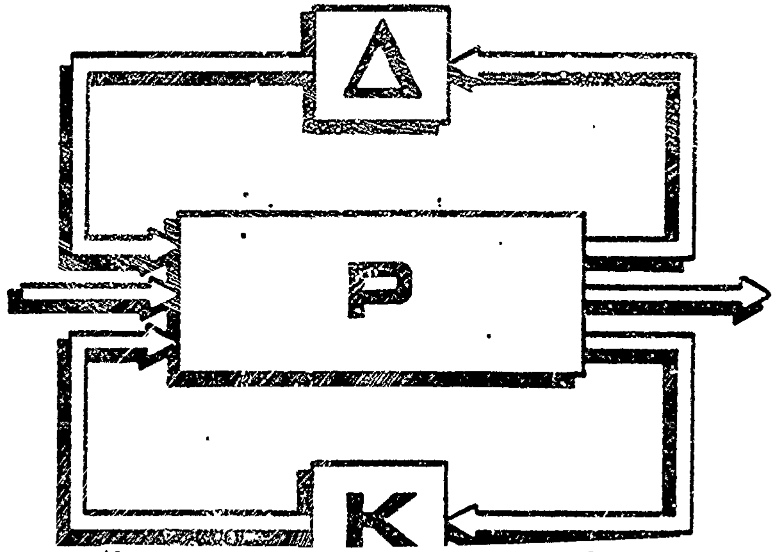
ONR / HONEYWELL WORKSHOP



ADVANCED TOPICS IN ROBUST CONTROL

N00014-82-C-0157
Vol. I

Advances in MULTIVARIABLE CONTROL

Lecture Notes
by
John Doyle

with contributions by
Cheng-Chih Chu
Bruce Francis
Pramod Khargonekar
Gunter Stein

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Honeywell

STATEMENT "A" per Dr. John Lavery
ONR/Code 1111MR, Title should read:
"Advanced Topics in Robust Control, Vol I'
Subtitle should read " Advances in Multi-variable Control" per same telecon.

27 September 1984

TELECON 8/10/90 VG

Dear Workshop Attendee:

Thank you for your interest in the ONR/Honeywell Multivariable Control Workshop. We have received your registration and are happy that you are able to attend.

Limousine service is available from the Minneapolis-St. Paul Airport to downtown. The Minneapolis/Downtown Limousine will drop at either the Holiday Inn or the Hyatt Regency. (We are sorry that the Concord Hotel, listed in your brochure, is no longer available.)

We expect the workshop to run from 8:30 a.m. to 5:00 p.m. each day. Lunch each day is provided as part of your registration fee.

A hospitality suite, hosted by Honeywell, will be held at the Hyatt Regency on Monday, October 8 from 5:30 - 7:30 p.m. to which you are all invited.

There will be a workshop banquet (also included in your registration fee) held at the Nicollet Island Inn on Tuesday evening, October 9. Featured speaker at the banquet will be Professor George Zames of McGill University. A sign-up sheet at the conference registration table will have menu selections. Bus service will be provided from the Hyatt Regency to the Nicollet Island Inn.

If you have any questions please feel free to contact Kelly Deedrick (612) 378-5716.

Sincerely,

John Doyle

John Doyle/Joe Wall

Workshop Coordinators
Honeywell Inc.
2600 Ridgway Parkway, Box 312
MN17-2370
Minneapolis, MN 55440

Best
Available
Copy

ONR/HONEYWELL WORKSHOP

AGENDA

Monday

8:00	Registration
8:30	Wall - Opening Remarks Stein - Problem Description and Motivation Break Doyle - Structured Singular Value Stein - Historical Perspective Discussion
12:00	Lunch
1:30	Stein - Overview of Synthesis Solutions Doyle - Constant Matrix Case Break Francis - Stabilization Francis & Doyle - Preview of Tuesday's Material Discussion
5:30	Hospitality Suite - Nicollet D1

Tuesday

8:30	Doyle - Factorizations Break Francis - H_∞ Approximations Discussion
12:00	Lunch
1:30	Doyle - Synthesis for the Structured Singular Value Economou - Design Example Break Doyle - Summary Glover
6:30	Banquet - Nicollet Island Inn

Wednesday

8:30	Helton Safonov Khargonekar
12:00	Lunch
1:30	Kimura Young Boyd Discusson

ONR7

Credits

Artwork by
John Doyle, Bruce Francis, Kathryn Lenz, Blaise Morton

Pagination by
Bruce Francis

Credits by
Bruce Francis

Typing by
Cheng-Chih Chu, John Doyle, Karen Pierce

PART ONE:

DOYLE'S NOTES

PREFACE

These notes are intended to supplement the tutorial lectures for the workshop by filling in a few details left out in the lectures. They are not complete, but are simply what was available at the time. We expect to have a draft of a complete set of notes for the material presented in the tutorial by the end of the year.

Many people contributed to these notes, although most are probably reluctant to admit it. They began as a very sketchy set of research notes that I developed, and have been somewhat rewritten by myself and the other authors. In particular, Cheng-Chih Chu wrote most of Section 2.3.3 on the Solution of the Algebraic Riccati Equation and helped throughout. Bruce Francis wrote most of Chapters 2.2 and 2.4 on Stabilization and Approximation. Pramod Khargonekar also helped throughout, but particularly with the method of proof used in Section 2.3.4 on Inner-Outer Factorization. Gunter Stein, of course, was the prime motivator of the whole tutorial. Whatever seems enlightening and clever about the presentation is probably due to the other authors. Any obfuscation is probably mine alone.

There are obviously lots of deficiencies in these notes. There are essentially no references, because we're trying to take credit for most of the control theory research of the last twenty years. Actually, we just didn't get to them. Ditto on most of the introductions to the chapters. The introductions were supposed to motivate the technical details in the subsequent sections, but hopefully the tutorial lectures will do that. There are probably lots of typos, and whole sections that are in the table of contents haven't been written.

John Doyle
Oct. 8/1984

Part 0. Preliminaries

Part 1. Analysis

Part 2. Synthesis Theory

0.0 Notation

SYMBOL	USAGE
$\cdot$	$\dot{x}(t) := \frac{d}{dt}x(t)$
$*$	1. $(x*y)(t) := \int_{-\infty}^{\infty} x(t-\tau)y(\tau)d\tau$
	2. A^* = complex-conjugate transpose of complex matrix A
	3. Γ^* = adjoint of operator Γ
1_+, δ	$1_+(t)$ = unit step function; $\delta(t)$ = unit impulse
s	$G(s)$ = two-sided Laplace transform of $g(t)$
$\perp$	$X^\perp$ = orthogonal complement of X
$\bar{\sigma}$	$\bar{\sigma}(A)$ = largest singular value of matrix A
ρ	$\rho(A)$ = spectral radius of matrix A

0.1.1 Function Spaces

Continuous time domain

$L_2(\mathbb{R},\mathbb{C}^{m\times n})$: Hilbert space of matrix-valued functions on $\mathbb{R}$, with inner product

$$<f,g> := \int_{-\infty}^{\infty} \operatorname{trace}\left[f(t)^* g(t)\right] dt.$$

$H_2(\mathbb{R},\mathbb{C}^{m\times n})$: subspace of functions zero for $t<0$.

$H_2(\mathbb{R},\mathbb{C}^{m\times n})^\perp$: subspace of functions zero for $t>0$.

P_{H_2} and $P_{H_2^\perp}$: . the orthogonal projections from $L_2(\mathbb{R},\mathbb{C}^{m\times n})$ onto $H_2(\mathbb{R},\mathbb{C}^{m\times n})$, $H_2(\mathbb{R},\mathbb{C}^{m\times n})^\perp$ respectively.

Continuous frequency domain

$j\mathbb{R}$: imaginary axis.

$L_2(j\mathbb{R},\mathbb{C}^{m\times n})$: Hilbert space of matrix-valued functions on $j\mathbb{R}$, with inner product

$$<F,G> := \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}\left[F(j\omega)^* G(j\omega)\right] d\omega.$$

$H_2(j\mathbb{R},\mathbb{C}^{m\times n})$: subspace of functions $F(s)$ analytic in Re $s>0$ and satisfying

$$\sup_{\sigma>0}\int_{-\infty}^{\infty} \operatorname{trace}\left[F(\sigma+j\omega)^* F(\sigma+j\omega)\right] d\omega < \infty$$

$L_\infty(j\mathbb{R},\mathbb{C}^{m\times n})$: Banach space of (essentially) bounded matrix-valued functions, with norm

$$\|F\|_\infty := \operatorname*{ess\,sup}_\omega \bar{\sigma}\left[F(j\omega)\right].$$

$H_\infty(j\mathbb{R},\mathbb{C}^{m\times n})$: subspace of functions $F(s)$ analytic and bounded in Re $s>0$.

$P_{H_2}, P_{H_2^\perp}$: the orthogonal projections from $L_2(j\mathbb{R},\mathbb{C}^{m\times n})$ onto $H_2(j\mathbb{R},\mathbb{C}^{m\times n})$, $H_2(j\mathbb{R},\mathbb{C}^{m\times n})^\perp$ respectively.

Prefix R denotes real-rational and the prefix B denotes the unit ball. The symbol $R_p^{(m\times n)}$ denotes proper real-rational matrices. Sometimes the spaces are abbreviated as $L_2(\mathbb{R})$, etc. or as L_2, etc. when context determines the arguments.

The Fourier transform yields the following [isometric] isomorphisms:

$$L_2(\mathbb{R}, \mathbb{C}^{m\times n}) \cong L_2(j\mathbb{R}, \mathbb{C}^{m\times n})$$
$$H_2(\mathbb{R}, \mathbb{C}^{m\times n}) \cong H_2(j\mathbb{R}, \mathbb{C}^{m\times n})$$
$$H_2(\mathbb{R}, \mathbb{C}^{m\times n})^\perp \cong H_2(j\mathbb{R}, \mathbb{C}^{m\times n})^\perp$$

The norms on these spaces are all denoted by $\|\cdot\|_2$.

A useful fact is that the norm of a matrix G in $L_\infty(j\mathbb{R}, \mathbb{C}^{m\times n})$ equals the norm of the corresponding multiplication operator

$$f \to Gf : L_2(j\mathbb{R}, \mathbb{C}^n) \to L_2(j\mathbb{R}, \mathbb{C}^m);$$

that is,

$$\|G\|_\infty = \sup\left\{\|Gf\|_2 : f \in L_2(j\mathbb{R}, \mathbb{C}^n),\ \|f\|_2 \le 1\right\}.$$

It also equals the norm of the operator restricted to $H_2(j\mathbb{R}, \mathbb{C}^n)$:

$$\|G\|_\infty = \sup\left\{\|Gf\|_2 : f \in H_2(j\mathbb{R}, \mathbb{C}^n),\ \|f\|_2 \le 1\right\}.$$

0.2.1 Controllability and Observability

Consider the system

$$\dot{x} = Ax + Bu \ , \quad x(0) = 0. \tag{1}$$

The system or the pair (A,B) is *controllable* if, for each time $t_1>0$ and final state x_1, there exists a (continuous) input $u(\cdot)$ such that the solution of (1) satisfies $x(t_1) = x_1$.

Theorem 1

The following are equivalent:

(i) (A,B) is controllable.

(ii) The matrix $\left[B, AB, A^2B, \ldots\right]$ has independent rows.

(iii) The matrix $\left[A-\lambda I, B\right]$ has independent rows for all λ in $\mathbb{C}$

(iv) The eigenvalues of $A+BF$ can be freely assigned by suitable choice of F.

The matrix A is said to be *stable* if all its eigenvalues satisfy $\mathrm{Re}\lambda<0$. The system, or the pair (A,B), is *stabilizable* if there exists an F such that $A+BF$ is stable.

Theorem 2

The following are equivalent:

(i) (A,B) is stabilizable.

(ii) The matrix $\left[A-\lambda I, B\right]$ has independent rows for all $\mathrm{Re}\lambda\geq 0$.

We will now consider the dual notions of observability and detectability with the system

$$\dot{x} = Ax \ , \quad x(0) = x_0$$

$y = Cx$.

The system, or the pair (C,A), is *observable* if, for every $t_1>0$, the function $y(t)$, $t \in [0,t_1]$, uniquely determines the initial state x_0.

Theorem 1':

The following are equivalent:

(i) (C,A) is observable.

(ii) The matrix $\begin{bmatrix} C \\ CA \\ CA^2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$ has independent columns.

(iii) The matrix $\begin{bmatrix} A-\lambda I \\ C \end{bmatrix}$ has independent columns for all λ in $\mathbf{C}$.

(iv) The eigenvalues of $A+HC$ can be freely assigned by suitable choice of H.

(v) (A',C') is controllable.

The system, or the pair (C,A), is *detectable* if $A+HC$ is stable for some H.

Theorem 2':

The following are equivalent:

(i) (C,A) is detectable

(ii) The matrix $\begin{bmatrix} A-\lambda I \\ C \end{bmatrix}$ has independent columns for all Re $\lambda \geq 0$.

(iii) (A',C') is stabilizable.

0.2.2 Transfer Functions

Consider the linear, time-invariant, ordinary differential equation described by

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

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^m$ is the input, and $y(t) \in \mathbb{R}^p$ is the output. The A, B, C, and D are appropriately dimensioned real matrices.

Associated with (1) is the convolution equation

$$\begin{aligned} y(t) &= (g * u)(t) \\ g(t) &= Ce^{At}B\,1_+(t) + D\delta(t) \end{aligned} \tag{2}$$

and, upon taking Laplace transforms, the resulting transfer function is

$$\begin{aligned} y(s) &= G(s)u(s) \\ G(s) &= C(sI-A)^{-1}B + D \end{aligned} \tag{3}$$

To expedite calculations involving transfer functions the notation

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \triangleq C(sI-A)^{-1}B + D \tag{4}$$

will be adopted. Note that $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is a real block matrix, not a transfer function. The product of two transfer functions is, of course, the cascade of the two systems or just the multiplication of two rational matrices. The convention will be adopted that the product of a matrix and a transfer function is a transfer function defined as

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \triangleq \left[\begin{array}{c|c} X_{11}A + X_{12}C & X_{11}B + X_{12}D \\ \hline X_{21}A + X_{22}C & X_{21}B + X_{22}D \end{array}\right]. \tag{5}$$

A similar convention holds for right multiplication by a matrix.

Suppose $G(s)$ is a real-rational transfer matrix which is *proper*, i.e., analytic at $s=\infty$. Then there exists a state-space model (A,B,C,D) such that

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]. \tag{6}$$

The quadruple (A,B,C,D) is called a *realization* of G. A realization is *minimal* if A has minimal dimension. It is a fact that a realization is minimal if and only if (A,B) is controllable and (C,A) is observable.

A basic object of study will be the transfer function and it will be assumed to have a realization. The next section describes standard operations on linear systems in terms of transfer functions and their realizations.

0.2.3 Operations on Linear Systems

1. Cascade

$$G_1 = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] \qquad G_2 = \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right]$$

$$\begin{aligned} G_1 G_2 &= \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right] \\ &= \left[\begin{array}{cc|c} A_1 & B_1 C_2 & B_1 D_2 \\ 0 & A_2 & B_2 \\ \hline C_1 & D_1 C_2 & D_1 D_2 \end{array}\right] = \left[\begin{array}{cc|c} A_2 & 0 & B_2 \\ B_1 C_2 & A_1 & B_1 D_2 \\ \hline D_1 C_2 & C_1 & D_1 D_2 \end{array}\right] \end{aligned}$$

Note: This realization may not be minimal.

2. Change of Variables

$$x \to \hat{x} = Tx$$

$$y \to \hat{y} = Ry$$

$$u \to \hat{u} = Pu$$

$$\begin{aligned} \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \to \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] &= \begin{bmatrix} T & 0 \\ 0 & R \end{bmatrix} \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \begin{bmatrix} T^{-1} & 0 \\ 0 & P \end{bmatrix} \\ &= \left[\begin{array}{c|c} TAT^{-1} & TBP \\ \hline RCT^{-1} & RDP \end{array}\right] \end{aligned}$$

3. State Feedback

$$u \to \hat{u} + Fx$$

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \to \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]\begin{bmatrix} I & 0 \\ F & I \end{bmatrix} = \left[\begin{array}{c|c} A+BF & B \\ \hline C+DF & D \end{array}\right]$$

4. Output Injection

$$\dot{x} = Ax + Bu \quad \to \quad \dot{\hat{x}} = A\hat{x} + Bu + Hy$$

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \to \begin{bmatrix} I & H \\ 0 & I \end{bmatrix}\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{c|c} A+HC & B+HD \\ \hline C & D \end{array}\right]$$

5. Transpose (Dual)

$$G \to G^T$$

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

6. Conjugate

$$G \to G^*$$

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \to \left[\begin{array}{c|c} -A^T & -C^T \\ \hline B^T & D^T \end{array}\right]$$

7. Inversion

Suppose $D^\dagger$ is a right (left) inverse of D. Then $G^\dagger = \left[\begin{array}{c|c} A-BD^\dagger C & -BD^\dagger \\ \hline D^\dagger C & D^\dagger \end{array}\right]$ is a right (left) inverse of G.

Proof: The right inverse case will be proven and the left inverse case follows by duality. Suppose $DD^{\dagger} = I$. Then

$$GG^{\dagger} = \left[\begin{array}{cc|c} A & BD^{\dagger}C & BD^{\dagger} \\ 0 & A-BD^{\dagger}C & -BD^{\dagger} \\ \hline C & DD^{\dagger}C & DD^{\dagger} \end{array}\right]$$

$$= \left[\begin{array}{cc|c} A & BD^{\dagger}C & BD^{\dagger} \\ 0 & A-BD^{\dagger}C & -BD^{\dagger} \\ \hline C & C & I \end{array}\right]$$

Conjugating the state by $\begin{bmatrix} I & I \\ 0 & I \end{bmatrix}$ on the left and $\begin{bmatrix} I & I \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} I & -I \\ 0 & I \end{bmatrix}$ on the right yields

$$GG^{\dagger} = \left[\begin{array}{cc|c} A & 0 & 0 \\ 0 & A-BD^{\dagger}C & -BD^{\dagger} \\ \hline C & 0 & I \end{array}\right]$$

$$= I$$

Corollary 7 Suppose $D^{\dagger}$ is a right inverse for D and let

$$\hat{G} = \left[\begin{array}{c|c} A-BD^{\dagger}C & -BZ \\ \hline D^{\dagger}C & Z \end{array}\right].$$

Then

$$G\hat{G} = DZ.$$

Corollary 7' Suppose $D^{\dagger}$ is a left inverse for D and let

$$\hat{G} = \left[\begin{array}{c|c} A-BD^{\dagger}C & -BD^{\dagger} \\ \hline ZC & Z \end{array}\right].$$

Then

$\hat{G}G = ZD.$

The following lemma characterizes the relationship between zeros of a transfer function and poles of its inverse.

8. **Lemma** Suppose $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ with D nonsingular. Then there exists (s_o, x_o) such that

$$(A - BD^{-1}C)x_o = s_o x_o \ , \quad Cx_o \neq 0$$

iff there exists $u_o \neq 0$ such that

$$G(s_o)u_o = 0$$

Proof

(if)

$G(s_o)u_o = 0$ implies that $G^{-1}(s)$ has a pole at s_o. Thus $\exists(s_o, x_o)$ such that $Cx_o \neq 0$ and

$$(A - BD^{-1}C)x_o = s_o x_o$$

(only if)

Set $u_o = -D^{-1}Cx_o \neq 0.$ Then $G(s_o)u_o = C(s_o I - A)^{-1}Bu_o + Du_o = Cx_o - Cx_o = 0$

QED

0.2.4 Linear Matrix Equations :

Property 1 : (Solution of Sylvester Equations)

Consider the Sylvester equation

$$AX + XB = C \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times m}$, $C \in \mathbb{R}^{n\times m}$ are given matrices.

Then, there exists a unique solution $X \in \mathbb{R}^{n\times m}$ if and only if $\mathrm{Re}[\lambda_i(A) + \lambda_j(B)] \neq 0, \quad \forall\, i = 1,\ldots\ldots,n$ and $j = 1,\ldots\ldots,m$.

Remark :

In particular, if $B = A^T$, (1) is called the "Lyapunov Equation" and the necessary and sufficient condition for the existence of unique solution will be that $\mathrm{Re}[\lambda_i(A) + \lambda_j(A)] \neq 0, \ \forall\, i,j = 1,\ldots\ldots,n$.

Property 2 : (Solution of Linear Equations)

Consider the linear equation

$$A\,X = B$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ are given matrices.

The following statements are equivalent :

(i) there exists a solution $X \in \mathbb{R}^{n\times m}$.

(ii) the columns of $B \in \mathit{Range}(A)$.

(iii) $\mathit{rank}\begin{bmatrix} A & B \end{bmatrix} = \mathit{rank}\begin{bmatrix} A \end{bmatrix}$.

(iv) $\mathrm{Ker}\,(A^T) \subset \mathrm{Ker}\,(B^T)$.

0.3.1 Gramians and Balanced Realizations

Suppose $G = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ where A is stable. Define the *controllability gramian* Y as

$$Y \triangleq \int_0^\infty e^{At} BB' e^{A't}\, dt$$

and the *observability gramian* as

$$X \triangleq \int_0^\infty e^{A't} C'C e^{At}\, dt.$$

By considering the corresponding matrix differential equations it is easily shown that Y and X satisfy the Lyapunov equations

$$AY + YA' + BB' = 0$$

$$A'X + XA + C'C = 0$$

Note that $Y \geq 0$ and $X \geq 0$. Furthermore, the pair (A,B) is controllable iff $Y > 0$ and (C,A) is observable iff $X > 0$.

Suppose the state is transformed by nonsingular T to $\hat{x} = Tx$ to yield the realization

$$G = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] \quad \text{with} \quad \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] = \left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array}\right].$$

Then the gramians transform as $\hat{Y} = TYT'$ and $\hat{X} = (T^{-1})'XT^{-1}$. Note that $\hat{Y}\hat{X} = TYXT^{-1}$ so the eigenvalues of the product of the gramians are invariant under state transformation.

Consider the similarity transformation T which gives the eigenvector decomposition

$$YX = T\Lambda T^{-1}\ , \qquad \Lambda = \mathrm{diag}(l_1,\ldots l_n)$$

Then columns of T are (possibly nonunique) eigenvectors of YX corresponding to the eigenvalues $\{\lambda_i\}$. It is shown in Lemma 1 at the end of this section that YX has real diagonal Jordan form and that $\Lambda \geq 0$. This is a consequence of $Y \geq 0$ and $X \geq 0$.

Although the eigenvectors are not unique, in the case of a minimal realization they can always be chosen such that

$$\hat{Y} = TYT' = \Sigma,$$
$$\hat{X} = (T^{-1})'XT^{-1} = \Sigma,$$

where $\Sigma = \mathrm{diag}(\sigma_1,\sigma_2,\ldots,\sigma_n)$ and $\Sigma^2 = \Lambda$. This new realization will be referred to as a balanced realization (also called internally balanced).

Suppose $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ is a balanced realization for G and can be partitioned as

$$\left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right]$$

with corresponding partitioning of the balanced gramian $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$. Suppose $\Sigma_1 = \mathrm{diag}(\sigma_1,\sigma_2,\ldots\sigma_r)$, $\Sigma_2 = \mathrm{diag}(\sigma_{r+1},\sigma_{r+2},\ldots,\sigma_n)$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} \geq \sigma_{r+2} \geq \cdots \geq \sigma_n$. Then it is immediate that the truncated system

$$G_r = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$$

is balanced since

$$A_{11}\Sigma_1 + \Sigma_1 A_{11}' + B_1 B_1' = 0$$

$$A_{11}'\Sigma_1 + \Sigma_1 A_{11} + C_1' C_1 = 0.$$

It can also be shown (Silverman and Pernebo) that a *minimal* realization for G_+ is stable, although in certain (non-generic) cases A_{11} may have uncontrollable or unobservable $j\omega$-axis eigenvalues.

Lemma 1 Product of Positive Semi-Definite Matrices is Similar to a Positive Semi-Definite Matrix

Proof: Let X and Y be positive semi-definite. First perform an orthogonal transformation so that

$$X \to \begin{bmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{bmatrix} \quad \Lambda_1 > 0 \text{ diagonal}, \quad Y \to \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{12}' & Y_{22} \end{bmatrix} \geq 0$$

By this transformation XY is similar to $\begin{bmatrix} \Lambda_1 Y_{11} & \Lambda_1 Y_{12} \\ 0 & 0 \end{bmatrix}$. Now

$$\begin{bmatrix} \Lambda_1 Y_{11} & \Lambda_1 Y_{12} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \Lambda_1^{1/2} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \Lambda_1^{1/2} Y_{11} \Lambda_1^{1/2} & \Lambda_1^{1/2} Y_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Lambda_1^{-1/2} & 0 \\ 0 & I \end{bmatrix},$$

and it is easy to find a matrix Z such that

$$\begin{bmatrix} \Lambda_1^{1/2} Y_{11} \Lambda_1^{1/2} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I & Z \\ 0 & I \end{bmatrix} \begin{bmatrix} \Lambda_1^{1/2} Y_{11} \Lambda_1^{1/2} & \Lambda_1^{1/2} Y_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I & -Z \\ 0 & I \end{bmatrix}$$

(Z exists because the columns of $\Lambda_1^{1/2} Y_{11}$ span the columns of $\Lambda_1^{1/2} Y_{12}$ owing to the fact that Y is positive semi-definite). The left hand side of this last equation is positive semi-definite and similar to XY. If $XY \neq 0$ it is possible to find a matrix T_1 such that

$$T_1^{-1}ABT_1 = \Lambda = \mathrm{diag}\left\{(\lambda_1, \ldots, \lambda_K, 0, \ldots, 0)\right\}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K > 0$.

Q.E.D.

