SENSOR AND ACTUATOR SELECTION FOR FAULT-TOLERANT CONTROL OF FLEXIBLE STRUCTURES

by

William Neil McCasland
B.S., United States Air Force Academy (1979)
S.M., Massachusetts Institute of Technology (1981)

SUBMITTED TO THE DEPARTMENT OF AERONAUTICS AND ASTRONAUTICS IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February, 1989

Signature of Author _____
Department of Aeronautics and Astronautics, September 15, 1988

Certified by _____
Professor Wallace E. Vander Velde
Thesis Supervisor, Professor of Aeronautics and Astronautics

Certified by _____
Professor Andreas H. von Flotow
Thesis Supervisor, Assistant Professor of Aeronautics and Astronautics

Certified by _____
Professor Richard H. Battin
Thesis Supervisor, Adjunct Professor of Aeronautics and Astronautics

Certified by _____
Dr. Eliezer Gai
Thesis Supervisor

Accepted by _____
Professor Harold Y. Wachman
Chairman, Departmental Graduate Committee

SENSOR AND ACTUATOR SELECTION FOR FAULT-TOLERANT CONTROL OF FLEXIBLE STRUCTURES

by

William Neil McCasland

Submitted to the Department of Aeronautics and Astronautics
on August 31, 1988
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

ABSTRACT

Control system design for structural dynamics requires instrumenting the structure with sensors and actuators, the effectiveness of which depends strongly on spatial position. For the long duration missions typical of spacecraft, the possibility of failures must be considered and coverage provided for with redundancy and a reconfiguration strategy. To support these, it is mandatory that sensors and actuators be located such that sufficient controllability and observability of the flexible dynamics is assured in the presence of failures. This thesis develops a systematic procedure for finding such configurations of sensors and actuators for large-scale systems.

Norms on the controllability and observability open loop Gramian are taken as an index for optimization. The control objectives are reflected by truncating the time interval over which the Gramian is defined, and by scaling the state basis with respect to a normalizing, or reference, signal path. A convex programming problem is posed by taking, as independent variables, the power gain (linear gain squared) at a set of discrete locations. While the resulting optimal distribution is, in general, non-integer, it serves as an upper bound and to delineate a search neighborhood for a fixed number of (integer) quantized sensors or actuators with failure combinations. The Gramian is an attractive expression for controllability and observability properties because, in addition to posing a reachable or detectable volume of state space, it is linear in the contributions of individual sensors or actuators so does not require computation in the main recursion.

Examples with a simply-supported beam and a large-scale model (ACOSS Model No 2) illustrate the operation of the algorithms developed in the thesis. With the beam model, the effect on the optimal distribution to changes in the basis scaling and relative weightings are illustrated. Results for the ACOSS model show the ability to select actuator locations which generally improve the performance of a LQ regulator over those locations determined in previous work with an algorithm based on the span of modal influence coefficients, and to mitigate the effect of failures on those locations.

Thesis Supervisor:	Professor Wallace E. Vander Velde
Title:	Professor of Aeronautics and Astronautics
Thesis Supervisor:	Professor Andreas H. von Flotow
Title:	Assistant Professor of Aeronautics and Astronautics
Thesis Supervisor:	Professor Richard H. Battin
Title:	Adjunct Professor of Aeronautics and Astronautics
Thesis Supervisor:	Dr. Eliezer Gai
Title:	Director of Control and Decision Systems, CSDL

ACKNOWLEDGEMENTS

I am fortunate and grateful to acknowledge the important contributions of several people and organizations, without which this work would not have been possible. To my committee, thanks are extended for their freely given time and suggestions for research, their leavening of my overall program, and their review of this thesis. In particular, to Professor Vander Velde, for showing the value of precision and thoroughness, to Professor Von Flotow for his expansive and provocative classes and suggestions, to Professor Battin for illustrating the elegance of mathematics, and to Dr. Gai for a focus on practicality and for providing much of the resources needed.

Mr. Jim Harrison freely extended the support of his division at the Draper Laboratory as if I were a member of his staff. I am particularly grateful to Eva Moody, who took up the production of the final form of this document, bearing out the details, the changes, and a lot of the drudgery.

The contributions of Dr. Dan Hegg, also on the Draper staff, ought to be acknowledged on the same line as that of my committee. In addition to providing invaluable technical and mathematical depth in the general area of research, he was solely responsible for my ability to exercise the ideas developed on the large scale ACOSS model in the time I had. I am also pleased to single out in that group Dr. John Dowdle for his counsel on systems and controls theory, and to Mr. Tim Henderson for further assistance with the ACOSS model.

The Draper Laboratory provided a superb environment, both by its staff and facilities, for this research. My thanks go to the former President, Brig. Gen. Bob Duffy for making the Laboratory available, to the memory of Mr. Bill Manlove, for carrying out the details, and to the current President, Maj. Gen. 'Jake' Jacobson for carrying it on. The Lab has been a welcome home for all of my graduate education.

The Fannie and John Hertz foundation directly supported all four years of my graduate work, allowing me the possibility of starting it when I did and of picking it up again after an interruption for military obligations. It is an honor to be associated with them.

Most of all, I am pleased to have the sustained enthusiasm and cheerfulness of my wife, Susan. She gave up the good life of her native Southern California to dwell in the land of the bean and the cod so I could pursue this program. Her careful proofing of and artwork in this thesis also deserve credit and my gratitude.

DEDICATION

This work is dedicated to the blue-suited soldiers of the cold war. While the country prospered, they gave up much of what they could have had to train for a war they never want. Though I am proud to wear their uniform, their duty has not been mine. This is to those who pull alert, log the hours in the sky, in the hole, on the scopes, in the hangers or on the ramp, with a jet or with a rifle, in the remotest of parts, so the rest of us may carry on. This is also to those who answered when their watch on the cold war was interrupted by a call to go fight. They went into bandit country and many never returned while we carried on. This is for those who fought because it was their duty, even when their country's heart was not with them. This is especially for those who, training for cold war or fighting a hot one, paid for answering their call and laid down their lives. And in particular, this is to honor the unfulfilled hopes and agenda of one such soldier, a civil engineer at heart, but aviator by duty, William H. McCasland, 1st Lieutenant, USAF.

William N. McCasland
Major, USAF

Table of Contents

Nomenclature

(A,B,C)	Matrix triple forming state space representation of a strictly proper linear system.
$\mathbf{b}$	Columns of control influence matrix, B
$\mathbf{c}^T$	Row of output influence matrix, C
W_O	Observability Gram matrix
W_C	Control Gram matrix
T	State transformation
σ,Σ	Second order modes, and diagaonal matrix of σ.
$\mathbf{x}$	State vector of system (A,B,C)
$\mathbf{y}$	Output vector of systems (A,B,C) and 2nd order modal.
$\mathbf{u}$	Control input vector of systems (A,B,C) and 2nd order modal representation
$\mathbf{z}$	Controlled variable output of systems (A,B,C) and 2nd order modal representation
$\mathbf{w}$	distubance input to system (A,B,C)
$\underline{\eta}$	Modal displacement vector of 2nd order structural system
n	dimesion of state vector, $\mathbf{x}$
n	dimension of η
m	dimension of control vector, $\mathbf{u}$
l	dimension of output vector, $\mathbf{y}$
M,K	Mass and Stiffness matricies of structural model
M, m_i	Total power and individual site power constraints
$\mathbf{q}$	physical displacement variables in structural model
ω_n ω_d	modal frequency, natural and damped respectively
ϕ, Φ	mode shape and modal matrix of structural model, generalized eigenvectors of (M,K) with eigenvalue ω_n^2.
Ω^2	diagonal matrix of ω_n^2
ζ	modal damping ratio
Z	diagonal matrix of ζ
J_1, J_2, J_3, J_{LQ}	Scalar performance index, norm of controllability/observability or LQG cost functional
ξ	independant variables representing sensor/actuators placement or authority/sensitivity distribution
$\mathbf{f}$	failure mode vector of parameters
Π	probability distribution of failure modes
λ, $\lambda(A)$	eigenvalue of matrix A
$\mathbf{v}$, V	eigenvector and matrix of eigenvectors
tr(A)	Trace of matrix A
|A|	Determinant of matrix A
P_O, P_C	Riccati equation solutions, observation and control
Q_x,R_u	State and control weighting for LQ regulator
Q_w,R_y	State disturbance and measurement noise intensities
G	Control feedback gains
H	Kalman gain matrix

$g(s)$, $G(s)$	Transfer function, scalar and matrix
$\mathbf{g}$	gradient of scalar index J
$E[\ ]$	Expectation operator
$Pr(\ )$	Probability of event ( )
T_f	Time of failure, random
$F_{Tf}(t)$, $p_{Tf}(t)$	Probability distribution and density functions of T_f
$\lambda_f(t)$	Failure rate
$r(t)$	Reliability function

Chapter 1. Introduction

Active control of flexible structures has been an area of fruitful research, but flight hardware applications remain rare. A general consensus of research in the field is that a control system which must contend with structural modes within its bandwidth must be significantly more complex, in terms of hardware and software, than current generation control systems, in which the rigid body dynamics dominate the bandpass. Complexity is usually detrimental to the high reliability requirement typical of flight vehicles; and in the case of space systems, detrimental to the requirement of long mission life with essentially no maintenance as well. Reliability and life are particularly stressing problems for spacecraft. After thirty years of flight operations, industry and government have developed a set of engineering practices and standards that allow confidence in meeting these demands with a new vehicle design, confidence high enough to embark on some programs with only one or two flight units. Contrast Ranger, which in 1964 returned the first close-up lunar photographs on the seventh mission after six consecutive failures with the Hubble space telescope or Magellan Venus mapper, with only one vehicle fabricated each. These practices are codified, in the forms of MIL-STD-1543A for military [1] and NASA Handbook 5300.4 for civil [2] programs, stating certain basic requirements such as no single-point-failures, and a whole system reliability analysis enumerating failure modes and effects. Examining the control system architectures typical for a flexible space structure in light of these standards reveals a number of issues; one of them - sensor and actuator selection - is the focus of this thesis.

1.1 Reliability Issues in Control of Flexible Space Structures

We first wish to distinguish between issues of reliability and of robustness. The latter is commonly interpreted as that desirable property of controllers that are, at a minimum, stable in the presence of certain bounded variations in the physical dynamics (stability robustness), and may satisfy performance objectives as well (performance robustness). These variations are commonly taken to be uncertainties, say in knowledge of modal frequencies, and possibly degradations of component performance within anticipated limits over mission life. Component failures, on the other hand, may be thought of as operations outside of anticipated limits. Failures are usually considered out of the scope of any robustness requirement; otherwise, excessively conservative or even unrealizable

design constraints may result. Failures are also characterized by an unpredictable time of occurrence, and, for active control system components, generally require some accommodating action, even if it is as simple as removing power from the failed component.

While passive redundancy techniques are particularly attractive for spacecraft because of their 'no-action required' operating idea, they are not generally applicable to active, power-consuming hardware such as that incorporated into flight -control systems. The two main steps of active redundancy management are:

-1 Failure detection and isolation (FDI),

-2 Reconfiguration to accommodate failure effects.

Each step has been a focus of research, and we note some results as applied to structural control.

The high order of structural dynamics imposes stressing requirements on failure-detection methods exploiting a dynamic model (analytic redundancy). We also desire, however, to avoid simple replication of components just to support a voting method for FDI. Reference [3] examines analytic redundancy methods for structural systems, noting how unmodeled dynamics adversely affects detection signatures. Continuing in this line, references [4,5] examine the Failure Detection Filter, which is a particular implementation of analytic redundancy. While it is clear from these that FDI may be a challenging problem for structural control, there are a variety of approaches available. Current practice for spacecraft relies heavily on monitoring instrumentation signals from the telemetry stream for broad ranges of anomalies, so including dynamic criteria such as those represented by analytic redundancy is certainly feasible. We will therefore assume, for this thesis, a functional FDI system exists.

Reconfiguration is likely to involve bringing a redundant unit on line to replace the function of a failed unit. If, for a sensor or actuator failure, the redundant unit is functionally equivalent, then reconfiguration involves power, telemetry, and command signals, but not changes to the control law. If, however, the redundant unit has functional differences, such as a different geometric alignment, then the reconfiguration will involve the control law algorithms as well. In particular for control of structural dynamics, incorporating a spare in a different location, or, if there is not a standby pool, just accom-

modating the failure, should require a re-tuning of the controller in some manner. With failure states modeled by Markov chains, and the Linear-Quadratic-Gaussian (LQG) cost functional for a performance index, the optimal control problem for 'jump-parameter' linear systems represents a reconfiguration strategy. Sworder [6], Siljak [7], Birdwell [8] and Mariton [9] have contributed to this line, the main results of which are a 'hedging' of optimal gains by the possibility of transition to another mode and an idea of stochastic stability that may include transitions through modes that are not necessarily stabilizable. This framework is a solution to the problem where the number of failure modes is relatively small and the system under control must continue operating across failure-mode transitions: a higher degree of autonomy than may be necessary for spacecraft. The 'control mixer' concept developed in the self-repairing flight control system program of the Air Force Flight Dynamics Lab takes the objective of approximating the forces and moments of the unfailed control surfaces with those remaining or degraded after failures [10]. A similar line of reasoning in the frequency domain is pursued by Weiss, et al. [11] They seek to recover the original loop transfer function through compensator redesign, possibly reflecting increased plant uncertainty with the robustness bounds that are tractable with LQG theory. Joshi [12,13,14] also takes the LQG control objective and investigates the influence of non-linearities confined to a conic sector (including failures) in the sensor or actuator signal path, developing an example for the special case of flexible dynamics as well. Joshi's idea is to rely on passive robustness for failure accommodation, generating stability guarantees but not explicitly optimizing or recovering performance. Vander Velde [15] sets down general guidelines for reconfiguration strategy, emphasizing the flexibility of eigenstructure assignment and the desirability of retaining the original control objectives of the unfailed system. A common characteristic of all of these is they essentially pose the question of how to accomplish a re-design, given some change in the plant has occurred. For aircraft, this question may need immediate answers [16], though a viable option may be establishing a reduced level of flight performance for return to base. For spacecraft, near-real-time action is generally simpler, since some form of independent, coarse-pointing, controller is a common feature for initial attitude acquisition. For coverage of failures, the primary mission control system may be shut off, and coarse pointing maintained with the secondary system until commanded differently. Considerable ground-based engineering resources may then be brought to bear on reconfiguration over periods

of time from hours to days. Return-to-base, however, is not an option, and even for those few spacecraft in low-altitude, low-inclination, orbits, the revisit interval for maintenance is years. It is mandatory, then, that reconfiguration strategies for spacecraft deliver the performance required of the mission, and that the remaining unfailed components allow such a reconfiguration.

The control system of modern spacecraft generally consists of a collection of dedicated sensors and actuators, and a flight computer which may have other functions as well. We will exclude here considerations of the reliability of the computer, since insights from dynamics and control theory are not directly relevant for that major component. For current-generation vehicles, the total number of flight-control components and operational permutations is small, so it is relatively easy to construct a reliability diagram of the control system. Single-point failures can be identified, redundant components incorporated, and the system reliability estimated as required by specification [1,2]. Much of the hardware for a control system for structural dynamics is similar, but there is an important dynamic difference. The vector space for rigid-body rotational motion is R^3, where there are clear physical and geometric properties. Structural dynamics, on the other hand, are approximated in a larger-dimensioned R^n, which is a truncation of an infinite-dimensional process. Control system effectors for rigid-body control commonly have only one or two directions of effectiveness, and, for pure torque devices, are not affected by location, so it is straightforward to insure an adequate level of redundancy for each of a roll, pitch, or yaw direction. For structural control, effectors may couple into all modes to some degree, and the coupling is a strong function of spatial location. It is not entirely obvious how to insure adequate measurement and control authority for important modes even without considering failures, and redundancy requirements are more obscure. We would like to use as few components as necessary. In the case of rigid-body control, geometric insights result in some elegant arrangement of sensors or actuators [17,18] for general, or 'm-out-of-n' sparing with some level of fault detection and isolation, but direct extension to the higher-ordered dynamics of flexible structures does not seem possible.

The advantage of general sparing becomes stronger as the number of components increases. It is possible to decouple reconfiguration considerations from sensor and actuator placement for structural control. We could determine, assuming no failures, the

number and placement of a configuration of sensors and actuators necessary for the control objectives, then provide whatever level of sparing is necessary for each component to insure at least one per site is functional over the mission life. This is a particularly expensive approach, as will be illustrated by the following example. Suppose we need 10 components (sensors or actuators) to function over the mission life with some specified probability; but the probability any one component operates for the full mission life is 0.9. Table 1-1 contrasts the level of individual redundancy (multiplicity of components at a site) necessary to meet a range of system reliability levels with the number of additional components required if any component could operate in place of any other.

Table 1-1. Comparison of Simple Redundancy and General Sparing.

	Simple Redundancy			General Sparing		
Spec. Reliability	Level	System Reliability	Components Total	No. of Spares	System Reliability	Components Total
0.9	2	0.904	20	3	0.966	13
0.99	3	0.990	30	4	0.991	14

It seems sensors and actuators for structural control should be well suited for general sparing if they are properly located since each one may interact with several (or all) of the important modes. The main focus of this research is developing an efficient method to find locations such that the degradation from failures is minimized. General sparing is implicitly assumed.

1.2 Thesis overview

Chapter 2 presents some results concerning sensor and actuator selection for distributed parameter systems in general, and specialized applications to structural dynamic systems. Algorithms based on finite-dimensional approximations to structural dynamics are also reviewed, but none are found entirely satisfactory when the combinatorial complexity of failures is introduced. Chapter 3 formulates an open loop controllability/observability optimization problem that is amenable to failure-induced complexity. Norms on the control-lability or observability Gramian are selected as an optimization parameter because the Gramian adequately reflects system properties, and its linearity permits incorporating failure modes with a minimum of computational effort. A

pair of algorithms are developed in Chapter 4: one is a purely combinatorial search and the second solves a continuous 'relaxation', providing a good starting point for the combinatorial algorithm as well as some means of limiting the search neighborhood. Applications examples are discussed in Chapter 5 for a simply-supported beam model, and then for a large scale-space optical structure. Chapter 6 summarizes the results and contributions of the thesis.

The main contribution of this research is a practical approach to placement of sensors and actuators accounting for possible failures. An unwieldy combinatorial problem is broken down into a sequence of economical steps; the key one is posing a 'relaxation' of the commonly adopted constraint that a sensor or actuator need be considered in integer quantities only. Application results demonstrate the ability to select actuator locations that are generally superior, in terms of the cost of a stochastic LQ regulator as an evaluation, than those by a method based on modal influence coefficients not considering failures, and the ability to minimize the impact of failures. The feasibility of general sparing, as opposed to simple redundancy, is illustrated.

Chapter 2. Background

In this chapter, we review previous work directly relevant to the problem of fault-tolerant sensor and actuator selection. With increasing focus on that particular problem, we consider first some significant results on sensor and actuator selection for distributed parameter systems (DPS) modeled by partial differential equations (PDE), then algorithms specialized for structural dynamics modeled by ordinary differential equations (ODE), and finally criteria for sensor and actuator selection considering possible failures.

2.1 Sensor and Actuator Selection for DPS Control

Control of structural dynamics is properly posed within the theoretical framework of distributed-parameter systems and their infinite-dimensional spaces. It is widely recognized that the spatial location of inputs and outputs, and possibly the temporal scheduling as well, are important design issues. Kubrusly and Malebranche [19] recently conducted a survey of over 50 papers on this topic. We will note some general observations from their survey, and examine the results of one particularly complete paper.

While applications were directed toward any of three objectives - system identification, state estimation, and optimal control - the primary theoretical distinction noted by Kubrusly and Malebranche was the decision of where to shift analysis from the infinite-dimensional space (L_2 or l_2) to a finite-dimensional space (R^N) by the 'N-modal' Fourier expansion of eigenfunctions [19]. For example, some authors carried out the approximating truncation first, then posed sensor or actuator criteria, while others defined infinite-dimensional controllers and approximated at a later stage. None of the works were concerned with practical issues such as computational costs or efficient algorithms, and none of them included examples involving more than four sensors and/or actuators. However, they develop a theoretical underpinning which offers some assurance that the problem has no fundamental pitfalls.

One of the latest and most complete papers in the survey was by Omatu and Seinfeld [20] and we will cite their main results. For a linear DPS of the form

$$dx(t) = A(t)x(t)dt + B(t,\xi_a)u(t)dt + dw(t) \qquad 2\text{-}1$$

state x evolves in a Hilbert space, H, with inner product $< , >$ and outer product, $\circ$. We will not precisely account for the properties of H and operators in H because we only wish

to comment here on some results, but the interested reader might consult [21] for a concise summary of such analysis. A(t) is a bounded differential operator, from and into H, while bounded operator $B(t,\xi_a)$ is parameterized by actuator (controller) spatial locations, ξ_a, and is a mapping from the finite dimension of control, $\mathbf{u}(t) \in R^m$, to H. Disturbance $\mathbf{w}(t)$ is an infinite-dimensional Wiener process with intensity

$$E\left[\mathbf{w}(t_1)\circ\mathbf{w}(t_2)\right] = \int_{t_0}^{\min t_1,t_2} Q_w(\tau)\, d\tau \,. \qquad 2\text{-}2$$

A finite number of measurements, $\mathbf{y}(t) \in R^l$, are taken from $\mathbf{x}(t)$ with $C(t,\xi_s)$, an operator from H to R^l, parameterized by sensor spatial locations, ξ_s,

$$d\mathbf{y}(t) = C(t,\xi_s)\mathbf{x}(t)dt + d\mathbf{v}(t) \qquad 2\text{-}3$$

and the measurements are corrupted by the finite-dimensioned Wiener process $\mathbf{v}(t)$ with intensity

$$E\left[\mathbf{v}(t_1)\circ\mathbf{v}(t_2)\right] = \int_{t_0}^{\min t_1,t_2} R_v(\tau)\, d\tau \,. \qquad 2\text{-}4$$

Omatu and Seinfeld [20] consider optimal control of the stochastic DPS, represented by equations 2-1 and 2-3, with respect to the LQG cost functional

$$J(\mathbf{u},\xi_a,\xi_s) = E\left\langle \mathbf{x}(t_f),Q_x\mathbf{x}(t_f)\right\rangle + \int_{t_0}^{t_f}\left\langle \mathbf{x}(t),Q_x(t)\mathbf{x}(t)\right\rangle + \left\langle \mathbf{u}(t),R_u(t)\mathbf{u}(t)\right\rangle dt \,. \qquad 2\text{-}5$$

The optimal control, in an extension of finite-dimensional results, is given by full-state feedback of the estimated state which evolves in H. The optimal feedback and observer gains involve infinite-dimensional Riccati operators, $P_c(t,\xi_a)$ and $P_o(t,\xi_s)$, respectively. Note that while the dimensions of $\mathbf{u}(t)$ and $\mathbf{y}(t)$ are both finite, the resulting model-based compensator is not.

The main results of reference [20] are existence theorems of an optimal actuator location, ξ_a^*, and an optimal sensor location, ξ_s^*, along with necessary and sufficient conditions for the optimal locations. Proof of existence centers on showing operators

$P_c(t,\xi_a)$ and $P_o(t,\xi_s)$ are continuous in ξ_a and ξ_s, respectively, with the further assumptions that ξ_a and ξ_s are compact and $B(t,\xi_a)$ and $C(t,\xi_s)$ are continuous with respect to the sensor or actuator locations. Sufficient conditions for optimal locations are

$$C^T(t,\xi_s^*)R_v^{-1}C(t,\xi_s^*) \geq C^T(t,\xi_s)R_v^{-1}C(t,\xi_s), \quad \forall\ \xi_s \tag{2-6}$$

$$B(t,\xi_a^*)R_u^{-1}B^T(t,\xi_a^*) \geq B(t,\xi_a)R_u^{-1}B^T(t,\xi_a), \quad \forall\ \xi_a \tag{2-7}$$

where the operator $(\ )^T$ represents the adjoint. The necessary condition takes the form

$$\alpha(t_0,t_f,\xi_a,\xi_s) \geq 0, \quad \forall\ \xi_a \neq \xi_a^*, \ \forall\ \xi_s \neq \xi_s^* \tag{2-8}$$

where

$$\alpha(t_0,t_f) = \int_{t_0}^{t_f} \mathrm{tr}\left\{ R_u(0)F_1(t,t_0,\xi_a^*)\left[B(t,\xi_a^*)R_u^{-1}B^T(t,\xi_a^*) - B(t,\xi_a)R_u^{-1}B^T(t,\xi_a)\right] F_1^T(t,t_0,\xi_a) \right\} dt +$$

$$\int_{t_0}^{t_f}\int_{t_0}^{t} \mathrm{tr}\left\{ F_1(t,\tau,\xi_a^*)\left[B(t,\xi_a^*)R_u^{-1}B^T(t,\xi_a^*) - B(t,\xi_c)R_u^{-1}B^T(t,\xi_a)\right] F_2(t,\tau,\xi_a,\xi_s) \right\} d\tau\, dt +$$

$$\int_{t_0}^{t_f}\int_{t_0}^{t} \mathrm{tr}\left\{ G_2(t,\tau,\xi_a,\xi_s)\left[C^T(t,\xi_s^*)R_v^{-1}C(t,\xi_s^*) - C^T(t,\xi_s)R_v^{-1}C(t,\xi_s)\right] G_1(t,\tau,\xi_s^*) \right\} d\tau\, dt \ . \tag{2-9}$$

Operators F_1, F_2, G_1, and G_2 are, in effect, weighting functions for the sufficient conditions, They involve the Riccati operators at optimal locations ξ^* as well as other locations ξ, the evolution operators of the filter and regulator dynamics, and the control effectiveness and measurement sensitivity maps, $B(t,\xi_a)$ and $C(t,\xi_s)$. For example,

$$F_1(t,\tau,\underset{\sim}{\xi}_a) = \Psi_c^T(t,\tau)P_c(t,\underset{\sim}{\xi}_a) \qquad 2\text{-}10$$

$$F_2(t,\tau,\underset{\sim}{\xi}_a,\underset{\sim}{\xi}_s) = P_c(t,\underset{\sim}{\xi}_a)\Psi_c(t,\tau)P_o(\tau,\underset{\sim}{\xi}_s)C^T(\tau,\underset{\sim}{\xi}_s)R_v^{-1}(\tau)C(\tau,\underset{\sim}{\xi}_s)P_o(\tau,\underset{\sim}{\xi}_s) \qquad 2\text{-}11$$

where Ψ_c is the evolution (state transition for ODE systems) operator of the LQ regulator at the optimal actuator locations, $\underset{\sim}{\xi}_a$*. Weighting functions G_1 and G_2 have a dual structure.

Further results include corollaries of the existence, necessary, and sufficient conditions for finite-dimensional (N-modal) approximations of the DPS system represented by equations 2-1 and 2-3. An example using the parabolic heat equation in one dimension places two sensors and actuators by checking the sufficient conditions, but neither a general algorithm nor an application of the necessary condition is suggested. The results are important because they are derived with respect to a measure of closed-loop performance for linear distributed parameter systems subject to stochastic disturbances and measurement noise, and are extended to a finite-dimensional approximation of the system. Perhaps the only significant theoretical point to add might be to show the sensor and actuator locations that are optimal with respect to the N-modal approximation converge to the optimal locations of the DPS. From an engineering point of view, however, the ability to compute optimal locations for large-scale systems is also necessary, so we turn next to research which has focused on algorithms with specific applications to structural systems.

2.2 Algorithms for Structural Dynamics

While for theoretical purposes, the step to specialization for structural dynamics is just focusing on a particular class of hyperbolic self-adjoint operators, an important engineering detail is the fact that, for practical purposes, dynamic analysis of flight structures requires finite-element methods. Only the simplest spatial domains are tractable with a direct application of assumed modes (or Fourier expansion), and for most structures, identification of independent spatial variables, $\underset{\sim}{\xi}$, with suitable continuity properties may be unacceptably difficult. A common pattern in the literature for sensor and actuator selection is to a suggest a criterion, several of which we will review in Chapter 3, and plot it over a simple structural model. Few researchers have stepped up to the difficulties imposed by complex spatial domains and the required finite-element approximations on a large scale. We review next results from two who have.

One approach to a discretized spatial domain is to consider only the finite number of nodes, or combinations of nodes, for sensor and actuator locations. In principle, there is no loss of generality, since the finite-element mesh can be selected as small as desired, and since devices with a distributed effect are reflected to the nodes as external point forces. This casts the problem, formally, as one of integer programming, but even for linear problems these are known to be intrinsically difficult. Since the effect of adding sensors or actuators to a structure is generally monotonic- that is, more is better from a controllability or observability point of view- sequential elimination (or addition) is a viable strategy for an algorithm. Hegg [22] and DeLorenzo [23] have developed algorithms for large-scale structures with sequential elimination, and implemented them as shown in Figure 2-1.

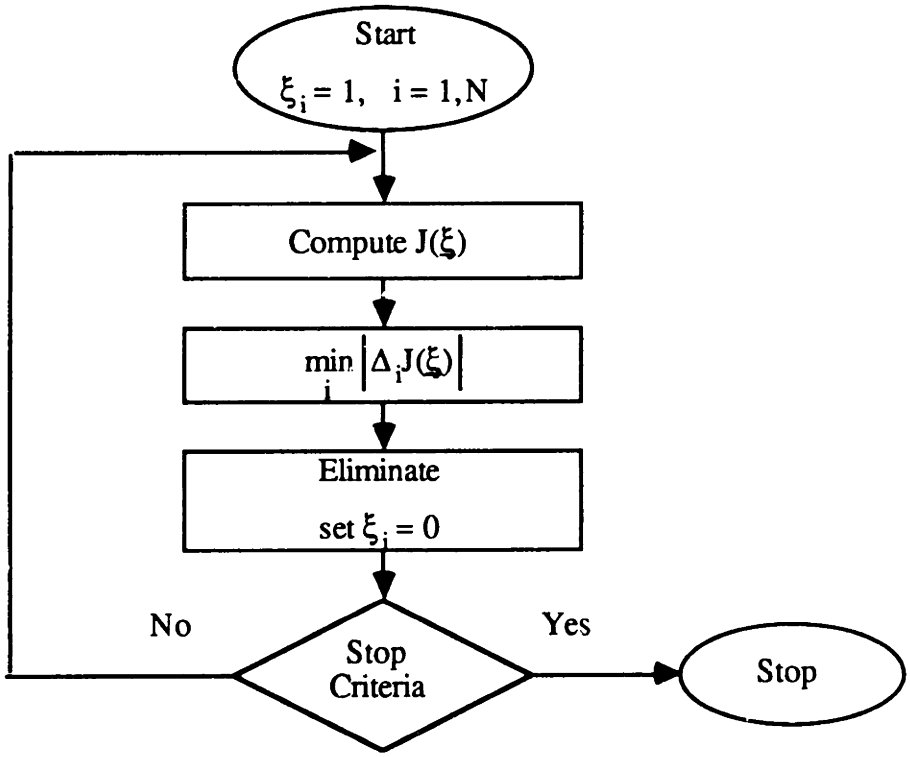


Figure 2-1. Sequential Elimination Algorithm.

In this case, components ξ_i are Boolean variables indicating presence or absence of a sensor or actuator on one of N potential sites. The algorithms start with a device in all possible sites and then identify the site which will cause the least reduction (ΔJ) in a performance index, $J(\xi)$. DeLorenzo [23] uses the LQ performance index, equation 2-5 for

for finite-dimensional systems, and selects a site for elimination by computing, for actuators,

$$\Delta_i J(\xi) = \left(\frac{\partial J}{\partial \mathbf{u}_i}\right)^T \mathbf{u}_i \qquad 2\text{-}12$$

exploiting the 'Component Cost Analysis' of Skelton [24] for the gradient of the LQ cost in direction of control $\mathbf{u}_i$. Hegg [22], on the other hand, considers columns of the control influence matrix, B, and their ability to span the sub-space of the modal influence vectors of a disturbance. We will revisit both of these particular performance indices in more detail in Chapter 3. For now, we note both Hegg and DeLorenzo demonstrated their algorithms on a version of the 'CSDL Model No. 2' from the ACOSS problem: Hegg considering as many as 485 potential sites and DeLorenzo about 10% of that number. Since we have Hegg's results in some detail, we use them to compare with the results of this thesis in Chapter 5. Neither Hegg nor DeLorenzo claimed their algorithm produced an optimal result with respect to any criterion, although each appeared to be a practical and useful heuristic approach. We should also note DeLorenzo included a modification for the case of noisy actuators, in which elimination of a site might actually improve (decrease) the LQ cost. Both developed corresponding criteria for sensor selections.

Some structures may present a particular axis or direction where a continuous spatial variable may be introduced and a hybrid approach, with numerical estimates of mode shapes at discrete nodes smoothed by a continuous interpolation, may work. An example of this by Johnson, et al. [25] considers placement of control surfaces for gust load alleviation along the wing of a transport aircraft. Cubic splines interpolated finite-element derived mode shapes so that modal influences of two pairs of ailerons were approximated as continuous functions of their locations in one dimension along the wings. The optimization criteria was the minimum LQG functional weighted for wing root stresses and control surface deflections of the ailerons, the locations of which were independent variables. Geometric simplifications such as this are not general, however, and likely to be particularly uncommon for flexible spacecraft.

If we now consider the possibility of failures, the combinatorial nature of failure enumeration makes sequential approaches, particularly elimination, expensive. To the

author's knowledge, there are no examples of sensor and actuator selection algorithms for large-scale structures which account for failures. There are, however, some useful results for measures of performance including failure effects, and we take these up next.

2.3 Criteria for Fault-Tolerant Sensor and Actuator Selection

Given some criteria for sensor or actuator placement, say $J(\xi)$, where ξ represents spatial locations, we may introduce an additional parameter, f_i, which represents a particular failure mode, and $p_i(t)$ to represent the probability the configuration, ξ, is in failure state f_i at time t. We may then compute a failure-weighted performance index with the expected value,

$$E[J] = \sum_i J(\xi, f_i)\, p_i(t) \qquad 2\text{-}13$$

which is defined at a particular time, but could be averaged (integrated) over a mission time interval. This approach, by Vander Velde and Carignan [26], also develops a particular performance index, J, based upon an optimal open-loop control or observation problem. Another criterion, the LQ optimal cost, was employed in a similar manner in reference [27]. We will examine both measures in Chapter 3 for algorithm suitability, but note here the application approaches and results.

With the expected value of the optimal LQ regulator as a criterion, Montgomery and Vander Velde [27] compared two configurations of five actuators each on a grid structure for time durations up to twice the mean-time-to-failure (MTF) of an individual component, i.e., $0 \le t/MTF \le 2$. One configuration was clearly superior if no failures were possible, e.g., $t = 0$, but for $t/MTF \ge .5$, the other one was preferable. Vander Velde and Carignan [26] fixed $t = MTF$ and considered placement of up to four actuators on 11 possible sites on a free-free beam, seeking the best locations via exhaustive search. Again, the optimal location was shown to change when the possibility of failures was included. For no failures, sites were favored that might only control a few modes, strongly influencing those they did, while if failures were admitted, sites that interacted to some degree with all modes were favored. Even for the limited number of sites and actuators considered, there are 330 ways to place four devices in 11 potential locations, and for each placement there are 12

failure modes (including all-fail) so the performance index must be computed 3,630 times in the search. Exhaustive enumeration quickly becomes impractical for larger problems.

In the next chapter we examine in more detail criteria for sensor and actuator placement, select a particular class, and formulate an optimization problem for algorithms. An important motivation is computational economy for applications to large-scale structures.

Chapter 3. Problem Formulation

From the rather broad goals of sensor and actuator selection for fault-tolerant active compensation, this chapter develops an optimization problem which strikes a balance between computational economy and precise representation of the actual objectives. We review structural dynamics and probabilistic failure modeling. Optimization requires an objective function; so several candidate classes are reviewed and a particular one, based upon open-loop system controllability and observability, is selected and motivated. Finally, the sensor/actuator selection process is posed as a non-linear integer programming problem with a 'relaxation' as a first step in its solution, which is reserved for the next chapter.

3.1 Failure modeling

Without considering the possibilities of component failures, most of the work reviewed so far may be thought of as a solution to the beginning-of-life (BOL) problem. That is, for some independent parameterization of the sensor/actuator suite, ξ, and the deterministic performance index, $J(\xi)$, the optimization of J over ξ is a useful result only for mission durations short enough to preclude planning for component failure, or BOL for conventional long-duration space systems. In order to account for failures, it is useful to introduce the additional parameter, $\mathbf{f}$, to represent distinct failure modes, and Π, the probability law of $\mathbf{f}$. The evolution of the probability law is often modeled with Markov chains. The now-stochastic $J(\xi,\mathbf{f})$ may be formed to express the goal of good performance in the presence of failures. Indeed, the definition of failure modes and their probability law can be a complex undertaking for even a relatively simple control system [28, 29]. At the top system level, reliability modeling includes FDI, reconfiguration, grouping of simple failure effects into 'modes', which are possibly distinguished by the history or order of failures and resulting actions, all in order to make a probabilistic estimate of system performance. For this application, however, we seek a number of simplifications.

We assume the problem of optimizing controllability and observability of a structural system may be done independently of the particular closed-loop control, FDI, and reconfiguration systems. The failure modes, $\mathbf{f}$, are restricted to the sensor/actuator

suite alone, and the evolution of its probability law is based on the reliability of the components. Furthermore, the nature of the failure modes for each component will be idealized as bi-modal, either fully operational or fully failed and turned off. A system-level failure mode is defined to be a unique list, **f**, of the operational status of sensors or actuators- 0 for failure, 1 for operational. We further define F to represent a particular set of failure modes, according to some criterion such as probability of occurrence. A bi-modal assumption is supported by considering the electro-mechanical hardware commonly envisioned as sensors and actuators for structural control. Unlike aerodynamic control surfaces, their influence upon the structural dynamics may be entirely eliminated (save for mass) by command and switching upon signal from a FDI system. Note that these simplifications on **f** do not restrict the modeling of its probability law, Π. In particular, there is no requirement that the components' individual reliabilities be independent.

Even with simplifying assumptions, the large number of components typical for a structural control system results in the possibility of even larger numbers of failure modes. In the example of Table 1-1, if 14 components are required at BOL to ensure 10 continue to function over the mission duration, there are 1471 combinations of up to 4 failures. This combinatorial nature of the problem adds significant weight to computational economy for a performance index and its parameterization, $J(\xi,\mathbf{f})$.

Several basic ideas useful for fault-tolerant systems are now defined [30-32]. For a simple component that is either failed or operating, the time to failure, T_f, is a random variable with a probability distribution function, $F_{Tf}(t) = Pr(T_f \le t)$. Another commonly used event is that failure has not occurred at time t, and its probability is the reliability of the component

$$r(t) = P_r (t < T_f) = 1 - F_{Tf} (t). \qquad 3\text{-}1$$

The failure time has a probability distribution (mass) function,

$$p_{Tf}(t) = \frac{d}{dt} F_{Tf}(t) = - \frac{d}{dt} r(t). \qquad 3\text{-}2$$

The expected value of the failure time is the mean time between failures (MTBF). For systems that are not maintained, such as spacecraft, a better term is mean time to failure (MTTF), which will be used here and is computed

$$\text{MTTF} \triangleq E[T_f] = \int_{-\infty}^{\infty} t\, p_{Tf}(t)\, dt . \qquad 3\text{-}3$$

Expressing the density in terms of the reliability, and integrating by parts, the MTTF may be calculated:

$$\text{MTTF} = \lim_{t \to \infty} \left[- t r(t) + \int_0^t r(\tau)\, d\tau \right] \qquad 3\text{-}4$$

which shows a finite MTTF is not defined for some reliability histories. Another important event to characterize is the time of failure conditioned on the functioning of the component at some earlier time, $0 < t_0 \leq t$. Applying the definition of conditional probability, this event may be related to the unconditional failure time:

$$P_r(T_f \leq t \mid T_f > t_0) = \frac{P_r(t_0 < T_f \leq t)}{P_r(T_f > t_0)} = \frac{F_{Tf}(t) - F_{Tf}(t_0)}{r(t_0)} . \qquad 3\text{-}5$$

Differentiation results in the conditional distribution, which may be expressed in terms of the (unconditional) reliability:

$$p(t \mid t_0) = \frac{\frac{d}{dt} F(t)}{r(t_0)} = \frac{-\frac{d}{dt} r(t)}{r(t_0)} . \qquad 3\text{-}6$$

The hazard, or failure, rate is the value of this conditional distribution for the special case $t_0 = t$, $\lambda_f(t) = p_{Tf}(t \mid t)$; and can be experimentally estimated in practice. It is often the starting point for reliability analysis, but does not itself have the properties of a probability density function. The reliability may be calculated from the failure rate by solving the differential equation

$$\lambda_f(t) = \frac{-\frac{d}{dt} r(t)}{r(t)} . \qquad 3\text{-}7$$

Subject to the initial condition $r(0) = 1$, the solution is

$$r(t) = \exp\left[- \int_0^t \lambda_f(\tau)\, d\tau \right] . \qquad 3\text{-}8$$

A constant failure rate is a commonly used simplification, which is often acceptable after a 'burn-in' or 'infant mortality' period and before wear-out phenomena dominate. For $\lambda_f(t) = \lambda_f$, reliability has the explicit form, $r(t) = e^{-\lambda_f t}$, and the MTTF definite integral (eqn. 3-8) evaluates to $MTTF = 1/\lambda_f$.

Included among the parameters to select for an index incorporating failures is the mission duration and the manner it is reflected in $J(\xi,\mathbf{f})$. Common practice is to specify system-level requirements at end of life (EOL) for those functions that generally degrade over the mission duration. Electrical power is a particularly good example, since both battery charging depth and solar-cell output fall steadily with time and use. Implied in the EOL-type specification is the assumption that the particular system will deliver performance up to or better than specification at all times prior to EOL. Another specification approach is to average performance over the life, or sections of the life, of the mission. This is done when there may be large fluctuations in performance during the mission, but not monotonic degradation. Examples include time between momentum desaturation maneuvers or high-level mission objectives, such as coverage of a multi-spacecraft network. With the assumptions that maintenance is not available during the mission and control system component failures are permanent, the EOL approach is adopted as most appropriate for sensors and actuators.

3.2 Measures of Controllability and Observability

The control authority of actuators and measurement sensitivity of sensors may be cast formally in terms of system controllability and observability. Optimization requires some scalar index, and in this section we develop and motivate a particular class that reflects the properties of the open-loop fixed-final-state control problem and an observation dual.

3.2.1 Gramians; Control and Measurement Interpretations

For a vector-valued function, $\mathbf{g}(t) = [g_1(t), \cdots, g_n(t)]^T$ to be linearly independent on the interval $[t_0, t_1]$, Gram's criterion [33], [34], is that the determinant of a matrix of inner products not vanish:

$$|W(t_1,t_0)| \neq 0, \quad \text{where} \quad W_{ij}(t_1,t_0) = \int_{t_0}^{t_1} g_i^H(\tau) g_j(\tau)\, d\tau\,. \qquad 3\text{-}9$$

The Gram matrix is symmetric and non-negative definite, so has real non-negative eigenvalues and orthogonal eigenvectors. In deterministic control applications, if we are interested in transferring the initial state, $\mathbf{x}(t_0)$, of system (A,B) to any terminal state, $\mathbf{x}(t_1)$, in finite time with a control of finite energy, then this is possible if and only if [35] the rows of the convolution kernel, $[e^{A(t_1-t)}B]$, are linearly independent on the interval $[t_0,t_1]$. Similarly, if we are interested in determining the initial state, $\mathbf{x}(t_0)$, of a system (A,C) from (noise-free) measurements $\mathbf{y}(t) = C\,\mathbf{x}(t)$ on a finite interval, $[t_0,t_1]$, then this is possible if and only if the columns of the free response, $[C\,e^{At}]$, are linearly independent over the observation interval. Furthermore, we may be interested in scaling or weighting the control and output signals to reflect particular objectives. We introduce positive definite weightings R_u and R_y for this purpose. In order to assess the linear dependence of the rows of $[e^{A(t_1-t)}B]$ or columns of $[C\,e^{At}]$, we may compute Gram matrices with all combinations of inner products:

$$W_c(t_1,t_0) = \int_{t_0}^{t_1} e^{A(t_1-\tau)} B R_u^{-1} B^T e^{A^T(t_1-\tau)} d\tau \;=\; \int_0^{t_1-t_0} e^{A\tau} B R_u^{-1} B^T e^{A^T\tau} d\tau\,, \qquad 3\text{-}10$$

and

$$W_o(t_1,t_0) = \int_0^{t_1-t_0} e^{A^T\tau} C^T R_y^{-1} C e^{A\tau} d\tau\,. \qquad 3\text{-}11$$

The second form of the Gram matrix for control results from a change of variables of integration. While the term 'Gramian' has as an older definition [33] the determinant of the Gram matrix, the control literature commonly applies the name to the matrix itself, and that convention is adopted here. The 'Controllability Gramian', then, is represented by equation 3-10 and the 'Observability Gramian' by 3-11. Weightings R_u and R_y are non-standard, but the effect is equivalent to including them directly in B or C with the standard Gramian forms: $B = BR_u^{-1/2}$ and $C = R_y^{-1/2}C$. Differentiation shows the Gramians are solutions to the Lyapunov equations with $t = t_1 - t_0$,

$$\frac{d}{dt}W_c(t) = AW_c(t) + W_c(t)A^T + BR_u^{-1}B^T, \quad W_c(0) = 0, \qquad 3\text{-}12$$

$$\frac{d}{dt} W_o(t) = A^T W_o(t) + W_o(t) A + C^T R_y^{-1} C, \quad W_o(0) = 0, \tag{3-13}$$

and for strictly stable systems, $\mathrm{Re}(\lambda_i(A)) < 0 \; \forall$ i, have bounded solutions in the limit as $t \to \infty$. The controllability of (A,B) and observability of (A,C) may be defined in terms of a full rank condition on the appropriate Gramian, and is equivalent to the more common use for time-invariant systems of the controllability matrix, $[A\,AB \cdots A^{n-1}B]$. Early work [35] referred to the (control) Gramian as the controllability matrix of the first kind, while their 'second kind' is the now commonly cited form. More useful to this work, however, is the manner in which the Gramians enter into open-loop optimal control and observation.

For the objective of sending $x(0) = 0$ to $x(t) = x_f$, the min-norm (L_2) control that accomplishes this is

$$u(\tau) = R_u^{-1} B^T e^{A^T(t-\tau)} W_c^{-1}(t)\, x_f \tag{3-14}$$

and the norm of the control is

$$\begin{aligned} \| u \|_R^2 &\triangleq \int_0^t u^T(\tau) R_u u(\tau)\, d\tau \\ &= \int_0^t [x_f^T W_c^{-1}(t) e^{A(t-\tau)} B R_u^{-1}] R_u [R_u^{-1} B^T e^{A^T(t-\tau)} W_c^{-1}(t)\, x_f]\, d\tau \\ &= x_f^T W_c^{-1}(t)\, x_f. \end{aligned} \tag{3-15}$$

In a dual manner, given observations of the free state response, $y(\tau) = Ce^{A\tau} x(0)$, the initial condition, $x(0) = x_0$ is observed by

$$x(t) = W_o^{-1}(t) \int_0^t e^{A^T\tau} C^T R_y^{-1} y(\tau)\, d\tau \tag{3-16}$$

and the output signal norm is

$$\| y \|_{R^{-1}}^2 = \int_0^t [x_0^T e^{A^T\tau} C^T] R_y^{-1} [C e^{A\tau} x_0] d\tau = x_0^T W_o(t)\, x_0. \tag{3-17}$$

The quadratic form in the norm expressions have a geometric interpretation as the reachable and observable sets, and the eigenstructure of the Gramians displays the properties of those sets. Let λ_i be a (real) eigenvalue of W, and $\mathbf{v}_i$ the corresponding eigenvector normalized to unit length, $\mathbf{v}_i^T\mathbf{v}_i = 1$. Note the eigenvectors are orthogonal by the symmetry of W. For a final state in the direction of an eigenvector for the control problem, $\mathbf{x}_f = \alpha \mathbf{v}_{cj}$, or for an initial state aligned with an observability Gramian eigenvector, $\mathbf{x}_0 = \alpha \mathbf{v}_{oj}$, the minimum control norm and signal norms are, respectively,

$$\| \mathbf{u} \|^2 = \mathbf{x}_f^T W_c^{-1}(t_f) \mathbf{x}_f = \alpha \mathbf{v}_{cj}^T W_c^{-1}(t_f) \alpha \mathbf{v}_{cj} = \frac{\alpha^2}{\lambda_{cj}} \qquad 3\text{-}18$$

$$\| \mathbf{y} \|^2 = \mathbf{x}_0^T W_o(t_f) \mathbf{x}_0 = \alpha \mathbf{v}_{oj}^T W_o(t_f) \alpha \mathbf{v}_{oj} = \alpha^2 \lambda_{oj}. \qquad 3\text{-}19$$

Therefore, the set reachable by a unit norm control is the interior of a hyperellipsoid with major axes of lengths $\lambda_c^{1/2}$ pointing in the directions of the eigenvectors of W_c. The set detectable with a signal output of at least unit norm is also a hyperellipsoid with axes aligned to the eigenvectors of the observability Gramian, but the lengths are $\lambda_o^{-1/2}$ and the set includes the exterior of the ellipsoid. The minimum eigenvalues and eigenvectors represent the most expensive direction to control and the least sensitive direction to detect for the respective cases. For 'good' controllability and observability properties, we desire larger eigenvalues of both Gramians distributed in the important directions of state space. This qualitative objective is made more precise later in sections 3.2.2 and 3.3.2.

It may be the case that the control objective is a fixed final point in the output space, say $\mathbf{y}_f$. The min-norm control driving $\mathbf{x}(0)$ to $\mathbf{x}(t)$ such that $\mathbf{y}(t) = \mathbf{y}_f = C\mathbf{x}(t)$ is:

$$\mathbf{u}(\tau) = R_u^{-1} B^T e^{A^T(t-\tau)} C^T [C W_c(t) C^T]^{-1} \mathbf{y}_f. \qquad 3\text{-}20$$

The (weighted) control norm is given by

$$\| \mathbf{u} \|_R^2 = \mathbf{y}_f^T [C W_c(t) C^T]^{-1} \mathbf{y}_f. \qquad 3\text{-}21$$

An output reachable set may be defined and studied with the eigenstructure of the Gramian CW_cC^T, which may be of considerably lower order than the full-state controllability

Gramian, particularly for structural systems. There is a dual to this case, and it may be physically motivated by considering the problem of determining the input which is known to be an impulse delivered at t=0, $\mathbf{u}(t) = \mathbf{u}_o\delta(t)$, from the output response of a system initially at rest. The initial state is then $\mathbf{x}(0) = B\mathbf{u}_o$, and the problem is a variation of the initial state observation case. The minimum-length initial input, calculated by the free response output is

$$\mathbf{u}_o = \left[B^T W_o(t) B\right]^{-1} B^T \int_0^t e^{A^T \tau} C^T R_y^{-1} \mathbf{y}(\tau)\, d\tau \qquad 3\text{-}22$$

and the (weighted) norm output is

$$\| \mathbf{y} \|^2_{R^{-1}} = \mathbf{u}_o^T \left[B^T W_o(t) B\right] \mathbf{u}_o . \qquad 3\text{-}23$$

The geometry of the 'input observable' set may be described by the Gramian $B^T W_o(t) B$.

A change in the state basis affects the structure of the Gramian. For the non-singular transformation $\mathbf{x} = T\mathbf{x}$, the system (A,B,C) for $\mathbf{x}$ is transformed to ($T^{-1}AT, T^{-1}B$, CT) for x. The Gramians are transformed:

$$W_c(t) = T^{-1} W_c(t) T^{-T} \qquad 3\text{-}24$$

$$W_o(t) = T^T W_o(t) T \qquad 3\text{-}25$$

which are congruent transformations. While the signs of the eigenvalues are preserved, their size and the directions of the eigenvectors are not. Note the output controllability Gramian (CW_cC^T) and input observability Gramian (B^TW_oB) are independent of the state basis, since they describe the linear dependences of input-output properties. This may be verified algebraically:

$$CW_cC^T = CT\left(T^{-1} W T^{-T}\right) T^T C = CW_cC^T \qquad 3\text{-}26$$

and likewise for B^TW_oB.

The explicit dependence on a terminal time can be settled by choosing an infinite horizon. In this case, the Gramians may be computed as the solutions to the algebraic Lyapunov equations:

$$0 = AW_c(t_\infty) + W_c(t_\infty)A^T + BR_u^{-1}B^T \qquad 3\text{-}27$$

$$0 = A^T W_o(t_\infty) + W_o(t_\infty)A + C^T R_y C \ . \qquad 3\text{-}28$$

For lightly damped structural systems, this choice has two significant shortcomings. They are best illustrated by an example of a lightly coupled, lightly damped two-mode system, such as the pair of identical pendula in Figure 3-1. There is a single input, a torque about the pivot of pendulum 1, and the coupling is represented by the spring, k. Viscous drag, b, at the pivots damps the system.

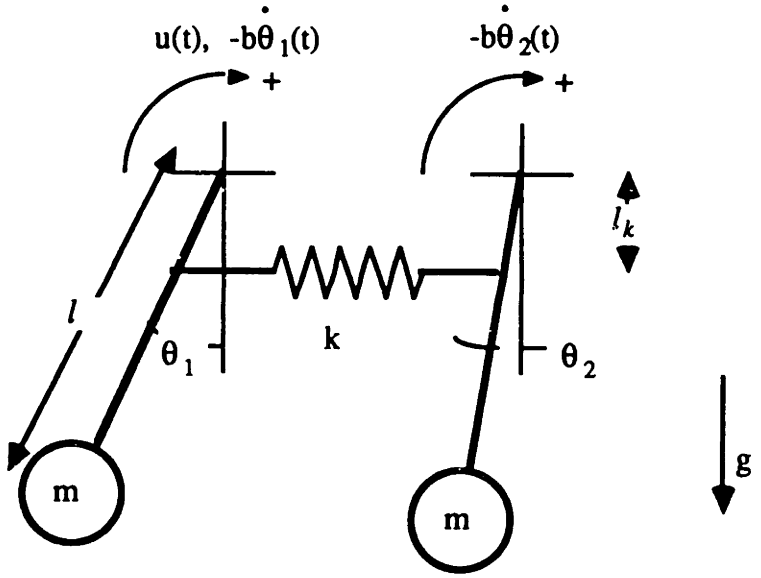


Figure 3-1. Coupled Pendula

Selecting as unit displacement the static deflection of the linearized equations from a unit torque input, $(mgl + kl_k^2)$, and as unit time the inverse of the undamped natural period, $[g/l + (k/m)(l_k/l)^2]^{1/2}$, the non-dimensional equations of motion may be written

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \ddot{\underline{\theta}} + 2 \begin{bmatrix} \zeta & 0 \\ 0 & \zeta \end{bmatrix} \dot{\underline{\theta}} + \begin{bmatrix} 1 & -\varepsilon \\ -\varepsilon & 1 \end{bmatrix} \underline{\theta} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathbf{u} \, . \qquad 3\text{-}29$$

The normal mode shapes provide a basis for decoupling the equations and are the solution to the eigenvalue problem

$$\omega_i^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{\Phi}_i = \begin{bmatrix} 1 & -\varepsilon \\ -\varepsilon & 1 \end{bmatrix} \underline{\Phi}_i \qquad 3\text{-}30$$

with eigenvalues

$$\omega_1 = \sqrt{1+\varepsilon}\,, \qquad \omega_2 = \sqrt{1-\varepsilon} \qquad 3\text{-}31$$

and corresponding normal mode shapes (eigenvectors)

$$\Phi = \left[\underline{\Phi}_1 \;\; \underline{\Phi}_2\right] = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}. \qquad 3\text{-}32$$

Note the effect of the coupling, ε, is to split the identical natural modes of the independent oscillators into two distinct system modes. With the change of basis, $\underline{\theta} = \Phi\, \underline{\eta}$, the modal equations are

$$\ddot{\underline{\eta}} + 2 \begin{bmatrix} \zeta & 0 \\ 0 & \zeta \end{bmatrix} \dot{\underline{\eta}} + \begin{bmatrix} 1+\varepsilon & 0 \\ 0 & 1-\varepsilon \end{bmatrix} \underline{\eta} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} u \qquad 3\text{-}33$$

which have as a first-order representation for $\mathbf{x}^T = [\, \eta_1 \; \dot{\eta}_1 \; \eta_2 \; \dot{\eta}_2 \,]$

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -(1+\varepsilon) & -2\zeta & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -(1-\varepsilon) & -2\zeta \end{bmatrix} \mathbf{x} + \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} u \;. \qquad 3\text{-}34$$

In this basis, the infinite-time controllability Gramian may be computed explicitly by solving the algebraic Lyapunov equation (3-27), $0 = \mathbf{A}\, \mathbf{W_c} + \mathbf{W_c}\, \mathbf{A}^T + \mathbf{b}\, \mathbf{b}^T$:

$$W_c = \begin{bmatrix} \frac{1}{8\zeta(1+\varepsilon)} & 0 & \frac{\zeta}{2(4\zeta^2+\varepsilon^2)} & \frac{\varepsilon}{4(4\zeta^2+\varepsilon^2)} \\ 0 & \frac{1}{8\zeta} & \frac{-\varepsilon}{4(4\zeta^2+\varepsilon^2)} & \frac{\zeta}{2(4\zeta^2+\varepsilon^2)} \\ \frac{\zeta}{2(4\zeta^2+\varepsilon^2)} & \frac{-\varepsilon}{4(4\zeta^2+\varepsilon^2)} & \frac{1}{8\zeta(1-\varepsilon)} & 0 \\ \frac{\varepsilon}{4(4\zeta^2+\varepsilon^2)} & \frac{\zeta}{2(4\zeta^2+\varepsilon^2)} & 0 & \frac{1}{8\zeta} \end{bmatrix} \qquad 3\text{-}35$$

We may see now the first difficulty with the infinite-time horizon: the Gramian is inversely proportional to the damping ratio and all non-zero terms grow without bound as $\zeta \to 0$. For structural dynamics, viscous damping is the least accurately known parameter of the linear model, and it is desirable to not allow damping such a direct influence on the results. The second problem has to do with the shape of the reachable set.

