an engineering design iteration is transferred to the computer in a fashion that permits the engineer to retain the constraints imposed by practical implementation considerations. In addition to reducing the effort associated with the design of complicated configurations, solutions to the design problems can emerge which might otherwise be overlooked, not anticipated, or not explored for lack of time. The design iteration sequence thus proceeds from the simplest configuration to the configuration of minimum complexity capable of meeting the operational performance specifications.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Automatic control Flight control Digital control Design technique Sampled data system Parameter optimization	Unclassified - Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	125	

*For sale by the National Technical Information Service, Springfield, Virginia 22161

TABLE OF CONTENTS

LIST OF SYMBOLS

Note: Vector quantities are denoted by the underline. Superscript T denotes transpose.

a_i	coefficient of the ith derivative term of the system characteristic equation
$\underline{\tilde{a}}$	augmented coefficient vector = $[a_0, a_1, \ldots, a_{n-1}, 1]^T$
a_z	aircraft z-axis linear acceleration
A,B,C,D	state equation matrices
b_i	coefficient of the ith order term of the numerator polynomial of the system transfer function
c	general matrix constant element
C*	linear combination of pitch rate and normal acceleration used as an aircraft handling qualities criterion; units - g's.
D(p)	denominator polynomial of a transfer function
g	acceleration of gravity
$G_{c[q_i q_o]}$	transfer function of component, c, relating the input quantity, q_i, to the output quantity, q_o; where no ambiguity result the subscripts are omitted
h	dummy integration variable
i(t)	error excitation function; that signal which, if applied to a component having the transfer function of the model, would cause the output of that component to be proportional to the error in the zero order states of the system and model
I	identity matrix
i,j,k,l, m,n	integer indices
N(p)	numerator polynomial of a transfer function
p	Laplace operator
P_i	i-th parameter

PILLIM	lower limit for the integral of the performance index
q	integer sample index
q_1	first bending mode displacement
R	residue coefficient in a partial fraction expansion
$S_{c[q_i q_o]}$	static sensitivity of component, c, relating an incremental change in the output, Δq_o, to an incremental change in the input, Δq_i, in steady-state; it is defined for that steady-state output derivative which is finite and non-zero for a step input
t	time
T	sample period
T_{ave}	time interval for normalizing the performance index
u	generalized input quantity
V	airplane airspeed
W	weighting matrix
W_x	roll angular velocity
W_z	yaw angular velocity
w	component of linear velocity along aircraft z-axis
$\underline{x}$	vector of state variables
$\underline{x}_0$	vector of initial values of the state variables
$\Delta\underline{x}_0$	vector of deviations between initial values of state variables of the system and model of order $(\ell+1)$, where ℓ is order of the model
α_0	airplane angle of attack
α_i	coefficient of the i-th derivative term of the model's characteristic equation
$\underline{\tilde{\alpha}}$	n-th order augmented coefficient vector = $[\alpha_0,\alpha_1,\ldots,\alpha_{\ell-1},1,0,\ldots]$ where n is order of the system, and ℓ is order of the model.
β_i	coefficient of the i-th order term of the numerator polynomial of the model's transfer function

δ	control deflection
Γ	convolution integral
λ_g	slope of the bending mode at the gyro location
Φ	state transition matrix
ϕ_{ij}	ij-th element of the state transition matrix
ϕ	Euler angle of roll; rotation about the x-body axis
τ	dummy integration variable
$\dot{\theta}$	airplane pitch rate
ω	frequency, (rad./sec.)
ω_n	undamped natural frequency of a second order component (rad./sec.)
ζ	damping ratio of a second order component

Subscripts

c	continuous elements
d	discrete elements
m	model
o	initial value
s	system

1. INTRODUCTION

1.1 General Objective

This report describes a design technique which, it is hoped, will be useful in the design of automatic control systems. It has been motivated by an interest in flight vehicle control systems, but is not limited to such design situations. The philosophical viewpoint that underlies the development is one that postulates that the best design is the simplest one that is capable of meeting the requirements of the operational use of the vehicle or other controlled member. The connotations of the term "simple" include the engineering considerations of cost, weight, reliability, maintenance, and whatever other factors that influence the decision of accepting a given design. The techniques being described here however, deal only with the system design aspects that synthesize a system configuration to the level of detail needed to meet static and dynamic performance specifications. As such, these are analytical techniques, and the definition of simplicity is then taken to be the use of the fewest sensors and feedback paths with practical constraints upon design parameter values, the least complicated information processing, and/or the use of sensors or other equipment which may be aboard the vehicle because of unrelated vehicle operational requirements. Although analytical optimization methods are employed, the concept of an optimum system only has validity as referring to whatever system the purchaser is willing to buy, and not to a system that is defined by a mathematical functional.

There is nothing startlingly new about such a concept. Indeed any successful operational system must have been designed under conceptionally similar ground rules. Hence there is some reservation about engaging in an activity which may only be adding still another technical report to the shelves of unread literature. But it may be that while some individual organizations may have developed similar design procedures in-house, their availability in the open literature may be severely restricted, and others may find features of the work described here to be adaptable to their own needs.

1.2 The Design Procedure

The design procedure is summarized more precisely in the following manner. It is assumed that the operational mission requirements specify the static and dynamic performance required for the automatic control system in such a manner that the response of the closed-loop system to a step input can be inferred. It is expected that such a specification will also give a tolerance boundary defining at least the minimum acceptable level of performance, and it may include a desired performance level together with acceptable deviations. The designer then decides upon the simplest design configuration that previous experience or practical considerations suggest can provide the control desired. If a closed-loop control system is desired, this will in general require at least one sensor and a means of combining signals and feeding an actuation signal to a control effector of some kind. One also denotes those design parameters which need to be numerically specified (e.g. an open-loop gain). The design method then determines the best performance achievable with this configuration and provides insight into what additional complexity should be provided in the next design iteration if the best performance of a given configuration fails to satisfy the operational requirements. Practical constraints upon the range of values of the design parameter are included.

The design is iterated until the necessary sufficient level of complexity is reached at which the performance specifications are satisfied.

The reader will recognize such a process as one of parameter optimization of a fixed configuration system. So what is new? Realizing the futility of claiming that one philosophical approach to design is in some sense better than another, one may be limited to answering that question by stating that the reader of the report may hope to find:

(a) a description of the reduction to practice of the parameter optimization technique for both continuous and digital control systems,

(b) illustration of the techniques by typical flight control system design examples,

(c) the use of a performance index with certain features that are easily relatable to physical interpretation and which has been found to result in satisfactory design of flight control systems

(d) an emphasis upon using in conjunction with the computer optimization program the useful classical control techniques to provide insight and interpretation of optimization design results some of which at times can be totally unexpected.

1.3 Relation to State Feedback Control

It is well known that one can design a linear feedback control system using the results of optimal control theory. The theoretical solution to the general control problem is the use of complete state vector feedback. If all the states are not directly measurable, various state estimators have been proposed to obtain signal representations of the missing states. While such a design procedure will result in control systems that have acceptable performance, the technique inherently results in the most complicated system configuration. For example, the control system for a flight vehicle which considered an 18th-order representation for the vehicle with 4 control effectors ended up with 72 feedback paths. One usually cannot accept such complexity, and although various designers have developed their own procedures and rules of thumb, there is little in the way of a straight-forward method for reducing the complexity. Performance indexes are also used, but there tends to be an arbitrariness about the choice of the performance index, and this re-introduces the trial and error nature to the design process which the theory had hoped to avoid.

This report does not intend to belittle the optimal control approach. Rather it emphasizes the alternative approach of starting with the simple and evolving toward the minimum complexity system, instead of starting from the most complex and having to reduce it to a practical embodiment. In the hands of capable designers either approach should be successful, and one cannot argue for one over the other except upon philosophical grounds. It was hoped that direct comparison designs could be investigated. Unfortunately those who have used optimal control to design systems which are more than trivial examples seldom report their results in the available literature in sufficient detail that the required system description data can be obtained (and without typographical errors). Thus, this report lacks such comparisons to its detriment.

1.4 Organization of the Report

The primary effort under this program has been the development of the required digital computer programming and its check-out through application to typical flight control design examples. It is only through extensive applications that one can uncover the subtle programming errors which exist in any complicated program. Correction of such "bugs" is often a time consuming process in itself. As was true in this case, this may necessitate taking entirely different approaches than were originally taken. All of that history of experience has not been reviewed in this report, and the programming described pertains only to the final version.

Because of its key importance a discussion of the performance index is presented in Section 2 together with the evolution of its present interpretation. The evaluation of the performance index and a summary description of the computer program is presented in Section 3. Typical design applications are presented in Section 4. A general summary and recommendations are presented in Section 5.

2. THE MODEL PERFORMANCE INDEX

In any parameter optimization process a criterion must be established that defines the optimum parameter set, and such terminology implies that there is one desired parameter set which is better than all others. In the design of automatic control systems there is no such thing as the optimum system which can be specified in any quantitative fashion. Rather, there are many designs which provide acceptable levels of the various capabilities of interest, and the choice between competing designs is made on the basis of many technical and nontechnical trade-offs. Even in the area of those performance characteristics which can be quantitatively determined, there is a tolerance band defining acceptable performance, which means that many different parameter sets are equally good.

To take advantage of the computational capacity of the digital computer however, it is convenient to establish a mathematical criterion of optimality. One such device is called the Performance Index and is customarily taken to be a quadratic functional. In using a performance index, one needs to establish somehow that when the performance index criterion is satisfied the system operational performance will meet its specifications. In light of the foregoing remarks, there is no exact transformation of operating specifications into the mathematical functional that one should use. Thus one can only try likely looking performance indices and examine the resulting designs for acceptability.

It was for the purpose of suggesting a performance index that would have a greater probability of leading to successful engineering designs that led

Rediess (reference 1) to formulate the Model Performance Index. This index has the advantage that under certain constraints placed upon the manner of specifying acceptable performance, the mathematical form of the performance index is established directly by the specifications of desired performance. This is a great advantage, and the use of the index has been shown to lead to acceptable system designs. It must be kept in mind however that it is only a design tool that serves as a means of expediting the overall design procedure, and one must be prepared to make design iterations so that the final design meets all of the design objectives.

Indeed, several modifications to Rediess's performance index have been made during the course of this investigation. These were made primarily to expedite the computational process, and they do not change the basic concept. This section of the report will define the Model Performance Index, discuss the modifications made and the reasons for them, and summarize the general features of its use.

2.1 Definition of the Model Performance Index

The Model Performance Index was proposed by Rediess in References 1 and 2 as a means of introducing engineering specifications into a mathematical optimization process of selecting system design parameters. He defined it as

$$(PI) = \int_0^{\infty} (\underline{\tilde{x}}^T \underline{\tilde{\alpha}}\ \underline{\tilde{\alpha}}^T \underline{\tilde{x}})\, dt + r(\Delta\underline{\tilde{x}}_0^T\ W\ \Delta\underline{\tilde{x}}_0) \qquad (2\text{-}1)$$

where

$\underline{\tilde{x}}$ is the augmented state vector representing the system's transient response; order (n+1), where n is order of the system

$\underline{\tilde{\alpha}}$ is the vector of coefficients of the characteristic equation of a model augmented by zero elements to the order (n+1)

r an arbitrary weighting factor

$\Delta\underline{\tilde{x}}_0$ is the vector of differences in state variables at zero time

W is an arbitrary weighting matrix which Rediess proposed to weight only the ℓ lower states, where ℓ is the order of the model.

The system's transient response, given by the lowest order state variable, x_1, is defined as the difference between the system's step function response and the steady state value of that response. The higher order states are the successive derivatives of the transient response up to and including the nth derivative, where n is the order of the system. Rediess showed that if one augments the system's state vector by adding the nth derivative as an additional state, a system's transient response as a function of time can be viewed geometrically as a trajectory lying within an (n+1)-dimension hyperplane in state space. This hyperplane is perpendicular to an (n+1)-dimension vector, $\underline{\tilde{a}}$, whose components are the n coefficients of the system's characteristic equation augmented so that the (n+1) component is 1.0. This merely recognizes that the step function transient response is given by the autonomous equation

$$\underline{\tilde{x}}^T\underline{\tilde{a}} = 0$$

when appropriate initial values for the states, $\underline{\tilde{x}}_0$, are specified. These initial values then include the effects of the step input and the zeroes, if any, of the system's transfer function. The trajectory plane is called the characteristic plane of the system.

The assumption is then made that the operational requirements for the control system can be transformed into a specification upon the dominant features of the step function response of the system and therefore of its unit impulse response. The Laplace transform of that unit impulse response thereby defines a transfer function of an equivalent system which would meet the operational requirements. This equivalent system can be called a "model for the system". Viewed itself as a dynamic system, the model's transient response is represented by the autonomous equation

$$\underset{\sim}{x}^T \underset{m}{\tilde{\alpha}} = 0$$

That response can be considered to be a trajectory in state space lying within the hyperplane that is perpendicular to the augmented coefficient vector of the model. That plane is the model's characteristic plane. Figure 2-1, taken from Ref. 2, illustrates these geometrical concepts for a second-order system and model.

The orientation of the system's characteristic trajectory plane is determined by its coefficient vector, which in turn is a function of the system design parameters. Thus the process of selecting a set of design parameter values is equivalent to orienting the system's characteristic plane in state space through modification of the components of the coefficient vector.

The second term in equation (2-1) includes the effect of the initial mismatch between system and model trajectories in state apace. For systems and models that have no zeroes, the initial values of the lowest order transient states of system and model will be -1.0, and the initial values of all higner order states will be zero. The initial difference in states, $\Delta\underset{\sim}{x}_0$ will then be identically zero, and minimizing the model performance index will involve only the integral term of equation (2-1).

The integrand in the integral term of equation (2-1) is the square of

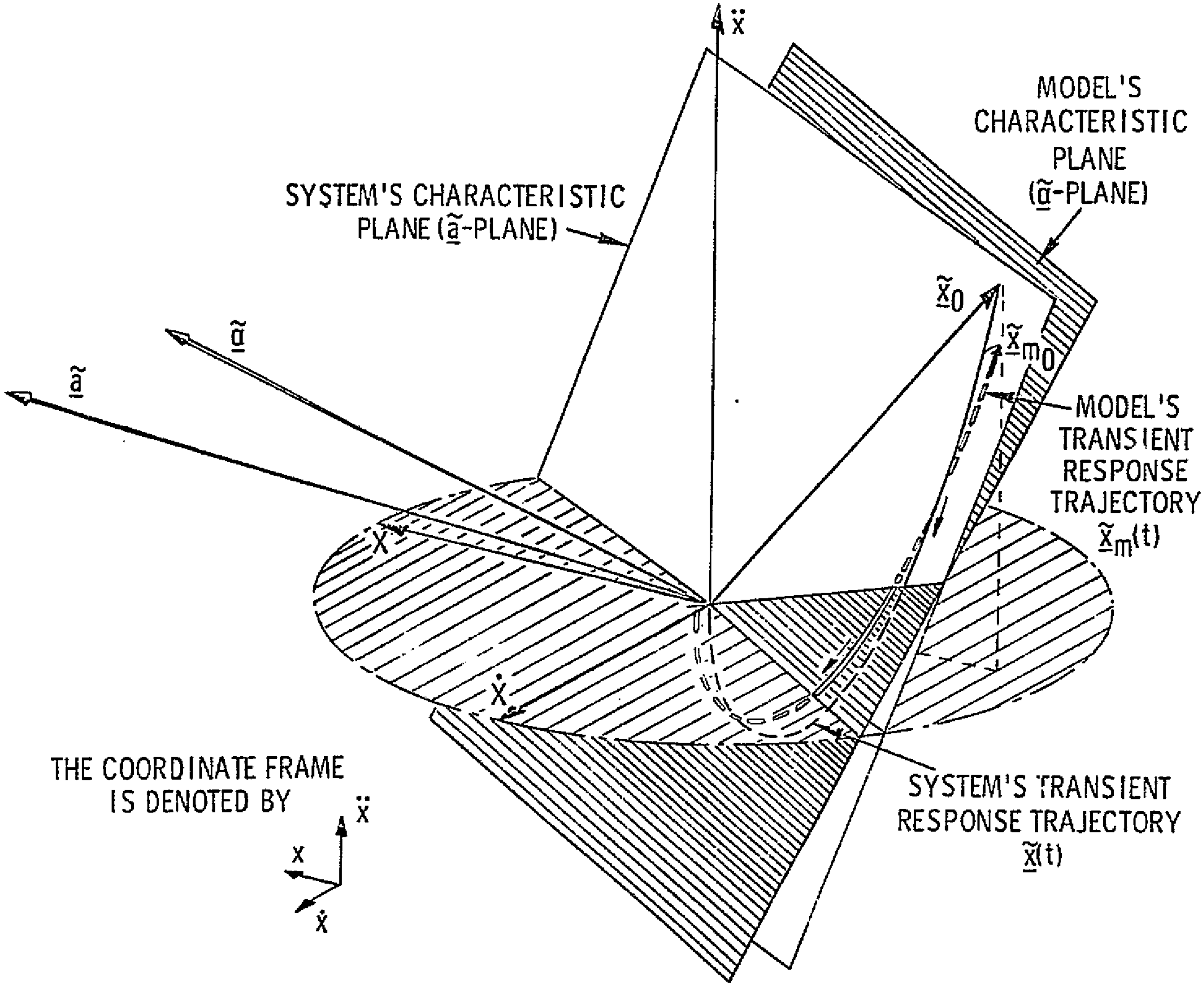


Fig. 2-1. Geometrical Representation of Approximating One Dynamical System by Another (taken from Ref. 1)

a linear combination of the state variables. The weights given to the various state variables are the coefficients of the characteristic equation of the reference model and are thus numerically specified once the model has been designed from the operational specifications. If the system and model had the same order, it is seen that if the coefficients of the system's characteristic equation were made equal to the model coefficients the integrand would become zero and the system and model would have identical transient responses. The integrand is the square of the quantity, $(\underline{\tilde{x}}^T\underline{\tilde{\alpha}})$, which geometrically is the projection distance of the system's trajectory onto the model's characteristic plane. Therefore minimizing (PI) is a process of orienting the system's characteristic plane so that the time averaged square value of the projection distance of the system's trajectory onto the model's characteristic plane is minimized.

If the model is of lower order than the system as is usually the case, minimizing (PI) tends to cause the lower order system states, up to the order of the model plus one, to approximate those of the model. In this case there will still be a mismatch of the trajectories at the initial time, and this may be large. If sufficient design freedom is available, the time duration of the mismatch can be made arbitrarily small, and the contribution of this initial portion of the phase trajectories to the (PI) will be neglibible. With only limited design freedom, the initial mismatch may bias the optimization procedure, and this effect could be a factor in one's choice of an acceptable model. Note that the initial response of the system or model is governed by its high frequency behavior.

When either or both of the system and model transfer functions have zeroes, the initial starting locations of the transient response vectors (or trajectory) in state space will differ even if the two characteristic planes are coincident. The added term in the (PI) in equation (2-1) permits one to trade-off the matching of the trajectory at the initial time and the projection distance in state space.

2.2 Importance of the Model

In the parameter optimization technique utilizing the Model Performance Index, the model plays the role of the specification of the desired system dynamic performance. It is assumed that one can translate the operational requirements for the system into a statement of the dominant characteristics that the step function response of an acceptable operating system would have. It is not necessary that the system should receive step function inputs in its operating environment. Since one is considering linear systems, it is also not necessary to start with the step function response characteristics, inasmuch as one can obtain the step function response from other mathematical representations, e.g., the frequency response or impulse response, by mathematical operations. One is, however, considering control systems for which one is expecting the steady-state output to be proportional to the input. That is, the input is looked upon as the desired output. The transient response of a system whose steady-state output would be zero is excluded in the present embodiment of the model performance index technique. (Note that this is merely a constraint of the present computer program which could be removed if desired).

Since the model in effect is an alternative way of stating the system specifications, its design should be independent of the system design choices. As always, the specifications should include all those dynamic characteristics that are important to operational success, but no more than those. Therefore if some particular dynamic effect is wanted, the model design must reflect that, but otherwise the model should be kept as simple as possible. In complicated systems the important dynamic interactions are not always known ahead of time, so the previous statements do not preclude the possibility of making model modifications as the design iteration cycle produces more information and understanding. This may be particularly true in the placement of model zeroes whose dynamics effects are somewhat more difficult to estimate than are those of poles.

The model cannot be selected arbitrarily. To do so is to sidestep the most important part of the design effort, the specification process. In particular it makes no sense to require the system to perform beyond its capabilities. For example, if there are known limitations to the amount of control power available which in turn limit the possible output acceleration, the model design should reflect that. If that is done, the problem of weighting the control effort in minimizing a performance index does not arise. The model design inherently provides the desired weighting. This can be seen from the optimal control solution leading to complete state vector feedback. If one can specify the locations of all the system eigenvalues, he then has sufficient information to determine the control effort involved for that desired system, and he can adjust the desired eigenvalue locations so as to avoid exceeding the specified maximum control effort. The state feedback solution will guarantee that the system responds in the same way, and hence use no more than the desired amount of control effort.

The model time response, including the time history of those output derivatives of importance, should be obtained and examined prior to the parameter optimization. This is particularly true if several poles and zeroes are involved, for one's intuitive feel for response characteristics deteriorates rapidly as the orders of the numerator and denominator of the transfer function increase.

In general if one is attempting to control the characteristics of several system modes, one needs to include those modes in the model together with other dominant response modes. A case in point arises in the case of using integral compensation for reducing static error. If one feeds the error signal to an integrator to form a parallel feedforward compensation as in Figure 2-2, the closed-loop transfer function will exhibit a corresponding pole and zero

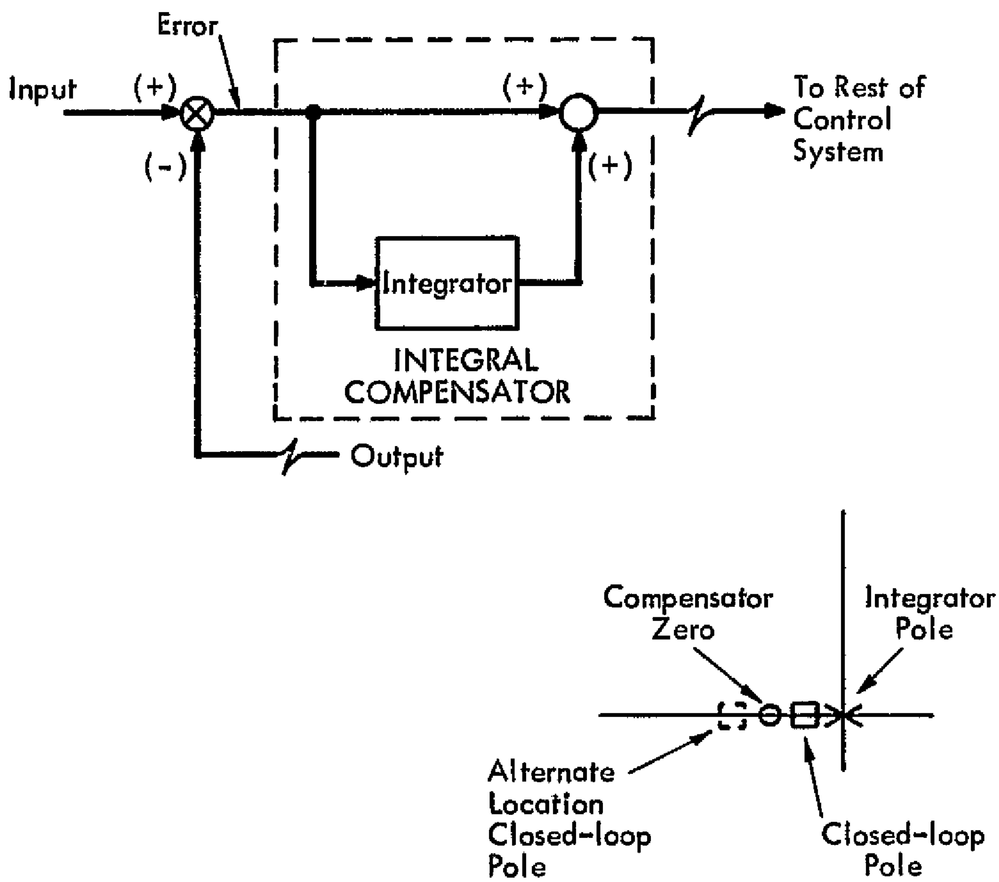


Fig. 2-2. Considerations of Integral Compensation in the Model Specification

at low frequency. The static error reduction takes place dominantly with the dynamic lag associated with this low frequency pole. One can usually design the integral compensation directly on the basis of the desired time available for reducing the static error to an acceptable level. If the uncompensated static error is large, the low frequency mode will also be very evident in the output response as a slow exponential approach to the final value. Assuming that such behavior is satisfactory under the operating environment, the low frequency mode should appear in the model. If that mode is omitted from the model, the parameter optimization process will use the design freedom available to cause the system to follow what in effect is a faster model. If the gain of the integrator is a design parameter, the optimization may use the gain to place the zero associated with the compensator so as to help provide high frequency response. This may give apparently different results than one intended and may or may not be satisfactory. An analogous situation is described in the illustrative design example of section 4.2.

In the case of digital systems, the program requires the Z-transform of the model. In general the Z-transform for both system and model will have a pole-zero excess of one corresponding to one sample delay. Thus $(\ell-k) = (n-m)$. Z-plane zeroes are more difficult to interpret than p-plane zeroes, and so it is even more necessary to obtain the transient model response to verify that the desired model is being used. A further caution is needed arising from the fact that the Z-transform can introduce zeroes on the negative real axis which have no direct counterpart in the p-plane. Since model zeroes will appear in the performance index as poles of $i(z)$, a model zero near $z = -1$ will have the unwanted effect of a low frequency pole in its contribution to $[i(t)]^2$. In such cases moving the zero near the origin on the positive real axis may have a negligible effect upon the model time response and avoid the difficulties in the performance index calculation. (See example 4 in section 4.3.) Finally, it is to be noted that there are several methods for obtaining a sample data representation for a continuous component (see References 4 and 5).

A further restriction upon the selection of the model is placed by the method of evaluating the performance index as will be discussed in section 3.1. Except under special circumstances, model zeroes in the right half p-plane or outside the unit circle in the z-plane are to be avoided.

2.3 Modifications to the Model Performance Index

As has been alluded to in the introduction of section 2, subsequent investigations have made minor modifications to the definition of the Model Performance Index. Palsson (Ref. 3) was motivated by a desire to remove the arbitrariness associated with the weighting factor, r, of the initial state term of equation (2-1). In so doing he derived an equivalent formulation of the performance index which has led to some computational advantages. But the effect of the initial states was still not adequately taken care of, and additional modification has been made during this investigation to accomplish this while retaining the Palsson formulation for its computational advantages. It is to be emphasized that these modifications, which are described in the following sections, do not make fundamental changes in the original concept of the Model Performance Index.

The advantage of using the Model Performance Index of Rediess is that it automatically provides the correct weighting factors on the states to be used in a quadratic performance index. These weighting factors are the coefficients of the model's characteristic equation. Even without zeroes and with the best matching, the initial state vectors differ when the order of the model is not the same as that of the system. Using phase variables as state variables is analogous to using an output quantity and its time derivatives. If the model is of lower order (ℓ) that that of the system (n), the model's ℓ-th state is nonzero and discontinuous at $t = 0$ for a step input, while the system's n-th state is nonzero and discontinuous. Thus one would not expect that the system's ℓ-th state could be made equal to the model's ℓ-th state at $t = 0$ by any design parameter choice.