Now consider two gramians X and Y. Let us suppose $XY \neq 0$, so that T_1 can be chosen as above:

$$T_1^{-1}XYT_1 = \Lambda = \mathrm{diag}\left\{(\lambda_1, \ldots, \lambda_K, 0, \ldots, 0)\right\}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K > 0$. Under this transformation the gramians become

$$Q = T_1^{-1}X(T_1^{-1})', \quad R = T_1'YT_1,$$

and $QR = \Lambda$. Because Q and R are symmetric, $RQ = \Lambda' = \Lambda = QR$ and so Q, R and Λ commute. Both Q and R must leave the eigenspaces of Λ invariant and so are of the form

$$Q = \mathrm{diag}\left\{Q_1, \ldots, Q_l, E\right\}$$

$$R = \mathrm{diag}\left\{\lambda_{j_1}Q_1^{-1}, \ldots, \lambda_{j_l}Q_l^{-1}, F\right\} .$$

where Q_i is a square matrix whose size equals the dimension of the λ_{j_i} eigenspace of Λ and $EF = 0$ where E and F are square matrices the size of the kernel of Λ. Of course, all the Q_i's are summetric so it is possible to find an orthogonal matrix

$$W = \mathrm{diag}\left\{W_1, \ldots, W_l, W_{l+1}\right\}$$

such that $W^{-1}Q(W^{-1})'$ and $W_{l+1}'FW_{l+1}$ are diagonal. Note that this same W gives a diagonal $W'RW$ and leaves Λ alone.

The transformation $T_2=T_1W$ diagonalizes both gramians. It is now obvious how to construct T_5 so that $T_5^{-1}X(T_5^{-1})'$ and $T_5'YT_5$ are diagonal and the controllable and observable portions are equal.

0.3.2 Inner Transfer Functions

Let $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$. Then G is inner if $G^*G=I$ and co-inner if $GG^*=I$. Note that G need not be square. Inner and co-inner are dual notions and are often called all-pass.

If $G \in R_p^{p\times m}$, $p>m$ is inner then any $G_\perp \in R_p^{(p-m)\times m}$ is called a complementary inner factor (CIF) if $\left[G \;\; G_\perp\right]$ is square and inner. The dual notion of complementary co-inner factor is defined in the obvious way.

The following lemma is useful in characterizing inner transfer functions in terms of a realization.

Lemma 1. Suppose $\exists\; X=X' \in \mathbb{R}^{n\times n}$ such that

i) $A'X + XA + C'C = 0$

ii) $B'X + D'C = 0$

Then $G^*G = D'D$.

Proof: Suppose that i) and ii) hold. Then conjugating the state of

$$G^*G = \left[\begin{array}{cc|c} A & 0 & B \\ -C'C & -A' & -C'D \\ \hline D'C & B' & D'D \end{array}\right]$$

by $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}$ on the left and $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$ on the right yields

$$G^*G = \left[\begin{array}{cc|c} A & 0 & B \\ -(A'X+XA+C'C) & -A' & -(XB+C'D) \\ \hline B'X+D'C & B' & D'D \end{array}\right]. \tag{1}$$

Now, applying i) and ii) yields

$$G^*G = \left[\begin{array}{cc|c} A & 0 & B \\ 0 & -A' & 0 \\ \hline 0 & B' & D'D \end{array}\right]$$

$$= D'D$$

By duality, we have the following

Lemma 1' Suppose $\exists Y=Y' \in \mathbb{R}^{n\times n}$ such that

i) $AY + YA^T + BB' = 0$

ii) $CY + DB' = 0$

Then $GG^*=DD'$.

These two lemmas immediately lead to one characterization of inner matrices in terms of their state space representation. Simply add the condition that $D'D=I$ ($DD'=I$) to lemma 1 (1') to get $G^*G=I$ ($GG^*=I$). Furthermore, by adding a few additional assumptions, the conditions in the lemmas become necessary as well as sufficient. This leads to the following complete characterization of stable inner transfer functions in terms of a minimal realization.

Suppose $G=\left[\begin{array}{cc} A & B \\ C & D \end{array}\right]$ is stable and minimal. Then the gramians X and Y satisfying

$$A'X + XA + C'C = 0 \tag{2}$$

$$AY + YA' + BB' = 0 \tag{3}$$

exist and are unique.

Corollary 1 G is inner iff

i) $B'X + D'C = 0$

ii) $D'D = I$

Corollary 1' G is co-inner iff

i) $CY + DB' = 0$

ii) $DD' = I$

Proof Sufficiency of i) and ii) follows immediately from the lemmas. For necessity, suppose $G^*G=I$. From 1) and 2) this implies that

$$\left[\begin{array}{c|c} A & B \\ \hline B'X+D'C & 0 \end{array}\right] = 0. \tag{4}$$

$$D'D = I \tag{5}$$

Since (A,B) is controllable, (4) implies that $B'X+D'C=0$. The co-inner case follows by duality.

This characterization of inner transfer functions plays a central role in the synthesis theory. It allows the construction of inner transfer functions by solving algebraic equations.

0.4.1 Linear Fractional Transformations

Suppose $P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \in R_p^{(p_1+p_2)\times(m_1+m_2)}$, $\Delta \in R_p$, $K \in R_p^{m_2 \times p_2}$. We will adopt the notation

$$\mathbf{F}_l(P,K) \triangleq P_{11} + P_{12}K(I-P_{22}K)^{-1}P_{21} \tag{1}$$

and

$$\mathbf{F}_u(P,\Delta) \triangleq P_{22} + P_{21}\Delta(I-P_{11}\Delta)^{-1}P_{12} \tag{2}$$

The linear fractional transformations (LFT) are illustrated in Figure 1. The l denotes that the second argument is fed back in the *lower* block, and the u denotes feedback in the upper block.

An important property of LFT's is that any interconnection of LFT's is again an LFT. Suppose $J = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}$. Then

$$\mathbf{F}_l(P,\mathbf{F}_l(J,Q)) = \mathbf{F}_l(T,Q) \tag{3}$$

$$\mathbf{F}_u(J,\mathbf{F}_u(P,\Delta)) = \mathbf{F}_u(T,\Delta) \tag{4}$$

where

$$\begin{aligned} T &= \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \\ &= \begin{bmatrix} P_{11}+P_{12}J_{11}(I-P_{22}J_{11})^{-1}P_{21} & P_{12}(I-J_{11}P_{22})^{-1}J_{12} \\ J_{21}(I-P_{22}J_{11})^{-1}P_{21} & J_{22}+J_{21}P_{22}(I-J_{11}P_{22})^{-1}J_{12} \end{bmatrix} \end{aligned} \tag{5}$$

Equations (3) and (4) are illustrated in Figure 2. Note that if $\mathbf{F}_l(J,Q)$ is a parametrization of a controller, $\mathbf{F}_l(T,Q)$ is affine if and only if $T_{22}=0$. This type of controller parametrization will play an important role in the synthesis theory.

Analysis

1. Introduction

This part of the notes describes recent results on the problem of analyzing the performance and robustness properties of systems. We believe that the approach described is providing the foundation for a new paradigm for control theory broader in scope and content than that of Classical or Modern Control Theory. An important aspect of this new paradigm is the treatment it gives to model uncertainty.

Modern Control Theory, the dominant paradigm for the past 20 years, has its basis in Stochastic Optimal Control and Estimation Theory. This theory essentially restricts model uncertainty to additive noise. The theory provides a methodology for analyzing the impact of noise on system performance and synthesizing to reduce that impact.

The inadequacies of this view of uncertainty became widely accepted in the late 1970's, as robustness to plant uncertainty became a major theme in the Modern Control Theory community. Ironically, this involved a renewed interest in the Classical Control paradigm which Modern Control displaced within the theoretical community (if not among practicing engineers). This new direction provided useful design tools, including Singular Value Analysis and Multivariable Loop Shaping.

While providing an important perspective, as well as practical techniques, the methods based on singular values still require rather restrictive assumptions about uncertainty. In particular, plant uncertainty must essentially be modelled as a single "unstructured perturbation."

The Structured Singular Value (SSV), μ, was developed several years ago to correct this deficiency in singular values. In the context of the general framework discussed in this memo, the SSV provides a very powerful mathematical tool for the analysis of complex systems. Indeed, we believe that this framework together with the SSV and the synthesis techniques discussed later, has the potential to form the basis for a new paradigm for control theory.

The remainder of this part of the notes describes the general framework for control system analysis and synthesis which includes all the viewpoints discussed as special cases. In particular, the assumptions about uncertainty required by each methodology are compared. In this context, the words analysis and synthesis have specific meanings.

Analysis is used to describe the process of determining whether a given system has the desired characteristics. In general, this may range from the use of mathematical tools to simulation to experimentation, although analysis is typically applied primarily to describe the former. **Synthesis**, on the other hand, is the process of finding a particular system component to achieve desired characteristics, which are typically expressed in terms of some analysis tools. **Analysis** and **synthesis** are just two aspects of the more general problem of engineering **design**.

The discussion which follows first considers analysis, then briefly touches on synthesis and ends with some illustrative examples. The next part on Synthesis Theory will take up that question in more detail.

1.2.1 General Framework

Various modelling assumptions will be considered and the impact of these assumptions on analysis and synthesis methods will be explored. Consider the diagram in Figure 1. This is the general framework to be considered. Models of this form are typically constructed from components which also have this form. The nominal model provides the basic interconnection structure between the signals, perturbations and controller, as shown. It has three inputs and outputs, each consisting of a vector of signals.

As typical examples, consider the following filtering and control problems. First, a simple filtering problem is given in the diagram in Figure 2. This may be rearranged as shown in Figure 3 to fit the general framework. In order to simplify the diagram, no perturbation was included.

A typical control problem might look like the diagram in Figure 4 where again, for simplicity, no perturbations are included. This too can be rearranged to fit the general framework, although the diagram is complicated.

Any system may be rearranged to fit the form of this general framework. Although the interconnection structure can become quite complicated for complex systems, many software packages are available which could be used to generate the interconnection structure from system components.

Note that uncertainty may be modelled in two ways, either as external inputs or as perturbations to the nominal nodel. The performance of a system is measured in terms of the behavior of the outputs or errors. The assumptions which characterize the uncertainty, performance and nominal model determine the analysis techniques which must be used.

The most fundamental assumption that is made throughout is that the nominal model is a finite dimensional ordinary differential equation and is linear and time invariant (LTIODE). The uncertain inputs are assumed to be either filtered white noise or weighted L_p signals. Performance is measured as weighted output variance or weighted output L_p norm. The perturbations are assumed to be themselves LTIODE's which are norm-bounded as input-output operators. Various combinations of these assumptions form the basis for all the standard linear systems analysis tools.

Given that the nominal model is an LTIODE, the interconnection system has the form

$$
\begin{aligned}
P &= \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} \\
&= \left[\begin{array}{c|ccc} A & B_1 & B_2 & B_3 \\ \hline C_1 & D_{11} & D_{12} & D_{13} \\ C_2 & D_{21} & D_{22} & D_{23} \\ C_3 & D_{31} & D_{32} & D_{33} \end{array}\right]
\end{aligned}
\tag{1}
$$

and the total system is a linear fractional transformation on the perturbation and the controller given by

$$
\begin{aligned}
e &= \mathrm{F}_u(\mathrm{F}_l(P,K),\Delta)\, u \\
&= \mathrm{F}_l(\mathrm{F}_u(P,\Delta),K)\, u
\end{aligned}
\tag{2}
$$

Since the focus of the current discussion is on analysis methods, the controller may be viewed as just another system component and absorbed into the interconnection structure. Thus the analysis framework reduces to the diagram in Figure 5 where

$$
\begin{aligned}
e &= \mathrm{F}_u(P,\Delta)\, u \\
&= \left[P_{22} + P_{21}\Delta(I-P_{11}\Delta)^{-1}P_{12}\right] u
\end{aligned}
\tag{3}
$$

Note that the P's in (2) and (3) are not necessarily the same. Table 1 and the discussion which follows summarize the various assumptions and resulting analysis and synthesis tools. In each case, stability of the nominal must be evaluated. Since P is assumed to have the state-space representation

$$P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] , \tag{4}$$

this may be done by checking that all eigenvalues of A lie in the open lhp. There are alternatives to this approach but, for simplicity, it will be assumed that the nominal plant, with controller is closed loop stable in the sense that all eigenvalues of A are in the open lhp.

Given nominal stability, the entries in the table may be interpreted as filling in the following general performance/robustness theorem:

General Analysis Theorem (GAT)

Given

Input Assumptions

and

Perturbation Assumptions

Then

Performance Specification

if and only if

Analysis Test

The details of each case will be considered in the following sections.

1.2.1

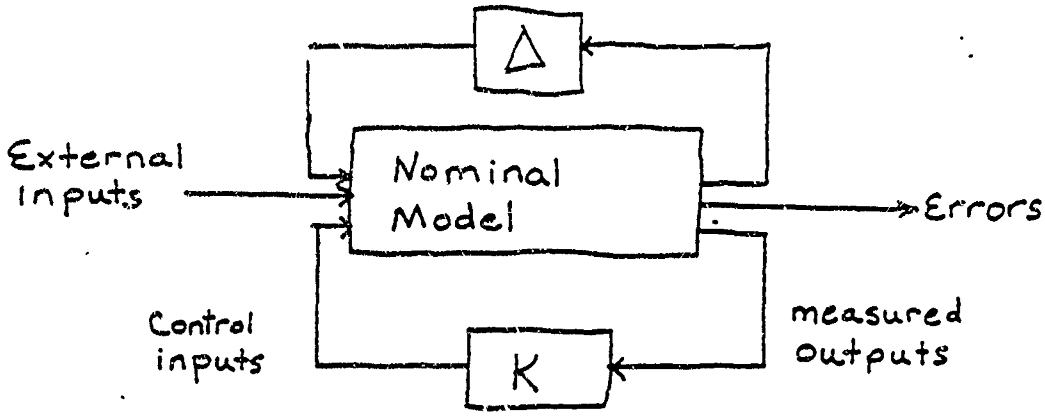


Figure 1

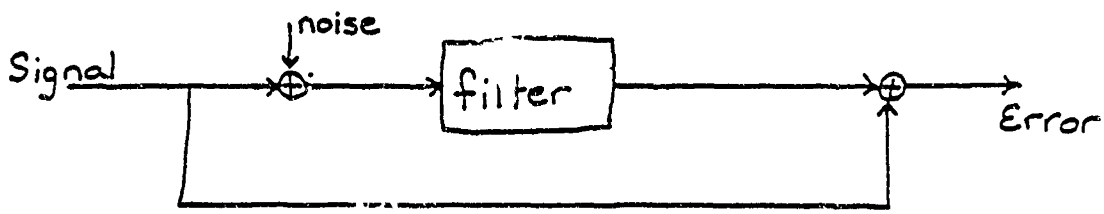


Figure 2

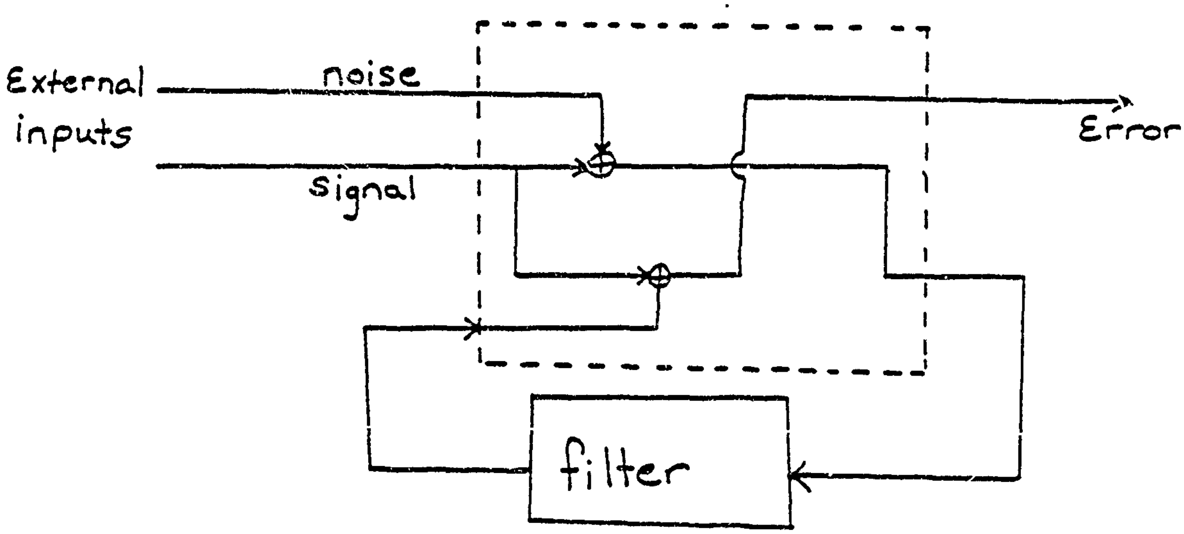


Figure 3

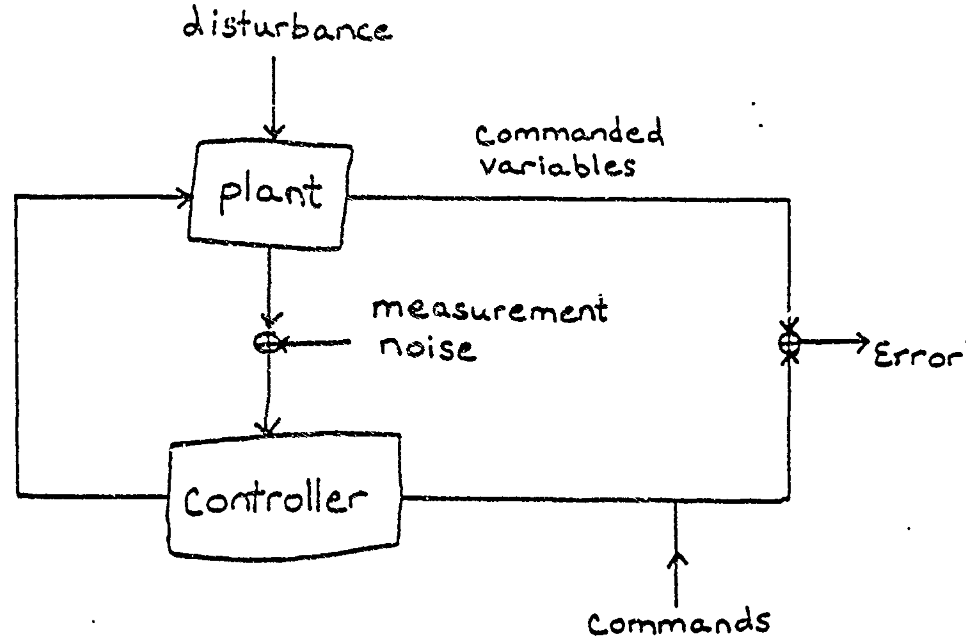
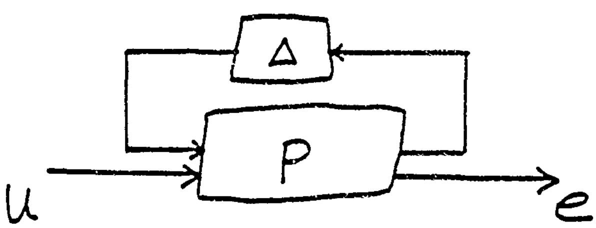


figure 4



$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

figure 5

1.2.2. Stochastic

Case 1a involves unit covariance white noise input with output variance as the evaluation criteria. Since no perturbation is allowed, the problem reduces to the diagram in Figure 1 and $E(e^T e) = \left\|P_{22}\right\|_2^2$. Note that colored noise or weighted variance could be used as shown in Figure 2. This reduces to the general case by absorbing the weights W_1 and W_2 into P_{22} as $P_{22} = W_2 G W_1$. In practice, it is essential to use weights to reflect spatial and frequency variations in inputs, perturbations and output specifications, but in every case, these weights may be absorbed into nominal model.

In Case 1b the input is an uncertain delta function, which is equivalent to uncertain initial conditions. The performance specification is the expected value of the L_2-norm of the output.

Case 1 forms the foundation of Stochastic Optimal Control Theory. Case 1a includes the standard linear stochastic filtering problem and Case 1b includes the standard linear quadratic optimal control problem. These are combined to obtain the full LQG problem, which is again Case 1a. These assumptions and resulting analysis methods have been the dominant paradigm in the control community for over 20 years.

The development of this paradigm has stimulated extensive research efforts and been responsible for important technological innovation, particularly in the area of estimation. The theoretical contributions include a deeper understanding of linear systems and improved computational methods for complex systems through state-space techniques. The major limitation of this theory is the lack of formal treatment of uncertainty in the plant itself. By allowing only additive noise for uncertainty, the stochastic theory ignored this important practical issue. Plant uncertainty is particularly critical in feedback systems.

1.2.2.

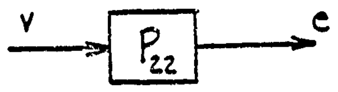


Figure 1

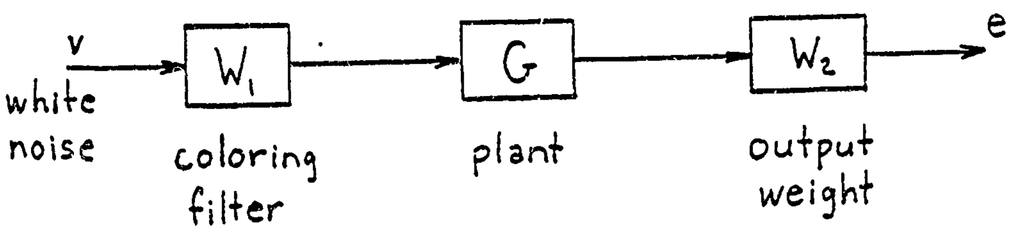


Figure 2

1.2.3 L_∞ Frequency Domain Methods

Case 2 involves an attempt to correct some of the deficiencies of Case 1 by moving to an unknown but bounded (in an L_2 sense) framework. This allows both types of plant uncertainty to be handled in a common framework, albeit in a limited manner.

Case 2a is an L_2 version of Case 1a. The input is constrained to lie in BL_2 as a time signal (unit ball in L_2) and the performance is specified in terms of the output's L_2 norm. With no perturbation, the analysis test involves simply the L_2 induced operator norm, i.e. L_∞ on the transfer function P_{22}.

The *GAT* in this case is

Theorem 1 $\quad \|e\|_2 \le 1$ for all $\|u\|_2 \le 1$

iff $\quad \|P_{22}\|_\infty \le 1$

Although this theorem is a trivial restatement of the definition of induced norm, it means that the analysis test is an exact characterization of the performance requirement.

Case 2b is significant departure from the previous three. It involves maintenance of stability in the presence of perturbations. The block diagram for $F_u'(P,\Delta)$ is shown in Figure 1. There are many ways to state the *GAT* for this case, depending on the desired notion of stability and assumptions on Δ. The distinctions are somewhat subtle, but are important from a theoretical point of view. Nevertheless, they do not significantly impact the application of the theory.

The Δ's are assumed to be LTIODE's, so that $\Delta \in RH_\infty$. The assumptions $\Delta \in \mathbb{C}$ or $\Delta \in \mathbb{C}H_\infty$ give the same result. The distributed case, $\Delta \in H_\infty$, causes some additional technical difficulties and is not the focus of these notes.

The following theorem treats internal stability

Theorem 2 $\mathbf{F}_u{}'(P,\Delta)$ is internally stable for all $\Delta \in BRH_\infty$

iff $\left\|P_{11}\right\|_\infty \le 1$

Note that input-output stability of $\mathbf{F}_u(P,\Delta)$ is not necessarily the same as internal stability. In particular, the following statement is not true:

Not–A–Theorem $\left\|e\right\|_2 < \infty$ for all $\left\|u\right\|_2 \le 1$ and $\Delta \in BRH_\infty$

iff $\left\|P_{11}\right\|_\infty \le 1$

Counterexample Suppose $\left\|P_{11}\right\|_\infty > 1$ but $P_{21} \equiv 0$.

From now on stability will mean *internal* stability, but be denoted by $\left\|e\right\|_2 < \infty$ in the table, even though this is definitely an abuse of notation. Note that generically this distinction between internal and i-o stability does not exist.

As in Case 1, it is essential to allow weights on inputs, outputs and perturbations. As before, these weights may be absorbed into the nominal model. This allows, without the loss of generality, the use of signals and perturbations which are in unweighted unit balls. Thus implementation of the analysis tools requires only a method for constructing interconnected systems and a method for evaluating the appropriate norm. The former applies to all cases, whereas the latter requres a different norm in each case.

Note that in Case 2 both uncertain inputs and uncertain plants can be handled with the same analysis tool. This approach is particularly useful for feedback problems where both types of uncertainty have significant impacts on system performance. Case 2 has attracted a great deal of research interest recently, and is currently the popular new paradigm in the multivariable control community. Although implicit in the methods of classical control and some more modern work (e.g. Zames' conic sector theory circa early 60's and Horowitz's 60's work), the approach did not gain wide attention until the late '70's.

The current interpretation is a consequence of research done in the late '70's. (Doyle and Stein, Safonov etc). This interpretation involves singular values as an analysis method and singular value loop shaping as a synthesis approach. The so-called LQG Loop Transfer Recovery (LQG/LTR) (Stein and Doyle) combines the synthesis methods of Case 1 with the analysis methods of Case 2 to produce a hybrid synthesis method. This gives an ad hoc approach to Case 2 that can be effective for many multivariable problems.

Another approach to synthesis for Case 2 is the so called H_∞ or L_∞/H_∞ methods introduced to the control community by Zames and Helton (with additional contributions by Francis, Pearson, Glover etc.). The L_∞/H_∞ methods for Case 2 are analogous to the L_2/H_2 methods of Case 1 with the exception that for Case 2 the L_∞ rather than L_2 norm is optimized. The solution to the general L_∞/H_∞ problem will be presented in the Synthesis part of these notes.

