A UNIFIED APPROACH TO STRUCTURE AND CONTROLLER DESIGN OPTIMIZATIONS

by

Kyong Been Lim

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Engineering Mechanics

APPROVED:

Dr. J. L. Junkins (Chairman)

Dr. S. L. Hendricks

Dr. D. Frederick

Dr. D. T. Mook

Dr. H. F. VanLandingham

August, 1986
Blacksburg, Virginia

A UNIFIED APPROACH TO STRUCTURE AND CONTROLLER
DESIGN OPTIMIZATIONS

by

Kyong Been Lim

(ABSTRACT)

A unified approach to structure and controller design optimization is examined. Difficult problems arise in a unified approach, namely, a high dimensioned design space, nonlinearity, complexity of constraints and many inequality constraints. As a candidate for overcoming the above problems, an optimization algorithm utilizing sequential linear programming and continuation methods is proposed.

The second part of this dissertation examines various ideas associated with both theory and practical issues arising in optimizing for eigenvalue sensitivity and stability robustness with respect to parameter variations or unstructured uncertainties. In particular, the time domain approach to stability robustness is pursued. It is found that a recently proposed stability robustness criteria of Patel and Toda is related to well known concepts of numerical conditioning of the eigenvalue problem and may be derived concisely using eigenvalue conditioning concepts. In addition, we review the more direct and perhaps less rigorous approach of dealing with uncertainties, namely modal insensitivity theory. The mathematical conditions for achieving modal insensitivity and eigenvalue placement simultaneously are reviewed along with a discussion of the practical merit of these ideas. As an alternative, we derive a scalar measure of eigenvalue sensitivity which

is a linearly predicted bound on weighted eigenvalue perturbation; we also introduce an algorithm for minimization of this index. Furthermore, the expressions for eigenvector derivatives are correctly derived for non-self-adjoint case. This latter contribution corrects errors present in at least two textbooks on the subject and serves to clear up confusion in the literature.

Finally, we use examples to demonstrate the design algorithm proposed here and numerically examine various designs arising from corresponding cost functions, using a specific configuration (a flexible free-free beam with an attached rigid body.) The numerical results confirm the conservatism of the stability robustness bound for highly structured perturbations but nevertheless clearly supports the hypothesis that maximizing the robustness measure significantly increases the true robustness of a closed loop system. The numerical results also indicate that maximizing the stability robustness measure is better (more efficient computationally and produces more robust designs) than minimizing the eigenvalue sensitivities directly for improving true stability robustness with respect to perturbations in the closed loop system matrix.

ACKNOWLEDGEMENT

The author wishes to express his deepest gratitude to his thesis advisor, Dr. J.L. Junkins, for suggesting the dissertation topic, providing invaluable guidance, financial support, encouragement, and many hours of valuable time during his research. Indeed, Dr. Junkins' remarkable personality, especially his general "for fun and profit" spirit, has made on the author, I believe, a lasting impression on the "finer points" of engineering research.

The author would like to extend his sincere appreciation to Dr. S.L. Hendricks, Dr. D.T. Mook, Dr. D. Frederick and Dr. H.F. VanLandingham for serving as members of the committee and providing valuable suggestions.

Thanks are also due to Beverly Shook for her efficient and expert typing and preparation of this manuscript and in particular, her patience in tolerating the author an innumerably large number of changes.

Finally, the author wishes to express warmest appreciation to his wife, and his mother, for their love, patience, understanding and many sacrifices.

DEDICATED TO MY MOTHER AND FOR THEIR LOVE

TABLE OF CONTENTS PAGE

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

LIST OF TABLES

1. INTRODUCTION

1.1 Motivation and Survey of Previous Work

During the last decade, the problem of the control of large flexible spacecraft has been the focus of intense research, as evidenced by [1-5]. Much more so than rigid spacecraft design, Large Space Structures (LSS) design is an extremely interdisciplinary subject drawing on structural and rigid body dynamics, mathematical modelling and approximation of Distributed Parameter Systems (DPS), optimization, estimation, control theory, numerical analysis and large-scale computation [6-8].

Among many complicating issues in the design of LSS has been the high-dimensionality of the LSS model required for sufficient accuracy. For example, a Finite Element Model (FEM), typically from a NASTRAN type of multi-purpose code, often produces a model consisting of several thousand degrees of freedom (DOF), even though one may require accurate convergence to a few of the dominant resulting modes. Approximate order reduction techniques can be used to obtain a lower order model, typically constrained by the maximum order of a solvable Riccati or Lyapunov equation (around 150 states or 75 modes [8]). Most "practical" control laws implemented to date have been based upon much more drastic order reductions (<20 modes), although the high order modes are retained to do verification simulations. However, when a FEM of a LSS or an equivalently complex dynamical system model is to be used for optimization studies, for example with respect to its geometry and/or material properties and/or control parameters, one can readily appreciate the high dimensionality posed by the family of optimization problems.

For some proposed space structures, for instance large optical telescopes [9], it is not difficult to imagine that some prescribed performance requirements may give rise to a scenario where an ad hoc nominal structure design is not "optimally" suited for a controller design or, perhaps, a feasible controller that satisfies a set of apriori desired constraints may not even exist. In light of the above hypothesized scenario, one may wish to investigate a "unified" approach whereby plants and controllers are designed simultaneously to achieve improved performance. Historically, it is suspected that the high dimensionality and similar difficulties of the separate structural design and control design optimization problems individually, may have discouraged research in the area of simultaneous structure and controller optimization. Also the technical breadth required to formulate approaches to simultaneous design optimization has been a significant obstacle. Not surprisingly, only very recently has the "unified" viewpoint attracted attention [10-14].

In [10], a nominal FEM, which is controlled by an optimal quadratic regulator is optimally modified (passive structural stiffness only) to minimize structural weight and improve damping factors of the closed-loop system. The optimization problem above was solved by a nonlinear programming technique. Haftka, et. al. [13], similar to [10], developed a sensitivity procedure to study the effect of small changes in a structure on an optimized control system. They have experimentally verified their proposed sensitivity procedure and concluded that small changes in a structure can indeed have a significant effect on its controlled performance.

To the author's knowledge, the first general optimization formulation to consider simultaneous structure and controller design is given in [11]. Here, a minimum modification strategy in conjunction with a continuation method [15] of handling difficult constraint equations is used for closed-loop eigenvalue optimizations for a direct linear output feedback system. In [12], extension of the technique developed in [11] is given, including "optimal" tuning of weight matrices for the quadradic regulator control law. Both [11] and [12] consider a 21 DOF model (of a rigid hub having four flexible cantilevered appendages) having 42 design gain elements, variable sensor and actuator locations, and variable stiffness, length and tip mass, for a total of 54 design variables. Although the results in [11] demonstrated an algorithm with convergence robustness to moderately high dimensionality and problems associated with nonlinearity, they suggest [11] more research and numerical studies to extend their ideas to truly high dimensioned systems. This dissertation work represents an extension of these ideas with the added elements of optimization using a nonlinear programming method and several significant new performance indices.

The physical problem that mainly motivated the research outlined here is the problem of simultaneous optimal design of structures and controllers of a large control system wherein real world applications would normally be subject to a large number of constraints, mostly inequalities on the design vector space. In this dissertation, we present an algorithm which is believed to be suitable for attacking the previously mentioned class of problems and in particular for computer aided design implementation.

The Simplex Algorithm is a very efficient and reliable method for solving general Linear Programming (LP) problems i.e. optimization problems having linear cost function and constraints. Since its invention in 1947 by Dantzig [16], extensive theoretical studies and applications have been done and today there are literally thousands of references to the Simplex algorithm and LP [17,18]. Consequently, Simplex algorithm software is easily available (although there are not many FORTRAN callable subroutines), reliable, and well founded theoretically. However, naturally occuring linear programming problems are rare in engineering applications and thus LP is not widely taught in traditional engineering curricula; as a result, LP has not played any meaningful role in nonlinear engineering research. In spite of the lack of popularity among dynamics and control engineers, reference [19] recognizes LP as a powerful approach for handling a large number of locally linear constraints and proposed solving nonlinear optimization problems by introducing a sequence of linear optimization problems. In [19], this approach is demonstrated by means of a truss beam FEM for the optimal sizing and placement of active and passive structural members for damping augmentation. Furthermore, a comparison is made between the sequential linear programming approach and more conventional nonlinear optimization approaches; the superiority of the former approach, at least for the class of problems considered here, is clearly demonstrated.

In a successful effort to improve the method as suggested originally in [19], a different formulation is presented in [20] (which involves the addition of a local maximum allowable step size constraint in the linear program). Reference [20] and the present development accomplishes transformation to the standard Linear Program without increasing the dimension of the local linear optimization problem, in contrast to Reference [19]. Furthermore, stability robustness, as measured by a norm of closed-loop eigenvalue sensitivity was minimized [20] with respect to structure/controller design vector while the closed-loop response characteristics were gradually improved by imposing both equality and inequality constraints on the eigenvalue locations. As a different application, in particular for structural redesign problems [21], sequential linear programming in conjunction with a continuation technique is used for a minimum weight design of a cantilever beam FEM subject to constraints in the design vector space and constraints on the structural modes. The above examples demonstrate the versatility and numerical robustness of the proposed algorithm; consistent with the earlier work on continuation methods [15].

We next focus on performance indices that provide some measure of stability robustness of a closed loop system with respect to inevitable "ignorance" (of the equations which perfectly model the actual system). The obvious practical significance of system robustness more than justifies the numerous research publications, especially in the area of control systems design at all levels of mathematical abstraction. In particular, performance and stability robust control of Multi Input Multi Output (MIMO) systems has been an area of intense research, especially in the use of Singular Value (SV) robustness measures as evident in the works of [23-29]. In general, the various robustness criteria proposed fall into two categories: FREQUENCY DOMAIN - perturbations of transfer functions and the use of generalized Nyquist stability criteria (examples: Doyle [23], Safonov [25]); TIME DOMAIN - perturbations of system stability matrix and the use of Lyapunov stability theorem (examples: Patel [30], Yedavalli [31]). As pointed out in [32], the majority of the effort to date has focused on SV to develop analysis tools in the frequency domain and only a relatively smaller amount of effort has been focused on the use of SV for control law synthesis as in [32-35]. The lack of design oriented methodology and research is even more pronounced for time domain measures. There is an urgent need to develop design methodology and the computational algorithms for robust control. The above stated need has motivated some of the research reported here, including development of associated sequential LP and continuation methods for use in developing algorithms for the systematic optimization of the above mentioned time-domain robustness criteria. Included also is a critical examination of the validity of various robustness criteria from the control designer's perspective. A more direct but perhaps less rigorous class of methods for dealing with the robustness problem (of maintaining near optimal performance and stability of a system in the presence of modelling errors/uncertainties/perturbations) is through sensitivity minimization. Howze and Cavin [36] first formulated a "modal insensitivity" condition and suggests an eigenstructure assignment approach for achieving modal insensitivity. It is clear that Howze and

Cavins' formulation was motivated by Moore [70] who in 1976 first clearly identified the freedom offered by state feedback beyond specifying closed-loop eigenvalues. Very recently Raman [37], and Raman and Calise [22], focus on design algorithms for optimal quadratic regulators with modal insensitivities. However, their use of equality constraints on the feedback gain matrix to impose zero modal sensitivity conditions will prove too restrictive for most practical applications of high-dimensioned systems. In the light of the latter problem, Lim and Junkins [20] proposed a direct minimization of the norm of eigenvalue sensitivities.

The favorable preliminary results reported in [20] motivated our search for a more suitable measure of stability robustness that would give directly quantifiable, maximum allowable perturbations and improve the applicability of this approach. Obviously, although both the modal insensitivity and direct minimization approaches do not guarantee stability in the presence of unknown model errors or even prescribed perturbations, they do seek out control configurations having low local stability sensitivity with respect to uncertainties.

In summary, the need for a dependable optimization tool that efficiently handles a large number of nonlinear inequality constraint conditions and a high dimensioned design space is critical, particularly for use in simultaneous design of structures and controllers of large systems. It is further important that nonlinear optimization iterations be done in such a way that convergence failures are informative, i.e., suggest restatements of the problem that would lead to a least compromised solution. In addition, methods are needed for incorporating performance and stability robustness tolerances based on recently developed robustness criteria for design of MIMO control systems.

1.2 Dissertation Outline

The range of issues and problems covered in this dissertation is very broad and is currently being pursued by many researchers. As a result, a significant portion of the research reported here involves careful study, interpretation and extensions of very recent work.

In chapter 2, we present the most specific and immediately useful contribution of this dissertation, namely, a new optimization algorithm involving sequential LP and a continuation method for attacking a class of optimization problems. The class of problems includes, but certainly is not restricted to the structure/controller optimization problems at the focus of this dissertation. With this tool in hand, we discuss the significance and the formulation of a simultaneous structure and controller design/optimization approach. This is followed by numerical examples involving the design of direct output feedback controllers by pole placement with both fixed and variable plants to demonstrate the methodology, its generality and its versatility.

In chapter 3, several issues associated with modal sensitivity are presented. First, a review of the recent and current literature on eigenstructure assignment and modal insensitivity theories is presented. This is followed by a discussion of a minimum eigenvalue sensitivity design approach involving the direct minimization of the magnitude of the eigenvalue sensitivities. In this context, an interpretation of the eigenvalue sensitivity matrix is given in terms of an operator norm. Furthermore, the well-established concepts of conditioning of eigenvalue problems are reviewed and applied to derive a criterion which guarantees asymptotic stability in a feedback controls context. The chapter ends with a re-examination of the general, non-self-adjoint eigenvalue problem and presents several new insights on bi-orthogonality and normalizations, and their role in correctly deriving eigenvector derivatives which are either incorrect, incomplete and/or misleading in the current literature.

Chapter 4 explores recently proposed stability robustness criteria in the time domain. It begins with a review of the Lyapunov stability theorem for time-invariant, discrete linear systems and then a robustness condition due to Patel and Toda is derived in a novel way. A new approach utilizing modal coordinates is presented for the derivation of a robustness measure which turns out to be directly related to the condition number of the corresponding eigenvalue problem. In addition, this chapter examines the stability robustness of optimal linear quadratic full state feedback controllers. The dependence of robustness measures on the dominant eigenvalues of the closed loop system and the weight matrices of the quadratic cost function is highlighted including its significance to pole placement problems.

Chapter 5 focuses on applications. A 20th-order dynamical system representing a free-free flexible beam with an attached rigid-body is considered. Formulations and comparisons of three designs, namely, designs of minimum sensitivity, minimum mass and maximum robustness, are presented. The design variables number 55 and includes 4 structural parameters, 3 actuator locations, and 48 control gain elements. In addition, 16 equality constraints on eigenvalues and 77 inequality constraints are used in the examples. The successful optimizations for this example provide an excellent basis for optimism vis-a-vis practical applications.

Finally, chapter 6 presents a summary of lessons learned and several concluding remarks.

2. A UNIFIED APPROACH TO STRUCTURE AND CONTROLLER DESIGN OPTIMIZATIONS

2.1 Introduction

In the conventional approach [10], structural design and structural control system design are essentially uncoupled as the interaction between the structures and controls designers has been very minimal in the past. Typically, the structural designer establishes a nominal structural design based on strength and stiffness requirements obtained from anticipated peak maneuver loads during the expected (or typical) operation of the control system. The structural engineer's concern is to design a lightweight structure that will satisfy the strength, stiffness and other performance requirements and traditionally, the designer of the control system has little input in the evolution of the basic structural design. An unfortunate consequence of the artificially uncoupled nature of the two disciplines may very well be difficulty in achieving desired closed-loop performance, this is especially true for marginally feasible missions. There does not presently exist systematic methods for modifying the nominal structure to enhance the performance of the controlled structure.

In this chapter, we present a general method for attacking a class of optimization problems frequently encountered in the design of structures and controllers. We begin by discussing, along the lines of [11], the parameterization of the structure and control system and highlighting the use of a global parameter vector which is comprised of both structural and controller parameters. We then focus on a conceptually simple (but nevertheless powerful!) method for handling constraints which overcomes several nontrivial problems that must be confronted in nonlinear optimization problems. The above method is commonly referred to as a continuation or homotopy method and has been implemented successfully in several problems as documented in [11,12,14,15,19-21]. Next, we reformulate the general nonlinear programming problem in terms of a sequence of linear programming problems which are solved by the Simplex algorithm. Finally we conclude by presenting a few numerical examples to fix ideas.

2.2 Parameterizations of Structures and Controllers

In this section, three standard control configurations are described which are used in the sequel to demonstrate the methodologies and algorithms developed in this dissertation for the design and optimization of structural and control systems.

A most fundamental and necessary step in any systems analysis and design involves the modelling and parameterizations of mathematical models of physical systems. This stage, generally referred to as "conceptual design", is usually followed by tuning of the design parameters (parameter optimization) to achieve apriori defined optimum conditions. In particular, we consider here the parameterization of structures and controllers for their simultaneous design in state space.

We consider the general second order linear system, often found in structural dynamics for describing structural vibrations

$$M\ddot{\xi}(t) + C\dot{\xi}(t) + K\xi(t) = Du(t) \qquad (2.2.1)$$

where M, C, and K represents the mass, damping and stiffness matrices respectively and D represents the control force distribution matrix. Eq. (2.2.1) can be re-written in the state-space form

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{2.2.2}$$

where

$$x(t) = \begin{Bmatrix} \xi(t) \\ \dot{\xi}(t) \end{Bmatrix} \quad ; \quad A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix} \quad ; \quad B = \begin{Bmatrix} 0 \\ M^{-1}D \end{Bmatrix}$$

We also consider here output measurements of the form

$$y(t) = Hx(t) \tag{2.2.3}$$

where y and H represents the output vector and measurement matrix respectively.

2.2.1 Direct Output Feedback

The simplest output feedback controller takes the linear form

$$u(t) = G\, y(t) \tag{2.2.4}$$

where G is a constant gain matrix. We see that by substituting Eqs. (2.2.3) and (2.2.4) into Eq. (2.2.2), we obtain the closed loop system

$$\dot{x}(t) = \bar{A}\, x(t) \tag{2.2.5}$$

where the closed loop system matrix is

$$\bar{A} = A + B\, G\, H\ . \tag{2.2.6}$$

The matrices introduced above can be assumed to be parameterized as

$$M = M(a) \qquad D = D(b)$$
$$C = C(a) \qquad H = H(c)$$
$$K = K(a) \qquad G = G(g)$$

where

a - structural design vector

b - actuator design/location vector

c - sensor design/location vector

g - control gain vector

The above structure and controller design parameters can be combined to form a global parameter vector

$$p = \mathrm{col}(a, b, c, g)$$

and closed loop system matrix can be written as

$$\bar{A}(p) = A(a) + B(a,b)\, G(g)\, H(c)\ .$$

Although it is obvious that open loop characteristics influence the closed loop characteristics, the design philosophy suggested here emphasizes the fact that open loop parameters could be tuned in precisely the same way control parameters are tuned and perhaps simultaneously for best results.

2.2.2 Dynamic Compensation Feedback

A more general output feedback controller is the dynamic compensator [44] of order s

$$u(t) = G_1 y(t) + G_2 z(t) \tag{2.2.7}$$

where

$$\dot{z}(t) = F_1 z(t) + F_2 y(t) \tag{2.2.8}$$

with the initial state assumed as $z(0) = 0$. Since the dynamic compensator state $z(t)$ can be obtained in principle by integrating Eq. (2.2.8), the control law of Eq. (2.2.7) can be seen as containing information on both current and past output signals.

By combining Eqs. (2.2.7) and (2.2.8) with Eqs. (2.2.2) and (2.2.3), we can write the closed loop system in the augmented form

$$\dot{q}(t) = \tilde{A} q(t) \tag{2.2.9}$$

where

$$\tilde{A} = \begin{bmatrix} A + B\,G_1 H & B\,G_2 \\ F_2 H & F_1 \end{bmatrix} \tag{2.2.10}$$

$$q(t) = \begin{Bmatrix} x(t) \\ z(t) \end{Bmatrix}$$

As evident from Eq. (2.2.10), the designer has under his control, in addition to parameters associated with direct output feedback controllers, the parameters

$$d = \mathrm{col}(d_1, d_2, d_3)$$

where

$$G_2 = G_2(d_1)$$

$$F_1 = F_1(d_2)$$

$$F_2 = F_2(d_3)$$

A main drawback of using dynamic compensators involves judicious selection of the augmented design parameters to maintain asymptotic stability of both the dynamic compensation states in addition to the original closed loop system. The above complication and an increase in the order of the system and the design vector space is the price paid for a more versatile set of closed loop characteristics. Through the dynamic compensator's "memory", one can achieve frequency shaping, in addition to damping of the controlled response.

2.2.3 Linear Quadratic Regulator (LQR) Output Feedback

This approach to feedback controller design is motivated by the simplicity of generating control commands directly by linear combinations of the available output instead of first reconstructing an estimate of the full state via a Kalman Filter [47] or state observer [48]. A derivation of the necessary conditions and the corresponding output feedback gains for optimum LQR using output feedback is given by Levine and Athans [49]. However, a more concise derivation with identical results is given by Mendel [50] and will be followed here. The problem can be stated as follows

$$\text{MINIMIZE} \qquad J = \frac{1}{2} \int_0^\infty [x^T(t)\ Qx(t) + u^T(t)\ Ru(t)]\ dt \qquad (2.2.11)$$

$$\text{SUBJECT TO} \qquad \dot{x}(t) = A\ x(t) + B\ u(t); \quad x(0) = x_0 \qquad (2.2.12)$$

$$y(t) = H\ x(t) \qquad (2.2.13)$$

where the controller is constrained to the form

$$u(t) = G\ y(t) \qquad (2.2.14)$$

and

$$Q \geq 0\ , \quad R > 0\ .$$

To convert the above problem to a static optimization problem (i.e. time dependent trajectory is absent), we use the following results (whose proof is given in [40]).

THEOREM 2.2.1

If the system

$$\dot{x}(t) = A\ x(t); \quad x(0) = x_0$$

is asymptotically stable, then

$$\frac{1}{2} \int_0^\infty x^T(t)\ R\ x(t)\ dt = \frac{1}{2}\ x^T{}_0 K x_0$$

where K satisfies the Lyapunov equation

$$A^TK + KA = -R \quad \text{and } R \geq 0.$$

To use the above theorem, we substitute Eqs. (2.2.13) and (2.2.14) into Eq. (2.2.11) to obtain

$$\text{MINIMIZE} \qquad J = \frac{1}{2}\int_0^\infty x^T(t)\, \bar{R}\, x(t)\, dt \tag{2.2.15}$$

$$\text{SUBJECT TO} \qquad \dot{x}(t) = \bar{A}\, x(t); \quad x(0) = x_0 \tag{2.2.16}$$

$$\bar{A} = A + BGH$$
$$\bar{R} = Q + H^TG^TRGH\ .$$

By applying Theorem 2.2.1 to Eqs. (2.2.15) and (2.2.16), we transform the above LQR problem to a constrained static optimization problem of the form

$$\text{MINIMIZE} \qquad J = \frac{1}{2}x^T{}_0Kx_0 = \frac{1}{2}\,\mathrm{trace}(Kx_0x_0^T) \tag{2.2.17}$$

$$\text{SUBJECT TO} \qquad \bar{A}^TK + K\bar{A} = -\bar{R} \tag{2.2.18}$$

$$\bar{A} = A + BGH$$
$$\bar{R} = Q + H^TG^TRGH\ .$$

We note that $\bar{R}$ is guaranteed to be positive semi definite since $Q \geq 0$ and $R > 0$ are initially assumed. From Eq. (2.2.17), we see the dependence of the the optimization problem on a given initial state. To eliminate this dependence on a specific initial state, and to account for the fact that an infinity of initial states are possible, the standard procedure is to take an average of the performance obtained for a linearly independent set of initial states. In particular, the initial state is assumed to be a vector of independent random variables belonging to distributions having equal (unit) variances, then

$$E[x_0 x_0^T] = I \tag{2.2.19}$$

We note that the equal variance assumption can always be made if we introduce an appropriate re-normalization of x. Under these assumptions, the modified cost function is

$$\begin{aligned}\bar{J} = E[J] &= E[\tfrac{1}{2}\ \text{trace}\ \{Kx_0x_0^T\}] \\ &= \tfrac{1}{2}\ \text{trace}\ \{K\ E[x_0x_0^T\ ]\} = \tfrac{1}{2}\ \text{trace}\ \{K\}\quad .\end{aligned}$$

The resulting static optimization problem can be written as

$$\text{MINIMIZE}\qquad \bar{J} = \tfrac{1}{2}\ \text{trace}\{K\} \tag{2.2.20}$$

$$\text{SUBJECT TO}\qquad \bar{A}^TK + K\bar{A} = -\bar{R} \tag{2.2.21}$$

$$\bar{A} = A + BGH$$
$$\bar{R} = Q + H^TG^TRGH$$

$$\text{Re}\ \lambda_i[\bar{A}] \leq -\alpha_i \quad ; \quad i=1,\ldots,n$$

where α_i represents the stability margin of i-th mode. Note that Eq. (2.2.19) is not overly restrictive; we can re-scale the components of x to make the unit, equal variance assumption appropriate.

The necessary conditions for the existence of an optimal output feedback solution can be obtained by introducing a Lagrange Multiplier matrix L and setting the gradients of the Lagrangian with respect to G, L and K to zero to obtain [49]

$$G = -R^{-1}B^TKLH^T(HLH^T)^{-1} \tag{2.2.22}$$

$$(A + BGH)L + L(A + BGH)^T + I = 0 \tag{2.2.23}$$

$$(A + BGH)^TK + K(A + BGH) + Q + H^TG^TRGH = 0 \tag{2.2.24}$$

From the above necessary conditions we can obtain the important special case when m = n (no. of outputs = order of system) and assuming the measurement matrix, H to be invertible, Eq. (2.2.22) becomes

$$G = -R^{-1}B^TKH^{-1}$$

and Eq. (2.2.24) reduces to

$$A^TK + KA - KBR^{-1}B^TK + Q = 0 \tag{2.2.25}$$

which is the standard Algebraic Riccati Equation (ARE) representing the necessary and sufficient conditions for optimality of full state feedback LQR. The above result is indeed expected since the assumption that the inverse of H exists implies that the full state is recoverable from the measured outputs and hence the initial output feedback law specializes to the full state feedback law.

It is important to note that Eqs. (2.2.22) to (2.2.24) represents only necessary conditions for optimal LQR output feedback. In addition, controllability and observability [57] conditions which are sufficient for the existence of Riccati equation solution will not guarantee stability for the output feedback controller. The above reasons explains the relative difficulty in designing numerical algorithms for computing optimal output feedback gains.

We observe from Eqs. (2.2.22) to (2.2.24) that in principle, the plant matrices, A, B, and H, the control gain matrix, G, the Lagrange multiplier matrix, L, and the quadratic weight matrices, Q and R can all be tuned simultaneously to satisfy various design constraints. Although currently no algorithms exist that allow all the above matrices to be varied simultaneously, there exists important special cases. In particular, reference [76] presents an iterative algorithm to solve for G, K and L matrices from Eqs. (2.2.22) to (2.2.24). It can be seen that this algorithm [76] can easily accommodate additional weight matrices to increase flexibility. On the other hand, reference [14] presents a robust algorithm for satisfying eigenvalue constraints by tuning a set of generalized quadratic weight matrices of optimal linear quadratic regulators with full state feedback while keeping all other parameters fixed.

2.3 A Continuation Approach for Imposing Constraints in Nonlinear Optimization

We consider all sets of constraint vectors that can be reduced to the forms $f(p) \le f^o$, $f(p) \ge f^o$, $f(p) = f^o$ and we indicate all three possibilities by the following notation

$$f(p) \quad \{ \leq, =, \geq \} \; f^0 \tag{2.3.1}$$

where f^0 denotes specified objectives and $f(p)$ represents constraint functions whose dependence on parameter p is assumed known and "well-behaved" (continuous and differentiable). For the class of problems including simultaneous structure and controller design, the set of constraints represented by Eq. (2.3.1) may include constraints on eigenvalues and eigenvectors, total structural weight, closed-loop stability requirements, structural geometry constraints, stability and performance robustness with respect to plant uncertainty requirements and direct upper and lower bounds on design variables. The complex nature of the constraints leads to various problems in the context of mathematical programming. First, the constraint objectives may be such that no solution exists, much less an optimal solution. Secondly, for nonlinear problems, in spite of a consistent set of constraints, it may be difficult to locate a starting or initial feasible solution. Indeed, an important element of engineering design is a satisfactory tradeoff between conflicting requirements. However, the above requires the designer to know the relative restrictiveness of individual constraints. The continuation method [15] resolves, at least to a significant degree, the above problems by (1) seeking out at least a feasible solution to a neighbor of the original problem (if indeed a feasible solution to the originally stated problem does not exist) and, (2) providing "arbitrarily good" initial guesses by starting each iteration with a neighboring converged solution.

The continuation method essentially involves replacing a subset of the original constraint objectives, f^0, typically consisting of compromisable and/or demanding constraints, by a sequential neighboring set of constraint objectives, $F(\gamma_i)$, where

$$F(\gamma_i) = (1 - \gamma_i)\, f(p^S) + \gamma_i f^0 \tag{2.3.2}$$

where p^S is an arbitrarily chosen starting design vector and γ_i is the scalar continuation parameter satisfying

$$0 = \gamma_0 \le \gamma_1 \le \gamma_2 \le \dots \le \gamma_N = 1 \quad . \tag{2.3.3}$$

For redesign problems, p^S is naturally taken to be the nominal system values [21]. The above convex combination of starting constraint and final constraint objectives as shown in Eq. (2.3.2) reveals the following facts:

$$\text{at} \quad \gamma = \gamma_0 = 0 \quad , \quad F(0) = f(p^S)$$
$$\text{at} \quad \gamma = \gamma_N = 1 \quad , \quad F(1) = f^0$$

i.e. if convergence is achieved at $\gamma = 1$, we recover the original objective constraint condition. It is also obvious but nevertheless important to note that if convergence is achieved at $\gamma = 1$ it will be done in N steps. In most nonlinear problems where linearizations about current values are assumed, of course, the step size of $\Delta\gamma$ can be chosen to validate the assumptions.

Finally, and perhaps most importantly, failure to reach the final ($\gamma = 1$) solution is usually softened by convergence to a neighboring solution. The active constraint set and gradient information of the final convergence provides a basis for intelligent revisions of the problem statement.

2.4 Optimization via Sequential Linear Programming

The following features motivate the use of optimization using Sequential Linear Programming (SLP) approach for the solution of mathematical programming formulations of simultaneous structure and controller design problems:

i) many design problems involve a large number of constraints, mostly inequalities,

ii) efficient and reliable computer codes using the Simplex algorithm are easily available,

iii) problem formulations are direct and simple, especially the constraint equations.

Perhaps further justification for our emphasis on the SLP approach can be attributed to its ease of use and competitiveness with respect to other leading optimization algorithms. Recently, Palacios-Gomez, et.al. [45,46] have demonstrated that nonlinear optimization by SLP compares favorably with other established methods such as the Generalized Reduced Gradient (GRG) method. In addition, they point out that at present, disagreement exists in the literature on the issue of which algorithms are superior and emphasize that performance is strongly dependent on the strategies used, the computer implementation and on parameter values used in the algorithm. Their work also includes a discussion of various strategies for implementation of SLP.

Let us consider the general nonlinear programming problem

$$\text{MAXIMIZE} \quad J(p) \tag{2.4.1}$$

$$\text{SUBJECT TO} \quad f(p) \;\{\leq, =, \geq\}\; f^0 \tag{2.4.2}$$

where p is a vector of design parameters, f(p) represents a given set of constraint functions, f^0 represents a set of desired constraint objectives and J represents the cost function. To transform the problem stated above to a sequential linear program, we simply linearize the equations about current parameters, p^c, to obtain

$$\text{MAXIMIZE} \quad \sum_{j=1}^{np} \left[\frac{\partial J}{\partial p_j}\bigg|_{p^c}\right]\Delta p_j + O(\Delta p^2) \tag{2.4.3}$$

$$\text{SUBJECT TO} \quad f(p^c) + \sum_{j=1}^{np} \left[\frac{\partial f}{\partial p_j}\bigg|_{p^c}\right]\Delta p_j + O(\Delta p^2) \tag{2.4.4}$$

$$\{\leq, =, \geq\}\; f^0$$

where $\quad p = p^c + \Delta p$

and "np" represents the number of design parameters. For the problem as stated in Eqs. (2.4.3) and (2.4.4), we point out that for a small class of problems where contraints of Eq. (2.4.4) generate a feasible region (the subspace of Δp where Eq. (2.2.4) is satisfied) and where all linearizations about a current point are reasonably valid, the problem just formulated may be solved iteratively using the Simplex algorithm until some type of numerical convergence is achieved. However, for general nonlinear problems (where initial guesses close to the optimum are not available and linearization assumptions about a nominal point do not hold over all of the feasible region), additional restrictions are needed for its numerical solution. Indeed, a demanding set of constraints may occasionally correspond to an inconsistent set, or equivalently, no solution may exist.

For the above reasons, we introduce constraints on the maximum and minimum allowable parameter corrections locally, i.e.

$$-\varepsilon \le \Delta p \le \varepsilon \tag{2.4.5}$$

and apply the continuation method of handling potentially difficult and compromisable constraints as described in Section 2.3. Equation (2.4.5) applies element-by-element, since Δp and ε are vectors. All elements of ε are assumed to be positive. Hence, by replacing the particular constraint objectives in Eq. (2.4.4) by a sequential neighboring set of constraint objectives, $F(\gamma_i)$ as given by Eq. (2.3.2), we rewrite the modified linear program at continuation step "i" as

$$\text{MAXIMIZE} \quad \sum_{j=1}^{np} \left[\frac{\partial J}{\partial p_j}\bigg|_{p^{i-1}}\right] \Delta p_j \tag{2.4.6}$$

$$\text{SUBJECT TO} \quad \sum_{j=1}^{np} \left[\frac{\partial f}{\partial p_j}\bigg|_{p^{i-1}}\right] \Delta p_j \quad \{\le, =, \ge\} \tag{2.4.7}$$

$$(1 - \gamma_i)\ f(p^s) + \gamma_i f^0 - f(p^{i-1})$$

$$-\varepsilon \le \Delta p \le \varepsilon \tag{2.4.8}$$

$$p^i = p^{i-1} + \Delta p \ ; \ p^0 \equiv p^s \quad i=1,\dots,N \tag{2.4.9}$$

We note here that the current parameter p^c has been replaced by p^{i-1} and p^0, (the parameter vector at step "0"), by an arbitrary starting parameter vector, p^s. It should also be pointed out here that the current step size bound, ε , may be adjusted judiciously to accommodate a tradeoff between satisfying local linearity assumption and permitting a solution to the neighbor constraint which depends on the constraint step change of $(\gamma_i - \gamma_{i-1})$. Some degree of numerical experimentation and artwork is usually necessary for an efficient implementation of the above method.

We next introduce a translational transformation of the form

$$y = \Delta p + \varepsilon \quad . \tag{2.4.10}$$

In terms of the nonnegative coordinate, y, we can rewrite the linear program of Eqs. (2.4.6) to (2.4.9) at step "i" as

$$\text{MAXIMIZE} \quad \sum_{j=1}^{np} \left[\frac{\partial J}{\partial p_j}\bigg|_{p^{i-1}}\right] y_j \tag{2.4.11}$$

$$\text{SUBJECT TO} \sum_{j=1}^{np} \left[\frac{\partial f}{\partial p_j}\Big|_{p^{i-1}}\right] y_j \ \{\le, =, \ge\} \ (1-\gamma_i)f(p^s) + \gamma_i f^o - f(p^{i-1}) + \sum_{j=1}^{np}\left[\frac{\partial f}{\partial p_j}\Big|_{p^{i-1}}\right]\varepsilon_j \qquad (2.4.12)$$

$$y \le 2\varepsilon \qquad (2.4.13)$$

$$p^i = p^{i-1} + y - \varepsilon \ ; \quad p^0 = p^s \qquad i=1,\dots,N \qquad (2.4.14)$$

y is nonnegative.

By adding "slack" and/or "surplus" variables which are themselves nonnegative, we see that the linear program above can be easily transformed to the "standard linear program" [17] form of

$$\text{MAXIMIZE} \quad C^T x \qquad (2.4.15)$$

$$\text{SUBJECT TO} \quad Ax = b \qquad (2.4.16)$$

x is nonnegative.

where x represents the augmented parameter vector consisting of y and the slack and surplus variables. The above linear program can be solved using standard Simplex codes.

2.5 Design Example: Mininum Sensitivity Eigenvalue Placement Using Output Feedback

2.5.1 Problem Formulation

In this section, we show details the of an example simultaneous design optimization problem. We seek to simultaneously assign closed loop eigenvalues and minimize eigenvalue sensitivities with respect to selected plant parameters (subject to various constraints on controller gains and stiffness elements). An idealized, simple structural model consisting of three lumped masses, springs and dashpots is chosen here to demonstrate the ideas introduced previously. A flow chart of the optimization strategy is given in Figure 2.1.

Before we begin detailed discussion of this example, we reiterate the practical feature of the algorithm that inequality constraints (which appear very frequently in practical design situations) are handled with ease. For contrast and comparison, one major issue which arises when using the Linear Quadratic Regulator (LQR) formulation in regulator designs is to penalize the indiscriminate use of control energy with respect to state error without explicitly introducing inequality constraints on controllers. Inequality constraints (control saturation bounds, for example) are typically enforced (in LQR designs) indirectly by ad hoc variations in the weight matrices. The present approach also imposes actuator saturation bounds indirectly, through inequality constraints on control gains. The following derivation closely parallels our original work in [20].

Let us consider the discrete mass-spring-dashpot system as shown in Figure 2.2. Two actuators at DOF's 1 and 3 and displacement and velocity sensors at DOF's 2 and 3 are assumed available for output feedback. The system equations are of the form given in Eqs. (2.2.2) and (2.2.3) where

$$x = \begin{Bmatrix} \xi \\ \dot{\xi} \end{Bmatrix} \qquad \text{(6x1) vector}$$

mass matrix,

$$M = \mathrm{diag}\{m_1,\ m_2, m_3\} = \mathrm{diag}\{10,\ 10,\ 25\}$$

damping matrix,

$$C = .02 \times \bar{K}$$

where $\bar{K}$ represents a constant stiffness matrix, chosen in this example as the starting value of the general stiffness matrix,

$$K = \begin{bmatrix} k_1+k_2 & -k_2 & 0 \\ -k_2 & k_2+k_3 & -k_3 \\ 0 & -k_3 & k_3 \end{bmatrix}$$

force distribution matrix,

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

measurement matrix,

$$H = \left[\begin{array}{c:c} \begin{matrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} & 0 \\ \hdashline 0 & \begin{matrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \end{array}\right]$$

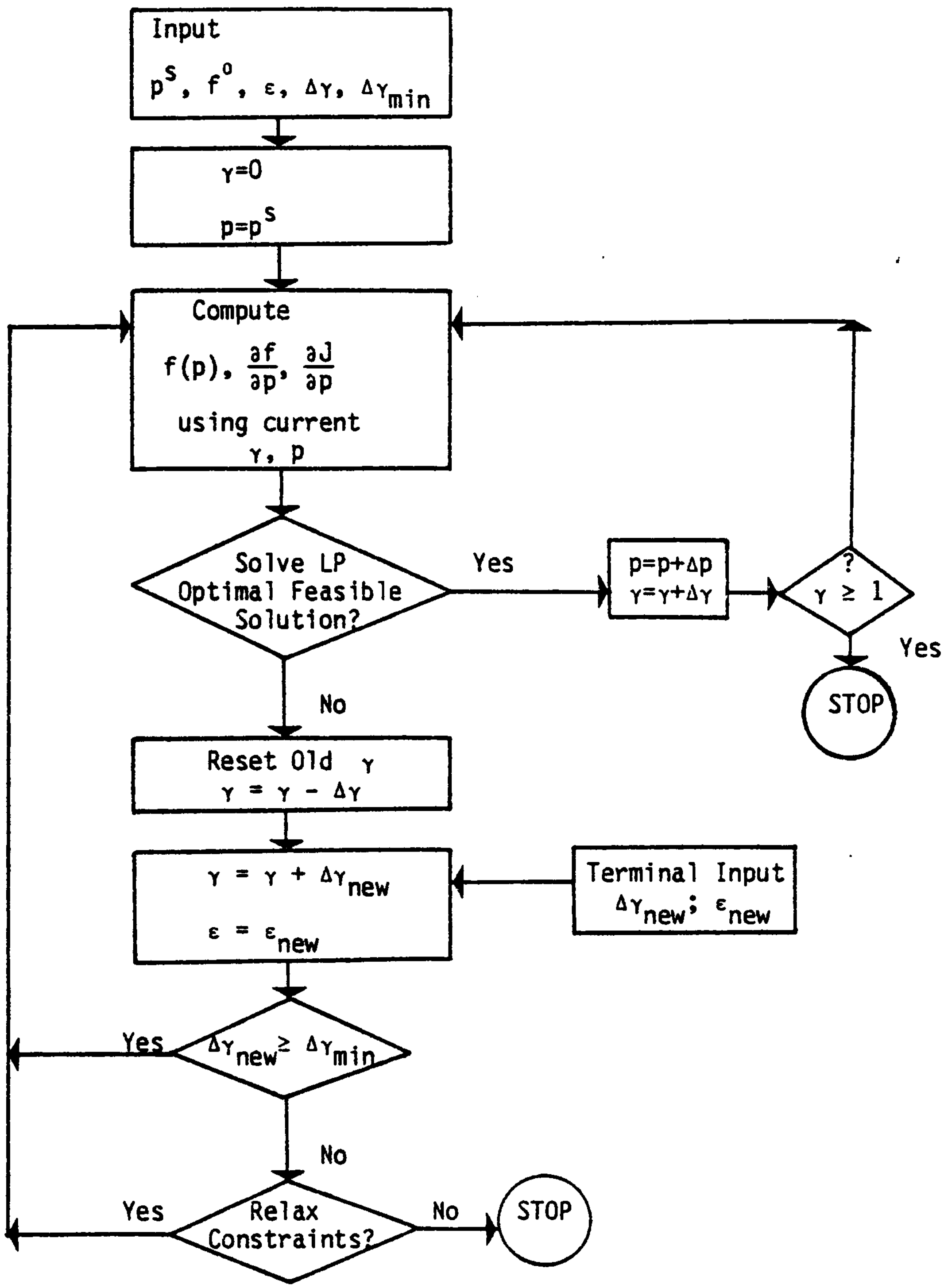


Figure 2.1 A Flow Chart Illustrating the Use of SLP and Continuation Method

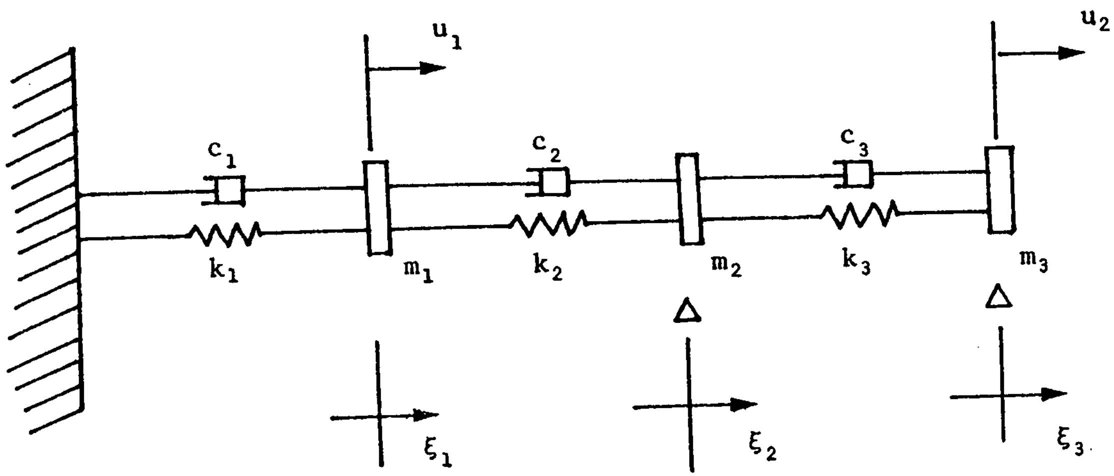


Figure 2.2 A 3-DOF Mass-Spring-Dashpot System

△ displacement and velocity sensors

and output gain matrix,

$$G = \begin{bmatrix} g_1 & g_2 & g_3 & g_4 \\ g_5 & g_6 & g_7 & g_8 \end{bmatrix} .$$

We next consider a direct measure of the sensitivity of closed-loop eigenvalues with respect to selected parameters, more specifically, with respect to three lumped masses. Let us then define the eigenvalue sensitivity measure (with respect to the mass properties) as

$$S \triangleq \sum_{j=1}^{3} \sum_{i=1}^{2n} \left| \frac{\partial \lambda_i}{\partial m_j} \right|^2 w_{ij} \tag{2.5.1}$$

where

$\dfrac{\partial \lambda_i}{\partial m_j}$ = i-th eigenvalue derivative with respect to j-th lumped mass

w_{ij} = weighting factor associated with the i-th eigenvalue sensitivity with respect to j-th lumped mass

Further assume that we are given the set of design variables, p_k ($k=1,\ldots,np$), which may or may not include the lumped masses. We linearly expand the sensitivity cost function (as shown in [20]) to obtain

$$S(p) = S(p^c) + 4 \sum_{k=1}^{np} S_k(p^c) \, \Delta p_k + \ldots \tag{2.5.2}$$

where

$$S_k(p^c) = \mathrm{Re}[\sum_{j=1}^{3} \sum_{i=1}^{n} \frac{\partial \lambda_i^*}{\partial m_j} \frac{\partial^2 \lambda_i}{\partial p_k \partial m_j} w_{ij}]_{p^c} \tag{2.5.3}$$

$$p = p^c + \Delta p$$

and "np" and "*" represents the number of design parameters and complex conjugate respectively. At a current design configuration, p^c, it is obvious from Eq. (2.5.2) that we need to minimize only

$$J(\Delta p) = \sum_{k=1}^{np} S_k(p^c)\, \Delta p_k \tag{2.5.4}$$

where

$$\Delta p = (\Delta p_1, \ldots, \Delta p_{np})^T$$

and S_k is given by Eq. (2.5.3). The first and second eigenvalue partial derivatives are required in Eq. (2.5.3) and are derived in detail in reference [67] and are repeated here for convenience

$$\frac{\partial \lambda_i}{\partial m_j} = \psi_i^T \frac{\partial \bar{A}}{\partial m_j} \phi_i \tag{2.5.5}$$

$$\frac{\partial^2 \lambda_i}{\partial p_k \partial m_j} = \psi_i^T \frac{\partial^2 \bar{A}}{\partial p_k \partial m_j} \phi_i +$$

$$\sum_{\substack{\ell=1 \\ \ell \neq i}}^{2n} \frac{\psi_i^T \frac{\partial \bar{A}}{\partial m_j} \phi_\ell \psi_\ell^T \frac{\partial \bar{A}}{\partial p_k} \phi_i + \psi_\ell^T \frac{\partial \bar{A}}{\partial m_j} \phi_i \psi_i^T \frac{\partial \bar{A}}{\partial p_k} \phi_\ell}{(\lambda_i - \lambda_\ell)} \tag{2.5.6}$$

where from Eq. (2.2.6)

$$\bar{A} = A + BGH$$

and ϕ, ψ , represents the right and left eigenvectors of the closed loop system. More generally, of course, one can include any subset of p (in addition to m_i) in the sensitivity measure of Eq. (2.5.1). Intuitively, one might select (and weight) the parameters depending upon apriori estimates of system parameter uncertainty. This gives rise to a "structured" sensitivity measure. In evaluating the final summation of Eq. (2.5.6), it is important to first evaluate the four quadratic products (such as $\psi_i^T \frac{\partial \bar{A}}{\partial m_j} \phi_\ell$), otherwise the operation count increases by orders of magnitudes.