Taking as an 'output' the physical displacements, θ, they may be written as a combination of the modal states:

$$\underline{\theta} = C\,\mathbf{x}, \text{ with} \quad C = [\,\phi_1\ 0\ \phi_2\ 0\,] = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \end{bmatrix}. \qquad 3\text{-}36$$

The control Gramian for $\underline{\theta}$ is then

$$W_{c\theta} = C\,W_{cx}\,C^T = \begin{bmatrix} \frac{1 + (\frac{\varepsilon}{2\zeta})^2 (1-4\zeta^2)/2}{4\zeta(1-\varepsilon^2)\left(1+(\frac{\varepsilon}{2\zeta})^2\right)} & \frac{\varepsilon}{8\zeta(1-\varepsilon^2)} \\ \frac{\varepsilon}{8\zeta(1-\varepsilon^2)} & \frac{\varepsilon^2\left(1+(\frac{1}{2\zeta})^2\right)}{8\zeta(1-\varepsilon^2)\left(1+(\frac{\varepsilon}{2\zeta})^2\right)} \end{bmatrix}. \qquad 3\text{-}37$$

Some limiting cases are of interest. As the coupling vanishes, all terms go to zero except the upper left, corresponding to θ_1, which approaches $1/(4\zeta)$. The θ_1 diagonal term (upper left) is always larger than the θ_2 term (lower right), but the proportional difference tends to zero as the damping decreases for any finite coupling:

$$\lim_{\zeta \to 0} \frac{W[1,1] - W[2,2]}{W[1,1]} = \frac{\zeta^2}{\zeta^2 - (\varepsilon/2)^2} = \begin{cases} 0, & \varepsilon > 0 \\ 1, & \varepsilon = 0 \end{cases} . \qquad 3\text{-}38$$

This says that for any coupling at all, the system is nearly as controllable in the θ_2 direction as in θ_1. Truncating the time horizon is necessary to express what is physically obvious in this example: part of the system is much easier to control than another. Selecting values of 1% for damping and 10% for coupling (ζ=0.01, ε = 0.1), the displacement Gramian is solved for several characteristic times. Figure 3-2 shows plots of the reachable set for unit norm control, that is plots of $1 = \underline{\theta}^T W^{-1} \underline{\theta}$.

As time increases, the reachable volume becomes 'rounder', and the Gramian reaches the steady-state solution

$$W_\theta = \begin{bmatrix} 13.1070 & 1.2626 \\ 1.2626 & 12.1455 \end{bmatrix} \qquad 3\text{-}39$$

which has eigenvalues $\lambda_{1,2}$ = (13.98, 11.28), so the lengths of the steady state axes are the square roots, 3.74 and 3.36. The envelope of the Gramian grows approximately as $(1 - e^{-2\zeta t})$ for large times. For t=50 we would expect the envelope to be about 63% of full size with axes of length 2.36 and 2.12, and this corresponds with Figure 3-2b.

For control intervals on the order of the characteristic time, the reachable volume clearly reflects the asymmetric distribution of control authority. Furthermore, a bounded control interval more closely represents the actual closed-loop objectives of command following or disturbance rejection than does the infinite horizon. A logical interval is the inverse of the desired closed-loop bandwidth. For this reason, and because of the sensitivity to the damping parameter, results based on the infinite horizon Gramian, which are common in the literature, may be misleading.

3.2.2 Scalar measures of Controllability and Observability

A reachable (or detectable) volume is an intuitively satisfying notion for controllability/observability analysis; however, we must still assign some scalar measure, J, to those sets to pose an optimization problem. This section reviews some candidate measures and motivates the selection of a particular class.

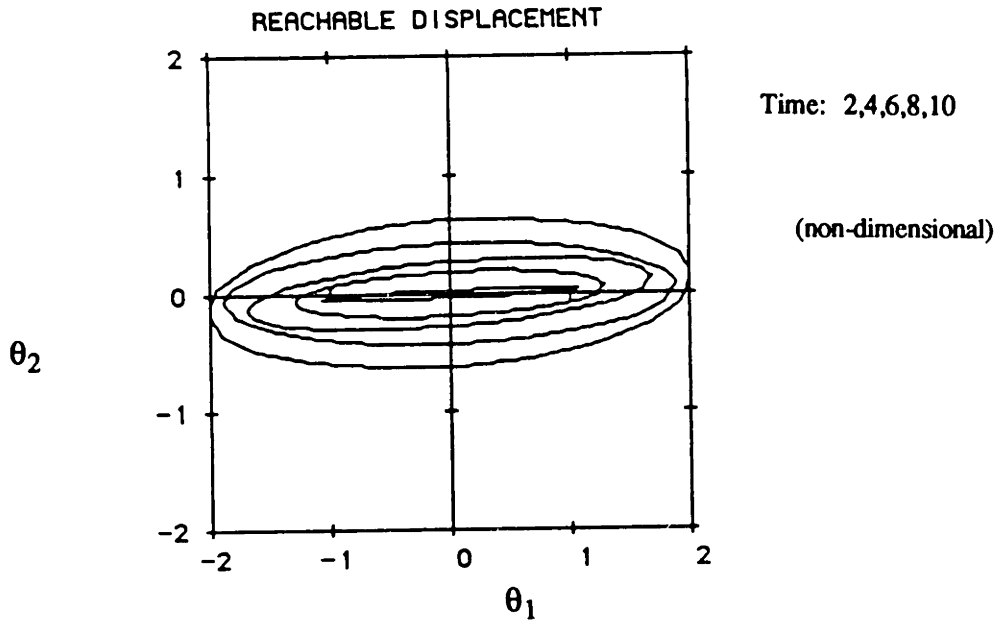


Figure 3-2a

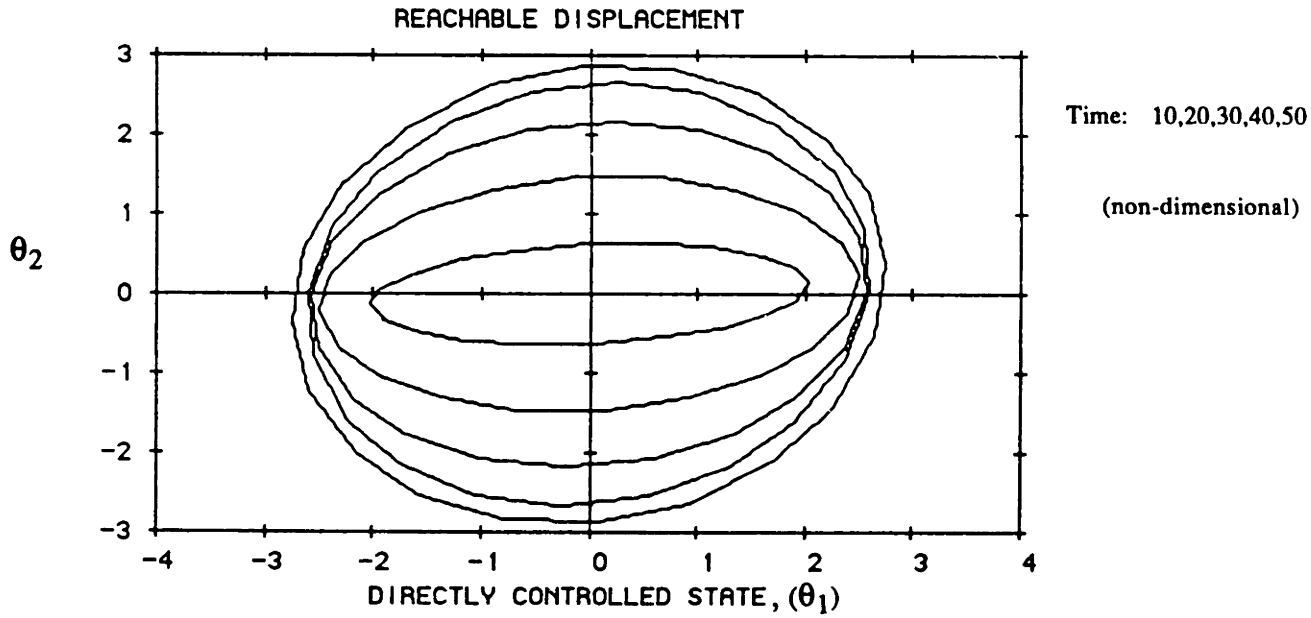


Figure 3-2b. Growth of Reachable Set.

The manner in which the eigenstructure of the controllability and observability Gramians describes the reachable and detectable volumes makes their use an obvious choice. Kalman, Ho, and Narendra originally suggested [35] the determinant and minimum eigenvalue of the Gramian as a scalar measure. Müller and Weber [36] amplified these ideas with the introduction of some additional requirements; and their results are reviewed now.

A scalar index, J(W), defined on a class of non-negative definite matrices W should have the following properties [36]:

$$J(W) = 0 \text{ for } \text{rank}(W) \leq n \tag{3-40}$$

$$J(W) > 0 \text{ for } \text{rank}(W) = n \tag{3-41}$$

$$J(\alpha W) = \alpha J(W) \text{ for scalar } \alpha \tag{3-42}$$

$$J(W_1) \geq J(W_2) + J(W_3) \text{ for } W_1 = W_2 + W_3. \tag{3-43}$$

These properties define a norm for W. The first two state J(W) is non-zero only for fully controllable/observable systems. Properties 3-42 and 3-43 result in J being concave for a combination of controls (sensors); that is

$$J[\alpha W_1 + (1-\alpha)W_2] \geq \alpha J(W_1) + (1-\alpha)J(W_2), \ \alpha \in [0,1] . \tag{3-44}$$

Concavity becomes particularly useful in developing an optimization algorithm.

The minimum eigenvalue of W_c or W_o satisfies all four properties, and has the geometric property of defining the volume of the largest sphere inscribed within the ellipsoidal reachable set, $1 = \mathbf{x}_f^T W_o^{-1} \mathbf{x}_f$. The spherical volume, V_s, is exactly $k\lambda_{min}^{n/2}$, where k depends on n (k= π for n=2 and k=4π/3 for n=3). The proportionality constant is not important, so we may define $V_s = (\lambda_{min})^{n/2}$. However, the exponent of n/2 is inconsistent with the desired linearity in property, equation 3-42, so we take as the first performance measure the minimum eigenvalue itself:

$$J_1 \triangleq \lambda_{min}(W) . \tag{3-45}$$

Another physically meaningful measure is the average control effort expended to drive the final state in any direction on the unit sphere, $\|x_f\|^2 = 1$. Kalman, et al. [35] compute an expression for this average, which is

$$\overline{\|\mathbf{u}\|^2} = \frac{\int_{\|x\|^2=1} x^T W^{-1} x \, dx}{\int_{\|x\|^2=1} dx} = \frac{1}{n} \mathrm{tr}(W^{-1}) . \qquad 3\text{-}46$$

We are interested in making the average small, so we take the inverse as a second performance index. By expressing

$$J_2 \hat{=} \frac{n}{\mathrm{tr}(W^{-1})} = \frac{n}{\sum_{i=1}^{n} \frac{|W^{(i)}|}{|W|}} \qquad 3\text{-}47$$

where $|W^{(i)}|$ are the principal minors of W, J_2 is shown [36] to satisfy all four requirements for a useful measure. Proving the convexity property, equation 3-43, requires [36], [37] Bergstrom's inequality and Minkowski's inequality of order -1.

The determinant of a Gramian has the useful geometric property that it is related to the ellipsoidal volume of the reachable set. Dropping the proportionality constant involving π again, we may define V_e as that volume:

$$V_e \hat{=} \prod_{i=1}^{n} \lambda_i^{1/2} = |W|^{1/2} . \qquad 3\text{-}48$$

V_e does not, however, satisfy the homogeneity property, equation 3-42, but can be made to do so with the exponent 1/n rather than 1/2. A third performance measure is then,

$$J_3 \hat{=} |W|^{1/n} . \qquad 3\text{-}49$$

The sign definite properties of equations 3-40 and 3-41 are plainly satisfied, and the convexity property is precisely another inequality of Minkowski, [36], [37].

The Lyapunov equations governing the evolution of the Gramians are linear. Therefore, if W_{ci} is the control Gramian of actuator $\mathbf{b}_1$, the system Gramian of m inputs is the sum,

$$W_c = \sum_{i=1}^{m} W_{ci} , \qquad 3\text{-}50$$

$$\text{since } BB^T = \sum_{i=1}^{m} \mathbf{b}_i \mathbf{b}_i^T \quad \text{for } B = [\mathbf{b}_1 \, \mathbf{b}_2 \ldots \mathbf{b}_m] . \qquad 3\text{-}51$$

While the scalar measures are non-linear functions of the eigenvalues, the linearity of the Gramian is computationally convenient if the number of candidate actuators is finite. It can

be advantageous to pre-compute and store the contributions from each site, W_i, since the number of possible combinations an algorithm might process can easily exceed the number of sites. This is particularly true when the combinatorial complexity of failures enters the problem.

It is certainly possible to define meaningful measures of controllability and observability without strictly adhering to the formal properties of a norm. Vander Velde and Carignan [26] define a 'degree of controllability' (D_c) that explicitly incorporates the 'out-of-roundness' of the recoverable set. The 'elliptical volume', V_e, is proportional to the volume of a 'recoverable' set, and V_s is the largest inscribed 'spherical volume':

$$D_c = \left[V_s + \frac{V_s}{V_e}(V_e - V_s) \right]^{1/n} \qquad 3\text{-}52$$

where the 'recoverable' volume refers to the set of initial states that can be driven to $x(t) = 0$ under an (L_2) optimal control. The control that does this is

$$u(\tau) = -R_u^{-1} B^T e^{-A^T \tau} W_{c*}^{-1}(t)\, x_0 . \qquad 3\text{-}53$$

The 'recoverable' control Gramian, W_{c*} is related to the 'reachable' form, W_c, by

$$W_{C*}(t) = e^{-At} W_c(t) e^{-A^T t} = \int_0^t e^{-A(t-\tau)} B R_u^{-1} B^T e^{-A^T (t-\tau)} d\tau \qquad 3\text{-}54$$

which is governed by the Lyapunov equation

$$-\frac{d}{dt} W_{C*}(t) = A W_{C*}(t) + W_{C*}(t) A^T - B R_u^{-1} B^T, \quad W_{C*}(0) = 0. \qquad 3\text{-}55$$

The norm of the optimal control is

$$\| u \|^2 = x_0^T W_{C*}^{-1}(t)\, x_0 \qquad 3\text{-}56$$

which defines a recovery region. Note that for strictly stable systems the recovery region becomes unbounded as $t \to \infty$ since the state decays exponentially to the origin and

$$\lim_{t \to \infty} W_{C*}^{-1}(t) = 0. \qquad 3\text{-}57$$

A 'degree-of-observability' (D_O) is defined based on the Fisher Information matrix, which is equivalent to a form of the observability Gramian. Note that these measures are a combination of J_1 and J_3 for W_{C^*}, but do not satisfy the desired homogeneity property of equation 3-42 precisely, i.e., $D_c(\alpha W) = \sqrt{\alpha} D_c(W)$ for scalar α. While this is just an inconvenience, D_C and D_O are more difficult measures to incorporate into an algorithm than the simpler J_1, J_2 or J_3 forms. The issue of 'roundness' is somewhat qualitative, and will be addressed in part by scaling the state basis.

There are other criteria upon which to base a recovery or reachable region. A recovery region defined by time and fuel optimal control is developed by Lindberg [38] based on Viswanathan, et al. [39]. The constraints

$$|u_i(t)| \leq 1 \quad i=1..m \tag{3-58}$$

$$\sum_{i=1}^{m} \int_0^t |u_i(\tau)|\, d\tau \leq F \tag{3-59}$$

where F represents here the 'fuel' limit, can be traded off between purely time-optimal, for $F \geq mT$, and fuel-optimal for large T, in the solution of the control via Pontryagin's minimum principle. While the recovery region does not generally have an analytical closed form, these researchers develop polytope approximations and define a scalar controllability measure as the smallest dimension in the recovery region. Lindberg [38] also considers the 'energy' optimal control and a scalar measure equivalent to $J_1 = \lambda_{min}(W_{c^*})$ for the recoverable form of the Gramian as previously discussed. The polytope approximations of the recovery region for lightly damped structural systems are not, however, easy to incorporate into a recursive algorithm.

The size of the models commonly required when dealing with structural systems provides a strong motivation for computational efficiency. Structural dynamics have a special form and may be represented by the second-order model

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{D}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{B}\mathbf{u}(t), \quad \mathbf{M}=\mathbf{M}^T > 0, \ \mathbf{K}=\mathbf{K}^T \geq 0 \tag{3-60}$$

$$\mathbf{y}(t) = \mathbf{C}_d\mathbf{q}(t) + \mathbf{C}_v\dot{\mathbf{q}}(t) \tag{3-61}$$

where $\mathbf{q}$ is the physical displacement.. Hughes and Skelton [40] have exploited the properties of this form to define controllability and observability measures for undamped,

damped, and gyric systems. These are extended [41] to rigid and multi-body modes as well. The modal form of the dynamics with the transformation $\mathbf{q}=\Phi\eta$ is

$$\ddot{\eta} + 2Z\Omega\dot{\eta} + \Omega^2\eta = \Phi^T B\,\mathbf{u} \qquad 3\text{-}62$$

$$\mathbf{y} = C_d\Phi\eta + C_v\Phi\dot{\eta} \qquad 3\text{-}63$$

$$\Phi^T B = [\mathbf{b}_1 \ \ldots \ \mathbf{b}_n]^T. \qquad 3\text{-}64$$

Note that $\mathbf{b}_i$ represents here the rows of $\Phi^T B$, and not the columns as defined elsewhere. Each row is, of course, m elements long for the m control signals in $\mathbf{u}$. For the case of distinct modal frequencies, the controllability index for each mode is defined as

$$J_i = \mathbf{b}_i\mathbf{b}_i^T \qquad i=1,n \qquad 3\text{-}65$$

and the system is controllable if and only if $J_i > 0$, $\forall$ i. An observability index is defined based on the column norms of $C\Phi$, again for each mode, and a generalization for identical frequencies is also developed. Skelton and Hughes [40] also suggest the set of controllability indices as suitable optimization parameters, but a scalar measure incorporating more than one mode is not defined. It does not follow, however, that good controllability properties for each mode imply good properties for the system. Consider again the coupled pendula example. For $\varepsilon > 0$, there are distinct eigenvalues, although closely spaced, and close spacing is certainly a common characteristic of structural dynamics. Each mode is individually as well controlled as the other by the single input, but the system as a whole is not. This near linear dependence between the modes is precisely what is captured by the Gramian.

Faced with a sensor and actuator selection problem of large scale in the ACOSS (Active Control Of Space Structures) program, Fogel [42] proposed an approach exploiting the directions represented by the vectors of modal influence coefficients. Including a disturbance input, $\mathbf{w}$, and a control objective vector $\mathbf{z}$, the structural dynamics in second order form is:

$$\ddot{\eta} + 2Z\Omega\dot{\eta} + \Omega^2\eta = \Phi^T B_a u + \Phi^T B_d w \qquad 3\text{-}66$$

$$y = [C_{Sd}\Phi]\eta + [C_{Sv}\Phi]\dot{\eta} \qquad 3\text{-}67$$

$$z = [C_z\Phi]\eta \qquad 3\text{-}68$$

$$C_z = \begin{bmatrix} c_1^T \\ c_2^T \end{bmatrix}. \qquad 3\text{-}69$$

Control of a two dimensional optical line of sight in the presence of a specified wide-band random disturbance is the particular ACOSS problem. The actuator selection problem is to include a relatively small number of actuators in B_a from a larger set, B_A, so there is good control authority in the directions (modes) strongly contributing to **z**. Sensor selection is based on a similar criterion so the rows of C_s 'match' the directions represented by the columns of B_d. For actuator selection, we seek a coefficient vector, $\mathbf{a}_i$, such that

$$c_{z_i} = [\Phi^T B_A]a_i \quad i = 1,2 \qquad 3\text{-}70$$

which almost certainly does not have an exact solution if the number of actuators is less than the number of modes. A criterion for minimization is then the error

$$J = \min_{a_i} \| c_{z_i} - [\Phi^T B_A]a_i \|^2. \qquad 3\text{-}71$$

With the least-squares solution for $\mathbf{a}_i$, the minimum value is

$$J = \| I - [\Phi^T B_A][\Phi^T B_A]^{\#} \|^2 \qquad 3\text{-}72$$

where # is the Moore-Penrose inverse. In their algorithm, Hegg and Kissel [22] exploit the properties of Householder transformations and 'QR' decompositions to identify columns of B_A that can be eliminated with minimal growth in J. Their algorithm starts with a large number of actuators, and proceeds with a sequential elimination process until a predetermined threshold in the error is exceeded for either LOS output influence coefficient vector. Admittedly heuristic, the method does not account for the spectral distribution of modes; but has been applied to a large scale model, the 'C.S. Draper Model No. 2', [22]. The setting of the threshold is also subjective, and the researchers noted a sharp growth in J

as the number of actuators (sensors) was reduced below the number of modes in the control design model.

All of the performance measures reviewed so far characterize the open-loop plant, although the control objective will ultimately be satisfied by closed-loop compensation. In general, multi-variable compensator design is a labor and engineering judgement intensive process not amenable to rapid recursion, say in an algorithm for sensor or actuator selection. A notable exception, if it can be properly posed, is the LQG approach.

3.2.3 LQG Design

Linear-Quadratic Gaussian (LQG) optimal methodology has become so ubiquitous in multi-variable control problems it is worthwhile to outline some of the basic relations between what are essentially open-loop properties, controllability and observability, and closed-loop performance, as reflected by the value of the LQG functional, J_{LQG}. A representation of the stochastic form of the cost for time invariant systems is

$$J_{LQG}(t_0,t_1) = E\left\{ \int_{t_0}^{t_1} \mathbf{x}^T(\tau)Q_x\mathbf{x}(\tau) + \mathbf{u}^T(\tau)R_u\mathbf{u}(\tau)\,d\tau + \mathbf{x}^T(t_1)P_{c1}\mathbf{x}(t_1) \right\}. \qquad 3\text{-}73$$

The optimal control and observer gains follow from the solutions to the independent Riccati equations:

$$-\dot{P_c}(t) = A^TP_c(t) + P_c(t)A - P_c(t)BR_u^{-1}B^TP_c(t) + Q_x \qquad P_c(t_1) = P_{c1} \qquad 3\text{-}74$$

$$\dot{P_o}(t) = AP_o(t) + P_o(t)A^T - P_o(t)C^TR_y^{-1}CP_o(t) + Q_w \quad P_o(0) = E[\mathbf{x}(0)\mathbf{x}^T(0)] \qquad 3\text{-}75$$

with Q_w the intensity of the state disturbance and R_y the intensity of the measurement noise. In the limiting steady state case ($t_0 \rightarrow -\infty$ and $t_1 \rightarrow \infty$), the value of the cost functional may be expressed

$$J_{LQG} = tr[P_cQ_w + P_oG^TR_uG] = tr[P_cHR_yH^T + P_oQ_x] \qquad 3\text{-}76$$

where G is the regulator gain matrix, $G = R_u^{-1}B^TP_c$, and H is the optimal observer (estimator) gain matrix, $H = P_oC^TR_y^{-1}$, [43]. The LQG cost is certainly a viable scalar measure by which to judge sensor and actuator placement, as is pointed out by the results

of Omatu and Seinfeld [20] reviewed in Chapter 2. For the finite-dimensional case, we may illustrate here some of the steps involved.

Suppose the control influence matrix, B, is continuously parameterized by a vector, ξ, representing actuators, $B(\xi)$. The gradient of J_{LQG} may be computed for each component in ξ by

$$\frac{\partial J_{LQG}(\xi)}{\partial \xi_i} = \mathrm{tr}\left[\frac{\partial P_c}{\partial \xi_i} HR_y H^T\right] = \sum_j \mathbf{h}_j^T \frac{\partial P_c}{\partial \xi_i} \mathbf{h}_j \qquad \mathbf{h}_j = \text{column j of } HR_y^{1/2} \qquad 3\text{-}77$$

where the matrix $\partial P_c/\partial \xi_i$ satisfies the equation obtained by differentiating the (control) algebraic Riccati equation

$$0 = [\frac{\partial}{\partial \xi_i} P_c][A - B(\xi)G] + [A - B(\xi)G]^T \frac{\partial}{\partial \xi_i} P_c - P_c\left[\frac{\partial}{\partial \xi_i}[B(\xi)R_u^{-1}B(\xi)^T]\right]P_c \quad . \quad 3\text{-}78$$

This is a Lyapunov equation and by inspection its solution is symmetric. The matrix [A-BG] is recognized as the closed-loop dynamics matrix and for fully controllable systems is known to have eigenvalues in the open left-half plane, hence solution $\partial P_c/\partial \xi_i$ is unique [33]. In order to compute the gradient of J with respect to ξ, one such Lyapunov equation must be solved for every component of ξ. The solution for the Riccati equation, P_c, is required as well, and in an algorithm cannot be summed from contributions of individual devices (columns of B) since that equation is non-linear. This may not be prohibitive for a problem of small dimensions that seeks only one solution, but if the combinatorial complexity of failures is introduced, the computational cost grows rapidly.

There are other issues that can reduce the practicality of the LQG cost functional as a measure of controllability or observability. It may be that minimization of J_{LQG} is not the actual control objective, but the LQG solution is used as a means to achieve some other property, such as loop transfer function shapes in the frequency domain. The results are highly sensitive to the particular weights chosen, and the weights are yet another set of parameters to specify, in addition to scaling and model selection. These parameters may not be well justified early in a design effort. For several reasons it may be more appropriate to measure the fundamental system properties of controllability and observability directly rather than against the LQG functional.

There is, however, a close relation between the controllability and observability Gramians and the important (non-negative definite, symmetric) Riccati equation solution. For the special case of Q = 0 (process noise for the estimator or state weighting for the filter) and $P(t_1)$ or $P(0) \rightarrow \infty$ (large terminal state weighting or no knowledge of the initial state) a version of the Gramian is exactly P^{-1}. Consider the estimator Riccati equation for this case ($Q_w = 0$ for $E[x(0)x^T(0)] \rightarrow \infty$) so $P_o^{-1}(0) = 0$

$$\dot{P}_o = AP_o + P_oA^T - P_oC^TR_y^{-1}CP_o \ . \qquad 3\text{-}79$$

With the relation

$$\frac{d}{dt}\left[P_o(t)^{-1}\right] = -P_o(t)^{-1}\left[\frac{d}{dt}P_o(t)\right]P_o(t)^{-1} \qquad 3\text{-}80$$

the inverse propagates by

$$-\frac{d}{dt}\left[P_o^{-1}\right] = P_o^{-1}A + A^TP_o^{-1} - C^TR_y^{-1}C \qquad P_o^{-1}(0) = 0 \qquad 3\text{-}81$$

which is the same equation that governs the 'final -time' observability Gramian,

$$P_o^{-1}(t) = W_{o*}(t) = e^{-A^Tt}W_o(t)e^{-At} \qquad 3\text{-}82$$

where $W_o(t)$ is the 'initial-time' observability Gramian (3-11). The distinction is motivated by the form of the deterministic observer equation. The 'final' state, $x(t)$, given observations $y(\tau)$, $\tau \in [0,t]$, and with no unknown inputs may be computed by

$$x(t) = W_{o*}^{-1}(t)\int_0^t e^{-A^T\tau}C^TR_y^{-1}y(t-\tau)\,d\tau \qquad 3\text{-}83$$

which is similar to the form for initial state $x(0)$ with $W_o(t)$ in equation 3-16. The final-time Gramian is also equivalent to the Fisher Information matrix, and it was with this interpretation Vander Velde and Carignan [26] defined their scalar 'degree of observability' for sensor placement. It is interesting to note how process noise modeled by Q_w requires feedback of residuals in the model-based estimator and changes the Lyapunov equation to a Riccati equation.

The dual relation for the regulator Riccati equation is

$$P_c^{-1}(t_1 - t) = W_{c*}(t) = \int_0^t e^{-A(t-\tau)} B R_u^{-1} B^T e^{-A^T(t-\tau)} d\tau \qquad 3\text{-}84$$

where t_1 is the terminal time in the regulator cost functional and W_{c*} is the 'recoverable' control Gramian (3-54).

This relationship with a special case of the Riccati equation leads to a handy method for computing the Gramians. The (control) Riccati solution may be obtained from the Hamiltonian system:

$$\frac{d}{dt}\begin{bmatrix} x(t) \\ x_c(t) \end{bmatrix} = \begin{bmatrix} A & -BR_u^{-1}B^T \\ Q_x & -A^T \end{bmatrix}\begin{bmatrix} x(t) \\ x_c(t) \end{bmatrix} = A_H x_H \qquad 3\text{-}85$$

where $x_c(t)$ is the co-state. Defining $\Psi(t)$ as the solution (transition) matrix to the Hamiltonian system (A_H) and the corresponding partitions,

$$\Psi(t) = e^{A_H t} = \begin{bmatrix} \Psi_{11}(t) \Psi_{12}(t) \\ \Psi_{21}(t) \Psi_{22}(t) \end{bmatrix} \qquad 3\text{-}86$$

the Riccati solution [43] may be expressed with partitions of $\Psi(t)$ and the terminal time state weighting, P_{c1}:

$$P_c(t) = \left[\Psi_{21}(t-t_1) + \Psi_{22}(t-t_1)P_{c1} \right]\left[\Psi_{11}(t-t_1) + \Psi_{12}(t-t_1)P_{c1} \right]^{-1} . \qquad 3\text{-}87$$

In the special case $Q_x = 0$ and $P_{c1}^{-1} \rightarrow 0$, the solution to the ('recoverable') controllability Gramian is

$$P_c^{-1}(t_1-t) = W_{c*}(t) = \Psi_{12}(-t)\Psi_{22}^{-1}(-t) \qquad 3\text{-}88$$

with

$$\Psi(t) = \begin{bmatrix} e^{At} & -\int_0^t e^{A(t-\tau)} B R_u^{-1} B^T e^{-A^T \tau} d\tau \\ 0 & e^{-A^T t} \end{bmatrix}. \qquad 3\text{-}89$$

Notice how the Gramian propagates forward in time as the inverse Riccati solution propagates backwards from the terminal time. Laub [44] has reported this approach for the 'reachable' form using $-A_H$.

For the general case of non-zero process noise and state weightings, Joseph and Bucy [45] develop bounds for the (estimator) Riccati solution:

$$[W_{CQ}^{-1}(\Delta) + W_{o*}(\Delta)]^{-1} \leq P_o(t) \leq [W_{o*}^{-1}(\Delta) + W_{CQ}(\Delta)] \qquad 3\text{-}90$$

where W_{CQ} is the 'controllability' Gramian of the process noise:

$$W_{CQ}(\Delta) = \int_0^{\Delta} e^{A\tau} Q_w e^{A^T\tau} d\tau \qquad 3\text{-}91$$

and W_{o*} is the Information matrix (or 'final-state' form of the observability Gramian). Integration limit Δ is any such than $0 \leq \Delta \leq t$. For the more general time-varying case in [45] the interval for the Gramians is $(t,t-\Delta)$ for any $\Delta \leq (t_1-t_0)$. A dual form exists for the control Riccati equation, $P_c(t_1 - t)$, with W_{c*} dual to W_{o*} and W_{OQ}, the 'observability' Gramian of the state weighting, dual to W_{CQ}. A useful implication of these bounds [46], is if the Riccati equation is too burdensome for incorporation into an algorithm, sensor and actuator optimization based upon a measure of the appropriate Gramian indirectly optimizes the respective Riccati solution, and hence the value of the LQG cost functional.

3.2.4 Optimization Parameter

Of the several scalar measures of controllability and observability reviewed, we select the class defined by Müller & Weber [36] computed with the Gramian eigenvalues for developing an algorithm. The several physical interpretations of the Gramian, including the relative linear dependence of the rows of $e^{At}B$ (or columns of Ce^{At}), the quadratic form representation of the reachable (detectable) volume, and the bounding of the closed-loop Riccati equations partially motivate this choice. Computational economy considerations complete the motivation, particularly for the case of combinatorial failure modes. No other measure considered was as satisfactory with respect to both physical meaning and economy.

Within the class selected, however, a further decision of the particular form remains. Indeed Müller & Weber [36] took up the question of posing the problem such

that the particular index selected (within their class) did not affect the sensor or actuator optimal locations. This situation occurs if at the optimum the Gramian is a hypersphere, which would be a fortunate accident in our problem. Intuition suggests the minimum eigenvalue measure, equation 3-45, is not as desirable as the others since it depends entirely upon one direction in the reachable/detectable volume. While both the average-energy based (3-47) and total-volume based (3-49) norms depend on all directions in the state space, a subtle problem exists with the latter and scaling. In the next section we will develop an approach to scaling the state basis so that control and sensing are equally important in all directions, motivated by the idea that a measure of controllability or observability is useful only if we accomplish such a 'calibration' of the basis. Suppose ξ^* represents some optimal location of actuators such that $J_3(\xi^*) = |W_c(\xi^*)|^{1/n}$ is maximized. Then suppose we rescale the state basis with a nonsingular transformation, T. While the value of the measure at ξ^* is changed by the rescaling, $J_3 \rightarrow |T|^{-2/n} |W_c(\xi^*)|^{1/n}$, ξ^* is still the maximizing argument. This negates the intended effect of scaling. We are left with the average-energy form, (3-47), but rather than adopting it exclusively, we will experiment with all three on the examples in Chapter 5, keeping in mind the theoretical shortcomings of the minimum eigenvalue and total-volume measures.

3.3 Model Selection

In addition to the usual issues of model selection for structural dynamics (dimension in particular) reviewed in this section, scaling the state basis is fundamental to qualitative analysis of controllability and observability properties. We have already pointed out how relative input and output weightings are dealt with by R_u and R_y, respectively, so a 'unit-norm' control or output signal is properly interpreted. Since controllability and observability are both defined pointwise with respect to the state, we must also necessarily settle on a similar interpretation for unit displacement in the state basis.

3.3.1 Scaling the State Basis

If there is sufficient meaning given to each component of the state it is possible to choose a scaling directly. This may very well be the case for rigid body modes. Following the reasoning in [26], if $|x_{i_{req}}|$ is the maximum excursion state component x_i is required to reach over the control interval, then the transformation, $\mathbf{x} = T\mathbf{x}$ with $T = \text{diag}\{|x_{i_{req}}|\}$, results in a basis for x where unit deviations are equally important in all directions. Notice

that directions requiring large excursions shrink in the scaled space, since $x = T^{-1}\mathbf{x}$. The control Gramian is scaled from W_c for $\mathbf{x}$ to W_c for x with $W_c = T^{-1}W_cT^{-T}$; the reachable envelope has been divided by $|x_{i_{req}}|$ in direction i. The sensitivity (gradient) of the selected scalar measures are all highest for the smallest eigenvalues (smallest dimensions in reachable volume), so optimization based on these measures will attempt to improve control authority in those directions more than in the directions of the larger eigenvalues. The scaling also defines a normalization surface, which contains the reachable volume in the original basis that is transformed to the unit sphere in the scaled basis. This surface is described by the matrix $T^TT = \text{diag}\{|x_{i_{req}}|^2\}$. Directly scaling for observability may be justified in a similar manner. Recall the detectable volume was that region in state space outside the ellipsoid defined by the Gramian which produced at least a unit-norm signal. If $|x_{i_{max}}|$ is the state displacement in direction i that is required to produce a unit-norm output, then that distance may be thought of as the maximum tolerable error in direction i. The scaling $\mathbf{x} = Tx$ where $T = \text{diag}\{|x_{i_{max}}|\}$ results in a basis of equal required sensitivity. The observability Gramian transforms $W_o = T^TW_oT$, so the ellipsoid contracts in directions requiring greater sensitivity, i.e., small values of $|x_{i_{max}}|$. Scalar measures based on the observability Gramian eigenvalues will have a steeper gradient for improved sensor sensitivity in these directions as a result. The scaling also describes a required detectable volume, which in the original basis is $T^{-T}T^{-1} = \text{diag}\{|x_{i_{max}}|^{-2}\}$, that scales to the unit sphere.

A shortcoming of direct scaling is the need to explicitly select the scale factor for each state. For structural dynamics, the states are commonly amplitudes of normal modes, but can be in a basis with no physical interpretation. It is not easy to imagine how to select a maximum modal amplitude excursion for each mode. We seek a scaling approach that is handier, with fewer parameters, but accomplishes the intent of direct scaling.

3.3.2 Normal Forms

With the motivation of model reduction, Moore [47] investigated several particular basis selections, focusing in particular on 'balanced realizations' that have equal and diagonal Gramians, $W_c = W_o = \Sigma = \text{diag}(\sigma_i)$. Balanced realizations are useful for displaying the relative importance of modes with respect to signal throughput, since in that realization the system may be thought of as equally controllable and observable. Two other

special forms are defined: 'input normal' with $W_C = I$, $W_O = \Sigma^2$, and 'output normal' with $W_O = I$, $W_C = \Sigma^2$. Similarity transforms for stable controllable and observable systems to these realizations are simple. Define matrices of orthonormal eigenvectors, V_C and V_O, and diagonal matrices of eigenvalues, Λ_C and Λ_O, such that

$$V_c^T W_c V_c = \Lambda_c \quad \text{and} \quad V_o^T W_o V_o = \Lambda_o \,. \tag{3-92}$$

Then with U, an orthogonal matrix of the eigenvectors, and Σ^2 ,a diagonal matrix of the respective eigenvalues resulting from applying the diagonalization transformation for the controllability Gramian to the observability Gramian,

$$U^T \left[\left(V_c \Lambda_c^{1/2} \right)^T W_o \left(V_c \Lambda_c^{1/2} \right) \right] U = \Sigma^2 \tag{3-93}$$

the state transformation given in [44]

$$T = V_c \Lambda_c^{1/2} U \Sigma^{-k} \tag{3-94}$$

results in the realizations of input normal for k=0, balanced for k=1/2, and output normal for k=1. The construction of this transformation starts by normalizing the control Gramian, with $V_C \Lambda_C^{1/2}$, then rotating the results of that transformation on the observability Gramian to principal axes with U, and finally scales both to equal shapes with $\Sigma^{-1/2}$. It is equally valid to construct the transform the other way, starting with the observability Gramian. The numbers σ_i^2 in diagonal Σ^2 are called the 'second-order modes' [47] and are system invariants. The product $W_C W_O$ (or $W_O W_C$) transforms to $T^{-1} W_C W_O T$ (or $T^{-1} W_O W_C T^1$) and these are similarity transforms, so the eigenvalues are invariant and are the σ_i^2. Indeed, if we are not interested in the controllability or observability properties separately, a balancing transform may be established by taking T as the modal matrix of $W_C W_O$. The σ_i are also the singular values of the Hankel operator [48] if they are defined for the infinite horizon interval, $t \in [0, \infty)$, so while all the above properties hold for a truncated interval, $t \in [0, t_f]$, σ_i may then not have any other system interpretations. Gregory studied balanced realizations for lightly damped structural systems in [49], and developed approximations for the required transforms that are asymptotically correct as damping vanishes. A useful result is that the basis for a balanced realization is a particular

form of modal coordinates, provided the frequency separation of modes is not too small, i.e., $|\omega_i-\omega_j| >> \max(\zeta_i,\zeta_j)\max(\omega_i,\omega_j)$ for all distinct modes, $i \neq j$. Despite the theoretical usefulness of balanced realizations, their direct exploitation for sensor and actuator placement is complicated by the explicit dependence of such realizations on both the input and output signal paths, which we are trying to form. There are, however, other possible signal paths with which we may define a balanced realization, such as disturbances and the regulated variables we seek to control.

Consider a structural dynamic system with two classes of inputs, actuators and disturbances (**u** and **w**), and two classes of outputs, sensors and regulated variables (**y** and **z**), as represented by Figure 3-3.

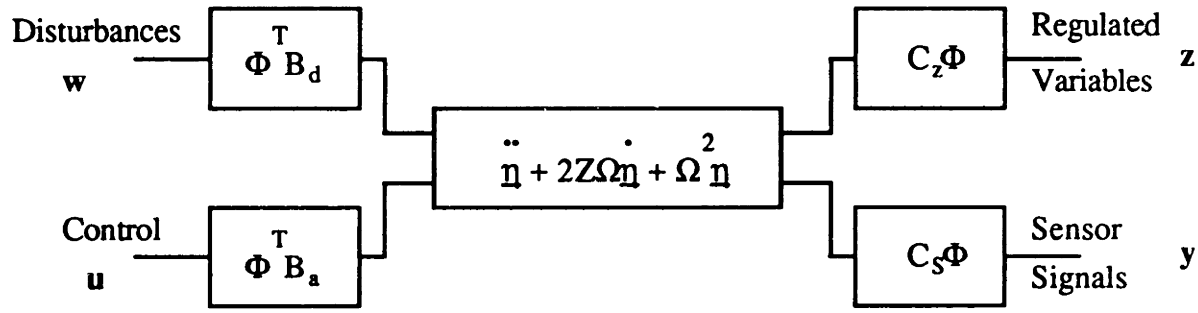


Figure 3-3. Structural Dynamic Signal Paths

This system is modeled by the equations

$$\ddot{\eta} + 2Z\Omega\dot{\eta} + \Omega^2\eta = [\Phi^T B_a]u + [\Phi^T B_d]w \qquad 3\text{-}95$$

$$y = [C_{Sv}\Phi]\dot{\eta} + [C_{Sd}\Phi]\eta \qquad 3\text{-}96$$

$$z = [C_{Zv}\Phi]\dot{\eta} + [C_{Zd}\Phi]\eta \ . \qquad 3\text{-}97$$

The overall design objective is to control the regulated variables to follow commands (including the regulator case of zero set-point) in the presence of disturbances. The sensor-actuator selection problem is to provide the control authority (B_a) and measurement sensitivity (C_S) so a compensator may be designed to accomplish the control objective. We assume here the disturbances and regulated variables are fixed and out-of-scope of the control engineer's design effort. Direct scaling of the state basis produces, in effect, a required , or reference, reachable control volume that is aligned with the original

state basis (principal axes) with dimensions $|x_{ireq}|$. The signal paths displayed in Figure 3-3 show two other choices for a reference reachable volume. The first is the reachable volume of the disturbances, since control authority ought to be distributed in some manner proportionally to the disturbances. The second is the inverse of the detectable volume of the regulated variables, which is a dynamic analogy to Hegg and Kissels' selection algorithm for actuators [22]. This may be thought of as normalizing the controllability Gramian by the shape of the observability Gramian, with the result that the optimization parameters in this basis have steeper gradients in the most sensitive output directions.

A scaling to 'input normal coordinates' by $T = V_c \Lambda_c^{1/2}$ effects the first approach, where V_c are the (normal) eigenvectors, and Λ_c the eigenvalues, of the disturbance controllability Gramian for the first order state $\mathbf{x}=[\eta^T \; d/dt\, \eta^T]^T$. In this basis, the disturbance Gramian represents the 'required volume' and is scaled to the unit hypersphere. This transformation is a generalization of direct scaling; the dimensions of the disturbance-reachable volume, $\lambda_c^{1/2}$, enter the transform as $|x_{ireq}|$ in the direct case, but there is now a rotation so the principal coordinates of the required reachable volume need not be aligned with the original basis.

Similarly, the scaling $T=V_o \Lambda_o^{1/2}$, where V_o and Λ_o are the eigenvectors and eigenvalues, respectively, of the regulated-variable observability Gramian, results in a basis that scales control effort by the degree it affects the regulated variables. The 'required volume' is represented in this scaling by the observability Gramian, and the label 'output scaled' is appropriate. This is a direct analogy to the idea implemented by Hegg & Kissel [22], except the use of the Gramian now accounts for dynamics. Note this basis is not 'output normal', for the observability Gramian is scaled to $T^T W_o T = \Lambda_o^2 \neq I$. Indeed, the basis 'output normal' results in a scaling that emphasizes adding control in directions least important to the output, which is exactly the opposite of the desired effect. This distinction motivates the order chosen for the balancing transformation, e.g., starting with the normalization with respect to the control Gramian, which is in the 'right' direction for the scaling objectives, rather than with the observability Gramian, which is not.

We may also wish to scale by some combination of the disturbance reachability and the regulated variables sensitivity. The 'balanced coordinate' basis is a natural starting point, and it is tempting at first to try $T = \Sigma^{1/2}$, since the controllability and observability Gramians in this basis are both Σ. The result, however, is equivalent to the previously-

defined 'input-normal' basis (plus a rotation, U) and the effect of the output weighting is exactly cancelled. We may see what to do by considering a signal from the input, to a point in state space, then to the output. Suppose over the control interval, $t \in [0,t_1]$, we command the final state to point in direction $x_f = \alpha e_i$ in the balanced basis subject to the constraint $\| u \|^2 = 1$, where e_i is a fundamental basis vector, with 1 in place i and 0 otherwise. We know the largest excursion reachable is $\sigma_i^{1/2}$ in this direction. Now if we start the observation interval and allow free response of the system, the output signal norm over $t \in [t_1,t_2]$, with (t_2-t_1) the integration interval of the observability Gramian, is $\| y \| = \sigma_i^{1/2}|x_f| = \sigma_i$. There is another factor of $\sigma_i^{1/2}$ to scale by. We adopt as a 'required volume' the product of W_c and W_o in the balanced realization, which is Σ^2, and we adopt the label 'throughput-scaled' for this basis. A transformation that implements this rule from the 'balanced' realization is $T=\Sigma$, and from any realization is

$$T = V_c \Lambda_c^{1/2} U \Sigma^{1/2} \qquad 3\text{-}98$$

where U is the rotation defined in the transformation to balanced coordinates. The required volume in the original basis is represented by

$$W_{req} = TT^T = V_c \Lambda_c^{1/2} U \Sigma U^T \Lambda_c^{1/2} V_c^T \qquad 3\text{-}99$$

and is observed to be close to a symmetric square root of $(W_c W_o)(W_c W_o)^T$ for the structural systems considered. That is, $W_{req} W_{req}^T \cong (W_c W_o)(W_c W_o)^T$; however this observation was not pursued to any formal bound. There are certainly other ways to construct a scaling weighting both inputs and outputs, but it is important to make the scaling directly proportional to both the disturbability and observability, and not inversely with respect to one as is done in construction of balanced realizations.

A similar process guides definition of observability scalings. In an extension of the direct scalings, we seek more general 'detectable surfaces' based on the disturbance to regulated variables signal path. Since the 'output-normal' realization is now in the appropriate direction for scaling, the balanced realization construction proceeds in the opposite order as for control scaling. The 'output-normal' transformation is $T=V_o \Lambda_o^{-1/2}$, so the square roots of the eigenvalues of the observability Gramian are inversely analogous to the direct scaling $|x_{imax}|$, the maximum acceptable error or excursion. Recall an initial state of length $\lambda_o^{-1/2}$ in the direction of the associated eigenvector produces a unit norm

output signal. Also in opposition to control authority scaling is the interpretation of 'input-normal' basis. In this basis, the disturbability volume dimensions, $\lambda_c^{1/2}$, would associate with the maximum permissible error,$|x_{imax}|$, and the proper relation is the inverse. Table 3-1 summarizes the several scalings discussed.

Table 3-1. Scaling Summary.

Scaling Type	Normal surface (W for $1=\mathbf{x}^T W \mathbf{x}$ original basis except as noted)	Transform from $\mathbf{x}$ original to x scaled basis $\mathbf{x} = Tx$
	Actuators	
Direct	$\mathrm{diag}\{\|x_{ireq}\|^2\}$	$\mathrm{diag}\{\|x_{ireq}\|\}$
Input Normal	W_c of Disturbance	$V_c\Lambda_c^{1/2}$
Throughput Scaled	Σ^2 in Balanced Realization	$V_c\Lambda_c^{1/2}U_1\Sigma^{1/2}$
Output Scaled	W_o of Regulated Variables	$V_o\Lambda_o^{1/2}$
	Sensors	
Direct	$\mathrm{diag}\{\|x_{imax}\|^{-2}\}$	$\mathrm{diag}\{\|x_{imax}\|\}$
Input Scaled	W_c of Disturbance	$V_c\Lambda_c^{-1/2}$
Throughput Scaled	Σ^2 in Balanced Realization	$V_o\Lambda_o^{-1/2}U_2\Sigma^{-1/2}$
Output Normal	W_o of Regulated Variables	$V_o\Lambda_o^{-1/2}$

where: $U_1^T[(V_c\Lambda_c^{1/2})^T W_o(V_c\Lambda_c^{1/2})]U_1 = \Sigma^2$; $U_1^T U_1 = I$

$U_2^T[(V_o\Lambda_o^{1/2})^T W_c(V_o\Lambda_o^{1/2})]U_2 = \Sigma^2$; $U_2^T U_2 = I$

Since the transformations must be non-singular, all require a completely controllable and/or observable system, and the eigenvalues of the normalizing Gramian, $\underline{\lambda}_c$ or $\underline{\lambda}_o$, are strictly positive. This topic is taken up again in section 3.3.3 on model order reduction.

The issue of scaling for the state basis can be dispensed with entirely for the case of point-wise output control or impulsive input observation. Recall the output controllability Gramian, $CW_c(t)C^T$, and the input observability Gramian, $B^TW_o(t)B$, do not depend upon the state basis. The dimensions of these Gramians can be considerably smaller than the state order, although their computation requires the full order W_c or W_o. We could base an algorithm entirely on output controllability or impulsive input observability, placing actuators to control the regulated variables and sensors to observe the disturbance inputs.

Such an approach leaves open, however, the possibility of placing actuators, for example, that while they do affect the regulated variables strongly, propagate to the output through structural modes which could well be orthogonal to the modes the disturbance affects. Pointwise control, state or output, is not the objective itself, but merely a computationally tractable method which has properties, we hope, that carry over into the actual closed-loop objectives. Therefore, this appealing simplification is rejected in favor of pointwise state control and estimation with attendant scaling requirements.

Use of a reference signal path to define a basis scaling significantly reduces, compared to direct scaling, the quantity of free parameters. There still remains, however, several choices and no clearly motivated manner of distinction. Intuition suggests 'throughput-scaling' should, by reflecting both the disturbance reachable set and regulated variables detectable set, be somehow superior to the others, but we do not have a precise way to measure that. Instead, we will try several scalings, along with other permutations of the overall approach, in the examples of Chapter 5 and note the results.

3.3.3 Model Order Selection

Creating a mathematical model of a structural system is a significant exercise in its own right before issues such as controllability, observability, state basis, and scalings can even be discussed. The computational steps are briefly reviewed here to provide a context for the more formal model order reduction problem. References for structural dynamic modeling and analysis include Meirovitch [50], and Strang and Fix [51].

A first result of a modeling process for structural analysis and response dynamics, as opposed to control synthesis, is a set of equations with at least the form

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \mathbf{B}\mathbf{u} \qquad \mathbf{M} > 0 \qquad \text{3-100}$$

$$\mathbf{y} = \mathbf{C}_d\mathbf{q} + \mathbf{C}_v\dot{\mathbf{q}} \qquad \mathbf{K} \geq 0 \qquad \text{3-101}$$

where $\mathbf{q}$ is a vector of physical displacements (translation and rotation) at a discrete set of node points, including possible rigid body displacements. Positive definite mass and positive semi-definite stiffness matrices M and K are taken as the result of finite element modeling of a spatial discretization of the structure by patch functions, such as elemental rods, beams and plates. Linearization of spinning masses results in skew-symmetric rate and displacement coefficient matrices, but their inclusion does not materially affect the

conclusions reached here, so we stick with the simpler case. Input **u**, including possibly distributed forces, are integrated over the patch functions and reflected as discrete nodal forces and torques with influence coefficients B for actuators and disturbances. Specified outputs, such as the regulated variables of the control problem, or sensor outputs are modeled in **y** as a combination of physical displacements, **q**, and rates. Computation of M, K, and the input and output influence matrices may be carried out by a code such as NASTRAN. The dimensions of **q** can be large, such as 6v where v is the number of structural nodes, and in the case of the Draper Model No. 2 used for this thesis, v = 59 nodes so 6v = 354. We form an eigenvalue problem from equation 3-100,

$$\omega_i^2 M\Phi_i = K\Phi_i \qquad 3\text{-}102$$

where ω_i are the real modal frequencies and Φ_i the mode shapes, which are orthogonal with respect to the mass and stiffness matrices and can be normalized such that:

$$\Phi_i^T M\Phi_j = 1\delta_{ij} \qquad \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \qquad 3\text{-}103$$

$$\Phi_i^T K\Phi_j = \omega_i^2\delta_{ij} \; . \qquad 3\text{-}104$$

Note that while there are up to 6v eigenvalues and vectors, only a small fraction, usually less than 1/2, are actually computed, and an even smaller fraction commonly demonstrate good correlation with hardware. The eigenvectors represent a basis for R^{6v}, and it is formally correct to express **q** as the expansion $\mathbf{q} = \Phi\eta$, where Φ is a square matrix of all 6v eigenvectors. However, what is done of practical necessity is to express **q** approximately in the subspace spanned by a truncated set of eigenvectors:

$$\mathbf{q}(t) \approx \sum_{i=1}^{n} \Phi_i \eta_i(t) = \Phi_1 \eta \quad v < 6v, \qquad \Phi_1 = [\Phi_1 \cdots \Phi_n] \; . \qquad 3\text{-}105$$

This first truncation may still employ all the eigenvectors actually computed. Truncation is by frequency, retaining the lower n modes, and the vector $\eta \in R^n$ is the modal amplitudes. In this subspace the structural dynamics may be represented by the uncoupled equations

$$\ddot{\eta} + \Omega^2 \eta = \Phi_1^T B \mathbf{u} \qquad 3\text{-}106$$

$$\mathbf{y} = C_d \Phi_1 \eta + C_v \Phi_1 \dot{\eta} + D\mathbf{u} \; . \qquad 3\text{-}107$$

The 'feedthrough' term may be thought of as the input-output correction for eliminating of part of the static stiffness, K, by modal truncation. The static deflection, **y**, from a steady load **u** to the system is $\mathbf{y} = CK^{-1}B\mathbf{u}$ (assuming an elimination of the rigid body modes resulting in K>0). Partitioning the modal matrix, Φ, into the n retained modes, $\Phi_1 \in R^{6v \times n}$ and the $6v$-n truncated modes, Φ_2, $\Phi = [\Phi_1 \; \Phi_2]$, and compatibly partitioning the (diagonal) eigenvalue matrix, the inverse stiffness matrix may be written:

$$K^{-1} = \Phi_1 \Omega_1^{-2} \Phi_1^T + \Phi_2 \Omega_2^{-2} \Phi_2^T \; . \qquad 3\text{-}108$$

The second term is the portion of the static gain lost by modal truncation, so a representation of D is $C\Phi_2[\Omega_2]^{-2}\Phi_2^T B$. Though modes Φ_2 may not be computed, D may be computed with equation 3-108, recognizing the columns of the matrix $[K^{-1}B]$ as the (force) influence coefficients, which may be computed by NASTRAN or other finite element coding,

$$D = C[K^{-1}B] - C\Phi_1 \Omega_1^{-2} \Phi_1^T B \; . \qquad 3\text{-}109$$

Including a feedthrough results in a plant that is not strictly proper, and is inconvenient for control synthesis, so the feedthrough is often dropped. Non-inclusion of D may be justified, however, if fidelity of static deflection is not as important as dynamic response. This is often an acceptable situation for disturbance rejection and vibration damping as a part of command following for a relatively small number of regulated variables, exploiting perhaps high DC gain in the loop as compensation. It may not be acceptable for static shape control where a limited number of measurements are used to infer distributed deflections. For feedback control synthesis, one may invoke a frequency separation argument to solve the dynamic problem separately from the steady state problem, which is further supported by the often distinct hardware implementations of nearly-static and dynamic devices. Controllability and observability properties pointwise

with respect to a state basis do not involve the feedthrough, and this research concentrates on the dynamic problem alone.

Structural dynamics analysis of spacecraft is required for launch loads and transient shock response whether a control-structure interaction problem exists or not. To compute the response to a random loading, usually specified in terms of a power spectral density, a damping estimate is necessary. A viscous damping model is convenient, since it enters as a linear term. If a damping matrix, introduced to equation 3-100, as a coefficient of dq/dt, is a linear combination of the mass and stiffness, say $\alpha_1 M + \alpha_2 K$, then the undamped normal modes used in the decoupling transformation also decouples the damped system, which is particularly convenient. A diagonal modal damping matrix, $Z = \mathrm{diag}(\zeta_i)$ may be computed by:

$$2Z\Omega = \Phi_1^T[\alpha_1 M + \alpha_2 K]\Phi_1 = \alpha_1 I + \alpha_2 \Omega^2 \qquad 3\text{-}110$$

so $\zeta_i = (\alpha_1/\omega_i + \omega_i\alpha_2)/2$. While some physical damping properties are adequately modeled in this way, the overall validity of linear and proportional damping is controversial and a subject of current research and engineering interest [52]. Damping mechanisms and modeling are also out of scope for this research. We assume some level of linear modal damping can be specified to adequately represent the important phenomena, say amplitude and phase change at a resonance. The shortcomings of linear damping assumptions, however, further motivates the use of the truncated time horizon Gramian. It is well defined for a range of damping including zero, and relatively insensitive to the details of the damping mechanism, while the infinite horizon case is not.

At this point, a first order model (A,B,C) and its associated transfer function, $G(s) = C(sI - A)^{-1}B$ may be computed and a formal model-reduction problem posed. A formulation is: Find a $G_r(s)$ of specified order that in some sense minimizes the error $\| G_r(s) - G(s) \|_p$, where p represents some norm. The first point is that a significant amount of 'model reduction' takes place for structural dynamics before this formal problem can be dealt with. The second is it is well posed only over a specific input-output signal path. An important path for control design is the one we do not have and are interested in forming through actuators and sensors. Finally, this statement of the model reduction problem does not have, to our knowledge, a general solution. Nevertheless, it is neces-

sesary to start somewhere and the disturbance to regulated variables signal path is specified and a natural point of departure. We review briefly some results in model order reduction.