As has already been noted, the presence of zeroes in the system and the model transfer functions can also cause the initial orientation of the state vector of the system in state space to differ from that of the model. If the model and system numerator orders were k and m respectively, the corresponding relationships then apply to the $(\ell-k)$th and the $(n-m)$th states respectively. Thus, even though the dominant response characteristics of the system and model may be approximately the same, the initial behavior of the derivatives of the two outputs may be quite different, and the differences are accentuated the higher the derivative one examines. If the model is of lower order than the system, projecting the system trajectory onto the model's characteristic plane in effect neglects the higher order states of order greater than $(\ell + 1)$ in the evaluation of the performance index.

When the initial state vectors are not the same, there is an unavoidable initial portion of the transient response during which the system states are being readjusted so that the lower order states approach those of the model. Unless there is some operational requirement that places emphasis upon this initial time behavior, one is usually looking for that parameter set which causes the trajectories of the model and the system to be close to one another in state space over most of the transient response time. To the extent that this occurs, the lower order states of the system will then be related to one another in the same manner as the corresponding states of the model are to each other. If one had sufficient design freedom in selecting parameters, one could obtain a good matching of the lower order states over most of the transient time response except for that initial period of time, which could be made arbitrarily short. The initial time behavior correlates with the high frequency characteristics of the system and model, and making the initial mismatch small in this case is equivalent to matching the low frequency modes while adjusting the system parameters to place the remaining modes at very high frequency where their contributions would be negligible.

Modification by Palsson. The investigation by Palsson (Ref. 3) derived the performance index in a manner that includes the effect of the initial transient state values automatically while retaining only the form of the integral term of equation (2-1) or

$$(PI) = \int_0^{\infty} (\underline{\tilde{x}}^T \underline{\tilde{\alpha}}\ \underline{\tilde{\alpha}}^T \underline{\tilde{x}})\, dt \qquad (2\text{-}2)$$

He showed that if one defines an augmented system which is formed by the original system to which is cascaded a component whose transfer function poles are the same as the zeroes of the reference model, and then uses as the model for the optimization of the augmented system one that has the poles of the reference model but no zeroes, the initial value term of equation (2-1) is identically zero. Since the added component is cascaded to the original system, its presence does not affect the stability or dynamic characteristics of the original system, nor will it contain any design parameters. That such an arrangement could lead to a good match between a system and the reference model can be seen for the case for which the order of the system is the same as that of the model. Assuming sufficient design freedom, for the augmented system to match the re-defined model having no zeroes, the parameters would be selected so that the system zeroes would cancel the poles of the cascaded component, and the system poles would become the same as the model poles. Since the cascaded poles are the real model zeroes, the system transfer function thus becomes the same as the model's, and the system and the original model are identical.

Palsson then went on to show that the output error response between system and model can be generated as the output of a component that has the same transfer function as the model when its input, or excitation function i(t), is given by

$$i(t) = \underline{\tilde{x}}^T \underline{\tilde{\alpha}} \qquad (2\text{-}3)$$

The square of this quantity is the integrand of the Model Performance Index. Therefore the optimization process can be thought of as minimizing

$$(PI) = \int_0^\infty [i(t)]^2 dt \qquad (2\text{-}4)$$

or minimizing the square of the excitation function that generates the error in the lowest order state. Palsson further showed that the Laplace transform of i(t) is

$$i(p) = \left[\frac{G_s(p)}{G_m(p)} - 1\right] u(p) \qquad (2\text{-}5)$$

where G_s is the system's closed loop transfer function

G_m is the model's transfer function.

If $G_s = G_m$, $i(p) = 0$, and there will be no excitation of error generating component and hence no error. In order for $i(t)$ to be well behaved, it is necessary that the ratio (G_s/G_m) not have more zeroes than poles. This in turn restricts the choice of the model so that the excess of poles over zeroes of the model is no greater than that for the system. Equation (2-5) permits one to consider the frequency response characteristics of the error excitation quantity. Equation (2-5) also leads to an alternative method of calculating the performance index as is discussed in Section 3.1.

2.4 Modifications Made During This Investigation

The Initial State Problem. When the available design freedom does not permit one to place the high frequency modes arbitrarily, the contribution of the initial transient behavior can be significant even though it may only persist over a relatively short duration of time. Rediess provided a means of specifying how much weight the designer wished to place upon the initial state effects through an arbitrary weighting matrix. Although simplifying the development somewhat, the approach taken by Palsson automatically placed strong weight upon the effects of the initial states. It has been found in examining design examples for which the closed-loop system transfer function exhibited many zeroes which were sensitive to the choice of design parameters that the Palsson approach did not provide sufficient flexibility. Unsatisfactory designs could result as the minimization of the performance index traded-off the low frequency effects of the dominant features of the response against the high frequency effects associated with the initial state values.

To see this more clearly, one can look at equation (2-4) again for the performance index.

$$(PI) = \int_0^\infty [i(t)]^2 dt$$

and the Laplace transform of i(t), equation (2-5),

$$i(p) = [G_s/G_m - 1]\ u(p)$$

Substituting general polynomials for the numerators and denominators of the transfer functions,

$$i(p) = \left[\underbrace{\left(\frac{b_m p^m + \ldots + b_1 p + b_0}{p^n + \ldots + a_1 p + a_0} \right)}_{\text{system}} \underbrace{\left(\frac{p^\ell + \ldots + \alpha_1 p + \alpha_0}{\beta_k p^k + \ldots + \beta_1 p + \beta_0} \right)}_{\text{model}} - 1 \right] u(p) \tag{2-6}$$

By specifying that the system and model have the same static sensitivity,

$$\frac{b_0}{a_0} = \frac{\beta_0}{\alpha_0} = 1$$

Considering the frequency response of this error excitation function, $i(j\omega)$, it is seen that for it to be zero the first term within the brackets of equation (2-6) should have a magnitude of 1.0 and a phase angle of zero throughout the frequency range. The static sensitivity specification insures this at zero frequency.

When it is not possible that the first term be unity at all frequencies, the fact that i(t) is an input function makes it difficult to determine how one would like to shape its frequency characteristics so that the transient time response will be satisfactory. The difficulty stems from the fact that i(t) is an excitation function which if applied to a component having the transfer function of the model causes the output to be the same as the error response. It is more difficult to specify an excitation function explicitly than an output quantity due to the frequency dependent effects of the component. Since the model transfer function will attenuate the high frequency components of i(t), some high frequency mismatch can be tolerated, but it is found that neither zero weighting of the initial states nor a full weighting results in a system design that is satisfactory, and therefore that some intermediate condition is required.

When $(n-m) = (\ell-k)$, that is, when the model has the same pole-zero excess as the system, the high frequency value of (G_s/G_m) approaches (b_m/β_k) which is the ratio of the root locus gain factors of the system and model. (The root locus gain factor is the ratio of the product of the poles to the product of the zeroes.) With a step input, by the initial value theorem

$$\underset{t \to 0}{i(t)} = \left(\frac{b_m}{\beta_k} - 1\right)$$

If the system has high frequency poles, and the model has low frequency zeroes, (b_m/β_k) can become very large. $[i(t)]_{t \to 0}$ will then be large and exhibit a narrow, high magnitude pulse at the initial time. Minimizing the performance index will then reduce the high frequency magnitude of $i(j\omega)$ usually at the expense of poorer low frequency behavior, and the performance index may have very large values even at its minimum point.

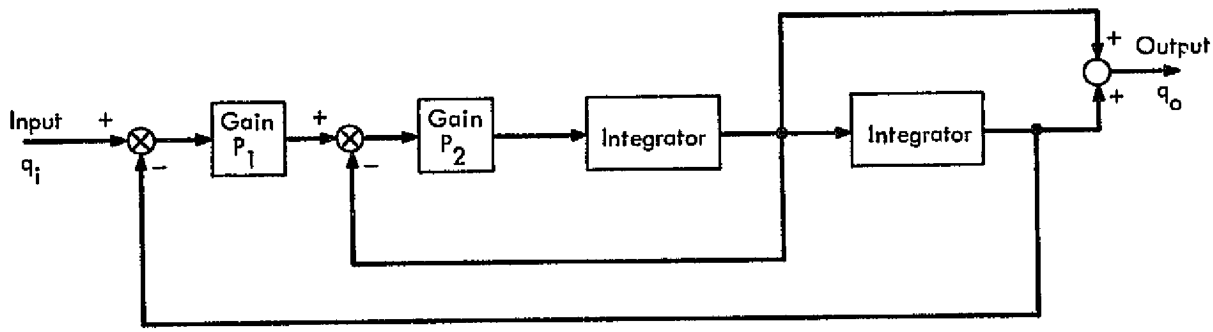
When the system's pole-zero excess is greater than that of the model, i(t) initially starts at a value of -1, but a high frequency mismatch then appears as large positive and/or negative swings of i(t) over the initial transient time, and these can contribute significantly to the performance index. Note that it is the high order states that contribute the large magnitude terms to i(t). If the pole-zero excess of the system is much greater than that of the model, projecting the system trajectory onto the model's plane effectively discards these bothersome states. When the operational requirements permit one to use such a model, the initial states contribute little to the performance index, and the Palsson approach is perfectly satisfactory.

If one were to use a model having a larger pole-zero excess than the system, $(\ell-k) > (n-m)$, equation (2-6) indicates that i(t) would then exhibit impulses. With $(\ell-k) = (n-m)$, the pulse response associated with the initial value of i(t) is approaching that condition. For similar reasons Palsson constrained the model choice so that $(\ell-k) \leq (n-m)$.

To summarize this discussion, one is forced to conclude that if the effect of the initial states is large, a degree of arbitrariness is introduced into the design procedure which is unavoidable. Those cases are those for which the operational requirements are such that $(\ell-k)$ is close to $(n-m)$, particularly those cases in which the system zeroes are strongly affected by the design parameters. Palsson's modification is therefore not satisfactory in general.

Rediess approached the problem of the effects of the initial sta e mismatch by adding an arbitrary weighting of the initial error states to the integral performance index (see equation (2-1)). The weighting factor to be used was obtained by trial and error upon examination of the system performance. The Palsson viewpoint has several conceptual advantages that facilitate the evaluation of the performance index. Hence it was desirable to modify the Palsson approach to achieve a similar means of weighting the initial state effects. Three such modifications have been examined. The first takes advantage of the fact that the higher order derivatives are neglected when $(n-m) > (\ell-k)$. The second considers modifications to $i(j\omega)$ in the frequency domain, and the third considers modification to $i(t)$ in the time domain. The first two were developed as a result of examining a simple design example, while the third perhaps better lends itself to more complicated design situations.

The example will be introduced at this point to motivate the remaining discussion. Consider the simple example shown in Figure 2-3. The system is



Closed-loop Transfer Function: $G(p) = \dfrac{(1+p)}{1 + \dfrac{2\zeta}{\omega_n}p + \dfrac{p^2}{\omega_n^2}}$

$2\zeta\omega_n = P_2 \qquad \omega_n^2 = P_1P_2$

Fig. 2-3. Simple Second-Order System Example

second order, and the closed-loop transfer function has a zero at -1.0 rad/sec. Assume that the only design freedom is the choice of the two gains, P_1 and P_2, so that the parameter choice can change the system closed-loop poles, but the closed-loop zero remains unaltered. Further consider that the system requirements state that the closed-loop step function response should approximate that of a second-order model with no zeroes, a damping ratio of 0.7, and undamped natural frequency of 1.0 rad/sec. Since the model has no zero, a perfect match cannot be expected. However, it is not unreasonable to investigate what parameter choice would best approximate such a response and thereby also determine how large a deviation between system and model would occur under the best design conditions. As this situation has been posed, the model has a pole excess of two, while the system only has a pole excess of one. The requirement that $(\ell-k) \leq (n-m)$ is not satisfied, and the performance index (Equation (2-4)) cannot be evaluated. It would appear at first glance that a zero added to the model at -100 rad/sec would permit the pole excess requirement to be met with only a negligible effect upon the model step response. A check of the step response of such a model indeed showed that at no time did the response differ by more than 0.5% of the steady-state value from that of the original model which is shown in Figure 2-4, and hence to the scale of that plot it would be barely distinguishable from the model response shown. At t = 0 the velocity would be 0.01 instead of zero. Using such a model the error excitation function, i(p), became

$$i(p) = \left\{ \frac{\omega_n^2 (p+1)}{(p^2 + 2\zeta\omega_n p + \omega_n^2} \cdot \frac{100(p^2 + 1.4p + 1)}{(p + 100)} - 1 \right\} \frac{1}{p} \tag{2-7}$$

where $\omega_n^2 = P_1 P_2$

$2\zeta\omega_n = P_2$

Applying the initial value theorem,

$$i(0) = 100\, \omega_n^2 - 1 \tag{2-8}$$

The factor of 100 was introduced by the "negligible" zero, and even though

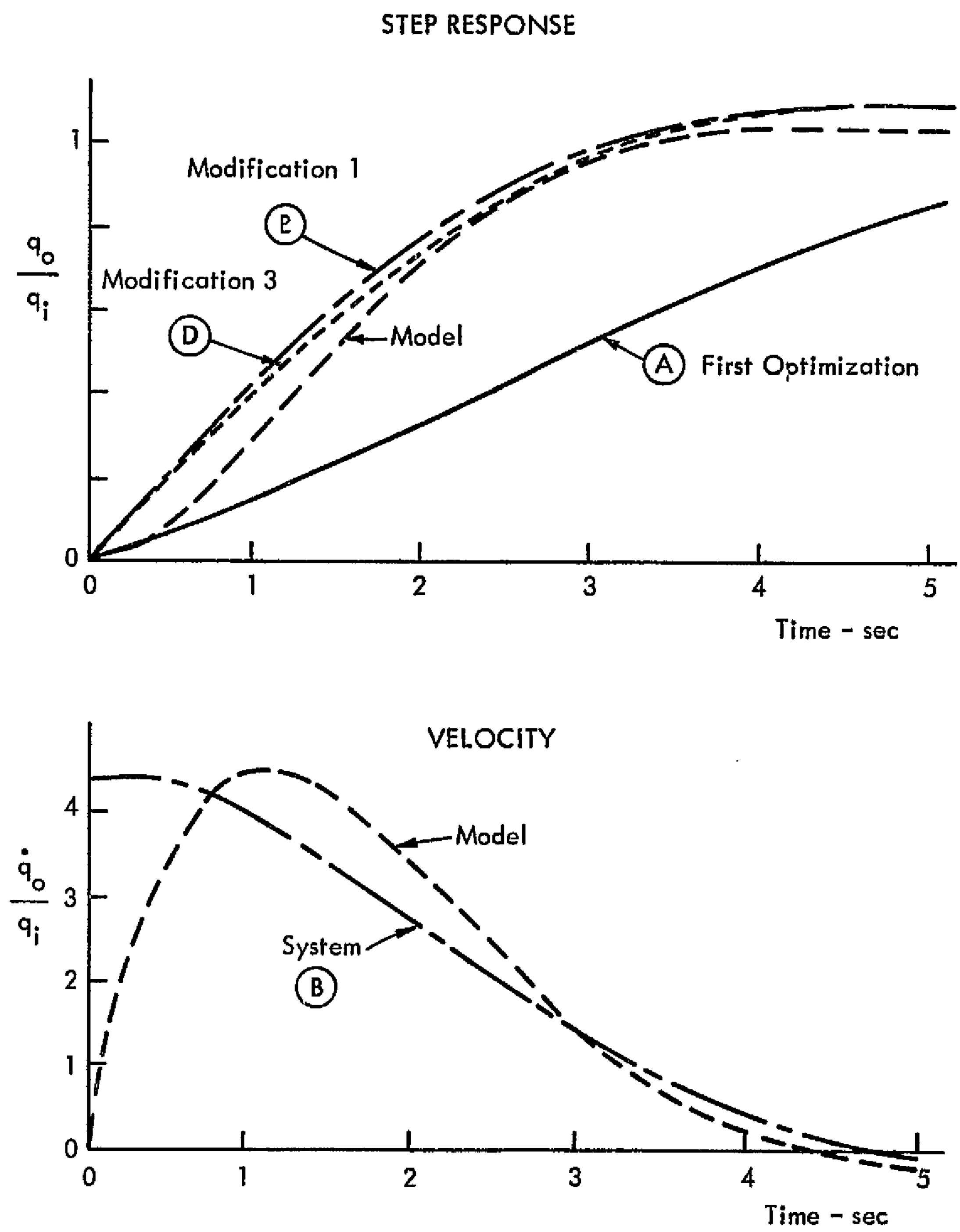


Fig. 2-4. Step Function Responses for System of Figure 2-3

this initial value would decay rapidly with the dynamics associated with the pole at -100 rad/sec, its contribution to the integral of the square of i(t) was significant. The minimum performance index using equation (2-4) occurred with $P_1 = 0.119$ and $P_2 = 0.382$, and its value was large. The system's step response for these optimum parameter values is also plotted on Figure 2-4 as curve A . The average error between system and model was of the order of 30% over the response time of the system and occurred predominantly as a low frequency, sluggish response. The system response time was approximately 15 seconds, and the initial value of the error excitation function was 11.

The example illustrates the difficulty introduced by the initial mismatch of the system and model state vectors. The optimized system displayed a closer match of the initial state vector at the expense of poorer subsequent matching of the state trajectory. In examining these results it could be argued that the difficulties arose from the introduction of a zero to the model in order to satisfy arbitrarily the pole-zero excess requirement. With $(\ell-k) = 1$, the zeroeth-, first-, and second-order states contribute to the performance index, and the zero at -100 caused the latter state to have a large initial value.

Modification 1. A modified procedure was suggested at this point to avoid using the arbitrary zero and still in effect reduce $(\ell-k)$ of the model. (Changing the model to a first order model was ruled out on the grounds that the model reflected requirements for the operational use of the system, and one therefore could not change its dominant characteristics.) Examining equation (2-5), it is seen that it makes no difference mathematically which transfer function is associated with the model and which with the system. One could imagine that the system is fixed and that one is adjusting parameters in the model so that the state trajectories match. Thus one can evaluate the performance index keeping $(\ell-k) \leq (n-m)$ by merely interchanging poles and zeroes of the two transfer functions. Using the pole-zero excess as a test, it is easy to instruct the program to make such an interchange. Doing so and using the original model, the program interchanged the role of system and model and optimized the choice of P_1 and P_2. The resulting parameter set was: $P_1 = 0.440$ and $P_2 = 0.868$. The optimum performance index value was smaller by a factor of 7, and the step response presented in Figure 2-4 as curve B is seen to be a much better match between system and model. The comparison between model and system velocity for this case is also shown. To see why such a design was rejected by the optimization program in the first example, one can examine what i(t) would have been using the second optimization results for the system but the original model with the "negligible" zero. Doing so, the initial value of i(t) was found to be 43, and that was also the high frequency gain $i(j\omega)$. The first optimization is seen to have reduced the high frequency gain to 11 and sacrificed low frequency behavior as evidenced by the step response of Figure 2-4. A satisfactory design was obtained in the second case by taking advantage of the fact that with $(\ell-k) < (n-m)$ in the interchanged situation, the highest output derivative was neglected in projecting the "system's" trajectory onto the "model's" characteristic plane. This interchange feature has been left in the program since it does permit one to design a system to approximate a model which had a greater pole-zero excess.

This alternative procedure has a different constraint however, and hence it does not in general remove all of the difficulties associated with the initial state effects. Since the performance index will diverge if any poles of i(p) lie in the right half plane, one cannot have right half plane zeroes in the model, for the model zeroes become poles of i(p). This further menas that if (ℓ-k) > (n-m) and the above model - system interchange is attempted, difficulties arise if the system can have right half plane zeroes. Such cases are common with aircraft control systems, and the airplane rudder coordination system examined in section 4.3 is an illustrative example. Although one could restrict the computer program so that right half plane system zeroes were not permitted as a result of the design parameter selection, it is not clear that the resulting airplane control system design would thereby be improved, and indeed it is conceivable that the system design would be unnecessarily complex. Therefore if system right half plane zeroes are to be permitted, another approach is needed in order to handle the initial state transient problem.

Modification 2. One notes from equations (2-5) and (2-6) that the performance index is zero if G_s/G_m is identically equal to unity. Since zero error is not possible nor is it necessary, one could state that satisfactory performance would also result if $G_s/G_m\Big]_{p=j\omega} = 1$ over the low frequency range of dominant importance to the operational use of the system. Since one is primarily interested in having the dominant response of the system approximate the model response, the low frequency region could be defined as some low multiple of the bandwidth of the model. One could then think of adding compensation terms to i(p) so that $G_s(j\omega)/G_m(j\omega) \approx 1$ over the frequency range of importance and so that little amplification occurred outside that frequency range. If this can be done by the addition of poles or zeroes outside the bandwidth of the model, then approximately the same dominant response would be expected. Examining equation (2-5), it is seen that one could either add poles to the system or zeroes to the model in order to attenuate the high frequency characteristics of i(p). To avoid dynamic interaction, poles added to the system would need to be cascaded outside of any feedback paths. It is probably conceptually easier to think of this compensation in the form of adding zeroes to the model rather than of adding poles to the system.

The magnitude of the frequency response of G_s/G_m is presented as curve A in Figure 2-5 for the first optimization result (with the zero at -100 added to the model). There is significant error at both low and high frequency. One can see however that an improvement of the low frequency portion over the bandwidth of the model by increasing ω_n of the system as in curve B would cause a much worse high frequency mismatch. In selecting the system of curve A , the optimization procedure had traded-off low frequency response in order to improve the high frequency response. Curve B was obtained using the system that resulted in the second optimization (modification 1) with the model having the zero at -100.

One notes that in the example cited the ratio of the compensation zero location to the model bandwidth was approximately 100, and this introduced the high frequency problem. In an iterative sequence of computer runs varying this ratio, it was found that using a zero whose characteristic frequency was five times the model bandwidth resulted in essentially the same optimum model-system match as did the model-system interchange technique. The frequency response corresponding to such a choice is shown in Figure 2-5 as curve C . The high frequency response was greatly attenuated. Thus, the technique of compensating i(p) so as to be approximately equal to 1.0 over a desired frequency range is one method of handling the initial state effect.

It requires a certain amoung of iterative cut and try as did the original Rediess technique.

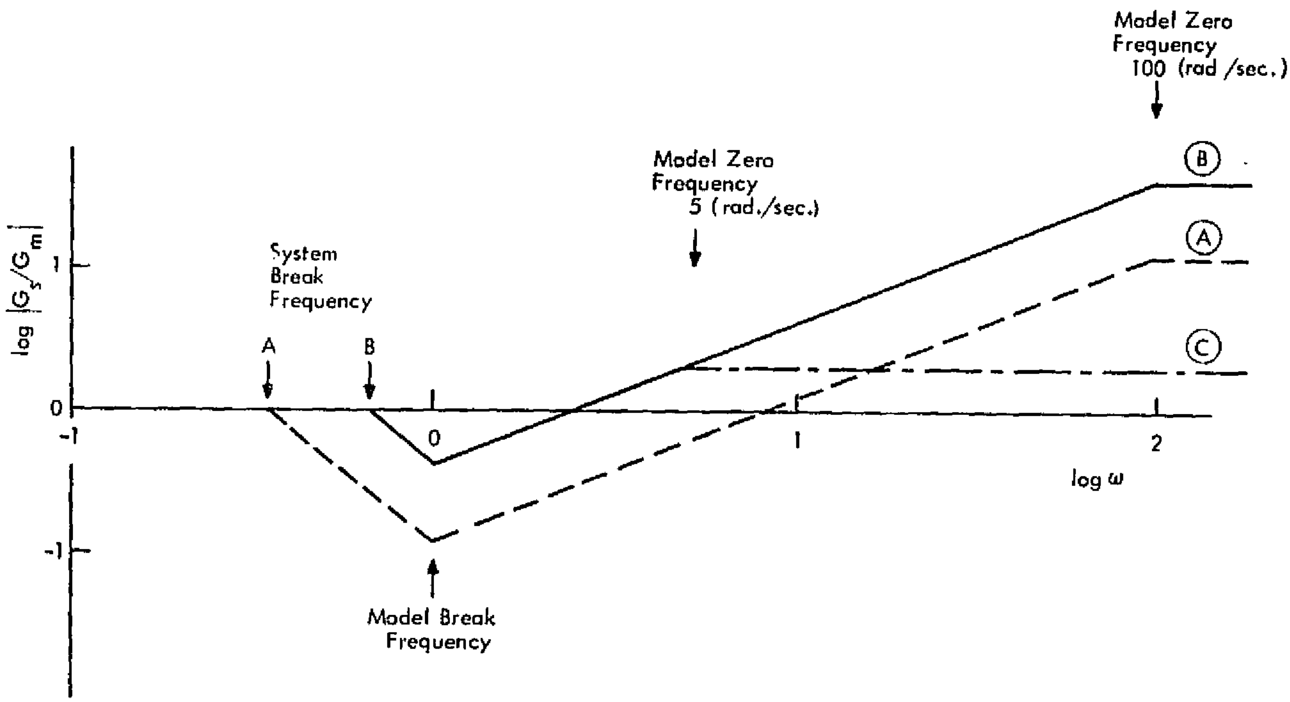


Fig. 2-5. Frequency Responses of G_s/G_m, Second Order Example

Modification 3. The Modification 3 alternative is to recognize that the unwanted frequency behavior of i(p) translates into the initial portion of the time history of i(t). One can screen out this initial contribution of i(t) to the performance index by starting the integration at a time somewhat greater than zero. This can be done by specifying a lower integration limit on the performance index so that

$$(PI) = \int_{PILLIM}^{\infty} [i(t)]^2 dt \qquad (2\text{-}9)$$

where PILLIM indicates the performance index lower limit. Selection of the lower limit is a trial and error process. In the example design that has been considered above, a lower limit of 0.02 seconds results in nearly the same system as that of modification 1 and the step response shown in Figure 2-4 as curve D. This lower limit was found by an iterative sequence of runs varying PILLIM until an acceptable design was achieved. In both modifications 2 and 3, one needs the constraint $(\ell-k) \leq (n-m)$. In modification 2 the compensation zero satisfies the constraint. In selecting the lower limit cited in the modification 3 example, the zero at -100 was used to satisfy the constraint.

The use of the lower limit technique is further illustrated by the airplane rudder coordination example discussed in section 4.3. There the initial state effects appeared with $(\ell-k) = (n-m)$, and no zeroes had to be added to the model. In that fairly complex example the lower integration choice was somewhat easier to use than was deciding how to compensate the frequency response of i(p) because of the presence of right half plane system zeroes. This is even more evident when considering digital systems. There, the frequency response would necessitate looking at the W-transform which is an added complication.

An Alternative Possibility. Finally, it is to be noted that using a quadratic performance index gives strong weight to the initial value of $[i(t)]^2$. Some preliminary work has indicated that using the absolute value of i(t) may accomplish the same thing as the lower limit on the integral. This would be a more difficult computational operation, but it would be worth investigating.

The modifications that have been made to the Model Performance index are thus seen to be alternative ways to take into account the initial state term of Equation (2-1). The iterative, trial and error nature of dealing with this term is still present. The modifications are recommended for computational convenience. The fundamental concept of the Model Performance Index remains unchanged.

2.5 Normalizing the performance index

The performance index as given so far has been expressed as

$$(PI) = \int_{PILLIM}^{\infty} [i(t)]^2 dt \qquad (2\text{-}10)$$

The objective of the optimization process is to determine that set of design parameters which minimizes (PI). In that sense the magnitude of the minimum (PI) is unimportant in determining the desired parameter set. To ignore the quantitative value of (PI) however would seem to discard relevant quantitative information descriptive of the performance capability of the control systems which should be helpful in comparing different system designs.

In the form given by equation (2-10), the magnitude of (PI) would vary with the characteristic frequency scaling, or bandwidth, of the system. That is, two systems whose respective poles and zeroes differed in frequency by a scale factor would have minimum (PI)'s which differed only by that scale factor, and compared to their respective scaled models would be considered to have the same performance capability. To remove the effect of time or frequency scaling, one needs to divide by some reference time, $T_{(ave)}$ over which i(t) could reasonably be expected to differ appreciably from zero. This could be some multiple of the step function response time of the model. To express the result as a time average, one would then want to take the square root.