Let us consider next the constraints on closed loop eigenvalues associated with their assignments. In particular, we seek to drive, if possible, the damping factors such that they exceed certain desirable values. In other words, we desire

$$\zeta_i(p) \geq \zeta_i^0 \quad ; \quad i=1,\ldots,nm$$

or

$$\zeta(p) \geq \zeta^0 \tag{2.5.7}$$

where ζ^0 denotes the objective (desired) damping factors and "nm" denotes the number of modes. In addition to damping constraints, we seek to drive damped natural frequencies ω_d to desired values. Symbolically, we can achieve the above by imposing in addition, the equality constraints

$$\omega_{d_i}(p) = \omega_{d_i}^{o} \quad ; \quad i=1,\ldots,nm$$

or

$$\omega_d(p) = \omega_d^{o} \qquad (2.5.8)$$

where ω_d^o represents objective (desired) damped frequencies. It should be pointed out that the constraints in Eq. (2.5.8) serves a dual purpose. They represent a second set of constraints on the closed loop eigenvalues. Clearly the constraints of Eqs. (2.5.7) and (2.5.8) may also be used to control eigenvalue trajectories and hence avoid the trajectories of all controlled eigenvalues from crossing each other or the real axis which may cause bifurcations or other unnecessary complications. It is obvious that the continuation form of the constraints as given by Eq. (2.3.2) when judiciously applied to Eq. (2.5.8) will lead to nearly parallel trajectories which are easily tractable.

The eigenvalue constraints of Eqs. (2.5.7) and (2.5.8) can be linearized about current parameters, p^c as

$$\sum_{k=1}^{np} \left[\frac{\partial \zeta}{\partial p_k} \bigg|_{p^c} \right] \Delta p_k \geq \zeta^{o} - \zeta \bigg|_{p^c} \qquad (2.5.9)$$

$$\sum_{k=1}^{np} \left[\frac{\partial \omega_d}{\partial p_k} \bigg|_{p^c} \right] \Delta p_k = \omega_d^{o} - \omega_d^{o} \bigg|_{p^c} \qquad (2.5.10)$$

where from (A.6) and (A.7), the sensitivities take the forms

$$\frac{\partial \zeta_i}{\partial p_k} = \frac{\mathrm{Im}\{\lambda_i\}}{|\lambda_i|^3} \; [\mathrm{Re}\{\lambda_i\} \; \mathrm{Im}\{\frac{\partial \lambda_i}{\partial p_k}\} - \mathrm{Im}\{\lambda_i\} \; \mathrm{Re} \; \{\frac{\partial \lambda_i}{\partial p_k}\} \;]$$

$$\frac{\partial \omega_{d_i}}{\partial p_k} = \mathrm{Im} \; \{\frac{\partial \lambda_i}{\partial p_k}\} \quad .$$

In addition to the above eigenvalue placement constraints, we allow here constraints involving explicit bounds on plant parameters and gain elements. Indeed, we may even include constraints associated with location of sensors and actuators, relative magnitudes of eigenvector elements or even stability robustness requirements. In short, we admit all equality and inequality constraints (other than eigenvalue constraints) which are expressible in the algebraic forms

$$z_E(p) = z_E^0$$

$$z_{IE}(p) \leq z_{IE}^0$$

or in their linearized forms

$$\sum_{k=1}^{np} [\frac{\partial z_E}{\partial p_k} \Big|_{p^c}] \; \Delta p_k = z_E^0 - z_E \Big|_{p^c} \tag{2.5.11}$$

$$\sum_{k=1}^{np} [\frac{\partial z_{IE}}{\partial p_k} \Big|_{p^c}] \; \Delta p_k \leq z_{IE}^0 - z_{IE}\Big|_{p^c} \tag{2.5.12}$$

For reasons discussed earlier in section 2.3, the continuation equation of Eq. (2.3.2) is applied to selected constraint objectives appearing in Eqs. (2.5.9) to (2.5.12). If all the above objectives are replaced by a sequential set of objectives (each convergence will provide starting estimates for the next step), the resulting linearized constraints can be written as

$$\sum_{k=1}^{np} [\frac{\partial \omega_d}{\partial p_k}\Big|_{p^{i-1}}] \Delta p_k = (1 - \gamma_i)\omega_d^s + \gamma_i\omega_d^o - \omega_d\Big|_{p^{i-1}} \tag{2.5.13a}$$

$$\sum_{k=1}^{np} [\frac{\partial \zeta}{\partial p_k}\Big|_{p^{i-1}}] \Delta p_k \geq (1-\gamma_i)\zeta^s + \gamma_i\zeta^o - \zeta\Big|_{p^{i-1}} \tag{2.5.13b}$$

$$\sum_{k=1}^{np} [\frac{\partial z_E}{\partial p_k}\Big|_{p^{i-1}}] \Delta p_k = (1 - \gamma_i)z_E^s + \gamma_i z_E^o - z_E\Big|_{p^{i-1}} \tag{2.5.13c}$$

$$\sum_{k=1}^{np} [\frac{\partial z_{IE}}{\partial p_k}\Big|_{p^{i-1}}] \Delta p_k \leq (1 - \gamma_i)z_{IE}^s + \gamma_i z_{IE}^o - z_{IE}\Big|_{p^{i-1}} \tag{2.5.13d}$$

where

$$p^i = p^{i-1} + \Delta p$$

and superscript "i" represents the i-th continuation step. What remains is the imposition of local step size bounds as given in Eq. (2.4.5) and the transformation to nonnegative coordinates through the use of Eq. (2.4.10). In summary, the linear program to solve at step "i" can be written in nonnegative coordinates as follows:

$$\text{MAXIMIZE} \quad \sum_{k=1}^{np} - S_k\Big|_{p^{i-1}} y_k \tag{2.5.14}$$

SUBJECT TO

$$\sum_{k=1}^{np} \left[\frac{\partial \omega_d}{\partial p_k}\Big|_{p^{i-1}}\right] y_k = (1 - \gamma_i)\omega_d^s + \gamma_i\omega_d^o - \omega_d\Big|_{p^{i-1}} + \sum_{k=1}^{np} \left[\frac{\partial \omega_d}{\partial p_k}\Big|_{p^{i-1}}\right] \varepsilon_k \tag{2.5.15a}$$

$$\sum_{k=1}^{np} \left[\frac{\partial \zeta}{\partial p_k}\Big|_{p^{i-1}}\right] y_k \geq (1 - \gamma_i)\zeta^s + \gamma_i\zeta^o - \zeta\Big|_{p^{i-1}} + \sum_{k=1}^{np} \left[\frac{\partial \zeta}{\partial p_k}\Big|_{p^{i-1}}\right] \varepsilon_k \tag{2.5.15b}$$

$$\sum_{k=1}^{np} \left[\frac{\partial z_E}{\partial p_k}\Big|_{p^{i-1}}\right] y_k = (1 - \gamma_i)z_E^s + \gamma_i z_E^o - z_E\Big|_{p^{i-1}} + \sum_{k=1}^{np} \left[\frac{\partial z_E}{\partial p_k}\Big|_{p^{i-1}}\right] \varepsilon_k \tag{2.5.15c}$$

$$\sum_{k=1}^{np} \left[\frac{\partial z_{IE}}{\partial p_k}\Big|_{p^{i-1}}\right] y_k \leq (1 - \gamma_i)z_{IE}^s + \gamma_i z_{IE}^o - z_{IE}\Big|_{p^{i-1}} + \sum_{k=1}^{np} \left[\frac{\partial z_{IE}}{\partial p_k}\Big|_{p^{i-1}}\right] \varepsilon_k \tag{2.5.15d}$$

$$y \leq 2\varepsilon \tag{2.5.16}$$

$$\Delta p = y - \varepsilon$$

$$p^{i} = p^{i-1} + \Delta p$$

In the sequel, optimizations with respect to three sets of design variables and associated constraints are given.

2.5.2 Controller Design with Fixed Plant

First, we consider minimum sensitivity eigenvalue placement with no explicit constraints on design parameters. The design vector consists of 8 gain elements

$$p = (g_1, \ldots, g_8)^T \qquad \text{(8x1) vector}$$

and all other matrices are constants. The linear program at step "i" then takes the form:

MAXIMIZE $\quad -S^T y$

SUBJECT TO

$$\underset{(11X8)}{\begin{bmatrix} I \\ \dfrac{-\partial\zeta}{\partial p}\Big|_{p^{i-1}} \end{bmatrix}} y \leq \begin{Bmatrix} 2\varepsilon \\ -(1-\gamma_i)\zeta^S - \gamma_i\zeta^0 + \zeta\Big|_{p^{i-1}} - \left|\dfrac{\partial\zeta}{\partial p}\Big|_{p^{i-1}}\right| \varepsilon \end{Bmatrix}$$

$$\underset{(3x8)}{\begin{bmatrix} \dfrac{\partial\omega_d}{\partial p}\Big|_{p^{i-1}} \end{bmatrix}} y = \left\{ (1-\gamma_i)\omega_d^S + \gamma_i\omega_d^0 - \omega_d\Big|_{p^{i-1}} + \left|\dfrac{\partial\omega_d}{\partial p}\Big|_{p^{i-1}}\right| \varepsilon \right\}$$

y nonnegative

$$\Delta p = y - \varepsilon$$

$$p^{i} = p^{i-1} + \Delta p$$

Table 2.1 Damped Frequencies and Damping Factors for Three Cases:
Fixed Plant (Section 2.5.2)
Fixed Plant + Inequality Constraint (Section 2.5.3)
Variable Plant + Inequality Constraint (Section 2.5.4)

	STARTING (γ=0)	OBJECTIVE	CONVERGED (γ=1.0) Section 2.5.2	CONVERGED (γ=.412) Section 2.5.3	CONVERGED (γ=.415) Section 2.5.4
ω_{d_1}	.103	.300	.300	.189	.185
ω_{d_2}	.348	.600	.600	.448	.451
ω_{d_3}	.554	.900	.900	.697	.696
ζ_1	.001	.500	.500	.196	.209
ζ_2	.003	.400	.400	.172	.163
ζ_3	.006	.300	.312	.126	.127

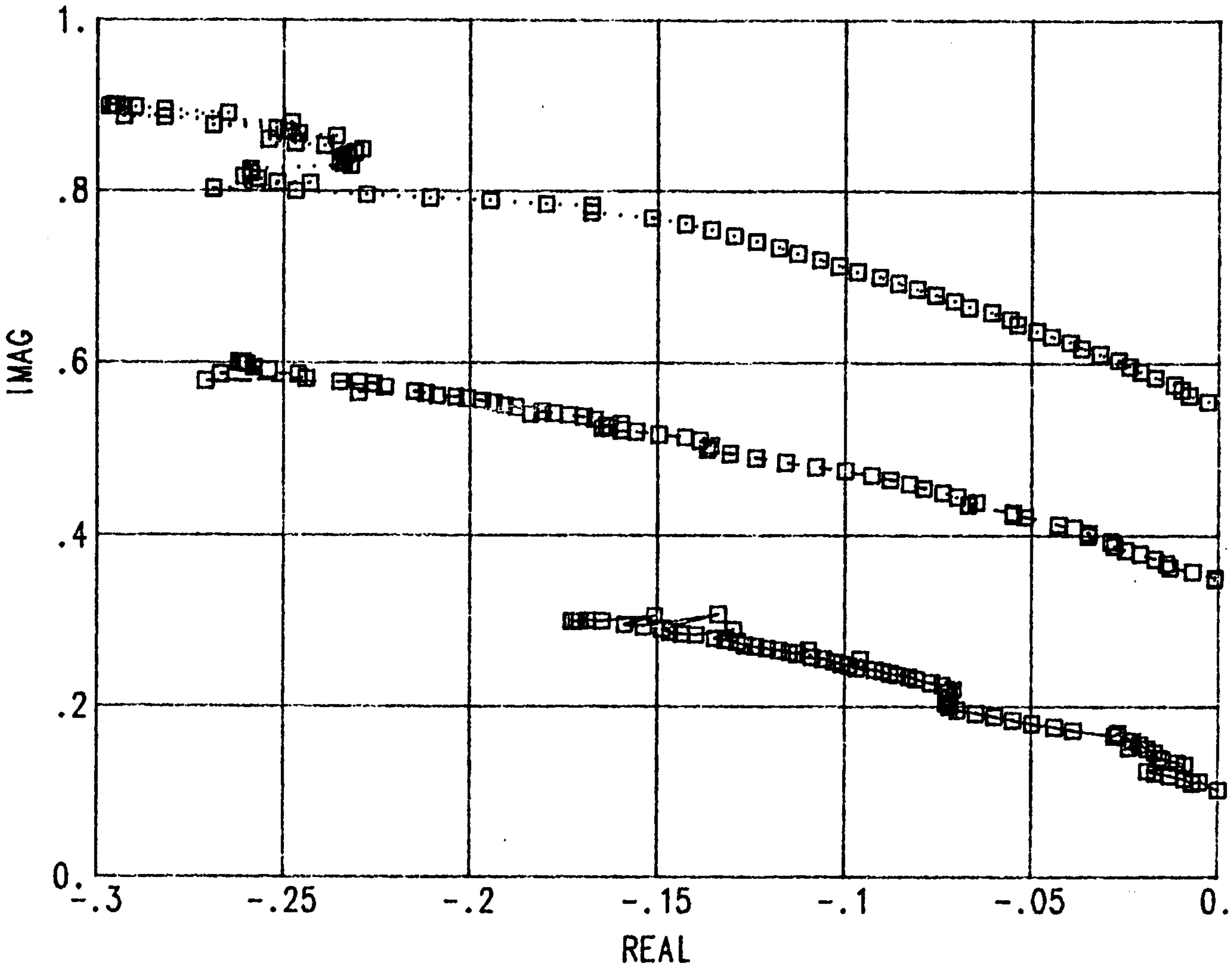


Figure 2.3a Eigenvalue Convergence Trajectories for Case with Fixed Plant

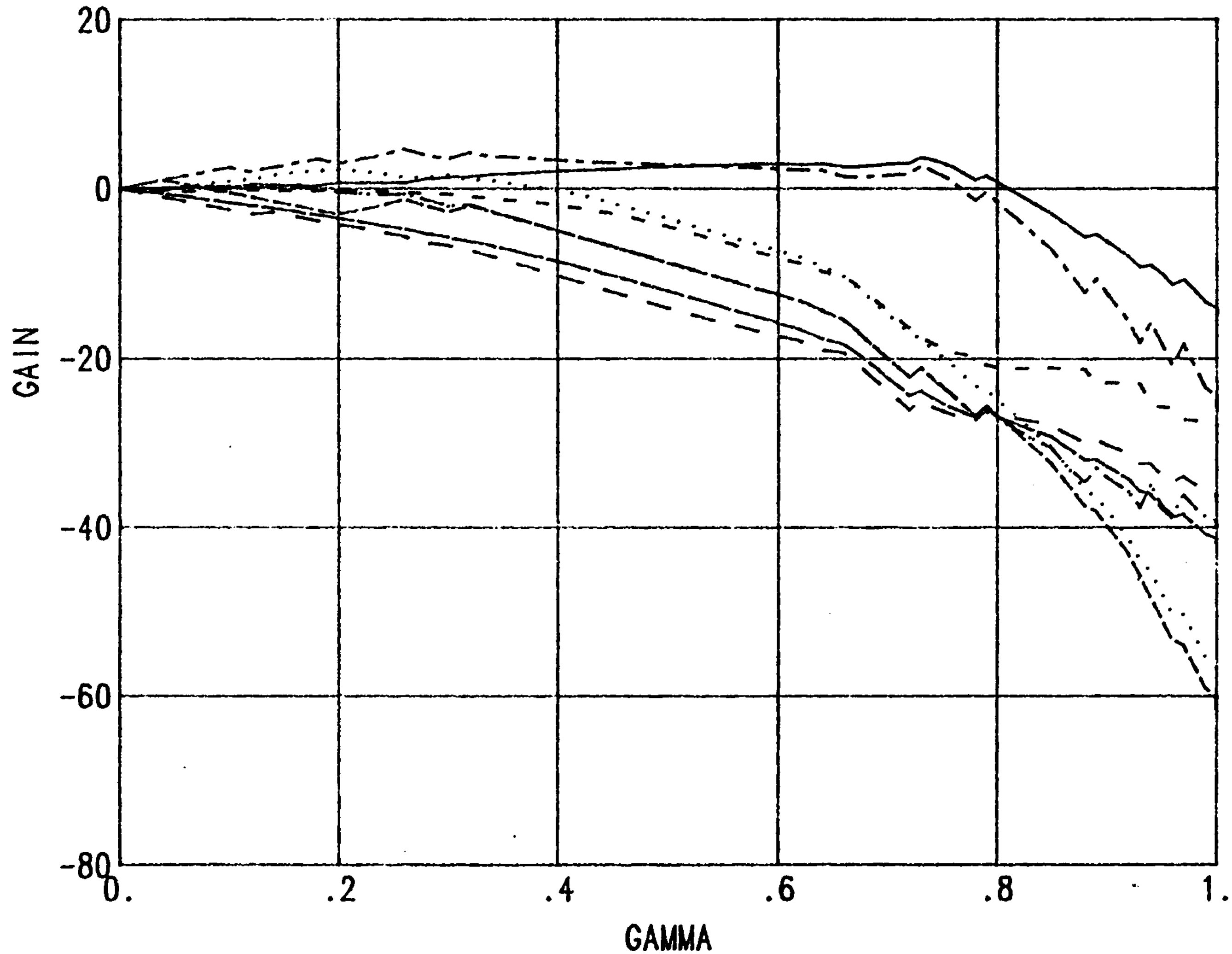


Figure 2.3b Gain Convergence Histories for Case with Fixed Plant

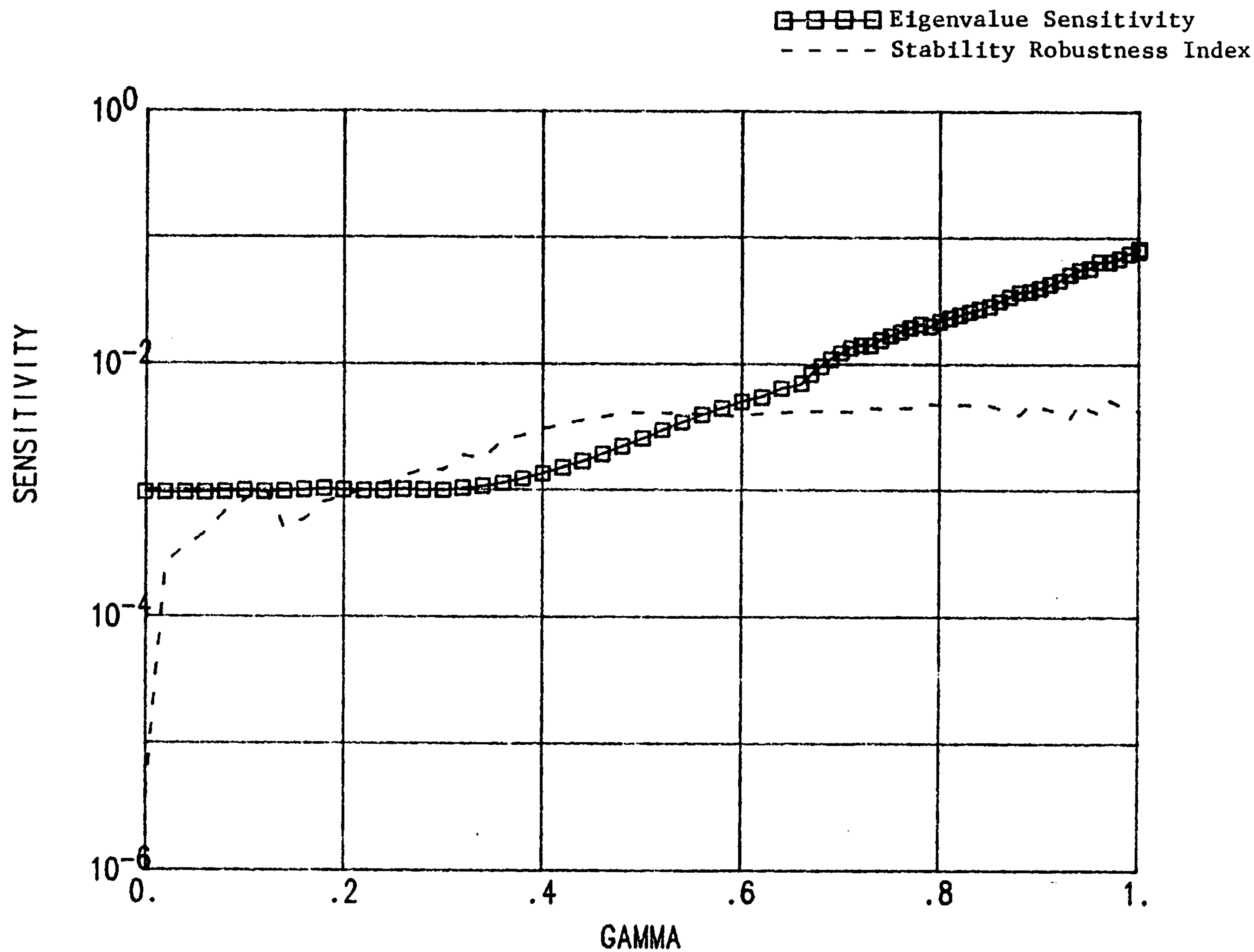


Figure 2.3c Eigenvalue Sensitivity (w.r.t. mass) Convergence Histories with Fixed Plant

The initial gains were chosen as zeros and the corresponding starting damping factors and damped frequencies computed accordingly (in this case the open loop values). Table 2.1 summarizes the starting, objective (desired) and converged values. It can be seen that all damped frequencies and damping factors have successfully converged to desired values (i.e. $\gamma = 1.0$). Figure 2.3a shows the eigenvalue trajectories during the continuation increments while Figures 2.3b and 2.3c show the gain histories and eigenvalue sensitivity cost functions respectively during the continuation increments from $\gamma = 0.0$ to $\gamma = 1.0$. From Figure 2.3a, we observe some erratic convergence behavior in the latter stages of the continuation history. Each square (symbol) represents a converged step. Note the degree of nonlinearity as the distance between the symbols vary at each nominal point. The final converged gains where

$$G = \begin{bmatrix} -14.6 & -27.8 & -57.8 & -39.7 \\ -25.8 & -41.8 & -61.5 & -36.2 \end{bmatrix}$$

It is interesting to note that all gain elements eventually converged to negative values. This means that the force actuators at degree of freedoms 1 and 3 always fed back with a negative linear combination of displacement and velocity measurements at degree of freedoms 2 and 3.

2.5.3 Controller Design with Fixed Plant and Inequality Constraints on Gain Elements

This case is identical to previous cases but with one significant difference; lower and upper bounds on gain elements are included to reflect practical constraints possibly arising from factors such as control saturation requirements or control energy limitations. By constraining all gain elements by lower and upper bounds of magnitudes g^{ℓ} and g^{u}, respectively, we have

$$g^{\ell} \leq p \leq g^{u}$$

or

$$y \leq g^{u} - p^{i-1} + \varepsilon$$

$$-y \leq -g^{\ell} + p^{i-1} - \varepsilon$$

The lower and upper bounds on gains were chosen in this example as

$$g^{\ell} = -10$$

$$g^{u} = 10 \ .$$

The linear program at step "i" then takes the form:

$$\text{MAXIMIZE} \qquad -S^{T}\, y$$

SUBJECT TO
(27x8)

$$\begin{bmatrix} I \\ -\frac{\partial\zeta}{\partial p}\Big|_{p^{i-1}} \\ I \\ -I \end{bmatrix} y \le \left\{ \begin{array}{c} 2\varepsilon \\ -(1-\gamma_i)\zeta^s - \gamma_i\zeta^0 + \zeta\Big|_{p^{i-1}} - [\frac{\partial\zeta}{\partial p}\Big|_{p^{i-1}}]\varepsilon \\ g^u - p^{i-1} + \varepsilon \\ -g^\ell + p^{i-1} - \varepsilon \end{array} \right\}$$

(3x8)

$$\left[\frac{\partial\omega_d}{\partial p}\Big|_{p^{i-1}}\right] y = \left\{(1-\gamma_i)\omega_d^s + \gamma_i\omega_d\Big|_{p^{i-1}} + [\frac{\partial\omega_d}{\partial p}\Big|_{p^{i-1}}]\,\varepsilon\right\}$$

y nonnegative,

$\Delta p = y - \varepsilon$

$p^i = p^{i-1} + \Delta p$.

The initial gains were chosen as zeros. Table 2.1 summarizes the starting, objective and converged values. We achieved convergence only up to $\gamma = .412$ and the desired objectives were not attained. This is indeed expected from the parameter histories for the unconstrained case given in Figure 2.3b, where it is seen that a bound on gains of magnitude 10 corresponds to a γ value of approximately 0.4.

Figure 2.4a shows the eigenvalue trajectories while Figures 2.4b and Figures 2.4c shows the gain histories and eigenvalue sensitivity cost function during the continuation increments from $\gamma = 0.0$ to $\gamma = 0.412$. The converged gains at $\gamma = 0.412$ were

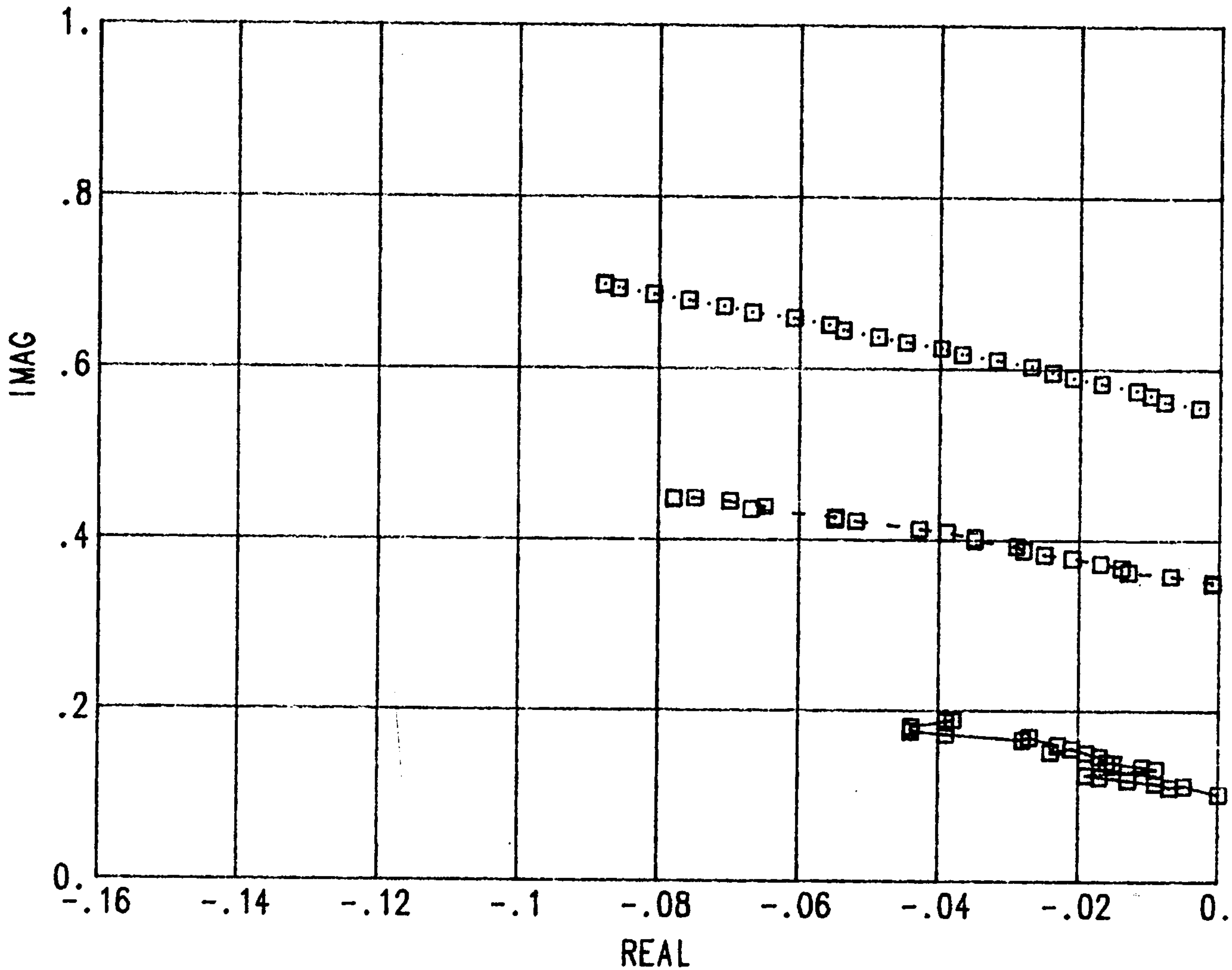


Figure 2.4a Eigenvalue Convergence Trajectories for Case with Fixed Plant and Inequality Constraints on Gains

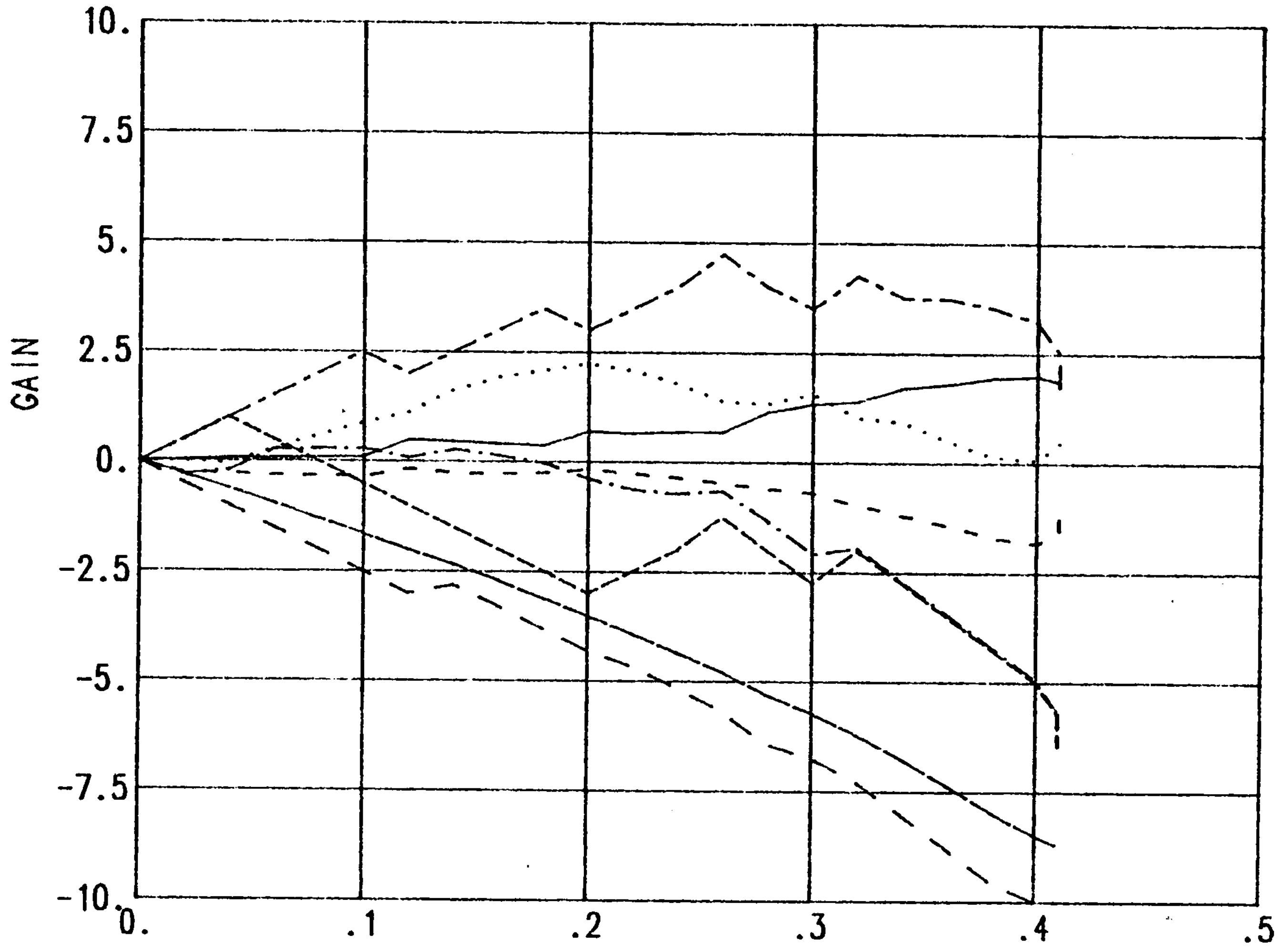


Figure 2.4b Gain Convergence Histories for Case with Fixed Plant and Inequality Constraints on Gains

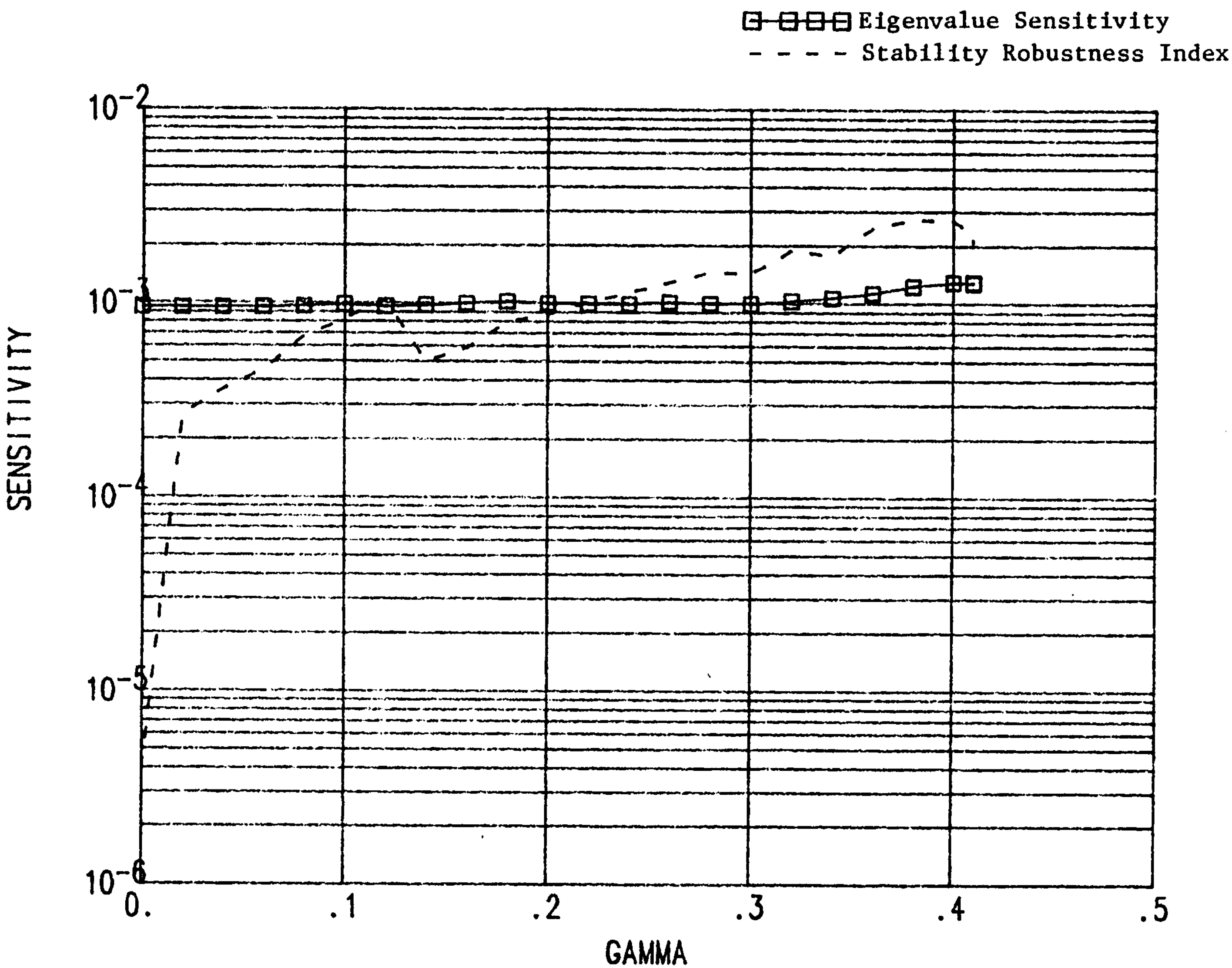


Figure 2.4c Eigenvalue Sensitivity (w.r.t. mass) Convergence Histories with Fixed Plant

$$G = \begin{bmatrix} 1.7 & -1.0 & .7 & -6.4 \\ 1.7 & -8.7 & -6.3 & -10.0 \end{bmatrix}.$$

It can be seen from the above that element (2,4) of the gain matrix reached the imposed lower bound. At this point no amount of change in the continuation step size (eigenvalue constraints) increased the level of convergence (γ). Clearly, the lower bound on gain element (2,4) is a critical constraint; a negative change in element (2,4) of the gain matrix (which violates the lower bound constraint) is required to make further progress.

2.5.4 Controller Design with Variable Plant and Inequality Constraints on Controller and Plant

This final case involves the simultaneous design of structural and controller parameters. The structural parameters considered here are the 3 stiffness elements. In addition, inequality constraints on both classes of parameters reflects the type of design restrictions one might encounter in practice. With the additional three structural parameters, we might expect to obtain better results than the previous fixed plant case as given in section 2.5.3.

The design vector thus consists of the 11 elements

$$p = (\underbrace{g_1, \ldots, g_8}_{\text{controller}},\ \underbrace{k_1,\ k_2,\ k_3}_{\text{structure}})^T$$

and the additional stiffness constraints imposed are

$$k^{\ell} \le k \le k^{u}$$

or

$$[0 \quad I]\ y \ \le\ k^u - [0 \quad I]\ \{p^{i-1} - \varepsilon\}$$

$$[0 \ -I]\ y \ \le\ -k^\ell + [0 \quad I]\ \{p^{i-1} - \varepsilon\}$$

where $k^\ell = .10$ and $k^u = 2$ are selected. The resulting linear program at step "i" takes the form:

MAXIMIZE $-S^T y$

SUBJECT TO
(36x11)

$$\begin{bmatrix} I \\ -\frac{\partial\zeta}{\partial p}\Big|_{p^{i-1}} \\ I \quad 0 \\ -I \quad 0 \\ 0 \quad I \\ 0 \ -I \end{bmatrix} y \le \left\{ \begin{array}{c} 2\varepsilon \\ -(1-\gamma_i)\zeta^s - \gamma_i\zeta^0 + \zeta\Big|_{p^{i-1}} - \left|\frac{\partial\zeta}{\partial p}\Big|_{p^{i-1}}\right| \varepsilon \\ [I \quad 0]\ \{g^u - p^{i-1} + \varepsilon\} \\ [I \quad 0]\ \{-g^\ell + p^{i-1} - \varepsilon\} \\ k^u - [0 \quad I]\ \{p^{i-1} - \varepsilon\} \\ -k^\ell + [0 \quad I]\ \{p^{i-1} - \varepsilon\} \end{array} \right\}$$

(3x11)

$$\left[\frac{\partial\omega_d}{\partial p}\Big|_{p^{i-1}}\right] y = \left\{(1-\gamma_i)\omega_d^s + \gamma_i\omega_d^0 - \omega_d\Big|_{p^{i-1}} + \left|\frac{\partial\omega_d}{\partial p}\Big|_{p^{i-1}}\right| \varepsilon \right\}$$

y nonnegative

$$\Delta p = y - \varepsilon$$

$$p^i = p^{i-1} + \Delta p \ .$$

The initial gains were again chosen as zeros and the corresponding starting damping factors and damped frequencies computed accordingly. Table 2.1 summarizes the starting, objective and converged values. We achieved convergence up to $\gamma = 0.415$ and thus the desired objectives (at $\gamma=1$) were not attained. This implies that the addition of stiffness elements as design variables did not significantly improve eigenvalue assignability over the previous case with fixed stiffness. In retrospect this is intuitively reasonable; we are attempting to move the eigenvalues left and varying stiffness parameters can be expected to have secondary effect on the real parts of the eigenvalues (neglecting weak coupling due to certain off-diagonal terms, the effect is zero).