The balanced realizations of Moore [47] are useful for reduction. Partition such a realization of fully controllable/observable system, (A,B,C), with $W_O=W_C=\Sigma^2$ as follows:

$$\frac{d}{dt}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}A_{11}A_{12}\\A_{21}A_{22}\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}+\begin{bmatrix}B_1\\B_2\end{bmatrix}u \qquad 3\text{-}111$$

$$y = C_1 x_1 + C_2 x_2 \qquad 3\text{-}112$$

then this partitioning may be thought of as representing two interconnected subsystems, (A_{11},B_1,C_1) and (A_{22},B_2,C_2). Moore has shown [47] for asymptotically stable systems any partitioning in this manner results in asymptotically stable subsystems. The 'second order modes' may be ordered so $\sigma_1^2 \geq \cdots \geq \sigma_n^2 > 0$, and Σ partitioned:

$$\Sigma = \begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix} \qquad 3\text{-}113$$

so

$$\Sigma_1 = \mathrm{diag}[\sigma_1^2, \cdots, \sigma_r^2] \qquad 3\text{-}114$$

$$\Sigma_2=\mathrm{diag}[\sigma_{r+1}^2, \cdots, \sigma_n^2]\,. \qquad 3\text{-}115$$

If there is an r such that $\sigma_r^2 >> \sigma_{r+1}^2$, then the subsystem (A_{11},B_1,C_1) is said to be 'internally dominant', and represents a viable reduced-order model. Moore's balancing algorithm requires starting with a fully controllable and observable system, which is not hard to do for structural dynamics in modal form. Any distinct mode is uncontrollable if its row of $[\Phi^T B]$ is identically zero and is unobservable if its column in $[C_d\Phi]$ (displacement) and $[C_v\Phi]$ (rate) are identically zero. For non-distinct modes, there is the additional requirement [40] the associated rows of $[\Phi^T B]$ and columns of $[C\Phi]$ be linearly independent. Elimination of all uncontrollable and unobservable modes results in positive definite controllability and observability Gramians, and the scalings discussed in 3.3.2 will not have any singularities. The degree of controllability and observability among the modes is then reflected in the σ_i. While such a reduction by balanced partitioning was not shown to have any particular optimality properties, Enns [53] showed the error is bounded:

$\| G_r(s) - G(s) \|_\infty \leq 2\ tr[\Sigma_2]$, where the infinity norm $\|F(s)\|_\infty$ is the maximum singular value of $F(j\omega)$ over all frequencies. Glover [48] studied all reduced-order $G_r(s)$ (of order r) that minimized the Hankel norm, $\|F(s)\|_H = \lambda_{max}^{1/2}(W_c W_o)$, where the Gramians of F(s) are defined over the infinite horizon. The H-optimal $G_r(s)$ includes a feedthrough, D, and the error is bounded by $\|G_r(s) - G(s)\|_H \leq \sigma_{r+1}[G(s)]$, where the σ_i are the Hankel operator singular values and/or the balanced realization eigenvalues of full order G(s). The H-optimal $G_r(s)$ also satisfies $\|G_r(s) - G(s)\|_\infty \leq tr[\Sigma_2]$, which is half the bound on the $G_r(s)$ resulting from partitioning the balanced realization, but does not include a feedthrough, while the H-optimal $G_r(s)$ does. In this case Σ_2 represents the equivalent partitioning of G(s), but it should be noted the H-optimal $G_r(s)$ is not simply the (A_{11},B_1,C_1) system. Enns [53] developed a recursive algorithm for $G_r(s)$ to minimize a weighted error $\|U(s)[G_r(s) - G(s)]Y(s)\|_\infty$. The choice of the infinity norm is motivated by stability robustness sufficient conditions, viewing the model error as the perturbation, and weightings U(s) and Y(s) emphasize modeling fidelity for frequencies in the crossover region. Weightings were constructed with a loop closing compensator, K(s): $U(s) = K(s)[I+G_r(s)K(s)]^{-1}$ or $Y(s)=[I+G_r(s)K(s)]^{-1}K(s)$, also with closed loop stability robustness motivations. While these details are not important here, Enns' computation involved frequency weighted Gramians, which are useful from a scaling point of view and will be discussed further in 3.3.4.

Lightly-damped modal dynamics imparts some additional structure to balanced realizations with useful results for model reduction, and have been studied by Gregory [49], Blelloch, et al., [54] and Jonckheere [55], the latter for the single-input-single-output (SISO) case. Asymptotically as $\zeta_j \rightarrow 0$ Gregory showed the second order modes (σ_j) approach identical pairs for each structural mode j, j=1,*n* where *n* is the number of structural modes,

$$\sigma_j^2 = \frac{[\mathbf{b}_j \mathbf{b}_j^T(\mathbf{c}_{vj}^T \mathbf{c}_{vj} + \mathbf{c}_{dj}^T \mathbf{c}_{dj}/\omega_j^2)]^{1/2}}{4\zeta_j\omega_j} \qquad 3\text{-}116$$

where $\mathbf{b}_j$ is a row of $[\Phi^T B]$; $\mathbf{c}_{dj}$ and $\mathbf{c}_{vj}$ the associated columns of the respective displacement and rate output matrices, $[C\Phi]$. If the modes are separated enough in frequency,

$$\frac{|\omega_i - \omega_j|}{\max(\zeta_i,\zeta_j)\,\max(\omega_i - \omega_j)} >> 1, \quad i \neq j \qquad 3\text{-}117$$

for all distinct modes, the transformation to balanced coordinates preserves the decoupling of the modes, again asymptotically as $\zeta_j \rightarrow 0$. The balancing transform, T, has the structure T = block diag[T_{jj}] j=1,*n* where T_{jj} is a 2x2 non-singular block for each mode, and depends on the particular basis the transform operates upon. The balanced Gramians have the form Σ = block diag[$\sigma_j I_{2x2}$] j=1,*n* illustrating the pairing mentioned. Model reduction is therefore equivalent to modal truncation with ordering by σ_j, which is not, of course, the same as ordering by frequency, ω_j. Blelloch, et al. [54] showed these results hold for second-order systems with skew-symmetric gyric terms as well. Gregory suggests select-ing modes by considering all four possible signal paths from disturbances and actuators to sensors and regulated variables, and retaining modes if they are significant in any path as measured by the σ_j computed for that particular path. The pairing of σ_j has a potentially significant practical implication in an algorithm for optimizing controllability and observability based on reachable and detectable volumes, which are generally of order n=2*n*. If the eigenvalues of W_o or W_c occur in pairs, then an ellipsoid of order *n* with distinct dimensions $\sigma_j^{1/2}$ conveys the same information about the reachable/detectable volume as the order 2*n* ellipsoids with paired dimensions.

In this work, we resolve the dilemma of requiring the signal path to be specified before model reduction in the following way: We construct a model by including modes in the balanced ordering of Moore [47] and Gregory [49], and the ordering is determined by the (fixed) disturbance to regulated-variables signal path. The model is then scaled by the procedures outlined in 3.3.2, and remains fixed for the optimization of a sensor or actuator signal path. No implication toward later modeling for compensator design is intended; it is more likely than not a revised model should be used then. Note this requires we always start with a fully controllable and observable model.

3.3.4 Frequency-Weighted Gramians

Construction of frequency weighted balanced realizations were motivated by Enns' idea that model reduction error is more important near the crossover frequency [53]. While that is not the same problem dealt with here, the idea that controllability and observability properties could be weighted explicitly by frequency, in addition to the scalings previously

discussed, is useful. Weighting transfer functions, U(s) at the input and Y(s) at the output, are assumed to be square, asymptotically stable, with representations (A_u,B_u,C_u,D_u) and (A_y,B_y,C_y,D_y), and are in the configuration shown in Figure 3-4 with system (A,B,C).

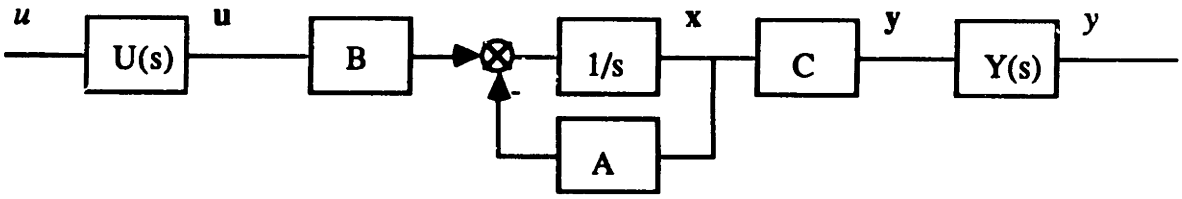


Figure 3-4. Input and Output Dynamic Weightings.

With state vectors $\mathbf{x}_u$ and $\mathbf{x}_y$ for U(s) and Y(s) and $\mathbf{x}$ for (A,B,C), and with the composite state $x^T=[\mathbf{x}^T\ \mathbf{x}_u^T\ \mathbf{x}_y^T]$ the connected systems are represented by (A,B,C) and

$$\frac{d}{dt}\begin{bmatrix}\mathbf{x}\\ \mathbf{x}_u\\ \mathbf{x}_y\end{bmatrix}=\begin{bmatrix}A & BC_u & 0\\ 0 & A_u & 0\\ B_yC & 0 & A_y\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{x}_u\\ \mathbf{x}_y\end{bmatrix}+\begin{bmatrix}BD_u\\ B_u\\ 0\end{bmatrix}u \qquad 3\text{-}118$$

$$y=\begin{bmatrix}D_yC & 0 & C_y\end{bmatrix}x. \qquad 3\text{-}119$$

For $W_c(t)$ and $W_o(t)$ representing the Gramian of the entire system, we are interested in the Gramians of the first partition alone, $W_c(t)$ and $W_o(t)$, which characterize the controllability of $\mathbf{x}$ through U(s) and observability of $\mathbf{x}$ through Y(s), respectively. We assume we do not introduce pole-zero cancellations, and the resultant loss of controllability or observability with the weightings. Writing with some significant abuse of notation, $G(\tau)=\pounds^{-1}\{(sI-A)^{-1}BU(s)\}$, and $F(\tau)=\pounds^{-1}\{Y(s)C(sI-A)^{-1}\}$ where $\pounds^{-1}$ denotes the inverse Laplace transform,

$$W_c(t)=\int_0^t G(\tau)G^T(\tau)d\tau=\begin{bmatrix}I & 0 & 0\end{bmatrix}W_c\begin{bmatrix}I\\ 0\\ 0\end{bmatrix}=W_{c\,11} \qquad 3\text{-}120$$

where the Gramian for the entire system is partitioned.

$$W_c = \begin{bmatrix} W_{c11} & W_{c12} & W_{c13} \\ W_{c12}^T & W_{c22} & W_{c23} \\ W_{c13}^T & W_{c23}^T & W_{c33} \end{bmatrix} \qquad 3\text{-}121$$

The observability Gramian we are interested in is simply the same partitioning:

$$W_o(t) = \int_0^t F^T(\tau)F(\tau)d\tau = [I\,0\,0]W_o\begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} = W_{o11} \quad . \qquad 3\text{-}122$$

We are interested in finding a balancing transform for the Gramians W_c and W_o defined with respect to x alone, which is not the same thing as balancing the entire system. A transform that does not affect U(s) or Y(s) at all has the form:

$$T = \begin{bmatrix} T_{11} & & \\ & I & \\ & & I \end{bmatrix} \qquad 3\text{-}123$$

which transforms the (controllability) Gramian of the connected system :

$$W_c \rightarrow T^{-1}W_cT^{-T} = \begin{bmatrix} T_{11}^{-1}W_{c11}T_{11}^{-T} & T_{11}^{-1}W_{c12} & T_{11}^{-1}W_{c13} \\ W_{c12}^T T_{11}^{-T} & W_{c22} & W_{c23} \\ W_{c13}^T T_{11}^{-T} & W_{c23}^T & W_{c33} \end{bmatrix} \qquad 3\text{-}124$$

so x Gramians W_c and W_o transform $W_c \rightarrow T_{11}^{-1}W_{c11}T_{11}^{-T}$ and $W_o \rightarrow T_{11}^{T}W_{o11}T_{11}$. We are interested in choosing T_{11} such that

$$T_{11}^{-1}W_{c11}T_{11}^{-T} = T_{11}^{T}W_{o11}T_{11} = \Sigma = \text{diag}\{\sigma_i\},\ \sigma_i > 0,\ i=1,n \qquad 3\text{-}125$$

which is precisely the same linear algebra problem as balancing the entire system, operating now on partitions W_{c11} and W_{o11} alone instead of the entire Gramian. We may compute the balancing transform as the modal matrix of $[W_{c11}W_{o11}]$, since

$$T_{11}^{-1}[W_{c11}W_{o11}]T_{11} = \Sigma^2 \qquad 3\text{-}126$$

represents an eigenvalue decomposition. Note the product $W_{c11}W_{o11}$ is not the [1,1] partition of full order $[W_cW_o]$. This has just been a book-keeping exercise showing that a part of a system can be balanced and scaled as desired and that pointwise state control or observation can be characterized with respect to that part alone. Since the controllability of **x** is independent of Y(s) and observability is independent of U(s), we need not incorporate both weightings if only one property or the other is of interest.

Weightings may be chosen to emphasize the directions in the reachable and detectable volumes corresponding to modes in the desired crossover frequency region, in parallel with Enns' original motivations. Alternatively, U(s) might represent a spectral shaping filter on the disturbance signals, and the disturbability Gramian is then computationally equivalent to the covariance of the system driven by correlated noise. Y(s) can represent output dynamics of sensors or the regulated variables. The attraction of 'externally' appending sensor or actuator dynamics is computational economy for an algorithm optimizing sensor and actuator authority, since we retain and operate only on the [1,1] partition of the connected systems. Even though the scalings previously developed do depend upon frequency insofar as each mode is at a discrete frequency, they do not depend explicitly upon frequency, and the additional freedom introduced by a dynamic input or output weighting may more accurately reflect the control problem at hand.

3.3.5 Order *n* vs. 2*n*

The first-order representation of a structural dynamic system (A,B,C) has a certain computation and data storage inefficiency from the $2n$ order of the state vector, where n is the number of modes in the model. Results of Jonckheere [55] and Gregory [49] show as damping diminishes, the second-order modes (σ_i) tend to pairs, and each pair associates with a structural mode. Measures of observability and controllability based on reachable volumes of a balanced realization, therefore, may need only include one σ_i since the measures are algebraic functions of these eigenvalues. While there is no savings in the computation of the Gramians in the first place, the control and sensor optimization algorithms developed in this work could benefit significantly by a 50% reduction in order. Exploiting this property requires the balancing transform for an unscaled model with a block diagonal dynamic (A) matrix to approximately have the structure,

$$T = \left[\left\{t_{11}\, t_{21}\right\} \cdots \left\{t_{1i}\, t_{2i}\right\} \cdots \left\{t_{1n}\, t_{2n}\right\}\right] \approx \text{diag}\left[T_{ii}\right] \qquad T_{ii} \in R^{2x2} \;\; i=1,n \; . \qquad 3\text{-}127$$

The off-diagonal blocks, T_{ij}, are not indentically zero, but satisfy $\| T_{ij} \| << \| T_{ii} \|$ for some matrix norm. The directions in each T_{ii} are not orthogonal, but do make each block non-singular. The balanced Gramian then has the structure.

$$\Sigma = \mathrm{diag}\left[\Sigma_i\right] \qquad \Sigma_i = \begin{bmatrix} \sigma_{1i} & 0 \\ 0 & \sigma_{2i} \end{bmatrix} \qquad \begin{matrix} \sigma_{1i} \approx \sigma_{2i} \\ i = 1,n \end{matrix} \tag{3-128}$$

The columns of T may be exchanged, including sign changes, without affecting the balancing properties of the transform, since the transform may be constructed with eigenvector matrices without regard to order. We may specify an ordering of the columns so there are two partitions, each with one column from each modal block, T_{ii}.

$$T = \left[\left\{ t_{11} \cdots t_{1i} \cdots t_{1n} \right\} \left\{ t_{21} \cdots t_{2i} \cdots t_{2n} \right\} \right] \tag{3-129}$$

The balanced Gramian now has the form

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \quad \begin{matrix} \Sigma_1 = \mathrm{diag}\left[\sigma_{1i}\right] \\ \Sigma_2 = \mathrm{diag}\left[\sigma_{2i}\right] \end{matrix} \quad i=1,n \quad \Sigma_1 \approx \Sigma_2 \tag{3-130}$$

so a norm on Σ_1 or Σ_2 is approximately equal to that based on Σ. The ability to partition T, however, hinges upon the relative linear independence of the individual modes. The approximate block structure is not at all apparent for Gramians computed over time intervals on the order of the modal periods rather than the infinite horizon used by Gregory [49] in his derivation.

The underlying reason lightly damped modal dynamics exhibit this structure is that within a mode during free response, the rate signal, $d/dt[\eta(t)]$, leads the displacement signal, $\eta(t)$, by $\zeta + \pi/2$, and ζ is small ($\zeta << \pi/2$) for the class of systems of interest. Within a modal block of a Gramian, with symbol W here, there are three different terms, represent-ing the three possible permutations of inner products of scalar rate and displacement signals over the time interval:

$$W(t) = \begin{bmatrix} \langle \dot{\eta}, \dot{\eta} \rangle & \langle \dot{\eta}, \eta \rangle \\ \langle \eta, \dot{\eta} \rangle & \langle \eta, \eta \rangle \end{bmatrix} \qquad \text{where } \langle f,g \rangle = \int_0^t f(\tau) g(\tau)\, d\tau \,. \tag{3-131}$$

For zero damping, the signals may be scaled and phased so $\eta(t) = \sin(\omega t)$ and $d/dt\ \eta(t) = \omega\cos(\omega t)$, and the three terms have the form

$$W_{11}(t) = \langle \dot{\eta}, \dot{\eta} \rangle = \int_0^t \omega^2 \cos(\omega\tau)^2 d\tau = \omega^2 \left[\frac{t}{2} + \frac{\sin(2\omega t)}{4\omega} \right] \qquad 3\text{-}132$$

$$W_{12}(t) = \langle \dot{\eta}, \eta \rangle = \int_0^t \omega \sin(\omega\tau)\cos(\omega\tau)\, d\tau = \frac{1}{2}\left[\sin(\omega t)\right]^2 \le \frac{1}{2} \ \forall\ t \qquad 3\text{-}133$$

$$W_{22}(t) = \langle \eta, \eta \rangle = \int_0^t \sin(\omega\tau)^2 d\tau = \frac{t}{2} - \frac{\sin(2\omega t)}{4\omega}. \qquad 3\text{-}134$$

The secular growth of the diagonal terms force the 'diagonal dominance' within a block. Between modal blocks, terms like $< \eta_i, \eta_j >$ and $< \dot{\eta}_i, \dot{\eta}_j >$, $i \neq j$, exhibit near-secular growth for $\omega_i \approx \omega_j$ over time intervals short compared to $1/(\omega_i - \omega_j)$. The cross terms, such as $< \eta_i, \dot{\eta}_j >$ and $< \dot{\eta}_i, \eta_j >$, integrate more like W_{12} above. Near-linear dependence between modes is captured with either $<\eta_i, \eta_j >$ or $< \dot{\eta}_i, \dot{\eta}_j >$, so we can define reachable and detectable volumes over the modal displacement or velocity subspaces alone. Furthermore, we saw, in the previous section, how to define a transform (balancing, scaling, or otherwise) that operates only on a certain part of the system. Therefore, we may partition the state first, then apply scaling transforms with a structure that preserves the partition, and compute a scalar controllability or observability measure with the associated order nxn partition of the $2n$x$2n$ full order of the Gramian. For example, suppose we cast the structural dynamics in the first order form:

$$\frac{d}{dt}\begin{bmatrix} \dot{\eta}(t) \\ \eta(t) \end{bmatrix} = \begin{bmatrix} -2Z\Omega & -\Omega^2 \\ I & 0 \end{bmatrix} \begin{bmatrix} \dot{\eta}(t) \\ \eta(t) \end{bmatrix} + \begin{bmatrix} \Phi^T B_a \\ 0 \end{bmatrix} u(t) \qquad 3\text{-}135$$

$$y(t) = C_v \Phi\, \dot{\eta}(t) + C_d \Phi\, \eta(t)\ . \qquad 3\text{-}136$$

Then for Gramians (control or observation) partitioned into equal dimensioned blocks,

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{bmatrix} \quad \begin{matrix} W \in R^{2n \times 2n} \\ W_{ij} \in R^{n \times n} \end{matrix} \qquad 3\text{-}137$$

the W_{11} block involves the product of rates $<\dot{\eta}, \dot{\eta}>$; the W_{12} block, the cross products $<\eta, \dot{\eta}>$; and the W_{22} block, the product of displacements $<\eta, \eta>$. If a non-singular transform has the structure imposed

$$T = \begin{bmatrix} T_{11} & 0 \\ 0 & I_{nxn} \end{bmatrix} \qquad 3\text{-}138$$

then the control and observation Gramians are transformed:

$$W_c = \begin{bmatrix} T_{11}^{-1} W_{11} T_{11}^{-T} & T_{11}^{-1} W_{12} \\ W_{12}^{T} T_{11}^{-T} & W_{22} \end{bmatrix} \qquad 3\text{-}139$$

$$W_o = \begin{bmatrix} T_{11}^{T} W_{11} T_{11} & T_{11}^{T} W_{12} \\ W_{12}^{T} T_{11} & W_{22} \end{bmatrix}. \qquad 3\text{-}140$$

We may construct T_{11} with the same method and for the same scaling objectives as were reviewed originally for the full state space using the [1,1] partition in the same manner as the full Gramian. The reachable and detectable volumes may be defined by the eigenstructure of this partition alone.

To summarize, the order-*n* partitioning may be entirely justified by defining the reachable and detectable volumes in terms of modal displacement or velocity alone, or some other combination that equally represents each mode in the partitioning. Such a definition is reasonable for lightly damped oscillatory systems, since the phase difference in rate and velocity signals is always in a narrow range: close to $\pi/2$ for all modes. Balanced coordinates are a special case that makes such a partitioning transparent to the scalar metrics selected in section 3.2.2 since each partition of the Gramian has nearly equal eigenvalues. As will be seen in the development of algorithms in Chapter 4, this partitioning is useful because the most expensive step in the recursion is solving a symmetric eigenvalue problem of the order of the reachable/detectable volume.

3.4 Optimization Problem Statement

After reviewing principles of reliability analysis, measures of controllability and observability, and model definition and scaling, we are almost in a position to pose the sensor/actuator optimization problem we wish to solve. The last issue to resolve is

selection of the independent variables, ξ. A natural and common idea is for ξ to represent spatial coordinates, since that is the form of the desired solution. For simple spatial domains, such as a simply-supported beam of length l with normal mode shapes, $\phi_i(\xi) = \sin(i\pi\xi/l)$, for $\xi \in [0,l]$, this is easy to do. For complex structures where a simple spatial dimension can be identified, such as the spanwise positioning of control surfaces for aeroelastic vibration suppression on an aircraft wing, successful algorithms can be operated with numerical approximations to mode shapes [25]. Consider the computation involved in determining the finite-dimension modal influence matrix, $[\Phi^T B_a]$. Assume the influence of control component j is separable in space and time and can be written $\beta(\xi)u_j(t)$, which captures a wide range of potential actuators since separability is a reasonable assumption for most hardware. Then for modes indexed by i, the i,j entry in $[\Phi^T B_a]$ is represented by

$$\left[\Phi^T B_a\right]_{i,j} = \int \phi_i(\xi)\beta_j(\xi)\, d\xi \qquad 3\text{-}141$$

where integration is over the spatial domain. Computation of these integrals is non trivial, and is carried out by structural analysis codes such as NASTRAN. Note that the spatial influence $\beta(\xi)$ is not restricted to point devices, although the integral simplifies nicely in practice to the mode shape or mode slope (curl) for point forces or point torques, respectively, at the modeling nodes. If ξ is the dependent variable, then in an algorithm we will want to compute the spatial gradient, which, for component k, is by Leibniz's rule:

$$\frac{\partial}{\partial \xi_k}\left[\Phi^T B_a\right]_{i,j} = \int \frac{\partial}{\partial \xi_k}\left[\phi_i(\xi)\beta_j(\xi)\right] d\xi \qquad 3\text{-}142$$

and is as difficult as computing $[\Phi^T B_a]$ in the first place. It is worthwhile to avoid inserting this kind of computation into a recursive algorithm. Furthermore, the architecture of the kind of spacecraft that present control structure interaction problems is notoriously non-continuous, defying geometric simplification, as can be seen by inspection of the ACOSS example used in Chapter 5. Even if the spatial gradient were computable, exploitation in an algorithm remains a demanding problem.

An alternative is to restrict ξ to a finite set of points, say $\xi \in \Xi$, $\Xi = \{\xi_1, \cdots \xi_N\}$. We may then re-define the independent variables, which were the vector-valued positions

themselves, to a set of integer scalars representing how many devices are at one site. The set Ξ is now $\Xi = \{ \xi_1, \cdots \xi_N \}$, with $0 \leq \xi_i$, ξ_i integer (or 0,1 if sites are restricted to one device each), and the spatial position is represented entirely by the index i. The number of potential sites, N, may be chosen as large as is necessary and computationally affordable. With failure modes, **f**, enumerated over a prescribed set, say $\mathbf{f} \in F$, as defined in section 3.1, the problem statement is:

$$\max_{\Xi} J[W(\xi), f] \qquad 3\text{-}143$$

subject to an integer constraint M, on the total number of devices:

$$\sum_{i=1}^{N} \xi_i \leq M \ . \qquad 3\text{-}145$$

J is one of the scalar measures defined by equations 3-35, 3-46, or 3-49, and is a function of the eigenvalues of Gramian $W(\xi)$, which in turn is a linear function of integer 'selection variables' ,ξ_i, and the Gramians of each site i alone, W_i,:

$$W(\xi) = \sum_{i=1}^{N} \xi_i W_i \ . \qquad 3\text{-}145$$

Failure weighting, as set down in section 3.1, is computed either as the expected value **or** the worst case:

$$J(W,\mathbf{f}) \cong \begin{cases} E[J] = \sum_{j=1}^{N_f} Pr(f_j)\, J(W,f_j) \\ \min_{F} J(W,\mathbf{f}) \qquad \mathbf{f} \in F \end{cases} \qquad 3\text{-}146$$

where the probability functions of the failure states are determined by the reliability model.

Even without failures, this is a difficult combinatorial problem, and is in the form of a non-linear integer program. Linear integer programs are known to be intrinsically difficult to solve exactly, but may be amenable to heuristic rules and approximate methods. A useful starting point for linear integer programs is the solution of the 'relaxation' problem

which is defined by removing the integer constraint on the independent variables. In the linear case exact solution methods exist, such as the Simplex algorithm. This idea is useful here as well.

We define a 'relaxation', without failure weightings, as

$$\max_{\xi} J[W(\xi)] \qquad \text{with } W(\xi) = \sum_{i=1}^{N} \xi_i W_i \tag{3-147}$$

$\xi \in R^N$ (continuous) and subject to the constraints

$$\sum_{i=1}^{N} \xi_i \le M, \qquad 0 \le \xi_i \le m_i \ . \tag{3-148}$$

Clearly the individual constraints must satisfy $m_i \le M$, and $m_i = 1$ is an important special case. This relaxation has a physical interpretation. Suppose, using actuators for example, the columns of B are computed for all N possible sites and are scaled to represent the influence of some 'unit' device. Then ξ_i represents a 'power gain' of each device, since for a linear gain, α_i, on each device the Gramian for site i is $\alpha_i^2 W_i$, where W_i is the Gramian computed with column $\mathbf{b}_i$ (of B) alone. The solution then represents the optimal sizing of actuator authority subject to a constraint on total power, as represented by the square of the linear gain. The selection of a 'power' weighting is rationalized in two ways: Like the selection of the L_2 functional norm for optimal control and estimation, the power gain is expedient for computation and algorithms. Also, the exponent, 2, is a better representation of the real hardware marginal costs in terms of weight and (electrical) power than a linear (exponent 1) model. For sensors, the analogy is not quite so direct: ξ_i may be thought of as representing sensitivity at a required quality. The relaxed problem has the useful property of no local maxima, which implies, like the Simplex algorithm for linear programs, the ability to compute the global optimum.

The scalar measures of the Gramian W are concave and homogeneous [36], so for any representative W_1 and W_2,

$$J[(1-\alpha)W_1 + \alpha W_2] \ge (1-\alpha)J[W_1] + \alpha J[W_2], \qquad \alpha \in [0,1] \ . \tag{3-149}$$

Now the set $W(\xi)$ is all possible linear combinations of the Gramians of each site, $W(i)$, subject to the constraints on ξ, $0 \le \xi(i) \le m_i$ and $\Sigma\xi(i) \le M$; and this set is convex. That is, if $W_1 = W(\xi_1)$ and $W_2 = W(\xi_2)$, where ξ_1 and ξ_2 are admissible (satisfy the constraints), then the convex combination of ξ_1 and ξ_2, $\xi_3 = (1-\alpha)\xi_1 + \alpha\xi_2$, $\alpha \in [0,1]$ also satisfies the constraints,

$$\sum_{i=1}^{N} \xi_3(i) = \sum_{i=1}^{N}\left[(1-\alpha)\xi_1(i) + \alpha\xi_2(i)\right] = (1-\alpha)\sum_{i=1}^{N}\xi_1(i) + \alpha\sum_{i=1}^{N}\xi_2(i) \le (1-\alpha)M + \alpha M = M$$

$$\sum_{i=1}^{N} \xi_3(i) \le M \qquad 3\text{-}150$$

$$\xi_3(i) = (1-\alpha)\xi_1(i) + \alpha\xi_2(i) \le (1-\alpha)m_i + \alpha m_i = m_i, \qquad i=[1,N]$$

$$0 \le \xi_3(i) \le m_i \; . \qquad 3\text{-}151$$

$W(\xi_3)$ is then a feasible Gramian. For notational economy, we will write $J(\xi)$ in lieu of $J[W(\xi)]$, since the underlying set of individual site Gramians, $W(i)$, do not enter directly into the argument, but form the 'instance' of the problem. We may now call upon a basic result of optimization theory: local maximal points of concave functions on convex sets are globally maximal [56]. Suppose ξ^* is a point at which $J(\xi^*)$ is globally (over all admissible ξ) maximal. Then assume we identify a local maximum at ξ_o in some neighborhood around ξ_o. On the line from ξ^* to ξ_o, concavity of the measure implies

$$J[(1-\alpha)\xi^* + \alpha\xi_o] \ge (1-\alpha)J[\xi^*] + \alpha J[\xi_o], \quad \alpha \in [0,1] \qquad 3\text{-}152$$

which, since $J[\xi^*] \ge J[\xi_o]$, contradicts the assumption ξ_o is a local maximum.

Finally, we note that Ξ, the feasible region in R^N for ξ, is compact (simple inequality boundaries), and J is a bounded, continuous function of the eigenvalues, λ, of $W(\xi)$. The eigenvalues, $\lambda[W(\xi)]$, are bounded continuous functions of parameters, ξ [57], so J is a bounded continuous function of ξ as well. We know from analysis, then, that $J(\xi)$ attains its maximum (and minimum) at some $\xi^* \in \Xi$. There is no guarantee of a unique maximum, however.

Returning to the original problem, with ξ_i constrained to the non-negative integers and a set of failure modes enumerated, we lose the convenience of continuity which

eliminated local maxima. Furthermore, to even enumerate the failures and state their probability distribution requires specifying the specific number of 'active' ($\xi_i > 0$) devices in the configuration, a number that is only constrained from 1 to N in the relaxed problem. It does not seem feasible to define a failure-weighted relaxation. However, the solution of the relaxed problem is useful for setting an upper bound on the solution of the no-failures integer problem, and for locating a neighborhood for the failure-weighted problem we are actually interested in. Algorithms for solving the relaxed problem and approximating the integer problems are the subject of the next Chapter.

Chapter 4. Algorithms

This chapter presents details of two algorithms: for the relaxed problem and for the combinatorial failure-optimum problem. Since the relaxed problem is continuous and convex, a straightforward application of a gradient-ascent approach worked well for the particular examples included in the next chapter. A conjugate-gradient modification accelerates terminal convergence, and some bookkeeping of active and inactive constraints is also required. With the additional constraints of integer device 'power', for each potential site (most often 0-1) and including failure combinations, the problem becomes purely combinatorial and is not known to be convex. The geometry of the relaxed problem along with the example results in Chapter 5 suggest the failure optimum solutions are not likely to be far from the relaxed solution, and a neighborhood search approach has been satisfactory. We start with the continuous formulation.

4.1 Relaxed Problem

At the end of Chapter 3, after settling modelling and scaling issues, we defined a continuous relaxed problem for N 'power variables', ξ, as:

$$\max_{\xi} J[W(\xi)], \qquad \xi \in R^N, \quad W(\xi) = \sum_{i=1}^{N} \xi_i W_i \tag{4-1}$$

subject to the total power, M, and individual, m_i, constraints

$$\sum_{i=1}^{N} \xi_i \le M, \qquad 0 \le \xi_i \le m_i . \tag{4-2}$$

The scalar metric J is computed from the eigenvalues of the Gramian W, and is either the minimum eigenvalue (J_1, eqn. 3-45), the inverse of the average of the eigenvalues (J_2, eqn. 3-47) of W^{-1}, or the geometric mean of the eigenvalues (J_3, eqn. 3-49).

4.1.1 Implemented Constraints

To pose a meaningful problem, some constraint on the size of ξ is necessary. The scalar, J, is homogeneous in ξ, $J(\xi_i W_i) = \xi_i J(W_i)\ \forall\ i$, and as will be seen in the

computation of the gradient, the unconstrained derivative of J with respect to each ξ_i is non-negative

$$\frac{d}{d\xi_i} J[W(\underline{\xi})] \geq 0 \quad \forall \ i \, . \tag{4-3}$$

The total power constraint is, therefore, always active, and we replace it with equality, so the constraints as actually implemented in the algorithm are

$$\sum_{i=1}^{N} \xi_i = M, \qquad 0 \leq \xi_i \leq m_i \, . \tag{4-4}$$

With one linear equality constraint, it is easy to set one component of $\underline{\xi}$, say ξ_d, as a dependent variable, and solve for it in terms of the remaining components and the total constraint:

$$\xi_d = M - \sum_{\substack{i=1 \\ i \neq d}}^{N} \xi_i \, . \tag{4-5}$$

The problem may be recast in N-1 dimensions and treated as unconstrained, except for the individual component inequality constraints that are at equality, which may be none or all for the non-zero components of $\underline{\xi}$. We partition the full (N) order $\underline{\xi}$ into three parts: the 'active' components within all of the inequality constraints, and two 'inactive' partitions for components at either the upper (m_i) or lower (0) inequality bounds:

$$\underline{\xi} = \left\{ \begin{bmatrix} \underline{\xi}_a \\ \underline{\xi}_{ul} \\ \underline{\xi}_{ll} \end{bmatrix} \begin{array}{l} N_a \text{ active components with } 0 < \xi_{a\,i} < m_i \\ N_{ul} \text{ inactive components at upper limit } \xi_{ul\,i} = m_i \\ N_{ll} \text{ inactive components at lower limit } \xi_{ll\,i} = 0 \, . \end{array} \right. \tag{4-6}$$

One of the components in the active partition is dependent to enforce the total power constraint, and the gradient computation is restricted to this subspace of dimension N_a - 1. The size of all three partitions varies during convergence, and any may have no members at all, although the sum of all components must equal the total constraint, M.

4.1.2 Steepest-Ascent Algorithm

The overall flow of the hill-climbing algorithm is diagrammed in Figure 4-1. The main steps are discussed in the following sections, not with the intent of providing a tutorial on classical parameter optimization, nor formal software documentation, but only to point out salient details so the results of this research may be repeated or applied to other problems. The implementation coding remained experimental, using the non-compiled "Control-C" instruction language; and we recommend any interested user apply some standard (compiling) programming language which allows easy matrix-vector manipulation. The only advanced function required is a symmetric eigenvalue-eigenvector routine. A reference for all of the techniques used in the relaxed algorithm is Luenberger [56].

4.1.2.1 Gradient Computation

The algorithm is principally a steepest-ascent approach modified with conjugate-gradient directions for terminal convergence. We state here the form of the gradient for the active partition, ξ_a, including the modification for a dependent component, and refer the interested reader to Appendix 1 for details. The primary linear algebra result called on is an expression for the first order perturbation of the eigenvalues of a symmetric matrix in terms of its eigenvectors.

For the minimum eigenvalue metric (J_1), the gradient is computed with at least $n^2(n+1)$ multiplies:

$$\nabla_{\xi} J_1 = \begin{bmatrix} v_1^T(W_1 - W_d)v_1 \\ \vdots \\ v_1^T(W_{Na} - W_d)v_1 \end{bmatrix} \qquad 4\text{-}7$$

where v_1 is the unit-norm eigenvector associated with the minimum eigenvalue, λ_1, of the Gramian $W(\xi)$. W_i is the Gramian of site i alone, and W_d is the Gramian of the site corresponding to the dependent component in ξ. Note that while not mathematically meaningful, a formal computation of the dependent component of the gradient is zero. Rather than removing ξ_d from ξ_a in the implementation, we include it and simply keep

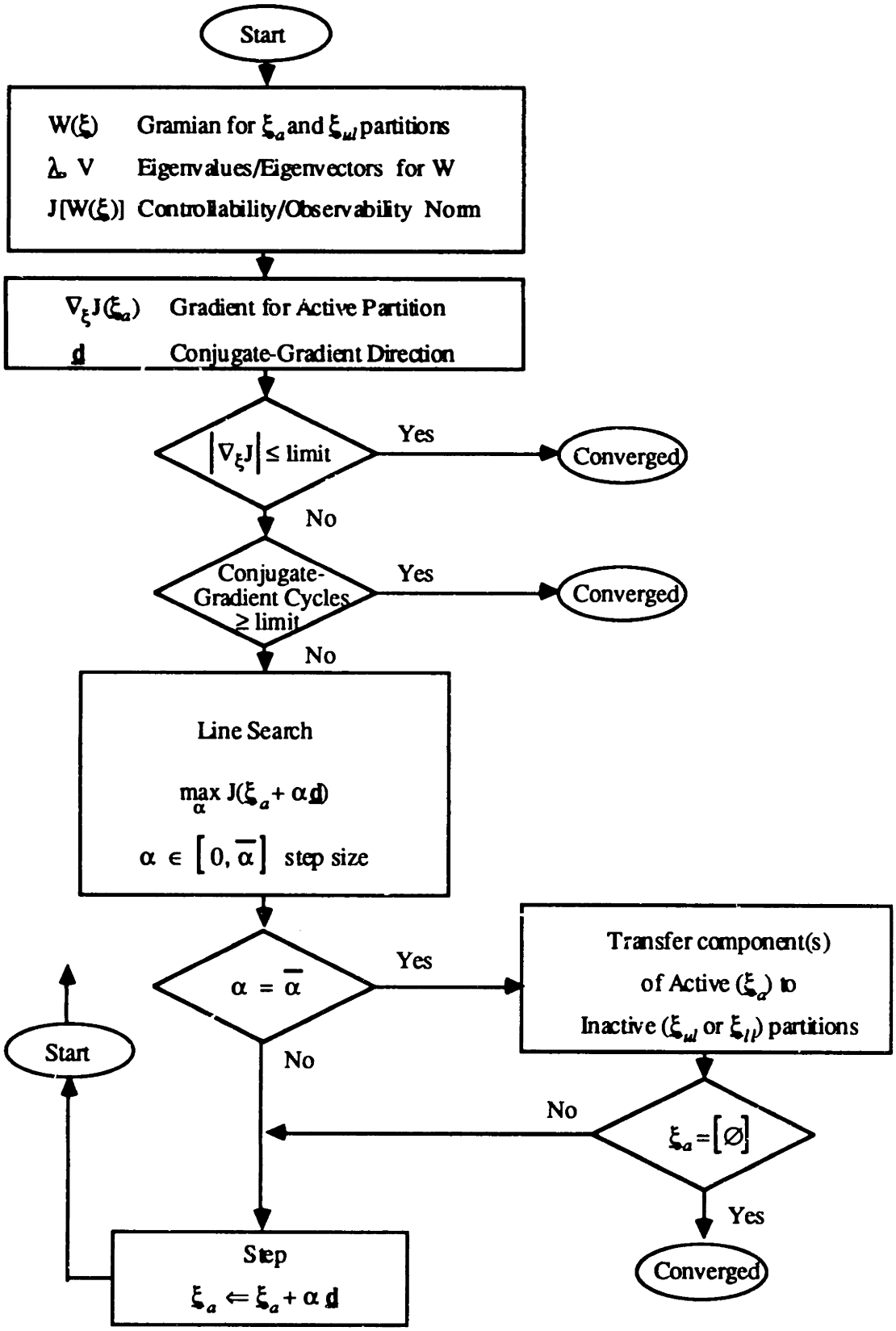


Figure 4-1A. Relaxed Algorithm.

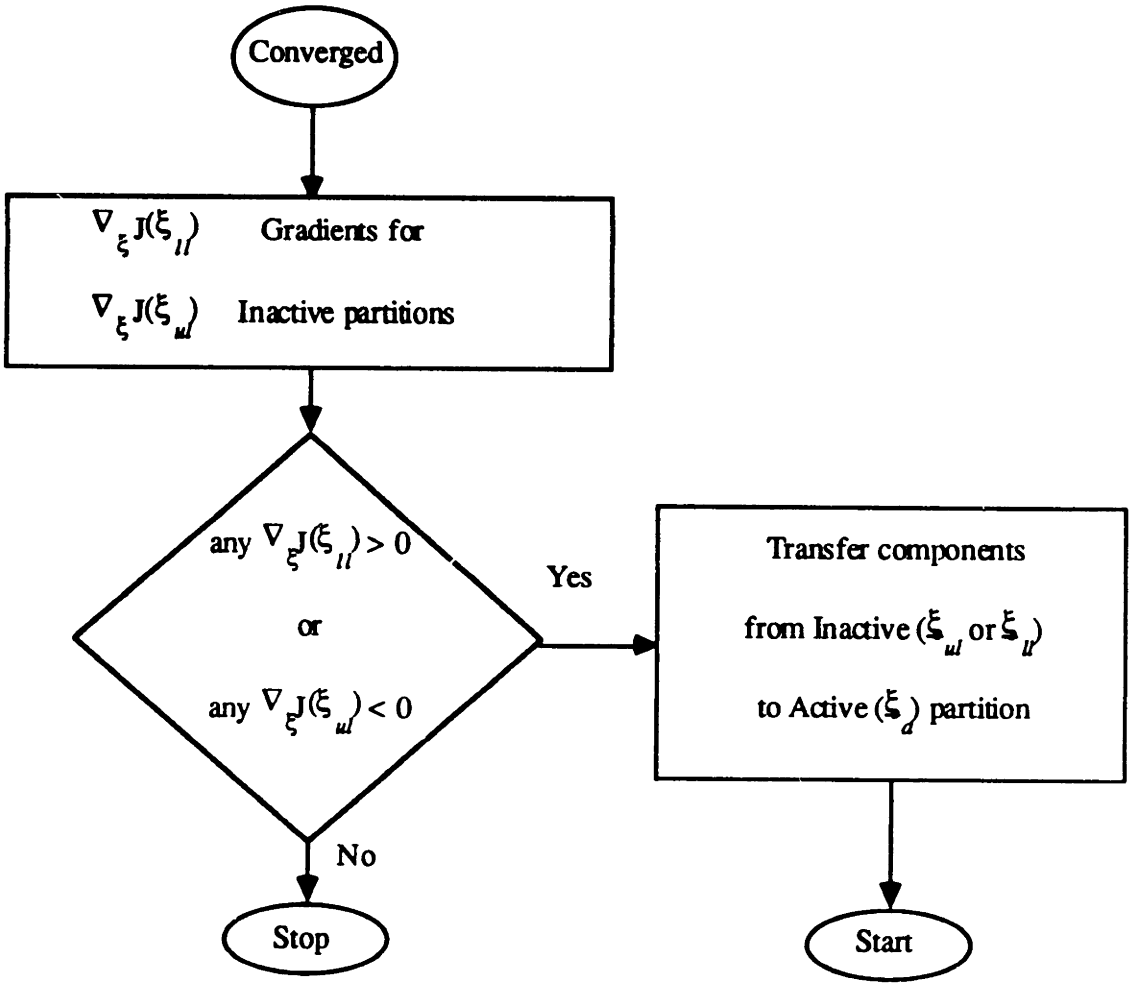


Figure 4-1B. Relaxed Algorithm.

track of the dependent component with a pointer. Since the other norms involve all n of the eigenvalues, λ_i, their expressions are more complex with at least an additional factor of n more multiplies, but share the basic form:

$$\nabla_\xi J_2 = \frac{J_2^2}{n} \begin{bmatrix} v_1^T(W_1 - W_d)v_1 & \cdots & v_n^T(W_1 - W_d)v_n \\ \vdots & & \vdots \\ v_1^T(W_{N_a} - W_d)v_1 & \cdots & v_n^T(W_{N_a} - W_d)v_n \end{bmatrix} \begin{bmatrix} \lambda_1^{-2} \\ \vdots \\ \lambda_n^{-2} \end{bmatrix} \quad 4\text{-}8$$

$$\nabla_{\xi} J_3 = \frac{J_3}{n} \begin{bmatrix} v_1^T(W_1 - W_d)v_1 & \cdots & v_n^T(W_1 - W_d)v_n \\ \vdots & & \vdots \\ v_1^T(W_{N_a} - W_d)v_1 & \cdots & v_n^T(W_{N_a} - W_d)v_n \end{bmatrix} \begin{bmatrix} \lambda_1^{-1} \\ \vdots \\ \lambda_n^{-1} \end{bmatrix} \quad 4\text{-}9$$

For all three cases, the unconstrained gradients may be obtained by deleting the W_d entry in the above expressions. Since all of the individual Gramians, W_i, and all the eigenvalues, λ_j, are non-negative definite, none of the components of the unconstrained gradient can be negative. The projection of the unconstrained gradient onto the plane of the total power constraint is useful in the algorithm if the active partition becomes completely depleted. If that occurs it becomes necessary to see if the projection for any of the inactive components points into the feasible region. If so, those components may be returned to the active list and, if not, the solution has been found at that particular corner. Otherwise we use the constrained form of the gradient (eqns. 4-8 and 4-9) in the algorithm and not the unconstrained form nor its projection.

If we were interested in the Hessian (matrix of second derivatives), we would need the second perturbation of the eigenvalues, which may be expressed as an expansion including the first perturbation of the eigenvectors. As shown in Appendix 1, each of the latter may be expressed in an orthogonal series of the other eigenvectors [34], but the computational burden has grown considerably over that of the gradient alone. We elected to try an algorithm without the Hessian, which worked, so a direct estimation or computation of the second derivatives was not pursued. The purpose of the Hessian is to accelerate convergence, especially near the solution. Another option for improving the terminal convergence of a steepest-ascent algorithm is the conjugate-directions method. Very little additional computation is required beyond that of the basic algorithm, and this modification was incorporated.

4.1.2.2 Conjugate-Gradient modification

The conjugate-directions method as implemented is a cycle of N_a -1 steps, starting with a pure steepest-ascent step. For this section only, we introduce subscript k as a counter for the steps within a cycle and $\mathbf{g}_k$ as a symbol for the gradient at step k. Vectors $\mathbf{g}$, ξ, and $\mathbf{d}$ lie in the N_a -1 dimensional space of the active partition without the dependent

variable. Employing the Polak-Ribiere approach, the conjugate gradient cycle is, as described in reference [56],

1) Start at ξ_0, compute $\mathbf{g}_0 = \nabla J(\xi_0)$ and set $\mathbf{d}_0 = \mathbf{g}_0$

2) For k = 1 to N_a - 1, do:

i) Line Search over α: $\max[J(\xi_{k-1} + \alpha\, \mathbf{d}_{k-1})]$ and set $\xi_k = \xi_{k-1} + \alpha \mathbf{d}_{k-1}$

ii) Compute $\mathbf{g}_k = \nabla J(\xi_k)$

iii) Set $\mathbf{d}_k = \mathbf{g}_k + \beta \mathbf{d}_{k-1}$, where $\beta = \dfrac{(\mathbf{g}_k - \mathbf{g}_{k-1})^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}}$ 4-10

3) Restart: at $k = N_a - 1$, replace ξ_0 with ξ_k, reset k = 0 and go to step 1)

The cycle starts whenever the line search terminates inside the feasible region, that is, no component of the active partition, including the dependent component, is at an inequality boundary. If during a cycle any component is driven into a boundary and removed from the active partition, the conjugate gradient cycle is restarted after the next steepest-ascent step remains within the feasible region. In the first several steps of the algorithm the line search almost always reaches an edge of the feasible region, and the conjugate gradient cycle does not start until most of the components of ξ have been placed into the inactive partitions. The Fletcher-Reeves form of the cycle was briefly experimented with, which has

$$\beta = \frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}} \qquad 4\text{-}11$$

and is otherwise identical. Convergence on this problem appeared to be superior with the Polak-Ribiere method, although a detailed comparison was not made. Polak-Ribiere's approach was employed for all of the examples.

4.1.2.3 Line Search

With a direction identified from the conjugate-gradient computation, the algorithm searches along that line for the maximum. Concavity of the metric helps here, insuring no spurious local maxima. Since the constraint plane is bounded, the interval to search is also

bounded. Because of the simple form of the equality constraint, it is expedient to compute the change in the dependent variable, $d\xi_a(d)$ from differential steps in the N_a -1 independent components, $d\xi_a(i)$, so with equation 4-5,

$$d\xi_a(d) = - \sum_{i=1,\, i\neq d}^{N_a} d\xi_a(i) . \qquad 4\text{-}12$$

Although the conjugate-gradient direction, **d**, is of dimension N_a -1, we may think of an augmented direction of dimension N_a which has a dependent component, **d**(d), constrained by equation 4-12, so we set

$$\mathbf{d}(d) = - \sum_{i=1,\, i\neq d}^{N_a} \mathbf{d}(i) . \qquad 4\text{-}13$$

This 'effective direction' component may replace the zero at the dependent variable's place in the conjugate-gradient direction for the purpose of the line search step only. It does not enter the update equation (4-10) for the next direction. The result is a search direction of the same dimension as the active partition, N_a, which lies entirely in the constraint plane of the total power. With search direction, **d**, augmented by the dependent component, we determine the maximum step length before encountering a boundary:

$$\bar{\alpha} = \max \alpha \Big| \quad 0 \le \xi_a(i) + \alpha \mathbf{d}(i) \le m_i , \quad i = 1, N_a . \qquad 4\text{-}14$$

The individual-component constraint values, m_i, are not necessarily equal and, if the fully-relaxed solution is sought, are each effectively just the total constraint, $m_i = M$. With the step size bounded, we search

$$\max_{\alpha} J(\alpha) \quad \text{with} \quad J(\alpha) = J\left[W(\xi_a) + \alpha\, W(\mathbf{d}_k)\right] \quad \text{and} \quad \alpha \in \left[0, \bar{\alpha}\right] . \qquad 4\text{-}15$$

Evaluation of $J(\alpha)$ over the interval requires repeated eigenvalue solutions. The method chosen for the search is three-point quadratic fit starting with J at both ends of the interval and the midpoint. The two points on either side of the quadratic fit maximum are chosen as the new interval, unless the maximum is computed to lie outside the original interval, then the corresponding half of the original interval is selected for the next step. The effect is to 'capture' the maximum between two of the search points, or else discover

that the maximum is at the boundary of the interval. The line search terminates when the proportional difference between all three values of J in the remaining interval is below a threshold. The α corresponding to the maximum J of the last three points is then selected as the result. Experience with example problems shows that the line search becomes quite flat as the global solution is approached, and a search threshold of 10^{-5} to 10^{-6}, proportional difference between any of the three values, is required.

The line search terminates with three possibilities, $\alpha = 0$, $\alpha = \bar{\alpha}$, or $0 < \alpha < \bar{\alpha}$. If $\alpha = \bar{\alpha}$, we know at least one component of the active partition has been driven into a constraint, and must be transferred to the appropriate inactive partition. If $\alpha = 0$, to the resolution of the line search, the solution has converged. This did not happen often in practice because another convergence criterion, the size of the gradient, would take effect first. Finally, if $0 < \alpha < \bar{\alpha}$, the algorithm continues for another step. If, after step(s) which did not converge within the line search interval, the line search for a step does converge within the interval, that step is taken as the first (pure steepest ascent) step in the conjugate gradient cycle and the conjugate directions modification is initiated on the next step.

4.1.2.4 Starting Points

Although the convex nature of the relaxed problem makes the starting point formally irrelevant, for computational economy it may be desirable to start close to the eventual solution. For small order problems, such as the simply-supported beam example in Chapter 5, a good starting point is to set all components equally to $\xi(i) = M/N$. Algorithm runs from such an initial distribution tend to spend the first several steps placing components into the zero inactive partition; after initial convergence very few components are later returned to the active partition. Larger problems, such as the ACOSS example, can be shortened by guessing a starting active partition, then setting equal each of those components. In this way the solution can converge quickly, but the guess is 'checked' at convergence by inspecting the gradient of all of the inactive components. Criteria for guessing are taken up in section 4.2.4 for the more expensive combinatorial algorithm.

4.1.2.5 Stopping Criteria.

There are two distinct criteria for stopping the algorithm and accepting the solution. First, the steepest-ascent steps for the active partition, as diagrammed in Figure

4-1A, must converge. Second, at convergence the gradient components of all the inactive components must point out of the feasible region. That is, for a gradient computed over all the partitions of ξ, those components in the lower-limit inactive partition, ξ_{ll}, must be non-positive and all the components in the upper inactive partition, ξ_{ul}, must be non-negative, as shown in Figure 4-1B. To prevent 'chatter', components are transferred to an inactive partition when they are driven into a constraint, but are not considered for removal from an inactive partition until the remaining active partition has converged. Convergence of the active partition is accepted by three possible tests, shown by the "Converged" label on Figure 4-1A. The primary test is if the magnitude of the gradient falls below some fraction of the first gradient computed in the algorithm. Ratios of 10^{-4} for starting points far from the expected solution to 10^{-2} for runs started near the solution have been successful. Since the active partition may converge several times in the course of one run, it is computationally expedient to initially try a coarse convergence test, then restart the algorithm near the solution with a finer test. Most of the algorithm's steps are spent sorting components into the inactive partitions if the starting point is an equal value in all components. The second possible 'good' convergence is if the line search drives the last two components (one independent, one dependent) of the active partition into the constraints. If none of the inactive components may be removed from their partition, this corner is the exact solution. A counter on the conjugate-gradient cycles prevents run-on of a poorly-converging instance, and three full cycles of N_a -1 steps were seldom reached. Slow convergence at that level could often be cured by a smaller line search criterion.

If the active partition ever became fully depleted, a slight modification was required to check the inactive partitions. Normally, one of the components in the active partition is identified as dependent, and the gradient for the inactive partitions may be computed, just as the gradient for the independent components of the active partition are. If there are no members in the active partition, we must first decide which components of the inactive partitions, if any, should be transferred to the active partition. To do this, we compute the unconstrained gradient, and project it into the total power constraint plane. The components of the projection are then used to identify members of the inactive partitions for transfer to the active partition.

Upon final convergence of the relaxed solution, the result may be interpreted as the optimal distribution of sensor/actuator authority. If the designer could vary the power

gain of actuators or the sensitivity of sensors, then the solution shows how to do that, possibly subject to the additional individual site constraints on the maximum proportion of the total weight (M) that may be accumulated at a single site (m_i). We do not believe, however, the controllability or observability norms used for the optimization are sharp enough indicators of potential closed loop performance to justify implementing such an exact interpretation of the result. We use the solution as a starting point for searching for the best possible distribution with respect to a failure-weighted norm under the additional constraint that the authority at each site is set at an integer level representing equal hardware units, with 0-1 an important special case.

4.2 Integer Solutions

With the addition of an integer constraint on the independent variables, $\xi_i = 0,1,2,..,$ m_i, the continuity and convexity that lead to the efficient solution of the relaxed problem are lost. Since the feasible region for ξ under integer constraints is a strict subset of the feasible region for the relaxed problem, the value of the optimal solution for the latter, say $J(\xi^*)$, represents an upper bound to the integer solution, $J(\xi^I)$. Furthermore, we hope the integer-constrained solution, ξ^I, will be close to ξ^*, where 'close' will be made more precise later. From linear integer programming we know that the solution may be quite far from the relaxed solution, but such instances are rare and the result of 'narrow' feasible regions. While all of the development to follow allows the integer constraint for each site, m_i, to range from 0 to the total constraint, M, we will concentrate in the application examples exclusively on the special case of $m_i = 1$, which implies M is the total number of effectors in the problem. This simplifies the algorithm since a site can then be moved between a '0' list and a '1' list. More than one unit at a site would require a bit more complicated bookkeeping.

4.2.1 Failure Weightings

In addition to introducing integer constraints to the relaxed problem, the other major modification is to compute the optimality index over failure modes. Section 3.4 set down two approaches to failure weightings: the expected value and the worst case over a specific set of failures, i. e.,

$$J(W,\mathbf{f}) = \begin{cases} E[J] = \sum_{j=1}^{N_f} Pr(f_j)\, J(W,f_j) \\ \min_{F} J(W,\mathbf{f}) \qquad \mathbf{f} \in F \end{cases} \tag{4-16}$$

A particular failure mode, f_j, is defined in Section 3.1 to be a unique list of the operational status of each active device (non-zero ξ_i), with 0 assigned for failure and 1 for operational, and are restricted to some feasible set, F. The probability of each failure mode, $Pr(f_j)$ results from the reliability model of the devices, and the controllability or observability index $J(W,f_j)$ is, for a particular failure mode, simply the sum of the Gramians of the operational devices in the suite:

$$W(\xi,\mathbf{f}_j) = \sum_{\substack{i=1,N \\ k=1,M}} \xi(i)\, \mathbf{f}_j(k)\, W_i \tag{4-17}$$

where index i for ξ ranges over all N potential sites and index k for f_j ranges over the *M* total devices in a particular configuration. Another representation that will be useful for estimating the change in J due to failure state f_j is to write $W(\xi,f_j)$ as the Gramian of the failures subtracted from the unfailed Gramian:

$$W(\xi,\mathbf{f}_j) = W(\xi) + \Delta W \tag{4-18}$$

$$\Delta W = -\sum_{\substack{i=1,N \\ k=1,M}} \xi(i)\, [1 - \mathbf{f}_j(k)]\, W_i\,. \tag{4-19}$$

The linearity of the Gramian makes computation of $W(\xi,f_j)$ trivial, but the computation of $J(W,f_j)$ requires the eigenvalues of $W(\xi,f_j)$ for every failure mode, f_j, since J is not linear in the component contributions, W_i. With the computation of $J(W,\mathbf{f})$ set down as an optimization parameter, we need a method to look for improvement, since the continuous gradient of the relaxed problem is no longer available.