The main objection to Case 2 is the restrictive assumptions about uncertainty (recall this was also the objection to Case 1). Although case 2 allows both uncertain inputs and perturbations, analysis can be performed for *either* individually *but not both* together. Thus a system can be shown to remain stable when perturbed and have acceptable response to uncertain inputs when $\Delta = 0$ but response when $\Delta \neq 0$ is not known. Only crude bounds can be obtained with the methods of Case 2.

An additional limitation of Case 2b is that all plant uncertainty must be modelled as a single norm-bounded perturbation. Typically, uncertainty is present throughout a system. Suppose that a system is built from components which are themselves uncertain and that component uncertainty is modelled as norm-bounded perturbations. This situaton can be rearranged to fit the general framework but the perturbation for the total system has structure. The problem of structured uncertainty is taken up in the next chapter.

1.2.3

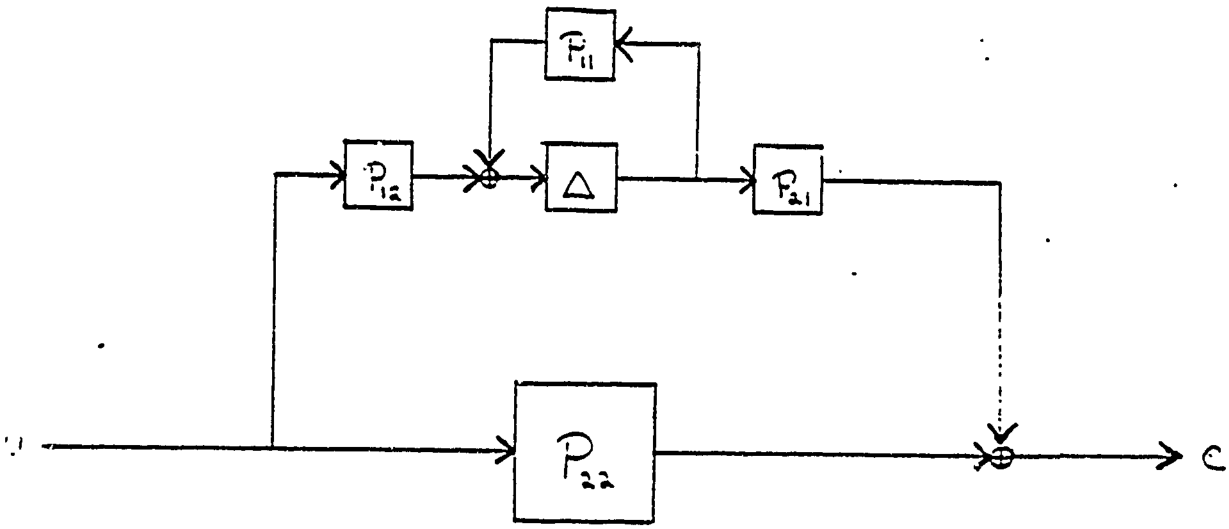


Figure 1

1.3.1 Introduction

This chapter considers the problem of stability with structured uncertainty and of performance in the presence of structured uncertainty. Typically, uncertainty is present throughout a system. Suppose that a system is built from components which are themselves uncertain and that component uncertainty is modelled as norm-bounded perturbations. This situaton can be rearranged to fit the general framework but the perturbation for the total system has structure. This can be seen schematically in Figure 1.

Note that the interconnection model P can always be chosen so that Δ is block diagonal, and by absorbing any weights, $\|\Delta\|_\infty < 1$. The results of Case 2b can be applied in two ways:

1) $\|P_{11}\|_\infty \leq 1$ implies stability, but *not* conversely. This can be arbitrarily conservative, in that stable systems can have arbitrarily large $\|P_{11}\|_\infty$.

2) Test for each Δ_i individually. This can be arbitrarily optimistic because it ignores interaction between the Δ_i.

The difference between the bounds obtained in 1) and 2) can be arbitrarily far apart. Only when they are close can conclusions be made about the general case with structured uncertainty.

These two limitations of Case 2 (and 1) have motivated much of the research described in these notes. The result is a new paradigm described in Case 3. The problem in Case 3 involves exactly that of structured uncertainty.

Consider the system in Figure 2. Stability and performance analysis of this system requires a new matrix function, the structured singular value (SSV), denoted by μ. Before proceeding with Case 3, a digression to discuss μ

will be taken. For details, see the reprints which accompany this writeup.

1.3.1

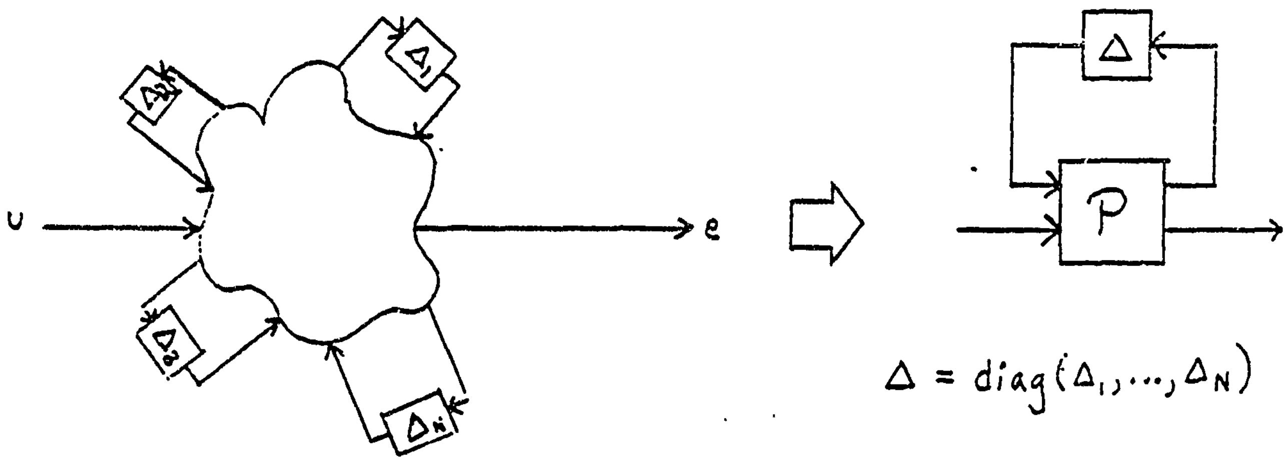
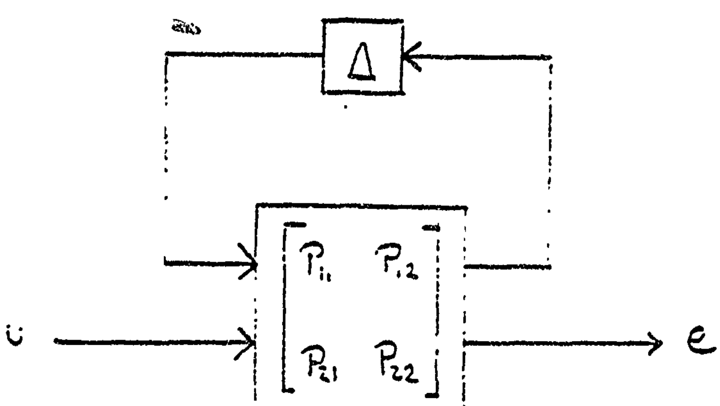


Figure 1



$\Delta = \mathrm{diag}(\Delta_1, \dots, \Delta_N)$

Figure 2

1.3.2 SSV for Constant Matrices

The problem is to test for $det(I-M\Delta)\neq 0$ for *sets* of Δ. Two standard results are

1) $det(I-M\Delta)\neq 0 \quad \forall\ \Delta \in \left\{\Delta \,\middle|\, \bar{\sigma}(\Delta)<1\right\}$

iff $\quad \bar{\sigma}(M) \leq 1$

2) $det(I-M\Delta)\neq 0 \quad \forall\ \Delta \in \left\{\lambda I \,\middle|\, \lambda \in \mathbb{C},\ |\lambda|<1\right\}$

iff $\quad \rho(M) \leq 1 \quad$ where $\rho(M) = \max_i |\lambda_i(M)|$

As a generalization, consider a function μ with the properties that $\mu(\alpha M)=|\alpha|\mu(M)$ and

3) $det(I-M\Delta)\neq 0 \quad \forall\ \Delta \in \left\{diag(\Delta_1,\Delta_2,\ldots,\Delta_n) \,\middle|\, \bar{\sigma}(\Delta_i)\leq 1\right\}$

iff $\quad \mu(M) < 1$

Obviously, μ is a function of M which depends on the *structure* of $\left\{ \Delta \right\}$. To be precise, a multi-index could be constructed which would specify the structure of $\left\{ \Delta \right\}$ and μ would depend on that index. For this informal discussion, just keep in mind this fact and assume that a structure is specified. Clearly $\bar{\sigma}$ and ρ are special cases of μ for particular structures as indicated above. Furthermore, for any structure

$$\rho(M) \leq \mu(M) \leq \bar{\sigma}(M). \tag{4}$$

Given these bounds, how important is μ? The answer can be clearly seen from the following examples:

Suppose $\Delta = \begin{bmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{bmatrix}$ and consider

1) $M = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \rho(M)=0 \quad \bar{\sigma}(M)=1$

$det(I-M\Delta)=1 \quad$ so $\quad \mu(M)=0$

2) $M = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad \rho(M)=0 \quad \bar{\sigma}(M)=1$

$det(I-M\Delta) = 1 + \dfrac{\delta_1-\delta_2}{2} \quad$ so $\quad \mu(M)=1$

Thus neither ρ nor $\bar{\sigma}$ provide useful bounds even in simple cases. The only time they do provide reliable bounds is when $\rho \approx \bar{\sigma}$. Thus better bounds on μ are needed to pursue the problem in Case 3.

For the rest of the discussion fix a structure for the Δ's as

$$\underline{X} = \left\{ diag\,(\Delta_1, \Delta_2, \ldots, \Delta_n) \right\}. \tag{5}$$

Then

$$\mu(M) = \left\{ \min_{\Delta \in \underline{X}} \left[\bar{\sigma}(\Delta) \,\middle|\, det\,(I-M\Delta)=0 \right] \right\}^{-1}. \tag{6}$$

This expression is little more than a definition of μ since the optimization problem implied by it is nonconvex, but it shows that μ exists as desired. To obtain useful properties of μ, some additional definitions are needed. Let

$$\underline{U} = \left\{ diag\,(U_1, U_2, \ldots, U_n) \,\middle|\, U_i^* U_i = I \right\} \tag{7}$$

$$\underline{D} = \left\{ diag\,(d_1 I, d_2 I, \ldots, d_n I) \,\middle|\, d_i \in \mathbb{R}- \right\} \tag{8}$$

where the sets $\underline{U}$ and $\underline{D}$ match the structure of $\underline{X}$. Note that the $\underline{U}$ and $\underline{D}$ leave $\underline{X}$ invariant in the sense that

1) $\Delta \in \underline{X}$, $U \in \underline{U}$ imply $\bar{\sigma}(\Delta U) = \bar{\sigma}(U\Delta) = \bar{\sigma}(\Delta)$

2) $\Delta \in \underline{X}$, $D \in \underline{D}$ imply $D\Delta D^{-1} = \Delta$

From these two properties and the definition above expression for μ, one immediately obtains

$$\max_{U \in \underline{U}} \rho(MU) \leq \mu(M) \leq \inf_{D \in \underline{D}} \bar{\sigma}(DMD^{-1})$$

The first important theorem about μ is

Theorem 1 $\max_{U \in \underline{U}} \rho(MU) = \mu(M)$

This theorem expresses μ in terms of familiar linear algebraic objects. Unfortunately, the implied optimization problem is nonconvex so it does not immediately yield a computational approach. The second important theorem is

Theorem 2 If $n \leq 3$ $\mu(M) = \inf_{D \in \underline{D}} \bar{\sigma}(DMD^{-1})$

This theorem states that if there are 3 or fewer blocks (no restriction on size), then $\mu(M)$ is just $\bar{\sigma}$ of a block diagonal similarity of M. Furthermore $\bar{\sigma}(DMD^{-1})$ is convex in D so that the infimum can be found by search over $n-1$ real parameters.

The theorem is not true for $n \geq 4$, but it is conjectured that $\inf_{D \in \underline{D}} \bar{\sigma}(DMD^{-1})$ still provides a reasonably tight bound for μ. Also, many problems of interest have 3 or fewer blocks so this provides a reasonable computational scheme.

Another important aspect of this theorem is that μ may be viewed as $\bar{\sigma}$ plus scaling. Thus the general synthesis methods recently developed to optimize the L_∞ norm (i.e. $\bar{\sigma}$) may be applied, via scalings, to optimize μ. This will be discussed more in the synthesis section. Now back to Case 3.

1.3.3 SSV Analysis of Systems

Abuse notation and define

$$\left\|M\right\|_{\mu} = \sup_{\omega} \mu(M(j\omega)). \tag{1}$$

Although $\left\|\cdot\right\|_{\mu}$ is not a norm, this will be convenient. Recall tate $\left\|M\right\|_{\mu}$ is a function of M which also depends on the assumed structure of the perturbations.

Case 3a involves stability in the presence of *structured* perturbations and the result is analogous with Case 2b. In fact, 3a reduces to 2b in the case that there is a single block in the perturbation. Suppose that $\Delta \in BRH_{\infty}$ and the Δ's have the structure $\Delta = diag(\Delta_1, \Delta_2, \Delta_3, \ldots, \Delta_n)$. The *GAT* for Case 3a is

Theorem 1 $F_u(P, \Delta)$ is internally stable for all *structured* $\Delta \in BRH_{\infty}$ **iff** $\left\|P_{11}\right\|_{\mu} \leq 1$

Case 3b puts everything together and is really the payoff for μ analysis. The problem is to check that $\left\|e\right\|_2 \leq 1$ is satisfied for all $\left\|u\right\|_2 \leq 1$ and all structured perturbations. Recall that from 2a and 2b that both stability with a single perturbation and performance with L_2 inputs involve the same test using $\left\|\cdot\right\|_{\infty}$ although on different parts of the system. This means that the system in Figure 1 has internal stability and $\left\|e\right\|_2 \leq 1$ for all $\left\|u\right\|_2 \leq 1$ and $\Delta = diag(\Delta_1, \Delta_2, \ldots, \Delta_n) \in BRH_{\infty}$ if and only if the system in Figure 2 has internal stability for all structured Δ and all $\Delta_{n+1} \in BRH_{\infty}$. This is exactly Case 3a with the structure $\tilde{\Delta} = diag(\Delta_1, \Delta_2, \ldots, \Delta_n, \Delta_{n+1})$. Using this structure for μ yields the following:

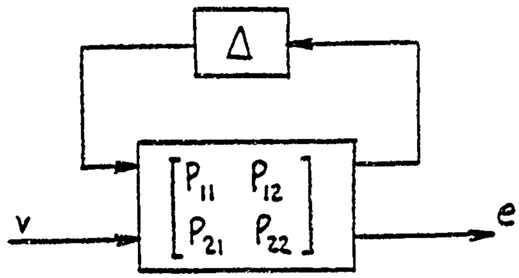
Theorem 2

$F_u(P,\Delta)$ is internally stable and $\|e\|_2 \le 1$ for all $\|u\|_2 \le 1$ and $\Delta = diag(\Delta_1, \Delta_2, \ldots, \Delta_n) \in BRH_\infty$

iff $\|P\|_\mu \le 1$

This is a remarkably useful theorem. It says that $\|P\|_\mu \le 1$ implies not only stability for all structured perturbations but also that $\|e\|_2 \le 1$ for all $\|u\|_2 \le 1$ and all structured perturbations. Furthermore, $\|P\|_\mu > 1$ implies that there exists a u with $\|u\|_2 \le 1$ and a structured Δ such that either $\|e\|_2 > 1$ or $F_u(P,\Delta)$ is internally unstable. This is the first general result which guarantees performance for a whole *set* of plants and gives an exact (nonconservative) analysis test.

1.3.3



$$\Delta = \mathrm{diag}(\Delta_1, \ldots, \Delta_n)$$

Figure 1

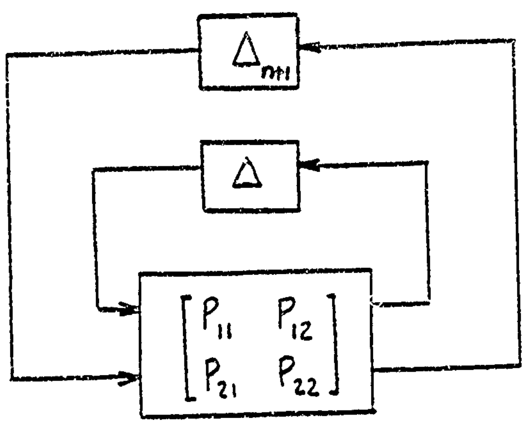


Figure 2

1.4.1 A Glimpse at Synthesis

This will be a sketchy outline of the new synthesis results. The details are somewhat complicated and are treated in Part 2 which is devoted to the synthesis theory. At this point, we simply want to point out how the analysis theory discussed in this part leads naturally to certain synthesis questions.

From the analysis results, we see that each case boils down to evaluating

$$\left\|P_{ij}\right\|_{\alpha} \qquad \alpha=2,\infty \text{ or } \mu \tag{1}$$

for some transfer function P_{ij}. Thus when the controller is put back into the problem, it involves just a simple linear fraction transformation as shown in the diagram in Figure 1. (Note: the P_{ij}'s here are not the same as the P_{ij}'s in the previous sections)

Each case then leads to the synthesis problem

$$\min_{K}\left\|\mathbf{F}_{l}(P,K)\right\|_{\alpha} \quad \text{for } \alpha=2,\infty, \text{ or } \mu \tag{2}$$

subject to internal stability of the nominal. Here

$$\mathbf{F}(P,K)=P_{11}+P_{12}K(I-P_{22}K)^{-1}P_{21}.$$

The solution of this problem for $\alpha=2$ and ∞ is the focus of Part 2 on Synthesis Theory. The solution presented there unifies the two approaches in a common synthesis framework. The $\alpha = 2$ case was already known and the results are simply a new interpretation. The $\alpha=\infty$ case had been solved only for special cases where P_{12} and P_{21} are square. Also, the existing solutions did not have computational schemes allowing their use on even moderately sized problems. These two limitations, especially the former, restricted the application of the pioneering H_∞ methods to fairly simple problems, such as sensitivity minimization. The new solution eliminates these two limitations.

Unfortunately, this new solution for the H_2 and H_∞ suffers from the same limitations imposed by restrictive assumptions about uncertainty as do the underlying analysis methods. While the SSV is a great improvement for analysis (Case 3), synthesis for the $\alpha=\mu$ case is not yet fully solved. Recalling that μ may be obtained by scaling and applying $\|\cdot\|_\infty$, a reasonable approach is to "solve"

$$\min_{K,D} \left\| DF(P,-K)D^{-1} \right\|_\infty \tag{3}$$

by iteratively solving for K and D. With either K or D fixed, the global optimum in the other variable may be found using the μ and H_∞ solutions described previously. Example designs have been done and this scheme seems to work well, but global convergence is not guaranteed. In fact, a counterexample has been constructed where (3) reaches a local minimum which is not global.

1.4.1

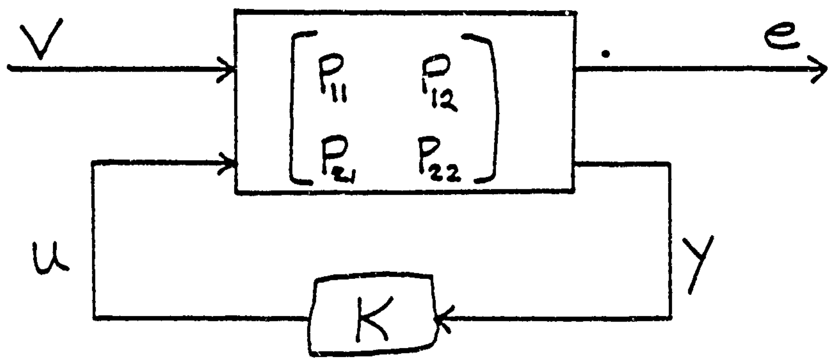


figure 1

2.1.1 Overview of Synthesis

From the previous part of these notes on analysis, we have seen that the synthesis problem in each case reduces to finding a controller K which achieves internal stability and solves

$$\min_{K \in R_p^{p_2 \times m_2}} \left\| F_l(P,K) \right\|_\alpha \quad \alpha = 2, \infty, \text{ or } \mu \tag{1}$$

where

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \in R_p^{(p_1+p_2)\times(m_1+m_2)} \quad , \quad P_{ij} \in R^{p_i \times m_j}$$

and

$$F_l(P,K) = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21}$$

We will restrict our attention for now to the $\alpha=2$ and ∞ cases of (1). Recall that the $\alpha=\mu$ case of (1) can be converted to the $\alpha=\infty$ case by scaling. The approach of these notes is to develop the $\alpha=2$ and ∞ cases in a parallel manner, emphasizing their common features.

We begin by considering the special case of (1) where all matrices are constants. This is an interesting problem in its own right and manages to capture the essential features of the general problem. While the $\alpha=2$ case is quite straightforward, the key step in the solution of (1) for $\alpha = \infty$ was first published in 1982 by Davis, Kahan, and Weinberger in their important paper on norm-preserving dilations.

The remainder of this part of the notes involves taking each step of the solution to (1) for the constant case and generalizing to the case of real-rational matrices. The difficulty arises from stability/causality considerations which are not present in the constant matrix case.

2.1.2 Constant Matrix Case

In this section, we will consider a special synthesis problem where all matrices are constants. The constant matrix case will allow us to study the synthesis problem in a simplified context, but one which parallels the rational case.

For constant matrices, the norms reduce to

$$\| P \|_\infty = \bar{\sigma}(P)$$
$$\| P \|_2 = \left[Tr(P^* P) \right]^{1/2}$$

Note that these definitions are not conventional, but they are convenient in allowing parallel development of the constant and rational cases.

Consider the constant matrix problem

$$\min_{K \in \mathbb{C}^{p_2 \times m_2}} \left\| \mathbf{F}_l(P,K) \right\|_\alpha \qquad \alpha = 2, \infty \tag{1}$$

where

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \in \mathbb{C}^{(p_1+p_2)\times(m_1+m_2)}, \quad P_{ij} \in \mathbb{C}^{p_i \times m_j},$$

and

$$\mathbf{F}_l(P,K) = P_{11} + P_{12}K(I-P_{22}K)^{-1}P_{21}.$$

Assume that $P_{12}^* P_{12} > 0$ and $P_{21} P_{21}^* > 0$.

The first step is to make the substitution of variables

$$K = \hat{Q}(I+P_{22}\hat{Q})^{-1}, \quad \hat{Q} = (P_{12}^* P_{12})^{-1/2} Q (P_{21} P_{21}^*)^{-1/2} \tag{2}$$

so

$$Q = (P_{12}^* P_{12})^{1/2} K(I-P_{22}K)^{-1}(P_{21} P_{21}^*)^{1/2} \tag{3}$$

Using the linear fractional representation notation,

$$K = \mathbf{F}_l(J, Q) \tag{4}$$

where

$$J = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} = \begin{bmatrix} 0 & (P_{12}^* P_{12})^{-1/2} \\ (P_{21} P_{21}^*)^{-1/2} & -(P_{21} P_{21}^*)^{-1/2} P_{22} (P_{12}^* P_{12})^{-1/2} \end{bmatrix}$$

With this substitution, we have

$$\begin{aligned} \mathbf{F}_l(P, K) &= \mathbf{F}_l(P, \mathbf{F}_l(J, Q)) \\ &= P_{11} + \left[P_{12}(P_{12}^* P_{12})^{-1/2}\right] Q \left[(P_{21} P_{21}^*)^{-1/2} P_{21}\right] \\ &= T_{11} + T_{12} Q T_{21} \end{aligned} \tag{5}$$

where the T_{ij} are defined in the obvious way. This parametrization has converted the nonlinear problem in (1) to one affine in the parameter Q. Note that $T_{12}^* T_{12} = I$ and $T_{21} T_{21}^* = I$. Thus we can find $T_\perp$ and $\tilde{T}_\perp$ such that both $\begin{bmatrix} T_{12} & T_\perp \end{bmatrix}$ and $\begin{bmatrix} T_{21} \\ \tilde{T}_\perp \end{bmatrix}$ are square and unitary.

Since both $\alpha = 2$ and ∞ norms are unitary invariant

$$\begin{aligned} \left\| T_{11} + T_{12} Q T_{21} \right\|_\alpha &= \left\| T_{11} + \begin{bmatrix} T_{12} & T_\perp \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} T_{21} \\ \tilde{T}_\perp \end{bmatrix} \right\|_\alpha \\ &= \left\| \begin{bmatrix} T_{12} & T_\perp \end{bmatrix}^* T_{11} \begin{bmatrix} T_{21} \\ \tilde{T}_\perp \end{bmatrix}^* + \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \right\|_\alpha \\ &= \left\| \begin{bmatrix} R_{11} + Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\alpha \end{aligned} \tag{6}$$

where

$$\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} = \begin{bmatrix} T_{12}^* T_{11} T_{21}^* & T_{12}^* T_{11} \tilde{T}_\perp^* \\ T_\perp^* T_{11} T_{21}^* & T_\perp^* T_{11} \tilde{T}_\perp^* \end{bmatrix} \tag{7}$$

Thus the problem in (1) reduces to

$$\min_{Q \in \mathcal{C}} \left\| \begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_{\alpha} \tag{8}$$

and solution of (8) yields a solution of (1) by solving (2) for K.

