NASA-CR-192179

DYNAMICS OF AEROSPACE VEHICLES

Final Technical Report

For

NASA Grant NAG 1-1086

On Research Performed For the

NASA Langley Research Center
Hampton, VA

Technical Monitor:
Mr. C. S. Buttrill

Principle Investigator

Dr. David K. Schmidt
Aerospace Research Center
College of Engineering and Applied Sciences
Arizona State University
Tempe, AZ 85287-8006

September 1, 1991

Final Technical Report

DYNAMICS OF AEROSPACE VEHICLES
(NAG 1 - 1086)

Project Period January 1, 1990 to January 3, 1991.

This constitutes the final technical report for the subject grant. The grant was administered through Arizona State University, and the principle investigator was Dr. David K. Schmidt. During this period, the principle investigator and graduate researchers supported by the project were all affiliated with Arizona State University's Aerospace Research Center in the College of Engineering and Applied Sciences, located in Tempe, Arizona 85287. This grant is a direct follow-on to grant number NAG 1-758, with the same principle investigator. Hence, it may be helpful to view these two grant activities as one research program.

The research focus of this grant was to address the modeling, including model reduction, of flexible aerospace vehicles, with special emphasis on models used in dynamic analysis and/or guidance and control system design. In the modeling, it is critical that the key aspects of the system being modeled that are of import in the application for the model (e.g., feedback system design) be captured in the model. In this work, therefore, aspects of the vehicle dynamics critical to control design were important.

In this regard, fundamental contributions were made in the areas of stability robustness analysis techniques, model reduction techniques, and literal approximations for key dynamic characteristics of flexible vehicles. All these areas are related. In the development of a model, approximations are always involved, so control systems designed using these models must be robust against uncertainties in these models. On the other hand, it is imperative the the systems analyst and designed be aware of the sources of possible uncertainty, and modeling itself is critical in developing this awareness.

The graduate researchers supported by the current grant are the following:

Mr. Brett Newman, Research Associate in ASU's Aerospace Research Center. Mr. Newman is also pursuing a Doctorate in Aerospace Engineering from Purdue University. This individual had key responsibilities in all phases of this research.

Mr. Juan Salas, Research Assistant in ASU's Aerospace Research Center. Mr. Salas is pursuing his M.S. graduate degree in Aerospace Engineering at ASU. This individual had responsibilities for numerical analysis in support of this grant activity.

Conference papers presented, reporting technical results from this (current) grant are as follows:

1) Newman, B. and Schmidt, D.K., "A Sequential Approach to Multivariable Stability Robustness Analysis," AIAA paper 91-2771, presented at, and appearing in the Proceeding of the 1991 Guidance, Navigation, and Control Conference, New Orleans, LA, August, 1991.

2) Newman, B. and Schmidt, D.K., "New Plant and Controller Order Reduction Results With Weighted Balancing," AIAA paper 91-2805, presented at, and

appearing in the Proceedings of the 1991 Guidance, Navigation, and Control Conference, New Orleans, LA, August, 1991.

The following papers have been, or are being submitted for publication in archival journals:

1) Newman, B. and Schmidt, D.K., "On The Control of Elastic Vehicles - Model Simplification and Stability Robustness," to be submitted to the IEEE Control Systems Magazine.

2) Schmidt, D.K. and Newman, B., "Multivariable Flight Control Synthesis and Literal Robustness Analysis for an Aeroelastic Vehicle," To be submitted to the Journal of Guidance, Control, and Dynamics.

3 and 4) Conference papers 1 and 2 above are also being submitted to the Journal of Guidance, Control, and Dynamics.

Copies of all the above papers are included in Appendices A through C, respectively, of this report.

Appendix A

PREV. ANN
91A 49724

AIAA-91-2771-CP

A Sequential Approach To Multivariable Stability Robustness Analysis

Brett Newman* and David K. Schmidt**
Aerospace Research Center
Arizona State University
Tempe, Arizona

Abstract

In sequential loop closure, the importance of evaluating the stability and stability robustness at the intermediate loop closures is well known, yet how the stability and stability robustness evaluated at the intermediate steps contribute to the stability and stability robustness of the overall feedback system must be developed. An analysis of the complete feedback system reveals the multivariable Nyquist contributions from the intermediate loop closures. It is also shown that the results greatly simplify if frequency separation exists between the intermediate loops. The analysis is presented with a two-step loop closure procedure using "inner" and "outer" loops which can be generalized to multi-step situations. The control of the longitudinal dynamics of an aircraft is addressed to further clarify and demonstrate the results.

Introduction

Consider the generic multivariable feedback loop in Figure 1 with responses y, control inputs u, response commands y_c, plant transfer function matrix G(s), and compensator transfer function matrix K(s). Usually, the compensator must stabilize all unstable modes present in the plant. Further, the compensator must ensure this stability in the presence of plant modeling errors.

Frequency domain criteria for stability and stability robustness such as multivariable Nyquist stability theory are well known and extensively documented.[1-4] These tools are directly applicable to a given multivariable compensator in the format of Figure 1. For example, stability and stability robustness are indicated by the Nyquist diagram corresponding to det [I + K(s)G(s)]. However, if the compensator is developed with a sequential loop closure strategy,[5,6] there exists a void concerning the relationship between the stability and stability robustness indicated after each loop closure and the stability and stability robustness of the complete feedback system.

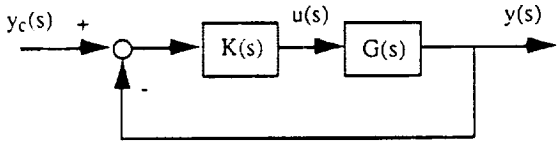
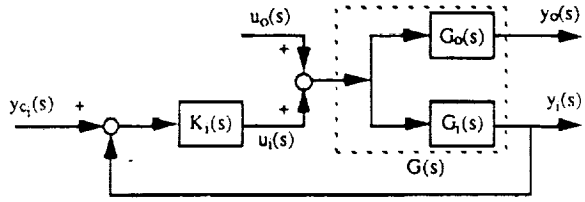


Figure 1. Generic Feedback Loop

For example, consider the two-step loop closures shown in Figures 2 and 3, where the "inner loops" consists of outputs y_i and the "outer loops" consists of outputs y_o. It is important to observe here that the inner and outer loops can be multivariable. Further, observe that the block diagram structure in Figure 2 can be manipulated into the more classical looking inner and outer loop structure depicted in Figure 4 if $K_i(s)$ and $G_i(s)$ are nonsingular. However, for ease of exposition the structure of Figure 2 will be considered.

With only the inner loops closed, system stability and robustness are indicated by the Nyquist diagram corresponding to det $[I + K_i(s)G_i(s)]$, while after the outer loops are closed, system stability and robustness are indicated by the Nyquist diagram corresponding to det $[I + K_o(s)G'_o(s)]$ where $G'_o(s)$ is the effective plant for the outer loop closure. Also, with only the inner loops closed, if unstable system modes remain (which are to be stabilized by the outer loops), then the standard multivariable stability robustness theory[3,4] applied at this intermediate step is not valid.

Inner Loop Closure :



Outer Loop Closure :

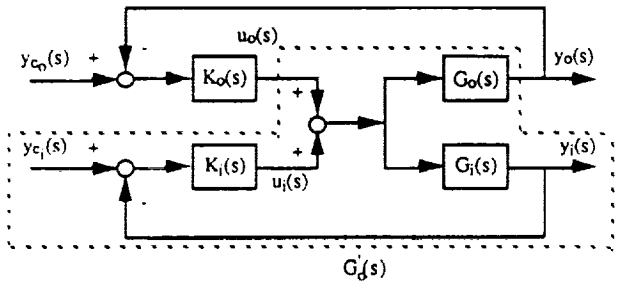
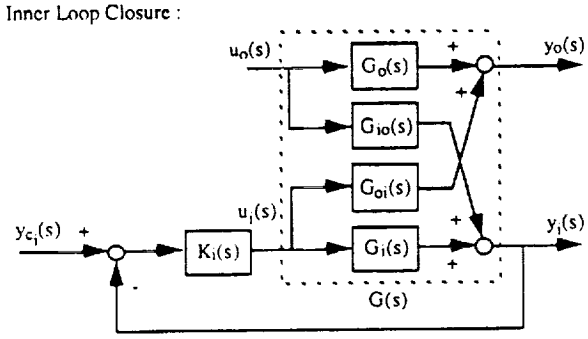


Figure 2. Two-Step Loop Closure

Inner Loop Closure :



Outer Loop Closure :

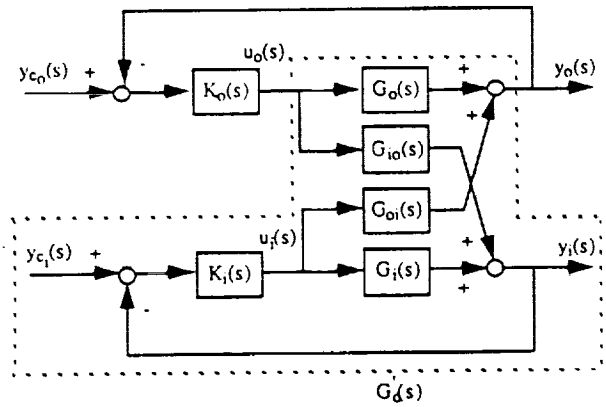


Figure 3. Two-Step Loop Closure

* Research Associate; Doctoral Student, School Of Aeronautics And Astronautics, Purdue University; Student Member AIAA

** Acting Director; Professor Of Mechanical And Aerospace Engineering; Associate Fellow AIAA

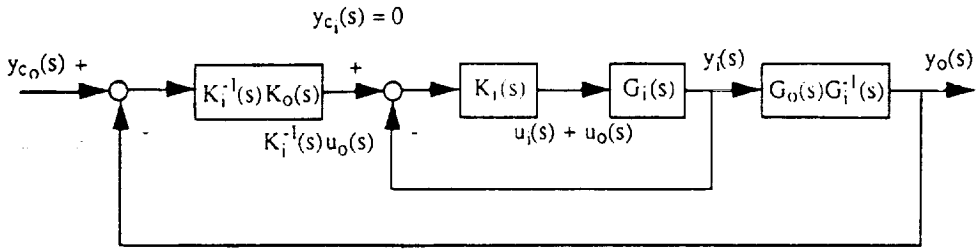


Figure 4. Classical Inner And Outer Loop Structure

The goal of this paper is to relate the stability and stability robustness evaluated at each stage of the (multivariable) sequential loop closure to that of the final feedback system, to offer insight, and to suggest a multivariable extension to this frequently used classical synthesis technique.

Nyquist Stability And Stability Robustness

Stability of the closed-loop system in Figure 1 is completely determined by the roots of the closed-loop characteristic polynomial $\phi_{cl}(s)$. The closed-loop characteristic polynomial is related to the open-loop characteristic polynomial, $\phi_{ol}(s)$, by the well known relationship[7]

$$\frac{\phi_{cl}(s)}{\phi_{ol}(s)} = \det[I + K(s)G(s)] \tag{1}$$

Application of the principle of the argument[8] to Eq. (1) yields

$$N(0,\det[I + K(s)G(s)],C_{RHP}) = Z - P \tag{2}$$

where the notation $N(0,\det [I + K(s)G(s)],C_{RHP})$ denotes the number of encirclements of the origin made by the Nyquist diagram (i.e., the mapping of $\det [I + K(s)G(s)]$ as s traverses the contour C_{RHP}, which encloses the entire right-half of the complex plane). Further, Z is the number of closed-loop poles (roots of $\phi_{cl}(s)$) inside C_{RHP}, and P is the number of open-loop poles (roots of $\phi_{ol}(s)$) inside C_{RHP}. For closed-loop asymptotic stability, no closed-loop poles may lie in the right-half plane, or

$$Z = 0 \quad \Rightarrow \quad N = -P \tag{3}$$

In other words, the Nyquist diagram must have the correct number of encirclements of the origin, namely -P.

The feedback loop in Figure 1 must also maintain stability in the presence of plant modeling errors. One common way to represent this error is with additive error $\Delta G(s)$ defined by

$$\Delta G(s) = G_T(s) - G(s) \tag{4}$$

where $G_T(s)$ denotes the "true" linear plant.

For the true feedback system, Eq. (1) becomes

$$\frac{\phi_{cl_T}(s)}{\phi_{ol_T}(s)} = \det[I + K(s)G_T(s)] \tag{5}$$

where $\phi_{cl_T}(s)$ and $\phi_{ol_T}(s)$ denote the true system's closed-loop and open-loop characteristic polynomials, respectively. Application of the principle of the argument[8] to Eq. (5) yields

$$N_T(0,\det[I + K(s)G_T(s)],C_{RHP}) = Z_T - P_T \tag{6}$$

where $N_T(0,\det [I + K(s)G_T(s)],C_{RHP})$ denotes the number of encirclements of the origin by the true system's Nyquist diagram, Z_T is the number of true closed-loop poles in the right-half plane, and P_T is the number of true open-loop poles in the right-half plane. For closed-loop asymptotic stability of the true system, none of it's closed-loop poles may lie in the right-half plane, or

$$Z_T = 0 \quad \Rightarrow \quad N_T = -P_T \tag{7}$$

It can be shown[4] that if

1. The nominal closed-loop system is asymptotically stable or N = -P (see Eq. (3)), and
2. The required number of encirclements of the origin is the same for both nominal and true closed-loop systems or P = P_T (see Eqs. (3) and (7)),

then a necessary and sufficient condition guaranteeing closed-loop asymptotic stability of the true system is

$$\det[I + K(s)G(s,\varepsilon)]|_{s \in C_{RHP}, 0 \le \varepsilon \le 1} \neq 0 \tag{8}$$

where $G(s,\varepsilon)$ is given as

$$G(s,\varepsilon) = G(s) + \varepsilon\Delta G(s) \tag{9}$$

Note that $\varepsilon = 0$ and $\varepsilon = 1$ corresponds to the nominal and true plants, respectively. The geometric concept associated with Eq. (8) is that under assumptions 1. and 2., if as the nominal Nyquist diagram is continuously warped to the shape of the true Nyquist diagram, the number of encirclements of the origin remains unchanged, closed-loop asymptotic stability of the true system is assured. In other words, to maintain stability in the presence of modeling errors the mapping $\det [I + K(s)G(s,\varepsilon)]|_{s \in C_{RHP}, 0 \le \varepsilon \le 1}$ must not pass through the origin. Two sufficient conditions, developed from Eq. (8), guaranteeing true closed-loop asymptotic stability are[3,4]

$$\underline{\sigma}[I + K(j\omega)G(j\omega)] > \bar{\sigma}[K(j\omega)\Delta G(j\omega)] \quad \text{for } 0 \le \omega \le \infty \tag{10}$$

and

$$\underline{\sigma}[I + (K(j\omega)G(j\omega))^{-1}] > \bar{\sigma}[E(j\omega)] \quad \text{for } 0 \le \omega \le \infty \tag{11}$$

where $s = j\omega$ and $E(j\omega)$ is the input multiplicative error

$$E(j\omega) = (K(j\omega)G(j\omega))^{-1}(K(j\omega)\Delta G(j\omega)) \tag{12}$$
$$= G^{-1}(j\omega)\Delta G(j\omega) \text{ if } K(j\omega) \text{ and } G(j\omega) \text{ are nonsingular}$$

Sequential Loop Closure

Sequential loop closure is defined here as the use of any appropriate synthesis technique to design loops in stages to yield the final multivariable control law. For example, classical control techniques can be used to close scalar loops one at a time, or of more interest here, modern multivariable control techniques can be used in stages. However, care must be taken because the selection and closure of a specific loop can both adversely affect the stability and performance already designed into previously closed loops, as well as influence the stability and performance in subsequent loops yet to be closed. Thus, the key to success is the selection and order of the loop closure and this is typically based upon a fundamental understanding of the plant dynamics. Specific examples of this approach can be found in Refs. 5 and 9. One situation where sequential loop closure is particularly effective is where frequency separation exists between each sets of loops. In this particular, but common situation, most modern multivariable synthesis methods would lead to undesirable results if used to close all loops simultaneously. This is due to the fact that the loop transfers are forced to be closely spaced at crossover, which yields strong coupling and destroys any frequency separation naturally present.

The analysis to follow is developed for the two-step loop closure depicted in Figures 2 and 3; however, the approach can be generalized to multi-step loop closure settings. The direct application of Nyquist stability and stability robustness theory to the complete feedback loop in Figures 2 and 3 offers very little information about the stability and stability robustness at each loop closure step which is of paramount importance during the synthesis. To obtain this information, Nyquist theory may be applied at each step in the loop closure process.

But first, it will be shown that the block diagram structures in Figures 2 and 3 are special cases of the structure in Figure 1. Consider the following partition of the system in Figure 1, or

$$y(s) = \begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} \qquad y_c(s) = \begin{bmatrix} y_{c_1}(s) \\ y_{c_2}(s) \end{bmatrix} \quad u(s) = \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} \tag{13}$$

$$K(s) = \begin{bmatrix} K_{11}(s) & K_{12}(s) \\ K_{21}(s) & K_{22}(s) \end{bmatrix} \qquad G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix}$$

$$\Delta G(s) = \begin{bmatrix} \Delta G_{11}(s) & \Delta G_{12}(s) \\ \Delta G_{21}(s) & \Delta G_{22}(s) \end{bmatrix}$$

The block diagram structure in Figure 2 is obtained by selecting

$$K_{21}(s) = 0 \qquad K_{22}(s) = 0 \tag{14}$$
$$G_{12}(s) = \Delta G_{12}(s) = 0 \qquad G_{22}(s) = \Delta G_{22}(s) = 0$$

leading to

$$K(s) = [K_i(s)\, K_o(s)] \qquad G(s) = \begin{bmatrix} G_i(s) \\ G_o(s) \end{bmatrix} \tag{15}$$
$$\Delta G(s) = \begin{bmatrix} \Delta G_i(s) \\ \Delta G_o(s) \end{bmatrix}$$

The block diagram structure in Figure 3 is obtained by selecting

$$K_{12}(s) = 0 \qquad K_{21}(s) = 0 \tag{16}$$

leading to

$$K(s) = \begin{bmatrix} K_i(s) & 0 \\ 0 & K_o(s) \end{bmatrix} \qquad G(s) = \begin{bmatrix} G_i(s) & G_{io}(s) \\ G_{oi}(s) & G_o(s) \end{bmatrix} \tag{17}$$

$$\Delta G(s) = \begin{bmatrix} \Delta G_i(s) & \Delta G_{io}(s) \\ \Delta G_{oi}(s) & \Delta G_o(s) \end{bmatrix}$$

Stability At Each Loop Closure

Let P_K and P_G denote the number of poles of K(s) and G(s), respectively, in the right-half plane, or

$$P = P_K + P_G \tag{18}$$

P_K can be separated into the number of compensator poles in the right-half plane in the inner loop compensation, P_{K_i}, and in the outer loop compensation, P_{K_o}, or

$$P_K = P_{K_i} + P_{K_o} \tag{19}$$

Further, P_G can be separated into the number of plant poles in the right-half plane to be stabilized with the inner loop, P_{G_i}, and with the outer loop, $P_{G_o'}$, or

$$P_G = P_{G_i} + P_{G_o'} \tag{20}$$

Applying Nyquist theory to the inner loop closure yields

$$N_i(0,\det[I + K(s)G_i(s)],C_{RHP}) = Z_i - P_i \tag{21}$$

where $N_i(0,\det [I + K(s)G_i(s)],C_{RHP})$ denotes the number of encirclements of the origin by the inner loop Nyquist diagram, Z_i is the number of closed-loop poles of the inner loop system in the right-half plane, and P_i is the number of open-loop poles of the inner loop system in the right-half plane. Although closed-loop stability of the complete feedback system is ultimately desired, requiring $Z_i = 0$ at this loop closure step is not necessary because the unstable poles represented by Z_i are to be stabilized by the outer loop. Using the notation in Eqs. (18) thru (20), Z_i and P_i are given as

$$Z_i = P_{G_o'} \qquad , \quad P_i = P_{K_i} + P_G \tag{22}$$

and the encirclement requirement in Eq. (21) becomes

$$N_i = - P_{K_i} - P_{G_i} \tag{23}$$

Next, applying Nyquist theory to the outer loop closure yields

$$N_o(0,\det[I + K(s)G_o'(s)],C_{RHP}) = Z_o - P_o \tag{24}$$

where $N_o(0,\det [I + K(s)G_o'(s)],C_{RHP})$ denotes the number of encirclements of the origin made by the outer loop Nyquist diagram, Z_o is the number of closed-loop poles of the outer loop system in the right-half plane, and P_o is the number of open-loop poles of the outer loop system in the right-half plane. Since this is the last loop closure, the requirement

$$Z_o = 0 \quad \Rightarrow \quad N_o = - P_o \tag{25}$$

is necessary for asymptotic stability of the complete system. Using the notation in Eqs. (18) thru (20), P_o is given as

$$P_o = P_{K_o} + P_{G_o'} \tag{26}$$

and the encirclement requirement in Eq. (25) becomes

$$N_o = - P_{K_o} - P_{G_o'} \tag{27}$$

In summary, for closed-loop asymptotic stability of the complete feedback system, the inner and outer loop Nyquist diagrams must have the correct number of encirclements of the origin, namely $-P_{K_i}-P_{G_i}$ and $-P_{K_o}-P_{G_o'}$, respectively.

To understand how the inner and outer loop encirclement requirements relate to the encirclement requirement for the complete feedback loop in Figure 1, consider $\det[I + K(s)G(s)]$ and the partitioning in Eq. (13), or

$$\det[I+KG] = \det\begin{bmatrix} I+K_{11}G_{11}+K_{12}G_{21} & K_{11}G_{12}+K_{12}G_{22} \\ K_{21}G_{11}+K_{22}G_{21} & I+K_{21}G_{12}+K_{22}G_{22} \end{bmatrix} \tag{28}$$

Using the identity for the determinate of a partitioned matrix[10] yields

$$\det[I+KG] = \det[I+K_{11}G_{11}+K_{12}G_{21}]\det[I+K_{21}G_{12}+K_{22}G_{22}-(K_{21}G_{11}+K_{22}G_{21})(I+K_{11}G_{11}+K_{12}G_{21})^{-1}(K_{11}G_{12}+K_{12}G_{22})] \tag{29}$$

For the block diagram structure in Figure 2, Eq. (29) reduces to

$$\det[I + KG] = \det[I + K_iG_i]\det[I + K_oG_o(I + K_iG_i)^{-1}] \tag{30}$$

From Figure 2, the effective transfer function between y_o and u_o with the inner loop closed is

$$G_o' = G_o(I + K_iG_i)^{-1} \tag{31}$$

Thus, the result in Eq. (30) becomes

$$\det[I + KG] = \det[I + K_iG_i]\det[I + K_oG_o'] \tag{32}$$

But from the block diagram in Figure 3, Eq. (29) reduces to

$$\det[I+KG] = \det[I+K_iG_i]\det[I+K_o\{G_o-G_{oi}(I+K_iG_i)^{-1}K_iG_{io}\}] \tag{33}$$

From Figure 3, the effective transfer function between y_o and u_o with the inner loop closed is

$$G_o' = G_o - G_{oi}(I + K_iG_i)^{-1}K_iG_{io} \tag{34}$$

Thus, Eq. (33) also becomes identical to Eq. (32).

The result in Eq. (32) is the key to relating the encirclement requirement at each step to the encirclement requirement for the complete feedback system. Using the conformal mapping identity for the product of two functions,[8] Eq. (32) yields

$$N = N_i + N_o \tag{35}$$

Thus, the number of encirclements of the origin by the Nyquist diagram for the complete feedback loop equals the sum of the number of encirclements for the inner and outer loop systems. Each loop closure contributes to the "unwrapping" of the origin. Further, by substituting Eqs. (23) and (27) into Eq. (35), it can be seen that closed-loop asymptotic stability as indicated by Eq. (3) is implied if the individual encirclement requirements for the inner and outer loop systems are achieved.

Just as in the case of single-loop closures, Eq. (32) can also be used to relate the Nyquist diagram at each "multivariable" step to the Nyquist diagram for the complete feedback system. Any point on a Nyquist diagram is a complex number with magnitude and phase. Thus, if one defines

$$\det[I + K(j\omega)G(j\omega)] = Me^{j\theta} \tag{36}$$

$$\det[I+K_i(j\omega)G_i(j\omega)] = M_ie^{j\theta_i} \qquad \det[I+K_o(j\omega)G_o'(j\omega)] = M_oe^{j\theta_o}$$

the magnitude and phase contributions from the inner and outer loop Nyquist diagrams to the complete Nyquist diagram are

$$M = M_iM_o \quad , \quad \theta = \theta_i + \theta_o \tag{37}$$

One common situation where matters are simplified is when frequency separation exists between the inner and outer loops. Suppose the inner loop crossover frequencies are in a higher frequency range and the outer loop crossovers all lie in a lower frequency range. Further, suppose the inner and outer loop shapes are as shown in Figure 5 where both loops are well attenuated above their respective crossover and the inner loop system is either type 1, 0, or -1.[11]

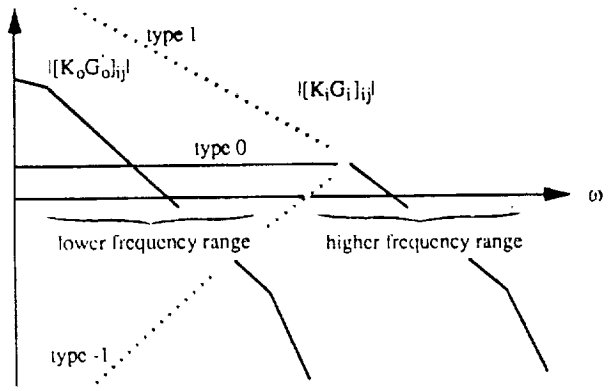


Figure 5. Inner And Outer System Loop Shapes

For frequencies well above the outer loop crossover

$$|[K_o(j\omega)G_o'(j\omega)]_{ij}| \ll 1 \tag{38}$$

and Eq. (32) becomes

$$\det[I + KG] \approx \det[I + K_iG_i] \tag{39}$$

indicating the Nyquist diagrams for the complete feedback system and the inner loop system are approximately identical. On the other hand, for frequencies well below the inner loop crossover

$$\begin{aligned} |[K_i(j\omega)G_i(j\omega)]_{ij}| &\approx \left|\frac{1}{j\omega}[\overline{KG}_i(0)]_{ij}\right| && \text{for type 1 } K_iG_i \\ |[K_i(j\omega)G_i(j\omega)]_{ij}| &\approx |[K_i(0)G_i(0)]_{ij}| && \text{for type 0 } K_iG_i \\ |[K_i(j\omega)G_i(j\omega)]_{ij}| &\approx |j\omega[\overline{KG}_i(0)]_{ij}| && \text{for type -1 } K_iG_i \end{aligned} \tag{40}$$

where $\overline{KG}_i(j\omega)$ is the remainder left over after $1/j\omega$ or $j\omega$ is factored from $K_i(j\omega)G_i(j\omega)$ and Eq. (32) becomes

$$\det[I + K(j\omega)G(j\omega)] \approx \det\left[\frac{1}{j\omega}\overline{KG}_i(0)\right]\det[I + K_o(j\omega)G_o'(j\omega)] \quad \text{for type 1 } K_iG_i$$

$$\det[I+K(j\omega)G(j\omega)] \approx \det[I+K_i(0)G_i(0)]\det[I+K_o(j\omega)G_o'(j\omega)] \quad \text{for type 0 } K_iG_i \tag{41}$$

$$\det[I + K(j\omega)G(j\omega)] \approx \det[I + K_o(j\omega)G_o'(j\omega)] \quad \text{for type -1 } K_iG_i$$

indicating the Nyquist diagrams for the complete feedback system and the outer loop system are approximately identical for type -1 K_iG_i, different by only a constant scale factor for type 0 K_iG_i, and different by a frequency dependent scale factor for type 1 K_iG_i.

Stability Robustness At Each Loop Closure

Let P_{G_T} denote the number of poles of $G_T(s)$ in the right-half plane, or

$$P_T = P_K + P_{G_T} \tag{42}$$

P_{G_T} can be separated into the number of true plant poles in the right-half plane to be stabilized with the inner loop, $P_{G_{i_T}}$, and with the outer loop, $P_{G_{o_T}'}$, or

$$P_{G_T} = P_{G_{i_T}} + P_{G_{o_T}'} \tag{43}$$

Applying Nyquist theory after the true inner loop closure yields

$$N_{i_T}(0,\det[I + K(s)G_{i_T}(s)],C_{RHP}) = Z_{i_T} - P_{i_T} \tag{44}$$

where $N_{i_T}(0,\det[I + K(s)G_{i_T}(s)],C_{RHP})$ denotes the number of encirclements of the origin by the true inner loop Nyquist diagram, Z_{i_T} is the number of true closed-loop poles of the inner loop system in the right-half plane, and P_{i_T} is the number of true open-loop poles of the inner loop system in the right-half plane. Again, requiring $Z_{i_T} = 0$ at this loop closure step is not necessary. Using the notation in Eqs. (19), (42), and (43), Z_{i_T} and P_{i_T} are given as

$$Z_{i_T} = P_{G_{o_T}'} \quad , \quad P_{i_T} = P_{K_i} + P_{G_T} \tag{45}$$

and the encirclement requirement in Eq. (44) becomes

$$N_{iT} = - P_{K_i} - P_{G_{iT}} \tag{46}$$

If

1. The nominal inner closed-loop system satisfies the encirclement requirement $N_i = - P_{K_i} - P_{G_i}$ (see Eq. (23)), and
2. The required number of encirclements of the origin is the same for both nominal and true inner closed-loop systems or $P_{G_i} = P_{G_{iT}}$ (see Eqs. (23) and (46)),

then the modeling errors $\Delta G_i(s)$ are guaranteed not to change the number of unstable poles when only the inner loop is closed if

$$\det[I + K_i(s)G_i(s,\varepsilon)]|_{s \in C_{RHP}, 0 \le \varepsilon \le 1} \neq 0 \tag{47}$$

Two sufficient conditions, developed from Eq. (47), guaranteeing the above are

$$\underline{\sigma}[I + K_i(j\omega)G_i(j\omega)] > \bar{\sigma}[K_i(j\omega)\Delta G_i(j\omega)] \quad \text{for } 0 \le \omega \le \infty \tag{48}$$

and

$$\underline{\sigma}[I + (K_i(j\omega)G_i(j\omega))^{-1}] > \bar{\sigma}[E_i(j\omega)] \quad \text{for } 0 \le \omega \le \infty \tag{49}$$

where $E_i(j\omega)$ is the inner loop input multiplicative error

$$E_i(j\omega) = (K_i(j\omega)G_i(j\omega))^{-1}(K_i(j\omega)\Delta G_i(j\omega)) \tag{50}$$
$$= G_i^{-1}(j\omega)\Delta G_i(j\omega) \text{ if } K_i(j\omega) \text{ and } G_i(j\omega) \text{ are nonsingular}$$

The validity and importance of considering stability robustness with only the inner loop closed may be unclear at this point, but it will be shown that the requirement in Eq. (47) is an integral part of the stability robustness requirement for the complete feedback loop.

