2/45

SANDIA REPORT

SAND95–0843 • UC–705
Unlimited Release
Printed April 1995

System Identification for Robust Control Design

Jeffrey L. Dohner

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550
for the United States Department of Energy
under Contract DE-AC04-94AL85000

Approved for public release; distribution is unlimited.

MASTER

SF2900Q(8-81)

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED 85

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
Office of Scientific and Technical Information
PO Box 62
Oak Ridge, TN 37831

Prices available from (615) 576-8401, FTS 626-8401

Available to the public from
National Technical Information Service
US Department of Commerce
5285 Port Royal Rd
Springfield, VA 22161

NTIS price codes
Printed copy: A03
Microfiche copy: A01

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Distribution
Category UC-705

SAND95-0843
Unlimited Release
Printed April 1995

System Identification for Robust Control Design

Jeffrey L. Dohner
Structural Dynamics and Vibration Control
Department 1434
Sandia National Laboratories

Abstract

System identification for the purpose of robust control design involves estimating a nominal model of a physical system and the uncertainty bounds of that nominal model via the use of experimentally measured input/output data. Although many algorithms have been developed to identify nominal models, little effort has been directed towards identifying uncertainty bounds. Therefore, in this document, a discussion of *both* nominal model identification *and* bounded output multiplicative uncertainty identification will be presented.

This document is divided into several sections. Background information relevant to system identification and control design will be presented. A derivation of eigensystem realization type algorithms will be presented. An algorithm will be developed for calculating the maximum singular value of output multiplicative uncertainty from measured data. An application will be given involving the identification of a complex system with aliased dynamics, feedback control, and exogenous noise disturbances. And, finally, a short discussion of results will be presented.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Acknowledgments

This work was funded by the National Center for Advanced Information Components Manufacturing (NCAICM). Management of this program was by the Advanced Research Program Agency (ARPA).

Special thanks goes to the program manager, Mr. Steve McBurnett, for his insight and support. Additional thanks goes to the many program participants from AT&T Federal Systems, METRON, Los Alamos National Laboratories, and Sandia National Laboratories.

Table of Contents

List of figures i
List of tables ii

1.0 Introduction 1
1.1 System identification for the purpose of control design 1
1.2 The role of model uncertainty in control design 2
1.3 A discussion of controllability and observability 7
1.3.1 Controllability in discrete systems 7
1.3.2 Observability in discrete system 8
1.3.3 The dual of the system 9

2.0 Eigensystem realization type algorithms 10
2.1 Markov parameters 10
2.2 Calculation of Markov parameters 10
2.2.1 Probe signal excitation 11
2.2.2 Data collection and averaging 13
2.2.3 Markov parameter estimation 13
2.3 Hankel matrix construction 13
2.4 The Eigensystem Realization Algorithm(ERA) 14
2.5 The Eigensystem Realization Algorithm with Data Correlation (ERA/DC) 15
2.6 State elimination using a balanced realization 16
2.7 Determining nominal model uncertainty bounds 17

3.0 An application - the magnetically levitated lithography stage 20
3.1 A model of the lithography fine stage 20
3.2 Nominal model accuracy for variations in feedback control and noise 27
3.3 Stability in the face of model uncertainty 28
4.0 A discussion of results 29
References 30
Appendix A 31

List of Figures

Figure 1.1 Active vibration control of a magnetically levitated lithography magnetically levitated lithography fine stage 2
Figure 1.2 Transfer function response of a vibrational system 3
Figure 1.3 An active control system 4
Figure 1.4 Performance, transition, and robustness bands 5
Figure 1.5a,b Output multiplicative uncertainty 5
Figure 2.1 Nominal model calculation 11
Figure 2.2 System excitation for the purpose of calculating output multiplicative uncertainty bounds 18
Figure 3.1a Continuous time fine stage dynamics 21
Figure 3.1b Simulink fine stage model 23
Figure 3.1c Simulink lithography machine model 24
Figure 3.1d Simulink system ID/vibration control model 25
Figure 3.2 Nominal model error for variations in control feedback and noise 27
Figure 3.3 Stability relative to model uncertainty 28
Figure A.1 A single input, single output feedback control system 31
Figure A.2 Mapping from the s domain to the F(s) domain 31
Figure A.3 Stable and unstable closed loop systems 33

List of Tables

Table 1: Response to arbitrary input, $u(i)$ 8
Table 2: Response to initial conditions, x_d 8
Table 3: Impulse response of a matrix first order system 10
Table 4: Natural frequencies and modal damping 22

1.0 Introduction

In this section, a brief discussion of controllability, observability, system duality and robust stability will be presented. Knowledge of this information is fundamental to future discussions.

1.1 System identification for the purpose of control design

There are many researchers, scientists, and engineers working in the area of system identification, and each of these individuals interprets the field of system identification relative to his own interests. Therefore, to establish commonality, a short discussion of system identification relative to control design follows.

Although there are many methods of control design, discussions in this document will focus on the use of popular, model based, modern, control design methods. In these methods, a first order mathematical model of system dynamics must be determined before a controller can be constructed. This mathematical model is determined from measured input/output data.

In Figure 1.1, a control design problem is illustrated. In this figure, a lithography fine stage[1] is instrumented with a set of actuators and sensors. The actuators apply force loads to the system, and the sensors measure some time derivative of vibratory displacement. This vibratory displacement is due to actuator force and exogenous disturbance force loadings.

Actuators are driven by power amplifiers which are driven by a set of Digital to Analog Converters (DAC's). The input to the DAC's is a stream of numerical input data, $u(i)$.

Sensors drive preamplifiers which drive amplifiers which drive anti-aliasing filters which drive Analog to Digital Converters (ADC's). The output of the ADC's is a stream of numerical output data $y(i)$.

Control is performed by numerically relating inputs, $u(i)$, to outputs, $y(i)$, via a control law. The physical device used to implement this control law is called a processor. Together, the actuators, sensors, amplifiers, filters, ADC's, DAC's and processor are referred to as the control system.

To derive a control law, a model of system dynamics from $u(i)$ to $y(i)$ must be found. This control law is designed to change system dynamics so as to achieve a desired objective. For example, in the Figure 1.1 lithography fine stage system, the control law is designed to minimize stage settling time. In most model based modern control design methods the model from $u(i)$ to $y(i)$ is required to be in first order state space form. That is,

$$x(i+1) = Ax(i) + Bu(i) \quad \textbf{(1.1a)}$$

$$y(i) = Cx(i) + Du(i) \quad \textbf{(1.1b)}$$

where $u(i) \in R^{n_{in}\times 1}$ is the input vector, $y(i) \in R^{n_{out}\times 1}$ is the output vector, $x(i) \in R^{n_{states}\times 1}$ is the state vector, $A \in R^{n_{states}\times n_{states}}$, $B \in R^{n_{states}\times n_{in}}$, $C \in R^{n_{out}\times n_{states}}$, and $D \in R^{n_{out}\times n_{in}}$ are state matrices, n_{in} is the number of DAC inputs, n_{out} is the number of

ADC outputs, n_{states} is the number of states, and i is a discrete time index. It is assumed that data is sampled at even intervals of time separated by the temporal increment Δt. Equation 1.1a,b is called the *nominal model.*

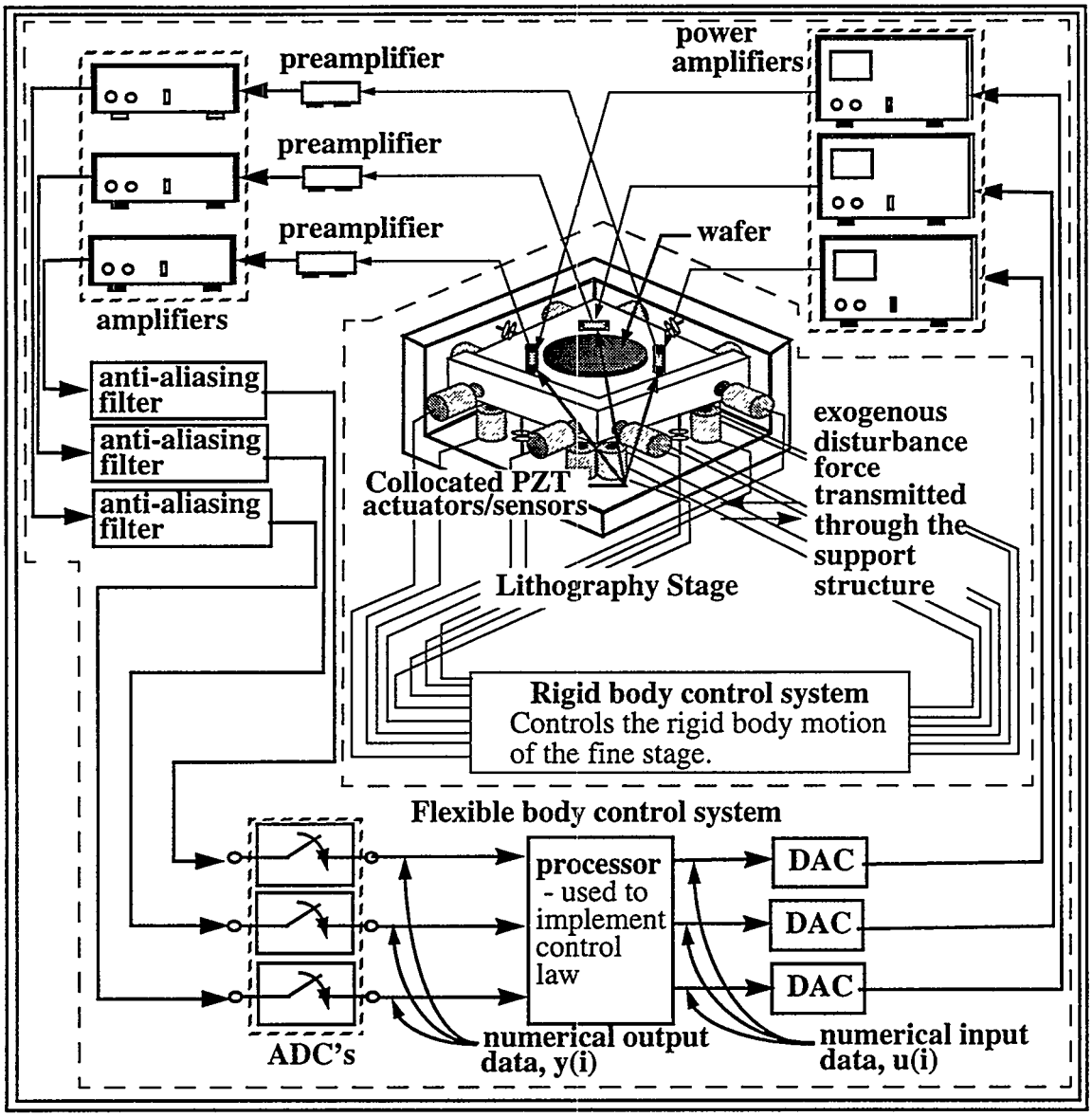


Figure 1.1 Active vibration control of a magnetically levitated lithography fine stage

1.2 The role of model uncertainty in control design

The equation 1.1ab nominal model represents a linear, factorizable, proper, time invariant system of finite order. A *linear* system is one for which the principle of superposition applies. That is, in a linear system, the response of the system to the sum of two distinct inputs will be the summed response of the system to each input acting alone. Non-linear systems dynamics such as backlash, saturation, and stiffness hardening will force system response not to behave in this fashion. *Factorizable* implies that the z domain transfer function between any DAC input and any ADC output can be represented in terms of poles

and zeros. Dynamics such as pure time delays cannot be represented using models of this form. *Proper* implies that the magnitude of the frequency domain transfer function between any input and any output is bounded as frequency approaches infinity. *Strictly proper* implies that this magnitude goes to zero as frequency approaches infinity. A strictly proper system is proper, but a proper system is not necessarily strictly proper. All real systems are strictly proper. *Time invariant* implies that state matrices are constant. And *finite order* implies that n_{states} is a positive integer less than infinity. No physical system dynamics can be represented exactly by a linear, factorizable, time invariant, finite order model, but, many physical systems can be *approximated* by linear, factorizable, proper, time invariant, finite order, nominal models. The degree to which a nominal model represents the physical system is quantified by an uncertainty bound.

A control law is designed from a nominal model, and that model, as stated above, never accurately represents the physical system to be controlled. Therefore, when the control law is implemented, it must act upon well modeled dynamics with certainty so as to produce effective system responses while acting upon poorly modeled dynamics with uncertainty so as not to produce instabilities and degraded performance.