4.2.2 k-Exchange Neighborhood Search

A combinatorial analogy to steepest ascent is the 'k-exchange' neighborhood search [58]. The 'k-exchange' neighborhood is defined as all combinations of ξ_i that differ from some nominal ξ_o in at most k entries, and those k range over a larger, possibly global, neighborhood. All members of the 'k-exchange' neighborhood are enumerated, the

optimization parameter computed, and the maximum identified. If the maximum is an improvement over the nominal, the maximum is adopted as the new nominal and the cycle repeated. If no improvement can be located in the local search, then the solution is said to be 'k-exchange' optimal. This process will not ensure, however, that the global optimum is ever reached, unless k is the same dimension as the solution. In our problem, where the number of potential sites (global dimension) is large-485 for the Draper Model No. 2 examples- and the desired number of effectors is relatively small- on the order of 10- then 'k-exchange' for any k > 1 becomes prohibitive. The k=1 neighborhood for $M = 10$ retained sites out of 485 may be as large as 4750, but the k=2 neighborhood has up to 20,263,500 members, and each member may still require failure enumeration. We use in an algorithm, therefore, only a '1-exchange' search.

We write the result of exchanging site i for site j (assuming $\xi(i) \neq 0$) as

$$W(\xi) = W(\xi_o) + \Delta W \qquad 4\text{-}20$$

$$\Delta W = W_j - W_i \ . \qquad 4\text{-}21$$

This structure, along with the similar representation for failures, is motivated by the desire to estimate, rather than compute directly, the eigenvalues of the Gramian resulting from failures or exchanges.

4.2.3 Eigenvalue Estimation for Exchange and Failures.

After summing the perturbation to the nominal Gramian, from either failures, an exchange, or a combination of both, computing the optimization norm, J(W), requires computing the eigenvalues of W. This is the expensive computation in the algorithm, and we would like to minimize the number of times it is required. If we have the eigenvalues, λ_o, and the orthonormal eigenvectors, V_o, of some nominal $W(\xi_o)$, and we are interested in the eigenvalues of W, λ, where $W = W(\xi_o) + \Delta W$, then, to first order,

$$\lambda \approx \lambda_o + V_o^T[\Delta W]V_o \ . \qquad 4\text{-}22$$

Appendix A illustrates the steps involved in this well-known result. For all n eigenvalues, a total of $n^2 + n$ multiplies and n additions are required, in contrast to approximately $2/3\ n^3$ operations [59] for the eigenvalues alone with a symmetric QR algorithm or about $5n^3$ if the eigenvectors are computed as well. This economy allows us, after the additional overhead of the eigenvector computation, to compute the failure-weighted measure and estimate the

effect of an exchange cheaper and faster than simply computing the eigenvalues directly at each stage. Since the 'step size' of the failure or exchange is fixed, there is no motivation for linearizing $J(\lambda)$, as is done in the gradient computation. $J(\lambda)$ is computed directly each time.

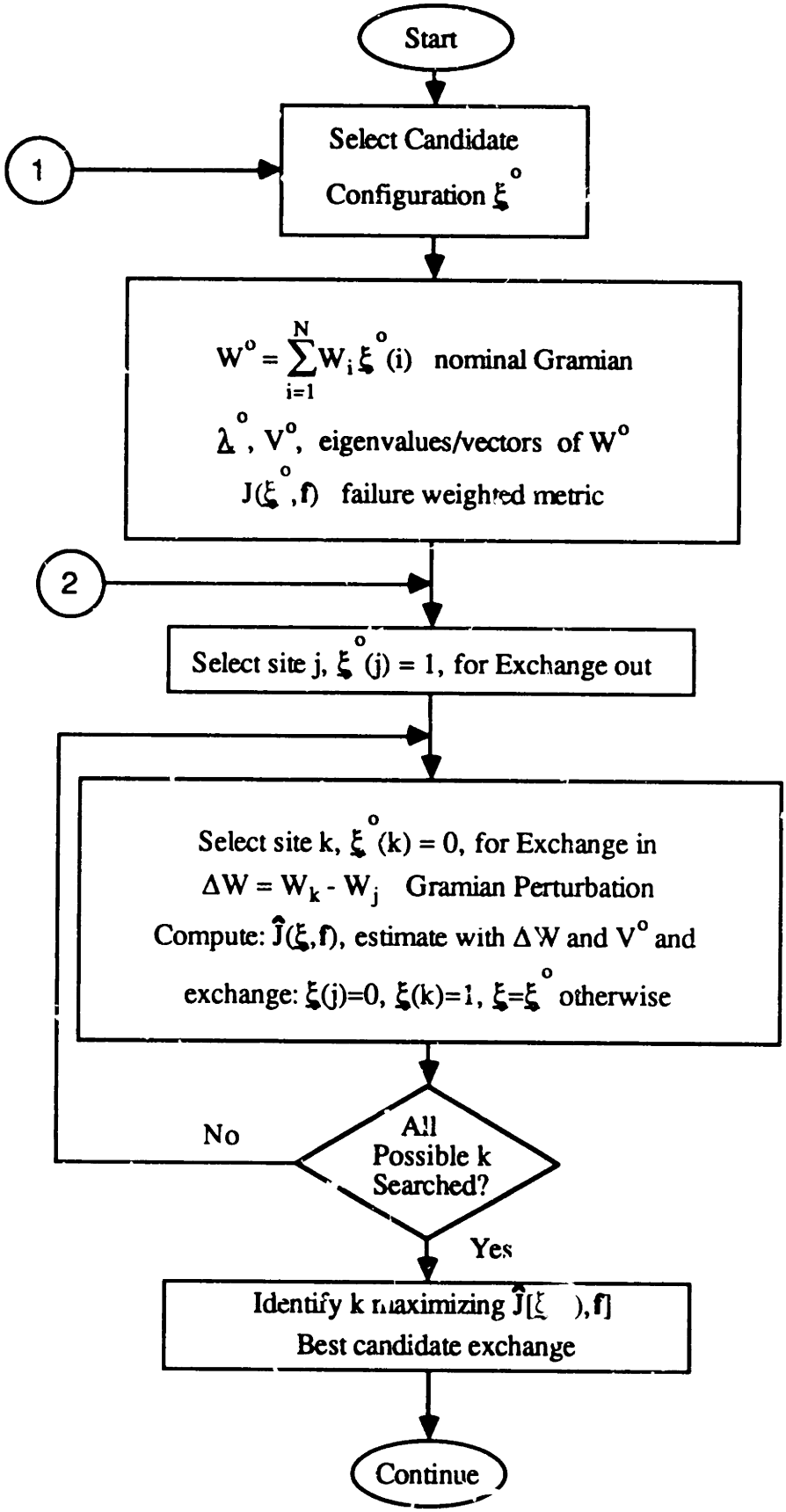
Norm $J(\xi)$ is concave in the relaxed solution, so we expect the '1-exchange' search to tc turn up more candidates for improvement of J than are correct. We modify the search to accept a new solution only if J computed with the actual eigenvalues is an improvement, but search over J computed with the estimated eigenvalues to identify the strongest candidate. The combinatorial algorithm may now be assembled with these pieces: Figure 4-2 shows the sequence. Additional details required are to specify (and limit) the global neighborhood and the stopping criteria.

4.2.4 Staring Points and the Global Neighborhood

Even with the computational savings of estimating the result of failures and exchanges with first-order eigenvalue perturbations, the '1-exchange' search can still be an expensive procedure, and examples with the Draper Model No. 2 required up to 10 times more run-time than to compute the relaxed solution. Two methods of mitigating the computational requirements are starting the search near the eventual solution (hence the relaxed solution), and restricting the number of candidates considered for a specific exchange to some fraction of the global dimension. Both techniques are heuristic.

After successful convergence of the relaxed algorithm, the solution is partitioned into three parts. For an individual upper limit constraint $m_i \leq 1$, the members of these partitions satisfy

$$\xi = \left\{ \begin{bmatrix} \xi_a \\ \xi_{ul} \\ \xi_{ll} \end{bmatrix} \begin{array}{l} N_a \text{ active components with } 0 < \xi_{a\,i} < 1 \\ N_{ul} \text{ inactive components at upper limit } \xi_{ul\,i} = 1 \\ N_{ll} \text{ inactive components at lower limit } \xi_{ll\,i} = 0 \ . \end{array} \right. \qquad 4\text{-}23$$


Start
1
Select Candidate
Configuration ξ^o
$W^o = \sum_{i=1}^{N} W_i \xi^o(i)$ nominal Gramian
λ^o, V^o, eigenvalues/vectors of W^o
$J(\xi^o, f)$ failure weighted metric
2
Select site j, $\xi^o(j) = 1$, for Exchange out
Select site k, $\xi^o(k) = 0$, for Exchange in
$\Delta W = W_k - W_j$ Gramian Perturbation
Compute: $\hat{J}(\xi, f)$, estimate with ΔW and V^o and
exchange: $\xi(j)=0$, $\xi(k)=1$, $\xi=\xi^o$ otherwise
No
All Possible k Searched?
Yes
Identify k maximizing $\hat{J}[\xi\ ), f]$
Best candidate exchange
Continue

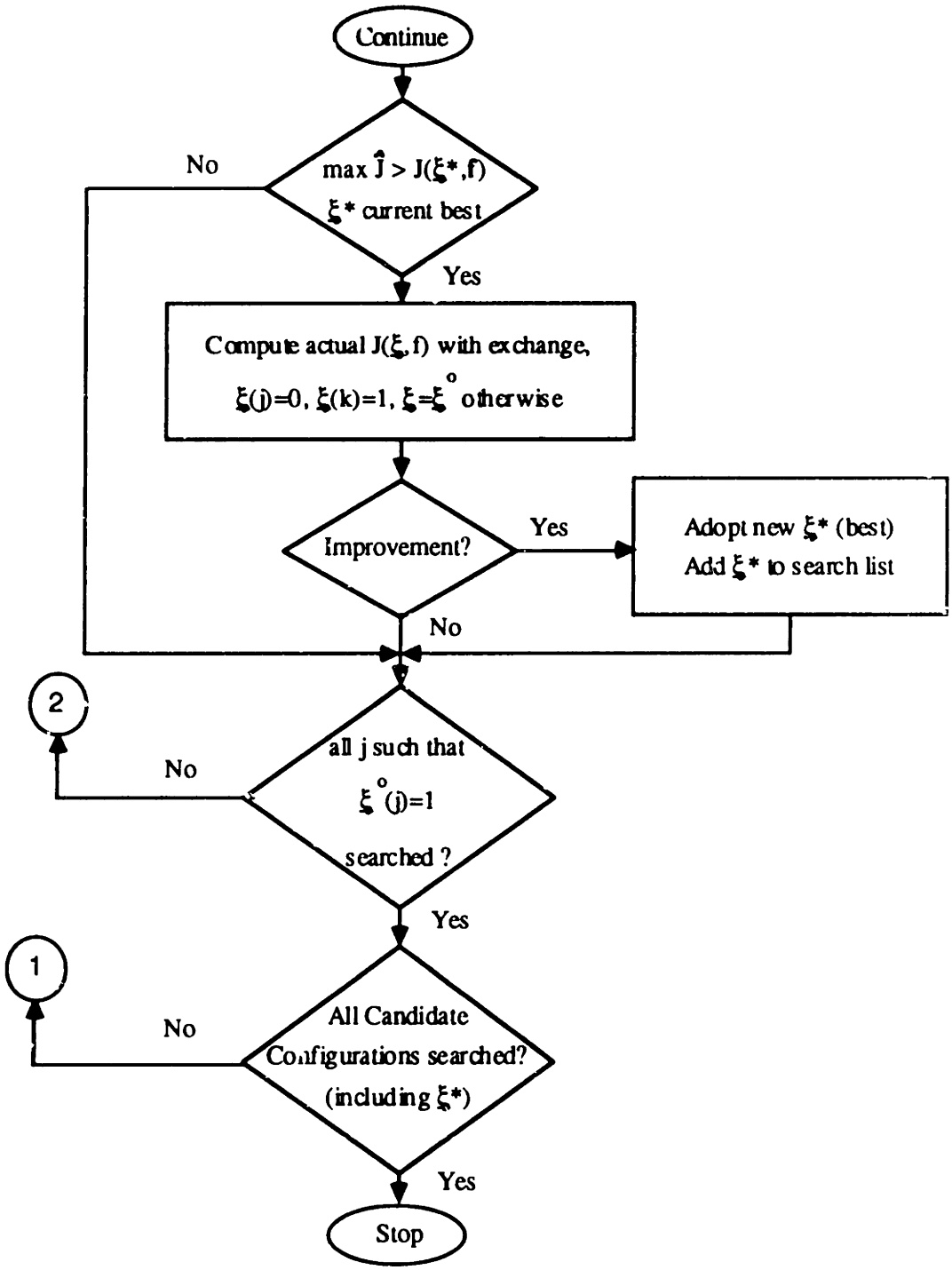


Figure 4-2. Combinatorial Algorithm

The upper and lower partitions clearly satisfy the integer constraints, so the question for starting points is how to lump the solution distributed over the active partition. There are $M - N_{ul}$ 'units' of authority to be placed in the N_{al} possible members of ξ_a, but lumped into entries of 0 or 1. Note $N_{al} > (M - N_{ul})$ strictly. In many instances N_a will be quite small, even zero, and nearby integer candidates may be seen by inspection. It is possible to enumerate all possible integer representations of the active partition and rank them in terms of 'distance' from the active partition. The 'distance' we would like is the resulting controllability/observability norm, but a 'distance' we can afford to compute for a large number of cases is the Euclidean geometric distance. The author wrote an algorithm that sequentially placed 0 or 1 in the N_a members of the active partition, and checked, via a least-squares fit of the places remaining, whether the geometric distance between the developing integer representation and the relaxed active partition could remain within an error bound. The closest candidates were taken as starting points for the failure-weighted algorithm.

Another approach for managing the computational burden is to restrict the search in the '1-exchange' step to less than the global dimension. This is motivated by the fact that many potential sites may be weak, and not likely to be included in any selection. Clearly we want to include those members of the active partition that were rounded to 0 in the integer representation, but this may not leave a very large list. One method of ranking the remaining sites is the controllability/observability index of each site alone. Experience with the Draper Model No. 2 example has shown no instances in which a site in the bottom half of individual rankings remains in the converged relaxed-solution. This method is potentially flawed, because the index for any site may be driven to zero if a single mode is unobservable or uncontrollable from that site, even though all the others might be strongly affected. Another possible ranking is the maximum eigenvalue, since it represents influence in the single strongest direction. Again, experience has shown no instances where a site ranked in the bottom half by maximum-eigenvalue remains in the solution. A third possibility is to inspect the components of the gradient projected onto the constraint plane after convergence, and eliminate those components in the zero inactive partition, ξ_{il}, whose gradient is most negative. Finally, all of these restrictions may be hedged by checking a '1-exchange' over the global list after the combinatorial search is believed to be concluded.

4.2.5 Stopping Criteria

Unlike the relaxed algorithm, the combinatorial search has a simple terminal criterion: When no improvement can be found over a '1-exchange' search, the algorithm stops. A check can be conducted with the optimal index of the relaxed solution, $J(\xi^*)$, to see the degradation induced by the integer constraints and/or failures. If the failure weighting is the 'worst case', then the relaxed metric can be linearly attenuated by the proportional failure level (e. g., if worse case is 2 failures out of 5, attenuate to 0.6 of initial value) to compare against the best possible performance without integer constraints at the lower level of total authority. Admittedly, it is hard to calibrate what is an acceptable level of degradation. It seems desirable that the performance should fall off proportionally slower than the number of devices fail, and this was commonly observed.

4.3 Overall Sequence

We have now developed all the pieces of the overall sensor/actuator selection process and may review them all as steps in a larger algorithm or process. Table 4-1 illustrates the sequence, referencing the sections in which each piece was developed. Note that the process is certainly not an automatic one, but it has reduced a difficult combinatorial optimization problem to a sequence of manageable and verifiable stages, each with an eye towards computational economy.

4.4 Computational Points

The single most important numerical requirement in this work is the ability to compute the truncated-horizon Gramian. The method chosen, as noted in reference [44] and Section 3.2.3 for 'recoverable' W_c^*, is to compute the transition matrix, $\Psi(t) = \exp(A_H t)$, for the Hamiltonian system:

$$A_H = \begin{bmatrix} -A & BR_u^{-1}B^T \\ 0 & A^T \end{bmatrix}. \qquad 4\text{-}24$$

With the partitions,

$$\Psi(t) = e^{A_H t} = \begin{bmatrix} \Psi_{11}(t) & \Psi_{12}(t) \\ \Psi_{21}(t) & \Psi_{22}(t) \end{bmatrix} \qquad 4\text{-}25$$

Table 4-1. Overall Sequence.

Step	Operation	Section
1.	Identify signal path for modeling and scaling. Disturbance to regulated variables, y(s)/w(s)	
2.	Compute σ_i and rank modes, select truncation	3.3.3
3.	Select time horizon and full ($2n$) or half (n) order reachable/detectable volume dimension	3.3.5
4.	Scale state basis	3.3.2
5.	Define relative device scalings and any dynamic input/output filters, U(s) and/or Y(s)	3.3.4
6.	Compute and store individual Gramians, W_i	3.2.3, 4.4
7.	Select optimization parameter, J(W)	3.2.3
8.	Specify constraints: total authority M, and individual site, m_i	4.1.1
9.	Compute relaxed solution	4.1
10.	Identify neighboring integer initial points	4.2.4
11.	Enumerate failure modes and select failure-weighted index	3.1, 4.2.1
12.	Specify search list (global or restricted)	4.2.4
13.	'1-Exchange' combinatorial search for failure-optimal solution	4.2.2, 4.2.3

the (controllability) Gramian may be computed:

$$W_c(t) = \int_0^t e^{A\tau} B R_u^{-1} B^T e^{A^T\tau} d\tau = \Psi_{22}^T(t)\,\Psi_{12}(t) \; . \qquad 4\text{-}26$$

The observability Gramian may be computed with (A^T,C^T) in place of (A,B). Since the Hamiltonian system has symmetric eigenvalues about the imaginary axis, there is some upper limit for the time interval, t, beyond which $\exp(A_H t)$ will not numerically converge. For the lightly damped modes over time intervals which remain short compared to the damping decay envelope, this has not been noticed as a problem. With the introduction of input and output frequency weightings, however, which may have real poles about the same radius from the complex origin as the structural modes, we may get stuck. A resolution, noted in reference [60] is to cut the time horizon in half and exploit the relations,

$$W_c(2t) = W_c(t) + e^{At}\, W_c(t)\, e^{A^T t} . \qquad 4\text{-}27$$

This may be done several times. Again, substituting A^T for A gives the relation for the observability Gramian. These are noted to be similar to the equations for the discrete time state covariance or estimate error covariance propagation.

Gawronski and Juang [61] recently developed a representation of the truncated time interval (and truncated frequency interval) Gramians, which for the former, requires only the algebraic Lyapunov equation solution and the state transition over the interval. While this appears a considerably improvement to the method used in this thesis, the author became aware of it too late to experiment with it.

Another, more minor, point has to do with the total volume metric, J_3. While we express

$$J_3(W) = \sqrt[n]{|W|} = \left[\prod_{i=1}^{n} \lambda_i\right]^{\frac{1}{n}} \qquad 4\text{-}28$$

if we were to actually compute it this way, we may find an underflow or overflow of the product of the eigenvalues. Taking the root of each eigenvalue first, then computing the product alleviates this difficulty:

$$J_3(W) = \prod_{i=1}^{n} \left[\lambda_i^{\frac{1}{n}}\right] . \qquad 4\text{-}29$$

The price paid, however, is increasing the number of n^{th} root (logarithm) calculations by a factor of n, and is worth avoiding if the condition of the Gramian allows it.

The numerical properties of the overall sensor/actuator selection algorithms have not been investigated, and an understanding of the limitations imposed would be a valuable addition to this work. Van-Loan's paper [60] on computing the matrix exponential integrals has a wealth of material in this direction.

Chapter 5. Application Examples

In this chapter we exercise the ideas developed in Chapter 3 and the algorithms of Chapter 4 on two flexible structural systems. The first, a simply-supported beam, is chosen for its combination of obvious modal dynamics and non-obvious optimal effector placements, while the second, the Draper Model No. 2 from the ACOSS program, was selected as a representation of the large-scale system class that motivated this research. Although at opposite extremes in terms of complexity, results from the two examples show two common themes of practical engineering significance. From among a large potential number of sites, the continuous relaxed solutions tends to converge toward relatively few sites, even though it is certainly feasible to distribute authority over all sites. Second, the failure-weighted search algorithm tends to stop at a distribution quite close to that resulting from the relaxed solution, indicating the relaxed solution is a good guess for the more computationally expensive combinatorial problem.

5.1 Lateral Vibration of a Beam

The simply-supported (pinned-pinned) Bernoulli-Euler beam model is useful for illustrating the controllability and observability programming algorithms for two reasons. The first reason is that the solution to the governing PDE is easily expressed analytically as an expansion of normal modes, and the second is that the best placement of sensors and actuators is not obvious for these boundary conditions. The latter is in contrast to a beam with a free end; the mode shapes are all at their maximum at the free end, making the placement of (lateral translational) devices trivial, without considering failures.

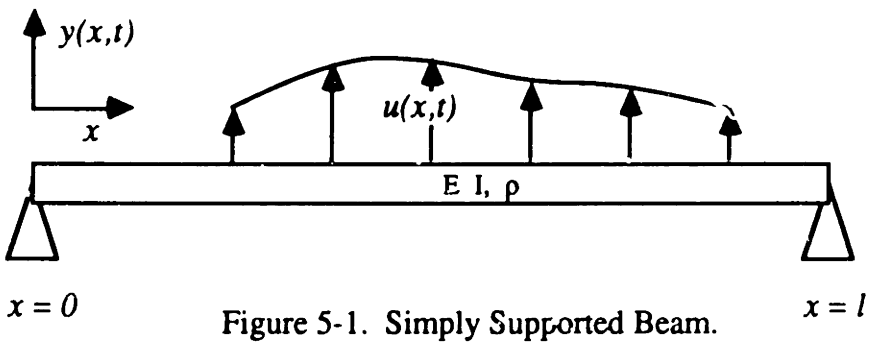


Figure 5-1. Simply Supported Beam.

5.1.1 Modeling

Figure 5-1 is a sketch of a beam of length l and mass per unit length ρ. We are interested in the lateral deflection, y, as a function of time and spatial dimension, x, and distributed forcing, $u(x,t)$. With the material elasticity modeled as usual by Young's modulus E and the cross section moment of inertia I, both constant for this example, the PDE for $y(x,t)$ is

$$[EI\ y(x,t)'']'' + u(x,t) = \rho\ \ddot{y}(x,t) \qquad 5\text{-}1$$

where y' represents the spatial derivative with respect to x, and $\dot{y}$ the temporal derivative with respect to t. The simply-supported ends impose the boundary conditions:

$$\begin{array}{ll} y(0,t) = y(l,t) = 0 & \text{deflections} \\ y''(0,t) = y''(l,t) = 0 & \text{moments} \end{array} \qquad 5\text{-}2$$

Assuming the solution is separable in space and time, $y(x,t) = \Phi(x)\,\eta(t)$, the unforced case ($u = 0$) admits the eigenvalue problems:

$$\frac{EI}{\rho}\frac{\Phi^{IV}(x)}{\Phi(x)} = -\omega^2 = \frac{\ddot{\eta}(t)}{\eta(t)} \quad \begin{cases} \Phi^{IV}(x) + \beta^4\Phi(x) = 0 \qquad \text{with } \beta^4 = \dfrac{\omega^2\rho}{EI} & 5\text{-}3 \\ \ddot{\eta}(t) + \omega^2\eta(t) = 0 & 5\text{-}4 \end{cases}$$

From the general solution of the spatial equation, we apply the boundary conditions and obtain the characteristic equation which has a sequence of solutions:

$$\sin(\beta l) = 0 \qquad \beta_i = \frac{\pi i}{l}, \quad i = 1,2,\cdots \qquad 5\text{-}5$$

For each β_i, the solutions to the spatial equation are the mode shapes, which are orthogonal with respect to mass and stiffness on the interval $x \in [0,l]$ and can be normalized:

$$\phi_i(x) = \alpha \sin(\beta_i x) = \alpha \sin\left(\frac{\pi i}{l} x\right) \qquad 5\text{-}6$$

$$\left\langle \phi_i(x), \rho\phi_j(x) \right\rangle \triangleq \int_0^l \rho\phi_i(x)\,\phi_j(x)dx = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad \text{with } \alpha = \sqrt{\frac{2}{\rho l}} \qquad 5\text{-}7$$

$$\left\langle \phi_i(x), EI\phi_j^{IV}(x) \right\rangle = \begin{cases} \omega_j^2 & i = j \\ 0 & i \neq j \end{cases} \qquad 5\text{-}8$$

The corresponding temporal solutions at discrete natural frequencies include real parameters a_{ji} that may be determined from initial conditions, if desired:

$$\eta_i(t) = a_{1i}\sin(\omega_i t) + a_{2i}\cos(\omega_i t) \qquad 5\text{-}9$$

$$\omega_i = \beta_i^2\sqrt{\frac{EI}{\rho}} = \left(\frac{\pi i}{l}\right)^2\sqrt{\frac{EI}{\rho}} \qquad 5\text{-}10$$

Returning to the forced response, we assume we may expand the particular solution in an orthogonal series of the normal modes:

$$y(x,t) = \sum_{i=1}^{\infty}\phi_i(x)\eta_i(t)\,. \qquad 3\text{-}11$$

Substituting this into the governing PDE, (eqn. 5-1), taking the inner product with one mode $\phi_j(x)$ at a time, and exploiting the orthogonality relations (eqns. 5-7 and 5-8) results in an infinite sequence of ODEs:

$$\ddot{\eta}_i(t) + \omega_i^2\eta_i(t) = \left\langle \phi_i(x), u(x,t) \right\rangle \qquad i = 1,2,\cdots \qquad 5\text{-}12$$

If we also assume the forcing, $u(x,t)$, is separable in space and time, $u(x,t) = b(x)u(t)$, the inner product on the right hand side of eqn. 5-12 simplifies to the spatial domain only:

$$\left\langle \phi_i(x), u(x,t) \right\rangle = \left\langle \phi_i(x), b(x) \right\rangle u(t) \qquad i = 1,2,\cdots \qquad 5\text{-}13$$

If, furthermore, the spatial influence of the forcing is concentrated at m discrete points, each with a corresponding component in the vector input $\mathbf{u}(t)$, then we may idealize $b(x)$ as

a vector of Dirac delta functions, $b^T(x) = [\delta(x-x_1) \cdots \delta(x-x_m)]$, and the inner product on the right side is particularly easy to evaluate:

$$\left\langle \phi_i(x), b(x) \right\rangle = [\phi_i(x_1) \ldots \phi_i(x_m)] \quad . \tag{5-14}$$

Another class of forcing inputs of interest is the 'point moment' device, Figure 5-2, which may be idealized as the limit of a couple applied by point forces around x_0:

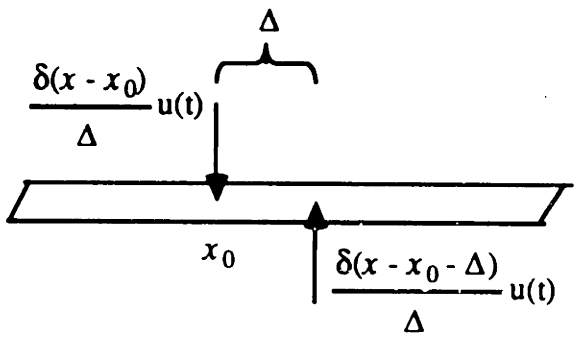


Figure 5-2. Point Moment

$$b(x) = \lim_{\Delta \to 0} \left[\frac{\delta(x - x_0 - \Delta) - \delta(x - x_0)}{\Delta} \right] \tag{5-15}$$

$$\left\langle \phi_j(x), b(x) \right\rangle = \lim_{\Delta \to 0} \left[\frac{\phi_j(x_0+\Delta) - \phi_j(x_0)}{\Delta} \right] = \frac{d}{dx}\phi_j(x)\Big|_{x=x_0} \tag{5-16}$$

The influence coefficients for torque inputs, then, are the mode slopes evaluated at the point of application.

Finally, we assume the dynamics may be adequately represented as a truncated modal expansion, introduce proportional damping, ζ_i, and write the model in the usual second order matrix form:

$$\begin{aligned} \ddot{\eta}(t) + 2Z\Omega\dot{\eta}(t) + \Omega^2\eta(t) &= B u(t) \\ y(t) &= C_1\dot{\eta}(t) + C_2\eta(t) \end{aligned} \qquad \Omega = \mathrm{diag}(\omega_i),\ Z = \mathrm{diag}(\zeta_i) \tag{5-17}$$

$$\omega_i = \left(\frac{\pi i}{l}\right)^2 \sqrt{\frac{EI}{\rho}} \qquad i = 1, \cdots, n \tag{5-18}$$

The forcing influence matrix is constructed by

$$B = [\mathbf{b}_1 \cdots \mathbf{b}_m] \qquad \mathbf{b}_j^T = \begin{cases} [\phi_1(x_j) \cdots \phi_n(x_j)] & \text{point force} \\ [\phi'_1(x_j) \cdots \phi'_n(x_j)] & \text{point torque} \end{cases} \tag{5-19}$$

$$\phi_i(x) = \sqrt{\frac{2}{\rho l}} \sin\left(\frac{\pi i x}{l}\right) \qquad i = 1, \cdots, n \ . \tag{5-20}$$

For a prescribed set of k general outputs, $\mathbf{y}(t)$, the output influence matrices are the coefficients of $\eta_{ii}(t)$ in the truncated modal expansion of the solution

$$C = \begin{bmatrix} \mathbf{c}_1^T \\ \vdots \\ \mathbf{c}_k^T \end{bmatrix} \qquad \mathbf{c}_j^T = \begin{cases} [\phi_1(x_j) \cdots \phi_n(x_j)] & \text{lateral displacement or rate at } x_j \\ [\phi'_1(x_j) \cdots \phi'_n(x_j)] & \text{rotational displacement or rate at } x_j \end{cases} \tag{5-21}$$

What we are not going to do with this example is to treat the m or k locations of actuators or outputs, x_j, as continuous independent variables, since this approach is impractical with a more complicated geometry. Instead, we will fix a finite set of sites over half the beam (exploiting the symmetry in the mode shapes), designate as the independent variables the squared 'gain' or 'power' of the devices, and compute the optimum distribution of authority for these sites. First, we will specify a disturbance and regulated variable to establish a signal path for scaling as described in Chapter 3, Section 3.2.

A point-torque disturbance at one end of the beam is interesting because it excites all the modes, and compared to a translational force type input, excites the higher modes more. In conjunction with this input, we will consider two regulated variable outputs: z_1, an angular displacement colocated with the torque input, and z_2, a lateral displacement at the center of the beam. Figure 5-3 illustrates disturbance, w(t), and the two outputs.

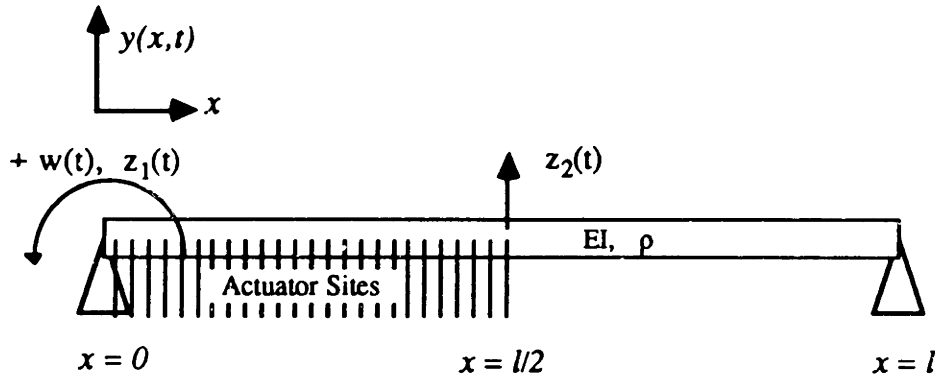


Fig. 5-3. Disturbance, Regulated Variables, and Actuator Sites.

We also fix an evenly-placed distribution of 25 potential sites for lateral force actuators over half the beam, and define new independent variables, ξ_j, j = 1, 25, as the power gain (square of the linear gain) at each actuator site. We may now rank order the modes in terms of the eigenvalues of the balanced Gramian (σ_i) for a z(s)/w(s) path. Since the lower modes are well separated in frequency, we may estimate σ_i, i =1, n with Gregory's [49] approximation. Considering first the $z_1(s)/w(s)$ (end rotation output) signal path the σ_i fall off as the square of the mode number:

$$\sigma_i \approx \frac{\sqrt{b_i^2 c_i^2}}{4\zeta_i\omega_i^2} = \frac{l}{2\zeta_i EI(i\pi)^2} \qquad \text{for } b_i = c_i = \frac{i\pi}{l}\sqrt{\frac{2}{\rho l}}\cos\left(\frac{\pi i x}{l}\right)\Big|_{x=0}. \qquad 5\text{-}22$$

The second signal path, $z_2(s)/w(s)$, excludes all the even modes because they have no displacement at the center of the beam, and we estimate

$$\sigma_i \approx \begin{cases} \dfrac{l^2}{2\zeta_i EI(i\pi)^3} & \text{for } b_{di} = \dfrac{i\pi}{l}\sqrt{\dfrac{2}{\rho l}}\ ,\ c_{zi} = (-1)^{i+1}\sqrt{\dfrac{2}{\rho l}} \quad \text{i odd} \\ 0, \quad \text{i even, since } c_{zi} = 0 \end{cases} \qquad 5\text{-}23$$

These estimates are accurate, by equation 3-117, [49] for modes, i, satisfying

$$\frac{i\zeta_i}{2} << 1 \qquad 5\text{-}24$$

so for the purpose of these examples, with 10 modes modeled, damping of 0.02 or less satisfies this inequality with a ratio of 10.

We plot the ratio of σ_i/σ_1 in the order of modal rankings by σ_i for a specific damping, $\zeta_i = 0.001$, in Figure 5-4. The ordering in this case is also the same as the frequencies, ω_i. Model reduction is here simply modal truncation at an upper frequency, and, for the case of the $z_2(s)/w(s)$ signal path, elimination of all even-numbered modes as well. It will be interesting to compare this to a similar plot for the Draper Model-2 large-scale example used later.

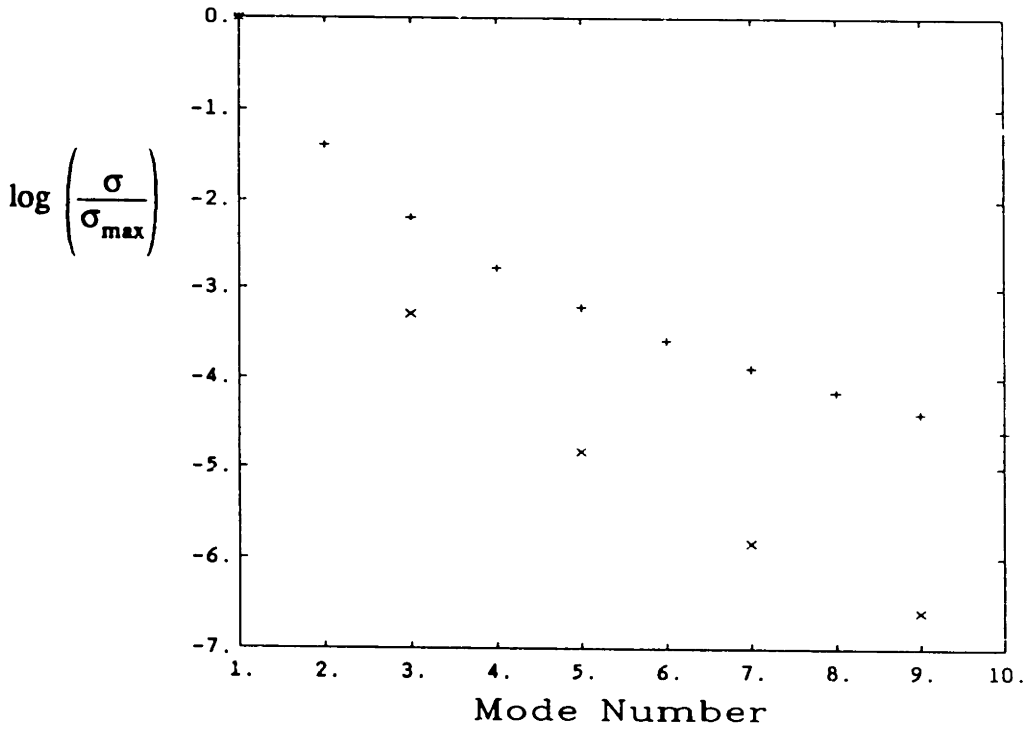


Figure 5-4. Mode Rankings for z_1/w (+), and z_2/w (X), reference signal paths

For convenience we fix parameters so

$$\frac{2}{\rho l} = 1 \quad \text{and} \quad \left(\frac{\pi}{l}\right)^2 \sqrt{\frac{EI}{\rho}} = 1 \qquad 5\text{-}25$$

and pick input and output scale factors (for the rotational output only):

$$R_w = R_{z1} = \left(\frac{\pi}{l}\right)^2 . \qquad 5\text{-}26$$

This makes the coefficients of the normal modes and their slopes unity, and also sets the coefficient in the sequence of modal frequencies, eqn. 5-18, to unity. Physically,

this could be the case of a square aluminum beam (density = 0.1 lbm/in^2, E = 10^7 psi = 12×10^7 slug/sec^2/in) of width h = 0.9417 in ($\rho = 0.1h^2/32.2$ slug/in, $I = h^4/12$ in^4) and l = 726 inches long. The parameters are selected here for convenience only, and will be less arbitrary in the Draper Model No. 2 example later. We will not use z_1 and z_2 at the same time in these cases, so their scale factors may also be different for convenience.

We truncate the order of the model at n=10 for computational purposes, but will experiment with smaller values. The nodal spacing for the 10th mode is $l/10$, which is five times longer than the actuator site spacing of $l/50$.

5.1.2. Relaxed Solutions

Exploiting the continuity introduced by selecting the 'power gain' of the discrete actuator sites as independent variables, we compute the optimal distribution of actuator authority for several cases of scalings, model order, controllability norms, and constraints. The example cases are grouped into three broad classes. The first class is actually an array that examines simultaneously the relaxed solution for several combinations of full- vs half-order Gramians, scalings, constraints, and measures of the Gramian. A particular case is selected for the second class of examples: the scaling signal path, model order, time horizon, and input frequency weighting are varied one at a time. In the third class, failures are introduced and the combinatorial neighborhood-optimal integer-constrained solution found for the expected value and worst-case failure weightings.

In order to provide some means of evaluating the results we introduce a LQG regulator which weights the output variable and control effort in the cost functional:

$$J_{LQi} = \lim_{t \to \infty} E\left\{ \frac{1}{t} \int_0^t \left[z(\tau)_i^2 + r_u \mathbf{u}^T(\tau)\mathbf{u}(\tau) \right] d\tau \right\} \quad i = 1,2 \ . \qquad 5\text{-}27$$

To compute the regulator for a particular actuator authority solution, ξ, we assemble an input matrix, B_a, from the up to N columns of the potential actuator sites weighted by the square roots of the non-zero ξ_i

$$B_a = [\ \cdots \sqrt{\xi_i}\mathbf{b}_i \ \cdots] \qquad \text{for } i \le N \text{ with } \xi_i > 0 \ . \qquad 5\text{-}28$$

The basis scalings are not used in the regulator evaluation model after a particular solution, ξ, is reached; however, the input and output constant scale factors, R_w and R_z (eqn. 5-26) are retained for numerical convenience. The solution to the regulator requires full state feedback, and the states are the modal displacements and velocities. For expediency, we evaluate the stochastic response of the closed-loop system to a wide-band disturbance of unit intensity assuming the modal signals are available for feedback. This represents an upper bound to the performance obtained with sensors and and estimator in the loop. The ratio of (scaled) output to control weighting was fixed by setting $r_u = 0.001$ for all cases; in physical units, this weights radians of tip rotation to pounds of lateral force in the ratio of $[r_u R_z]^{1/2} = \pi/l\sqrt{r} \approx 1.37 \times 10^{-4}$ for the physical parameters mentioned. This resulted in a little added stiffness and closed-loop damping, ζ, in the range 0.03 - 0.15, up from the open-loop of $\zeta = 0.001$ for all modes.

The first step is to compute scalings based on the disturbance to regulated variables signal path, per Table 3-1, and we compute the 'input-normal', 'throughput-scaled', and 'output-scaled' transformations and store these. Next we compute the controllability Gramian of each actuator site for the system in modal coordinates over a truncated interval and store them. The interval is selected for the first class of examples as one period of the lowest mode, which results in the highest mode integrating over 100 cycles, and is about one time constant for the steady-state value of the Gramian in the direction of the ninth mode. The scalings will be applied individually to each Gramian before running the relaxed algorithm, rather than to the system model before computing the Gramians. This allows the expensive computation of all the Gramians to be done once for a particular time horizon.

The purpose of the first class of cases is to experimentally evaluate the effect of the scaling rule, the particular optimization parameter measuring the Gramian, and the validity of the hypothesis that the reachable volume may be adequately characterized in n-dimensional space rather than $2n$, where n is the number of modes in the model. The disturbance torque to collocated rotation output , $z_1(s)/w(s)$, signal path defines the scaling normalizations. The total actuator power is constrained to $M \leq 5$ in all cases, which only serves to make the relaxed solution comparable, in terms of power, to a set of five identical

'unit' actuators. A total of 36 cases of the relaxed solution are computed, corresponding to all possible combinations of the parameters detailed in Table 5-1.

Table 5-1 Cases

Parameter	Selections	Number
Gramian order	n, $2n$	2
State Basis Scaling	Input Normal Throughput Scaled Output Scaled	3
Optimization	J_1: $\lambda_{min}(W)$ J_2: $n/tr(W^{-1})$ J_3: $\|W\|^{1/n}$	3
Individual Power Constraint	$m_i/M \leq 1$ $m_i/M \leq 0.5$	2
Total Cases		$\Pi = 36$

In all cases we start the algorithm for the relaxed solution with actuator power distributed uniformly across all 25 sites at a level so that the sum equals the total power constraint, which is $\xi_i = M/25 = 0.2$. As a point of departure, we evaluate the LQG regulator for this solution and the resulting stochastic response of the closed-loop system, and compare it, in Table 5-2, to the open-loop response

Table 5-2. Open-Loop and All-Actuators Closed-Loop Response.

root-mean-square:	Closed-Loop	Open-Loop	
J_{LQ} (rmJ)	1.076	19.7	
Regulated Variable (rmZ) for z_1	0.8029	19.7	(angle scaled by R_z)
Control Effort (rmU)	22.659	0	(lbf)

Again, for physical units, all of the above are with respect to a disturbance torque of intensity $R_w^{-1/2}$ in-lbf, rms, which is approximately 231 in-lb rms, for the particular parameters identified. The overall mean-square open-loop response, then, is $19.7(R_w R_z)^{1/2}$ rad, rms, or about 3.7×10^{-4} rad for the particular parameters noted.

The two most important features looked for in these results is 1) the ability of the relaxed solution to distribute actuator authority in a manner that produces good regulator performance, and 2) the ability of the order-n solution to produce solutions (ξ) close to the order-$2n$ solutions. Hopefully, these two effects occur together, and in fact they do. We will not review all cases here, but describe the general trend and show selected cases. The strongest single observation is that the 'output-scaled' cases were all unable to produce any correlation between the order-n and order-$2n$ solutions, except for the determinant cases where the scaling does not matter. The second major observation was that there was a greater difference between the use of the minimum-eigenvalue norm (J_1) and the trace norm (J_2) than there was between the 'input-normal' and 'throughput-scaled' cases. Another observation was that scaling did not in fact affect any of the solutions based on the determinant norm. The only reason these cases were run at all is because that is how this particular detail was discovered in the first place.

The first half of the cases were under a total power constraint only ($M \leq 5$, so $m_i \leq M$); that is, the constraint on any particular site was just the total constraint. Three distinct distributions of actuator power turned up. Figures 5-5 through 5-7 illustrate an example of each of these patterns, and Table 5-3 lists the parameters for each group. The LQG regulator evaluation line shows the root-mean-square value of the LQ cost functional, 'rmJ', regulated variable 'rmZ', and control effort 'rmU'. The 'station' variable represents the equally-distributed actuator sites, with station 25 at center-beam and station 1 at $l/50$ from the end, as shown in Figure 5-3.

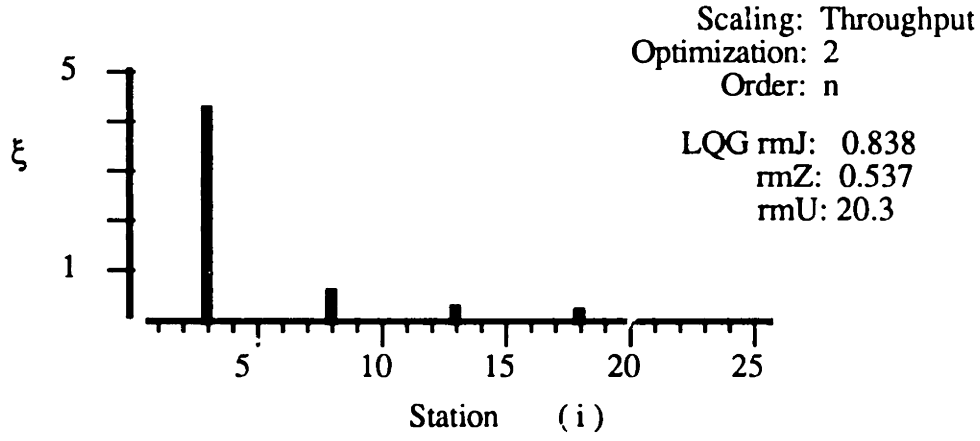


Figure 5-5. Group 1 Example.

The first group (Figure 5-5) produced the best regulators of the cases considered, so a particular case- 'throughput-scaled', J_2 (trace) norm for order n- was selected for further experiments. Notice that all three groups were able to enhance performance by concentrating actuator authority from the starting even distribution. The best level of improvement (for Group 1) was only about a 22% reduction from the starting solution, but the number of actuators was reduced from 25 to 4 in most cases in the group and down to only the two strongest sites of Figure 5-5 in one case

Table 5-3. Parameters for Groups

Group	Basis Scaling	Gramian Order	optimization	Number of Cases
1	Input	n & $2n$	2	
	Throughput	n & $2n$	1 & 2	
	Output	n	2	7
2	Input	n & $2n$	1	
	Output	n	1	3
3	All	n & $2n$	3	6

Common to all Groups:

Constraints:	$M \leq 5$ and $m_i \leq 5$
Model Order	n = 10
Scaling Signal Path:	$z_1(s)/w(s)$

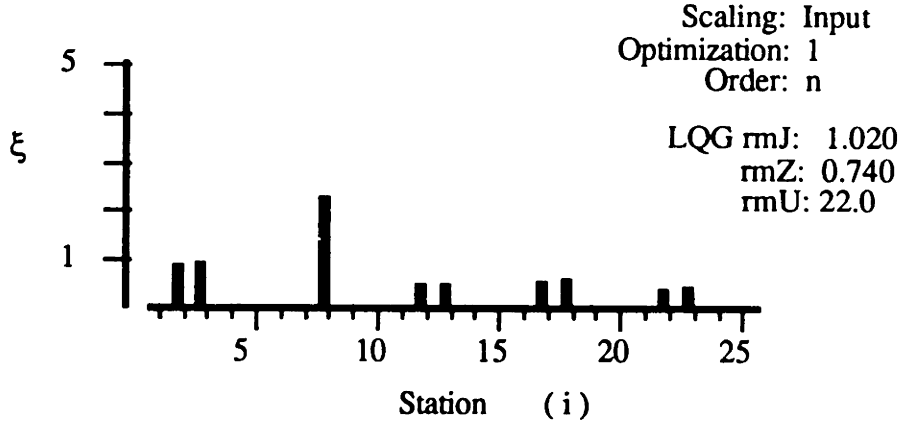


Figure 5-6. Group 2 Example.

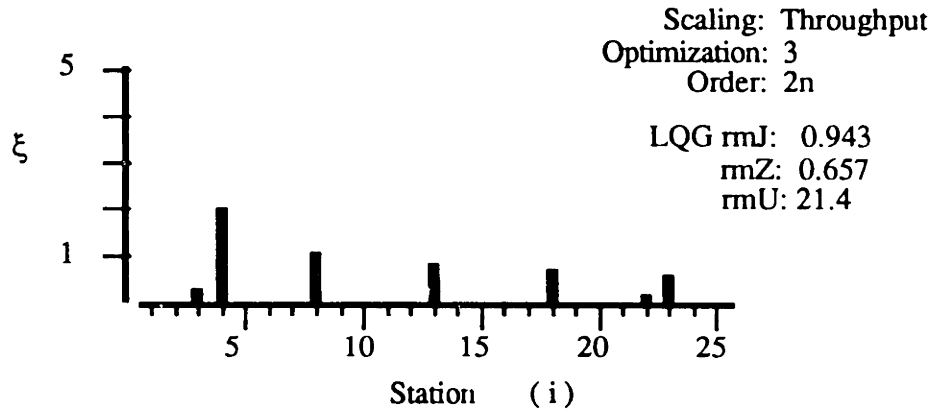


Figure 5-7. Group 3 Example.

While these cases do not conclusively show this particular combination of parameters to be superior to all other, there are several heuristic arguments suggesting just that. 'Throughput scaling' weights the input and output of the disturbance to regulated variables signal path, and that is a more thorough representation of the problem than the input or output alone. The J_2 (trace) norm is a function of all the dimensions in the reachable volume rather than just one for J_1 (minimum λ). Furthermore, the stochastic LQG cost is computed as a weighted trace of the mean-square closed-loop system response which is the solution of a Lyapunov equation closely related to that of the Gramian. Since we are able to produce essentially the same solutions with the algorithm operating on the half-order (modal-velocity) reachable volume, we will continue to exploit that for computational economy. The determinant norm (J_3) results of Figure 5-7 are interesting because they are independent of any (non-singular) scaling, and also add confidence to the use of the order-*n* approach. It appears that solution generally distributes authority on the peaks of mode 10, which has nodes at stations 5, 10, 15, 20 and 25. Toward the center of the beam (station 25) the nodes of mode 9 are closer to the peaks of mode 10, so the distributions tend to favor the ends, since the particular scaling was chosen to emphasize the high modes more than the low ones.

Introducing an individual power constraint of $m_i \leq 1$ to every potential actuator site modifies the solutions of Figure 5-5, and for the case selected ('throughput-scaled', J_2, and order-*n*) Figure 5-8 plots the new solution over the shadow of the previous solution. Notice the LQG performance is still superior to the Figures 5-6 and 5-7 class (Groups 2 and 3) solutions. Appropriate distribution of actuator authority can decrease both control

effort and regulated variable excursions, rather than the trade-off between them as we are accustomed to when dealing with a regulator problem with fixed inputs (or outputs).

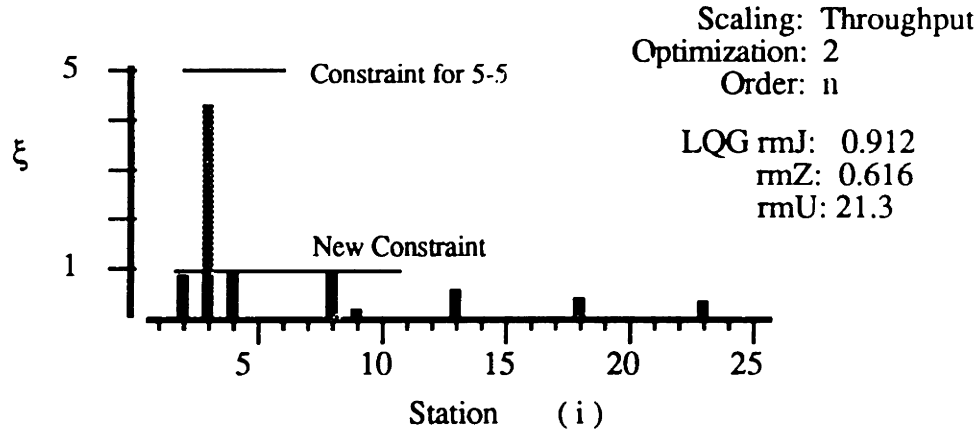


Figure 5-8 Individually Constrained Solution.

The emphasis on the highest modes is also clear here. Recall that this is deliberate by selecting a torque input and rotational output for the scaling signal path. To illustrate how we might alter the results, we carry out three examples: reducing the order of the model, introducing a frequency-dependent input weighting, and changing the scaling signal path. First we reduce the order of the model from 10 modes to 6 then 3, and compute the relaxed solution with the individual (and total) power constraint set at five. Figure 5-9 shows the results; predictably, the actuator authority is placed where it will emphasize the highest modes in the model.

Since this research is focused on structural dynamics applications, we cannot forget the fundamental limitations of model order. The torque-to-rotation scaling signal path used so far was deliberately selected to emphasis high modes, and it did so. As the last examples showed, the solution is very sensitive to the highest mode in the model. Recall, too, that the disturbance was not band limited. We can emphasize intermediate frequencies by introducing a frequency-dependent input weighting in the signal path as shown in Figure 3-4. For an example, set

$$U(s) = \frac{\mathbf{u}(s)}{u(s)} = \frac{\omega_c}{s + \omega_c} \qquad \omega_c = 10 \text{ rad/sec} \ . \tag{5-29}$$

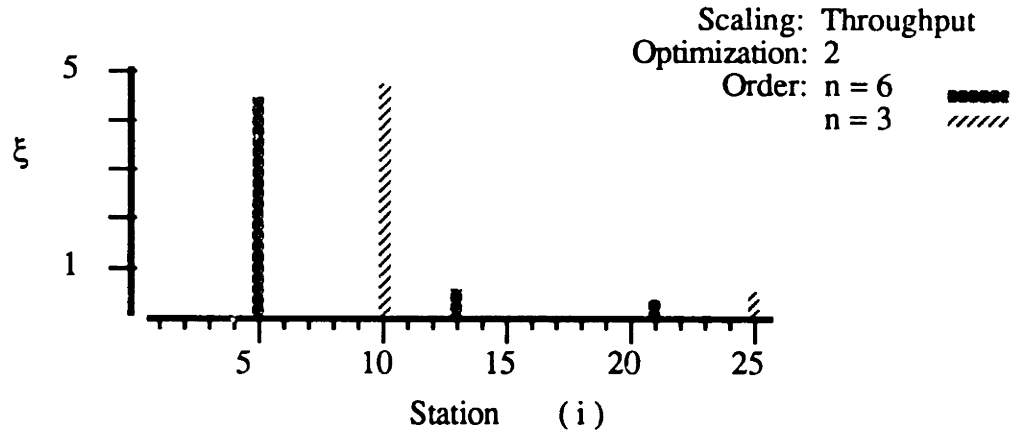


Figure 5-9. Reduced Models.

This changes the reachable volume of the disturbance, contracting it progressively more for modes above ω_c, mode 3 in this example, than for modes below ω_c. The detectable volume of the regulated variable is unaffected. Recomputing the 'throughput-scaled' transformations based on the new reachable volume of the disturbance through U(s) and the unaltered detectable volume of z_1 results in a shift of optimal actuator authority toward lower modes. We could accomplish approximately the same objective by introducing a direct scaling based upon the magnitude response of the desired filtering at each modal frequency, but would then lose the direct dynamic interpretation of the filter. Figure 5-10 plots the new solution with 10 modes in the model.

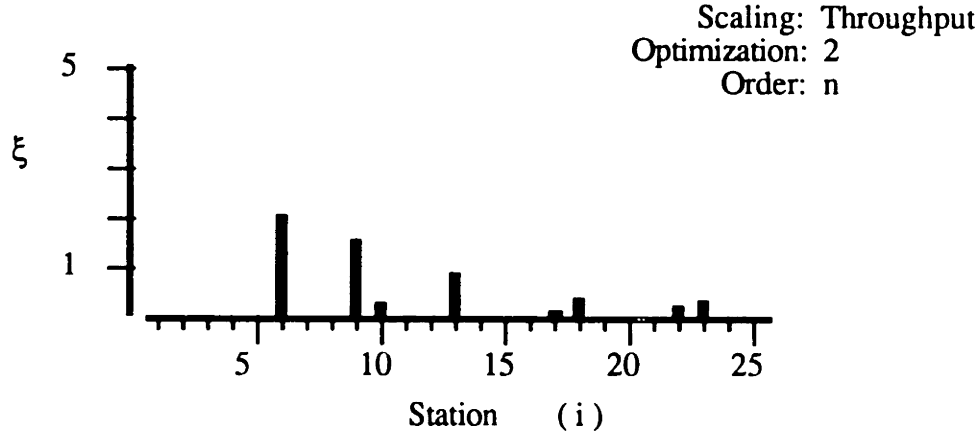


Figure 5-10. Input Frequency Weighted Solution.

Changing the scaling signal path from the rotational output to the displacement at the middle of the beam ($z_2(t) = y(l/2,t)$) results in the elimination of all even-ordered modes from the model and a predictable shift in actuator authority. The optimal location with respect to all controllability norms tried (J_1, J_2, J_3) and all scalings (input, throughput, and output) was at the center of the beam (station 25). All the odd modes have a peak at that point, and it is the only such point. Since this is not too surprising these cases are not pursued further or plotted. Note that the fact the determinant norm (J_3) also produced a mid-beam solution is not in conflict with the claim that the solutions for this norm are unaffected by scaling. There is no non-singular scaling which will eliminate half the modes from the model.