Next, applying Nyquist theory to the true system, after the outer loop closure yields

$$N_{o_T}(0,\det[I + K_o(s)G'_{o_T}(s)],C_{RHP}) = Z_{o_T} - P_{o_T} \tag{51}$$

where $N_{o_T}(0,\det[I + K_o(s)G'_{o_T}(s)],C_{RHP})$ denotes the number of encirclements of the origin by the true system's outer loop Nyquist diagram, Z_{o_T} is the number of closed-loop poles in the right-half plane after the outer loop closure, and P_{o_T} is the number of open-loop poles in the right-half plane of the true system before the outer loop closure. For the block diagram structure in Figure 2, $G'_{o_T}(s)$ is defined as

$$G'_{o_T} = G_{o_T}(I + K_iG_{iT})^{-1} \tag{52}$$

which accounts for modeling errors in both $G_i(s)$ and $G_o(s)$, while for the block diagram structure in Figure 3, $G'_{o_T}(s)$ is defined as

$$G'_{o_T} = G_{o_T} - G_{oiT}(I + K_iG_{iT})^{-1}K_iG_{ioT} \tag{53}$$

which accounts for modeling errors in $G_i(s)$, $G_{io}(s)$, $G_{oi}(s)$, and $G_o(s)$. Again, the requirement

$$Z_{o_T} = 0 \quad \Rightarrow \quad N_{o_T} = - P_{o_T} \tag{54}$$

is necessary for true asymptotic stability of the complete feedback system. Using the notation in Eqs. (19), (42) and (43), P_{o_T} is given as

$$P_{o_T} = P_{K_o} + P_{G'_{o_T}} \tag{55}$$

and the encirclement requirement in Eq. (54) becomes

$$N_{o_T} = - P_{K_o} - P_{G'_{o_T}} \tag{56}$$

If

1. The nominal outer closed-loop system is asymptotically stable or $N_o = -P_{K_o} - P_{G'_o}$ (see Eq. (27)), and
2. The required number of encirclements of the origin is the same for both nominal and true outer closed-loop systems or $P_{G'_o} = P_{G'_{o_T}}$ (see Eqs. (27) and (56)),

then true closed-loop asymptotic stability is guaranteed if

$$\det[I + K_o(s)G'_o(s,\varepsilon)]|_{s \in C_{RHP}, 0 \le \varepsilon \le 1} \neq 0 \tag{57}$$

where for the block diagram structure in Figure 2, $G'_o(s,\varepsilon)$ is defined as

$$G'_o(s,\varepsilon) = G_o(s,\varepsilon)(I + K_i(s)G_i(s,\varepsilon))^{-1} \tag{58}$$

while for the block diagram structure in Figure 3, $G'_o(s,\varepsilon)$ is defined as

$$G'_o(s,\varepsilon) = G_o(s,\varepsilon) - G_{oi}(s,\varepsilon)(I + K_i(s)G_i(s,\varepsilon))^{-1}K_i(s)G_{io}(s,\varepsilon) \tag{59}$$

Two sufficient conditions guaranteeing Eq. (57) are

$$\underline{\sigma}[I + K_o(j\omega)\tilde{G}'_o(j\omega,\varepsilon)] > \bar{\sigma}[K_o(j\omega)\Delta\tilde{G}'_o(j\omega,\varepsilon)] \tag{60}$$
$$\text{for } 0 \le \omega \le \infty \text{ and } 0 \le \varepsilon \le 1$$

and

$$\underline{\sigma}[I + (K_o(j\omega)\tilde{G}'_o(j\omega,\varepsilon))^{-1}] > \bar{\sigma}[\tilde{E}'_o(j\omega,\varepsilon)] \tag{61}$$
$$\text{for } 0 \le \omega \le \infty \text{ and } 0 \le \varepsilon \le 1$$

where for the block diagram structure in Figure 2, the effective outer loop plant and additive error $\tilde{G}'_o(s,\varepsilon)$ and $\Delta\tilde{G}'_o(s,\varepsilon)$, respectively, are defined as

$$\tilde{G}'_o(s,\varepsilon) = G_o(s)(I + K_i(s)G_i(s,\varepsilon))^{-1} \tag{62}$$
$$\Delta\tilde{G}'_o(s,\varepsilon) = \Delta G_o(s)(I + K_i(s)G_i(s,\varepsilon))^{-1}$$

while for the block diagram structure in Figure 3, $\tilde{G}'_o(s,\varepsilon)$ and $\Delta\tilde{G}'_o(s,\varepsilon)$ are defined as

$$\tilde{G}'_o(s,\varepsilon) = G_o(s) - G_{oi}(s,\varepsilon)(I + K_i(s)G_i(s,\varepsilon))^{-1}K_i(s)G_{io}(s,\varepsilon)$$
$$\Delta\tilde{G}'_o(s,\varepsilon) = \Delta G_o(s) - G_{oi}(s,\varepsilon)(I + K_i(s)G_i(s,\varepsilon))^{-1}K_i(s)G_{io}(s,\varepsilon) \tag{63}$$

and where $\tilde{E}'_o(j\omega,\varepsilon)$ is the effective outer loop input multiplicative error

$$\tilde{E}'_o(j\omega,\varepsilon) = (K_o(j\omega)\tilde{G}'_o(j\omega,\varepsilon))^{-1}(K_o(j\omega)\Delta\tilde{G}'_o(j\omega,\varepsilon)) \tag{64}$$
$$= \tilde{G}'^{-1}_o(j\omega,\varepsilon)\Delta\tilde{G}'_o(j\omega,\varepsilon) \text{ if } K_o(j\omega) \text{ and } \tilde{G}'_o(j\omega,\varepsilon) \text{ are nonsingular}$$

Although conceptually the same as the standard singular value robustness tests, Eqs. (60) and (61) are more complicated because of the modeling errors present in more than one location in the feedback system. Unfortunately, the dependence upon ε can not be eliminated in a simple manner.

To understand how the inner and outer loop Nyquist diagram warpings relate to the warping of the complete feedback loop in Figure 1, consider $\det[I + K(s)\tilde{G}(s,\varepsilon)]$ and the partitioning in Eq. (13). Similar to the development in Eqs. (28) thru (34), $\det[I + K(s)G(s,\varepsilon)]$ can be expressed as

$$\det[I+K(s)G(s,\varepsilon)] = \det[I+K_i(s)G_i(s,\varepsilon)]\det[I+K_o(s)G'_o(s,\varepsilon)] \tag{65}$$

The result in Eq. (65) is the key to relating the warping of the Nyquist diagram at each step to the warping for the complete feedback system. The complete nominal feedback system is robust against modeling errors as indicated by Eq. (8) if the warpings for the inner and outer loops satisfy Eqs. (47) and (57), respectively. Note that although achieving the individual singular value robustness requirements in Eqs. (48) and (60) or (49) and (61) implies the requirement in Eq. (8), it does not necessarily imply that the the requirements in Eqs. (10) and (11), respectively, are satisfied.

Eq. (65) can also be used to relate the Nyquist diagram warping at each step to the warping for the complete feedback loop. The magnitude and phase of a point on the Nyquist diagram warping for the complete feedback loop is given by the the corresponding points on the inner and outer loop Nyquist diagram warpings similar to the idea given in Eq. (37).

For the important special case involving sufficient frequency separation between the inner and outer loops, suppose that the loop shapes shown in Figure 5 are not significantly altered by the inclusion of the modeling errors. In other words, for frequencies well above the outer loop crossover

$$|[K_o(j\omega)G_o'(j\omega,\varepsilon)]_{ij}| \ll 1 \tag{66}$$

and Eq. (65) becomes

$$\det[I + K(s)G(s,\varepsilon)] \approx \det[I + K_i(s)G_i(s,\varepsilon)] \tag{67}$$

indicating the warping of the Nyquist diagram for the complete feedback system and the inner loop system are approximately identical. On the other hand, for frequencies well below the inner loop crossover

$$\begin{aligned}
|[K_i(j\omega)G_i(j\omega,\varepsilon)]_{ij}| &\approx |\frac{1}{j\omega}[\overline{KG}_i(0,\varepsilon)]_{ij}| && \text{for type 1 } K_iG_i \\
|[K_i(j\omega)G_i(j\omega,\varepsilon)]_{ij}| &\approx |[K_i(0)G_i(0,\varepsilon)]_{ij}| && \text{for type 0 } K_iG_i \\
|[K_i(j\omega)G_i(j\omega,\varepsilon)]_{ij}| &\approx |j\omega\,[\overline{KG}_i(0,\varepsilon)]_{ij}| && \text{for type -1 } K_iG_i
\end{aligned} \tag{68}$$

where $\overline{KG}_i(j\omega,\varepsilon)$ is the remainder left over after $1/j\omega$ or $j\omega$ is factored from $K_i(j\omega)G_i(j\omega,\varepsilon)$ and Eq. (65) becomes

$$\det[I+K(j\omega)G(j\omega,\varepsilon)] \approx \det[\frac{1}{j\omega}\overline{KG}_i(0,\varepsilon)]\det[I+K_o(j\omega)G_o'(j\omega,\varepsilon)]$$

for type 1 K_iG_i

(69)

$$\det[I+K(j\omega)G(j\omega,\varepsilon)] \approx \det[I+K_i(0)G_i(0,\varepsilon)]\det[I+K_o(j\omega)G_o'(j\omega,\varepsilon)]$$

for type 0 K_iG_i

$$\det[I + K(j\omega)G(j\omega,\varepsilon)] \approx \det[I + K_o(j\omega)G_o'(j\omega,\varepsilon)]$$

for type -1 K_iG_i

indicating the warping of the Nyquist diagram for the complete feedback system and the outer loop system are approximately identical for type -1 K_iG_i, different by only an ε dependent scale factor for type 0 K_iG_i, and different by a frequency and ε dependent scale factor for type 1 K_iG_i.

Example

The example to be considered involves the longitudinal flight control of a large, flexible aircraft. Controlled inputs consist of elevator deflection δ_E and canard deflection δ_C while responses of interest include the pitch rate measured at two locations on the fuselage, q_1 and q_2, and the surge velocity u. The model for the aircraft dynamics is 12th order and the state space description is given in the Appendix. Frequency responses for elevator deflection are shown in Figures 6 thru 8. The open-loop eigenvalues consist of

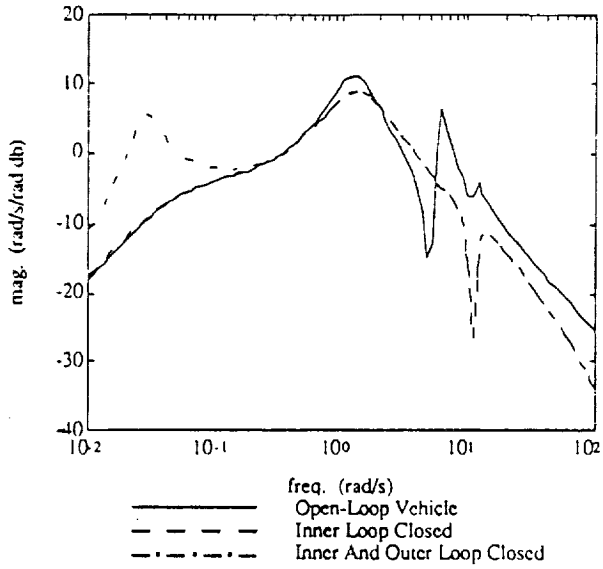


Figure 6. $q_1(s)/\delta_E(s)$ And $q_1(s)/\delta(s)$ Frequency Responses

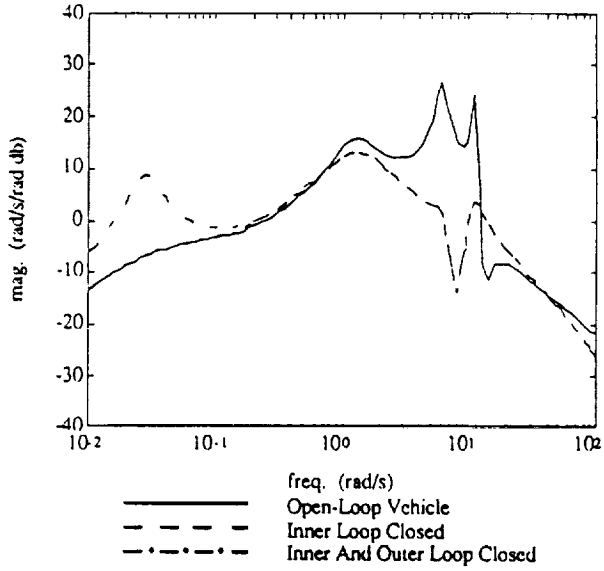


Figure 7. $q_2(s)/\delta_E(s)$ And $q_2(s)/\delta(s)$ Frequency Responses

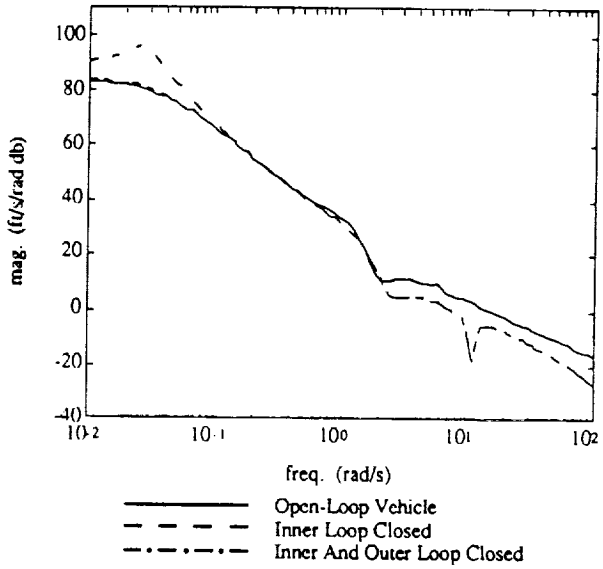


Figure 8. $u(s)/\delta_E(s)$ And $u(s)/\delta(s)$ Frequency Responses

0.033	phugoid mode
-0.043	phugoid mode
-0.45±j1.2	short period mode
-0.44±j6.0	1st aeroelastic mode
-0.22±j11.0	2nd aeroelastic mode
-0.36±j11.0	3rd aeroelastic mode
-2.6±j13.0	4th aeroelastic mode

From Figures 6 thru 8 and the open-loop eigenvalue data, observe the low damping of the short period and aeroelastic modes, the significant aeroelastic contributions to the pitch rate responses, and the unstable phugoid mode.

The flight control design objectives are to increase the damping of the short period and aeroelastic modes, reduce the aeroelastic contributions to the pitch rate responses, and stabilize the phugoid mode. With the existing frequency separation between the phugoid mode and the other modes, the flight control synthesis will be accomplished in a two-step approach as indicated in Figure 2. The inner loop closure consists of angular rates q_1 and q_2 feed back to δ_E and δ_C, respectively, while the outer loop closure consists of speed u feed back to δ_E.

The inner loop compensation was synthesized in Ref. 9 and is briefly reviewed here. First, the q_2/δ_C loop is closed to improve the 1st aeroelastic mode damping. Next, a δ_E to δ_C crossfeed is introduced to reduce 1st aeroelastic mode excitations from δ_E. Finally, the q_1/δ_E loop is closed to improve the short period damping. Ref. 9 neglected the 2nd and higher aeroelastic modes, thus a notch filter is introduced here at 11 rad/s to reduce the significant 2nd aeroelastic mode contribution and a low pass filter with a bandwidth of 60 rad/s is introduced for attenuation of higher frequency aeroelastic modes. Inclusion of the notch and low pass filters introduced approximately 15 deg of phase lag at the 1st aeroelastic mode frequency. With this, the inner loop compensator is

$$K_I(s) = \frac{60}{s+60}\,\frac{s^2+.47s+116}{s^2+3.2s+116}\begin{bmatrix} -.05 & 0 \\ .075 & .05 \end{bmatrix}\left(\frac{\text{rad}}{\text{rad/s}}\right) \qquad (70)$$

and the block diagram structure is shown in Figure 9, where δ represents the pilot stick inputs.

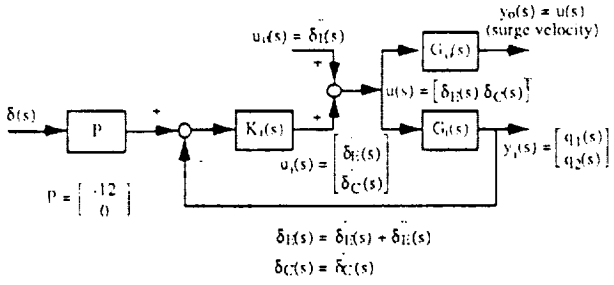


Figure 9. Inner Loop Block Diagram

Figures 6 thru 8 show the effect of the inner loop closure on the frequency responses and the intermediate closed-loop eigenvalues are

0.031	phugoid mode
-0.042	phugoid mode
-0.69±j1.1	short period mode
-0.75±j6.1	1st aeroelastic mode
-0.22±j11.	2nd aeroelastic mode
-0.36±j11.	3rd aeroelastic mode
-2.6±j13.	4th aeroelastic mode
-1.7±11.	compensator mode
-60.	compensator mode

Note the increased short period and 1st aeroelastic mode damping as well as the more rigid-body like pitch rate frequency responses. Also note the relatively unaffected phugoid characteristics.

The Nyquist diagram corresponding to $\det[I + K_I(s)G_I(s)]$ is shown in Figure 10 and note that $N_i = 0$, $Z_i = 1$, and $P_i = 1$. Further, the intermediate singular value "robustness" (see Eq. (49)) is plotted in Figure 11.

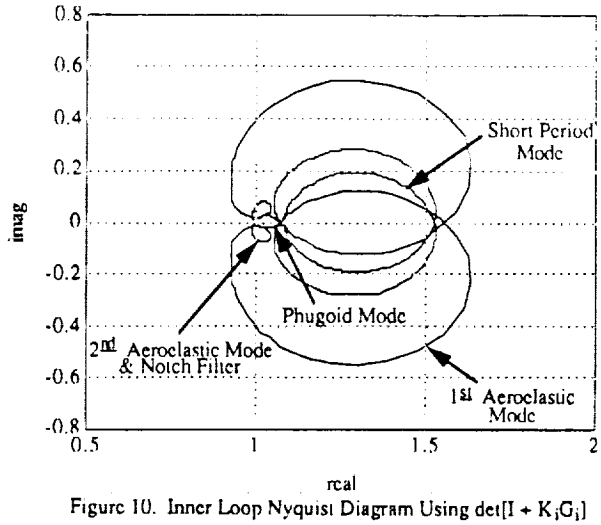


Figure 10. Inner Loop Nyquist Diagram Using $\det[I + K_IG_I]$

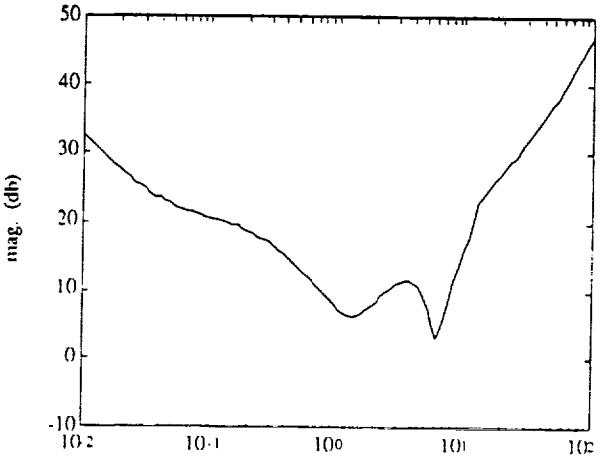


Figure 11. Inner Loop Singular Value Robustness Characteristic $\underline{\sigma}[I + (K_IG_I)^{-1}]$

Now, the outer loop consists of constant gain feedback of speed (u), or the outer loop compensator is

$$K_O(s) = 0.0001\ \left(\frac{\text{rad}}{\text{ft/s}}\right) \qquad (71)$$

and the block diagram is shown in Figure 12. Figures 6 thru 8 show the effect of the outer loop closure on the frequency responses and the final closed-loop eigenvalues are

-0.0067±0.027	phugoid mode
-0.69±j1.1	short period mode
-0.75±j6.1	1st aeroelastic mode
-0.22±j11.	2nd aeroelastic mode
-0.36±j11.	3rd aeroelastic mode
-2.6±j13.	4th aeroelastic mode
-1.7±11.	compensator mode
-60.	compensator mode

Note the stabilized phugoid mode and the relatively unaffected higher frequency modes.

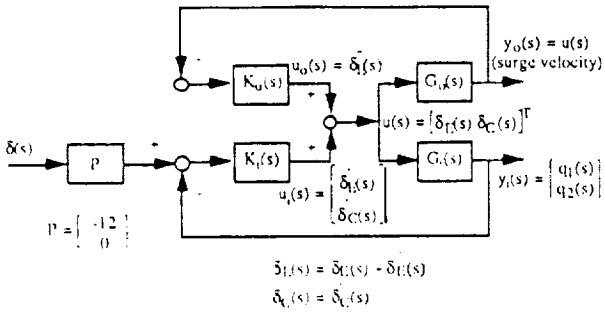


Figure 12. Inner And Outer Loop Block Diagram

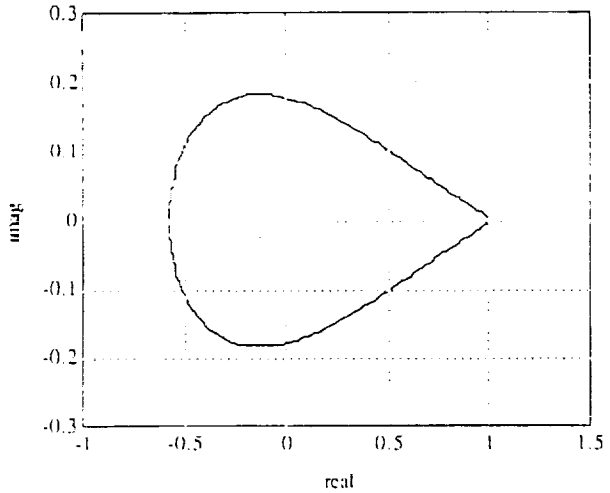


Figure 13. Outer Loop Nyquist Diagram Using $\det[I + K_o G_o']$

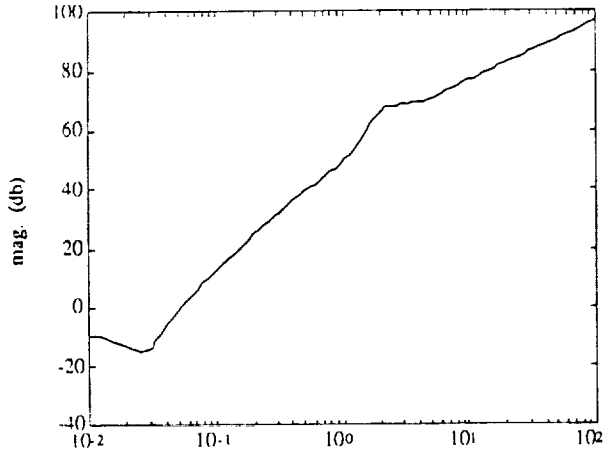


Figure 14. Outer Loop Singular Value Robustness Characteristic $\underline{\sigma}[I + (K_o G_o')^{-1}]$

The Nyquist diagram corresponding to $\det[I + K_o(s)G_o'(s)]$ is shown in Figure 13 and it is seen that $N_o = -1$, $Z_o = 0$, and $P_o = 1$. Further, the singular value robustness (see Eq. (61)) with $\varepsilon = 0$ for this scalar loop is plotted in Figure 14.

If the block diagram structure in Figure 12 is recast into the format of Figure 1, then the Nyquist diagram corresponding to the complete feedback system or $\det[I + K(s)G(s)]$ is shown in Figure 15. Note the aggregate of the inner and outer loop Nyquist diagrams closely matches the complete Nyquist diagram. It is insightful to see how each loop closure contributes to the shape of the overall Nyquist diagram and the required number of encirclements of the origin. Further, the singular value robustness (see Eq. (11)) is shown in Figure 16.

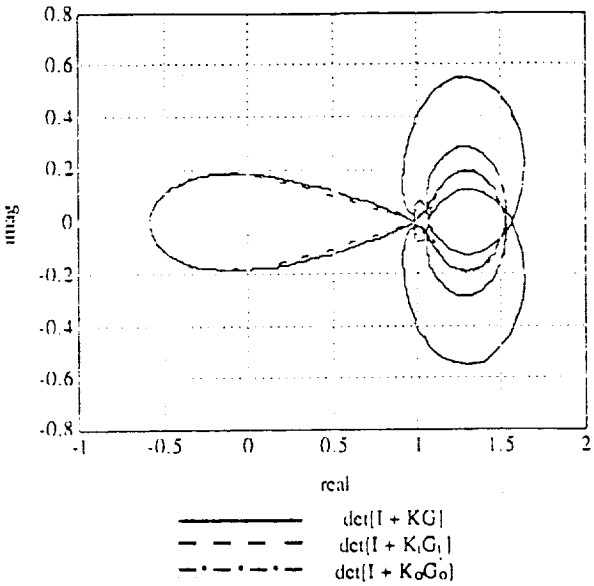


Figure 15. Nyquist Diagram Comparisons

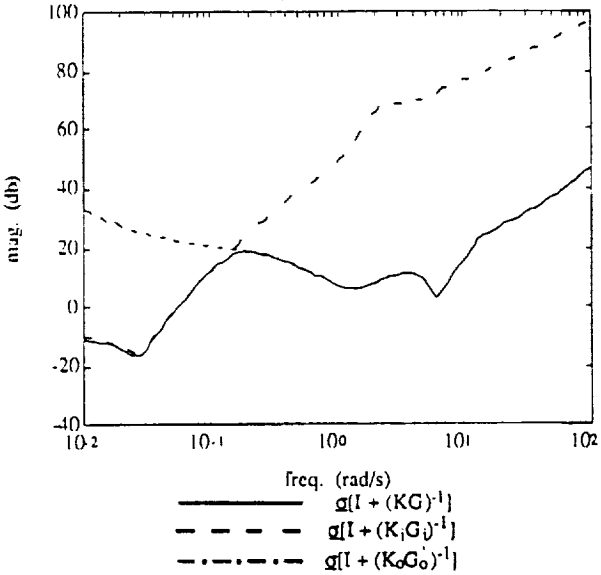


Figure 16. Singular Value Robustness Test Comparisons

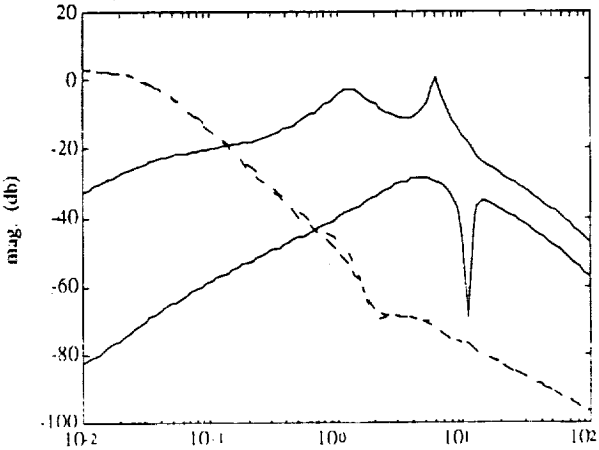


Figure 17. Inner And Outer Loop Shapes

Note the contributions from the inner and outer loop to the complete feedback loop. Here, the match occurs because of several special features in the loop shapes as shown in Figure 17. With the block diagram structure in Figure 2 or 12, the loop gain is given as

$$KG = K_iG_i + K_oG_o \tag{72}$$

Further, for this example, K_oG_o and K_oG_o' are approximately equal as indicated in Figure 17. Therefore, as seen from Figure 17, the higher frequency range match occurs because K_oG_o is sufficiently attenuated relative to K_iG_i, or

$$KG \approx K_iG_i \tag{73}$$

Also, as seen from Figure 17, the lower frequency range match occurs because K_iG_i is type -1 (small K_iG_i relative to K_oG_o), or

$$KG \approx K_oG_o' \tag{74}$$

Conclusions

It has been shown how the nominal, multivariable Nyquist diagram and it's continuous warping to the true shape for the overall feedback system is related to the contributions from the inner and outer loops. The encirclement requirement of the overall feedback system to assure nominal asymptotic stability is converted to the encirclement requirements for the inner and outer loops. Further, to assure robustness against modeling errors, the requirement of avoiding the origin, when the Nyquist diagram is warped from the nominal shape to the true shape, is converted to similar requirements for the inner and outer loops. The implications for analysis and design are that the overall stability and robustness characteristics can be decomposed into contributions from the inner and outer loops, which can offer guidance in feedback design.

Acknowledgements

This research was supported by NASA Langley Research Center under Grant NAG1-758. Mr. D. Arbuckle has served as the technical monitor. This support is appreciated.

References

1. Nyquist, H., "Regeneration Theory," Bell System Technical Journal, Vol. 11, January, 1932

2. MacFarlane, A. G. J. and Postlethwaite, I., "The Generalized Nyquist Stability Criterion And Multivariable Root Loci," International Journal Of Control, Vol. 25, No. 1, January, 1977

3. Doyle, J. C. and Stein, G., "Multivariable Feedback Design: Concepts For A Classical/Modern Synthesis," IEEE Transactions On Automatic Control, Vol. AC-26, No. 1, February, 1981

4. Lehtomaki, N. A., "Practical Robustness Measures In Multivariable Control System Analysis," Ph.D. Dissertation, Department Of Electrical Engineering And Computer Science, Massachusetts Institute Of Technology, May, 1981

5. McRuer, D., Ashkenas, I., and Graham, D., *Aircraft Dynamics And Automatic Control*, Princeton University Press, 1973

6. Maciejowski, J. M., *Multivariable Feedback Design*, Addison-Wesley, 1989

7. Hsu, C. and Chen, C., "A Proof Of The Stability Of Multivariable Feedback Systems," Proceedings Of The IEEE, Vol. 56, No. 11, November, 1968

8. Churchill, R. V. and Brown, J. W., *Complex Variables And Applications*, McGraw-Hill,1984

9. Newman, B. and Schmidt, D. K., "Multivariable Flight Control Synthesis And Literal Robustness Analysis For An Aeroelastic Vehicle," Proceedings Of The AIAA Guidance, Navigation, And Control Conference, Portland, Oregon, August, 1990

10. Nobel, B. and Daniel, J. W., *Applied Linear Algebra*, Prentice-Hall, 1977

11. Ogata, K., *Modern Control Engineering*, Prentice-Hall, 1970

Appendix

The aircraft model is

$$\dot{x} = Ax + Bu$$
$$y = Cx \tag{75}$$

where

$$y = \begin{bmatrix} q_1\ (\text{rad/s}) \\ q_2\ (\text{rad/s}) \\ u\ (\text{ft/s}) \end{bmatrix} \qquad u = \begin{bmatrix} \delta_E\ (\text{rad}) \\ \delta_C\ (\text{rad}) \end{bmatrix} \tag{76}$$

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$A_{11} = \begin{bmatrix} -1.451e\text{-}2 & 1.935e+1 & -3.220e+1 & -1.907e+0 & 0 & 0 \\ -1.487e\text{-}4 & -4.285e\text{-}1 & 0 & 1.025e+0 & -4.223e\text{-}3 & -1.445e\text{-}4 \\ 0 & 0 & 0 & 1.e+0 & 0 & 0 \\ -1.090e\text{-}4 & -3.430e+0 & 0 & -8.335e\text{-}1 & -6.625e\text{-}2 & -3.814e\text{-}3 \\ 0 & 0 & 0 & 0 & 0 & 1.e+0 \\ 4.227e\text{-}1 & -1.072e+3 & 0 & -7.935e+1 & -3.536e+1 & -6.028e\text{-}1 \end{bmatrix}$$

$$A_{12} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 4.487e\text{-}2 & 1.755e\text{-}3 & 2.170e\text{-}3 & 1.036e\text{-}4 & -2.053e\text{-}3 & -3.391e\text{-}5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -5.160e\text{-}2 & 2.950e\text{-}2 & 8.544e\text{-}2 & 4.126e\text{-}3 & -3.777e\text{-}2 & -1.190e\text{-}3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -6.475e+0 & -1.647e\text{-}1 & 2.554e+1 & 5.435e\text{-}1 & 8.633e+0 & -1.171e\text{-}1 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 2.773e\text{-}2 & 3.552e+1 & 0 & 1.858e\text{-}1 & 5.801e+0 & 1.395e\text{-}1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -6.653e\text{-}2 & 1.475e+2 & 0 & 4.567e+0 & 2.880e+0 & -7.236e\text{-}2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 3.078e\text{-}5 & 1.014e\text{-}2 & 0 & 9.965e\text{-}5 & 6.696e\text{-}3 & -1.954e\text{-}4 \end{bmatrix}$$

$$A_{22} = \begin{bmatrix} 0 & 1.e+0 & 0 & 0 & 0 & 0 \\ -1.766e+2 & -5.054e+0 & 2.714e+0 & 6.710e\text{-}2 & 4.643e+0 & 6.646e\text{-}2 \\ 0 & 0 & 0 & 1.e+0 & 0 & 0 \\ 1.425e+1 & 2.243e+0 & -1.156e+2 & -4.246e\text{-}1 & 1.425e+0 & -1.507e\text{-}1 \\ 0 & 0 & 0 & 0 & 0 & 1.e+0 \\ -3.999e\text{-}2 & 1.440e\text{-}3 & 2.933e\text{-}3 & -1.641e\text{-}4 & -1.211e+2 & -7.226e\text{-}1 \end{bmatrix}$$

$$B = \begin{bmatrix} 1.477e+1 & 0 \\ -6.384e\text{-}2 & -1.248e\text{-}2 \\ 0 & 0 \\ -5.321e+0 & 8.392e\text{-}1 \\ 0 & 0 \\ -9.230e+2 & -6.211e+2 \\ 0 & 0 \\ -8.841e+1 & -1.107e+1 \\ 0 & 0 \\ 2.529e+2 & -4.571e+1 \\ 0 & 0 \\ 4.445e\text{-}2 & -1.521e+0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 0 & 1.e+0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1.e+0 & 1.e+0 & 0 \\ 0 & 0 & 0 \\ 0 & -2.100e\text{-}2 & 0 \\ 0 & 0 & 0 \\ 0 & -2.300e\text{-}2 & 0 \\ 0 & 0 & 0 \\ 0 & -3.200e\text{-}2 & 0 \\ 0 & 0 & 0 \\ 0 & -3.400e\text{-}2 & 0 \end{bmatrix}^T$$

Appendix B

AIAA-91-2805-CP

New Plant And Controller Order Reduction Results With Weighted Balancing

Brett Newman* and David K. Schmidt**
Aerospace Research Center
Arizona State University
Tempe, Arizona

Abstract

Frequency weighted internally balanced (FWIB) truncation is briefly reviewed. A previous frequency response error analysis for FWIB truncation is extended and an exact error bound for the special case of order reduction by one state is presented in terms of the controllability-observability measure used in selecting the coordinate to truncate, as well as two additional frequency dependent variables. The two additional variables are shown to be small when the controllability-observability measure is small, justifying the reduction technique based only upon consideration of the controllability-observability measures. An approximate error bound for the general case of order reduction by more than one state, under the assumption that only small controllability-observability measures are truncated, is presented. FWIB residualization is presented and a frequency response error analysis yields results similar to that found for FWIB truncation. Numerical examples are given to support the error analysis results, as well as to stress that FWIB truncation and residualization can be used in a coordinated manner to achieve higher accuracy than that achievable from either technique used alone.