For example, Figure 1.2 shows the transfer function of a typical structural dynamic system. At low frequencies, the modes of the system are well spaced and easily identified. At high frequencies there are an infinite number of uncertain modes. No numerical model can represent the dynamics of all these high frequency modes. Therefore, a compromise between identification and control is made. At low frequencies, the control system acts upon well modeled dynamics with certainty to produce effective system responses, and at high frequencies the control system acts upon poorly modeled dynamics with uncertainty so as not to produce instabilities. At high frequencies the control system is robustly stable.

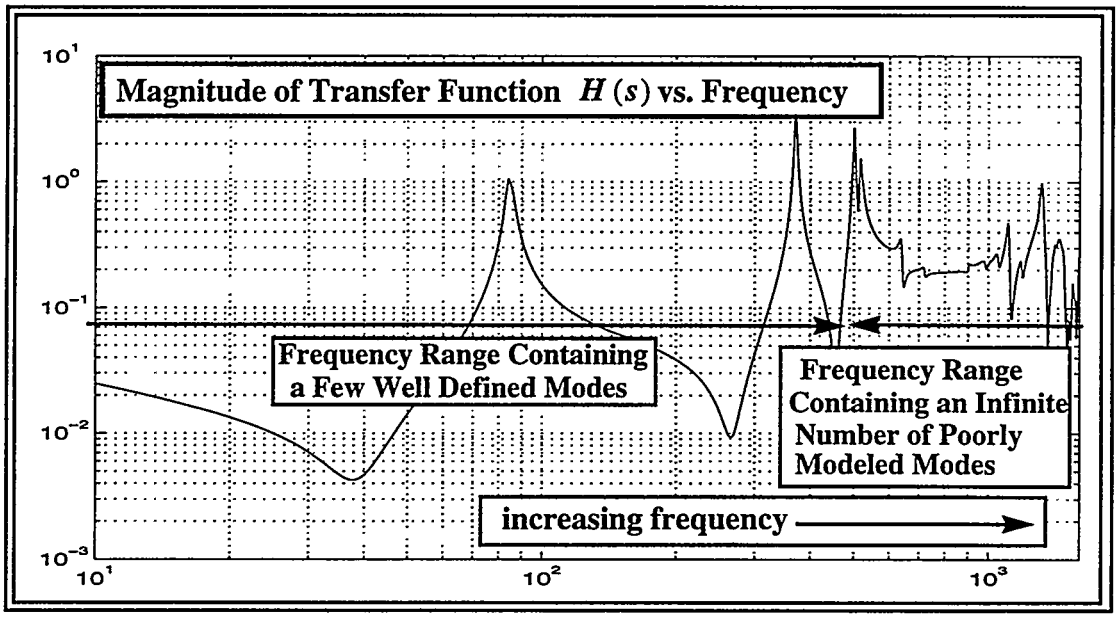


Figure 1.2 Transfer function response of a vibrational system

Robust stability is usually imposed by gain stabilizing the closed loop system. Using Figure 1.3, gain stabilization can be explained for a Single Input, Single Output, (SISO), system.

In this figure numerical input data, $u(i)$, and the numerical output data, $y(i)$, have both been transformed into the z domain as $\mathbf{u}(z)$ and $\mathbf{y}(z)$ respectively [2], and DAC/ADC, input/output dynamics are represented in the z domain by the function

$$\mathbf{H}(z) = \mathbf{y}(z)/\mathbf{u}(z) .$$

The output of $\mathbf{H}(z)$, $\mathbf{y}(z)$, is subtracted from a zero reference signal, $\mathbf{r}(z) = 0$ to produce an error signal, $\mathbf{e}(z)$. This error signal is input to the control law, $\mathbf{G}(z)$, whose output is $\mathbf{u}(z)$. The control law is designed such that $\mathbf{y}(z)$ follows $\mathbf{r}(z)$. That is, the closed loop system produces command following. For the case when $\mathbf{r}(z) = 0$, good command following implies good disturbance rejection to any exogenous disturbance, $w(j\omega)$.

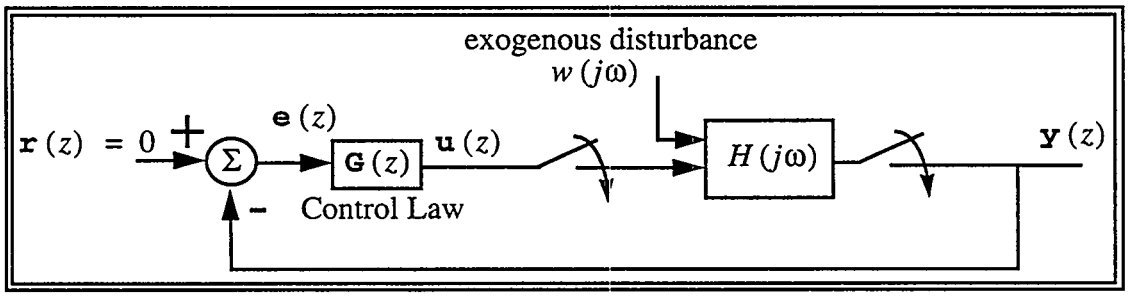


Figure 1.3 An active control system

Gain stabilization simply reduces the magnitude of $\mathbf{G}\left(z = e^{j\omega\Delta t}\right)$ to be near zero over poorly modeled frequency bands*. Therefore, little or no control effort is used over these bands and no change in stability from open loop stability can occur due to uncertainty.

Robustness, achieved through gain stabilization, has a significantly effect on performance. Achieving robust stability through gain stabilization is the same as driving the magnitude of the closed loop transfer function towards zero over poorly modeled frequency bands. That is, if $T(j\omega)$ is the closed loop transfer function, then

$$T(j\omega) = \frac{\mathbf{y}\left(z = e^{j\omega\Delta t}\right)}{\mathbf{r}\left(z = e^{j\omega\Delta t}\right)} = \frac{\mathbf{G}\left(z = e^{j\omega\Delta t}\right)\mathbf{H}\left(z = e^{j\omega\Delta t}\right)}{1 + \mathbf{G}\left(z = e^{j\omega\Delta t}\right)\mathbf{H}\left(z = e^{j\omega\Delta t}\right)} \tag{1.2a}$$

must be zero or small over the frequency band of uncertain dynamics since $\mathbf{G}(z)$ is zero or small over this band.

Good performance implies good command following and good command following implies that $\mathbf{e}\left(z = e^{j\omega\Delta t}\right)$ is zero over the bandwidth of well modeled dynamics. If $S(j\omega)$ is the transfer function from $\mathbf{r}\left(z = e^{j\omega\Delta t}\right)$ to $\mathbf{e}\left(z = e^{j\omega\Delta t}\right)$, then

$$S(j\omega) = \frac{\mathbf{e}\left(z = e^{j\omega\Delta t}\right)}{\mathbf{r}\left(z = e^{j\omega\Delta t}\right)} = \frac{1}{1 + \mathbf{G}\left(z = e^{j\omega\Delta t}\right)\mathbf{H}\left(z = e^{j\omega\Delta t}\right)} \tag{1.2b}$$

*This is a very simplified explanation. More detail is given at the end of this section and in Appendix A.

must be zero or small over this band. From the above equations it can be seen that $T(j\omega)$ and $S(j\omega)$ cannot both be zero over all frequencies since $T(j\omega) + S(j\omega) = 1$. Thus, when robustness is high, performance is low, and when performance is high, robustness is low.

For "real" dynamics neither $T(j\omega)$ nor $S(j\omega)$ can be made to go to zero over an infinitesimally small $\Delta\omega$. Therefore, as shown in Figure 1.4, three frequency bands exist. In the first band $|S(j\omega)|$ is small and $|T(j\omega)|$ is near unity. This is the performance band. In the second band, $|T(j\omega)|$ and $|S(j\omega)|$ are neither small nor near unity. This is the transition band. In the third band, $|T(j\omega)|$ is small and $|S(j\omega)|$ is near unity. This is the robustness band.

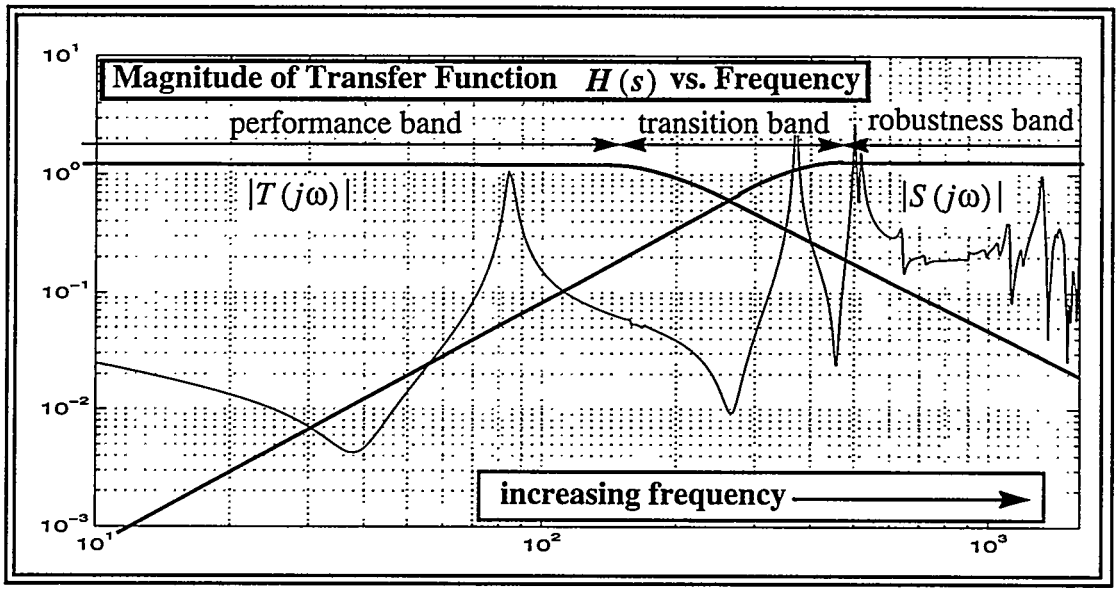


Figure 1.4 Performance, transition, and robustness bands

To determine the amount of uncertainty which can be included in the nominal model before instability occurs, model uncertainty must be represented mathematically and quantified. One representation of this uncertainty is shown in Figure 1.5.

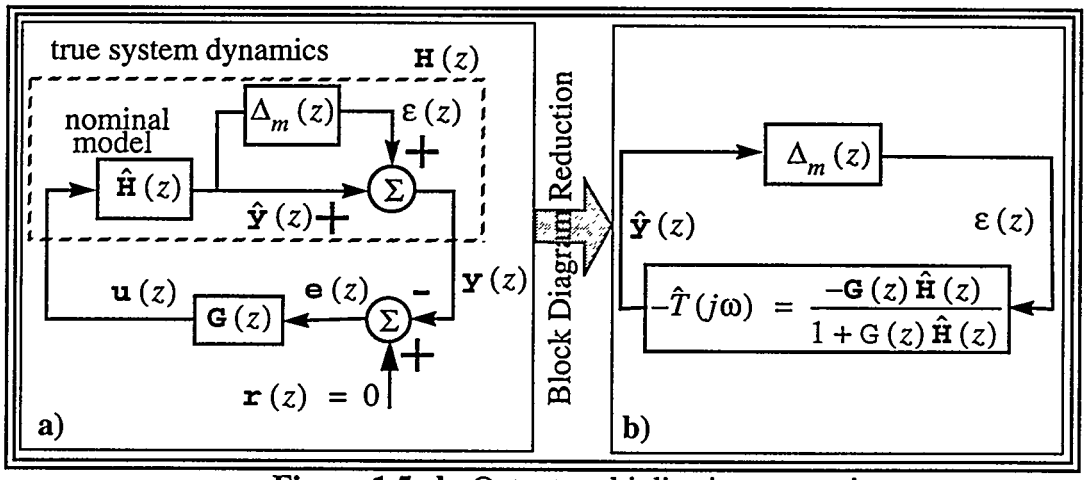


Figure 1.5a,b Output multiplicative uncertainty

Figure 1.5a is the Figure 1.3 system with $\mathbf{H}(z) = (1+\Delta_m(z))\hat{\mathbf{H}}(z)$ where $\Delta_m(z)$ is called the output multiplicative uncertainty of the nominal model, $\hat{\mathbf{H}}(z)$. Figure 1.5b is the loop reduction of Figure 1.5a.

If both $\hat{T}(j\omega)$ and $\Delta_m(z)$ are stable*, then from the Nyquist stability theorem (derived in appendix A), a bound for stability is given by

$$|\hat{T}(j\omega)| < \frac{1}{\left|\Delta_m\left(z = e^{j\omega\Delta t}\right)\right|}. \qquad \textbf{(1.3a)}$$

Equation 1.3a can be understood from the standpoint of energy production. If both $\hat{T}(j\omega)$ and $\Delta_m(z)$ are stable systems, then the net storage of energy in each of these systems is bounded. Moreover, if the net production of energy from any system input to any system output is also bounded, then the system is stable since total energy (stored energy + produced energy) will also be bounded. Therefore, for a SISO system, if $|\hat{T}(j\omega)|\left|\Delta_m\left(z = e^{j\omega\Delta t}\right)\right|$ is less than one for any ω and if both $\Delta_m(z)$ and $\hat{T}(j\omega)$ are stable, then the system must be stable. This is a statement of equation 1.3a.