Before turning to the combinatorial failure-optimal problem there is another important parameter to investigate: the time interval over which all the Gramians are defined. We had mentioned in Chapter 3 that the interval should probably be related to the inverse of the desired closed-loop bandwidth. Returning to our example of Figure 5-5, in an attempt to characterize the effective bandwidth, the open-loop and closed-loop poles for the evaluation regulator are plotted in Figure 5-11, and the minimum and maximum singular values of the loop transfer function, $G_{LQ} = G(sI-A)^{-1}B_a$, are plotted in Figure 5-12. Not unexpectedly, since there was no explicit attempt to shape the singular values of the forward loop, it does not have a well-defined bandwidth and, indeed, the 'crossover' region is not well defined. We could modify the state weighting and generate a control law that does have a narrow crossover region, but the physical interpretation of the output-weighted regulator and the ability to make a simple rms comparison between cases would be lost. Rather than introduce a new method of comparison, we will leave the evaluation regulator formulation as it is, and experimentally vary the time horizon over which the Gramians are defined. The ill-defined crossover band of Figure 5-12 suggests a significant variation of the horizon is necessary to note changes in the solution, and this in fact turned out to be true. In particular, we compute the Gramians for an infinite horizon and then for a shorter one. Numerical problems occurred in the scaling transformations for time horizons of 0.01 and 0.1 because the disturbance Gramian became effectively singular. An interval of 1.0 was successful, and is somewhat shorter than the nominal interval of 6.28 (one period of the lowest mode). The optimal actuator authority distributions for these two

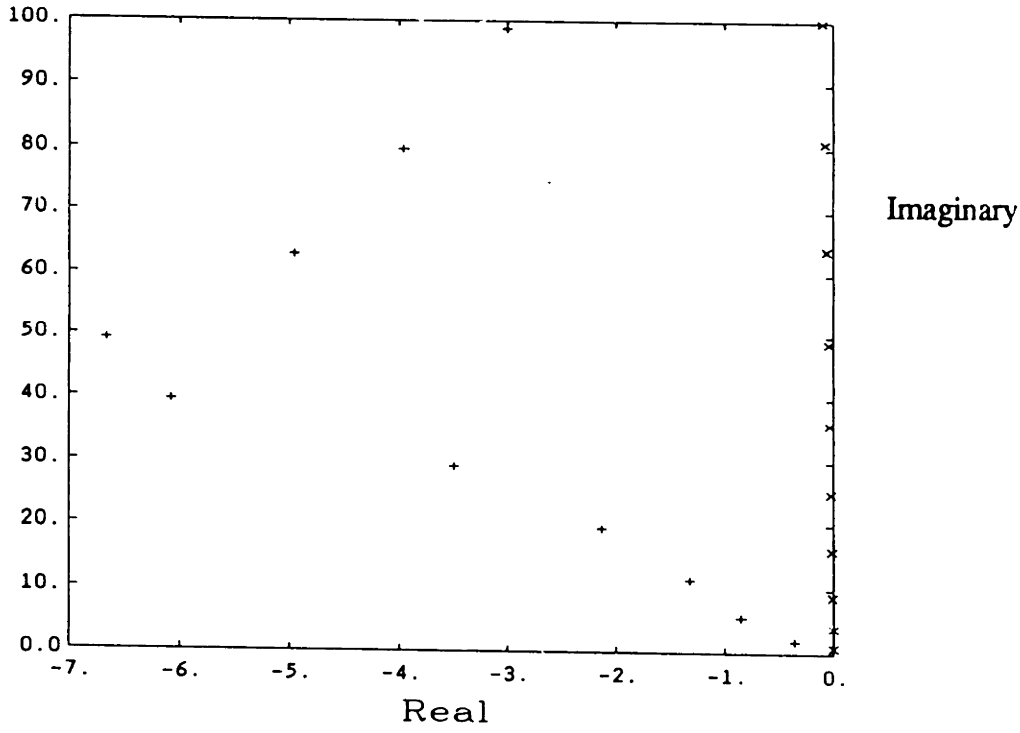


Figure 5-11. Open Loop (X) and Closed Loop (+) Poles

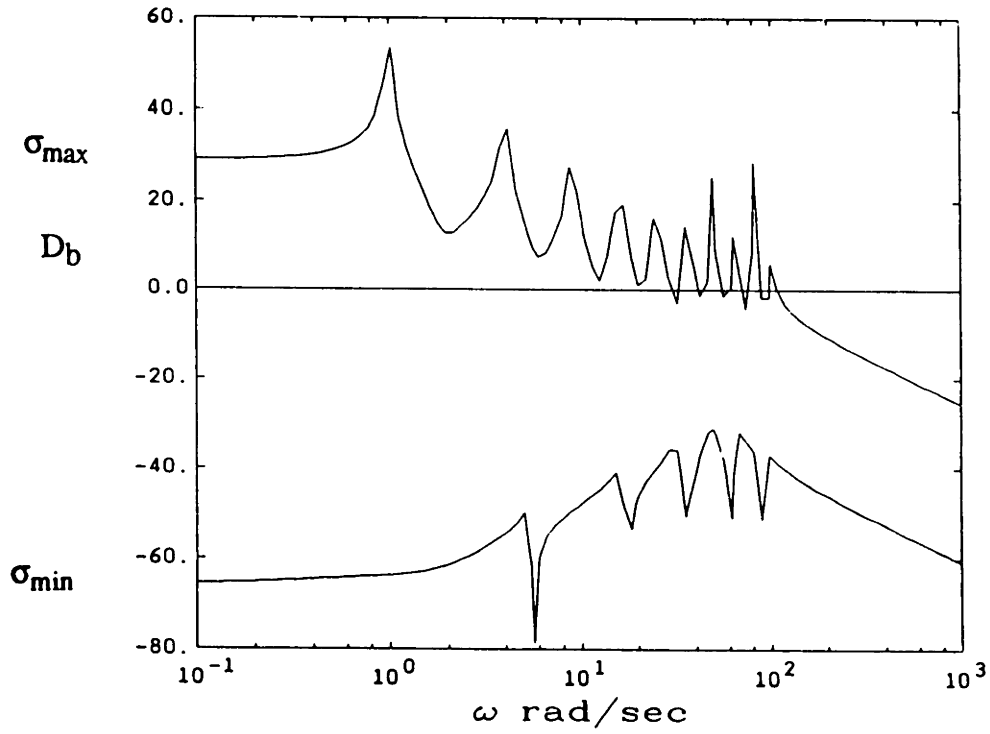


Figure 5-12. Singular values of $G_{LQ}(j\omega)$

cases were computed with the J_2 norm, 'throughput scaling', order-n Gramians, and are plotted in Figure 5-13. The infinite-horizon case shifts emphasis toward lower modes, and the resulting regulator is poorer than the examples in Figure 5-5. The shorter interval produces a result that does not make sense physically, since the peak for mode 10 is between stations 2 and 3. Most likely, the order-n assumption breaks down for the short time horizon . To investigate, we compute the spread of the second order modes ($\underline{\sigma} = \underline{\lambda}(W_c W_o)$) for the full-order ($2n$) to the half-order ($n$) Gramians of the scaling signal path and compare them in Table 5-4. We see that, while for the infinite and 2π time horizons the half-order block has comparable extreme dimensions to that of the full order, in the short-horizon case it does not.

Table 5-4. Full- and Half-Order σ over Three Time Intervals

$\sigma_{max}/\sigma_{min}$ for $z_1(s)/w(s)$ signal path

	Gramian Order	
Time Horizon	$2n$	n
1.0	2.25×10^8	5.5736
6.28	3.4786	3.0471
∞	1.11×10^4	1.00×10^4

The infinite-horizon example supports the assertion that some truncation of the interval is a better way to represent the problem, and generally reflects the fact that the lower modes become relatively more controllable over a longer interval. The short horizon case, however, points out numerical problems and limitations that are not well understood.

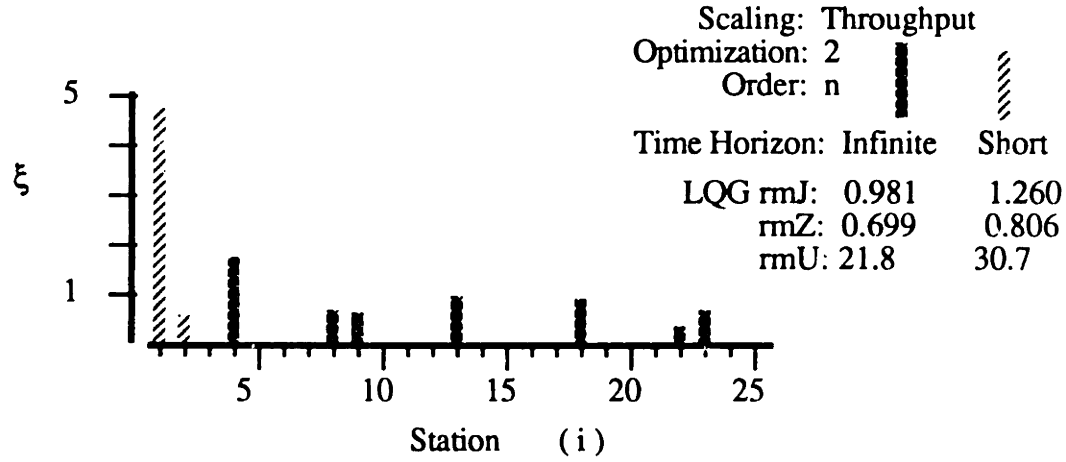


Figure 5-13. Infinite and Short Time Horizons.

5.1.3 Fault Tolerant Solutions

Turning finally to the fault-tolerant integer solution, we start with the constrained solution of Figure 5-8. Retaining the scaling and optimization parameter of that case, Table 5-5 lists additional specifications.

Table 5-5. Fault-Tolerant Problem

Number of actuators at BOL:	5
Required at EOL:	3 (94% probability)
Individual Reliability:	0.8

Enumeration of Failure Modes:

Number Failed	Probability	Distinct failure modes
0	0.3277	1
1	0.4096	5
2	0.2048	10
≥3	0.0579	16

The reliability of each actuator is assumed to be independent of the others. For a constant hazard rate, the end-of-life reliability represents a MTTF 4.5 times longer than the 'mission' duration. Failure modes of three or more failures are lumped and the controllability norm is assigned to zero for the expected-value computation. The worst-case is defined with respect to the failure modes with two failures.

The combinatorial search algorithm is started with four candidate configurations of actuators that are close to the distribution in Figure 5-8. These starting points and their geometric distance from the distribution in Figure 5-8 are listed in Table 5-6. Increasing the allowable distance to 1.2 resulted in 21 candidates, which was more than desired for computational-cost reasons.

Table 5-6. Initial Search Points

Configuration (Station)					Distance (from Fig 5-8)
2	3	4	8	13	0.5885
2	3	4	8	18	0.9189
2	3	4	8	23	1.0097
2	3	4	8	9	1.1526

The algorithm was executed twice, once for each of the failure weightings. In the case of the expected value weighting, the search of the '1-exchange' neighborhoods revealed no improvements and identified the first starting point as the best. For the worst-case weighting, one improved distribution was located and added to the list; the '1-exchange' search around it revealed no improvements. The two solutions differed by one actuator. The algorithm also keeps up with the best no-failure solution encountered, and it happened to be the same as the best expected-value solution. Also, since the algorithm had searched the neighborhood around that candidate, we know in this case that the best no-failure solution is also at least as strong a combinatorial optimum as the failure weighted solution. Table 5-7 lists the two solutions and the three methods of weighting the optimization norm, J_2, for each. Also listed are the failure-weighted regulator stochastic performances to a unit intensity wideband signal at the disturbance. To compute the closed-loop performance over failures, the regulator gains and the stochastic response are computed for every failure mode (16 in each case) and then the worst case identified or the results for all (failure) modes are weighted by their individual probability for the overall expected value. This presumes, in effect, reconfiguration after failure by resolving the same LQG optimization problem.

Table 5-7. Failure Weighted Results

	Best for Expected Value (and no-failures)			Best for Worst-case		
Solution (stations):	[2 3 4 8 13]			[2 3 4 9 13]		
Trace Norm, J_2	(relative to Fig 5-8)					
No Failures:	0.9951			0.9845		
Expected Value:	0.7646			0.7570		
Worst-Case:	0.4893			0.5019		
LQG Response	(rms, relative to Fig 5-8)					
	J_{LQ}	z_1	u	J_{LQ}	z_1	u
No Failures:	0.9977	0.9950	1.0000	1.0021	0.9938	1.0089
Expected Value:	1.0013	1.0058	0.9969	1.0063	1.0070	1.0052
Worst-Case:	1.2822	1.3915	1.1831	1.2940	1.4045	1.1938

Comparing the expected-value and worst-case controllability norm between the two results shows that the algorithm did correctly differentiate the two. The worst-case failure

mode for both solutions in terms of either the controllability norm or LQ cost was the loss of actuators at stations 3 and 4. There is a monotonic degradation from the relaxed solution of Figure 5-8 through the best no-failure integer, expected value, and worst-case best distributions (ratios of J_2 norm are less than one). The closed-loop response, however, does not follow this ordering exactly, nor do we have any theoretical reason to believe it should. Ratios of greater than one for LQ cost indicate poorer performance. In fact, the 'Expected-Value' solution of the right column is superior to the 'Worst-Case' solution on the left for all three closed-loop evaluations, and it is better with no failures than the Figure 5-8 relaxed distribution. However, even without knowing the best solution exactly for either the controllability norm or a closed-loop cost, these are likely to be very close to the global optimum, as we can show by a Monte-Carlo numerical experiment.

There are a total of 53,130 combinations of five actuators placed on 25 sites with at most one device per site. We construct a pseudo-random sample of these by generating sets of five random variables, each distributed uniformly over the interval [0, 25] and rounded up to the next integer. Sets with any duplicate integers are eliminated, as are any completely duplicate sets. A total of 100 sample actuator configurations were generated this way, and the controllability norm and regulator stochastic response were computed for each over the failure modes listed in Table 5-5. This required a total of 1600 Riccati equation solutions for the regulator and the same number of Lyapunov equation solutions for the response. The scaled Gramians were combined in as many combinations from a precomputed file, as is done in the algorithms, but the eigenvalues were computed for each combination to provide the exact value of the controllability norm.

The significant result is that none of the 100 samples in this Monte-Carlo exercise produced a better actuator distribution, either in terms of the closed-loop regulator or open-loop controllability norm, than the algorithm's solutions of Table 5-7. We might have expected the latter, but the former is an indication the overall approach is worthwhile. Figures 5-14 through 5-16 are histograms of the regulator cost functional for no failures, expected value and worst-case failures. The bins are equal increments, and range from a low of the best 'Expected-Value' solution's cost to twice that. Most of the cases fell in bins 2 and 3, with costs higher than the candidate best by 1.2 to 1.4 for all failure weighting methods. Figures 5-17 and 5-18 illustrate the expected value of the regulated variable (end

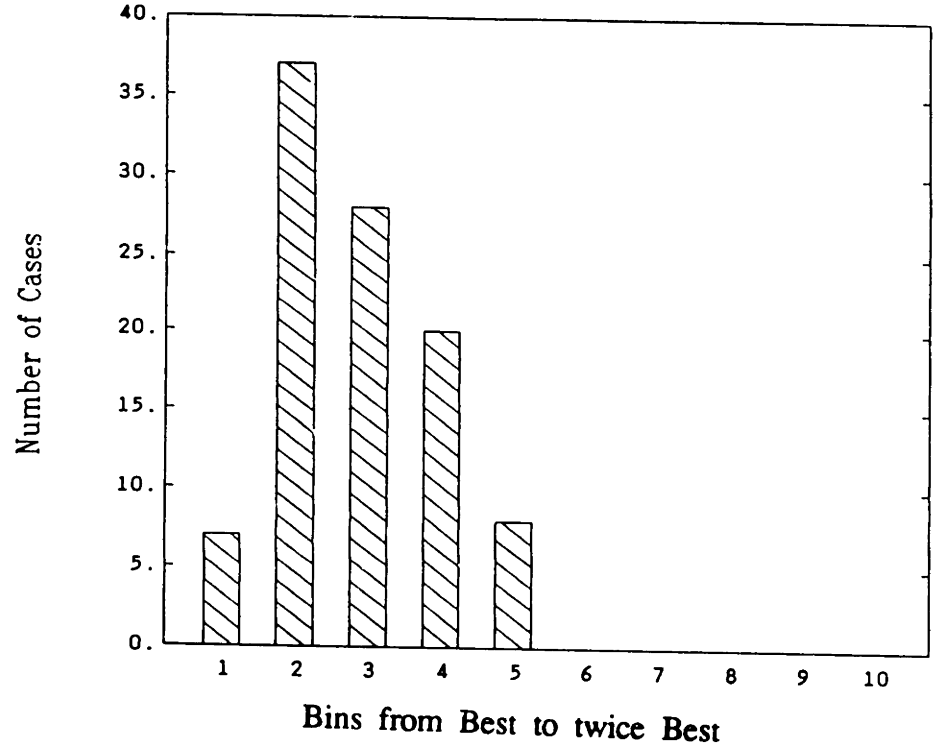


Figure 5-14. J_{LQ} for No-Failures

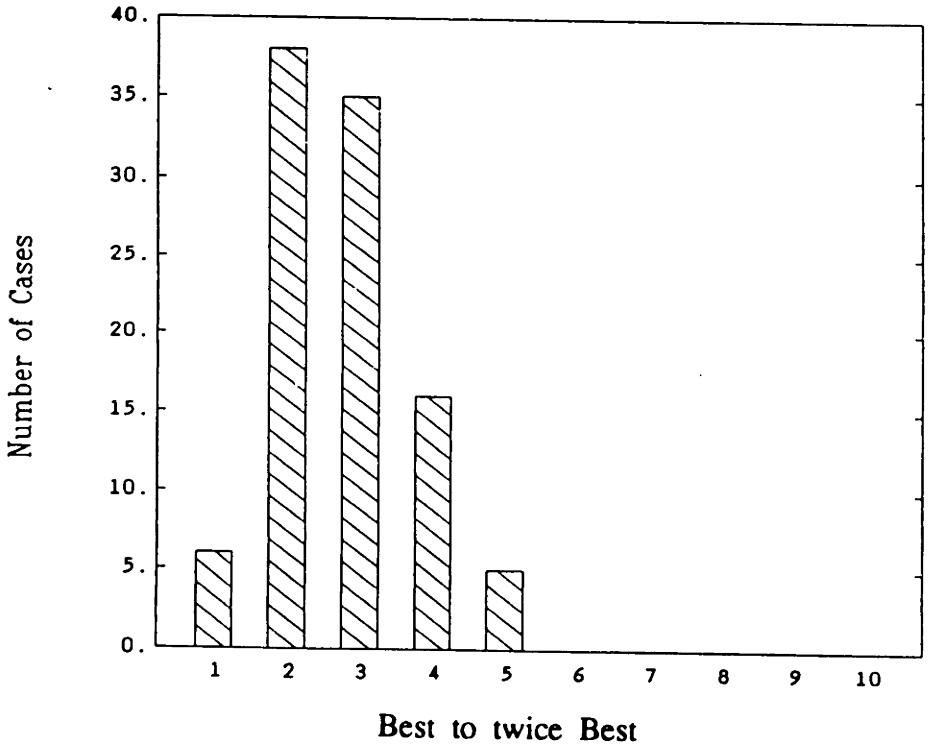


Figure 5-15. J_{LQ}, Expected Value over Failures

rotation) and control effort. The output is the dominant term in the cost functional, so its distribution in both cases is similar. The rms control effort ranged from 0.975 to 1.18 that of the candidate 'best' solution, which would fit into bin 2 on Figure 5-18. The open-loop norm, J_2, was distributed entirely below that of the best found for all failure weightings, and Figures 5-19 and 5-20 show the expected value and worst-case histograms for the 100 samples. Note the bins in Figure 5-19 range from half of the 'Expected-Value' solution's J_2 up to that value, while the worst-case bins in Figure 5-20 are from zero to the best 'Worst-Case' solution at the upper end.

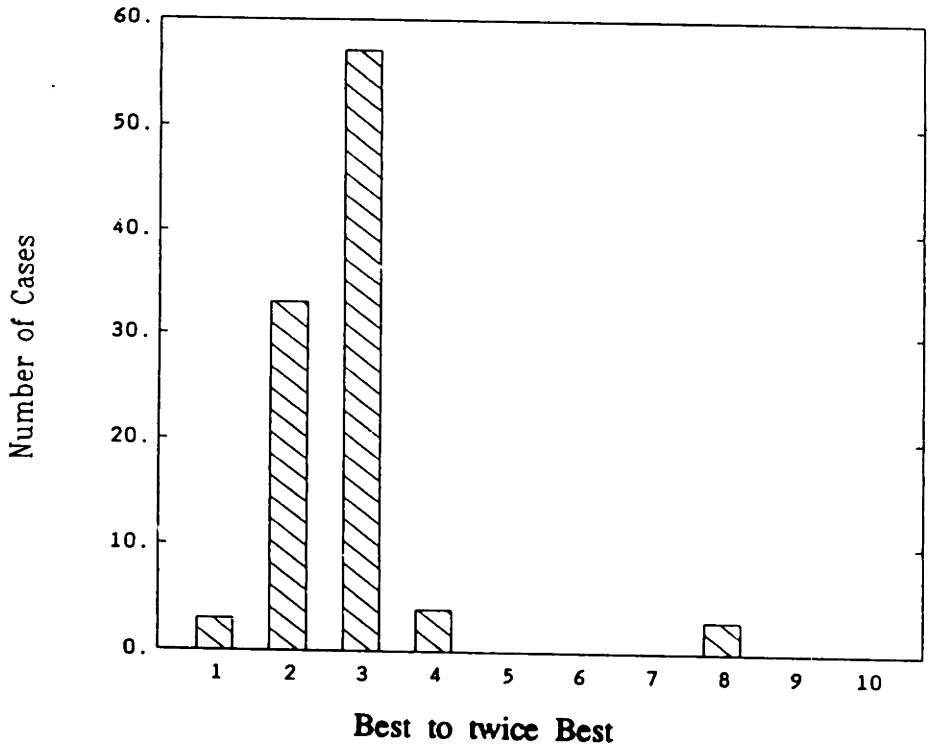


Figure 5-16. J_{LQ} Worst Case Failures

While these results do not prove anything rigorously, they are a strong indication that the algorithm's solutions are close to the global optimum, and this is often the best practical solution for combinatorial problems. Furthermore, the Monte-Carlo experiment took about five times more CPU time than did the algorithms to reach the solutions of Table 5-5 in the first place, including computing the Gramian file. This is a crude comparison because all of the coding was implemented in the non-compiled 'CTRL-C' instruction language, which excessively penalizes the multiple iterations through failure modes. About 100 minutes of VAX-8650 CPU time was required for the Monte-Carlo

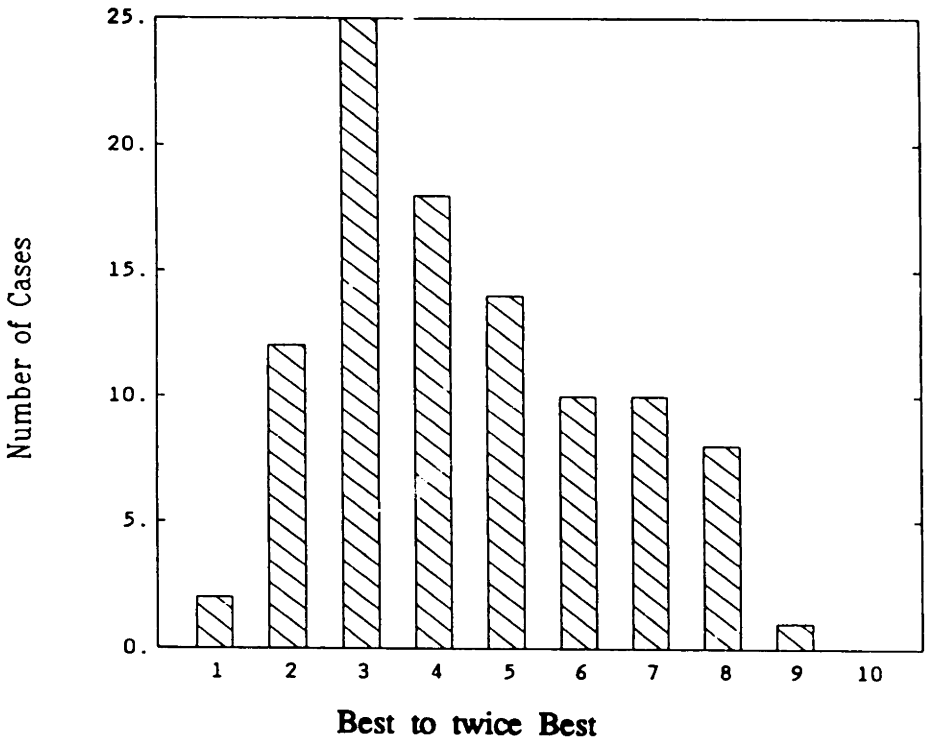


Figure 5-17. z_1 (t), rms, Expected Value over Failures

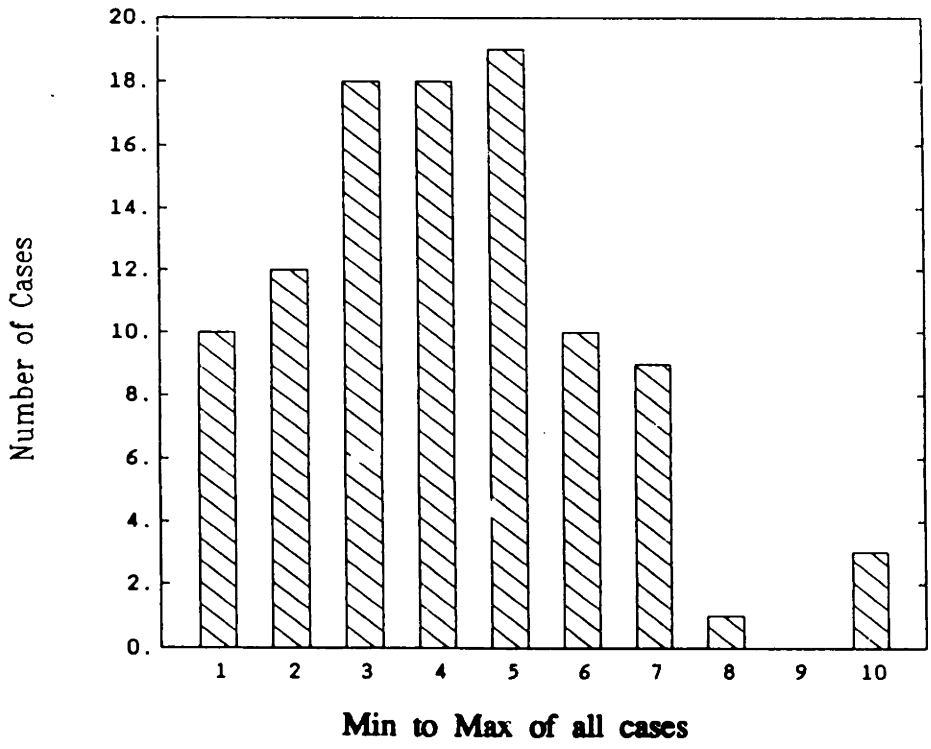


Figure 5-18. Expected Value of Control, rms

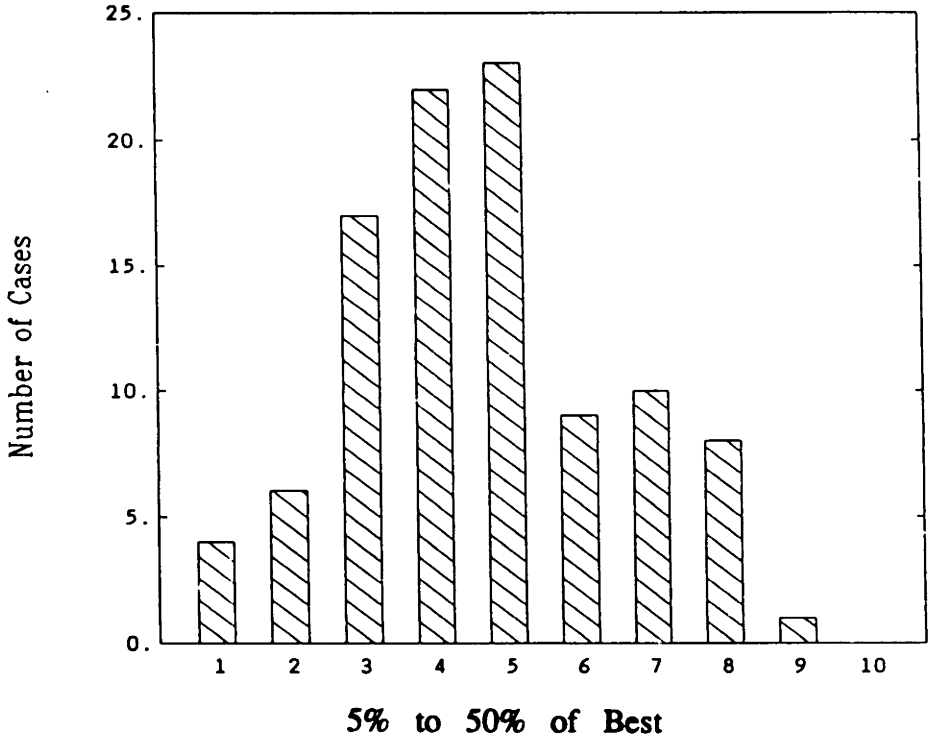


Figure 5-19. J_2 norm, Expected Value over Failures

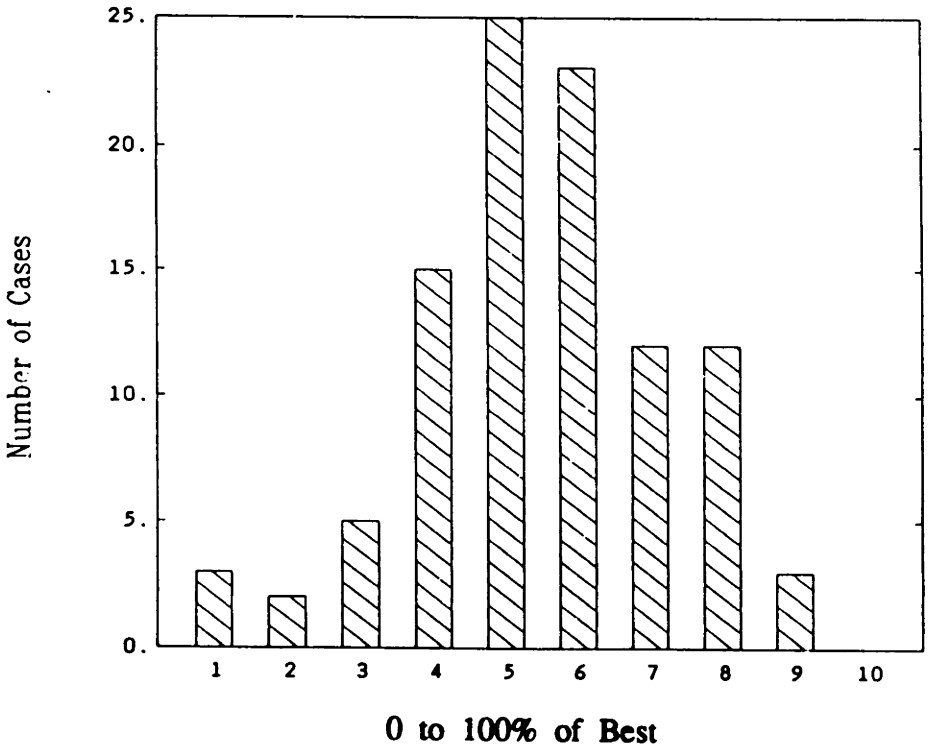


Figure 5-20. J_2 norm, Worst Case Failures

run, compared to 18 minutes for the algorithms, 15 of which were spent on the combinatorial search for the failure weighted solution.

We turn next to a large-scale structural model more representative of the kind of flight vehicle that motivated this research.

5.2 Draper Model No. 2

As a part of DARPA's (Defense Advanced Research Projects Agency) Active Control of Space Structures (ACOSS) program, the C. S. Draper Laboratory developed a model of a hypothetical optical system requiring active structural control [62]. Later, under the Vibration Control of Space Structures (VCOSS) program, the model was revised [63] to include additional kinematic and rigid-body detail in the optical elements and the equipment bay (Rev 1), and two variations of stiffness and weight; a lighter and less stiff version (Rev 2) and a heavier and stiffer version (Rev 4). These models were disseminated to participants in DARPA's programs and to academia; thus they represent a reasonably well known system which has many of the features that make control of lightly damped structures a difficult engineering problem. Revision 1 was selected for these examples; Figure 5-21 is an illustration of the overall configuration.

The particular control problem posed in the VCOSS program was regulation of the line-of-sight (LOS) pointing (x and y rotations) and maintenance of the focal length (defocus) in the z direction in the presence of a broadband disturbance acting upon the structure at nodes 37 and 46, as shown in Figure 5-21. The disturbances are statistically independent with a power spectrum as in Figure 5-22.

Specifications for the LOS were also imposed, and are listed in Table 5-8 along with the open-loop stochastic response of the model to the disturbance of Figure 5-22. We assume a modal damping ratio of $\zeta_i = 0.001$ for all modes in order to compare results to those of Hegg and Kissel [66].

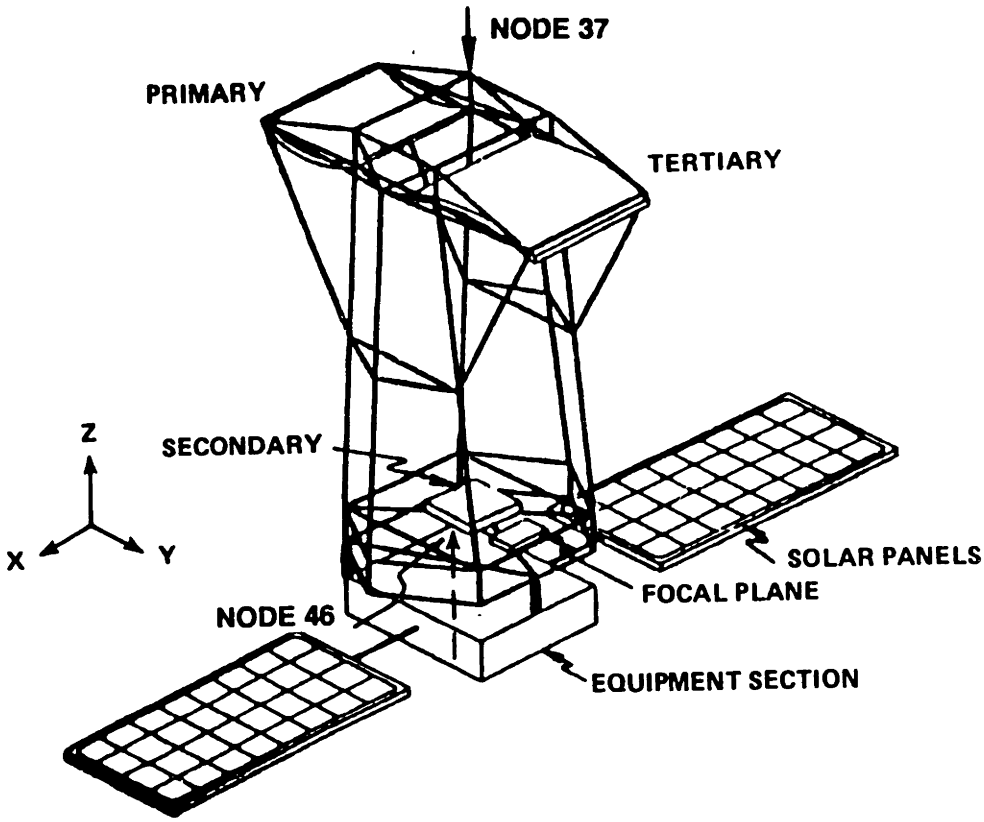


Figure 5-21. ACOSS Model No. 2

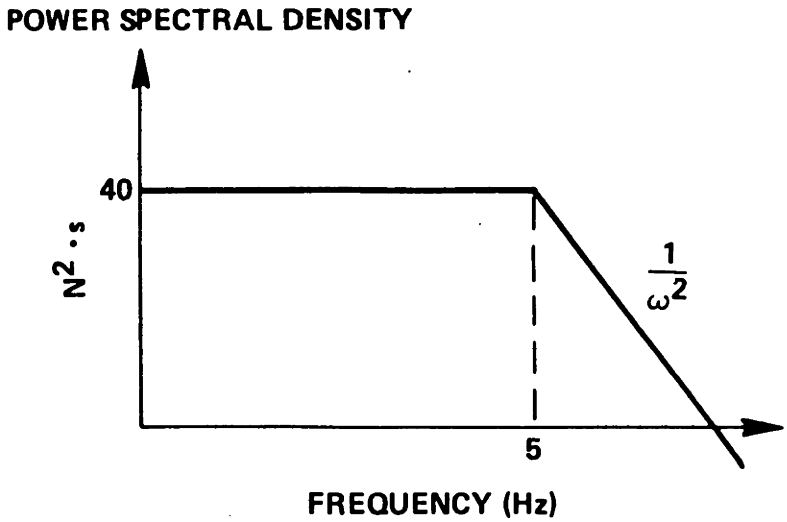


Figure 5-22. Disturbance Spectrum

Table 5-8. Control Requirements and Open-Loop Response

	Transverse μ-rad, rms		Defocus mm, rms
	x	y	z
Requirement:	0.05	0.05	1
Open-Loop:	75.5	771.4	6.7×10^{-3}

Note that while the structure presents more than adequate stiffness for elongation along the optical axis, as seen by the small defocus response, 64 to 84 Db of disturbance rejection is required for pointing. We next summarize some modeling details in preparation for computing sensor and actuator distributions.

5.2.1 Modeling

The Model No. 2 is essentially a truss structure supporting several discrete rigid bodies which comprise most of the mass. Joints on the truss, however, are solid, so member forces may include shears and moments as well as tension and compression. Overall dimensions are about 28 meters from the upper support truss (for the primary and tertiary) to the equipment section, and 52 meters tip to tip across the solar panels. The rigid masses supported by the truss include three optical (mirror) segments, a focal plane assembly, and the lower equipment bay. There are a total of 58 nodes in the model, each with six degrees of freedom. Total mass is approximately 9337 kg, of which 1023 kg are structural and 8314 kg are non-structural and rigid. Reference [63] contains further details, including the NASTRAN run deck for producing a finite-element model of the structure. From a finite-element analysis, mode shapes and eigenvalues are computed. Table 5-9 lists the first several frequencies, and sketches of their corresponding mode shapes are in Appendix B.

Table 5-9. Low Frequency Eigenvalues of Draper Model No. 2 (Rev 1)

System Mode Number	Frequency (Hertz)	System Mode Number	Frequency (Hertz)
1-6	0 (rigid body)	14	0.673
7	0.148	15	0.960
8	0.282	16	1.092
9	0.319	17	1.839
10	0.335	18	1.844
11	0.468	19	1.889
12	0.583	20	1.990
13	0.601	21	2.060

Also computed by the NASTRAN code are the modal influence and output matrices for several signal paths. We will require these matrices for the disturbance input, $[\Phi^T B_d]$, the line-of-sight output, $[C_z\Phi]$, and a large class of potential actuators, $[\Phi^T B_a]$, and/or sensors, $[C_s\Phi]$. Three classes of transducers are considered in this example, all discrete-point devices for computational expedience. At each of the 58 structural nodes are three orthogonal translation (external force) devices, which could be idealizations of proof mass actuators or accelerometers. Three orthogonal rotation (torque) devices are also modeled at each node, which could represent momentum wheels, control moment gyros (CMG) or instrumentation gyros. There are 174 devices each in the rotational and translational class. A third class of devices is the linear extension of the 134 beams distributed through the truss. Sensors in this class would measure the relative displacement (strain) between the nodes linked by a beam, or actuators would generate a tension or compression across the member. We will reconsider the issue of sensor and actuator dynamics in the context of relative weightings between the three classes in 5.2.2, but for these examples all devices are idealized as broad-band. The influence coefficients, and later the Gramians, for all 485 candidate devices are computed and stored. Appendix B lists the first significant figures for the first 18 flex modes of some of the input and output matrices we will use with this example. The numerical data is furnished to allow a partial check of the model by a researcher interested in verifying the results to follow, but the model should be transferred on magnetic tape or generated from an original NASTRAN run with the data in reference [63] to be accurate.

For the LOS regulated variables,

$$z = \begin{bmatrix} \text{LOS x (rad)} \\ \text{LOS y (rad)} \\ \text{defocus z (m)} \end{bmatrix} \qquad 5\text{-}30$$

an output scaling, R_z as described in section 3.2.1, is introduced so the effective output units are in terms of the specified performance listed in Table 5-8.

$$R_z = \begin{bmatrix} (0.05 \times 10^{-6}\ \text{rad})^2 & & \\ & (0.05 \times 10^{-6}\ \text{rad})^2 & \\ & & (1 \times 10^{-3}\ \text{m})^2 \end{bmatrix} \qquad 5\text{-}31$$

The effect of this weighting, since the defocus is already well below the specification listed in Table 5-8, is to sharply attenuate the impact of the defocus performance on the basis scaling.

The spectrum of the disturbance, Figure 5-22, motivates a frequency-dependent input weighting, U(s), (section 3.3.4) which, with a unit intensity 'white' process input, produces the required spectrum. A two-state realization that does this is given by (A_u, B_u, C_u) with parameters

$$A_u = \begin{bmatrix} -\omega_c & 0 \\ 0 & -\omega_c \end{bmatrix} \quad B_u = \begin{bmatrix} \sqrt{40}\,\omega_c & 0 \\ 0 & \sqrt{40}\,\omega_c \end{bmatrix} \quad C_u = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \omega_c = 10\,\pi\ \frac{\text{rad}}{\text{sec}}\ . \qquad 5\text{-}32$$

We are now in a position to produce an economical model by selecting modes according to their significance in the disturbance to (weighted) regulated variables signal path. Since we are not interested in rigid-body control for this example; we first delete the six corresponding zero-eigenvalue modes. A total of 50 flexible modes, ranging in frequency from 0.148 Hz to 32.83 Hz, and the corresponding mode shapes were computed by NASTRAN, and we would like to rank order all of them. While we could try writing a first-order system of order 102, (including two states for the input shaping filter), and computing the 'second order modes', $\Sigma^2 = \text{eig}[W_c W_o]$, as described in section 3.3.3, the software available would not manage a task of that size. Gregory's approximation, equation 3-116, is not directly applicable either, since it does not account for a frequency-

dependent weighting. What we can do, however, is represent the signal path of interest with input and output scalings, in the form of Figure 5-23, and assume (temporarily) that the modes are completely independent. We may check the confidence of the independence assumption (over an infinite horizon) with a check of Gregory's bound (equation 3-117).

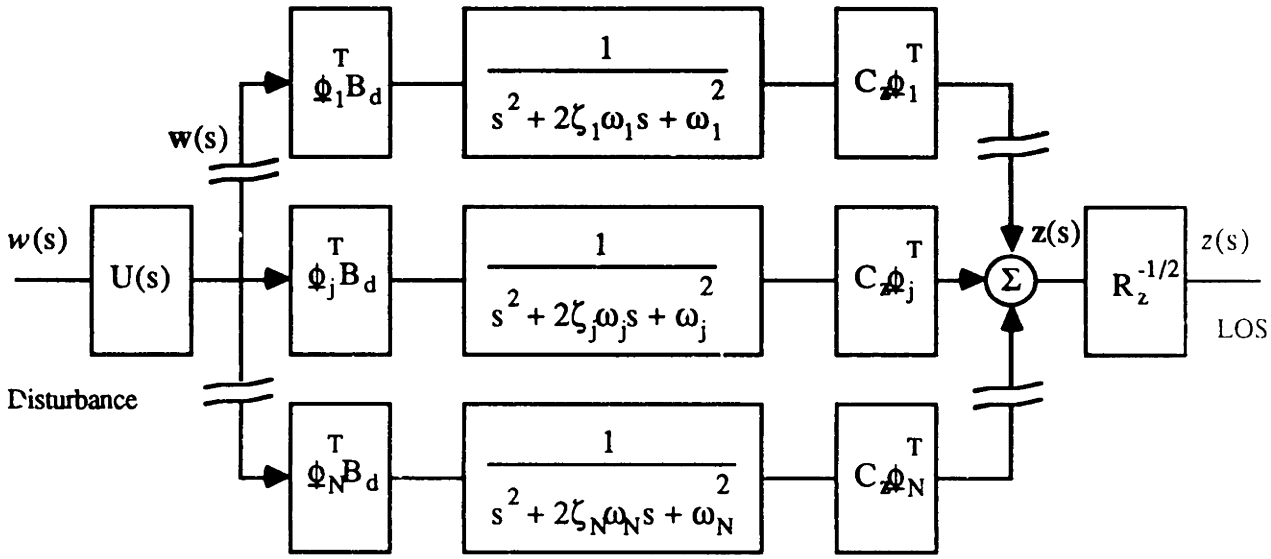


Figure 5-23. Scaling Signal Path.

We may now compute the controllability and observability Gramians of the signal path through each mode separately, which is inexpensive since the order of the system for each channel is four, two states for the mode and two for U(s). A realization of a modal channel is:

$$\frac{d}{dt}\begin{bmatrix}\dot{\eta}_j \\ \eta_j \\ x_u\end{bmatrix} = \begin{bmatrix} -2\zeta_j\omega_j & -\omega_j^2 & \Phi_j^T B_d \\ 1 & 0 & 0 \quad 0 \\ 0 & & A_u \end{bmatrix}\begin{bmatrix}\dot{\eta}_j \\ \eta_j \\ x_u\end{bmatrix} + \begin{bmatrix}0 \\ 0 \\ \\ B_u\end{bmatrix} w(t) \qquad 5\text{-}33$$

$$z(t) = R_z^{-1/2} C_z \Phi_j \eta_j(t)$$

where A_u and B_u are realizations of input filter U(s) from equations 5-32 and R_z is the output scaling, eqn. 5-31. With the ordering of equation 5-33 for the four-element state of each modal channel, we partition the corresponding infinite-horizon Gramians so the upper left, or [1,1] bloc of dimension 2x2, corresponds to the modal variables, $[\dot{\eta}\ \eta]^T$:

$$W = \begin{bmatrix} W_{11} W_{12} \\ W_{12}^T W_{22} \end{bmatrix} \begin{matrix} \left\{ \begin{bmatrix} \dot{\eta} \\ \eta \end{bmatrix} \right. \\ \left\{ x_u \text{ (dim = 2)} \right. \end{matrix} \qquad 5\text{-}34$$

We may then estimate the σ_i for each mode by computing

$$\Sigma_i^2 = \text{eig}(W_c[1,1]\ W_o[1,1]) \qquad 5\text{-}35$$

and, as noted by Gregory [49], within a lightly damped mode the two eigenvalues (σ_{i1} and σ_{i2}) are very nearly equal, so we take their geometric mean as a modal index. As a check on the validity of such an independent ranking, we also compute Gregory's bound (equation 3-117) between adjacent modes. With a significant change in interpretation, but very slight change in computation, we may interpret the upper left partition of the controllability Gramian, $W_c[1,1]$, which corresponds to the modal variables, as the steady-state covariance from a broadband disturbance with the spectrum of Figure 5-22. Removing the output scaling, R_z, we compute the regulated variable output covariance by

$$E\left[z_j z_j^T\right] = \begin{bmatrix} 0 & C_z \Phi_j \end{bmatrix} W_c[1,1] \begin{bmatrix} 0 & C_z \Phi_j \end{bmatrix}^T \qquad 5\text{-}36$$

where z_j is interpreted here as the contribution of mode j to the regulated variables z.

Table 5-10 contains a ranking of the flexible modes by both criteria, the deterministic σ_i ('second order modes' or 'Hankel singular values') and the stochastic contribution of each mode to the total (rss of x and y directions) line-of-sight error response.

Table 5-10. Modal Rankings.

Rank	Flex Mode Number	σ_i	Rank	Flex Mode Number	LOS (μ-rad)
1	1	19.705×10^4	1	1	600.97
2	18	3.193	2	18	455.85
3	7	1.508	3	17	105.90
4	6	1.196	4	7	92.65
5	17	8.495×10^3	5	16	77.04
6	16	6.206	6	6	72.43
7	10	3.919	7	10	32.47
8	2	2.901	8	8	15.91
9	4	2.820	9	4	12.94
10	8	2.446	10	2	12.22
11	3	1.130	11	3	5.06
12	5	2.195×10^2	12	33	5.02
13	9	1.186	13	34	4.29
14	33	1.644	14	49	3.28
15	34	1.333	15	46	3.21
21	15	6.33×10^1	30	15	0.72
29	14	2.20	34	14	0.25
35	11	8.55×10^0	35	11	0.09
38	12	0.49	38	12	5.3×10^{-3}
48	13	9.00×10^{-19}	49	13	0

Modal truncation will be based on the σ_i criterion, as discussed in section 3.2.2, but it is interesting to note the similarity to the covariance contribution Hegg used [66] for mode selection. We plot $\log[\sigma_i/\sigma_{max}]$ in the order listed in Table 5-10 along with Gregory's bound (eqn. 3-116) in Figure 5-24. We see clearly there are two logical truncations, one at 11 flex modes and the other at about 36. The dotted line in Fig. 5-24, showing Gregory's bound, is much smaller than unity (0 on the log scale) for the modes through the candidate truncation points and assures us that the independent estimates of the (infinite horizon) σ_i are accurate. That is, if we had the computational power to determine the σ_i's for all 50 modes simultaneously, we would find similar results for all but a couple of modes. We elect to truncate the model at the first 11 modes listed in Table 5-10, noting this is not a contiguous, but an interlaced, model by frequency. It is also interesting to note that with truncation at 11 modes, we would include the same modes whether we selected them by the σ_i criterion or their contribution to the LOS error, even though the ordering is not exactly the same.

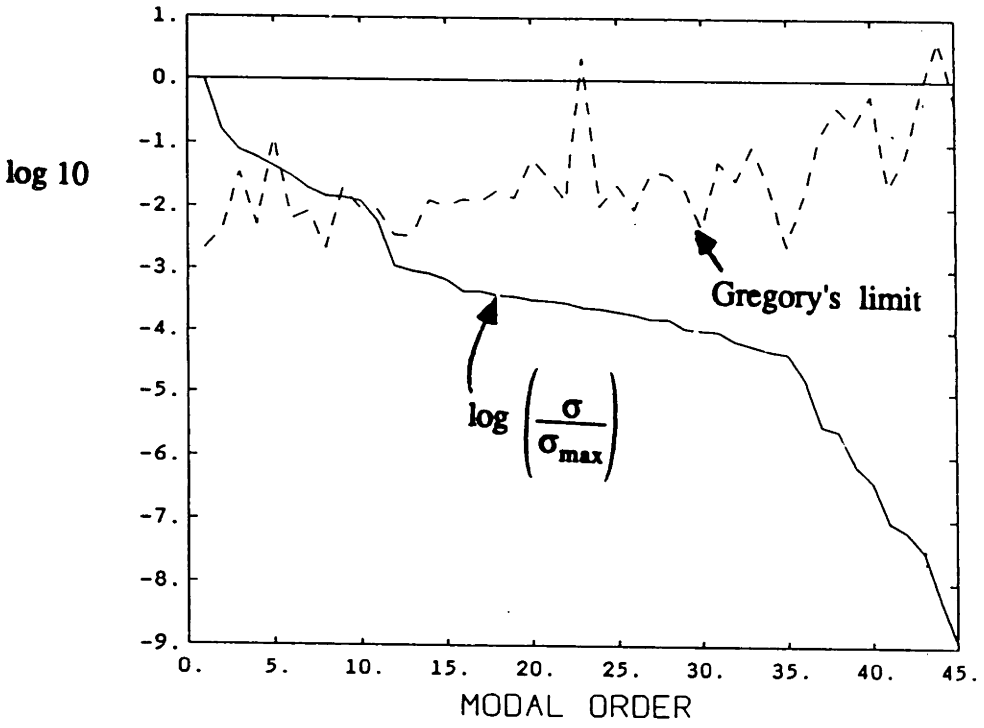


Figure 5-24. Second Order Modes with Gregory's limit.

Figure 5-24 can be compared to Figure 5-4 from the beam example. Notice the modal contributions to the signal path fall off more quickly for the beam by mode number, but less so by frequency. The significant differences between the two structures is that the modal density, by frequency, is far higher for the ACOSS structure than for the beam; and that the ordering by σ_i and ω_i is the same for the beam, but as shown in Table 5-8 is not for the ACOSS structure. The modal density and the fact that the second-most important mode is of relatively high order (18th flex mode) are among the primary causes of the control challenge this system presents.

With a model, a scaling signal path, and a global class of sensors and/or actuators specified, we are prepared scale the state basis, and find optimal distributions of effector authority.

5.2.2 Relaxed Solutions

One of the first results from the relaxed problem and the global class of 485 actuators was finding a strong distinction in overall authority between the three device

classes. The translational force devices are the strongest and, if they are admitted to the problem, no other type of device will remain in the converged solution. The extensional force devices, in contrast, were the weakest and, if any other class was admitted, no extensional devices would remain. Therefore, we initially exercise the algorithm for each of the three classes of actuators alone. The model basis is scaled to 'throughput-scaled' in accordance with Table 3-1, and the average energy (J_2) norm operating upon the half-order (n) Gramian is selected for the optimization parameter. The time horizon over which all Gramians are integrated is set at one period of the first flex mode, or 6.76 seconds. For each class of actuators, we first compute the fully relaxed solution with a total power constraint of unity, $M \leq 1$, then compute one or two further constrained solutions holding the individual constraints to unity, $m_i \leq 1$, while raising the total power to some higher level. The levels are not chosen completely arbitrarily, but correspond to the total number of devices in some similar selections by Hegg and Kissel [66] by their algorithm (section 3.2.2) for the same Model No. 2. Recall that the absolute magnitudes of the constraints are not as important as their proportions. Figures 5-25 through 5-33 illustrate the configuration of these solutions, and Tables 5-11 through 5-13 list the actuators in the solutions with their distributions of authority for the several constraint cases imposed.