When viewed as Rediess developed the performance index,

$$(PI) = \int_0^{\infty} \frac{\underset{\sim}{x}^T \underset{\sim}{\alpha}\ \underset{\sim}{\alpha}^T \underset{\sim}{x}\ dt}{||\underset{\sim}{\alpha}||} \tag{2-11}$$

The quantity, $(\underset{\sim}{x}^T\underset{\sim}{\alpha})/||\underset{\sim}{\alpha}||$, is the projection "distance" from the phase trajectory to the model characteristic plane. However the components of the state vector

$$\underset{\sim}{x}^T = [x, \dot{x}, \ddot{x}, \ldots, x^M]$$

do not have the same physical dimensions, and hence the concept of "distance" has an ambiguous meaning. The quantitative value of the performance index is distorted by the scaling selected.

For systems and models without zeroes and for a model of lower order than that of the system, i(t) has an initial value of -1.0 and a final value of zero. Thus the root-mean-square time average of i(t) relative to 1.0 is an approximate interpretation of the numerical value of the performance index. The presence of zeroes has been seen to distort the initial time response, some portion of which needs to be discarded in the (PI) evaluation by selecting PILLIM. Hence one could take as the performance index

$$(PI) = \left\{\frac{1}{T_{(ave)}} \int_{PILLIM}^{\infty} [i(t)]^2 dt\right\}^{\frac{1}{2}} \times 100 \tag{2-12}$$

in per cent and expect that for a desirable matching of the model and system, the value of (PI) from equation (2-12) would not exceed 5 to 10%.

Further work should be done in this area so that the quantitative value of the performance index will have a more recognizable significance. Since one is comparing system and model state space trajectories in model space, the order of the model should also affect the magnitude of the minimum performance index. Just how these aspects enter and affect the comparison of different system designs is still unclear. If the quantitative significance could be clearly established, the magnitude of the (PI) could be used to define acceptable performance deviations and thereby establish parameter tolerance margin specifications similar to the gain margin or phase margin specification that deal with stability margins. One could then investigate the use of compensation to widen the parameter variation over which acceptable performance would be obtained and thereby decrease parameter sensitivity.

2.6 Multi-Input Case

The multiple input design situation has not been examined in any depth in this investigation. If the various inputs are related in operational use, no particular difficulty arises. An example is the lateral control of an airplane. The airplane response depends upon what inputs are fed to the aileron and to the rudder, but if one postulates that the rudder input is to be coordinated with the aileron input to achieve an acceptable rolling maneuver response, one can look upon the overall system as receiving a single roll command input. If there is more than one output for a system, one may use several models and associate different sets of the design parameters with a particular output and its modelled response. The airplane lateral control provides an example in that one may use a roll angle model in optimizing feedback paths to the aileron while using yaw angular velocity in optimizing feedback paths to the rudder. In these cases an iterative design procedure is encountered in which first one parameter set and then the other is optimized, and the process repeated until a satisfactory design evolves. (Such a process could of course be made automatic on the computer if any advantage would thereby accrue.) As has been emphasized previously, the design of the model employed is of prime importance in achieving practical design results. Under the assumption of linearity, of course, any non-linear coupling effects for a multi-output case have been excluded from consideration here.

3. COMPUTATIONAL TECHNIQUES

To perform the parameter optimization, extensive numerical computations must be made which makes mandatory a computer program. A FORTRAN IV program has been written for this purpose. There are two versions, one for continuous system design and one for digital system design. This section describes the method of evaluating the performance index, the system representation, the digital control system techniques, and the general features of the program. The detailed program listing has not been included, but it can be made available upon request.

3.1 Evaluation of the Performance Index

Continuous Systems. Following the viewpoint of Palsson, the unnormalized Model Performance Index is

$$(PI) = \int_0^{\infty} [i(t)]^2 dt \tag{3-1}$$

where i(t) is the input to a component having the same transfer function as the model and whose output will be the same as the error between the model and system outputs. The performance index can be evaluated explicitly from the inverse Laplace transform of i(p). The latter can be written

$$i(p) = \left[\frac{G_s(p)}{G_m(p)} - 1\right]\frac{1}{p} \tag{3-2}$$

Since the input to the system is taken to be a unit step, $u(p) = \frac{1}{p}$. Then i(p) can be expressed as the sum of its partial fractions as

$$i(p) = \sum_{j=1}^{(n+k)} \frac{R_j}{p-\tilde{p}_j} \tag{3-3}$$

where

n = number of poles of the system

k = number of zeroes of the model

$$\tilde{p}_j = \begin{cases} \text{pole of the system for } j = 1 \text{ to } n \\ \text{zero of the model for } j = (n+1) \text{ to } (n+k) \end{cases}$$

and

$$R_j = \left\{(p-\tilde{p}_j)\cdot i(p)\right\}_{p=\tilde{p}_j} \tag{3-4}$$

Note that under the assumption that the system and model have the same static static sensitivity, the term associated with the input pole at the origin is identically zero. Taking the inverse Laplace transform

$$i(t) = \sum_{j=1}^{n+k} R_j e^{p_j t} \tag{3-5}$$

If the real part of p_j is less than zero, corresponding to a stable system,

$$\int_0^\infty [i(t)]^2 dt = \int_0^\infty \sum_{j=1}^{n+k} \sum_{i=1}^{n+k} R_j R_i \, e^{(p_j+p_i)t} \, dt \tag{3-6}$$

$$(PI) = -\sum_{j=1}^{n+k} \sum_{i=1}^{n+k} \frac{R_j R_i}{p_j+p_i} \tag{3-7}$$

Therefore by calculating the residues knowing the poles and zeroes of the system and the model transfer functions, the performance function is readily calculated from equation (3-7).

Digital Systems. In the case of digital systems the discrete version of the performance index becomes

$$(PI) \triangleq \sum_{q=0}^{\infty} [i(q)]^2 \cdot T \qquad (3\text{-}8)$$

where T is the sample period

$i(q) \triangleq i(qT)$, q = integer

In a manner analogous to that of the continuous system, the performance index can be evaluated from the inverse Z-transform of i(z)

$$i(z) = \left[\frac{G_s(z)}{G_m(z)} - 1\right] u(z) \qquad (3\text{-}9)$$

where $G_s(z)$ = system Z-transfer function

$G_m(z)$ = model Z-transfer function

$u(z)$ = Z-transform of a unit step, $\frac{z}{z-1}$

Expanding i(z) in its partial fractions

$$\frac{i(z)}{z} = \frac{G_s(z)}{G_m(z)} \frac{1}{(z-1)} - \frac{1}{z-1}$$

$$\frac{i(z)}{z} = \sum_{j=1}^{n+k} \frac{R_j}{z - \tilde{p}_j} \qquad (3\text{-}10)$$

where n = number of poles of the system

k = number of zeroes of the model

$$\tilde{p}_j = \left\{\begin{matrix} \text{system pole for } j = 1 \text{ to } n \\ \text{model zero for } j = (n+1) \text{ to } (n+k) \end{matrix}\right\} \text{ located in the z-plane}$$

and

$$R_j = \left\{ (z-\tilde{p}_j) \frac{G_s(z)}{G_m(z)} \frac{1}{(z-1)} \right\}_{z=\tilde{p}_j} \tag{3-11}$$

Under the assumption that system and model have the same static sensitivity, the (z-1) term vanishes.

Performing the inverse transform of equation (3-10),

$$i(qT) = \sum_{j=1}^{n+k} R_j (\tilde{p}_j)^q \tag{3-12}$$

$$(PI) = T \sum_{q=0}^{\infty} [i(qT)]^2$$

$$= T \sum_{q=0}^{\infty} \left(\sum_{j=1}^{n+k} R_j (\tilde{p}_j)^q \right) \left(\sum_{i=1}^{n+k} R_j (\tilde{p}_i)^q \right)$$

$$= T \sum_{j=1}^{n+k} \sum_{i=1}^{n+k} R_i R_j \left(\sum_{q=0}^{\infty} (\tilde{p}_i \tilde{p}_j)^q \right) \tag{3-13}$$

For stable systems $|\tilde{p}_j| < 1$

and

$$(PI) = T \sum_{j=1}^{n+k} \sum_{i=1}^{n+k} \frac{R_i R_j}{1-\tilde{p}_i \tilde{p}_j} \tag{3-14}$$

Thus, by calculating the residues from equation (3-11), one can obtain the performance index from equation (3-14).

3.2 Placing a Lower Limit upon the Integration in the Model Performance Index

If instead of equation (3-1), one defines the performance index as

$$(PI) = \int_a^{\infty} [i(t)]^2 dt \tag{3-15}$$

$$= - \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{R_i R_j}{(p_i + p_j)} e^{(p_i+p_j)a} \tag{3-16}$$

Thus, to use a lower integration limit, the exponential term is added inside the summation.

The analogous case for discrete systems is

$$(PI) = T \sum_{j=1}^{n+k} \sum_{i=1}^{n+k} R_i R_j \left(\sum_{q=a}^{\infty} (p_i p_j)^q \right) \tag{3-17}$$

where the lower limit is aT and a is an integer. Then changing the index of the final summation,

$$(PI) = T \sum_{j=1}^{n+k} \sum_{i=1}^{n+k} R_i R_j \ (p_i p_j)^a \left(\sum_{q=0}^{\infty} (p_i p_j)^q \right)$$

for $|p_i p_j| < 1.0$

$$(PI) = T \sum_{j=1}^{n+k} \sum_{i=1}^{n+k} \frac{R_i R_j (p_i p_j)^a}{1 - p_i p_j} \tag{3-18}$$

3.3 The Computer Program

Design synthesis through parameter optimization is feasible only with the use of a large digital computer. The programming effort devoted to the development of a suitable computer program was by far the largest component of the work performed under this grant. While the calculations are not particularly complex theoretically, there are many arithmetic operations required, and the inaccuracies associated with the finite word length of the computer required many programming modifications and software design iterations before a reasonably useful program was obtained. These accuracy difficulties increase as the order of the system being designed increases, and a program that produces satisfactory results for low order systems may still break down as ever higher order examples are attempted. The present program has been checked out for example designs of continuous systems of as high as 18th-order and of digital systems as high as 10th-order with satisfactory results. It is dimensioned for 30th-order systems. Computation time increases rapidly with order of the system as one would expect.

General Program Organization. Two separate programs, both of which use some of the same subroutines, have been developed -- one for design of continuous systems and one for digital systems. Figure 3.1 presents a generalized flow diagram for the computer program applicable to both programs. The basic function of the program is to evaluate the performance index for a specified set of design parameters and to vary the design parameters to search for that set which minimizes the performance index. One can specify the maximum number of iterations, (ITMAX), that one wishes a given computer run to use. This permits one to examine the design results in situations in which the convergence to the minimum is very slow to see if the design is adequate for practical purposes without excessive computation cost. By setting (ITMAX) to zero, the performance index can also be evaluated for a specific set of design parameters since only the initial calculation will be made before the stopping condition is encountered.

To evaluate the performance index, the development presented in section 3.1 is used. Thus one needs to form the closed-loop transfer function of the system

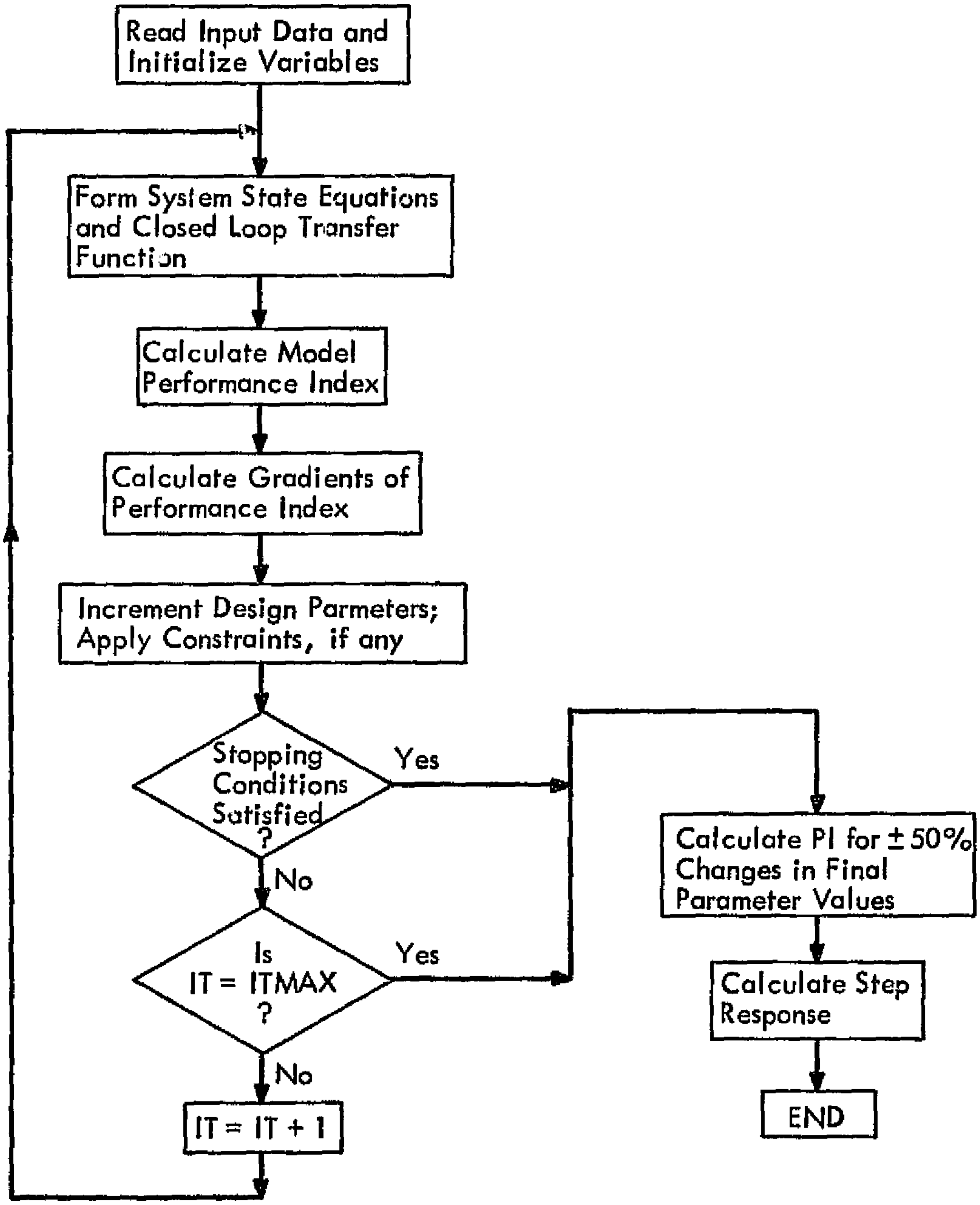


Fig. 3-1. Generalized Flow Diagram for the Parameter Optimization Program

for the specified values of the design parameters. There are of course several ways of obtaining the transfer function. This program first forms the system state equation for the closed-loop system in the standard form

$$\dot{\underline{x}} = A\,\underline{x} + B\,u$$

$$\underline{y} = C\,\underline{x} + D\,u$$

Taking the Laplace transform, the eigenvalues of the A-matrix are then found, and these are the poles of the system transfer function. The zeroes of the desired transfer function are obtained in a similar manner as the eigenvalues of a matrix resulting from algebraic manipulation of the numerator determinant of the transfer function (see section 3.5). Having the closed loop transfer function, the performance function is evaluated from equation (3-16). This requires computation of the step function response residues.

While it is possible to obtain an analytical expression for the gradient of the performance index, it was considered that a simple approximation would be adequate and less expensive in computation time. The gradient is the partial derivative of the performance index with respect to a particular design parameter. This has been approximated as

$$\frac{\partial(PI)}{\partial P_i} \cong \frac{\Delta(PI)}{\Delta P_i}$$

where $\Delta(PI)$ is the incremental change in the performance index obtained for an incremental change, ΔP_i, in the parameter. The parameter change used is ±1% (see section 3.6 for accuracy considerations).

The minimum performance index point is searched for using a modified gradient search algorithm. While some of the more elaborate search algorithms were investigated*, no significant advantage was found from their use. This results from the fact that for practical flight control design situations the convergence time with this algorithm is rapid enough that with present day computation speeds the computation costs may be no more than input-output charges. This of course would vary with the charging policy of the particular computation facility being used. In any event, there would be no inherent difficulty in using a different search algorithm with the rest of the program if one desired to do so. From the gradient values, the incremental changes (or steps) to the parameter are computed subject to any user supplied constraints, and the whole process is repeated with the new parameter values. The process continues until a stopping condition is encountered or until the maximum number of iterations is reached.

After a stopping condition is reached, the final parameters are individually changed by 50% and the performance index is calculated. This provides an indication of parameter sensitivity. The step function response of the optimized system is then computed and tabulated.

The continuous system and the discrete system programs differ only in the detailed differences between the continuous and the discrete representations of the system. For example, the model is specified by its Laplace transform poles and zeroes for the continuous system, while for the discrete system the Z-transform poles and zeroes are specified.

3.4 Continuous System Representation

Since the design parameters that are to be optimized can occur at any point

* For example, those available in the IBM Scientific Subroutine Package.

in the system, a flexible method of system representation is required. The program achieves that in a manner essentially the same as reported in reference 6. Each continuous component is represented by its transfer function. Each transfer function in turn can be represented as a series cascade of first- or second-order components with first- or second-order numerators. In addition summation points and gain elements are needed. For a multi-input component such as the flight vehicle, it is possible to specify its state equations rather than its transfer function as input data. One then prepares a block diagram arranging the elemental units or blocks in any configuration one wishes. All signal paths are numerically labelled, or assigned subscripts, as are the state variables, if any, associated with any element. Identifying names, input and output signal subscripts, and state variable subscripts for the components are read as part of the data input. Using summation points, any arrangement of feedback of feedforward paths can be specified. The design parameters can be any of the gain elements or the time constants or dynamic parameters of the elemental blocks. During the search process constraints can be placed upon the parameters. For example if the pole and zero of a lead compensator were design parameters, a maximum value for the ratio of pole to zero could be specified if that were desirable. For digital flight control systems, the design parameters can occur in either the digital or in the continuous section of the system.

3.5 Digital Systems

System Configuration. In a digital flight control system some or all of the information processing takes place in a digital computer. Analog to digital conversion is first needed to convert the analog signals from the sensor system into digital format for computer processing. The output from the computer is generally the set of input command signals needed to actuate the control surface servos. The latter are analog devices, and therefore a digital to analog signal conversion is also needed.

The system synthesis program that has been developed in this investigation permits one to specify a general control system configuration of the type shown by the functional block diagram of Figure 3.2 In that figure the information

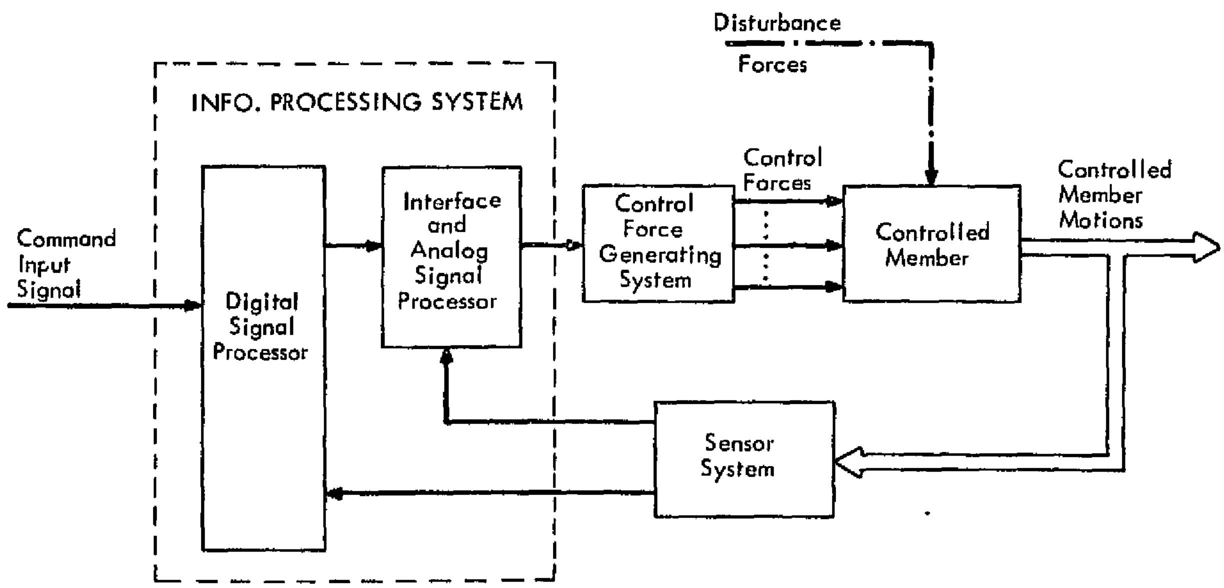


Fig. 3-2. General Digital Control System

processing system which connects the sensor system to the control force generating system includes both analog and digital processing sections. To suggest further the functional nature of the various devices, Figure 3.. provides a somewhat more detailed elaboration of Figure 3.2. Various filters are indicated to denote signal processing and summation which can occur in either the analog

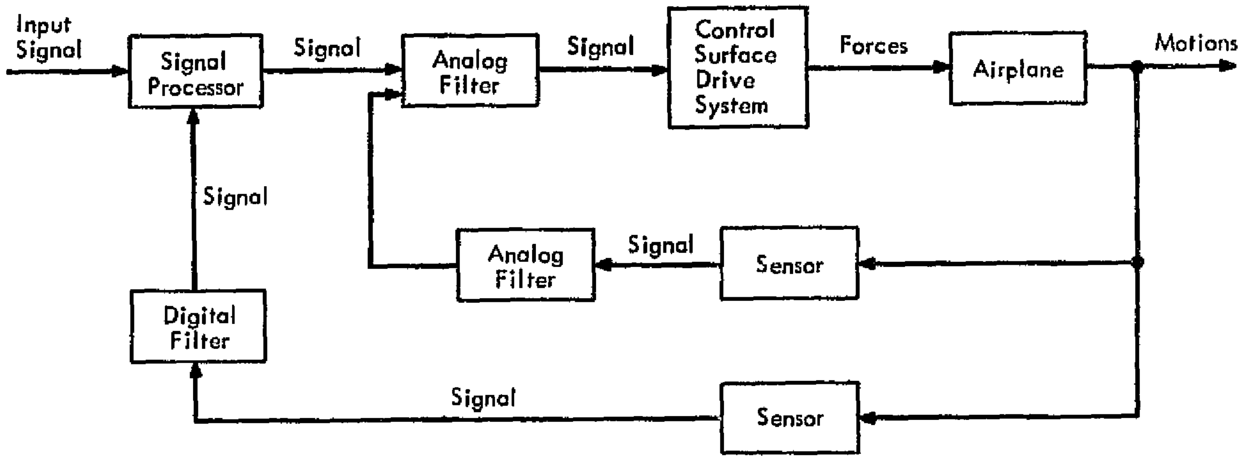
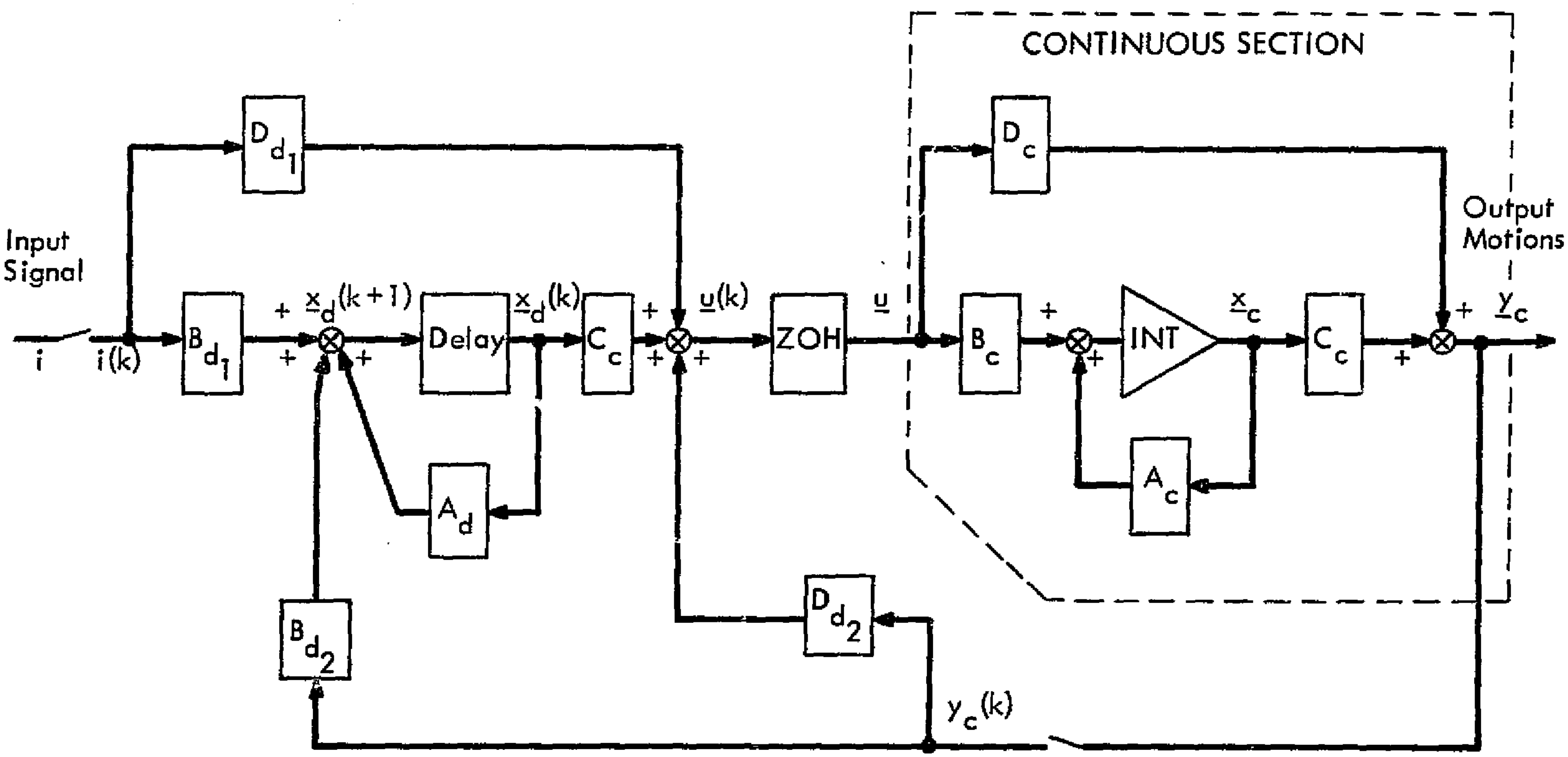


Fig. 3-3. Typical Aircraft Digital Control System

or the digital sections, although there may not be the physical separation of equipment shown by the diagram.

State Variable Representation. From a mathematical modelling standpoint the system is described by the relationship of various continuous state variables and various discrete state variables. Symbolically, the mathematical block diagram of Figure 3.4 suggests this, although the correlation with the physical devices is thereby obscured. The continuous system elements are represented by an n_c-dimensioned state vector, $\underline{x}_c$, and the discrete system elements by an n_d-dimensioned state vector, $\underline{x}_d$. The input to the system enters the digital section, and the inputs to the continuous element section are considered to be the outputs of zero-order hold devices at the digital to analog interface. Only one sampling rate is assumed.