The $\alpha = 2$ case can be solved immediately from (8) since

$$\left\| \begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_2^2 = \left\| R_{11}+Q \right\|_2^2 + \left\| \begin{bmatrix} 0 & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_2^2 . \tag{9}$$

Thus

$$Q_{opt} = -R_{11} = -T_{12}^* T_{11} T_{21}^*$$

and

$$\min_{Q \in \mathcal{C}} \left\| \begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} 0 & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_2 \tag{10}$$

The simplicity of the $\alpha=2$ case is responsible for much of its appeal. Optimization in this norm reduces to projection since L_2 is a Hilbert space. This holds as well for the rational matrix problem.

The $\alpha=\infty$ case is somewhat more complicated since L_∞ is not a Hilbert space and the minimization in (8) cannot be solved by projection. Fortunately, L_∞ arises as the space of linear operators on the Hilbert space L_2, and (8) can be treated as a dilation problem. The next section focuses on matrix dilation problems.

2.1.3 Matrix Dilation Problems

Consider the optimization problem

$$\min_X \left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\|_\infty \qquad (1)$$

where X, B, C, A are constant matrices of compatible dimensions. This is a restatement of (2.7) for the $\alpha=\infty$ case. The matrix $\begin{bmatrix} X & B \\ C & A \end{bmatrix}$ is a *dilation* of its submatrices as indicated in the following diagram:

$$\begin{array}{ccc} \begin{bmatrix} X & B \\ C & A \end{bmatrix} & \underset{d}{\overset{c}{\rightleftarrows}} & \begin{bmatrix} B \\ A \end{bmatrix} \\ \\ d \quad c & & d \quad c \\ \\ \begin{bmatrix} C & A \end{bmatrix} & \underset{d}{\overset{c}{\rightleftarrows}} & \begin{bmatrix} A \end{bmatrix} \end{array} \qquad (2)$$

In this diagram, c stands for the operation of *compression* and d stands for *dilation*. Compression is always norm decreasing; sometimes dilation can be made to be norm preserving. Norm preserving dilations are the focus of this section.

The simplest matrix dilation problem occurs when solving

$$\min_X \left\| \begin{bmatrix} X \\ A \end{bmatrix} \right\|_\infty \qquad (3)$$

Although (3) is a much simplified version of (1), we will see that it contains all the essential features of the problem. Letting γ_0 denote the minimum norm in (3), it is immediate that $\gamma_0 = \|A\|_\infty$. The following theorem characterizes all solutions to (3).

Theorem 1: For $\forall \gamma \geq \gamma_0$,

$$\left\| \begin{bmatrix} X \\ A \end{bmatrix} \right\|_\infty \leq \gamma \tag{4}$$

iff $\exists Y$ with $\|Y\|_\infty \leq 1$ such that

$$X = Y(\gamma^2 I - A^* A)^{1/2} \tag{5}$$

Proof:

$$\left\| \begin{bmatrix} X \\ A \end{bmatrix} \right\|_\infty \leq \gamma$$

iff

$$X^* X + A^* A \leq \gamma^2 I$$

iff

$$X^* X \leq (\gamma^2 I - A^* A)$$

iff

$$\|Xu\| \leq \|(\gamma^2 I - A^* A)^{1/2} u\| \quad \forall u$$

iff

$$X = Y(\gamma^2 I - A^* A)^{1/2} \quad \text{for some} \quad \|Y\|_\infty \leq 1$$

This theorem implies that, in general, (3) has more than one solution. This is in contrast to the $\alpha = 2$ case. The solution $X = 0$ is the central solution but others are possible unless $A^* A = \gamma^2 I$. A more restricted version of the theorem is

Corollary 1: For $\gamma > \gamma_o$,

$$\left\| \begin{bmatrix} X \\ A \end{bmatrix} \right\|_\infty \le \gamma \tag{6}$$

iff

$$\left\| X(\gamma^2 I - A^* A)^{-1/2} \right\|_\infty \le 1. \tag{7}$$

The corresponding dual results are

Theorem 1' For $\forall \gamma \ge \gamma_o$

$$\left\| \begin{bmatrix} X & A \end{bmatrix} \right\|_\infty \le \gamma \quad \text{iff} \quad \exists Y,\ \left\| Y \right\|_\infty \le 1 \tag{8}$$

such that

$$X = (\gamma^2 I - AA^*)^{1/2} Y \tag{9}$$

Corollary 1' For $\gamma > \gamma_o$

$$\left\| \begin{bmatrix} X & A \end{bmatrix} \right\|_\infty \le \gamma \tag{10}$$

iff

$$\left\| (\gamma^2 I - AA^*)^{-1/2} X \right\|_\infty \le 1 \tag{11}$$

Now, returning to the problem in (1), let

$$\gamma_o = \min_X \left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\|_\infty \tag{12}$$

The following theorem (Parrott) will play a central role in the synthesis theory. The proof is a straightforward application of Theorem 1 and 1'.

Theorem 2

$$\gamma_0 = \max\left\{\left\|\begin{bmatrix} C & A\end{bmatrix}\right\|_\infty, \left\|\begin{bmatrix} B \\ A\end{bmatrix}\right\|_\infty\right\} \tag{13}$$

Proof: Denote by $\hat{\gamma}$ the right hand side of the equation (13). Clearly, $\gamma_0 \geq \hat{\gamma}$ since compressions are norm decreasing. That $\gamma_0 \leq \hat{\gamma}$ will be shown by using Theorem 1 and 1'.

From Theorem 1 we have that $B = Y(\hat{\gamma}^2 I - A^*A)^{1/2}$ for some Y such that $\|Y\|_\infty \leq 1$. Similarly, Theorem 1' yields $C = (\hat{\gamma}^2 I - AA^*)^{1/2} Z$ for some Z with $\|Z\|_\infty \leq 1$.

Let $\hat{X} = -YA^*Z$. Then

$$\begin{aligned}
\left\|\begin{bmatrix} \hat{X} & B \\ C & A\end{bmatrix}\right\|_\infty &= \left\|\begin{bmatrix} -YA^*Z & Y(\hat{\gamma}^2 I - A^*A)^{1/2} \\ (\hat{\gamma}^2 I - AA^*)^{1/2} Z & A\end{bmatrix}\right\|_\infty \\
&\leq \left\|\begin{bmatrix} -A^* & (\hat{\gamma}^2 I - A^*A)^{1/2} \\ (\hat{\gamma}^2 I - AA^*)^{1/2} & A\end{bmatrix}\right\|_\infty \\
&= \hat{\gamma}
\end{aligned}$$

Since

$$\begin{bmatrix} -A^* & (\hat{\gamma}^2 I - A^*A)^{1/2} \\ (\hat{\gamma}^2 I - AA^*)^{1/2} & A\end{bmatrix}\begin{bmatrix} -A & (\hat{\gamma}^2 I - AA^*)^{1/2} \\ (\hat{\gamma}^2 I - A^*A)^{1/2} & A^*\end{bmatrix} = \begin{bmatrix} \hat{\gamma}^2 I & 0 \\ 0 & \hat{\gamma}^2 I\end{bmatrix}$$

Thus $\hat{\gamma} \geq \gamma_0$, so $\hat{\gamma} = \gamma_0$.

This theorem gives one solution to (12) and an expression for γ_0. As in (3), there may be more than one solution to (12), although Theorem 2 only exhibits one. Theorem 3 considers the problem of parametrizing all solutions. The solution $\hat{X} = -YA^*Z$ is the "central" solution analogous to $X = 0$ in (3). The next corollary is an alternative statement of Theorem 2 using the form of (2.4) for the problem

$$\gamma_0 = \min_Q \left\| R + UQV \right\|_\infty \tag{14}$$

where $U^*U = I$ and $VV^* = I$

Corollary 2

$$\gamma_0 = \max \left\{ \left\| U_\perp^* R \right\|_\infty , \left\| R V_\perp^* \right\|_\infty \right\} \tag{15}$$

The following theorem (Davis, Kahan, and Weinberger) parametrizes all solutions to (1). The proof is omitted, but is similar to Theorem 2 and involves application of Theorem 1 and 1'.

Theorem 3 Suppose $\gamma \geq \gamma_0$. The solutions X such that

$$\left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\|_\infty \leq \gamma \tag{16}$$

are exactly those of the form

$$X = -YA^*Z + \gamma(I-YY^*)^{1/2} W (I-Z^*Z)^{1/2} \tag{17}$$

where W is an arbitrary contraction ($\left\|W\right\|_\infty \leq 1$) and Y and Z solve the linear equations

$$\begin{aligned} B &= Y(\gamma^2 I - A^*A)^{1/2} \\ C &= (\gamma^2 I - AA^*)^{1/2} Z. \end{aligned} \tag{18}$$

The following corollary gives an alternative version of Theorem 3.

Corollary 3 For $\gamma > \gamma_o$,

$$\left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\|_\infty \le \gamma \tag{19}$$

iff

$$\left\| (I - YY^*)^{-1/2} (X + Y^* A Z)(I - Z^* Z)^{-1/2} \right\|_\infty \le \gamma \tag{20}$$

where

$$Y = B(\gamma^2 I - A^* A)^{-1/2}$$
$$Z = (\gamma^2 I - AA^*)^{-1/2} C \tag{21}$$

There are many alternative characterizations of solutions to (19), although the formulas in (20) and (21) seem to be the simplest.

For the problem in (14), the following equivalences apply for all $\gamma > \gamma_o$:

$$\| R + UQV \|_\infty \le \gamma \tag{22}$$

iff

$$\left\| \begin{bmatrix} RV^* + UQ & RV_\perp^* \end{bmatrix} \right\|_\infty \le \gamma \tag{23}$$

iff

$$\left\| \begin{bmatrix} R_1 + UQ & A \end{bmatrix} \right\|_\infty \le \gamma \tag{24}$$

where $R_1 = RV^*$ and $A = RV_\perp^*$

iff

$$\left\| (\gamma^2 I - AA^*)^{-1/2} \begin{bmatrix} R_1 + UQ \end{bmatrix} \right\|_\infty \le 1 \tag{25}$$

(by Corollary 1′)

iff

$$\left\| \left[R_2 + U_2 Q \right] \right\|_\infty \le 1 \tag{26}$$

where $R_2 = (\gamma^2 I - AA^*)^{-1/2} R_1$

$U_2 = (\gamma^2 I - AA^*)^{-1/2} U.$

To complete this, simply factor U_2 to extract a unitary factor, and apply the dual of (22)-(26) to (26). Although the formulas get messy, (22) can be solved in this manner.

In each of these cases, Theorem 2, Theorem 3, their corollaries, and the solution described above, the general case reduces almost immediately to application of Theorem 1 or Corollary 1. Thus, when it is convenient, we will consider (4) rather than (16) and (26) rather than (22). This will simplify the discussion of the rational case without introducing any loss of generality.

2.1.4 Summary of Constant Problem

The rational matrix problem in equation (1.1) can be solved in a manner which parallels the treatment of the constant case in the last two sections. This generalization is the focus of the next three chapters on synthesis. To reinforce the similarity between the constant and rational case, we will now review the key steps from the previous two sections and preview their generalizations to the rational case.

Consider the diagram in Figure 1. This summarizes the steps in the constant matrix problem (2.1). The main steps are as follows:

1) **Parametrization:** Make the substitution $K = F_l(J, Q)$ so that

$$\begin{aligned} F_l(P, K) &= F_l(P, F_l(J, Q)) \\ &= F_l(T, Q) \\ &= T_{11} + T_{12} Q T_{21} \end{aligned} \qquad (1)$$

is affine. Additionally, we want $T_{12}^* T_{12} = I$ and $T_{21} T_{21}^* = I$.

2) **Unitary Invariance:** Find $T_\perp$ and $\tilde{T}_\perp$ so that $\begin{bmatrix} T_{12} & T_\perp \end{bmatrix}$ and $\begin{bmatrix} T_{21} \\ \tilde{T}_\perp \end{bmatrix}$ are square and unitary.

Pre- and post-multiply by $\begin{bmatrix} T_{12} & T_\perp \end{bmatrix}^*$ and $\begin{bmatrix} T_{21} \\ \tilde{T}_\perp \end{bmatrix}^*$ to yield

$$\begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \qquad (2)$$

where

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

$$= \begin{bmatrix} T_{12}^* \\ T_\perp^* \end{bmatrix} T_{11} \begin{bmatrix} T_{21}^* & \tilde{T}_\perp^* \end{bmatrix} \tag{3}$$

Recall that without loss of generality, we may assume the $T_{11}\tilde{T}_\perp^* = 0$ so that (2) becomes

$$\begin{bmatrix} R_{11}+Q \\ R_{21} \end{bmatrix}. \tag{4}$$

3) **Projection / Dilation:** At this point the $\alpha=2$ and $\alpha=\infty$ cases differ. For $\alpha=2$, the problem reduces, by projection, to

$$\min_Q \left\| R_{11} + Q \right\|_2 \tag{5}$$

which has the unique solution $Q=-R_{11}$.

The $\alpha=\infty$ case must be treated using the matrix dilation theory of the previous section. Recall that, in general, the solution is not unique. From Theorem 3.1, all solutions to

$$\left\| \begin{bmatrix} R_{11}+Q \\ R_{21} \end{bmatrix} \right\|_\infty \le \gamma \quad \text{for} \quad \gamma \ge \left\| R_{21} \right\|_\infty \tag{6}$$

are of the form

$$Q = -R_{11} + Y(\gamma^2 I - R_{21}^* R_{21})^{1/2} \tag{7}$$

for some $\left\| Y \right\|_\infty \le 1$. Corollary 3.1 gave the alternative characterization that

$$\left\| \begin{bmatrix} R_{11}+Q \\ R_{21} \end{bmatrix} \right\|_\infty \le \gamma \quad \text{for} \quad \gamma > \left\| R_{21} \right\|_\infty \tag{8}$$

if and only if

$$\left\| (R_{11}+Q)(\gamma^2 I - R_{21}^* R_{21})^{-1/2} \right\|_\infty \le 1. \tag{9}$$

It is this latter characterization which will be used in the rational case. Note that $Q=-R_{11}$ is one solution to (6) and (8).

4) **Recovery of the optimal K:** This is obtained by simply computing K from the formula $K=F_l(J,Q)$ used in step 1) to parametrize the problem.

2.1.4 Figure 1

$$\min_K \left\|F_l(P,K)\right\|_\alpha \quad \text{where} \quad F_l(P,K) = P_{11} + P_{12}K(I-P_{22}K)^{-1}P_{21}$$

$$\Big\downarrow \; \begin{array}{l} \text{parametrization} \\ K=F_l(J,Q) \end{array}$$

$$\min_Q \left\|T_{11} + T_{12}QT_{21}\right\|_\alpha \quad \text{where} \quad T_{12}^{*}T_{12} = I \qquad T_{21}T_{21}^{*} = I$$

$$\Big\downarrow \; \text{unitary invariance}$$

$$\min_Q \left\|\begin{matrix} R_{11}-Q & R_{21} \\ R_{12} & R_{22} \end{matrix}\right\|_\alpha$$

$$\begin{array}{l} \alpha=2 \\ \text{projection} \end{array} \Big\downarrow \qquad \Big\downarrow \begin{array}{l} \alpha=\infty \\ \text{dilation} \end{array}$$

$$\min_Q \left\|G-\hat{Q}\right\|_\alpha$$

$$\Big\downarrow$$

$$Q_{opt}$$

$$\Big\downarrow \; K_{opt}=F_l(J,Q_{opt})$$

$$K_{opt}$$

2.1.5 Rational Matrix Generalization

The steps in the rational case closely parallel the constant case, as shown in Figure 1. Most of the work in the remaining chapters is devoted to generalizing these steps from constants to rationals. The source of all the difficulty in the rational case comes from the requirement for internal stability, or equivalently, causality. Without this the rational case would reduce to the constant case at each frequency, and could be solved using the results of the previous two sections.

We will now briefly outline the steps required to solve the rational case and preview the upcoming chapters.

1) **Parametrization:** Find J so that the substitution $K=F_l(J,Q)$ yields

$$\begin{aligned} F_l(P,K) &= F_l(P,F_l(J,Q)) \\ &= F_l(T,Q) \\ &= T_{11} + T_{12}QT_{21} \end{aligned} \tag{10}$$

with the additional requirement that $T \in H_\infty$ and

$$\begin{aligned} F_l(P,K) &\quad \text{internally stable} \\ \text{iff} &\quad Q \in H_\infty \end{aligned} \tag{11}$$

This parametrizes all stabilizing K's in terms of a stable $Q \in H_\infty$ in addition to providing an affine parametrization of all stable $F_l(P,K)$. This parametrization (Youla) is developed in Chapter 2 on Stabilization.

A further requirement is that T_{12} and T_{21} be inner, that is $T_{12}^* T_{12}=I$ and $T_{21}T_{21}^*=I$. Methods for obtaining the particular parametrizations which achieve this are developed in Chapter 3 on Factorization.

2) **Unitary Invariance:** Find $T_\perp$ and $\tilde{T}_\perp$ so that $\begin{bmatrix} T_{12} & T_\perp \end{bmatrix}$ and $\begin{bmatrix} T_{21} \\ \tilde{T}_\perp \end{bmatrix}$ are *square* and inner (also Chapter 3). Then pre- and post-multiply by $\begin{bmatrix} T_{12} & T_\perp \end{bmatrix}^*$ and $\begin{bmatrix} T_{21} \\ \tilde{T}_\perp \end{bmatrix}^*$ to yield

$$\begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \tag{12}$$

where

$$R = \begin{bmatrix} T_{12}^* \\ T_\perp^* \end{bmatrix} T_{11} \begin{bmatrix} T_{21}^* & \tilde{T}_\perp^* \end{bmatrix}$$

Again, to simplify the presentation suppose that $T_{11}\tilde{T}_\perp^*=0$ so that (12) becomes

$$\begin{bmatrix} R_{11}+Q \\ R_{21} \end{bmatrix}. \tag{13}$$

3) **Projection / Dilation:** At this point the $\alpha=2$ and $\alpha=\infty$ cases again differ. For $\alpha=2$, the problem reduces, by projection, to

$$\min_{Q \in H_\infty} \left\| R_{11} + Q \right\|_2 \tag{14}$$

But since $R_{11} \in L_\infty$, $Q=-R_{11}$ would not correspond to a stable solution. The unique solution is yet another projection

$$Q_{opt} = P_{H_2}(R_{11}) \tag{15}$$

where P_{H_2} denotes projection onto H_2. When viewed appropriately, these two projections can be seen as a single projection onto a subspace of $L_2(j\mathbb{R}, \mathbb{C}^{p_1 \times m_1})$.

The $\alpha=\infty$ case is again treated as a dilation problem. Since, generically

$$\min_{Q \in H_\infty} \left\| \begin{bmatrix} R_{11}+Q \\ R_{21} \end{bmatrix} \right\|_\infty > \hat{\gamma} \triangleq \left\| R_{21} \right\|_\infty \tag{16}$$

it is convenient to use the characterization in Corollary 3.1 and (8)-(9). Recall that

$$\left\| \begin{bmatrix} R_{11}+Q \\ R_{21} \end{bmatrix} \right\|_\infty \leq \gamma \quad \text{for} \quad \gamma > \hat{\gamma} \tag{17}$$

iff

$$\left\| (R_{11}+Q)(\gamma^2 I - R_{21}^* R_{21})^{-1/2} \right\|_\infty \leq 1 \tag{18}$$

The key to proceeding in the rational case is to find $M \in RH_\infty$ such that $M^{-1} \in RH_\infty$ and $M^* M = (\gamma^2 I - R_{21}^* R_{21})$. If we use the symbol $(\gamma^2 I - R_{21}^* R_{21})^{1/2}$ to denote this M, then (18) makes sense in the rational case. Finding M involves spectral factorization and is treated in Chapter 3.

Given $M \in RH_\infty$ with the desired properties, (18) reduces to

$$\left\| G + \hat{Q} \right\|_\infty \leq 1 \tag{19}$$

where $G = R_{11} M^{-1} \in RH_\infty$ and $\hat{Q} = Q M^{-1}$. Solving (19) for $\hat{Q} \in RH_\infty$ solves (18) for $Q \in RH_\infty$. Note that $Q = \hat{Q} M$ is in RH_∞ if $\hat{Q}$ is, since $M \in RH_\infty$ by construction.

The final step in the rational case then involves solving (19) for $\hat{Q} \in RH_\infty$. This is a standard mathematical problem of approximating an L_∞ matrix by an H_∞ matrix. This turns out to be yet another dilation problem but in a somewhat different guise than those treated in the constant case. The solution of (19) is the focus of Chapter 4 on Best Approximation.

4) **Recovery of the optimal K:** Just as in the constant case $K_{opt} = F_l(J, Q_{opt})$. This K_{opt} will stabilize $F_l(P, K_{opt})$ since the parametrization in Step 1) insured that Q stable lead to internal stablility of

$$F_1(P,K)=F_1(P,F_1(J,Q)).$$

2.1.5 **Figure 1**

$$\min_{K \in R_p} \left\{ \left\| F_l(P,K) \right\|_\alpha \;\middle|\; \text{internal stability} \right\}$$

$$\Big\downarrow \; \begin{matrix} \text{parametrization} \\ K = F_l(J,Q) \end{matrix}$$

$$\min_{Q \in H_\infty} \left\| T_{11} + T_{12} Q T_{21} \right\|_\alpha \quad \text{where} \quad T_{12}^* T_{12} = I \qquad T_{21} T_{21}^* = I$$

$$\Big\downarrow \; \text{unitary invariance}$$

$$\min_{Q \in H_\infty} \left\| \begin{matrix} R_{11} - Q & R_{21} \\ R_{12} & R_{22} \end{matrix} \right\|_\alpha$$

$$\begin{matrix} \alpha = 2 \\ \text{projection} \end{matrix} \Big\downarrow \qquad \Big\downarrow \begin{matrix} \alpha = \infty \\ \text{dilation} \end{matrix}$$

$$\min_{\hat{Q} \in H_\infty} \left\| G - \hat{Q} \right\|_\infty , \quad G \in L_\infty$$

$$\begin{matrix} \alpha = 2 \\ \text{projection} \end{matrix} \Big\downarrow \qquad \Big\downarrow \begin{matrix} \alpha = \infty \\ \text{dilation} \end{matrix}$$

$$Q_{opt}$$

$$\Big\downarrow \; K_{opt} = F_l(J, Q_{opt})$$

$$K_{opt}$$

2.2.2 Internal Stability

In this section P and K are fixed proper transfer matrices. The block diagram of Figure 1 represents the two equations

$$\begin{bmatrix} e \\ y \end{bmatrix} = P \begin{bmatrix} v \\ u \end{bmatrix}, \quad u = Ky. \tag{1}$$

Partition P accordingly:

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}. \tag{2}$$

It is convenient to introduce two fictitious external signals, w_1 and w_2, as in Figure 1a.

Suppose the signals v, w_1, and w_2 are specified and that u in Figure 1a is well-defined. Then so are e and y. Thus it makes sense to define the system in Figure 1a to be *well-posed* provided the transfer matrix from $\begin{bmatrix} v \\ w_1 \\ w_2 \end{bmatrix}$ to u exists and is proper.

Lemma 1. The system is well-posed if and only if

$$I - K(\infty)P_{22}(\infty) \quad \text{is invertible.} \tag{3}$$

Proof. Figure 1a implies the equations

$$u = w_1 + KY + Kw_2$$
$$y = P_{21}v + P_{22}u$$

and these in turn imply that

$$(I - KP)u = w_1 + KP_{21}v + Kw_2.$$

Thus well-posedness is equivalent to the condition that $(I - KP)^{-1}$ exists and is proper.

QED

It is straightforward to show that (3) is equivalent to either of the following two conditions:

$$\begin{bmatrix} I & -K(\infty) \\ -P_{22}(\infty) & I \end{bmatrix} \quad \text{is invertible ;} \tag{4}$$

$$I - P_{22}(\infty)K(\infty) \quad \text{is invertible .} \tag{5}$$

The well-posedness condition is simple to state in terms of state-space realizations. Introduce minimal realizations of P and K:

$$P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] \tag{6}$$

$$K = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right]. \tag{7}$$

Note that the partition in (6) corresponds to that in (2), i.e.,

$$P_{ij} = \left[\begin{array}{c|c} A & B_j \\ \hline C_i & D_{ij} \end{array}\right]. \tag{8}$$

Then $P_{22}(\infty)=D_{22}$ and $K(\infty)=\hat{D}$, so for example, from (4) well-posedness is equivalent to the condition that

$$\begin{bmatrix} I & -\hat{D} \\ -D_{22} & I \end{bmatrix} \quad \text{is invertible.} \tag{9}$$

Well-posedness will be assumed for the rest of this section. Let x and $\hat{x}$ denote the state vectors for P and K respectively, and write the system equations in Figure 1 with v set to zero and e ignored:

$$\dot{x} = Ax + B_2u \tag{10a}$$

$$y = C_2x + D_{22}u. \tag{10b}$$

$$\dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}y \tag{10c}$$

$$u = \hat{C}\hat{x} + \hat{D}y. \tag{10d}$$

The system of Figure 1 is *internally stable* provided the origin $(x,\hat{x}) = (0,0)$ is asymptotically stable. To get a concrete characterization of internal stability, solve equations (10b) and (10d) for u and y:

$$\begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} I & -\hat{D} \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & \hat{C} \\ C_2 & 0 \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}.$$

(Note that the inverse exists from (9)). Now substitute this into (10a) and (10c) to get

$$\frac{d}{dt}\begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \tilde{A}\begin{bmatrix} x \\ \hat{x} \end{bmatrix}$$

where

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix} + \begin{bmatrix} B_2 & 0 \\ 0 & \hat{B} \end{bmatrix} \begin{bmatrix} I & -\hat{D} \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & \hat{C} \\ C_2 & 0 \end{bmatrix}.$$

Thus internal stability is equivalent to the condition that $\tilde{A}$ has all its eigenvalues in the open left half-plane.

It is routine to verify that the above definition of internal stability depends only on P and K, not specific realizations of them. The following result is standard.