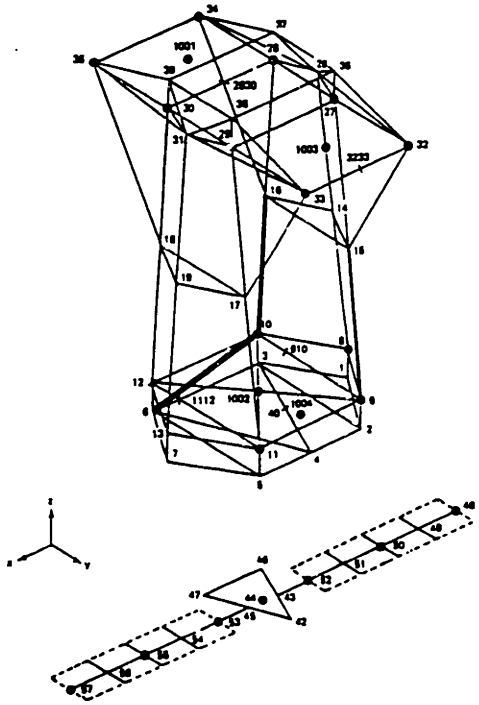


Figure 5-25. Extensional Actuators, Fully Relaxed Distribution

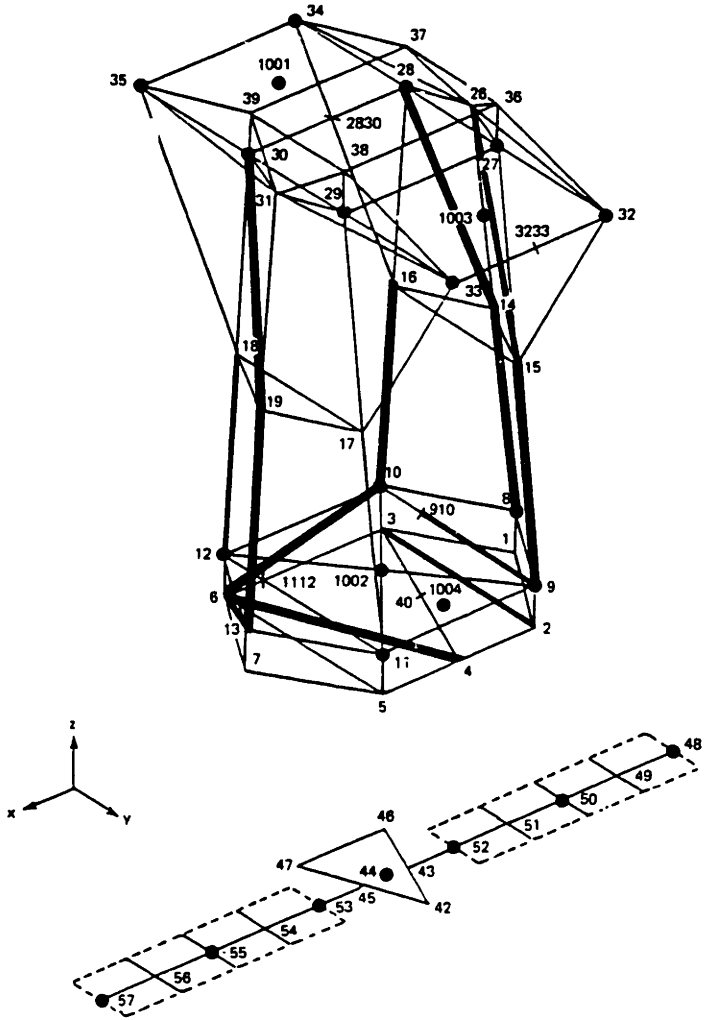


Figure 5-26. Extensional Actuators, '1/10' Constrained

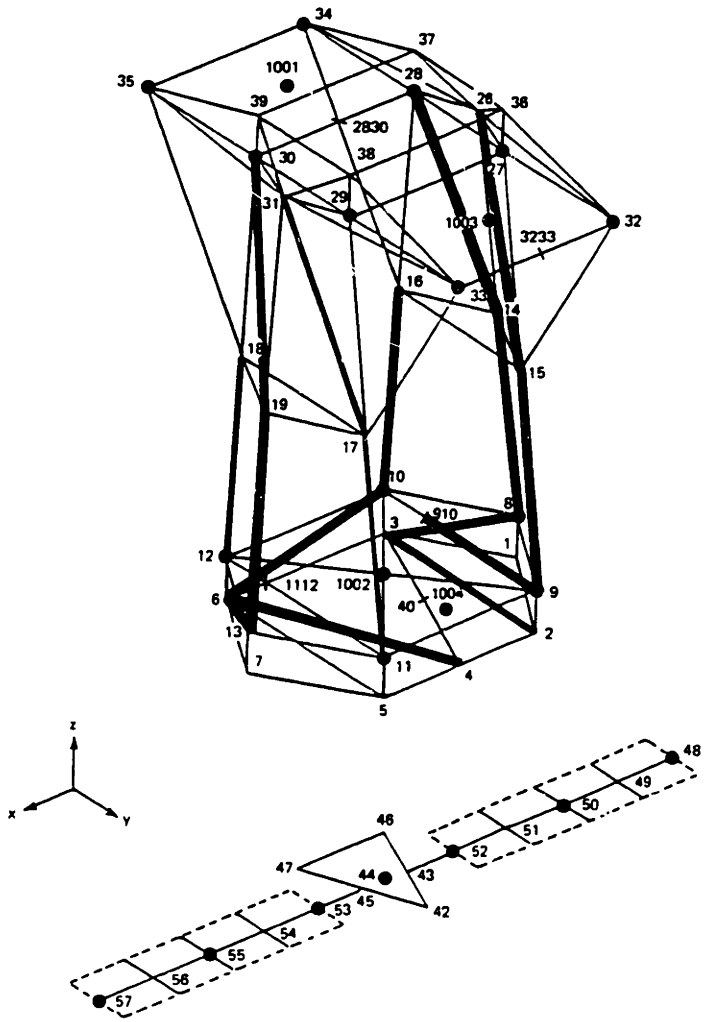


Figure 5-27. Extensional Actuators, '1/14' Constrained

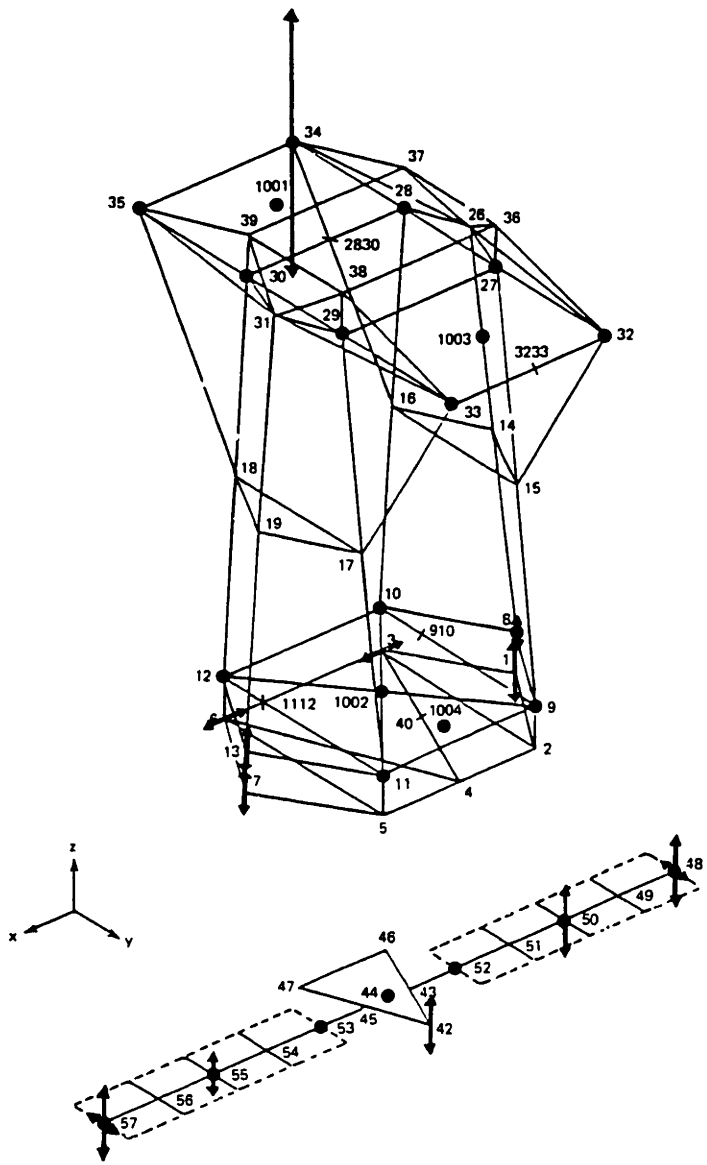


Figure 5-28. Translational Actuators, Fully Relaxed Distribution

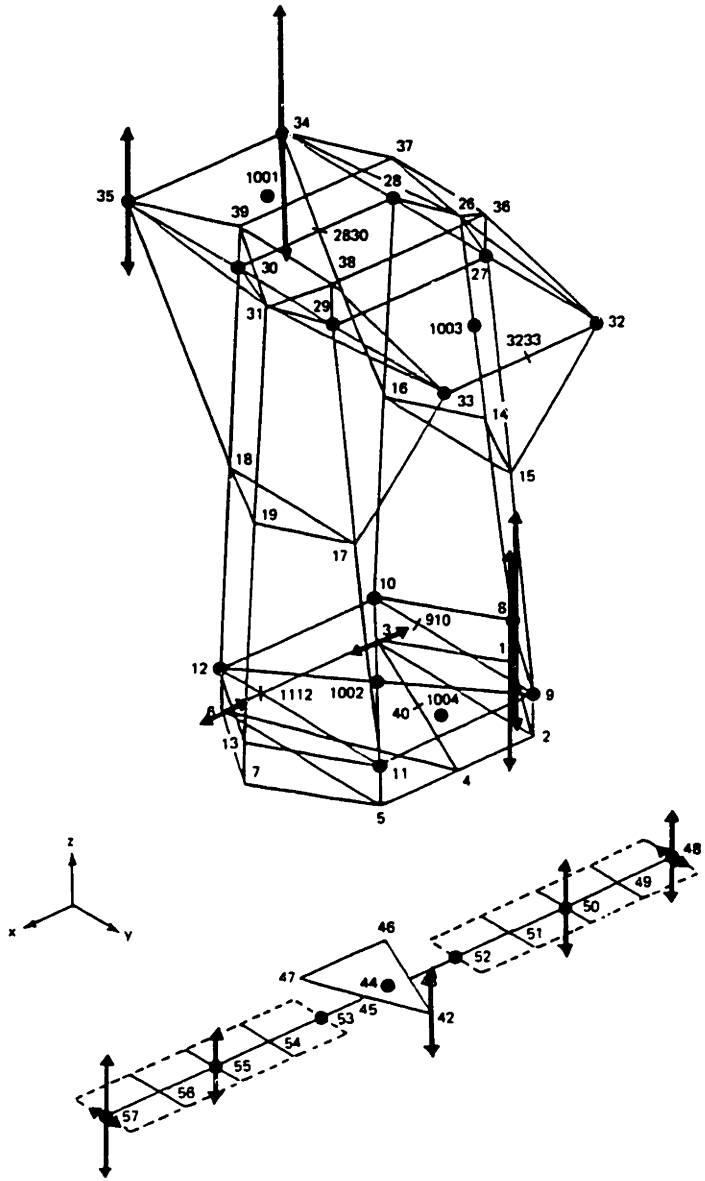


Figure 5-29. Translational Actuators, '1/5' Constrained

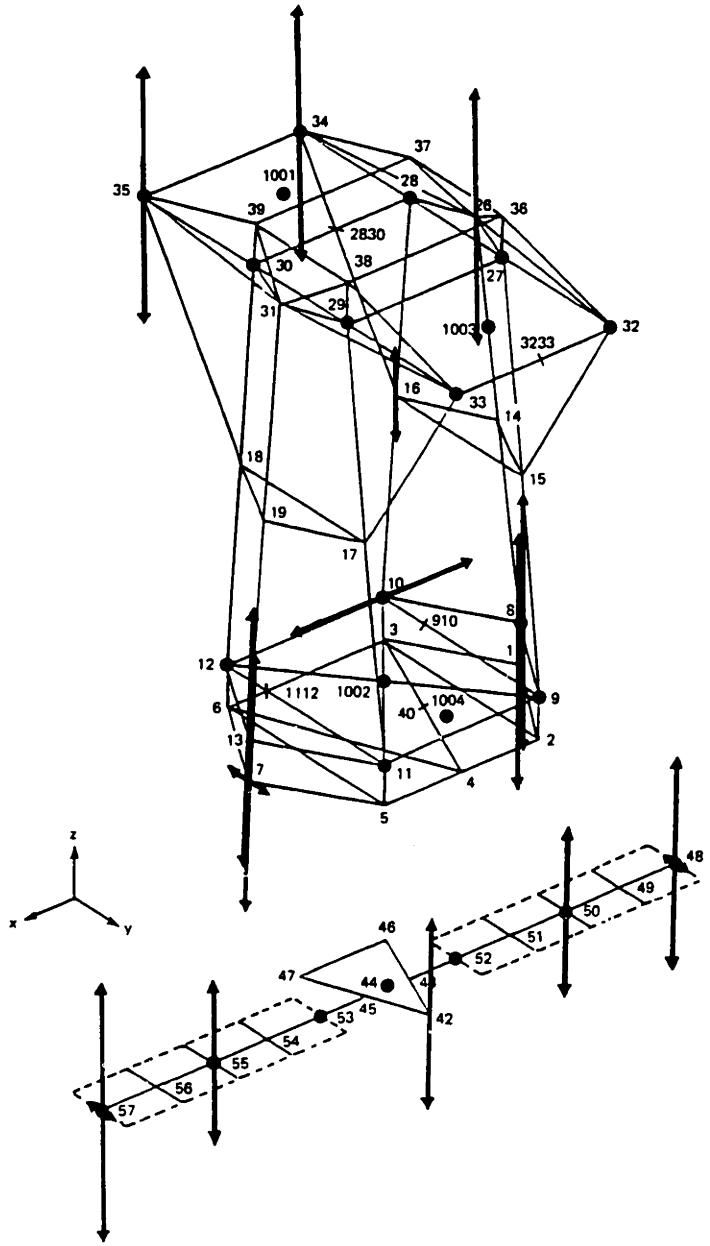


Figure 5-30. Translational Actuators, '1/12' Constrained

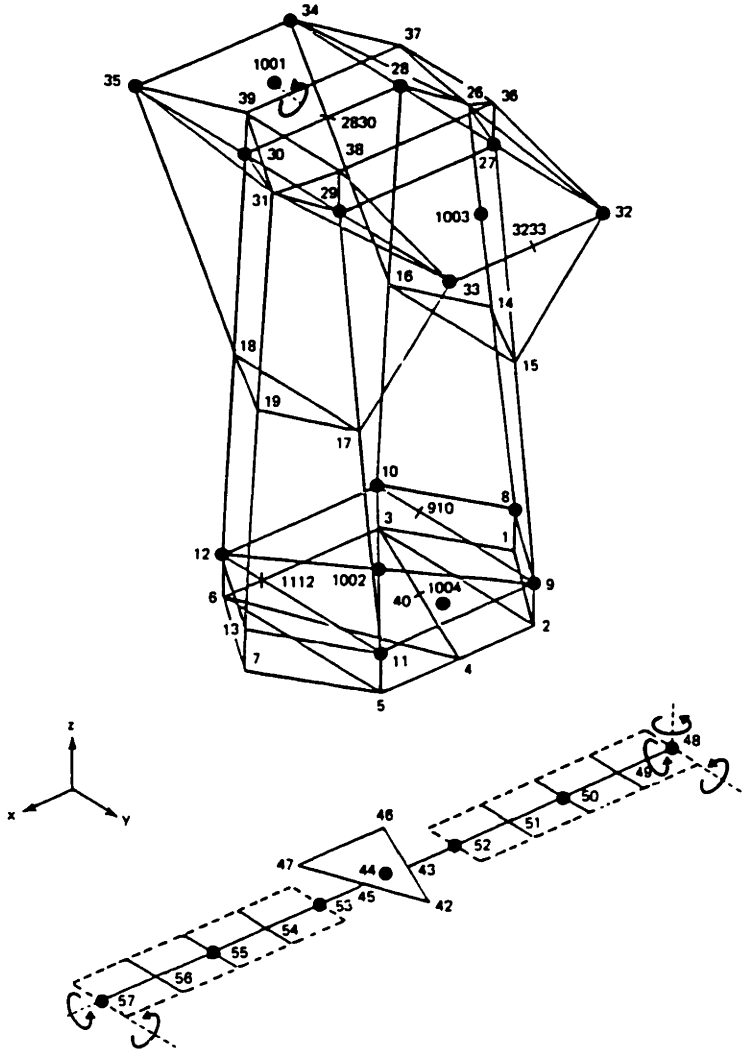


Figure 5-31. Rotational Actuators, Fully Relaxed Distribution

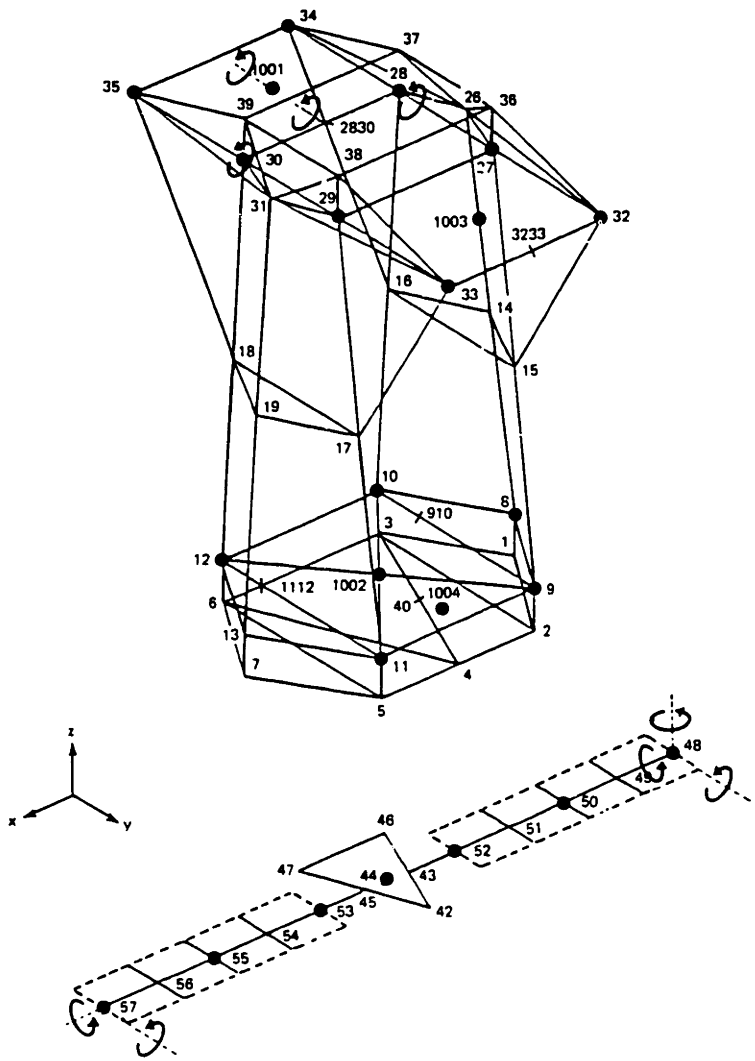


Figure 5-32. Rotational Actuators, '1/5' Constrained

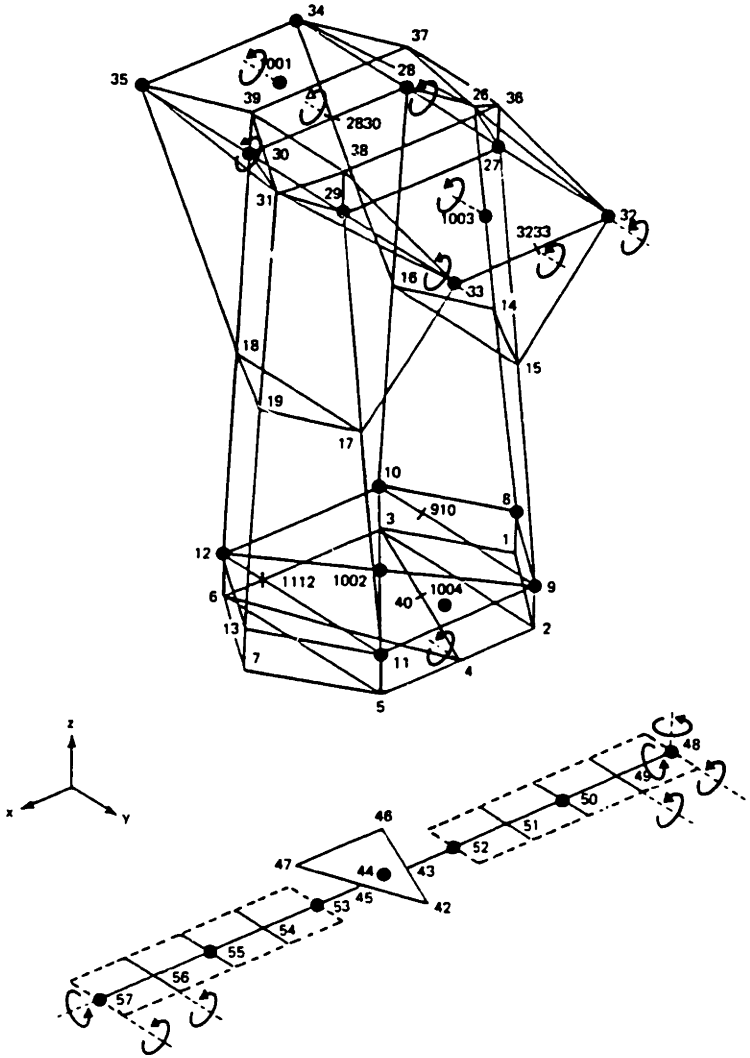


Figure 5-33. Rotational Actuators, '1/11' Constrained

Table 5-11. Extensional Actuator Solutions.

		Figures:	5-25	5-26	5-27
Constraints:		Total Power M:	1	10	14
		Individual m_i:	(n/a)	1	1
Actuator No (i)	Nodes			Distribution (ξ_i)	
25	6 10		.734	1	1
68	10 16		.170	1	1
70	9 15		.096	1	1
7	4 6			1	1
66	8 14			1	1
75	13 19			1	1
78	14 28			1	1
86	19 30			1	1
82	15 26			.596	1
23	6 13			.462	1
28	9 910			.361	1
18	3 8				1
3	2 3			.363	.949
73	12 18			.219	.564
88	17 31				.352
71	11 17				.136
		Total Devices in Soln, $N_a + N_{ul}$:	3	13	16

A significant result emerges from the fully-relaxed (individual constraints m_i not active) solutions for all classes, and that is the optimal distribution of authority is concentrated in relatively few devices. There are 137 potential extensional actuators, but only three remain in the fully relaxed solution of Figure 5-25. Somewhat more translational and rotational actuators are included in their respective fully-relaxed solutions (Figures 5-28 & 5-31, but still only a fraction of the 174 total of each class. This is a particularly practical result, since we would like to be able to control the structure with as few as necessary. Notice that while the structure has some symmetry, the disturbance entering the reference signal path is not symmetric, hence symmetric mode shapes are not equally weighted. As we bring into effect proportionally smaller individual site constraints, the total number of devices increases, but not dramatically. Notice also as the total power increases, devices reach the individual constraints and new ones enter from the inactive partition. A few translation devices apparently defy this expected behavior: devices 260 and 266 as shown in Table 5-12, in particular. Inspecting the figures corresponding to the extensional solutions shows a strong preference for the members linking the upper and lower support trusses. This is not too surprising, since there are mass concentrations on both ends of the main truss, which the optical line of sight traverses several times. We will put off discussing the actual controllability norms associated with these solutions for the

moment, but suffice it to say now that the degradation from the fully relaxed solution to the 1/14 constrained (M=14, m_i=1) solution is about 74% of the initial value.

Most of the structural modes retained as significant in the disturbance-to-LOS signal path are either solar panel motions or modes of the lower equipment bay with the stiffer upper truss across the relatively soft isolator springs between them, as can be seen in the sketches in Appendix B. There were no strong extensional devices for the solar panel modes, so they appeared largely ignored. However, there are translational actuator sites that do affect those modes strongly and they emerge in the next series of solutions, listed in Table 5-12, with corresponding Figures 5-21 through 5-23. Another interesting observation about these solutions is the strong emphasis in the 'z' direction, which supports control of modes across the isolators. There also appears to be a roughly even split between authority for the lower equipment bay with its solar panels and the trusses. Again, we will put off more quantitative discussions for the moment.

Table 5-12. Translational Actuator Solutions.

		Figures:	5-28	5-29	5-30
Constraints:		Total Power M:	1	5	12
		Individual m_i:	(n/a)	1	1
Actuator No(i)	Node	Direction		Distribution(ξ_i)	
281	34	Z	.528	1	1
303	57	Z	.074	.461	1
254	1	Z	.060	.850	1
260	7	Z	.021		1
266	13	Z	.016		1
261	8	Z	.003	.828	1
282	35	Z		.512	1
273	26	Z			1
294	48	Z	.077	.291	.811
147	10	X			.767
288	42	Z	.054	.279	.712
296	50	Z	.077	.279	.620
301	55	Z	.022	.225	.601
140	3	X	.035	.146	
143	6	X	.030	.114	
269	16	Z			.328
202	7	Y			.124
236	48	Y	.001	.010	.033
245	57	Y	.001	.005	.005
Total Devices in Soln, $N_a + N_{ul}$:			14	13	17

Distributions for rotational actuators are shown in Figures 5-31, 5-32, and 5-33, which accompany Table 5-13. The emphasis on the solar panel tips appears again. Note the preponderance of authority in the 'y' direction. In the upper support truss the solutions have found the primary and tertiary mirror masses (nodes 1001 and 1003), which may not be physically admissible. We could go through the model and purge any inadmissible sites then rerun the solutions without them, but did not do so for any of the examples presented here.

Table 5-13. Rotational Actuator Solutions.

		Figures:	5-31	5-32	5-33
	Constraints:	Total Power M:	1	5	11
		Individual m_i:	(n/a)	1	1
Actuator No(i)	Node	Direction		Distribution (ξ_i)	
421	1001	Y	.492	1	1
419	57	Y	.216	1	1
410	48	Y	.164	.908	1
426	2830	Y		.487	1
391	28	Y		.483	1
393	30	Y		.483	1
418	56	Y			1
411	49	Y			1
373	4	Y			1
352	48	X	.082	.273	.604
361	57	X	.027	.273	.604
423	1003	Y			.326
468	48	Z	.018	.095	.216
395	32	Y			.084
396	33	Y			.084
427	3233	Y			.084
	Total Devices in Soln, $N_a + N_{ul}$:		6	9	16

Returning to the issue of relative strength among the three classes of devices, we know it is possible to add a weighting among the classes in order to reflect their relative effectiveness. After all, an extensional member device reacting entirely against the structure ought to be a stronger device than an equivalent 'unit' translational force device, which is idealized as reacting against inertial space. In the current scaling we are in effect stating than 1 Newton of force reacting between two nodes comes at the same price as 1 Newton reacting against free space, and at the same price as 1 Nt-m of torque also reacting against free space. Furthermore, any relative weighting ought also to depend upon frequency,

since the translational and rotational devices are idealizations of hardware that has no DC authority, while the extensional class very well might.

References [64], [65] and [66] contain summaries of a broad range of actuator dynamics and their incorporation into a structural dynamics model. Actuators are divided into two broad classes: transmission and reaction. The distinction stems from the mechanics exploited to generate the control force. Reaction devices directly force some small mass, and the reaction force against the structure is the control force, while transmission devices drive some displacement which elastically transmits a control force. Examples of reaction devices are translational proof-mass actuators or reaction wheels. Figure 5-34 illustrated the signal flow of this class.

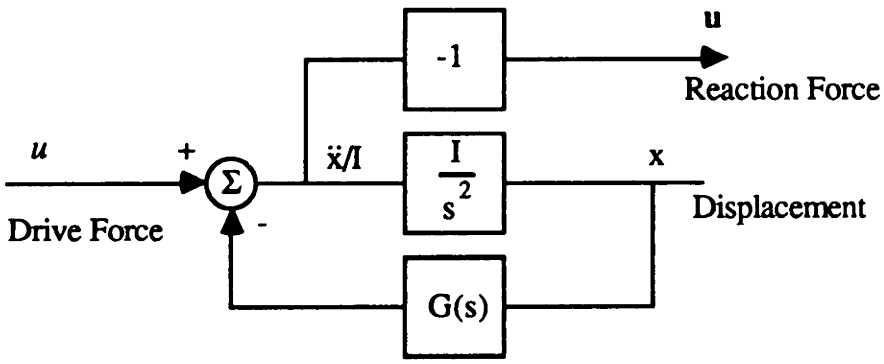


Figure 5-34. Reaction Actuator.

The driven mass has inertia, I, and feedback G(s) represents a combination of device-particular dynamics, such as damping elasticity and non-linear limits, as well as local active compensation. The main point of Figure 5-34 is that the reaction force from accelerating the actuator mass, which always has physical limits on displacement and rate, is the output force, and none can be generated at DC. Note also there is not intrinsically any high-frequency roll-off from the signal path in Figure 5-34, but this would be present, of course, from other dynamics such as electrical driving circuits. For transmission devices, on the other hand, a displacement is important and we diagram the signal flow of this class in Figure 5-35.

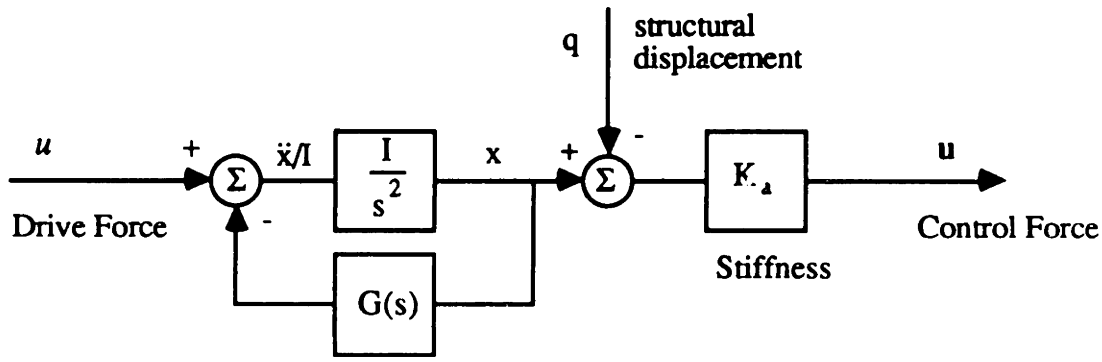


Figure 5-35. Transmission Actuator.

Examples of transmission devices include jack screws, cable-tendon devices, and piezo-electric actuators, the latter having particularly small moving mass, and all of which react against the structure. A structural displacement signal, q, enters Figure 5-35 and represents the fact that the control force depends upon relative displacement. The actuator stiffness, K_a, may be a part of the structure, but the main point is there is authority at DC and a mechanically-induced high-frequency roll-off. Control-moment gyros are a particularly important class of device which do not quite fit either classification. They are closer to a transmission device than to a reaction device, with the interpretation of Figure 5-35 that the gimbal rate and structural rates replace the mass and structural displacements, and the actuator stiffness represents the rotor momentum. More like a reaction device, however, is the fact that a CMG does not have DC authority, but this may be captured in the actuator stiffness by making it a function of the gimbal angle, specifically $K_a = \mathbf{h} \cos(x)$, where $x = 0$ represents the nominal gimbal angle with rotor momentum $\mathbf{h}$ perpendicular to the CMG output axis.

The role of local compensation, G(s), in Figures 5-34 and 5-35, is often overlooked in analyses of sensor and actuator dynamics for structural control. It is not likely that the control signal from the vehicle will be the actual drive voltage for an actuator, which 'sees' the electro-mechanical dynamics of the device itself. Flight hardware incorporates local compensation to make the unit appear as a simple linear gain over a specific bandwidth, and the high-frequency roll-off may also be significantly sharper than that of a simple model. For control synthesis, then, it may well be justified to model sensors and actuators as all-pass constant gains, but evaluation then should include the

locally-compensated device dynamics as well as the most significant unmodeled structural modes. This does not imply, however, that local compensation eliminates the usefulness of frequency-dependent weightings for sensor and actuator selection. The fundamental physical differences of the device classes or different point-designs within a single class imply favoring some frequencies over others. We could construct frequency-dependent weightings, such as shown in Figure 5-36, for devices applicable to the examples of this section.

These weightings may be derived from actuator dynamics, but may reflect other information as well. The roll-on rates of the reaction devices (and the CMGs) model excursion limits of the reacting mass, not linear dynamics. The steeper roll-off of the CMG represents the effect of a gear train commonly used to drive the gimbal. With a family of plots such as in Figure 5-36, we may construct approximating input or output filters, U(s)

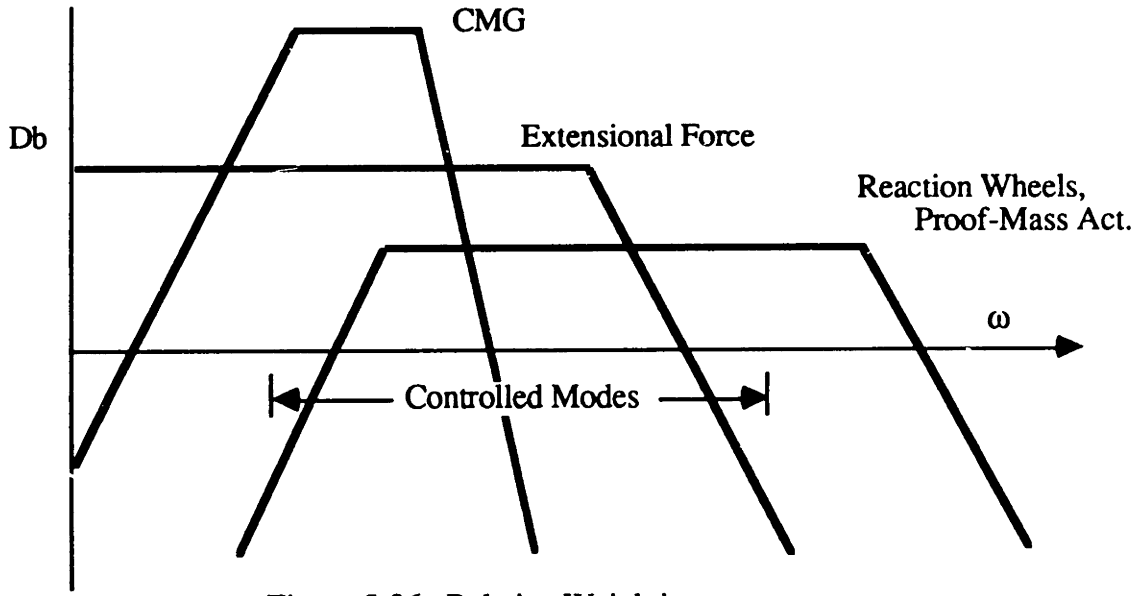


Figure 5-36. Relative Weightings.

and Y(s), respectively, and append them to the sensor or actuator signal paths as shown Figure 3-4. This was done for the reference signal path from which we derived basis scalings, with U(s) representing a spectral filter for the disturbance, but it may also be done, with different U(s), for computing the individual site Gramians.

Deciding what the relative weights ought to be is a bit of an aside, so we continue by considering a simple case with the weights assigned in proportion to the controllability norm of the relaxed solution for each class alone, and no frequency dependence. That is,

the controllability norm of the best solution from Table 5-12 (translational force devices) is about 780,000 that of the fully relaxed solution of the extensional devices, so we weight (up) all the extensional devices by the precise ratio. Likewise, the translational solution's norm is about 37 times larger than that of the rotational devices, so the latter are weighted by that amount. Weightings are actually applied to the individual Gramians, or equivalently the square roots of the weights are applied to the corresponding influence coefficients. We expect the result will be to bring all classes of devices into the solution, and this is in fact the case, as detailed in Table 5-14 and Figure 5-37. Notice that the strongest devices from the fully-relaxed solutions of each class are not present in Figure 5-37, but the emphasis on solar panels and main truss extension and bending remains. The effect of including all classes, albeit with some positive weightings, is a controllability norm 4.1 times larger than that of the fully-relaxed translational solution alone, which was the class retaining unity weighting in this example. We would expect exploitation of the different classes to yield better solutions than a single class, but interpretation requires more engineering meaning behind the relative weights rather than just to create an interesting result as in this case.

Table 5-14. Combined Actuator Solutions.

Figure: 5-37

Constraints: Total Power M: 100

Actuator No (i)	Node(s)	Direction	Class	Distribution (ξ_i)
236	48	Y	Translation	60.39
71	11	17	Extension	14.01
294	48	Z	Rotation	13.37
303	57	Z	Extension	10.62
361	57	X	Rotation	1.19
68	10	16	Extension	0.42

Total Devices in Soln, N_a: 6

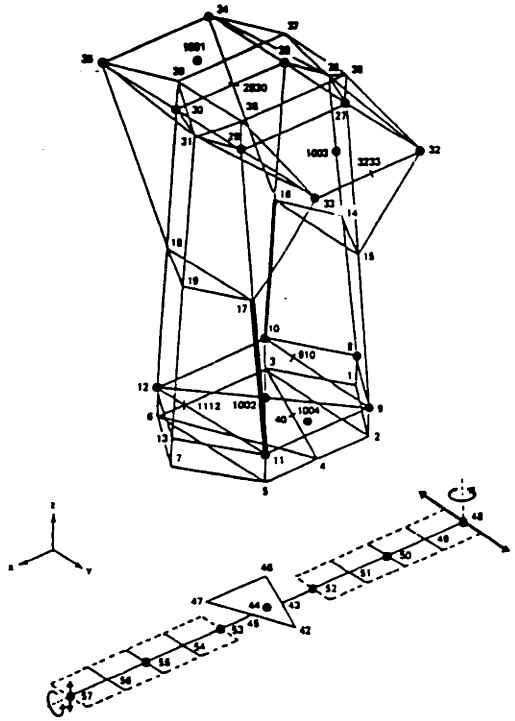


Figure 5-37. Fully Relaxed Distribution for All Classes

We would like some evaluation of these solutions, other than their respective controllability norms, before proceeding with the expensive step of combinatorial search for fault-tolerant configurations. We turn to the LQ regulator for such a check, motivated as in the beam examples by its simplicity and the fact the stochastic performance of the full-state feedback regulator is an upper bound to the performance of an LQG model-based compensator operating on noisy measurements. Furthermore, we have the opportunity to compare results with actuator selections by Hegg and Kissel [66] using Fogel's [42], [22] modal-influence coefficient algorithm, as reviewed in section 3.2.2. Five of their selections are extracted for comparison here- 'Examples A, B, C, E-5 and E-11', in their nomenclature- and are plotted on Figures 5-38 through 5-42, respectively. For each of their examples, Hegg and Kissel computed the full state feedback regulator based on LOS output weighting and uniform control weighting, with a cost functional in the form

$$J_{LQ} = \lim_{t \to \infty} E\left\{ \frac{1}{t} \int_0^t \left[z^T(\tau) Q_z z(\tau) + r_u u^T(\tau) u(\tau) \right] d\tau \right\} \qquad 5\text{-}37$$

where Q_z = diag $[q_x\ q_y\ q_z]$ for the three components of the optical line of sight, and r_u is a scalar. For each of their examples, Hegg and Kissel experimented with the state weightings until the stochastic response of the system to the disturbance spectrum of Figure

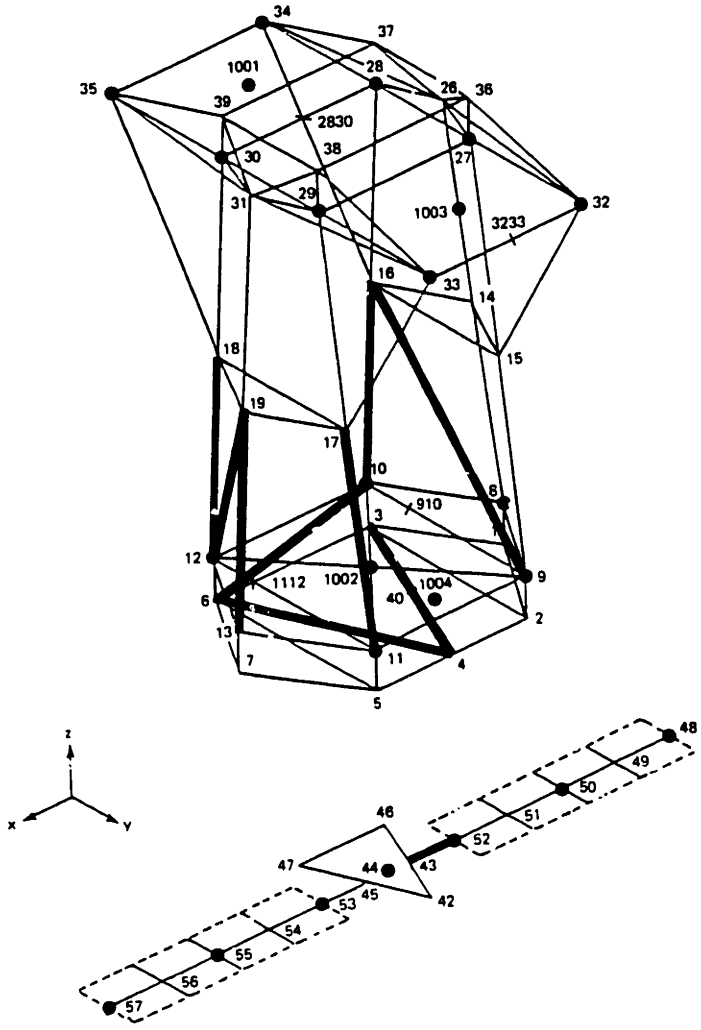


Figure 5-38. HK, Example A

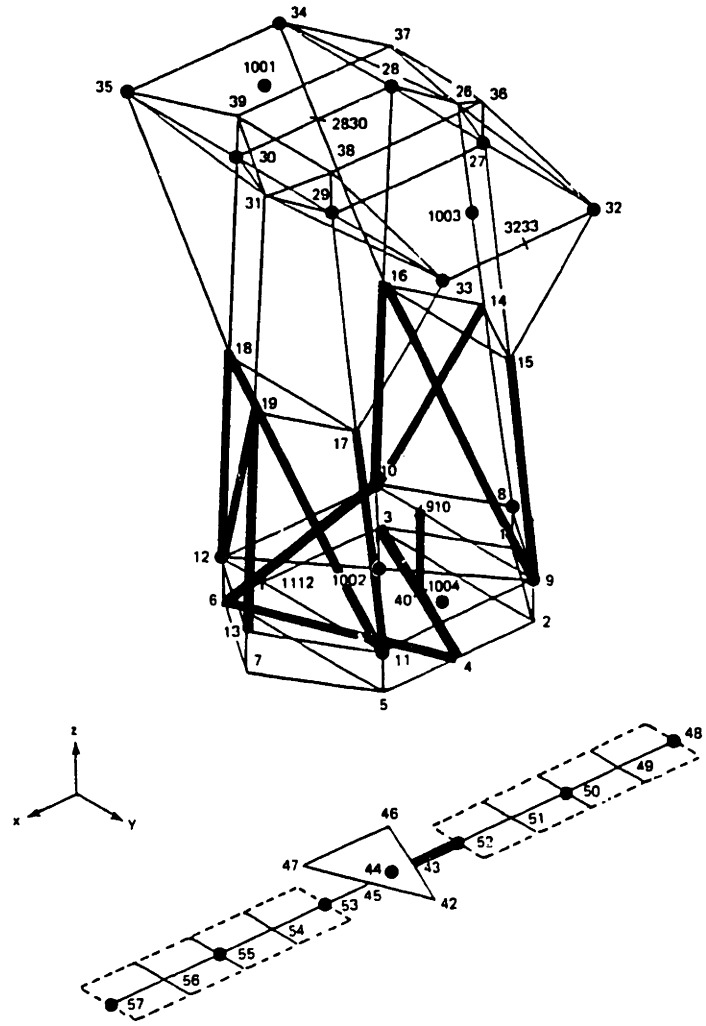


Figure 5-39. HK, Example B

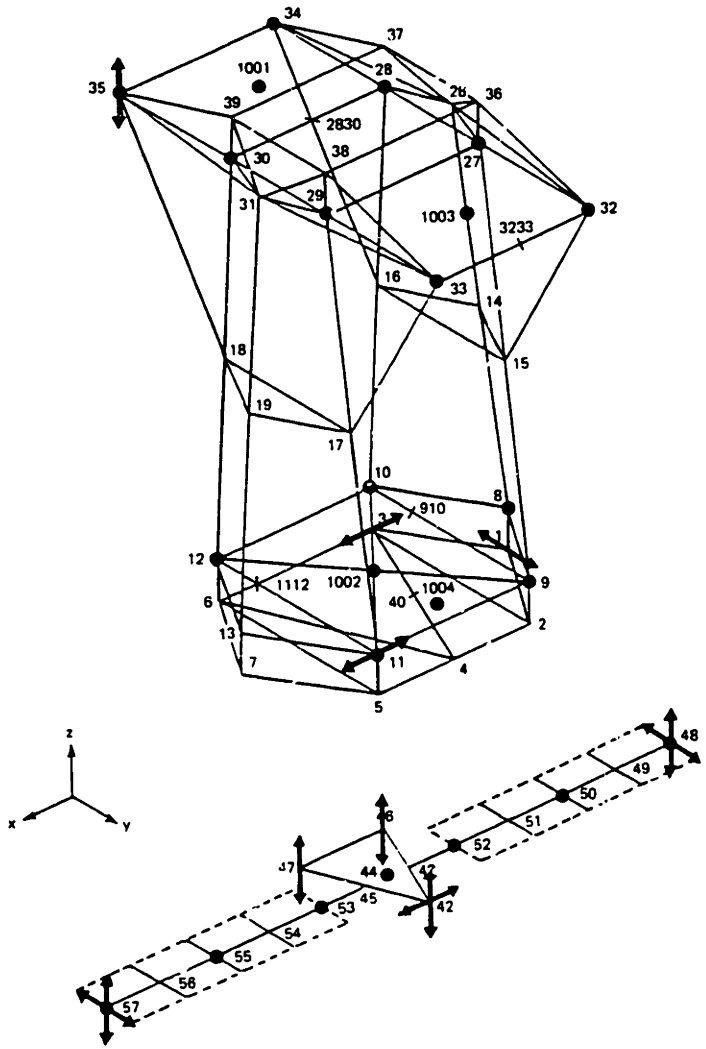


Figure 5-40. HK, Example C

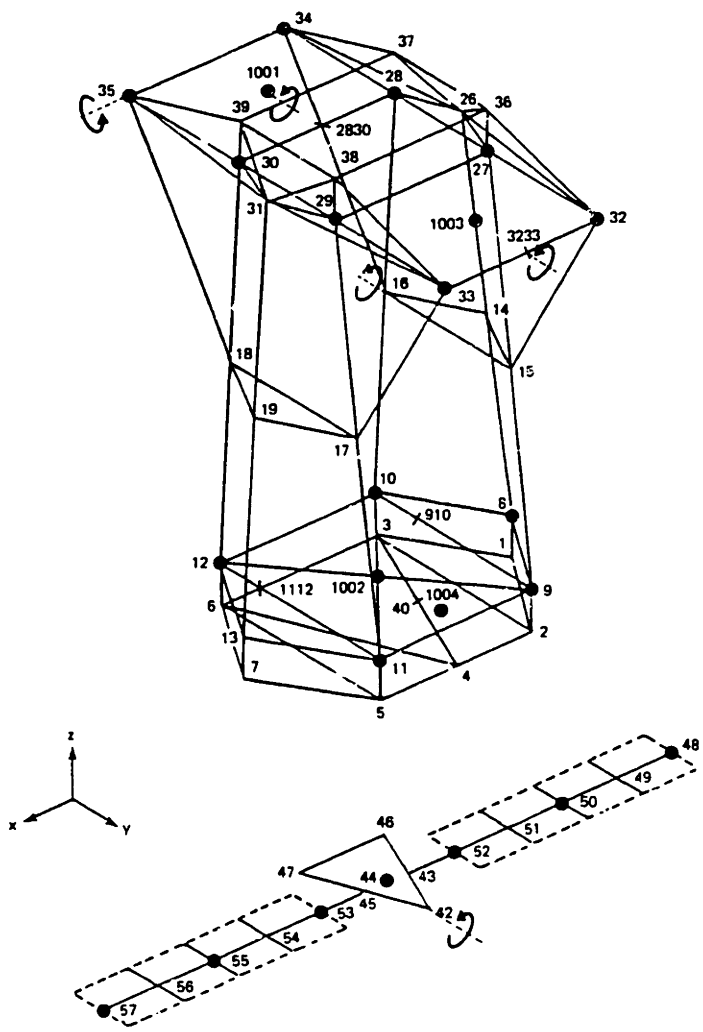


Figure 5-41. HK, Example E-5

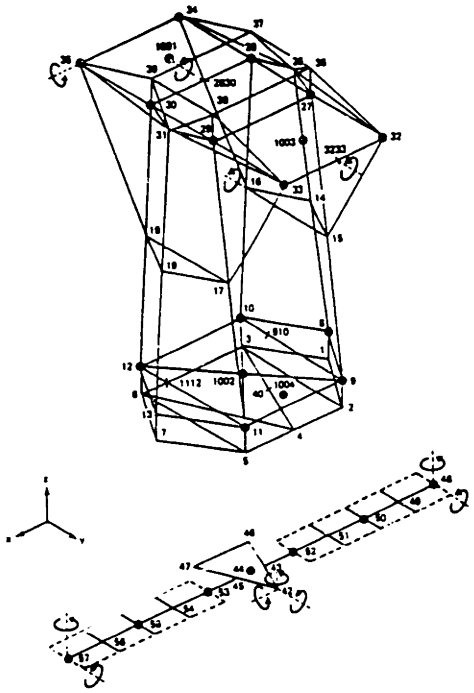


Figure 5-42. HK, Example E-11

5-15 met the pointing requirement of Table 5-8, of which the transverse requirement of 0.05 μrad was the stressing specification. Later they experimented with more modes in the evaluation model than the design model and more sophisticated (and realistic) compensators, such as those resulting from the LQG/LTR methodology, and considered a host of other important concerns with control of flexible structures. We will use their LQ regulator results as a point of departure for experimenting with different actuator configurations, specifically the various relaxed solutions. We fix their LQ weights for a class, and compare the closed-loop pointing performance, including control effort and the overall LQ cost. Only the 11 controlled modes are included in the regulators. Several reasons justify this. First, the purpose of the comparison is the relative performance of different actuator configurations. Suppose we design the regulator with respect to the controlled modes only, then evaluate it against some larger model. The unmodeled modes cannot be destabilized, since they are not feedback in the idealization of the regulator. They will respond to control signals (control spillover) and to the disturbance, and the latter response is the same for any actuator configuration. Some checking reveals the dominant term for these examples in the LQ cost (eqn. 5-37) is the control signal, and the value of the cost varies only a few percent as more modes are introduced into the evaluation model in the order of Table 5-10. More important, however, is the fact that the relative ordering

between the different actuator configurations does not change by including more modes. Second, we wish to conduct some Monte-Carlo experiments and some evaluations which enumerate every failure mode, and are interested in computational economy. We are assuming, in this case, that the mode selection done by the reference disturbance-to-LOS signal path is an appropriate choice for regulator design, and in this case it appears to be so. If not, it is certainly possible to revisit the mode selections, using the new actuator-to-LOS signal path as well as the disturbance-to-LOS signal path for mode rankings. Finally, while a more realistic evaluation might be feedback from the linear estimator and physical signals rather than direct modal feedback, this would be a significantly longer sequence of steps for which the LQ regulator represents an upper bound of performance.

Table 5-15 is a summary of LQ regulator closed-loop comparisons of the several relaxed solutions of Tables 5-11, -12, and -13, and five comparable cases, 'A', 'B', 'C', 'E-5', and 'E-11', from reference [66], plus a few other actuator selections we will discuss momentarily. The first columns list the total and individual device power constraints, which while strictly applicable to the relaxed solutions only, show that they match the total authority of the examples of Hegg and Kissel. The third column shows how the solutions differ; with continuous admissible power levels there are almost always more active devices in the relaxed solutions than the minimum. The output weightings chosen by Hegg and Kissel are listed, and their control weight, $r_u = 10^{-17}$, was fixed for all cases. Notice how, in order tc obtain roughly the specified pointing performance, the output weights are reduced for the weaker actuator classes. The stochastic pointing performance and control effort are listed in the central columns in physical units. The LQ cost, J_{LQ}, is not comparable across classes of devices because different weightings were used for each class, so the relative cost with respect to the best solution among those listed within a class is displayed. In the same manner, the relative reachable volume controllability norms are listed in the right-hand columns with one last caveat, they are also scaled by the total power, M, as well as with respect to the best (always the fully relaxed) solution in the class. Dividing the norms by M allows comparison of configurations with differing total power within a class. For a fixed distribution, the absolute size of the norms always increases linearly with the total power. We list both the average energy related norm (J_2) and the total volume norm (J_3) for the extensional cases, because one solution using the

Table 5-15. Summary of Closed Loop Comparisons

Figure 5-		Constraints M	Constraints m_i	Number of Actuators N_a+N_{ul}	LQ Weights LOS Output q_x	q_y	q_z	Closed Loop LQR perfomance, rms x	y	defocus	control	J_{LQ}	Controllability Norm $[n/tr(W^{-1})]/M$	$[\|W\|^{1/n}]/M$
								10^{-2} μ-rad		μ-m	10^3 Nt	(relative within device class)		
	Extentional Actuators						Spec:	5.00	5.00	1000				
38	Hegg/Kissle A	10	1	10	10^3	10^1	10^{-4}	2.33	1.57	1.89	3.89	1.15	.516	.673
25	Fully Relaxed	10	n/a	3				11.53	3.72	3.41	6.39	1.44	1	.650
26	1/10 Constrained	10	1	13				3.77	1.59	2.37	4.64	1.37	.799	.686
	Det. Norm, 1/10 Constr.	10	1	10				1.96	.995	2.23	3.54	1.04	.660	.896
46	Monte-Carlo (LQ & J_3)	10	1	10				5.64	3.81	1.79	5.72	1.71	.253	.401
	Monte-Carlo (J_2)	10	1	10				6.08	3.44	2.80	5.96	1.76	.327	.366
39	Hegg/Kissel B	14	1	14				2.05	1.15	1.97	3.39	1	.440	.616
25	Fully Relaxed	14	n/a	3				8.72	2.74	3.41	5.44	1.62	1	.650
27	1/14 Constrained	14	1	16				2.73	1.05	2.37	3.90	1.15	.742	.620
	Translational Actuators										10^0 Nt			
40	Hegg/Kissel C	12	1	12	10^0	10^{-1}	10^{-5}	.557	.857	.954	35.73	2.54	.221	
28	Fully Relaxed	12	n/a	14				.350	.212	.836	15.16	1.08	1	
30	1/12 Constrained	12	1	17				.242	.106	.716	14.08	1	.963	
47	Monte-Carlo best	12	1	12				.146	.208	1.06	17.34	1.23	.515	
	Rotational Actuators										10^2 Nt-m			
41	Hegg/Kissel E-5	5	1	5	10^1	10^{-1}	10^{-5}	5.51	4.94	2.94	2.60	2.21	.0039	
31	Fully Relaxed	5	n/a	6				4.27	8.83	1.22	2.03	1.73	1	
32	1/5 Constrained	5	1	9				3.73	5.06	1.20	1.85	1.57	.956	
42	Hegg/Kissel E-11	11	1	11				1.82	4.91	0.85	1.36	1.14	.510	
31	Fully Relaxed	11	n/a	6				2.41	4.21	1.22	1.40	1.18	1	
33	1/11 Constrained	11	1	16				1.98	2.24	1.22	1.23	1.04	.912	
48	Monte Carlo (LQ)	11	1	11				1.01	5.22	1.26	1.20	1	.478	
	Monte Carlo (J_2)	11	1	11				1.60	3.89	0.86	1.26	1.05	.479	

latter is included. For a total power constraint of $M \leq 10$ and individual constraints of $m_i \leq 1$, the optimal distribution with respect to the total volume norm happenes to be in a 'corner', resulting in an all-integer solution with an empty active partition ($N_a = 0$ and $N_{ul} = 10$) at convergence. This is a curiosity, but the reason the solution was included is that it happens to be better, with respect to the LQ cost, than any of the other $M = 10$ class configurations encountered, so it is of interest to inspect that controllability norm (J_3) for the other sets of 10 extensional actuators as well. The design and evaluation model for all cases was the same 11 modes selected by the disturbance-to-LOS signal path ranking, and, since the state weighting is defined in terms of physical outputs, the LQ results are independent of the basis. A reader comparing Table 5-15 with the original results by Hegg and Kissel (Table 6-1 in [66]) will note the LOS pointing reported by the latter is not exactly the same. We suspect the discrepancy stems from the use of larger evaluation models than 11 modes, but are not certain, and the differences are small. With this somewhat lengthy list of details in mind we may now discuss some findings from these comparisons.

One question worth posing from the results on Table 5-15 is whether or not the additional computational cost of using the Gramians of dimension n^2 instead of influence vectors of dimension n is worthwhile in terms of producing superior solutions. The answer, at least with respect to the particular LQ regulator used for the evaluation is 'sometimes, but not always'. The translational and rotational cases illustrate how actuator distributions of Tables 5-12 and 5-13 allow better line-of-sight pointing with lower control effort than the comparable selections of Hegg and Kissel. In the case of the translational configurations of Table 5-12 (and Figure 5-30), an overall LQ cost reduction of a factor of 2.5 is obtained with respect to Example 'C'. This was the largest margin noted with respect to any of the selections by Hegg and Kissel [66]. For the extensional actuator class, however, distributions based on the average energy (J_2) norm are not any better, although the total volume (J_3) optimal distribution is. There was no general superiority observed for solutions using that norm for other actuator classes, however.

Under the 'Controllability Norm' columns of Table 5-15 we may see the relative measures of the reachable volume within the three actuator classes. The 'fully-relaxed' solutions are, of course, always the best within a class, and the comparisons are

normalized to that solution. Note how, as the proportion of the total power allowed at an individual site is lowered, i.e., as the ratio of m_i/M decreases, the optimal size of the norm falls. The worst-case loss imposed by the individual constraints noted was 26% when any extensional actuator was restricted to 1/14 of the total power, and the minimum impact was for the translational class. While we knew the controllability norms could not increase as the feasible region of actuator authority is constrained, the resulting regulator performance, in all cases noted here, actually improved (J_{LQ} reduced) as the individual site constraints tighten. We may also see that all of the selections by Hegg and Kissel had controllability norms from 22% to 52% that of the relaxed solutions, with one notable exception. The five-rotational actuator selection 'E-5' had a controllability norm of less than 1% that of the fully-relaxed solution. The resulting regulator performance was quite a bit closer, though still inferior to the performance of the fully-relaxed distribution. Apparently the controllability norms do not follow the LQ regulator performance particularly closely. In an effort to see a bit more into just how the two parameters correlate we turn again to a Monte-Carlo experiment.