It is convenient to represent the combined system by a set of discrete



Discrete State Variables = $\underline{x}_d$ Continuous State Variables = $\underline{x}_c$

Fig. 3-4. Mathematical Block Diagram of a General Digital Control System difference equations in which the variables denote the values of the system quantities at equally spaced time intervals. To obtain such a set of equations, consider the continuous section of the system of Figure 3.4 to be described by the differential equation

$$\left(\frac{d\underline{x}_c}{dt}\right) = A_c\underline{x}_c + B_c\underline{u} \tag{3-19}$$

with initial condition $\underline{x}_c(0) \equiv \underline{x}_{c_0}$

$\underline{u}$ = rth-dimension vector input to the continuous section

$\underline{x}_c$ = n_cth-dimension vector of state variables of the continuous section

A_c = $n_c \times n_c$ - matrix of coefficients

B_c = $n_c \times r$ - matrix of coefficients

Since $\underline{u}$ is the output of the zero order holds,

$$\underline{u}(t) = \underline{u}(k), \text{ for } kT \le t < (k+1)T \tag{3-20}$$

where T is the sampling period of the digital section and the simplification in notation, $u(kT) \equiv u(k)$, has been used. The various signal paths are given by

$$\underline{y}_c = C_c\underline{x}_c + D_c\underline{u} \tag{3-21}$$

where the matrices are appropriately dimensioned depending upon the number of y signals that are of interest.

The digital section is described by the difference equations

$$\underline{x}_d(k+1) = A_d\underline{x}_d(k) + B_{d_1}i(k) + B_{d_2}\underline{y}_c(k) \tag{3-22}$$

$$\underline{u}(k) = C_d\underline{x}_d(k) + D_{d_1}i(k) + D_{d_2}\underline{y}_c(k) \tag{3-23}$$

where $\underline{x}_d$ = n_d-th dimension vector of state variables of the digital section

$\underline{u}(k)$ = output signal vector of the digital section which is the input to the zero order holds at the digital-analog interface

The D_{d_2} matrix represents those output feedback paths which are processed in the digital information section and modified only in magnitude but not in phase. Those which also receive dynamic compensation are represented in the B_{d_2} matrix. The input signal, i(t), is chosen to be a step function so that

$$i(t) = \begin{cases} 0 & , t < 0 \\ 1.0 & , t > 0 \end{cases} \tag{3-24}$$

Discrete Representation of the Continuous Section. It is convenient in analyzing this overall system to obtain a discrete representation of the continuous section giving the state variable values at the sampling time of the digital section. Making use of the state transition matrix

$$\underline{x}_c(k+1) = \Phi_c \underline{x}_c(k) + \Gamma_c \underline{u}(k) \tag{3-25}$$

where

$$\Phi_c = e^{A_c T} \tag{3-26}$$

$$\Gamma_c = \int_0^T e^{A_c(T-\tau)} B_c d\tau \tag{3-27}$$

Letting $\underline{x}$ denote the combined state vector

$$\underline{x} = \begin{bmatrix} \underline{x}_c \\ \underline{x}_d \end{bmatrix} \tag{3-28}$$

these equations can be solved to obtain the discrete state equations for the overall system.

Substituting equation (3-23) into (3-21),

$$\underline{y}_c = C_c\underline{x}_c + D_c\{C_d\underline{x}_d + D_{d_1} i + D_{d_2} \underline{y}_c\} \tag{3-29}$$

or

$$\underline{y}_c = \{I - D_c D_{d_2}\}^{-1}\{C_c\underline{x}_c + D_c C_d \underline{x}_d + D_c D_{d_1} i\} \tag{3-30}$$

and

$$\underline{u} = C_d\underline{x}_d + D_{d_1} i + D_{d_2}\{I - D_c D_{d_2}\}^{-1}\{C_c\underline{x}_c + D_c C_d\underline{x}_d + D_c D_{d_1} i\} \tag{3-31}$$

These can be substituted into (3-22) and (3-21) to obtain discrete state equations of the form

$$\left.\begin{aligned} \underline{x}(k+1) &= \Phi\,\underline{x}(k) + \underline{\Gamma}\; i(k) \\ \underline{y}(k+1) &= C\;\underline{x}(k) + \underline{D}\; i(k) \end{aligned}\right\} \tag{3-32}$$

The state variables, $\underline{x}$, and the signal quantities, $\underline{y}$, of equation (3-32) may or may not be directly observable physical quantities. This depends upon the particular mathemaical modelling which is employed, and that is often chosen on the basis of expediting the mathematical analysis. The system output quantities of interest to the designer therefore may be some of the y_i signal points or may be combinations of them. In any event the output quantities will be linear combinations of the state variables and of the input to the system.

Various techniques are available for obtaining the state equations (3-32) for the complete system. If the development presented previously is followed, a matrix inversion occurs in equation (3-30) when D_c is not zero. A typical case is a normal acceleration control system for an airplane. The lift due to elevator deflection produces an output indication proportional to the control input. The computer program used in this study avoided the matrix inversion by forming the discrete state equations (3-25). The entire continuous section can then be represented as a single component of a system which now is made up of discrete components. There may be several signal paths leading into and/or out of that component. The overall discrete system representation is then reduced to obtain the overall state equations.

In obtaining the discrete version of the continuous section, equation (3-25) shows that the state transition matrix, $\boldsymbol{\Phi}$, and the convolution integral, Γ, are needed. Several methods of evaluating the state transition matrix were examined. The best numerical accuracy was obtained, at the expense of greater computation time, by an analytical evaluation of the time response through the inverse Laplace transform. Since a zero-order hold is assumed, the control, $\underline{u}$, is constant over one sample period. Hence Γ involves the integral of the state transition matrix which is easily found in the same procedure.

Letting

$$\boldsymbol{\Phi}(t) = e^{At}$$

$$\boldsymbol{\Phi}(p) = (pI - A)^{-1} \qquad \text{an } n \times n \text{ matrix}$$

Then $e^{AT} = \mathscr{L}^{-1}\{\boldsymbol{\Phi}(p)\}\Big|_{t=T}$

also $\displaystyle\int_0^T e^{Ah}dh = \mathscr{L}^{-1}\{\frac{1}{p}\boldsymbol{\Phi}(p)\}\Big|_{t=T}$

The (ij)th element of $\boldsymbol{\Phi}(p)$ can be denoted

$$\phi_{ij}(p) \triangleq \frac{N_{ij}(p)}{D(p)}$$

$$= \sum_{k=1} \frac{R_{ijk}}{(p-p_k)} \qquad (3\text{-}32)$$

where $N_{ij}(p)$ = numerator polynomial of the (ij)th element

$D(p)$ = system's characteristic polynomial

p_k = system poles, or eigenvalues of the matrix A.

From the definition of the inverse of a matrix,

$$N_{ij}(p) = (-1)^{i+j} \times \text{determinant of } (\text{adjoint } (pI-A))_{ji}$$

where the (ji)th element is needed so that the transpose of the cofactor matrix will be obtained.

For purposes of programming for computer evaluation, the needed (ji)th determinant can be evaluated by taking the matrix (pI-A) and replacing the ith column by a column vector having its jth element equal to 1.0 and all others zero. The expansion of this determinant will be the cofactor of the (ji)th elements.

In the matrix (pI-A), the Laplace operator occurs only in the diagonal elements which are of the form $(pI-a_{ii})$. Thus the cofactors of the diagonal elements, i = j, will be of order (n-1) while those of the off-diagonal elements will be of order (n-2) or less. By determinant operations each of the determinants needed for the $N_{ij}(p)$ terms can be arranged in the form

$$\begin{vmatrix} c_{11} & -c_{12} & \cdots & \cdots \\ 0 & p-c_{22} & -c_{23} & \cdots \\ 0 & -c_{32} & p-c_{33} & \cdots \\ \vdots & \vdots & \vdots & \end{vmatrix} = c_{11} \begin{vmatrix} p-c_{22} & -c_{23} & \cdots \\ -c_{32} & p-c_{33} & \cdots \\ \vdots & & \end{vmatrix}$$

This latter determinant is of the form $c_{11}(pI-C)$, and the roots of the polynomial $N_{ij}(p)$ are given by the eigenvalues of C. The inverse transform of equation (3-32) is

$$\phi_{ij}(t) = \sum_{k=1}^{n} R_{ijk}\, e^{jp_k t} \tag{3-33}$$

whence

$$\phi_{ij}(T) = \sum_{k=1}^{n} R_{ijk}\, e^{jp_k T} \tag{3-34}$$

To evaluate the ϕ_{ij}, one then needs the residues, R_{ijk}. These are readily evaluated from the poles and zeroes of ϕ_{ij} as

$$R_{ijk} = (p-p_k)\,\phi_{ij}(p)\Big|_{p=p_k} \tag{3-35}$$

The poles are obtained as the eigenvalues of the A matrix, and the zeroes are the eigenvalues of the new matrices, C_{ij}. The appropriate gain factor must also be calculated in the process. When the continuous section contains integrators, some of the ϕ_{ij} elements may have multiple poles at the origin. If there are s poles at the origin, the partial fraction expansion will also involve the derivatives of

$$p^s\ \phi_{ij}(p)$$

up to the order (s-1). The program presently permits a maximum of two poles at the origin in the continuous section.

The residues needed for the integral of the transition matrix are obtained by dividing each of the above residues by its associated pole, since

$$\int_0^T e^{At}\,dt = \mathscr{L}^{-1}\{\frac{1}{p}\,\Phi(p)\}\Big|_{t=T}$$

$$= \sum_{k=1}^{n} \frac{R_{ijk}}{p_k}\,(e^{jp_k T} - 1) \tag{3-36}$$

This method of evaluating Φ and Γ thus involves calculating n^2 sets of numerator polynomials, the associated zeros and residues and results in long computation times for high order systems. The accuracy has been acceptable, however, for the aircraft flight control systems examined to date. Note that if there are no variable optimization parameters in the continuous section of the control system, the computation of the discrete representation of the continuous section need only be performed once during an optimization run. In that case, which may be the more common case for future flight control systems, the computation time penalty remains a small fraction of the total computation time.

3.6 Numerical Inaccuracy Difficulties

The method of evaluating the model performance index that is being used was arrived at after several alternative methods encountered inaccuracy difficulties that were not successfully surmounted. It cannot be claimed however that the present technique is inherently more accurate, for perhaps a more sophisticated programming capability than was available would have solved the problems.

The fundamental impediment to achieving satisfactory optimization results is the numerical inaccuracy of the calculation of the value of the performance index and its gradient. In all of the methods so far investigated, one encounters at some stage the equivalent of summing a set of large numbers whose sum is a very small number. In the process many significant numbers are lost, and one eventually encounters the limit of significant figures imposed by the finite word length of the digital computer. Thus, even though one may need only 2 or 3 significant figures in the values of the final optimized parameter set, a much higher level of precision is required in the intermediate calculations. To a large extent this results from the use of a gradient search algorithm. One can encounter regions of parameter space where the gradients are small even though one is not near enough to the optimum point. If the gradient inaccuracies predict the wrong algebraic sign, then a corresponding parameter step will proceed away from the optimum, and to avoid excessive computer time usually some sort of stopping condition is encountered.

It is estimated that the present program calculates the performance index with an inaccuracy of one part in 10^5. Since the gradients are computed as first order differences due to incrementing the parameters, one desires to use small increments. One percent parameter increments are used. If the parameter change produces a performance index change of less than the estimated inaccuracy, 5% increments are used. If this also fails, the gradient is set to zero, and no change is made in that parameter for the next iteration. The maximum component term in the performance index computation is printed out so that one can check the accuracy of the listed values.

The programming used by Palsson involved matrix inversion whose inaccuracies ultimately caused it to be abandoned (see Ref. 7) in favor of the present technique.

Thus the numerical inaccuracy associated with making the very large number of arithmetic operations must be constantly guarded against. It will no doubt cause trouble with very large order systems. Rather than resort to more elegant computational procedures, it is felt that simpliciations to the design techniques could be made that would proceed in the direction of making sure that negligible effects would either remain negligible or would be removed altogether so that the resulting system designs would possess the maximum utility. This is an area for future development.

The numerical inaccuracy problem is compounded by the difficulty in assessing a physical meaning to the value of the performance index (see Section 2.4). If one could be assured that a certain value of the performance index meant that the system response (under some agreed upon definition of the word "response") approximated that of the model to within say 10%, then one would only need to carry the optimization to that level without too much concern for what the exact mathematical optimum might be. However, the Model Performance Index weights a number of states of the system that depend upon the order of the model, and it has not been shown that two designs with the same performance index value exhibit the same degree of model-system matching. The normalization discussed in Section 2.4 is an attempt in that direction, but results are inconclusive.

Thus, one is plagued by the uncertainty in knowing whether the optimum reached on a computer run is the best one can achieve or merely the one limited by computational inaccuracies. As in any design task, it is recommended that one perform a parameter sensitivity analysis on the final design to insure that a practical system has been obtained.

4. DESIGN EXAMPLES

This chapter presents several design examples of the use of this parameter optimization technique. All of these are flight vehicle control systems. The first illustrates the reduction in system complexity possible with a parameter optimization approach compared with the state variable feedback approach for an airplane pitch damper system. The second example is the design of a digital C* fly-by-wire system in which nine flight conditions have been considered. Next a lateral-directional aircraft flight control system is examined to illustrate the use of two design parameter sets to satisfy separately the roll and sideslip requirements. Both continuous and digital systems are considered. Finally the design of a compensation filter to phase-stabilize a body bending mode for a launch booster vehicle is considered. The highest order system is tenth order. Sensor dynamics and control effector dynamics have been illustrated. The effect of sampling frequency is briefly examined in the lateral control system example.

4.1 Pitch Damper Example

The design of a pitch rate damper system for an airplane provides an example of the design simplification that a parameter optimization approach permits in comparison with the state variable feedback solution of the optimal control technique.

Optimal Control Design. The well known theoretical result from optimal control theory states that one needs to measure all of the state variables and feed all of these signal indications modified by suitable gains to each of the control effectors. For a high order system such a system configuration is unnecessarily complex since one can never specify the desired performance in sufficient detail that the so called "optimum" solution can be defined. In general, performance specifications only express acceptable performance tolerance boundaries within which the values of performance measures should lie.

With the design freedom that state variable feedback provides, one can place the eigenvalues of the system anywhere in the complex plane one desires. If one knows what the closed-loop zeroes will be, one can specify the desired closed-loop transfer function and solve for the feedback gains needed to match it exactly. One can denote the desired transfer function as a model, and one of the requirements of the model specification in this case is that its order be the same as that of the system.

Consider a pitch rate control system for the F-8 airplane. For a flight condition of Mach 1.1 at 5180 meter altitude, the state equations for the airplane are given in Appendix A as flight condition 2. The control system using state variable feedback then consists of sensors to indicate the state variables, amplifiers to adjust feedback gains and add signals, an elevator servo to position the elevator control surface, and the airplane as shown in Figure 4.1. The corresponding block diagram is shown in Figure 4.2. The elevator servo is assumed to be a first order lag whose state equation then is

$$\dot{\delta} = -12.5\ \delta + 12.5\ \delta_c \tag{4-1}$$

where δ_c is the input signal to the servo. Since the elevator deflection, δ, is another state variable, provision must be made for a feedback path from δ as shown in Figure 4.2. In the figure the loop gains are represented as the static sensitivities, S_i. These represent overall loop calibrations including

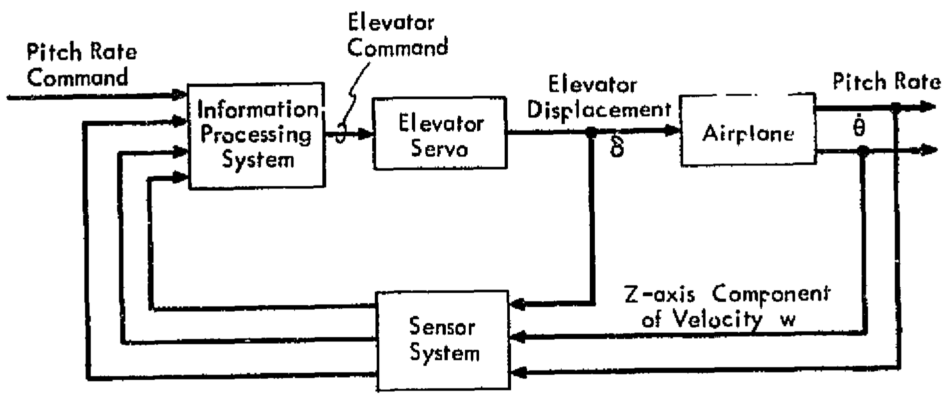


Fig. 4-1. Functional Block Diagram. State Variable Feedback Pitch Damper System.

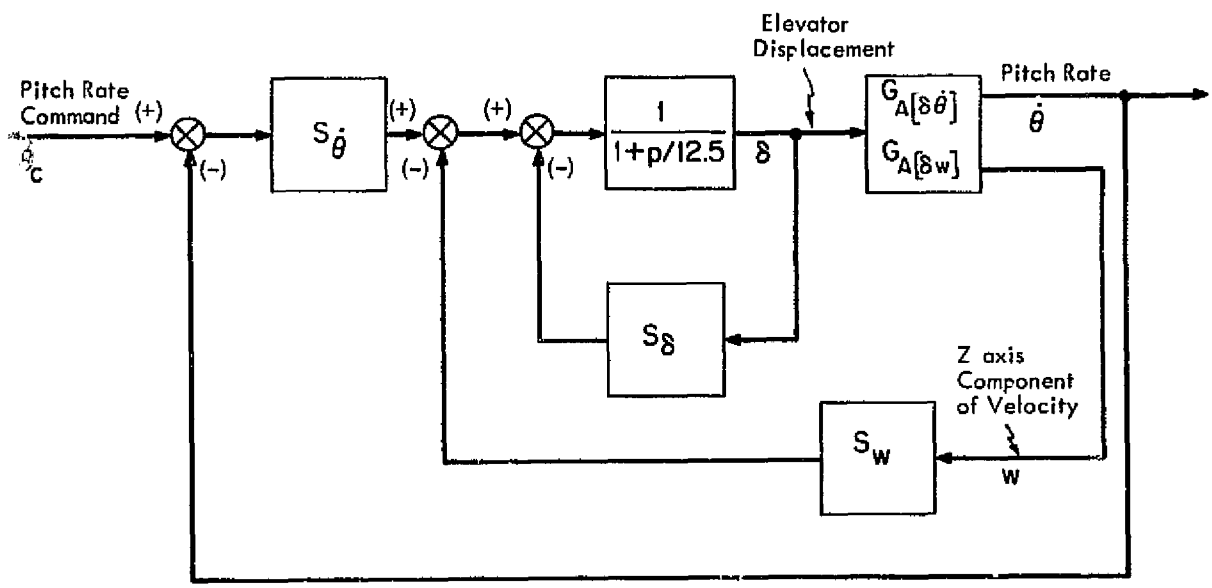


Fig. 4-2. Mathematical Block Diagram. State Variable Feeaback Pitch Damper System.

the static sensitivities of corresponding sensors, amplifiers, and servo. The input to the system has been assumed to have been calibrated as an equivalent command pitch angular rate, $\dot{\theta}_c$. The feedback signals are assumed to be subtracted at the signal summation points. The complete system state vector, $\underline{x}$, then is

$$\underline{x} = \begin{bmatrix} w \\ \dot{\theta} \\ \delta \end{bmatrix} \tag{4-2}$$

Although one can solve for the feedback gain matrix by matrix operations, it is easier for this simple system to use block diagram reduction making use of the airplane transfer functions. The Laplace transformation of the state equations yields the transfer functions

$$G_{A[\delta,w]} = \frac{-179.8(p + 162.8)}{p^2 + 2.196p + 71.73} \quad (\text{m/s})/\text{rad.} \tag{4-3}$$

$$G_{A[\delta,\dot{\theta}]} = \frac{-41.73(p + 1.722)}{p^2 + 2.196p + 71.73} \quad (\text{sec})^{-1} \tag{4-4}$$

The closed-loop transfer function then becomes

$$\frac{\dot{\theta}}{\dot{\theta}_c} = \frac{-521.6(p + 1.722)S_{\dot{\theta}}}{p^3 + a_2p^2 + a_1p + a_0} \tag{4-5}$$

where

$$\left.\begin{aligned} a_2 &= 2.196 + 12.5\ S'_{\delta} \\ a_1 &= 71.73 + 27.45\ S'_{\delta} - 1128\ S_w - 521.6\ S_{\dot{\theta}} \\ a_0 &= 896.6\ S'_{\delta} - 183500\ S_w - 898.5\ S_{\dot{\theta}} \\ S'_{\delta} &= 1 + S_{\delta} \end{aligned}\right\} \tag{4-6}$$

One needs to specify a model transfer function. It is known that the airplane pitch rate zero will also be a closed-loop zero, and hence it can be a model zero. From experience one expects the pitch rate feedback to increase the oscillatory mode's damping ratio, probably increase the natural frequency somewhat, and increase the break frequency of the real model. Typical values of these quantities for the model might be

$\zeta_m = 0.5$, $\omega_m = 9.0$ rad./sec.

Real pole at -15 rad/sec

As a practical system only the oscillatory mode characteristics would be of dominant importance, so the choice or the real mode location has been arbitrary. Since that choice will affect the values of the feedback gains, there is an element of trial and error in the design procedure which is not readily apparent from the theoretical development of the optimal control theory.

Using the suggested model, the model transfer functions became

$$G_m = \frac{705.3(p + 1.722)}{p^3 + 24p^2 + 216p + 1215} \qquad (4\text{-}7)$$

One can then compare the system and model transfer functions, equate polynomial coefficients, and solve for the required static sensitivities. Hence

$S_\delta = 0.744$

$S_w = 2.84 \times 10^{-3}$ rad./(m/s)

$S_{\dot{\theta}} = -0.191$ sec.

The model step function response is presented in Figure 4.3 as the dashed curve. The system step response will of course be identical to it.

One can understand what the feedback configuration is accomplishing by

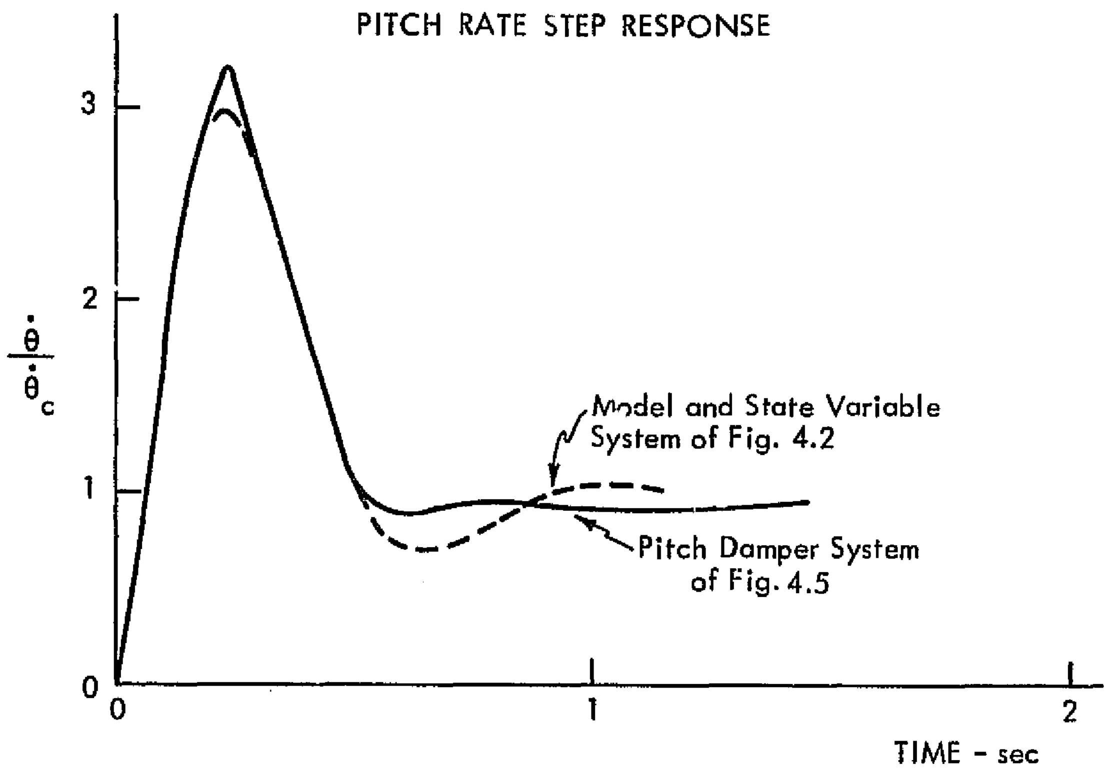


Fig. 4-3. Step Function Response of the Pitch Damper Systems.

examining the variation in the system open-loop modes using root locus diagrams. Considering the inner loop first, the servo feedback shown in Figure 4.2 increased the bandwidth of the servo from 15 to 21.8 (rad/sec). Considering the w loop next, one notes that the gain value that was specified was positive, which from Figure 4.2 indicates that a positive angle of attack (positive w) would call for a negative δ (trailing edge up). In servo loop parlance this is positive feedback, negative open-loop gain, and tends to be destabilizing in the sense that it reduced the inherent airplane aerodynamic static stiffness and hence the natural frequency of the airplane mode. This was necessary since the subsequent closure of the pitch rate loop increased the natural frequency. The root locus of Figure 4.4 shows this. The real mode changed very little and is not shown. The final loop closure is also shown on Figure 4.4, and it placed the closed-loop poles in the desired locations, primarily through an increase in damping ratio of the oscillatory mode. The frequency of the real mode was reduced to that specified by the model by the final loop closure.

A Parameter Optimization Design. The state variable feedback configuration required three sensors with the associated summing amplifiers unless some form of state estimator were to be used. The dynamics of the sensors have been neglected in the above analysis. In a practical implementation of a pitch damper system, Figure 4.5, the input to the system would be the pilot's mechanical input to the elevator servo. This would be summed mechanically with the output of a series servo driven by the feedback signal configuration. Only a pitch rate gyro is needed as a sensor, and to reduce the effects of the damper system upon the pilot's stick force characteristics a high-pass filter is customarily required to null the steady-state value of the feedback signal. Figure 4.6 is a mathematical block diagram of such a system. One needs to select the feedback static sensitivity and the value of the time constant for the high-pass filter. The rate gyro is assumed to be second-order and the series servo dynamics are negligible. Note that a state variable feedback analysis would have required additional feedback paths for the states associated with the gyro and the filter.