Introduction

Models developed from the governing physical principles are often of high dynamic order,[1,2] complicating the direct use of the model in the intended application. For example, control law synthesis is a common application for dynamic models. However, many modern linear control synthesis techniques produce a controller with dynamic order at least equal to the plant dynamic order.[3,4] This is unacceptable for controller implementation. Thus, order reduction of dynamic models is of extreme importance.

Here it is assumed the system that is modeled will be actively controlled in a feedback loop such as in Figure 1, with responses y(s), control inputs u(s), response commands $y_c(s)$, plant transfer function matrix G(s), and compensator transfer function matrix K(s). A reduced order model for the plant, $G_R(s)$ (or for the compensator, $K_R(s)$), should preserve the key frequency domain characteristics of the higher order model,[5-7] and an order reduction technique specifically tailored for this task is frequency weighted internally balanced (FWIB) truncation.[5]

In this technique, coordinates reflecting small measures of weighted controllability-observability are truncated based upon the engineering argument that this procedure will yield a reduced order model that matches the frequency response of the higher order model in the critical frequency range, as numerous examples have demonstrated.[8,9] As of yet, however, there exists no rigorous theoretical justification for this technique, that guarantees a small error in the critical frequency range, similar to the result discovered for unweighted internally balanced (IB) truncation.[5,10] Enns[5] did consider a frequency response error analysis of this technique, but his result was left so cumbersome that its utility was limited. The first goal of this paper is to extend the frequency response error analysis of Reference 5 for FWIB truncation, so as to develop a theoretical justification for this weighted technique.

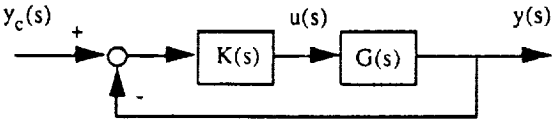


Figure 1. Generic Feedback Loop

Recall from classical truncation and residualization theory that truncation is most appropriate for eliminating lower frequency dynamics, while residualization is most appropriate for eliminating higher frequency dynamics, relative to the frequency range of interest.[8] Reference 11 recently considered this point and showed that all the properties existing for IB truncation also exist for IB residualization, including an upper bound on the frequency response error identical to that for IB truncation. In light of these results, a second goal of this paper is to establish FWIB residualization as an acceptable order reduction technique, to be used in conjunction with FWIB truncation.

Truncation And Residualization With FWIB States

Consider a finite dimensional, linear, time invariant state space model representing the higher order plant in Figure 1, or

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{1}$$

Also consider an input weighting filter

$$\begin{aligned} \dot{x}_{wi}(t) &= A_{wi}x_{wi}(t) + B_{wi}\delta(t) \\ u(t) &= C_{wi}x_{wi}(t) + D_{wi}\delta(t) \end{aligned} \tag{2}$$

and an output weighting filter

$$\begin{aligned} \dot{x}_{wo}(t) &= A_{wo}x_{wo}(t) + B_{wo}y(t) \\ \gamma(t) &= C_{wo}x_{wo}(t) + D_{wo}y(t) \end{aligned} \tag{3}$$

cascaded with the higher order model in Eq. (1), as shown in Figure 2, where $G_{wi}(s)$ and $G_{wo}(s)$ are the corresponding weighting transfer function matrices, respectively. The input weighting filter is used to adjust the frequency response such that $\delta(j\omega)$ to $y(j\omega)$ is approximately the same as $u(j\omega)$ to $y(j\omega)$ in the frequency range of interest, and is well attenuated outside the frequency range of interest, while the output weighting filter is used to adjust the frequency response such that $u(j\omega)$ to $\gamma(j\omega)$ is approximately the same as $u(j\omega)$ to $y(j\omega)$ in the frequency range of interest, and is well attenuated outside the frequency range of interest.

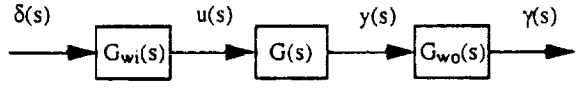


Figure 2. Frequency Weighted Model

The weighted controllability grammian X and the weighted observability grammian Y for the higher order model in Eq. (1) are defined as[5,6]

$$\begin{aligned} X &= \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{X}(j\omega)\overline{X}^{*}(j\omega)d\omega \\ &\text{with} \quad \overline{X}(j\omega) = (j\omega I - A_X)^{-1}B_X \\ Y &= \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{Y}^{*}(j\omega)\overline{Y}(j\omega)d\omega \\ &\text{with} \quad \overline{Y}(j\omega) = C_Y(j\omega I - A_Y)^{-1} \end{aligned} \tag{4}$$

$$A_X = \begin{bmatrix} A & BC_{wi} \\ 0 & A_{wi} \end{bmatrix} \qquad B_X = \begin{bmatrix} BD_{wi} \\ B_{wi} \end{bmatrix}$$

* Research Associate; Doctoral Student, School Of Aeronautics And Astronautics, Purdue University; Student Member AIAA
** Acting Director; Professor Of Mechanical And Aerospace Engineering; Associate Fellow AIAA

$$A_Y = \begin{bmatrix} A & 0 \\ B_{wo}C & A_{wo} \end{bmatrix} \qquad C_Y = \begin{bmatrix} D_{wo}C & C_{wo} \end{bmatrix}$$

Further, if A, A_{wi}, and A_{wo} are asymptotically stable, then X and Y are the unique, positive semidefinite solutions to

$$\begin{aligned} A_X X + X A_X^* + B_X B_X^* &= 0 \\ A_Y^* Y + Y A_Y + C_Y^* C_Y &= 0 \end{aligned} \tag{5}$$

FWIB states $\hat{x}$, which decompose the system such that weighted controllability and observability are balanced, are related to the state vector x in Eq. (1) by the transformation[5]

$$x(t) = T\hat{x}(t) \tag{6}$$

where T is given as

$$T = VW \tag{7}$$

with V decomposing $X_{11}Y_{11}$, X_{11}, and Y_{11} as

$$\begin{aligned} X_{11}Y_{11} &= V\Sigma^2 V^{-1} \qquad X_{11} = V\Sigma_c^2 V^T \qquad Y_{11} = V^{-1^T}\Sigma_o^2 V^{-1} \\ \Sigma &= \Sigma_c \Sigma_o \qquad\qquad \Sigma = \mathrm{diag}\{\sigma_i\} \;\; \sigma_i \ge 0 \\ \Sigma_c &= \mathrm{diag}\{\sigma_{c_i}\} \;\; \sigma_{c_i} \ge 0 \qquad \Sigma_o = \mathrm{diag}\{\sigma_{o_i}\} \;\; \sigma_{o_i} \ge 0 \end{aligned} \tag{8}$$

(note that the σ_i's, σ_{c_i}'s, and σ_{o_i}'s are all real numbers) and W defined as

$$W = \mathrm{diag}\{w_i\} \qquad w_i = \begin{cases} \left(\dfrac{\sigma_{c_i}}{\sigma_{o_i}}\right)^{1/2} & \text{if } \sigma_{c_i} \neq 0 \text{ and } \sigma_{o_i} \neq 0 \\ 1 & \text{if } \sigma_{c_i} = 0 \text{ or } \sigma_{o_i} = 0 \end{cases} \tag{9}$$

Further, the partitions X_{11} and Y_{11} of the weighted controllability and observability grammians are transformed such that they are equal and diagonal, or

$$\hat{X}_{11} = \hat{Y}_{11} = \Sigma \tag{10}$$

A key result from References 5 and 10 concerning order reduction is that the σ_i's are measures of the controllability-observability of the corresponding coordinates in $\hat{x}$ for the weighted system. Therefore, states with larger values of σ_i contribute more, in effect, to the weighted frequency responses. (This is the claim that has not been rigorously justified with a frequency response error analysis.)

Now assume the higher order model in Eq. (1) is FWIB and suppose the higher order and reduced order models have dynamic order n and n_R, respectively. Further, suppose the σ_i's are ordered from smallest to largest. Partitioning of $X_{11}Y_{11}$ as follows,

$$\begin{aligned} X_{11}Y_{11} = \Sigma^2 &= \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & \Sigma_2^2 \end{bmatrix} \\ \Sigma_1 &= \mathrm{diag}\{\sigma_i\} \quad i = 1, \ldots, n - n_R \\ \Sigma_2 &= \mathrm{diag}\{\sigma_i\} \quad i = n - n_R + 1, \ldots, n \\ 0 &\le \sigma_1 \le \cdots \le \sigma_n \end{aligned} \tag{11}$$

leads to a partitioning of Eq. (1), or

$$\begin{aligned} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(t) \\ y(t) &= [C_1 C_2] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + Du(t) \end{aligned} \tag{12}$$

where states x_1 are less controllable-observable and states x_2 are more controllable-observable in the weighted, balanced system. Truncation of the states x_1 leads to the reduced model[5]

$$\begin{aligned} \dot{x}_2(t) &= A_{22}x_2(t) + B_2 u(t) \\ y(t) &= C_2 x_2(t) + Du(t) \end{aligned} \tag{13}$$

Suppose the σ_i's in Eq. (8) are now ordered from largest to smallest. Partitioning of $X_{11}Y_{11}$ as follows,

$$\begin{aligned} X_{11}Y_{11} = \Sigma^2 &= \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & \Sigma_2^2 \end{bmatrix} \\ \Sigma_1 &= \mathrm{diag}\{\sigma_i\} \quad i = 1, \ldots, n_R \\ \Sigma_2 &= \mathrm{diag}\{\sigma_i\} \quad i = n_R + 1, \ldots, n \\ \sigma_1 &\ge \cdots \ge \sigma_n \ge 0 \end{aligned} \tag{14}$$

leads to a partitioning as in Eq. (12), where states x_1 are more controllable-observable and states x_2 are less controllable-observable in the weighted, balanced system. Residualization of the states x_2 (provided A_{22} is nonsingular) leads to the reduced model[12]

$$\begin{aligned} \dot{x}_1(t) &= (A_{11} - A_{12}A_{22}^{-1}A_{21})x_1(t) + (B_1 - A_{12}A_{22}^{-1}B_2)u(t) \\ y(t) &= (C_1 - C_2A_{22}^{-1}A_{21})x_1(t) + (D - C_2A_{22}^{-1}B_2)u(t) \end{aligned} \tag{15}$$

FWIB Truncation Error Analysis

Let the transfer function error from order reduction be defined as

$$E(s) = G(s) - G_R(s) \tag{16}$$

The magnitude of the individual elements of $E(j\omega)$ in the frequency range of interest are an important measure of the accuracy of the reduced order model.[5,6] A closely related measure is the individual elements of the weighted frequency response error defined as

$$E_w(j\omega) = G_{wo}(j\omega)E(j\omega)G_{wi}(j\omega) \tag{17}$$

Note that if $G_{wi}(j\omega)$ and $G_{wo}(j\omega)$ leave the frequency response unaffected in the frequency range of interest and provide high attenuation otherwise, then $E_w(j\omega)$ and $E(j\omega)$ are essentially the same in the frequency range of interest, and $E_w(j\omega)$ is small otherwise. Finally, the maximum singular value of $E_w(j\omega)$, denoted as $\bar{\sigma}[E_w(j\omega)]$, is an upper bound on the magnitude of the elements of $E_w(j\omega)$.[7]

It can be shown that the weighted truncation error can be expressed as[5,12]

$$E_w(j\omega) = \tilde{C}_w(j\omega)\Delta^{-1}(j\omega)\tilde{B}_w(j\omega) \tag{18}$$

where

$$\begin{aligned} \tilde{B}_w(j\omega) &= \tilde{B}(j\omega)G_{wi}(j\omega) \qquad \tilde{C}_w(j\omega) = G_{wo}(j\omega)\tilde{C}(j\omega) \\ \tilde{B}(j\omega) &= B_1 + A_{12}\Phi(j\omega)B_2 \qquad \tilde{C}(j\omega) = C_1 + C_2\Phi(j\omega)A_{21} \\ \Phi(j\omega) &= (j\omega I - A_{22})^{-1} \qquad \Delta(j\omega) = j\omega I - A_{11} - A_{12}\Phi(j\omega)A_{21} \end{aligned} \tag{19}$$

Further, the maximum singular value of the weighted error can be expressed as

$$\bar{\sigma}^2[E_w] = \bar{\lambda}\left[\Delta^{-1}\tilde{B}_w\tilde{B}_w^*\Delta^{-1^*}\tilde{C}_w^*\tilde{C}_w\right] \tag{20}$$

where $\bar{\lambda}$ denotes the maximum eigenvalue. By expanding the products $\tilde{B}_w(j\omega)\tilde{B}_w^*(j\omega)$ and $\tilde{C}_w^*(j\omega)\tilde{C}_w(j\omega)$ and using the transformed, partitioned Eq. (5), it can be shown that[5,12]

$$\tilde{B}_w\tilde{B}_w^* = \Delta\{\Sigma_1 + (X_{12})_1(-j\omega I - A_{wi}^*)^{-1}C_{wi}^*\tilde{B}^*\} \quad (21)$$
$$\{\Sigma_1 + \tilde{B}C_{wi}(j\omega I - A_{wi})^{-1}(X_{21})_1\}\Delta^*$$
$$\tilde{C}_w^*\tilde{C}_w = \Delta^*\{\Sigma_1 + (Y_{12})_1(j\omega I - A_{wo})^{-1}B_{wo}\tilde{C}\}$$
$$\{\Sigma_1 + \tilde{C}^*B_{wo}^*(-j\omega I - A_{wo}^*)^{-1}(Y_{21})_1\}\Delta$$

where $(X_{21})_1$ and $(Y_{12})_1$ are partitions of X_{21} and Y_{12}, respectively, induced by Eq. (12), or

$$X_{21} = [\ (X_{21})_1 \quad (X_{21})_2\ ] \qquad Y_{12} = \begin{bmatrix} (Y_{12})_1 \\ (Y_{12})_2 \end{bmatrix} \quad (22)$$

After substituting Eq. (21) into Eq. (20), $\bar{\sigma}[E_w(j\omega)]$ can be rewritten as[5,12]

$$\bar{\sigma}^2[E_w] = \bar{\lambda}\{\{\Sigma_1 + M^* + \Delta^{-1}(\Sigma_1 + M)\Delta^*\} \quad (23)$$
$$\{\Sigma_1 + N + \Delta^{-1^*}(\Sigma_1 + N^*)\Delta\}\}$$

where

$$M(j\omega) = \tilde{B}(j\omega)C_{wi}(j\omega I - A_{wi})^{-1}(X_{21})_1 \quad (24)$$
$$N(j\omega) = (Y_{12})_1(j\omega I - A_{wo})^{-1}B_{wo}\tilde{C}(j\omega)$$

Note that the terms $M(j\omega)$ and $N(j\omega)$ are functions of the frequency weightings.

For $n - n_R = 1$, then $\Sigma_1 = \sigma_1$, $\Delta(j\omega) = \delta(j\omega)$, $M(j\omega) = m(j\omega)$, and $N(j\omega) = n(j\omega)$ all become scalars, and Eq. (23) yields[12]

$$\bar{\sigma}^2[E_w] = \{1 + b^{-1}b^*a\}\{1 + c^{-1}c^*a^*\}(\sigma_1 + m^*)(\sigma_1 + n) \quad (25)$$

where

$$a(j\omega) = \delta^{-1}(j\omega)\delta^*(j\omega) \quad (26)$$
$$b(j\omega) = \sigma_1 + m^*(j\omega) \qquad c(j\omega) = \sigma_1 + n(j\omega)$$

provided $\sigma_1 \neq -m^*(j\omega)$ and $\sigma_1 \neq -n(j\omega)$ for all frequencies. (Note that σ_1 is a real number, while $m(j\omega)$ and $n(j\omega)$ are complex numbers, making it unlikely that $\sigma_1 = -m^*(j\omega)$ and $\sigma_1 = -n(j\omega)$.) Since the right-hand side of Eq. (25) represents a positive semidefinite matrix $(E_w(j\omega)E_w^*(j\omega))$, taking the absolute value does not alter the equality, or

$$\bar{\sigma}^2[E_w] = |1 + b^{-1}b^*a| \cdot |1 + c^{-1}c^*a^*| \cdot |\sigma_1 + m| \cdot |\sigma_1 + n| \quad (27)$$
$$\leq \{1 + |b^{-1}b^*| \cdot |a|\}\{1 + |c^{-1}c^*| \cdot |a|\}\,|\sigma_1 + m| \cdot |\sigma_1 + n|$$

Substitution of the equalities $|a(j\omega)| = 1$, $|b^{-1}(j\omega)b^*(j\omega)| = 1$, and $|c^{-1}(j\omega)c^*(j\omega)| = 1$ into Eq. (27) yields the error bound for order reduction by one state[12]

$$\bar{\sigma}[E_w(j\omega)] \leq 2\{\,|\sigma_1 + m(j\omega)| \cdot |\sigma_1 + n(j\omega)|\,\}^{1/2} \quad (28)$$

Observe that the structure of this error bound is quite similar to that for IB truncation[5] and clearly reduces to the IB result when the weighting filters are selected to be unity (i.e., $m(j\omega) = n(j\omega) = 0$).

Analysis Of $m(j\omega)$ And $n(j\omega)$

Recall that FWIB truncation is based upon the engineering premise that states corresponding to small σ_i's contribute little to the frequency response in the frequency range of interest, and thus can be eliminated. Does the result in Eq. (28) imply that one should eliminate the state corresponding to the smallest value for $|\sigma_i + m(j\omega)| \cdot |\sigma_i + n(j\omega)|$? Numerous examples,[8,9] demonstrating that FWIB truncation yields an accurate reduced order model in the frequency range of interest, suggest that the current procedure of considering only the relative sizes of the σ_i's is sufficient in the sense that, states corresponding to smaller values of σ_i also correspond to the smaller values for $|\sigma_i + m(j\omega)| \cdot |\sigma_i + n(j\omega)|$, and this is to be shown next.

$\bar{X}(j\omega)$ and $\bar{Y}(j\omega)$ from Eq. (4) can be partitioned as

$$\bar{X}(j\omega) = \begin{bmatrix} \bar{X}_G(j\omega)G_{wi}(j\omega) \\ \bar{X}_{wi}(j\omega) \end{bmatrix} \quad (29)$$

with $\bar{Y}(j\omega) = [\ G_{wo}(j\omega)\bar{Y}_G(j\omega)\ \ \bar{Y}_{wo}(j\omega)\ ]$

$$\bar{X}_G(j\omega) = (j\omega I - A)^{-1}B \qquad \bar{Y}_G(j\omega) = C(j\omega I - A)^{-1} \quad (30)$$
$$\bar{X}_{wi}(j\omega) = (j\omega I - A_{wi})^{-1}B_{wi} \quad \bar{Y}_{wo}(j\omega) = C_{wo}(j\omega I - A_{wo})^{-1}$$

Observe that the ij^{th} element of $\bar{X}_G(j\omega)$ is the frequency response of the i^{th} state x_i from a unit impulse in the j^{th} input u_j and the ij^{th} element of $\bar{Y}_G(j\omega)$ is the frequency response of the i^{th} output y_i from a unit initial condition in the j^{th} state x_j. Also, the ij^{th} element of $\bar{X}_{wi}(j\omega)$ is the frequency response of the i^{th} state x_{wi_i} from a unit impulse in the j^{th} input δ_j and the ij^{th} element of $\bar{Y}_{wo}(j\omega)$ is the frequency response of the i^{th} output γ_i from a unit initial condition in the j^{th} state x_{wo_j}. Further, partition $\bar{X}_G(j\omega)$, $\bar{Y}_G(j\omega)$, $\bar{X}_{wi}(j\omega)$, and $\bar{Y}_{wo}(j\omega)$ as

$$\bar{X}_G(j\omega) = \begin{bmatrix} \bar{X}_{G_1}(j\omega) \\ \vdots \\ \bar{X}_{G_n}(j\omega) \end{bmatrix} \quad \bar{X}_{wi}(j\omega) = \begin{bmatrix} \bar{X}_{wi_1}(j\omega) \\ \vdots \\ \bar{X}_{wi_p}(j\omega) \end{bmatrix} \quad (31)$$

$$\bar{Y}_G(j\omega) = [\ \bar{Y}_{G_1}(j\omega) \quad \cdots \quad \bar{Y}_{G_n}(j\omega)\ ]$$
$$\bar{Y}_{wo}(j\omega) = [\ \bar{Y}_{wo_1}(j\omega) \cdots \bar{Y}_{wo_q}(j\omega)\ ]$$

where p and q are the dynamic orders of $G_{wi}(j\omega)$ and $G_{wo}(j\omega)$, respectively.

Using the notation in Eq. (29) and the transformed Eq. (4),

$$\Sigma = \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{X}_G(j\omega)G_{wi}(j\omega)G_{wi}^*(j\omega)\bar{X}_G^*(j\omega)d\omega \quad (32)$$
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{Y}_G^*(j\omega)G_{wo}^*(j\omega)G_{wo}(j\omega)\bar{Y}_G(j\omega)d\omega$$
$$X_{21} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{X}_{wi}(j\omega)G_{wi}^*(j\omega)\bar{X}_G^*(j\omega)d\omega$$
$$Y_{12} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{Y}_G^*(j\omega)G_{wo}^*(j\omega)\bar{Y}_{wo}(j\omega)d\omega$$

Finally, from Eq. (32) and using the notation in Eq. (31),

$$\sigma_i = \frac{1}{2\pi}\int_{-\infty}^{\infty} \|\bar{X}_{G_i}(j\omega)G_{wi}(j\omega)\|_2^2 d\omega \quad (33)$$
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \|G_{wo}(j\omega)\bar{Y}_{G_i}(j\omega)\|_2^2 d\omega$$

$$|(X_{21})_{1_{ij}}| \le \frac{1}{2\pi}\int_{-\infty}^{\infty} \|\bar{X}_{wi_i}(j\omega)\|_2 \, \|\bar{X}_{G_j}(j\omega)G_{wi}(j\omega)\|_2 d\omega \quad (34)$$

$$|(Y_{12})_{1_{ij}}| \le \frac{1}{2\pi}\int_{-\infty}^{\infty} \|G_{wo}(j\omega)\bar{Y}_{G_i}(j\omega)\|_2 \, \|\bar{Y}_{wo_j}(j\omega)\|_2 d\omega \quad (35)$$

First, observe from Eq. (33) that for the smaller values of σ_i, one can expect the $\|\bar{X}_{G_j}(j\omega)G_{wi}(j\omega)\|_2$'s and $\|G_{wo}(j\omega)\bar{Y}_{G_i}(j\omega)\|_2$'s in Eqs. (34) and (35) to be small, since the integrand in Eq. (33) is always nonnegative.

Second, consider the special but common case of IB low pass and high pass weighting filters with 40 db/dec attenuation as shown in Figure 3. In this case, the $\|\bar{X}_{wi_i}(j\omega)\|_2$'s and $\|\bar{Y}_{wo_j}(j\omega)\|_2$'s in Eqs. (34) and (35) have maximum value near unity, depending upon the filter bandwidth, λ. To show this, let A_w, B_w (a 2×1 matrix), C_w (a 1×2 matrix), and D_w be the state space description of $g_w(s)$, and then

$$\bar{X}_{wi}(j\omega) = \bar{X}_w(j\omega) \cdot I \quad \text{with} \quad \bar{X}_w(j\omega) = (j\omega I - A_w)^{-1}B_w \quad (36)$$
$$\bar{Y}_{wo}(j\omega) = \bar{Y}_w(j\omega) \cdot I \quad \text{with} \quad \bar{Y}_w(j\omega) = C_w(j\omega I - A_w)^{-1}$$

where for the low pass filter

$$\bar{X}_w(j\omega) = \frac{(0.59/\lambda^{1/2})}{\{(j\omega/\lambda)+1\}^2}\begin{bmatrix}(j\omega/\lambda)+2.4\\(j\omega/\lambda)-0.41\end{bmatrix} \quad (37)$$

$$\bar{Y}_w(j\omega) = \frac{(0.59/\lambda^{1/2})}{\{(j\omega/\lambda)+1\}^2}\begin{bmatrix}(j\omega/\lambda)+2.4\\-\{(j\omega/\lambda)-0.41\}\end{bmatrix}^T$$

while for the high pass filter

$$\bar{X}_w(j\omega) = \frac{(1/\lambda^{1/2})}{\{(j\omega/\lambda)+1\}^2}\begin{bmatrix}1.4\{(j\omega/\lambda)+0.41\}\\0.25\{(j\omega/\lambda)-2.4\}\end{bmatrix} \quad (38)$$

$$\bar{Y}_w(j\omega) = \frac{(1/\lambda^{1/2})}{\{(j\omega/\lambda)+1\}^2}\begin{bmatrix}-1.4\{(j\omega/\lambda)+0.41\}\\0.25\{(j\omega/\lambda)-2.4\}\end{bmatrix}^T$$

Observe from Eqs. (37) and (38) that the elements of $\bar{X}_w(j\omega)$ and $\bar{Y}_w(j\omega)$, in a frequency response sense, consist of two real poles with equal time constants, a real zero with nearly the same time constant, and a Bode gain inversely proportional to $\lambda^{1/2}$. Therefore, in this special case, the $\|\bar{X}_{wi_i}(j\omega)\|_2$'s and $\|\bar{Y}_{wo_j}(j\omega)\|_2$'s in Eqs. (34) and (35) have maximum value near unity, depending upon the filter bandwidth, λ. Note that low pass and high pass weighting filters with attenuation rates other than 40 db/dec indicate similar results.

$G_{wi}(j\omega)$ or $G_{wo}(j\omega) = g_w(j\omega) \cdot I$
where $g_w(j\omega)$ is a scalar

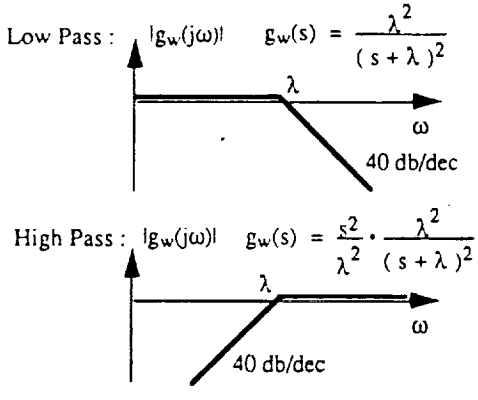


Figure 3. Low Pass And High Pass Weighting Filters

The reason for this result is explained by noting first that

$$g_w(j\omega) = \underbrace{C_w(j\omega I - A_w)^{-1}B_w}_{\bar{X}_w(j\omega)} + \underbrace{D_w}_{\bar{Y}_w(j\omega)} \quad (39)$$

Now, since the state space description of $g_w(s)$ is IB, the elements of B_w and C_w have equal magnitude or $(B_w)_i = \pm (C_w)_i$ (i.e., the situation where either B_w or C_w is large and the other is small is excluded).[13] Further, since $g_w(j\omega)$ provides unity magnitude in the weighted frequency range and high attenuation otherwise, then from Eq. (39) either of the following three situations can occur:

1. large C_w and B_w and small $(j\omega I - A_w)^{-1}$
2. intermediate C_w, B_w, and $(j\omega I - A_w)^{-1}$
3. small C_w and B_w and large $(j\omega I - A_w)^{-1}$

depending upon the magnitude of A_w, or the filter bandwidth λ.

Therefore, the $\|\bar{X}_{wi_i}(j\omega)\|_2$'s and $\|\bar{Y}_{wo_j}(j\omega)\|_2$'s in Eqs. (34) and (35) are constrained to have maximum value near unity, depending upon the filter bandwidth λ.

Although strictly a conjecture, it is felt that, using this same explanation, the $\|\bar{X}_{wi_i}(j\omega)\|_2$'s and $\|\bar{Y}_{wo_j}(j\omega)\|_2$'s in Eqs. (34) and (35) will be constrained to have maximum value near unity, depending upon the filter bandwidth, for any general weighting filter that provides near unity magnitude in the weighted frequency range and high attenuation otherwise.

Based upon the above observations, small σ_i's imply small $(X_{21})_1$ and $(Y_{12})_1$ from Eqs. (34) and (35), and hence small $M(j\omega)$ and $N(j\omega)$ from Eq. (24), at least for the special but common case of low pass or high pass weighting filters. Therefore, if a state corresponding to a small σ_i is truncated, then one is also inherently eliminating a state corresponding to a small $|\sigma_i + m(j\omega)| \cdot |\sigma_i + n(j\omega)|$, and the weighted frequency response error will be bounded according to Eq. (28).