Lemma 2. Consider a minimal realization of P as in (6). There exists a proper K achieving internal stability iff (A,B_2) is stabilizable and (C_2,A) is detectable.

The latter stabilizability and detectability conditions are *assumed* throughout the remainder of this chapter. Since

$$P_{22} = \left[\begin{array}{c|c} A & B_2 \\ \hline C_2 & D_{22} \end{array}\right], \tag{11}$$

equations (10) constitute a state-space representation of the system in Figure 2. Although the realization in (11) is not necessarily minimal, it is stabilizable and detectable, and these are enough to yield the following result.

Lemma 3. The system in Figure 1 is internally stable iff the one in Figure 2 is.

The next section contains a parametrization of all K's which achieve internal stability for the system in Figure 2. To simplify notation, define

$$G := P_{22}, \quad B := B_2, \quad C := C_2, \quad D := D_{22}.$$

Then (A,B) is stabilizable, (C,A) is detectable, and the system under study is that in Figure 3.

The above notion of internal stability is defined in terms of state-space realizations of G and K. It is important and useful to characterize internal stability from an input/output point of view. For this, consider the feedback system in Figure 4. This system is described by:

$$\begin{bmatrix} I & -K \\ -G & I \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}. \tag{10}$$

Now it is intuitively clear that if the system in Figure 4 is internally stable then for all bounded inputs (v_1, v_2), the outputs (e_1, e_2) are also bounded. The following lemma shows that this idea lends to an input/output characterization of internal stability.

Lemma 4. The system in Figure 4 is internally stable if and only if $(I-GK)$ is invertible and the transfer matrix

$$\begin{bmatrix} I & -K \\ -G & I \end{bmatrix}^{-1} = \begin{bmatrix} I+K(I-GK)^{-1}G & K(I-GK)^{-1} \\ (I-GK)^{-1}G & (I-GK)^{-1} \end{bmatrix} \tag{11}$$

between (v_1,v_2) and (e_1,e_2) belongs to RH_∞.

Proof. As above let (A,B,C,D) and $(\hat{A},\hat{B},\hat{C},\hat{D})$ be stabilizable and detectable realizations of G and K respectively. Then the state-space equations for the system in Figure 4 are:

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & \hat{B} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & \hat{C} \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} D & 0 \\ 0 & \hat{D} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$e_1 = v_1 + y_2, \quad e_2 = v_2 + y_1.$$

The last two equations can be rewritten as

$$\begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

Now suppose that this system is internally stable. Then (7) implies that $(I-D\hat{D}) = (I-GK)(\infty)$ is invertible. Hence $(I-GK)$ is invertible. Further, since the eigenvalues of

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & \hat{B} \end{bmatrix} \begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix}$$

are in the open left half plane, it follows that the transfer matrix in (10) from (v_1,v_2) to (e_1,e_2) is in RH_∞.

Conversely, suppose that $(I-GK)$ is invertible and the transfer matrix in (10) is in RH_∞. Then, in particular, $(I-GK)^{-1}$ is proper which implies that $(I-GK)(\infty) = (I-D\hat{D})$ is invertible. Therefore

$$\begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix}$$

is nonsingular. Now routine transfer function calculations give,

$$\begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \left[I + \begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix}(sI-\tilde{A})^{-1}\begin{bmatrix} B & 0 \\ 0 & \hat{B} \end{bmatrix}\right]\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

Since the transfer matrix from (v_1, v_2) to (e_1, e_2) belongs to RH_∞, it follows that

$$\begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix}(sI-\tilde{A})^{-1}\begin{bmatrix} B & 0 \\ 0 & \hat{B} \end{bmatrix}$$

belongs to RH_∞. Finally, since (A, B, C) and $(\hat{A}, \hat{B}, \hat{C})$ are stabilizable and detectable, it follows that the eigenvalues of $\tilde{A}$ are in the open left half plane.

QED

We note that to check internal stability it is necessary (and sufficient) to check that each of the four transfer matrices in (11) are in RH_∞. It is not difficult to construct examples of G and K such that any three of the four transfer matrices in (11) are in RH_∞ while the fourth one is unstable.

2.2.3 Parametrization of All Stabilizing Controllers

Two matrices $N, M \in RH_\infty$ with the same number of columns are *right-coprime* if the combined matrix $\begin{bmatrix} M \\ N \end{bmatrix}$ has a left inverse in RH_∞. That is, there exists $X, Y \in RH_\infty$ such that $XM + YN = I$. This is often called a Bezout or Diophantine equation. An alternative definition is that two matrices in RH_∞ are right-coprime if every common right divisor in RH_∞ is invertible in RH_∞. This can be shown to be equivalent to the above definition in terms of a left inverse, but we will not use this fact.

It is a fact that every $G \in R_p$ (proper, real-rational) has a right-coprime factorization $G = NM^{-1}$ where $N, M \in RH_\infty$ are right coprime. Similarly, there exist left coprime factorizations (lcf), defined in the obvious way by duality. The proof of the existence of such coprime factorizations will be given in the next section with explicit realizations for the factorizations. In this section, we will see how these factorizations can be used to obtain a parametrization of all stabilizing controllers.

Begin with rcf's and lcf's of G and K in Figure 4:

$$G = NM^{-1} = \tilde{M}^{-1}\tilde{N} \tag{1}$$

$$K = UV^{-1} = \tilde{V}^{-1}\tilde{U} \tag{2}$$

Lemma 1. Consider the system in Figure 4. The following conditions are equivalent:

1. The feedback system is internally stable.
2. $\begin{bmatrix} M & U \\ N & V \end{bmatrix}$ is invertible in RH_∞.
3. $\begin{bmatrix} \tilde{V} & -\tilde{U} \\ -\tilde{N} & \tilde{M} \end{bmatrix}$ is invertible in RH_∞.

Proof: As we saw in Lemma 2.3 of the last section, internal stability is equivalent to the condition that

$$\begin{bmatrix} I & -K \\ -G & I \end{bmatrix}^{-1} \in RH_\infty$$

or, equivalently,

$$\begin{bmatrix} I & K \\ G & I \end{bmatrix}^{-1} \in RH_\infty. \tag{3}$$

Now

$$\begin{aligned} \begin{bmatrix} I & K \\ G & I \end{bmatrix} &= \begin{bmatrix} I & UV^{-1} \\ NM^{-1} & I \end{bmatrix} \\ &= \begin{bmatrix} M & U \\ N & V \end{bmatrix} \begin{bmatrix} M^{-1} & 0 \\ 0 & V^{-1} \end{bmatrix} \end{aligned}$$

so that

$$\begin{bmatrix} I & K \\ G & I \end{bmatrix}^{-1} = \begin{bmatrix} M & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} M & U \\ N & V \end{bmatrix}^{-1}.$$

Since the matrices

$$\begin{bmatrix} M & 0 \\ 0 & V \end{bmatrix}, \begin{bmatrix} M & U \\ N & V \end{bmatrix}$$

are right-coprime, (3) holds iff

$$\begin{bmatrix} M & U \\ N & V \end{bmatrix}^{-1} \in RH_\infty.$$

This proves the equivalence of conditions 1 and 2. The equivalence of 1 and 3 is proved similarly.

QED

We shall see in the next section how to find explicit realizations for $N, M, \widetilde{N}, \widetilde{M}, U_0, V_0, \widetilde{U}_0, \widetilde{V}_0$ and such that (1) holds and

$$\begin{bmatrix} \widetilde{V}_0 & -\widetilde{U}_0 \\ -\widetilde{N} & \widetilde{M} \end{bmatrix} \begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{5}$$

The above lemma says that

$$K_0 \triangleq U_0 V_0^{-1} = \widetilde{V}_0^{-1} \widetilde{U}_0$$

then qualifies a particular controller achieving internal stability. All stabilizing controllers can be expressed in terms of K_0 and a parameter Q, as shown in the following:

Theorem 1. The set of all proper controllers achieving internal stability is parametrized by the formula

$$K = K_0 + \widetilde{V}_0^{-1} Q (I + V_0^{-1} N Q)^{-1} V_0^{-1} \tag{6}$$

where Q ranges over RH_∞ such that $(I + V_0^{-1} N Q)(\infty)$ is invertible.

Proof: Assume K has the form indicated.

Define

$$U \triangleq U_0 + MQ, \quad V \triangleq V_0 + NQ$$
$$\widetilde{U} \triangleq \widetilde{U}_0 + Q\widetilde{M}, \quad \widetilde{V} \triangleq \widetilde{V}_0 + Q\widetilde{N}$$

then

$$\begin{aligned} \begin{bmatrix} \widetilde{V} & -\widetilde{U} \\ -\widetilde{N} & \widetilde{M} \end{bmatrix} \begin{bmatrix} M & U \\ N & V \end{bmatrix} &= \begin{bmatrix} \widetilde{V}_0 + Q\widetilde{N} & -(\widetilde{U}_0 + Q\widetilde{M}) \\ -\widetilde{N} & \widetilde{M} \end{bmatrix} \begin{bmatrix} M & U_0 + MQ \\ N & V_0 + NQ \end{bmatrix} \\ &= \begin{bmatrix} I & -Q \\ 0 & I \end{bmatrix} \begin{bmatrix} \widetilde{V}_0 & -\widetilde{U}_0 \\ -\widetilde{N} & \widetilde{M} \end{bmatrix} \begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix} \begin{bmatrix} I & Q \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} I & -Q \\ 0 & I \end{bmatrix} \begin{bmatrix} I & Q \\ 0 & I \end{bmatrix} \qquad \text{from (5)} \end{aligned}$$

$$= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{7}$$

Thus K achieves internal stability by lemma 1.

Conversely, suppose K is proper and it achieves internal stability. Introduce rcf and lcf of K as in (2).

Then by lemma 1, $Z \triangleq \tilde{M} V - \tilde{N} U$ is invertible in RH_∞. Define Q by the equation

$$U_0 + MQ = UZ^{-1}, \tag{8}$$

so

$$Q = M^{-1}(UZ^{-1} - U_0) \tag{9}$$

Then

$$\begin{aligned} V_0 + NQ &= V_0 + NM^{-1}(UZ^{-1} - U_0) \\ &= V_0 + \tilde{M}^{-1}\tilde{N}(UZ^{-1} - U_0) && \text{from (1)} \\ &= \tilde{M}^{-1}(\tilde{M} V_0 - \tilde{N} U_0 + \tilde{N} UZ^{-1}) \\ &= \tilde{M}^{-1}(I + \tilde{N} UZ^{-1}) && \text{from (5)} \\ &= \tilde{M}^{-1}(Z + \tilde{N} U)Z^{-1} \\ &= \tilde{M}^{-1}\tilde{M} VZ^{-1} \\ &= VZ^{-1} && (10) \end{aligned}$$

Thus,

$$\begin{aligned} K &= UV^{-1} \\ &= (U_0 + MQ)(V_0 + NQ)^{-1} \\ &= U_0 V_0 + (M - U_0 V_0^{-1} N) Q (I + V_0^{-1} NQ)^{-1} V_0^{-1} && (11) \end{aligned}$$

from (prelim?). Then, since

$$(M - U_0 V_0^{-1} N) = (M - \tilde{V}_0^{-1} \tilde{U}_0 N) = \tilde{V}_0^{-1}(\tilde{V}_0 M - \tilde{U}_0 N) = \tilde{V}_0^{-1}$$

we have that

$$K = U_0 V_0 + \tilde{V}_0^{-1} Q (I + V_0^{-1} N Q)^{-1} V_0^{-1}. \tag{12}$$

To see that Q belongs to RH_∞ observe first from (9) and (10) that MQ and NQ both do. Right-coprimeness of N and M then implies that $Q \in RH_\infty$.

Finally, since V and Z evaluated at $s=\infty$ are both invertible, so is $V_0 + NQ$, from (10), hence so is $I + V_0^{-1} NQ$.

QED

Define the rational matrix

$$J \triangleq \begin{bmatrix} K_0 & \tilde{V}_0^{-1} \\ V_0^{-1} & -V_0^{-1} N \end{bmatrix} \tag{13}$$

and consider a controller K given by formula (3). Then the controller equation

$$\begin{aligned} u &= \mathbf{F}_l(J, Q) y \\ &= \left[K_0 + \tilde{V}_0^{-1} Q (I + V_0^{-1} N Q)^{-1} V_0^{-1} \right] y \end{aligned}$$

is equivalent to the triple of equations

$$\begin{aligned} u &= K_0 y + \tilde{V}_0^{-1} y_1 \\ u_1 &= V_0^{-1} y - V_0^{-1} N y_1 \\ y_1 &= Q u_1 \end{aligned}$$

The block diagram corresponding to this triple is in Figure 5. We conclude that every stabilizing controller can be represented as $K = \mathbf{F}_l(J, Q)$, as in Figure 5, for some parameter Q, which is constrained only to be stable and proper and to make K proper.

The next section gives an explicit state-space realization of one choice of the interconnection matrix J.

2.2.4 Realization of J :

Recall from Section 2 that we have

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

where (A,B) is stabilizable and (C,A) is detectable. To obtain a right-coprime factorization of G, choose a matrix F such that $A+BF$ is stable.

Lemma 1. A stabilizing state feedback F yields rcf $G = NM^{-1}$ where

$$\begin{bmatrix} M \\ N \end{bmatrix} := \left[\begin{array}{c|c} A+BF & B \\ \hline F & I \\ C+DF & D \end{array}\right]. \tag{1}$$

Proof: That $G = NM^{-1}$ follows from:

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

$$= \left[\begin{array}{cc|c} A & BF & B \\ 0 & A+BF & B \\ \hline C & DF & D \end{array}\right] \qquad \text{(cascade of two systems)}$$

$$= \left[\begin{array}{cc|c} A+BF & BF & B \\ 0 & A & 0 \\ \hline C+DF & DF & D \end{array}\right] \qquad \text{(by change of basis in the state-space)}$$

$$= \left[\begin{array}{c|c} A+BF & B \\ \hline C+DF & D \end{array}\right] \qquad \text{(deletion of uncontrollable part)}$$

$$= N.$$

That N and M are right-coprime will follow from (3) below.

QED

$$\begin{bmatrix} M \\ N \end{bmatrix} := \left[\begin{array}{c|c} A+BF & B \\ \hline F & I \\ C+DF & D \end{array}\right]. \tag{1a}$$

is also a realization of an *rcf* of G.

By duality, to get a left-coprime factorization of G, take H such that $A+HC$ is stable.

Lemma 1'. A stabilizing output injection H yields *lcf* $G=\tilde{M}^{-1}\tilde{N}$ where

$$\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} := \left[\begin{array}{c|cc} A+HC & H & B+HD \\ \hline C & I & D \end{array}\right]. \tag{2}$$

The next step is to specify U_0, V_0, $\tilde{U}_0$, $\tilde{V}_0$ to satisfy

$$\begin{bmatrix} \tilde{V}_0 & -\tilde{U}_0 \\ -\tilde{N} & \tilde{M} \end{bmatrix}\begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{3}$$

The idea behind the choice of these matrices is as follows. Using observer theory, find a controller K_0 achieving internal stability. Perform factorizations

$$K_0 = U_0 V_0^{-1} = \tilde{V}_0^{-1}\tilde{U}_0$$

analogous to the ones just performed on G. Then Lemma 3.1 implies that the left-hand side of (3) must be invertible in RH_∞. We shall see that, in fact, (3) is satisfied.

The equations for K_0 are

$$\dot{\hat{x}} = A\hat{x} + Bu + H(C\hat{x} + Du - y)$$

$$u = F\hat{x},$$

that is,

$$K_o := \left[\begin{array}{c|c} A + BF + HC + HDF & -H \\ \hline F & 0 \end{array}\right]. \tag{4}$$

Define

$$\hat{A} := A + BF + HC + HDF \ , \ \hat{B} := -H$$

$$\hat{C} := F \ , \ \hat{D} := 0$$

$$\hat{F} := C + DF \ , \ \hat{H} := -(B + HD).$$

Following (1) and (2), define

$$\begin{bmatrix} V_o \\ U_o \end{bmatrix} := \left[\begin{array}{c|c} \hat{A} + \hat{B}\hat{F} & \hat{B} \\ \hline \hat{F} & I \\ \hat{C} + \hat{D}\hat{F} & \hat{D} \end{array}\right] = \left[\begin{array}{c|c} A+BF & -H \\ \hline C+DF & I \\ F & 0 \end{array}\right] \tag{5}$$

$$\begin{bmatrix} \tilde{V}_o & \tilde{U}_o \end{bmatrix} := \left[\begin{array}{c|cc} \hat{A} + \hat{H}\hat{C} & \hat{H} & \hat{B} + \hat{H}\hat{D} \\ \hline \hat{C} & I & \hat{D} \end{array}\right] = \left[\begin{array}{c|cc} A+HC & -(B+HD) & -H \\ \hline F & I & 0 \end{array}\right]. \tag{6}$$

Using the above definitions we have that

$$\begin{bmatrix} M & U_o \\ N & V_o \end{bmatrix} = \left[\begin{array}{c|cc} A+BF & B & -H \\ \hline F & I & 0 \\ C+DF & D & I \end{array}\right]. \tag{8}$$

$$\begin{bmatrix} \tilde{V}_o & -\tilde{U}_o \\ -\tilde{N} & \tilde{M} \end{bmatrix} = \left[\begin{array}{c|cc} A+HC & -(B+HD) & H \\ \hline F & I & 0 \\ C & -D & I \end{array}\right]. \tag{28c}$$

and the following theorem holds.

Theorem 2.

Equation (3) is satisfied.

Proof: Verification of (3) is immediate using (7), (8), and the inversion formula for systems (prelim).

A realization of J is now immediate. Substitution of (1), (4), (5), and (6) into (3.13) leads after simplification to

$$J = \left[\begin{array}{c|cc} A+BF+HC+HDF & -H & B+HD \\ \hline F & 0 & I \\ -(C+DF) & I & -D \end{array}\right]. \tag{9}$$

Let's recap. We began with the (stabilizable and detectable) realization

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

We chose F and H so that $A+BF$ and $A+HC$ were stable. Define J by (9). Then the proper K's achieving internal stability are precisely those representable as in Figure 5, where $Q \in RH_\infty$ and

$$I + DQ(\infty) \text{ is invertible}$$

(The last condition is equivalent to the one

$$(I + V_0^{-1}NQ)(\infty) \quad \text{is invertible}$$

which is required as per Theorem 1).

This representation result has an interesting interpretation : every internal stabilization amounts to adding stable dynamics to the plant and then stabilizing the extended plant by means of an observer. The precise statement is as follows; for simplicity of the formulas, only the case of strictly proper G and K is treated.

Theorem 2.

Assume G and K are strictly proper and the system in Figure 3 is internally stable. Then G can be embedded in a system

$$\left[\begin{array}{c|c} A_e & B_e \\ \hline C_e & 0 \end{array}\right],$$

where

$$A_e = \begin{bmatrix} A & 0 \\ 0 & A_2 \end{bmatrix}, \quad B_e = \begin{bmatrix} B \\ 0 \end{bmatrix}, \quad C_e = \begin{bmatrix} C & 0 \end{bmatrix} \tag{10}$$

and A_a is stable, such that K has the form

$$K = \left[\begin{array}{c|c} A_e + B_e F_e + H_e C_e & -H_e \\ \hline F_e & 0 \end{array}\right], \tag{11}$$

where $A_e + B_e F_e$ and $A_e + H_e C_e$ are stable.

Proof. K is representable as in Figure 5 for some Q in RH_∞. For K to be strictly proper, so must Q be (see (3.6)). Take a minimal realization of Q:

$$Q = \left[\begin{array}{c|c} A_a & B_a \\ \hline C_a & 0 \end{array}\right].$$

Since $Q \in RH_\infty$, A_a is stable. Let x and x_a denote state vectors for J and Q respectively, and write the equations for the system in Figure 5 :

$$\begin{aligned} \dot{x} &= (A + BF + HC)x - Hy + By_1 \\ u &= Fx + y_1 \\ u_1 &= -Cx + y \\ \dot{x}_a &= A_a x_a + B_a u_1 \\ y_1 &= C_a x_a \end{aligned}$$

These equations yield

$$\begin{aligned} \dot{x}_e &= (A_e + B_e F_e + H_e C_e)x_e - H_e y \\ u &= F_e x_e , \end{aligned}$$

where

$$x_e := \begin{bmatrix} x \\ x_a \end{bmatrix}, \quad F_e := \begin{bmatrix} F & C_a \end{bmatrix}, \quad H_e := \begin{bmatrix} H \\ -B_a \end{bmatrix}$$

and A_e, B_e, C_e are as in (10).

QED

2.2.5 Closed-Loop Transfer Matrix

Theorem 1 provides a parametrization, in terms of Q, of all proper K's which achieve internal stability in Figure 1. The goal in this section is to express the transfer matrix from v to e in terms of Q.

A stabilizing K is representable as in Figure 5. Substitution of the block diagram in Figure 5 into that in Figure 1 leads to the one in Figure 6. Elimination of the signals u and y leads to Figure 7 for a suitable transfer matrix T. Thus all closed-loop transfer matrices are representable as in Figure 7. It remains to give a realization of T.

We must first put back the original notation which was simplified at the end of Section 2. Let

$$P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] \tag{1}$$

be a minimal realization of P, and choose F and H so that $A+B_2F$ and $A+HC_2$ are stable.

Lemma 4.

$$T = \left[\begin{array}{cc|cc} A+B_2F & -B_2F & B_1 & B_2 \\ 0 & A+HC_2 & B_1+HD_{21} & 0 \\ \hline C_1-D_{12}F & -D_{12}F & D_{11} & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right] \tag{2}$$

Proof. With the original notation we have from (..4.9) that

$$J = \left[\begin{array}{c|cc} A+B_2F+HC_2+HD_{22}F & -H & B_2+HD_{22} \\ \hline F & 0 & I \\ -(C_2+D_{22}F) & I & -D_{22} \end{array}\right] \tag{3}$$

Partition J and T accordingly:

$$T_{12} = \left[\begin{array}{c|c} A+B_2F & B_2 \\ \hline C_1+D_{12}F & D_{12} \end{array}\right] \tag{5b}$$

$$T_{21} = \left[\begin{array}{c|c} A+HC_2 & B_1+HD_{21} \\ \hline C_2 & D_{21} \end{array}\right] \tag{5c}$$

$$T_{22} = 0.$$

In Figure 7 the governing equations are therefore

$$\begin{aligned} e &= T_{11}v + T_{12}y_1 \\ u_1 &= T_{21}v \\ y_1 &= Qu_1, \end{aligned}$$

so that

$$e = (T_{11} + T_{12}QT_{21})v.$$

In summary, we have

Theorem 3. The set of all closed-loop transfer matrices from v to e achievable by an internally stabilizing proper controller is equal to

$$\left\{T_{11} + T_{12}QT_{21} : Q \in RH_\infty, \quad I + D_{22}Q(\infty) \text{ invertible}\right\}.$$

The important points to note are that the closed-loop transfer matrix is simply an affine function of the controller parameter matrix Q and that the coefficient matrices T_{ij} have very simple realizations, namely, as in (5).

$$T_{12} = \left[\begin{array}{c|c} A+B_2F & B_2 \\ \hline C_1+D_{12}F & D_{12} \end{array}\right] \tag{5b}$$

$$T_{21} = \left[\begin{array}{c|c} A+HC_2 & B_1+HD_{21} \\ \hline C_2 & D_{21} \end{array}\right] \tag{5c}$$

$$T_{22} = 0.$$

In Figure 7 the governing equations are therefore

$$\begin{aligned} z &= T_{11}v + T_{12}y_1 \\ u_1 &= T_{21}v \\ y_1 &= Qu_1, \end{aligned}$$

so that

$$z = (T_{11} + T_{12}QT_{21})v.$$

In summary, we have

Theorem 3. The set of all closed-loop transfer matrices from v to z achievable by an internally stabilizing proper controller is equal to

$$\left\{T_{11} + T_{12}QT_{21} \; : \; Q \in RH_\infty, \quad I + D_{22}Q(\infty) \text{ invertible}\right\}.$$

The important points to note are that the closed-loop transfer matrix is simply an affine function of the controller parameter matrix Q and that the coefficient matrices T_{ij} have very simple realizations, namely, as in (5).

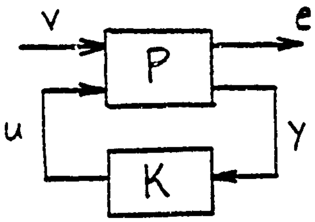


Figure 1

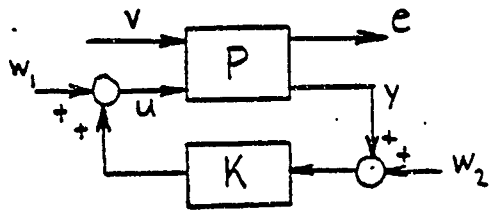


Figure 1a

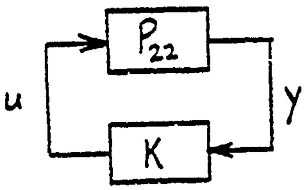


Figure 2

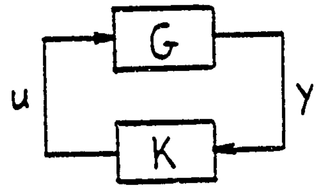


Figure 3

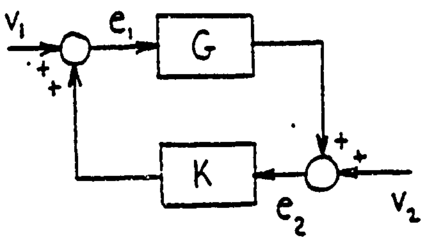


Figure 4

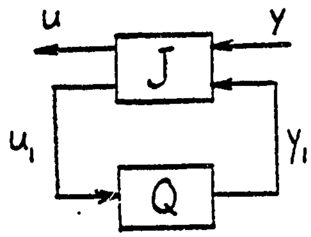


Figure 5

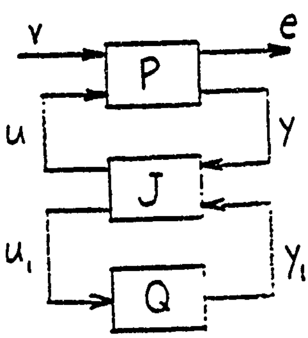


Figure 6

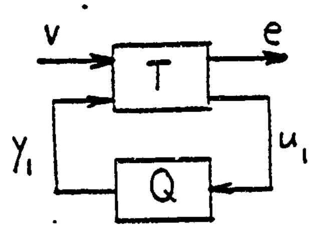


Figure 7

2.3.2 Riccati Equations and Factorizations

Consider the Algebraic Riccati Equation,

$$E'X + XE - XWX + Q = 0 \qquad \text{(ARE)}$$

where

$$E,\ W,\ Q \ \in\ \mathbb{R}^{n\times n},\ \ W = W' \geq 0 \ \text{and}\ Q = Q'$$

with the associated Hamiltonian matrix

$$A_H = \begin{bmatrix} E & -W \\ -Q & -E' \end{bmatrix} \qquad \text{(Hamiltonian)}$$

The following theorem and corollary characterizes the relationship between spectral factorization, Riccati equations, and decomposition of Hamiltonians.