Using parameter optimization and the same model as previously specified, values of the gain and the time constant were found to be

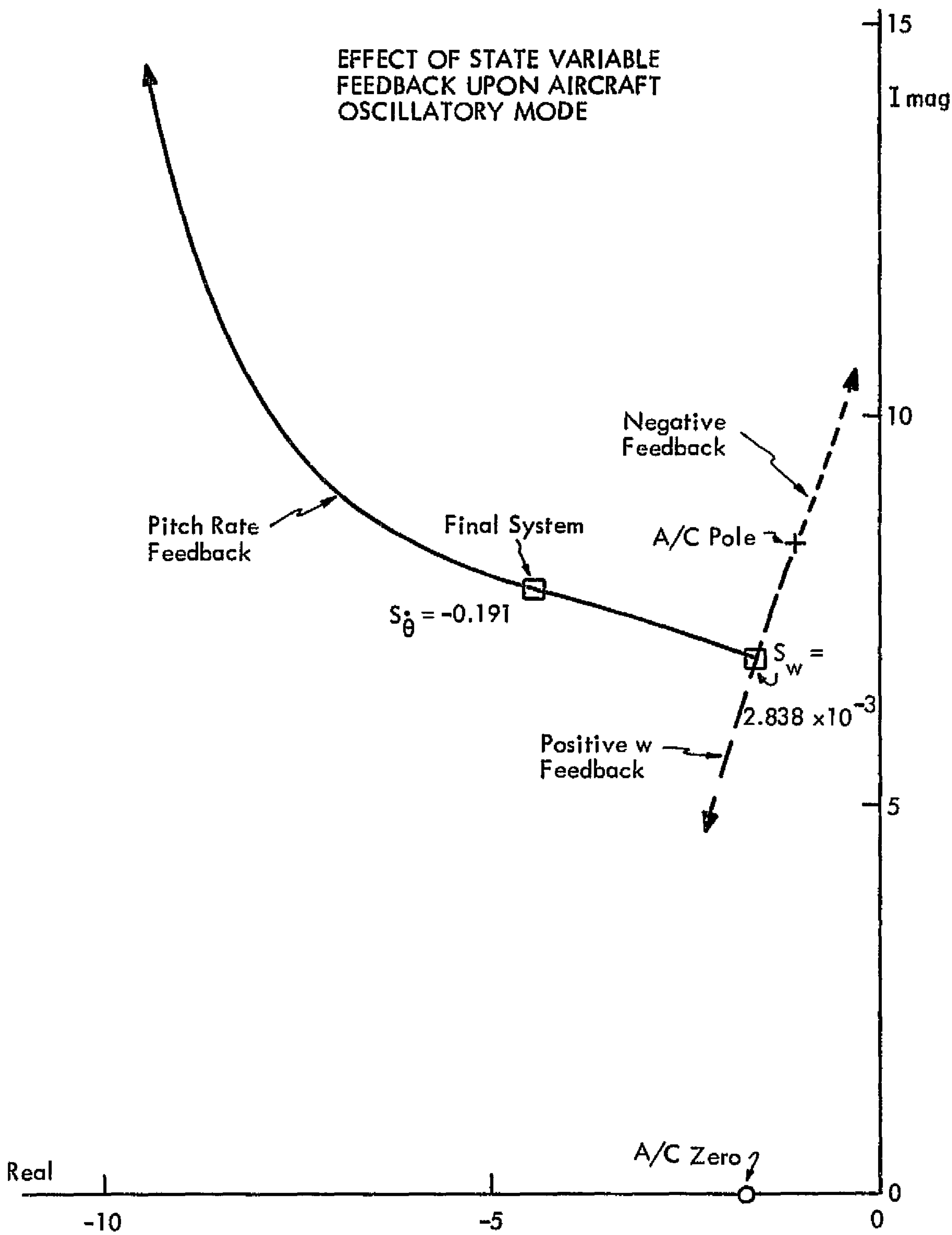


Fig. 4-4. Root Locus. State Variable Feedback Pitch Damper System.

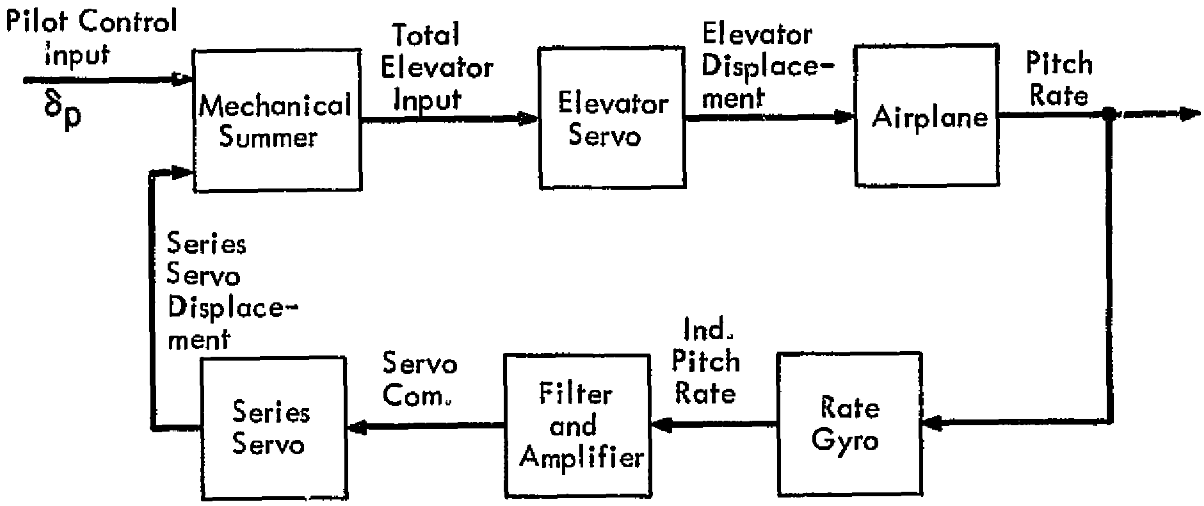


Fig. 4-5. Functional Block Diagram. Pitch Damper System.

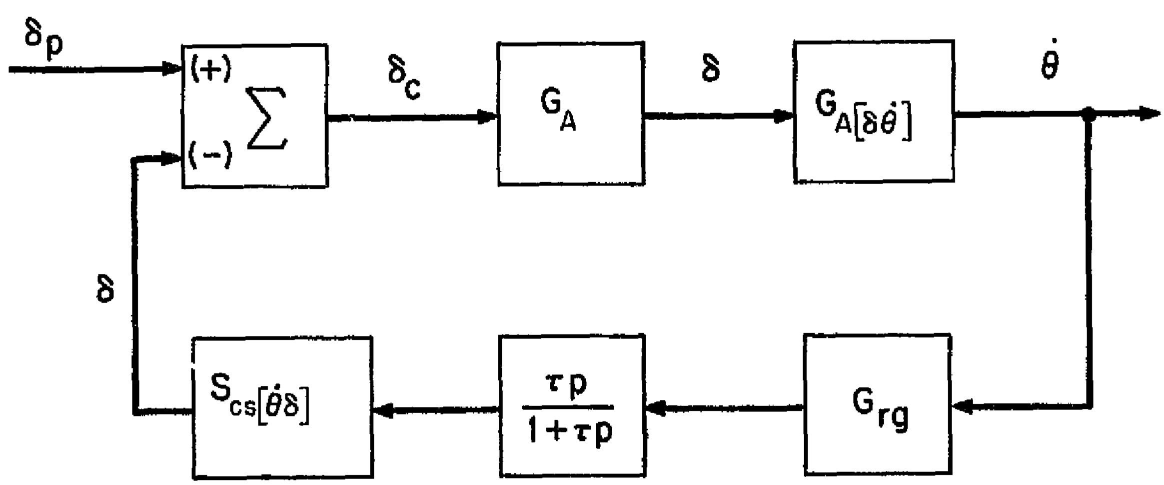


Fig. 4-6. Mathematical Block Diagram. Pitch Damper System.

$S_{cs[\dot{\theta},\delta]} = 0.0370$ sec.

$\tau = 0.187$ sec.

Comparison of the system responses for the two designs is presented in Figure 4.3. The two are not identical, but the differences have no practical significance. This design is simpler and includes several practical requirements omitted in the first design. The root locus for this single loop configuration is presented in Figure 4.7 with the optimized closed-loop poles indicated as small squares.

It would have been difficult to anticipate that such a closed-loop pole-zero configuration would have provided as good a matching of system and model as it did. If one could have done so and used it to specify the model, then of course the state variable feedback technique would have led to the same design. The parameter optimization approach permitted one to examine the performance capabilities and limitations of the simplest system configuration that previous experience had indicated might be satisfactory rather than presenting one with the most complex solution which then usually would need to be made practical to implement through some process of simplification.

In the example cited, the simple configuration was adequate. Other design considerations might have specified some constraint on the design parameters such as a value for the desired filter time constant. For example, if one specified that τ must be 1.0 second, the parameter optimization showed that the best value of loop gain still did not result in a satisfactory system from the standpoint of too low damping of the oscillatory airplane mode. Thus one would need to provide compensation, and the natural next step was to investigate the use of a lead compensation filter. The parameter optimization of the loop gain and lead filter time constants readily led to a

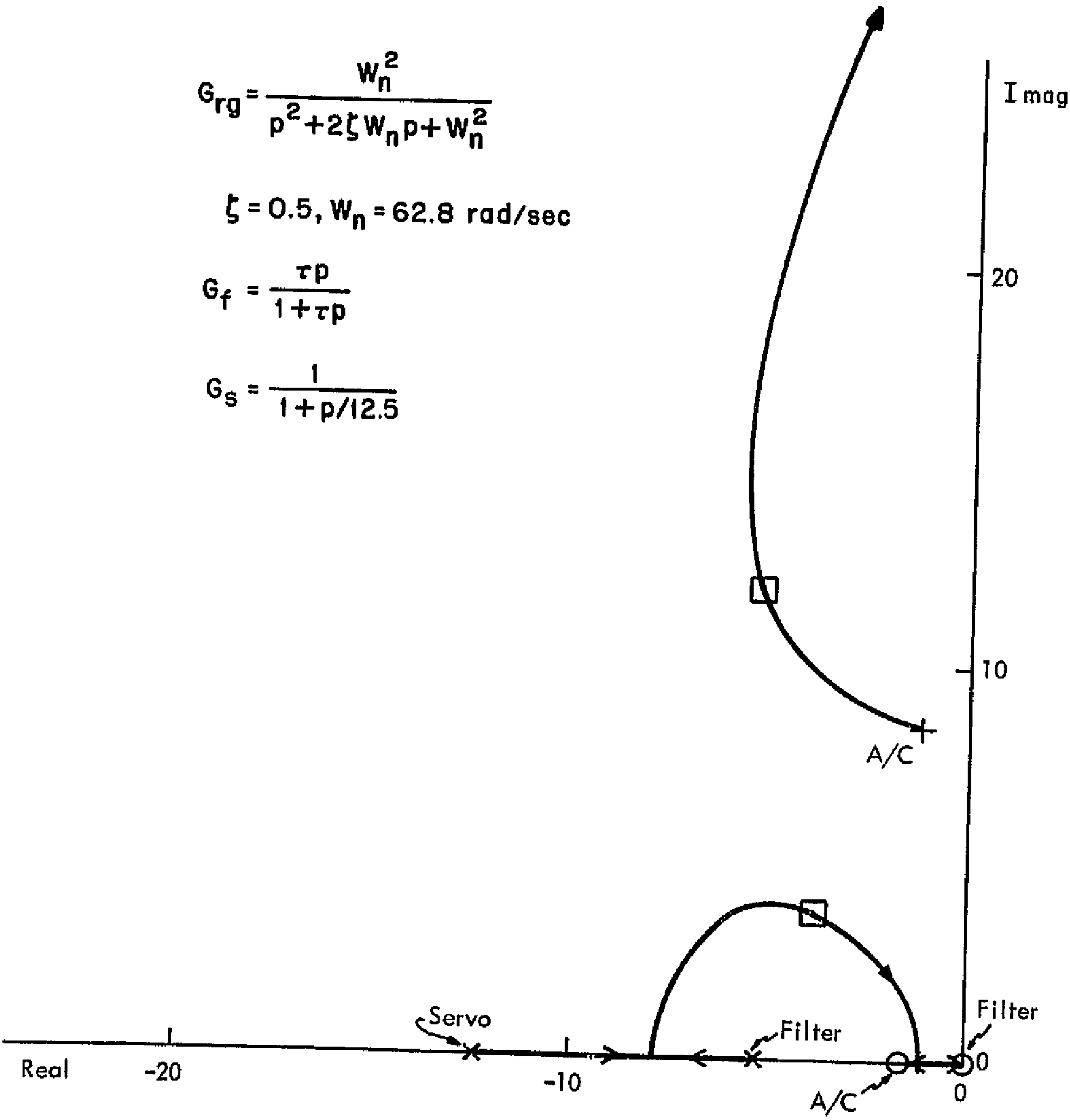


Fig. 4-7. Root Locus. Pitch Damper System. Final Design.

system having essentially similar time response characteristics as obtained with the first design. The program permits one to perform the optimization specifying parameter constraints, such as a maximum time constant or a maximum ratio of pole and zero of the lead filter if one so desires.

4.2 C* Control System

Another example of an aircraft longitudinal control system is a C* fly-by-wire system. Using a suitable controller, the pilot generates a command signal which becomes the input to a full authority automatic control system. The system produces a pitch angular velocity and normal acceleration proportional to the input command signal. A linear combination of these quantities defines the C* quantity as

$$C^* = -0.102\, a_z + k\, \dot{\theta} \tag{4-8}$$

where

a_z = the z-axis component of the incremental acceleration in m/sec^2

$\dot{\theta}$ = incremental pitch rate in rad/sec

k = arbitrary constant expressing the relative weighting of the two terms; in this example k = 10.062 g-sec

The dimensional unit for C* is numbers of g, the acceleration of gravity.

C* has been proposed as a handling qualities criterion (Ref. 8) using a tolerance boundary for a step function C* time response as shown in Figure 4.8. If one accepts this C* specification as the desired performance for the control system, one would specify a model whose step response was contained within the boundaries shown. A further specification was assumed, for illustrative purposes, which stated that the closed-loop static sensitivity of the system in terms of C* per unit input command signal should remain constant over the

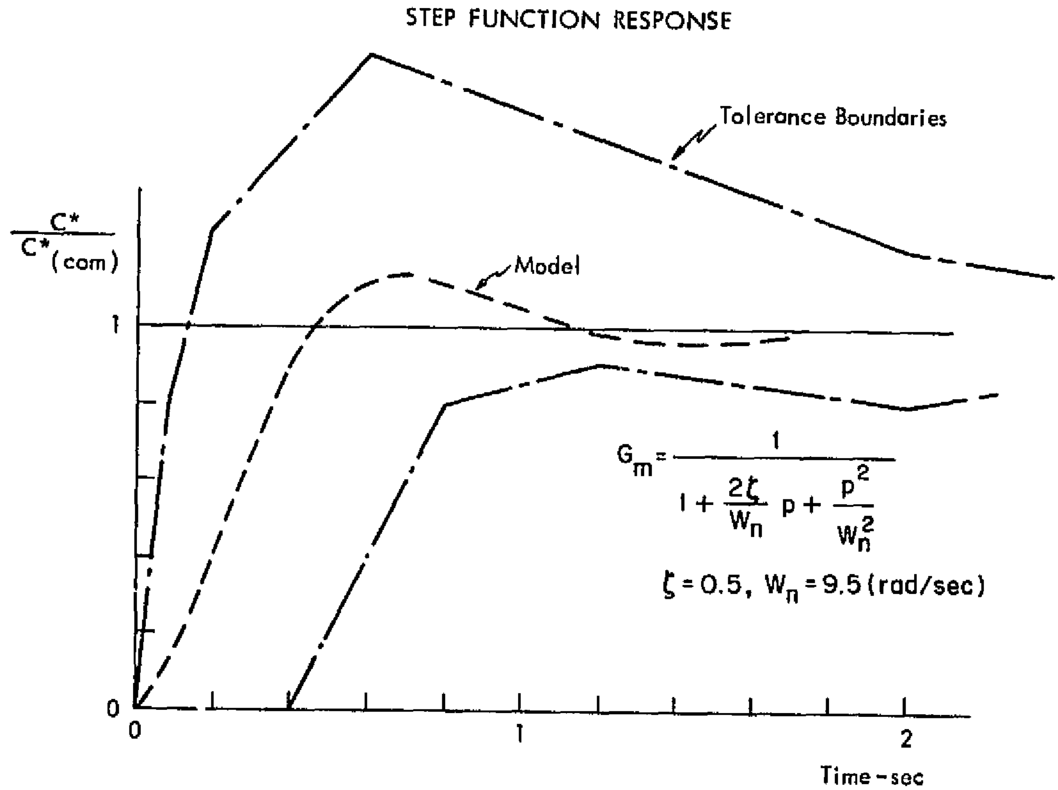


Fig. 4-8. Specification for Allowable C* Step Function Response to Meet Handling Qualities Requirements and Selected System Model Response

flight envelope. It is noted that the latter specification is not required in order to use the parameter optimization technique discussed in this report. One could, for example, have required a constant stick force per g specification. If one had not placed a specification upon the variation of static sensitivity with flight condition, somewhat simpler systems would have been possible. Since a steady-state C* requires that both pitch rate and normal acceleration are constant, a system that controls either of these quantities could also be calibrated in terms of C*, and the C* criterion could be used as the design requirement. The pitch damper system previously discussed is one such system. The parameter optimization technique could then have been used to examine the performance capabilities of the simpler configuration so as to achieve a design of minimum complexity. An example design sequence of this type was reported in Reference 9.

Returning to the present example, the static sensitivity requirement necessitated that the system have a C* error summation point and integral compensation of the error. Thus two sensors were required so that C* could be generated. Since both a_z and $\dot{\theta}$ were then available to the system information processing subsystem, these quantities could also be used to form feedback compensation loops. Accordingly the system configuration presented by the block diagram of Figure 4.9 was specified.

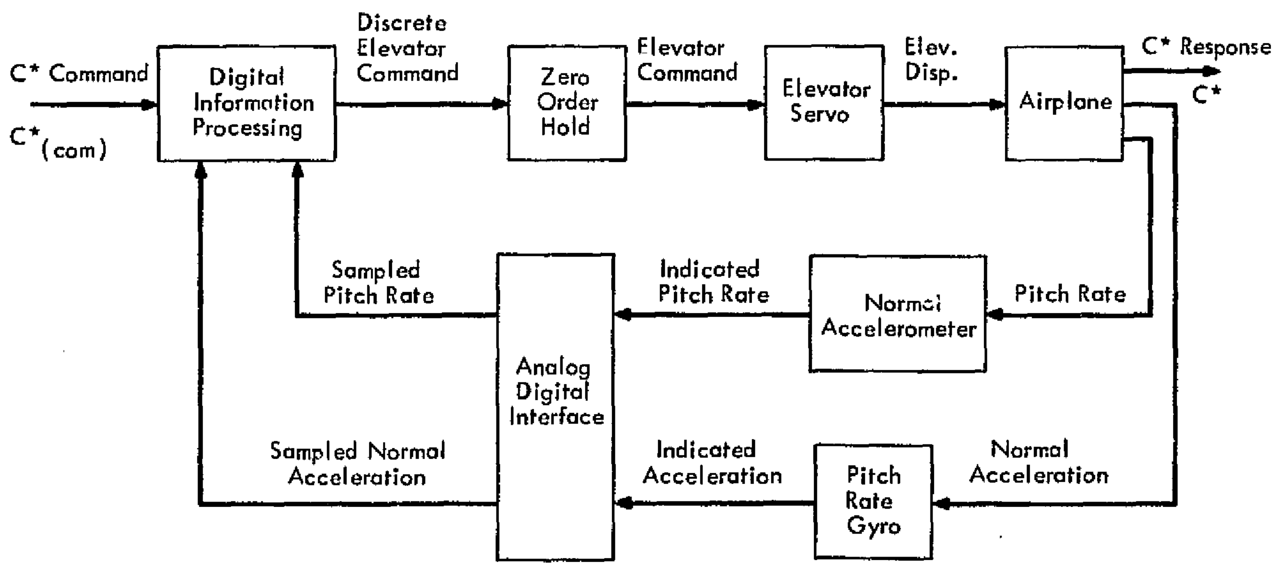


Fig. 4-9. Functional Block Diagram. Digital C* Control System.

This design example is used to illustrate a digital flight control application. A continuous version of the same system has also been examined, and for sampling frequencies of the order of 30 samples/second the performances of the sample data system and the continuous system design were virtually indistinguishable.

The mathematic block diagram for the system is shown in Figure 4.10. The separation of discrete and continuous system elements is denoted by the digital/analog interface dotted line. Sensor dynamic lags have been neglected. The optimization program permits one to examine the use of digital compensation filters. The recommended procedure is to begin the design sequence with the simplest system and to add complexity only as required. Thus the simplest filter embodiment is just a constant gain, and for this airplane that was sufficient as will be seen. The trapezoidal integration algorithm was used for the integration element shown for the C* error compensation (Ref. 5).

C* was formed by combining signals obtained from a normal accelerometer and a rate gyro. The C* error signal was formed and integral compensation was used to eliminate any steady-state error. Pitch rate and normal acceleration signals were gain-compensated and summed with the compensated error signal to produce the elevator servo input command signal. All of the signal processing was assumed to take place in a digital computer, and thus sampling devices were assumed at the outputs of the sensors and the pilot command generating controller.

The design parameters whose values needed to be determined were the open-loop static sensitivites, S_i, of the three control loops and the static sensitivitiy of the compensation integrator. The forward loop gain has been located within the acceleration inner loop, so that the design parameter modifying the accelerometer signal is the sensitivitiy ratio, $(SR)_{a_z}$, of the open-loop calibration sensitivities, $S_{a_z} = \left(\frac{\delta_e}{a_z}\right)_{OL}$ and $S_{C*} = \left(\frac{\delta_e}{C*}\right)_{OL}$. The latter are the

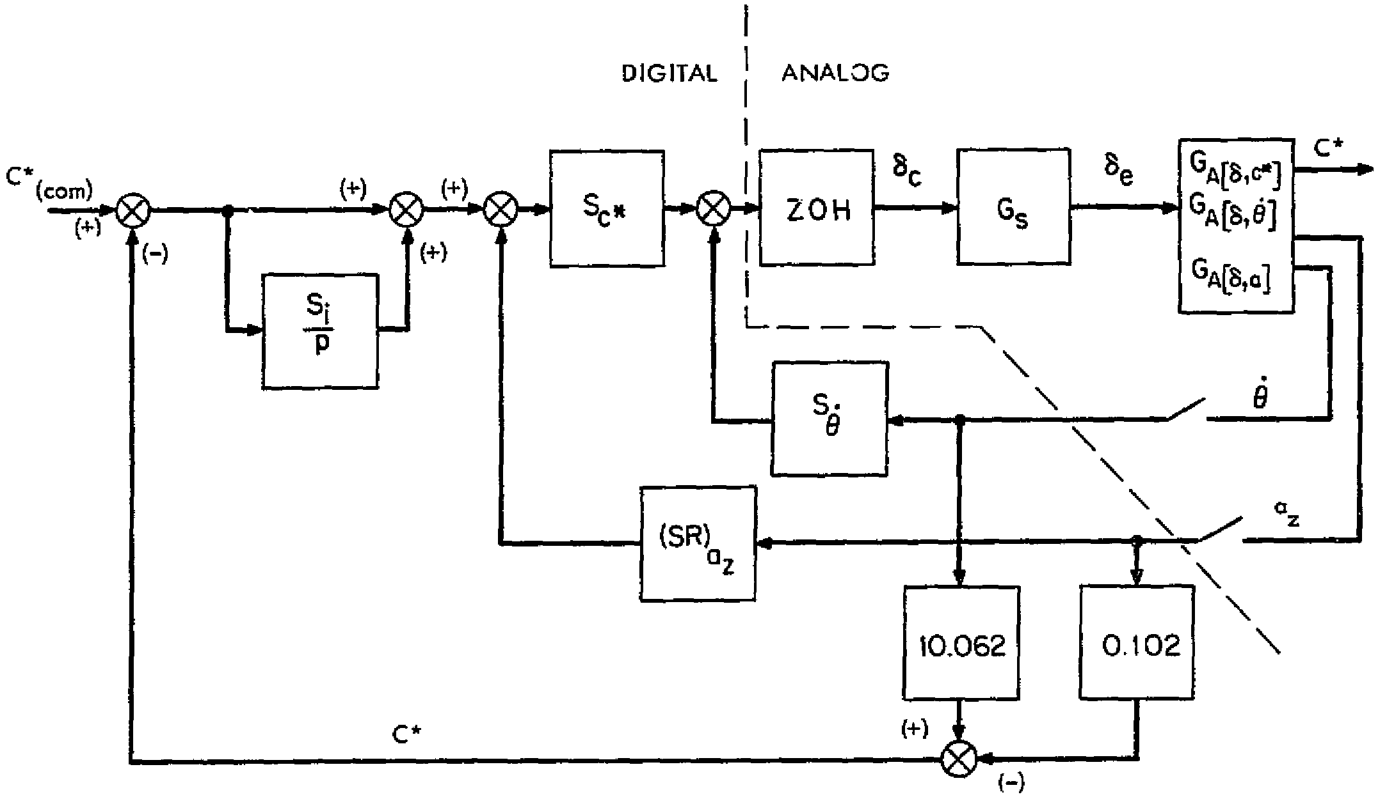


Fig. 4-10. Mathematical Block Diagram. C* Control System.

overall loop calibrations of the cascade of elements from the aircraft output motion through to the elevator surface deflection with the other loops open.

The model was chosen so that its step response was contained within the tolerance boundaries shown in Figure 4.8. Since the dominant system mode was the airplane longitudinal oscillation and also since it was desirable to keep the model simple, a second-order model with no zeroes was selected. Since overshoot in C* response is expected and indeed desirable, a model damping ratio of 0.5 was chosen resulting in approximately an 18% overshoot. To locate the time for the peak overshoot an undamped natural frequency of 9.5 (rad/sec) was chosen. The model poles in the complex plane then were located at ($-4.75 \pm 8.23j$). The digital program required specification of the Z-plane transfer function for the model. This was obtained from the p-plane transfer function using the integrating operators of Reference 5 and a sample period of 0.03 sec. This gave

$$G_m(z) = \frac{0.21192z}{z^2 - 1.63360z + 0.84522} \qquad (4\text{-}9)$$

Poles: $z = 0.81680 \pm 0.21849j$

The model step response is shown in Figure 4.8.

Nine flight conditions for the F-8 airplane were investigated. These are identified in Appendix A, Table A-1. The aircraft state equations are also presented in Appendix A. The design procedure followed was: (1) optimize the 4 parameters at each of the flight conditions; (2) examine the optimized parameter sets to ascertain whether the flight condition variation could be suppressed; (3) fix those parameters where step 2 indicates that is feasible; (4) re-optimize the remaining parameters at the various flight conditions.

Section 2.2 referred to the effects of integral error compensation of the type being proposed in this investigation. The integrator supplied the open-loop pole at the origin needed to insure a zero static error in C*. With the design freedom that the four parameters permitted in this example, the parameter optimization acted so as to eliminate the need for the integrator. It did this by causing the airplane to become an integrator by calling for a regenerative a_z feedback loop. The effect of this sign of a_z feedback was to reduce the static stability of the airplane, causing the natural frequency of the oscillatory poles to decrease until they became two real poles, one of which migrated to the origin. See a similar trend in the root locus for the pitch damper example of Figure 4.4. Stability of the overall system was assured by the $\dot{\theta}$ and C* feedback loops. When that inner loop adjustment took place, the compensation integration gain became a very insensitive parameter as one would expect. Hence the parameter optimization preferred such a solution incorporating a relatively high regenerative a_z feedback.

That solution however may not be attractive. The amount of a_z feedback had to vary with flight condition, since the airplane poles were a function of flight condition. Although it was not investigated, it would seem that it would be possible that the effects of partial system failures could leave the pilot with a neutrally stable or even an unstable airplane. Therefore, as a further design modification the a_z feedback was constrained to a value which the previous results indicated would still keep the inner a_z loop stable by itself. This was done by selecting the ratio of the a_z feedback gain to the forward loop gain to be 0.5. When the remaining three parameters were re-optimized the low $\bar{q}$ flight condition required an integrator gain of 1.26 sec^{-1} while at some of the other flight conditions at which the aircraft was more of an integrator due to the a_z feedback, relatively little integrator gain was needed. Since the C* specification boundaries permitted a step function overshoot, the integrator gain was fixed at 1.26 sec.$^{-1}$ so that the worse flight condition was covered. The remaining conditions were then examined to determine if the optimization of the remaining two parameters would satisfy the specifications.