Figure 2.5a shows the eigenvalue trajectories while Figures 2.5b, c and d shows the gain and stiffness histories and eigenvalue sensitivity cost function respectively. The converged gains and stiffness parameters at $\gamma = 0.415$ were

$$G = \begin{bmatrix} 2.0 & -3.9 & -1.8 & -10.0 \\ .9 & -10.0 & -5.6 & -9.9 \end{bmatrix}$$

$$k_1 = .78 \ ; \quad k_2 = .85 \ ; \quad k_3 = 1.00$$

Comparing the converged eigenvalue trajectories of Figures 2.3a, 2.4a and 2.5a we observe that they are almost identical. Of course, this is the result of specifying the same objective eigenvalue trajectories. The gain histories of Figures 2.3b and 2.4b are identical

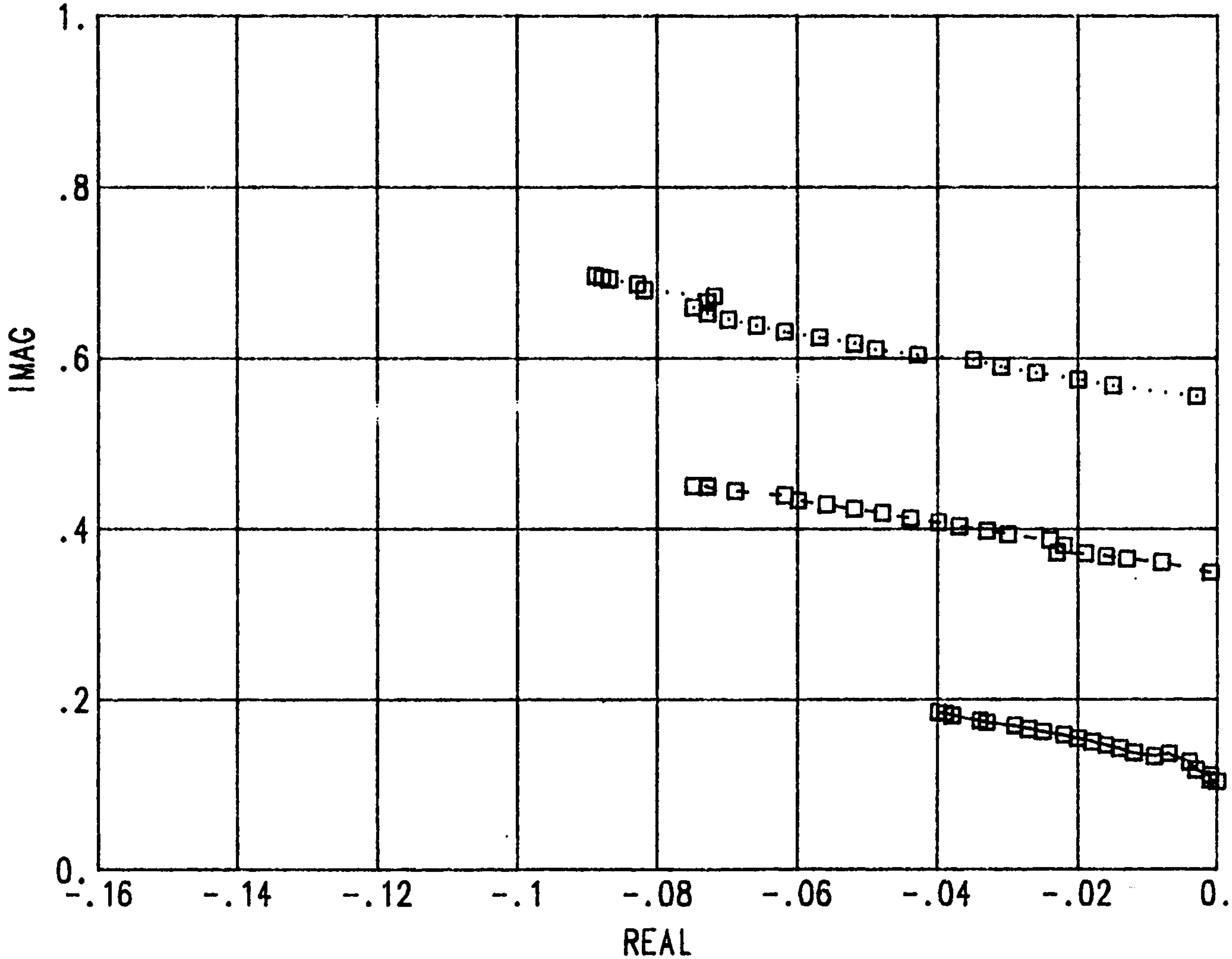


Figure 2.5a Eigenvalue Convergence Trajectories for Case with Variable Plant and Inequality Constraints on Gains and Stiffness

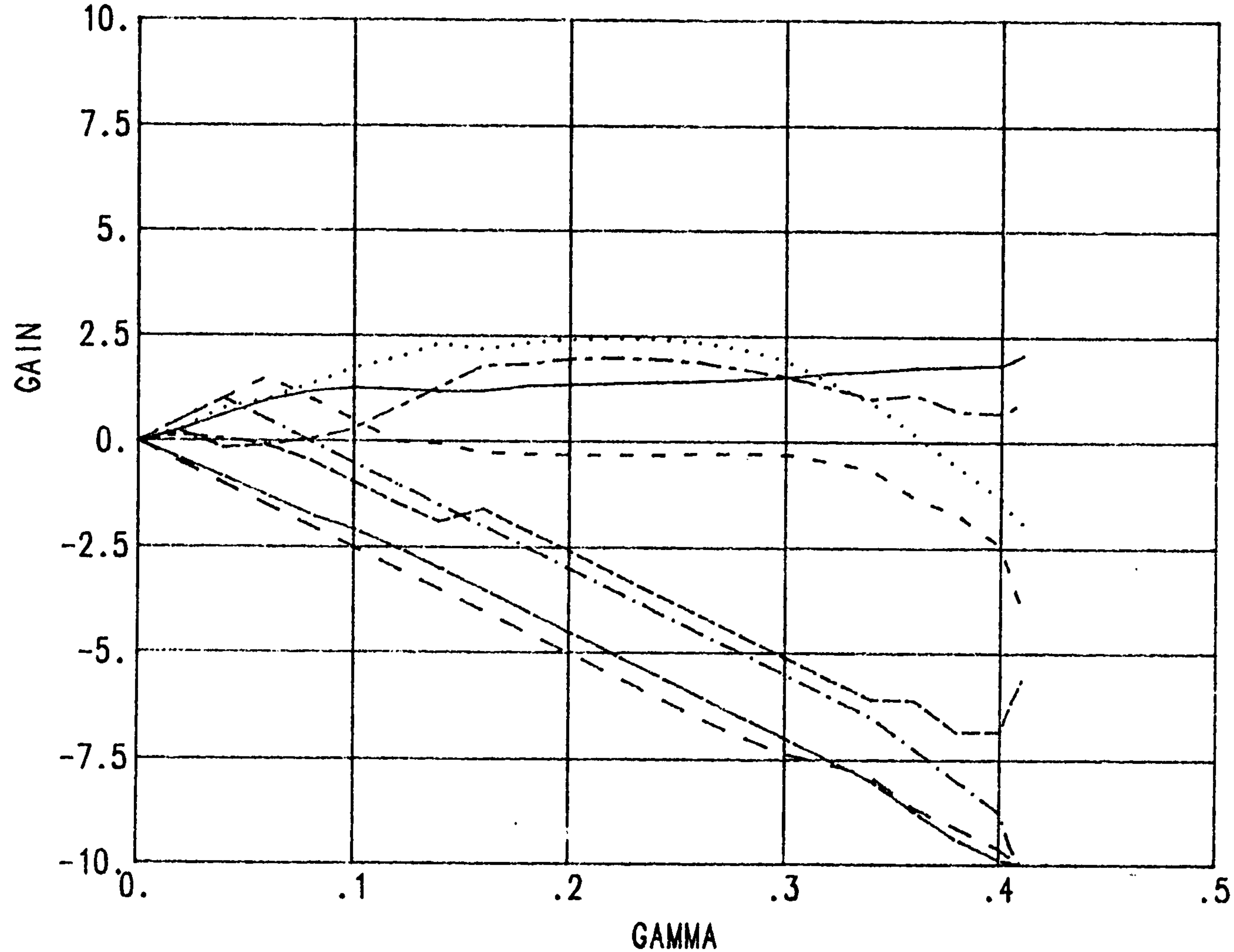


Figure 2.5b Gain Convergence Histories for Case with Variable Plant and Inequality Constraints on Gains and Stiffness

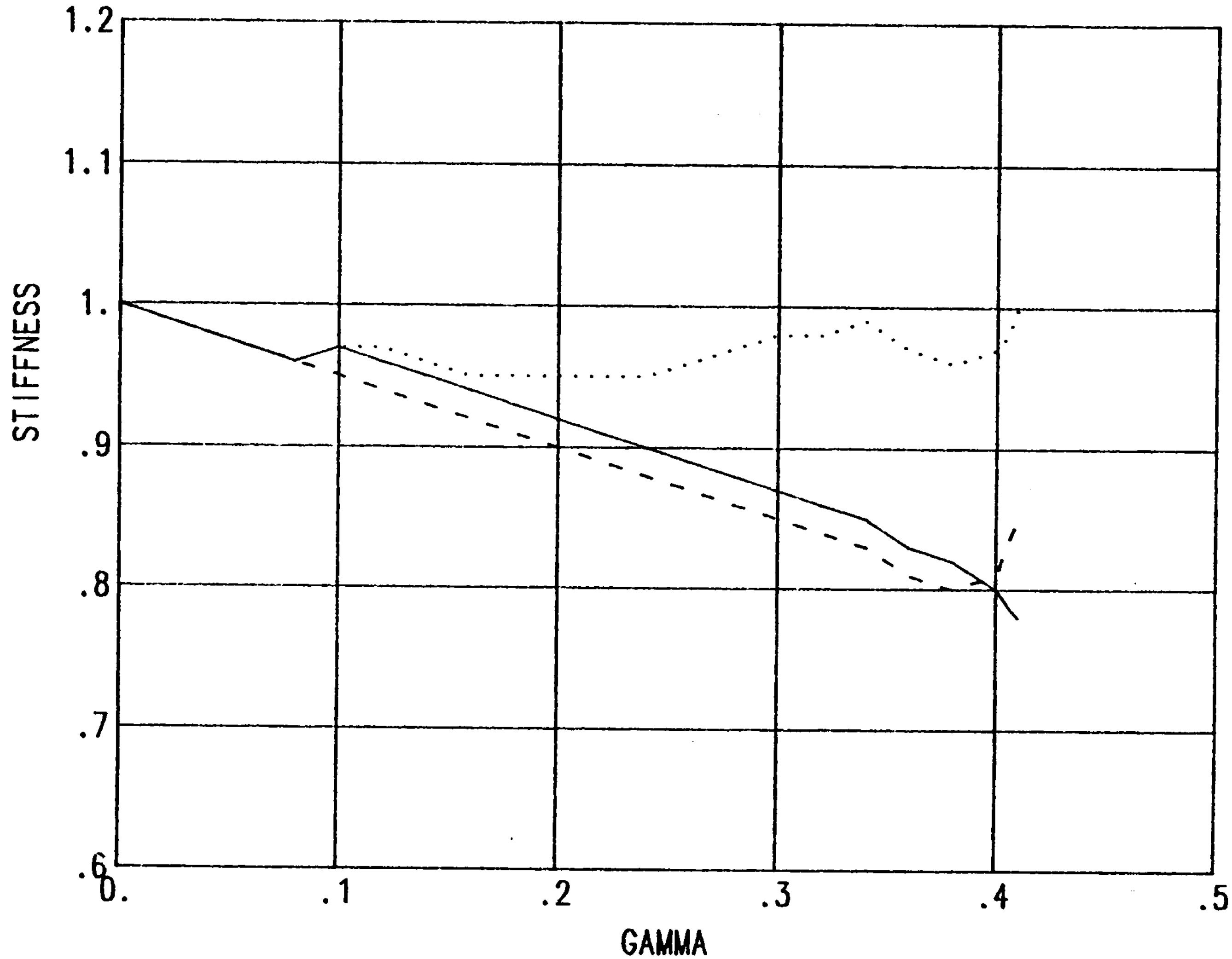


Figure 2.5c Stiffness Convergence Histories for Case with Variable Plant and Inequality Constraints on Gains and Stiffness

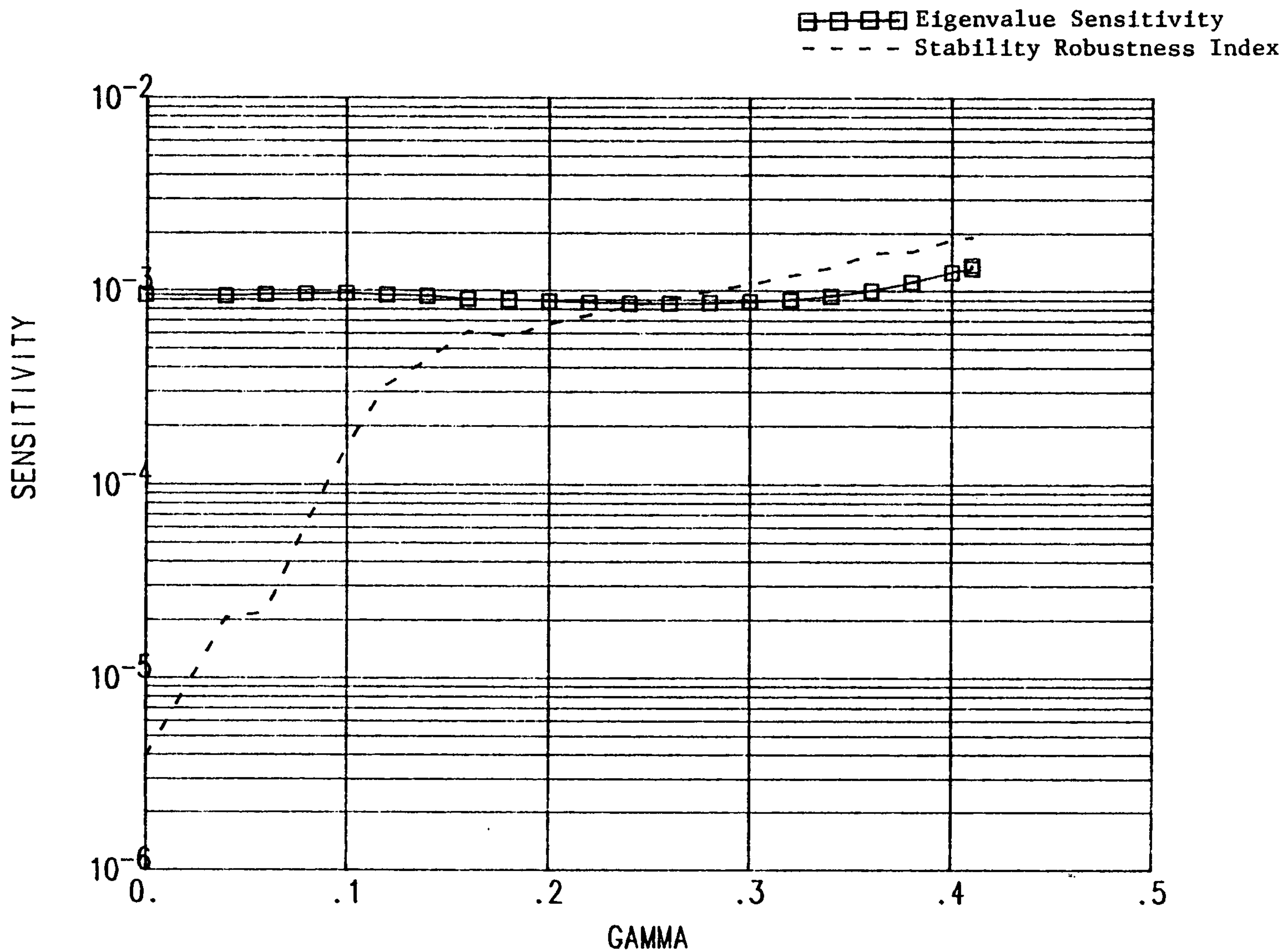


Figure 2.5d Eigenvalue Sensitivity (w.r.t. mass) Convergence Histories for Case with Variable Plant and Inequality Constraints on Gains and Stiffness

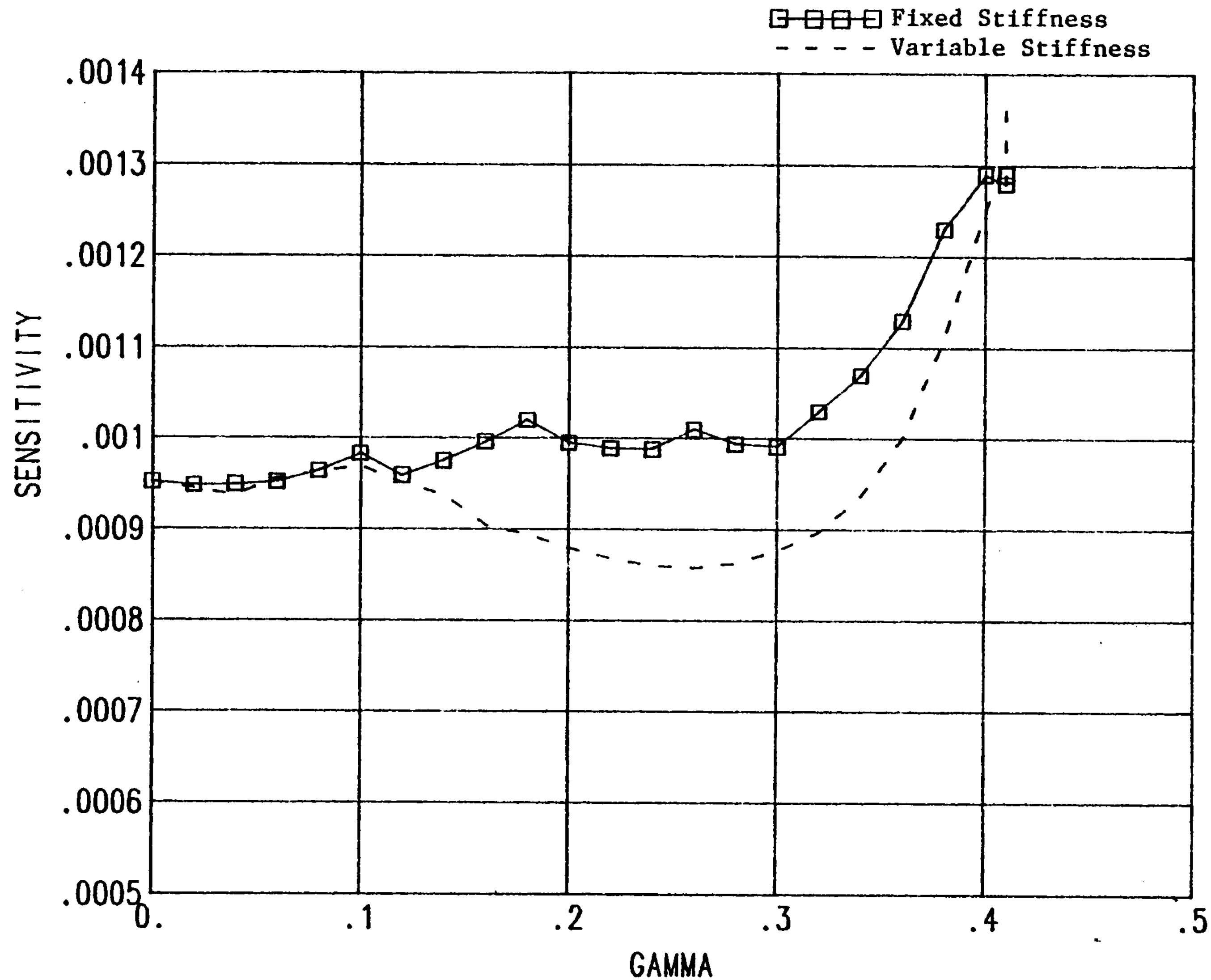


Figure 2.6 Improvement in Eigenvalue Sensitivity with Addition of Stiffness Parameters

except for the effect of constraints on gain elements in the latter case. The latter result is expected since the design variables, equality constraints and the cost functions are identical for the two cases; the additional inequality constraints on the gains only restricts further progress in the design space and has no effect on the shape of the path in the design space. On the other hand, Figure 2.5b differs from Figures 2.3b and 2.4b since the addition of stiffness parameters to the design vector fundamentally modifies the dimension of the design space, hence its trajectories. It is interesting to observe (Figure 2.5c) that all stiffness elements decreased although the frequencies were driven upwards.

The converged eigenvalue sensitivity cost function of Figures 2.3c, 2.4c and 2.5d displays a similar trend; the sensitivities remained approximately constant until $\gamma \cong .3$ where it begins to grow exponentially. Curiously enough, the corresponding stability robustness index (see section 4.3) increased almost monotonically and very rapidly near the beginning. The above result implies that our direct measure of eigenvalue sensitivity and the measure of stability robustness do not have a "one-to-one" relationship. This truth is demonstrated again for a more elaborate example in Chapter 5.

2.5.5 Summary of Numerical Results

The following are the main points demonstrated in the above examples:

- o Constraining gains leads to implicit limitations on eigenvalue assignability.

- o Although constraints on gains may prevent convergence to desired conditions, the algorithm converges to a nearest neighbor consistent with a least compromised specification for the constraints. This property is highly advantageous over existing algorithms for pole placement for which inequality constraints cannot be conveniently incorporated.
- o Inclusion of stiffness elements as design variables may not significantly improve damping assignability but may significantly improve the eigenvalue sensitivity cost function as evident in Figure 2.6.

2.6 Concluding Remarks

We have derived here a fairly general design algorithm which uses sequential linear programming and continuation methods. A simple example is used for illustration. The simultaneous design of a structure and an output feedback controller of a sixth-order dynamical system demonstrates the simplicity of the proposed algorithm. Given any parameterized system, the gradient of the constraints and cost functions are needed to set up a linear program. The validity of local linearizations is enhanced by a constraint on the maximum allowable parameter change at each step and also by adjusting the continuation step size. It is important that the maximum allowable parameter change used at each step be compatible with the continuation step size chosen, since a relatively small step size bound may eliminate the possibility of a feasible region to a set of neighboring constraints whose distance is directly related to the continuation step size. On the other hand, too large a step size bound may destabilize convergence as the linearity assumptions are sufficiently violated. Currently, we have no way to choose "optimal" parameter and continuation step sizes that would allow the largest possible continuation step that is consistent with the linearizations. However, a reasonably successful rule of thumb has been found workable; we use some percentage, (say 10%) of the non-dimensionalized range allowed for the parameter values. In the algorithm implemented here, an interactive execution of the computer program permitted the input of a new step size as needed at any continuation step to enhance convergence.

A major advantage of sequential linear programming optimization is that at each continuation step, an optimal solution, if it exists, can be found very efficiently using the simplex method. Efficiency ensues from the truth that only a finite number of feasible possibilities ("basic" solutions) exist and the simplex algorithm finds the optimal solution systematically through a matrix reduction technique. A second advantage is the flexibility to implicitly handle equality and inequality constraints on both the design variables and functions thereof, without additional programming or "bookkeeping".

In summary, the combination of sequential linear programming and continuation methods is believed to be very powerful basis for attacking high-dimensioned, nonlinear, inequality constrained optimization problems. In addition, the two methods can be combined naturally without elaborate programming.

3. MODAL SENSITIVITY THEORY

3.1 Introduction

An important issue in the design of control systems is the effect of uncertainties and perturbations on the controlled performance of the dynamical system. Discounting outright human blunders and other complications such as catastrophic failures of various types of hardware, we list below several not unexpected sources of perturbations and uncertainties:

- o modelling errors due to linearizations or other assumptions for the sake of simplifying analysis
- o parameter variations due to age or changes in structural properties induced by reconfiguration of the environment,
- o unmodelled external disturbances
- o measurement and actuator uncertainties

It is well known from Single-Input, Single-Output (SISO) theory [40] that with feedback, control systems that are minimally sensitive to input perturbations can be designed. In general, we can define "robust controllers" as the class of controllers that preserve certain properties under perturbations and uncertainties of the system. Since stability is a necessary condition for other peformance criteria of a control system, it is clear that maintaining stability of closed loop system under uncertainties is the most important design requirement.

For problems involving only infinitesimally small perturbations, small first derivatives of the properties of interest, with respect to the suspected uncertain parameters will usually be a useful criteria for designing robust controllers. In particular, in designing regulator control systems, "small" modal (eigenvalue and eigenvector) sensitivities are usually desirable. In section 3.2, we review the fact that transient reponse of a closed loop system is completely characterized by its modal properties. This means that the ability to assign eigenvalues and shape eigenvectors (eigenstructure assignment) and to control their degree of sensitivities (modal sensitivity) must play a key (explicit or implicit) role in designing robust feedback controllers for regulator problems. Consequently, section 3.3 reviews the main concepts associated with eigenstructure assignment, modal insensitivity and the problem of simultaneously assigning eigenvalues with modal insensitivity using constant gain output feedback controllers. It is seen that the above approach for simultaneously assigning eigenvalues while imposing modal insensitivity constraints is non-trivial from both theoretical and computational viewpoints. As an alternative, we examine in section 3.4 the matrix of eigenvalue sensitivities in an effort to capture modal sensitivity in a scalar performance index. Although eigenvector derivative measures are not of central concern here, a scalar sensitivity measure which is directly related to a linearly predicted bound on weighted eigenvalue perturbation is obtained.

Another approach to describing the sensitivity of eigensystems is through the well-established concept of conditioning of eigenvalue problems in numerical analysis. The formulation involves the use of matrix and vector norms to measure the size of changes in the solution due to changes in the given matrix. In short, a condition number can be defined that represents the degree of ill-conditioning of the nominal eigensystem and is related directly to an upper bound on the perturbation of the eigenvalues due to a unit norm change in the system matrix. The eigenvalue bounding equations are applied to guarantee asymptotic stability to arrive at a stability robustness criteria which incidentally turns out to correspond exactly to a robustness criteria derived earlier by Patel and Toda [30] using the Lyapunov stability theorem. This is indeed an interesting equivalence.

Finally, in section 3.6, we digress slightly to present an important and fundamental result. We begin by reviewing carefully the eigenvalue problem for non-self-adjoint systems and then re-examine the derivation of eigenvector derivatives by earlier authors. We obtain here a correct expression for the eigenvector derivatives. We find the non-self-adjoint eigenvector derivative is not normal to the eigenvector as is explicitly assumed in most of the available literature.

3.2 Transient Response

The details of this section are well known but are included here for completeness and since they form the basis for an understanding of the significance of eigenstructure assignment in shaping transient response for linear systems. The following exposition closely follows the work in [39].

Let us consider a system described by n-first order, linear differential equations

$$\dot{x}(t) = A\, x(t) \quad ; \quad x(0) = x_0 \tag{3.2.1}$$

We note here that a closed-loop feedback system will be of the form in Eq. (3.2.1). The solution to Eq. (3.2.1) is well known [40,41] and can be written as

$$x(t) = e^{At} x_0 \tag{3.2.2}$$

where e^{At} represents the matrix exponential of At. Although Eq. (3.2.2) is a perfectly valid representation of the solution, an alternate form can be written in terms of the eigenvalues, eigenvectors and the initial conditions. To this end, the eigenvalue equations are written as

$$Ax_i = \lambda_i x_i \quad , \quad i=1,\ldots,n \tag{3.2.3}$$

where λ_i and x_i represent the i-th eigenpair. For simplicity, we assume that the eigenvalues are distinct, which implies that a set of n linearly independent eigenvectors can always be found. We define next the right modal matrix, X, as

$$X = [x_1,\ldots,x_n] \tag{3.2.4}$$

and transform coordinates to modal space by

$$x(t) = X\, z(t) \tag{3.2.5}$$

and rewrite Eq. (3.2.1) as

$$\dot{z}(t) = X^{-1}AX\, z(t) \quad ; \quad z_0 = X^{-1}x_0 \tag{3.2.6}$$

We note that the inverse of the modal matrix exists since we have restricted attention to the case for which we have a linearly independent set of eigenvectors. Furthermore, the distinct eigenvalue assumption guarantee diagonalizability through the similarity transformation

$$X^{-1}AX = \mathrm{diag}\{\lambda_i,\ldots,\lambda_n\} \triangleq \Lambda \tag{3.2.7}$$

Hence in modal space, the solution of Eq. (3.2.6) can be written as

$$z(t) = e^{\Lambda t}z_0 \tag{3.2.8}$$

where

$$\begin{aligned} e^{\Lambda t} &= \mathrm{diag}\{e^{\lambda_1 t},\ldots,e^{\lambda_n t}\} \\ z_0 &= X^{-1}x_0 \end{aligned} \tag{3.2.9}$$

Returning to the original coordinate system, our solution may now be written as

$$x(t) = Xe^{\Lambda t}X^{-1}x_0 \tag{3.2.10}$$

Thus $e^{At} = Xe^{\Lambda t}X^{-1}$.

The contribution of individual modes to the total response can be more readily seen by using the following relations

$$Xe^{\Lambda t} = [x_1,\ldots,x_n] \;\; \mathrm{diag}(e^{\lambda_1 t},\ldots,e^{\lambda_n t})$$

$$= [e^{\lambda_1 t}x_1,\ldots,e^{\lambda_n t}x_n]$$

and

$$X^{-1} = Y^T \triangleq [y_1,\ldots,y_n]^T$$

where $y_1,\ldots,y_n$ denotes the left eigenvectors which are normalized by

$$Y^T X = I \; . \tag{3.2.11}$$

So, Eq. (3.2.10) can be rewritten as

$$x(t) = [e^{\lambda_i t}x_1,\ldots,e^{\lambda_n t}x_n]\begin{Bmatrix} y_1^T \\ \vdots \\ y_n^T \end{Bmatrix} x_0$$

or

$$x(t) = \sum_{i=1}^{n} \alpha_i e^{\lambda_i t} x_i \quad ; \quad \alpha_i \triangleq y_i^T x_0 \tag{3.2.12}$$

We note that Eq. (3.2.12) is often referred to as the "modal expansion theorem". We also note that every transient solution of the initial value problem Eq. (3.2.1) depends on three quantities:

(i) eigenvalues, which determine the decay/growth rate of the response,

(ii) eigenvectors, which determine the shape of the response,

(iii) initial condition, which determines the degree to which each mode participates in the response.

In summary, it is clear that eigenstructure assignment is simply a direct approach for altering the system transient response. The important properties of closed loop stability and the "shape" of the transient response is conveniently related to the eigenvalues and eigenvectors. This explains the popularity and significance of eigenstructure assignment algorithms in the controls literature.

3.3 Eigenstructure Assignment and Modal Insensitivity

3.3.1 Eigenstructure Assignment

From the previous section on transient response, we see that eigenstructure assignment plays a key role in shaping transient response and in fact, any response concerning the alteration of transient response must change the eigenstructure of the system. It is obvious that the most direct approach to attaining a particular eigenstructure is through a judicious choice of the plant parameters. In the areas of structural dynamics where vibrational characteristics are of interest, eigenstructure assignment techniques can be used for modifying existing vibratory systems to satisfy new modal constraints [21,42]. On the other hand, for a class of control design problems, eigenstructure assignment can be achieved through feedback and in some cases a simultaneous tuning of plant parameters and feedback gains may be justified or even be necessary. Indeed, for the past two decades, a great deal of research has been devoted to developing practical algorithms for eigenstructure assignment using state or output feedback or with dynamic compensation [73; chap.6]. However, the underlying basic theory which provides the framework for the various eigenstructure assignment algorithms existing today is credited to a few pioneering works as represented in [38,69-71].

In this and the next two sections, we review the basic concepts of eigenstructure assignment, modal insensitivity and their interdependence through a choice of an eigenvector set.

Perhaps a good way to overview the development of the main theoretical results is through a chronological order. We begin by stating Bertram's results. According to Kalman [38], Bertram first showed in 1959 that if a system is controllable, then every self-conjugate set of eigenvalues can be realized by a unique gain vector for single input case. This was followed by Wonham [69] who in 1967 extended the controllability result from single input to multi-input systems. Perhaps Wonham's result can be regarded as the major contribution in modern control theory. This is plausible because, since his paper appeared, there have been literally hundreds of papers written on multi-input pole placement, including the extensions to output feedback and dynamic compensation. Wonham's result can be stated formally as follows:

THEOREM 3.3.1

The system $\dot{x}(t) = A\,x(t) + B\,u(t)$ is controllable iff for every self-conjugate set of scalars $\{\lambda_i^d\}$, $i=1,\ldots,n$, there exists a real $(m \times n)$ matrix K such that $(A + BK)$ has $\{\lambda_i^d\}$, $i=1,\ldots,n$, as its eigenvalues.

It is important to note several facts regarding Wonham's result, (i) for multiinput case, the state feedback gains for a given set of eigenvalues is not unique. This is a "blessing in disguise" since nonuniqueness allows the possibility of the designer to satisfy other closed loop properties, such as minimizing the sensitivity of system performance, or maximization of stability robustness due to modelling errors, and imposing an assortment of other performance constraints. This flexibility continues to be exploited by researchers focusing on pole placement algorithms. (ii) Wonham's result applies to a full state feedback which is frequently undesirable (or unachievable) from an applications viewpoint; however, other workers have extended his results to output feedback, (iii) in most other papers following Wonham's fundamental result, the problem of eigenvector assignment, or the shaping of transient response has not been addressed or at best partially resolved.

The previously mentioned unresolved issues of nonuniqueness of gains and the possibility of eigenvector assignment using state feedback was addressed by Moore [70] in 1976. He identified the freedom offered by state feedback beyond specification of the closed loop eigenvalues for the case of distinct eigenvalues. In particular, Moore derived the necessary and sufficient conditions for the existance of a state feedback gain matrix which yields prescribed eigenvalues and eigenvectors. By doing so, he has in fact characterized the class of all closed loop eigenvector sets attainable for a given set of eigenvalues. His result, stated as a theorem is shown next followed by a proof because it includes a useful procedure for computing the gain matrix directly.

THEOREM 3.3.2

Let $\{\lambda_i\}$, i=1,...,n, be a self-conjugate set of distinct complex numbers and matrices A and B represent plant matrices, as defined in Theorem 3.3.1. For each λ_i, let us also define an nx(n+m) matrix

$$S_{\lambda_i} \triangleq [\lambda_i I - A \vdots B]$$

and denote the corresponding matrix of basis vectors of dimension $(n+m) \times \nu_i$ as

$$R_{\lambda_i} = \begin{bmatrix} N_{\lambda_i} \\ M_{\lambda_i} \end{bmatrix}$$

for the null space of matrix S_{λ_i} or Ker $\{S_{\lambda_i}\}$.

Then, there exists a real (m x n) matrix K such that

$$(A + BK)x_i = \lambda_i x_i \ , \quad i=1,\ldots,n$$

iff

(i) $x_i \in C^n$ are linearly independent

(ii) $x_i = x_j^*$ whenever $\lambda_i = \lambda_j^*$

(iii) $x_i \in \text{span}\ \{N_{\lambda_i}\}$

If K exists and rank$\{B\} = m$, then K is unique.

PROOF OF THEOREM 3.3.2

Necessary Condition:

The necessity of (i) and (ii) follows directly from matrix theory, i.e., if the system matrix (A + BK) is real, the eigenvalues and eigenvectors will be self-conjugate sets. If the eigenvalues are also assumed distinct, the eigenvector set will always be linearly independent. The condition in (iii) arises as follows:

$$(A + BK)x_i = \lambda_i x_i$$

$$\Longrightarrow (\lambda_i I - A)x_i - BKx_i = 0$$

$$[\lambda_i I - A \vdots B] \begin{bmatrix} x_i \\ -Kx_i \end{bmatrix} = 0$$

and since R_{λ_i} is defined as a matrix of basis vectors for

Ker $\{S_{\lambda_i}\}$, i.e. the null space of $[\lambda I - A \vdots B]$,

$$\Longrightarrow \begin{Bmatrix} x_i \\ -Kx_i \end{Bmatrix} \varepsilon \ \text{span} \begin{bmatrix} N_{\lambda_i} \\ M_{\lambda_i} \end{bmatrix}$$

$$\Longrightarrow \quad x_i \quad \varepsilon \quad \text{span} \quad \{N_{\lambda_i}\}$$

Q. E. D.

<u>Sufficient Condition:</u>

Assume that the set $\{x_i\}$, i=1,...,n, satisfy conditions (i) to (iii). From (iii), there exists a vector z_i (real or complex) such that

$$x_i = N_{\lambda_i} z_i \quad ; \quad i=1,\ldots,n. \tag{3.3.1}$$

Also by definition of S_{λ_i} and R_{λ_i} any vector, ξ_i , such that

$$\xi_i = R_{\lambda_i} z_i$$

satisfies

$$S_{\lambda_i} \xi_i = 0$$

$$\Longrightarrow \quad [\lambda_i I - A \quad B] \begin{bmatrix} N_{\lambda_i} \\ M_{\lambda_i} \end{bmatrix} z_i = 0$$

$$\Longrightarrow \quad (\lambda_i I - A) N_{\lambda_i} z_i + BM_{\lambda_i} z_i = 0$$

$$\Longrightarrow \quad (\lambda_i I - A) x_i + BM_{\lambda_i} z_i = 0 \ .$$

We observe here that if a K can be chosen such that

$$M_{\lambda_i} z_i = -Kx_i \ ; \quad i=1,\ldots,n \tag{3.3.2}$$

then

$$(\lambda_i I - A) x_i - BKx_i = 0$$

$$[\lambda_i I - (A + BK)] x_i = 0 \ .$$

To compute an (m x n) matrix K, we arrange Eqs. (3.3.2) in matrix form

$$[Kx_1,\ Kx_2,\ldots,Kx_n] = -[M_{\lambda_1}z_1,\ldots,M_{\lambda_n}z_n]$$

$$\Longrightarrow \quad KX = -[M_{\lambda_1}z_1,\ldots,M_{\lambda_n}z_n] \qquad (3.3.3)$$

Since X is chosen to be linearly independent, we can always invert it to solve for K from Eq. (3.3.3). It can also be shown that if rank{B} = m, the columns of N_{λ_i} in Eq. (3.3.1) will be linearly independent [39] and can be solved uniquely. This implies that the gain matrix, K will be unique if matrix B has full rank. At this stage, we note carefully that this "uniqueness" is with respect to an apriori specified set of eigenvectors that satisfies Eqs. (3.3.1).

What remains in the proof is to show that the computed K will always be real if the conditions (i) to (iii) are satisfied. For the case when each desired λ_i is real, then all, x_i and R_{λ_i} will also be real so that K computed from Eq. (3.3.3) will be real. For the complementary case, if the desired eigenvalues are a set of self-conjugate complex scalars, we need only to prove the case of one conjugate pair since for additional pairs, similar arguments apply. First, rewrite Eq. (3.3.3) as

$$K[x_1,\ldots,x_n] = [w_1,\ldots,w_n] \qquad (3.3.4)$$

where

$$w_i \triangleq -M_{\lambda_i}z_i \ .$$

For the case of one conjugate pair, let us assume for convenience

$$\lambda_1 = \lambda_2^*$$

$$\Longrightarrow \quad x_1 = x_2^* \quad \text{from (ii)}$$

$$\Longrightarrow \quad z_1 = z_2^* \quad \text{from (3.3.1)}$$

$$\text{since} \quad N_{\lambda_2} = (N_{\lambda_1})^* \text{ and } M_{\lambda_2} = (M_{\lambda_1})^* \ .$$

Therefore, Eq. (3.3.4) can be rewritten as

$$K[\mathrm{Re}\{x_1\} + j\mathrm{Im}\{x_1\} : \mathrm{Re}\{x_1\} - j\mathrm{Im}\{x_1\} : x_3,\ldots,x_n]$$

$$= [\mathrm{Re}\{w_1\} + j\mathrm{Im}\{w_1\} : \mathrm{Re}\{w_1\} - j\mathrm{Im}\{w_1\} : w_3,\ldots,w_n] \qquad (3.3.5)$$

Postmultiplying Eq. (3.3.5) by a nonsingular matrix

$$Q = \left[\begin{array}{cc|c} \frac{1}{2} & -j\frac{1}{2} & 0 \\ \frac{1}{2} & j\frac{1}{2} & \\ \hline 0 & & I \end{array}\right]$$

we obtain the real matrix equation

$$K[\mathrm{Re}\{x_1\} : \mathrm{Im}\{x_1\} : x_3,\ldots,x_n] = [\mathrm{Re}\{w_1\} : \mathrm{Im}\{w_1\} : w_3,\ldots,w_n]$$

so that K will be guaranteed real. In general, when $\{\lambda_i\}$, $i=1,\ldots,n$, is a self-conjugate set, we have

$$K[\mathrm{Re}\{x_1\} : \mathrm{Im}\{x_1\} : \cdots : \mathrm{Re}\{x_{n/2}\} : \mathrm{Im}\{x_{n/2}\}]$$

$$= [\mathrm{Re}\{w_1\} : \mathrm{Im}\{w_1\} : \cdots : \mathrm{Re}\{w_{n/2}\} : \mathrm{Im}\{w_{n/2}\}]$$

and K will obviously be real. Q. E. D.

It should be noted that the above approach for calculating the gains requires prior knowledge of a set of basis for the null space of S_{λ_i}, $i=1,\ldots,n$. Reference [72] provides an algorithm for the computation of such a basis and is claimed to be ideally suited to digital implementation. In the sequel in reference [77] the distinct eigenvalue requirement in Theorem 3.3.2 is relaxed. At this point, we should also note the alternative class of approaches for eigenstructure assignment developed by Bhattacharyya and deSouza [79] which do not require the computation of the above basis for null space. Instead, their algorithm requires the solution of a cleverly rearranged matrix eigenvalue equation which is called a Sylvester equation. The simplicity and directness of the "Sylvester" eigenstructure assignment algorithm is attractive and appears promising as a design tool. The above algorithm has also been suggested for use in robust and well-conditioned eigenstructure assignment by an iterative procedure outlined in reference [80].

We see from Moore's theorem that a choice of a set of eigenvalues implicitly defines a subspace of attainable eigenvector sets i.e. by condition (iii) or Eq. (3.3.1). Since a set of eigenvectors determines its modal sensitivities (see section 3.5), we observe that the eigenvalue placement problem cannot be considered independent of the problem of reducing modal sensitivities, at least for the case of constant state feedback. As we shall see in the next section, the conditions for modal insensitivity of every mode may not be realizable due to the constraints imposed by eigenvalue placement.