In a manner similar to that used for evaluating the fault-tolerant actuator distributions for the beam, we select three random sets of actuators. Each set has 100 unique configurations of actuators, and the number and class of actuators within a set are comparable to examples 'A', 'C', and 'E-11' of Hegg and Kissel [66]; that is, configurations are of 10 extensional, 12 translational and 11 rotational devices. The probability of selecting any particular site for inclusion in any particular configuration was uniform, although configurations with any duplicate sites were later eliminated, as would have been any entirely identical configurations, had the latter occurred. Using the output weights of Table 5-15, the closed-loop performance of the LQ regulator and the open-loop controllability norms were computed for all 100 configurations within a set. Unlike the earlier Monte-Carlo experiment, these do not include actuator failures. We plot the controllability norm against the regulator cost in Figures 5-43, 5-44, and 5-45. Specifically, the average energy norm (J_2) in the 'throughput-scaled' basis is plotted, which was the parameter used for most of the solutions presented. Normalization of both axes is with respect to the best case encountered in the particular set, and the $\log_{10}$ scale was employed for the controllability norms of the translational and rotational sets for

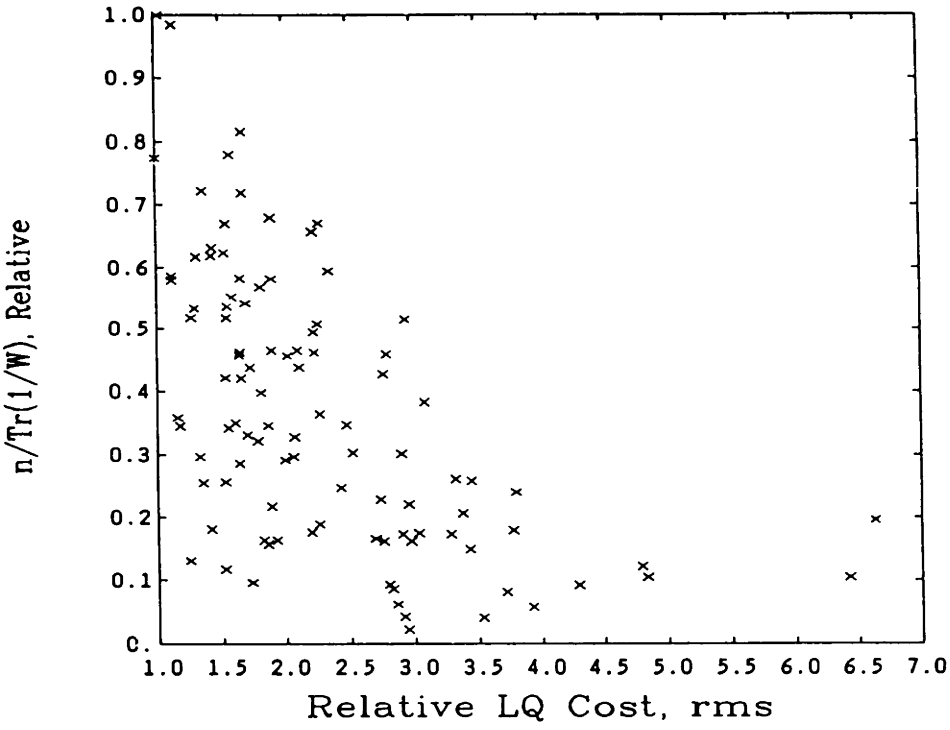


Figure 5-43. Extensional Actuators, Throughput Scaled

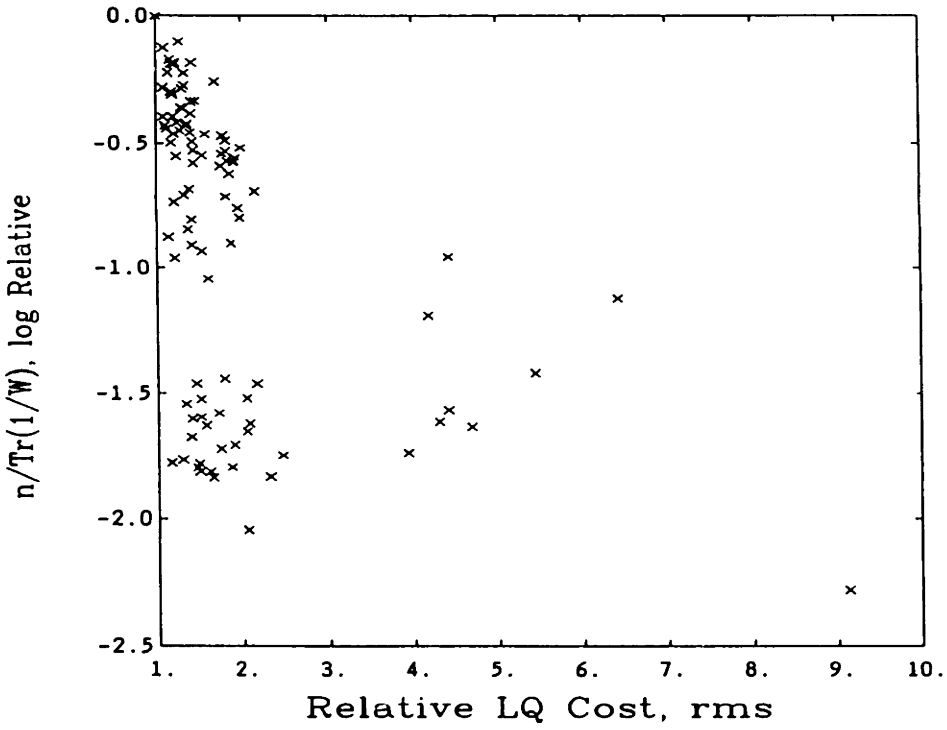


Figure 5-44. Translational Actuators, Throughput Scaled

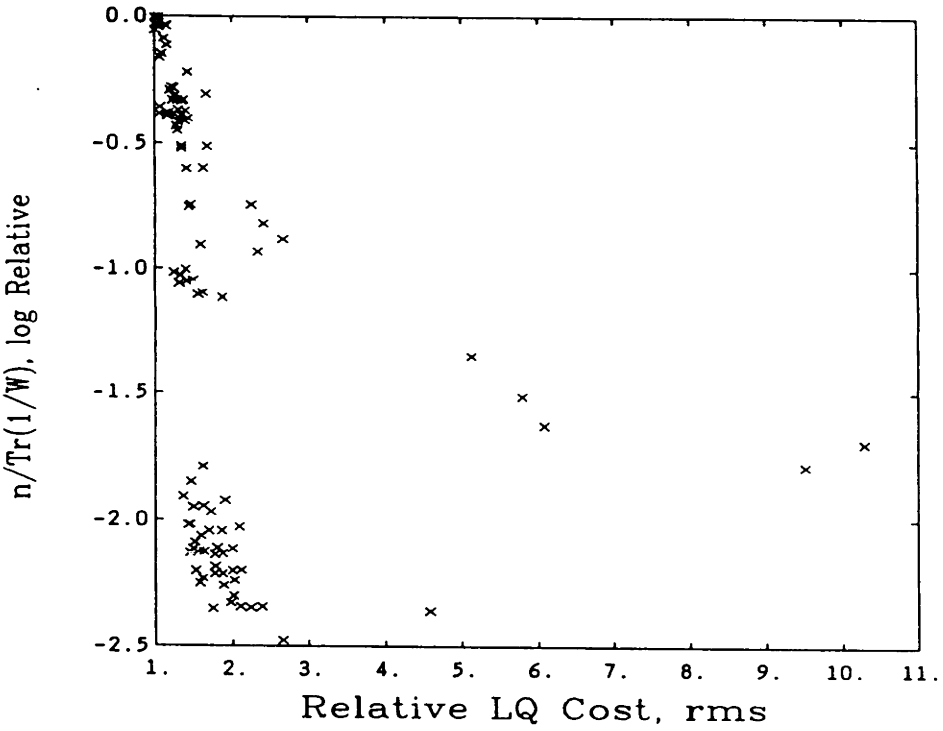


Figure 5-45. Rotational Actuators, Throughput Scaled

illustration only. Notice that while there is no clear functional relation between the LQ cost and the controllability norms, nor had we a theoretical expectation of one, there does appear to be a useful correlation. The relatively dense populations of instances in the lower left corner of these plots indicate 'missed-opportunities', configurations which, while having a low controllability norm, still allow a good regulator. The empty upper right corner is more important, however, illustrating how a good controllability measure appears to insure against a bad LQ regulator. The best configurations, with respect to the LQ cost, are illustrated in Figures 5-46, 5-47, and 5-48, and a summary of these is also included on Table 5-15. None of the randomly selected extensional or translational configurations produced better regulators than either the fully-relaxed or constrained distributions; however, for the rotational class, 3 of the 100 produced a lower LQ cost than the corresponding best relaxed solution (with 1/11 individual constraints). On the other hand, 11 of the rotational actuator configurations resulted in lower LQ cost than example 'E-11' by Hegg and Kissel, and nearly all, 96 of the 100, translation configurations had lower cost than example 'C'. It seems if the LQ cost (for no failures) is truly the parameter to be

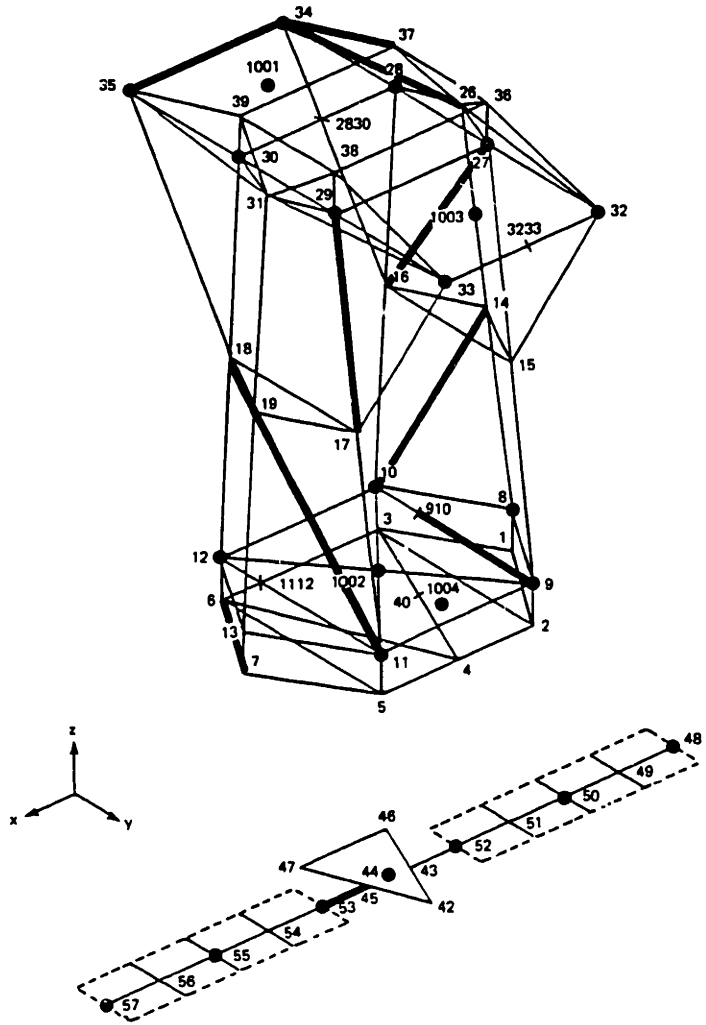


Figure 5-46. Extensional Actuators Monte-Carlo Best

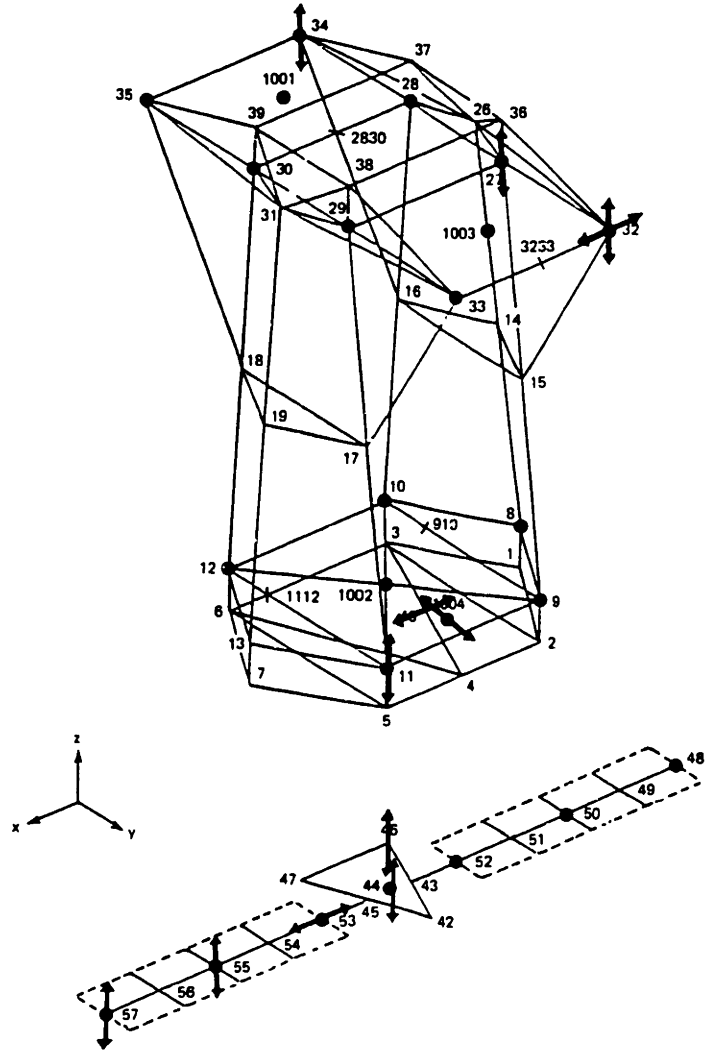


Figure 5-47. Translational Actuators, Monte-Carlo Best

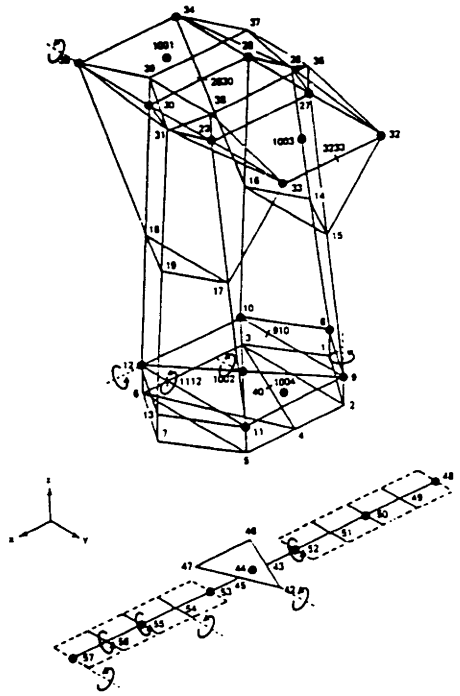


Figure 5-48. Rotational Actuators, Monte-Carlo Best

optimized, a little random searching may turn up about as good or better a configuration for this system as either the approach in this thesis or Fogel's algorithm [42] as implemented by Hegg and Kissel [22,66], although the latter fared worse against the Monte-Carlo experiment. Notice, also, that with respect to the controllability norms, the best of the random samples ranged from about 40% to 50% of the theoretical best represented by the individually-constrained solutions. Of course, if the LQ cost is the parameter, then one could construct an algorithm with it as noted in section 3.2.3.

It is relatively easy to change the basis of the controllability Gramian for any particular actuator configuration and recompute the norms. Since the output-weighted LQ regulator is independent of basis, this provides a means to examine the effect of the basis scaling and the norm selected. This was in fact done, and a few cases are reviewed here. While the correlation between LQ cost and the J_2 controllability norm may appear coarse in Figures 5-43 through 5-45, it is possible to make it much worse with an inappropriate basis scaling. For example, in developing several candidate scalings in Chapter 3 we pointed out the basis called 'output-normal' (with $W_C=\Sigma^2$ and $W_O=I$) was quite inappropriate for actuator optimization. Figure 5-49 is a plot similar to 5-45 except the controllability norm has been computed in the 'output-normal' basis. Notice the loss of correlation and presence of some 'upper right corner' instances of configurations with a large

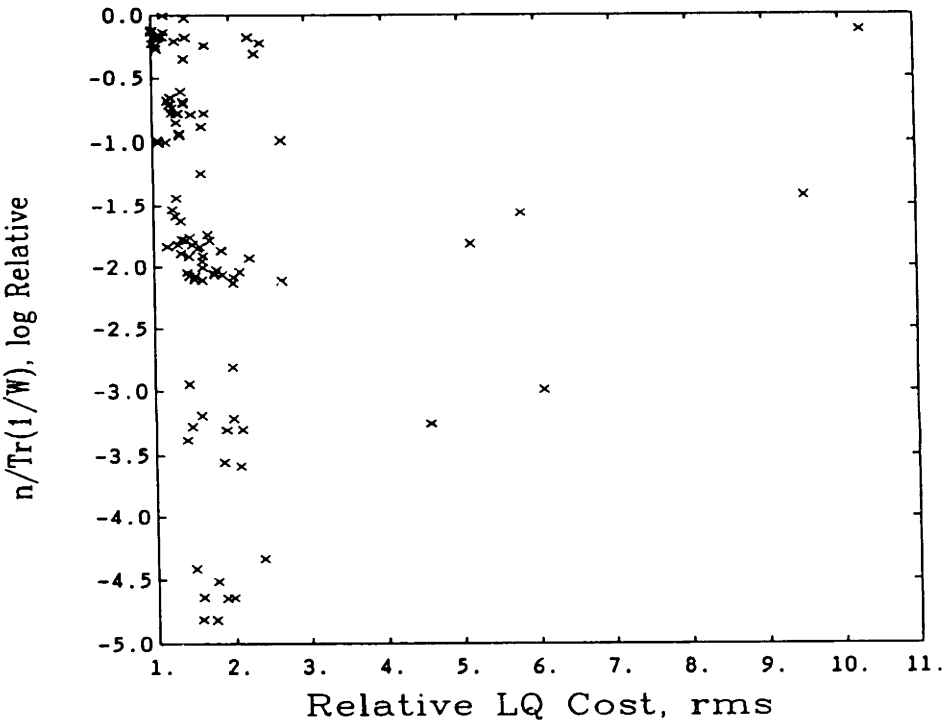


Figure 5-49. Rotational Actuators, Output Normal Basis

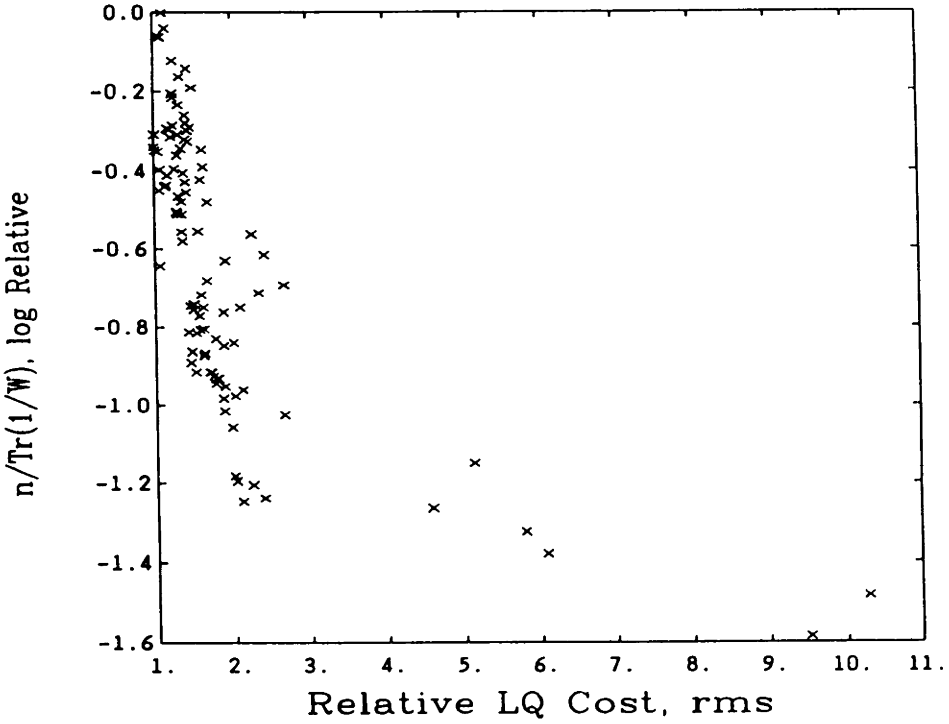


Figure 5-50. Rotational Actuators, Output Scaled

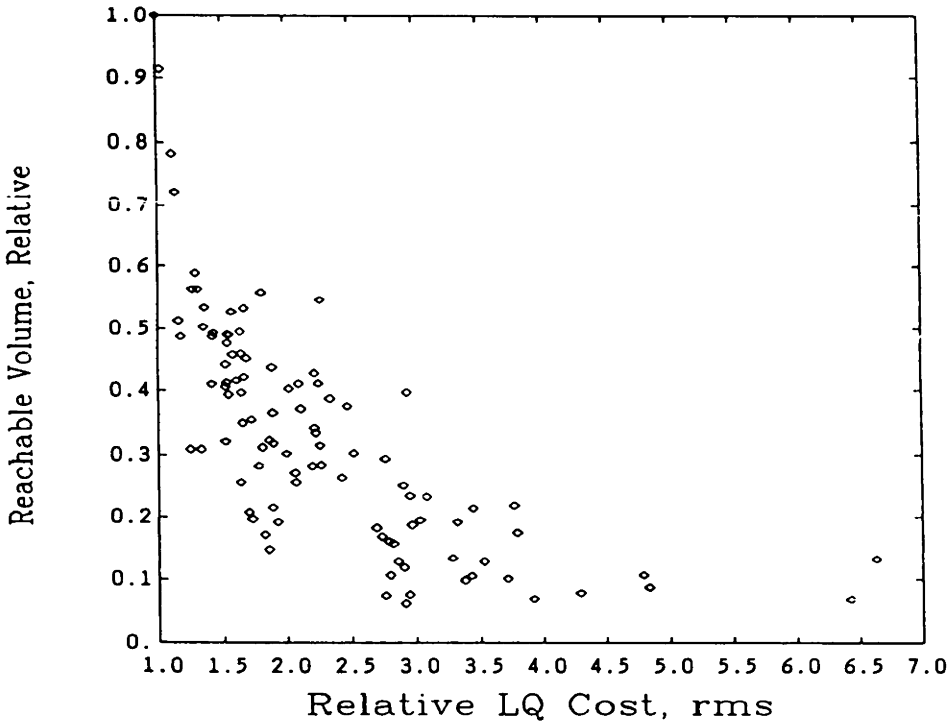


Figure 5-51 Extensional Actuators, Throughput Scaled

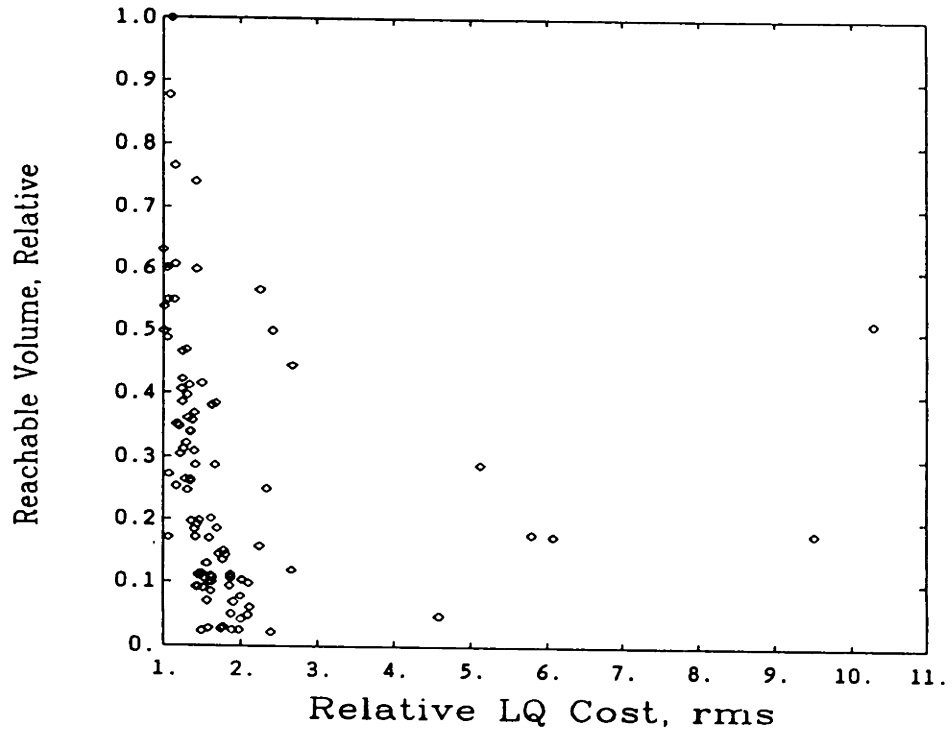


Figure 5-52. Rotational Actuators, Throughput Scaled

controllability norm but particularly poor regulator performance. The basis 'output scaled', which selects the disturbance observability Gramian as the normalizing surface, results in a plot, Figure 5-50, as well correlated as that of 'throughput-scaled'. Indeed, all of the basis scalings of Table 3-1 suggested for actuator selections (input normal, throughput scaled and output scaled) produced reasonably well correlated plots with the minimum eigenvalue (J_1) and average energy (J_2) norms, the latter appearing slightly sharper. Several inappropriate basis scalings were also examined, such as the unscaled modal variables, the 'balanced' basis, and 'output normal', and most of them produced poor to bad correlations. One of the most nicely correlated plots was for the total volume (J_3) norm and the extensional actuators, as shown in Figure 5-51. As can be seen from Table 5-15, the optimal distribution for this norm also produced the best regulator performance in its class, and Figure 5-51 indicates that norm is a particularly sharp indicator for extensional actuators. Recall, however, that the J_3 norm is unaffected by basis scaling, and while Figure 5-51 was actually computed in the 'throughput-scaled' basis, the plot looks the same for any other basis scaling. For the class of rotational actuators, the J_3 norm is not as sharp as the J_2 norm, as can be seen by comparing Figures 5-52 to 5-45, and no rescaling can change it. Considering all of the combinations of scaling, controllability norms, and actuator classes examined, no selection of norm and basis appeared consistently better than the average energy norm (J_2) and 'throughput scaling'. This result was also observed in the beam examples.

We have spent some time examining the relaxed solutions because they are relatively inexpensive to compute and allow the development of some confidence in their relation to one form of closed-loop compensation. All of the configurations used for comparison, either those of Hegg and Kissel or from the random selections, were 0-1 integer-constrained, so to make the comparisons entirely even it would have been appropriate to search for the best 0-1 integer-constrained solutions from the relaxed solutions. This is a relatively expensive step, however, and it is an attractive feature of the overall approach of this thesis is that some useful insights can be obtained before initiating a combinatorial algorithm.

5.2.3 Fault Tolerant Solutions

We turn next to considering two examples of fault-tolerant and 0-1 integer-constrained actuator selections: the 10-element extensional and 5-element rotational classes. We will search for distributions that are best for the worst-case of up to two failures. In the case of the five rotational actuators, the component reliability and end-of-life (EOL) requirements could be similar to those detailed in Table 5-5. We have not carefully addressed the question of determining just what the minimum number of functional components need be at EOL, but assume it has been determined, perhaps with a sequence of individual site constraints on the fully relaxed solution coupled with some other evaluation such as was done with the LQ regulator for Table 5-15. For the extensional case, requiring a minimum of eight out of 10 to be EOL functional could be the result of a system reliability requirement of 0.95 with independent actuators of reliability 0.9113, which in turn is representative of components with a MTTF 11 times longer than the mission duration. We elect to consider only the worst-case failure weighting here, because these are expensive searches. Furthermore, with the expected-value weighting and for failure levels a small fraction of the total, the optimal distribution tends to follow the no-failures distribution closely. In the case of the five rotational actuators there are 10 failure modes of two failures to be examined for each potential exchange, and 45 modes for the extensional case.

Starting configurations are derived from the '1/5-constrained' relaxed solution of Table 5-13 for the rotational class, and the '1/10-constrained' relaxed solution of Table 5-11 for the extensional class. The non-integer components of the active list are lumped into candidate 0-1 distributions, and the Euclidean distance from the relaxed solution used as a metric. Since the search algorithm is '1-exchange', we also eliminate the starting candidates that are '1-exchange' apart; that is, we require at least two sites to be different between all starting configurations. A limited search neighborhood is specified based upon the size of the components of the projected gradient of the relaxed solution for the lower inactive partition, ξ_{II}, and the top several sites according to their individual ranking by the controllability norm or maximum eigenvalue. The latter indicates the relative strength of a site in its single strongest direction. These criteria are quite heuristic, but after identifying a candidate optimum from the limited neighborhood search, the solution can be verified over

a larger neighborhood. In no case was further improvement noted over the initial search. We are interested in two distinct types of solutions, the first is the worst-case failure weighted configuration, labeled 'F-Opt', and the second is the no-failure but 0-1 integer-constrained solution for which we adopt the label 'I-Opt'. The 'F-Opt' solution is also 0-1 integer-constrained, of course. The 'I-Opt' configuration is a by-product of the algorithm in the sense that the algorithm keeps track of the best no-failure configuration encountered, but does not modify the search in any way for it. We may, and did in fact, run the algorithm with no failure modes to explicitly search for the 'I-Opt' configuration. Table 5-16 summarizes some of the results of the combinatorial searches, including the VAX 8650 CPU time required with the non-compiled 'CTRL-C' instruction language coding. While the total time expended could (and should) be significantly reduced with compiled code, the relative time between the cases is interesting and the price paid for searching over failure modes is apparent. Figures 5-53 and 5-54 illustrate these configurations, noting the devices that appear uniquely in the 'F-Opt' or 'I-Opt' result.

Table 5-16. Combinatorial Solutions.

Class:	5 Rotational		11 Extensional		
	F-Opt	I-Opt	F-Opt	I-Opt	
No. of Starting Configurations:	3	14	4		
Initial:					
Neighborhood Size	37	49	33		
CPU time	21 min	16 min	150 min		
Verification:					
Neighborhood Size	149	149	79	137	
CPU time	20 min	8 min	74 min	13 min	
Node/Direction	Site	Site	Site	Site	Nodes
57 X		361		3	2 3
30 Y	393		7	7	4 6
48 Y	410	410	23		6 13
57 Y	419	419	25		6 10
1001 Y	421	421		27	8 10
2830 Y		426	66	66	8 14
57 Z	477		68	68	10 16
			70	70	9 15
			73		12 18
			75	75	13 19
			78	78	14 28
				82	15 26
			86	86	19 30

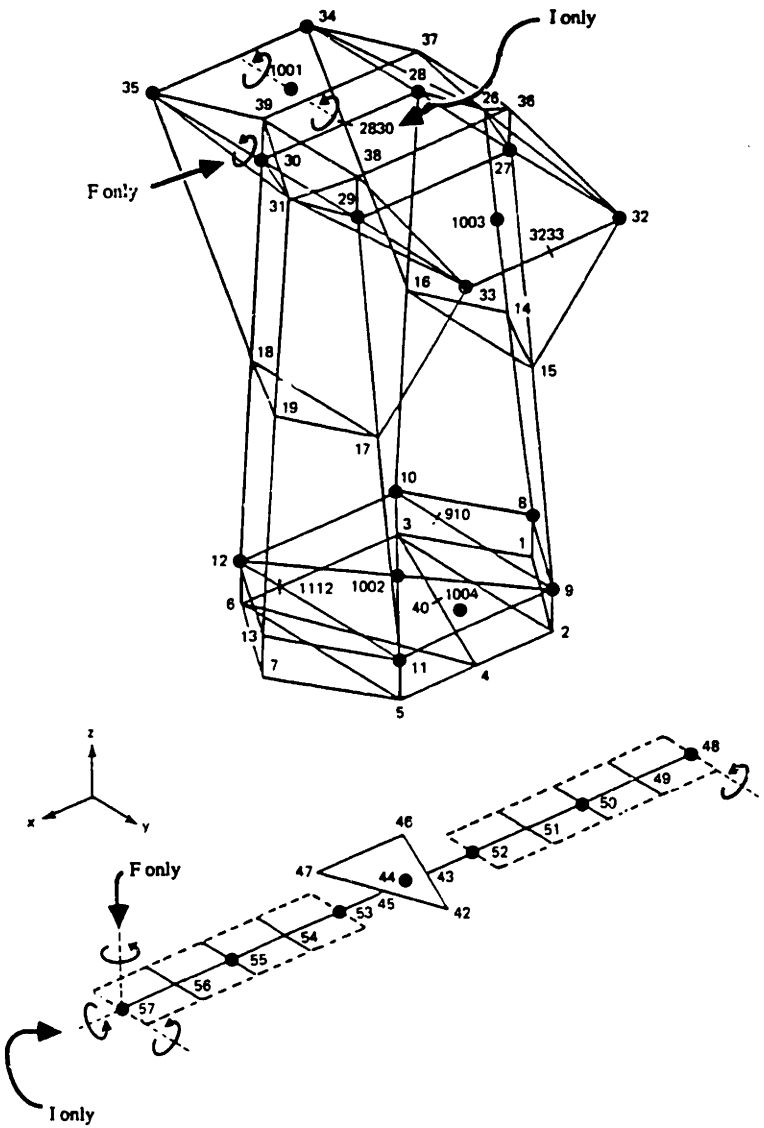


Figure 5-53. Rotational Actuators 'F-Opt' and 'I-Opt' Configuratiens

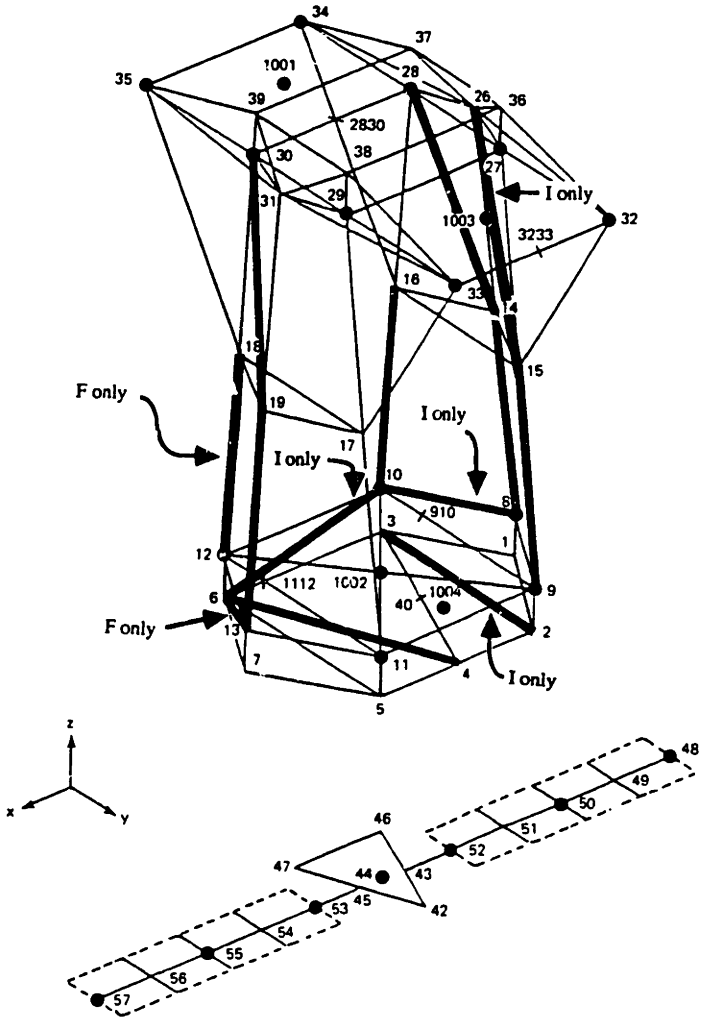


Figure 5-54. Extensional Actuators 'I-Opt' and 'F-Opt' Configurations

All four of the solutions shown were identified only '1-exchange' distant from one of the initial starting configurations. The '1-exchange' neighborhood around each of them revealed no further improvement. While this might have been expected for the no-failures 'I-Opt' solution, it was unexpected good fortune for the failure-weighted cases and confirms the usefulness of the relaxed algorithm for limiting the search of the combinatorial algorithm. For the rotational actuators the no-fail best that was noted during the initial search for the 'F-Opt' solution was about 28% poorer in terms of the controllability norm than the upper limit represented by the relaxed solution, so a separate search was conducted explicitly for a no-failures 'I-Opt' solution. A better one turned up with about half that degradation and is listed in Table 5-16. On the other hand, the best no-fail solution encountered during the search of the extensional actuators was only 0.2% below the theoretical limit, and was simply verified over a larger neighborhood.

For an evaluation of these solutions, we use again the LQ regulator as was done for the several relaxed solutions. The weightings are fixed for each class of devices, and the gains are recomputed for each failure mode as a simple reconfiguration scheme. We compare examples 'A' and 'E-5' by Hegg and Kissel to the 'I-Opt' and 'F-Opt' solutions for each class, and, since we have it, the best (with respect to LQ cost and no failures) from the Monte-Carlo search with the extensional actuators. Table 5-17 summarizes the LQ cost and average-energy controllability norm (J_2) for these cases, noting the no-failure performance and worst-case with respect to two failures. The LQ cost is listed with respect to the no-failures best case considered here, but the controllability norm is listed with respect to the upper limit represented by the appropriately constrained relaxed solution from Table 5-15. That is, the theoretical upper limit for the configurations of five rotational actuators is the '1/5-constrained' relaxed optimum and the corresponding limit for the extensional actuators is the norm associated with the '1/10-constrained' solution.

For a better look at the effect of failures on the several configurations, we plot in Figure 5-55 the controllability norm against the LQ cost for each failure mode, and in the case of the rotational actuators include the five 1-failure modes as well. Figure 5-56 is the corresponding plot for the extensional devices. Also included on each plot is the unfailed mode, which may be identified by the point lying furthest toward the the top left for each

configuration. A $\log_{10}$ scaling is necessary for the rotational actuators, and the strong degradations indicate three point torque devices are too few to control the structure to the specified levels, but that detail does not detract from the point of the example.

Table 5-17. Performance with Failures.

	J_{LQ} relative to no-failures best		Controllability Norm relative to upper limit	
Solution	No-Failures	Worst-Case	No-Failures	Worst-Case
Hegg & Kissel-A	1	1.42	.646	.302
F-Opt	1.14	1.89	.988	.682
I-Opt	1.20	2.25	.998	.497
Monte Carlo (J_2) best	1.49	4.72	.317	.147
Hegg & Kissel-E5	1.60	255	.00413	3.55×10^{-8}
F-Opt	16.7	38.4	.716	.00396
I-Opt	1	38.4	.863	.00264

Both the 'I-Opt' and 'E-5' configurations have a single particularly important torquer, the loss of which causes a significant degradation in performance with respect to either the J_2 or J_{LQ} measures. This may be seen in Figure 5-55 by the widely split groups of points for these configurations. In the case of 'E-5', the strong site is at node 35 on the upper support truss in the X direction, and for the 'I-Opt' configuration it is at node 57 on the solar panel tip, also in the X direction. With a similar configuration, differing for only two of the five devices, the 'F-Opt' set is able to improve the worst-case controllability norm by about a factor of 1.5, but with a price of about a 17% degradation in the norm for no failures. The coarse correlation between the controllability norm and LQ cost is apparent in Figure 5-55. While only the single worst failure mode of the 'I-Opt' and 'F-Opt' configurations degrades, with respect to J_2, to that of 'E-5' for no failures, if ordered instead by LQ cost, the 'F-Opt' configuration for no failures is markedly poorer, by a factor of 10, than either the 'I-Opt' or E-5 configurations.

With proportionally fewer failures, the extensional actuators exhibit far milder degradations over failures, so a linear scale is used in Figure 5-56. Notice again how the 'F-Opt' solution has improved the worst-case performance of the 'I-Opt' configuration, both with respect to J_2, which is a direct result of the search, and with respect to the LQ cost indirectly. All of the 'F-Opt' failed configurations are stronger, with respect to J_2,

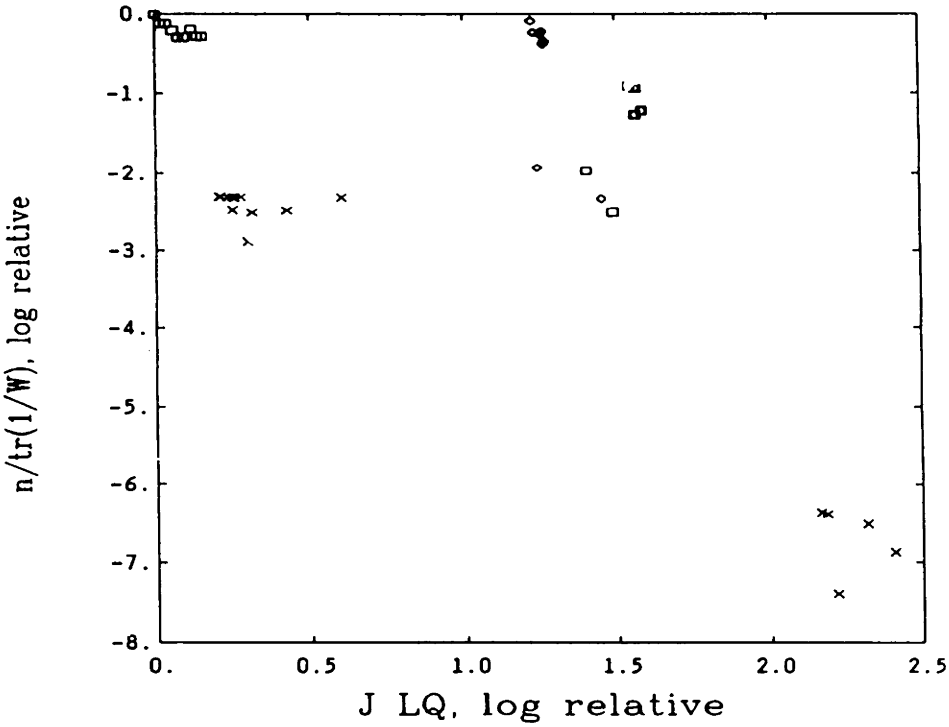


Figure 5-55 Rotational Actuators, up to 2 Failures
HK E-5 (X), F-Opt (◊), I-Opt (□).

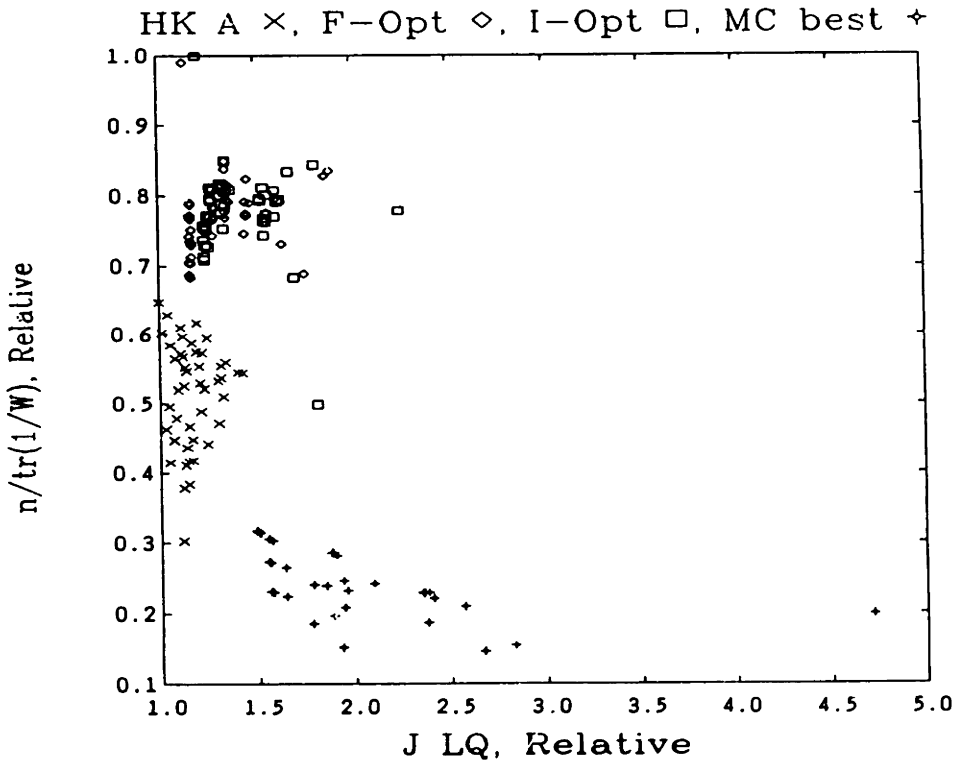


Figure 5-56 Extensional Actuators, up to 2 Failures
HK A (X), F-Opt (◊), I-Opt (□), Monte Carlo best (+)

than configuration 'A' for no failures, however, the latter offers slightly lower LQ cost than either the 'F-Opt' or 'I-Opt' configuration. This example is included in part to illustrate an open issue in this research: the relationship between the open-loop controllability and observability properties as an indicator of closed-loop performance. All three configurations- 'A' by Hegg and Kissel, 'F-Opt', and 'I-Opt'- are superior to the best of the small Monte-Carlo search, indicating they may well be close to the best possible for the LQ regulator. Note that if we are interested in fault-tolerant performance, random searching becomes quite unattractive, since each case must be evaluated over an exhaustive enumeration of failure modes.

What has not been particularly obvious in examining the individual actuators in a configuration is spotting a clear indication of why some were included and others eliminated in the 'F-Opt' configurations. One consistent trend emerges if we compare the rankings of the sites with respect to the J_2 norm and with respect to the maximum eigenvalue, both defined on the individual site Gramian W_i. The maximum eigenvalue is an indication of the strength of an actuator in one particular direction regardless of any poorly-controlled directions, while the J_2 norm, with which we are now familiar, is only large if all directions are well controlled. We list in Table 5-18 the sites in each configuration and these rankings, putting the two differing sites at the top of the list.

Table 5-18. Comparison of 'I-Opt' and 'F-Opt' Configurations

	Extensional			Ranking		Rotational			Ranking	
	Actuator	Nodes		J_2	λ_{max}	Actuator	Node/Direction		J_2	λ_{max}
'I-Opt' only:	3	2	3	1	26	361	57	X	158	70
	82	15	26	38	7	427	3233	Y	148	127
'F-Opt only:	23	6	13	18	49	393	30	Y	107	162
	73	12	18	11	35	477	57	Z	10	4
Common:	7	4	6	10	2	410	48	Y	2	10
	25	6	10	91	16	419	57	Y	1	9
	66	8	14	60	34	421	1001	Y	116	166
	68	10	16	88	29					
	70	9	15	5	5					
	75	13	19	71	39					
	78	14	28	64	10					
	86	19	30	50	9					

Considering only the two differing sites, on the average the 'F-Opt' solution includes sites with better J_2 rankings than λ_{max} rankings. For the extensional actuators the 'F-Opt' configuration uniquely includes two actuators with an average J_2 rank of 14.4 and average λ_{max} rank of 42 while the 'I-Opt' configuration includes two sites with an average J_2 rank of 19.5 and average λ_{max} rank of 16.5. The same change in ordering may be seen for the rotational configurations. This is consistent with physical intuition. For a fault-tolerant configuration, actuators that control well in all directions (lower J_2 rank) should be favored over actuators that are strong in single directions (lower λ_{max} rank) which in a configuration with no failures ('I-Opt') can collectively generate more authority than a collection of devices each required to have good individual authority in all directions. The hazy view of these distinctions, as illustrated by the broad range of rankings shown in Table 5-18, is even less distinct when inspecting modal influence coefficients, and underscores the value of a systematic method such as developed in this thesis.

5.2.4 Sensor Placements

So far all of the examples have dealt strictly with actuator selections. In part this is because the procedure is identical for sensors, and in part to avoid adding yet another major variation to what is already a lengthy series of examples. For lightly damped structural systems, it may highly desirable for compensator robustness that actuators have co-located sensors [67], insuring a 'strictly positive real' forward path. Using Popov's concept of 'hyperstability', Stieber relaxed the requirement of co-location somewhat, and produced a more general criteria for the relative location between actuators and sensors [68]. We will stick here with co-location as representing a particular constraint on sensor locations after actuator locations have been fixed. Under some conditions, the relaxed solution for sensors will in fact be exactly the same distribution as for actuators. The conditions are that the actuator and sensor be duals, that is $\mathbf{b} = \mathbf{c}^T$, and the relative weightings, including any with frequency dependencies, also be dual. This is not the case for the Draper Model 2 examples, since the input weighting filter, U(s) for the $\mathbf{z}(s)/\mathbf{w}(s)$ signal path, does not have an output dual. It is even further from the case if sensors measure directly linear combinations of modal displacement, $\eta(t)$, and not rate, which would be the dual to force type actuators. Also, there may be measurements available, such as the of the LOS itself,

that cannot physically have an actuator dual. Suppose co-location is physically possible, but displacement measurements are taken, not rate. It is interesting to consider what the observability loss from enforced co-location of displacement sensors to optimal actuator locations might be, so the relaxed solutions for extensional, translational and rotational displacement sensors were computed and are listed in Table 5-19 along with illustrations of their distributions in Figures 5-57, 5-58, and 5-59.

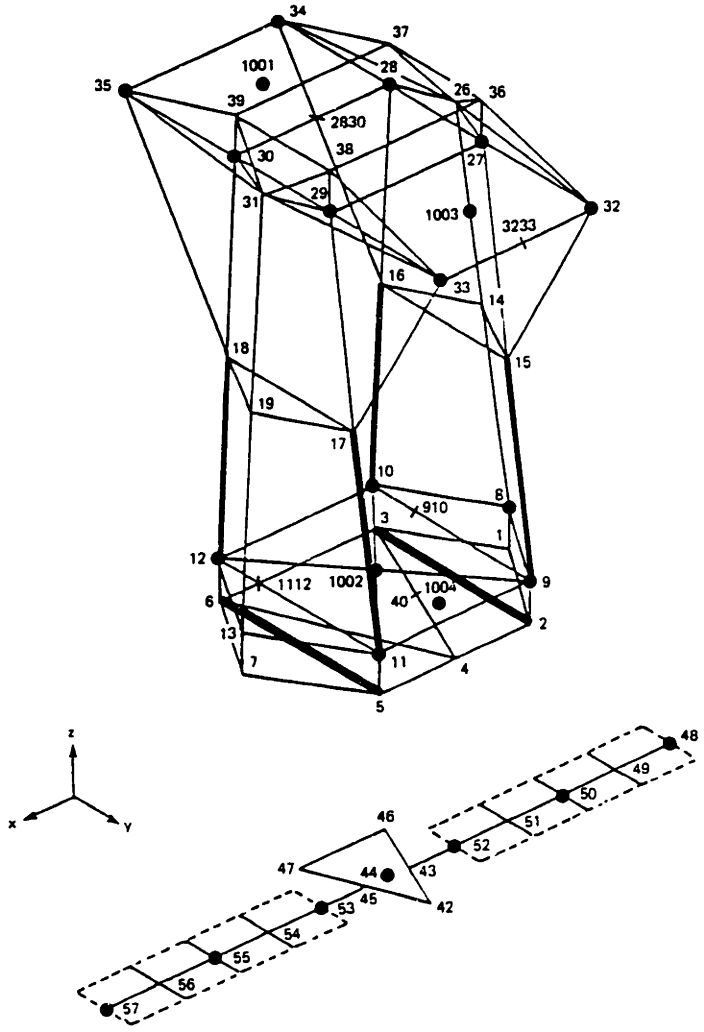


Figure 5-57. Extensional Sensors Fully Relaxed Distribution

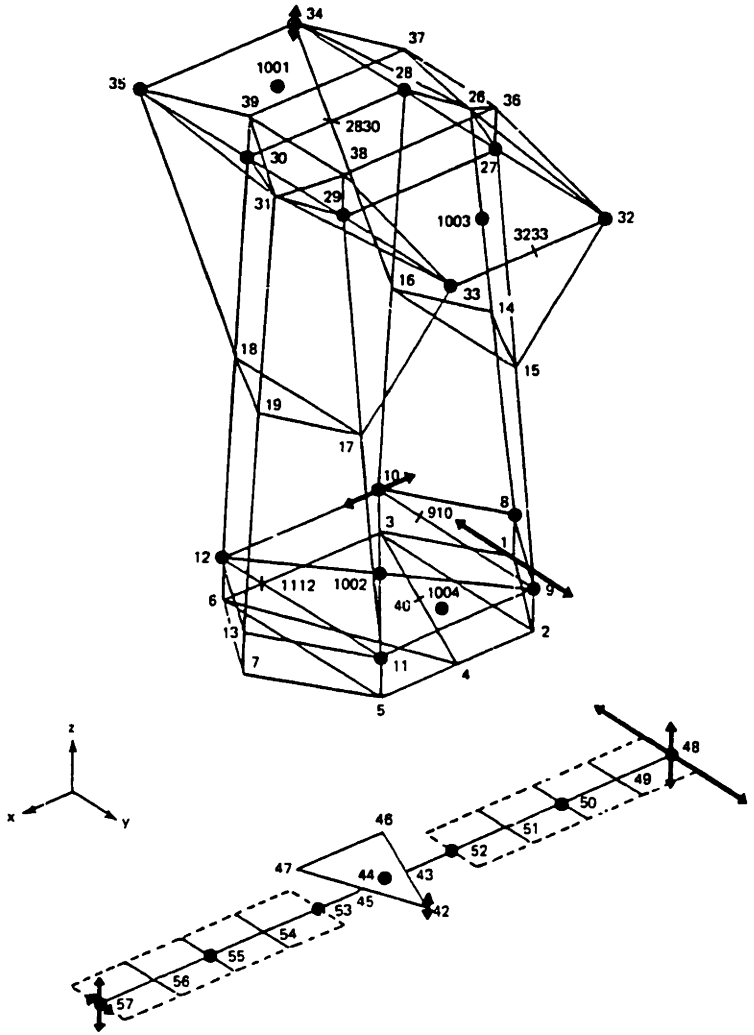


Figure 5-58. Translational Sensors Fully Relaxed Distribution

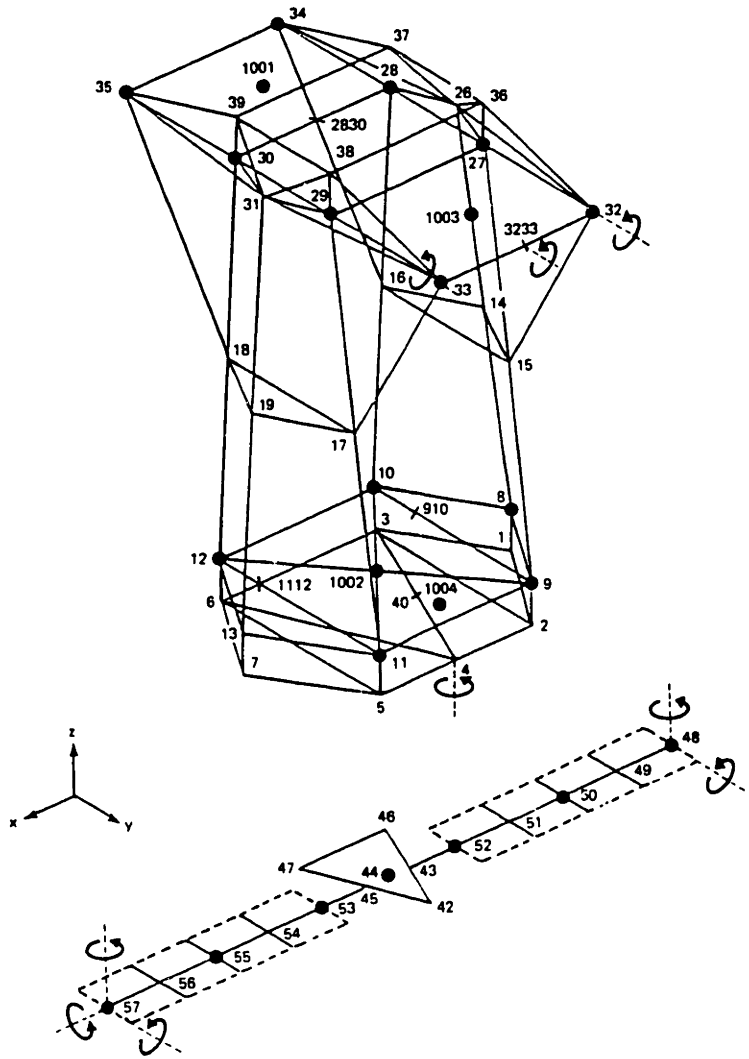


Figure 5-59. Rotational Sensors Fully Relaxed Distribution

Table 5-19. Displacement Sensor Optimal Distributions.

Extensional, Fig. 5-57			Translational, Fig. 5-58			Rotational, Fig. 5-59			
Site	Nodes	%	Site	Node/Direction	%	Site	Node/Direction		%
9	5 6	32.2	236	48 Y	35.1	477	57	Z	32.6
3	2 3	30.0	196	1 Y	25.5	468	48	Z	30.3
71	11 17	20.1	147	10 X	13.0	410	48	Y	17.8
70	9 15	14.6	294	48 Z	11.2	419	57	Y	10.3
73	12 18	2.6	303	57 Z	7.8	431	4	Z	4.9
68	10 16	0.5	245	57 Y	3.4	395	32	Y	1.3
			281	34 Z	2.8	396	33	Y	1.3
			288	42 Z	1.2	427	3233	Y	1.3
						361	57	X	0.2
Relative J_2:	5.54×10^{-9}				1				6.030×10^{-3}

Notice again the translational class is the strongest (most sensitive), followed by the rotational and extensional classes. We have not modified the relative scaling of the physical units here, which are extension across a beam in meters, nodal displacement with respect to inertial space in meters, and nodal rotation with respect to inertial space in rad/sec. Compared to the relative sizes of the fully relaxed norms for actuators, extensional sensors are poorer than extensional actuators, but physically the extensional class may not have the same attraction for sensors as it does for actuators. While many of the same sites show up in Table 5-19 as in the relaxed solutions for the actuators, the distributions are clearly different. We then compute the J_2 observability norm for sensor configurations with authority distributed identically to that of the relaxed actuator configurations of Figures 5-18, 5-21 and 5-24. A summary of the degradations is listed in Table 5-20.

Table 5-20. Observability Degradation from Co-locations

Class	Relaxed Sensor Dist.	Degradation if Co-located (relative to fully relaxed)	Relaxed Actuator Dist.
Extensional	Fig 5-57	0.2019	Fig 5-25
Translational	Fig 5-58	0.0215	Fig 5-28
Rotational	Fig 5-59	0.0687	Fig 5-31

If we must have co-located (displacement) sensors and actuators, we may face a significant loss of observability. The implication of this in a feedback loop has not been explored, but there are some other facets of the overall problem that affect sensor

placement. Assuming an actuator selection has been made, we now have an additional input path through the actuators that we may care to reflect in the basis scaling, so we may wish to revise the basis scaling for sensor selection. Sensors suitable for instrumenting structural dynamics are physically smaller and lighter than functionally dual actuators, so it is likely to be the case that there may be more of them than actuators. If so, we may constrain some to be co-located with the actuators, then optimize the distribution of the remainder. The only change required in the algorithm is to add a fixed bias observability Gramian representing the co-located constrained sensors. Examining issues such as these should provide fruitful avenues of research.

5.3 Summary

This chapter has exercised the overall sensor and actuator selection algorithms developed in Chapters 3 and 4 on two types of structural models. The first, the simply-supported beam, has obvious mode shapes and allowed a clear view into the effect of varying the particular optimization problem being solved. The second illustrated the utility of the approach on a large-scale system. In both cases an evaluation of the distributions of actuator authority revealed a coarse, but useful, correlation with the stochastic LQ regulator cost. Also, in both cases, the failure-weighted combinatorial search terminated near the relaxed solution, illustrating the value of the latter to reduce the cost of the former.

Chapter 6. Conclusions

6.1 Thesis Summary

This thesis reports applied research toward the general problem of reliable active control of space structures over missions of long duration. We reviewed, in Chapter 1, current approaches to FDI and reconfiguration as might be applicable to a structural control system and identified the underlying problem of sensor and actuator selection as a critical design step. While FDI asks what has failed, and a reconfiguration strategy asks how best to employ hardware after failures, sensors and actuators must be positioned on the structure such that, for some specified level of failures, sufficient controllability and observability remains to allow successful reconfiguration. Chapter 2 reviewed some results in sensor and actuator placement for control of distributed parameter systems, noting that algorithms for the special case of large-scale structural dynamic systems were few. Considering the possibility of failures added the need to evaluate a placement criterion over an enumeration of the possible failure modes. While no examples of large-scale algorithms that do account for failures were previously known to the author, two examples were reviewed of suggested criteria with results showing the possibility of failures does alter the best locations.