As an additional simplification, in the final design iteration the rate feedback path was eliminated (parameter set to zero), and this left only one design parameter to optimize. The variation of the optimized forward loop gain with flight condition is summarized in Figure 4.11 as a plot of $(S_{C*})^{-1}$

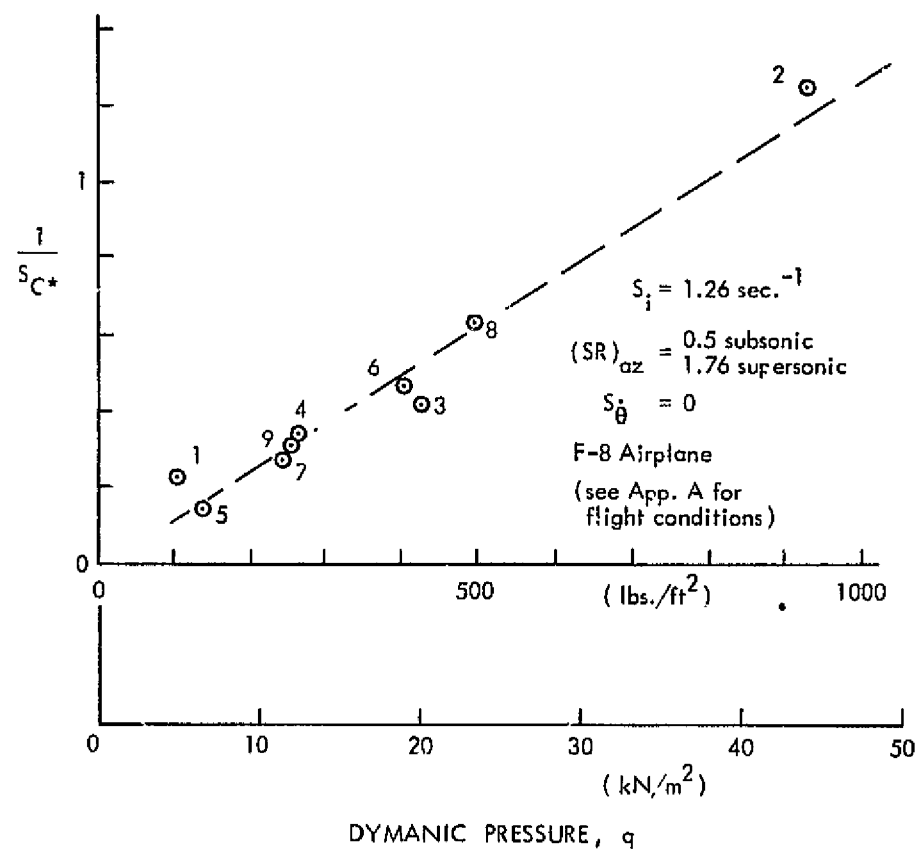


Fig. 4-11. Variation of Forward Loop Static Sensitivity with Flight Condition for the Optimized C* Control System.

versus dynamic pressure. The step function responses of the resultant designs for the various subsonic flight conditions are presented in Figure 4.12. For the supersonic flight conditions, the a_z feedback sensitivity ratio was increased to 1.76. The step responses for the three supersonic flight conditions are presented in Figure 4.13. The performance specifications were satisfied with a relatively simple control system. The responses exhibited greater overshoot than that of the model at most flight conditions due to the closed-loop zero introduced by the integrator compensations but that only brought the responses closer to the mid-values of the tolerance boundaries, and that would be satisfactory.

The design iteration was stopped at this stage. Perhaps further design simplification would have been possible, since the margin of the responses relative to the specification boundaries was fairly large and could have been traded-off against reducing the need for a two-level a_z gain variation subsonic to supersonic. Linearizing the forward loop gain variation with dynamic pressure of Figure 4.11 would be easy to implement, and a final check of the resulting performance would need to be made. Note that the parameter optimization approach permitted one to include several engineering design constraints in a straightforward manner. By correlating the optimization results with classical design tools such as the root locus, one can generate clearer understanding of why the parameter selection gives the performance achieved than is apparent from the computer output of an optimization run alone.

4.3 F-8 Lateral-Directional System

<u>General Features of Lateral-Directional Control</u>. The control of roll angle of an airplane provided an example of a control system that utilized two control

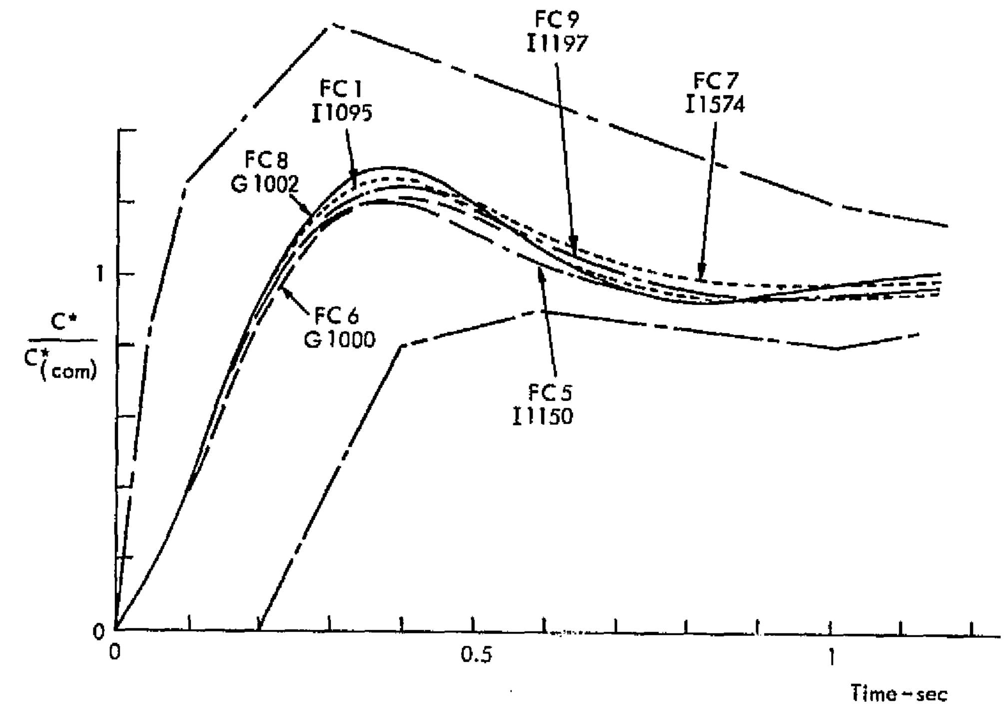


Fig. 4-12. Step Function Responses. C* Flight Control System. Subsonic Flight Conditions.

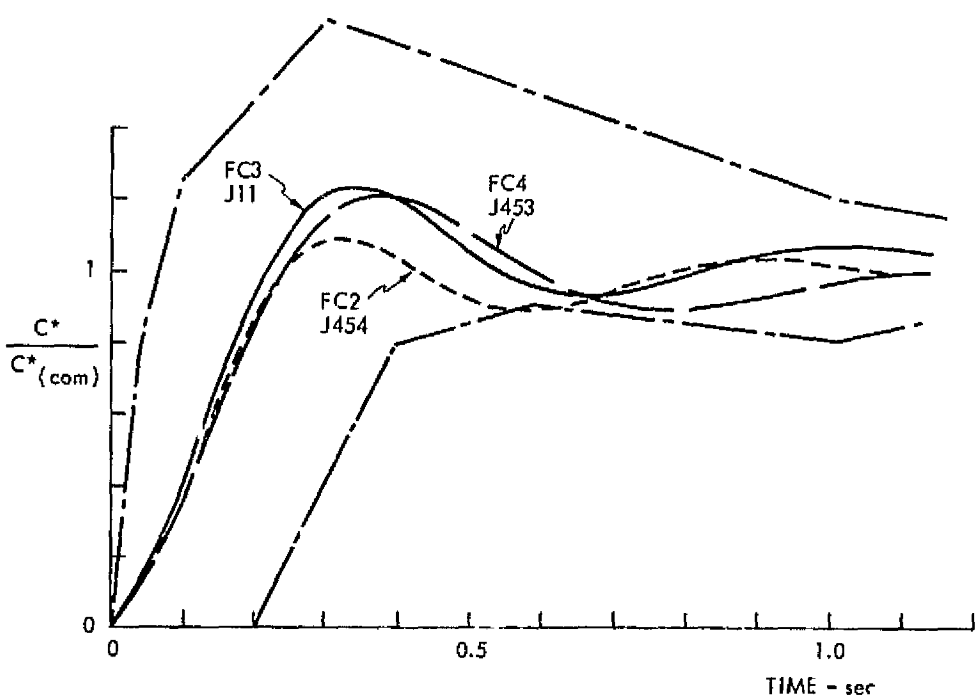


Fig. 4-13. Step Function Responses. C* Flight Control System. Supersonic Flight Conditions.

effectors, the aileron and the rudder. The primary method of generating the forces needed to change the direction of the airplane's velocity vector in the horizontal plane is to roll the airplane about its longitudinal axis so that there will be horizontal component of lift. This horizontal force component will then produce an angular velocity of the velocity vector, that is, a rate of change of the direction of the flight path traversed by the airplane's center of mass. The primary control effector for exerting the rolling moments needed to roll the airplane is the aileron. In addition to translation, the airplane has angular degrees of freedom, and the angular heading of the airplane's longitudinal body axis may deviate from the direction of the velocity vector thereby generating an angle of sideslip. The geometrical quantity of importance in the directional control of the airplane is this sideslip angle, which is defined as the angle between the velocity vector and the projection of the velocity vector onto the airplane's plane of symmetry. The velocity vector referred to here is the velocity of the airplane with respect to the air mass.

For reasons of obtaining acceptable handling qualities, it is desirable to perform rolling maneuvers in such a way that sideslip is kept small. If sideslip were kept zero, the velocity vector would remain in the plane of symmetry as the airplane rolled. To the pilot the airplane would appear to roll about the direction of flight. If there were a tendency for the airplane to roll about an axis inclined to the velocity vector by some angle of attack, sideslip would then be generated as the rolling motion converted angle of attack into sideslip. To the pilot in that case the airplane's direction of flight may appear initially to turn opposite to the direction of the turn he wished to establish. For this reason and because sideslip generates side forces which are less comfortable to the crew and passengers, sideslip is objectionable. While minimizing sideslip is usually considered to be the criterion for rudder coordination, there is evidence that the attendant improvement in yaw angular velocity is the more important consideration to a pilot (Ref. 10). The primary means for minimizing sideslip is the deflection of the rudder control. The rudder dominantly produces yawing moments which can be used to generate the yaw angular velocity components needed to minimize sideslip when the aileron input alone would cause the airplane to roll about an axis not aligned with the velocity vector. The process of deflecting the rudder in conjunction with the aileron is termed rudder coordination.

For small angles of roll the steady-state turning angular velocity is proportional to the roll angle. If one aligns a set of body axes so that the longitudinal x-axis is parallel to the velocity vector in trimmed flight (a stability axis set), the yaw angular velocity component found along the Z-axis will then be proportional to roll angle. If sideslip is kept zero during the rolling transient and altitude is maintained, that component of yaw rate is the only one present, since the rolling motion takes place about an axis perpendicular to the yaw axis.

The example to be presented is primarily an illustration of the use of the parameter optimization technique to design a rudder coordination subsystem for an F-8 fighter airplane lateral autopilot. In the process the roll angle control system was also designed. The airplane has two control inputs, the aileron and the rudder deflections. Each input produces responses in sideslip, roll angle, and yaw angular velocity, and therefore one is dealing with a multi-input, multi-output controlled member. Synthesis of such a control system is difficult, though certainly not impossible, by classical automatic control system design techniques. From a transfer function standpoint, closing a feedback loop from any output, i, to the first control surface changes not only the closed-loop poles, but also changes the zeroes of the transfer function relating a second output, j, to the second control surface input because of the cross coupling involved. While it is not necessary to call for a design resulting in completely non-interacting controls, advantage can be taken of the fact that the ailerons have been designed dominantly to produce rolling moments while the rudder has been designed to produce yawing moments. By the very nature of the characteristics of an airplane, if the rudder is doing an adequate job of minimizing sideslip, the rolling characteristics are dominantly affected by feedback to the aileron alone. Thus certain feedback paths to the aileron can be identified with roll angle control while other feedback paths to the rudder are associated with the rudder coordination system.

The natural response modes of the airplane are the spiral mode, the roll subsidence mode, and the lateral oscillation. Large excitation of the lateral oscillation during rolling maneuvers is undesirable. This mode usually is underdamped, and typically yaw rate is fed to the rudder to increase the damping ratio of the mode. Finally it is also operationally desirable to minimize the steady-state sideslip which may exist due to steady bias yawing torques that may arise from a variety of sources.

<u>Design Procedure</u>. With these operational requirements in view, a design procedure is suggested as follows: (a) provide a yaw rate feedback path to the rudder to increase the damping ratio of the lateral oscillation to an acceptable level, e.g., 0.5. (This is readily done from a root locus analysis inasmuch as it is a very simple feedback loop.) (b) select a roll angle

response model; (c) provide roll angle and possibly roll rate feedback paths to the aileron and use the parameter optimization program to select an initial set of open-loop static sensitivities (gains); (d) provide a rudder coordination signal path structure and using the same roll model as in step (b), use the parameter optimization program to select an initial set of coordination system parameters considering roll angle command as the input and yaw rate (stability axis) as the output quantity; (e) using the coordination system parameters found in step (d), reoptimize the roll angle control system parameters; (f) using the new roll angle parameters, re-optimize the rudder coordination system; (g) continue the iteration as needed. As the coordination system performance improves, it is found that the interaction between the two sub-systems decreases, and convergence to an acceptable design is rapid. It is then found that the rudder coordination system parameters primarily affect the locations of the zeroes of the closed loop system transfer function relating yaw rate to the roll angle input command with little effect upon the poles. The roll angle parameter optimization changes the pole locations of the transfer function relating roll angle to the roll angle command and, depending upon the compensation structure specified, may or may not change the closed-loop zeroes.

The above discussion lays the background for selection of the model used in the parameter optimization program for designing a rudder coordination system. The model expresses the desired transient behavior of the system. If one were to state as an operational requirement that sideslip should be kept zero, there would be no desired transient response, and one could select coordination design parameters so as to minimize the integral squared sideslip response. As is seen in the examples presented below, such a system design is not necessarily the best from an operational point of view. An alternative expression of desirable characteristics based on the previous discussion is to state that the yaw angular velocity, as seen in a stability axis coordinate frame, be proportional to the roll angle as the airplane maneuvers to establish a commanded roll (or bank) angle. The model in this case is a model of the desired yaw angular velocity transient response, and it can be obtained from a specification of the roll angle response characteristics. Such a choice for design optimization leads to a better angular velocity response with less excitation of the lateral oscillatory mode (Dutch roll) at the expense of a somewhat greater magnitude of peak sideslip than does minimizing the root mean square sideslip.

A Continuous System Design. The above procedure was applied for two flight conditions for the F-8 airplane. A control system configuration was specified which resulted in the functional block diagram for the system of Figure 4.14. Roll angle and roll rate were fed to the aileron without additional dynamic compensation. Yaw rate was fed to the rudder to improve the damping of the lateral oscillation. To avoid a steady-state rudder deflection during turns, the indicated yaw rate signal was modified by a first-order high-pass filter. To provide rudder coordination a compensated aileron to rudder interconnect signal path was provided. Various compensation dynamics are possible, but perhaps the simplest is one whose transfer function has a single pole, a single zero, and a static sensitivity. To reduce steady-state sideslip due to steady-state yawing moments, the integral of sideslip was provided as a slow acting sideslip trim path. Use of sideslip to achieve high frequency sideslip control is unattractive due to the difficulty in measuring dynamic sideslip and the reduction of damping of the lateral oscillation that it

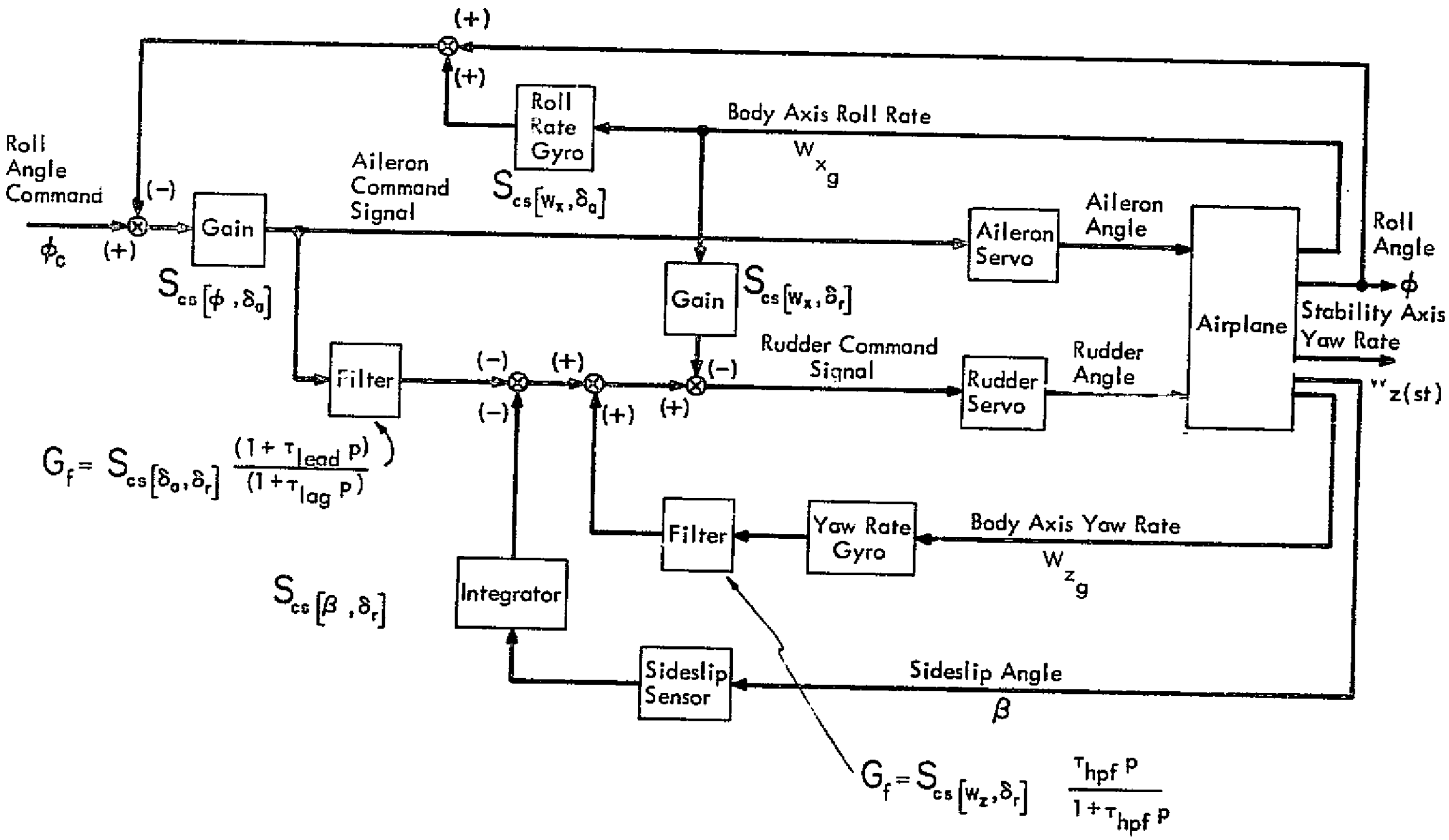


Fig. 4-14. Functional Block Diagram. Lateral-Directional Flight Control System.

produces. The roll and yaw rate signals were assumed to be obtained from rate gyros whose input axes were aligned with an orthogonal set of body axes whose X-axis was displaced from the stability axis coordinate frame X-axis by the trim angle of attack. The dynamic equations of motion for the airplane found in Appendix A are written in this body axis coordinate frame.

The aileron servo was modelled as a second-order lag with undamped natural frequency of 34.64 rad./sec. and a damping ratio of 1.0103. The rudder servo was a first order lag with the pole at -25.0 rad./sec. The complete system was 10th order.

The yaw angular velocity, as seen in a stability axis coordinate frame (see Figure A-1), has components of both of the body axis roll and yaw rates as given by equation (4-10)

$$W_{z(st)} = -W_x \sin \alpha_0 + W_z \cos \alpha_0 \qquad (4\text{-}10)$$

The steady-state yaw rate is

$$W_{z(st)ss} = (g/v)\, \phi_{ss} \qquad (4\text{-}11)$$

where ϕ_{ss} is the steady-state Euler angle, ϕ, measured as a rotation about the X body axis, and V is the true airspeed. Equation 4.2 shows that at high speed the steady-state stability axis yaw rate, which is also the desired maximum yaw rate, will be small relative to the maximum roll rate that can result during the rolling maneuver. In a poorly coordinated maneuver, the roll rate term in equation (4-10) may initially exceed the yaw rate term giving rise to the negative stability axis yaw rate that pilots find objectionable. Accordingly, it was also decided to provide a roll rate to rudder path, and let the parameter optimization determine whether or not it would be beneficial as part of the rudder coordination system.

Model Selection. The input to the control system as shown in Figure 4.14 is the roll angle command signal. For a step function input the system is called upon to establish a steady-state roll angle proportional to the command with a dynamic transient response approximately the same as the step function response of the selected model. In order to have a desirable yaw angular velocity behavior, the time response of the normalized yaw angular velocity as seen in a stability axis coordinate frame would also be approximately the same as that of the roll angle. The steady-state values differ, of course, as seen in equation (4-11). One desires to keep sideslip acceptably low during the maneuver and to reduce the excitation of the airplane's lateral oscillatory mode.

Since the objective of using this example was to illustrate the use of the parameter optimization program for designing a rudder coordination system, the selection of the model for the roll angle command system was arbitrary. The model transfer function selected was

$$G_m(p) = \frac{1}{1 + \frac{2\zeta}{\omega_n} p + \frac{p^2}{\omega_n^2}} \qquad (4\text{-}12)$$

$\zeta = 0.7 \qquad \omega_n = 1.0$ (Rad/sec)

This provided the step function response presented in Figure 4 16 and exhibited a response time of approximately 3 seconds.

Design Parameters. The design parameters to be selected were the open-loop static sensitivities and the time constants of the aileron to rudder interconnect path compensation filter. Since the static sensitivities of the airplane were fixed, the open-loop static sensitivities were specified by expressing the sensitivity of the portion of the loop from the ouput quantity of the airplane through to the control surface displacement. When more than one quantity was fed to a given control surface, the calibration specified for a particular path assumed the other paths to that point were open. The design parameters were

Aileron control paths

$$S_{cs[\phi,\delta_a]} = \left(\frac{\delta_a}{\phi}\right)_{ss} \qquad S_{cs[W_x,\delta_a]} = \left(\frac{\delta_a}{W_x}\right)_{ss}$$

Rudder control paths

$$S_{cs[W_z,\delta_r]} = \left(\frac{\delta_r}{W_z}\right)_{ss} \qquad S_{cs[\int\beta,\delta_r]} = \left(\frac{\delta_r}{\int\beta dt}\right)_{ss}$$

$$S_{cs[\delta_a,\delta_r]} = \left(\frac{\delta_a}{\delta_r}\right)_{ss} \qquad S_{cs[W_x,\delta_r]} = \left(\frac{\delta_r}{W_x}\right)_{ss}$$

$$\tau_{hpf}\ ,\quad \tau_{(lead)}\ ,\quad \tau_{(lag)}$$

Of these nine parameters the two aileron path parameters were selected by the parameter optimization of the roll angle response. Four of the rudder path parameters were used in the optimization of the stability axis yaw rate response. The control system design was not sensitive to the characteristics of the yaw rate and the integrated sideslip paths to the rudder, and these paths were designed arbitrarily and held fixed. Thus the parameters used to specify the rudder coordination system were

$$S_{cs[\delta_a,\delta_r]},\ S_{cs[W_x,\delta_r]},\ \tau_{(lead)},\ \tau_{(lag)}$$

For the two flight conditions investigated, the optimization placed the lead zero at rather high frequency relative to the bandwidth of the model, and so in the final design iterations the system was further simplified by eliminating the lead zero and reducing the number of optimization parameters to three.

In optimizing the rudder coordination system parameters, the output quantity of interest was the yaw angular velocity as found in a stability axis coordinate frame. This quantity was formed using equation (4-10). In attempting to minimize sideslip during a rolling maneuver, one is basically asking the rudder to operate so as to suppress one of the aircraft's degrees of freedom. This is to a large extent a process of adjusting the zeroes of the roll input command to yaw angular velocity transfer function. It is not surprising that the zero locations are very sensitive to the rudder coordination design parameters. Indeed the zeroes migrated very rapidly over both the left and right halves of the complex plane as the parameters were varied.

From the discussion of the performance index of Section 2.3, the zeroes strongly affect the initial response of the error excitation function, i(t). Thus if the performance index lower integration limit was zero, the performance index was also sensitive to the design parameters, and the slopes of the performance index in parameter space were very large and sensitive. Although the system was 10th order, if the aileron to rudder interconnect path contained a zero, the pole-zero excess of the system roll command to yaw rate transfer function was only two. Without the interconnect path zero, the excess was three. Thus the model excess could only be two or three respectively. The pole-zero configuration of i(p) was such that the high frequency gain was so large that even the minimum value of the average performance index could be of the order of 10 rather than 0.10. This caused two adverse effects. Most importantly, the optimization traded-off low frequency response behavior against the high frequency initial state response, usually resulting in poor sideslip response and excessive excitation of the airplane's lateral oscillation. Secondly, the very large (PI) values due to these high frequency terms increased the requirements upon the significant figures to be retained to maintain acceptable computational accuracy. As was discussed in Section 2.3, one needs to trade-off the effects of the initial mismatch of system and model state vectors versus matching of the state space trajectories during the rest of the transient response. This can be done through an iterative variation of the lower integration limit of the performance index. After three trials, a lower integration limit of 0.05 seconds, which was approximately (1/60) of the model and system response time, was sufficient to lead to an acceptable system design.

Design results. Two flight conditions were investigated: M = 1.6 at 12192 m. altitude and M = 0.56 at 6096 m. altitude. The former was a test case probably never to be encountered since it corresponded to a roll during a pull-out maneuver at the extreme corner of the airplane's flight envelope. The second was a moderate Mach number cruise condition.