Finally, we conclude this section by stating an important result due to Srinathkumar [71]. This involves the problem of eigenstructure assignment using output feedback. His main result can be stated as follows:

THEOREM 3.3.3

Given the controllable and observable system

$$\dot{x}(t) = A\,x(t) + B\,u(t)$$

$$y(t) = C\,x(t)$$

with matrices B of $(n \times m)$ and C of $(r \times n)$ having full ranks,

- o $\max(m,r)$ closed-loop eigenvalues can be assigned,
- o $\max(m,r)$ eigenvectors, or reciprocal vectors by duality, can be partially assigned with $\min(m,r)$ entries in each vector arbitrarily chosen, using output feedback, $u(t) = F\,y(t)$.

For the proof of the above theorem, we refer the reader to reference

[71]. We now make a few remarks pertaining to Srinathkumar's result. First, we note that the number of actuators (m) and sensors (r) plays a direct role in eigenstructure assignability and we emphasize the relative difficulty of eigenvector assignment over eigenvalue assignment if m and r differ widely. Finally, we state from past numerical experience that the above result, while very enlightening, is not too obvious or easily realizable in practice.

3.3.2 Modal Insensitivity

The problem of designing control systems under varying plants and external disturbances is very important. An approach to attack these problems is to incorporate sensitivity constraints into controller designs. However, the exact type of sensitivity constraints to impose usually depends on the control system objectives. For example, if the primary goal of a control system is to track a given signal closely, an appropriate sensitivity function might take the form of a mean square integral of trajectory sensitivities [37,74] over a given period of time. For regulator problems with feedback, a useful measure of sensitivity are the sensitivities of closed loop eigenvalues and eigenvectors with respect to potentially varying parameters.

In the above context, reference [36] formally defines "modal insensitivity" as the condition when the closed-loop eigenvalues and eigenvectors are insensitive to infinitesimally small variations in plant parameters. As we shall see in the sequel, the conditions for modal insensitivity can be obtained easily from either first order perturbation equations of the eigenvalue problem or from explicit expressions for eigenvalue and eigenvector derivatives.

In this section, we state and prove the results originally given in [36]. It should be pointed out however that the proofs presented here uses explicit formulas for eigenvalue and eigenvector derivatives (derived in section 3.6), while the proof as given in [36] utilizes first order perturbation equations. In addition, it will be shown that conditions for modal insensitivity as originally given in [36] should be slightly modified in light of implicit normalizations overlooked by previous authors. The following results assume that the eigenvalues are all distinct and the eigenvectors form a basis for spanning the n-dimensional complex space.

First, we examine the conditions for zero eigenvalue sensitivity.

THEOREM 3.3.4

A necessary and sufficient condition for

$$\frac{\partial \lambda_i}{\partial \rho} = 0 \tag{3.3.6}$$

is

$$\frac{\partial A}{\partial \rho} x_i \in \text{span } \{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n\} . \tag{3.3.7}$$

PROOF OF THEOREM 3.3.4

<u>Necessary Condition:</u>

Since $\{x_1, \ldots, x_n\}$ is a basis for n-space, $\frac{\partial A}{\partial \rho} x_i$ can be expanded as

$$\frac{\partial A}{\partial \rho} x_i = \sum_{j=1}^{n} \alpha_{ij} x_j$$

where α_{ij} $(j=1,\ldots,n)$ are the coordinates in modal space. By using the eigenvalue derivative expression given in Eq. (3.6.15),

$$\frac{\partial \lambda_i}{\partial \rho} = \frac{1}{s_i} y_i^T \frac{\partial A}{\partial \rho} x_i \quad ; \quad s_i = y_i^T x_i$$

$$= \frac{1}{s_i} y_i^T \sum_{j=1}^{n} \alpha_{ij} x_j$$

$$= \alpha_{ii} \quad ; \quad \text{since } y_i^T x_j = \delta_{ij} s_i$$

Therefore Eq. (3.3.6) ===> $\alpha_{ii} = 0$.

<u>Sufficient Condition:</u>

$$(3.3.7) \implies \frac{\partial A}{\partial \rho} x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \alpha_{ij} x_j$$

$$\implies \frac{\partial \lambda_i}{\partial \rho} = \frac{1}{s_i} y_i^T \sum_{\substack{j=1 \\ j \neq i}}^{n} \alpha_{ij} x_j$$

$$= 0 \qquad \text{Q.E.D.}$$

The next theorem defines "modal insensitivity". We point out that the second equation of (3.3.8) is slightly different from the modal insensitivity condition as given in reference [36; theorem 3].

THEOREM 3.3.5

A necessary and sufficient condition for modal insensitivity of the i-th mode,

$$\frac{\partial \lambda_i}{\partial \rho} = 0$$

$$\frac{\partial x_i}{\partial \rho} = 0 \qquad (3.3.8)$$

is

$$\frac{\partial A}{\partial \rho} x_i = 0 \; . \qquad (3.3.9)$$

PROOF OF THEOREM 3.3.5

Necessary Condition:

e.v.p.

$$Ax_i = \lambda_i x_i$$

$$\Longrightarrow \frac{\partial A}{\partial \rho} x_i + A\frac{\partial x_i}{\partial \rho} = \frac{\partial \lambda_i}{\partial \rho} x_i + \lambda_i \frac{\partial x_i}{\partial \rho}$$

$$(3.3.8) \Longrightarrow \frac{\partial A}{\partial \rho} x_i = 0$$

Sufficient Condition:

From Eqs. (3.6.15), (3.6.18)

$$(3.3.9) \Longrightarrow \frac{\partial \lambda_i}{\partial \rho} = \frac{1}{s_i} y_i^T \frac{\partial A}{\partial \rho} x_i = 0$$

and

$$\alpha_{ij} = 0 \quad ; \quad j=1,\ldots,n \; ; \quad j \neq i$$

$$\Longrightarrow \alpha_{ii} = 0$$

$$\Longrightarrow \frac{\partial x_i}{\partial \rho} = 0 \qquad \text{Q.E.D.}$$

We see from theorem 3.3.5 that for modal insensitivity of all modes with respect to s-parameters, the conditions that the eigenvectors must satisfy are

$$TX \triangleq \overset{(nxs)xn}{\begin{bmatrix} \frac{\partial A}{\partial \rho_1} \\ \vdots \\ \frac{\partial A}{\partial \rho_s} \end{bmatrix}} \overset{nxn}{X} = \overset{(nxs)xn}{0} \qquad (3.3.9)$$

From matrix theory [p.33; 51], we know that the dimension of the null space, ν , must satisfy

$$\nu = n - \gamma$$

where n and γ are the number of columns and the rank of matrix T respectfully. For a general nontrivial problem, we expect matrix T to have full rank which implies a null space of zero dimension. This means that eigenvectors that satisfy modal insensitivity condition of Eq. (3.3.10), do not exist in general, much less for the case with additional eigenvalue placement constraints on the eigenvector subspace. The above difficulty has been recognized in reference [37] where it is pointed out that in many practical situations, it is certainly not necessary to require the insensitivity of an entire eigenvector and even every modelled eigenvalue. We note here that this is very evident in the field of large flexible structures controls [1-6,8] where the structural models are usually based on finite element theory and suffers a lack of accuracy in predicting higher frequency modes of the "modelled set" [75]. In reference [37], the conditions for modal insensitivity of selected eigenvalues and specific elements of eigenvectors are derived. This leads to a design freedom in the selection of state or output constant feedback gain matrices and is approached in [37,14] by seeking the "optimal" quadratic regulator.

3.3.3 Eigenstructure Assignment with Modal Insensitivity

In the previous two sections, we have seen the conditions for eigenstructure assignment and modal insensitivity independently. We now examine the problem of finding constant feedback gains that simultaneously assigns eigenvalues and achieves modal insensitivity.

The high level of abstraction required in the following derivations is unfortunate but necessary for development of the general characteristics of the problem. It should also be noted that the implicit computational requirements for the following approach is nontrivial! The development here parallels the work in [37].

We begin by recalling from a generalization of theorem 3.3.2 to output feedback [39] while noting that full state feedback is a special case. It states that the set of attainable eigenvectors associated with r prespecified eigenvalues can be characterized by

$$\begin{Bmatrix} x_i \\ w_i \end{Bmatrix} \in \quad \text{span} \begin{bmatrix} N_{\lambda_i} \\ M_{\lambda_i} \end{bmatrix} \tag{3.3.11}$$

where

$$w_i = -KHx_i \qquad i = 1,\ldots,r \tag{3.3.12}$$

and r, K and H represents the number of output, output gain matrix and measurement matrix respectively. The assumption that the eigenpairs are self-conjugate and (r x r) matrix HX is invertible is implicit in the above. We also recall from theorem 3.3.5 that the conditions for modal insensitivity for ν modes with respect to a single parameter, ρ, are

$$\frac{\partial \bar{A}}{\partial \rho} x_i = 0 \quad ; \quad i=1,\ldots,\nu \tag{3.3.13}$$

where $\bar{A}$ represents the closed loop system matrix,

$$\bar{A} = A + BKH \ .$$

Suppose we desire that ν ($\leq r$) modes be made insensitive. Then, for the class of problems where the uncertain parameter appears only in the plant matrices A, B and H, Eq.(3.3.13) becomes

$$\left(\frac{\partial A}{\partial \rho} + \frac{\partial B}{\partial \rho} KH + BK \frac{\partial H}{\partial \rho}\right) x_i = 0 \qquad (3.3.14)$$

$$i=1,\ldots,\nu$$

A sufficient condition

$$\begin{bmatrix} \frac{\partial A}{\partial \rho} & -\frac{\partial B}{\partial \rho} \\ \frac{\partial H}{\partial \rho} & 0 \end{bmatrix} \begin{Bmatrix} x_i \\ w_i \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \qquad (3.3.15)$$

$$i=1,\ldots,\nu \ .$$

We next define a basis for the null space of Eq. (3.3.15) so that all solutions of Eq. (3.3.15), i.e. all eigenvector sets satisfying modal insensitivity condition with respect to a single parameter, must satisfy

$$\begin{Bmatrix} x_i \\ w_i \end{Bmatrix} \in \text{span} \begin{bmatrix} P \\ Q \end{bmatrix} \qquad (3.3.16)$$

$$i=1,\ldots,\nu$$

where P and Q are appropriately partitioned matrices of basis vectors for the null space.

With the above definitions, we deduce that to simultaneously assign r eigenvalues while attaining modal insensitivity of ν ($\leq r$) modes, the following conditions must be satisfied by the eigenvectors:

$$\begin{Bmatrix} x_i \\ w_i \end{Bmatrix} \in \operatorname{span}\left\{ \begin{bmatrix} N_{\lambda_i} \\ M_{\lambda_i} \end{bmatrix} \cap \begin{bmatrix} P \\ Q \end{bmatrix} \right\} \qquad (3.3.17)$$

$$i=1,\dots,\nu$$

$$\begin{Bmatrix} x_i \\ w_i \end{Bmatrix} \in \operatorname{span} \begin{bmatrix} N_{\lambda_i} \\ M_{\lambda_i} \end{bmatrix} \qquad (3.3.18)$$

$$i=\nu+1,\dots,r \ .$$

The intersection space appearing in Eq. (3.3.17) is called "modal insensitivity subspace" for i-th mode in [37]. A direct approach for computing the above subspace as given in [36] is to find sets of basis vectors,

$$\begin{bmatrix} L_{1_i} \\ L_{2_i} \end{bmatrix} \quad \text{of rank } a^i$$

for

$$\operatorname{Ker} \begin{bmatrix} \lambda_i I-A & B \\ \dfrac{\partial A}{\partial \rho} & -\dfrac{\partial B}{\partial \rho} \\ \dfrac{\partial H}{\partial \rho} & 0 \end{bmatrix} \qquad (3.3.19)$$

$$i=1,\dots,\nu$$

to place ν eigenvalues and attain modal insensitivity. By the previous definitions of basis vectors for Eqs. (3.3.17) and (3.3.18), the eigenvectors must now satisfy

$$\left.\begin{aligned} x_i &= L_{1_i} u_i \\ w_i &= L_{2_i} u_i \end{aligned}\right\} i=1,\ldots,\nu \tag{3.3.20}$$

and

$$\left.\begin{aligned} x_i &= N_{\lambda_i} v_i \\ w_i &= M_{\lambda_i} v_i \end{aligned}\right\} i=\nu+1,\ldots r \tag{3.3.21}$$

where u_i $(i=1,\ldots\nu)$ and v_i $(i=\nu + 1,\ldots,r)$ are arbitrary vectors of dimensions a^i and b^i respectively where

$$a^i = \text{rank}\begin{bmatrix} L_{1_i} \\ L_{2_i} \end{bmatrix} \quad ; \quad i=1,\ldots,\nu$$

and

$$b^i = \text{rank}\begin{bmatrix} N_{\lambda_i} \\ M_{\lambda_i} \end{bmatrix} \quad ; \quad i=\nu+1,\ldots,r .$$

Since x_i and w_i are related by Eq. (3.3.12), the feedback gain matrix K can be computed by substituting Eqs. (3.3.20) and (3.3.21) into Eq. (3.3.12) to obtain

$$-KHL_{1_i} u_i = L_{2_i} u_i \quad ; \quad i=1,\ldots,\nu$$

$$-KHN_{\lambda_i} v_i = M_{\lambda_i} v_i \quad ; \quad i=\nu+1,\ldots,r$$

$$\Longrightarrow \quad KH\,[L_{11}u_1,\ldots,L_{1\nu}u_\nu,\ N_{\lambda_{\nu+1}}v_{\nu+1},\ldots,N_{\lambda_r}v_r]$$

$$= -[L_{21}u_1,\ldots,L_{2\nu}u_\nu,\ M_{\lambda_{\nu+1}}v_{\nu+1},\ldots,M_{\lambda_r}v_r] \tag{3.3.22}$$

In summary, the problem of assigning r eigenvalues and attaining modal insensitivity of ν ($\leq r$) modes using constant output feedback reduces to the following steps:

(a) select r eigenvalues to be placed which are self-conjugate,

(b) compute a set of basis vectors to span "modal insensitivity subspace" for ν modes, i.e. Eq. (3.3.19),

(c) compute a set of basis vectors to span "modal assignability space" of $(r-\nu)$ modes as given by Eq. (3.3.18),

(d) select r eigenvectors by selecting u_i and v_i in Eqs. (3.3.20) and (3.3.21)

(e) compute gain matrix K from Eq. (3.3.22).

We observe from step (e) that the output feedback gain matrix, K will exist provided that the $(r \times r)$ matrix $H\,[x_1,\ldots,x_r]$ is nonsingular, i.e., the eigenvectors should not be unobservable. Step (d) assures us simultaneous modal insensitivity and eigenvalue assignment. In addition, we observe that we may still have some freedom left in the choice of eigenvectors in step (d). This remaining freedom on the choice of eigenvectors can obviously be used for satisfying other design requirements. In particular, [37] utilizes this freedom to select an "optimal" output feedback gain configuration in the linear quadratic sense which amounts to imposing some penalty for indiscriminate use of control. This is indeed a useful criteria since a random selection of the eigenvectors may lead to unrealistically large values in elements of the feedback gain matrix. We also note that the modal insensitivity imposed here is with respect to a single parameter. It is clear then that the problem of eigenstructure assignment with modal insensitivity with respect to several parameters should be challenging indeed! There is ample room for new analytical algorithm development in this area.

Finally, it should be mentioned that various complicating but nevertheless important issues such as sufficient conditions for existence of a modal insensitivity subspace, imposing partial modal insensitivity, conditions for modal insensitivity for the case of non-distinct eigenvalues and improvements with using dynamic compensators are also discussed in reference |37|.

3.4 Eigenvalue Sensitivity Norm

In this section, we consider the problem of quantifying the sensitivities of eigenvalues due to parameter variations. This section originated from the author's desire to explore the significance of the eigenvalue sensitivity matrix vis-a-vis robust control, eigenvalue placement/sensitivity, and related issues. We assume here that first derivatives of eigenvalues with respect to specified parameters are continuous. Eigenvalue sensitivity is important in many real-world situations where imperfect mathematical models (which are usually parameterized) must be used, and a design goal is to minimize the sensitivity of an apriori specified transient response performance with respect to a selected set of uncertain parameters.

Although both eigenvalues and eigenvectors determine the total transient response of linear systems, eigenvalue locations completely determine the relative stability of a system. The shape of the transient response as determined by the eigenvectors is of secondary importance only since the sensitivity of stability is of concern here. Hence, eigenvector sensitivity will not be considered directly in the following discussions. However, the two are related by a choice of an eigenvector set as clearly evident from developments in section 3.6 (c.f., Eq. 3.6.15).

For a general n-th order system with ℓ parameters considered uncertain, we assume that we have available a matrix of $(n \times \ell)$ partial derivatives representing the sensitivities of every eigenvalue with respect to every uncertain parameter. We consider here the minimization of a weighted norm of the eigenvalue sensitivity matrix with respect to a set of design variables. However, as required in any standard optimization problem, a scalar index which captures the relative weights and hence the overall eigenvalue sensitivity is highly desirable. One such direct measure of sensitivity is the quadratic

$$J(p) \triangleq \sum_{i=1}^{n} \sum_{j=1}^{\ell} \left| \frac{\partial \lambda_i}{\partial p_j} \right|^2 w_{ij} \tag{3.4.1}$$

where w_{ij} are nonnegative weighting factors associated with the sensitivity of the i-th eigenvalue with respect to j-th parameter. From the definition, it is obvious that the minimum J can take is zero. For purposes of direct minimization, Eq. (3.4.1) can easily be linearized about a nominal design point, p^c , to obtain

$$J(p) = J(p^c) + 2 \sum_{j=1}^{\ell} S_j(p^c) \cdot \Delta p_j \tag{3.4.2}$$

where

$$S_j(p_c) = \text{Re} \sum_{k=1}^{\ell} \sum_{i=1}^{n} \left(\frac{\partial\lambda_i}{\partial p_k}\right)^* \left(\frac{\partial^2\lambda_i}{\partial p_k \partial p_j}\right) \cdot w_{ik} \tag{3.4.3}$$

$$p = p^c + \Delta p .$$

The algebraic details necessary to arrive at Eqs. (3.4.2) and (3.4.3) are given in [20]. Furthermore, reference [20] demonstrates the feasibility of using such an index of sensitivity and its local linearizations to attain a minimum sensitivity configuration while simultaneously placing eigenvalues.

Next, we examine closely the significance of a matrix of eigenvalue sensitivities using concepts from matrix operator norm theory. We begin by writing the linearly predicted changes in n-eigenvalues due to small changes in ℓ-parameters as

$$\begin{Bmatrix} \Delta\lambda_1 \\ \vdots \\ \Delta\lambda_n \end{Bmatrix} = \begin{bmatrix} \dfrac{\partial\lambda_1}{\partial p_1} & \cdots & \dfrac{\partial\lambda_1}{\partial p_\ell} \\ \dfrac{\partial\lambda_n}{\partial p_1} & \cdots & \dfrac{\partial\lambda_n}{\partial p_\ell} \end{bmatrix} \begin{Bmatrix} \Delta p_1 \\ \vdots \\ \Delta p_\ell \end{Bmatrix}$$

or $$\Delta\lambda = G \, \Delta p \tag{3.4.4}$$

where G represents the matrix of eigenvalue sensitivities

$$\underset{(n\times\ell)}{G} = \begin{bmatrix} \frac{\partial\lambda_1}{\partial p_1} & \cdots & \frac{\partial\lambda_1}{\partial p_\ell} \\ \vdots & \ddots & \vdots \\ \frac{\partial\lambda_n}{\partial p_1} & \cdots & \frac{\partial\lambda_n}{\partial p_\ell} \end{bmatrix}$$

Let us next define a weighted vector norm of eigenvalue change as

$$\|\Delta\lambda\|^W \triangleq (\Delta\lambda^H W \Delta\lambda)^{\frac{1}{2}} \tag{3.4.5}$$

where the weight matrix W can be taken as a designer specified positive definite symmetric matrix which weighs the sensitivity of individual modes. We also choose to normalize our parameter changes by

$$\delta_i = \frac{\Delta p_i}{|p_i|} \quad ; \quad i=1,\ldots,\ell \tag{3.4.6}$$

where p_i represents some nominal value for i-th parameter or in general, weights for the sensitivity of individual parameters, and rewrite Eq. (3.4.6) as a linear transformation

$$\Delta p = \Theta\ \delta \tag{3.4.7}$$

where

$$\Theta \triangleq \text{diag}\ \{\ |p_1|,\ldots,|p_\ell|\ \}$$

We note that δ represents an ℓ-vector of non-dimensionalized or normalized parameter change.

By using Eqs. (3.4.7) and (3.4.4), we obtain from Eq. (3.4.5) the following:

$$\begin{aligned}(\|\Delta\lambda\|^{W})^2 &= \Delta\lambda^H\ W\ \Delta\lambda \\ &= \delta^H\ \Theta^H\ G^H\ W\ G\ \Theta\ \delta \\ &= \delta^H\ Q\ \delta \qquad (3.4.8)\end{aligned}$$

where

$$Q \triangleq \Theta^H\ G^H\ W\ G\ \Theta\ .$$

Since W was assumed symmetric, Q is hermitian and by Rayleigh's principle [53]

$$\lambda_{min}(Q) \le \frac{\delta^H\ Q\ \delta}{\delta^H\ \delta} \le \lambda_{max}(Q)$$

and an upper bound for Eq. (3.4.8) is

$$(\|\Delta\lambda\|^{W})^2 \le \lambda_{max}(Q)\ \delta^H\ \delta \qquad (3.4.9)$$

By writing Eq. (3.4.9) for any normalized perturbation vector $\bar{\delta}$ where

$$\bar{\delta}^H\ \bar{\delta} \le \delta^H\ \delta$$

we obtain

$$\begin{aligned}(\|\Delta\lambda\|^{W})^2 &\le \lambda_{max}(Q)\ \bar{\delta}^H\ \bar{\delta} \\ &\le \lambda_{max}(Q)\ \delta^H\ \delta\end{aligned}$$

By induction, we conclude that $\lambda_{max}(Q)$ represents an upper bound on the square of weighted eigenvalue error norms, $\|\Delta\lambda\|^{W}$, for all normalized perturbations δ satisfying

$$\delta^{H} \delta \leq 1 \quad .$$

Investigating further, we note that since W is assumed to be symmetric and positive definite, it can be decomposed by Cholesky factorization [53]

$$W = LL^{T} \tag{3.4.10}$$

As a result of Eq. (3.4.10), Q of Eq. (3.4.8) can be rewritten as

$$Q = (L^{T}G\Theta)^{H} (L^{T}G\Theta) \geq 0$$

so that

$$\begin{aligned} \lambda_{max}(Q) &= \bar{\sigma} (L^{T}G\Theta) \\ &= \|L^{T}G\Theta\|_2 \end{aligned} \tag{3.4.11}$$

where $\bar{\sigma}$ represents the maximum singular value. In words, Eq. (3.4.11) represents a convenient scalar index for a weighted and normalized measure of eigenvalue sensitivity. For a typical application, L and Θ may be specified by the designer while the elements of G may be iteratively driven so as to minimize the overall sensitivity measure in Eq. (3.4.11). We emphasize here the physical significance of L and Θ terms.

For the special case where W and Θ are identity matrices, the above sensitivity measure reduces to the simpler form

$$\lambda_{max}(Q) = \bar{\sigma}(G) = \lambda_{max}(G^H G) \; .$$

It should be emphasized that although both Eqs. (3.4.1) and (3.4.11) represents sensitivity indices, only Eq. (3.4.11) is directly related to a linearly predicted bound on weighted and normalized eigenvalue perturbation. It should also be mentioned that the above derivations are analogous to the concept of matrix operator norms (see for example p. 163- of [54]).

In summary, Eq. (3.4.11) represents a convenient scalar index that is suitable for the use as a cost function in the problem of eigenvalue sensitivity minimization. Some new interpretations of eigenvalue sensitivity matrix are also established.

3.5 Conditioning of Eigenvalue Problems

3.5.1 Basic Theory

In numerical analysis, whenever "small" changes in the data can lead to "large" changes in the solution, a problem is said to be "ill-conditioned"; otherwise it is said to be "well-conditioned". Exactly how changes are measured and what "small" and "large" mean will vary with the problem and the choice of vector and matrix norms used. In general, the degree of ill-conditioning of a problem can be quantified by a scalar number related directly to an upper bound on the perturbation of the solution called a "condition number" [53,54]. We carefully note here that a condition number represents the upper bound only, meaning that a moderate condition number implies well-conditioning but the converse is not true, i.e., a large condition number does not guarantee ill-conditioning. In other words, for some cases, an upper bound may be too conservative and may not necessarily be a realistic estimate of actual errors (this familiar tune can also be heard in bounds associated with robust control, see Chapter 4).

In the area of matrix computations and numerical analysis, condition numbers for both inversion and eigenvalue problems are well established. For our purposes, we focus on the conditioning of a general eigenvalue problem. In particular, we would like to consider here the sensitivities of eigenvalues with respect to perturbations in the original matrix. We limit ourselves to the "condition" of the closed loop eigenvalues only since they completely characterize the asymptotic stability of a feedback control system.

We now review two results from matrix theory (see p. 292-, |54|) that establishes the concept of conditioning of the eigenvalues:

THEOREM 3.5.1

Suppose an (nxn) matrix A has a linearly independent set of n eigenvectors (i.e. A is "non-defective") associated with eigenvalues $\lambda_1,\ldots,\lambda_n$. If μ and v, where $\|v\| = 1$, represents an approximate eigenpair of A, then at least one eigenvalue of A satisfies

$$|\lambda_i - \mu| \le \|r\| \; \|P\| \; \|P^{-1}\| \tag{3.5.1}$$

where

$r \triangleq Av - \mu v$ = error vector of (μ, v) pair

$P \triangleq |x_1,\ldots,x_n|$ = true modal matrix of A .

Furthermore, if P is unitary, then at least one eigenvalue of A satisfies

$$|\lambda_i - \mu| \le \|r\|_2 \ . \qquad (3.5.2)$$

PROOF OF THEOREM 3.5.1

Since A was assumed to have a linearly independent set of n eigenvectors, P, we can write

$$A = P\Lambda P^{-1} \quad \text{where} \quad \Lambda = \text{diag}\ (\lambda_1,\ldots,\lambda_n) \ .$$

From definition of error vector

$$\begin{aligned} r &= P\Lambda P^{-1}v - \mu PP^{-1}v \qquad (3.5.3) \\ &= P\ (\Lambda - \mu I)P^{-1}v \end{aligned}$$

For the trivial case where μ equals one of the eigenvalues of A, clearly Eq. (3.5.1) is automatically satisfied. If μ does not equal one of eigenvalues of A, $(\Lambda - \mu I)$ is nonsingular and Eq. (3.5.3) can be inverted to get

$$v = P(\Lambda - \mu I)^{-1}P^{-1}r \ .$$

By taking any operator norm of both sides and using normalization, $\|v\| = 1$

$$1 = \|v\| \ \le \ \|P\|\ \|(\Lambda - \mu I)^{-1}\|\ \|P^{-1}\|\ \|r\| \ . \qquad (3.5.4)$$

Since $(\Lambda - \mu I)$ is diagonal

$$(\Lambda - \mu I)^{-1} = \text{diag}\ \left(\frac{1}{\lambda_1 - \mu}, \ldots, \frac{1}{\lambda_n - \mu}\right)$$

so that for any $(1,2,\infty)$ matrix operator norms,

$$\|(\Lambda - \mu I)^{-1}\| \ = \ \max_i |\lambda_i - \mu|^{-1} = [\min_i |\lambda_i - \mu|]^{-1}. \qquad (3.5.5)$$

From Eqs. (3.5.4) and (3.5.5)

$$\min_i |\lambda_i - \mu| \leq \|r\| \, \|P\| \, \|P^{-1}\|$$

thus verifying Eq. (3.5.1). If P is unitary, then so is P^{-1} and since the 2-norms of all unitary matrices equals unity, condition Eq. (3.5.2) follows. Q.E.D.

The condition number for the eigenvalue problem is defined as

$$c(P) \triangleq \|P\| \, \|P^{-1}\|$$

and we note that it depends only on the unperturbed A matrix. It is obvious from the above theorem that the condition number as defined above is directly related to an upper bound on eigenvalue estimate errors. We see from the above that the larger the condition number, the more difficult it will be to estimate the true eigenvalues.

For the purpose of further clarifying the concept of conditioning, we restate from [58] a fundamental but important distinction between the conditioning of a problem and the stability of a numerical algorithm. The former, which is of interest here, is an inherent property of a given problem whereas the latter is a property of an algorithm designed (not unique!) to solve the given problem numerically. It should be clear that "numerically stable" algorithms do not resolve the ill-conditioning of a problem but is structured to avoid introducing more sensitivity to perturbation than is already inherent in the given problem. For further details on this interesting problem, the following references are recommended [53,58,59].

The previous theorem provides an upper bound on the error of an eigenvalue estimate corresponding to an approximating eigenpair. It can be observed that if the condition number is large, i.e. "ill-conditioned", then the approximating eigenpair will probably not be a good estimate of the unperturbed eigenvalue. To arrive at an error bound on the eigenvalues due to errors in matrix A, and thus be more useful for our purposes, we state a second theorem (see p. 294-, [54]) which is based on the previous theorem:

THEOREM 3.5.2

Let A be nondefective and the eigenvalue and eigenvector matrices be written as

$$\Lambda = \text{diag}\,(\lambda_1,\ldots,\lambda_n)$$

and

$$P = [x_1,\ldots,x_n]\ .$$

If μ and v (where $\|v\| = 1$) is an eigenpair of a perturbed matrix, A+E, then, at least one eigenvalue of A satisfies

$$|\lambda_i - \mu| \le \|E\|\ \|P\|\ \|P^{-1}\| \tag{3.5.6}$$

Furthermore, if A is normal, then at least one eigenvalue of A satisfies

$$|\lambda_i - \mu| \le \|E\|_2\ . \tag{3.5.7}$$

PROOF OF THEOREM 3.5.2

Since μ and v (where $\|v\| = 1$) are defined as an eigenpair of A+E,

$$(A + E)v = \mu v$$

$$Av - \mu v = -Ev \triangleq r\ .$$

Since A was assumed nondefective, we can apply Theorem 3.5.1 so that at least one eigenvalue of A satisfies

$$|\lambda_i - \mu| \le \|-Ev\| \ \|P\| \ \|P^{-1}\|$$

$$\le \|E\| \ \|P\| \ \|P^{-1}\| \ .$$

If A is normal, P and P^{-1} are unitary, and since the 2-norms of unitary matrices equals unity, the condition of Eq. (3.5.7) follows. Q.E.D.

As a consequence of Theorem 3.5.2, the following holds:

$$\begin{aligned} \min_i |\lambda_i - \mu_1| &\le \|E\| \ c(P) \\ &\ \vdots \\ \min_i |\lambda_i - \mu_n| &\le \|E\| \ c(P) \end{aligned} \qquad (3.5.8)$$

where $\lambda_1,\ldots,\lambda_n$ and $\mu_1,\ldots,\mu_n$ represents the eigenvalues of A and A+E respectively. Geometrically, Eq. (3.5.8) states that the distance from any particular perturbed eigenvalue to its nearest unperturbed eigenvalue is bounded equally for every perturbed eigenvalue. We observe that in the limit when $\|E\|$ approaches zero,

$$\min_i |\lambda_i - \mu_j| \le |\lambda_j - \mu_j| \qquad j=1,\ldots,n$$

i.e. disks of perturbations will be disjoint. It is then clear that the special case when A has distinct eigenvalues and E is sufficiently small, i.e.

$$\|E\| \le \tfrac{1}{2} \min_{\substack{i,j \\ i \ne j}} \frac{|\lambda_i - \lambda_j|}{c(P)}$$

all disks of perturbations about unperturbed eigenvalues will be disjoint and have the same radii, i.e.

$$|\lambda_i - \mu_i| \leq \|E\| \; c(P) \qquad i=1,\ldots,n \tag{3.5.9}$$

For the above special case, the bound in Eq. (3.5.9) have a very simple geometrical interpretation, i.e., we can imagine circles of equal radii error bounds about each unperturbed eigenvalue such that the resulting eigenvalues due to any matrix perturbation in A bounded by $\|E\|$ will remain in a disk of equal radii about the unperturbed eigenvalues. The significance of the condition number is clear, it directly establishes the radius of uncertainty within which all eigenvalues are perturbed due to an error E in the A matrix.

3.5.2 A Stability Robustness Criteria

Let us consider the use of eigenvalue bounds, as previously derived in the context of eigenvalue problem conditioning in Theorems 3.5.1 and 3.5.2, to obtain asymptotic stability bounds. If we assume that the nominal (unperturbed) system, A, is asymptotically stable, its eigenvalues will all lie on the left side of the imaginary axis in eigenspace. The problem then is to find the upper bound on $\|E\|$ which guarantees that all perturbed eigenvalues remain in the left half plane. Let us consider all E satisfying

$$\|E\| \; \|P\| \; \|P^{-1}\| \; < \; \min_k |\mathrm{Re}\lambda_k(A)| \quad . \tag{3.5.10}$$

Then by Eq. (3.5.8) of Theorem 3.5.2,

$$\min_i |\lambda_i - \mu_j| < \min_k |Re\lambda_k(A)| \qquad j=1,\ldots,n \ . \qquad (3.5.11)$$

The condition Eq. (3.5.11) states that the distance from every perturbed eigenvalue to the closest unperturbed eigenvalue is always less than the perpendicular distance of "dominant" unperturbed eigenvalue to the imaginary axis. This implies that all the perturbed eigenvalues will remain in the left half plane, thus maintaining asymptotic stability. From Eq. (3.5.10), it follows that for all perturbation, E satisfying

$$\|E\| < \frac{\min_k |Re\lambda_k(A)|}{c(P)} \qquad (3.5.12)$$

the system, A+E remains asymptotically stable. It is now obvious from Eq. (3.5.12) that its right hand side represents a measure of stability robustness.

We conclude this section by pointing out that the robustness measure of Eq. (3.5.12) can be arrived at from a completely different approach using Lyapunov stability theorem and several other lemma, as will be shown in chapter 4. However, the derivation presented here is relatively more concise and conceptually more straightforward than Patel and Toda's derivation as given in reference [30]. This indeed contributes to the understanding of the relationship between Patel and Todas' stability robustness criteria and the well established eigenvalue conditioning concepts in matrix theory and numerical analysis.

3.6 A Re-examination of Eigenvector Derivatives

3.6.1 The Eigenvalue Problem

The usefulness of modal sensitivities for analysis and design of engineering systems is well known. Some specific application includes identification of dynamical systems [60,61], re-design of vibratory systems [13,21,63] and design of control systems by pole placement [11,12,14,20]. In the above algorithms, eigenvalue and eigenvector derivatives with respect to design parameters are often required. In the literature the development of the derivatives appears in the following order: Wilkinson [62] present a clear derivations of first order perturbation equations for a general real matrix. Wittrick [65] derived the first derivatives of eigenvalues while Fox and Kapoor [66] extended it to include the first derivatives of eigenvectors for real symmetric systems. The first general expressions for eigenvalue and eigenvector derivatives (using modal expansion approach) for non-self-adjoint systems appears to be given by Plaut and Huseyin [67]; in fairness to these authors, their formulation is correct but incomplete. In essence, the derivation presented here provides the completion of their developments. It should be mentioned here that the first correct expression for eigenvector derivatives using the modal expansion appears to be the results published by Rogers [81]. Additionally, we show in this section a simple relationship between the left and right eigenvector derivative expansion coefficients. Another purpose of this section is to help clear up some confusion in the literature as well be evident in the sequel.

This section is most concerned with a correction of an assumption made in the several completed derivations of equations for the eigenvector derivatives (e.g. as found in [62,64,68]) and resolving an indeterminancy remaining in [67]. A careful re-examination of the role played by the normalizations and bi-orthogonality for a general non-self-adjoint eigenvalue problem leads to a unique generalization and correction of equations for eigenvector derivatives derived in previous papers and books cited. We begin here by reviewing the familiar eigenvalue problem. Let us write the right and left eigenvalue problems as

$$\text{right:} \quad Ax_j = \lambda_j x_j \quad ; \quad j=1,\ldots,n \tag{3.6.1}$$

$$\text{left:} \quad y_i^T A = \lambda_i y_i^T \quad ; \quad i=1,\ldots,n \ . \tag{3.6.2}$$

Here A is an nxn real matrix, x,y, denote nx1 complex-valued eigenvectors, and λ is a generally complex eigenvalue of A. Premultiplying Eq. (3.6.1) by y_i^T and postmultiplying Eq. (3.6.2) by x_j, we get

$$y_i^T A x_j = \lambda_j y_i^T x_j \tag{3.6.3}$$

$$y_i^T A x_j = \lambda_i y_i^T x_j \ . \tag{3.6.4}$$

By subtracting Eq. (3.6.4) from Eq. (3.6.3) we get

$$0 = (\lambda_j - \lambda_i)\, y_i^T x_j \;. \tag{3.6.5}$$

Restricting ourselves to the class of problems where the eigenvalues are assumed distinct, Eq. (3.6.5) leads to the biorthogonality property

$$y_i^T x_j = 0 \quad ; \quad i \neq j \;. \tag{3.6.6}$$

For i = j, the left and right eigenvectors can be normalized such that

$$y_i^T x_i = s_i \quad ; \quad i=1,\ldots,n \tag{3.6.7}$$

where s_i represents chosen normalization constants, commonly set to unity. At this point we note the fact that biorthogonality of Eq. (3.6.6) is a property of the eigensystem which is completely independent of an apriori normalizations (of x_i or y_j) which is also arbitrary.

In matrix notation, Eqs. (3.6.4), (3.6.6) and (3.6.7) can be written as the pair of matrix equations

$$Y^T A X = S\Lambda \tag{3.6.8}$$

$$Y^T X = S \tag{3.6.9}$$

where

$$S = \mathrm{diag}\,(s_1,\ldots,s_n)$$

$$\Lambda = \mathrm{diag}\,(\lambda_1,\ldots,\lambda_n)$$

and X, Y are nxn right and left modal matrices (containing x_j and y_i, respectively as columns). We note that from Eq. (3.6.9), we could solve for X or Y since

$$Y = (X^{-1})^T S \tag{3.6.10}$$

$$X = (Y^{-1})^T S \ . \tag{3.6.11}$$

From the above equations, it is apparent that after solving the right eigenvalue problem and then imposing arbitrary normalizations on the columns x_j of X, we still need to choose S (i.e. second set of n-normalization constants) to solve for Y from Eq. (3.6.10) uniquely. Similarly if the left eigenvalue problem was solved instead, then X is computed from Eq. (3.6.11). In other words, the two sets of eigenvectors, X and Y, requires two independent, arbitrary sets of normalizations for their unique representations. We stress here the fact that simply requiring that S=I, as is common practice, does not uniquely scale X and Y for non-self-adjoint systems. Henceforth, we let the two sets of normalizations be represented by

$$x_i^T x_i = f_i \quad ; \quad i=1,\ldots,n \tag{3.6.12}$$

$$y_i^T x_i = s_i \quad ; \quad i=1,\ldots,n \tag{3.6.13}$$

where f_i and s_i, i=1,...,n, are fixed normalization constants.

3.6.2 Eigenvalue Derivatives

For completeness, we rederive the eigenvalue derivatives with respect to a scalar parameter, ρ . First, we take partial derivatives of Eq. (3.6.1) to get

$$\frac{\partial A}{\partial \rho} x_j + A \frac{\partial x_j}{\partial \rho} = \frac{\partial \lambda_j}{\partial \rho} x_j + \lambda_j \frac{\partial x_j}{\partial \rho} \tag{3.6.14}$$

We then premultiply the above equation by y_j^T to obtain

$$y_j^T \frac{\partial A}{\partial \rho} x_j + y_j^T A \frac{\partial x_j}{\partial \rho} = \frac{\partial \lambda_j}{\partial \rho} y_j^T x_j + \lambda_j y_j^T \frac{\partial x_j}{\partial \rho}$$

$$y_j^T \frac{\partial A}{\partial \rho} x_j + y_j^T (A - \lambda_j I) \frac{\partial x_j}{\partial \rho} = \frac{\partial \lambda_j}{\partial \rho} y_j^T x_j$$

so that by Eq. (3.6.2), the eigenvalue derivative expression takes the form

$$\frac{\partial \lambda_j}{\partial \rho} = \frac{1}{s_j} (y_j^T \frac{\partial A}{\partial \rho} x_j) \tag{3.6.15}$$

where s_j is the j-th normalization constant chosen in Eq. (3.6.13).

3.6.3 Eigenvector Derivatives

Here we derive an expression for eigenvector derivatives along the lines of reference [13], by using $\{x_1, \ldots, x_n\}$ as basis vectors

$$\frac{\partial x_i}{\partial \rho} = \sum_{j=1}^{n} \alpha_{ij} x_j \ . \tag{3.6.16}$$

Since the eigenvalues were assumed distinct, then the eigenvectors are linearly independent, so these may be used as a set of basis vectors to spanning the complex n-dimensional space. To obtain expansion coefficients , α_{ij},i,j = 1,...,n, we begin by substituting Eq. (3.6.16) into Eq. (3.6.14) and premultiply by y_k^T to get

$$y_k^T \frac{\partial A}{\partial \rho} x_i + \sum_{j=1}^{n} \alpha_{ij}\lambda_j y_k^T x_j = \frac{\partial \lambda_i}{\partial \rho} y_k^T x_i + \lambda_i \sum_{j=1}^{n} \alpha_{ij} y_k^T x_j$$

and by using biorthogonality conditions, we obtain

$$y_k^T \frac{\partial A}{\partial \rho} x_i + \alpha_{ik}\lambda_k y_k^T x_k = \frac{\partial \lambda_i}{\partial \rho} y_k^T x_i + \lambda_i \alpha_{ik} y_k^T x_k$$

$$(\lambda_k - \lambda_i)\alpha_{ik} y_k^T x_k = \frac{\partial \lambda_i}{\partial \rho} y_k^T x_i - y_k^T \frac{\partial A}{\partial \rho} x_i \quad .$$

Since A was assumed to have distinct eigenvalues, we solve for α_{ik} as

$$\alpha_{ik} = - \frac{1}{(\lambda_k - \lambda_i)s_k} \left(y_k^T \frac{\partial A}{\partial \rho} x_i\right) \quad ; \quad k \neq i \qquad (3.6.17)$$

For the case of k=i, α_{ii} can be computed from the normalization condition of Eq. (3.6.12). By taking its partial derivatives and using Eq. (3.6.16) we find

$$\frac{\partial x_i^T}{\partial \rho} x_i = 0$$

$$\Longrightarrow \sum_{j=1}^{n} \alpha_{ij} x_j^T x_i = 0$$

$$\Longrightarrow \quad \alpha_{ii} x_i^T x_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} \alpha_{ij} x_j^T x_i = 0$$

which leads to the expression

$$\alpha_{ii} = -\frac{1}{f_i} \sum_{\substack{j=1 \\ j \neq i}}^{n} \alpha_{ij} x_j^T x_i \tag{3.6.18}$$

where f_i is the i-th normalization constant chosen in Eq. (3.6.12). It can be deduced from Eq. (3.6.18) that the diagonal terms, α_{ii}, becomes identically zero if:

(i) if A has orthogonal eigenvectors (for example: A is real and symmetric) or

(ii) if the i-th eigenvector is insensitive with respect to a particular parameter except in the i-th eigenvector direction.