Theorem 1 Let A,B,P,S,R be matrices of compatible dimensions such that $P=P'$, $R=R'>0$, with (A,B) stabilizable and (P,A) detectable. Then the following statements are equivalent.

a) The parahermitian rational matrix

$$\Gamma(s) = \begin{bmatrix} B'(-sI-A')^{-1} & I \end{bmatrix}\begin{bmatrix} P & S \\ S' & R \end{bmatrix}\begin{bmatrix} (sI-A)^{-1}B \\ I \end{bmatrix}$$

satisfies

$$\Gamma(j\omega) > 0 \quad \text{for all} \quad 0 \leq \omega \leq \infty$$

b) For $E=A-BR^{-1}S'$, $W=BR^{-1}B'$ and $Q=P-SR^{-1}S'$, there exists a unique real $X=X'$ such that

$$E'X + XE - XWX + Q = 0$$

and $E - BR^{-1}B'X$ is stable.

c) The Hamiltonian matrix

$$A_H = \begin{bmatrix} A-BR^{-1}S' & -BR^{-1}B' \\ -P+SR^{-1}S' & -(A-BR^{-1}S')' \end{bmatrix}$$

has no $j\omega$-axis eigenvalues.

Corollary 1 If the conditions in Theorem 1 are satisfied then $\exists M \in R_p$ such that $M^{-1} \in RH_\infty$ and

$$\Gamma = M^* R M$$

A particular realization of one such M is

$$M = \left[\begin{array}{c|c} A & B \\ \hline -F & I \end{array}\right]$$

where $F = -R^{-1}(S' + B'X)$.

Proof: (a)→(c) Let

$$\Gamma(s) = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] \triangleq \left[\begin{array}{cc|c} A & 0 & B \\ -P & -A' & -S \\ \hline S' & B' & R \end{array}\right] \tag{2}$$

Then $A_H = \hat{A} - \hat{B}\hat{D}^{-1}\hat{C}$. Suppose A_H has an eigenvalue on the $j\omega$-axis. Then $\exists \omega_o,\ x_o = (x_1', x_2')'$ such that

$$A_H x_o = j\omega_o x_o \tag{3}$$

or

$$(j\omega_o I - A)x_1 = -BR^{-1}(S'x_1 + B'x_2) \tag{4}$$

$$(j\omega_o I + A')x_2 = -(P - SR^{-1}S')x_1 + SR^{-1}B'x_2 \tag{5}$$

Suppose

$$0 = \hat{C}x_0 = S'x_1 + B'x_2 \tag{6}$$

then from (4) and (5)

$$(j\omega_0 I - A)x_1 = 0 \tag{7}$$

$$(j\omega_0 I + A')x_2 = -Px_1 \tag{8}$$

Since (7) implies $x_1^*(j\omega_0 I + A') = 0$, from (8) we have $x_1^* P x_1 = 0$. This implies, along with (7) that (P,A) is not detectable. Hence $\hat{C}x_0 \neq 0$. Now Lemma 0.2.3.8 implies that there exists $u_0 \neq 0$ such that $\Gamma(j\omega)u_0 = 0$. This contradicts the hypothesis that $\Gamma(j\omega) > 0$. Hence (a) $\rightarrow$ (c).

(b) $\rightarrow$ (a) Suppose $\exists X = X'$ such that $E - BR^{-1}B'X = A - BR^{-1}(S' + B'X)$ is stable. Let $F = -R^{-1}(S' + B'X)$ and

$$M = \left[\begin{array}{c|c} A & B \\ \hline -F & I \end{array}\right]$$

It is easily verified by use of the Riccati equation for X and routine algebra that $\Gamma = M^* R M$ so

$$\Gamma^{-1}(s) = M^{-1}(s) R^{-1} (M'(-s))^{-1}$$

Now

$$M^{-1} = \left[\begin{array}{c|c} A+BF & B \\ \hline F & I \end{array}\right]$$

So $M^{-1} \in RH_\infty$. Thus $\Gamma^{-1} \in RL_\infty$ and for all $0 \leq \omega \leq \infty$

$$\Gamma^{-1}(j\omega) = M^{-1}(j\omega) R (M'(-j\omega))^{-1} > 0$$

Hence $\Gamma(j\omega) > 0$ and (b) $\rightarrow$ (a).

(c) $\rightarrow$ (b) This is proven as part of Theorem 3.1 in the next section.

QED

The next section focuses on the solution of the Riccati equation and completes the proof of Theorem 1.

2.3.3 Solution of the Algebraic Riccati Equation :

Consider once again the Algebraic Riccati Equation,

$$E^T X + XE - XWX + Q = 0 \qquad \text{(ARE)}$$

where

$$E, W, Q \in \mathbb{R}^{n \times n}, \; W = W^T \geq 0 \text{ and } Q = Q^T$$

with the associated Hamiltonian matrix

$$A_H = \begin{bmatrix} E & -W \\ -Q & -E^T \end{bmatrix} \qquad \text{(Hamiltonian)}$$

Our main interest is to find the unique real symmetric stabilizing solution such that the matrix $(E - WX)$ is asymptotically stable. For simplicity we will use "solution" of the ARE to mean a real symmetric one. The ARE considered here is more general than the ARE which arises in linear quadratic optimal control and Kalman-Bucy filtering theory in that there is no assumption on the definiteness of the matrix Q.

An important property of the Hamiltonian matrix A_H is that the distribution of its eigenvalues (denoted as $\Lambda(A_H)$) is symmetric with respect to both the real and imaginary axes, i.e., if $\lambda \in \Lambda(A_H)$ with multiplicity k, so is $\bar{\lambda}$, $-\lambda$, and $-\bar{\lambda}$. Therefore, Λ can be partitioned as Λ_1 and Λ_2 so that $\lambda \in \Lambda_1$ with multiplicity k implies that $\bar{\lambda} \in \Lambda_1$ and $-\lambda, -\bar{\lambda} \in \Lambda_2$ all with the same multiplicity.

One connection between the ARE and A_H can be seen by assuming that X is a solution to ARE and conjugating A_H by $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}$ to yield

$$\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix} \begin{bmatrix} E & -W \\ -Q & -E^T \end{bmatrix} \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$$
$$= \begin{bmatrix} E-WX & -W \\ -(E^TX+XE-XWX+Q) & -(E-WX)^T \end{bmatrix}$$
$$= \begin{bmatrix} E-WX & -W \\ 0 & -(E-WX)^T \end{bmatrix}.$$

This puts A_H in block upper triangular form and clearly exhibits a particular partitioning of the eigenvalues of A_H with respect to the imaginary axis. For example, if $E-WX$ has all its eigenvalues in $\mathbb{C}_-$, then $-(E-WX)^T$ has all its poles in $\mathbb{C}_+$. Thus, the solution of ARE which stabilizes $E-WX$ yield a decomposition of A_H into stable and unstable parts.

This section will explore the conditions under which the desired solution of ARE exists. There is a considerable literature addressing the theory of ARE and it is not the purpose of this report to give a detailed treatment of this subject. We will simply present and prove the results which are relevant to the factorization theorems in this report.

Now, we are going to state the main theorem of this section which gives the necessary and sufficient conditions for the existence of a unique stabilizing solution of (ARE). Without loss of generality, we will assume that $W = GG^T$.

Theorem 1 :

The stabilizability of (E,G) and $\mathrm{Re}[\lambda_i(A_H)] \neq 0$ ($\forall$ $i = 1,2,\ldots\ldots,2n$) is necessary as well as sufficient for the existence of a unique stabilizing solution of (ARE).

Remark :

The unique stabilizing solution of Theorem 1 will be denoted by Ric (A_H). Note that this theorem is more general than Theorem 2.1 from the previous section since no detectability assumptions are made. The following theorem will play an important role in the next section in obtaining complementary inner factors.

Theorem 2 :

If $Q = H^T H \geq 0$ in (ARE) and X is its solution, then $\mathrm{Ker}(X) \subset \mathrm{Ker}(H)$.

The remainder of this section is devoted to proofs of Theorems 1 and 2.

Lemma 1 : (Potter, Martensson)

Let the columns of the matrix $\begin{bmatrix} Y \\ Z \end{bmatrix} \in \mathbb{R}^{2n \times n}$ ($Y, Z \in \mathbb{R}^{n \times n}$) be the eigenvectors or generalized eigenvectors of A_H corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then

(i) the matrix $Y^T Z$ is symmetric.

(ii) if Y^{-1} exists, then $X = ZY^{-1}$ is the solution of (ARE) such that the matrix $(E - WX)$ has the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

[Proof] :

The proof was first given by Potter (1966) and generalized by Martensson (1971).

Lemma 2 :

There exists at most one stabilizing solution to (ARE).

[Proof] :

Let X_1 and X_2 both be the stabilizing solutions to (ARE),

$$E^T X_1 + X_1 E - X_1 W X_1 + Q = 0 \tag{1}$$

$$E^T X_2 + X_2 E - X_2 W X_2 + Q = 0 \tag{2}$$

subtracting (2) from (1) yields

$$E^T(X_1 - X_2) + (X_1 - X_2)E - X_1 W X_1 + X_2 W X_2 = 0$$

which may be rewritten as

$$(E - WX_1)^T (X_1 - X_2) + (X_1 - X_2)(E - WX_2) = 0 \tag{3}$$

Since both X_1 and X_2 are stabilizing solutions, we have

$$\mathrm{Re}[\lambda_i(E - WX_1)] < 0 \qquad \forall\ i = 1,2,\ldots\ldots,n$$

and

$$\mathrm{Re}[\lambda_j(E - WX_2)] < 0 \qquad \forall\ j = 1,2,\ldots\ldots,n$$

From Property 1, we conclude that (3) has a unique solution

$$(X_1 - X_2) = 0\,, \quad \text{or} \quad X_1 = X_2.$$

Q.E.D.

Now, we are going to state the main theorem of this section which gives the necessary and sufficient conditions for the existence of a unique stabilizing solution of (ARE). Without loss of generality, we will assume that $W = GG^T$.

Theorem 1 :

The stabilizability of (E, G) and $\mathrm{Re}[\lambda_i(A_H)] \neq 0$ ($\forall\ i = 1,2,\ldots\ldots,2n$) is necessary as well as sufficient for the existence of a unique stabilizing solution of (ARE).

[Proof] :

(Sufficiency) :

Suppose (E, G) is stabilizable and $\mathrm{Re}[\lambda_i(A_H)] \neq 0$ for all i. Let the columns of the matrix $\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ $(Y_1, Y_2 \in \mathbb{R}^{n\times n})$ be the eigenvectors or generalized eigenvectors corresponding to n eigenvalues with negative real parts and J be the corresponding (real) Jordan block, i.e.,

$$A_H \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} J$$

or

$$EY_1 - GG^T Y_2 = Y_1 J \tag{4}$$

$$-QY_1 - E^T Y_2 = Y_2 J \tag{5}$$

We will first prove that the matrix Y_1 is nonsingular. Suppose Y_1 is singular and $\mathrm{Ker}(Y_1)$ denotes the null space of Y_1. Then $\forall\, v \in \mathrm{Ker}(Y_1)$, we have $Y_1 v = 0$ and

$v^T Y_2^T \times (4) \times v$:

$$\Rightarrow \qquad v^T Y_2^T E Y_1 v - v^T Y_2^T G G^T Y_2 v = v^T Y_2^T Y_1 J v$$

Since $Y_1^T Y_2$ is symmetric (from Lemma 2-1),

$$v^T Y_2^T Y_1 J v = v^T Y_1^T Y_2 J v = (Y_1 v)^T Y_2 J v = 0$$

$$\Rightarrow \qquad G^T Y_2 v = 0 \tag{6}$$

$(4) \times v$:

$$\Rightarrow \qquad EY_1 v - GG^T Y_2 v = Y_1 J v$$

$$\Rightarrow \qquad Y_1 J v = 0$$

$$\Rightarrow \qquad Jv \in \mathrm{Ker}(Y_1)$$

It is clear that $Ker(Y_1)$ is invariant under J and is spanned by some subset of the (generalized) eigenvectors of J. Therefore, $\exists$ eigen-pair $(\hat{\lambda}, \hat{v})$ of J such that $J\hat{v} = \hat{\lambda}\hat{v}$ and $\hat{v} \in Ker(Y_1)$. Then

(5) $\times\ \hat{v}$:

$$\Rightarrow \quad -QY_1\hat{v} - E^T Y_2\hat{v} = Y_2 J\hat{v}$$

$$\Rightarrow \quad -E^T Y_2\hat{v} = \hat{\lambda} Y_2\hat{v}$$

$$\Rightarrow \quad E^T(Y_2\hat{v}) = (-\hat{\lambda})(Y_2\hat{v}) \tag{7}$$

$$\Rightarrow \quad (-\hat{\lambda}) \text{ is an eigenvalue of } E^T (\text{or } E)$$

Furthermore, we know that $\mathrm{Re}(-\hat{\lambda}) > 0$. By the assumption of stabilizability, we conclude that $(-\hat{\lambda})$ must be controllable. Thus, from Theorem0.2.1.1(iii) (PBH rank tests), the matrix $[\ -\hat{\lambda}I - E \mid G\ ]$ must have full rank n.

But, from (6) and (7), we have

$$(Y_2\hat{v})^T [\ -\hat{\lambda}I - E \mid G\ ] = 0$$

This is a contradiction, and therefore, Y_1 must be nonsingular and, from Lemma 2-1, we know $X = Y_2 Y_1^{-1}$ is the stabilizing solution. Applying Lemma 2-2, the uniqueness of stabilizing solution is guaranteed.

(**Necessity**) :

If X is a unique stabilizing solution, then $(E - GG^T X)$ is asymptotically stable. This implies that (E, G) must be stabilizable. Because of the symmetry of $\Lambda(A_H)$ along the imaginary axis, we conclude that $Re[\lambda_i(A_H)] \neq 0$ for $i = 1, 2, \ldots\ldots, 2n$.

Q.E.D.

Corollary 1 :

If $Q \geq 0$, then the stabilizing solution $X \geq 0$.

[Proof] :

$$E^T X + XE - XWX + Q = 0$$

$$(E - WX)^T X + X(E - WX) = -(XWX + Q)$$

$(E - WX)$ is asympotically stable since X is the stabilizing solution. Thus, the solution of the above Lyapunov equation can be written as

$$X = \int_0^\infty e^{(E-WX)^T t}(XWX + Q)\, e^{(E-WX)t}\, dt$$

Since the matrix $(XWX + Q) \geq 0$, $X \geq 0$ is concluded.

Q.E.D.

Remark :

The proof of Theorem 1 was first given by Kučera in 1972 with the assumption $Q \geq 0$. In fact, this assumption is not necessary.

Theorem 2 :

If $Q = H^T H \geq 0$ in (ARE) and X is its solution, then Ker(X) $\subset$ Ker(H).

[Proof] :

Since X is a solution of (ARE), we have

$$E^T X + XE - XWX + H^T H = 0 \tag{8}$$

Let $u \in$ Ker(X) ($Xu = 0$). Then

$u^T \times (8) \times u$:

$\Rightarrow \quad u^T A^T Xu + u^T XAu - u^T XWXu + u^T H^T Hu = 0$

$\Rightarrow \quad u^T H^T H u = 0$

$\Rightarrow \quad Hu = 0$

$\Rightarrow \quad u \in \mathrm{Ker}(H)$

Hence, we conclude that $\mathrm{Ker}(X) \subset \mathrm{Ker}(H)$.

Q.E.D.

2.3.4 Inner-Outer and Spectral Factorization :

In this section, the special form of coprime factorizations required to reduce the general H_∞ optimal control problem to a best approximation problem will be developed. In particular, explicit realizations are given for coprime factorizations $G = NM^{-1}$ with inner numerator N (Theorem 1) and inner denominator M (Theorem 3); and for the complementary inner factor $N_\perp$ which completes the inner numerator to make $\left[N \; N_\perp\right]$ square and inner (Theorem 2). The theorems will be stated for right coprime factorizations (rcf) with the duals for lcf's following just as for the general case of coprime factorization developed earlier.

For the following theorems, it is assumed that $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in R_p^{p\times m}$ and the realization is minimal. We will denote by $R^{1/2}$ $(R \geq 0)$ the symmetric matrix such that $R^{1/2}R^{1/2} = R$ and use "$D_\perp$" for any orthogonal complement of D so that $\left[DR^{-1/2} \; D_\perp\right]$ (with $R = D'D$) is square and orthogonal.

Recall from Corollary 0.3.3.1 that $N = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right]$ is inner if and only if

i) $\hat{B}'X + \hat{D}'\hat{C} = 0$ (1)

ii) $\hat{D}'\hat{D} = I$ (2)

where the observability gramian X solves

$$\hat{A}'X + X\hat{A} + \hat{C}'\hat{C} = 0 \tag{3}$$

From Lemma 2.4.1 a stabilizing state feedback F yields rcf $G = NM^{-1}$ where

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A+BF & BZ \\ \hline F & Z \\ C+DF & DZ \end{array}\right]. \tag{4}$$

and Z can be any nonsingular matrix. To obtain a rcf with N inner, we simply need to use equations (1)-(4) to solve for F and Z. This yields the following theorem:

Theorem 1 :

Assume $p \geq m$. Then, there exists a *rcf* $G = NM^{-1}$ with N inner if and only if $G^*G > 0$ on the $j\omega$-axis, including at ∞. This factorization is unique up to a constant unitary multiple.

A particular realization for the factorization is

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A+BF & BR^{-1/2} \\ \hline F & R^{-1/2} \\ C+DF & DR^{-1/2} \end{array}\right] \in RH_\infty^{(m+p)\times m} \tag{5}$$

where

$$R = D'D > 0$$

$$F = -R^{-1}(B'X + D'C) \tag{6}$$

and

$$X = \mathrm{Ric}\begin{bmatrix} A-BR^{-1}D'C & -BR^{-1}B' \\ -C'D_\perp D_\perp' C & -(A-BR^{-1}D'C)' \end{bmatrix} \geq 0 \tag{7}$$

[Proof] :

(only if) :

Suppose $G = NM^{-1}$ is a *rcf* and $N^*N = I$. Then $G^*G = (NM^{-1})^*(NM^{-1}) = (M^{-1})^*M^{-1} > 0$ on the $j\omega$-axis since $M \in RH_\infty$.

(if) :

The if part will be proven by showing that (1)-(4) lead directly to the above realization of the *rcf* of G with inner numerator. That $G = NM^{-1}$ is an *rcf* follows immediately from (4) once it is established that F is a stabilizing state feedback. Using the notation

$$N = \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix} = \left[\begin{array}{c|c} A+BF & BZ \\ \hline C+DF & DZ \end{array}\right] \tag{8}$$

we will use (1)-(3) to get N inner and $A+BF$ stable. From (2) we have that $Z=R^{-1/2}U$ where $R=D'D>0$ and U is any orthogonal matrix. Take $U=I$. Equation (1) implies that

$$R^{-1/2}B'X + R^{-1/2}D'(C+DF) = 0$$

so solving for F yields

$$F = -R^{-1}(B'X + D'C) \tag{9}$$

Then equation (3) yields

$$\begin{aligned} 0 &= \hat{A}'X + X\hat{A}' + \hat{C}'\hat{C} \\ &= (A+BF)'X + X(A+BF) + (C+DF)'(C+DF) \\ &= (A-BR^{-1}D'C-BR^{-1}B'X)'X + X(A-BR^{-1}D'C-BR^{-1}B'X) \\ &\qquad + (C-DR^{-1}B'X-DR^{-1}D'C)'(C-DR^{-1}B'X-DR^{-1}D'C) \\ &= (A-BR^{-1}D'C)'X + X(A-BR^{-1}D'C) - XBR^{-1}B'X + C'D_\perp D_\perp' C \end{aligned} \tag{10}$$

since $D_\perp D_\perp' = I - DR^{-1}D'$. Thus $X=\mathrm{Ric}\,[A_H]$, where

$$A_H = \begin{bmatrix} A-BR^{-1}D'C & -BR^{-1}B' \\ -C'D_\perp D_\perp' C & -(A-BR^{-1}D'C)' \end{bmatrix} \tag{11}$$

That $X=\mathrm{Ric}\,(A_H)$ exists such that $A+BF$ is stable follows from Theorem 2.1 as follows. Let

$$\Gamma = G^*G$$

$$= \begin{bmatrix} B'(-sI-A')^{-1} & I \end{bmatrix} \begin{bmatrix} C'C & C'D \\ D'C & D'D \end{bmatrix} \begin{bmatrix} (sI-A)^{-1}B \\ I \end{bmatrix} \tag{12}$$

That $\Gamma = G^*G > 0$ is true by assumption. To satisfy Theorem 2.1 we must have $(A\ ,\ B)$ stabilizable and $(P\ ,\ A)$ detectable, but this is immediate since the realization for G was assumed minimal. Thus, Theorem 2.1 ensures that $X = \mathrm{Ric}\ (A_H)$ exists such that $A+BF$ is stable.

The uniqueness of the factorization follows from coprimeness and N inner. Suppose that $G = N_1M_1^{-1} = N_2M_2^{-1}$ are two right coprime factorizations and that both numerators are inner. By coprimeness, these two factorizations are unique up to a right multiple which is a unit in $RH_\infty^{m \times m}$. That is, there exists a unit $\Theta \in RH_\infty^{m \times m}$, such that $\begin{bmatrix} M_1 \\ N_1 \end{bmatrix} \Theta = \begin{bmatrix} M_2 \\ N_2 \end{bmatrix}$. Clearly, Θ is inner since $\Theta^*\Theta = \Theta^* N_1^* N_1 \Theta = N_2^* N_2 = I$. The only inner units in RH_∞ are constant matrices, and thus the desired uniqueness property is established. Note that the nonuniqueness is contained entirely in the choice of a particular square root of R.

Q.E.D.

In a similar manner equations (1)-(3) can be used to obtain the complementary inner factor (CIF) in the following theorem.

Theorem 2 :

If $p > m$ in Theorem 1, then there exists a CIF $N_\perp \in RH_\infty^{p \times (p-m)}$ such that the matrix $\begin{bmatrix} N & N_\perp \end{bmatrix}$ is square and inner. A particular realization is $N_\perp = \begin{bmatrix} A+BF & -X^\dagger C'D_\perp \\ C+DF & D_\perp \end{bmatrix}$ where X and F are from Theorem 1 and $X^\dagger$ is the pseudo-inverse of X.

[Proof] :

The proof consists of verifying directly that $\begin{bmatrix} N & N_\perp \end{bmatrix}$ is inner using the above realization for $N_\perp$ and the realization for N from Theorem 1. Using the notation

$$\begin{bmatrix} N & N_\perp \end{bmatrix} = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] = \left[\begin{array}{c|cc} A+BF & BR^{-1/2} & -X^{\dagger}C'D_\perp \\ \hline C+DF & DR^{-1/2} & D_\perp \end{array}\right] \tag{13}$$

and the fact that $\mathrm{Ker}(X) \subset \mathrm{Ker}(D_\perp' C)$ (Theorem 3.2), equations (1)-(3) follow immediately. Thus $\begin{bmatrix} N & N_\perp \end{bmatrix}$ is inner.

Theorem 3 :

There exists a *rcf* $G = NM^{-1}$ such that $M \in RH_\infty^{m \times m}$ is inner if and only if G has no poles on the $j\omega$-axis. A particular realization is

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A+BF & B \\ \hline F & I \\ C+DF & D \end{array}\right] \in RH_\infty^{(m+p)\times m} \tag{14}$$

where

$$F = -B'X \tag{15}$$

and

$$X = \mathrm{Ric} \begin{bmatrix} A & -BB' \\ 0 & -A' \end{bmatrix} \geq 0 \tag{16}$$

[Proof] :

The proof is essentially the same as for Theorem 1. The details are straightforward and are omitted.

Remarks :

(1) The minimality condition in Theorem 3 can be weakened to (A,B) stabilizable and A has no eigenvalues on the $j\omega$-axis and the theorem still holds.

(2) If $G \in RH_\infty^{p\times m}$ in Theorem 1, then M is a unit in RH_∞ and M^{-1} is "outer". In this case, $G = N(M^{-1})$ is called "inner-outer factorization" (IOF).

(3) Dual results for all factorizations can be obtained when $p \le m$. In these factorizations, output injection using the dual Riccati solution replaces state feedback to obtain corresponding left factorizations.

In the following theorem, we may assume that $G(s)$ is stable without loss of generality. Any $G \in RL_\infty$ may be factored using the dual of Theorem 3 to obtain a stable numerator $\tilde{N}$ such that $\tilde{N}^*\tilde{N} = G^*G$.