In Chapter 3, we examine several possible performance criteria for sensor and actuator locations. Norms on the controllability and observability Gramians were selected because they strike a balance between capturing essential features of functional linear dependence of the dynamics and computational economy. Since controllability and observability is characterized by the Gramians with respect to a particular state basis, the basis must be properly scaled. A novel approach to scaling using a reference signal path, such as from the disturbances to regulated variables, was proposed, along with frequency dependent (dynamic) input and output weightings. The potential sensor and actuator locations were restricted to a finite number of discrete sites, in keeping with the use of finite-element methods for practical analysis of structural dynamics. However, the resulting integer programming problem is still difficult, and an exact solution is not known to the author. Relaxing the constraint that a sensor or actuator must be considered in integer quantities (typically 0 or 1) to considering a non-negative, continuous quantity of authority or sensitivity resulted in an attractive convex programming problem. Chapter 4

developed an algorithm for this 'relaxed' problem, the solution of which was a starting point and upper bound for a combinatorial search for the failure-weighted, 0-1 integer-constrained solution we desired. A simple '1-exchange' search algorithm for the latter problem was also described in Chapter 4.

Two different structural models were taken for example results in Chapter 5: a simple system in the form of a pinned-pinned beam, and a large scale structure- the Draper Model No. 2 from DARPA's ACOSS program. With the beam model, we were able to experiment with a wide range of the free parameters in the problem, such as the particular basis scaling or the form of the optimization parameter. The solutions were easily interpreted in terms of the mode shapes, and illustrated how changes in the posing of the problem shift the optimal distribution. The large-scale example, on the other hand illustrated the practicality of the overall approach, considering sensor/actuator configurations on the order of 10 devices from up to 485 potential sites. Comparisons with actuator selections by another method generally showed the approach developed here was superior when evaluated by the cost of a stochastic regulator, though not in every case. While the correlation between the controllability and observability measures of optimality and the LQ stochastic cost was somewhat coarse, optimal configurations with respect to the former were always among the best with respect to the latter. Finally, the ability to improve the performance of a configuration of devices against the possibility of failures was demonstrated. The solution of the relaxed problem tended to include sites that provided good authority to all modes, so the changes made when explicitly searching with respect to failure-weighted performance tended to be minor. Changes made for fault-tolerant performance tended to exchange out devices contributing strongly to only a few modes and replace them with devices that were more evenly distributed in authority.

6.2 Conclusions and Contribution

The main contribution of this thesis is a practical, sequential approach to a difficult combinatorial problem that must be addressed for reliable, long-duration missions of actively controlled structures. The most important step in the overall approach is posing a 'relaxed' problem which is easy to solve and provides guidance, in terms of an upper limit and a domain for search, to a more expensive combinatorial optimization algorithm. By expressing actuator authority or sensor sensitivity in terms of controllability or

observability Gramians, we obtain a strong system interpretation of the optimization objective, which is closely related to techniques of model reduction. A minor contribution is the development an expedient approach to reflecting overall control objectives into the open-loop measures of controllability and observability, in part by appropriate scaling of the state basis with respect to a reference signal path. The ability to count on general sparing of sensors and actuators, as opposed to simple redundancy, while not guaranteed by any results of this thesis, appears reasonable.

6.3 Suggestions for Further Research

This thesis opened more issues than it settled. Suggestions for future work can be grouped into two broad themes. The first is given that combinatorial programming based on norms of Gramians is a reasonable approach for sensor and actuator selections, how might algorithms be improved in terms of speed and efficiency; while the second is more system-theoretic in nature, questioning how to better reflect the design issues and, in particular, modelling uncertainty.

6.3.1 Incorporation of Modeling Uncertainty

Given parameter variations in mass, M, and stiffness, K, references [69,70] develop estimates for the second moments of the eigenvalues and mode shapes from the eigenvalue problem,

$$\omega_i^2 M\underline{\phi}_i = K\underline{\phi}_i \qquad 6\text{-}1$$

and the second-moments of interest are

$$E\left[\underline{\omega}^T\underline{\omega}\right] = \Delta_\omega \qquad 6\text{-}2$$

$$E\left[\underline{\phi}_i^T\underline{\phi}_j\right] = \Delta_{\phi ij} \quad i=1,n;\ j=1,n\ . \qquad 6\text{-}3$$

Assume there are ν structural nodes represented in equation 6-1, each with six degrees of freedom; then the second moment for each pair of mode shapes is a 6νx6ν matrix, while for the frequencies, it is an n x n matrix where n is the number of modes in the model, and $n << \nu$, typically. Expressions 6-2 and 6-3 are by nc means trivial to obtain, but they are only intermediate results if we are interested in the statistics of the Gramians. Mehta [71]

has studied the distribution of the eigenvalues of Hermitian matrices with random elements, and developed expressions for the joint density of the eigenvalues with considerably fewer degrees of freedom than the numbers of elements in the matrix. However, even if we settled on approximating a random perturbation to the nominal Gramian as a matrix of random elements, the issue of the distribution of the eigenvectors still remains. The latter are particularly important in this problem because the condition number of the Gramians at an individual site tend to be large, 10^5 to 10^8 in the ACOSS problem, so minor changes in their directions make a significant difference in how any one contributes to a configuration of sensors or actuators. The important idea is seeking to differentiate between configurations of sensors or actuators. If we weight them in some manner related to relative uncertainty, then we may be able to obtain useful results without dramatically increasing the complexity of the problem.

One idea is to form a pair of simple (diagonal) scaling transforms for the structural dynamics in modal form, such as equation 3-62. In the first scaling, we weight the directions in modal space inversely with uncertainty in frequency. Referring to the 'direct-scaling' discussion in section 3.3.1, if for controllability we take as a required reachable state excursion of mode i

$$\eta_{i_{req}} = \frac{1}{\delta\omega_{ii}} \qquad 6\text{-}4$$

where $\delta\omega^2$ is the variance of modal frequency ω about its mean,

$$\delta\omega_{ii}^2 = E\left[\left(\omega_i - E(\omega_i)\right)^2\right] \qquad 6\text{-}5$$

or, for observation if we take as a maximum tolerable error of that mode for unit norm output,

$$\eta_{i_{max}} = \delta\omega_{ii} \qquad 6\text{-}6$$

then for both cases we may scale the modal coordinates, $\eta \rightarrow T_\omega \eta$, with

$$T_\omega = \text{diag}\left\{ \delta\omega_{ii} \right\} . \qquad 6\text{-}7$$

The effect of scaling 6-7 is to seek more control authority or measurement sensitivity in the directions which have lower uncertainty in frequency. Since this is a 'direct scaling', based on physical insights in a particular basis (modal coordinates), it must be implemented before any other scaling, such as one based on a reference signal path, is applied. Uncertainty in the mode shapes is probably more important for sensor and actuator effectiveness than uncertainties in the frequencies, and it is not generally reasonable to assume uncertainty in frequency is well correlated with uncertainty in mode shape. A second transform, unique to each device site, is suggested to reflect mode shape uncertainty.

The mode shape, ϕ_i, for the discrete eigenvalue problem of equation 6-1, has dimension 6v, and each element represents a displacement or rotation of a node for that particular mode. For each point type actuator, such as translation force or rotational torque, we can associate a particular element in ϕ_i, say $\phi_i(k)$, and the corresponding diagonal element from $\Delta\phi_{ii}$, equation 6-3. Defining $\delta^2\phi_i(k)$ as the variance about the mean, as in equation 6-5 for frequency, we may define an effective transform,

$$T_k = \text{diag}\left\{ \delta\phi_{ii}(k) \right\} \quad i=1,n \tag{6-8}$$

which is generally unique for each site, k. Some device types, such as the extension force class of actuators, would actually involve more than one entry from ϕ_i and $\Delta\phi_{ii}$, but in principle several entries could be combined for an uncertainty weight for each particular device. Equation 6-8 is the simplest case; we actually need a T_k for each device. This transform is not properly a state transform, but a relative weighting for each site which is applied as if it were a site-specific state transform, which again has the effect of attenuating the authority of sites if they have high uncertainty is some mode shapes. With transform 6-7 and relative weightings 6-8, we may seek optimal distributions which should be desensitized to the modes and sites with high uncertainty, in proportion to the uncertainty. An attractive feature of this idea is that once all the weightings are determined and applied, the problem is not any more complicated, and the algorithms of Chapter 4 may be applied without modification. We may not even have, however, estimates of the modal uncertainties in as much detail as represented by the second moments, equations 6-2 and

6-3, but may be able to specify some extreme limits for a few important parameters in the model.

Suppose as a result of structural analysis we have a small collection of extreme-case parameters, such as the upper and lower limits of stiffness, K_{max} and K_{min}, or the heaviest and lightest mass matrices, M_{max} and M_{min}, or some similar extremes of damping. Indeed, estimates of 'worst case' extremes such as these are often all the information available to structural dynamicists early in the development of a flight program, which is the stage sensor and actuator placement is likely determined. One possible method of exploiting such information is to determine several distinct models, each one representing a particular combination of the extreme cases. The performance index for a particular distribution, ξ, can be determined for each particular model, j,

$$J(\xi,j) = J\left(\sum_{i=1}^{N} \xi_i W_i^j\right). \qquad 6\text{-}9$$

This could be implemented by computing and storing the individual site Gramians, W_i^j, for each site i and for each particular model j. We could then seek distributions to maximize J, as usual, but for the worst-case j,

$$\xi^* = \arg\left\{\max_{\xi} \min_{j} J(\xi,j)\right\}. \qquad 6\text{-}10$$

In the relaxed algorithm, for instance, J could be computed for each model, j, but the gradient computed and step taken based upon the worst one. In the combinatorial algorithm, the search would examine the effect of an exchange over not just failures, but over different models. Any investigator who wishes to experiment with this idea, should note the suggestion to compute the relaxed solution for each model alone, then apply a relative weighting scalar to all the site Gramians from each particular model such that the optimal J has the same value for all models. If this is not done, then simply because the reachable and detectable volumes are smaller for stiffer, heavier, or more damped models, the solution will likely depend entirely upon that particular combination alone.

If there are very many model combinations to examine, the computational effort and expense grows rapidly, and we have already noticed in the examples of Chapter 5 how the combinatorial algorithm is computationally more burdensome than the continuous relaxed

algorithm, so another area of useful research would be directed toward more efficient algorithms, particularly the combinatorial algorithm.

6.3.2 Algorithms

If we accept the idea that the combinatorial optimization problem of equation 3-146 or 6-10 is a valid representation for sensor and actuator placement, then it seems that somewhat more efficient algorithms should be possible than the simple '1-exchange' approach employed in this thesis. It was a disappointment that no more headway was made in this area than it was. Much of the literature of combinatorial optimization appears centered on certain standard problems, and many of them are linear forms[58]. While we exploited the linearity of the Gramian to avoid having to recompute it at every step, the norms used as an optimization index are not linear, so heuristics developed on other problems, such as for integer linear programming, did not seem much help. The problem looks like it might work well with the 'simulated-annealing' idea of taking occasional non-improving steps. An important first step in further work on algorithms is coding in an efficient language. Computing a large number of matrix quadratic forms is required for exchanges, and at least once a step, solving a symmetric eigenvalue/eigenvector problem. Having an estimate of the eigenvalues of the Gramian of a new configuration as a perturbation of the previous step ought to allow a faster converging eigenvalue routine than starting with no knowledge of the eigenvalues.

6.3.3 Theoretic Insights

This suggestion for further research is the compliment of the previous one; that is, given it is relatively easy to optimize the locations of sensors and actuators with norms on Gramians, what are system-theoretic interpretations, if any, beyond those of energy-optimal open-loop control or no-disturbance state estimation discussed in section 3.2.1? Bucy and Joseph's [45] bounding of the Riccati equation solution, equation 3-90, is an example, but there ought to be others outside of the context of LQG theory. Good control and measurements of important modes ought to make the forward-loop signal path proportionally stronger, and hence compensation less sensitive to modeling errors and unmodeled modes in particular. In section 3.2.1 we gave an illustration of why truncation of the integration interval of the Gramian is important to capture relative linear dependencies between closely-spaced modes. Truncation of time or frequency intervals has a broader

context, and reference [61] examines this for model reduction tools, which should also offer some background for deeper insights into our problem.

Appendix A

Derivation of Gradients

Each of the scalar controllability or observability metrics exploited as an optimization parameter are functions of the eigenvalues (λ) of the appropriate Gramian.

$$J_1 = \lambda_{min}(W) \qquad J_2 = \frac{n}{tr(W^{-1})} = \frac{n}{\sum_{i=1}^{n} \frac{1}{\lambda_i}} \qquad J_3 = \sqrt[n]{|W|} = \prod_{i=1}^{n} \lambda_i^{\frac{1}{n}}$$

where Gramian W is a linear combination of components

$$W(\xi) = \sum_{i=1}^{N} \xi_i W_i \ .$$

The main mathematical tool required is eigenvalue perturbations for self-adjoint systems, and a reference is by Courant and Hilbert, [34]. Gramians are real symmetric non-negative definite matrices, so they have real non-negative eigenvalues and a full complement of (n) orthogonal eigenvectors which may be normalized to unit length.

$$W v_i = \lambda_i v_i \quad i=1,n \qquad v_i^T v_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

We perturb W by δW, and are interested in the resulting perturbations to λ and v. Writing the eigenvalue problem for the perturbed matrix,

$$\tilde{W}\,\tilde{v}_i = \tilde{\lambda}_i\,\tilde{v}_i$$
$$(W + \delta W)(v_i + \delta v_i) = (\lambda_i + \delta\lambda_i)(v_i + \delta v_i) \ .$$

Retaining terms to first order in the perturbations and removing the original eigenvalue problem,

$$W\,\delta v_i + \delta W\, v_i = \lambda_i\,\delta v_i + \delta\lambda_i\, v_i \ . \qquad \text{A-1}$$

For the eigenvalue perturbation, we take the inner product of each side of equation A-1 with v_i and obtain the needed result:

$$\left(\lambda_i v_i^T\right)\delta v_i + v_i^T \delta W\, v_i = \lambda_i v_i^T \delta v_i + \delta\lambda_i v_i^T v_i$$

$$v_i^T \delta W\, v_i = \delta\lambda_i \ . \qquad \text{A-2}$$

For distinct eigenvalues, a similar result may be obtained for the eigenvector perturbations. Since the eigenvectors span n-space, we may express the perturbation of any eigenvector as a linear combination with coefficients **a** of the other eigenvectors. Excluding a perturbation in the same direction as the eigenvector simply reflects the fact that length may be indeterminate for an eigenvector:

$$\delta v_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} a(j)\, v_j \ . \qquad \text{A-3}$$

Substituting equation A-3 into A-1 and taking the inner product with all other eigenvectors allows for solution for the coefficients. Substituting those back into A-3 results in

$$\delta v_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{v_j^T \delta W\, v_i}{(\lambda_i - \lambda_j)} v_j \ . \qquad \text{A-4}$$

For non-distinct eigenvalues see references [34] and [72]. Computationally, the eigenvector perturbations are at least n-1 times more expensive to obtain than the eigenvalue perturbations. We do not need them for the gradient computations, but would if we wanted the Hessian, thus the motivation for a gradient-only algorithm.

Since the J_1 norm is the minimum eigenvalue, its gradient is particularly simple. For component i of ξ, the perturbation to the nominal Gramian W from differential $d\xi_i$ is $W_i d\xi_i$, since the nominal is a linear combination of the N contributions. The i^{th} component of the gradient is then:

$$\frac{\partial J_1}{\partial \xi_i} = v_1^T W_i\, v_1 \qquad \text{A-5}$$

where v_1 is the eigenvector associated with λ_{min}. The other norms are functions of all the eigenvalues, and an application of the chain rule results in

$$\nabla_\xi J_2 = \left(\nabla_\xi \lambda^T\right)\left(\nabla_\lambda J_2\right) \quad \text{and} \quad \nabla_\xi J_3 = \left(\nabla_\xi \lambda^T\right)\left(\nabla_\lambda J_3\right) .$$

For the average energy-based metric (J_2) a component of the gradient with respect to the eigenvalues is

$$\frac{\partial J_2}{\partial \lambda_j} = \frac{n}{\lambda_j^2 \left(\sum \frac{1}{\lambda}\right)^2} = \frac{J_2^2}{n \lambda_j^2} \qquad \text{A-6}$$

and a component for the same gradient of the total-volume metric, J_3, is:

$$\frac{\partial J_3}{\partial \lambda_j} = \frac{\lambda_j^{1/n - 1}}{n}\left(\prod_{i=1,\, i\neq j}^{n} \lambda_i^{\frac{1}{n}} \right) = \frac{\prod_{i=1}^{n} \lambda_i^{\frac{1}{n}}}{n \lambda_j} = \frac{J_3}{n \lambda_j} . \qquad \text{A-7}$$

The gradient of λ^T with respect to the independent variables, ξ, is an N x n matrix with elements

$$\left[\nabla_\xi \lambda^T\right]_{i,j} = \frac{\partial \lambda_j}{\partial \xi_i} = v_j^T W_i v_j . \qquad \text{A-8}$$

For our applications, the active partition of ξ is not independent, but must satisfy the total power constraint:

$$\sum_{i=1}^{N} \xi_i = M .$$

We may select one component, ξ_d, as dependent and solve for it in terms of the others:

$$\xi_d = M - \sum_{i=1,\, i\neq d}^{N} \xi_i .$$

The nominal Gramian, expressed in terms of the remaining independent components, is

$$W(\xi) = \xi_d W_d + \sum_{i=1,\, i\neq d}^{N} \xi_i W_i = \sum_{i=1}^{N} \xi_i (W_i - W_d) + M W_d \ . \qquad \text{A-9}$$

With respect to an independent perturbation, $d\xi_i$, the perturbation to the nominal Gramian is now $(W_i - W_d)d\xi_i$. The gradient in the constraint plane and as employed by the algorithm may be obtained from the previous results by substituting $(W_i - W_d)$ for W_i in equation A-8. For J_2 and J_3 they are:

$$\nabla_\xi J_2 = \frac{J_2^2}{n} \begin{bmatrix} v_1^T(W_1 - W_d)v_1 & \cdots & v_n^T(W_1 - W_d)v_n \\ \vdots & & \vdots \\ v_1^T(W_N - W_d)v_1 & \cdots & v_n^T(W_N - W_d)v_n \end{bmatrix} \begin{bmatrix} \lambda_1^{-2} \\ \vdots \\ \lambda_n^{-2} \end{bmatrix} \qquad \text{A-10}$$

$$\nabla_\xi J_3 = \frac{J_3}{n} \begin{bmatrix} v_1^T(W_1 - W_d)v_1 & \cdots & v_n^T(W_1 - W_d)v_n \\ \vdots & & \vdots \\ v_1^T(W_N - W_d)v_1 & \cdots & v_n^T(W_N - W_d)v_n \end{bmatrix} \begin{bmatrix} \lambda_1^{-1} \\ \vdots \\ \lambda_n^{-1} \end{bmatrix} \qquad \text{A-11}$$

The simple manner in which this constraint is incorporated into the gradient computation is a direct consequence of the fact the constraint is linear in the independent variables. If we had chosen the linear gain of the devices rather that the power (linear gain squared) the constraint equation would have been a quadratic form. This further motivates the selecting the 'power gains' as independent parameters for the relaxed algorithm.

Appendix B

This section lists the essential numerical data which comprises the ACOSS 'Draper No. 2' structural model used in Chapter 5. For the NASTRAN run deck, and details on the construction of this model, consult references [62] and [63]. The purpose of the listings here is to allow another investigator to roughly verify his own version of this model.

We assume the structural dynamics are in modal form,

$$\ddot{\eta} + \Omega^2 \eta = \left[\Phi^T B_a\right] u$$

$$z = \left[C_z \Phi\right] \eta$$

where input **u** is in units of Newtons for forces and Newton-meters for torques, and displacement outputs are in units of meters for translation and radians for rotation. The mode shapes, ϕ, are normalized with respect to the mass and stiffness matrices as usual, $\Phi^T M \Phi = I$ and $\Phi^T K \Phi = \Omega^2$. The optical line-of-sight is taken as the regulated variables, **z**, and has three elements,

$$z = \begin{bmatrix} \text{LOS - X, rad} \\ \text{LOS - Y, rad} \\ \text{Defocus, Z, meters} \end{bmatrix}$$

Its output coefficient matrix, $[C_z \Phi]$ is listed later and reference [63] contains details on developing those coefficients. For brevity here, we restrict the number of modes to the first 18 flexible modes, which are listed in Table B-1 in slight variation of the order presented in Table 5-10.

Table B-1

Flex Mode	Freq (Hz)	Rank in Disturbance to LOS Signal Path, out of 50	Flex Mode	Freq (Hz)	Rank
1	0.1480	1	10	1.0922	7
2	0.2824	2	11	1.8387	35
3	0.3187	11	12	1.8435	38
4	0.3352	9	13	1.8892	48
5	0.4678	12	14	1.9904	29
6	0.5835	4	15	2.0601	21
7	0.6006	3	16	2.4516	6
8	0.6734	10	17	2.4718	5
9	0.9601	13	18	3.2421	2

Figures B-1 through B-11 plot the corresponding mode shapes of the 11 modes incorporated into the dynamic model. Most of the interleaving deleted modes were solar panel modes, although one was the X-rotation of the equipment bay.

A total of 485 sensors or actuators were considered, 137 axial extension devices and 174 each of point-translational and point-rotational devices. We list next the enumeration of these devices; that is the nodes the axial members connect and the nodes and directions of action for the point devices.

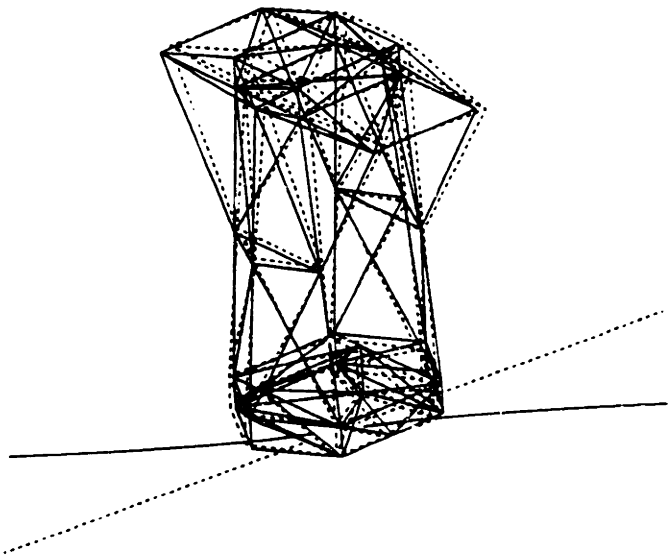


Figure B-1. Mode 1, 0.148 Hz, Equip. bay Y-rotation.

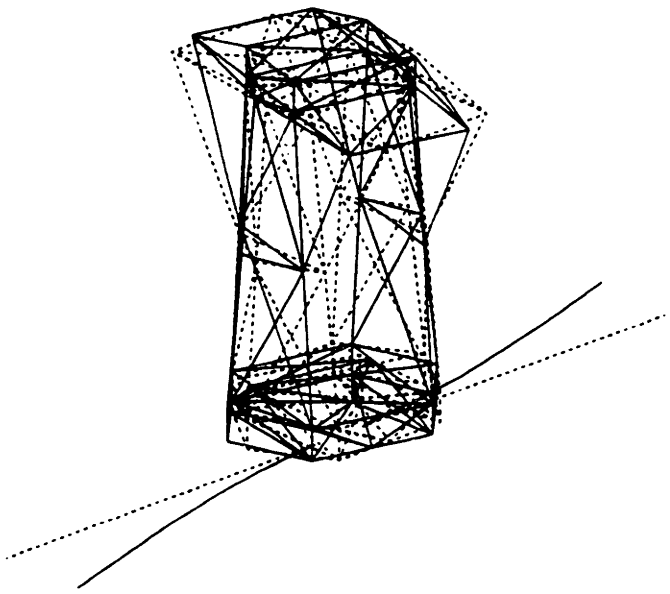


Figure B-2. Mode 2, 0.282 Hz, Equip. bay Z-rotation.

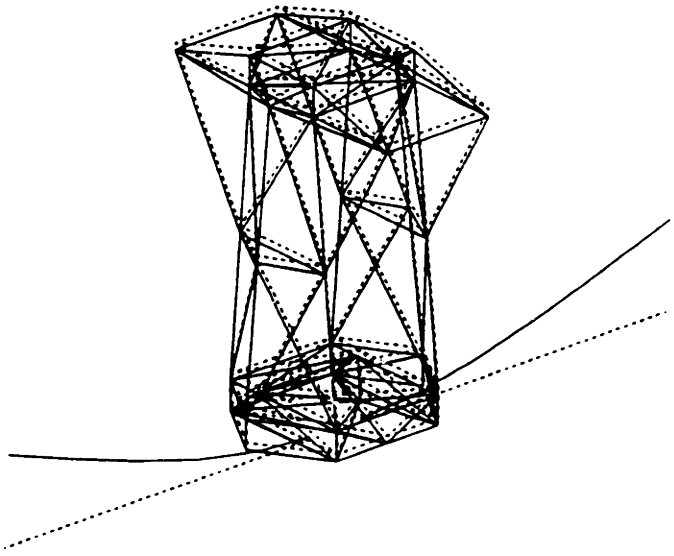


Figure B-3. Mode 3, 0.319 Hz, 1st Solar Panel in X-Z plane.

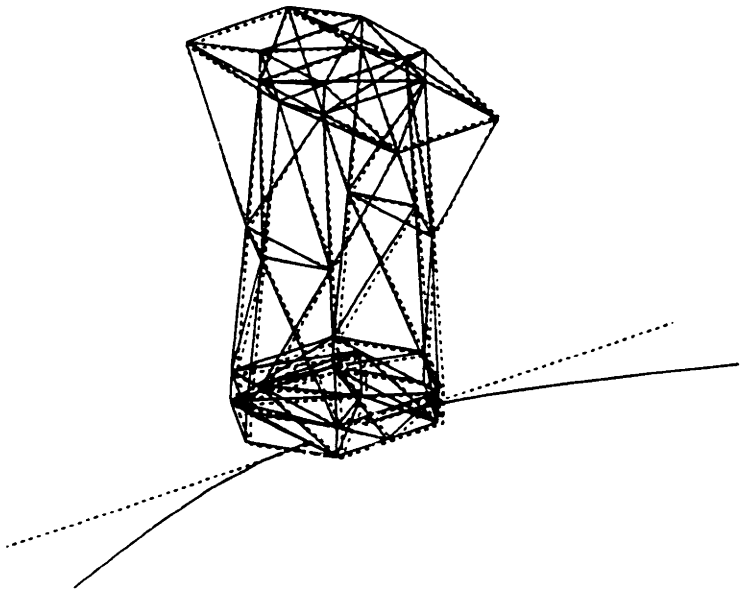


Figure B-4. Mode 4, 0.335 Hz, 1st Solar Panel in X-Y plane.

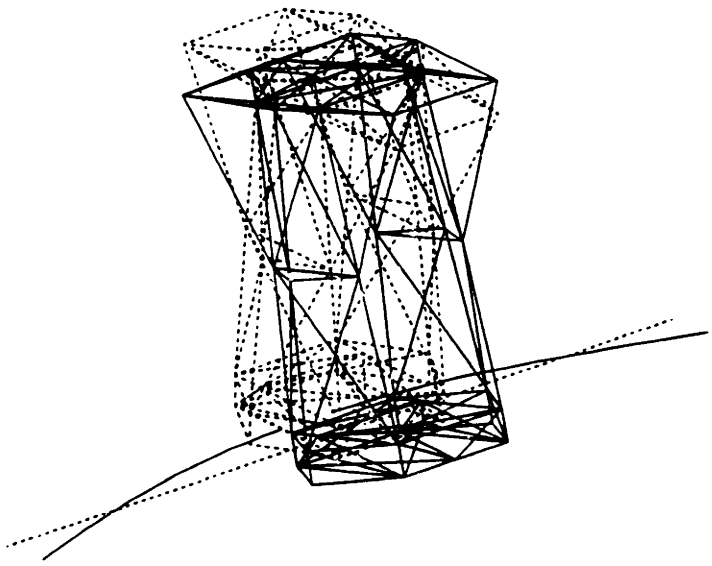


Figure B-5. Mode 6, 0.583 Hz, Equip. bay Y-translation.

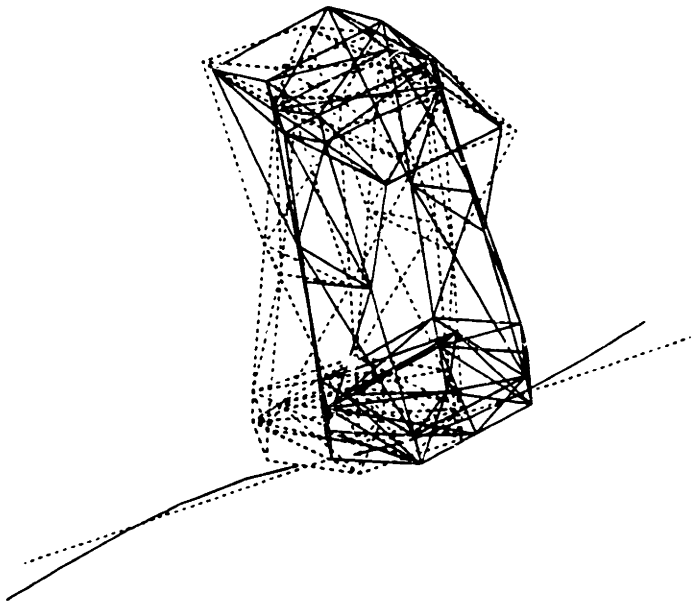


Figure B-6. Mode 7, 0.601 Hz, Equip. bay X-translation.

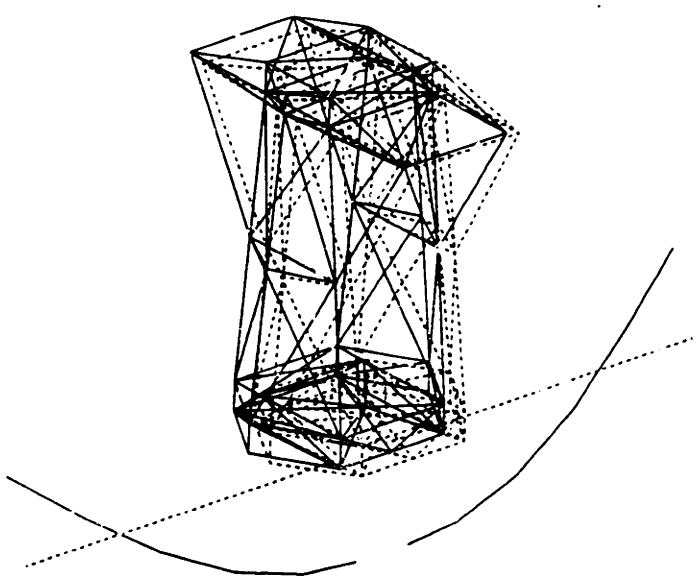


Figure B-7. Mode 8, 0.673 Hz, 2nd Solar Panel in X-Z plane.

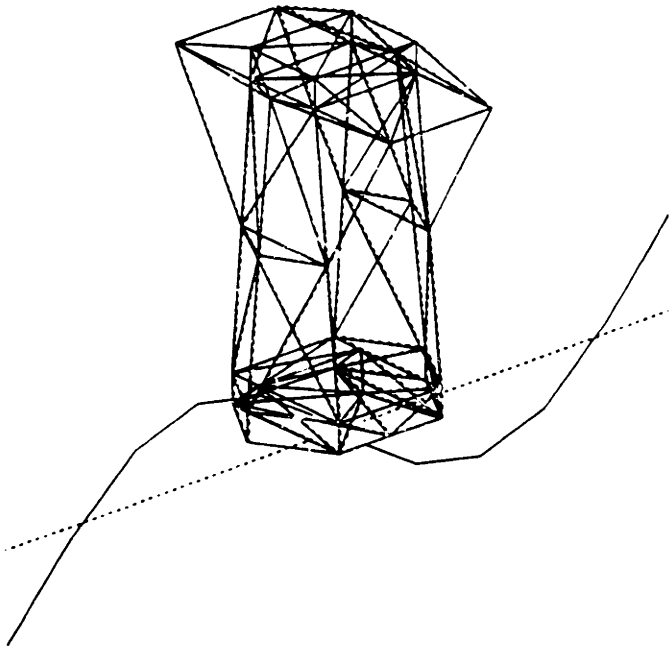


Figure B-8. Mode 10, 1.092 Hz, 2nd Equip. bay Y-Rotation.

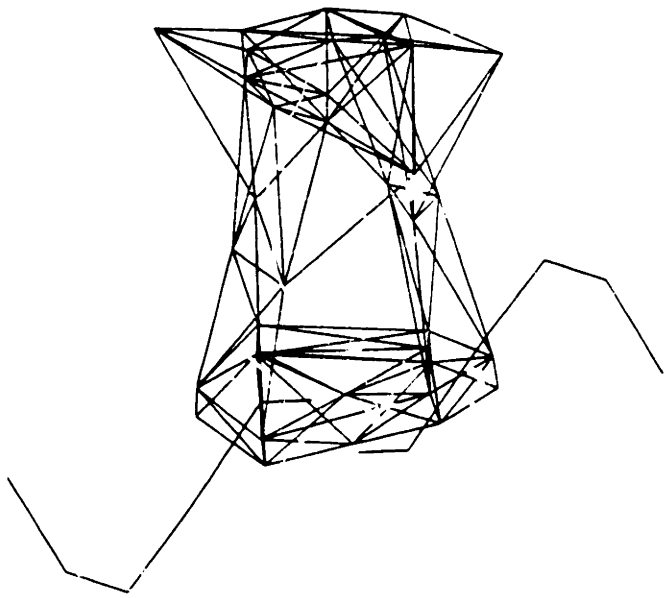


Figure B-9. Mode 16. 2.452 Hz, Torsion.

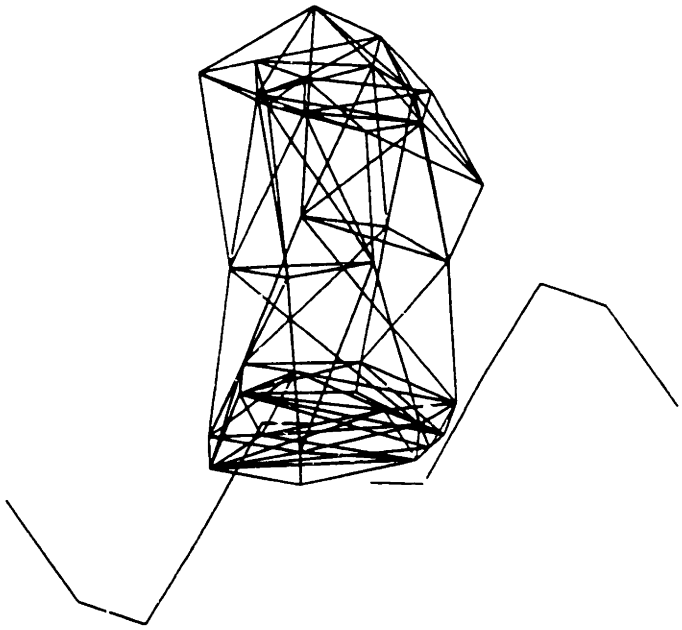


Figure B-10. Mode 17, 2.472 Hz, 1st Bending

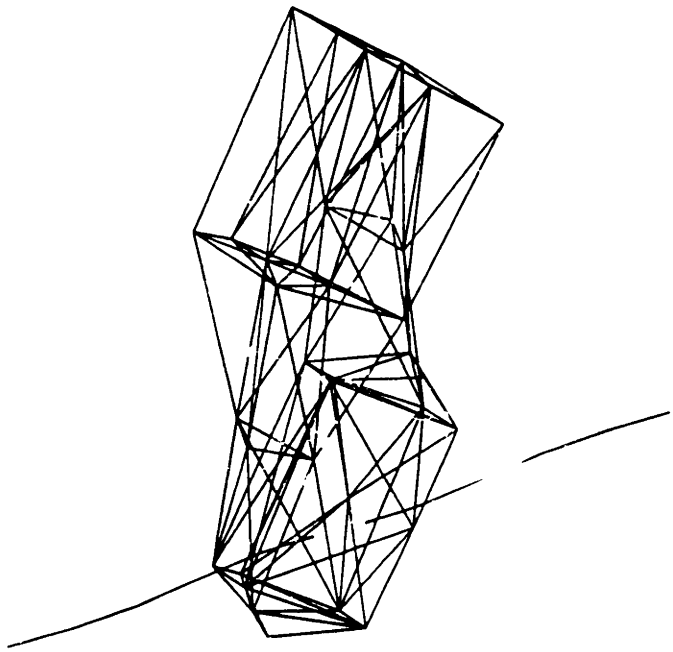


Figure B-11. Mode 18, 3.242 Hz, Upper Truss rotation.

Table B-2 Extensional Devices

Device	Nodes	Device	Nodes	Device	Nodes
1	1 2	47	29 30	93	26 32
2	1 3	48	29 31	94	27 32
3	2 3	49	30 31	95	27 33
4	2 4	50	27 29	96	29 33
5	3 4	51	27 30	97	31 33
6	4 5	52	28 2830	98	32 3233
7	4 6	53	27 36	99	26 34
8	3 6	54	28 37	100	28 34
9	5 6	55	30 39	101	30 34
10	5 7	56	29 38	102	30 35
11	6 7	57	29 36	103	31 35
12	1 8	58	27 37	104	34 35
13	2 9	59	28 39	105	32 36
14	3 10	60	30 38	106	33 38
15	5 11	61	36 37	107	34 37
16	5 11	62	37 39	108	35 39
17	7 13	63	39 38	109	26 37
18	3 8	64	36 38	110	26 36
19	2 8	65	37 38	111	31 39
20	4 9	66	4 14	112	31 38
21	4 11	67	10 14	113	48 49
22	5 13	68	10 16	114	49 50
23	6 13	69	16 9	115	50 51
24	1112 3	70	9 15	116	51 52
25	6 10	71	11 17	117	52 43
26	8 9	72	11 18	118	45 53
27	8 10	73	12 18	119	53 54
28	9 910	74	12 19	120	54 55
29	9 40	75	13 19	121	55 56
30	910 40	76	13 17	122	56 57
31	11 40	77	14 26	123	8 157
32	1112 40	78	14 28	124	6 40
33	9 11	79	16 28	125	2 40
34	10 12	80	16 27	126	2830 30
35	11 1112	81	15 27	127	3233 33
36	11 13	82	15 26	128	3 40
37	12 13	83	17 29	129	5 40
38	14 15	84	18 29	130	910 1112
39	14 16	85	18 30	131	2 910
40	16 15	86	19 30	132	3 910
41	17 18	87	19 31	133	5 1112
42	17 19	88	17 31	134	6 1112
43	18 19	89	15 32	135	12 910
44	26 27	90	16 34	136	910 10
45	26 28	91	17 33	137	1112 12
46	27 28	92	18 35		

Table B-3. Translational Devices

Node	Device # for Direction: X	Y	Z	Node	Device # for Direction: X	Y	Z
1	138	196	254	36	167	225	283
2	139	197	255	37	168	226	284
3	140	198	256	38	169	227	285
4	141	199	257	39	170	228	286
5	142	200	258	40	171	229	287
6	143	201	259	42	172	230	288
7	144	202	260	43	173	231	289
8	145	203	261	44	174	232	290
9	146	204	262	45	175	233	291
10	147	205	263	46	176	234	292
11	148	206	264	47	177	235	293
12	149	207	265	48	178	236	294
13	150	208	266	49	179	237	295
14	151	209	267	50	180	238	296
15	152	210	268	51	181	239	297
16	153	211	269	52	182	240	298
17	154	212	270	53	183	241	299
18	155	213	271	54	184	242	300
19	156	214	272	55	185	243	301
26	157	215	273	56	186	244	302
27	158	216	274	57	187	245	303
28	159	217	275	910	188	246	304
29	160	218	276	1001	189	247	305
30	161	219	277	1002	190	248	306
31	162	220	278	1003	191	249	307
32	163	221	279	1004	192	250	308
33	164	222	280	1112	193	251	309
34	165	223	281	2830	194	252	310
35	166	224	282	3233	195	253	311

Table B-4. Rotational Devices

Node	Device # for Direction: X	Y	Z	Node	Device # for Direction: X	Y	Z
1	312	370	428	36	341	399	457
2	313	371	429	37	342	400	458
3	314	372	430	38	343	401	459
4	315	373	431	39	344	402	460
5	316	374	432	40	345	403	461
6	317	375	433	42	346	404	462
7	318	376	434	43	347	405	463
8	319	377	435	44	348	406	464
9	320	378	436	45	349	407	465
10	321	379	437	46	350	408	466
11	322	380	438	47	351	409	467
12	323	381	439	48	352	410	468
13	324	382	440	49	353	411	469
14	325	383	441	50	354	412	470
15	326	384	442	51	355	413	471
16	327	385	443	52	356	414	472
17	328	386	444	53	357	415	473
18	329	387	445	54	358	416	474
19	330	388	446	55	359	417	475
26	331	389	447	56	360	418	476
27	332	390	448	57	361	419	477
28	333	391	449	910	362	420	478
29	334	392	450	1001	363	421	479
30	335	393	451	1002	364	422	480
31	336	394	452	1003	365	423	481
32	337	395	453	1004	366	424	482
33	338	396	455	1112	367	425	483
34	339	397	456	2830	368	426	484
35	340	398	456	3233	369	427	485

For each of the devices listed in Tables B-2, B-3 and B-4, corresponds a column of modal influence coefficients in the matrix $[\Phi^T B_a]$, which is too lengthy to be listed here.

Finally, we list the output coefficients for the optical line-of-sight, $[C_z\Phi]$, and the modal influence coefficients of the disturbance, $[\Phi^T B_d]$. As developed in reference [42], the disturbance acts on node 37 on the upper support truss, and on node 46 on the equipment bay, and at each node are taken to act in the X, Y, and Z directions simultaneously. The influence coefficients at node 37, then, is the sum of the influence coefficients of actuators 168, 226, and 284, while at node 46 the disturbance is equivalent to devices 176, 234, and 292.

Table B-5. Line-of-Sight and Disturbance Influence Coefficients

	$[C_z\Phi]$			$[\Phi^T B_d]$	
Flex	Line-of-sight Component			Structural Node	
Mode	X	Y	Defocus	37	46
1	2.9443D-08	-3.8978D-04	4.9407D-07	8.8943D-03	-1.0599D-02
2	1.0187D-06	1.6846D-04	2.9026D-07	-8.6572D-05	-1.7154D-03
3	-2.9944D-05	-2.6540D-07	3.5972D-06	-4.7018D-03	-9.4432D-04
4	-2.3175D-04	6.6579D-10	-1.9974D-06	1.0715D-03	-1.3330D-03
5	-3.5017D-06	-7.8167D-07	1.1903D-05	-9.2474D-03	1.4010D-02
6	8.3206D-04	6.7386D-07	2.1978D-05	-6.1456D-03	-2.7354D-04
7	2.6354D-06	-7.2596D-04	4.2508D-06	6.5602D-03	6.7649D-03
8	-4.5593D-05	8.5189D-07	-9.3845D-06	-1.6338D-03	3.0606D-02
9	3.2269D-06	1.1370D-04	1.8749D-06	6.5796D-04	-1.8033D-03
10	-6.2449D-07	2.1420D-04	7.3255D-07	-2.2441D-04	-2.7890D-02
11	-1.2146D-05	-1.1460D-07	-2.5343D-06	4.1176D-05	3.1676D-03
12	1.3614D-06	-2.2848D-07	2.4190D-06	-8.4283D-05	1.6087D-03
13	1.1808D-14	5.6450D-16	3.9406D-18	8.8476D-14	3.5105D-13
14	-1.2235D-05	-5.3384D-07	2.0708D-06	7.9942D-05	-9.5648D-03
15	5.7025D-06	3.2102D-04	3.5795D-06	2.7832D-04	-1.0963D-03
16	-3.8700D-05	-2.2725D-03	-2.5783D-05	-1.4606D-03	-2.2776D-02
17	3.8312D-05	3.0975D-03	3.1076D-05	4.2141D-04	-2.3335D-02
18	3.4799D-06	-1.9824D-02	-1.4602D-04	2.5158D-02	-1.3257D-03

References

1. MIL-STD-1543A (USAF), "Reliability Program Requirements for Space and Missile Systems", 25 Jun 1982.

2. NASA NHB 5300.4, "Reliability Program Provisions for Aeronautical and Space System Contractors", Jan 87.

3. Vander Velde, W. E., "Detection of Component Failures in Flexible Spacecraft Control Systems" in 'Dynamics and Control of Large Structures', *Proceedings of the 4th VPI&SU/AIAA Symposium*, Blacksburg, VA, 6-8 Jun 1983, pp. 481-496.

4. Carignan, C. F., "Filter Failure Detection for Systems with Large Space Structure Dynamics", SM Thesis, Department of Aeronautics and Astronautics, MIT, Cambridge, MA, 1986.

5. San Martin, A. M., and Vander Velde, W. E., "Design of Robust Failure Detection Filters", *Proceedings of the American Control Conference*, Paper TP2, Seattle, WA, 18-20 Jun 1986.

6. Sworder, D. D., "Control of Systems Subject to Sudden Changes in Character", *Proceedings of the IEEE*, Vol 64, No. 8, 1976, pp. 1219-1225.

7. Siljak, D. D., "Reliable Control Using Multiple Control Systems", *International Journal of Control*, Vol. 31, No. 2, 1980, pp. 303-329.

8. Birdwell, J. D., "On Reliable Control Systems Designs", PhD Thesis, MIT Electronic Systems Laboratory Report ESL-TH-821, 1978.

9. Mariton, M., "On Controllability of Linear Systems with Stochastic Jump Parameters", *IEEE Transactions on Automatic Control*, Vol. AC-31, No. 7, July 1986, pp. 680-683.

10. "Self-Repairing Digital Flight Control System", Final Report of Contract F33615-80-C-3100, General Electric, Jul 1984.

11. Weiss, J. L., Looze, D. P., Eterno, J. S., Grunberg, D. B., "Initial Design and Evaluation of Automatic Restructurable Flight Control System Concepts", NASA LRC Contractor Report 178064, Jun 1986.

12. Joshi, S. M., "Design of Failure-Accommodating Multi-Loop LQG Regulators", *Proceedings of 24th Conference on Decision and Control*, Ft. Lauderdale, FL, Dec 1985, pp. 876-877.

13. Joshi, S. M., "Failure-Accommodating Control of Large Flexible Spacecraft", Proceedings of the American Control Conference, Seattle, WA, June, 1986, pp. 156-160.

14. Joshi, S. M., "Design of Failure-Accommodating Multiloop LQG-Type Controllers", *IEEE Transactions on Automatic Control*, Vol. AC-32, No. 8, Aug 1987, pp. 740-741.

15. Vander Velde, W. E., "Control System Reconfiguration", *Proceedings of the American Control Conference*, San Diego, CA, Jun 6-8, 1984, pp. 1741-1745.

16. Wagner, E. A., "On-Board Automatic Aid and Advisor for Pilots of Control-Impaired Aircraft", PhD thesis, Department of Aeronautics and Astronautics, MIT, Cambridge, MA, Feb 1988.

17. Potter, J. E. and Suman, M. C., "Thresholdless Redundancy Management with Arrays of Skewed Instruments", AGARDograph-224, *Integrity in Electronic Flight Control Systems*, 1977, pp. 15-1 to 15-25.

18. Daly, K. C., Gai, E., Harrison, J. V., "Generalized Likelihood Test for FDI in Redundant Sensor Configurations", *Journal of Guidance, Control, and Dynamics*, Vol. 2, No. 1 Jan-Feb 1979, pp. 9-17.

19. Kubrusly, C. S., and Malebranche, H., "Sensors and Controllers Location in Distributed Systems - A Survey", *Automatica*, Vol 21, No. 2, 1985, pp. 117-128.

20. Omatu, S., Seinfeld, J. H., "Optimization of Sensor and Actuator Locations in a Distributed Parameter System", *Journal of the Franklin Institute*, Vol. 315, No. 5/6, 1983, pp. 407-421.

21. Leigh, J. R., Functional Analysis and Linear Control Theory. Academic Press, New York, 1980.

22. Hegg, D. R., Kissel, G. J., "Wideband Disturbance Accommodation in Precision Flexible Space Structures", in *Proceedings of the NASA/JPL Workshop on Identification and Control of Flexible Space Structures*, San Diego, CA, Jun 84; Vol II.

23. DeLorenzo, M. L., "Selection of Noisy Sensors and Actuators for Regulation of Linear Systems", Ph.D. Thesis, School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN, May 1983.

24. Skelton, R. E., and Yousuff, A., "Component Cost Analysis of Large-Scale Systems", in *Control and Dynamic Systems*, Vol 18, Academic Press, New York, 1982.

25. Johnson, T. L., Athans, M., and Skelton, G. B., "Optimal Control-Surface Locations for Flexible Aircraft", *IEEE Transactions on Automatic Control*, Vol. AC-16, No. 4, Aug 1971, pp. 320-331.

26. Vander Velde, W. E., Carignan, C. R., "Number and Placement of Control System Components Considering Possible Failures", *Journal of Guidance, Control, and Dynamics*, Vol. 7, No. 6, Nov-Dec 1984, pp. 703-709.

27. Montgomery, R. C., and Vander Velde, W. E., "Reliability Considerations in the Placement of Control System Components", *Journal of Guidance, Control, and Dynamics*, Vol. 8, No., 3, May-Jun 1985, pp. 411-413.

28. Harrison, J. V., Daly, K. C., Gai, E., "Reliability and Accuracy Prediction for a Redundant Strapdown Navigator," *AIAA Journal of Guidance and Control*, Vol. 4, No. 4, Sep 81, pp. 523-529.

29. Gai, E., Adams, M., "Measures of Merit for Fault-Tolerant Systems," CSDL-P-1752, 1983.

30. Anderson, R. T., Reliability Design Handbook, Report RDH-376, Reliability Analysis Center, Rome Air Development Center, Griffis AFB, NY, 1976.

31. Daly, Gai, Vander Velde, et al., "Fault-Tolerant Control Systems," unpublished course notes, MIT.

32. Papoulis, A., Probability, Random Variables, and Stochastic Processes. McGraw Hill, New York, 1965.

33. Gantmacher, F. R., The Theory of Matrices. Chelsea Publishing Co., New York, N.Y., 1977.

34. Courant, R., and Hilbert, D., Methods of Mathematical Physics. John Wiley and Sons, New York, 1953.

35. Kalman, R. E., Ho, Y. C., Narendra, K. S., "Controllability of Linear Dynamical Systems," *Contributions to Differential Equations*, Vol 1, No. 2, 1961, pp. 182-213.

36. Müller, P. C., Weber, H. I., "Analysis and Optimization of Certain Qualities of Controllability and Observability for Linear Dynamical Systems", *Automatica*, Vol. 8, 1972, pp. 237-246.

37. Beckenbach, E. F., Bellman, R., Inequalities. Springer-Verlag, New York, 1965.

38. Lindberg, R. E., "Actuator-Placement Considerations for the Control of Large Space Structures," Naval Research Laboratory Report 8675, May 1983.

39. Viswanathan, C. N., Longman, R. W., Likens, P. W., "A Definition of the Degree of Controllability-A Criterion for Actuator Placement," in *Dynamics and Control of Large Flexible Spacecraft; Proceedings of the Second Symposium*, Blacksburg, VA., Jun 21-23, 1979, pp. 369-384.

40. Hughes, P. C., Skelton, R. E., "Controllability and Observability of Linear Matrix-Second-Order Systems", *Journal of Applied Mechanics*, Vol 47, June 1980, pp. 415-420.

41. Hughes, P. C., Skelton, R. E., "Controllability and Observability for Flexible Spacecraft," *Journal of Guidance and Control*, Vol 3, No. 5, Sep-Oct 1980, pp. 452-459.

42. Fogel, E., Hegg, D. R., et al., 'ACOSS-Eleven Second Semiannual Technical Report, Volume 2: Active Controller Designs", Rome Air Development Center Report RADC-TR-82-295, Charles Stark Draper Laboratory, Cambridge, MA, Nov 1982; also Report CSDL-R-1583.

43. Kwakernaak, H., Sivan, R., Linear Optimal Control Systems. John Wiley & Sons, Inc., New York. 1972.

44. Laub, A. J., "Computation of 'Balancing' Transformation," Proceedings Joint Automatic Control Conference, San Francisco, CA, Aug 1980.

45. Bucy, R. S., Joseph, Peter D., Filtering for Stochastic Processes with Applications to Guidance. John Wiley & Sons, New York, 1968.

46. Mehre, R. K., "Optimization of Measurement Schedules and Sensor Designs for Linear Dynamic Systems," *Transactions on Automatic Control*, Vol AC-21, No.1, Feb 1976.

47. Moore, B., "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction," *IEEE Transactions on Automatic Control*, Vol AC-26, No. 1, Feb 31.

48. Glover, K., "All Optimal Hankel-norm Approximations of Linear Multivariable Systems and their L^{∞}-error Bounds", *International Journal of Control*, Vol. 39, No. 6, 1984, pp 1115-1193.

49. Gregory, C. Z., "Reduction of Large Flexible Spacecraft Models Using Internal Balancing Theory," *Journal of Guidance, Control and Dynamics*, Vol. 7, No. 6, Nov-Dec 1984, pp. 725-732.

50 Meirovitch, L. Analytical Methods in Vibrations. Macmillan, NY, 1967.

51. Strang, G., Fix, G., An Analysis of the Finite Element Method. Prentice-Hall, NJ, 1973.

52. Rogers, L., (Editor), "Vibration Damping 1984 Workshop Proceedings", AFWAL-TR-84-3064, Nov 1984.

53. Enns, D., "Model Reduction for Control System Design", Ph.D. Dissertation, Dept. of Aeronautics and Astronautics, Stanford University, June 1984.

54. Blelloch, P. A., Mingori, D. L., and Wei, J. D., "Perturbation Analysis of Internal Balancing for Lightly Damped Mechanical Systems with Gyroscopic and Circulatory Forces," *AIAA Journal of Guidance and Control*, Vol. 10, No. 4, 1987, pp. 406-410.

55. Jonckheere, E. A., "Principal Component Analysis of Flexible Systems–Open-Loop Case," *IEEE Transactions on Automatic Control*, Vol. AC-30, 1985, pp. 690-693.

56. Luenberger, D. G., Linear and Nonlinear Programming. Addison-Wesley, Reading, MA, 1984.

57. Franklin, J. N., Matrix Theory. Prentice-Hall, New Jersey, 1968

58. Papadimitriou, C. H., & Steiglitz, K., Combinatorial Optimization: Algorithms and Complexity. Prentice-Hall, Englewood Cliffs, NJ 1982.

59. Golub, Gene H., and Van Loan, Charles F., Matrix Computations. Johns Hopkins University Press, Baltimore, MD, 1983.

60. Van Loan, Charles F., "Computing Integrals Involving the Matrix Exponential," *IEEE Transactions on Automatic Control*, Vol. AC-23, No. 3, June 1978.

61. Gawronski, W., and Juang, J., "Gramians and Model Reduction in Limited Time and Frequency Intervals", in *Proceedings of AIAA Guidance, Navigation and Control Conference*, 15-17 Aug, 1988, Minneapolis, MN, pp 275-285.

62. Henderson, T., "Active Control of Space Structures (ACOSS) Model 2", Report C-5437, Charles Stark Draper Laboratory, Cambridge, MA, Sep 1981.

63. Henderson, T., "Modifications to ACOSS Model No. 2 Design; Technical Report, Data Base (Final)", Rome Air Development Center Report RADC-TR-83-56, Charles Stark Draper Laboratory, Cambridge, MA, Mar 1983; also Report CSDL-R-1585.

64. Hegg, D. R., McClamroch, N. H., et al., "ACOSS-Eleven Third Interim Report, Vol. 2: Active Control", Rome Air Development Center Report RADC-TR-83-158, Charles Stark Draper Laboratory, Cambridge, MA Dec 1982; also Report CSDL-R-1598, Dec 1982.

65. McClamroch, N. H., "Modeling and Control of Large Flexible Structures Using Electromechanical Actuators", in *Dynamics and Control of Large Structures, Proceedings of the 4th VPI&SU/AIAA Symposium*, Blacksburg, VA, 6-8 Jun 1983, pp. 421-438.

66. Hegg, D.R., Kissel, G. J., McClamroch, N. H., et al., "ACOSS-Eleven Final Technical Report", Rome Air Development Center Report RADC-TR-85-165, Charles Stark Draper Laboratory, Cambridge, MA, Sep 1985; also Report CSDL-R-1721.

67. Bhaya, A., Desoer, C. A., "On the Design of Large Flexible Space Structures", *IEEE Transactions on Automatic Control*, Vol. AC-30, No. 11, Nov 1985, pp. 1118-1120.

68. Stieber, M. E., "Sensors, Actuators, and Hyperstability of Structures", in *Proceedings of AIAA Guidance, Navigation and Control Conference*, August 15-17, 1988, Minneapolis, MN, pp 51-56.

69. Collins, J. D., and Thomson, W. T., "The Eigenvalue Problem for Structural Systems with Statistical Properties", *AIAA Journal*, Vol. 7, No. 4, Apr 1969, pp. 642-648.

70. Schiff, A. J., and Bogdanoff, J. L., "An Estimator for the Standard Deviation of a Natural Frequency", *Journal of Applied Mechanics*, Jun 1972, pp. 535-544.

71. Mehta, M. L., Random Matricies and the Statistical Theory of Energy Levels. Academic Press, NY, 1967.

72. Lim, K. B., Junkins, J. L., and Wang, B. P., "Re-examination of Eigenvector Derivatives", *Journal of Guidance and Control*, Vol. 10, No. 6, Nov-Dec 1987, pp. 581-587.