We can also observe from Eqs. (3.6.17) and (3.6.18) that the off-diagonal and diagonal coefficients α_{ik} and α_{ii} depend on the normalization constants s_k and f_i respectively.

To obtain left eigenvector derivatives, we similarly let

$$\frac{\partial y_i}{\partial \rho} = \sum_{j=1}^{n} \gamma_{ij} y_j \ . \tag{3.6.19}$$

By taking partial derivatives of Eq. (3.6.2), substituting Eq. (3.6.19) and postmultiplying by x_k, we obtain

$$\sum_{j=1}^{n} \gamma_{ij} y_j^T A x_k + y_i^T \frac{\partial A}{\partial \rho} x_k = \frac{\partial \lambda_i}{\partial \rho} y_i^T x_k + \lambda_i \sum_{j=1}^{n} \gamma_{ij} y_j^T x_k \ .$$

By using Eqs. (3.6.1) and (3.6.6), we get

$$\gamma_{ik}\lambda_k y_k^T x_k + y_i^T \frac{\partial A}{\partial \rho} x_k = \frac{\partial \lambda_i}{\partial \rho} y_i^T x_k + \lambda_i \gamma_{ik} y_k^T x_k$$

$$\Longrightarrow \quad (\lambda_k - \lambda_i)\gamma_{ik} y_k^T x_k = \frac{\partial \lambda_i}{\partial \rho} y_i^T x_k - y_i^T \frac{\partial A}{\partial \rho} x_k$$

$$\Longrightarrow \quad \gamma_{ik} = - \frac{1}{(\lambda_k - \lambda_i) s_k} \left(y_i^T \frac{\partial A}{\partial \rho} x_k \right) \quad ; \quad k \neq i \qquad (3.6.20)$$

By comparing Eqs. (3.6.17) and (3.6.20), we observe the antisymmetric property

$$\gamma_{ik} = - \alpha_{ki} \quad ; \quad k \neq i \qquad (3.6.21)$$

For k=i, γ_{ii} and α_{ii} are related through the normalization condition of Eq. (3.6.13). Thus by taking the partial derivative of Eq. (3.6.13), we get

$$\frac{\partial y_i^T}{\partial \rho} x_i + y_i^T \frac{\partial x_i}{\partial \rho} = 0$$

and by using expansions of Eqs. (3.6.16) and (3.6.19)

$$\sum_{j=1}^{n} \gamma_{ij} y_j^T x_i + y_i^T \sum_{j=1}^{n} \alpha_{ij} x_j = 0$$

$$\gamma_{ii} y_i^T x_i + \alpha_{ii} y_i^T x_i = 0$$

$$\gamma_{ii} = - \alpha_{ii} \, . \qquad (3.6.22)$$

From Eqs. (3.6.21) and (3.6.22), we observe the nice relationship between the left and right eigenvector derivative expansion coefficients:

$$[\gamma] = - [\alpha]^T \tag{3.6.23}$$

provided, of course, that the normalization constants, s_k, are chosen the same values.

The above procedure can be extended to the generalized eigenvalue problem of the form

$$Ax_i = \lambda_i B x_i$$

$$A^T y_i = \lambda_i B^T y_i \quad ; \quad i=1,\ldots,n$$

with biorthogonality and normalizations

$$y_k^T B x_i = \delta_{ik} \quad ; \quad i,k=1,\ldots,n$$

$$x_i^T B x_i = 1 \quad ; \quad i=1,\ldots,n \quad .$$

After some algebra analogous to the above developments, it can be shown that the eigenvalue and eigenvector derivatives take the form

$$\frac{\partial \lambda_i}{\partial \rho} = y_i^T \left(\frac{\partial A}{\partial \rho} - \lambda_i \frac{\partial B}{\partial \rho}\right) x_i$$

$$\frac{\partial x_i}{\partial \rho} = \sum_{k=1}^{n} \alpha_{ik} x_k$$

$$\frac{\partial y_i}{\partial \rho} = \sum_{k=1}^{n} \gamma_{ik} y_k$$

where

$$\alpha_{ik} = \begin{cases} y_k^T\left(\frac{\partial A}{\partial \rho} - \lambda_i \frac{\partial B}{\partial \rho}\right) x_i & ; \quad i \neq k \\ -\frac{1}{2}\left[x_i^T \frac{\partial B}{\partial \rho} x_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} \alpha_{ij} x_j^T (B + B^T) x_i\right] & ; \quad i = k \end{cases} \tag{3.6.24}$$

$$\gamma_{ik} = \begin{cases} \dfrac{y_i^T \left(\frac{\partial A}{\partial \rho} - \lambda_i \frac{\partial B}{\partial \rho}\right) x_k}{(\lambda_i - \lambda_k)} & ; \quad i \neq k \\ -y_i^T \frac{\partial B}{\partial \rho} x_i - \alpha_{ii} & ; \quad i = k \ . \end{cases} \tag{3.6.25}$$

It is interesting to note that for the special case where A and B are real and symmetric, the eigenvector derivative coefficients reduce, as a consequence of orthogonality, to the to the forms originally derived in [66]. Furthermore, if B is in addition constant (with respect to ρ's variation), for example the standard symmetric eigenvalue problem where B=I, then the diagonal coefficients vanish and this conforms to the standard symmetric case found in Refs. [62] and [64]. However, if A is nonsymmetric, the eigenvectors are not generally orthogonal and the diagonal coefficients from Eqs. (3.6.24) and (3.6.25) are clearly nonzero even when B is a constant. This crucial result directly contradicts the reasoning in [68] where the diagonal coefficients are stated to be "arbitrary" and hence the "most convenient" choice is a zero, which is clearly in error! This result is also interesting in the context of a first order eigenvector perturbation analysis since it implies that in the expansion of a particular eigenvector, the contribution of the change in the eigenvector in the same direction cannot be assumed zero in general for non-self-adjoint problems, as done commonly, for example in [64] (but subsequently corrected in [82]). We mention here that the dissertation reported in [83] notes this "ambiguity" of normalizing both left and right eigenvector derivatives and chooses to weigh the left and right diagonal terms equally.

In the next section, we present a numerical example to demonstrate the errors incurred in neglecting the diagonal terms in the evaluation of eigenvector derivatives for a non-self adjoint problem.

3.6.4 Numerical Example and Concluding Remarks

We demonstrate here using a (5 x 5) real matrix, the calculation of right and left eigenvector derivatives with respect to a single parameter. The randomly chosen matrix is

$$A(\rho) = \begin{bmatrix} 1.7 & 3.2 & 4 & 2 & -1 \\ 3.6 & .2 & (2+\rho) & -.9 & .1 \\ 1.7 & 0 & -3.5 & 9.4 & 2.9 \\ -.6 & 2 & 5 & 0 & -.3 \\ 0 & -3.9 & 1.2 & -4.5 & 1.1 \end{bmatrix}$$

where ρ is the scalar parameter appearing linearly only in location (2,3) of matrix A. The nominal matrix was chosen as the above matrix A when $\rho = 0$. It follows that the partial derivative of A with respect to ρ is all zero except at location (2,3) where it equals unity.

The nominal right and left eigenvalue problems of Eqs. (3.6.1) and (3.6.2) were solved and the eigenvectors normalized by Eqs. (3.6.12) and (3.6.13) to unity. By using the above normalized nominal eigenvectors, the right eigenvector derivatives were computed using Eqs. (3.6.16) to (3.6.18) and the left eigenvector derivatives using Eqs. (3.6.19), (3.6.20) and (3.6.22). Furthermore, the right and left eigenvector derivatives were computed for two cases, namely, "with" ($\alpha_{ii} \neq 0$, $\gamma_{ii} \neq 0$, $i=1,\ldots,5$) and "without" ($\alpha_{ii} = 0$, $\gamma_{ii} = 0$, $i=1,\ldots,5$) the diagonal terms in the eigenvector expansion equations of (3.6.16) and (3.6.19).

To provide a basis for comparison, the above computed eigenvector derivatives are compared to each other and to eigenvector derivatives computed using finite-differences. The i-th finite-difference eigenvector derivatives are computed as

$$\frac{\partial x_i}{\partial \rho}(\rho_0, \Delta\rho) \cong \frac{x_i(\rho) - x_i(\rho_0)}{\Delta\rho}$$

where

$$\Delta\rho = \rho - \rho_0.$$

Since the accuracy of the above approximation depends on the choice of step size, $\Delta\rho$, a graph of normalized error (10^{-n} corresponds approximately to n-digit accuracy) incurred in eigenvector derivatives computed by finite-differencing was plotted with respect to step size as shown in Figure 3.1. It can be seen from the figure that the "optimum" step size for computing the finite-difference derivative is about $\Delta\rho$ = .3e-5 and the finite-difference derivative and the derivative computed by the formula "with" diagonal terms correlate to within six decimal places. The finite difference derivatives can be calculated to five or better digits for all $\Delta\rho$'s in the interval $10^{-7} < \Delta\rho < 10^{-4}$. From the other experiments, depending upon the local behavior of the eigenvectors and machine word length, we can usually employ this experimental approach to find a $\Delta\rho$ range which gives five to six digit confirmation of Eqs. (3.6.16)-(3.6.22). Of course the importance of the analytical partial derivatives lies in the fact that we do not require the finte difference approximation and its associated numerical pitfalls and experimentation.

Table 3.1 shows the right and left eigenvector derivatives computed by three different ways, namely, "without" diagonal terms, "with" diagonal terms and by finite difference with step size .3e-5. It can be concluded that the eigenvector derivatives computed using the formula "with" diagonal terms show very accurate agreement (identical, to within small erros in the sixth digit) with the optimized finite difference approximations. The error between the finite difference results and the "without" results are essentially the missing diagonal terms; these frequently occur in the second and third digits. Furthermore, neglecting the diagonal terms in eigenvector derivative calculation often leads to serious errors, as is evident in Table 3.1.

Table 3.1 Eigenvector Derivatives

FINITE-DIFFERENCE EIGENVECTOR DERIVATIVES (STEP SIZE = .3E-05)/

EIGENVECTOR DERIVATIVES (WITH DIAGONAL TERMS)

RIGHT

Re(x_1)	Im(x_1)	Re(x_2)	Im(x_2)	Re(x_3)	Im(x_3)	Re(x_4)	Im(x_4)	Re(x_5)	Im(x_5)
-.0201	.0000	.0106	.0000	-.0221	.0271	-.0221	-.0271	-.0247	.0000
.0871	.0000	-.0206	.0000	-.0020	-.0160	-.0020	.0160	.0159	.0000
.0084	.0000	.0094	.0000	-.0002	.0015	-.0002	-.0015	-.0159	.0000
-.0184	.0000	.0042	.0000	-.0010	-.0120	-.0010	.0120	-.0075	.0000
.0255	.0000	.0039	.0000	.0102	.0156	.0102	-.0156	-.0051	.0000

LEFT

Re(y_1)	Im(y_1)	Re(y_2)	Im(y_2)	Re(y_3)	Im(y_3)	Re(y_4)	Im(y_4)	Re(y_5)	Im(y_5)
.0072	.0000	.0020	.0000	.0089	.2608	.0089	-.2608	.0192	.0000
-.0055	.0000	.0069	.0000	-.0028	-.1874	-.0028	.1874	.0042	.0000
-.0172	.0000	-.0639	.0000	-.0210	.0142	-.0210	-.0142	.0454	.0000
.0258	.0000	-.0539	.0000	.0534	.0423	.0534	-.0423	.0807	.0000
.0081	.0000	-.0238	.0000	.0152	.1654	.0152	-.1654	.0496	.0000

EIGENVECTOR DERIVATIVES (WITHOUT DIAGONAL TERMS)

RIGHT

Re(x_1)	Im(x_1)	Re(x_2)	Im(x_2)	Re(x_3)	Im(x_3)	Re(x_4)	Im(x_4)	Re(x_5)	Im(x_5)
-.0143	.0000	.0015	.0000	-.0109	.0144	-.0109	-.0144	-.0434	.0000
.0907	.0000	-.0262	.0000	-.0527	-.0221	-.0527	.0221	.0010	.0000
-.0147	.0000	.0063	.0000	.0130	.0043	.0130	-.0043	-.0095	.0000
-.0067	.0000	.0006	.0000	.0191	.0004	.0191	-.0004	-.0056	.0000
.0341	.0000	.0102	.0000	-.0490	-.0089	-.0490	.0089	.0161	.0000

LEFT

Re(y_1)	Im(y_1)	Re(y_2)	Im(y_2)	Re(y_3)	Im(y_3)	Re(y_4)	Im(y_4)	Re(y_5)	Im(y_5)
.0020	.0000	.0095	.0000	.0449	.1439	.0449	-.1439	.0212	.0000
-.0021	.0000	.0132	.0000	.0037	-.0944	.0037	.0944	.0129	.0000
.0031	.0000	-.0500	.0000	-.0162	.0180	-.0162	-.0180	.0138	.0000
.0046	.0000	-.0358	.0000	.0148	.1128	.0148	-.1128	.0363	.0000
.0016	.0000	-.0185	.0000	.0535	.1205	.0535	-.1205	.0205	.0000

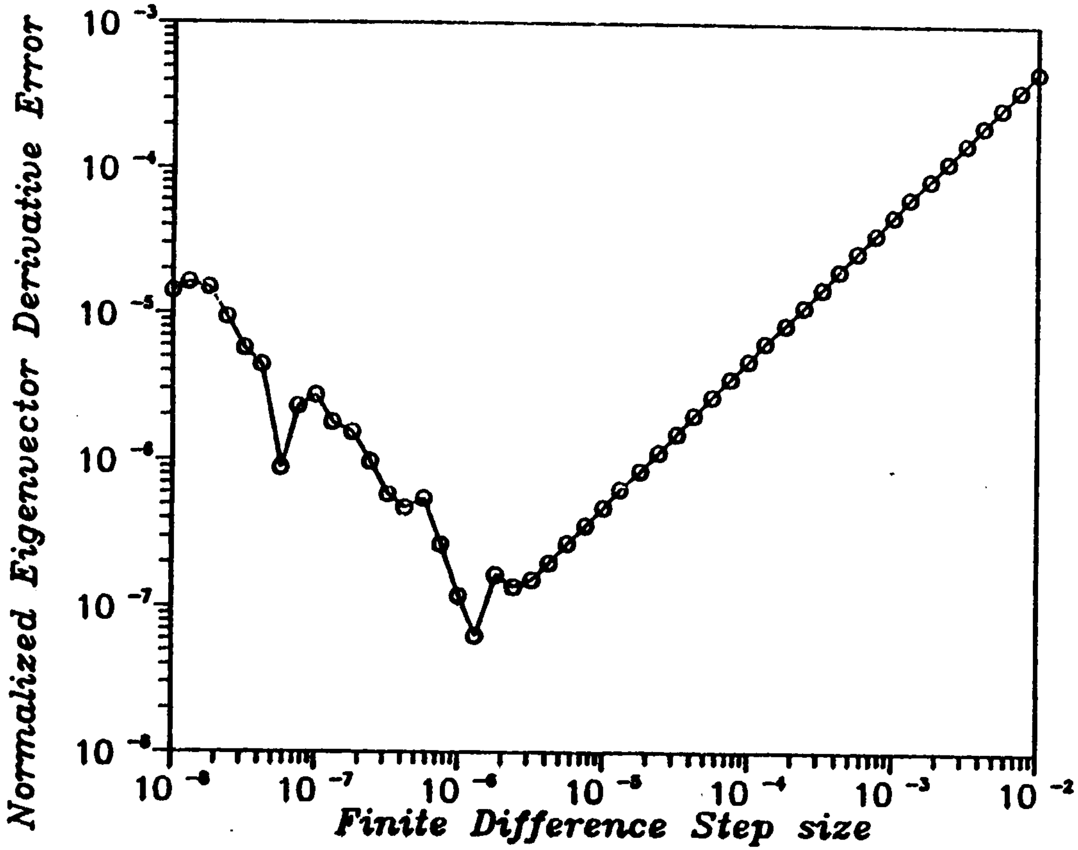


Figure 3.1 Normalized error of finite difference right eigenvector derivative for mode 1

In summary, the first order change in an eigenvector has a nonzero projection onto the eigenvector for non-self-adjoint systems.

3.7 Concluding Remarks

In this chapter, we have examined two main approaches for simultaneously placing eigenvalues and attaining low sensitivities, namely, modal insensitivity and minimum sensitivity approaches. From a theoretical perspective, the former method of utilizing null space basis vectors to impose modal insensitivity and place eigenvalues appears more attractive than the latter approach due to its mathematical rigor and elegance of the former approach. On the other hand, from the practical and implementation viewpoints, the latter approach appears more attractive for two reasons: (i) in situations where zero modal sensitivities (modal insensitivity) are not attainable, the former becomes inapplicable while for the latter approach, some minimal sensitivity configuration can be anticipated, and (ii) inequality and equality constraints, reflecting various physical constraints cannot be easily accomodated in the former approach. However, additional numerical experiments are needed to further evaluate the applicability and usefulness of the above approaches for real physical problems.

It should also be mentioned that in the null space approach, the remaining freedom (if it exists in a given problem) in choosing a set of feasible eigenvectors has not been sufficiently explored. The problem of developing a numerical algorithm to implement the above optimization problem will indeed be interesting.

A convenient scalar index designed for use as a cost function in eigenvalue sensitivity optimization is derived. By using concepts of matrix operator norms, the scalar index and its relationship to a linearly predicted bound on weighted and normalized eigenvalue perturbation is established. Unfortunately, numerical studies have not yet been done to examine its usefulness.

In a different context from the central theme of this chapter, we have derived modal expansion coefficients of eigenvector derivatives for non-self-adjoint systems. The eigenvector derivatives are shown to have generally nonzero projections onto all eigenvectors. Additionally, these basis coefficients for left and right eigenvector derivatives are found to be related by a simple expression.

Finally, a more concise derivation using concepts of eigenvalue conditioning of a stability robustness measure, originally derived by Patel and Toda, has been presented. This result provides additional rigor and insight to Patel and Toda's robustness measure.

4. STABILITY ROBUSTNESS THEORY IN TIME DOMAIN

4.1 Introduction

In the previous chapter, we discussed various ideas associated with eigenvalue sensitivity including scalar indices that presumably quantify the degree of sensitivity with respect to a chosen set of parameters. As with any sensitivity formulation, the eigenvalue sensitivity indices discussed above are strictly speaking, local measures. This implies that for a given nominal configuration with low sensitivity, there is no guarantee that even a small finite perturbation will not destabilize the nominal system. Indeed this is the major weakness of sensitivity methods. As a result, there is a pressing need for methods that can rigorously guarantee, within the framework of mathematical modelling, various properties of a control system under finite "ignorance". For convenience, we shall make a distinction between "sensitivity theory" and "robustness theory" with latter admitting finite plant perturbations.

In this chapter, we concentrate on robustness theory in the time domain, or more precisely, the effect of perturbating of vectors and matrices in state space equations on system stability. The main goals are to examine the basic principles that underlie a class of robustness measures and also to provide some new connections to well established concepts of numerical conditioning in matrix theory and eigenvalue placement problem. The earlier work of Patel, Toda and Sridhar [56,30] forms the historical basis of the ideas examined here. It is believed that results reported here extend their earlier work and provide several new insights.

We begin by a review of a Lyapunov stability theorem for time invariant linear systems. This is followed by a discussion on the problem of imposing absolute stability in the design of feedback controllers using the Lyapunov equation. In a different context, Lyapunov theory and matrix norm theory are used to arrive at a stability robustness measure first introduced by Patel and Toda [30]. We next derive a second measure of robustness in modal coordinates. It is believed that this new derivation substantially simplifies Patel and Todas' original derivation. In addition, the present derivation arrives at new relationships between the Patel-Toda robustness measure, the condition number, and the limit of robustness in a more direct fashion. The last section examines the robustness of the standard optimal linear quadratic regulator. The derivation presented here differs from that given in reference [56] as will be self evident. Finally, an equation relating the quadratic weight matrices and the Riccati and force distribution matrices is derived which represents conditions for "optimal" stability robustness with respect to Patel and Todas' measure for optimal linear quadratic regulators.

4.2 Lyapunov Stability Theorem

For linear, time invariant systems, a convenient means of testing for asymptotic stability is by using Lyapunov Stability Theorem. This theorem provides a necessary and sufficient condition for guaranteeing asymptotic stability. This is done by using the generalization of the concept of energy and its monotonic decay to a minimum zero. In other words, if the generalized energy of the state along any trajectory of the system decreases with time, then the state itself must approach zero as time increases. The Lyapunov theorem can be formally stated as [51]:

THEOREM 4.2

The zero state of the linear, time invariant system

$$\dot{x}(t) = Ax(t) \tag{4.2.1}$$

is asymptotically stable if
for every $\{Q: Q \geq 0,\ Q \text{ hermitian}\}$
there exists $\{P: P > 0,\ P \text{ hermitian}\}$
where P satisfies the Lyapunov matrix equation

$$A^H P + PA = -Q \ .$$

The above theorem implies that if the system matrix, A is asymptotically stable and Q is positive semidefinite, then the solution, P, from the Lyapunov equation must be positive definite. However, it is important to note that the theorem does not imply that if A is asymptotically stable and P is some chosen positive definite matrix, then, the computed Q is positive definite. This important fact leads to a design limitation in using Lyapunov equation directly for guaranteeing asymptotic stability of closed loop system. This limitation is evident from Figure 4-1 where a set relationship between a stable A, P and Q matrices are depicted. We see that for a given positive definite matrix P, the computed Q matrix may not be positive semidefinite even though A is asymptotically stable. In other words, selecting a positive definite matrix P and constraining matrix A (to be such that the computed right hand side be negative semi definite) may be too restrictive for practical implementation. On the other hand, solving the Lyapunov equation for P and determining its definiteness provides only a check for absolute stability. Of course, in principle, additional constraints

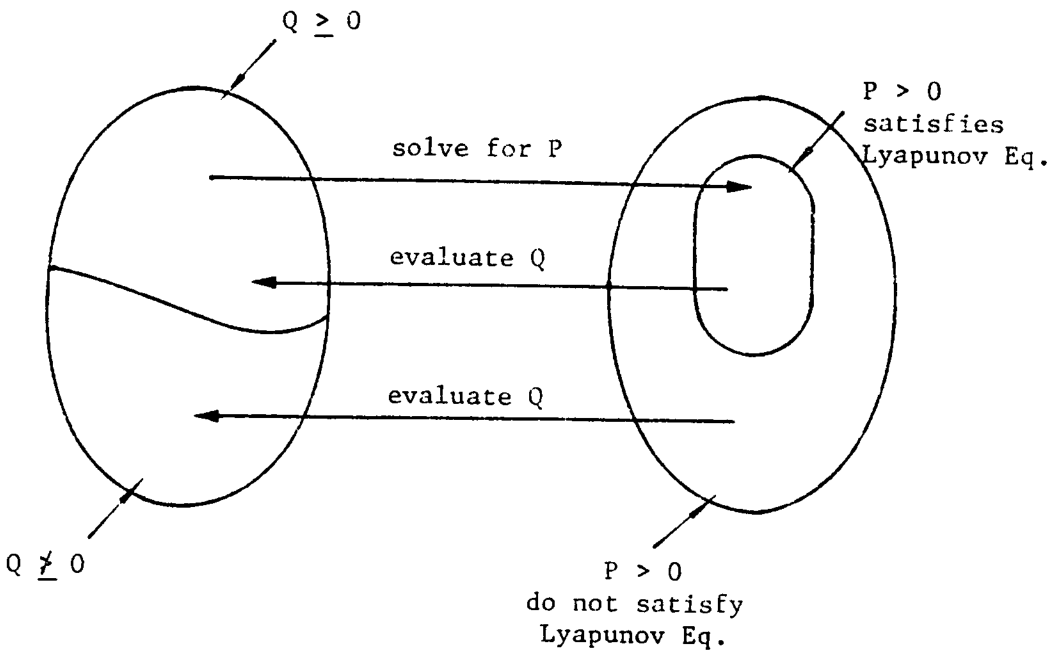


Figure 4.1 Relationship between A, P and Q matrices in Lyapunov equation

may be imposed on the solution of the Lyapunov equation to guarantee its positive definiteness but this results in a bilinear equation which substantially increases the level of numerical difficulty. To summarize, a nontrivial approach to imposing stability constraints using Lyapunov theorem is to select a positive definite Q, impose $\frac{n(n+1)}{2}$ constraints of Lyapunov equation and the positive definiteness of matrix P.

To prove the above theorem, we consider the time history of a Lyapunov function, V(x),

$$V(x) = x^H(t)\ P\ x(t), \qquad P > 0.$$

By taking the time derivative along the trajectory of Eq. (4.2.1), we get

$$\frac{dV}{dt} = \frac{d}{dt}\ [x^H(t)Px(t)]$$

$$= x^H(t)\ [A^HP + PA]\ x(t)$$

$$= -x^H(t)Qx(t)$$

where $\quad A^HP + PA \triangleq -Q \quad , Q=Q^H.$

We observe that for V(x) to monotonically decrease, $\frac{dV}{dt}$ must be zero[1]

---

[1] if $\frac{dV}{dt}$ is not identically zero along any nontrivial solution.

or negative, or equivalently $Q \geq 0$. Therefore, if we can find a positive definite P and positive semidefinite Q such that the Lyapunov equation is satisfied, then every trajectory of Eq. (4.2.1) will approach zero as time increases, i.e. A is asymptotically stable. Figure 4.2 gives a geometrical interpretation of Lyapunov theorem.

In the next two sections, the Lyapunov theorem will be used for deriving various stability robustness criteria and measures.

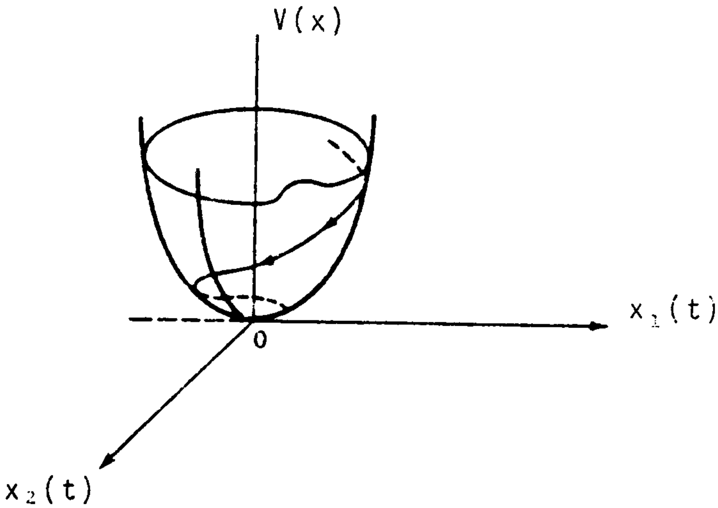
4.3 Robustness Measure of Patel and Toda

In this section, the robustness measure first introduced by Patel and Toda [30] is considered in detail. The results are derived using Lyapunov stability theorem and matrix norm theory. Due to their importance, a few theorems and lemmas will be stated and proved.

Let us consider the system described by

$$\dot{x}(t) = A\, x(t) + f(x(t),t) \tag{4.3.1}$$

where the uncertainty and/or perturbations on the system are assumed representable by f. Furthermore, we assume here that exact expressions for f are unknown and only bounds on f are known. We note in Eq. (4.3.1) that A may represent a closed loop system matrix and that the dependence of f on controls may be eliminated by substituting for a specific control strategy. The problem here is this: given that A is asymptotically stable, how can we obtain bounds on f such that the system maintains its stability?



Observe: $V(x) \geq 0$

$\dot{V} \leq 0$

$\|x\| \rightarrow 0$

$\|V(x)\| \rightarrow 0$ as $t \rightarrow \infty$

Figure 4.2 Geometrical Interpretation of Lyapunov Theorem

First, we choose as Lyapunov function

$$V(x) = x^T P x \tag{4.3.2}$$

where $P > 0$, is the solution of Lyapunov equation

$$A^T P + PA = -2Q \tag{4.3.3}$$

Since the unperturbed system, A is assumed asymptotically stable, a P matrix of Eq. (4.3.3) exists by the Lyapunov theorem for any positive semidefinite matrix Q. Taking the time derivative of the Lyapunov function and using Eq. (4.3.1), we obtain

$$\begin{aligned} \dot{V} &= \dot{x}^T P x + x^T P \dot{x} \\ &= x^T (A^T P + PA) x + 2f^T P x \\ &= -2x^T Q x + 2f^T P x \end{aligned}$$

Therefore, $\dot{V} \le 0$, i.e. the system described by Eq. (4.3.1) remains stable, if f satisfies

$$f^T P x \le x^T Q x \quad . \tag{4.3.4}$$

The condition of Eq. (4.3.4) is clearly inconvenient for applications and we seek instead a bound on the vector norm of perturbation f. By Rayleigh's principle [54], we obtain the lower bound

$$x^T Q x \ge \min \lambda(Q) \, \|x\|_2^2 \tag{4.3.5}$$

and an upper bound

$$
\begin{aligned}
f^T P x &\le |f^T P x| \\
&\le \|f\|_2 \, \|Px\|_2 \\
&\le \|f\|_2 \, \|P\|_2 \, \|x\|_2 \quad .
\end{aligned}
\qquad (4.3.6)
$$

We conclude from the bounds of Eqs. (4.3.5) and (4.3.6) that Eq. (4.3.4) is satisfied if $\|f\|_2$ satisfies

$$
\frac{\|f\|_2}{\|x\|_2} \le \frac{\min \lambda(Q)}{\|P\|_2} \quad . \qquad (4.3.7)
$$

From Eq. (4.3.7), we see the need for the condition

$$
f(0,t) = 0.
$$

The results of the above derivations can be summarized as follows: Given an asymptotically stable system, A, the system defined by Eq. (4.3.1) remains stable if f satisfies Eq. (4.3.7) and where P and Q (positive definite and semidefinite matrices respectively) satisfies Eq. (4.3.3). We make several observations here: first, we note that some level of conservatism in robustness measure is introduced by using the lower and upper bounds of Eqs. (4.3.5) and (4.3.6) to obtain Eq. (4.3.7) and second, it is obvious from Eq. (4.3.7) that the robustness measure depends on the choice of matrix Q, which hitherto was assumed arbitrary except for its positive semidefiniteness. The question naturally follows as to the choice of Q which maximizes the robustness measure.

In fact, this maximization with respect to Q amounts to reducing the conservatism of the robustness measure. In the sequel, the results of Patel and Toda [30] are presented that resolves the above problem.

LEMMA 4.3.1

Let $\overline{P}$ and P be the unique positive definite solutions of the Lyapunov equations

$$A^T\overline{P} + \overline{P}A = -2\overline{Q} \tag{4.3.8}$$

$$A^TP + PA = -2Q \tag{4.3.9}$$

where

$$\overline{Q} = qQ \tag{4.3.10}$$

and q is a positive scalar. Then,

$$\frac{\min \lambda(\overline{Q})}{\max \lambda(\overline{P})} = \frac{\min \lambda(Q)}{\max \lambda(P)} \triangleq \mu(Q)$$

PROOF OF LEMMA 4.3.1

Using Eqs. (4.3.10), (4.3.8) can be rewritten as

$$A^T(\frac{1}{q}\,\overline{P}) + (\frac{1}{q}\,\overline{P})A = -2Q \tag{4.3.11}$$

By uniqueness of solutions of Eqs. (4.3.9) and (4.3.11)

$$\frac{1}{q}\,\overline{P} = P$$

$$\Longrightarrow \quad \max \lambda(\overline{P}) = \max \lambda(qP) = q \max \lambda(P)$$

$$\Longrightarrow \quad \frac{\min \lambda(\overline{Q})}{\max \lambda(\overline{P})} = \frac{\min \lambda(qQ)}{q \cdot \max \lambda(P)} = \frac{\min \lambda(Q)}{\max \lambda(P)}$$

Q.E.D.

The above lemma, which proves that μ is independent of scaling, will be used in the following lemma to show that μ is maximum when Q is the identity matrix.

LEMMA 4.3.2

$\mu(Q)$ is a maximum when $Q = I$.

PROOF OF LEMMA 4.3.2

Let $\bar{Q} = qQ$, $Q > 0$

where $q = \dfrac{1}{\min \lambda(Q)}$

so that

$$\min \lambda(\bar{Q}) = 1 \; . \tag{4.3.12}$$

The solution of

$$A^T\bar{P} + \bar{P}A = -2\bar{Q} \tag{4.3.13}$$

can be written as [55]

$$\bar{P} = 2 \int_0^\infty e^{A^T t} \, \bar{Q} \, e^{At} dt \tag{4.3.14}$$

and by using Eq. (4.3.12)

$$\bar{\mu} \triangleq \frac{\min \lambda(\bar{Q})}{\max \lambda(\bar{P})} = \frac{1}{\max \lambda(\bar{P})} \; .$$

Consider a second system

$$A^T\tilde{P} + \tilde{P}A = -2I \tag{4.3.15}$$

where the solution can similarly be written as

$$\tilde{P} = 2 \int_0^\infty e^{A^T t} \, I \, e^{At} dt \tag{4.3.16}$$

and

$$\tilde{\mu} \triangleq \frac{\min \lambda(I)}{\max \lambda(\tilde{P})} = \frac{1}{\max \lambda(\tilde{P})} .$$

From Eqs (4.3.14) and (4.3.16), we can write

$$\overline{P} - \tilde{P} = 2 \int_0^\infty e^{A^T t}(\overline{Q} - I)e^{At}dt .$$

$$\begin{aligned}
\text{Eq. (4.3.12)} \implies \lambda(\overline{Q}) &\geq 1 \\
\implies \overline{Q} - I &\geq 0 \\
\implies \overline{P} - \tilde{P} &\geq 0 \text{ ; since } e^{At},\ e^{A^T t} \text{ nonsingular} \\
\implies \max \lambda(\overline{P}) &\geq \max \lambda(\tilde{P}) \\
\implies \tilde{\mu} \geq \overline{\mu} &, \text{ for every pos. } q \text{ and pos.def.} Q
\end{aligned}$$

Q.E.D.

The results of the previous derivations and lemmas can be summarized in the following theorem [30]:

THEOREM 4.3.1

Given an asymptotically stable system A, the perturbed system

$$\dot{x}(t) = A\, x(t) + f(x(t),t)$$

maintains its stability if f satisfies

$$\frac{\|f\|_2}{\|x\|_2} \leq \frac{1}{\max \lambda(P)} \triangleq \mu \tag{4.3.17}$$

where $P > 0$ is the solution of Lyapunov equation

$$A^T P + PA = -2I . \tag{4.3.18}$$

It is clear from above theorem that μ is a measure of stability robustness of the asymptotically stable unperturbed system, A. The implicit dependence of μ on A is also evident from the following lemma whose proof is given in [30]:

LEMMA 4.3.3

If A is asymptotically stable and P satisfies

$$A^TP + PA = -2I$$

then

$$\frac{1}{\max \lambda(P)} \leq -\max\,[\mathrm{Re}\,\lambda(A)] \qquad (4.3.19)$$

and equality holds when A is a "normal" matrix.

The above lemma gives a bound on achievable robustness with this kind of measure. It is apparent that the ideas of "robustness" and "stability margin" are closely related. In particular, for pole placement designs where eigenvalues are specified, the robustness measure, μ , can in principle be maximized to $-\max[\mathrm{Re}\lambda(A)]$ by seeking a normal matrix with the desired eigenvalues, provided no additional constraints are imposed. The above problem is certainly non-trivial in practice and there is currently no algorithm in the literature to accomplish this. Incidentally, the general form of lemma 4.3.3 when $Q \neq I$ is given in reference [52]. For the important case of linear perturbations of the form

$$f(x(t),t) = E(t)x(t)$$

we can easily apply theorem 4.3.1 to arrive at the following results from matrix norm theory:

THEOREM 4.3.2

The system

$$\dot{x}(t) = (A + E(t))x(t)$$

remains stable if E(t) satisfies any of the following conditions

a) $\|E(t)\|_2 \leq \mu$ (4.3.20a)

b) $\|E(t)\|_F \leq \mu$ (4.3.20b)

c) $\max\limits_{i,j} \|E_{ij}\| \leq \mu/n$ (4.3.20c)

where μ is defined by Eq. (4.3.17) and Eq. (4.3.18) and n is the order of the system. We note the following relationships[‡]:

$$\|\cdot\|_2 \triangleq \|\cdot\|_S = \bar{\sigma}[\cdot] = \text{maximum singular value of } [\cdot]$$

$$\leq \sqrt{\sum_i \sigma_i^2} = \|\cdot\|_F \triangleq \sqrt{\sum_j \sum_i E_{ij}^2}$$

$$\leq n \max_{i,j} |E_{ij}|$$

Since $\|\cdot\|_S$ norm is the smallest norm, it is the "best" choice, i.e., the least conservative of the three conditions. However, $\|\cdot\|_S$ is defined as

$$\|\cdot\|_S = \bar{\sigma}[\cdot] = \max_x \frac{\|[\cdot]x\|_2}{\|x\|_2}$$

so that it is more difficult to evaluate $\|\cdot\|_S$ than $\|\cdot\|_F$ or $n \max\limits_{i,j} |E_{ij}|$.

‡ $\|\cdot\|_S$, $\|\cdot\|_2$, $\|\cdot\|_F$ denotes "spectral", "2" and "Frobenius" norms respectively in matrix theory nomenclature, while σ_i denotes the i-th singular value of some matrix $[\cdot]$.

4.4 Relationship Between Robustness and Conditioning of Eigenvalue Problem

For a given linear, asymptotically stable system, an upper bound on the maximum tolerable perturbation that maintains stability is derived in the last section. The motivation is to obtain a criteria that guarantees stability for some bounded finite perturbation in the neighborhood of a nominal point in the context of closed loop control system matrices. As pointed out in Chapter 3, it is very interesting to observe that the above robustness criteria are not unrelated to the concept of conditioning of the corresponding eigenvalue problem. In contrast, we note that the motivations associated with the concept of conditioning originates from the field of numerical analysis and matrix computations [53,54,58] where robust solutions of algebraic equations with respect to numerical errors are of central importance. It should be emphasized that this source of numerical errors are from finite precision (and range) nature of computer arithmetic, which could be considered infinitesimal when compared to mathematical model errors, external disturbances or parameter variations. In any case, the two resulting equations are shown below to be intimately related.

We now present a derivation of a stability robustness measure which depends explicitly on the condition number of an eigenvalue problem. Consider the system

$$\dot{x}(t) = A\, x(t) + f(x(t),t) \tag{4.3.1}$$

and let us confine ourselves to the case where the constant matrix A has distinct eigenvalues so that there exists a full set of linearly independent eigenvectors, $\{x_1,\ldots,x_n\}$. The modal transformation is written as

$$x(t) = X\,\eta(t) \tag{4.4.1}$$

and the biorthonormality

$$Y^T X = I \tag{4.4.2}$$

$$Y^T A X = \Lambda \tag{4.4.3}$$

where X and Y denote modal matrices corresponding to right and left eigenvalue problems. By using Eq. (4.4.1), we can rewrite Eq. (4.3.1) in modal coordinates as

$$\dot{\eta}(t) = \Lambda\,\eta(t) + g(t) \tag{4.4.4}$$

where

$$g(t) \triangleq Y^T f(t)\ .$$

We now proceed along the lines of Section 4.3, except in modal coordinates. By defining a Lyapunov function, and taking its time derivative along the trajectory of Eq. (4.4.4), we get

$$V(\eta) = \eta^H(t)\,P\,\eta(t) \quad , \quad P > 0$$

$$\dot{V}(\eta) = \dot{\eta}^H(t)\,P\,\eta(t) + \eta^H(t)\,P\,\dot{\eta}(t)$$

$$= \eta^H[\Lambda^H P + P\Lambda]\eta + 2\text{Re}\,\{\eta^H Pg\}$$

$$= -2\eta^H Q\eta + 2\text{Re}\,\{\eta^H Pg\}$$

where

$$\Lambda^H P + P\Lambda \triangleq -2Q\ . \qquad (4.4.5)$$

By Lyapunov theorem, the system represented by Eq. (4.4.4) remains asymptotically stable if

$$\dot{V} \leq 0$$

$$\text{or Re}\,\{\eta^H Pg\} \leq \eta^H Q\eta\ . \qquad (4.4.6)$$

Since the terms in Eq. (4.4.6) have the bounds

$$\text{Re}\,\{\eta^H Pg\} \leq \|\eta\|_2\ \|P\|_2\ \|g\|_2$$

and

$$\eta^H Q\eta \geq \min\lambda(Q)\ \|\eta\|_2^2$$

therefore, Eq. (4.4.6) is satisfied if the perturbation in modal coordinates, g satisfies

$$\|\eta\|_2\ \|P\|_2\ \|g\|_2 \leq \min\lambda(Q)\ \|\eta\|_2^2$$

or

$$\frac{\|g\|_2}{\|\eta\|_2} \leq \frac{\min \lambda(Q)}{\|P\|_2} = \frac{\min \lambda(Q)}{\max \lambda(P)} \tag{4.4.7}$$

where P and Q are related by Eq. (4.4.5). From lemma 4.3.2, the bound in Eq. (4.4.7) is maximum when Q is the identity matrix. Also, since the modal forces and coordinates are related to the physical forces and coordinates by

$$\|g\|_2 \leq \|Y^T\|_2 \, \|f\|_2$$

and

$$\|x\|_2 \leq \|X\|_2 \, \|\eta\|_2 \quad ,$$

the stability condition of Eq. (4.4.7) is now satisfied if

$$\frac{\|Y^T\|_2 \, \|f\|_2}{\|x\|_2 / \, \|X\|_2} \leq \frac{1}{\max \lambda(P)}$$

$$\Longrightarrow \qquad \frac{\|f\|_2}{\|x\|_2} \leq \frac{1}{\|Y^T\|_2 \, \|X\|_2 \;\; \max \lambda(P)} \tag{4.4.8}$$

where P satisfies

$$\Lambda^H P + P\Lambda = -2I \; . \tag{4.4.9}$$

The solution for P in Eq. (4.4.9) can be easily obtained if written in index notation,

$$\sum_{k=1}^{n} \Lambda_{ik}^{*} P_{kj} + \sum_{k=1}^{n} P_{ik}\Lambda_{kj} = -2\delta_{ij} \quad ; \quad i,j=1,\ldots,n$$

$$\Longrightarrow \quad P_{ij}(\lambda_i^* + \lambda_j) = -2\delta_{ij}$$

$$\Longrightarrow \quad P_{ij} = \frac{-2\delta_{ij}}{\lambda_i^* + \lambda_j} = \begin{cases} 0 & ; \ i \neq j \\ -\dfrac{1}{\mathrm{Re}\{\lambda_i\}} & ; \ i = j \end{cases}$$

so,

$$P = \begin{bmatrix} \dfrac{1}{-\mathrm{Re}\{\lambda_1\}} & & 0 \\ & \ddots & \\ 0 & & \dfrac{1}{-\mathrm{Re}\{\lambda_n\}} \end{bmatrix}$$

where for convenience (without loss of generality) we let

$$0 < -\mathrm{Re}\{\lambda_1\} < -\mathrm{Re}\{\lambda_2\} < \ldots . < -\mathrm{Re}\{\lambda_n\}$$

i.e. λ_1 is assumed closest to the imaginary axis in eigenspace. Therefore,

$$\max \lambda(P) = \frac{1}{-\mathrm{Re}\{\lambda_1\}} = \frac{1}{-\max\limits_i [\mathrm{Re}\ \lambda_i(A)]}$$

so that the stability condition of Eq. (4.4.8) becomes

$$\frac{\|f\|_2}{\|x\|_2} \leq \frac{-\max\limits_i [\mathrm{Re}\ \lambda_i(A)]}{c(X)} \tag{4.4.10}$$

where $c(X) = \|Y^T\|_2 \ \|X\|_2$

$= \|X^{-1}\|_2 \ \|X\|_2 \ = \$ condition number.