Theorem 4 : (Spectral Factorization)

Assume $G(s) \in RH_\infty^{p\times m}$ and $\gamma > \|G(s)\|_\infty$. Then, there exists a $M \in RH_\infty^{m\times m}$ with stable inverse such that $M^*M = \gamma^2 I - G^*G$ with

$$M = \left[\begin{array}{c|c} A & B \\ \hline -R^{1/2}K_C & R^{1/2} \end{array}\right]$$

where

$$R = \gamma^2 I - D'D > 0$$

$$K_C = -R^{-1}(B'X - D'C)$$

$$X = \mathrm{Ric}\begin{bmatrix} A+BR^{-1}D'C & -BR^{-1}B' \\ C'(I+DR^{-1}D')C & -(A+BR^{-1}D'C)' \end{bmatrix}$$

[Proof] :

Let

$$\Gamma = \gamma^2 I - G^* G = \begin{bmatrix} B'(-sI-A')^{-1} & I \end{bmatrix} \begin{bmatrix} -C'C & -C'D \\ -D'C & R \end{bmatrix} \begin{bmatrix} (sI-A)^{-1}B \\ I \end{bmatrix}$$

where $R = \gamma^2 I - D'D$.

Since $\gamma > ||G||_\infty$, $\Gamma(j\omega)>0$. The minimality of the realization of $G(s)$ guarantees that (A, B) is controllable and $(-C'C, A)$ is observable. Thus, from Corollary 2.1, there exists $M(s) \in R_p^{m \times m}$ such that $\Gamma = M^* M$ and a particular realization is

$$M = \left[\begin{array}{c|c} A & B \\ \hline -R^{1/2}K_c & R^{1/2} \end{array}\right]$$

where

$$K_c = -R^{-1}(B'X - D'C)$$

and

$$X = \mathrm{Ric}\begin{bmatrix} A+BR^{-1}D'C & -BR^{-1}B' \\ C'(I+DR^{-1}D')C & -(A+BR^{-1}D'C)' \end{bmatrix}.$$

Since G is stable, we conclude that $M \in RH_\infty^{m \times m}$.

Q.E.D.

2.3.5 Parametrizing the Optimal Controller and the H_2 solution :

This section combines the results of Youla's parametrization and the coprime factorization to parameterize all stabilizing controllers in a way that is convenient for solving optimal L_2 and L_∞ control problems.

Let

$$P = \begin{bmatrix} P_{11} & P_{21} \\ P_{12} & P_{22} \end{bmatrix} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

Suppose that neither P_{12} nor P_{21} has transmission zeros on the $j\omega$-axis (including ∞) and, without loss of generality, $D_{12}{}^T D_{12} = I$ and $D_{21} D_{21}{}^T = I$. Under these assumptions, let $D_\perp = (D_{12})_\perp$ and $\tilde{D}_\perp = (D_{21})_\perp$ that is, $\begin{bmatrix} D_{12} & D_\perp \end{bmatrix}$ and $\begin{bmatrix} D_{21}{}^T & \tilde{D}_\perp{}^T \end{bmatrix}$ are orthogonal matrices. Then, factor P as before with F and H given as follows :

$$F = -(D_{12}{}^T C_1 + B_2{}^T X)$$

$$X = \text{Ric} \begin{bmatrix} A - B_2 D_{12}{}^T C_1 & -B_2 B_2{}^T \\ -C_1{}^T D_\perp D_\perp{}^T C_1 & -(A - B_2 D_{12}{}^T C_1)^T \end{bmatrix}$$

and

$$H = -(B_1 D_{21}{}^T + Y C_2{}^T)$$

$$Y = \text{Ric} \begin{bmatrix} (A - B_1 D_{21}{}^T C_2)^T & -C_2{}^T C_2 \\ -B_1 \tilde{D}_\perp{}^T \tilde{D}_\perp B_1{}^T & -(A - B_1 D_{21}{}^T C_2) \end{bmatrix}$$

Then, $N_{12}{}^* N_{12} = I$ and $\tilde{N}_{21} \tilde{N}_{21}{}^* = I$. Also, let $N_\perp$ and $\tilde{N}_\perp$ be CIF's so that

$$\begin{bmatrix} N_{12} & N_\perp \end{bmatrix} = \left[\begin{array}{c|cc} A + B_2 F & B_2 & -X^\dagger C_1{}^T D_\perp \\ \hline C_1 + D_{12} F & D_{12} & D_\perp \end{array}\right]$$

$$\begin{bmatrix} \tilde{N}_{21} \\ \tilde{N}_\perp \end{bmatrix} = \left[\begin{array}{c|c} A + HC_2 & B_1 + HD_{21} \\ \hline C_2 & D_{21} \\ -\tilde{D}_\perp B_1{}^T Y^\dagger & \tilde{D}_\perp \end{array}\right]$$

Letting

$$T = \begin{bmatrix} (N\tilde{V})_{11} & N_{12} \\ \tilde{N}_{21} & 0 \end{bmatrix} = \left[\begin{array}{cc|cc} A+B_2F & -B_2F & B_1 & B_2 \\ 0 & A+HC_2 & B_1+HD_{21} & 0 \\ \hline C_1+D_{12}F & -D_{12}F & D_{11} & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right]$$

reduces

$$\min_{K \in R_p} \left\{ \left\| F_l(P,K) \right\|_\alpha \,\middle|\, F_l(P,K) \text{ stable} \right\}$$

to

$$\min_{Q \in RH_\infty} \left\{ \left\| F_l(T,Q) \right\|_\alpha \right\}$$

$$= \min_{Q \in RH_\infty} \left\{ \left\| (N\tilde{V})_{11} + N_{12}Q\tilde{N}_{21} \right\|_\alpha \right\}$$

The optimal K may be recovered from Q.

Because both the $\|\cdot\|_2$ and $\|\cdot\|_\infty$ norms are unitary invariant, an alternative expression is possible. For any $Q \in RH_\alpha$ ($\alpha = 2, \infty$), we have

$$\left\| (N\tilde{V})_{11} + N_{12}Q\tilde{N}_{21} \right\|_\alpha$$

$$= \left\| \begin{bmatrix} N_{12} & N_\perp \end{bmatrix}^* \left((N\tilde{V})_{11} + N_{12}Q\tilde{N}_{21} \right) \begin{bmatrix} \tilde{N}_{21} \\ \tilde{N}_\perp \end{bmatrix}^* \right\|_\alpha$$

$$= \left\| \begin{bmatrix} N_{12}^*(N\tilde{V})_{11}\tilde{N}_{21}^* + Q & N_{12}^*(N\tilde{V})_{11}\tilde{N}_\perp^* \\ N_\perp^*(N\tilde{V})_{11}\tilde{N}_{21}^* & N_\perp^*(N\tilde{V})_{11}\tilde{N}_\perp^* \end{bmatrix} \right\|_\alpha$$

$$= \left\| \begin{bmatrix} G_{11}+Q & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \right\|_\alpha$$

where $G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \begin{bmatrix} N_{12}^* \\ N_\perp^* \end{bmatrix} (N\tilde{V})_{11} \begin{bmatrix} \tilde{N}_{21}^* & \tilde{N}_\perp^* \end{bmatrix}$

The $\alpha = 2$ case is particularly simple. Since

$$\left\|\begin{bmatrix} G_{11}+Q & G_{12} \\ G_{21} & G_{22} \end{bmatrix}\right\|_2 = \left(\left\|G_{11}+Q\right\|_2^2 + \left\|\begin{bmatrix} 0 & G_{12} \\ G_{21} & G_{22} \end{bmatrix}\right\|_2^2 \right)^{\frac{1}{2}}$$

the optimal Q is seen immediately to be

$$Q_{opt} = \left\{G_{11}\right\}_+$$

Unfortunately, the $\alpha = \infty$ case is more complicated and will be investigated in the next section. To obtain an explicit expression for Q_{opt}, we need to compute G.

Note that $(N\tilde{V})_{11} = \begin{bmatrix} N_{11} & N_{12} \end{bmatrix} \begin{bmatrix} \tilde{V}_{11} \\ \tilde{V}_{21} \end{bmatrix}$. It is convenient to compute

$$\begin{bmatrix} N_{12}^* \\ N_\perp^* \end{bmatrix} \begin{bmatrix} N_{11} & N_{12} \end{bmatrix} = \begin{bmatrix} N_{12}^* N_{11} & I \\ N_\perp^* & 0 \end{bmatrix}$$

and

$$\begin{bmatrix} \tilde{V}_{11} \\ \tilde{V}_{21} \end{bmatrix} \begin{bmatrix} \tilde{N}_{21}^* & \tilde{N}_\perp^* \end{bmatrix}$$

seperately.

Claim 1 :

$$\begin{bmatrix} N_{12}^* \\ N_\perp^* \end{bmatrix} \begin{bmatrix} N_{11} \end{bmatrix} = \left[\begin{array}{c|c} -(A+B_2F)^T & (C_1-D_{12}F)^T D_{11} + XB_1 \\ \hline -B_2^T & D_{12}^T D_{11} \\ D_\perp^T C_1 X^T & D_\perp^T D_{11} \end{array}\right]$$

[Proof] :

$$\begin{bmatrix} N_{12}^* \\ N_\perp^* \end{bmatrix} \begin{bmatrix} N_{11} \end{bmatrix} = \left[\begin{array}{cc|c} A+B_2F & 0 & B_1 \\ -(C_1+D_{12}F)^T(C_1+D_{12}F) & -(A+B_2F)^T & -(C_1+D_{12}F)^T D_{11} \\ \hline D_{12}^T(C_1+D_{12}F) & B_2^T & D_{12}^T D_{11} \\ D_\perp^T C_1 & -D_\perp^T C_1 X^T & D_\perp^T D_{11} \end{array}\right]$$

conjugating the states by $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}$, we have

$$\begin{bmatrix} N_{12}^* \\ N_\perp^* \end{bmatrix} \begin{bmatrix} N_{11} \end{bmatrix} = \left[\begin{array}{cc|c} A+B_2F & 0 & B_1 \\ 0 & -(A+B_2F)^T & -(C_1+D_{12}F)^T D_{11} - XB_1 \\ \hline D_{12}^T(C_1+D_{12}F)+B_2^T & B_2^T & D_{12}^T D_{11} \\ 0 & -D_\perp^T C_1 X^T & D_\perp^T D_{11} \end{array}\right]$$

Since $D_{12}^T(C_1+D_{12}F)+B_2^T X = D_{12}^T C_1 + B_2^T X + F = 0$, the claim is verified.

Q.E.D.

Claim 2 :

$$\begin{bmatrix} \tilde{V}_{11} \\ \tilde{V}_{21} \end{bmatrix} \begin{bmatrix} \tilde{N}_{21}^* & \tilde{N}_\perp^* \end{bmatrix} = \left[\begin{array}{c|cc} -(A+HC_2)^T & -C_2^T & Y^T B_1 \tilde{D}_\perp^T \\ \hline (B_1+HD_{21})^T & D_{21}^T & \tilde{D}_\perp^T \\ FY & 0 & 0 \end{array}\right]$$

[Proof] :

$$\begin{bmatrix} \tilde{V}_{11} \\ \tilde{V}_{21} \end{bmatrix} \begin{bmatrix} \tilde{N}_{21}^* & \tilde{N}_\perp^* \end{bmatrix} = \left[\begin{array}{cc|cc} A+HC_2 & -(B_1+HD_{21})^T(B_1+HD_{21}) & (B_1+HD_{21})D_{21}^T & B_1\tilde{D}_\perp^T \\ 0 & -(A+HC_2)^T & C_2^T & -Y^T B_1 \tilde{D}_\perp^T \\ \hline 0 & -(B_1+HD_{21})^T & D_{21}^T & \tilde{D}_\perp^T \\ F & 0 & 0 & 0 \end{array}\right]$$

conjugating by $\begin{bmatrix} I & Y \\ 0 & I \end{bmatrix}$,

$$\begin{bmatrix} \tilde{V}_{11} \\ \tilde{V}_{21} \end{bmatrix} \begin{bmatrix} \tilde{N}_{21}^* & \tilde{N}_\perp^* \end{bmatrix} = \left[\begin{array}{cc|cc} A+HC_2 & 0 & 0 & 0 \\ 0 & -(A+HC_2)^T & C_2^T & -Y^T B_1 \tilde{D}_\perp^T \\ \hline 0 & -(B_1+HD_{21})^T & D_{21}^T & \tilde{D}_\perp^T \\ F & FY & 0 & 0 \end{array}\right]$$

which verifies the claim.

Q.E.D.

Putting these results together yields

$$\begin{bmatrix} N_{12}^* \\ N_\perp^* \end{bmatrix} \begin{bmatrix} N_{11} \end{bmatrix} \begin{bmatrix} \tilde{V}_{11} \\ \tilde{V}_{21} \end{bmatrix} \begin{bmatrix} \tilde{N}_{21}^* & \tilde{N}_\perp^* \end{bmatrix}$$

$$
= \left[\begin{array}{c|c} -(A+B_2F)^T & (C_1+D_{12}F)^T D_{11} + XB_1 \\ \hline B_2^T & D_{12}^T D_{11} \\ D_\perp^T C_1 X^T & D_\perp^T D_{11} \end{array}\right] * \left[\begin{array}{c|cc} -(A+HC_2)^T & -C_2^T & Y^T B_1 \tilde{D}_\perp^T \\ \hline (B_1+HD_{21})^T & D_{21}^T & \tilde{D}_\perp^T \\ -FY & 0 & 0 \end{array}\right]
$$

Note that this is the cascade of two systems with all of their poles in $\mathbb{C}_+$. Thus, projection onto $H_2 + \mathbb{C}$ leaves only the constant term. Therefore, in the L_2 case :

Theorem :

$$Q_{opt} = D_{12}^T D_{11} D_{21}^T$$

2.4.2 Hankel Operators

Let $G(s)$ be a strictly proper transfer matrix which is analytic in Re $s \le 0$, i.e., totally unstable. The *Hankel operator* associated with G will be denoted by Γ_G and is defined as follows. Let

$$P_{H_2^\perp} : L_2(j\mathbb{R}) \to H_2(j\mathbb{R})^\perp \tag{1}$$

denote the orthogonal projection. Then

$$\begin{aligned} \Gamma_G &: H_2(j\mathbb{R}) \to H_2(j\mathbb{R})^\perp \\ \Gamma_G f &:= P_{H_2^\perp} G f. \end{aligned}$$

There is a corresponding operator in the time domain. Let $g(t)$ denote the inverse Laplace transform of $G(s)$ and let $P_{H_2^\perp}$ also denote the orthogonal projection

$$P_{H_2^\perp} : L_2(\mathbb{R}) \to H_2(\mathbb{R})^\perp \tag{2}$$

(Context distinguishes the two projections (1) and (2).) The time-domain Hankel operator is

$$\begin{aligned} \Gamma_g &: H_2(\mathbb{R}) \to H_2(\mathbb{R})^\perp \\ \Gamma_g f &:= P_{H_2^\perp}(g * f). \end{aligned}$$

Thus

$$(\Gamma_g f)(t) = \begin{cases} \int_0^\infty g(t-\tau) f(\tau) d\tau, & t < 0 \\ 0, & t \ge 0. \end{cases}$$

Since the Fourier transform establishes the isomorphisms

$$L_2(\mathbb{R}) \cong L_2(j\mathbb{R})$$
$$H_2(\mathbb{R}) \cong H_2(j\mathbb{R})$$
$$H_2(\mathbb{R})^{\perp} \cong H_2(j\mathbb{R})^{\perp}$$

we have that

$$\|\Gamma_G\| = \|\Gamma_g\|.$$

The norm of Γ_g can be computed by state-space methods starting from a minimal realization of G,

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right].$$

Let Y and X be the controllability and observability gramians,

$$Y := \int_{-\infty}^{0} e^{At} BB' e^{A't}\, dt\ , \quad X := \int_{-\infty}^{0} e^{A't} C'Ce^{At}\, dt\ .$$

Lemma 1. $\Gamma_g^*\Gamma_g$ and YX have the same nonzero eigenvalues. In particular

$$\|\Gamma_g\| = \left[\rho(YX)\right]^{\frac{1}{2}}.$$

Proof. For this proof only, drop the subscript g on Γ_g and define $\underline{g}(t) := g(-t)'$. The adjoint operator of Γ is

$$\Gamma^* : H_2(\mathbb{R})^{\perp} \to H_2(R)$$
$$\Gamma^* \;\; = P_{H_2^{\perp}}(\underline{g} * h).$$

Let $\sigma^2 \neq 0$ be an eigenvalue of $\Gamma^*\Gamma$ and let $f \in H_2$ be a corresponding eigenvector. Then

$$\Gamma^*\Gamma f = \sigma^2 f. \tag{3}$$

Define

$$h := \frac{1}{\sigma}\Gamma f$$

so that

$$\Gamma f = \sigma h \qquad (4a)$$

$$\Gamma^* h = \sigma f. \qquad (4b)$$

(Vectors (f, h) satisfying these equations form a Schmidt pair for Γ.) Since

$$g(t) = Ce^{At}B, \quad t<0$$

$$\underline{g}(t) = B'e^{-A't}C', \quad t>0$$

we get from (4) that

$$\int_0^\infty Ce^{A(t-\tau)}Bf(\tau)d\tau = \sigma h(t), \quad t<0$$

$$\int_{-\infty}^0 B'e^{A'(\tau-t)}C'h(\tau)d\tau = \sigma f(t), \quad t>0,$$

or, equivalently,

$$Ce^{At}v = \sigma h(t), \quad t<0 \qquad (5a)$$

$$B'e^{-A't}w = \sigma f(t), \quad t>0 \qquad (5b)$$

where

$$v := \int_0^\infty e^{-A\tau}Bf(\tau)d\tau$$

$$w := \int_{-\infty}^0 e^{A'\tau}C'h(\tau)d\tau.$$

Now premultiply (5a) by $e^{A't}C'$ and integrate from $-\infty$ to 0, and premultiply (5b) by $e^{-At}B$ and integrate from 0 to ∞. This yields

$$Xv = \sigma w$$

$$Yw = \sigma v.$$

Finally, we get

$$YXv = \sigma^2 v, \tag{6}$$

showing that σ^2 is an eigenvalue of YX. The reverse argument leads from (6) to (3).

QED

We shall be concerned with approximating G by a stable transfer function, i.e., one analytic in Re $s \geq 0$, where the approximation is with respect to the L_∞ norm. Here we establish only that the distance in $L_\infty(j\mathbb{R})$ from G to the nearest matrix in $H_\infty(j\mathbb{R})$ equals $\|\Gamma_g\|$.

Theorem 1.

$$\inf \left\{ \|G-Q\|_\infty : Q \in H_\infty(j\mathbb{R}) \right\} = \|\Gamma_g\| \tag{8}$$

and *the infimum is achieved.*

The remainder of this section is devoted to a proof of Theorem 1: only section 5 requires some of the material to be presented next.

The inequality

$$\inf \left\{ \|G-Q\|_\infty : Q \in H_\infty(j\mathbb{R}) \right\} \geq \|\Gamma_G\|$$

is easy to establish. Fix Q in $H_\infty(j\mathbb{R})$. Then

$$\begin{aligned} \|G-Q\|_\infty &= \sup \left\{ \|(G-Q)f\|_2 : f \in H_2(j\mathbb{R}),\ \|f\|_2 \leq 1 \right\} \\ &\geq \sup \left\{ \|P_{H_2^\perp}(G-Q)f\|_2 : f \in H_2(j\mathbb{R}),\ \|f\|_2 \leq 1 \right\} \\ &= \sup \left\{ \|P_{H_2^\perp}Gf\|_2 : f \in H_2(j\mathbb{R}),\ \|f\|_2 \leq 1 \right\} \end{aligned}$$

$$= \left\|\Gamma_G\right\|.$$

Take the infimum over Q.

It is convenient at this stage to bring in L_2 and H_2 with respect to the unit disk:

T: unit circle.

$L_2(T, \mathbb{C}^{m \times n})$: Hilbert space of matrix-valued functions on T, with inner product

$$<F,G> := \frac{1}{2\pi}\int_0^{2\pi} \text{trace}\left[F(e^{j\vartheta})^* G(e^{j\vartheta})\right] d\vartheta.$$

$H_2(T, \mathbb{C}^{m \times n})$: subspace of functions $F(z)$ analytic in $|z|<1$ and satisfying

$$\sup_{0 \le r < 1} \int_0^{2\pi} \text{trace}\left[F(re^{j\vartheta})^* F(re^{j\vartheta})\right] d\vartheta < \infty.$$

$L_\infty(T, \mathbb{C}^{m \times n})$: Banach space of (essentially) bounded matrix-valued functions, with norm

$$\left\|F\right\|_\infty := \operatorname{ess\,sup}_\vartheta \ \bar{\sigma}\left[F(e^{j\vartheta})\right].$$

$H_\infty(T, \mathbb{C}^{m \times n})$: subspace of functions analytic and bounded in $|z|<1$.

Map the right half-plane Re $s \ge 0$ onto the unit disk $|z| \le 1$ via

$$z = \frac{s-1}{s+1}, \quad s = \frac{1+z}{1-z}, \tag{9}$$

and define

$$\tilde{G}(z) := G(s)|_{s=\frac{1+z}{1-z}}. \tag{10}$$

Since G in analytic in Re $s \ge 0$, including the point at ∞, $\tilde{G}$ is analytic in

$|z| \geq 1$, i.e., $\tilde{G} \in L_\infty(T)$. Associated with $\tilde{G}$ is a Hankel operator, $\Gamma_{\tilde{G}}$, defined as follows. Again, let $P_{H_2^\perp}$ denote the orthogonal projection

$$P_{H_2^\perp} : L_2(T) \rightarrow H_2(T)^\perp.$$

Then

$$\begin{aligned} \Gamma_{\tilde{G}} &: H_2(T) \rightarrow H_2(T)^\perp \\ \Gamma_{\tilde{G}} f &= P_{H_2^\perp} \tilde{G} f. \end{aligned} \tag{11}$$

Lemma 2. $\left\|\Gamma_G\right\| = \left\|\Gamma_{\tilde{G}}\right\|$

Proof. Define the function

$$\psi(s) = \sqrt{2}/(s+1).$$

The relation between a point $j\omega$ on the imaginary axis and the corresponding point $e^{j\vartheta}$ on the unit circle is, from (9),

$$e^{j\vartheta} = \frac{j\omega-1}{j\omega+1}.$$

This yields

$$\begin{aligned} d\vartheta &= -\frac{2}{\omega^2+1} d\omega \\ &= -|\psi(j\omega)|^2 d\omega. \end{aligned}$$

This implies that the mapping

$$\tilde{f} \rightarrow \psi f : H_2(T) \rightarrow H_2(j\mathbb{R}),$$

where $\tilde{f}(z) = f(s)\big|_{s=\frac{1+z}{1-z}}$, is an isomorphism. Similiarly,

$$\tilde{f} \rightarrow \psi f : H_2(T)^\perp \rightarrow H_2(j\mathbb{R})^\perp$$

is an isomorphism; note that if $\tilde{f} \in H_2(T)^\perp$, then $\tilde{f}=0$ at $z=\infty$, so that $f=0$ at

$s = -1$, and hence ψf is analytic in Re $s < 0$.

The lemma now follows from the commutative diagram

$$H_2(T) \xrightarrow{\Gamma_{\tilde{G}}} H_2(T)^{\perp}$$

$$H_2(j\mathbb{R}) \xrightarrow{\Gamma_G} H_2(j\mathbb{R})^{\perp}$$

Q.E.D.

There is a matrix representation of the operator $\Gamma_{\tilde{G}}$. Let the power series expansion of $\tilde{G}$ be

$$\tilde{G}(z) = \sum_{i=-\infty}^{\infty} z^i G_i$$

(Actually the sum only ranges from $i = -\infty$ to $i = 0$.) In (11) let the power series expansion of f be

$$f(z) = \sum_{0}^{\infty} z^i f_i$$

and let that of $h := P_{H_2^{\perp}} \tilde{G} f$ be

$$h(z) = \sum_{-\infty}^{-1} z^i h_i .$$

Then (11) is equivalent to the equation

$$\begin{bmatrix} h_{-1} \\ h_{-2} \\ h_{-3} \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} G_{-1} & G_{-2} & \cdots \\ G_{-2} & G_{-3} & \cdots \\ G_{-3} & G_{-4} & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} . \tag{12}$$

The matix in (12) is the familiar Hankel matrix of the transfer matrix $\tilde{G}(z)$.

Proof of Theorem 1. In view of Lemma 2, it suffices to show there exists a matrix Q in $H_\infty(T)$ such that

$$\left\|\tilde{G}-Q\right\|_\infty = \left\|\Gamma_{\tilde{G}}\right\|. \tag{13}$$

Let the power series expansion of Q be

$$Q(z) = \sum_0^\infty z^i Q_i.$$

The left side of (13) equals the norm of the operator

$$f \to (\tilde{G}-Q)f : H_2(T) \to L_2(T).$$

The matrix representation of this operator is

$$\begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ G_1-Q_1 & G_0-Q_0 & \cdots \\ G_0-Q_0 & G_{-1} & \cdots \\ G_{-1} & G_{-2} & \cdots \\ G_{-2} & G_{-3} & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \end{bmatrix}. \tag{14}$$

The idea in the construction of a Q to satisfy (13) is to select $Q_0, Q_1, \ldots$ in turn to minimize the norm of (14). First, choose Q_0 to minimize

$$\left\|\begin{bmatrix} G_0-Q_0 & G_{-1} & \cdots \\ G_{-1} & G_{-2} & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \end{bmatrix}\right\|.$$

By Parrott's Theorem, the minimum equals the norm of the Hankel matrix in (12), i.e., $\|\Gamma_{\tilde{G}}\|$. Next, choose Q_1 to minimize

$$\left\| \begin{bmatrix} G_1 - Q_1 & G_0 - Q_0 & \cdots \\ G_0 - Q_0 & G_{-1} & \cdots \\ G_{-1} & G_{-2} & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdots \end{bmatrix} \right\|.$$

Again, the minimum equals $\left\| \Gamma \mathfrak{z} \right\|$. Continuing in this way gives a suitable Q.