The yaw damper loop was designed using the root locus presented in Figure 4.15. The high-pass filter time constant was specified to be 1.0 second, a value that in current design practice is found to be reasonable. The integral of sideslip feedback gain to the rudder was set at 0.1 sec^{-1} on the basis of specifying an acceptable time for the reduction of steady-state sideslip bias. As a further check at the end of the optimization procedure, these three design parameters were added as optimization parameters

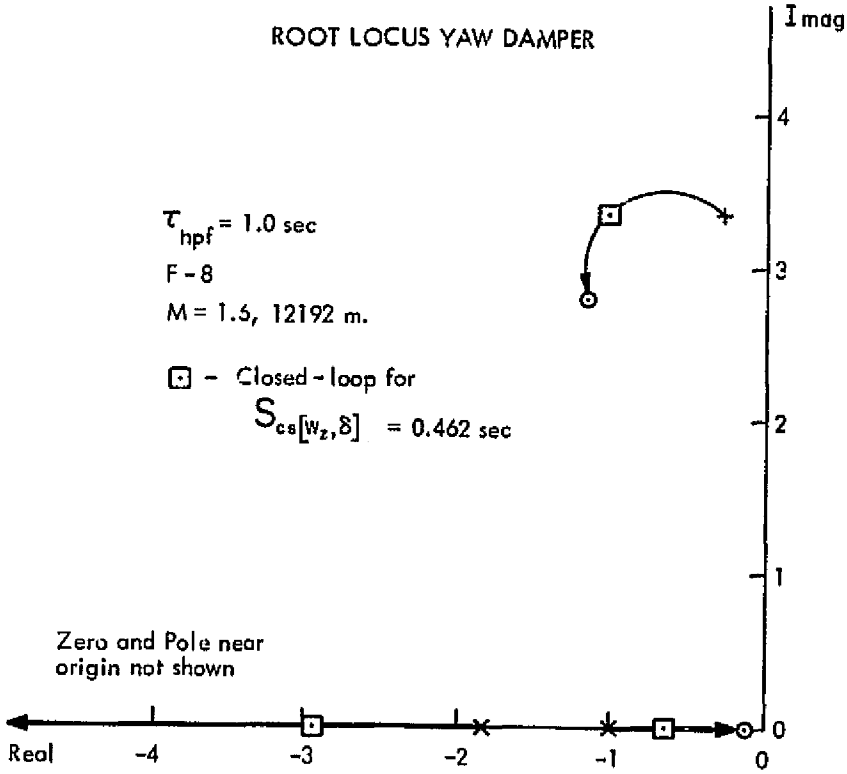
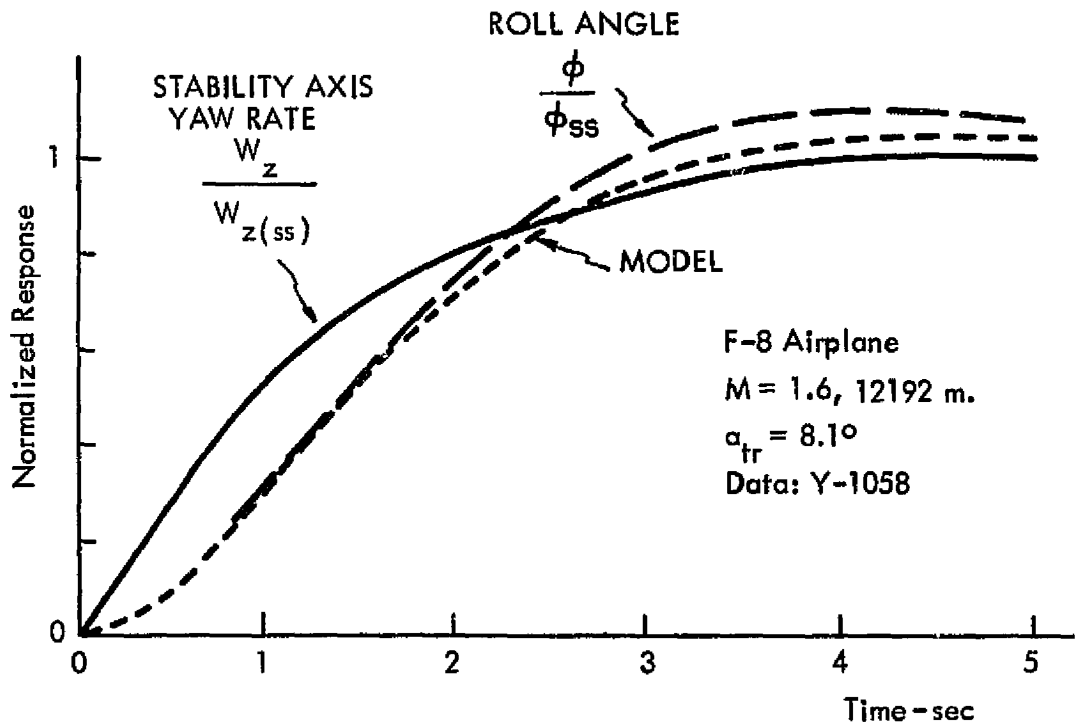


Fig. 4-15. Root Locus Design of the Yaw Damper Loop.

and found to be insensitive.

The performance of the final design is presented in Figure 4.16. The optimized parameter set is listed on the figure. The roll angle response was acceptably close to that of the model. Although the normalized yaw rate response deviated somewhat from the roll angle and resulted in the negative sideslip shown, there was little excitation of the lateral oscillation and no negative initial swing of the yaw rate. The maximum sideslip was only 0.5% of the commanded roll angle which is certainly acceptable. The most sensitive rudder coordination system parameter was the static sensitivity of the aileron to rudder path. Figure 4.16 presents the sideslip variation due to ±25% changes in this gain. The system sensitivity would appear to be



$S_{cs}[\phi_1,\delta_a] = 0.127,\quad S_{cs}[W_x,\delta_a] = 0.274$ sec, $\quad S_{cs}[\delta_a,\delta_r] = -0.666$

$\tau_{(lead)} = 0 \quad , \quad \tau_{(lag)} = 0.0180$ sec, $\quad S_{cs}[W_x,\delta_r] = 0.0671$ sec

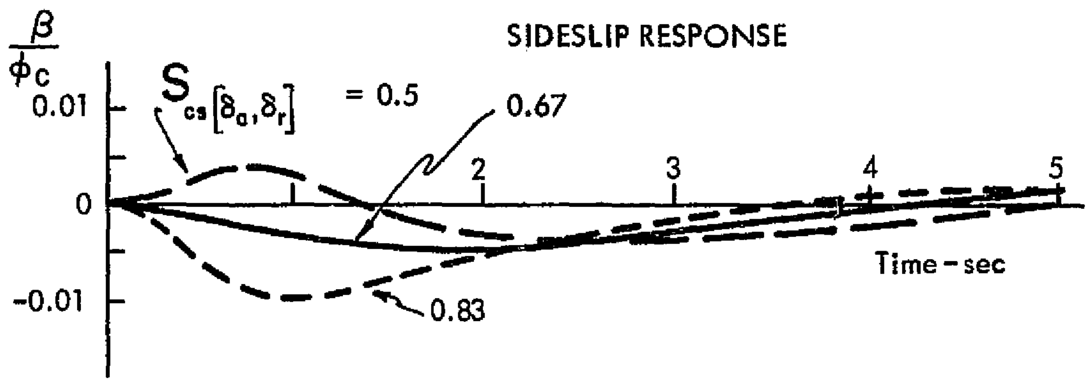


Fig. 4-16. Step Function Response. Lateral-Directional System. Continuous System Design.

acceptable.

Comparison of an RMS Sideslip Design. The method of evaluating the performance index (see Section 3.1) makes it easy to use as a performance index the integral of the square of the sideslip response. This was done for comparison purposes, and the response of Figure 4.17 resulted. The normalized yaw angular velocity was closer to the airplane roll angle response and the resultant sideslip was smaller than for the Model Performance Index design. However there was a much greater unwanted excitation of the lateral oscillation and an initial negative swing of the yaw angular velocity. Since the Model Performance Index evaluated the contributions of several output derivatives, it did a better job in reducing the oscillatory component of the response at the expense of greater sideslip. The optimization procedure thus presented one with a clear trade-off between sideslip response and yaw angular velocity response as the performance capability of a practical lateral control system configuration. The results were also indicating that to reduce the sideslip further would involve feedback of more states, a complication that for this application would not be warranted.

Change in Flight Condition. The system configuration obtained for the Mach 1.6 flight condition was then reoptimized for the 6096 m. altitude, Mach 0.56 condition using the same model. The system parameter values that resulted and the transient responses are presented in Figure 4.18. The system performance is comparable to that obtained with the Mach 1.6 condition.

Digital System. If one wished to use a digital computer for the information processing in such a lateral control system, the same design procedure leads

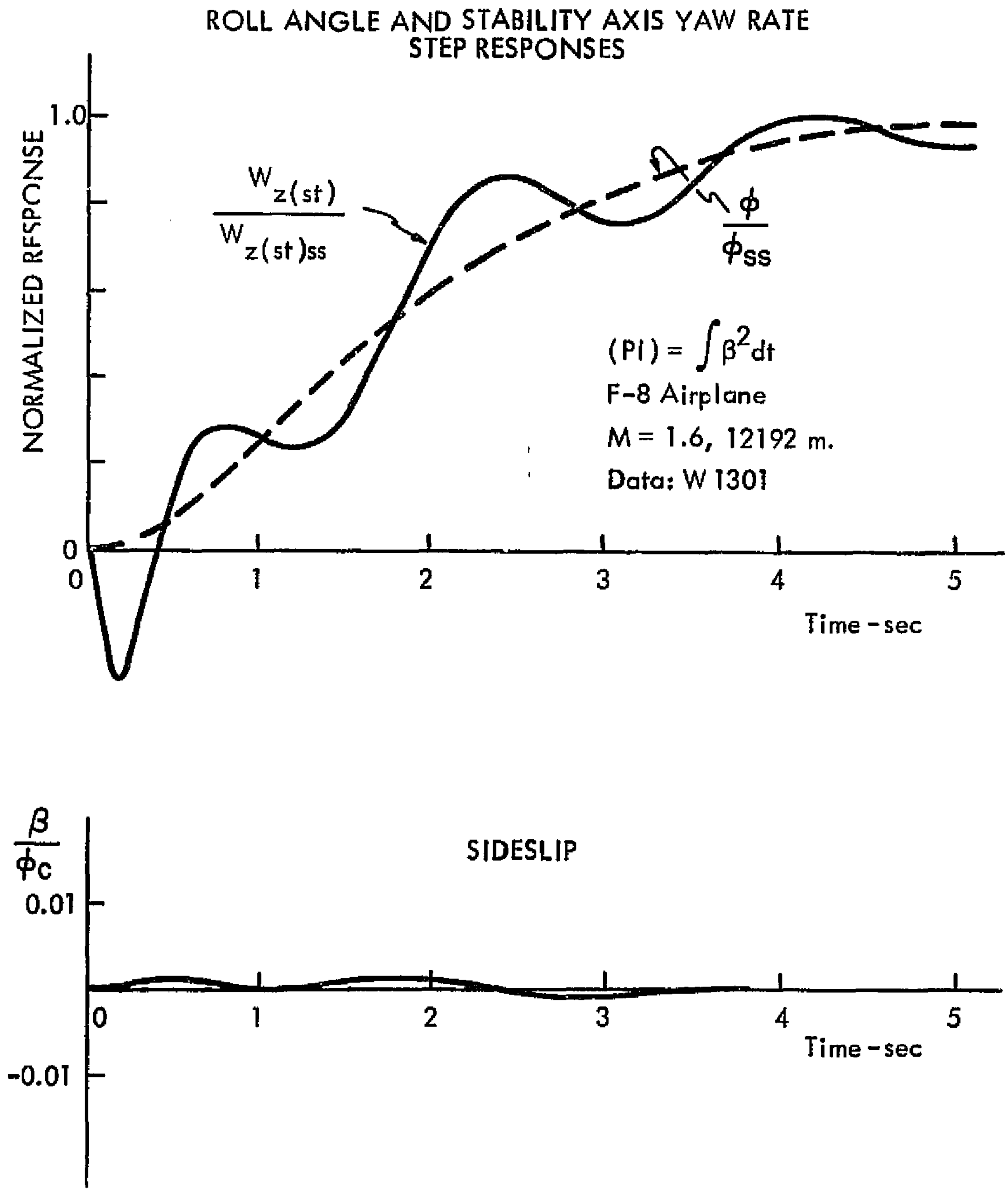
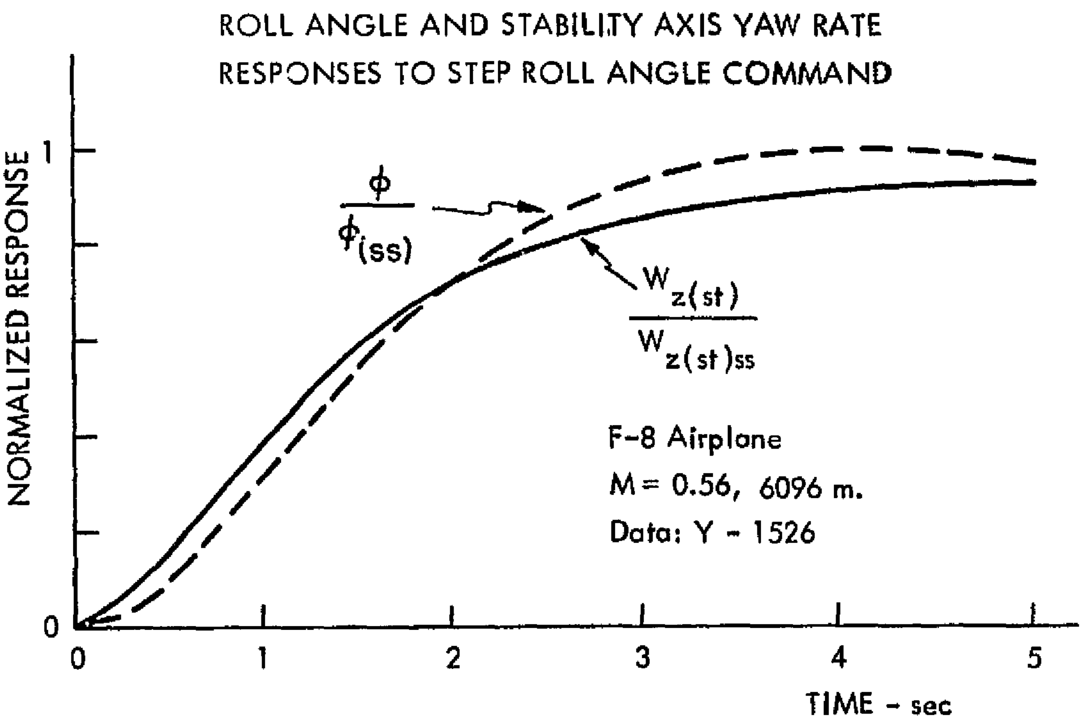


Fig. 4-17. Performance of Lateral-Directional System when Rudder Coordination Designed to Minimize the Integral Squared Sideslip.



$S_{cs}[\phi,\delta_a] = 0.0597$, $S_{cs}[w_x,\delta_a] = 0.0598$ sec., $S_{cs}[\delta_a,\delta_r] = -0.551$

$\tau_{(lead)} = 0$, $\tau_{lag} = 0.0153$ sec. , $\tau_{(hpf)} = 1.0$ sec.

$S_{cs}[w_z,\delta_r] = 0.461$

Fig. 4-18. Performance of the Optimized Lateral Control System at the Mach 0.56 Flight Condition.

to the system configuration shown in Figure 4.19. In this case the input data for describing the model was the Z-transfer function of the model. There is a chance for some ambiguity at this point, inasmuch as there are several ways of representing a desired continuous transfer function by a discrete algorithm (see Ref. 4).

To illustrate the difficulties one can encounter in specifying the Z-transform of the model, one can examine the straightforward Z-transform of the same second-order model that was used for the continuous sytem including a zero-order hold. One obtains for a sampling period of 0.03 seconds

$$G(z) = \frac{0.43808 \times 10^{-3}(z + 0.98596)}{z^2 - 1.95803z + 0.95890} \tag{4-13}$$

The zero, which is introduced by the sampling process, is difficult to interpret physically. If $i(z)$ is formed using equation (3-9), this model zero becomes a pole of $i(z)$ on the negative real axis near the point $z = -1$. Such a pole contributes an exponential sequence to $i(z)$ at the sample instants with successive values alternating in algebraic sign. When $[i(t)]^2$ is formed, this contribution appears as an exponential response mode similar to a very low frequency mode at $z = +1$. Such a mode contributes significantly to the time response, and hence to the performance index and can distort the optimization process. If one examined the model step response, it was seen that moving this objectionable model zero to the origin had very little effect upon the time response. The effect upon $i(t)$ then is that of a very high frequency negligible system pole. The model transfer function then was

$$G(z) = \frac{0.8700 \times 10^{-4}\, z}{z^2 - 1.95803z + 0.95890} \tag{4-14}$$

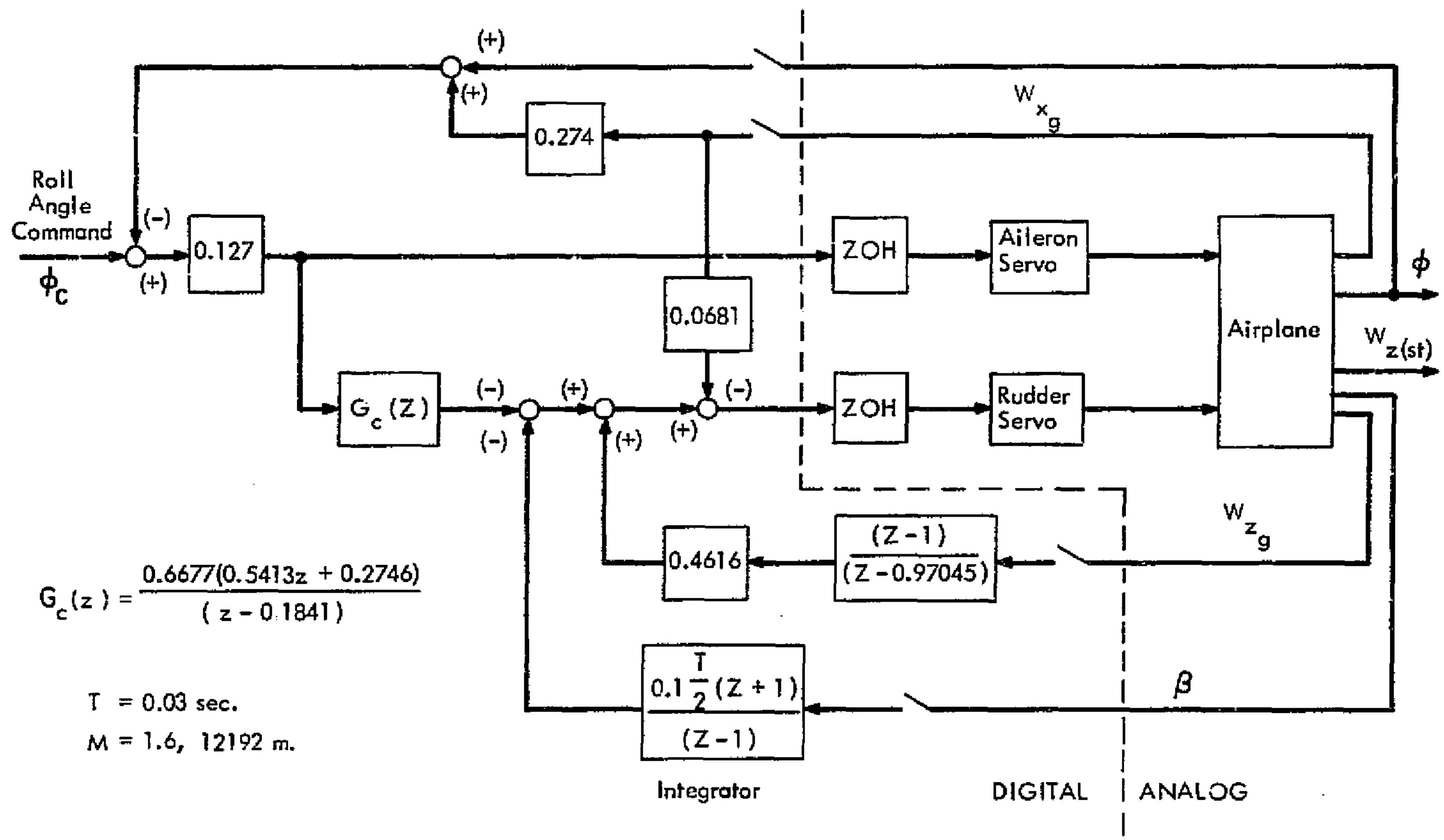


Fig. 4-19. Mathematical Block Diagram for a Digital-Lateral-Directional Flight Control System. Parameter Values Correspond to the Optimized Set.

with poles at

$$z = 0.97902 \pm 0.020986j$$

Note that this is the same form of the transfer function that would have resulted from using the integrating operators that were used in specifying the model for the digital C* system example of Section 4.2.

Similarly the form of the aileron to rudder compensation path is more complex in the Z-plane, since the sample data representation of a first-order pole introduces a zero as well as a pole. The general form for such a filter then is

$$G_c(z) = \frac{c_1(z + b_0)}{(z + a_0)}.$$

The high-pass filter in the yaw rate to rudder path became (for T = 0.03 sec)

$$G_f(z) = S_{cs[W_z\delta_r]} \frac{(z - 1)}{(z - 0.97045)}$$

and the integrator for the sideslip path was taken to be

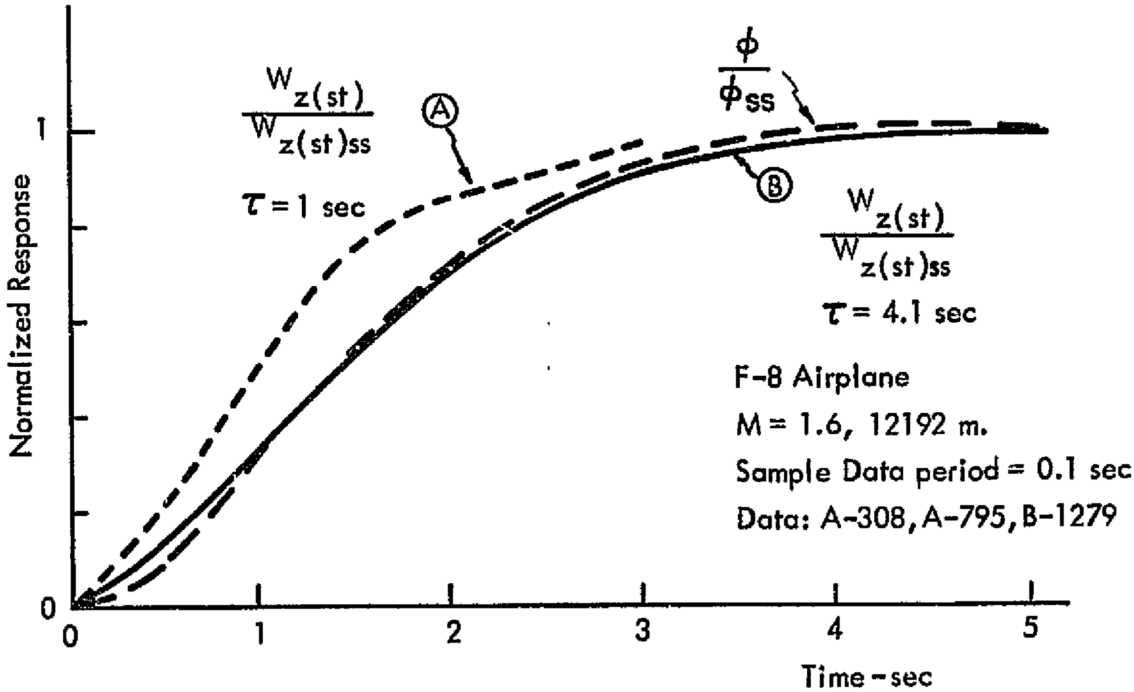
$$0.1 \frac{T}{2} \frac{(z + 1)}{(z - 1)}$$

Specifying the aileron to rudder compensation filter pole and zero and the same loop gains as were used in the continuous system to be design parameters, the optimized system of Figure 4.19 resulted. The performance index lower limit was specified as 2 sample periods or 0.06 sec compared with 0.05 sec for the continuous system. The step function time response was indistinguishable from that shown in Figure 4.18 which indicated that for this sampling frequency, the continuous and the discrete systems exhibited essentially the same performance.

Reduced Sampling Frequency. The effect of sampling frequency was also investigated. Since a different set of design parameters might provide better performance if a different sampling frequency were used, the parameter optimization was repeated using a lower sampling frequency. The sample period was 0.1 sec, a factor of 3 chang . If the time constant of the yaw rate high-pass filter were kept equal to 1.0 second, the response shown in Figure 4.20 as curve A resulted. Since the damping ratio of the lateral oscillation mode for this design was approximately 0.3 and the response shows a fair amount of excitation of that model, the optimization was re-run adding the yaw damper gain and filter time constant as design variables. For this lower sampling frequency, the optimization program preferred a filter time constant of 4.1 sec and gave excellent matching of roll angle and yaw rate as shown in Figure 4.20 as curve B . The system design parameters are listed on the figure. For the lowered sample frequency, the aileron to rudder compensation path pole was placed in the left half plane rather than the right half plane and the magnitude of the zero was greater than 1.0. The other design parameters were not greatly different from those for the higher sample frequency. If the larger yaw rate filter time constant and the lower sampling frequency were not objectionable from other operational specifications, the performance of the lateral coordination system would be considered to be better at the lower sampling frequency than at the higher one.

4.4 First Stage Launch Vehicle Control System

In the design of a pitch attitude control system for a rocket launch



$S_{cs}\left[\phi_1, \delta_a\right] = 0.125$ $\qquad S_{cs}\left[W_x, \delta_a\right] = 0.209$ sec $\qquad S_{cs}\left[W_x, \delta_r\right] = 0.0671$ sec

$$G(z) = 0.06383\left(\frac{0.6606z + 0.7555}{z + 0.4161}\right) \qquad G_{hpf}(z) = \frac{(z - 1.0)}{(z - 0.97613)}$$

$G_{cs}\left[W_z, \delta_r\right] = 0.4336$ sec

Fig. 4-20. Performance of the Digital Lateral Control System Optimized for a Lower Sample Frequency.

vehicle, the system specifications may consist primarily of a minimum forward loop gain together with a desired gain margin. The gain may be specified by the allowable engine deflection caused by wind shears encountered during first stage flight through the atmosphere. Inasmuch as dynamic maneuvering requirements are minimal for such a vehicle, there may exist rather broad tolerances upon acceptable dynamic response characteristics. Hence in selecting a suitable model one can design the low frequency characteristics from consideration of a simplified analysis of the vehicle rigid body mode. Rate information to provide damping for the rigid body mode is assumed to be obtained from a rate gyro and the attitude information from an inertial platform. From such an analysis the required angle and rate feedback loop gains can be obtained for an assumed variation in atmospheric wind shear and practical damping requirements. The resulting low frequency pitch control system can be taken as the model to be used for design of bending mode compensation. Depending upon the dynamic behavior of sensors and actuators, this rigid body design could be further refined if needed by specifying minimum values for the loop gains and using the parameter optimization program to establish other compensation parameters.