For linear perturbations

$$\|f\|_2 = \|Ex\|_2 \le \|E\|_2 \; \|x\|_2$$

and the stability condition of Eq. (4.4.10) becomes

$$\|E\|_2 \le \frac{-\max\limits_i \,[\mathrm{Re}\ \lambda_i(A)]}{c(X)} \quad . \tag{4.4.11}$$

We note the following relations from a property of operator norms [53]:

$$c(X) = \|X^{-1}\| \; \|X\| \ge \|X^{-1}X\| = \|I\| = 1 \; .$$

It is clear from the above that the minimum value of a condition number is unity. In the limit when A is normal, X is a unitary matrix, i.e.

$$\|X\|_2 = \|X^{-1}\|_2 = 1$$

so that $\qquad c(X) = 1$.

We also observe that Eqs. (4.4.10) and (4.4.11) corresponds to Eqs. (4.3.17) and (4.3.20a) respectively. In the limit when A is normal, the robustness measures all have the same value as the right hand side of Eq. (4.3.19).

It can be concluded from Eqs. (4.4.10) and (4.4.11) that within the set of all A matrices with the same max[Re λ (A)], the maximum stability robustness condition corresponds to the minimum condition number of the eigenvalue problem, a consistent and an intuitively pleasing result! In fact, it can be seen from chapter 3 that the stability condition of Eq.

(4.4.11) can be derived purely from the eigenvalue conditioning viewpoint. A significance of the above results is that the problem of minimizing the condition number by eigenvector shaping is equivalent to the problem of maximizing a stability robustness measure of Patel and Toda. Also, a robust (in the Patel-Toda sense) control law will automatically have a well-conditioned closed loop eigenvalue problem. An alternative derivation leading to Eq. (4.4.10) is given in [30].

4.5 Robustness of Optimal Linear Quadratic Regulators

In this section, the stability robustness of the optimal Linear Quadratic State Feedback (LQSF) regulator is reviewed as originally derived in [56] but in a slightly different fashion. In essence, matrix bounds on the perturbations (modelling errors and/or parameter variations) in the system matrices, for a standard optimal LQSF regulator that would maintain closed loop stability are derived. The class of perturbation considered includes the general nonlinear, time-varying case. Furthermore, the special case of linear perturbations is examined in detail.

The steady state LQSF problem with prescribed degree of stability [57] can be stated as follows:

$$\text{minimize} \qquad J = \int_0^\infty e^{2\alpha t}[x^T Q x + u^T R u]dt \tag{4.5.1}$$

$$\text{subject to} \quad \dot{x}(t) = Ax(t) + Bu(t) \tag{4.5.2}$$

$$\text{where} \qquad Q \geq 0, \quad R > 0$$

and α (≥ 0) prescribing the degree of stability. Assuming the pair (A,B) to be controllable, and all states available for state feedback, the above optimization problem yields the optimal, constant gain state feedback control law

$$u(t) = - R^{-1}B^{T}P\, x(t) \tag{4.5.3}$$

where P>0 is the solution of the Algebraic Riccati Equation (ARE)

$$(A + \alpha I)^{T}P + P(A + \alpha I) - PBR^{-1}B^{T}P + Q = 0 \tag{4.5.4}$$

where I is an (nxn) identity matrix. From Eqs. (4.5.2) and (4.5.3), the closed loop system of the optimal regulator is

$$\dot{x}(t) = (A - BR^{-1}B^{T}P)\, x(t) \quad .$$

The problem of interest here is to derive the bounds on perturbation vector, f, that would maintain closed loop stability of the system

$$\dot{x}(t) = (A - BR^{-1}B^{T}P)\, x(t) + f(x(t),t) \tag{4.5.5}$$

4.5.1 Bounds for General Perturbation

Let us derive a bound on f using Lyapunov theorem. We choose as the Lyapunov function

$$V(x) = x^T P x \tag{4.5.6}$$

where P is the Riccati matrix, i.e. the solution of ARE Eq. (4.5.4). By taking the time derivative of Eq. (4.5.6) and substituting Eq. (4.5.5), we obtain

$$\dot{V}(x) = x^T(\bar{A}^T P + P\bar{A})x + 2f^T P x \tag{4.5.7}$$

where

$\bar{A} = A - BR^{-1}B^T P$; closed loop system matrix.

We proceed now to impose $\dot{V}(x) \le 0$ in Eq. (4.5.7), i.e.,

$$x^T(\bar{A}^T P + P\bar{A})x + 2f^T P x \le 0 \tag{4.5.8}$$

By using a rearranged ARE of Eq. (4.5.4) of the form

$$\bar{A}^T P + P\bar{A} = -(D + 2\alpha P) \tag{4.5.9}$$

where

$D \triangleq Q + PBR^{-1}B^T P$

the stability condition of Eq. (4.5.8) can be rewritten as

$$f^T P x \le \frac{1}{2} x^T (D + 2\alpha P) x \; . \tag{4.5.10}$$

To obtain bounds on the vector norm of f, we use Rayleigh's Principle

$$x^T(D + 2\alpha P)x \geq \min \lambda(D + 2\alpha P) \; \|x\|_2^2$$

and
$$\begin{aligned} f^T Px &\leq |f^T Px| \\ &\leq \|f\|_2 \; \|Px\|_2 \\ &\leq \|f\|_2 \; \|P\|_2 \; \|x\|_2 \end{aligned}$$

so that the stability condition of Eq. (4.5.10) is automatically satisfied if f satisfies

$$\|f\|_2 \; \|P\|_2 \; \|x\|_2 \;\leq\; \frac{1}{2} \min \lambda(D + 2\alpha P) \cdot \|x\|_2^2$$

or

$$\frac{\|f\|_2}{\|x\|_2} \;\leq\; \frac{1}{2} \, \frac{\min \lambda(D + 2\alpha P)}{\max \lambda(P)} \;. \tag{4.5.11}$$

The numerator of the right hand side of Eq. (4.5.11) can be further simplified by using the eigenvalue property for symmetric matrices

$$\min \lambda(D) + \min \lambda(2\alpha P) \;\leq\; \min \lambda(D + 2\alpha P)$$

so that Eq. (4.5.11) is satisfied if f satisfies

$$\begin{aligned} \frac{\|f\|_2}{\|x\|_2} &\leq \frac{1}{2} \, \frac{\min \lambda(D)}{\max \lambda(P)} + \alpha \, \frac{\min \lambda(P)}{\max \lambda(P)} \\ &\triangleq \mu \;. \end{aligned} \tag{4.5.12}$$

A less direct but nevertheless equivalent derivation of Eq. (4.5.12) is presented in reference [56]. We now summarize the above results in a theorem:

THEOREM 4.5.1

The optimal LQSF regulator closed loop system with perturbation, f, where

$$\dot{x}(t) = (A - BR^{-1}B^TP)\ x(t) + f(x(t),t)$$

remains stable if f satisfies

$$\frac{\|f\|_2}{\|x\|_2} \leq \mu$$

where

$$\mu = \frac{1}{2}\ \frac{\min \lambda(D)}{\max \lambda(P)} + \alpha\ \frac{\min \lambda(P)}{\max \lambda(P)}$$

$$D \triangleq Q + PBR^{-1}B^TP$$

and P satisfies ARE of Eq. (4.5.4).

From the above robustness measure, we observe that the quadratic weights, Q and R, plays a major role in determining robustness and must be chosen judiciously especially since various other properties such as eigenvalue assignability, control energy and state error magnitudes are also directly related to these matrices. This point is also evident in the numerical results of reference [14].

For the special case, $\alpha = 0$,

$$\mu = \frac{1}{2}\ \frac{\min \lambda(D)}{\max \lambda(P)}$$

where P satisfies

$$\bar{A}^T P + P\bar{A} = -D \triangleq -(Q + PBR^{-1}B^TP)$$

and we conclude from lemma 4.3.2 that μ is maximum when $D = I$, i.e. if the weights are chosen such that

$$Q + PBR^{-1}B^TP = I. \tag{4.5.13}$$

We see that for the special case of $\alpha=0$, Eq. (4.5.13) provides the conditions for which robustness of an optimal linear quadratic regulator is "optimal" with respect to Patel and Todas' measure.

4.5.2 Bounds for Linear Perturbations

Let us next consider the special but important case where the perturbation is assumed linear

$$f = E\, x(t) + F\, u(t)\ . \tag{4.5.14}$$

By adding Eq. (4.5.14) to Eq. (4.5.2) we get

$$\dot{x}(t) = (A + E)\, x(t) + (B + F)\, u(t) \tag{4.5.15}$$

where matrices E (nxn) and F (nxm) may represent modeling errors and/or parameter variations in the plant (A,B). By using the optimal control law of Eq. (4.5.3), the perturbation of Eq. (4.5.14) becomes

$$f = (E - FR^{-1}B^TP)\ x$$

$$= (E + FK)\ x \qquad (4.5.16)$$

where

$$K = -\ R^{-1}B^TP$$

and the closed loop system takes the form

$$\dot{x}(t) = [(A + E) + (B + F)K]\ x(t)\ . \qquad (4.5.17)$$

The problem now is to obtain bounds on E and F that will guarantee stability of Eq. (4.5.17). To use theorem 4.5.1 directly, we consider the following relations:

$$\frac{\|f\|_2}{\|x\|_2} = \frac{\|(E + FK)x\|_2}{\|x\|_2} \le \|E + FK\|_2 \le \|E\|_2 + \|F\|_2\ \|K\|_2\ .$$

Using theorem 4.5.1, the closed loop system Eq. (4.5.17) will maintain stability if E and F satisfies

$$\|E\|_2 + \|F\|_2\ \|K\|_2 \le \mu \qquad (4.5.18)$$

where μ is given in Eq. (4.5.12).

More convenient forms for the perturbations E and F in the left hand side of Eq. (4.5.18) can be obtained by using $\|\cdot\|_F$ instead of $\|\cdot\|_2$; i.e. since

$$\|E\|_2 + \|F\|_2 \, \|K\|_2 \leq \|E\|_F + \|F\|_F \, \|K\|_2 \qquad (4.5.19)$$

Eq.(4.5.18) is satisfied if

$$\|E\|_F + \|F\|_F \, \|K\|_2 \leq \mu \qquad (4.5.20)$$

The usefulness of Frobenious (or Euclidean) norm over spectral (or 2-norm) matrix norms is that the former are much easier to evaluate then the latter, at the expense of added conservatism by using Eq. (4.5.19).

A yet another convenient form of the bound for linear perturbation takes the form

$$\delta_E + \sqrt{\frac{m}{n}} \, \|K\|_S \delta_F \leq \frac{\mu}{n} \qquad (4.5.21)$$

where

$$|E_{ij}| \leq \delta_E \quad ; \quad i,j=1,\ldots,n$$

$$|F_{ik}| \leq \delta_F \quad ' \quad i=1,\ldots,n \quad ; \quad k=1,\ldots,m \quad .$$

The form in Eq. (4.5.21) easily follows from Eq. (4.5.18) and the inequalities

$$\|E\|_S \leq \|E\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} E_{ij}^2} \leq \sqrt{n^2 \max_{i,j} |E_{ij}|^2} = n\delta_E$$

$$\|F\|_S \le \|F\|_F = \sqrt{\sum_{j=1}^{m} \sum_{i=1}^{n} F_{ij}^2} \le \sqrt{m \cdot n \max_{i,j} |F_{ij}|^2} = \sqrt{m \cdot n}\ \delta_F \ .$$

4.5.3 Dependence of Bounds on Dominant Eigenvalue

In this section, relationships between perturbation bounds and optimal closed loop eigenvalues are derived, following ideas in reference [56]. In essence, it is shown that the following result represents a special case of lemma 4.3.3 where matrix A represents the optimum LQSF closed loop stability matrix. Let z be the eigenvector corresponding to the "dominant" eigenvalue of $\bar{A}$. By pre- and postmultiplying the ARE of Eq. (4.5.9), we get

$$z^{*}\bar{A}^{T}Pz + z^{*}P\bar{A}z = -z^{*}Dz - 2\alpha z^{*}Pz \ . \tag{4.5.22}$$

We can use the eigenvalue equations

$$\bar{A}z = \lambda z$$

$$z^{*}\bar{A}^{T} = \lambda^{*}z^{*}$$

where λ is the "dominant" eigenvalue and rewrite Eq. (4.5.22) as

$$\lambda^{*}z^{*}Pz + \lambda z^{*}Pz = -z^{*}Dz - 2\alpha z^{*}Pz$$

$$\Longrightarrow \qquad (\lambda^{*} + \lambda)z^{*}Pz = -z^{*}Dz - 2\alpha z^{*}Pz$$

$$\Longrightarrow \quad 2\max[\mathrm{Re}\,\lambda(\bar{A})]z^{*}Pz = -z^{*}Dz - 2\alpha z^{*}Pz$$

$$\Longrightarrow \quad -\max[\mathrm{Re}\,\lambda(\bar{A})] = \frac{z^{*}Dz}{2z^{*}Pz} + \alpha \qquad (4.5.23)$$

Since D and P are symmetric, by Rayleigh's Principle we have

$$\min \lambda(D) \le \frac{z^{*}Dz}{z^{*}z} \le \max \lambda(D)$$

$$\min \lambda(P) \le \frac{z^{*}Pz}{z^{*}z} \le \max \lambda\ (P)$$

and a lower bound on the first term of right hand side in Eq. (4.5.23) can be used to obtain

$$-\max\ [\mathrm{Re}\,\lambda(\bar{A})] \ge \frac{\min \lambda(D)}{2\max \lambda(P)} + \alpha$$

$$\ge \frac{\min \lambda(D)}{2\max \lambda\ (P)} + \alpha\frac{\min \lambda(P)}{\max \lambda(P)} \equiv \mu$$

since

$$\frac{\min \lambda(P)}{\max \lambda(P)} \le 1$$

i.e. $\quad \mu \le -\max\ [\mathrm{Re}\,\lambda(\bar{A})]$ (4.5.24)

In words, Eq. (4.5.24) says that the robustness measure or perturbation bounds for stability for an optimum LQSF closed loop system, as given by theorem 4.5.1, is bounded by the dominant optimum closed loop eigenvalue. Clearly, Eq. (4.5.24) can be considered a special case of lemma 4.3.3.

4.6 Concluding Remarks

We have examined in detail the basic principles underlying the derivation of stability robustness measures or bounds in state space. Like most mathematical models of real physical systems, the state space model is not unique. This raises the question of the validity of robustness measures that are based upon system matrix perturbations. On the other hand, robustness measures arrived at by using perturbations of transfer function have received much more attention, especially by control theory researchers from electrical engineering. In the context of the control of flexible structures, this popularity may be attributed to the following reasons: (i) transfer function matrices are unique for real physical systems and (ii) transfer function matrix perturbation bounds could include unmodelled higher modes and even spillover effects since the dimension of transfer function matrices are independent of the order of the underlying dynamical system. However, we point out that the practical significance of the above factors have not yet been substantiated by successful applications. In addition, we note that robustness bounds of transfer function matrices are difficult to implement due to their dependence on frequency. In summary, it is believed here that the results reported in this chapter have significantly increased the credibility of "time domain robustness measures", by developing these concepts from Lyapunov concepts and then making new connections to matrix norm and condition number concepts.

A major problem, which is fairly well known in the literature on robust control, is the conservatism of robustness criteria or measures that admit general or unstructured perturbations. This conservatism is rather democratic, in that it afflicts all of the time and frequency domain robustness norms known to this author. This conservatism is apparently more evident for problems having system matrices that are highly parameterized, or have significant internal structure. This problem is germane to all known multi-input, multi-output robustness criteria or measures. Consequently, the goal to obtain tighter perturbation bounds, for a given class of problems with known internal structures, by utilizing the internal structure, is indeed an important sequel problem that needs to be researched to increase the practical applicability of stability/robustness theory.

5. APPLICATIONS

5.1 Introduction

In this chapter, we focus on applying the design optimization algorithm developed in chapter 2 to extremize three different cost functions. The cost functions examined here namely, a scalar measure of eigenvalue sensitivities with respect to specified parameters, the total mass of the structure, and a stability robustness measure, all have obvious physical significance. A hypothetical structure is chosen here for our design study, the model details are presented in Appendix B. The structure consists of a free-free flexible beam with a rigid body attached to the center of the beam by a pin-joint and a torsional spring. Allowing only external torques and noting the coupling effect of the internal torsional spring, we see that the center of mass of the total structure can be assumed fixed in inertial space and the system has only a single rotational rigid body mode (see Appendix B). This structure can be seen as a planar model of a flexible satellite consisting of a rigid main body and a gimbaled flexible appendage.

We consider here the simultaneous design of structures and a direct output feedback controller for the flexible structure with the attached rigid body. The general transient response includes both rigid body and elastic motions. We assume here that these design optimizations are a part of a preflight design study in which a nominal structure is given with the corresponding open loop characteristics known. The goal is to simultaneously tune both the nominal structural design and the controller design parameters to move closed loop eigenvalues to desired regions in eigenspace, in some optimal manner as defined by a cost function. As a consequence of the particular parameterization, a total of 55 design variables (made up of 4 structural parameters, 3 actuator locations and 48 gain elements) are used to drive 10 damping factors and 6 damped frequencies to desired locations. Additional inequality constraints on the structural parameters are included to reflect various physical constraints.

In essence, the three cases presented differ only in their cost functions. Although many other cost functions merit investigation, we specifically choose the above mentioned functions because of their physical siginificance in both structures and control fields. We shall compare minimum mass designs (a structural cost function) with minimum closed loop eigenvalue sensitivity and maximum stability robustness designs (control cost functions). In addition, the latter two designs provide interesting comparisons.

Tables 5.1a and 5.1b show the nominal design variables and other fixed parameters for the structure. The open loop damping factors and frequencies corresponding to the nominal design is given in Table 5.2.

The dynamic model considered is a 20-th order system and the real system matrix gives in general, 10 conjugate pairs of eigenvalues and eigenvectors. Generally speaking, only the lower frequency modes will be accurately modeled due to truncation effects. However, we elect to simplify the present discussion by ignoring truncation errors. We consider here, and in the following two sections, the problem of driving 10 damping factors (which are almost negligible in the nominal design) to desired values while constraining the first 6 damped frequencies to

Table 5.1a Nominal Design Variables

DESIGN VARIABLE	SYMBOL	NOMINAL VALUE
actuator 2 location	a_1	5m
actuator 3 location	a_2	10m
actuator 4 location	a_3	15m
stiffness of torsional spring	k	500 n-m/rad
thickness of flexible beam	t_F	.1m
Young's modulus of beam	E	$.1482 \times 10^9$ N/m^2
mass density of rigid body	ρ_R	300 kg/m^3
output gain elements:		
G(1,1), G(2,1), G(3,1), G(4,1)		-1
All other elements		0

Table 5.1b Fixed structural parameters

PARAMETER	SYMBOL	FIXED VALUE
width of rigid body	w_R	1m
thickness of rigid body	t_R	3m
depth of rigid body	d_R	2m
width of flexible beam	w_F	20m
depth of flexible beam	d_F	1m
mass density of flexible beam	ρ_F	1799 kg/m^3
sensor 1 location	s_1	3m
sensor 2 location	s_2	7m
sensor 3 location	s_3	13m
sensor 4 location	s_4	17m

Table 5.2 Open loop and desired closed loop damped frequencies and damping factors.

MODE	OPEN LOOP		DESIRED CLOSED LOOP	
#	ω_d(rad/s)	ζ	ω_d^o(rad/s)	ζ^o
1	.0056	.4819E-10	.1	.7
2	.2803	.1402E-5	.3	.1
3	.3443	.1718E-5	.45	.1
4	1.241	.6204E-5	1.0	.05
5	1.768	.8839E-5	1.5	.05
6	3.981	.1990E-4	4.0	.05
7	5.004	.2502E-4	$>\omega_{d_6}^o + .1$	.02
8	8.295	.4147E-4	unconstrained	.02
9	9.902	.4951E-4	unconstrained	.02
10	14.34	.7171E-4	unconstrained	.02

some location (with the exception of the first mode). The first mode is the only rigid body mode in the model and we would like to increase its damping most significantly while changing its frequency to some positive number (so the controlled vehicle will have a preferred pointing direction). Table 5.2 also shows the desired damping factors and frequencies. In the next three sections, we present the problem formulations for the three different cost functions.

5.2 Optimal Eigenvalue Placement Designs for a Flexible Structure with Attached Rigid-Body Using Output Feedback

5.2.1 Minimum Eigenvalue Sensitivity Design

Since lower frequency modes generally dominate transient response for structural systems, we choose here to minimize the eigenvalue sensitivity of the lower 6 modes out of a possible total of 10 modes, with respect to an assumed set of 5 relatively uncertain parameters consisting of

$$V = \begin{Bmatrix} a_1 \\ a_2 \\ a_3 \\ k \\ t_F \end{Bmatrix} = \begin{bmatrix} \text{actuator location 2} \\ \text{actuator location 3} \\ \text{actuator location 4} \\ \text{torsional spring constant} \\ \text{thickness of flexible beam} \end{bmatrix}$$

The eigenvalue sensitivity cost function can now be written as

$$S = \sum_{j=1}^{5} \sum_{i=1}^{6} \left| \frac{\partial \lambda_i}{\partial v_j} \right|^2 w_{ij}$$

where w_{ij} represents the relative weight of the sensitivity of the i-th eigenvalue with respect to the j-th parameter of vector V. The weights are taken as the square of the magnitude of the nominal parameters. This amounts to a nondimensionalization of the sensitivity measure and weighting all terms of the cost function equally. The set of design parameters that are used to minimize the above sensitivity measure consists of 55 variables as given in Table 5.1a. The constraints are summarized as follows:

actuator location constraints

$$a_i^{\ell} \le a_i \le a_i^{u} \;; \quad i=1,2,3. \tag{5.2.1}$$

plant parameter constraints

$$\begin{aligned} & k^{\ell} \le k \\ & t_F^{\ell} \le t_F \le t_F^{u} \\ & E^{\ell} \le E \\ & \rho_R^{\ell} \le \rho_R \le \rho_R^{u} \end{aligned} \tag{5.2.2}$$

eigenvalue constraints

$$\begin{aligned} & \zeta_i = \zeta_i^{o} \;; \quad i=1,\ldots,10 \\ & \omega_{d_i} = \omega_{d_i}^{o} \;; \quad i=1,\ldots,6 \\ & \omega_{d_i} \ge \omega_{d_6} + \Delta\omega_{76} \end{aligned} \tag{5.2.3}$$

local step size constraints

$$-\varepsilon_i \le \Delta p_i \le \varepsilon_i \;; \quad i=1,\ldots,55 \tag{5.2.4}$$

where Δp_i and ε_i represents the i-th parameter change and the corresponding scalar bounds respectively. As evident from above

Table 5.3 Lower, upper and local step size bounds on design parameters.

PARAMETER		
lower bounds on actuator location	a_1^ℓ, a_2^ℓ, a_3^ℓ	0m
upper bounds on actuator location	a_1^u, a_2^u, a_3^u	20m
lower bounds on spring stiffness	k^ℓ	5 N-m/rad
lower bound on beam thickness	t_F^ℓ	.01m
upper bound on beam thickness	t_F^u	2m
lower bound on beam stiffness	E^ℓ	$.1480 \times 10^9$ N/m^2
lower bound on rigid body density	ρ_R^ℓ	50 kg/m^3
upper bound on rigid body density	ρ_R^u	1000 kg/m^3
local step size bounds		
actuator location	Δa_1, Δa_2, Δa_3	.1m
spring stiffness	Δk	30 N-m/rad
beam thickness	Δt_F	.01m
beam stiffness	ΔE	$.1 \times 10^5$ N/m^2
rigid body density	$\Delta \rho_R$	40 kg/m^3
gain elements i=1,...,4; j=1,...,12	$\Delta G(i,j)$	10
frequency separation between mode 7 and 6	$\Delta \omega_{76}$	.1 rad/s

equation, an inequality constraint on the frequency of the 7-th mode was imposed because it was found to be necessary (for this particular problem) to avoid an algorithm related pitfall when the unconstrained higher frequency mode cross the trajectory of the 6 lower frequency constrained modes. The lower and upper bounds on the actuator locations, structural parameters and step size are given in Table 5.3.

The sequential linear programming approach outlined in Chapter 2 can now be applied, the linear program at step-i has the form

$$\text{MAXIMIZE} \quad -\iota^T y$$

SUBJECT TO

$$\begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ I \end{bmatrix} y \geq \begin{Bmatrix} b_1 + A_1 \varepsilon \\ b_2 + A_2 \varepsilon \\ b_3 + A_3 \varepsilon \\ 2\varepsilon \end{Bmatrix} \tag{5.2.5}$$

$$\begin{bmatrix} A_4 \\ A_5 \end{bmatrix} y = \begin{Bmatrix} b_4 + A_4 \varepsilon \\ b_5 + A_5 \varepsilon \end{Bmatrix} \tag{5.2.6}$$

y is nonnegative,

where

$$\underset{(6\text{x}55)}{A_1} = \begin{bmatrix} (3\text{x}3) & (3\text{x}52) \\ -I & 0 \\ & \\ (3\text{x}3) & (3\text{x}52) \\ I & 0 \end{bmatrix} ; \quad \underset{(6\text{x}1)}{b_1} = \begin{Bmatrix} a_1 \\ a_2 \\ a_3 \\ w_F - a_1 \\ w_F - a_2 \\ w_F - a_3 \end{Bmatrix}$$

$$\underset{(1\text{x}55)}{A_2} = \left[\left(\frac{\partial \omega_{d_6}}{\partial \rho_1} - \frac{\partial \omega_{d_7}}{\partial \rho_1} \right), \ldots, \left(\frac{\partial \omega_{d_6}}{\partial \rho_{55}} - \frac{\partial \omega_{d_7}}{\partial \rho_{55}} \right) \right] ; \quad b_2 = \omega_{d_7} - \omega_{d_6} - \Delta\omega_{76}$$

$$\underset{A_3}{(6x55)} \left[\begin{array}{c:cccc:c} & -1 & & & & \\ & & -1 & & & \\ (6x3) & & 1 & & & (6x48) \\ 0 & & & -1 & & 0 \\ & & & & -1 & \\ & & & & 1 & \end{array}\right] \; ; \; \underset{b_3}{(6x1)} = \left\{\begin{array}{c} -k^l + k \\ -t_F^l + t_F \\ t_F^u - t_F \\ -E^l + E \\ -\rho_R^l + \rho_R \\ \rho_R^u - \rho_R \end{array}\right\}$$

$$\underset{A_4}{(10x55)} = \left[\frac{\partial \zeta}{\partial \rho}\right] \; ; \; \underset{b_4}{(10x1)} = \{(1-\gamma)\zeta^s + \gamma\zeta^u - \zeta\}$$

$$\underset{A_5}{(6x55)} = \left[\frac{\partial \omega_d}{\partial \rho}\right] \; ; \; \underset{b_5}{(6x1)} = \{(1-\gamma)\omega_d^s + \gamma\omega_d^u - \omega_d\}$$

After solving the above linear program at step-i, the parameters can be updated using

$$\rho^{i+1} = \rho^i + y - \varepsilon \quad .$$

The above linear program is solved by a FORTRAN subroutine called "ZX4LP" which is available in IMSL |78|. The execution times on a CYBER 170 computer for each continuation step is given in Table 5.4. We can

Table 5.4 Computer execution times for various subproblems.

SUBPROBLEM	CP secs. (CYBER 170)
real, nonsymmetric left and right eigenvalue problems (20x20) using IMSL routine EIGRF	2.67
Solution of linear program (55 variables, 83 constraints) using IMSL routine ZX4LP	22.73
Compute eigenvalue derivatives $\dfrac{\partial \lambda_i}{\partial p_j}$; $i=1,\ldots,6$, $j=1,\ldots,55$	5.05
Compute second eigenvalue derivatives $\dfrac{\partial^2 \lambda_i}{\partial p_k \partial v_j}$; $i=1,\ldots,6$, $j=1,\ldots,5$, $k=1,\ldots,55$	159.06

see that the computation of second derivatives of the eigenvalues represents most of the computing effort despite the use of analytical derivatives; finite-difference derivatives requires much more computing time. After some numerical experimentation, a starting continuation step size of $\Delta\gamma = .005$ and gradually increasing to $\Delta\gamma = .1$ was found to be suitable for our particular problem. After each new increment by $\Delta\gamma$, the value of γ (which corresponds to a percentage enforcement of the eigenvalue relocation eigenvalue constraints) is kept fixed while the local parameter step size is decreased by one half and the linear program resolved. This is done sequentially until no discernable improvement in cost function is observed.

5.2.2 Maximum Stability Robustness Design

We consider here the optimization of the stability robustness measure defined by Eqs. (4.3.17), (4.3.18) and (4.3.20a). The robustness measure we seek to maximize is

$$\mu = \frac{1}{\max \lambda[P]} \triangleq \frac{1}{\bar{\lambda}} \tag{5.2.7}$$

where

$$A^T P + PA = -2I \tag{5.2.8}$$

and A represents the nominal closed loop system matrix. Since A is asymptotically stable, by the Lyapunov stability theorem P is symmetric and positive definite and therefore μ is a well defined positive number. To formulate the linear program, we expand the denominator of the above cost function about the current point p^c so that,

$$\bar{\lambda} = \bar{\lambda}\Big|_{p^c} + \sum_{j=1}^{\ell} \frac{\partial\bar{\lambda}}{\partial p_j}\Big|_{p^c} \Delta p_j + O(\Delta p^2)$$

where the sensitivity of the maximum eigenvalue of P with respect to the j-th parameter, p_j is

$$\frac{\partial\bar{\lambda}}{\partial p_j} = \bar{u}^T \frac{\partial P}{\partial p_j} \bar{u} \tag{5.2.9}$$

and $\bar{u}$ is the real eigenvector corresponding to the maximum eigenvalue $\bar{\lambda}$, which satisfies the real, symmetric eigenvalue problem,

$$P\bar{u} = \bar{\lambda}\,\bar{u} \quad . \tag{5.2.10}$$

The sensitivity of $^{\text{n}}$ as required in Eq. (5.2.9) can be obtained by taking the partial derivative of the Lyapunov Eq. (5.2.8) to get

$$A^T \frac{\partial P}{\partial p_j} + \frac{\partial P}{\partial p_j} A = -\left(\frac{\partial A^T}{\partial p_j} P + P \frac{\partial A}{\partial p_j}\right) \tag{5.2.11}$$

From the above equation we see that $\frac{\partial P}{\partial p_j}$ always exists and is unique for any right hand side since all the eigenvalues of A is assumed to have negative real parts (see p. 416 of [51]).

To summarize, the linear program for the optimization of stability robustness can be written as

MINIMIZE $\sum_{j=1}^{\ell} (\bar{u}^T \frac{\partial P}{\partial p_j} \bar{u})\Big|_{p^c} y_j$

SUBJECT TO Eqs. (5.2.5), (5.2.6)
Eqs. (5.2.8), (5.2.10), (5.2.11)
y is nonnegative

5.2.3 Minimum Mass Design

The total mass of the structural system is used as the cost function to be minimized. The total mass of the combined structure can be written as

$$
\begin{aligned}
M &= M_{flexible} + M_{rigid} \\
&= t_F d_F w_F \rho_F + t_R d_R w_R \rho_R
\end{aligned}
$$

where the only design variables affecting the total mass is the thickness of the flexible beam, t_F, and the mass density of the rigid body, ρ_F. The above cost function can be readily linearized about a current design point to obtain a linear cost function with eigenvalue and design variable constraints as given in Eqs. (5.2.1) to (5.2.4).

5.2.4 Numerical Results

Figures 5.1 to 5.3 illustrate the convergence histories of weighted eigenvalue sensitivity, stability robustness and total mass of the structural system respectively for the three different designs as discussed previously. It can be observed that for all three cases, convergence to the desired eigenvalue constraints is complete

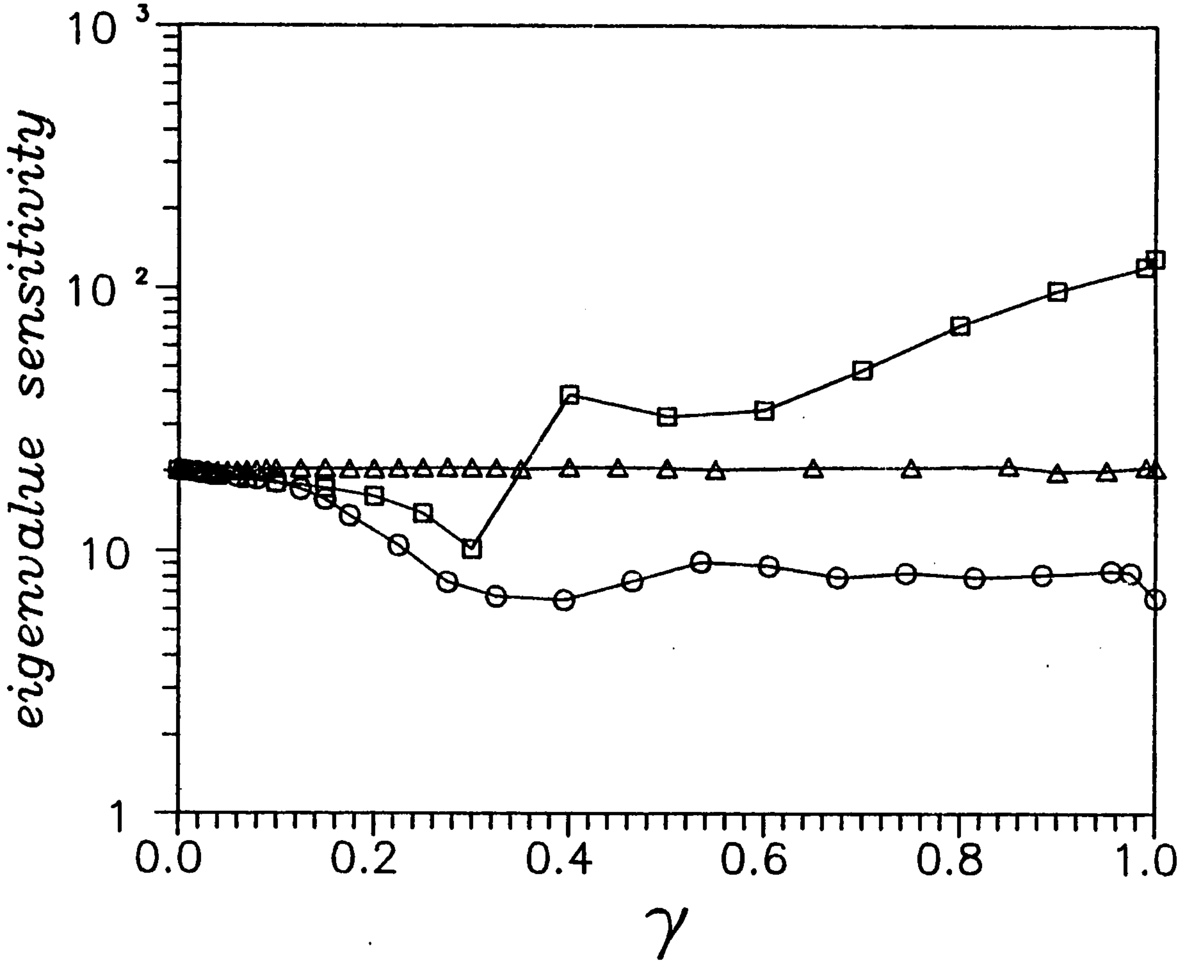


Figure 5.1 Convergence histories of eigenvalue sensitivity index
○ Minimum Sensitivity Design
□ Minimum Mass Design
△ Maximum Robustness Design

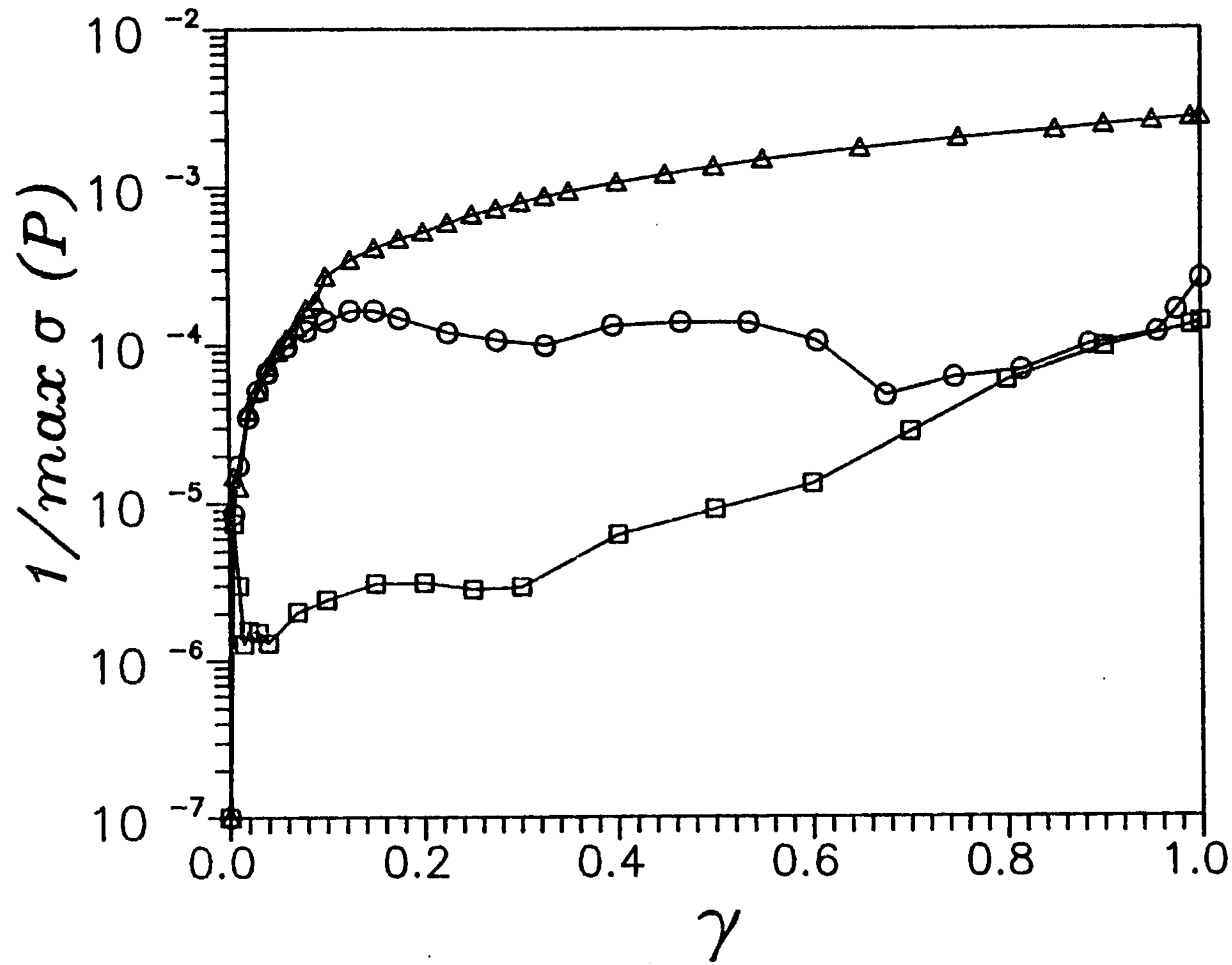


Figure 5.2 Convergence histories of stability robustness index
○ Minimum Sensitivity Design
□ Minimum Mass Design
△ Maximum Robustness Design

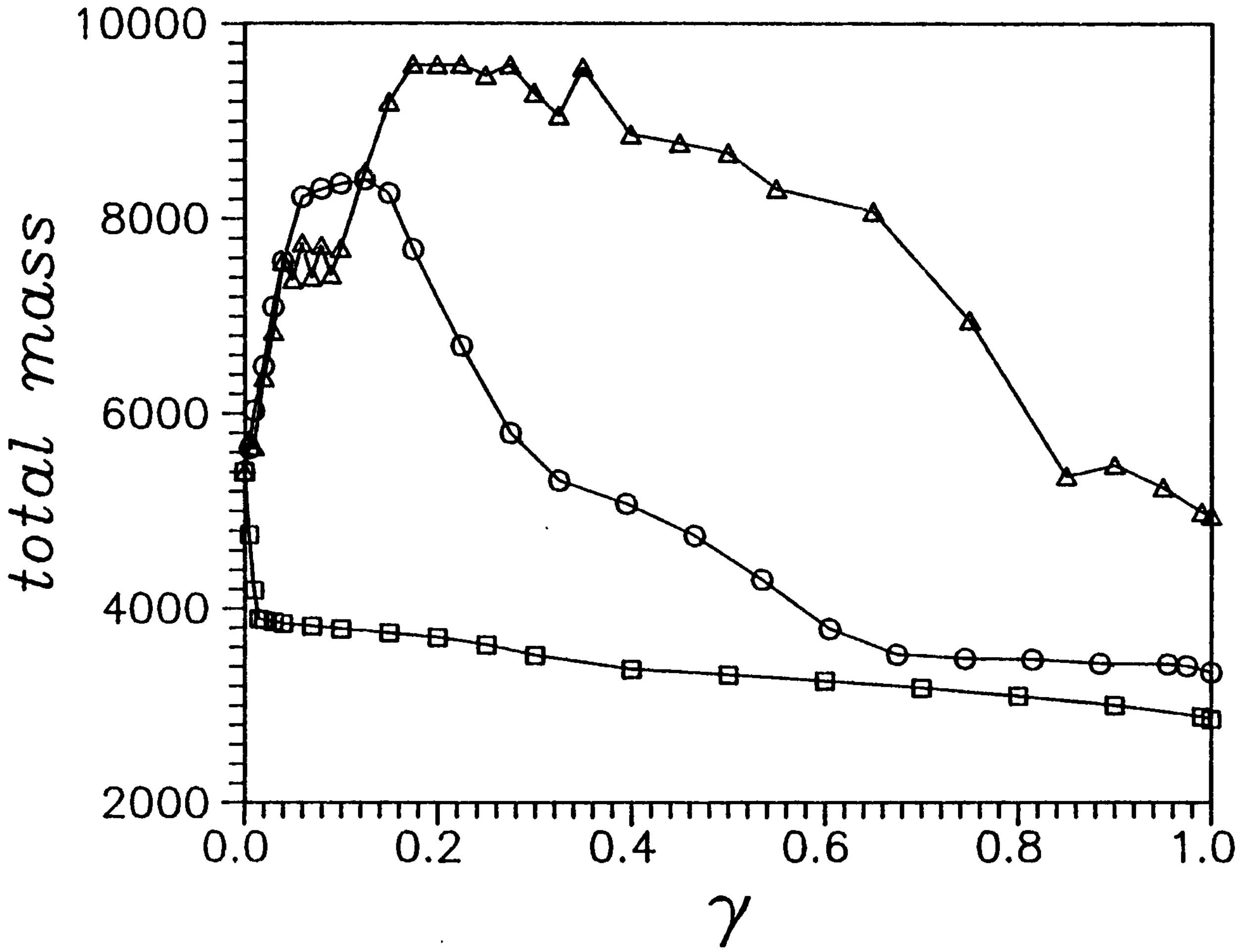


Figure 5.3 Convergence histories for total mass
○ Minimum Sensitivity Design
□ Minimum Mass Design
△ Maximum Robustness Design

(i.e., γ=1.0 is achieved). From Figure 5.1, we observe that initially (γ<.1), the minimum mass design results in eigenvalue sensitivity comparable to minimizing eigenvalue sensitivity directly and for values of γ>.1, the eigenvalue sensitivities fluctuated in an unpredictable manner far above the values of minimum sensitivity case. It is interesting to note that maximizing robustness did not influence eigenvalue sensitivity significantly. It should be recalled here that the eigenvalue sensitivity, as defined here, depends on the particular choice of parameters and the eigenvalues chosen, i.e., the sensitivity of the first 6 eigenvalues (out of 10) with respect to a subset of 5 parameters (out of 55). In addition, the selection of relative magnitudes of the weights associated with each sensitivity component is important and significantly influences the numerical results. However, as is common with most applications involving selection of weights, the results depend heavily on the particular application and the analyst's insight (or lack thereof).