Equation 1.3a can be extended to Multiple Input/Multiple Output (MIMO) systems by the use of singular values[3]. If equation 1.3a is valid for every possible combination of loops in a MIMO system, the system will be stable. Singular values are used to bound the maximum magnitudes of every possible combination of loops.

For a system, $\mathbf{R}(z)$ where $v(z) = \mathbf{R}(z)r(z)$, the square root of the ratio of energy out of the system over energy into the system due to the i^{th} input is given by

$$\Psi_i(z) = \frac{v_i^H(z)\,v_i(z)}{r_i^H(z)\,r_i(z)} \le \sigma_i^2 \quad \text{where } z = e^{j\omega\Delta t},\ \mathbf{R}^H(z)\mathbf{R}(z) = \phi^H\Lambda\phi,\ \phi^H\phi = I,$$

$\Lambda = diag\left(\sigma_1^2, \sigma_2^2, \sigma_3^2 \ldots\right)$, $\sigma_1^2 \ge \sigma_2^2 \ge \sigma_3^2 \ldots$, $r(z) = \left[r_1^H(z)\ r_2^H(z)\ \ldots\ r^H{}_{n_{in}}(z)\right]^H$, and

$v(z) = \sum_i^{n_{in}} v_i(z)$. Therefore, $\sigma_1^2(j\omega)$ is an upper bound on the ratio of the sum energy out of $\mathbf{R}(z)$ over energy into $\mathbf{R}(z)$ due to any $r_i(z)$. Moreover, it must also be a bound on the ratio of energy out of any output over energy in to any input. Thus, $\sigma_1(j\omega)$ is a bound on the magnitude of the transfer functions contained in $\mathbf{R}(z)$. The parameter $\sigma_1(j\omega)$ is called the maximum singular value of $\mathbf{R}(z)$. Functionally, it is written as $\sigma_1(j\omega) = \bar{\sigma}(\mathbf{R}(z))$.

Therefore, for a MIMO system, if

$$\bar{\sigma}(\hat{T}(j\omega)) < \frac{1}{\chi(\Delta_m(z))} \qquad \textbf{(1.3b)}$$

* If both $\mathbf{H}(z)$ and $\hat{\mathbf{H}}(z)$ are stable, then $\Delta_m(z)$ is stable.

where $\chi(\Delta_m(z))$ is an energy production bound from any input to any output of $\Delta_m(z)$, then the net production of energy around *any* loop in the MIMO loop must be decreasing, and if $\hat{T}(j\omega) = \hat{\mathbf{H}}(z)\mathbf{G}(z)(I+\hat{\mathbf{H}}(z)\mathbf{G}(z))^{-1}$ and $\Delta_m(z)$ are also stable, then the closed loop system is stable.

1.3 A discussion of controllability and observability

When using feedback control, inputs, $u(i)$, force the system in such a way that outputs, $y(i)$, are driven to a desired response. To do this, the system must be both controllable and observable. For time invariant systems, state controllability implies that there exist some input $u(i)$ which, over some time, $N\Delta t$, will drive the states of the system, $x(i)$, to some desired state, x_d. State observability implies that by measuring $y(i)$ over some past time, $N\Delta t$, the state of the system can be determined exactly at $N\Delta t$ time in the past*. The following subsections discuss the mathematical conditions required for controllability and observability.

These conditions will be used in sections 2.3 to 2.5 to derive algorithms for the purpose of identifying a nominal model.

1.3.1 Controllability in discrete systems

There are two parts to the solution of a linear matrix difference equation. The first part of the solution is the homogeneous solution and the second part is the particular solution. The particular solution is the response of the system to input $u(i)$ for zero initial state conditions (i.e. $x(0) = \varnothing$). This input will produce a particular state response, $x_p(i)$, which can be deduced from equation 1.1a. The homogeneous state response, $x_h(i)$, is the state response for zero input, and some initial state condition, x_i. For discrete systems $x_h(i) = A^N x_i$ and $x_p(i) = x(i) - A^N x_i$.

Controllability implies that there is a set of $u(i)$'s which will drive the state response, $x(i)$, to a desired state response, x_d, in a finite amount of time, $N\Delta t$. If $u(i)$ is the set of inputs required to drive $x(i)$ to x_d in $N\Delta t$ time, and if x_i is known, then $u(i)$ can be calculated from equation 1.1a. Table 1 gives $x(i)$ for arbitrary $u(i)$ as a function of i and initial state x_i. As can be seen from this table, to determine a $u(i)$ that will produce a x_d at $N\Delta t$, the set of equations,

$$x_d - A^N x_i = \left[B|AB|A^2B|A^3B|\ldots A^NB\right]\left[u^T(N), u^T(N-1), \ldots u^T(1)\right]^T \qquad \text{(1.4a)}$$

must be solvable. In order for a solution to exist, the matrix $\left[B|AB|A^2B|A^3B|\ldots A^NB\right]$ must be of rank n_{states}. If it is not, then there is no $u(i)$ which will drive $x(i)$ to x_d in

*This is only true for time invariant system. If the system time invariant, this is the definition of reconstructibility.

$N\Delta t$ time. When N is greater than or equal to $n_{states}-1$, the matrix $\left[B|AB|A^2B|A^3B|\ldots A^NB\right] = W$ is called the controllability matrix. If the controllability matrix has rank less than n_{states}, there will be no input which will drive $x(i)$ to x_d in any amount of time. In this situation, the system is said to be uncontrollable.

i	$x(i) - A^N x_i$	$u(i)$
0	$Bu(0)$	$u(0)$
1	$ABu(0) + Bu(1)$	$u(1)$
2	$A^2Bu(0) + ABu(1) + Bu(2)$	$u(2)$
3	$A^3Bu(0) + A^2Bu(1) + ABu(2) + Bu(3)$	$u(3)$
4	$A^4Bu(0) + A^3Bu(1) + A^2Bu(2) + ABu(3) + Bu(4)$	$u(4)$
:	:	:
N	$A^NBu(0) \ldots A^3Bu(N-3) + A^2Bu(N-2) + ABu(N-1) + Bu(N)$	$u(N)$

Table 1: Response to arbitrary input, $u(i)$

1.3.2 Observability in discrete systems

The homogeneous response of the system to an initial condition is tabulated in table 2. From table 2, the initial condition, x_i, can be found by solving the set of equations

$$\left[y^T(0), y^T(1), y^T(2), \ldots y^T(N)\right]^T = \left[C^T|A^TC^T|A^{2^T}C^T|\ldots A^{N^T}C^T\right]^T x_i \qquad \textbf{(1.4b)}$$

where $N \geq n_{states} - 1$. For this solution to exist the matrix $\left[C^T|A^TC^T|A^{2^T}C^T|\ldots A^{N^T}C^T\right]^T$ must have rank n_{states}. The matrix $\left[C^T|A^TC^T|A^{2^T}C^T|\ldots A^{N^T}C^T\right]^T = V$ is called the observability matrix. If the rank of the observability matrix is n_{states} then the system is observable and the state response at some time $N\Delta t$ in the past can be deduced from past inputs, $u(i)$, and outputs, $y(i)$..

i	$y(i)$	$x(i) - x_p(u(N), u((N-1)\ldots))$	$u(i)$
0	Cx_d	x_d	0's
1	CAx_d	Ax_d	0's
2	CA^2x_d	A^2x_d	0's
:	:	:	:
N	CA^Nx_d	A^Nx_d	0's

Table 2: Response to initial condition, x_d

1.3.3 The dual of the system

Notice that if the observability matrix has rank n_{states}, the matrix $\left[C^T \middle| A^T C^T \middle| A^{2^T} C^T \middle| \ldots A^{N^T} C^T\right]$ must also have rank n_{states}. This matrix is identical to the controllability matrix of the system

$$x(i+1) = A^T x(i) + C^T y(i) \tag{1.5a}$$

$$u(i) = B^T x(i) + D^T y(i) \ . \tag{1.5b}$$

Moreover, if the controllability matrix of the equation 1.1a,b system has rank n_{states}, then the matrix $\left[B|AB|A^2B|\ldots A^N B\right]^T$ must also have rank n_{states}. This is the observability matrix of the equation 1.5a,b system. The equation 1.5a,b system is called the dual of the equation 1.1a,b system. The dual system is controllable if the original system is observable, and the dual is observable if the original system is controllable. Therefore, observability of a system can be checked by determining the controllability of its dual, and the controllability of a system can be check by determining the observability of its dual.

The dual of the system will be used in section 2.6 to define model reduction parameters.

2.0 Eigensystem realization type algorithms

In this section, the Eigensystem Realization Algorithm (ERA) and the Eigensystem Realization Algorithm with Data Correlations (ERA/DC) will be derived. Both of these algorithms will produce first order state space realizations* of a system from measured input/output data. Uncertainty bounds, required for robust control design, will also be derived.

2.1 Markov parameters

Markov parameters are used in both the ERA and the ERA/DC algorithms. The Markov parameters of a system are derived from a redistribution of system impulse response data. *Therefore, they can be physically measured.* Table 3 gives the impulse responses of the equation 1.1a,b system.The Markov parameters, $Y(i)$, are D, CB, CAB, CA^2B ...etc. In compact form they are given by

$$Y(0) = D \tag{2.1a}$$

$$Y(i+1) = CA^iB \tag{2.1b}$$

i	$Y(i)$	$x(i)$	$u(i)$
0	D	0's	
1	CB	B	
2	CAB	AB	
3	CA^2B	A^2B	
:	:	:	:
N	CA^NB	A^NB	

Table 3: Impulse response of a matrix first order system

2.2 Calculation of Markov parameters

To measure the Markov parameters of a physical system, the system must be excited by a probe signal excitation, the response of the system must be collected and averaged to minimize the effects of noise, estimates of cross and auto spectral response matrices must be produced, transfer function matrices must be calculated, and impulse response functions determined. Figure 2.1 shows this situation when a feedback controller, $\mathbf{G}(z)$, is present.

*a nominal model

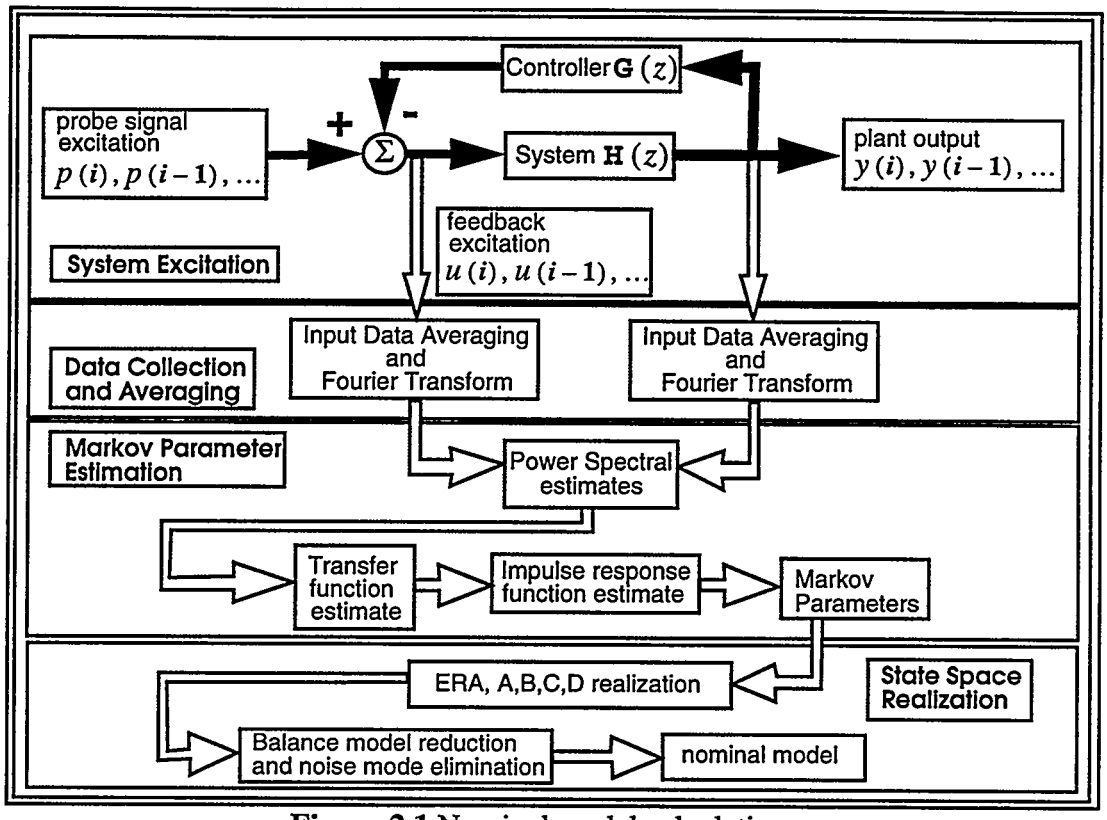


Figure 2.1 Nominal model calculation

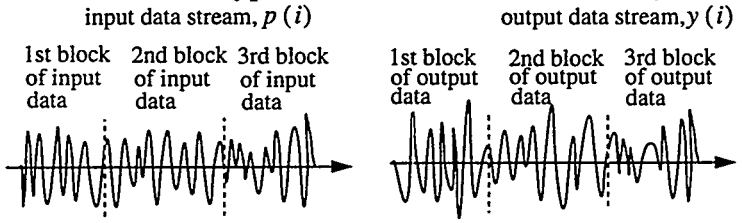
2.2.1 Probe signal excitation

To determine Markov parameters, the physical system, $\mathbf{H}(z)$, is excited by a probe signal excitation, $p(i)$. Some of the characteristics of probe signal excitations are;