Approximate Bound For The General Case

Unfortunately, the error bound for order reduction by one state can not be applied successively to obtain an error bound for the general case of order reduction by more than one state. This is because the reduced order model from truncation is not FWIB. This is seen from the transformed, partitioned Eq. (5), or

$$A_{X_R}X_R + X_RA^*_{X_R} + B_{X_R}B^*_{X_R} = R_X \quad (40)$$
$$A^*_{Y_R}Y_R + Y_RA_{Y_R} + C^*_{Y_R}C_{Y_R} = R_Y$$

where

$$A_{X_R} = \begin{bmatrix}A_{22} & B_2C_{wi}\\0 & A_{wi}\end{bmatrix} \qquad B_{X_R} = \begin{bmatrix}B_2D_{wi}\\B_{wi}\end{bmatrix} \quad (41)$$

$$A_{Y_R} = \begin{bmatrix}A_{22} & 0\\B_{wo}C_2 & A_{wo}\end{bmatrix} \qquad C_{Y_R} = [\,D_{wo}C_2 \;\; C_{wo}\,]$$

$$X_R = \begin{bmatrix}\Sigma_2 & (X_{12})_2\\(X_{21})_2 & X_{22}\end{bmatrix} \qquad Y_R = \begin{bmatrix}\Sigma_2 & (Y_{12})_2\\(Y_{21})_2 & Y_{22}\end{bmatrix}$$

$$R_X = \begin{bmatrix}0 & -A_{21}(X_{12})_1\\-(X_{21})_1A^*_{21} & 0\end{bmatrix} \qquad R_Y = \begin{bmatrix}0 & -A^*_{12}(Y_{12})_1\\-(Y_{21})_1A_{12} & 0\end{bmatrix}$$

For the reduced order model to be FWIB, the residual terms R_X and R_Y would have to equal zero, and they are clearly not zero, in general.

Here, one could rebalance the reduced order model and continue by eliminating one state at a time with rebalancing until the desired reduced order model is obtained; however, a different approach is considered here. Based upon the previous development in Eqs. (29) thru (39), note the residual terms R_X and R_Y will be small if a truly small σ_i is eliminated. In this case, the model resulting from order reduction by one state is nearly FWIB.

Denote $m_i(j\omega)$ and $n_i(j\omega)$ as the variables corresponding to $m(j\omega)$ and $n(j\omega)$, respectively, in Eq. (24) for successive order reductions by one state without rebalancing. Also denote $E_{w_i}(j\omega)$ as the weighted frequency response error for each successive order reduction by one state, or

$$\bar{\sigma}[E_{w_1}(j\omega)] \le 2\,\{\,|\sigma_1 + m_1(j\omega)| \cdot |\sigma_1 + n_1(j\omega)|\,\}^{1/2} \qquad (42)$$
$$\bar{\sigma}[E_{w_i}(j\omega)] \le 2\,\{\,|\sigma_i + m_i(j\omega)| \cdot |\sigma_i + n_i(j\omega)|\,\}^{1/2} \quad \text{for} \quad i = 2, \ldots, n - n_R$$

Now, the frequency response error $E_w(j\omega)$, for the general case of a reduction from n to n_R in one step, is related to the errors $E_{w_i}(j\omega)$ by

$$E_w(j\omega) = \sum_{i=1}^{n-n_R} E_{w_i}(j\omega) \qquad (43)$$

Taking the singular value of Eq. (43) and pulling the summation outside the singular value yields

$$\bar{\sigma}[E_w(j\omega)] \le \sum_{i=1}^{n-n_R} \bar{\sigma}[E_{w_i}(j\omega)] \qquad (44)$$

Finally, substitution of Eq. (42) into Eq. (44) leads to the approximate error bound for the general case, or[12]

$$\bar{\sigma}[E_w(j\omega)] \le 2 \sum_{i=1}^{n-n_R} \{\,|\sigma_i + m_i(j\omega)| \cdot |\sigma_i + n_i(j\omega)|\,\}^{1/2} \qquad (45)$$

Since the reduced order models obtained by eliminating one state at a time without rebalancing are nearly FWIB, the argument put forth in Eqs. (29) thru (39), that states corresponding to smaller values of σ_i also correspond to the smaller values of $|\sigma_i + m_i(j\omega)| \cdot |\sigma_i + n_i(j\omega)|$, is applicable here. Therefore, by eliminating the states corresponding to the smaller values of σ_i, one is also inherently eliminating the states corresponding to the smaller values of $|\sigma_i + m_i(j\omega)| \cdot |\sigma_i + n_i(j\omega)|$, and the weighted frequency response error is approximately bounded according to Eq. (45).

FWIB Residualization Error Analysis

Attention is now turned to the frequency response error analysis for FWIB residualization. Much of the analysis and notation appearing in this section parallels the frequency response error analysis given in the previous section for FWIB truncation. However, the reader is warned that the notation in this section represents features of the residualization technique, which are distinct from that of the truncation technique.

It can be shown that the weighted residualization error can be expressed as[12]

$$E_w(j\omega) = \tilde{C}_w(j\omega)\{\,\Delta^{-1}(j\omega) - \Delta'^{-1}(j\omega)\,\}\tilde{B}_w(j\omega) \qquad (46)$$

where

$$\tilde{B}_w(j\omega) = \tilde{B}(j\omega)G_{wi}(j\omega) \qquad \tilde{C}_w(j\omega) = G_{wo}(j\omega)\tilde{C}(j\omega) \qquad (47)$$
$$\tilde{B}(j\omega) = B_2 + A_{21}\Phi(j\omega)B_1 \qquad \tilde{C}(j\omega) = C_2 + C_1\Phi(j\omega)A_{12}$$
$$\Phi(j\omega) = (j\omega I - A_{11})^{-1}$$
$$\Delta(j\omega) = j\omega I - A_{22} - A_{21}\Phi(j\omega)A_{12}$$
$$\Delta'(j\omega) = -A_{22} - A_{21}\Phi(j\omega)A_{12}$$

Further, the maximum singular value of the weighted error can be expressed as

$$\bar{\sigma}^2[E_w] = \bar{\lambda}\,[\{\Delta^{-1} - \Delta'^{-1}\}\tilde{B}_w\tilde{B}_w^*\{\Delta^{-1^*} - \Delta'^{-1^*}\}\tilde{C}_w^*\tilde{C}_w] \qquad (48)$$

By expanding the products $\tilde{B}_w(j\omega)\tilde{B}_w^*(j\omega)$ and $\tilde{C}_w^*(j\omega)\tilde{C}_w(j\omega)$ and using the transformed, partitioned Eq. (5), it can be shown that[12]

$$\tilde{B}_w\tilde{B}_w^* = \Delta\{\Sigma_2 + (X_{12})_2(-j\omega I - A_{wi}^*)^{-1}C_{wi}^*\tilde{B}^*\} \qquad (49)$$
$$\{\Sigma_2 + \tilde{B}C_{wi}(j\omega I - A_{wi})^{-1}(X_{21})_2\}\Delta^*$$
$$\tilde{C}_w^*\tilde{C}_w = \Delta^*\{\Sigma_2 + (Y_{12})_2(j\omega I - A_{wo}^*)^{-1}B_{wo}^*\tilde{C}^*\}$$
$$\{\Sigma_2 + \tilde{C}B_{wo}^*(-j\omega I - A_{wo}^*)^{-1}(Y_{21})_2\}\Delta$$

where $(X_{21})_2$ and $(Y_{12})_2$ are partitions of X_{21} and Y_{12} as in Eq. (22). After substituting Eq. (49) into Eq. (48), $\bar{\sigma}[E_w(j\omega)]$ can be rewritten as[12]

$$\bar{\sigma}^2[E_w] = \bar{\lambda}\,[\{\Delta^{-1}(\Sigma_2+M)\Delta^* - \Delta'^{-1}(\Sigma_2+M)\Delta'^* + j\omega\Delta'^{-1}(M-M^*)\}$$
$$\{\Delta^{-1^*}(\Sigma_2+N^*)\Delta - \Delta'^{-1^*}(\Sigma_2+N^*)\Delta' + j\omega\Delta'^{-1^*}(N-N^*)\}] \qquad (50)$$

where

$$M(j\omega) = \tilde{B}(j\omega)C_{wi}(j\omega I - A_{wi})^{-1}(X_{21})_2 \qquad (51)$$
$$N(j\omega) = (Y_{12})_2(j\omega I - A_{wo})^{-1}B_{wo}\tilde{C}(j\omega)$$

Note that the terms $M(j\omega)$ and $N(j\omega)$ are functions of the frequency weightings.

For $n - n_R = 1$, then $\Sigma_2 = \sigma_n$, $\Delta(j\omega) = \delta(j\omega)$, $\Delta'(j\omega) = \delta'(j\omega)$, $M(j\omega) = m(j\omega)$, and $N(j\omega) = n(j\omega)$ all become scalars, and Eq. (50) yields[12]

$$\bar{\sigma}^2[E_w] = \{a + p\}\{a^* + q\}(\sigma_n + m)(\sigma_n + n^*) \qquad (52)$$

where

$$a(j\omega) = \delta^{-1}(j\omega)\delta^*(j\omega) - \delta'^{-1}(j\omega)\delta'^*(j\omega) \qquad (53)$$
$$p(j\omega) = j\omega\delta'^{-1}(j\omega)[1 - \{\sigma_n + m(j\omega)\}^{-1} \cdot \{\sigma_n + m^*(j\omega)\}]$$
$$q(j\omega) = j\omega\delta'^{-1^*}(j\omega)[\{\sigma_n + n^*(j\omega)\}^{-1} \cdot \{\sigma_n + n(j\omega)\} - 1]$$

provided $\sigma_n \neq -m(j\omega)$ and $\sigma_n \neq -n^*(j\omega)$ for all frequencies. (Note that σ_n is a real number while $m(j\omega)$ and $n(j\omega)$ are complex numbers, making it unlikely that $\sigma_n = -m(j\omega)$ and $\sigma_n = -n^*(j\omega)$.) Since the right-hand side of Eq. (52) represents a positive semidefinite matrix ($E_w(j\omega)E_w^*(j\omega)$), taking the absolute value does not alter the equality, or

$$\bar{\sigma}^2[E_w] = |a + p| \cdot |a^* + q| \cdot |\sigma_n + m| \cdot |\sigma_n + n| \qquad (54)$$
$$\le \{|a| + |p|\}\{|a| + |q|\}\,|\sigma_n + m| \cdot |\sigma_n + n|$$

Substitution of the inequality $|a(j\omega)| \le 2$ into Eq. (54) yields the error bound for order reduction by one state[12]

$$\bar{\sigma}[E_w(j\omega)] \le k(j\omega)\,\{\,|\sigma_n + m(j\omega)| \cdot |\sigma_n + n(j\omega)|\,\}^{1/2} \qquad (55)$$

where

$$k(j\omega) = \{\,(2 + |p(j\omega)|) \cdot (2 + |q(j\omega)|)\,\}^{1/2} \qquad (56)$$

Observe that the structure of this error bound is quite similar to that for IB residualization[11] and in fact reduces to the IB result when the weighting filters are selected to be unity (i.e., $m(j\omega) = n(j\omega) = 0$ and $p(j\omega) = q(j\omega) = 0$). Note however, the scalar multiplying the σ_n term is now frequency dependent.

An analysis similar to that given in Eqs. (29) thru (39) for FWIB residualization[12] shows that states corresponding to smaller values of σ_i also correspond to the smaller values for $|\sigma_i + m(j\omega)| \cdot |\sigma_i + n(j\omega)|$. Therefore, if a state corresponding to a small σ_i is residualized, then one is also inherently eliminating a state corresponding to a small $|\sigma_i + m(j\omega)| \cdot |\sigma_i + n(j\omega)|$, and the weighted frequency response error is bounded according to Eq. (55).

Analysis Of $|p(j\omega)|$ And $|q(j\omega)|$

One concern in the above argument is the values of $|p(j\omega)|$ and $|q(j\omega)|$, especially as ω tends to infinity, since these terms contain $j\omega$ in the numerator as seen from Eq. (53). A direct analytical calculation of the maximum values of $|p(j\omega)|$ and $|q(j\omega)|$ is not practical; however, note the following observations from Eq. (53). When ω tends to zero, $p(j\omega)$ and $q(j\omega)$ tend to zero, or[12]

$$\lim_{\omega \to 0} p(j\omega) = 0 \qquad \lim_{\omega \to 0} q(j\omega) = 0 \tag{57}$$

When ω tends to infinity, the limits of $p(j\omega)$ and $q(j\omega)$ from Eq. (53) are indeterminate because $j\omega$ tends to infinity and both the terms $1 - \{\sigma_n + m(j\omega)\}^{-1} \cdot \{\sigma_n + m^*(j\omega)\}$ and $\{\sigma_n + n^*(j\omega)\}^{-1} \cdot \{\sigma_n + n(j\omega)\} - 1$ tend to zero. By using l'Hopital's rule,[12]

$$\lim_{\omega \to \infty} p(j\omega) = 2 + \frac{B_2 D_{wi} D_{wi}^* B_2^*}{\sigma_n A_{22}} \qquad \lim_{\omega \to \infty} q(j\omega) = 2 + \frac{C_2^* D_{wo}^* D_{wo} C_2}{\sigma_n A_{22}} \tag{58}$$

For intermediate values of ω, a direct numerical calculation of $p(j\omega)$ and $q(j\omega)$ for specific examples reveals that $|p(j\omega)|$ and $|q(j\omega)|$ are typically no larger than the limits in Eq. (58).

Approximate Bound For The General Case

Unfortunately, the error bound for order reduction by one state again can not be applied successively to obtain an error bound for the general case of order reduction by more than one state. This is because the reduced order model from residualization is not FWIB. This is shown by multiplying the first of the transformed, partitioned Eq. (5) by Z_1 on the left and Z_1^* on the right while multiplying the second of the transformed, partitioned Eq. (5) by Z_2^* on the left and Z_2 on the right where

$$Z_1 = \begin{bmatrix} I & -A_{12}A_{22}^{-1} & 0 \\ 0 & 0 & I \end{bmatrix} \qquad Z_2 = \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & 0 \\ 0 & I \end{bmatrix} \tag{59}$$

This leads to

$$A_{X_R}X_R + X_R A_{X_R}^* + B_{X_R}B_{X_R}^* = R_X \qquad A_{Y_R}^* Y_R + Y_R A_{Y_R} + C_{Y_R}^* C_{Y_R} = R_Y \tag{60}$$

where

$$A_{X_R} = \begin{bmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{21}) & (B_1 - A_{12}A_{22}^{-1}B_2)C_{wi} \\ 0 & A_{wi} \end{bmatrix} \tag{61}$$

$$A_{Y_R} = \begin{bmatrix} (A_{22} - A_{12}A_{22}^{-1}A_{21}) & 0 \\ B_{wo}(C_1 - C_2A_{22}^{-1}A_{21}) & A_{wo} \end{bmatrix}$$

$$B_{X_R} = \begin{bmatrix} (B_1 - A_{12}A_{22}^{-1}B_2)D_{wi} \\ B_{wi} \end{bmatrix}$$

$$C_{Y_R} = \begin{bmatrix} D_{wo}(C_1 - C_2A_{22}^{-1}A_{21}) & C_{wo} \end{bmatrix}$$

$$X_R = \begin{bmatrix} \Sigma_1 & (X_{12})_1 \\ (X_{21})_1 & X_{22} \end{bmatrix} \qquad Y_R = \begin{bmatrix} \Sigma_1 & (Y_{12})_1 \\ (Y_{21})_1 & Y_{22} \end{bmatrix}$$

$$R_X = \begin{bmatrix} R_{X_{11}} & A_{12}A_{22}^{-1}(X_{12})_2 A_{wi}^* \\ A_{wi}(X_{21})_2 A_{22}^{-1^*} A_{12}^* & 0 \end{bmatrix}$$

$$R_{X_{11}} = (B_1 - A_{12}A_{22}^{-1}B_2)C_{wi}(X_{21})_2 A_{22}^{-1^*} A_{12}^* + A_{12}A_{22}^{-1}(X_{12})_2 C_{wi}^*(B_1^* - B_2^* A_{22}^{-1^*} A_{12}^*)$$

$$R_Y = \begin{bmatrix} R_{Y_{11}} & A_{21}^* A_{22}^{-1^*}(Y_{12})_2 A_{wo} \\ A_{wo}^*(Y_{21})_2 A_{22}^{-1} A_{21} & 0 \end{bmatrix}$$

$$R_{Y_{11}} = (C_1^* - A_{21}^* A_{22}^{-1^*} C_2^*)B_{wo}^*(Y_{21})_2 A_{22}^{-1} A_{21} + A_{21}^* A_{22}^{-1^*}(Y_{12})_2 B_{wo}(C_1 - C_2A_{22}^{-1}A_{21})$$

For the reduced order model to be FWIB, the residual terms R_X and R_Y would have to equal zero and they are clearly not zero, in general.

As shown for FWIB truncation, the residual terms will be small if a truly small σ_i is eliminated.[12] In this case, the model resulting from order reduction by one state is nearly FWIB. Denote $m_i(j\omega)$, $n_i(j\omega)$, and $\delta_i'(j\omega)$ as the variables corresponding to $m(j\omega)$, $n(j\omega)$ and $\delta'(j\omega)$, respectively, in Eqs. (47) and (51) for successive order reductions by one state without rebalancing, and define $p_i(j\omega)$, $q_i(j\omega)$, and $k_i(j\omega)$ as

$$p_i(j\omega) = j\omega\delta_i'^{-1}(j\omega)[1 - \{\sigma_i + m_i(j\omega)\}^{-1} \cdot \{\sigma_i + m_i^*(j\omega)\}] \tag{62}$$

$$q_i(j\omega) = j\omega\delta_i'^{-1^*}(j\omega)[\{\sigma_i + n_i^*(j\omega)\}^{-1} \cdot \{\sigma_i + n_i(j\omega)\} - 1]$$

$$k_i(j\omega) = \{ (2 + |p_i(j\omega)|) \cdot (2 + |q_i(j\omega)|) \}^{1/2}$$

Also denote $E_{w_i}(j\omega)$ as the weighted frequency response error for each successive order reduction by one state, or

$$\bar{\sigma}[E_{w_n}(j\omega)] \le k_n(j\omega) \{ |\sigma_n + m_n(j\omega)| \cdot |\sigma_n + n_n(j\omega)| \}^{1/2} \tag{63}$$
$$\bar{\sigma}[E_{w_i}(j\omega)] \le k_i(j\omega) \{ |\sigma_i + m_i(j\omega)| \cdot |\sigma_i + n_i(j\omega)| \}^{1/2} \quad \text{for} \quad i = n_R + 1, \ldots, n - 1$$

Now, the frequency response error $E_w(j\omega)$, for the general case of a reduction from n to n_R in one step, is related to the errors $E_{w_i}(j\omega)$ by

$$E_w(j\omega) = \sum_{i = n_R + 1}^{n} E_{w_i}(j\omega) \tag{64}$$

Taking the singular value of Eq. (64) and pulling the summation outside the singular value yields

$$\bar{\sigma}[E_w(j\omega)] \le \sum_{i = n_R + 1}^{n} \bar{\sigma}[E_{w_i}(j\omega)] \tag{65}$$

Finally, substitution of Eq. (63) into Eq. (65) leads to the approximate error bound for the general case, or[12]

$$\bar{\sigma}[E_w(j\omega)] \le \sum_{i = n_R + 1}^{n} k_i(j\omega) \{ |\sigma_i + m_i(j\omega)| \cdot |\sigma_i + n_i(j\omega)| \}^{1/2} \tag{66}$$

Since the reduced order models obtained by eliminating one state at a time without rebalancing are nearly FWIB, the argument that states corresponding to smaller values of σ_i also correspond to the smaller values of $|\sigma_i + m_i(j\omega)| \cdot |\sigma_i + n_i(j\omega)|$ is applicable here. Therefore, by eliminating the states corresponding to the smaller values of σ_i, one is also inherently eliminating the states corresponding to the smaller values of $|\sigma_i + m_i(j\omega)| \cdot |\sigma_i + n_i(j\omega)|$, and the weighted frequency response error is approximately bounded according to Eq. (66).

Examples

Consider the model given in the Appendix describing the stable longitudinal dynamics of a large, flexible aircraft, similar to that studied in Reference 8. The model is 12th order with phugoid, short period, and four aeroelastic modes. Control inputs consist of elevator deflection δ_E and canard deflection δ_C while responses of interest are pitch rate q' and vertical acceleration a_z', from sensors located near the cockpit.

Suppose an accurate reduced order model is desired in the frequency range above 3 rad/s. A 5th order model is obtained by FWIB truncation (from 12th to 5th order in one step) using an input weighting filter with unity magnitude above 3 rad/s and 40 db/dec attenuation below 3 rad/s. The frequency responses of the reduced order and higher order models are shown in Figures 4 and 5, indicating the 5th order model accurately reflects the dynamics of the higher order model in the weighted frequency range as desired.

To investigate the assertion that a reduced order model, obtained by the elimination of a single state corresponding to a small σ_i, is nearly FWIB, consider the following from Eq. (40).

$$\underbrace{A_{22}(X_{12})_2}_{Z_1}+\underbrace{B_2C_{wi}X_{22}}_{Z_2}+\underbrace{(X_{12})_2A^*_{wi}}_{Z_3}+\underbrace{B_2D_{wi}B^*_{wi}}_{Z_4} = \underbrace{-A_{21}(X_{12})_1}_{R_{X_{12}}} \tag{67}$$

Table 1 contains the average values of Z_1 thru Z_4 and $R_{X_{12}}$ for the reduced order models obtained by the truncation of one state at a time, based solely upon the σ_i's and without rebalancing, leading to the 5th order model. Observe that the residual term $R_{X_{12}}$ is small relative to the terms Z_1 thru Z_4, making the reduced order model essentially FWIB.

Attention is now turned to the assertion that elimination of a state corresponding to a small σ_i inherently eliminates a state corresponding to a small $|\sigma_i + m(j\omega)| \cdot \sigma_i$. (Note $n(j\omega) = 0$ for no output weighting and $|\sigma_i| = \sigma_i$.) Table 2 contains the maximum values of $|\sigma_i + m(j\omega)| \cdot \sigma_i$ for the reduced order models obtained by the truncation of one state at a time, based solely upon the σ_i's and without rebalancing, leading to the 5th order model. Observe that at each step, the state with the smallest σ_i also corresponds to the state with the smallest $\max_\omega |\sigma_i + m(j\omega)| \cdot \sigma_i$ denoted by the underline, supporting the reduction algorithm based only upon the σ_i's.

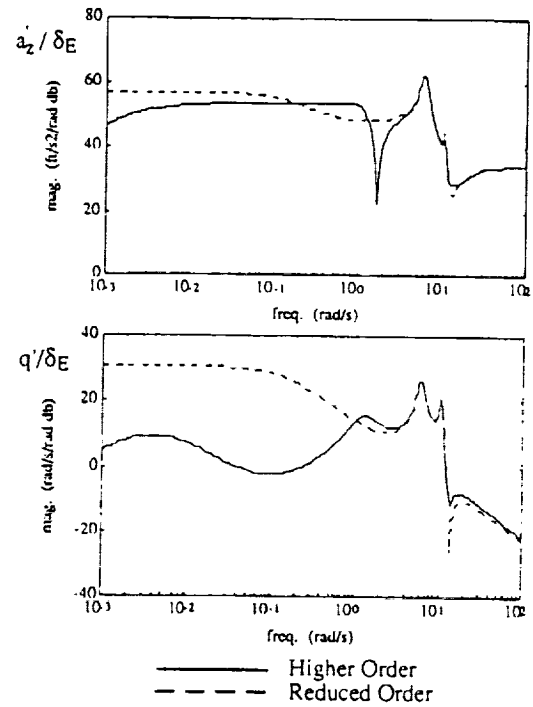


Figure 4. Frequency Responses From FWIB Truncation

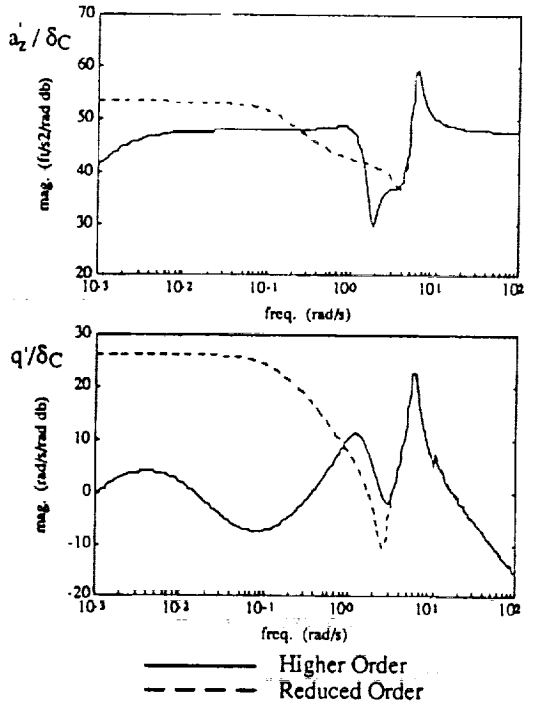


Figure 5. Frequency Responses From FWIB Truncation

Table 1. $R_{X_{12}}$ Data For FWIB Truncation

n_R	11	10	9	8	7	6	5
$\bar{R}_{X_{12}}$	0.000031	0.016	0.00086	0.015	0.13	0.034	0.42
$\bar{Z}_1$	3.8	4.2	4.7	5.2	5.6	6.1	6.3
$\bar{Z}_2$	6.2	6.9	7.6	8.4	9.2	10	12
$\bar{Z}_3$	3.9	4.3	4.8	5.3	6.0	7.0	7.8
$\bar{Z}_4$	12	13	14	16	17	19	22

$$\bar{R}_{X_{12}} = \frac{1}{n_R p}\sum_{i=1}^{n_R}\sum_{j=1}^{p}|(R_{X_{12}})_{ij}|$$

$$\bar{Z}_i = \frac{1}{n_R p}\sum_{j=1}^{n_R}\sum_{k=1}^{p}|(Z_i)_{jk}|$$

Table 2. $|\sigma_i + m(j\omega)| \cdot \sigma_i$ Data For FWIB Truncation

n_R		11	10	9	8	7	6	5
$\max_\omega \lvert\sigma_i + m(j\omega)\rvert \cdot \sigma_i$	1	0.0000017						
	2	0.0078	0.0079					
	3	0.0079	0.0079	0.0078				
	4	3.5	3.5	3.5	3.5			
	5	49	49	49	49	47		
	6	180	180	180	180	170	70	
i	7	2400	2400	2400	2400	2400	2400	2300
	8	64000	64000	64000	64000	64000	54000	74000
	9	200000	200000	200000	200000	200000	240000	130000
	10	23000	23000	23000	23000	23000	21000	21000
	11	400000	400000	400000	400000	400000	410000	400000
	12	510000	510000	510000	510000	510000	510000	510000

Now suppose a different reduced order model is required in the frequency range below 3 rad/s. A 5th order model is obtained by FWIB residualization (from 12th to 5th order in one step) using a weighting filter with unity magnitude below 0.5 rad/s and 40 db/dec attenuation above 0.5 rad/s. The frequency responses of the reduced order and higher order models are shown in Figures 6 and 7, indicating the 5th order model accurately reflects the dynamics of the higher order model in the weighted frequency range as desired.

Again, to test the assertion that a reduced order model, obtained by the elimination of a single state corresponding to a small σ_i, is nearly FWIB, consider the following from Eq. (60)

$$\underbrace{A_R\Sigma_1 + \Sigma_1 A_R^*}_{Z_1} + \underbrace{B_R C_{wi}(X_{21})_1 + (X_{12})_1 C_{wi}^* B_R^*}_{Z_2} = \underbrace{B_R C_{wi}(X_{21})_2 A_{22}^{-1^*} A_{12}^* + A_{12}A_{22}^{-1}(X_{12})_2 C_{wi}^* B_R^*}_{R_{X_{11}}} \tag{68}$$

$$\underbrace{A_R(X_{12})_1}_{Z_3} + \underbrace{B_R C_{wi} X_{22}}_{Z_4} + \underbrace{(X_{12})_1 A_{wi}^*}_{Z_5} = \underbrace{A_{12}A_{22}^{-1}(X_{12})_2 A_{wi}^*}_{R_{X_{12}}}$$

where $D_{wi} = 0$ for this weighting, and A_R and B_R are the reduced order system matrices defined by Eq. (15). Table 3 contains the average values of Z_1 thru Z_5 as well as $R_{X_{11}}$ and $R_{X_{12}}$ for the reduced order models obtained by the residualization of one state at a time, based solely upon the σ_i's and without rebalancing, leading to the 5th order model. Observe that the residual terms $R_{X_{11}}$ and $R_{X_{12}}$ are small relative to the terms Z_1 and Z_2 and Z_3, Z_4, and Z_5, respectively, making the reduced order model almost FWIB.

Furthermore, to test the assertion that elimination of a state corresponding to a small σ_i inherently eliminates a state corresponding to a small $|\sigma_i + m(j\omega)| \cdot \sigma_i$, Table 4 contains the maximum values of $|\sigma_i + m(j\omega)| \cdot \sigma_i$ for the reduced order models obtained by the residualization of one state at a time, based solely upon the σ_i's and without rebalancing, leading to the 5th order model. Observe that at each step, except for n_R = 11 and 5, the state with the smallest σ_i also corresponds to the state with the smallest $\max_\omega |\sigma_i + m(j\omega)| \cdot \sigma_i$ denoted by the underline, supporting the residualization algorithm based only upon the σ_i's. (The values of $\max_\omega |\sigma_i + m(j\omega)| \cdot \sigma_i$ for i = 11 and 12 and n_R = 11 are nearly in the correct sequence.) Finally,

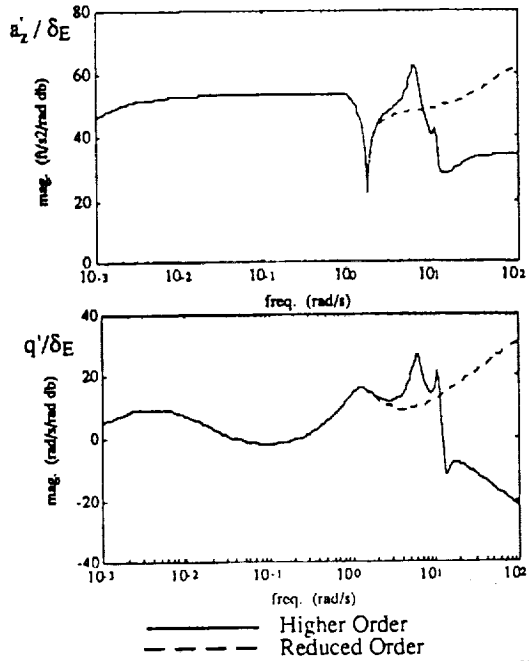


Figure 6. Frequency Responses From FWIB Residualization

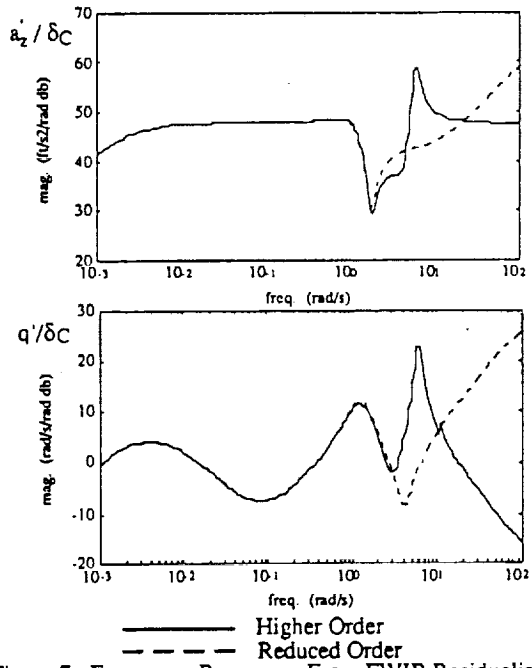


Figure 7. Frequency Responses From FWIB Residualization

Table 3. $R_{X_{11}}$ And $R_{X_{12}}$ Data For FWIB Residualization

n_R	11	10	9	8	7	6	5
$\bar{R}_{X_{11}}$	0.0032	0.00029	0.042	0.0084	0.81	0.010	30
$\bar{Z}_1$	78	87	280	110	920	130	2700
$\bar{Z}_2$	78	87	280	110	920	130	2700
$\bar{R}_{X_{12}}$	0.0012	0.0000016	0.011	0.00010	0.040	0.00035	0.92
$\bar{Z}_3$	8.6	8.4	24	8.3	51	8.4	150
$\bar{Z}_4$	8.5	8.4	24	8.2	51	8.3	150
$\bar{Z}_5$	0.22	0.24	0.27	0.30	0.34	0.39	0.41

$$\bar{R}_{X_{11}} = \frac{1}{n_R^2} \sum_{i=1}^{n_R} \sum_{j=1}^{n_R} |(R_{X_{11}})_{ij}|$$

$$\bar{R}_{X_{12}} = \frac{1}{n_R p} \sum_{i=1}^{n_R} \sum_{j=1}^{p} |(R_{X_{12}})_{ij}|$$

$$\bar{Z}_i = \frac{1}{n_R^2} \sum_{j=1}^{n_R} \sum_{k=1}^{n_R} |(Z_i)_{jk}| \quad \text{for } i=1,2$$

$$\bar{Z}_i = \frac{1}{n_R p} \sum_{j=1}^{n_R} \sum_{k=1}^{p} |(Z_i)_{jk}| \quad \text{for } i=3,4,5$$

Table 4. $|\sigma_i + m(j\omega)| \cdot \sigma_i$ And $|p_i(j\omega)|$ Data For FWIB Residualization

n_R		11	10	9	8	7	6	5
$\max_\omega \lvert\sigma_i + m(j\omega)\rvert \cdot \sigma_i$	12	0.0000035						
	11	0.0000027	0.000025					
	10	0.0033	0.0033	0.0033				
	9	0.13	0.13	0.13	0.14			
	8	1.5	1.5	1.5	1.5	1.5		
i	7	55	55	55	55	55	91	
	6	930	930	930	930	930	930	920
	5	120	120	120	120	120	120	120
	4	4200	4200	4200	4200	4200	4200	4200
	3	3100000	3100000	3100000	3100000	3100000	3100000	3100000
	2	230000	230000	230000	230000	230000	230000	230000
	1	62000	62000	62000	62000	62000	62000	62000
$\max_\omega \{ \lvert p_{n_R+1}(j\omega)\rvert \}$		2.0	1.4	2.0	1.8	1.2	1.7	2.0

Table 4 also contains the peak values of $p_i(j\omega)$, which appear in the frequency dependent factors multiplying the σ_i terms in Eq. (66). Note the peak values are approximately 2 or less, as predicted.