When the structural bending modes are added, one may have to add additional compensation in order to achieve the desired gain margin. This example looked at the use of the parameter optimization program in designing a bending compensation filter for this case. A functional block diagram for the control system is shown in Figure 4.21. The corresponding mathematical block diagram is presented in Figure 4.22. The gyro sensors have been assumed to be located at the same fuselage station. The gyros sensed both the rigid body

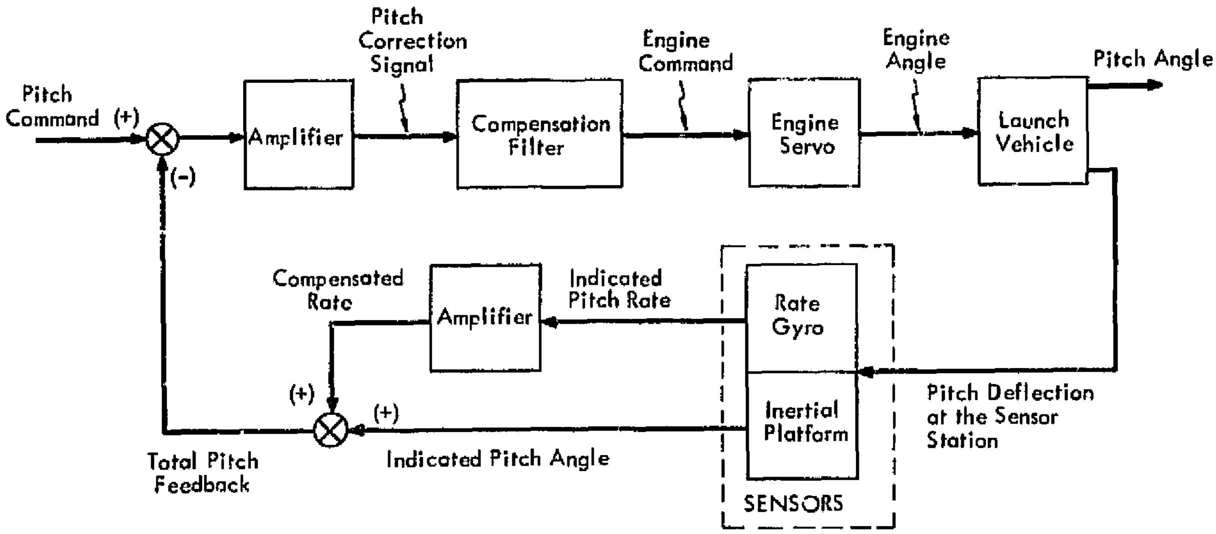
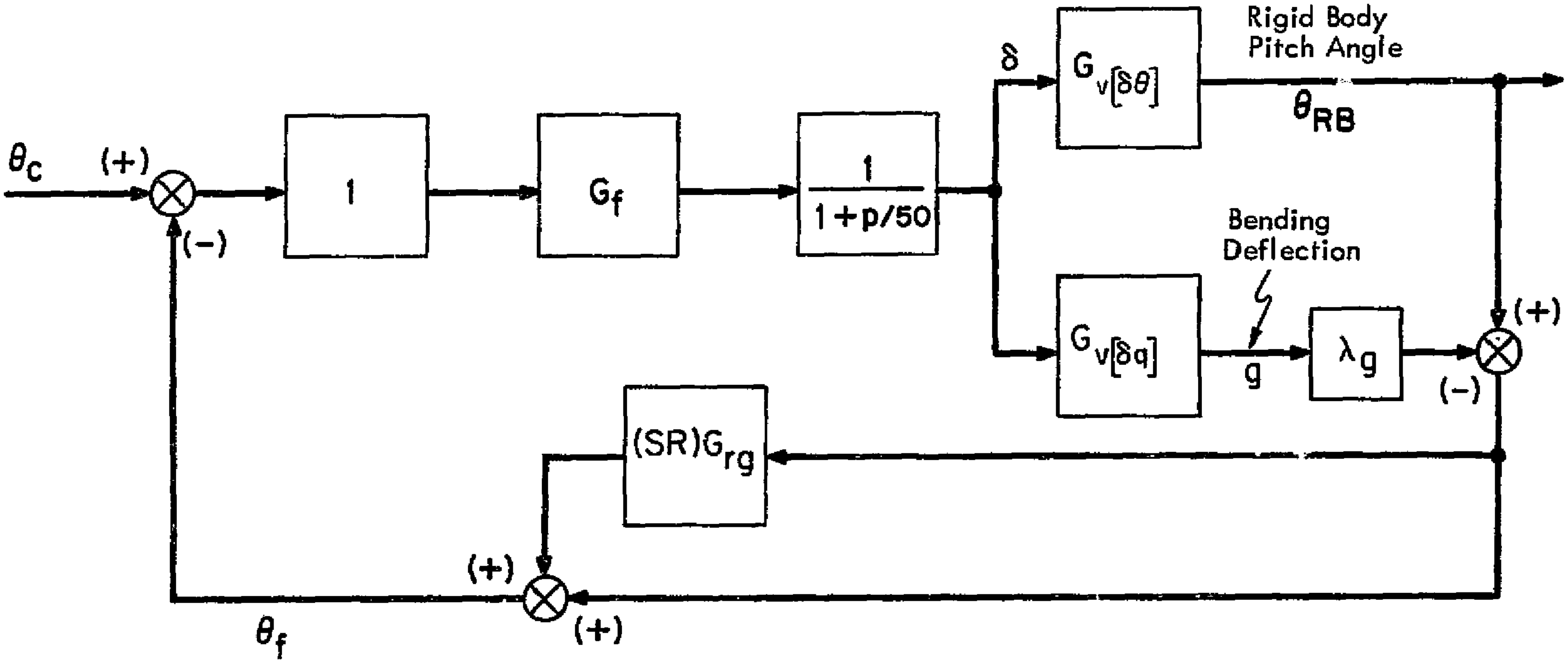


Fig. 4-21. Functional Block Diagram. Launch Vehicle Pitch Attitude Control System.

motion, θ_{RB}, and the component of the structural bending corresponding to the bending mode slope at the sensor location, $\lambda_g q_1$. The ratio of the rate gyro path static sensitivity to that for the attitude gyro was expressed as the sensitivity ratio, (SR), and the required value was taken as 1.0 sec. The system specifications were assumed to lead to a required attitude loop gain of 1.0, expressed as the ratio of the steady-state engine deflection, δ, to a steady-state pitch angle, θ. The rate gyro was assumed to be represented as a second-order component with damping ratio of 0.7 and undamped natural frequency of 120 rad/sec. The engine actuator was a first-order lag with a pole at -50 rad/sec. At this flight condition the uncontrolled vehicle was aerodynamically unstable. The transfer functions relating the rigid body pitch angle and the bending mode displacement, q_1, to the engine deflection are listed on Figure 4.22. The "tail-wags-dog" zeroes are included.



$$G_{rg} = \left(\frac{p}{1 + \frac{2(0.7)}{120} p + \frac{p^2}{(120)^2}} \right) \qquad G_f = \frac{1}{1 + \frac{2\zeta}{W_n} p + \frac{p^2}{W_n^2}}$$

$$G_{v[\delta\theta]} = \frac{-7.621(1+p/0.0780)(1-p/z_1)(1-p/z_1)}{(1-p/0.0210)(1+p/1.57)(1-p/1.496)} \qquad G_{v[\delta q]} = \frac{-0.333(1+p^2/w_{z_1}^2)}{1+5\times10^{-4}p+2.5\times10^{-3}p^2}$$

$$z_1 = 0.0123 + 173j \qquad w_{z_1} = 54$$

Fig. 4-22. Mathematical Block Diagram. Launch Vehicle Pitch Attitude Control System.

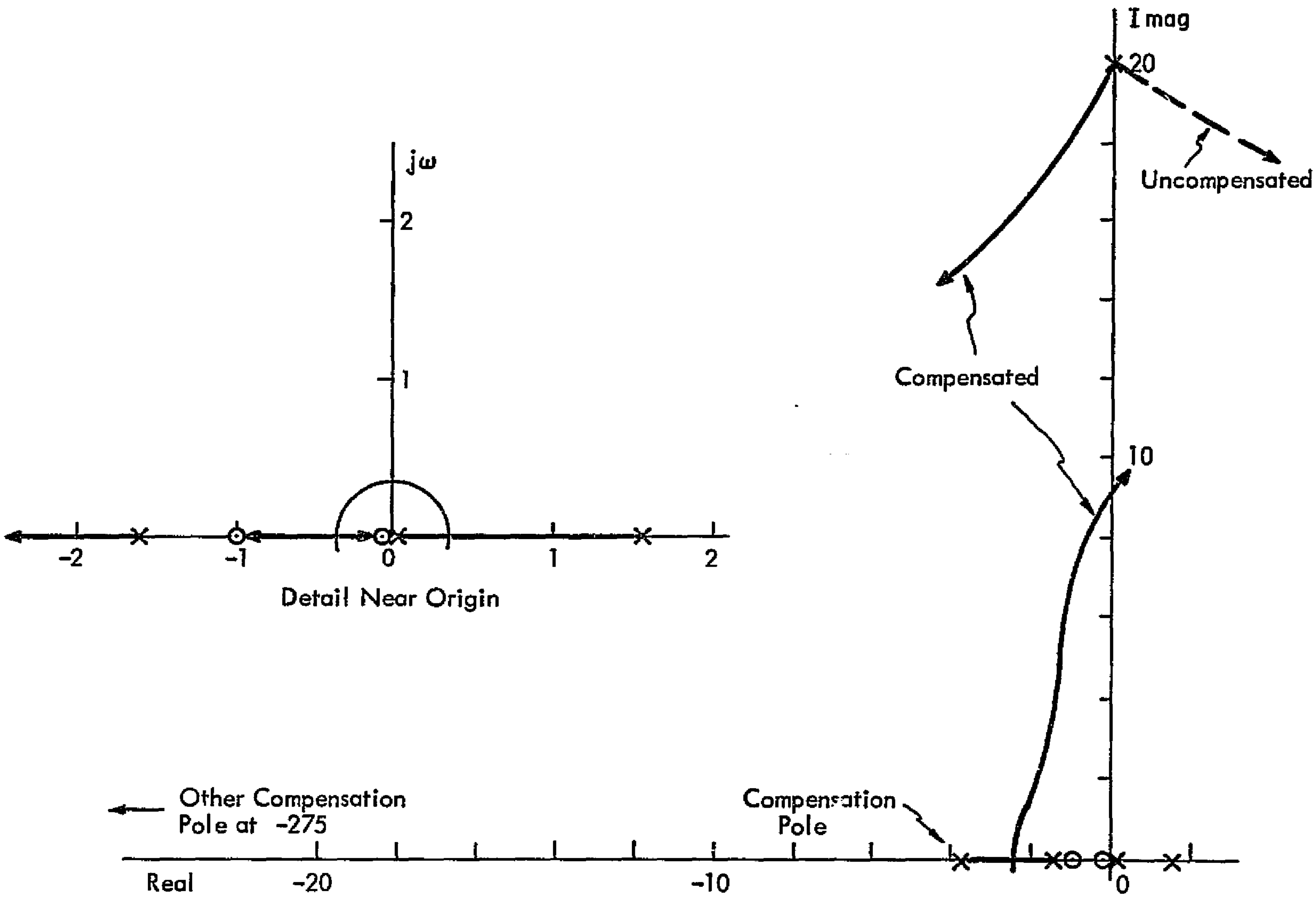


Fig. 4-23. Root Locus, for Final Design, Launch Vehicle Control System.

Figure 4.23 presents root loci for this control system. The dominant rigid body mode was the low frequency oscillatory mode which resulted from the attitude feedback loop. This mode was adequately compensated by the zero introduced by the rate gyro loop. (Note that the loci shown in Figure 4.23 include the effects of the high frequency modes although they were outside the scale range of the figure.) The first bending mode has a frequency of approximately 20 rad/sec and the figure shows by the dashed line that without additional compensation that mode is unstable.

Previous experience had shown that the bending mode could be phase-stabilized using a second-order lag filter placed in the forward path of the pitch loop. The filter's natural frequency and damping ratio were taken as optimization parameters. The order of the overall system was 10. Initial parameter values that produced at least a stable system were specified. The parameter optimization was first performed using the required pitch attitude loop gain of 1.0 using as the model the low frequency rigid body mode that the simplified analysis predicted.

$$G_m = \frac{(1 + p/0.078)}{(1 + p/0.2)(1 + p/0.55)}$$

This optimization selected a bending mode compensation filter with damping ratio of 1.4 and undamped natural frequency of 24.5 rad/sec.

Increasing the Gain Margin. This design had a gain margin of 1.2 which was considered to be too low. The performance index does not include any indication of gain margin. Therefore it was decided to use the program to redesign the filter using a higher pitch attitude loop gain thereby insuring a larger gain margin at the desired lower loop gain. The forward loop gain was accordingly increased to 1.5, and the optimization was repeated. The model was also adjusted to reflect the change in rigid body mode characteristics. The filter design that resulted had a damping ratio of 4.3 with undamped natural frequency of 32.2 rad/sec. With this high a value for the damping ratio, the poles of this filter were real at -3.9 and -275 rad/sec. Therefore the filter was dominantly a first-order lag rather than the second-order lag that had been originally anticipated. The parameter optimization thus led to a simpler compensator design than one would have specified on the basis of previous experience. At a forward loop gain of 1.0, this system design exhibited a gain margin of 2.4, and the root locus is shown in Figure 4.23 as the compensated branches. The bending mode branches remained in the left half plane, and the critical mode arose from the compensation pole coupling with one of the booster poles. At the design open loop gain, this mode had a damping ratio of 0.35. There was very little change in the low frequency gain margin (2.1), since that was determined essentially by the aerodynamic characteristics of the vehicle. For wind gust inputs the rigid body response was satisfactory. Although it was not done, the design iteration could have been continued in the same manner to further refine the gain margin if desired.

5. CONCLUDING REMARKS

The results of the analytical design studies performed in this investigation show that a digital computer parameter optimization program provides an effective technique for the design of linear automatic control systems. It permits one to start with a simple and practical system configuration and from it evolve a system of acceptable performance with a minimum of complexity. Such a simple system is defined to be the "optimum" one. The inverse problem of having to simplify a complex configuration which the state vector feedback approaches introduce is thereby avoided.

Rarely do the operational requirements upon system performance specify more than the dominant dynamic characteristics of the nominal system response to operating inputs. In addition, acceptable operational use often results even for deviations in performance from the nominal within rather broad tolerance limits. In such cases very simple compensation networks or similar devices can usually suffice. However selection of the values of their design parameters may not be obvious if the system is complex, since the rules of thumb that apply to simple low order systems are difficult to extend to high order systems whose transfer functions contain several zeroes. The computer optimization procedure described in this report solves this difficulty by permitting one to use his knowledge of the characteristics of low order systems in the design of a reference model and to use the computer to perform the extensive computations needed to find that parameter set which yields the best "matching" of the dominant characteristics of the response of the complex system to that of the simpler model. The computer program accomplishes this task by searching for that set of parameters which minimizes a mathematical functional called the performance index. Therefore one has to make the further stipulation that when the performance index is minimized the desired matching between system and model responses indeed results. If that is true, the optimization procedure provides an efficient method of design iteration which in effect is stating that it discards all of those parameter combinations which would give poorer performance. Having obtained the desired parameter set one can usually obtain further insight into why that set is accomplishing its purpose through root locus and/or frequency response analyses permitting one to examine also such design features as sensitivity, gain margin, etc.

Because the required response characteristics cannot be specified exactly, as has already been noted, there is no analytical transformation of the specifications that defines the mathematical functional to use. One is forced to use that performance index which experience shows will result in acceptable designs. The Model Performance Index is recommended as the quadratic functional to use for the optimization process. The advantage of this index is that it automatically provides the state variable weighting factors that should be used, and these are the coefficients of the characteristic equation of a reference model. Its use has led to practical system designs with acceptable performance. Simple performance specifications can be used, but these must be in a form that can be translated into a tolerance boundary within which the step function response of the system should lie. The model is selected so that its step response meets the performance specification. The model can be of any order subject to the constraint that the excess of its poles over its zeroes is no greater than that of the closed-loop system when the system exhibits right half plane zeroes. In general one needs to model only the dominant modes of response of the system, and the model is typically of much lower order than the system. Various system response characteristics can be incorporated into the model design, if desired. In particular, the maximum control effort limitations can be incorporated through specification of the model's acceleration.

Not all of the properties of the Model Performance Index have been established to date. The importance to be attached to the quantitative value of the index is one such area. If that could be more clearly defined, the value of the performance index could be used to establish parameter tolerance ranges for acceptable performance capabilities and establish specification parameters similar to the widely used gain margin and phase margin specifications. A clearer picture of parameter sensitivity using the performance index would then be possible.

It was found to be necessary to modify the Model Performance Index as given in references 1 and 3 for these design situations in which there is a large discrepancy at the initial time between the lower order states up to the order of the model. This is primarily important for systems having zeroes in the transfer function. If the model is of different order than the system, it is impossible to match the state trajectories at $t = 0$. The optimization process in attempting to achieve an overall trajectory match can improve the initial portion of the trajectory by permitting poorer matching at subsequent times. This is equivalent to improving the high frequency response at the expense of the lower frequency characteristics, which usually is an unsatisfactory trade-off. The modification used was to place a lower limit upon the performance index integral thereby ignoring the initial portion of the trajectory. The selection of the lower limit at present is still a trial and error procedure.

The parameter optimization technique permits one to include all of the component characteristics without complicating the feedback structure. These include actuator and sensor dynamics and certain filter characteristics chosen for engineering design reasons other than dynamic performance. The order of the system affects the process primarily in increasing the cost of the computation. The controlled member, or plant, may have zeroes in either the left or right half plane as is typical with flight vehicles. The design variables can be gains, filter time constants, undamped natural frequencies or damping ratios of components. Practical constraints can be placed upon the allowable values of the design parameters in a straightforward manner. Digital information processing can be considered, and the design constants in the digital section can be either gains or the coefficients of the difference equation (or Z-transform) for compensation filters. Within the dimension specification of storage arrays in the program, any arrangement of feedback paths and feed-forward paths to a multiple-input controlled member can be considered. Any signal path can be taken as the output.

In essence the analysis effort of the engineering design iteration steps has been computerized in a fashion that permits the engineer to retain the constraints imposed by practical implementation considerations. Aside from reducing the manual effort associated with the design process of complicated configurations, those solutions to the problem can emerge which might have been overlooked or not anticipated or not explored for lack of time. This is particularly true of systems whose transfer functions contain zeroes, since it becomes more difficult to correlate the various pole-zero arrangements in terms of the resulting transient time response. This can result in a simpler system than one would have expected, and that is an important advantage of a computer optimization technique.

The question of system design for stochastic inputs has not been treated in this report. Conceptually there would be no change other than a change in system specification and a corresponding change in performance index. To the extent that one could consider the specification of acceptable performance in the presence of stochastic inputs as a specification upon the desired frequency response of the system, one can incorporate such specifications in the design of the reference model. The same design procedure would then result. Undoubtedly one would specify the model design over a larger frequency range than would otherwise be necessary. This is an area for future investigation however.

APPENDIX A

DATA FOR EXAMPLE DESIGNS

1. F-8 Airplane Longitudinal Data

The longitudinal aerodynamic stability data for the F-8 airplane are tabulated in Table A-1 for the nine flight conditions investigated. Mass and geometric data are as follows:

Mass	9994 kg	(648.8 slug)
Pitch Moment of Inertia	118640 kg - m^2	(87490 slug-ft^2)
Chord	3.59 m	(11.8 feet)
Wing Area	34.87 m^2	(374.9 $(ft)^2$)
Tail Length	4.8 m	(15.7 feet)
Sensor Location	4.57 m	(15 feet)

Stability axes were used for writing the equations of motion. The state equations are

$$\dot{\underline{x}} = \begin{bmatrix} \dot{w} \\ \ddot{\theta} \end{bmatrix} = A\,\underline{x} + B\,\delta_e$$

The A and B matrices for the nine flight conditions are:

(w in in m/sec; θ and δ are in radian)

F.C.1

$$A = \begin{bmatrix} -0.560 & 277.4 \\ -0.0256 & -0.146 \end{bmatrix} \qquad B = \begin{bmatrix} -18.29 \\ -9.26 \end{bmatrix}$$

F.C.2

$$A = \begin{bmatrix} -2.16 & 351.7 \\ -0.2037 & -0.0362 \end{bmatrix} \qquad B = \begin{bmatrix} -90.19 \\ -42.0 \end{bmatrix}$$

F.C.3

$$A = \begin{bmatrix} -1.05 & 326.4 \\ -0.102 & -0.0150 \end{bmatrix} \qquad B = \begin{bmatrix} -40.60 \\ -17.8 \end{bmatrix}$$

F.C.4

$$A = \begin{bmatrix} -0.653 & 324.6 \\ -0.0623 & -0.0100 \end{bmatrix} \qquad B = \begin{bmatrix} -24.99 \\ -11.6 \end{bmatrix}$$

F.C.5

$$A = \begin{bmatrix} -0.562 & 154.8 \\ -0.0325 & -0.305 \end{bmatrix} \qquad B = \begin{bmatrix} -12.59 \\ -6.22 \end{bmatrix}$$

F.C.6

$$A = \begin{bmatrix} -1.360 & 187.9 \\ -0.05315 & -0.7758 \end{bmatrix} \qquad B = \begin{bmatrix} -36.88 \\ -17.8 \end{bmatrix}$$

F.C.7

$$A = \begin{bmatrix} -0.820 & 189.6 \\ -0.0312 & -0.4142 \end{bmatrix} \qquad B = \begin{bmatrix} -22.34 \\ -10.85 \end{bmatrix}$$

F.C.8

$$A = \begin{bmatrix} -1.32 & 268.6 \\ -0.04856 & -0.6956 \end{bmatrix} \qquad B = \begin{bmatrix} -49.53 \\ -23.8 \end{bmatrix}$$

F.C.9

$$A = \begin{bmatrix} -0.720 & 252.1 \\ -0.02986 & -0.378 \end{bmatrix} \qquad B = \begin{bmatrix} -22.56 \\ -12.3 \end{bmatrix}$$

2. F-8 Lateral Equations

The data for the lateral-directional control system example of Section 4.3 4.3 are presented here. The state equations are written for a body axis coordinate frame whose X-axis is inclined to the trimmed velocity vector by the angle of attack, α_0. The roll angle, ϕ, is the Euler angle rotation about the airplane X axis, and its rate of change therefore contains components of the angular velocity of the aircraft along both the X and Z axes as in Fig. A-1. Again

$$\dot{\underline{x}} = A\,\underline{x} + B\,\underline{u}$$

where

$$\underline{x} = \begin{bmatrix} \phi \\ W_x \\ W_z \\ \beta \end{bmatrix} \qquad \underline{u} = \begin{bmatrix} \delta_a \\ \delta_r \end{bmatrix}$$

Two flight conditions were considered. (All angle units are radians.) At Mach = 1.6, 12192 m. altitude:

$$\alpha_0 = 8.1°$$

$$A = \begin{bmatrix} 0 & 1.0 & 0.1423 & 0 \\ 0 & -1.9066 & 0.33441 & -67.403 \\ 0 & -0.18778 & -0.41197 & 1.777 \\ 0.0208 & 0.1417 & -0.990 & -0.2384 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 \\ 13.387 & 5.2363 \\ -0.047038 & -2.3726 \\ 0.0010 & 0.010 \end{bmatrix}$$

At Mach = 0.56, 6706 m. altitude:

$\alpha_0 = 7.75°$

$$A = \begin{bmatrix} 0 & 1.0 & 0.1361 & 0 \\ 0 & -2.625 & 1.91 & -29.80 \\ 0 & -0.0759 & -0.426 & 2.650 \\ 0.0555 & 0.1359 & -0.9974 & -0.2173 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 \\ 27.0 & 6.13 \\ 1.420 & -3.55 \\ 0.002315 & 0.04222 \end{bmatrix}$$

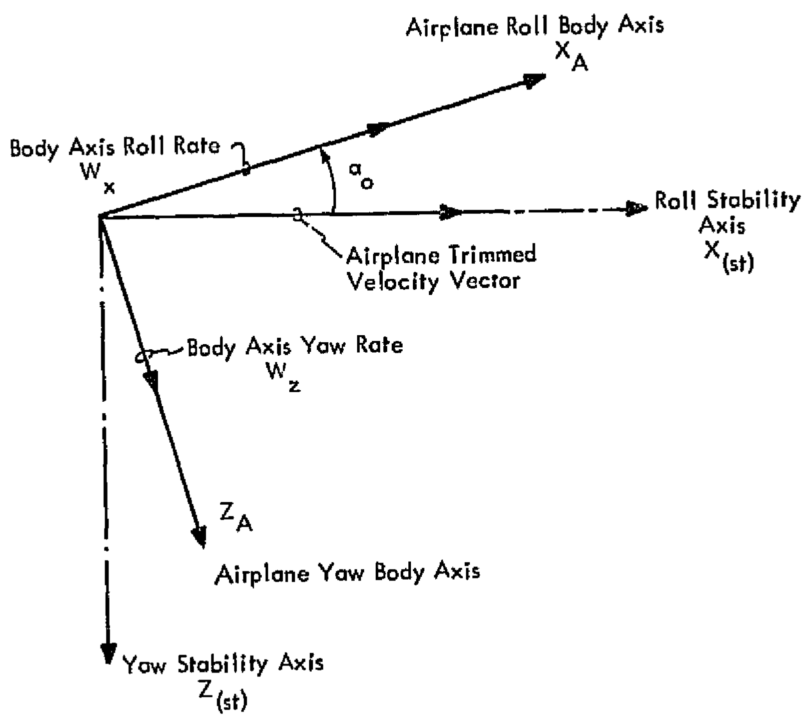


Fig. A-1. Coordinate Axis for Aircraft Equation of Motion.

TABLE A-1 Aerodynamic Data for the F-8 Airplane, Longitudinal Axis

F8 NON DIMENSIONAL LONGITUDINAL STABILITY DERIVATIVES

F.C.	1	2	3	4	5	6	7	8	9
C_D	0.031	0.0415	0.0430	0.0480	0.0410	0.0195	0.0230	0.0175	0.0210
C_{D_M}	0.115	0.116	0.116	0.116	-0.0167	-0.0175	-0.018	0.115	0.115
$\partial T/\partial u$	0.649	-0.1	1.47	0.649	0.50	-0.67	0.3	3.68	1.47
C_L	0.284	0.0775	0.143	0.224	0.452	0.11	0.240	0.11	0.23
C_{L_M}	0.120	0.117	0.116	0.116	0.144	0.109	0.109	0.08	0.123
C_{D_α}	0.328	0.106	0.217	0.335	0.29	0.29	0.29	0.072	0.210
C_{L_α}	4.864	4.866	4.866	4.866	3.80	3.80	3.81	4.30	4.30
C_{m_M}	-0.145	0.068	-0.002	-0.144	-0.077	-0.07	-0.002	0.01	-0.145
C_{m_α}	-0.745	-1.55	-1.55	-1.55	-0.715	-0.492	-0.477	-0.525	-0.59
$C_{L_{\delta_e}}$	0.573	0.58	0.58	0.58	0.55	0.55	0.550	0.604	0.532
$C_{m_{\delta_e}}$	-0.967	-0.882	-0.84	-0.99	-0.895	-0.88	-0.87	-0.96	-0.964
C_m	1.16	2.7	2.7	2.7	-0.283	-0.233	-0.200	-0.070	-0.070
C_{m_q}	-3.45	-2.85	-2.85	-2.85	-3.50	-3.47	-3.45	-4.1	-4.1
Altitude (km.)	13.7	5.18	10.7	13.7	7.62	1.22	6.10	6.10	10.7

TABLE A-1 (continued)

F.C.	1	2	3	4	5	6	7	8	9
Mach No.	0.94	1.1	1.1	1.1	0.5	0.56	0.6	0.85	0.85
α_0(deg)	4.7	1.5	2.4	3.7	7.7	2.7	4.6	2.3	3.9
δ_{e_0}(deg)	-4.5	-1.26	-2.9	-5.0	-4.68	-2.22	-2.6	-1.93	-2.6
V(m/s)	277	352	326	325	155	188	190	267	252
q(kN/m^2)	4.84	44.6	20.2	12.52	6.56	19.20	11.7	23.6	12.1

REFERENCES

1. Rediess, Herman A.; A New Model Performance Index for the Engineering Design of Control Systems. Ph.D. Thesis, Massachusetts Institute of Technology, 1969.

2. Rediess, Herman A., and Whitaker, H. Philip; A New Model Performance Index for Engineering Design of Flight Control Systems. Journal of Aircraft, Vol. 7, No. 6, Nov.-Dec., 1970, pp. 542-549.

3. Palsson, Thorgeir; Parameter Uncertainties in Control System Design. Sc.D. Thesis, Massachusetts Institute of Technology, 1971.

4. Slater, G.L.; A Unified Approach to Digital Flight Control Algorithm. American Institute of Aeronautics and Astronautices, Paper No. 74-884, August 1974.

5. Hung, J.W., et al.; Investigation of Numerical Techniques as Applied to Digital Flight Control Systems. U.S. Air Force Flight Dynamics Laboratory Report AFFDL-TR-66-68, 1967.

6. Whitaker, H. Philip; The System Description Program. Measurement Systems Laboratory, Massachusetts Institute of Technology, Report RE-79 (Revision 1), 1972.

7. Baram, Yoram; An Analysis of the Accuracy of a Parameter Optimization Technique. Measurement Systems Laboratory, Massachusetts Institute of Technology, Report TE-54, 1974.

8. Malcom, L.G., and Tobie, N.N.; New Short Period Handling Quality Criteria for Fighter Aircraft. The Boeing Company, Document No. D6-17841T/N, October 1965.

9. Whitaker, H.P., Baram, Y., and Cheng, Y.; Annual Report for NASA Grant NGL-22-009-548. Measurement Systems Laboratory, Massachusetts Institute of Technology, Report RE-89, Nov. 1973.

10. Stapleford, R.L., Klein, R.H., and Hoh, R.H.; Handling Qualities Criteria for the Space Shuttle Orbiter During the Terminal Phase of Flight. NASA Contractor Report, NASA CR-2017, April 1972.