1) *Root mean square to peak level-* The root mean square level of the probe signal excitation is proportional to the total amount of energy placed into the system for the purpose of identification. Since peak excitation levels are bounded by the clipping of amplifiers, DAC's or actuators, a signal with a high root mean square to peak level will be able to place more energy into the system than one with a low root mean square to peak level. Therefore, for set noise levels, higher root mean square to peak probe signal levels imply a higher signal to noise ratio.
2) *Distortion-* As stated in section 1.2, all systems have some non-linear, time variant dynamics associated with them. Nevertheless, when determining the transfer functions of a system, it is assumed that the system is linear and time invariant. Therefore, transfer functions estimates will be distorted. Different types of probe signals will produce different levels of distortion.
3) *Leakage-* Leakage is due to the finite truncation of a sampled, infinite duration response. This truncated sampled response is used in the Discrete Fourier Transform (DFT) algorithm to produce approximate system spectral responses in the z domain. Different types of probe signals will produce different levels of leakage.

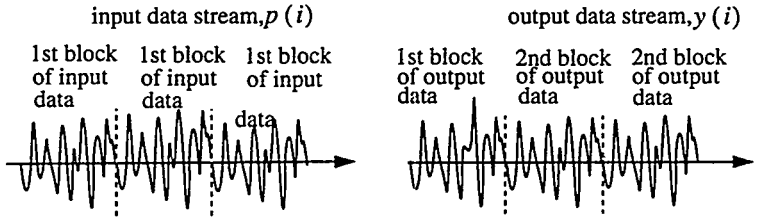
Data is collected from a system in data blocks. A block of data is a sampling of the input signal, $u(i)$, and the output signal, $y(i)$, for a sequential, finite number of evenly space time steps. Probe signal excitations are defined relative to these data blocks. An explanation of some possible types of *random* probe signal excitations follow. Similar type of excitations are available for sinusoidal excitations.

1) *Pure random excitations*- This type of excitation is a continuously random excitation.



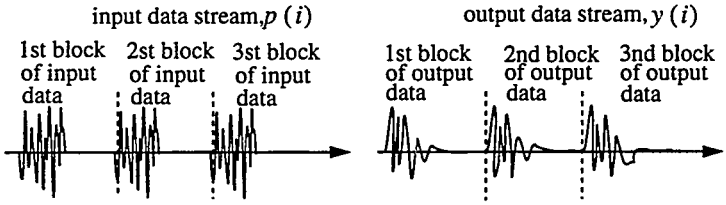
Therefore, no two blocks of noiseless data are identical. Leakage is a serious problem with this type of excitation, but can be minimized by the use of time windows[4].

2) *Pseudo random excitation*- In this type of excitation, the same random input is repeated every block. Therefore, the output stored in all blocks after the first block is the same, and leakage from one block is absorbed into the next so that the net leakage effect in all blocks, except the first block, is minimized.



3) *Periodic random excitation*- Periodic random excitation is a series of back to back pseudo random excitations. In periodic random excitation, a random block input is repeated a set number of times and then another random block input is repeated the same number of times, and so on.

4) *Burst random excitation*- In this type of excitation, a random excitation excites the system for a fraction of the block length and no excitation excites the system for the rest of the block. This allows the output response to die output before the end of the block and therefore, leakage is minimized. Burst random excitations have low root



mean square to peak levels.

The type of probe signal excitation used to measure Markov parameters will be dependent upon the type of system that is being identified and hardware data acquisition limitations. For example, in acoustic systems, sinusoidal excitations are often used due to there high root mean square to peak levels. Nevertheless, for structural mechanical systems which include rubber mounts or bolted connections, random excitations tend to produce less distortion.

2.2.2 Data collection and averaging

To estimate the transfer functions of a system, data must be collected and stored. If this data is corrupted by noise, averaging must also be performed. Depending on the type of excitation, averaging can occur in the time domain or after DFTing, in the z domain.

2.2.3 Markov parameter estimation

Transfer functions can be estimated by using a H_1, a H_2 or a H_v calculation. The derivation of the H_1, H_2, and H_v calculation is given in reference 5. If $\hat{\mathbf{H}}(z) = C(zI+A)^{-1}B+D$ is the nominal model, $U_m(z)$ is the z transform of the m^{th} input ($u(i)$) data block, and $Y_m(z)$ is the z transform of the m^{th} output ($y(i)$) data block, then an H_1 calculation is given by

$$\hat{\mathbf{H}}(z) = \sum_m Y_m(z)\, U_m^H(z) \left(\sum_m U_m(z)\, U_m{}^H(z) \right)^{-1} \tag{2.2a}$$

and an H_2 calculation is given by

$$\hat{\mathbf{H}}(z) = \sum_m Y_m(z)\, Y_m^H(z) \left(\sum_m U_m(z)\, Y_m{}^H(z) \right)^{-1} \tag{2.2b}$$

where the superscript H stand for conjugate transpose. The calculation of H_v is given in the reference 5. A H_v calculation is computationally more intensive than a H_1, or a H_2 calculation, but it is also more accurate. A H_2 computation requires that $\sum_m U_m(z)\, Y_m{}^H(z)$ is invertible.

If $Y(k)$ is the inverse z transform of $C(zI+A)^{-1}B+D$, then $Y(k)$ is the k^{th} Markov parameter of the system. Inverse and forward z transforms are computed by using the inverse and forward DFT.

2.3 Hankel matrix construction

Markov parameters are assembled into a matrix called a Hankel matrix, $H(k)$. This matrix is used in ERA and ERA/DC to form a realization. This matrix is formed by stacking Markov parameters into a matrix as

$$H(k) = \begin{bmatrix} Y(k+1) & Y(k+2) & Y(k+3) & \circ & \circ \\ Y(k+2) & Y(k+3) & Y(k+4) & \circ & \circ \\ Y(k+3) & Y(k+4) & Y(k+5) & \circ & \circ \\ \circ & \circ & \circ & \circ & \\ \circ & \circ & \circ & & \circ \end{bmatrix}. \tag{2.3a}$$

The Markov parameters are arranged in this fashion since

$$H(k) = \begin{bmatrix} CA^{k}B & CA^{k+1}B & CA^{k+2}B & \circ & \circ \\ CA^{k+1}B & CA^{k+2}B & CA^{k+3}B & \circ & \circ \\ CA^{k+2}B & CA^{k+3}B & CA^{k+4}B & \circ & \circ \\ \circ & \circ & \circ & \circ & \\ \circ & \circ & \circ & & \circ \end{bmatrix} = VA^{k}W. \qquad \textbf{(2.3b)}$$

where, from equation 1.4a,b, V and W are the controllability and observability matrices of the system.

2.4 The Eigensystem Realization Algorithm (ERA)

Once the Markov parameters of the system have been measured, a nominal model can be determined by using ERA or ERA/DC and model reduction. In this section and the next, the derivation of ERA and ERA/DC is presented. In section 2.6 a singular value model reduction method with be described.

The singular value decomposition of a real valued matrix $M \in R^{m \times l}$ is given by

$$M = PdQ^{T}.$$

where

$$\text{if } m \geq l \text{ then } d = diag(\sigma_1, \sigma_2, \sigma_3 \ldots \sigma_l) \in R^{l \times l},\ P \in R^{m \times l},\ Q^{T} \in R^{l \times l},$$

$$P^{T}P = Q^{T}Q = QQ^{T} = I_l, \text{ and } I_l \text{ is an identity matrix,}$$

and

$$\text{if } m \leq l \text{ then } P \in R^{m \times m},\ d = diag(\sigma_1, \sigma_2, \sigma_3 \ldots \sigma_m) \in R^{m \times m},\ Q^{T} \in R^{m \times l},$$

$$Q^{T}Q = P^{T}P = PP^{T} = I_m, \text{ and again } I_m \text{ is an identity matrix[3].}$$

As in section 1.2, the values $\sigma_1, \sigma_2, \sigma_3 \ldots \sigma_l$ are called singular values. These values are real and positive and are usually sorted such that $\sigma_1 \geq \sigma_2 \geq \sigma_3 \ldots \geq \sigma_l \geq 0$.

The singular value decomposition of $H(0)$ gives $H(0) = PdQ^{T} = VW$. Therefore, *one possibility* for W and V is given by

$$\left[B|AB|A^{2}B|A^{3}B|\ldots A^{N}B\right] = W = d^{\frac{1}{2}}Q^{T} \qquad \textbf{(2.4a)}$$

$$\left[C^{T}|A^{T}C^{T}|A^{2^{T}}C^{T}|\ldots A^{N^{T}}C^{T}\right]^{T} = V = Pd^{\frac{1}{2}} \qquad \textbf{(2.4b)}$$

Notice two aspects of the above equations. First, this selection of V and W is not unique. This is reasonable considering that C, B, or A are not unique for a given set of dynamics. Second, the number of rows or columns of d must equal the number of columns in V and the number of rows in W, and this row or column dimension must be equal to n_{states}.

Also notice that if the singular value decomposition of $H(0)$ produces singular values equal to zero for all $n_i < n \leq n_{states}$, rows and columns of P, d, and Q associated with these zero singular values can be truncated and the product PdQ^T will still exactly equal $H(0)$. Thus, the loss of states associated with these zero singular values does not change modeled dynamics. By eliminating these non-contributing states, a model with a minimal number of states will be produced. A model which contains a minimal number of states is called a *minimal realization.*

Minimal realizations exist for idealized analytical systems, but seldom for real systems. This is due to the fact that real systems are seldom linear, factorizable, time invariant and of finite order. Therefore, they cannot be represented exactly by equation 1.1 and, as a result, produce a continuous distribution of non-zero singular values.

By knowing V, W, n_{in} and n_{out}, the B and C matrices can be determined as

$$B = WE_B \tag{2.5a}$$

$$C = E_C V \tag{2.5b}$$

where E_B and E_C are matrices of zeros and ones used to filter B and C from the 2.4a, and 2.4b observability and controllability matrices. Substituting equations 2.4a,b into equation 2.2, solving for A^k, and letting $k = 1$ gives

$$A = d^{-\frac{1}{2}} P^T H(1) Q d^{-\frac{1}{2}}. \tag{2.5c}$$

Equation 2.1a, and 2.5a,b,c is the ERA algorithm[6]. ERA is implemented by calculating Markov parameters from experimental data, assembling the $H(0)$ and $H(1)$ Hankel matrices, performing a singular value decomposition of $H(0)$, and using equation 2.1a, and 2.5a,b,c to calculate a D, B, C, and A.

2.5 The Eigensystem Realization Algorithm with Data Correlation (ERA/DC)

In both ERA and ERA/DC a Hankel matrix is used to produce a system realization. Nevertheless, in ERA a singular value decomposition on $H(0)$ is performed, whereas in ERA/DC a singular value decomposition on $R(0) = H(0)H^T(0)$ is performed when $H(k)$ has more columns than rows, and on $R(0) = H^T(0)H(0)$ when $H(k)$ has more rows than columns. This results in the decomposition of a matrix which is dimensionally small. Therefore, for large variations in the number of rows and columns of the Hankel matrix, ERA/DC is numerically more efficient. The matrix, $R(k) = H(k)H^T(0)$, is called the correlation matrix.