It is important to realize that in the above two examples, the regions of interest consist of the high and low frequency ranges, implying that in each case, the states to be eliminated should be either truncated or residualized, respectively.[8] In fact, for the first example, a 5th order model obtained from FWIB residualization does a poor job of matching the dynamics of the higher order model in the weighted frequency range, while in the second example, a 5th order model obtained from FWIB truncation does a poor job of matching the dynamics of the higher order model in the weighted frequency range.

To demonstrate that FWIB truncation and residualization may be used in coordinated manner to achieve higher accuracy than that attainable from either technique used alone, suppose an accurate reduced order model is desired in the 1 to 10 rad/s frequency range. An input weighting filter with unity magnitude in the 1 to 10 rad/s frequency range and 40 db/dec attenuation otherwise is used. One 4th order model is obtained by FWIB truncation. Another 4th order model is obtained by a combination of FWIB truncation and residualization. In this technique with the σ_i's ordered from smallest to largest as in Eq. (11), the states x_1 and x_6 are truncated and states x_2, x_3, x_4, x_5, x_7, and x_8 are residualized. By performing a modal analysis on the FWIB model, or by eliminating one state at a time and observing which mode is essentially eliminated, it can be found that states x_1 and x_6 are associated with a low frequency (phugoid) mode while states x_2, x_3, x_4, x_5, x_7, and x_8 are associated with modes at high frequency relative to the frequency range of interest. The frequency responses of the reduced order and higher order models are shown in Figures 8 and 9.

As seen in Figures 8 and 9, both 4th order models accurately reflect the dynamics of the higher order model in the weighted frequency range as desired. However, note in the q'/δ_E and q'/δ_C frequency responses that the reduced order model from FWIB truncation/residualization has improved accuracy in the 0.1 to 1 rad/s frequency range relative to the reduced order model from FWIB truncation. Note this is achieved at the expense of less accuracy in the q'/δ_E frequency response around 40 to 100 rad/s. These results demonstrate that even with FWIB truncation and residualization, the user should still be aware of the classical order reduction knowledge that truncation is most appropriate for eliminating lower frequency dynamics while residualization is most appropriate for eliminating higher frequency dynamics.[8]

Conclusions

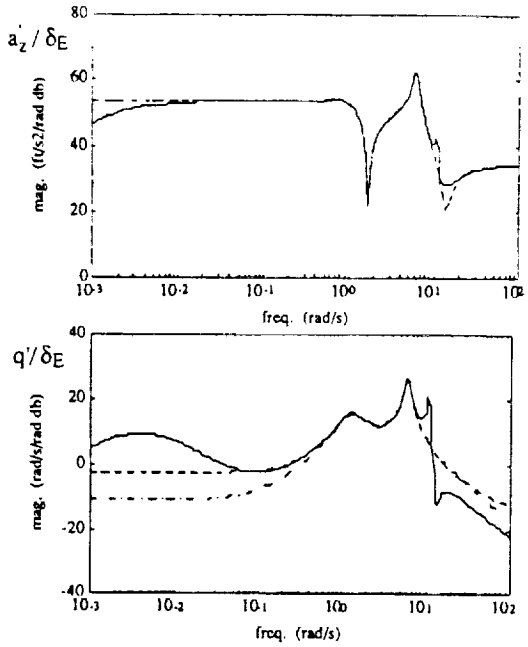
The real importance of a frequency response error analysis is not for an a priori assessment of the numerical reduction accuracy, but rather in gaining insight, offering guidance, and giving justification for the technique. The FWIB truncation error analysis presented here gives support for the truncation of coordinates based solely upon the weighted controllability-observability measures, and explains how these measures contribute to an upper bound on the frequency response error. FWIB residualization has been shown to be an equally valid order reduction technique possessing frequency response error properties similar to that for FWIB truncation. Finally, FWIB truncation and residualization can be used in a coordinated manner, consistent with classical truncation and residualization, to obtain higher accuracy than that achievable from either technique used alone.

Acknowledgements

This research was supported by NASA Langley Research Center under Grant NAG1-758. Mr. D. Arbuckle has served as the technical monitor. This support is appreciated.

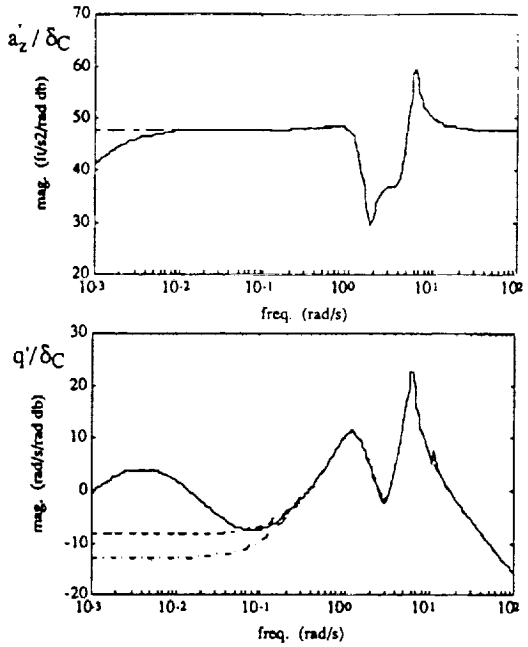
References

1. Bisplinghoff, R. L. and Ashley, H., *Principles Of Aeroelasticity*, Dover Publications, 1962



Higher Order
Reduced Order From FWIB Trun./Resid.
Reduced Order From FWIB Truncation

Figure 8. Frequency Responses From A Combination Of FWIB Truncation And Residualization



Higher Order
Reduced Order From FWIB Trun./Resid.
Reduced Order From FWIB Truncation

Figure 9. Frequency Responses From A Combination Of FWIB Truncation And Residualization

2. Likins, P. W., "Dynamics And Control Of Flexible Space Vehicles," NASA-TR-32-1329, January, 1970

3. Kwakernakk, H and Sivan, R., "*Linear Optimal Control Systems*," John Wiley & Sons, 1972

4. Doyle, J. C., et al, "State-Space Solutions To Standard H_2 And H_∞ Control Problems," IEEE Transactions On Automatic Control, Vol. AC-34, No. 8, August, 1989

5. Enns, D., "Model Reduction For Control System Design," Ph.D. Dissertation, Department Of Aeronautics And Astronautics, Stanford University, June, 1984

6. Bacon, B. J. and Schmidt, D. K., "Multivariable Frequency-Weighted Order Reduction," Journal Of Guidance, Control, And Dynamics, Vol. 12, No. 1, January-February, 1989

7. Bacon, B. J. and Schmidt, D. K., "Fundamental Approach To Equivalent Systems Analysis," Journal Of Guidance, Control, And Dynamics, Vol. 11, No. 6, November-December, 1988

8. Newman, B. and Schmidt, D. K., "Numerical And Literal Aeroelastic-Vehicle-Model Reduction For Feedback Control Synthesis," To appear in the Journal Of Guidance, Control, And Dynamics, 1991

9. Enns, D. F., "Model Reduction With Balanced Realizations: An Error Bound And A Frequency-Weighted Generalization," Proceedings Of The 23rd IEEE Conference On Decision And Control, Las Vegas, Nevada, December, 1984

10. Moore, B. C., "Principal Component Analysis In Linear Systems: Controllability, Observability, And Model Reduction," IEEE Transactions On Automatic Control, Vol. AC-26, No. 1, February, 1981

11. Liu, Y. and Anderson, B. D. O., "Singular Perturbation Approximation Of Balanced Systems," International Journal Of Control, Vol. 50, No. 4, October, 1989

12. Newman, B., Ph.D. Dissertation, School Of Aeronautics And Astronautics, Purdue University, 1991

13. Fernando, K. V. and Nicholson, H., "On The Structure Of Balanced And Other Principal Representations Of SISO Systems," IEEE Transactions On Automatic Control, Vol. AC-28, No. 2, February, 1983

Appendix

The aircraft model is

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{69}$$

where

$$y = \begin{bmatrix} q' \ (rad/s) \\ a_z' \ (ft/s^2) \end{bmatrix} \qquad u = \begin{bmatrix} \delta_E \ (rad) \\ \delta_C \ (rad) \end{bmatrix} \qquad A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{70}$$

$$A_{11} = \begin{bmatrix} -1.451e\text{-}2 & 1.935e+1 & -3.220e+1 & -1.907e+0 & 0 & 0 \\ -1.487e\text{-}4 & -4.285e\text{-}1 & 0 & 1.025e+0 & -4.223e\text{-}3 & -1.445e\text{-}4 \\ 0 & 0 & 0 & 1.e+0 & 0 & 0 \\ 1.105e\text{-}4 & -3.430e+0 & 0 & -8.335e\text{-}1 & -6.625e\text{-}2 & -3.814e\text{-}3 \\ 0 & 0 & 0 & 0 & 0 & 1.e+0 \\ 4.227e\text{-}1 & -1.072e+3 & 0 & -7.935e+1 & -3.536e+1 & -6.028e\text{-}1 \end{bmatrix}$$

$$A_{12} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 4.487e\text{-}2 & 1.755e\text{-}3 & 2.170e\text{-}3 & 1.036e\text{-}4 & -2.053e\text{-}3 & -3.391e\text{-}5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -5.160e\text{-}2 & 2.950e\text{-}2 & 8.544e\text{-}2 & 4.126e\text{-}3 & -3.777e\text{-}2 & -1.190e\text{-}3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -6.475e+0 & -1.647e\text{-}1 & 2.554e+1 & 5.435e\text{-}1 & 8.633e+0 & -1.171e\text{-}1 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 2.773e\text{-}2 & 3.552e+1 & 0 & 1.858e\text{-}1 & 5.801e+0 & 1.395e\text{-}1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -6.653e\text{-}2 & 1.475e+2 & 0 & 4.567e+0 & 2.880e+0 & -7.236e\text{-}2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 3.078e\text{-}5 & 1.014e\text{-}2 & 0 & 9.965e\text{-}5 & 6.696e\text{-}3 & -1.954e\text{-}4 \end{bmatrix}$$

$$A_{22} = \begin{bmatrix} 0 & 1.e+0 & 0 & 0 & 0 & 0 \\ -1.766e+2 & -5.054e+0 & 2.714e+0 & 6.710e\text{-}2 & 4.643e+0 & 6.646e\text{-}2 \\ 0 & 0 & 0 & 1.e+0 & 0 & 0 \\ 1.425e+1 & 2.243e+0 & -1.156e+2 & -4.246e\text{-}1 & 1.425e+0 & -1.507e\text{-}1 \\ 0 & 0 & 0 & 0 & 0 & 1.e+0 \\ -3.999e\text{-}2 & 1.440e\text{-}3 & 2.933e\text{-}3 & -1.641e\text{-}4 & -1.211e+2 & -7.226e\text{-}1 \end{bmatrix}$$

$$B = \begin{bmatrix} 1.477e+1 & 0 \\ -6.384e\text{-}2 & -1.248e\text{-}2 \\ 0 & 0 \\ -5.321e+0 & 8.392e\text{-}1 \\ 0 & 0 \\ -9.230e+2 & -6.211e+2 \\ 0 & 0 \\ -8.841e+1 & -1.107e+1 \\ 0 & 0 \\ 2.529e+2 & -4.571e+1 \\ 0 & 0 \\ 4.445e\text{-}2 & -1.521e+0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 1.501e\text{-}2 \\ 0 & -3.330e+2 \\ 0 & 0 \\ 1.e+0 & 5.034e+1 \\ 0 & -6.585e+0 \\ -2.100e\text{-}2 & 2.019e\text{-}2 \\ 0 & -1.190e+1 \\ -2.300e\text{-}2 & -1.954e+0 \\ 0 & -5.232e+0 \\ -3.200e\text{-}2 & -7.069e\text{-}2 \\ 0 & -2.894e+0 \\ -3.400e\text{-}2 & -1.672e\text{-}2 \end{bmatrix}^T$$

$$D = \begin{bmatrix} 0 & 0 \\ 5.201e+1 & -2.445e+2 \end{bmatrix}$$

Appendix C

AIAA 89-3558

ON THE CONTROL OF ELASTIC VEHICLES- MODEL SIMPLIFICATION AND STABILITY ROBUSTNESS

David K. Schmidt
Department of Mechanical And Aerospace Engineering, Arizona State University

And

Brett Newman
School of Aeronautics and Astronautics
Purdue University

AIAA Guidance, Navigation and Control Conference

August 14-16, 1989 / Boston, MA

On The Control Of Elastic Vehicles - Model Simplification And Stability Robustness

David K. Schmidt*
Department of Mechanical and Aerospace Engineering
Arizona State University
and
Brett Newman**
School of Aeronautics and Astronautics
Purdue University

Abstract

Quantitative criteria are presented for model simplification, or order reduction, such that the reduced order model may be used to synthesize and evaluate a control law, and the stability and stability robustness obtained using the reduced-order model will be preserved when controlling the full-order system. The error introduced due to model simplification is treated as modeling uncertainty, and some of the results from multivariable robustness theory are brought to bear on the model simplification problem. A numerical procedure developed previously is shown to lead to results that meet the necessary criteria. The procedure is applied to reduce the model of a flexible aircraft. Also, the importance of the control law itself, in meeting the modeling criteria, is underscored. An example is included that demonstrates that an apparently robust control law actually amplifies modest modeling errors in the critical frequency region, and leads to undesirable results. The cause of this problem is identified to be associated with the canceling of lightly-damped transmission zeroes in the plant.

---

* Professor, Associate Fellow AIAA
** Graduate Student, Student Member AIAA

Whether the engineer is developing a system model for dynamic analysis, control law synthesis, or simulation, a simple low-order model with the requisite validity is desirable for a variety of practice reasons. The question arises, therefore, as to how to obtain such a simple yet valid model. Even more fundamental is the question of what model characteristic are important such that one may strive to retain them. Although the initial question has been addressed for some time, from the attention still paid to model and controller order reduction (c.f. Refs. 1,2), it appears that the issues still remain unresolved.

In Refs 3-6, some previous offerings on the subject are presented. In this paper, discussion will continue, in the attempt to expand on some of the earlier results, to further clarify the theoretical basis behind the proposed methodology, and to reveal some important aspects of not only model-simplification, but also control-law synthesis for elastic vehicles.

1. Criteria for Modeling

The objective in model simplification, as with all system modeling, is to develop a fundamental understanding of the system in question. For the model to be useful, it should predict to the required engineering accuracy the behavior of the actual system. Note that it does not have to predict with perfect accuracy, and that is not possible anyway. The required accuracy depends on the application for which the model is intended.

In this paper, as in Refs. 3 - 6, the intended application of the model is to predict the behavior of the system when it is subject to feedback action, as shown, for example, in Fig. 1. Clearly, then, the critical characteristics of the actual system that must be adequately captured by the model are those characteristics important in a feedback system. (Note that the feedback action could represent an automatic control system, as well as that of a human, or manual controller.) Finally, the existence of a sufficiently valid, although perhaps complex model for the system is assumed to be available - admittedly a big assumption. Further, if this model is infinite-dimensional and/or non-linear, it is assumed that a locally linearized, finite-dimensional model may be obtained. The original (complex) model will be denoted as $\underline{G}$, while the linear model will be denoted as G.

As a result of any simplification process, differences between the more-accurate model and the simple model arise. Or conceptually, if G_R is a simpler model for G, the model-simplification error may be considered to be $\Delta G = G - G_R$. These errors are key to the research presented here. In contrast, model-simplification errors arising due to the development of G, or $\Delta\underline{G} = \underline{G} - G$, will be considered only indirectly.

The critical question then is what errors ΔG are critical, or should be minimized, and what procedure will do so? The answer to the first part of the question could be that ΔG's critical in a feedback loop should be minimized. Further, if these ΔG's are interpreted more generally as model uncertainty rather that model-reduction error, the recent research on multi-variable robustness theory may be bought to bear on the model-simplification problem. This is the main idea in this research.

2. Robustness and Model Reduction

In this section, some key results from robust control theory will be noted, and they will be interpreted in the context of the model reduction problem.

With reference to the system shown in Fig. 2, G_R is the transfer-function-matrix representation of this simplified model, $\Delta G(s)$ in the analogous representation of the model-simplification error, and the full-order linear model is $G = G_R + \Delta G$. Likewise, K(s) is the matrix of control compensators, perhaps to be designed using G_R. Clearly in this context, one desires that the K(s) so obtained will control the "true" G(s) as predicted through the use of G_R. Attention is now turned to exposing the critical ΔG's via multi-variable Nyquist theory.[7]

Let $\Phi(s)$ be an analytic function of the complex variable s, and let the number of zeros of $\Phi(s)$ in the open right half of the complex plane be denoted as z. Then the Principle of the Argument states that

$$\underset{R\to\infty}{N} (O, \Phi(s), D_R) = z$$

or the number (N) of clockwise encirclements of the origin made by the image of the contour D_R, under the mapping of $\Phi(s)$, as s travels clockwise around D_R , equals z. Here D_R is the "Nyquist D contour" that encloses the entire right-half of the complex plane. Clearly, with regards to stability, the $\Phi(s)$ of interest is the closed-loop characteristic polynomial of the feedback system, denoted by $\Phi_{CL}(s)$.

Now, as shown in Ref. 8, and elsewhere, and referring to Fig. 1, for example,

$$\begin{aligned} \Phi_{CL}(s) &= \Phi_{OL}(s) \det [I + GK] \\ &= \Phi_{OL}(s) \det [I + KG] \end{aligned} \qquad (1)$$

where $\Phi_{OL}(s)$ is the characteristic polynomial of the open-loop system KG(s) or GK(s). That is, if either the transfer function matrix GK(s) or KG(s) has the state-space realization

$$\begin{aligned} \dot{x} &= A_{GK}x + B_{GK}e \\ y &= C_{GK}x \end{aligned}$$

then $\Phi_{OL}(s) = \det[sI - A_{GK}]$, and the zeros of $\Phi_{OL}(s)$ are the open-loop poles of the system. Note that Eqn. 1 may therefore be re-written as

$$\Phi_{OL}(s) = \det[sI - A_{GK}] \det [I + C_{GK} [sI - A_{GK}]^{-1} B_{GK}]$$

Now if the number of right-half-plane zeros of $\Phi_{OL}(s)$ is p, then the number of right-half-plane zeros of $\det [I + C_{GK} [sI - A_{GK}]^{-1} B_{GK}]$ must be z-p. Furthermore, from the Principle of the Argument

$$\underset{R\to\infty}{N}(O, \det[I + C_{GK}(sI - A_{GK})^{-1} B_{GK}], D_R) = z-p$$

Consequently, if p is known, z may be deduced from

$$\begin{aligned} z &= p + (z - p) \\ &= p + [\underset{R\to\infty}{N}(O, \det[I + C_{GK}(sI - A_{GK})^{-1} B_{GK}], D_R)] \end{aligned}$$

or closed-loop stability is determined from knowledge of p and the examination of the Nyquist contour for det[I + GK] or det [I + KG]. Therefore, the closed-loop system is stable if and only if the Nyquist contour for det[I + GK](= det[I +KG]) encircles the origin counterclockwise exactly p times.

Of course the determination of z is possible from other means, and the real utility of the above fact is in defining the concept of relative stability, and in identifying factors that are critical to closed-loop system stability. These issues are of special import here.

Consider the model error, or uncertainty, to be ΔG (as in Fig. 2), and assume that K is such that KG_R leads to a stable closed-loop system with good stability margins. (Note this assumption should always be true as it involves a key objective in determining K(s) using G_R to begin with.) Then if (assumption 1) the number of right-half-plane poles of KG (= p) is identical to the number of right-half-plane poles of KG_R ($=p_R$), K will stabilize G if and only if (assumption 2)

$$\underset{R\to\infty}{N}(O, \det[I + GK], D_R) = \underset{R\to\infty}{N}(O, \det[I + G_RK], D_R)$$

or the number of encirclements of the origin made by the Nyquist contours associated with G and with G_R are identical.

Stability is guaranteed as follows:
Let z = no. of unstable closed-loop poles of the KG loop.
z_R = no. of unstable closed-loop poles of the KG_R loop
p, p_R - defined above

Then to show stability (or z = 0), note that if (assumption 1) $p_R = p$, then

$$z = z_R - (z_R - p_R) + (z - p)$$

By the assumption KG_R leads to a stable system, $z_R = 0$, and from assumption 2, $(z_R - p_R) = (z-p)$. Hence, z = 0.

This now establishes in a meaningful way, qualitative criteria for model simplification, the simplification must at least lead to ΔG's such that assumption 1 and 2 are satisfied. But the criteria goes further. Not only must stability of the KG loop be assured (i.e., z = 0) but the margins "designed" into KG_R should carry over to the closed-loop system associated with KG. Otherwise, the K so designed would not be satisfactory. It is for this reason that any model reduction technique that just assures stability of the full-order closed-loop system may not be good enough!

To satisfy assumption 2, or to assure that the number of encirclements of the origin is unchanged due to ΔG, requires that [9]

$$\det [\, I + G_R K + \varepsilon \Delta G K] \neq 0 \quad \forall\, \omega > 0,\ \varepsilon \in [0,1] \tag{2}$$

In other words, if as the Nyquist contour for $\det[I + G_R K]$ is continually warped to that for $\det [I + GK]$ the origin is never intersected, the number of encirclements of the origin cannot change. Furthermore, Eqn. 2 is assured if (c.f., Ref. 9)

$$\bar{\sigma} (\Delta G K) < \underline{\sigma} [\, I + G_R K] \qquad \forall\, \omega > 0 \tag{3}$$

Finally, it is known that an alternative to Eqn. 3 is

$$\bar{\sigma} (E_m) < \underline{\sigma} [I + (G_R K)^{-1}] = \underline{\sigma} \{[G_R K (I + G_R K)^{-1}]^{-1}\} \qquad \forall\, \omega > 0 \tag{4}$$

where $E_m = G_R^{-1} \Delta G$

The above expressions (Eqns. 2 - 4) may be extended by breaking the frequency domain $(0 \leq \omega < \infty)$ into the domains $(0 \leq \omega \leq \omega^*)$ and $(\omega^* < \omega < \infty)$. Note that these domains are non-intersecting. Now it can be argued that Eqn. 2 will be satisfied if

$$\det [\, I + G_R K + \varepsilon \Delta G K] \neq 0 \qquad (0 \leq \omega \leq \omega^*) \quad (0 \leq \varepsilon \leq 1) \tag{5}$$

and

$$\det [\, I + G_R K + \varepsilon \Delta G K] \neq 0 \qquad (\omega^* < \omega < \infty) \quad (0 \leq \varepsilon \leq 1) \tag{6}$$

Further, Eqn. 5 is assured if Eqn. 3 is satisfied for $\omega \leq \omega^*$, while satisfying Eqn. 4 for $\omega > \omega^*$ assures that Eqn. 6 is satisfied. Hence, in such a situation, Eqn. 2 is satisfied.

By Eqns. 3 and 4, quantitative criteria on critical ΔG's are established. Further, the overall strategy for model simplification becomes apparent, and the interaction between model simplification and control law synthesis is underscored. Regarding the later, it should be clear that the allowable ΔG's (those that do not destroy closed-loop stability of the full-order system controlled by K(s)) depend on K itself. In other words, designing a "good" K(s) increases that allowable ΔG, while designing a bad one may put very strict limitations on the allowable ΔG, and hence model accuracy. The former K(s) is robust, the latter is not.

Regarding the model simplification strategy, then, first observe the right side of Eqn. 3. When $\underline{\sigma} (G_R K) >> 1$, $\underline{\sigma} [I + G_R K] \approx \underline{\sigma} (G_R K)$. Conversely, when $\bar{\sigma}(G_R K) << 1$, $\underline{\sigma} [I + G_R K] \approx 1$. Finally, the $\underline{\sigma} [I + G_R K]$ will take on its minimum value in the frequency range where $\sigma_i (G_R K) \approx 1$. The frequency range where the latter occurs is of course the (multi-variable) gain crossover region. Consequently, it is this frequency range where the ΔG must be the smallest, and this can be assured if each element of the ΔG matrix is small in this frequency range.

Also, noting the above discussion, Eqn. 3 may be satisfied by rather large ΔG in any frequency range where $\underline{\sigma}\,[I + G_R K]$ is large, and this will occur when $\underline{\sigma}\,(G_R K)$ is large. If K is designed to give a good classical Bode loop shape, $\underline{\sigma}\,(G_R K)$ will be large for frequencies below crossover.[9]

Now consider Eqn. 4. When $\bar{\sigma}\,(G_R K) << 1$, $\underline{\sigma}\,(G_R K)^{-1} >> 1$, and $\underline{\sigma}\,\{I + (G_R K)^{-1}] \approx \underline{\sigma}\,(G_R K)^{-1} >> 1$. Hence the allowable ΔG may also be rather large in this case. Further, if K yields a good loop shape, or is well attenuated, at high frequencies, $\underline{\sigma}\,(G_R K)$ will be small for frequencies above crossover. So clearly, the ΔG must be smallest in the region of multi-variable crossover, while if K yields a good bode loop shape, rather large ΔG elsewhere may be acceptable and Eqns. 3 and 4 may be satisfied. The above discussion is summarized in Fig. 3.

The final issue to be addressed is that of satisfying assumption 1, or the number of unstable poles of KG_R must be identical to the number of unstable poles of KG. First note that this is equivalent to requiring the number of unstable poles of G and G_R to be the same, since only one K is involved. Then observe that the poles of G are the poles of $G_R + \Delta G$, which consists of the poles of G_R plus the poles of ΔG. Hence to satisfy assumption 1, ΔG must be stable.

Attention will now turn to some additional criteria arising from performance considerations rather that from stability robustness. The system to be considered is that shown in Fig. 4. The vector of responses Y(s) is given by

$$Y = [I + (G_1 + \Delta G_1)K]^{-1} (G_1 + \Delta G_1)K(Y_c - N) + [I + (G_1 + \Delta G_1)K]^{-1} (G_2 + \Delta G_2) D$$

Here G_1 is the reduced-order model for the response of G to control inputs, where G_2 is the reduced-order model for the response of G to disturbances being considered. ΔG_1 and ΔG_2 are the analogous model-simplification errors.

The first observation to be made is that stability and stability robustness depends on G_1 and ΔG_1, not on G_2 and ΔG_2. Note that the poles of $(G_1 + \Delta G_1)$ are the poles of the "true" plant G, as are the poles of $G_2 + \Delta G_2$. Hence if K stabilizes G, which will be assured if G_1 and ΔG_1 satisfy the criteria developed previously, K must therefore stabilize $(G_2 + \Delta G_2)$. This is significant since some (stable) poles of G may be approximately cancelled by some zeroes for the transfer functions governing responses to control inputs, but not cancelled in those governing responses to disturbances. Cancelling these poles to obtain G_1 has raised questions by some as to whether those poles so cancelled could lead to problems later in analysis. The answer appears to be that they will not if G_2 is obtained such that those poles are retained. But from the above discussion on stability, the only reason to keep these poles in G_2 (that by assumption are not approximately cancelled) is such that the disturbance-rejection performance predicted using G_2 (when designing K, for example) will be reasonably accurate.

Finally, noting that the disturbance response due to ΔG_2 is

$$Y_{D_2} = [I + (G_1 + \Delta G_1)K]^{-1} \Delta G_2 D$$

for good performance prediction (Y_{D_2} small), ΔG_2 should tend to be small whenever D is large and $(G_1 + \Delta G_1)K$ is small. But here again, if K is designed to obtain a "good loop," it will be designed such that G_1K (and by implication $(G_1 + \Delta G_1)$ K) will be large over the frequency range where D is large. Consequently, this should not pose stringent requirements on ΔG_2.

In ending this section, it is worth noting that assuming K is designed properly has been critical. By doing so, one takes advantage of one of the basic advantages of a good feedback system, reduction in sensitivity to plant (or plant model) variations. This allows the development of a modeling procedure that focuses on the really critical problem of obtaining a good model in the crossover region.

3. Methodology and Sample Results

The procedure offered was discussed in detail in Ref. 5, and the computational technique is summarized again in Table 1. The technique is a frequency weighted internally-balanced approach, with stable factorization in the case of an unstable plant G. The stable factorization procedure sets the unstable subsystem of G aside via partial fraction expansion, leaving the remaining subsystem G_s stable. This stable subsystem is then reduced, such that a stable reduced order model G_{R_s} is guaranteed. The unstable subsystem is then rejoined with G_{R_s} to obtain the final reduced-order model G_R. By this procedure, the number of unstable poles of G are preserved. In fact the unstable poles in G are exactly retained in G_R.

The internally balanced technique [10] requires the frequency-weighting extension[5] since the basic technique leads to small model-simplification errors ΔG where the elements of G have large magnitude, which is not necessarily the crossover region. Further, a very poor model may be obtained where the elements of G have small magnitude. As will be shown later, this can be totally unacceptable.

In Ref. 11, a frequency-weighted approach was also suggested, but the weighting required the knowledge of the compensator K, obtained using the full-order plant. Since designing a simple K using the simpler plant G_R is the typical design objective, the above weighting is undesirable. In Ref. 5, it was noted that simply adding a weighting filter obtainable by inspection of the Bode plots of G and knowledge of the desired crossover frequency range let to excellent results. This filter is easily discarded after G_R is retained. In the example presented later, it will be shown that this approach again appears quite acceptable.

The key to the concept is the knowledge of the fact that the internally balanced approach yields a small ΔG where the elements of G have large magnitude. Heuristically, if a filter W(s) is used such that W(s)G(s) has large magnitude in the required frequency range, and if WG reduced such that WG_R is obtained, then G_R will have the desired properties.

As the example, consider an elastic aircraft identical to the configuration investigated in Refs. 3 and 6. This configuration is of reasonably conventional geometry with a low-aspect ratio swept wing, conventional tail, and canard. A numerical model for the longitudinal dynamics is

Table 1 Frequency Weighted Internally Balanced Reduction

Given: System state space description A, B, C and weighting filter state space description A_w, B_w, C_w.