In Figure 5.2 the stability robustness measures are plotted for the three designs. As expected, the maximum robustness design gives the highest robustness measures. It is interesting to observe that the minimum sensitivity design results in a robustness history paralleling that of maximum robustness design history for γ's at initial stages (γ<.1), and remains essentially constant at 10^{-4} thereafter. Interestingly this verifies the usefulness of minimizing eigenvalue sensitivity for robustness optimizations when the eigenvalues are close to the imaginary axis. On the other hand, for the minimum mass design, the stability robustness index remains significantly below optimal robustness values. However, the stability robustness index gradually increases with γ for minimum mass design although its corresponding eigenvalue sensitivity becomes very large. This is probably due to the increase in stability margin (distance from imaginary axis) with increasing γ. Incidentally, this supports the previous observation that robustness and sensitivity are not related "one to one".

The total mass histories as given in Figure 5.3 clearly shows the large differences in total mass of the different designs. However, the trend of the total mass for minimum sensitivity and maximum robustness designs are similar inspite of their large differences in their absolute values. Table 5.5 shows the performance indices at starting and final converged conditions. The improvements in sensitivity, total mass and robustness are clearly evident. The condition number and an alternate robustness index is also shown for additional comparison.

To further evaluate the stability robustness and eigenvalue sensitivity of the various designs, an alternate stability robustness index, originally introduced by Patel and Toda [30] and rederived elegantly in this dissertation (see section 3.5.2), is computed and plotted along with the condition number of the eigenvalue problem in Figures 5.4 and 5.5 respectively. From Figure 5.4, the maximum robustness design is clearly seen to be the most robust in terms of the alternative stability robustness index (Eq. 3.5.12) and in fact, Figures 5.4 and the corresponding Figure 5.2 are very similar. Furthermore, Figure 5.5 shows that the condition numbers corresponding to the maximum robustness design case had the smallest (optimal) values and decreased monotonically to an asymptotic value. Interestingly enough, similar to

Table 5.5 Performance Indices at Starting and Final Converged Conditions

		MINIMUM SENSITIVITY DESIGN	MINIMUM MASS DESIGN	MAXIMUM ROBUSTNESS DESIGN
$\gamma = 0$	eigenvalue sensitivity	20.08	20.08	20.08
	total mass	5398	5398	5398
	robustness $(1/\bar{\sigma})$	.2512E-7	.2512E-7	.2512E-7
	robustness $-\max(\mathrm{Re}(\lambda))/c(X)$	.6516E-15	.6516E-15	.6516E-15
	condition number $c(X)$	419.3	419.3	419.3
$\gamma = 1$	eigenvalue sensitivity	6.64	130.6	20.70
	total mass	3336	2853	4952
	robustness $(1/\bar{\sigma})$	.2589E-3	.1384E-3	.2750E-2
	robustness $-\max(\mathrm{Re}(\lambda))/c(X)$	.6650E-4	.9157E-4	.3140E-3
	condition number $c(X)$	453.3	329.2	95.99

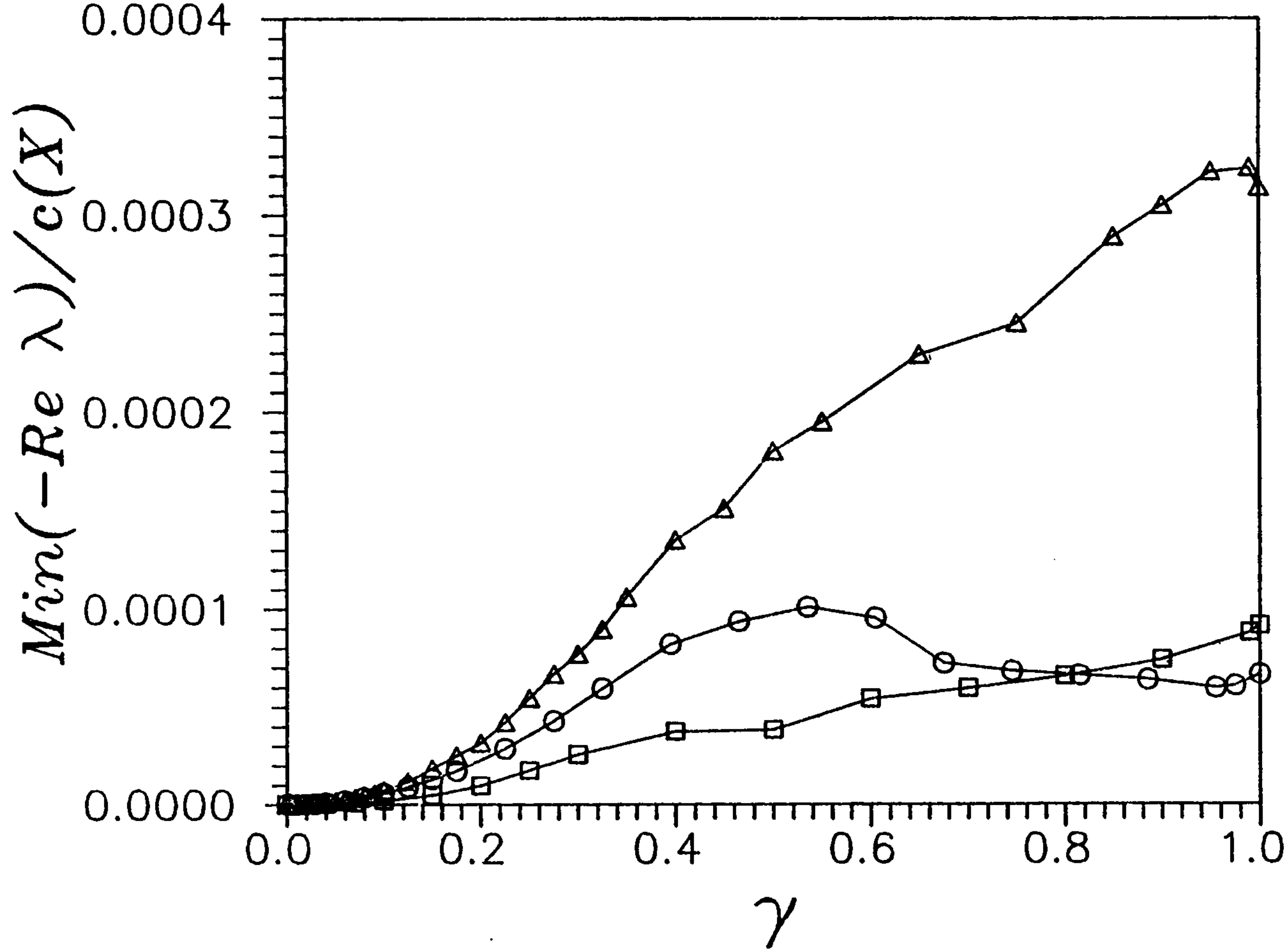


Figure 5.4 Convergence histories of stability robustness index
○ Minimum Sensitivity Design
□ Minimum Mass Design
△ Maximum Robustness Design

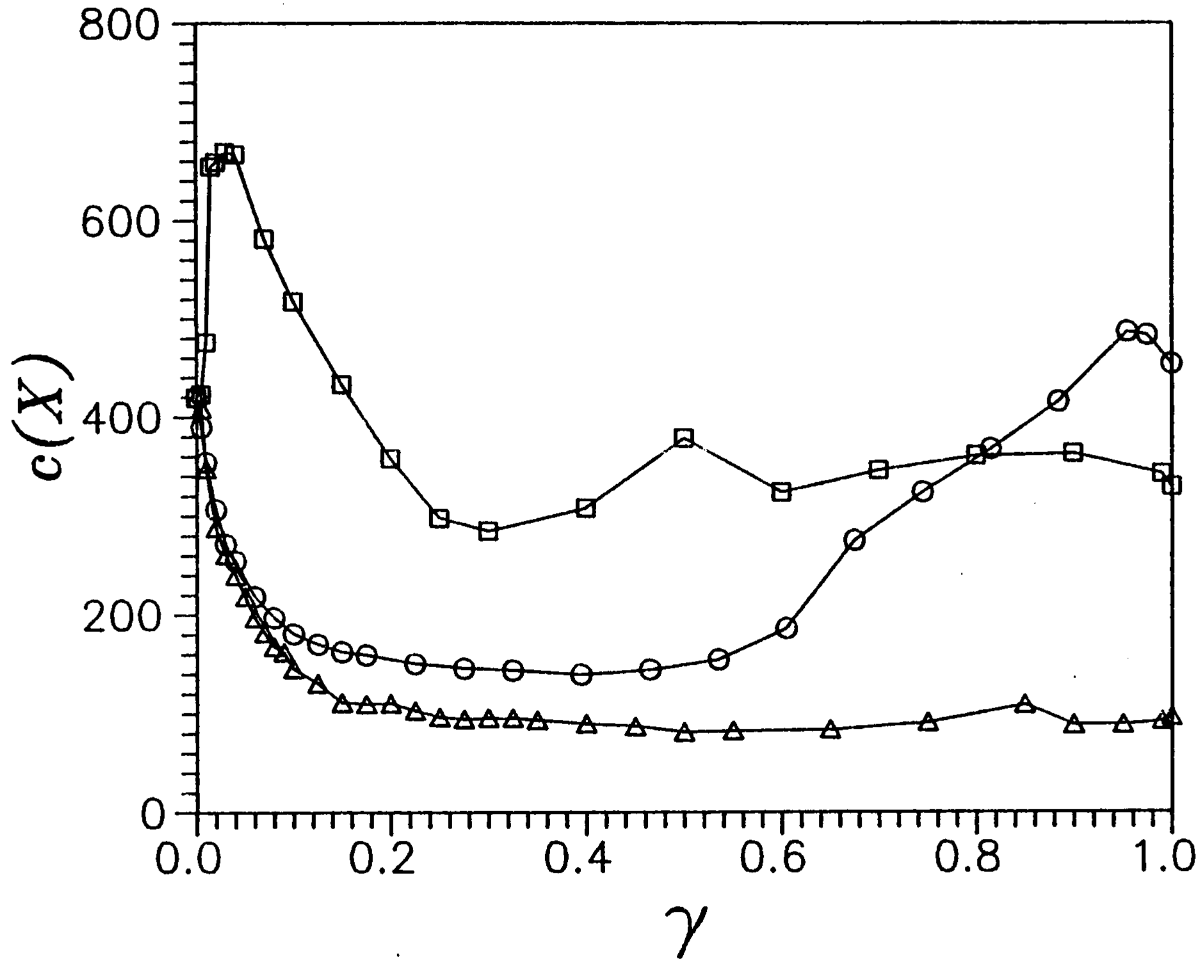


Figure 5.5 Convergence histories of eigenvalue condition number
○ Minimum Sensitivity Design
□ Minimum Mass Design
△ Maximum Robustness Design

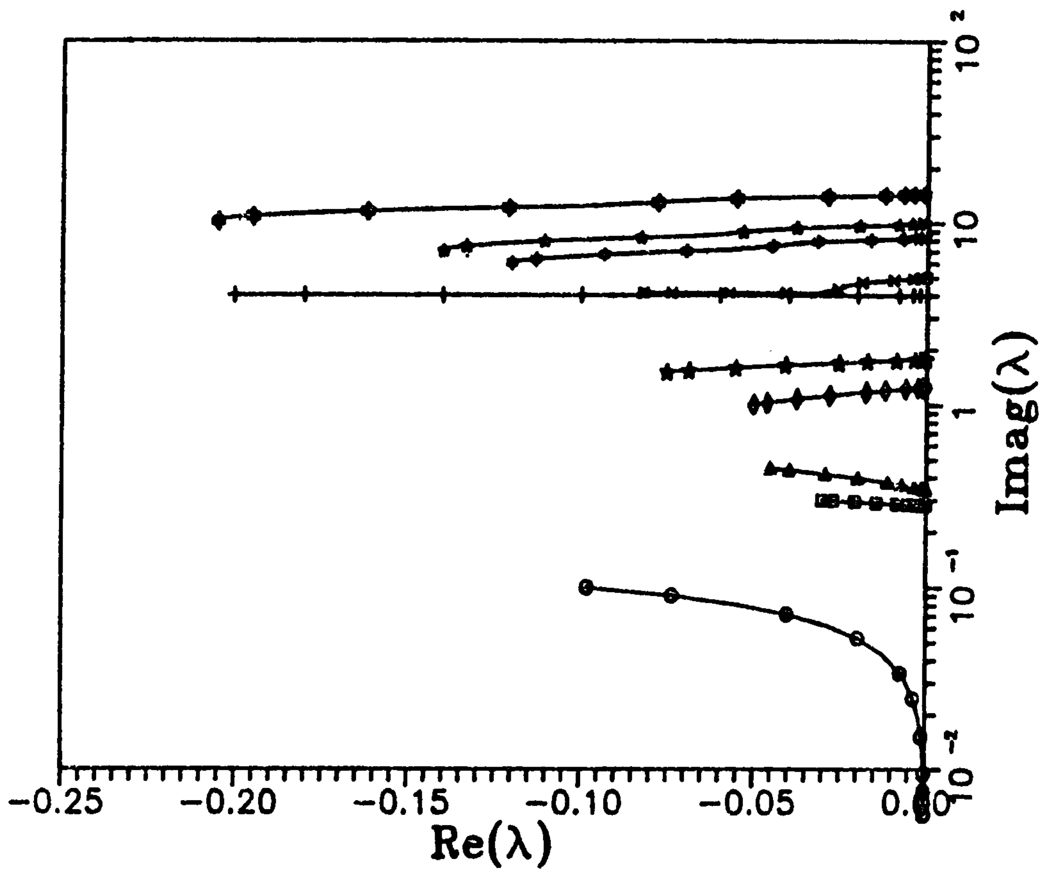


Figure 5.8 Closed loop eigenvalue trajectories for minimum eigenvalue sensitivity design

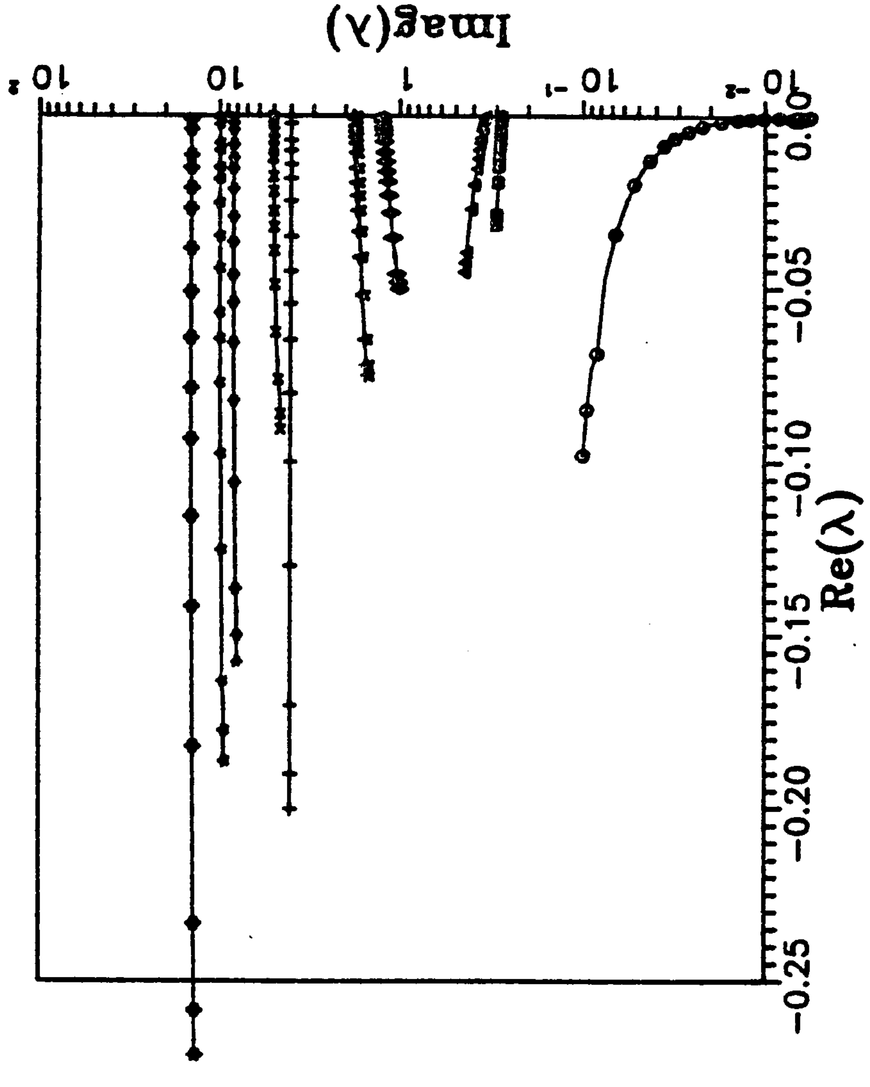


Figure 5.7 Closed loop eigenvalue trajectories for maximum stability robustness design

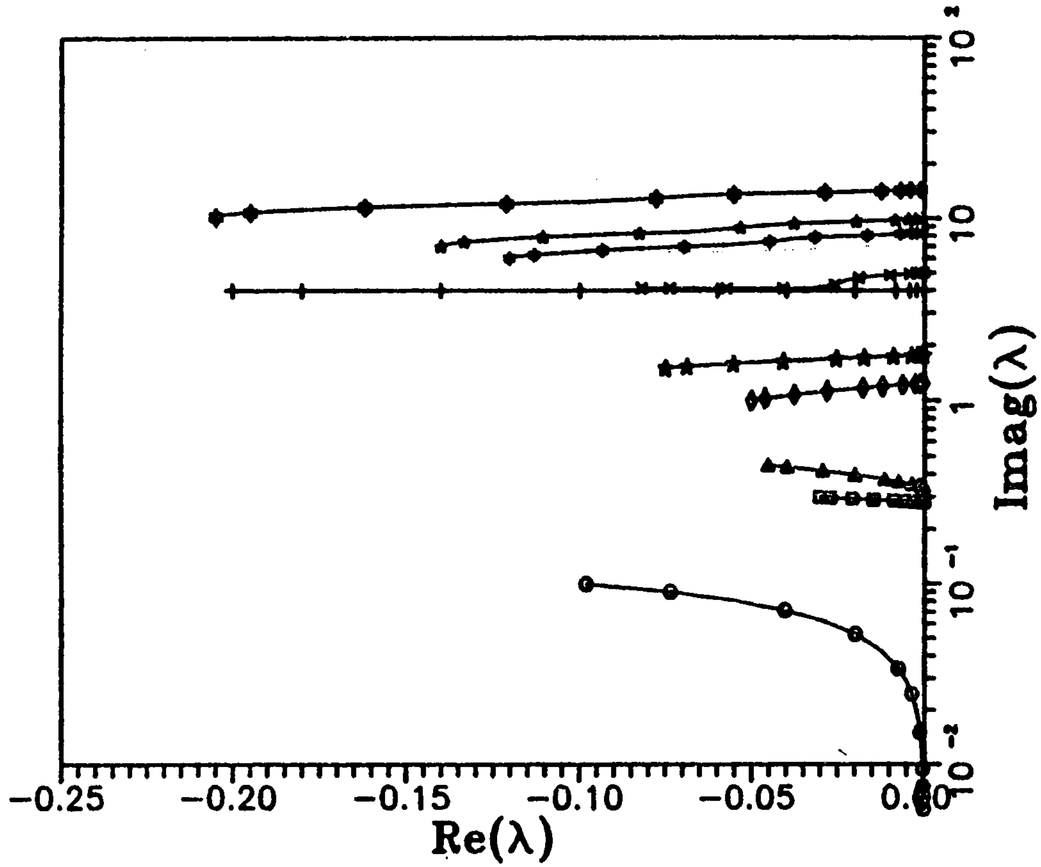


Figure 5.8 Closed loop eigenvalue trajectories for minimum mass design

Table 5.6 Converged Parameters for the Three Optimal Designs

CASE I: Minimum Eigenvalue Sensitivity
CASE II: Maximum Stability Robustness Design
CASE III: Minimum Mass Design

CASE	a_1	a_2	a_3	k	t_f	E	ρ_R
I	4.924	7.878	14.72	1443	.0841	.1480E+9	51.22
II	4.797	8.924	15.48	1481	.0939	.1496E+9	262.1
III	4.087	9.795	14.75	2315	.0709	.1494E+9	50.0

GAIN MATRICES:

$$G_I = \begin{bmatrix} 654.2 & -461.4 & 155.7 & -117.6 & -567.4 & -142.4 & -104.8 & 399.5 & -528.3 & 260.5 & 314.3 & 880.6 \\ -154.2 & -132.9 & 844.6 & -552.6 & -661.6 & -362.1 & -806.8 & 785.8 & 581.0 & -160.4 & 199.4 & -845.5 \\ -614.0 & 54.8 & -316.7 & -179.3 & 524.0 & 283.6 & -320.5 & 124.2 & 261.1 & 354.4 & 232.1 & 191.8 \\ -730.5 & 200.5 & 328.1 & -228.3 & 514.2 & -935.8 & -813.7 & 1060.0 & 539.3 & 323.1 & 193.1 & 221.0 \end{bmatrix}$$

$$G_{II} = \begin{bmatrix} 137.2 & -944.7 & 523.0 & 202.8 & 720.3 & -573.3 & -1292.0 & -288.3 & -291.2 & 251.1 & 146.2 & 68.9 \\ -530.1 & -874.1 & 811.2 & 442.1 & 822.8 & -1223.0 & -757.5 & 397.4 & -508.8 & -631.3 & 159.6 & -322.2 \\ -478.7 & -710.1 & -1443.0 & -1545.0 & -103.1 & 694.3 & -1202.0 & 410.2 & 1115.0 & 483.6 & 133.5 & -582.4 \\ 363.7 & -497.1 & -32.5 & 7.3 & -1335.0 & -1445.0 & -269.0 & -223.3 & 45.5 & 197.3 & 176.6 & -967.7 \end{bmatrix}$$

$$G_{III} = \begin{bmatrix} 1038.0 & -411.2 & 980.8 & 93.0 & -319.0 & -535.0 & -934.0 & -374.8 & 455.1 & -1032.0 & 54.5 & -1146.0 \\ -71.2 & -157.3 & 753.3 & -1051.0 & -703.4 & -282.2 & 823.4 & -928.3 & 285.2 & 37.9 & -549.3 & 321.6 \\ -611.7 & -198.3 & 302.2 & 210.1 & 10.9 & 1218.0 & -1106.0 & 1165.0 & 13.9 & 251.5 & 516.7 & -654.6 \\ -707.5 & -145.7 & 469.1 & -57.3 & 1194.0 & -1166.0 & -404.0 & -683.8 & -170.9 & 220.0 & 558.2 & 353.1 \end{bmatrix}$$

the observation made earlier from Figures 5.1 and 5.2, Figures 5.4 and 5.5 show a close relationship between condition number and the alternate stability robustness index for optimal robustness and sensitivity designs for small γ's. The similarity between sensitivity and robustness designs withrespect to condition number for small γ ($\lesssim .1$) is evident in Figure 5.5. Since the similarity also holds with respect to robustness index (see Figure 5.2) for the same range of γ ($\lesssim .1$), it appears that the eigenvalue condition number may serve as a stability robustness index when the system is marginally stable, i.e., for small γ. This observation is not unexpected since the condition number does not contain stability margin information. In other words, the condition number is essentially an eigenvalue sensitivity measure and as such, it is strongly related to stability robustness only when the eigenvalues lie close to the imaginary axis. It should be noted that the condition number does not appear to be a reliable indicator of stability robustness relative to the robustness measures of Patel and Toda (in this example) due to the asymptotic behavior for $\gamma \gtrsim .2$ (see Figure 5.5) while the other two robustness indices both show a steady increase in robustness over all γ (see Figures 5.2 and 5.4).

Figures 5.6 to 5.8 depict the closed loop eigenvalue trajectories of all 10 modes for the three designs. Much of the parallelness in the eigenvalue trajectories can be attributed to the continuation method of handling eigenvalue constraints; the smooth convergence and the numerical robustness in the trajectories is evident from these figures.

5.3 Evaluation of "True" Stability Robustness for the Three Designs

From Theorem 4.3.2 of Chapter 4, it is shown that the robustness measure of Patel and Toda represents a lower bound on the norm of system matrix perturbation for guaranteeing closed loop stability. It can be seen that the three designs correspond to three different levels of robustness at $\gamma=1.0$, as shown in Figure 5.2. For the dual purpose of verifying Patel and Toda's robustness criteria and to obtain an estimate of its conservatism when applied to parameter perturbation problems, we compute true stability limits by sweeping individual parameters while rigorously (nonlinearly) calculating the corresponding eigenvalues locii to detect the actual onset of instability. Table 5.6 shows the converged design variables at $\gamma=1.0$ where the parametric sweeping begins. In Tables 5.7 to 5.9, the upper and lower stability limits of individual parameters (determined by the first eigenvalue crossing the imaginary axis) are shown along with the 2-norms of the corresponding perturbation matrix (i.e., the nominal closed loop system matrix minus the perturbed matrix). It can be concluded from these tables that the maximum robustness design (Table 5.8) tolerates the largest amount of parameter perturbations, i.e., the maximally robust design is the most robust of the three designs. On the other hand, the minimum mass design generally tolerates the least amount of parameter perturbation, i.e. the least robust of the three designs. It can also be concluded that perturbations in structural parameters are generally less tolerable than gain element perturbations among the set of parameters considered here.

The relative robustness of the maximum robustness design can also be seen from the fact that there is essentially no lower (positive)

Table 5.7 True Stability Limits for Individually Perturbed Parameters at γ=1.0 of Minimum Sensitivity Design

LOWER LIMIT PERTURBATION $\|E\|_2$	LOWER LIMIT CRITICAL MODE #	LOWER STABILITY LIMIT p^{ℓ}	NOMINAL VALUE p°	UPPER STABILITY LIMIT p^{u}	UPPER LIMIT CRITICAL MODE #	UPPER LIMIT PERTURBATION $\|E\|_2$	min $\|p-p^{\circ}\|$ for instability
.3756E+1	10	4.136	a_1=4.924	5.514	8	.2117E+1	.590
.1772E+1	9	7.169	a_2=7.878	10.08	1	.6381E+1	.709
.3960E+1	3	13.69	a_3=14.72	15.30	9	.1994E+1	.580
.2251E+0	3	1360	k=1443	1616	2	.5627E+0	83
.1280E+2	6	.0801	t_F=.0841	.1557	7	.3275E+3	.0040
.1759E+2	6	.1272E+9	E=.1480E+9	.4144E+9	3	.2346E+3	.208E+8
.1330E+1	6	38.92	ρ_R=51.22	71.70	3	.1298E+1	12.3
.2512E+0	2	471.0	G(1,1)=654.2	713.0	3	.8374E-1	58.8
.7752E+0	1	-108.9	G(1,3)=155.7	311.4	2	.4360E+0	155.7
.1201E+1	2	-78.58	G(2,8)=785.8	1139	3	.4806E+0	353.2
.2454E+1	5	-783.3	G(3,9)=261.1	522.2	7	.6311E+0	261.2
.4409E+0	7	57.93	G(4,11)=193.1	1641	3	.5144E+1	135

Table 5.8 True Stability Limits For Individually Perturbed Parameters at γ=1.0 of Maximum Robustness Design

LOWER LIMIT PERTURBATION $\|E\|_2$	LOWER LIMIT CRITICAL MODE #	LOWER STABILITY LIMIT p^{ℓ}	NOMINAL VALUE p°	UPPER STABILITY LIMIT p^{u}	UPPER LIMIT CRITICAL MODE #	UPPER LIMIT PERTURBATION $\|E\|_2$	min $\|p-p^{\circ}\|$ for instability
.4591E+1	9	4.029	a_1=4.797	6.236	5	.5554E+1	.768
.4944E+1	8	8.299	a_2=8.924	9.994	9	.6392E+1	.625
.3662E+1	2	14.86	a_3=15.48	16.25	4	.5528E+1	.620
-	-	<0	k=1481	6220	7	.1094E+2	1480
.3157E+2	2	.0845	t_F=.0939	.1455	3	.2302E+3	.0094
.5537E+2	2	.1032E+9	E=.1496E+9	>.6133E+10	-	-	.464E+8
.1271E+1	4	78.6	ρ_R=262.1	655.2	2	.5213E+0	183.5
.5208E+0	1	-1372	G(1,1)=137.2	411.6	1	.8333E-1	274.4
.2317E+0	2	156.9	G(1,3)=523	706.0	2	.1158E+0	183
.4190E+1	2	-3179	G(2,8)=397.4	3974	2	.4190E+1	3576
.1990E+1	9	-55.7	G(3,9)=1115	1449	7	.4975E+0	334
.1466E+1	7	-441.5	G(4,11)=176.6	1942	3	.4399E+1	618

Table 5.9 True Stability Limits for Individually Perturbed Parameters at γ=1.0 of Minimum Mass Design

LOWER LIMIT PERTURBATION $\|E\|_2$	LOWER LIMIT CRITICAL MODE #	LOWER STABILITY LIMIT p^ℓ	NOMINAL VALUE p^*	UPPER STABILITY LIMIT p^u	UPPER LIMIT CRITICAL MODE #	UPPER LIMIT PERTURBATION $\|E\|_2$	min $\|p-p^*\|$ for instability
.1968E+1	7	3.800	a_1=4.087	4.740	10	.4870E+1	.287
.2837E+1	2	9.207	a_2=9.795	9.951	2	.6270E+0	.156
.3325E+1	6	14.01	a_3=14.75	15.11	8	.1729E+1	.360
.4040E+0	3	2199	k=2315	3241	2	.3535E+1	116
.1773E+2	1	.0638	t_F=.0709	.0730	6	.5236E+1	.0021
.2525E+2	1	.1105E+9	E=.1494E+9	.1613E+9	6	.7364E+1	.119E+8
.4399E+1	8	30	ρ_R=50	53	6	.3895E+1	3
.1375E+1	3	0	G(1,1)=1038	1141	3	.1375E+0	103
.3790E+0	2	833.6	G(1,3)= 980.8	1216	3	.6633E+0	147.2
.4795E+0	3	-1253	G(2,8)=-928.3	-649.8	7	.4475E+0	278.5
.2164E+1	8	-612.4	G(3,9)= 13.9	108.5	7	.3462E+0	94.6
.5275E+0	6	424.2	G(4,11)= 558.2	893.1	2	.1256E+1	134

stability limit on torsional stiffness, k, and upper stability limit on Young's modulus, E (see Table 5.8). Physically, this means that the effective stiffness introduced by the maximally robust control law is large enough to dominate the torsional stiffness. Furthermore, the maximum robustness design has more (4) gain elements that have individual perturbation ranges which include zero gains. This implies, granting that actuator failures can be modelled approximately by corresponding gain elements having zero values, that the maximum robustness design may have the best chance of the three designs studied, to remain closed loop stable for significant subsets of actuator and/or sensor failures/malfunctions.

The conservatism of the robustness measure, given by Eq. (4.3.20a) of Theorem 4.3.2, in guaranteeing closed loop stability for a bounded perturbation can be observed from the data in Tables 5.7 to 5.9. We first note that all twelve individually perturbed matrix norms, as given in the above tables (corresponding to upper and lower stability limits), are substantially greater than the predicted lower bounds that guarantee stability. In some cases, several orders of magnitude of conservatism is evident. This result, in essence, demonstrates the validity of the sufficient condition of Patel and Toda's stability robustness criteria. It should be noted that the above results are apparently typical, the predicted matrix norm bounds that guarantee stability is quite often a few orders of magnitude smaller than true stability limits. This large degree of conservatism is not unexpected for this type of robustness measure since a single scalar measure of perturbation magnitude in a multi-dimensioned parameter space (in our case, of dimension 55) is bound to be highly conservative, not to mention the heavily structured nature of the parameterized perturbations introduced here, whereas unstructured perturbations are assumed in the Patel/Toda theory. Nevertheless, we conclude from the above results that a significant numerical difference in the above robustness measure corresponds to a significant difference in the actual stability robustness of the closed loop system, even for parameterized perturbations. The above numerical results confirm the usefulness of the robustness measure as an objective function for robustness optimization, but does not resolve the problem that this measure (as well as all known robustness measures) is overly conservative if used as a predictive measure of the size of the perturbation which will lead to actual instability. The fortunate paradox is that maximization of this conservative robustness measure is demonstrated to be very effective in substantially increasing the system's true robustness.

6. CONCLUSIONS

6.1 Work in Retrospect

This dissertation addressed a few problems associated with unified design and optimization of structures and controllers of large flexible structures. In particular, we have focused on two main aspects, namely, a general design algorithm for conveniently handling many constraints efficiently and the problem of designing for stability robustness with respect to uncertainty in the closed loop system matrix.

Due to the high dimensionality of the dynamical systems under consideration and the need to impose many inequality constraint conditions in the optimization problem, an efficient and reliable optimization algorithm, such as the Simplex algorithm, is highly desirable. Thus, the nonlinear optimization problem is converted to a sequence of linear optimization problems so that the Simplex algorithm can be applied. Furthermore, the continuation method is used in conjunction with constraints on local step size bound on design variables to enhance convergence. Numerical results indicate that eigenvalue constraints used here for the purpose of pole placement works well although a few issues remain unresolved, such as the question of the "optimum" continuation step size and local parameter step size bounds. The proposed algorithm has been successfully applied to two significant problems, one of which involves a significant order dynamical model of a flexible structure with an attached rigid body with many constraints and the other, an optimal redesign of a cantilever beam finite element model as cited earlier.

Two sensitivity approaches for the design of closed loop systems under perturbations have been reviewed in detail. It was shown that the conditions for simultaneously satisfying eigenvalue placement and modal insensitivity constraints are in general restrictive and frequently unachievable although theoretically elegant and rigorous. However, it should be mentioned that if the above conditions of eigenvalue placement and modal insensitivity can be satisfied for a particular problem, the numerical effort does not involve any type of iterative calculation and the gain matrix can be easily computed (without the need for more expensive iterations involving first and second order eigenvalue derivative calculations). In other words, once a basis for the modal insensitivity subspace is computed, a simple linear combination of these basis vectors results in an eigenvector set that simultaneously place eigenvalues and have zero eigenvalue and eigenvector sensitivities. The corresponding gains can then be computed by a single matrix inversion. The alternative method of directly minimizing a weighted magnitude of eigenvalue sensitivity by mathematical programming is proposed and demonstrated here. In addition, a convenient scalar index designed for use as a cost function in eigenvalue sensitivity optimization is presented. By using concepts from matrix operator norm theory, a relationship between the scalar index and a linearly predicted bound on weighted eigenvalue perturbation is established.

Another important class of methods for dealing with closed loop stability under system uncertainties is referred to as robustness theory. Here the fundamental difference (with respect to sensitivity approach) is that it guarantees closed loop stability provided the perturbations are smaller than certain bounds. The robustness theory of Patel and Toda has been reviewed in detail. It is shown here that one of two robustness measure originally formulated by Patel and Toda can be derived more concisely using concepts from eigenvalue conditioning theory. This clearly demonstrates the physical importance and the theoretical soundness of their robustness measures. In addition, the robustness measures of an optimal linear quadratic regulator with full state feedback has been reviewed. This included a simpler derivation of the robustness condition and a derivation of a new condition on the weight matrices for "optimal" robustness.

We have presented and demonstrated a design algorithm and a stability robustness measure, which have not received much historical attention by researchers in dynamics and controls fields. Three different cost functions (total mass, stability robustness and eigenvalue sensitivity) have been successfully optimized with respect to a unified set of design parameters which incuded structural and control parameters and actuator locations. It was found that some similarity in the convergence histories of cost function exist between the minimum eigenvalue sensitivity and maximum robustness designs although their absolute values differed significantly. The attractive practical consequences of optimizing the robustness criteria was verified in spite of its conservatism when used as a predictive bound on allowable perturbation norms in the system matrix. It can be concluded that a significant difference in the robustness measure does indeed correspond to a significant difference in the actual stability robustness of the closed loop system.

Finally, it should be remarked that a major portion of the theoretical work reported here involved the unification and interpretation of various known results in the context of a particular class of problems. This is indeed the author's intention and is a major contribution of this work.

A summary of the contributions of this dissertation is as follows:

- o Presents a general sequential linear programming algorithm for unified structure and controller optimization problems. This algorithm is general and applies to a large family of optimization problems, especially those having many inequality constraints.
- o Presents a unified discussion of the main theorems associated with eigenvalue/eigenvector placement and modal insensitivity.
- o Presents new results and insights on existing eigenvalue sensitivity measure using matrix operator norm concepts.
- o Generalizes the eigenvector derivative formula to non-self-adjoint system. This corrects errors present in several texts and papers and generally serves to illuminate the relationship between the two independent eigenvector normalizations and the associated eigenvector derivatives. A simple relationship between the left and right eigenvector derivative expansion coefficients are found.
- o Relates Patel and Todas' stability robustness measures to eigenvalue conditioning concepts in computational matrix theory. A new derivation of this robustness criteria is presented using conditioning concepts only.

- o Numerically demonstrates the validity and usefulness of Patel and Todas' stability robustness criteria. Some similarity is shown between designs found minimizing an eigenvalue sensitivity measure and maximizing the Patel and Toda robustness measure.

6.2 Directions for Further Research

The research reported here should be extended as follows:

- o The direct eigenvalue placement approach (using a judicious choice on eigenvector sets via apriori computed basis vectors of null spaces), does not require eigenvalue derivative calculations. Thus, this approach may require less numerical work for computing gain matrices. However, plant parameters may not be as conveniently included as design variables in a unified approach. It would be interesting to study this seemingly more direct approach to pole placement combined with sequential linear programming and continuation methods to tune plant parameters, sensor/actuator parameters and so on, versus the methods presented herein.

- o To obtain stability robustness measures with respect to structural uncertainties that are minimally conservative. This amounts to a need for more accurate models to represent actual uncertainties and perturbations of "real" systems, and the need to develop robustness measures which can accommodate "structure" in the plant perturbations. Familiarity with realistic mathematical models of each physical system seems imperative for attacking the above problem.

- The problem of finding an "optimal" or largest possible continuation step size along with the associated local step size bound on parameter changes. This is largely an artistic issue but is nonetheless important vis-a-vis efficient implementations.

REFERENCES

[1] Proceedings of the First VPI&SU/AIAA Symposium on Dynamics and control of Large Flexible Spacecraft, Blacksburg, VA, ed. by L. Meirovitch, June 1977.

[2] Proceedings of the Second VPI&SU/AIAA Symposium, 1979.

[3] Proceedings of the Third VPI&SU/AIAA Symposium, 1981.

[4] Proceedings of the Fourth VPI&SU/AIAA Symposium, 1983.

[5] Proceedings of the Fifth VPI&SU/AIAA Symposium, 1985.

[6] Balas, M.J., "Trends in LSS Control Theory: Fondest Hopes, Wildest Dreams," IEEE Trans. on AC, Vol. AC-27, No.3, June 1982, pp.522-35.

[7] Ashley, H., et.al., "Effects of Structural Flexibility on Spacecraft Control Systems," NASA SP-8016, April 1969.

[8] Kosut, R.L., "Robust Control of Flexible Spacecraft," Journal of Guidance and Control, Vol.6, No.2, April 1983, pp. 104-11.

[9] Bekey, I., and Naugle, J.E., "Just Over the Horizon in Space," Astronautics and Aeronautics, May 1980, pp.64-76.

[10] Khot, N.S., et.al., "Optimal Structural Modifications to Enhance the Optimal Active Vibration Control of Large Flexible Structures," 26th Structures, Structural Dynamics, and Materials Conference, Orlando, FL, April 15-17, 1985.

[11] Bodden, D.S. and Junkins, J.L., "Eigenvalue Optimization Algorithms for Structural/Control Design Iterations," ACC, San Diego, CA, June 6-8, 1984.

[12] Junkins, J.L., Bodden, D.S., and Turner, J.D., "A Unified Approach to Structure and Control System Design Iterations," Fourth International Conference on Applied Numerical Modelling, Tainan, Taiwan, December 27-29, 1984.

[13] Haftka, R.T., et.al., "Sensitivity of Optimized Control Systems to Minor Structural Modifications," 26th Structures, Structural Dynamics & Materials Conference, Orlando, FL, April 15-17, 1985.

[14] Junkins, J.L. and Rew, D.W., "A Simultaneous Structure/Controller Design Iteration Method," ACC, Boston, MA, June 1985.

REFERENCES

[15] Dunyak, J.P., Junkins, J.L., and Watson, L.T., "Robust Nonlinear Least Squares Estimation Using the Chow-Yorke Homotopy Method," Journal of Guidance and Control, Vol.7, No.6, Nov-Dec 1984, pp.752-4.

[16] Dantzig, G.B., Linear Programming and Extensions, Princeton University Press, 1963.

[17] Hadley, G., Linear Programming, Addison-Wesley Publishing Co., Inc., Reading, MA, 1962.

[18] Gass, S.I., Linear Programming, New York: McGraw-Hill, 1958.

[19] Horta, L.G., Juang, J-N. and Junkins, J.L., "A Sequential Linear Optimization Approach for Controller Design," AIAA Paper 85-1971-CP, AIAA Guidance, Navigation & Control Conference, Snowmass, CO, August 19-21, 1985.

[20] Lim, K.B. and Junkins, J.L., "Minimum Sensitivity Eigenvalue Placement via Sequential Linear Programming," Proceedings of the Mountain Lake Dynamics and Control Institute, Virginia Polytechnic Institute, ed. by J.L. Junkins, Mountain Lake, VA, June 9-11, 1985, pp.122-144.

[21] Lim, K.B. and Junkins, J.L., "Optimal Redesign of Dynamic Structures via Sequential Linear Programming," Fourth International Modal Analysis Conference, Feb. 3-6, 1986, Los Angeles, CA, pp. 1615-1620.

[22] Raman, K.V. and Calise, A.J., "Design of an Optimal Output Feedback Control System with Modal Insensitivity," AIAA Paper 84-1940, AIAA Guidance and Control Conference, Seattle, WA, August 20-22, 1984.

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

[24] Kosut, R.L., Salzwedel, H. and Emami-Naeini, A., "Robust Control of Flexible Spacecraft," Journal of Guidance and Control, Vol.6, No.2, March-April 1983, pp.104-111.

[25] Safonov, M.G., Laub, A.J. and Hartmann, G.L., "Feedback Properties of Multivariable Systems: The Role and Use of the Return Difference Matrix," IEEE Trans. on AC, Vol. AC-26, No.1, Feb. 1981, pp.47-65.

[26] Lehtomaki, N.A., Sandell, N.S., Jr. and Athans, M., "Robustness Results in Linear Quadratic Gaussian Based Multivariable Control Designs," IEEE Trans on AC, Vol. AC-26, No.1, Feb. 1981, pp.75-92.