If $H(k)$ has more columns than rows, this decomposition become

$$R(0) = PdQ^T = H(0)H^T(0) = VWW^TV^T = VW_c$$

where $W_c = WW^TV^T$. Therefore, letting $V = Pd^{\frac{1}{2}}$, and $W_c = d^{\frac{1}{2}}Q^T$ gives

$$C = E_C P d^{\frac{1}{2}}, \tag{2.6a}$$

and since $R(k) = H(k)H^T(0) = VA^kWW^TV^T = VA^kW_c = Pd^{\frac{1}{2}}A^k d^{\frac{1}{2}}Q^T$, then

$$A = d^{-\frac{1}{2}}P^T H(1)\, Q d^{-\frac{1}{2}}. \tag{2.6b}$$

Moreover, since $H(0) = VW = Pd^{\frac{1}{2}}W$ then $W = d^{-\frac{1}{2}}P^T H(0)$ and

$$B = d^{-\frac{1}{2}}P^T H(0)\, E_B \quad . \tag{2.6c}$$

On the other hand, if $H(0)$ has more rows than columns then the matrix $R(k) = H(0)^T H(k)$ is used. Following a similar approach gives

$$B = d^{\frac{1}{2}}Q^T E_B, \tag{2.6d}$$

$$C = E_C H(0)\, Q d^{-\frac{1}{2}}, \tag{2.6e}$$

and equation 2.6b for A. Equation 2.1a, and 2.6a,b,c,d,e is the ERA/DC algorithm[7].

2.6 State elimination using a balanced realization

As stated in the previous section, if a state has a zero singular value, then it does not contribute to system response and can be eliminated from the model. Nevertheless, no real world system has states with zero singular values. Therefore, the contribution of a state to system response must be ranked by testing system dynamics.

If impulsive loads are applied to nominal model inputs (i.e. $u(i) = I\delta_i$) then state response for $i > 0$ will be given by the columns of the controllability, matrix, W. Therefore, the correlated state response matrix is given by

$$\mathbf{P} = \sum_{i=1}^{N} x(i)\, x^T(i) = WW^T .$$

This matrix, $\mathbf{P}$, is called the controllability grammian. From equation 2.4a,

$$\mathbf{P} = WW^T = d^{\frac{1}{2}}Q^T Q d^{\frac{1}{2}} = d$$

since $Q^T Q = I$. The matrix d is a diagonal matrix containing ranked singular values. Therefore, if the k^{th} singular value is large, then $\sum_{i=1}^{N}\left(\sum_{j=1}^{n_{in}} x^2{}_{kj}(i)\right)$ is large where $x_{kj}(i)$

is the k^{th} state response for an impulse into the j^{th} input. Vice versa, if the k^{th} singular value is small, then the k^{th} state response is also small. Thus, if the contributions of all states to all outputs were equal, then those states with small singular values could be eliminated from the model since they would contribute little to system response.

Unfortunately, the contribution of all states to all outputs is not equal. Therefore, the observability of the system must be determined. This can be determined by working with the dual system described in section 1.3.3. If impulsive loads (i.e. $u(i) = I\delta_i$) are applied to the inputs of the nominal model dual, state response will be equal to the transpose of the block rows of the observability matrix, V. The correlated state matrix response of the dual system is given by,

$$\mathbf{Q} = \sum_{k=0}^{N} x(k)x^T(k) = V^T V \quad .$$

This matrix, $\mathbf{Q}$, is called the observability grammian. From equation 2.4b, $\mathbf{Q}$ is also equal to d.

Therefore, for ERA, the controllability and observability grammians are equal and diagonal. This is called a balanced realization[8,9]. In this type of realization, the contribution of the states to the output are equally ranked to the response of the states to the input. Thus, states which do not contribute significantly to system response will have "small" singular values and dynamics which can be eliminated from the realization.

ERA produces a balanced realization, whereas ERA/DC does not produce a balanced realization. In ERA/DC, $\mathbf{P} = d^{\frac{1}{2}} P^T R(0) P d^{\frac{1}{2}}$, and $\mathbf{Q} = d$ when $H(k)$ has more columns than rows, and $\mathbf{P} = d$, and $\mathbf{Q} = d^{\frac{1}{2}} Q^T R(0) Q d^{\frac{1}{2}}$ when $H(k)$ has more rows than columns. Therefore, it is not obvious that a realization obtained through ERA/DC is balanced. Nevertheless, the singular values in ERA/DC can be related to a system response for which it is balanced. Thus, in a fashion similar to ERA, model reduction by the deletion of small singular values is usually performed.

2.7 Determining nominal model uncertainty bounds

A nominal model can be constructed by using ERA or ERA/DC to form a realization (sections 2.4 and 2.5), and by using singular values to perform model reduction (section 2.6). This nominal model will represent most, but not all, physical system dynamics. Therefore, without checking for instabilities caused by nominal model uncertainty, a controller designed from this nominal model will not necessarily produce closed loop stability. To assure stability in the presence of model uncertainty, a measured bound on nominal model uncertainty is required.

Equation 1.3b was derived from the MIMO representation of Figure 1.5b. This equation is a mathematical statement which, if satisfied, will assure the stability of the closed loop

system in the presence of uncertainty, $\Delta_m(z)$. To use equation 1.3b, the bound, $\chi(\Delta_m(z))$, must be found. This is performed by exciting each input into $\Delta_m(z)$, $\hat{y}(z)$, separately and measuring the resulting outputs of $\Delta_m(z)$, $\varepsilon(z)$. The bound, $\chi(\Delta_m(z))$, is the maximum ratio of energy out over energy in on a frequency by frequency basis. This can be calculated from power spectrums of $\hat{y}(z)$ and $\varepsilon(z)$.

The difficulty with using this approach is that both $\hat{y}(z)$ and $\varepsilon(z)$ are parameters internal to the system and therefore, cannot be directly excited or measured. Therefore, they must be indirectly excited or calculated from measured data. The parameter $\hat{y}(z)$ is excited by determining a filter $\mathbf{F}(z)$ such that $\hat{\mathbf{H}}(z)\mathbf{F}(z)$ is equal to a diagonal matrix with non-zero magnitude response. If there is no control feedback, then for a random excitation, $\mathrm{P}(z) = [p_1(z), p_2(z), \ldots]^T$ where $p_n(z) = 0$ for all $n \neq m$, $\hat{y}_n(z) = 0$ for all $n \neq m$ where $\hat{y}(z) = [\hat{y}_1(z), \hat{y}_2(z), \ldots]^T = \hat{\mathbf{H}}(z)\mathbb{F}(z)\mathrm{P}(z)$. Therefore an input, $\mathbf{u}(z) = \mathbf{F}(z)\mathrm{P}(z)$, can be constructed which will excite only one input to $\Delta_m(z)$ at a time. The output of $\Delta_m(z)$, $\varepsilon(z)$ can then be calculated from the equation

$$\varepsilon(z) = \mathbf{y}(z) - \hat{\mathbf{y}}(z) \tag{2.7a}$$

where

$$\hat{\mathbf{y}}(z) = \hat{\mathbf{H}}(z)\mathbf{u}(z). \tag{2.7b}$$

The block diagram of this system is shown in Figure 2.2.

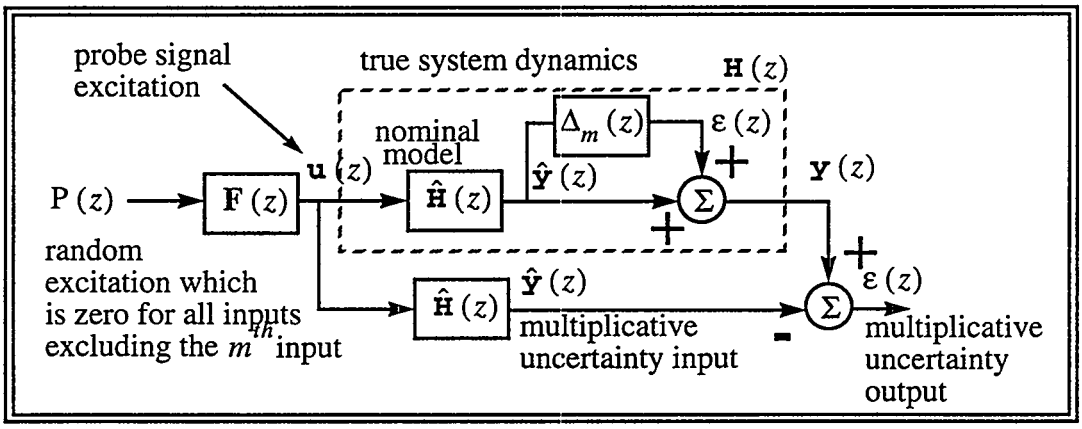


Figure 2.2 System excitation for the purpose of calculating output multiplicative uncertainty

One way to calculate a filter $\mathbf{F}(z)$ is to solve a series of control problems which will suppress all $\hat{y}_n(z)$ for all $n \neq m$. If the dynamics of $\hat{y}_m(z)$ are not strongly coupled to the dynamics of all other $\hat{y}_n(z)$, then $\hat{y}_m(z)$ will be non-zero. These control problems can be solved by using a Linear Quadratic Regulator (LQR)[12] solution for each element in $\hat{y}_n(z)$. If $H(z) = C(zI + A)^{-1}B + D$ then LQR can be used to determine a matrix $\mathcal{F}_m$

such that $\lim_{\gamma \to 0} \lim_{K \to \infty} \frac{1}{K} \sum_{i=0}^{K-1} \underset{\sim}{\hat{y}}_m (i+1)^T \underset{\sim}{\hat{y}}_m (i+1) + \gamma u^T (i) u (i)$ is minimized for $u (i) = -\mathcal{F}_m u (i)$ where $\underset{\sim}{\hat{y}}_m (i)$ is equal to $\hat{y} (i)$ except for the deletion of its m^{th} row. This problem is solved for every row of $\hat{y} (i)$. Therefore, the resulting filter, $\mathbf{F} (s)$, is given by $\mathbf{F} (s) = [\mathbf{F}_1 (z), \mathbf{F}_2 (z) \ldots]$ where $\mathbf{F}_m (z) = -\mathcal{F}_m (zI - A - B\mathcal{F}_m)^{-1} B_m + D_m$, B_m is the m^{th} column of B, and D_m is the m^{th} column of D.

3.0 An application - the magnetically levitated lithography stage

In this section, the identification and control of the flexible body dynamics of a magnetically levitated lithography fine stage is described. The stage, shown in Figure 1.1, consists of an aluminum platen suspended in space by a distribution of electro-magnets. A rigid body controller is used to control the rigid body position of the stage. The bandwidth of this rigid body controller determines the speed at which the stage can be moved in space. Nevertheless, this bandwidth is limited by the destabilization of flexible body modes. Therefore, a flexible body controller was constructed to reduce this destabilization so that rigid body bandwidth, and therefore speed, could be increased

3.1 A model of the Figure 1.1 system

In this section, a model of the Figure 1.1 system will be presented. System identification will be performed on this model. This model contains attributes which are expected to fully and realistically test the limitations of any system identification algorithm. These attributes are; i) modally dense flexible body dynamics forced by rigid body motion, ii) non-linearities due to actuator saturation effects, iii) feedback control, iv) aliased modal dynamics, and v) external noise excitations correlated with plant dynamics.

A model of the Figure 1.1 pictorial was constructed using Simulink[10]. Block diagrams of this model are shown in Figures 3.1a to 3.1d.The innermost dynamics of this model are shown in Figure 3.1a. In this figure a 60th order continuous time state space system representing the rigid and flexible body dynamics of the platen is shown cascaded with piezo-electric actuator and anti-aliasing filter dynamics.

This 60th order continuous time first order system was constructed from a 20,000 DOF Finite Element (FE) model*. This FE model had sixteen actuator force inputs, $f_r(t) \in R^{16x1}$, corresponding to sixteen electromagnetic actuators, three piezoelectric force inputs, $f_p(t) \in R^{3x1}$, corresponding to the three piezoelectric actuators[11], six capacitive sensor displacement outputs, $d_c(t) \in R^{6x1}$, and three collocated piezoelectric velocity outputs, $v_p(t) \in R^{3x1}$. This large order FE model was reduced down to a 30 DOF FE model using component mode synthesis, and modal truncation. The result was then transformed into the 60th order continuous time state space model shown in Figure 3.1a. The modal frequencies and damping for the reduced FE model is shown in table 4. Damping was assumed to vary linearly with frequency between 0.1% critical damping at 0.97kHz to 5.0% critical damping at 18.01kHz.

*This model was developed by Garth Reese of department 1434 Sandia.