Find: r^{th} order system

Step 1: Solve for X and Y

$$\begin{bmatrix} A & BC_w \\ 0 & A_w \end{bmatrix} \begin{bmatrix} X & X_{12} \\ X_{21} & X_{22} \end{bmatrix} + \begin{bmatrix} X & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} A^T & 0 \\ C_w^T B^T & A_w^T \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & B_w B_w^T \end{bmatrix} = 0$$

$$\begin{bmatrix} A^T & 0 \\ C_w^T B^T & A_w^T \end{bmatrix} \begin{bmatrix} Y & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} + \begin{bmatrix} Y & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} A & BC_w \\ 0 & A_w \end{bmatrix} + \begin{bmatrix} CC^T & 0 \\ 0 & 0 \end{bmatrix} = 0$$

Step 2: Find T and Σ where $XY = T\Sigma^2 T^{-1}$, $T = [T_r\ T_{n-r}]$, $T^{-T} = [U_r, U_{n-r}]$

$$\Sigma^2 = \begin{bmatrix} \Sigma_r^2 & 0 \\ 0 & \Sigma_{n-r}^2 \end{bmatrix} \quad \text{where}$$

$$\Sigma_r = \text{diag}(v_{c_i} v_{o_i}) \quad i = 1, \ldots, r$$

$$\Sigma_{n-r} = \text{diag}(v_{c_i} v_{o_i}) \quad i = r+1, \cdots, n$$

$$v_{c_1} v_{o_1} \geq \cdots \geq v_{c_n} v_{o_n} \geq 0$$

Step 3 : r^{th} order system is

$$A_r = U_r^T A T_r$$

$$B_r = U_r^T B$$

$$C_r = C T_r$$

available from the above references. Both rigid-body modes and four elastic modes (resulting in a 11^{th} order model) are included. The in-vacuo vibration frequencies are 6.3, 7.0, 10.6, and 11.0 rad/s, and are representative for a supersonic/hypersonic cruise vehicle. These frequencies, furthermore, are all near the anticipated frequencies at crossover for the control systems to be designed.

Control inputs are elevator deflection δ_E and canard deflection δ_C, while the disturbance is the perturbation in angle of attack due to atmospheric turbulence α_g. Selected responses are vertical acceleration a_z' measured at the cockpit and pitch rate q measured at the antinode of the first bending mode. Therefore, the flight and structural mode control loops in the context of Figure 4, might correspond to the following, for example

$$Y = [a_z' \; q]^T$$

$$U = [\delta_E \; \delta_C]^T$$

$$D = \alpha_g$$

Obtaining the reduced order model G_1 was the subject of Ref. 6. An anticipated crossover frequency range (for G_1K) was assumed as 1 to 10 rad/s. In that reference, it was also noted that a fourth-order for G_1 was sought based on the observation that the full order model has two oscillatory models in this frequency range.

Attention is now turned to the requirements for G_2. As a realistic example, the Dryden gust spectrum for turbulence is used to describe the disturbance. A fourth-order model for G_2 is sought based on the observation that the full order model has two oscillatory models in the frequency range where the spectrum of D is largest. This frequency range is coincidentally also 1 to 10 rad/s.

The reduced order models for G_1 and G_2 were then obtained simultaneously from the frequency-weighted internally-balanced reduction technique[5] which was specifically developed to meet the criteria in Section 3. The frequency-weighting filter used was a band pass filter of unity magnitude in the 1 to 10 rad/s frequency range with 40 db/dec roll off on either side of this frequency range.

Table 2 contains the reduced order state space matrices A, B, C and D. Figures 5 through 10 show the reduced order and full order frequency response magnitudes for G_1 and G_2. Observe that the reduced order model accuracy approximates the full order model in the 1 to 10 rad/s frequency range as desired. To complete this example, a simple control law, consisting of three constant gains was synthesized using the model G_1. The synthesis objective was to augment the damping of the first aeroelastic mode with acceleration feedback to the canard, to augment the short period damping with pitch-rate feedback to the elevator, and to provide some response decoupling with a cross feed from the elevator to the canard. The resulting control law is of the following form

$$\begin{bmatrix} \delta_C \\ \delta_E \end{bmatrix} = \begin{bmatrix} K_1 & K_2K_3 \\ 0 & K_3 \end{bmatrix} \begin{bmatrix} a_z' \\ q \end{bmatrix} + \begin{bmatrix} 1 & K_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \delta_{C_{com}} \\ \delta_{E_{com}} \end{bmatrix}$$

Actuation effects were modeled with simple first-order lags, with corner frequencies at 20 r/s for both the canard and the elevator.

Table 2 Reduced Order Model

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] =$$

$$\left[\begin{array}{cccc|cc|c} -.9932 & .8294 & -.0138 & -.0507 & -31.67 & 14.48 & 13.59 \\ -2.013 & -.0137 & .0121 & .0329 & 35.92 & -21.42 & -24.38 \\ -5.593 & -.6638 & -.3175 & -9.658 & -593.7 & -420.0 & 700.1 \\ 4.934 & .2098 & 3.739 & -.5171 & -281.4 & -175.2 & 342.5 \\ \hline .0665 & -.03471 & .0017 & .0015 & 0 & 0 & 0 \\ 8.762 & .7218 & .9287 & -2.038 & 52.01 & -244.5 & 333.0 \end{array}\right]$$

$$y = \begin{bmatrix} q \ (r/s) \\ az' \ (ft/s) \end{bmatrix} \qquad u = \begin{bmatrix} \delta_E \ (rad) \\ \delta_C \ (rad) \end{bmatrix} \qquad D = \alpha_g \ (rad)$$

Shown in Fig. 11 is the plot of Eqn. 3, while Eqn. 4 is shown in Fig. 12. Note that although this control law did not result in high gain (large G, K) at low frequencies, Eqn. 3 was still satisfied below crossover region. Conversely, Eqn. 4 is satisfied, although barely, in the frequency range above crossover. Hence, from the argument in Section 2, if ω^* in Eqns. 5 and 6 is in the crossover region, stability is assured. For reference, the pitch-rate to elevator transfer function is

$$\frac{q(s)}{\delta_{E_C}(s)} = \frac{50\ (0.33)[.13,4.84][.01,10.6][.03,11.0][.21,13.](45.)}{[.53,1.81][.15,4.78][.02,10.8][.03,11.][.19,13.3](19.)(69.)}$$

4. An Additional Criteria

As noted in Section 3, the ΔG arising from the model simplification must satisfy stringent criteria in the crossover region, and if Eqn. 3 and/or 4 (or 5 and 6) is satisfied, closed-loop stability is assured. To be discussed here is the fact that the controller K should not be such that small ΔG is amplified such that $\bar{\sigma}$ (ΔGK) becomes large. It will be shown by example that this can easily occur where the magnitudes of G (or of the g_{ij}'s) are small. Hence, the example will demonstrate why obtaining a good model in this situation is important (recall that unweighted balanced reduction has a problem here), and some implications regarding control-law synthesis will also arise.

Consider the simple scalar plant

$$g(s) = \frac{(s^2 + .04s + 1.^2)}{s(s^2 + .032s + 0.8^2)}$$

The plant is stable and minimum phase, so a robust control law should be obtainable. Using LQG/LTR or H_∞, for example, the following compensator could be obtained.

$$k(s) = \frac{8(s^2 + .032s + 0.8^2)}{(s^2 + .04s + 1.^2)(s + 8)}$$

It can be easily verified that the loop shape kg is very good, yielding infinite gain margin, 90 degree phase margin, and good roll off above 8 rad/s.

Now assume that the "true" plant is

$$g_{true} = \frac{0.69\,(s^2 + .048s + 1.2^2)}{s(s^2 + .032s + 0.8^2)}$$

or the numerator "frequency" is in error by 20% (1.0 → 1.2). Note that this could occur, for example, if a vibration mode shape was slightly off in the modeling. Shown in Fig. 13 is the plot of Eqn. 3 for this example, and clearly $\bar{\sigma}(\Delta gk) > \underline{\sigma}(1 + gk)$ at 1 r/s (the designed crossover frequency). Further, a quick check would show the kg_{true} loop to be unstable. But the $|\Delta g| = |g - g_{true}|$ (not shown) would be found to be rather modest at ω = 1r/s, with much larger $|\Delta g|$ at lower frequencies. The problem could be interpreted as one of the control law k amplifying the $|\Delta g|$ at ω = 1 r/s, and this is confirmed from the plot of $|k(j\omega)|$ in Fig. 14.

Stability of the kg_{true} loop would result, and Eqn. 3 satisfied, if the $|k(j\omega)|$ at ω = 1 rad/sec were simply reduced. This is accomplished with the following compensator

$$k_{mod}(s) = \frac{8(s^2 + .032s + 0.8^2)}{(s^2 + 1.2s + 1.^2)(s + 8)}$$

or the damping of the complex compensator poles is increased, and the plant model zeroes close to the imaginary axis are not exactly cancelled. Clearly the loop shape with this compensator is not as "optimal" as the original, but this control law is more robust against this Δg.

Noting that the problem arose with a modeling error that is associated with lightly-damped zeroes, the critical Δg was at a frequency (ω = 1 r/s), where $|g(j\omega)|$ was relatively small as shown in Fig. 15. Hence, obtaining a good model at this frequency is important. Furthermore, by attempting to cancel those lightly-damped zeroes in the plant, the original controller was very sensitive to their location. Increasing the damping of the compensator poles, as in a classical notch filter, made the loop more robust against the uncertainty in the location of these plant zeroes. (Incidentally, this can be accomplished with a modified LTR procedure, as noted in Ref. 12 and in another paper in preparation.)

As a final remark, it is observed that lightly-damped zeroes in the compensator are different from similar zeroes in the plant since through the design and implementation of the compensator, the location of its zeroes may be more accurately defined.

5. Conclusions

Quantitative criteria are presented for model (or controller) simplification. The reduced order model (or controller) must well approximate the full-order system in the (multivariable) crossover region for stability, and stability robustness, to be assured. Bounds on the model-simplification error were noted, and if the bounds are satisfied, stability is assured. It was also

noted that the model reduction criteria were functions of the control law, and by synthesizing a robust control law, the criteria could be easier to satisfy.

A numerical procedure, consisting of stable factorization with weighted balancing of coordinates has been shown, by example, to meet the above criteria. The example involved reducing an eleventh order linear model of an elastic aircraft to obtain a fourth-order model leading to the desired six transfer functions.

Finally, another example demonstrated the importance of obtaining good agreement between the full- and reduced-order model in the crossover region, even where the transfer function (or functions) have relatively small magnitude. Furthermore, the example demonstrated that an apparently robust controller could in fact amplify small errors, and lead to unstable results. The problem would occur with any control law that had the effect of cancelling lightly-damped transmission zeroes of the plant model.

6. Acknowledgement

This research was supported by NASA Langley Research Center under Grant NAG1-758. Mr. Douglas A. Arbuckle has served as the technical monitor. This support is appreciated.

7. References

1. Liu, Y. and Anderson, B.D.O., "Controller Reduction Via Stable Factorization and Balancing," Int. Jrn. of Auto. Cont., Vol. 44, No. 2, 1986.

2. Anderson, B.D.O., "Controller Reduction," Plenary Session, 1987 ACC, Minneapolis.

3. Waszak, M. R. and Schmidt, D. K., "Flight Dynamics of Aeroelastic Vehicles," Journal of Aircraft, Vol. 25, No. 6, June, 1988.

4. Bacon, B. J. and Schmidt, D. K., " A fundamental Approach to Equivalent Systems Analysis," Journal of Guidance, Control, and Dynamics, Vol. 11, No. 6, Nov. - Dec., 1988.

5. Bacon, B. J. and Schmidt, D. K., "Multivariable Frequency-Weighted Order Reduction," Journal of Guidance, Control, and Dynamics, Vol 12, No. 1, Jan. - Feb., 1989.

6. Schmidt, D. K. and Newman, B., "Modeling, Model Simplification, and Stability Robustness With Aeroelastic Vehicles," 1988 AIAA Guidance And Control Conference, Minneapolis.

7. Stein, G., Lecture Notes on Control System Design, MIT, 1979.

8. Kwakernaak, H. and Sivan, R., Linear Optimal Control Systems, New York: Wiley-Interscience, 1972.

9. Doyle, J. C. and Stein, G., "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," IEEE Transactions Automatic Control, Vol. AC-26, Feb., 1981.

10. Moore, B.C., "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," IEEE Transactions On Automatic Control, Vol. AC-26, February 1981.

11. Enns, D.F., "Model Reduction for Control System Design," Ph.D. dissertation, Department of Aeronautics and Astronautics, Stanford University, June, 1984.

12. Schmidt, D.K., "Research on Control System Design to Meet Aircraft Handling Qualities Requirements," Final Report for MCAIR, July, 1989.

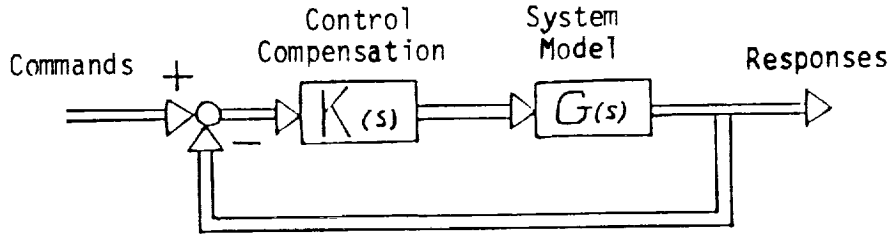


Figure 1

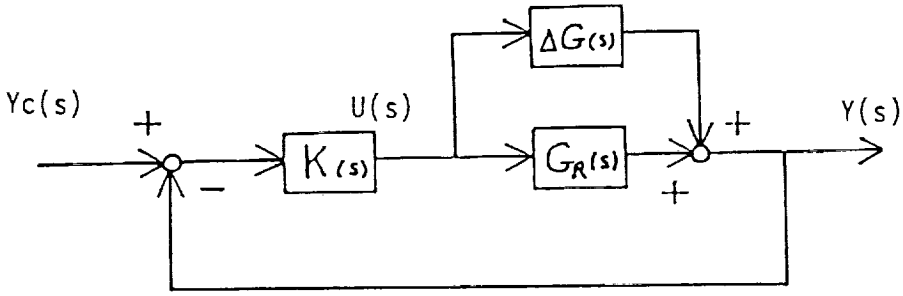


Figure 2

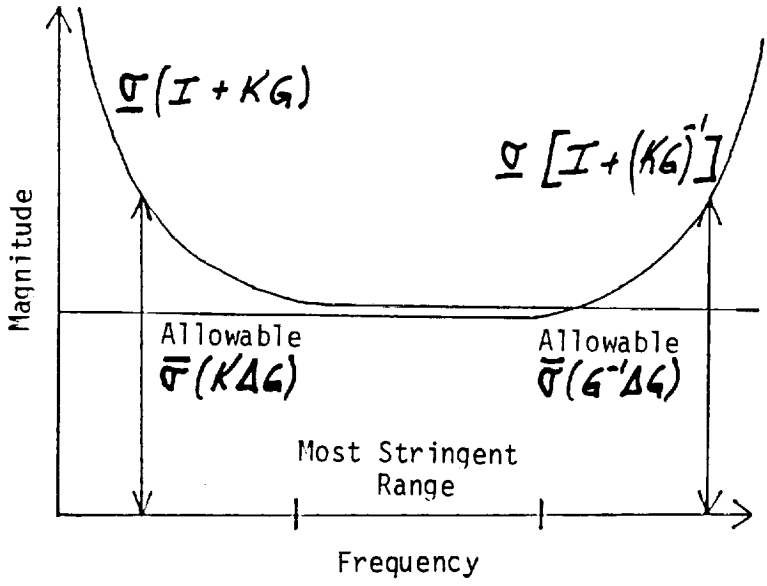


Figure 3

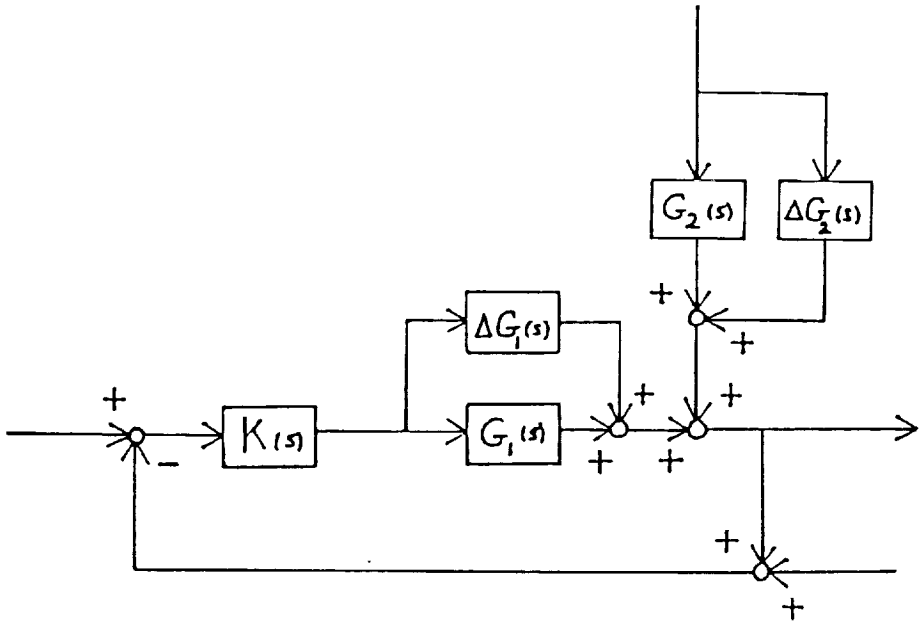


Figure 4

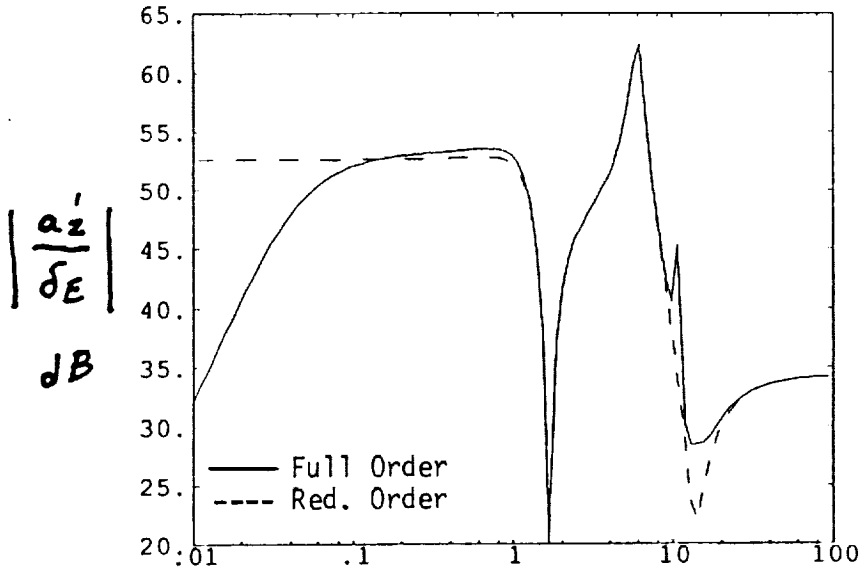


Figure 5

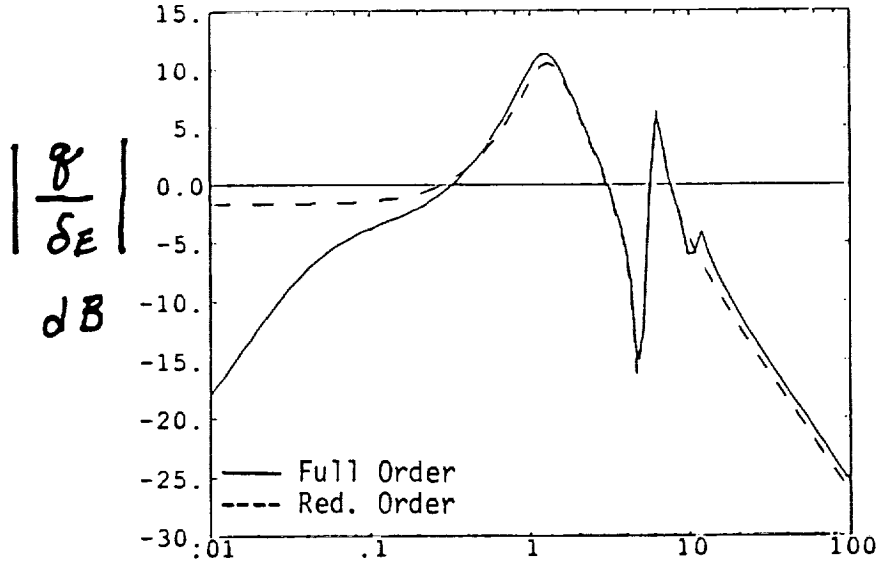
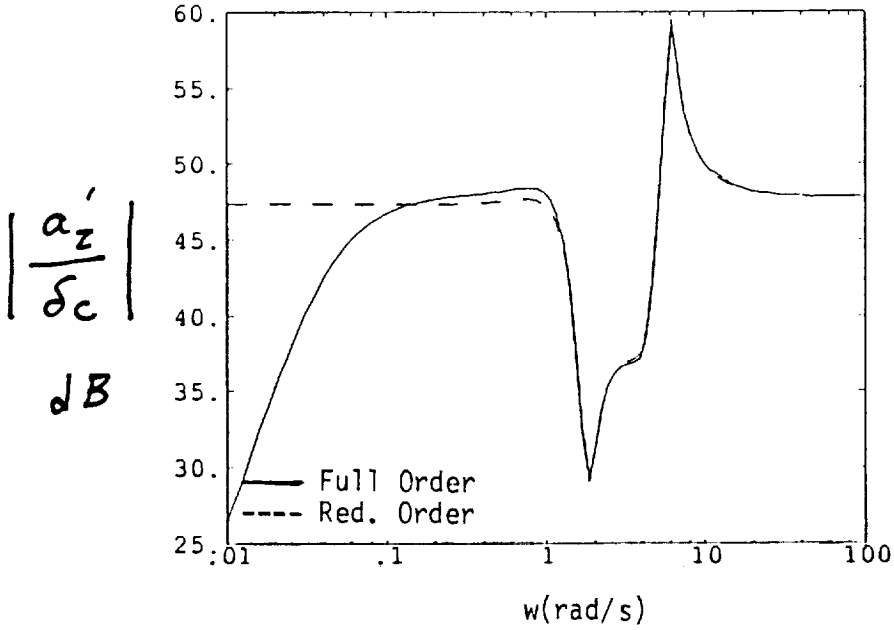
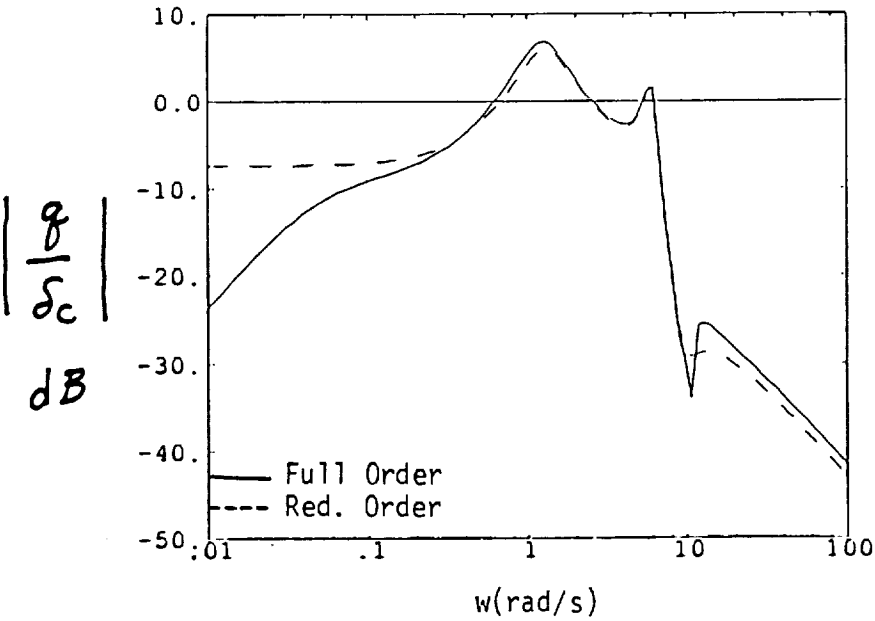


Figure 6



$\left| \frac{a_z'}{\delta_c} \right|$ dB

Figure 7



$\left| \frac{q}{\delta_c} \right|$ dB

Figure 8

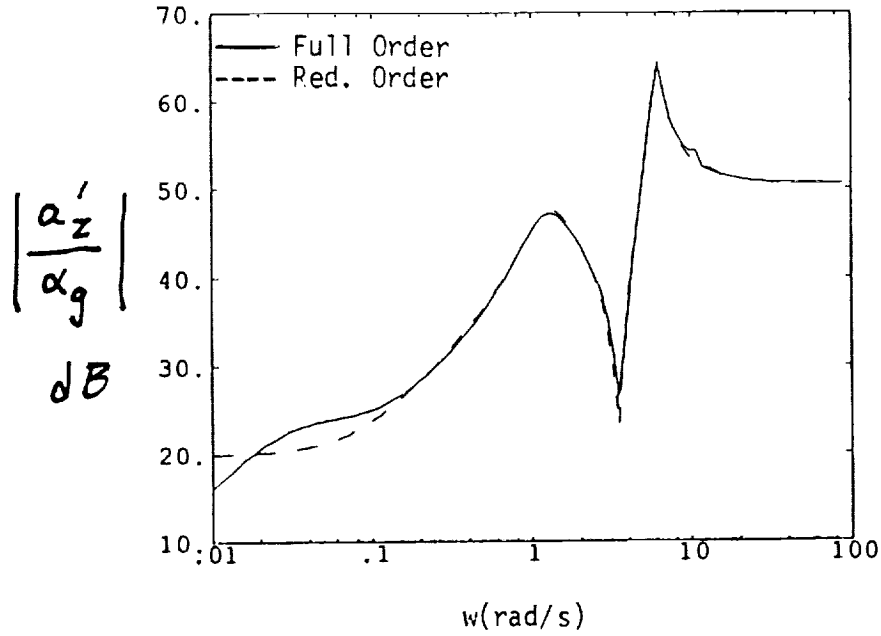


Figure 9

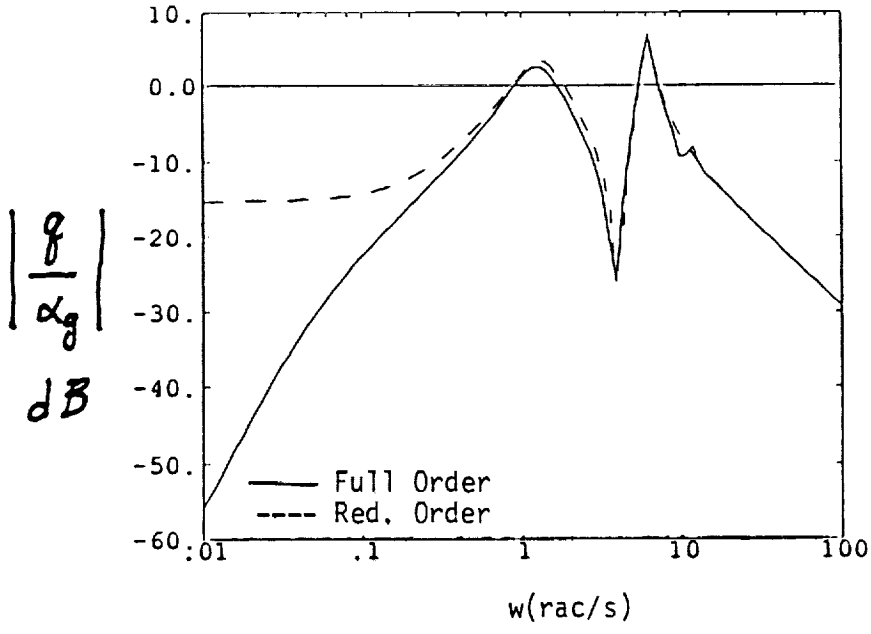


Figure 10

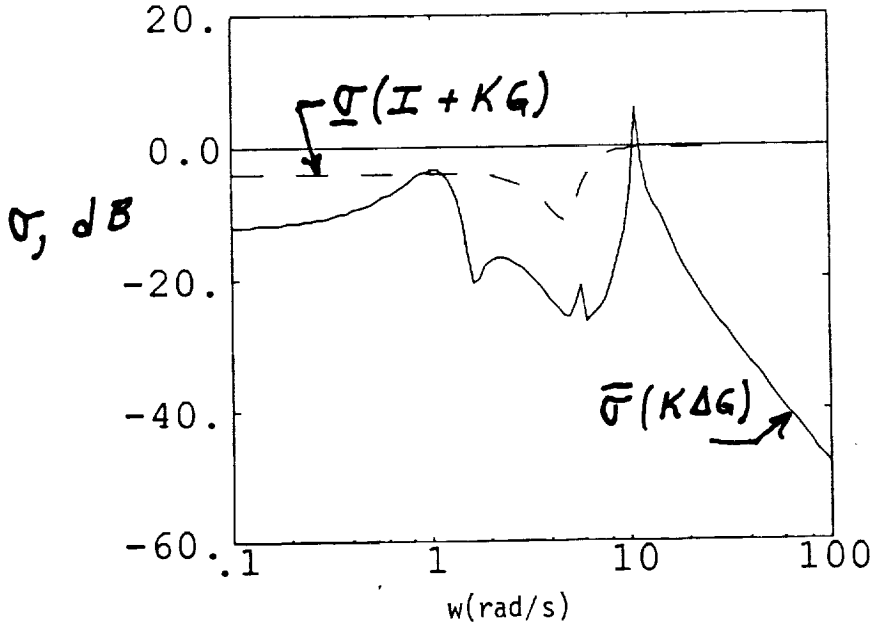


Figure 11

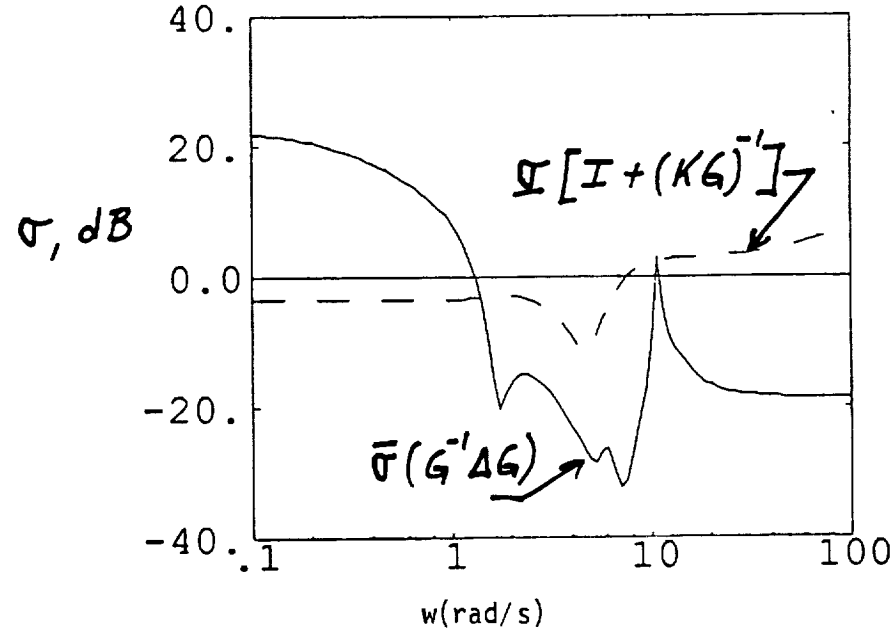


Figure 12

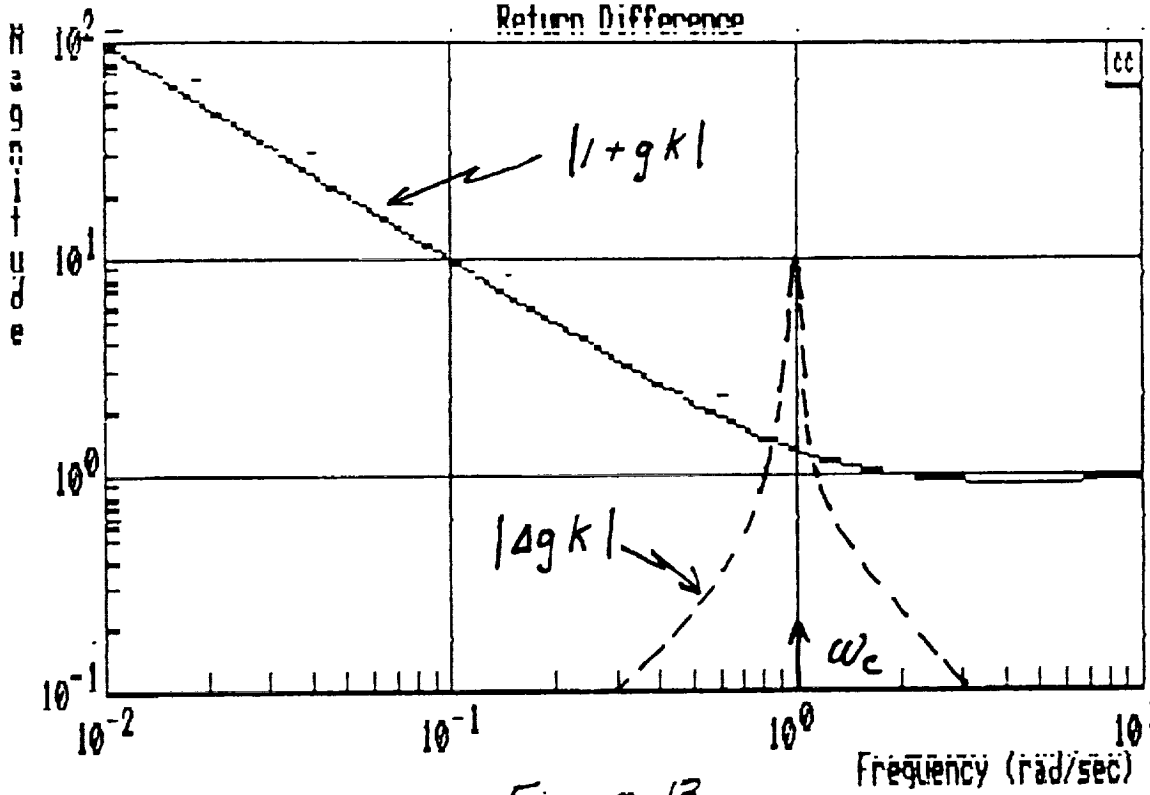


Figure 13.