REFERENCES

[27] Freudenberg, J.S., Looze, D.P. and Cruz, J.B., "Robustness Analysis Using Singular Value Sensitivities,"International Journal of Control, Vol.35, No.1, 1982, pp.95-116.

[28] Hefner, R.D. and McKenzie, C.H. "Robust Controller Design Using Frequency Domain Constraints,"Proceedings of the 4th VPI&SU Symposium on Dynamics and Control of Large Spacecraft, Blacksburg, VA, ed. L. Meirovitch, June 1983, p.549.

[29] Sandell, N.R., "Robust Stability of Multivariable Feedback Systems," Proceedings of 1978 Allerton Conference, Urbana, ILL, 1978, pp.471-9.

[30] Patel, R.V. and Toda, M., "Quantitative Measures of Robustness for Multivariable Systems," Proceedings of JACC, San Francisco, TP8-A, 1980.

[31] Yedavalli, R.K., Banda, S.S. and Ridgely, D.B., "Time-Domain Stability Robustness Measures for Linear Regulators," Journal of Guidance and Control, Vol.8, No.4, July-August 1985, pp.520-24.

[32] Newsom, J.R. and Mukhopadhyay, "A Multiloop Robust Controller Design Study Using Singular Value Gradients," Journal of Guidance and Control, Vol.8, No.4, July-August 1985, pp.514-19.

[33] Stein, G., "LQG-Based Multivariable Design: Frequency Domain Interpretation,"Multivariable Analysis and Design Techniques, AGARD Lecture Series 117, September 1981, pp.5.1-5.9.

[34] Doyle, J.C., "Multivariable Design Techniques Based on Singular Value Generalizations of Classical Control,"Multivariable Analysis and Design Techniques, AGARD Lecture Series 117, September 1981, pp.4.1-4.9.

[35] Safonov, M.G. and Chen, B.S., "Multivariable Stability Margin Optimization with Decoupling and Output Regulation," Proceedings of the IEEE, Vol.129, Pt.D, No.6, November 1982, pp.276-282.

[36] Howze, J.W. and Cavin, R.K., "Regulator Design with Modal Insensitivity," IEEE Trans. on AC, Vol. AC-24, No.3, June 1979, pp.466-9.

[37] Raman, K.V., "Modal Insensitivity with Optimality," Ph.D. Dissertation, Drexel University, Philadelphia, PA, June 1984.

[38] Kalman, R.E., Falb, P.L. and Arbib, M.A., Topics in Mathematical System Theory, New York: McGraw-Hill, 1969.

REFERENCES

[39] Andry, A.N. Jr., Shapiro, E.Y. and Chung, J.C., "Eigenstructure Assignment for Linear Systems," IEEE Trans. on AES, Vol.AES-19, No.5, September 1983, pp.711-728.

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

[41] Junkins, J.L., Optimal Estimation of Dynamical Systems, Sijthoff & Noordhoff, 1978.

[42] Elliot, K.B. and Mitchell, L.D., "Realistic Structural Modifications: Part I Theoretical Development," Third International Modal Analysis Conference, January 1985, Orlando, FL, pp.471-476.

[43] Stein, G. and Doyle, J.C., "Singular Values and Feedback: Design Examples," Proceedings of 1978 Allerton Conference, Urbana, IL, 1978.

[44] Johnson, T.L. and Athans, M., "On the Design of Optimal Constrained Dynamic Compensators for Linear Constant Systems," IEEE Trans. on AC, December 1970, pp.658-660.

[45] Palacios-Gomez, F., Lasdon, L. and Engquist, M., "Nonlinear Optimization by Successive Linear Programming," Management Science, Vol.28, No.8, October 1982, pp.1106-1120.

[46] Palacios-Gomez, F., "The Solution of Nonlinear Optimization Problems Using Successive Linear Programming," Ph.D Dissertation, The University of Texas, Austin, TX, August 1980.

[47] Kalman, R.E. and Bucy, R.S., "New Results in Linear Filtering and Prediction Theory," Trans. of ASME, Journal of Basic Engineering, Vol.83, March 1961, pp.95-108.

[48] Luenberger, D.G., "Observers for Multivariable Systems," IEEE Trans of AC, Vol. AC-11, April 1966, pp.190-197.

[49] Levine, W.S. and Athans, M., "On the Determination of the Optimal Constant Output Feedback Gains for Linear Multivariable Systems," IEEE Trans on AC, Vol. AC-15, No.1, February 1970, pp.44-48.

[50] Mendel, J.M., "A Concise Derivation of Optimal Constant Limited State Feedback Gains," IEEE Trans. on AC, Vol. AC-19, August 1974, pp.447-448.

[51] Chen, C.T., Introduction to Linear Systems Theory, Holt, Rinehart & Winston, Inc., 1970.

REFERENCES

[52] Shapiro, E.Y., "On the Lyapunov Matrix Equation," IEEE Trans. on AC, October 1974, pp.594-6.

[53] Stewart, G.W., Introduction to Matrix Computations, New York: Academic Press, 1973.

[54] Noble, B. and Daniel, J.W., Applied Linear Algebra, Englewood Cliffs, NJ: Prentice-Hall, 1977.

[55] Bellman, R., Introduction to the Mathematical Theory of Control Processes, Vol.1, New York: Academic Press, 1967.

[56] Patel, R.V., Toda, M. and Sridhar, B., "Robustness of Linear Quadratic State Feedback Designs in the Presence of System Uncertainty," IEEE Trans. on AC, Vol. AC-22, No.6, December 1977, pp.945-949.

[57] Anderson, B.D.O. and Moore, J.B., Linear Optimal Control, Englewood Cliffs, NJ: Prentice-Hall, 1971.

[58] Klema, V.C. and Laub, A.J., "The Singular Value Decomposition: Its Computation and Some Applications," IEEE Trans. on AC, Vol. AC-25, No.2, April 1980, pp.164-176.

[59] Forsythe, G.E., Malcolm, M.A. and Moler, C.B., Computer Methods for Mathematical Computations, Englewood Cliffs, NJ: Prentice Hall, 1977.

[60] Collins, J.D., et.al., "Statistical Identification of Structures, AIAA Journal, Vol.12, No.2, February 1974, pp.185-190.

[61] Berman, A. and Flannelly, W.G., "Theory of Incomplete Models of Dynamic Structures," AIAA Journal, Vol.9, No.8, August 1971, pp.1481-1487.

[62] Wilkinson, J.L., The Algebraic Eigenvalue Problem, Oxford, England: Oxford University Press, 1965, p.68-.

[63] Chou, Y-F. and Chen, J-S., "Structural Dynamics Modification via Sensitivity Analysis," Third International Modal Analysis Conference, January 1985, Orlando, FL, pp.483-9.

[64] Meirovitch, L., Computational Methods in Structural Dynamics, Sijthoff & Noordhoff, 1980, p.102-.

[65] Wittrick, W.H., "Rates of Change of Eigenvalues with Reference to Buckling and Vibration Problems,"Journal of the Royal Aeronautical Society, Vol.66, No.621, September 1962, pp.590-1.

REFERENCES

[66] Fox, R.L. and Kapoor, M.P., "Rates of Change of Eigenvalues and Eigenvectors," AIAA Journal, Vol.6, No.12, December 1968, pp.2426-9.

[67] Plaut, R.H. and Huseyin, K., "Derivatives of Eigenvalues and Eigenvectors in Non-Self-Adjoint Systems," AIAA Journal, Vol.11, No.2, February 1973, pp.250-1.

[68] Porter, B. and Crossley, T.R., Modal Control: Theory and Applications, Taylor & Francis Ltd., London, 1972.

[69] Wonham, W.M., "On Pole Assignment in Multiinput, Controllable Linear Systems," IEEE Trans. on AC, Vol. AC-12, 1967, pp.660-665.

[70] Moore, B.C., "On the Flexibility Offered by State Feedback in Multivariable Systems Beyond Closed-Loop Eigenvalue Assignments," IEEE Trans. on AC, Vol. AC-21, 1976, pp.689-692.

[71] Srinathkumar, S., "Eigenvalue/Eigenvector Assignment Using Output Feedback," IEEE Trans. on A.C., Vol. AC-23, No.1, 1978, pp.79-81.

[72] Porter, B. and D'Azzo, J.J., "Algorithm for Closed-Loop Eigenstructure Assignment by State Feedback in Multivariable Linear Systems," International Journal of Control, 1978, Vol.27, No.6, pp.943-947.

[73] Patel, R.V. and Munro, N., Multivariable System Theory and Design, Pergamon Press, Oxford, 1982.

[74] Kriendler, E., "On the Definition and Application of the Sensitivity Function," Journal of Franklin Institute, Vol.285, 1968, pp.26-36.

[75] Meirovitch, L., Baruh, H. and Oz, H., "A Comparison of Control Techniques for Large Flexible Systems," Journal of Guidance and Control, Vol.6, No.4, July-August 1983, pp.302-310.

[76] Moerder, D.D. and Calise, A.J., "Convergence of a Numerical Algorithm for Calculating Optimal Output Feedback Gains," IEEE Trans. on AC, Vol. AC-30, No.9, September 1985, pp.900-3.

[77] Klien, G. and Moore, B.C., "Eigenvalue-Generalized Eigenvector Assignment with State Feedback," IEEE Trans. on AC, February, 1977, pp. 140-1.

[78] IMSL Library Reference Manual, International Mathematical and Statistical Library, Inc., 1982.

REFERENCES

[79] Bhattacharyya, S.P. and E. deSouza, "Pole Assignment via Sylvester's Equation," Systems and Controls Letters, Vol. 1, No. 4, 1982, pp. 261-3.

[80] Cavin, R.K. III and Bhattacharyya, S.P., "Robust and Well-Conditioned Eigenstructure Assignment via Sylvester's Equation," Optimal Control Applications & Methods, Vol. 4, 1983, pp. 205-12.

[81] Rogers, L.C., "Derivatives of Eigenvalues and Eigenvectors," AIAA Journal, Vol. 8, No. 5, May 1970, pp. 943-4.

[82] Meirovitch, L. and Ryland, G., "A Perturbation Technique for Gyroscopic Systems with Small Internal and External Damping," Journal of Sound and Vibration, Vol. 100, No. 3, 1985, pp. 393-408.

[83] Ryland, G., "Perturbation Methods for Slightly Damped Gyroscopic Systems," Ph.D. Dissertation, VPI&SU, September 1982, Chapter 2.

[84] Meirovitch, L., Analytical Methods in Vibrations, Macmillan Company, New York, 1967.

APPENDIX A

DERIVATIVES OF DAMPING FACTORS AND DAMPED FREQUENCIES

The eigenvalues of the system A(p) can be written as

$$\begin{aligned}\lambda_i &= \mathrm{Re}\{\lambda_i\} + j\,\mathrm{Im}\{\lambda_i\} \\ &= -\zeta_i\omega_i + j\,\omega_{d_i}\end{aligned} \tag{A.1}$$

where

$$\omega_{d_i} \triangleq \omega_i\sqrt{1-\zeta_i^2} \quad ; \quad \text{damped natural frequency}$$

and ω_i, ζ_i represents the undamped natural frequency and damping factors of i-th mode respectively. The derivative of damping factors can be derived from the relation

$$\zeta_i(p) = -\frac{\mathrm{Re}\{\lambda_i\}}{|\lambda_i|} \tag{A.2}$$

by taking the derivative of above equations with respect to parameter p, to get

$$\frac{\partial\zeta_i}{\partial p} = -\frac{\frac{\partial}{\partial p}\mathrm{Re}\{\lambda_i\}}{|\lambda_i|} + \frac{\mathrm{Re}\{\lambda_i\}}{|\lambda_i|^2}\frac{\partial}{\partial p}|\lambda_i| \tag{A.3}$$

To compute $\frac{\partial}{\partial p}|\lambda_i|$, we write

$$|\lambda_i|^2 = (\mathrm{Re}\{\lambda_i\})^2 + (\mathrm{Im}\{\lambda_i\})^2$$

$$\Longrightarrow \frac{\partial}{\partial p}|\lambda_i| = \frac{1}{|\lambda_i|}\left[\mathrm{Re}\{\lambda_i\}\ \mathrm{Re}\{\frac{\partial \lambda_i}{\partial p}\} + \mathrm{Im}\{\lambda_i\}\ \mathrm{Im}\{\frac{\partial \lambda_i}{\partial p}\}\right] \qquad (A.4)$$

Using (A.4), (A.3) can be rewritten as,

$$\frac{\partial \zeta_i}{\partial p} = \frac{-\mathrm{Re}\{\frac{\partial \lambda_i}{\partial p}\}}{|\lambda_i|} + \frac{\mathrm{Re}\{\frac{\partial \lambda_i}{\partial p}\}}{|\lambda_i|^3}(\mathrm{Re}\{\lambda_i\})^2$$

$$+ \frac{\mathrm{Re}\{\lambda_i\}\mathrm{Im}\{\lambda_i\}\mathrm{Im}\{\frac{\partial \lambda_i}{\partial p}\}}{|\lambda_i|^3} \qquad (A.5)$$

but since the first term on R.H.S. of (A.5) can be written as

$$\frac{-\mathrm{Re}\{\frac{\partial \lambda_i}{\partial p}\}}{|\lambda_i|^3}\left[(\mathrm{Re}\{\lambda_i\})^2 + (\mathrm{Im}\{\lambda_i\})^2\right]$$

(A.5) takes the final form,

$$\frac{\partial \zeta_i}{\partial p} = \frac{-\mathrm{Re}\{\frac{\partial \lambda_i}{\partial p}\}\ (\mathrm{Im}\{\lambda_i\})^2}{|\lambda_i|^3} + \frac{\mathrm{Re}\{\lambda_i\}\ \mathrm{Im}\{\lambda_i\}\ \mathrm{Im}\{\frac{\partial \lambda_i}{\partial p}\}}{|\lambda_i|^3}$$

$$= \frac{\mathrm{Im}\{\lambda_i\}}{|\lambda_i|^3}\left[\mathrm{Re}\{\lambda_i\}\ \mathrm{Im}\{\frac{\partial \lambda_i}{\partial p}\} - \mathrm{Im}\{\lambda_i\}\ \mathrm{Re}\{\frac{\partial \lambda_i}{\partial p}\}\right] \qquad (A.6)$$

Finally from (A.1) the damped frequency sensitivity can be written as

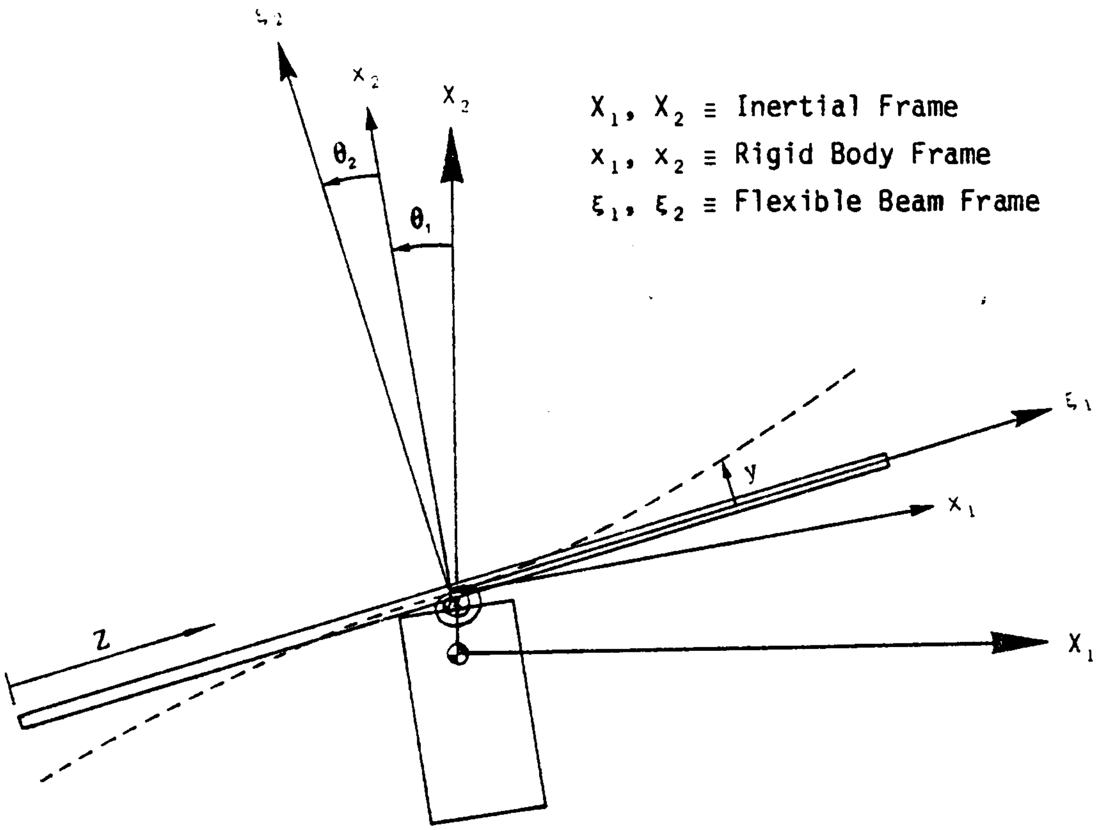
$$\frac{\partial \omega_{d_i}}{\partial p} = \frac{\partial}{\partial p}\ \mathrm{Im}\{\lambda_i\}$$

$$= \mathrm{Im}\{\frac{\partial \lambda_i}{\partial p}\} \qquad (A.7)$$

APPENDIX B

DERIVATION OF EQUATION OF MOTION FOR FLEXIBLE BEAM WITH ATTACHED RIGID BODY

The structure under consideration is shown in Figure A. It consists of a free-free beam with a rigid body attached to its center by means of a pin joint and a torsional spring. The material property of the beam is assumed to be isotropic, homogeneous with a uniform thickness distribution. In the sequel, the equation of motion for the above distributed system is derived using the extended Hamilton's principle and its discretization by means of the assumed modes method [84].



$$d = \frac{\frac{t_R}{2}}{1 + M_F/M_R} \qquad r = \frac{t_R}{2} - d$$

M_F ≡ Total Mass of Flexible Beam = $\rho_F(w_F\ d_F\ t_F)$

M_R ≡ Total Mass of Rigid Body = $\rho_R(w_R\ d_R\ t_R)$

I_F ≡ Mass Moment of Inertia of Flexible Beam about its c/m = $\frac{1}{12}\ \rho_F t_F d_F\ w_F^3$

I_R ≡ Mass Moment of Inertia of Rigid Body about its c/m = $\frac{1}{12}\ \rho_R d_R (t_R w_R^3 + t_R^3 w_R)$

Figure A Flexible beam with attached rigid body

Derivation of Total Kinetic Energy:

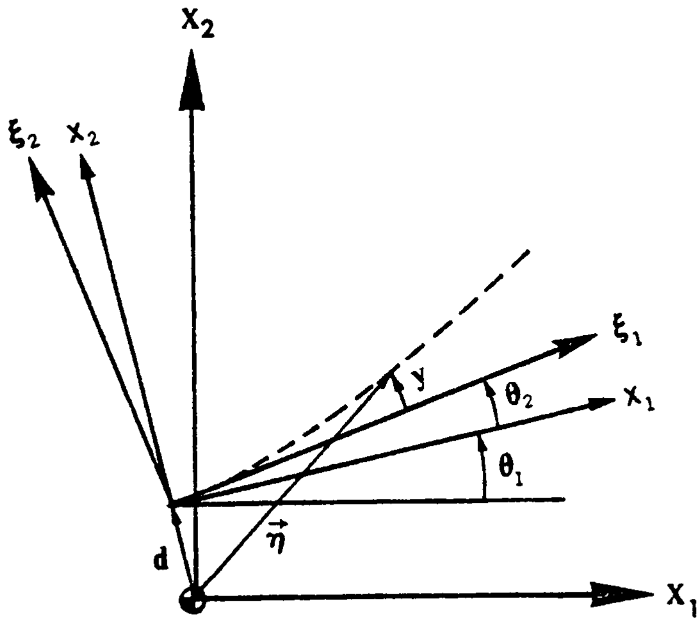
$$T = T_R + T_F \qquad \begin{array}{l} R \sim \text{Rigid Body} \\ F \sim \text{Flexible Beam} \end{array} \tag{B.1}$$

$$T_R = \frac{1}{2}(I_R + M_R r^2)\, \dot{\theta}_1^2 \tag{B.2}$$

$$T_F = \int_0^L \frac{1}{2}\, \rho_F A\, \dot{\vec{\eta}}(z) \cdot \dot{\vec{\eta}}(z)\, dz \tag{B.3}$$

where

$\dot{\vec{\eta}}(z)$ is the instantaneous inertial velocity of infinitesimal beam element at location z.



$$\vec{\eta} = d\hat{X}_2 + \left(z - \frac{L}{2}\right)\hat{\xi}_1 + y\hat{\xi}_2 \tag{B.4}$$

where

"^" denotes unit vectors

from geometry,

$$\hat{x}_2 = \sin\theta_2\ \hat{\xi}_1 + \cos\theta_2\ \hat{\xi}_2 \tag{B.5}$$

Therefore, (B.4) can be written in terms of only ξ_1, ξ_2 coordinates by using (B.5) to get,

$$\vec{\eta} = (d\ \sin\theta_2 + z - \frac{L}{2})\hat{\xi}_1 + (d\ \cos\theta_2 + y)\hat{\xi}_2 \tag{B.6}$$

To obtain time derivative w.r.t. inertial frame, we can use "transport theorem" to get,

$$\dot{\vec{\eta}} = (\frac{d\dot{\vec{\eta}}}{dt})_{\xi\text{-frame}} + \vec{\Omega}_{\xi/X} \times \vec{\eta}$$

where

$$\vec{\Omega}_{\xi/X} \equiv \text{angular velocity of } \xi\text{-frame w.r.t. } X\text{-frame}$$

$$= (\dot{\theta}_1 + \dot{\theta}_2)\hat{\xi}_3$$

$$\dot{\vec{\eta}} = (d\ \cos\theta_2\ \dot{\theta}_2)\hat{\xi}_1 + (-d\ \sin\theta_2\ \dot{\theta}_2 + \dot{y})\hat{\xi}_2$$

$$+ (\dot{\theta}_1 + \dot{\theta}_2)\hat{\xi}_3 \times [(d\ \sin\theta_2 + z - \frac{L}{2})\hat{\xi}_1 + (d\ \cos\theta_2 + y)\hat{\xi}_2]$$

Terms arising from the cross product are

$$[(\dot{\theta}_1 + \dot{\theta}_2)(d\ \sin\theta_2 + z - \frac{L}{2})]\hat{\xi}_2 - [(\dot{\theta}_1 + \dot{\theta}_2)(d\ \cos\theta_2 + y)]\hat{\xi}_1$$

$$\Longrightarrow \dot{\vec{\eta}} = [d\ \cos\theta_2\ \dot{\theta}_2 - (\dot{\theta}_1 + \dot{\theta}_2)(d\ \cos\theta_2 + y)]\hat{\xi}_1$$

$$+ [-d \sin\theta_2 \, \dot\theta_2 + \dot y + (\dot\theta_1 + \dot\theta_2)(d \sin\theta_2 + z - \tfrac{L}{2})]\hat\xi_2$$

$$\dot{\vec\eta} \cdot \dot{\vec\eta} = d^2\cos^2\theta_2\dot\theta_2^2$$

$$-2d \cos\theta_2 \, \dot\theta_2(\dot\theta_1 + \dot\theta_2)(d \cos\theta_2 + y)$$

$$+ (\dot\theta_1 + \dot\theta_2)^2(d \cos\theta_2 + y)^2$$

$$+ d^2\sin^2\theta_2\dot\theta_2^2$$

$$+ \dot y^2$$

$$+ (\dot\theta_1 + \dot\theta_2)^2(d \sin\theta_2 + z - \tfrac{L}{2})^2$$

$$- 2d \sin\theta_2 \, \dot\theta_2\dot y$$

$$- 2d \sin\theta_2 \, \dot\theta_2(\dot\theta_1 + \dot\theta_2)(d \sin\theta_2 + z - \tfrac{L}{2})$$

$$+ 2 \dot y(\dot\theta_1 + \dot\theta_2)(d \sin\theta_2 + z - \tfrac{L}{2})$$

$$= d^2\dot\theta_2^2 + \dot y^2 + (\dot\theta_1 + \dot\theta_2)^2[(d \cos\theta_2 + y)^2 + (d \sin\theta_2 + z - \tfrac{L}{2})^2]$$

$$-2d\dot\theta_2(\dot\theta_1 + \dot\theta_2)[\cos\theta_2(d\cos\theta_2 + y) + \sin\theta_2(d\sin\theta_2 + z - \tfrac{L}{2})]$$

$$- 2d \sin\theta_2 \, \dot\theta_2\dot y$$

$$+ 2\dot{y}\dot{\theta}_1(d\ \sin\theta_2 + z - \frac{L}{2}) + 2\dot{y}\dot{\theta}_2 d\ \sin\theta_2$$

$$+ 2\dot{y}\dot{\theta}_2(z - \frac{L}{2})$$

$$\dot{\vec{\eta}} \cdot \dot{\vec{\eta}} = d^2\dot{\theta}_2^2 + \dot{y}^2$$

$$+ (\dot{\theta}_1 + \dot{\theta}_2)^2\{d^2 + y^2 + (z - \frac{L}{2})^2 + 2d[\cos\theta_2\ y + \sin\theta_2(z - \frac{L}{2})]\}$$

$$-2d\dot{\theta}_2(\dot{\theta}_1 + \dot{\theta}_2)[d + \cos\theta_2\ y + \sin\theta_2\ (z - \frac{L}{2})]$$

$$+ 2\dot{y}(\dot{\theta}_1 + \dot{\theta}_2)(z - \frac{L}{2}) + 2\dot{y}\dot{\theta}_1 d\ \sin\ \theta_2$$

Since

$$\theta_1 = \theta_1(t)$$

$$\theta_2 = \theta_2(t)$$

$$y = y(z,t)\ ,$$

the kinetic energy of flexural beam is

$$T_F = \int_0^L \frac{1}{2}\ \rho A\dot{\vec{\eta}} \cdot \dot{\vec{\eta}}\ dz$$

$$= d^2\dot{\theta}_2^2 \int_0^L \frac{1}{2}\ \rho A\ dz \qquad \leftarrow\text{--}A_1$$

$$+ \int_0^L \frac{1}{2}\ \rho A\dot{y}^2 dt$$

$$+ (\dot{\theta}_1 + \dot{\theta}_2)^2 d^2 \int_0^L \frac{1}{2} \rho A \, dz \qquad \leftarrow\text{--}A_2$$

$$+ (\dot{\theta}_1 + \dot{\theta}_2)^2 \int_0^L \frac{1}{2} \rho A [y^2 + (z - \frac{L}{2})^2] dz \qquad \leftarrow\text{--}D$$

$$+ (\dot{\theta}_1 + \dot{\theta}_2)^2 2d \cos\theta_2 \int_0^L \frac{1}{2} \rho A y \, dz \qquad \leftarrow\text{--}C_1$$

$$+ (\dot{\theta}_1 + \dot{\theta}_2)^2 2d \sin\theta_2 \int_0^L \frac{1}{2} \rho A (z - \frac{L}{2}) dz \qquad \leftarrow\text{--}B_1$$

$$- 2d^2 \dot{\theta}_2 (\dot{\theta}_1 + \dot{\theta}_2) \int_0^L \frac{1}{2} \rho A \, dz \qquad \leftarrow\text{--}A_3$$

$$- 2d \dot{\theta}_2 (\dot{\theta}_1 + \dot{\theta}_2) \cos\theta_2 \int_0^L \frac{1}{2} \rho A y dz \qquad \leftarrow\text{--}C_2$$

$$- 2d \dot{\theta}_2 (\dot{\theta}_1 + \dot{\theta}_2) \sin\theta_2 \int_0^L \frac{1}{2} \rho A (z - \frac{L}{2}) dz \qquad \leftarrow\text{--}B_2$$

$$+ 2(\dot{\theta}_1 + \dot{\theta}_2) \int_0^L \frac{1}{2} \rho A \dot{y} (z - \frac{L}{2}) dz$$

$$+ 2\dot{\theta}_1 d \sin\theta_2 \int_0^L \frac{1}{2} \cancel{\rho A \dot{y} dz}$$

0, ξ_2-axis c/m of beam is constant w.r.t. time.

where

$$A = A_1 + A_2 + A_3$$

$$= d^2 (\dot{\theta}_2^2 + \dot{\theta}_1^2 + \dot{\theta}_2^2 + 2\dot{\theta}_1 \dot{\theta}_2 - 2\dot{\theta}_2 \dot{\theta}_1 - 2\dot{\theta}_2^2) \int_0^L \frac{1}{2} \rho A dz$$

$$= d^2 \; \dot{\theta}_1^2 \int_0^L \frac{1}{2} \rho A dz$$

$$= \frac{1}{2} M_F d^2 \dot{\theta}_1^2 \quad ; \quad M_F \triangleq \int_0^L \rho A dz \sim \text{total mass of beam}$$

$$B = B_1 + B_2$$

$$= (\dot{\theta}_1^2 + \dot{\theta}_2^2 + 2\dot{\theta}_1\dot{\theta}_2 - \dot{\theta}_1\dot{\theta}_2 - \dot{\theta}_2^2) 2d \sin\theta_2 \int_0^L \frac{1}{2} \rho A (z - \frac{L}{2}) dz$$

$$= (\dot{\theta}_1^2 + \dot{\theta}_1\dot{\theta}_2) 2d \sin\theta_2 \int_0^L \frac{1}{2} \rho A (z - \frac{L}{2}) dz$$

0, if c/m of beam is at middle or if beam has uniform mass distribution.

$$C = C_1 + C_2$$

$$= (\dot{\theta}_1^2 + \dot{\theta}_2^2 + 2\dot{\theta}_1\dot{\theta}_2 - \dot{\theta}_1\dot{\theta}_2 - \dot{\theta}_2^2) 2d \cos\theta_2 \int_0^L \frac{1}{2} \rho A y dz$$

$$= (\dot{\theta}_1^2 + \dot{\theta}_1\dot{\theta}_2) 2d \cos\theta_2 \int_0^L \frac{1}{2} \rho A y dz$$

0, if beam ξ_2-axis c/m location is at center of beam.

$$D = \frac{1}{2} (\dot{\theta}_1 + \dot{\theta}_2)^2 I_F$$

where

$$I_F \triangleq \int_0^L \rho A [y^2 + (z - \frac{L}{2})^2] \equiv \text{mass moment of inertia of beam w.r.t. c/m of beam}$$

In short, by neglecting the effect of flexural motion on the location of c/m of the beam, the kinetic energy expression for the flexural beam reduces to

$$T_F = \int_0^L \frac{1}{2} \rho A \dot{y}^2 \, dz \qquad \leftarrow\text{-- flexural contribution to kinetic energy}$$

$$+ \frac{1}{2} M_F d^2 \, \dot{\theta}_1^2 \qquad \leftarrow\text{-- rotational kinetic energy due to rotation of beam c/m.}$$

$$+ \frac{1}{2} I_F \, (\dot{\theta}_1 + \dot{\theta}_2)^2 \qquad \leftarrow\text{-- beam rotational kinetic energy w.r.t. beam c/m.}$$

$$+ 2(\dot{\theta}_1 + \dot{\theta}_2) \int_0^L \frac{1}{2} \rho A \dot{y}\left(z - \frac{L}{2}\right) dz \qquad \leftarrow\text{-- kinetic energy due to flexural-rigid body coupling.}$$

The total kinetic energy is

$$\begin{aligned} T &= T_R + T_F \\ &= \int_0^L \frac{1}{2} \rho A \dot{y}^2 dz + 2(\dot{\theta}_1 + \dot{\theta}_2) \int_0^L \frac{1}{2} \rho A \dot{y}\left(z - \frac{L}{2}\right) dz \\ &\quad + \frac{1}{2}(I_R + M_R r^2 + M_F d^2)\dot{\theta}_1^2 \\ &\quad + \frac{1}{2} I_F(\dot{\theta}_1 + \dot{\theta}_2)^2 \end{aligned} \tag{B.7}$$

The potential energy is due to elastic strain-energy and internal spring which elastically couples the rigid-body coordinate θ_2 and flexural slope at $z = z_S$ (spring location chosen here as $\frac{L}{2}$), i.e.

$$V = V_{\text{flexural strain}} + V_{\text{spring}}$$

$$= \frac{1}{2} \int_0^L EI \left[\frac{\partial^2 y}{\partial z^2}\right]^2 dz + \frac{1}{2} k\left(\Theta_2 + \left.\frac{\partial y}{\partial z}\right|_{z_s}\right)^2 \qquad (B.8)$$

Virtual work due to external torques,

$$u(t) = (u_1, u_2, u_3, u_4)^T$$

can be written as

$$\delta W = u_1 \; \delta\Theta_1$$

$$+ u_2 \; \delta\left(\Theta_1 + \Theta_2 + \left.\frac{\partial y}{\partial z}\right|_{z=a_1}\right)$$

$$+ u_3 \; \delta\left(\Theta_1 + \Theta_2 + \left.\frac{\partial y}{\partial z}\right|_{z=a_2}\right)$$

$$+ u_4 \; \delta\left(\Theta_1 + \Theta_2 + \left.\frac{\partial y}{\partial z}\right|_{z=a_3}\right) \qquad (B.9)$$

where a_1, a_2, a_3 represents actuator locations along the beam. It is assumed above that actuator u_1 is attached to the rigid body.

Using the method of assumed modes, we now represent the flexural motion by

$$y(z,t) = \sum_{i=1}^{N} \phi_i(z) q_i(t) = \phi^T(z) \;\; q(t) \qquad (B.10)$$

where

$N \equiv$ order of flexural d.o.f.(= 8 chosen)

$q_i \equiv$ flexural generalized coordinates.

$\phi_i \equiv$ assumed elastic mode shapes.

For the elastic mode shapes required above we choose here the free-free uniform beam eigenfunctions. We also note that the translational and rotational rigid-body modes of free-free beam are not included in ϕ_i since Θ_2 represents the beam rigid-body rotational motion and the c/m of beam is assumed a constant distance away from system c/m, which eliminates translational d.o.f.

We now define the generalized coordinates of system of dimension 10,

$$x(t) = [\Theta_1(t),\ \Theta_2(t),\ q_1(t),\ldots,q_8(t)]^T \tag{B.11}$$

The mass matrix can be deduced from total kinetic energy expressed in generalized coordinates, i.e., by substituting Eq. (B.10) into Eq. (B.7) to get

$$\begin{aligned} T = {} & \frac{1}{2}\,\dot{q}^T(t)\int_0^L \rho A\,\phi(z)\,\phi^T(z)\,dz\,\dot{q}(t) \\ & + (\dot{\Theta}_1(t) + \dot{\Theta}_2(t))\int_0^L \rho A(z - \tfrac{L}{2})\,\phi^T(z)dz\ \ \dot{q}(t) \\ & + \frac{1}{2}\,(I_R + I_F + M_R r^2 + M_F d^2)\,\dot{\Theta}_1^2 + I_F\dot{\Theta}_1\dot{\Theta}_2 + \frac{1}{2}\,I_F\dot{\Theta}_2^2 \end{aligned}$$

or

$$T = \frac{1}{2}\,\dot{x}^T \begin{bmatrix} J & I_F & m_{\Theta q}^T \\ I_F & I_F & m_{\Theta q}^T \\ m_{\Theta q} & m_{\Theta q} & m \end{bmatrix} \dot{x} \tag{B.12}$$

where

$$\underset{}{\overset{(8x8)}{m}} = \int_0^L \rho A[\phi(z)\ \phi^T(z)]dz$$

$$\overset{(8x1)}{m_{\Theta q}} = \int_0^L \rho A(z - \frac{L}{2})\ \ \phi(z)dz$$

$$J = I_R + I_F + M_R r^2 + M_F d^2$$

The stiffness matrix can be deduced from total potential energy expressed in terms of generalized coordinates, i.e., by substituting (B.10) into (B.8),

$$V = \frac{1}{2} q^T(t) \int_0^L EI\ [\frac{d^2\phi}{dz^2}\ (\frac{d^2\phi}{dz^2})^T]\ dz\ q(t) + \frac{1}{2} k\ \Theta_2^2(t)$$

$$+ k\Theta_2(t) \frac{d\phi}{dz}^T \Big|_{z_s} q(t) + \frac{1}{2} kq(t)\ [\frac{d\phi}{dz}(\frac{d\phi}{dz})^T] \Big|_{z_s} q(t)$$

or

$$V = \frac{1}{2} x^T \begin{bmatrix} 0 & 0 & 0 \\ 0 & k & s_{\Theta q}^T \\ 0 & s_{\Theta q} & s \end{bmatrix} x \qquad \text{(B.13)}$$

where

$$\overset{(8x8)}{s} = \int_0^L EI[\frac{d^2\phi}{dz^2}\ (\frac{d^2\phi}{dz^2})^T]dz + k[\frac{d\phi}{dz}(\frac{d\phi}{dz})^T]\Big|_{z_s}$$

$$\overset{(8\times1)}{s_{\theta q}} = k \left.\frac{d\phi}{dz}\right|_{z_s}$$

We observe that the spring, k, elastically couples flexible motion and θ_2 coordinates.

The generalized forces due to external torques can be obtained from virtual work expression by substituting (B.10) into (B.9) to get

$$\delta W = (u_1 + u_2 + u_3 + u_4)\delta\theta_1 + (u_2 + u_3 + u_4)\delta\theta_2$$

$$+ \left(u_2 \left.\frac{d\phi}{dz}^T\right|_{z=a_1} + u_3 \left.\frac{d\phi}{dz}^T\right|_{z=a_2} + u_4 \left.\frac{d\phi}{dz}^T\right|_{z=a_3}\right)\delta q$$

$$= Q^T \delta x$$

where

$$\overset{(10\times1)}{Q} = \begin{Bmatrix} u_1 + u_2 + u_3 + u_4 \\ u_2 + u_3 + u_4 \\ u_2 \left.\frac{d\phi}{dz}\right|_{z=a_1} + u_3 \left.\frac{d\phi}{dz}\right|_{z=a_2} + u_4 \left.\frac{d\phi}{dz}\right|_{z=a_3} \end{Bmatrix}$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 \\ \hline 0 & 1 & 1 & 1 \\ \hline \underline{0} & \left.\frac{d\phi}{dz}\right|_{z=a_1} & \left.\frac{d\phi}{dz}\right|_{z=a_2} & \left.\frac{d\phi}{dz}\right|_{z=a_3} \end{bmatrix} \begin{Bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{Bmatrix}$$

$$= \overset{(10x4)}{D} \quad \overset{(4x1)}{u}$$

Measurement (Output) Equations:

We assume measurements of flexural displacements, $y(s_1,t)$, $y(s_2,t)$, $y(s_3,t)$ and $y(s_4,t)$, on beam locations $z=s_1$, s_2, s_3 and s_4 respectively and rigid-body angular displacements of θ_1 and θ_2. In addition, we assume measurements of all corresponding velocities. The displacement measurements can then be written as

$$\overset{(6x1)}{y_1(t)} = \left\{ \begin{array}{c} \theta_1(t) \\ \theta_2(t) \\ \hline y(s_1,t) \\ \vdots \\ y(s_4,t) \end{array} \right\} = \left\{ \begin{array}{c} \theta_1(t) \\ \theta_2(t) \\ \hline \phi^T(s_1)\; q(t) \\ \vdots \\ \phi^T(s_4)\; q(t) \end{array} \right\}$$

$$= \left[\begin{array}{cc|c} 1 & 0 & \overset{(2x8)}{0} \\ 0 & 1 & \\ \hline \multicolumn{2}{c|}{\overset{(4x2)}{0}} & \phi^T(s_1) \\ & & \vdots \\ & & \phi^T(s_4) \end{array} \right] \left\{ \begin{array}{c} \theta_1(t) \\ \theta_2(t) \\ \hline q(t) \end{array} \right\} = H' x(t)$$

Since velocity measurements are assumed collocated, they can be written as

$$y_2(t) = \left\{ \begin{array}{c} \dot{\theta}_1(t) \\ \dot{\theta}_2(t) \\ \hline \dot{y}(s_1,t) \\ \vdots \\ \dot{y}(s_4,t) \end{array} \right\} = \quad \cdots \quad = H' \dot{x}$$

so,

$$y = \begin{Bmatrix} y_1 \\ y_2 \end{Bmatrix} = \begin{bmatrix} H' & 0 \\ 0 & H' \end{bmatrix} \begin{Bmatrix} x \\ \dot{x} \end{Bmatrix}$$

Summary of Equations

$$M\ddot{x} + Kx = Du$$

$$y = \begin{Bmatrix} y_1 \\ y_2 \end{Bmatrix} = \begin{bmatrix} H' & 0 \\ 0 & H' \end{bmatrix} \begin{Bmatrix} x \\ \dot{x} \end{Bmatrix}$$

where

$$M = \left[\begin{array}{cc|c} J & I_F & m_{\theta q}^T \\ I_F & I_F & m_{\theta q}^T \\ \hline m_{\theta q} & m_{\theta q} & m \end{array}\right] \sim \text{Mass Matrix}$$

$$K = \left[\begin{array}{cc|c} 0 & 0 & 0^T \\ 0 & k & s_{\theta q}^T \\ \hline 0 & s_{\theta q} & s \end{array}\right] \sim \text{stiffness matrix}$$

$$\underset{(10\times4)}{D} = \left[\begin{array}{cccc} 1 & 1 & 1 & 1 \\ \hline 0 & 1 & 1 & 1 \\ \hline \underline{0} & \left.\frac{d\phi}{dz}\right|_{z=a_1} & \left.\frac{d\phi}{dz}\right|_{z=a_2} & \left.\frac{d\phi}{dz}\right|_{z=a_3} \end{array}\right] \text{Force Distribution Matrix}$$

$$\underset{H}{6\text{x}10} = \left[\begin{array}{cc|c} 1 & 0 & \underline{0}^T \\ 0 & 1 & \underline{0}^T \\ \hline \multicolumn{2}{c|}{0} & \phi^T(s_1) \\ & & \vdots \\ & & \phi\,(s_4) \end{array}\right] \sim \text{Measurement Matrix}$$

$$\underset{x(t)}{10\text{x}1} = \left\{\begin{array}{c} \theta_1(t) \\ \theta_2(t) \\ \hline q_1(t) \\ \vdots \\ q_8(t) \end{array}\right\} \sim \text{generalized displacements}$$

$$\underset{u(t)}{(4\text{x}1)} = \left\{\begin{array}{c} u_1(t) \\ u_2(t) \\ u_3(t) \\ u_4(t) \end{array}\right\} \sim \text{control vector}$$

The assumed modes used were translated uniform, free-free beam eigenfunctions (see p. 164 of [84]). The characteristic equation and the eigenfunctions for the free-free beam assumed modes are as follows:

$$g(\beta L) = \cosh\beta L \cos\beta L - 1 = 0$$

where

$$\beta^4 = \frac{m}{EI}\,\omega^2$$

$$Y_i(z) = A_i[B_i(\sin\beta_i z + \sinh\beta_i z) - C_i(\cos\beta_i z + \cosh\beta_i z)]$$

where

$$A_i = \text{normalization constant}$$

$$B_i = \cos\beta_i L - \cosh\beta_i L \qquad i=1,\ldots,8 \; .$$

$$C_i = \sin\beta_i L - \sinh\beta_i L$$

The vita has been removed from the scanned document