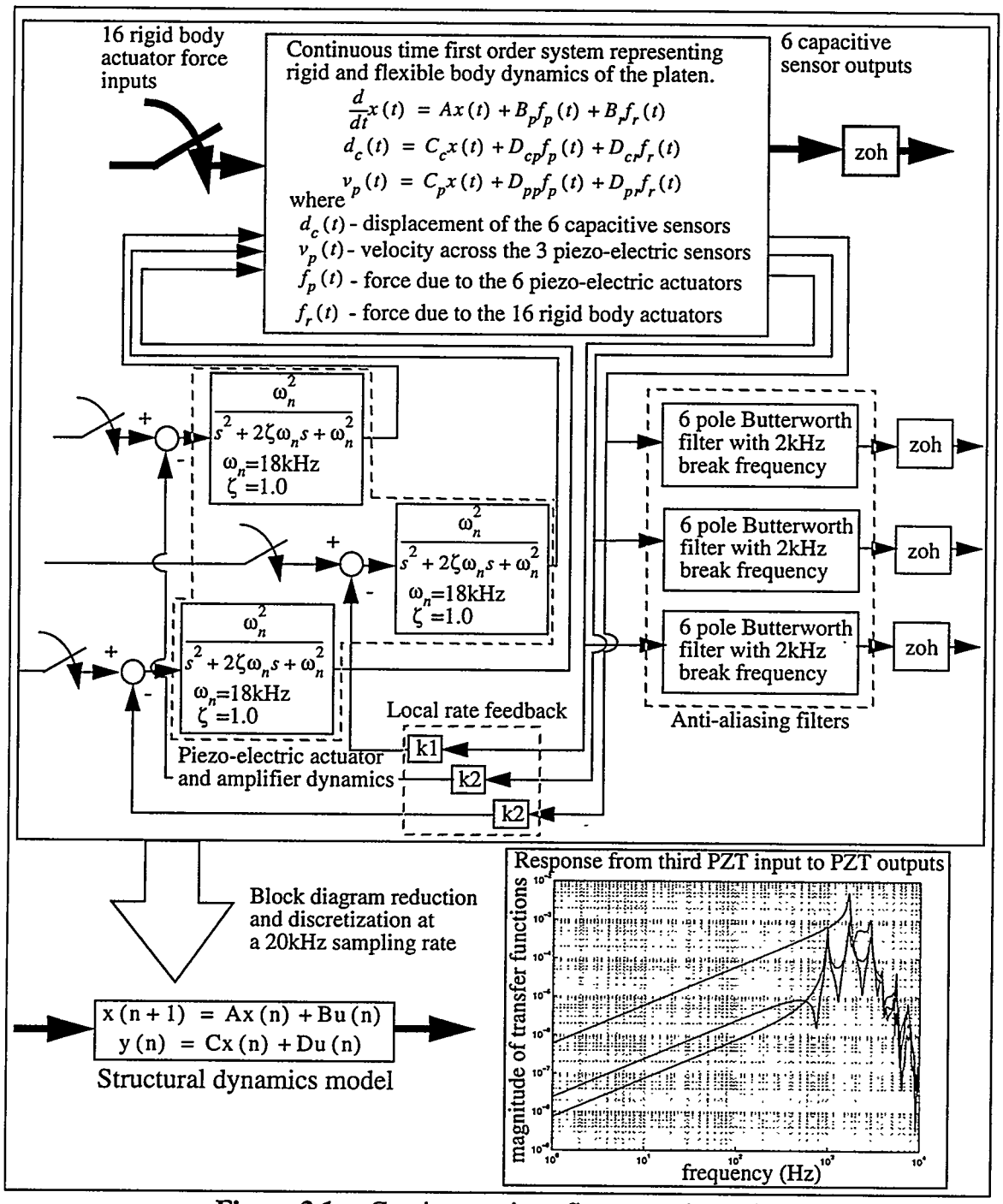


Figure 3.1a Continuous time, fine stage dynamics

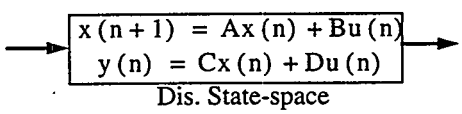
mode #	natural frequency (kHz)	percent damping (%)	mode #	natural frequency (kHz)	percent damping (%)
1-6	0.00	0.00	19	7.11	1.86
7	0.97	0.10	20	7.47	1.97
8	1.66	0.30	21	7.81	2.06
9	1.78	0.33	22	8.47	2.25
10	2.87	0.65	23	9.49	2.54
11	3.18	0.73	24	10.20	2.75
12	3.39	0.79	25	10.33	2.75
13	3.84	0.93	26	13.22	3.62
14	3.94	0.95	27	14.03	3.86
15	4.33	1.07	28	15.28	4.21
16	5.22	1.32	29	16.45	4.55
17	5.51	1.40	30	18.01	5.00
18	5.80	1.49			

Table 4: Natural frequencies and damping

The 60th order continuous time state space model was cascaded with a set of second order filters with natural frequencies at 18KHz, and damping ratios at 1.0. These filters represented the dynamics of the piezoelectric actuators and amplifiers. The 60th order model was also cascaded with a set of 6 pole Butterworth anti-aliasing filters with break frequencies at 2kHz. After a low authority rate feedback controller was applied to damp high frequency modes, the cascaded model was then transformed into discrete time using a sampling frequency of 20kHz. The reduced order, cascaded, discrete time state space model was 84th order.

A magnitude plot of the transfer functions from the third piezoelectric amplifier input to all piezoelectric velocity sensor outputs is shown in Figure 3.1a. Notice that the high frequency response is attenuated by the use of the anti-aliasing filters, and that there are many modes above the break frequency of the anti-aliasing filters. Moreover, from table 4, notice that there are 7 modes with natural frequencies above the Nyquist frequency.

Figure 3.1b shows the next highest level of Simulink modeling. The discrete time representation of the reduced order, augmented model shown in Figure 3.1a is contained in the block represented by



where the vector, u (n) $\in R^{19x1}$, is an input vector whose first sixteen elements contain the sixteen force inputs from the electromagnetic actuators and whose last three elements contain the three piezoelectric actuator amplifier inputs. The vector, y (n) $\in R^{9x1}$, is a discrete output vector whose first six elements are the six capacitive sensor displacement outputs and whose last three elements are the collocated piezoelectric velocity outputs. The parameter, n, is a discrete time increment, and A $\in R^{84x84}$, B $\in R^{84x19}$, C $\in R^{9x84}$, and D $\in R^{9x19}$ are state matrices.

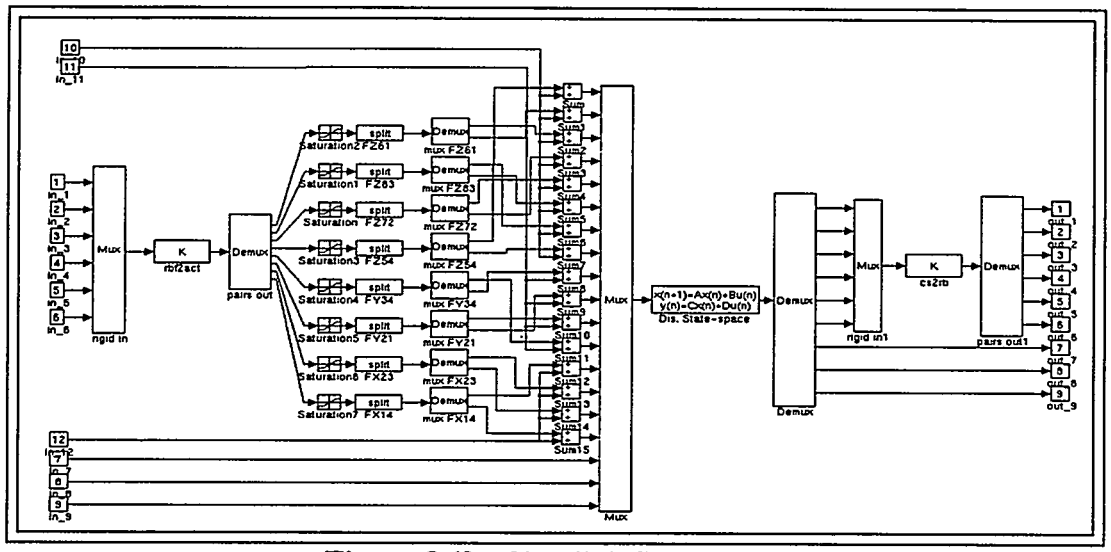
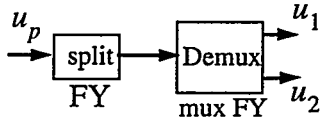


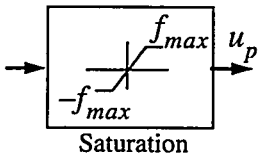
Figure 3.1b Simulink fine stage model

The sixteen electromagnetic actuators are used to move the fine stage in rigid body directions. These actuators were mono-directional and therefore were paired in order to produce bi-directional rigid body motion. This actuator pairing was modeled by the blocks



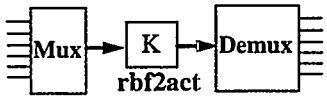
where u_p is the force produced by a pair of actuators, u_1 is the force produced by the first actuator in the pair, and u_2 is the force produced by the second actuator in the pair. If $u_p > 0$, $u_1 = u_p$ and $u_2 = 0$, otherwise $u_2 = u_p$ and $u_1 = 0$ where the second actuator pulls in the opposite direction of the first.

Saturation effects are also included in this model and are represented by the block

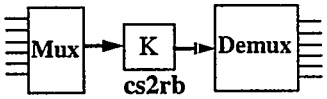


where f_{max} is the maximum force which an electromagnetic actuator can apply to the platen. For $10\mu m$ step excitations, it was assumed that $f_{max} = 15lb_f$*.

Saturation blocks are driven by the transformation blocks,



which relates rigid body forces and torques to forces applied to the eight actuator pairs. Moreover, capacitive measurements drive the transformation blocks,



which relates displacements at each capacitor to three rigid body translations and three rigid body rotations to form six rigid body outputs.

Figure 3.1b is called the fine stage model. The fine stage model was imbedded into another model called the lithography machine model. The lithography machine model is shown in Figure 3.1c where the masked block labeled "fine stage" is given by Figure 3.1b.

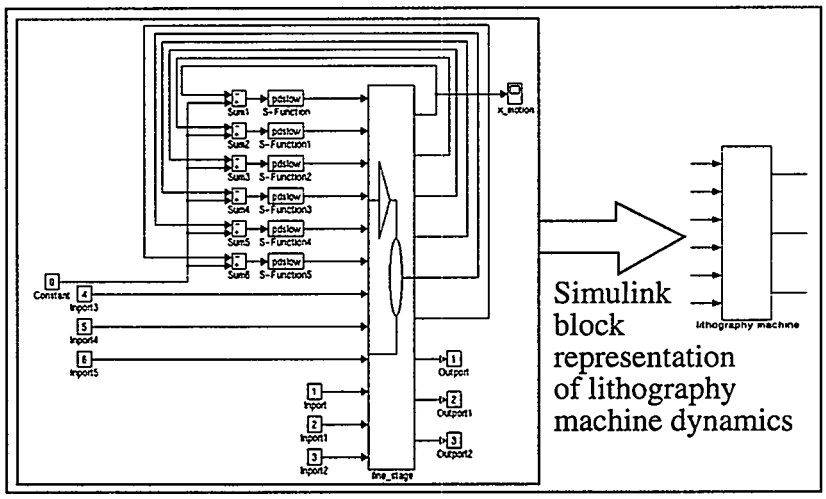


Figure 3.1c Simulink lithography machine model

In Figure 3.1c, the blocks labeled "PD slow" are discrete, Proportional and Derivative (PD) rigid body controllers. These controllers control the six rigid body degrees of freedom of the fine stage. These controllers run at a 200Hz sampling rate (a tenth of the sampling rate of the flexible body controller). If the fine stage had no flexible body dynamics, the rigid body controllers would change the dynamics of the system in each rigid body direction to that of a second order system with a 2Hz bandwidth and a 0.707 damping ratio. This low bandwidth was required to hold the magnetically levitated fine stage in place while system identification is being performed. Once flexible body dynamics have been identified and a

*A complete description of the rigid body controller is not provided in this document due to its confidentiality. Therefore, assumptions as to saturation levels are required.

flexible body controller designed and implemented, rigid body control bandwidth will be moved up to as large of a value as possible.

The Figure 3.1c lithography machine model was include into the Figure 3.1d system ID/ vibration control model. This figure represents the highest level of model of the Figure 1.1 system.