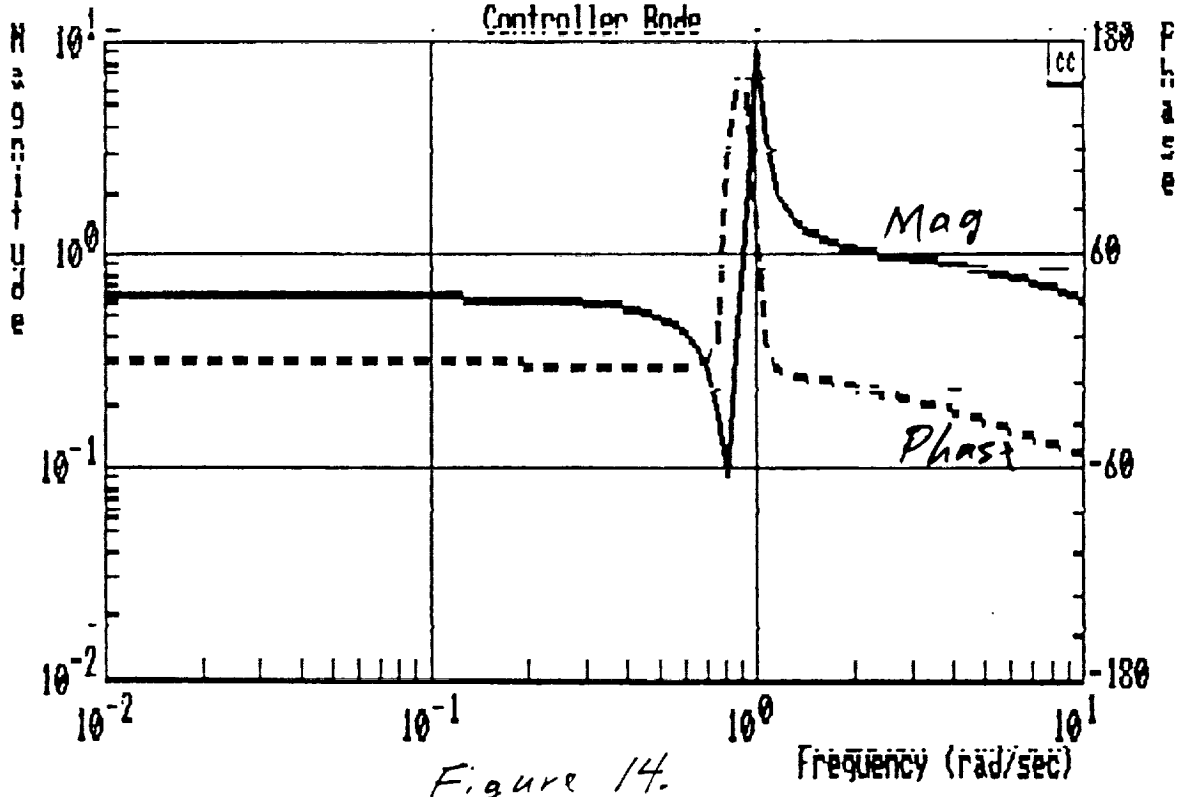


Figure 14.

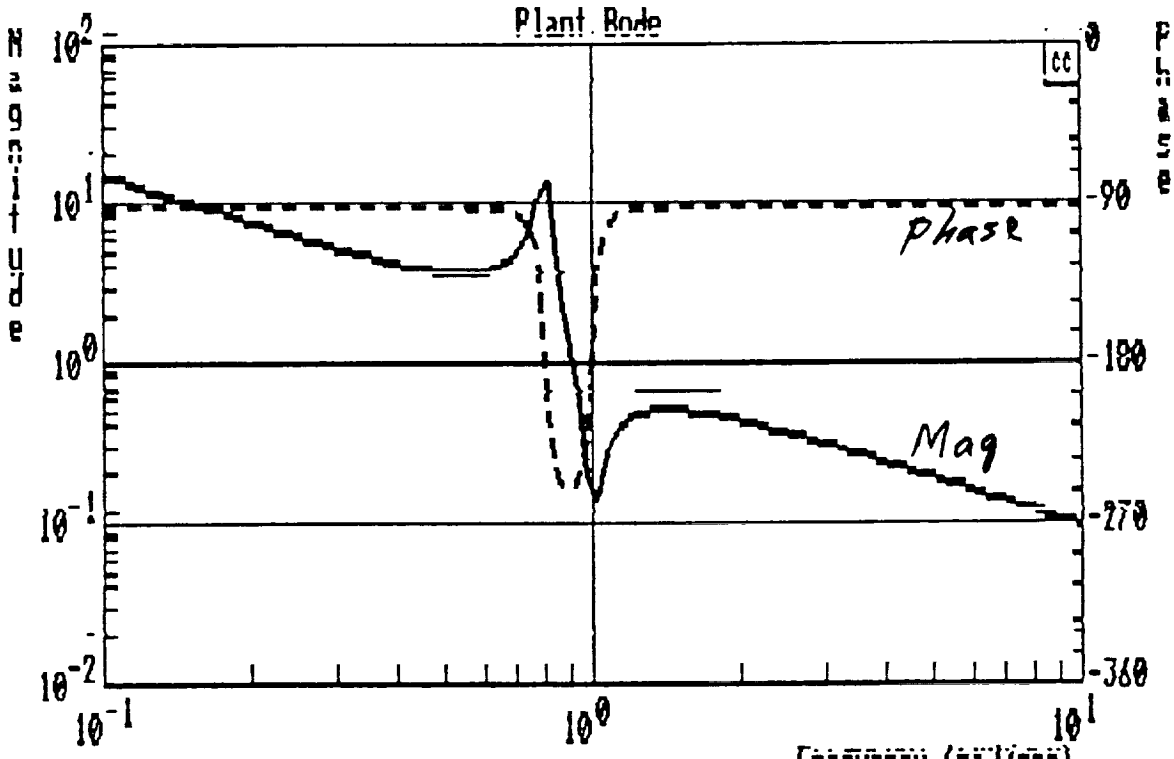

Plant Bode
Magnitude
Phase
10^{2}
10^{1}
10^{0}
10^{-1}
10^{-2}
0
-90
-180
-270
-360
Phase
Mag
10^{-1}
10^{0}
10^{1}
Frequency (rad/sec)

Figure 15.

Appendix D

AIAA-90-3446-CP

MULTIVARIABLE FLIGHT CONTROL SYNTHESIS AND LITERAL ROBUSTNESS ANALYSIS FOR AN AEROELASTIC VEHICLE

David K. Schmidt* and Brett Newman**
Department of Mechanical and Aerospace Engineering
Arizona State University
Tempe, AZ 85287-6106

Abstract

The vehicle to be augmented is representative of a large supersonic transport, with first fuselage aeroelastic mode frequency at six rad/sec, very close to the two rad/sec short-period mode. An integrated flight- and aeroelastic-mode control law is synthesized using a previously developed model-following synthesis approach. This technique, designed to yield a desired closed-loop rather than an open-loop loop shape, involves a specific LQR formulation leading to the model-following state-feedback gains. Then the use of asymptotic loop transfer recovery is utilized to obtain the compensation that recovers the LQR robustness properties, and which leads to an output-feedback control law. A classically designed control law is also developed for comparison purposes. The resulting closed-loop systems are then evaluated in terms of their performance and multivariable stability robustness, measured in terms of the appropriate singular values. This evaluation includes the use of approximate literal expressions for those singular values, expressed in terms of literal expressions for the poles and zeros in the vehicle transfer-function matrix. It is found that the control laws possess roughly equivalent performance and stability robustness, and the characteristecs limiting this robustness are traced to some specific loop gains and the frequency and damping of the open-loop aeroelastic mode dipole. Furthermore, closed-form literal expressions for these characteristics are presented in terms of the stability derivatives of the vehicle. Insight from such an analysis would be hard to obtain from a strictly numerical proceedure.

1. Introduction

The supersonic and hypersonic capabilities of advanced aerospace vehicles and the use of extremely light metalic or composite materials in them can lead to vehicles with significant dynamic coupling between the rigid-body and elastic motions. Ref. 1 and 2, for example, specifically addressed this coupling at the earliest stage of system modeling and flight-control synthesisin.

Augmentation of an aeroelastic vehicle's open-loop dynamics via feedback is often necessary to provide sufficient levels of stability and performance (e.g., handling qualities). Feedback is used to stabilize the attitude and or aeroelastic responses (such as static aerodynamic instability or flutter) or just augment damping. Crossfeeds may also be used to improve the dynamic responses. And the control-law must ensure this stability and performance in the presence of vehicle modeling errors (i.e., robustness). For aeroelastic vehicle applications, modeling errors can arise from uncertainty in the aerodynamic model and neglected high-frequency structural modes both leading to uncertainty in the pole/zero locations in the vehicle transfer functions, for example.Such control objectives have been noted in the literature[3-9].

If possible, the vehicle model (used in control synthesis) should aid in the understanding and thereby provide insight regarding the vehicle physics, exposing key dynamic characteristics and their causes. This can be achieved by developing literal expressions for the vehicle transfer functions (gains, zeros, and poles) in terms of vehicle model parameters, such as stability and control derivatives or vibrational characteristics, which have their genesis in the fundamental vehicle geometric shape and structural layout.[2,9,10] Models of this type can be an extremely powerful tool in open-loop or closed-loop design.[11]

The control synthesis for an aeroelastic vehicle, and the systems' analyses specifically using a literal model, is the subject of this paper. An aeroelastic vehicle model is briefly presented and deficiencies in the vehicle dynamics are noted. Control objectives are stated and sufficient conditions ensuring an acceptable design are given. A new approach to implicit model following (IMF) control synthesis[12,13] is briefly discussed and applied to the vehicle model. A classical control synthesis approach is also considered for the purposes of comparison. The resulting compensators and closed-loop systems are analyzed with a literal model to expose sources of system characteristics that limit the closed-loop system stability robustness. It will be shown, for example, that major among these critical characteristics are the frequency and damping of the vehicles first aeroelastic mode dipole, and closed-

---

* Professor of Engineering; Assoc. Fellow AIAA

** Research Associate; also Doctoral Candidate, School of Aero and Astro, Purdue Univ.; Student Member AIAA

form expressions for these terms are presented in terms of the vehicle stability derivatives.

2. The Vehicle Model For Feedback Synthesis

The configuration to be considered (from Refs. 2 and 10) is a large supersonic aircraft of reasonably conventional geometry with a low-aspect ratio swept wing, conventional tail, and canard. Controlled inputs consist of elevator δ_E and canard (located near the cockpit) deflection δ_C. The reference flight condition is level flight at Mach 0.6 and altitude 5,000 ft.

The complete non-linear modelling of this vehicle was the subject of Ref. 2, and the development of low order linear models for control synthesis was considered in Ref. 10. A fourth order state space realization and the corresponding transfer functions for this linear model are given in Tables 1 and 2. This model involves the small perturbation longitudinal dynamics of the effective short period and first aeroelastic modes. The responses of interest are the rigid-body angle of attack α, rigid-body pitch rate q, and pitch rate q' measured at the cockpit. Here, rigid-body α and q are the angle of attack and pitch rate associated with the vehicle mean axes. An approximate measurement of q can be obtained from a rate gyro located at the anti-node of the first elastic structural mode, and then lowpass filtering of the higher-frequency modes. The effects of such filtering will not be specifically addressed, but it would add additional phase loss in the loops, which is considered in the robustness analysis

Table 1. Elastic Aircraft Model

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

$$y = \begin{bmatrix} \alpha \text{ (rad)} \\ q \text{ (rad/s)} \\ q' \text{ (rad/s)} \end{bmatrix}, \quad u = \begin{bmatrix} \delta_E \text{ (rad)} \\ \delta_C \text{ (rad)} \end{bmatrix}$$

$$A = \begin{bmatrix} -0.517 & 3.85 & 0.150 & 4.24 \\ -9.39 & -0.318 & -0.523 & -4.67 \\ 0.0438 & 0.0164 & -0.0128 & -2.06 \\ -0.0591 & -0.0165 & 0.764 & -0.986 \end{bmatrix}$$

$$B = \begin{bmatrix} -292. & -182. \\ -598. & -424. \\ 53.7 & -31.2 \\ -38.4 & 17.7 \end{bmatrix}$$

$$C = \begin{bmatrix} 0.000480 & -0.0000247 & -0.0188 & -0.0286 \\ 0.00147 & 0.00170 & -0.0264 & 0.0549 \\ -0.0222 & -0.0213 & -0.0372 & 0.0687 \end{bmatrix}$$

Table 2. Elastic Aircraft Transfer Functions

$\alpha(s)/\delta_E(s) = -0.036(s - 0.018 \pm j4.9)(s + 150.)/d(s)$ rad/rad

$q(s)/\delta_E(s) = -5.0(s + 0.36)(s + 0.11 \pm j4.9)/d(s)$ rad/s/rad

$q'(s)/\delta_E(s) = 15.(s + 0.040)(s - 2.9)(s + 4.0)/d(s)$ rad/s/rad

$\alpha(s)/\delta_C(s) = 0.0044(s + 1.8 \pm j9.0)(s + 200.)/d(s)$ rad/rad

$q(s)/\delta_C(s) = 0.80(s + 0.33)(s + 1.3 \pm j9.1)/d(s)$ rad/s/rad

$q'(s)/\delta_C(s) = 15.(s + 0.056)(s + 0.73 \pm j2.9)/d(s)$ rad/s/rad

where $d(s) = (s + 0.47 \pm j1.2)(s + 0.44 \pm j6.0)$

A fourth-order model was developed to accurately approximate the appropriate frequency responses of a twelfth-order model, in the anticipated critical frequency range of 1 to 10 rad/s. Figures 1 thru 3 show some of these frequency responses from elevator input δ_E. (It is noted that the next significant unmodeled aeroelastic mode frequency is above 13 rad/sec.) From Table 2 and Figures 1 thru 3, the major open-loop dynamic deficiency is the level of damping of the short period and aeroelastic modes. Furthermore, the aeroelastic mode contributes significantly to the vehicle's dynamic responses.

3. Classical Control Synthesis

A classical design approach consists of sequential single loop closures, using root loci, and relying upon knowledge of the physics of the elastic aircraft for synthesis strategy.

Consider a 2 × 2 system from Table 2 with the following notation.

$$q(s) = g_{11}(s)\delta_E(s) + g_{12}(s)\delta_C(s) \qquad (1)$$
$$q'(s) = g_{21}(s)\delta_E(s) + g_{22}(s)\delta_C(s)$$

First, the q'/δ_C loop is closed to improve the aeroelastic mode damping. Recall q' and δ_C are a co-located sensor and actuator pair near the cockpit. The control law $\delta_C = \delta_C' - k_{22}q'$ yields

$$q = \left(g_{11} - \frac{k_{22}g_{12}g_{21}}{1 + k_{22}g_{22}}\right)\delta_E + g_{12}\left(1 - \frac{k_{22}g_{22}}{1 + k_{22}g_{22}}\right)\delta_C' \qquad (2)$$
$$q' = \frac{g_{21}}{1 + k_{22}g_{22}}\delta_E + \frac{g_{22}}{1 + k_{22}g_{22}}\delta_C'$$

The root locus for $1 + k_{22}\frac{n_{22}}{d}$, where n_{ij} and d are the numerator and denominator polynomials, respectively, of g_{ij}, is shown in Figure 4. A gain of k_{22} = 0.05 rad/rad/s increases the aeroelastic mode damping by over 60% of the open-loop value.

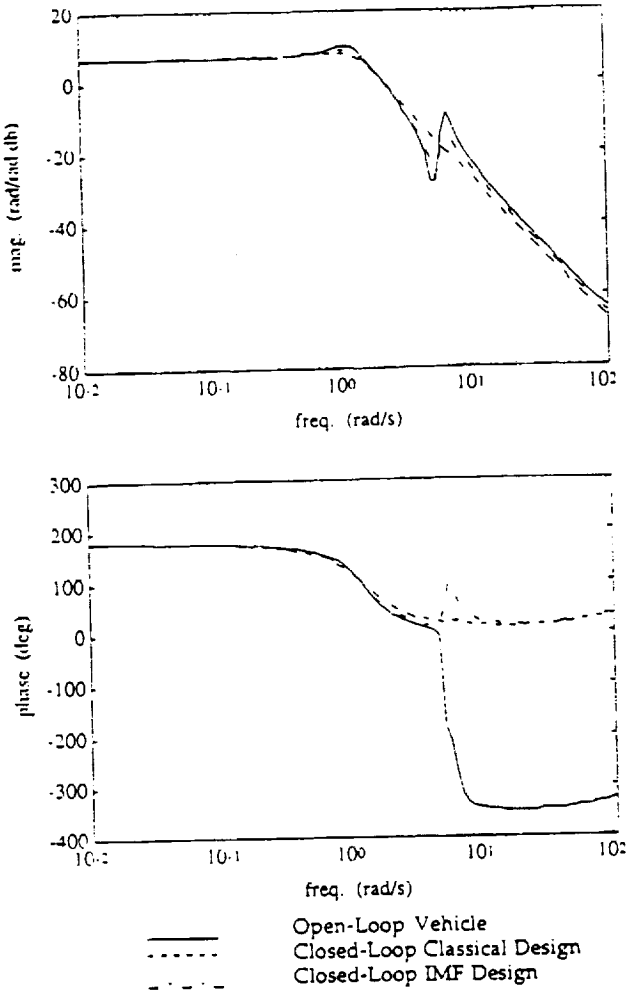


Figure 1. $\alpha(s)/\delta_T(s)$ And $\alpha(s)/\delta(s)$ Frequency Responses

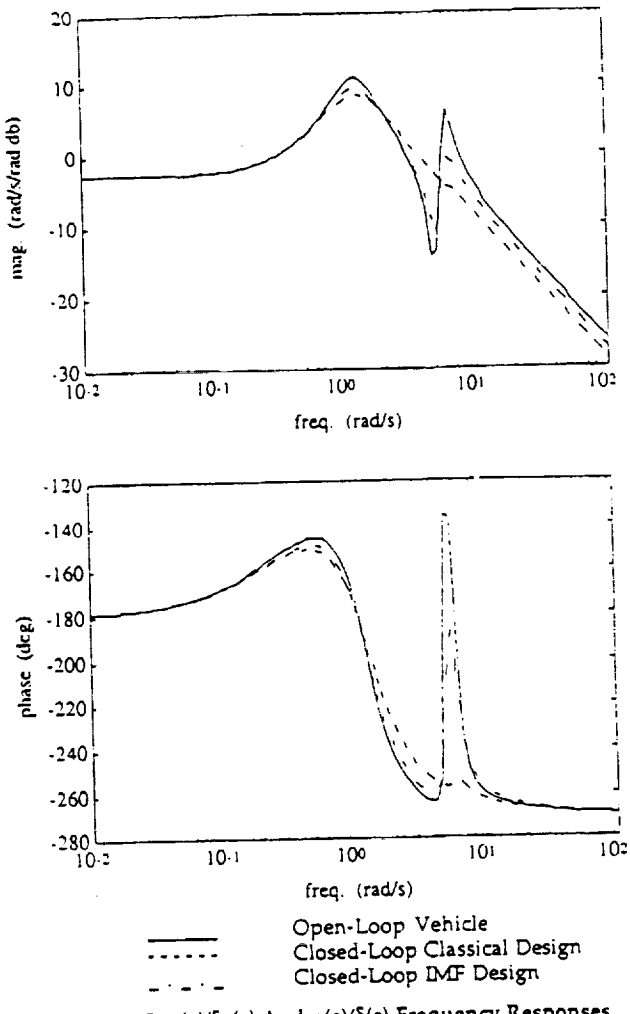


Figure 2. $q(s)/\delta_E(s)$ And $q(s)/\delta(s)$ Frequency Responses

An elevator-to-canard crossfeed is now introduced to reduce aeroelastic mode excitation from the elevator. Interconnecting "up canard" with "up elevator" will reduce aeroelastic mode deflections from the elevator because the fuselage mode shape[2] is similar to the fundamental bending mode shape of a slender beam.

The crossfeed $\delta_C' = k_{cf}\,\delta_E$ yields

$$q = \frac{g_{11} + k_{22}(g_{11}g_{22} - g_{12}g_{21}) + k_{cf}\,g_{12}}{1 + k_{22}g_{22}}\,\delta_E \qquad (3)$$

This can be simplified with the identity[7]

$$\det[G] = \det\begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = g_{11}g_{22} - g_{12}g_{21} = \frac{\psi_G}{d} \qquad (4)$$

where ψ_G is the transmission zero polynomial corresponding to the plant in eq. (1).

$$\psi_G(s) = 89(s + 0.081)(s + 0.46) \qquad (5)$$

Substitution of eq. (4) into eq. (3) yields

$$q = \frac{n_{11} + k_{22}\psi_G + k_{cf}\,n_{12}}{d + k_{22}n_{22}}\,\delta_E \qquad (6)$$

It is now evident the crossfeed has the effect of moving the zeros of the q/δ_E transfer function (with the q'/δ_C loop closed) from $n_{11} + k_{22}\psi_G$ to n_{12}. The root locus for $1 + k_{cf}\frac{n_{12}}{n_{11} + k_{22}\psi_G}$ is shown in Figure 5. A gain of $k_{cf} = -1.5$ rad/rad results in almost perfect pole-zero cancellation for the aeroelastic dipole in the effective q/δ_E transfer function.

Finally, the effective q/δ_E loop is closed to furter improve the short period damping. The control law $\delta_E = p\delta - k_{11}q$ yields

$$q = \frac{p(g_{11} + k_{22}(g_{11}g_{22} - g_{12}g_{21}) + k_{cf}\,g_{12})}{1 + k_{22}g_{22} + k_{11}(g_{11} + k_{22}(g_{11}g_{22} - g_{12}g_{21}) + k_{cf}\,g_{12})}\,\delta \qquad (7)$$

where p is the gain on the pilot input δ. The root locus

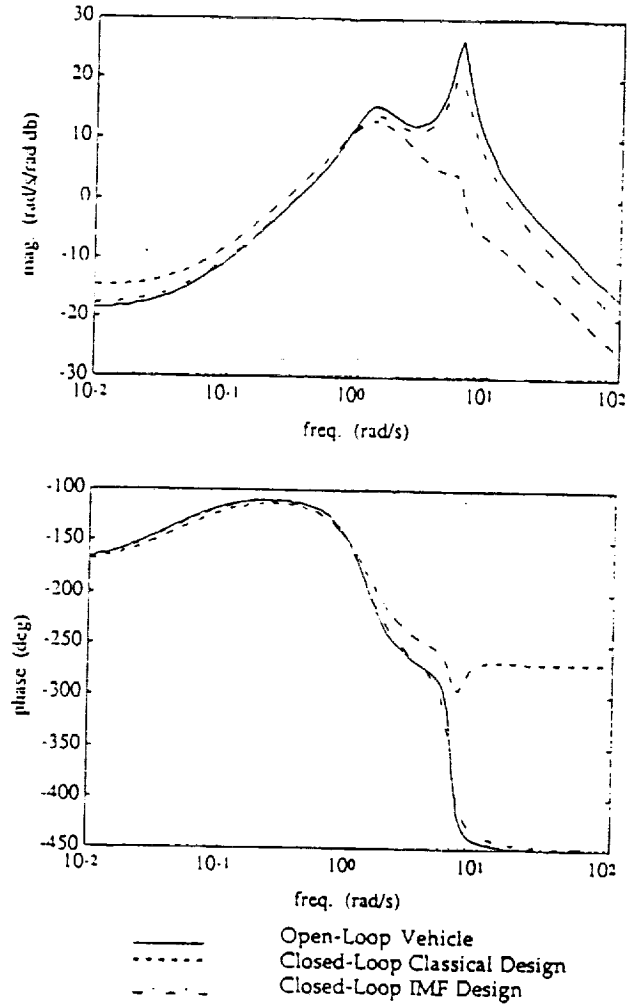
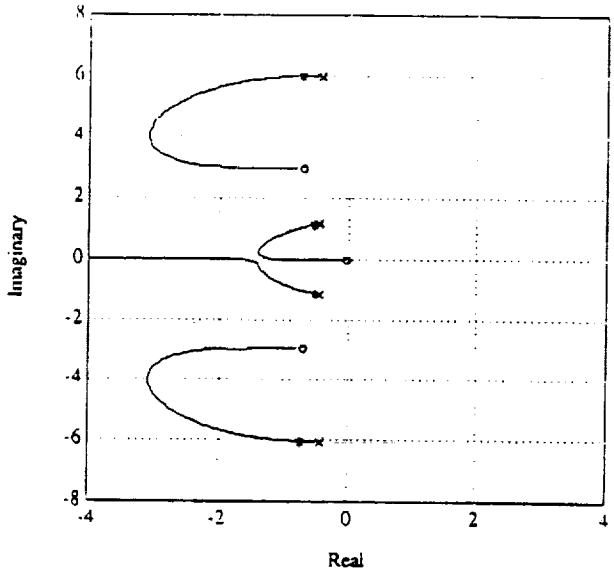


Figure 3. $q'(s)/\delta_E(s)$ And $q'(s)/\delta(s)$ Frequency Responses

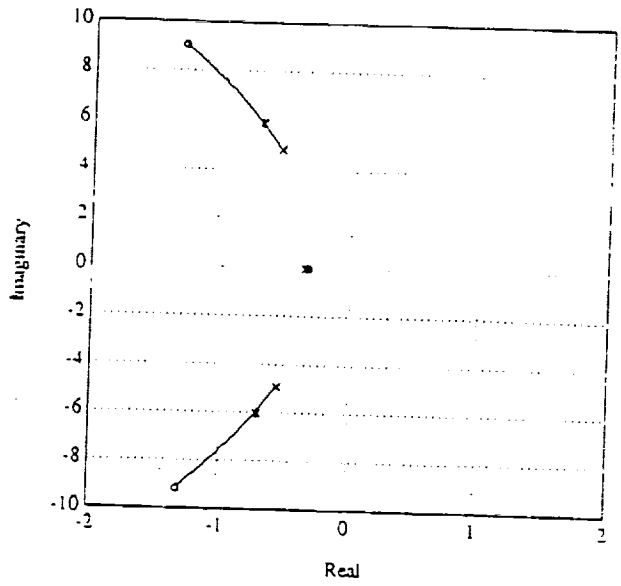


x : -0.47±j1.2
-0.44±j6.0

o : -0.056
-0.73±j2.9

* (k_{22} = 0.05) : -0.55±j1.1
-0.75±j6.0

Figure 4. Root Locus For q'/δ_C Loop Closure



x : -0.36
-0.56±j4.9

o : -0.33
-1.3±j9.1

* (k_{cf} = -1.5) : -0.35
-0.71±j5.9

Figure 5. Zero Root Locus For δ_E To δ_C Crossfeed

for $1 + k_{11}\dfrac{n_{11} + k_{22}\psi_C + k_{cf}\,n_{12}}{d + k_{22}n_{22}}$ is shown in Figure 6, as well as the final closed-loop pole locations for a gain of k_{11} = -0.05 rad/rad/s.

With some block diagram manipulation, the closed-loop system can be represented as in Figure 7. Table 3 contains the effective closed-loop transfer functions corresponding to the pilot command δ, while Figures 1 thru 3 show the corresponding frequency responses for the augmented vehicle. Short-period damping has improved from ζ_{sp} = 0.36 (see Table 2) to ζ_{sp} = 0.54 (a 50% increase), while the first aeroelastic mode damping has improved from ζ_{f_1} = 0.073 (see Table 2) to ζ_{f_1} = 0.12 (a 64% increase). These improvements are apparent in the closed-loop frequency responses.

Significant improvement in the rigid-body (α and q) frequency response shapes is also achieved. Besides improved short period damping, the aeroelastic mode pole-zero "saw tooth" located near 6 rad/s in Figures 1 and 2 are virtually eliminated when compared to the corresponding open-loop behavior. This is a result of improved closed-loop pole-zero cancellations (see Table 3) as desired in the classical control synthesis. Or the aeroelastic mode has been rendered undisturbable from pilot input.