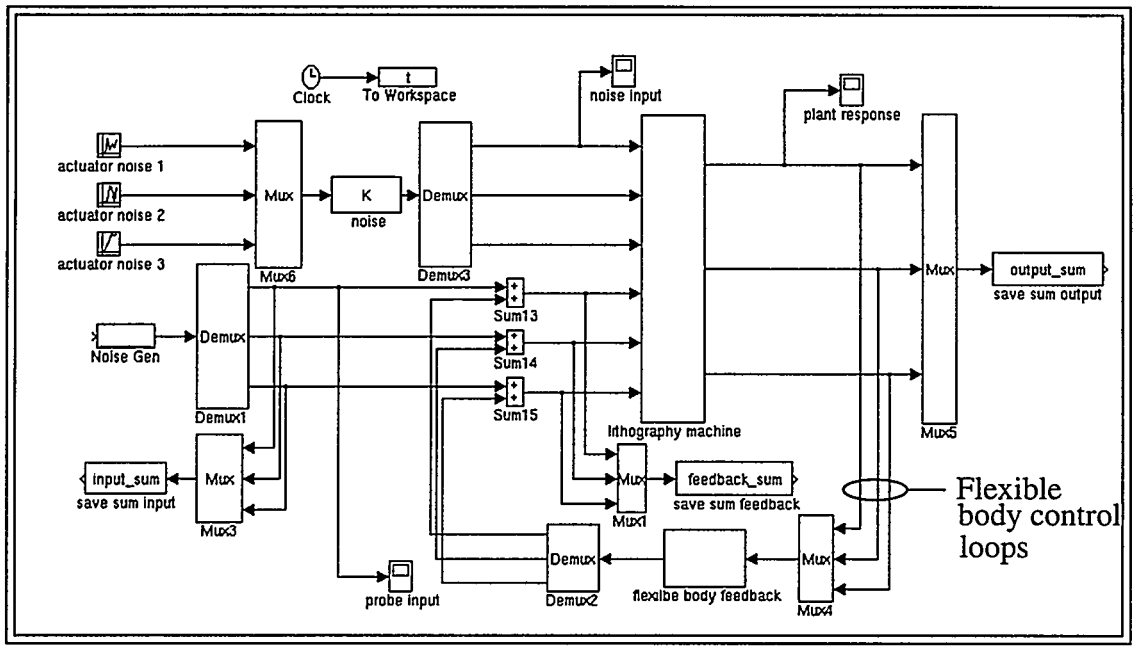


Figure 3.1d Simulink system ID/vibration control model

In this model, a set of control loops, called the flexible body control loops, relate piezoelectric actuator velocities to piezoelectric actuator amplifier inputs through the block labeled "flexible body feedback". In this document, two types of controllers can be used in this model. The first controller is a rate feedback controller, and the second controller is a Linear Quadratic Gaussian (LQG) controller. System identification is performed while one of these flexible body controllers are operating.

In rate feedback and LQG control, it is assumed that the controller takes the form,

$$x_c(i+1) = A_c x_c(i) + B_c y(i)$$

$$u(i) = C_c x_c(i) + D_c y(i) .$$

In this document, the rate feedback controller is derived from a Linear Quadratic Regulator (LQR) solution[12].

The LQR problem statement is;

given that $$x(i+1) = Ax(i) + Bu(i)$$

$$y(i) = Cx(i) + Du(i)$$

where $u(i)$ is a vector of force excitations and $y(i)$ is a vector of collocated velocity feedbacks,

find an F where $-Fx(i) = u(i)$ such that

$$\lim_{K\to\infty}\frac{1}{K}\sum_{i=0}^{K-1}\left(x(i+1)\,Qx^{T}(i+1)+u(i)\,Ru^{T}(i)\right)$$

is minimized for given Q and R.

To derive a rate feedback controller, the realization is placed into modal coordinates and Q and R are matrices chosen to weight modes of greatest importance. Then, from a single Riccati equation LQR solution F is determined. With F a least squares estimate of K is determined such that $K \sim FC$ where K is diagonal and positive semidefinite.The final solution is $A_c = 0$, $B_c = 0$, $C_c = 0$, and $D_c = -K$.

The LQG problem statement is;

given
$$x(i+1) = Ax(i) + Bu(i) + Gw(i)$$
$$y(i) = Cx(i) + Du(i) + Hv(i)$$

where $w(i)$ and $v(i)$ are white noise Gaussian processes of known statistics,

find a compensator (A_c, B_c, C_c, D_c) such that

$$E\left\{\lim_{K\to\infty}\frac{1}{K}\sum_{i=0}^{K-1}\left(x(i+1)\,Qx^{T}(i+1)+u(i)\,Ru^{T}(i)\right)\right\}$$

is minimized for given Q and R.

Again, the solution involves placing the realization into modal coordinates and choosing the Q and R matrices to weight modes of greatest importance. Then a two Riccati equation solution is used to solve for the optimal compensator. A more complete description of LQG is given reference 12.

The system ID/vibration control model shown in Figure 3.1d was augmented with a number of blocks for the purpose of performing system identification. Blocks labeled "Noise Gen", "save sum input", save sum output", and "save sum feedback" were added to produce periodic random probe signals excitations and to save averaged measured response data for the purpose of identifying system dynamics from piezoelectric amplifier inputs to piezoelectric velocity sensor outputs.

Noise was added to the system at the rigid body actuators. This noise not only excited the flexible body dynamics but also the rigid body dynamics. Therefore, a low bandwidth rigid body controller was needed to limit rigid body modal response while system identification was being performed. Signal to Noise Ratios (SNR's) at the PZT outputs were set to -6, 0, or 6 dB*.

*SNR=$20\log_{10}\frac{\text{root mean square level of signal}}{\text{root mean square level of noise}}$. These number are completely unrealistic for a lithography machine stage. Nevertheless, they used with an alternative applications in mind.

3.2 Accuracy of solution with variations in feedback and noise

Figure 3.2 contains a comparison of test data from the Figure 3.1 model for variations in noise and control. Each element in Figure 3.2 is calculated from the metric Γ. This metric is defined from the impulse responses of the exact solution and the impulse responses of the realization. If $y_{km}(i)$ is the impulse response of the exact solution between the k^{th} input and the m^{th} output, and $\hat{y}_{km}(i)$ is the impulse response of the realization between the k^{th} input and the m^{th} output, then

$$\Gamma = \frac{\sum_{k=1}^{n_{in}} \sum_{m=1}^{n_{out}} \sqrt{\sum_{i=1}^{N} e^2_{km}(i)}}{\sum_{k=1}^{n_{in}} \sum_{m=1}^{n_{out}} \sqrt{\sum_{i=1}^{N} y^2_{km}(i)}} \cdot 100$$

where $e_{km}(i) = y_{km}(i) - \hat{y}_{km}(i)$.

Increasing Rate Control Feedback Effort	Increasing External Noise Excitation (dB)	Initial Model Order Selection Scheme: min. SV	Initial Model Order Selection Scheme: Accum. Sum	Increasing LQG Control Feedback Effort	Increasing External Noise Excitation (dB)	Initial Model Order Selection Scheme: min. SV	Initial Model Order Selection Scheme: Accum. Sum
none	inf	0.72	0.80		inf	0.72	0.80
	6	10.61	6.54		6	10.61	6.54
	0	14.46	11.79		0	14.46	11.79
	-6	21.07	24.51		-6	21.07	24.51
low	inf	1.10	1.28		inf	0.99	1.33
	6	6.96	5.31		6	5.53	5.34
	0	8.83	13.65		0	11.00	10.57
	-6	16.54	19.25		-6	17.60	15.74
high	inf	8.05	8.23		inf	1.12	1.43
	6	12.06	10.88		6	5.37	6.09
	0	14.04	15.38		0	10.52	8.96
	-6	21.90	28.03		-6	14.69	13.72

Figure 3.2 Nominal model error for variations in control and noise

To construct Figure 3.2, the Figure 3.1 model was excited with periodic random excitation and output data was stored in 1024 point blocks which were averaged 325 times. An H_1 calculation was then used to produce transfer functions which were used to produce Markov parameters which were used to produce a nominal model realization. The singular values of the realization were truncated by using two methods. In the first method (min.

SV), only states associated with singular values whose value divided by the maximum singular value was greater than 0.0001 were retained in the model. In the second method (Accum. Sum), only those states associated with singular values which when summed with all larger singular values divided by the sum of all singular values produced a number less than 0.999, where retained in the model.

Variations in Figure 3.2 show that for higher noise levels, nominal model identification error increases, but for higher controller feedback, little variation in error occurs. Any of the values in Figure 3.2 could have been made smaller by producing more averages.

3.3 Stability in the face of model uncertainty

Figure 3.2 shows nominal model error for variations in noise and control. A controller can be designed from this nominal model and nominal stability can be checked. Nevertheless, to assure closed loop stability in the face of model uncertainty, the maximum singular value of the multiplicative uncertainty must be determined. If one over this singular value is greater than the complementary sensitivity function at every frequency, then the closed loop system will be stable. If it is not greater, the system can be either stable or unstable. Figure 3.3 shows the maximum singular value of the multiplicative uncertainty of a nominal model identified for the Figure 3.1 model with a SNR of -6dB, and a series of complementary sensitivity functions defined by various LQG controllers. The stability of each closed loop system is also shown. Notice that for the complementary sensitivity functions which are less than one over the uncertainty bound, the system is stable. This agrees with equation 1.3b..

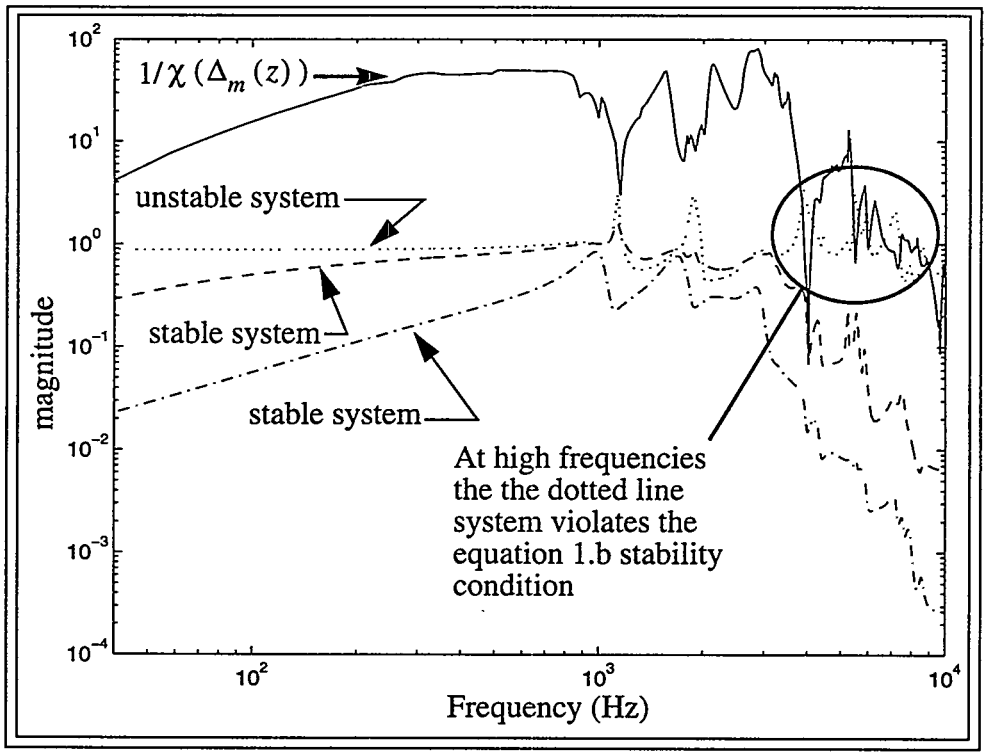


Figure 3.3: Stability relative to model error

4.0 Discussion of results

In this document, a discussion of some of the aspects of system identification as they relate to robust control design were presented. A few comments are given below.

In section 2.6, it was shown that ERA produces a realization which can be used to order the importance of states relative to an impulsive input. Nevertheless, the selection of model order relative to this ordering is subjective, since the state response to an impulsive load may be a poor indicator of the importance of that state to closed loop performance. The form of the realization which produced this ordering was called a balanced realization.

Also in section 2.6, it was stated that ERA/DC does not produce a balanced realization, yet model reduction relative to its singular values is still performed. Just how this model reduction relates to control performance is a research issue.

In section 3.2, it was shown that system identification could be performed in the presence of noise and feedback control. It should be noted again, that any of the elements in the Figure 3.2 matrix could have been lowered by further averaging.

In section 3.3, it was shown that for a system without control feedback, uncertainty bounds could be measured which can be used to assure stability in the face of model uncertainty. These uncertainty bounds determine frequencies in the closed loop system where energy can be added without causing instability, or where energy can be removed to impose stability. *In essence, this uncertainty bound shows the control designer how to shape the loop on the system so as to obtain the best closed loop stable response.* Determining these uncertainty bounds for system with outputs which have closely coupled dynamics, and systems with control feedback, is a matter of research.

This report represents the final aspects of work performed on the NCAICM phase I program on automated system identification.

References

[1] Sandia Memorandom,'Maglev fine stage control design requirements', J.L. Dohner, T. Smith, J. Jordon, Dec. 16, 1994

[2] K. Ogata, **Modern Control Engineering**, Prentice-Hall, Inc., Englewood Cliffs, N.J. chapter 13. This is a simple introduction to z-transforms.

[3] J.M. Maciejowski, **Multivariable Feedback Design,** Addison-Wesley Publishing Company, Reading Massachusetts, see chapter 3.

[4] A.B. Oppenheim, R.W. Schafer, **Digital Signal Processing**, Prentice-Hall inc. Englewood Cliffs, NJ, see section 5.5.

[5] R.J. Allemang, **Vibrations: Analytical and Experimental Modal Analysis**, course notes, UC-SDRL-CN-20-263-662 Department of Mech., Ind., and Nuclear Engr., Univerisity of Cincinnati, March 1994, pp. 5-6 to 5-27

[6] J.N. Juang, and R.S. Pappa, "An Eigensystem Realization Algorithm for Modal Parameter Identification and Model Reduction", *Journal of Guidance, Control and Dynamics*, Vol. 8, No. 5, Sept.-Oct. 1985, pp 620-627

[7] J.N. Juang, J.E. Cooper, J.R. Wright, "An Eigensystem Realization Algorithm Using Data Correlations (ERA/DC) for Modal Parameter Identification," *Journal of Control Theory and Advanced Technology*, Vol. 4, No. 1, March 1988, pp. 5-14.

[8] J.N. Juang, J.S. Lew, "Integration of System Identification and Robust Controller Designs for Flexible Structures in Space," Paper No. AIAA-90-3467-CP, *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Portland, Oregon, August 1990, pp. 1361-1375 (this is the reference that is easiest to follow)

[9] B.C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction", *IEEE Transaction on Automatic Control*, Vol. AC-26, No. 1, Feb. 1981, pp. 17-32 (this is the more popular reference)

[10] **SIMULINK User's Guide**, The MathWorks, Inc. Cochituate Place, 24 Prime Park Way, Natick, MA 01760, (508) 653-1415, 1992

[11] J.M. Redmond, G.M. Reese, G.G. Parker, "Locations and Orientations of PZT Actuators for Platen/Mirror Structure", Sandia Internal Memorandum. This report gives a description of piezo-actuator placement.

[12] H. Kwakernaak, R. Sivan, **Linear Optimal Control Systems**, Wiley-interscience, New York, 1972

Appendix A Nyquist Stability

The Nyquist stability theorem is used to determine the robustness, and stability of a single input, single output system in the presence of feedback control. Figure A.1 shows a closed loop system. The closed loop response of this system is given by $T(s) = \frac{G(s)H(s)}{1 + Gs)H(s)}$. The characteristic equation of this system is obtained from the denominator of the closed loop transfer function, $1 + G(s)H(s) = 0$. If the zeros of this equation are in the right hand plane, the homogeneous response of this system to any input will be unbounded, and therefore, the system will be unstable. Nevertheless, looking at the zeros of the characteristic equation gives limited information about closed loop robustness. For this purpose, the Nyquist stability theorem is used.

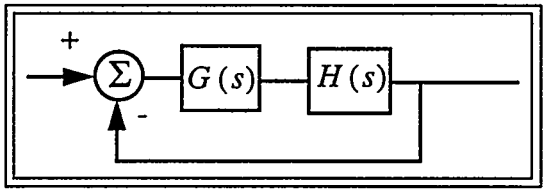


Figure A.1 A single Input, Single Output Feedback Control System

Before the Nyquist stability theorem can be derived, a simple relationship used in the mapping of complex functions must first be examined. This relationship states that the number of clockwise encirclements about zero of a closed locus of points in the s domain mapped into the $F(s)$ domain will be equal to the number of encircled poles of $F(s)$ minus the number of encircled zeros of $F(s)$. Figure A.2 illustrates this. The function $F(s)$ has two poles and one zero in the closed S domain contour. If this contour of S values is mapped into the $F(s)$ domain, then it will also produce a closed $F(s)$ domain contour which will encircle the zero $F(s)$ location only once in a clockwise fashion.

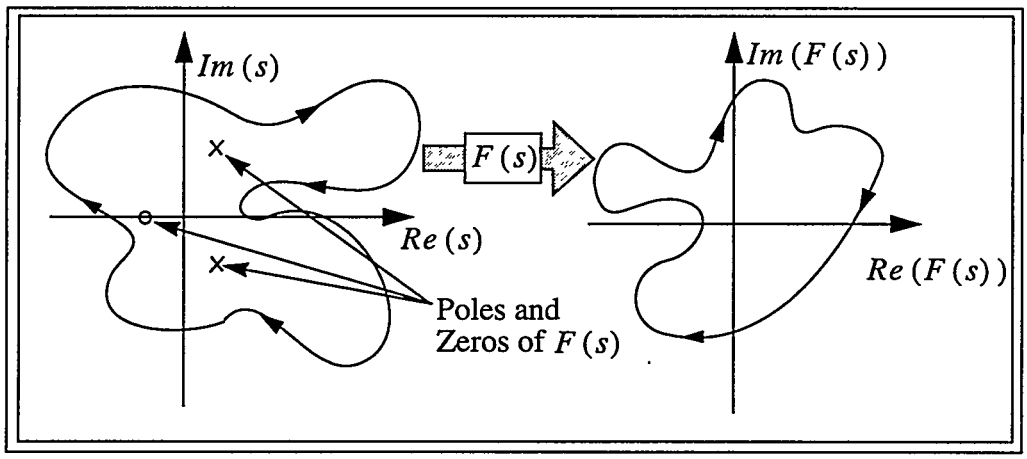


Figure A.2 Mapping from the s Domain to the $F(s)$ Domain

This can be proven by determining an expression for $\oint_S d\angle(F(s))$ where $\oint_S$ is an integral around a closed contour in the s domain. Since

$$\frac{d}{ds}\ln(F(s)) = \frac{1}{F(s)}\frac{d}{ds}F(s) = \frac{d}{ds}\ln|F(s)| + j\frac{d}{ds}\angle(F(s)) \text{ [10]},$$

$$\oint_S d\ln(F(s)) = \oint_S \frac{1}{F(s)}dF(s) = \oint_S d\ln|F(s)| + j\oint_S d\angle(F(s)),$$

and since $\oint_S d\ln|F(s)|$ must always be real, $\oint_S d\angle(F(s)) = Im\left(\oint_S \frac{1}{F(s)}dF(s)\right)$.

But, from Residue theory [11],

$$\oint_S \frac{1}{F(s)}dF(s) = \oint_S\left(\frac{1}{F(s)}\frac{d}{ds}F(s)\right)ds = -j2\pi\sum Res\left(\frac{1}{F(s)}\frac{d}{ds}F(s)\right),$$

$$\text{therefore, } \oint_S d\angle(F(s)) = -2\pi\sum Res\left(\frac{1}{F(s)}\frac{d}{ds}F(s)\right).$$

If $F(s)$ is factorizable,

$$\text{then } F(s) = \frac{\prod_{i=1}^{N_z}(s-z_i)^{K_i}}{\prod_{i=1}^{N_p}(s-p_i)^{M_i}} \text{ and } \frac{1}{F(s)}\frac{d}{ds}F(s) = \sum_{i=1}^{N_z}\frac{K_i}{(s-z_i)} - \sum_{i=1}^{N_p}\frac{M_i}{(s-p_i)}$$

where z_i is the i^{th} zero, p_i is the i^{th} pole, K_i is the multiplicity of the i^{th} zero, and M_i is the multiplicity of the i^{th} pole.

$$\text{From above } \sum Res\left(\frac{1}{F(s)}\frac{d}{ds}F(s)\right) = n_z - n_p$$

where n_p is the number of poles of $F(s)$ encircled by the closed contour in s and n_z is the number of zeros of $F(s)$ encircled by the closed contour in s. Thus, if n_c is the number of clockwise encirclements of zero in the $F(s)$ domain, then

$$-\oint_S d\angle(F(s)) = 2\pi n_c = 2\pi(n_z - n_p) \qquad \text{A(2)}$$

Equation A(2) states that the number of clockwise encirclements of zero in the $F(s)$ domain is equal to the number of zeros minus the number of poles of $F(s)$ encircled in the s domain.

In Nyquist, the number of zeros of $1 + G(s)H(s)$ in the right half s plane is determined from A(2). If this number is greater than zero, then the system is unstable. Thus, as shown in figure A3, the closed s domain contour should encircle the entire right half s plane.

Moreover, notice that the same results can be obtained by choosing $F(s)$ to equal $1 + G(s)H(s)$ and counting encirclements, n_c, about zero, as can be obtained by choosing $F(s)$ to equal $G(s)H(s)$ and counting encirclement, n_c, about -1. Therefore, for simplicity, $F(s)$ is chosen to equal $G(s)H(s)$, and the number of encirclements about -1 are counted. By using equation A(2), the number of zeros in the right half s plane is given by

$$n_z = n_p + n_c \qquad \text{A(3)}$$

where n_c is the number of encirclements about the -1 location, and since the poles of $1 + G(s)H(s)$ are the same as the poles of $G(s)H(s)$, n_p is simply the number of poles of $G(s)H(s)$ in the right half s plane. Equation A(3) is the Nyquist Stability theorem. More specifically, from Nyquist, if $G(s)H(s)$ is strictly proper and has no poles in the right half s plane, then, for the closed loop system to be stable, the Nyquist diagram of $G(s)H(s)$ should not encircle -1. Figure A3 shows this situation for both a stable and unstable system

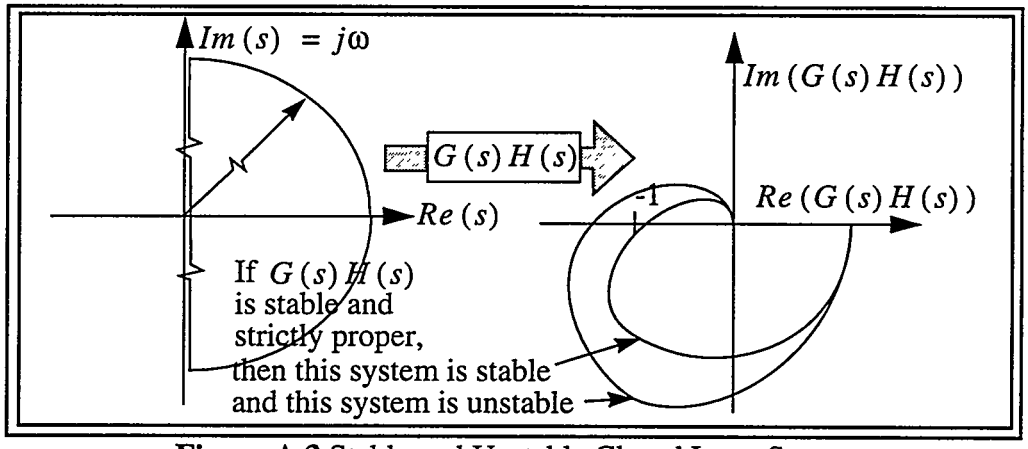


Figure A.3 Stable and Unstable Closed Loop Systems

As was discussed in section 1.2, Figure 1.5b shows a closed loop system in feedback with multiplicative uncertainty. Notice that the stability of this system can also be determined by Nyquist. In particular, if both Δ_m and $T(s)$ are both stable, and strictly proper, then the Nyquist diagram of $\Delta_m(j\omega)T(j\omega)$ must not encircle -1, or, as a conservative bound on stability, $|\Delta_m(j\omega)||T(j\omega)| < 1$. This leads directly to equation 1.3a

Distribution:

MS0321	1400	E.H.Barsis
MS0439	1434	J.L. Dohner (3)
MS0439	1434	D.R. Martinez
MS0439	1434	G.G. Parker
MS0439	1434	J.M. Redmond
MS0439	1434	D.J. Segalman
MS0841	1434	C.R. Dohrmann
MS0828	1502	P.J. Hommert
MS0441	1503	J.H. Biffle
MS0443	1517	H.S. Morgan
MS0439	1518	R.K. Thomas
MS0501	2338	J.D. Jordan
MS0501	2338	S.M. Kohler
MS0501	2338	T.G. Smith
MS0557	2741	T.J. Baca
MS0557	2741	J.J. Allen
MS0557	2741	R. Rodeman
MS0557	2741	P.S. Barney
MS0954	2903	E.G. DiBello
MS1033	6111	W.N. Sullivan
MS0860	9122	R.J. Kipp
MSJ576	LANL	J. Fales
MSC931	LANL	N. Hunter
MS9018	8523-2	Central Technical Files (1)
MS0899	13414	Technical Library(5)
MS0619	12615	Print Media
MS0100	7613-2	DOE/OSTI (2)

Dr. A. Jenab
Metron Inc.
11911 Freedom Drive
Suite 800
Reston VA 22090-5603

Dr. D. Armoza
Advanced Information Technology
Naval Research Lab-code 5585
Washington DC 20375-5337

Pat Draper
AT&T/ATS Suite 300
1919 South Eads St.
Arlington VA 22202

PRC/ARPA