4. IMF Control Synthesis[12,13]

A newly developed technique for the synthesis of flight-control laws will now be outlined. Although LQR and LTR concepts are used in the formulation of the algorithm, this approach is fundamentaly different

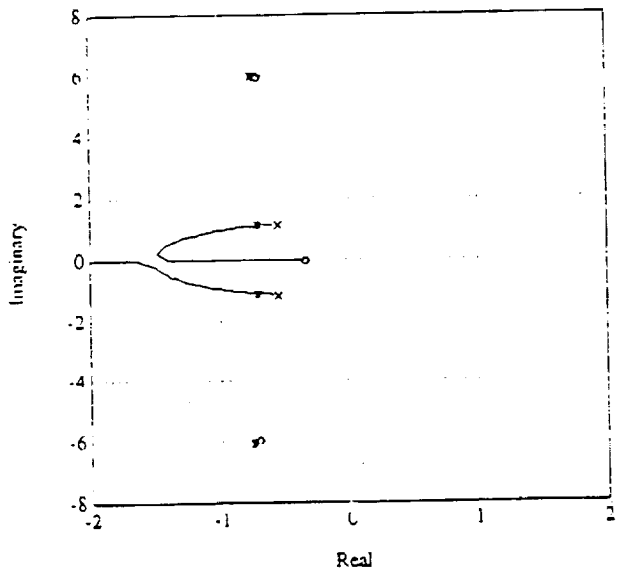


Figure 6. Root Locus For q/δ_E Loop Closure

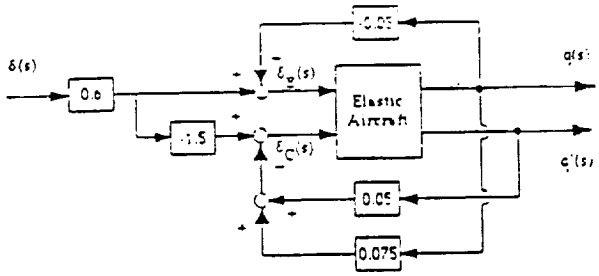


Figure 7. Classically Designed Closed-Loop System

from LQG/LTR methodology[4]. LQG/LTR addresses the problem of obtaining specified open-loop shapes, while the approach taken here is to synthesize a control law that yields a desired closed-loop shape.

The system to be controlled is represented as

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad (8)$$
$$y(t) = Hx(t)$$

and the model of the desired dynamics to be followed is

$$\dot{x}_m(t) = A_m x_m(t) + B_m \delta(t) \qquad (9)$$
$$y_m(t) = H_m x_m(t)$$
$$\dot{\delta} = -100.\,\delta$$

where δ represents the input from the pilot. The error vector

$$e(t) = y(t) - y_m(t) \qquad (10)$$

is constrained to be governed by stable, homogeneous dynamics

$$\dot{e}(t) = -Ge(t) \qquad (11)$$

where G is to be selected in the synthesis process. The model-following control law is obtained by solving the LQR problem with the following objective function.

$$J = \int_0^\infty [(\dot{e} + Ge)^T Q(\dot{e} + Ge) + u^T R u]\, dt \qquad (12)$$

If the product HB is square and invertible, and the same for $H_m B_m$, and if G is chosen as $G = -H_m A_m H_m^{-1}$, then perfect model following is achieved asymptotically as R in Eqn. 7 approaches the null matrix. If this is the case, then the closed-loop poles approach the model poles (for G as defined above) and any open-loop plant finite transmission zeros (or their stable miror image).[12]

The solution to this problem iss the first step of the control law synthesis, yielding the state-feedback control law

$$u = Kx + K_\delta \delta \qquad (13)$$

For the elastic aircraft model in Table 1, rigid-body angle of attack and pitch rate α and q, are the responses selected for model following, so that the handling characteristics will be improved. Also it is desirable that the response approximate that of a rigid vehicle. With this selection, the open-loop plant transmission zeros are located at -23. 1/s and 35. 1/s, and HB is square and invertible.

The model of the desired dynamics is chosen to be

Table 3. Closed-Loop Transfer Functions

Classical Control Synthesis

$\alpha(s)/\delta(s) = 3.7(s + 0.70 \pm j5.9)(s + 160)/d(s)$ rad/rad

$q(s)/\delta(s) = 0.025(s + 0.35)(s + 0.71 \pm j5.9)/d(s)$ rad/s/rad

$q'(s)/\delta(s) = 5.1(s + 0.049)(s + 1.0 \pm j6.6)/d(s)$ rad/s/rad

where $d(s) = (s + 0.70 \pm j1.1)(s + 0.75 \pm j6.0)$

IMF Control Synthesis

$\alpha(s)/\delta(s) = -0.0062(s + 0.22 \pm j5.1)(s + 150)/d(s)$ rad/rad

$q(s)/\delta(s) = -0.87(s + 0.36)(s + 0.34 \pm j5.1)/d(s)$ rad/s/rad

$q'(s)/\delta(s) = 2.0(s + 0.042)(s - 3.6)(s + 4.5)/d(s)$ rad/s/rad

where $d(s) = (s + 0.56 \pm j1.1)(s + 0.73 \pm j5.8)$

$$\alpha_m(s)/\delta(s) = -3.5/(s + 0.89 \pm j0.91)$$

$$q_m(s)/\delta(s) = -3.3(s + 0.36)/(s + 0.89 \pm j0.91)$$

Note that the short-period mode is well damped. With this selection, H_m is square and also invertible. The short-period poles will approach those of the model, the aeroelastic mode poles will move toward the plant transmission zeros (or their stable mirror images via the optimal control formulation), and the α and q time and frequency responses will be shaped to better approximate those of a rigid vehicle.

With the state feedback gains K and K_δ so determined, compensators will now be synthesized using the loop transfer recovery procedure[4,7,13], which will then yield the output-feedback loop structure in Figure 9. Although the α and q responses were those used in the model-following step, they are not the measurements to be used for feedback. The feedback measurements are the same used for the classical design, q' and q. This selection leads to minimum phase transmission zeros, for the loop-transfer recovery, located at -0.081 and -0.46 1/s.

Figure 8 shows the resulting feedback compensators, prefilter stick gains, and closed-loop structure after the loop-transfer recovery proceedure is completed, and some straight-forward pole-zero cancellations are performed on the compensators. Note the compensators consist of relatively simple lead-lag and lag-lead filters of second order. Table 3 contains the effective closed-loop transfer functions corresponding to the pilot command δ, while Figures 1 thru 3 show the corresponding frequency responses. Short period damping has improved from ζ_{sp} = 0.36 to ζ_{sp} = 0.45 (25% increase) while the first aeroelastic mode damping has improved from ζ_{f_1} = 0.073 to ζ_{f_1} = 0.12 (64% increase).

These improvements are also apparent in the closed-loop frequency responses. Besides improved short-period damping, the aeroelastic mode pole-zero "saw tooth" located near 6 rad/s in the angle-of-attack and pitch-rate responses in Figures 1 and 2 is reduced by roughly 10 db, when compared to the corresponding open-loop response. This is a result of improved

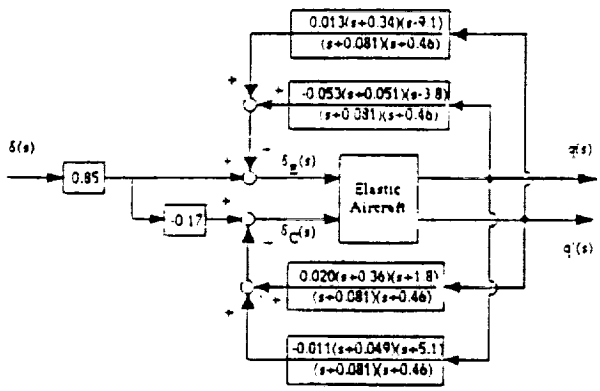


Fig.8, IMF Designed Closed-Loop System

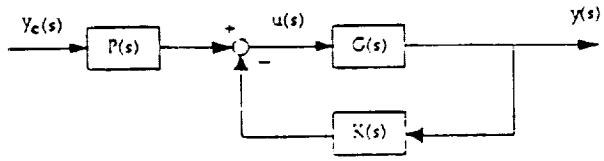


Fig. 9, Generic Closed-Loop System

closed-loop pole-zero cancellations (see Table 3) as desired in the IMF control synthesis (i.e., following a rigid-body model).

5. Robustness Analysis

Now consider the generic feedback loop structure in Figure 9, which is a generalization of the closed-loop systems in Figures 7 and 8, with response vector y, control inputs u, commands y_c , and plant, compensator, and prefilter transfer function matrices G(s), K(s), and P(s), respectively. The feedback compensation in Figure 9 is assumed to be synthesized with a design model G(s), but the "true" plant transfer function is taken to be G'(s). Specifically, consider generic phase loss in each input channel to the plant, or let

$$G'(s) = G(s)\,(e^{-\tau s} I) \qquad (14)$$

This phase loss can represent, for example, unmodeled high-frequency dynamics originating from structural modes, actuators, sensors, etc. Rewriting G'(s) as

$$G'(s) = G(s)(I + E(s)) \qquad (15)$$

it can be shown that

$$E(s) = (e^{-\tau s} - 1)\, I \qquad (16)$$

where E(s) is the so called plant input multiplicative error.[4]

The "true" closed-loop system poles are roots of the "true" characteristic equation, obtained from

$$\det[I + K(s)G(s)(I + E(s))] = 0 \qquad (17)$$

If the nominal closed-loop system is stable and the required number of encirclements of the critical point in Nyquist stability theory is the same for both nominal and "true" systems, then a sufficient condition, developed from Eq. (17), guaranteeing closed-loop stability under E(s)[4] is

$$\bar{\sigma}[E(j\omega)] < \underline{\sigma}[I + (K(j\omega)G(j\omega))^{-1}] \quad , \quad 0 \le \omega \le \infty \qquad (18)$$

Eq. (18) is an indication of the system's multivariable stability robustness margin.

Figure 10 indicates the stability robustness of the classically designed closed-loop system, with the effect of multiplicative error due to generic phase loss in each input channel displayed as well. Note from Figure 10 the characteristic limiting the stability robustness is the dip in $\underline{\sigma}[I + (KG)^{-1}]$ near 6 rad/s. In fact, the phase loss allowed using this criteria is limited to $\tau \le 0.3$ s.

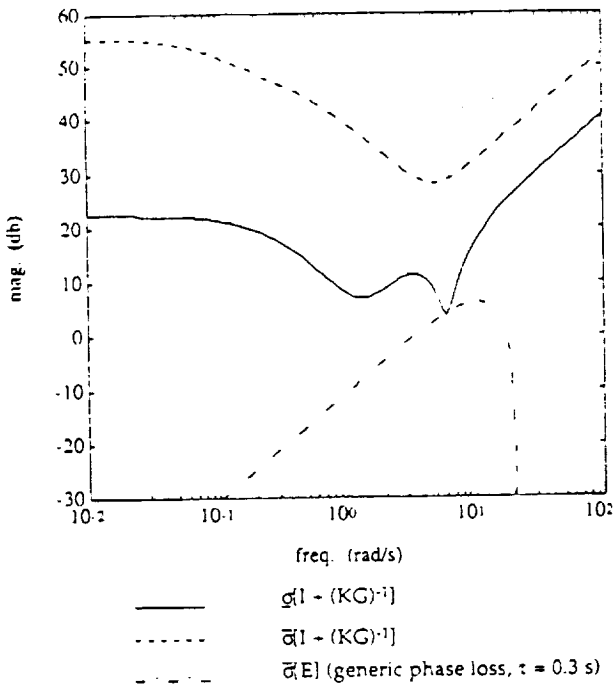


Figure 10. Classical Design Stability Robustness Analysis

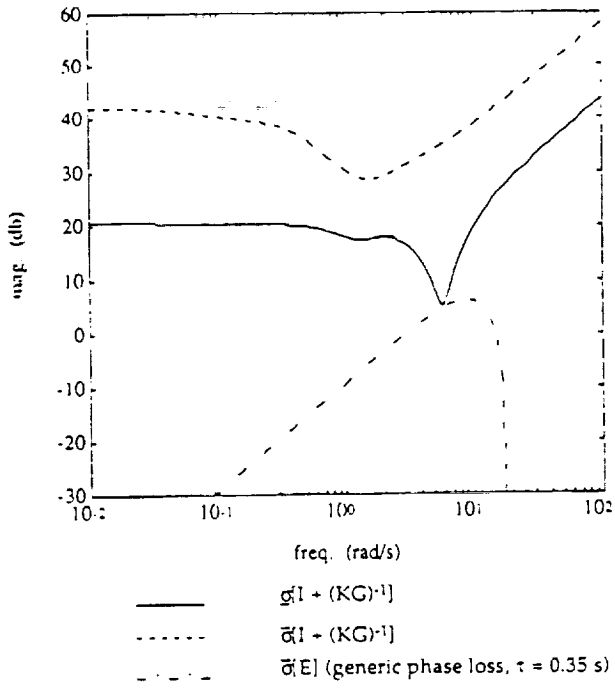


Figure 11. IMF Design Stability Robustness Analysis

Figure 11 indicates the stability robustness properties of the IMF design, again with the effect of generic phase loss displayed. Note again a similar characteristic limiting the robustness of this loop. Here the allowable phase loss is $\tau \leq 0.35$ s, only slighty better than the previous result.

The question now turns to the causes for this limiting characteristic. Literal expressions for the vehicle transfer function poles and zeros in Table 2 are available from Ref. 10 for further analysis. Before this, however, a literal expression for $\underline{\sigma}[I + (KG)^{-1}]$ is necessary. The approach to be taken here is similar to that presented in Ref. 11.

With reference to Figure 9, consider a 2 × 2 closed-loop system with

$$K(j\omega) = \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}, \qquad G(j\omega) = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

$$I + (K(j\omega)G(j\omega))^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{19}$$

where

$$\begin{aligned} a_{11} &= 1 + \frac{k_{21}g_{12} + k_{22}g_{22}}{\Delta} \\ a_{12} &= -\frac{k_{11}g_{12} + k_{12}g_{22}}{\Delta} \\ a_{21} &= -\frac{k_{21}g_{11} + k_{22}g_{21}}{\Delta} \\ a_{22} &= 1 + \frac{k_{11}g_{11} + k_{12}g_{21}}{\Delta} \end{aligned} \tag{20}$$

$$\Delta = \det[KG] = [k_{11}k_{22} - k_{12}k_{21}]\,[g_{11}g_{22} - g_{12}g_{21}]$$

The minimum and maximum singular values of $I + (KG)^{-1}$ are given as

$$\underline{\sigma}[I + (KG)^{-1}] = \underline{\lambda}^{1/2}[(I + (KG)^{-1})(I + (KG)^{-1})^*] \tag{21}$$

$$\bar{\sigma}[I + (KG)^{-1}] = \bar{\lambda}^{1/2}[(I + (KG)^{-1})(I + (KG)^{-1})^*]$$

where $\underline{\lambda}$ and $\bar{\lambda}$ denote the minimum and maximum eigenvalues, respectively. $\underline{\lambda}$ and $\bar{\lambda}$ solve

$$\det[\lambda I - (I + (KG)^{-1})(I + (KG)^{-1})^*] = \lambda^2 - (\underline{\lambda} + \bar{\lambda})\lambda + \underline{\lambda}\bar{\lambda} = 0 \tag{22}$$

where

$$\begin{aligned} \underline{\lambda} + \bar{\lambda} &= |a_{11}|^2 + |a_{12}|^2 + |a_{21}|^2 + |a_{22}|^2 \\ \underline{\lambda}\bar{\lambda} &= |a_{11}a_{22} - a_{12}a_{21}|^2 \end{aligned} \tag{23}$$

If $\underline{\lambda} \ll \bar{\lambda}$, then from Eq. (23), $\underline{\lambda}$ is approximately given as

$$\underline{\lambda} \approx \frac{|a_{11}a_{22} - a_{12}a_{21}|^2}{|a_{11}|^2 + |a_{12}|^2 + |a_{21}|^2 + |a_{22}|^2} \tag{24}$$

or

$$\underline{\sigma}[I + (KG)^{-1}] \approx \frac{|a_{11}a_{22} - a_{12}a_{21}|}{(|a_{11}|^2 + |a_{12}|^2 + |a_{21}|^2 + |a_{22}|^2)^{1/2}} \tag{25}$$

From inspection of Figures 10 and 11, it can be seen that the condition $\underline{\lambda} \ll \bar{\lambda}$ (or $\underline{\sigma} \ll \bar{\sigma}$) is reasonably satisfied.

Substitution of Eq. (20) into the numerator and denominator of Eq. (25. yields

$$|a_{11}a_{22} - a_{12}a_{21}| =$$

$$|\frac{1}{\Delta}| \cdot |1 + k_{11}g_{11} + k_{12}g_{21} + k_{21}g_{12} + k_{22}g_{22} + \Delta| \quad (26)$$

$$|a_{11}|^2 + |a_{12}|^2 + |a_{21}|^2 + |a_{22}|^2 =$$

$$|\frac{1}{\Delta}|^2 \cdot (|k_{21}g_{12} + k_{22}g_{22} + \Delta|^2 + \quad |k_{11}g_{12} + k_{12}g_{21}|^2 +$$

$$|k_{21}g_{11} + k_{22}g_{21}|^2 + |k_{11}g_{11} + k_{12}g_{21} + \Delta|^2)$$

These can be further simplified with the following observation[7]

$$1 + k_{11}g_{11} + k_{12}g_{21} + k_{21}g_{12} + k_{22}g_{22} + \Delta =$$

$$\det [I + KG] = \frac{\Phi_{cl}}{\Phi_{ol}} \quad (27)$$

$$\Delta = \det [KG] = \frac{\Psi_K \Psi_G}{\Phi_{ol}}$$

where Φ_{cl} and Φ_{ol} are the system's closed-loop and open-loop characteristic polynomials, respectively, and Ψ_K and Ψ_G are the compensator and plant transmission zero polynomials, respectively. With the notation

$$k_{ij}g_{pq} = \frac{n_{k_{ij}} n_{g_{pq}}}{\Phi_{ol}} \quad , \qquad i,j,p,q = 1,2$$

where $n_{k_{ij}}$ and $n_{g_{pq}}$ are the numerator polynomials of k_{ij} and g_{pq} ,respectively, substitution of Eqs. (26) and (27) into Eq. (25) yields the following literal expression for $\underline{\sigma}[I + (KG)^{-1}]$.

$$\underline{\sigma}[I + (KG)^{-1}] = \frac{|\Phi_{cl}|}{(|n_{11}|^2 + |n_{12}|^2 + |n_{21}|^2 + |n_{22}|^2)^{1/2}} \quad (28)$$

$$n_{11} = n_{k_{21}} n_{g_{12}} + n_{k_{22}} n_{g_{22}} + \Psi_K \Psi_G$$

$$n_{12} = n_{k_{11}} n_{g_{12}} + n_{k_{12}} n_{g_{22}}$$

$$n_{21} = n_{k_{21}} n_{g_{11}} + n_{k_{22}} n_{g_{21}}$$

$$n_{22} = n_{k_{11}} n_{g_{11}} + n_{k_{12}} n_{g_{21}} + \Psi_K \Psi_G$$

Observe that the "zeros" of $\underline{\sigma}[I + (KG)^{-1}]$ are nothing more than the closed-loop poles, while the "poles" of $\underline{\sigma}[I + (KG)^{-1}]$ depend on the plant's and compensator's transfer-function zeros as well as their transmission zeros. This result was first noted in Ref. 11, but the transmission zeros $\Psi_K \Psi_G$ were related to the so called coupling numerators.

Now consider the classically designed closed-loop system shown in Figure 7. Here

$g_{11} = q(s)/\delta_E(s)$ $\qquad$ $k_{11} = n_{k_{11}} = \delta_E/q$

$g_{12} = q(s)/\delta_C(s)$ $\qquad$ $k_{12} = n_{k_{12}} = 0$

$g_{21} = q'(s)/\delta_E(s)$ $\qquad$ $k_{21} = n_{k_{21}} = \delta_C/q$

$g_{22} = q'(s)/\delta_C(s)$ $\qquad$ $k_{22} = n_{k_{22}} = \delta_C/q'$

$\Psi_G(s) = 89(s + 0.081)(s + 1.46)$ $\qquad$ $\Psi_K = k_{11}k_{22} - k_{12}k_{21}$

with g_{ij} available from Table 2 and k_{ij} available from Figure 7. Substitution of the above quantities into Eq. (28) yields

$$\underline{\sigma}[I+(KG)^{-1}] = \frac{|(j\omega + 0.70 \pm j1.1)(j\omega + 0.75 \pm j6.0)|}{0.94\,|(j\omega + 0.22)(j\omega + 3.1 \pm j3.8)|} \quad (29)$$

It is evident that the augmented first aeroelastic mode poles, denoted

$$s^2 + (2\zeta\omega)_{cl_2}s + (\omega^2)_{cl_2} = s^2 + 1.5s + 37 = s + 0.75 \pm j6.0 \quad (30)$$

and their low damping are responsible for the previously discussed critical stability robustness feature near 6 rad/s in Figure 10.

From the classical design (see Section 3) and Figure 4, these poles are primarily a function of the q'/δ_C loop closure. With increasing q'/δ_C root locus gain k_{22} , these augmented aeroelastic mode poles originate from their open-loop locations

$$s^2 + (2\zeta\omega)_{f_1}s + (\omega^2)_{f_1} = s^2 + 0.88s + 36 = s + 0.44 \pm j6.0 \quad (31)$$

and migrate towards their corresponding aeroelastic mode zeros in the q'/δ_C transfer function (see Table 2), denoted as

$$s^2 + {}_{f_1}(2\zeta\omega)_{q'}^{\delta_C}s + {}_{f_1}(\omega^2)_{q'}^{\delta_C} = s^2 + 1.5s + 8.9 = s + 0.73 \pm j2.9 \quad (32)$$

Yielding the closed-loop locations in Eq. (30) for the selected value of k_{22} .

From Ref. 10, the open-loop natural frequency and damping terms of the aeroelastic mode poles and zeros in Eqs. (31) and (32) are approximately given by

$$(\omega^2)_{f_1} = (\omega_1^2 - F_{1\eta_1}) + \frac{(1 + \frac{Z_q}{V_T})M_{\eta_1}F_{1_\alpha}}{(\omega_1^2 - F_{1\eta_1})}$$

$$= \quad 35 \quad + \quad 2.0 \quad (33)$$

$$(2\zeta\omega)_{f_1} = (2\zeta_1\omega_1 - F_{1\eta_1}) + \frac{M_{\eta_1}F_{1_q} + [\frac{Z_{\eta_1}}{V_T} + (1 + \frac{Z_q}{V_T})M_{\dot{\eta}_1}]F_{1_\alpha}}{(\omega_1^2 - F_{1\eta_1})}$$

$$= \quad 0.62 \quad + \quad 0.35$$

$${}_{f_1}(\omega^2)_{q'}^{\delta_C} = \frac{(\omega_1^2 - F_{1\eta_1})M_{\delta_C}}{M_{\delta_C} - \phi_1'(x)F_{1\delta_C}} - \frac{M_{\eta_1} + \phi_1'(x)(1 + \frac{Z_q}{V_T})M_\alpha}{\phi_1'(x)}$$

$$= \quad 2.0 \quad - \quad (-6.5) \quad (34)$$

$${}_{f_1}(2\zeta\omega)_{q'}^{\delta_C} = \frac{(2\zeta_1\omega_1 - F_{1\eta_1})M_{\delta_C} + \phi_1'(x)M_qF_{1\delta_C}}{M_{\delta_C} - \phi_1'(x)F_{1\delta_C}} - \frac{\phi_1'(x)(1 + \frac{Z_q}{V_T})\frac{Z_\alpha}{V_T}M_\alpha F_{1\delta_C}}{(\omega_1^2 - F_{1\eta_1})M_{\delta_C}}$$

$$= \quad 0.82 \quad - \quad (-0.67)$$

with the following numerical values:

$\frac{Z_\alpha}{V_{T_1}} = -0.416$ ft/s^2 $\quad (1 + \frac{Z_q}{V_{T_1}}) = 1.03$

$\frac{Z_{\eta_1}}{V_{T_1}} = -0.00267$ 1/s $\quad M_\alpha = -3.33$ 1/s^2

$M_q = -0.830$ 1/s $\quad M_{\eta_1} = -0.0655$ 1/s^2

$M_{\dot{\eta}_1} = -0.00390$ 1/s $\quad M_{\delta_E} = 0.809$ 1/s^2

$F_{1_\alpha} = -1.040$ 1/s^2 $\quad F_{1_q} = -78.4$ 1/s

$F_{1_{\delta_C}} = -631$ 1/s^2 $\quad (\omega_1^2 - F_{1_{\eta_1}}) = 34.8$ 1/s^2

$(2\zeta_1\omega_1 - F_{1_{\dot{\eta}_1}}) = 0.621$ 1/s $\quad \phi_1'(x) = 0.021$ ft/ft

The above parameters are functions of the flight velocity V_T; rigid-body and aeroelastic aerodynamic stability derivatives Z_i, M_i, and F_{1_i}; first in vacuo elastic mode shape, vibration frequency ω_1 and damping ratio ζ_1. These vehicle parameters appear explicitly in the linear equations of motion for the elastic aircraft[10] listed below.

$$\dot{\alpha} = \frac{Z_\alpha}{V_{T_1}}\alpha + (1 + \frac{Z_q}{V_{T_1}})q + \frac{Z_{\eta_1}}{V_{T_1}}\eta_1 + \frac{Z_{\dot{\eta}_1}}{V_{T_1}}\dot{\eta}_1 + \frac{Z_{\delta_E}}{V_{T_1}}\delta_E + \frac{Z_{\delta_C}}{V_{T_1}}\delta_C$$

$$\dot{q} = M_\alpha\alpha + M_q q + M_{\eta_1}\eta_1 + M_{\dot{\eta}_1}\dot{\eta}_1 + M_{\delta_E}\delta_E + M_{\delta_C}\delta_C$$

$$\ddot{\eta}_1 = F_{1_\alpha}\alpha + F_{1_q}q - (\omega_1^2 - F_{1_{\eta_1}})\eta_1 - (2\zeta_1\omega_1 - F_{1_{\dot{\eta}_1}})\dot{\eta}_1 + F_{1_{\delta_E}}\delta_E + F_{1_{\delta_C}}\delta_C \qquad (35)$$

$$q' = q - \phi_1'(x)\dot{\eta}_1$$

As seen from Eq. (33), the frequency of the open-loop aeroelastic mode poles is primarily due to the elastic mode structural frequency and aerodynamic stiffness (i.e., $(\omega_1^2 - F_{1_{\eta_1}})$). Also, the inherent low damping in this mode is primarily due to the elastic mode structural and aerodynamic damping (i.e., $(2\zeta_1\omega_1 - F_{1_{\dot{\eta}_1}})$). However, note also that approximately 1/3 of the total damping is due to aerodynamic coupling between the rigid and elastic degrees of freedom. It is now clear which key vehicle and compensator parameters contribute to the critical stability robustness properties of this closed-loop system.

Now consider the IMF design closed-loop system shown in Figure 8. Here

$g_{11} = q(s)/\delta_E(s)$ $\quad k_{11} = \delta_E(s)/q(s)$

$g_{12} = q(s)/\delta_C(s)$ $\quad k_{12} = \delta_E(s)/q'(s)$

$g_{21} = q'(s)/\delta_E(s)$ $\quad k_{21} = \delta_C(s)/q(s)$

$g_{22} = q'(s)/\delta_C(s)$ $\quad k_{22} = \delta_C(s)/q'(s)$

$\psi_C(s) = 89(s + 0.081)(s + 0.46)$

$\psi_K(s) = 0.00091(s+0.060)(s+0.35\pm j0.21)(s-1.9)$

with g_{ij} available from Table 2 and k_{ij} available from Figure 8. Substitution of the above quantities into Eq. (28) yields

$$g[I + (KG)^{-1}] = \frac{1(j\omega + 0.56 \pm j1.1)(j\omega + 0.73 \pm j5.8)}{0.661(j\omega + 0.83 \pm j1.0)(j\omega + 4.6)} \qquad (36)$$

It is evident that again the augmented first aeroelastic mode poles

$$s^2 + (2\zeta\omega)_{cl_2}s + (\omega^2)_{cl_2} = s^2 + 1.5s + 34 = s + 0.73 \pm j5.8$$

and their low damping are responsible for the critical stability robustness feature near 6 rad/s in Figure 11. From the IMF design (see Section 4), these poles originate at their open-loop location and migrate toward the transmission zeros (or their stable mirror image) defined through the model-following formulation, as the control weighting in the loss function is reduced (or the loop gains are increased). Although literal approximations for these transmission zeros are still being developed, the above expressions for the open-loop aeroelastic poles again reveal the major source of these critical characteristics.

6. Conclusions

An integrated flight- and aeroelastic-mode control law was synthesized for a very flexible supersonic vehicle, using a previously developed model-following synthesis approach. This technique, designed to yield a desired closed-loop rather than an open-loop loop shapes, involves a specific LQR formulation leading to the model-following state-feedback gains. Then the use of asymptotic loop transfer recovery is utilized to obtain the compensation that recovers the LQR robustness properties, and which leads to an output-feedback control law. A classically designed control law was also developed for comparison purposes, and parallels between the results obtained with the two approaches are observed.

The resulting closed-loop systems were evaluated in terms of their performance and multivariable stability robustness, measured in terms of the appropriate singular values. This evaluation utilized approximate literal expressions for those singular values, expressed in terms of literal expressions for the poles and zeros of the vehicle transfer functions. It was found that both control laws possessed equivalent performance and stability robustness, and the characteristics limiting this robustness were in both cases traced to some specific step in the synthesis process, as well as the locations of critical open-loop poles and zeros (or transmission zeros). Furthermore, closed-form literal expressions for these characteristics were presented in terms of the stability derivatives of the vehicle. The insight gained from this analysis is considered invaluable to the control system designer, and unavailable from strictly numerical analysis.

7. Acknowledgements

This research was supported by NASA Langley Research Center under Grant NAG1-758. Mr. Douglas Arbuckle and Mr. Carey Buttrill have served as technical monitors. Thanks also goes out to Mr. John Schierman for his advice concerning the IMF

procedure, to Mr. Shawn Molodow for development of the classical controller, and to Prof. Bong Wie for several fruitful conversations.

8. References

1. Bisplinghoff, R. L. and Ashley, H., "Principles Of Aeroelasticity," Dover Publications, 1962
2. Waszak, M. R. and Schmidt, D. K., "Flight Dynamics Of Aeroelastic Vehicles," *Journal Of Aircraft*, Vol. 25, No. 6, June, 1988
3. Schmidt, D. K. and Foxgrover, J. A., "Multivariable Flight Control Synthesis Approaches To Meet Handling Qualities Objectives," Proceedings Of The AIAA Guidance And Control Conference, Seattle, Wash., Aug. 1984
4. Doyle, J. C. and Stein, G., "Multivariable Feedback Design: Concepts For A Classical/Modern Synthesis," *IEEE Transactions On Automatic Control*, Vol. AC-26, Feb., 1981
5. McRuer, D., Ashkenas, I., and Graham, D., "Aircraft Dynamics And Automatic Control," Princeton University Press, 1973
6. Blakelock, J. H., "Automatic Control Of Aircraft And Missiles," John Wiley & Sons Inc., 1965
7. Kwakernaak, H. and Sivan, R., "Linear Optimal Control Systems," John Wiley & Sons, 1972
8. Doyle, J.C., et al, "State-Space Solutions To Standard H_2 And H_∞ Control Problems," Proceedings Of The American Control Conference, Atlanta, GA, June, 1988
9. Pearce, B.F., et al, "Analytical Study of Approximate Longitudinal Transfer Functions For A Flexible Airframe," ASD - TDR - 62 - 279, June, 1962.
10. Schmidt, D. K. and Newman, B., "Modeling, Model Simplification, And Stability Robustness With Aeroelastic Vehicles," AIAA-88-4079-CP, Proceedings Of The AIAA Guidance, Navigation, And Control Conference, Minneapolis, Minn. Aug.15-17, 1988. (To appear in the *Journal of Guidance, Control, and Dynamics.)*
11. McRuer, D.T., Meyers, T.T., and Thompson, P.M., "Literal Singular-Value-Based Flight Control System Design Techniques," Vol.12, No. 6, *Journal of Guidance, Control, and Dynamics*, Nov-Dec, 1989.
12. Anderson, M.R. and Schmidt, D.K., "The Significance Of Error Dynamics In Model-Following For Flight Control Design," AIAA-87-2311, Proceedings Of The AIAA Guidance, Navigation, And Control Conference, Monterey, Calif., Aug., 1987. (To appear in the *Journal of Guidance, Control, and Dynamics.)*
13. Schmidt, D.K., Control System Design Research to Meet Aircraft Handling Qualities Requirements, Phase II," Final Report to McDonnell Aircraft Co., Performed at Arizona State Univ., Dept. of Mech. and Aero. Engr., Tempe, AZ, July, 1989.