Q.E.D.

2.4.4 Best Approximation

The transfer matrix $R(s)$ is real-rational, proper and anti-stable, i.e., analytic in Re $s \le 0$. The objective is to find $Q(s)$, real-rational, proper, and stable, such that $\|R-Q\|_\infty$ is minimum, i.e., equals $\|\Gamma\|$. The constant term of R can be absorbed into Q, so we can assume R is strictly proper. Furthermore, by adding rows or columns of zeros, we can assume R is square.

Let

$$R = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \qquad \text{(minimal)}$$

$$\sigma := \|\Gamma\|$$

$$Q = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] \qquad \text{(to be found)}$$

Then

$$\begin{aligned} G &:= R - Q \\ &= \left[\begin{array}{cc|c} A & 0 & B \\ 0 & \hat{A} & \hat{B} \\ \hline C & -\hat{C} & -\hat{D} \end{array}\right] \\ &=: \left[\begin{array}{c|c} A_e & B_e \\ \hline C_e & D_e \end{array}\right]. \end{aligned}$$

So Q is optimal if $\hat{A}$ is stable and $\|G\|_\infty = \sigma$.

Lemma 5. *Suppose* $(\exists P_e)$ s.t.

$$A_e P_e + P_e A_e' + B_e B_e' = 0 \tag{5}$$

$$C_e P_e + D_e B_e' = 0 \tag{6}$$

$$D_e D_e' = \sigma^2 I \tag{7}$$

$$\hat{P} \geq 0 \quad \text{where} \quad P_e = \begin{bmatrix} \cdots & \cdots \\ \cdots & \hat{P} \end{bmatrix}. \tag{8}$$

Suppose also that $(\hat{A},\hat{B})$ is stablizable. Then Q is optimal.

Proof. Lemma 4 together with (5) and (6) $\Rightarrow$

$$G(s)G(-s)' = \sigma^2 I$$

$$\Rightarrow \|G\|_\infty = \sigma. \text{ Also, (5) } \Rightarrow$$

$$\hat{A}\hat{P} + \hat{P}\hat{A}' + \hat{B}\hat{B}' = 0.$$

Thus Lemma 3 $\Rightarrow$ $\hat{A}$ stable.

Q.E.D.

Recap: objective is to construct $A,B,C,\hat{A},\hat{B},\hat{C},\hat{D},P_e$ s.t.

i) $R = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ (minimal)

ii) $A_e P_e + P_e A_e' + B_e B_e' = 0$ (5)

iii) $C_e P_e + D_e B_e' = 0$ (6)

iv) $D_e D_e' = \sigma^2 I$ (7)

v) $\hat{P} \geq 0$ (8)

vi) $(\hat{A},\hat{B})$ stabilizable (9)

Construction.

Step 1. Find a balanced realization of R:

$$R = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right].$$

Thus controllability gramian

= observability gramian

$$= -\begin{bmatrix} \sigma I_r & 0 \\ 0 & \Sigma \end{bmatrix} =: P$$

where $\sigma > \left\|\Sigma\right\|$, i.e., r = multiplicity of σ. Partition A, B, C accordingly

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}$$

Step 2. Choose $\hat{D}$ s.t.

$$\hat{D}B_1' + \sigma C_1 = 0 \tag{10}$$

$$\hat{D}\hat{D}' = \sigma^2 I. \tag{11}$$

Step 3. Set

$$\hat{B} = -(\sigma^2 I - \Sigma^2)^{-1}(\Sigma B_2 + \sigma C_2'\hat{D}) \tag{12}$$

$$\hat{A} = (-A_{22} + B_2\hat{B}')' \tag{13}$$

$$\hat{C} = C_2\Sigma + \hat{D}B_2' \tag{14}$$

$$\hat{P} = \Sigma(\sigma^2 I - \Sigma^2)^{-1} \tag{15}$$

$$P_e = \begin{bmatrix} -\sigma I & 0 & 0 \\ 0 & -\Sigma & -I \\ 0 & -I & \hat{P} \end{bmatrix} \tag{16}$$

Verification of (5) - (9).

From Step 1 we have that

$$AP + PA' + BB' = 0$$

$$A'P + PA + C'C = 0.$$

These yield the following equations:

$$-\sigma A_{11} - \sigma A_{11}' + B_1B_1' = 0 \tag{17}$$

$$-A_{12}\Sigma - \sigma A_{21}' + B_1B_2' = 0 \tag{18}$$

$$-A_{22}\Sigma - \Sigma A_{22}' - B_2B_2' = 0 \tag{19}$$

$$-\sigma A_{11}' - \sigma A_{11} + C_1'C_1 = 0 \tag{20}$$

$$-A_{21}'\Sigma - \sigma A_{12} + C_1'C_2 = 0 \tag{21}$$

$$-A_{22}'\Sigma - \Sigma A_{22} + C_2'C_2 = 0. \tag{22}$$

To see that Step 2 is possible, i.e., $\hat{D}$ exists, observe from (17) and (20) that

$$B_1B_1' = C_1'C_1.$$

Hence there exists a unitary matrix U such that

$$UB_1' + C_1 = 0.$$

Take $\hat{D} = \sigma U$.

To verify (5), it suffices to show that the blocks in positions (1,3), (2,3), and (3,3) of the 3x3 block matrix $A_e P_e + P_e A_e' + B_e B_e'$ are all zero.

The (1,3)-block equals

$$\begin{aligned}
-A_{12} + B_1\hat{B}' &= -A_{12} + B_1(B_2'\Sigma + \sigma\hat{D}C_2)(\Sigma^2 - \sigma^2 I)^{-1} && \text{from (12)}\\
&= -A_{12} + (B_1B_2'\Sigma - \sigma C_1'C_2)(\Sigma^2 - \sigma^2 I)^{-1} && \text{from (10)}\\
&= -A_{12} + && \text{from (18);(21)}\\
&\quad \left[(A_{12}\Sigma + \sigma A_{21}')\Sigma - \sigma(A_{21}'\Sigma + \sigma A_{12})\right](\Sigma^2 - \sigma^2 I) && \text{from (18);(21)}\\
&= 0.
\end{aligned}$$

The (2,3)-block equals zero immediately from (13). Finally, the (3,3)-block equals:

$$\begin{aligned}
\hat{A}\hat{P} + \hat{P}\hat{A}' + \hat{B}\hat{B}' &= (-A_{22}' + \hat{B}B_2')\hat{P} + \hat{P}(-A_{22} + B_2\hat{B}') + \hat{B}\hat{B}' && \text{from (13)}\\
&= -A_{22}'\Sigma(\sigma^2 I - \Sigma^2)^{-1} - (\sigma^2 I - \Sigma^2)^{-1}\Sigma A_{22} && \text{from (12);(15)}\\
&\quad -(\sigma^2 I - \Sigma^2)^{-1}(\Sigma B_2 + \sigma C_2'\hat{D})B'\Sigma(\sigma^2 - \Sigma^2)^{-1}\\
&\quad -(\sigma^2 I - \Sigma^2)^{-1}\Sigma B_2(B_2'\Sigma + \sigma\hat{D}'C_2)(\sigma^2 I - \Sigma^2)^{-1}
\end{aligned}$$

$$+(\sigma^2 I-\Sigma^2)^{-1}(\Sigma B_2+\sigma C_2'\hat{D})(B_2'\Sigma+\sigma\hat{D}'C_2)(\sigma^2 I-\Sigma^2)$$

$$= (\sigma^2 I-\Sigma^2)^{-1}(-\sigma^2 A_{22}'\Sigma+\Sigma^2 A_{22}'\Sigma-\Sigma B_2 B_2'\Sigma$$

$$-\sigma^2\Sigma A_{22}+\Sigma A_{22}\Sigma^2+\sigma^2 C_2'C_2)(\sigma^2 I-\Sigma^2)^{-1}$$

$$= 0. \qquad \text{from (19);(22).}$$

Next is the verification of (6). We have

$$C_e P_e + D_e B_e' = \left[-\hat{D}B_1'-\sigma C_1,\ -\hat{D}B_2'-C_2\Sigma+\hat{C},\ -\hat{D}\hat{B}'-C_2-\hat{C}\hat{P}\right]$$

$$= \left[0,\ 0,\ -\hat{D}\hat{B}'-C_2-\hat{C}\hat{P}\right] \qquad \text{from (10) and (14).}$$

And

$$\hat{D}\hat{B}' + C_2 + \hat{C}\hat{P} = 0$$

by substitution from (12), (14), and (15).

Equation (7) is immediate from (11), and (8) follows from the definition (15) of $\hat{P}$.

It remains to prove (9). Suppose there exists λ, Re $\lambda \geq 0$, such that

$$\left[\hat{A}-\lambda I,\ \hat{B}\right]$$

doesn't have independent rows, i.e.,

$$x^*\left[\hat{A}-\lambda I,\ \hat{B}\right] = 0$$

for some $x \neq 0$. From (13) we get

$$x^*(-A_{22}'+\hat{B}B_2') = \lambda x^*, \quad x^*\hat{B} = 0,$$

so that

$$(-A_{22})x = \bar{\lambda}x.$$

This implies that $(-A_{22})$ is unstable, which is not possible: stability of $(-A)$ implies that of $(-A_{22})$.

2.4.5 General Distance Formula

Consider the problem of minimizing

$$\left\| \begin{bmatrix} R_{11}-Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty$$

where

$$R := \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

is real-rational, strictly proper, and analytic in Re $s \le 0$, and Q is required to be real-rational, proper, and analytic in Re $s \ge 0$. This section contains a formula for the minimum in terms of the norm of a certain operator Γ. Note that the minimum is the distance

$$\text{dist}\,(R, \begin{bmatrix} RH_\infty & 0 \\ 0 & 0 \end{bmatrix})$$

from R to the set of all matrices of the form

$$\begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}, \quad Q \in RH_\infty.$$

The matrix R induces a multiplication operator on $L_2 \oplus L_2$, i.e.,

$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \to \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}.$$

Define Γ via

$$\Gamma : H_2 \oplus L_2 \to H_2^\perp \oplus L_2$$

$$\Gamma \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} P_{H_2^\perp} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}.$$

Theorem 1

$$\mathrm{dist}\,(R, \begin{bmatrix} RH_\infty & 0 \\ 0 & 0 \end{bmatrix}) = \|\Gamma\|$$

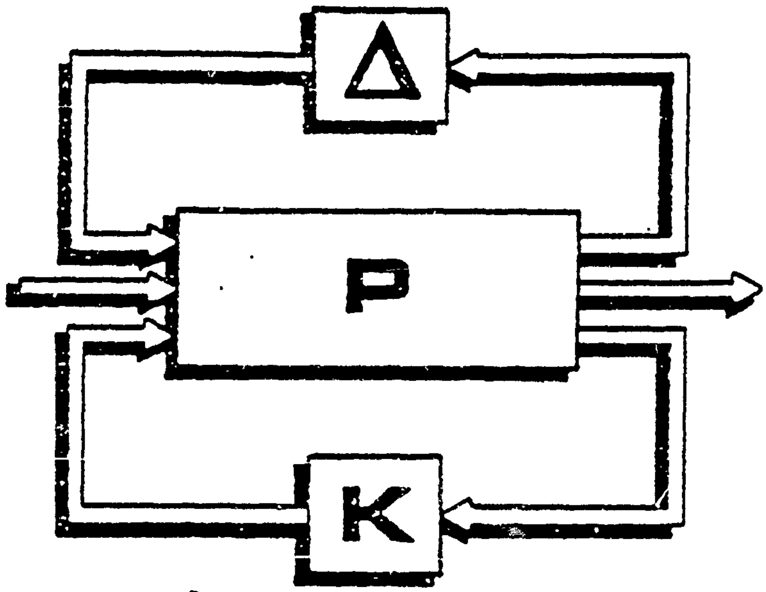
The proof of this theorem is a straightforward application of the theorem of Parrott / Davis - Kahan - Weinberger, as in Section 2. It is possible to write down a matrix representation of Γ, but an efficient numerical procedure for computing its norm has not yet been developed.

PART TWO:

STEIN'S NOTES

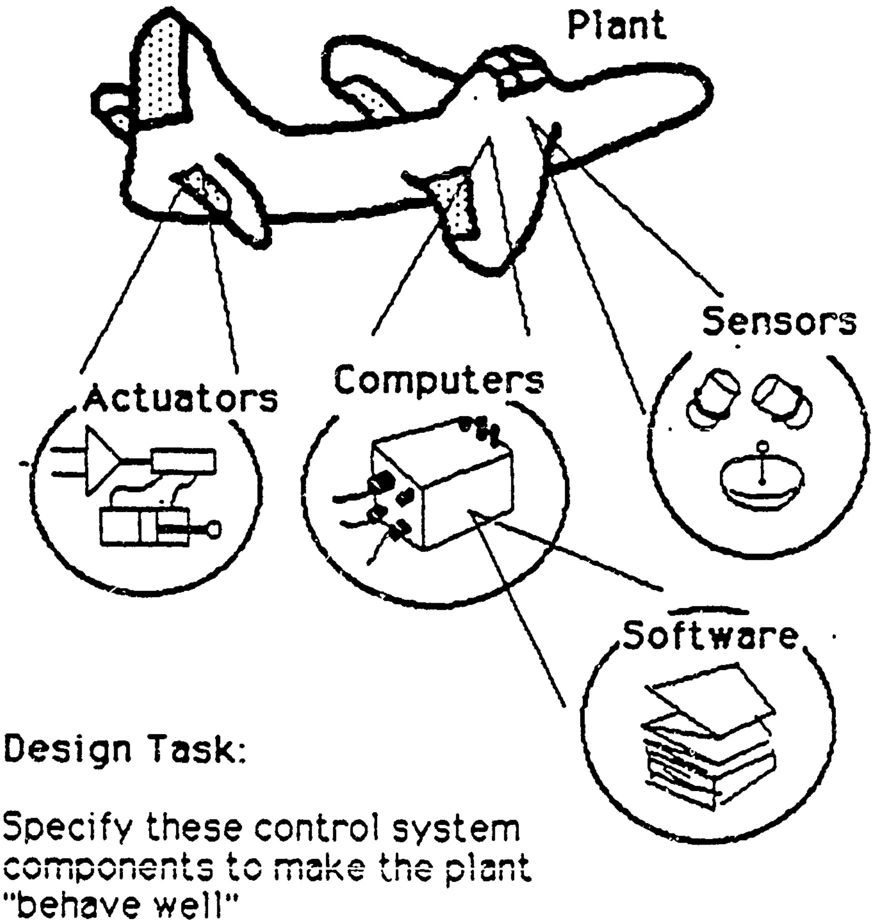
[illegible]

ONR/HONEYWELL WORKSHOP



Problem Description and Motivation

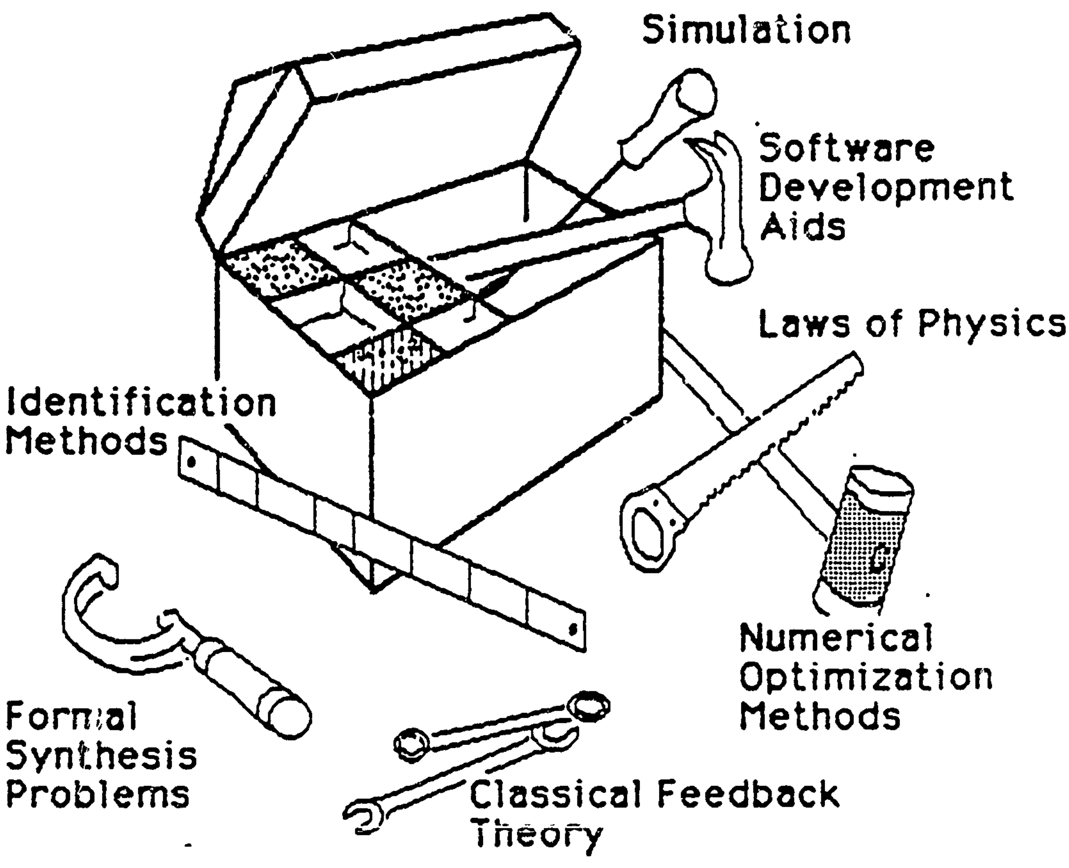
Real Control Design Problems



Design Task:

Specify these control system components to make the plant "behave well"

Control Designers Need Many Tools



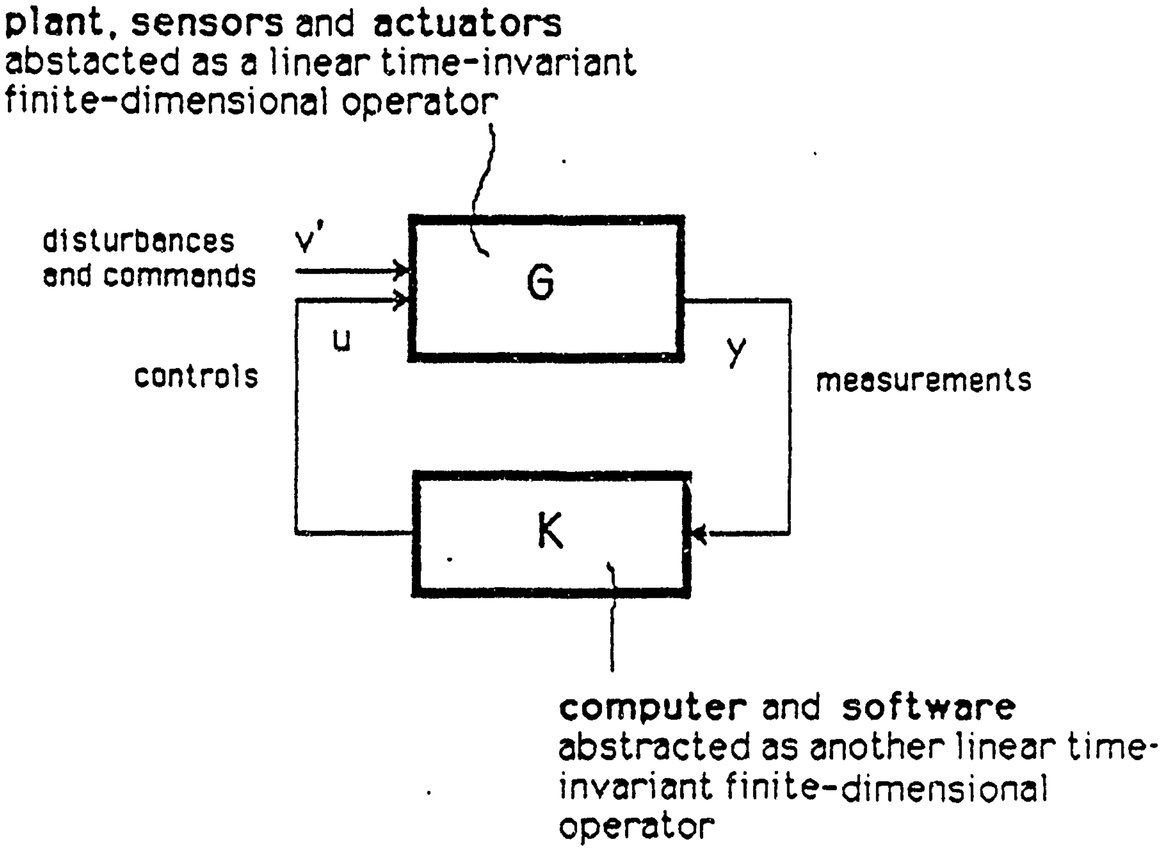
Formal Synthesis Problems

- Some Element of Design Abstracted as a Formal Mathematics Problem
- Solutions Known Mathematically and Computable Practically
- Properties of Solutions Well-Understood and Desirable

--The closer the abstracted problem matches the real design task, the more powerful it is as a design tool

--No abstracted problem matches the real thing perfectly

Our Formal Synthesis Problem

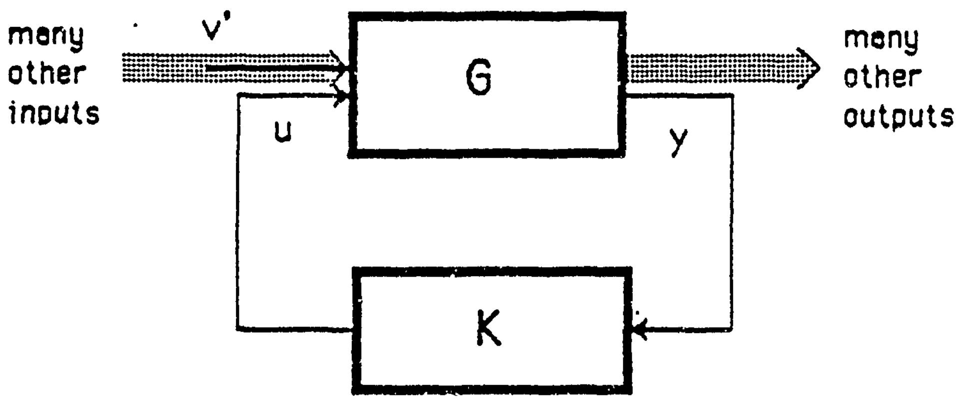


Design Task

Specify $K(y)$ to

- achieve stability
- optimize performance
- provide robustness

Stability



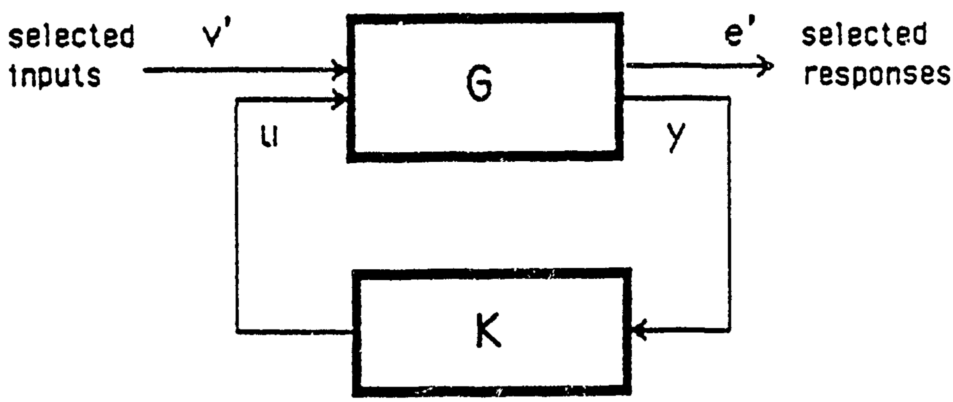
No bounded input should produce unbounded outputs

$\Longrightarrow$ we must achieve **internal stability**

(all poles in the left hand plane)

(stabilizable/detectable design models

Performance

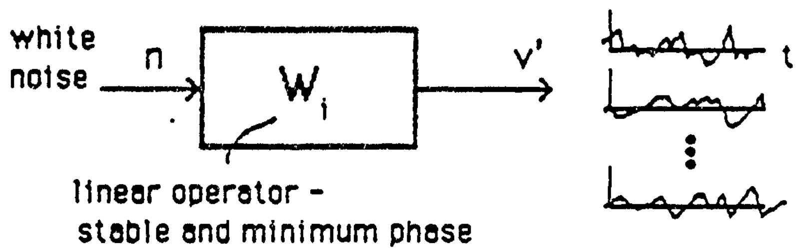


Inputs in a Specified Class should produce Responses with Specified Optimal Properties

Performance Objectives for A Familiar Example

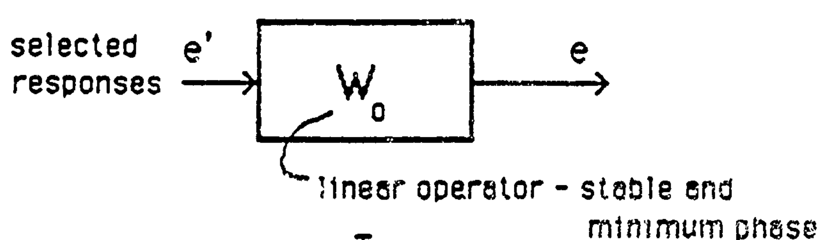
- Input Class

Sample Functions of a Stationary Random Process with Specified Autocorrelation



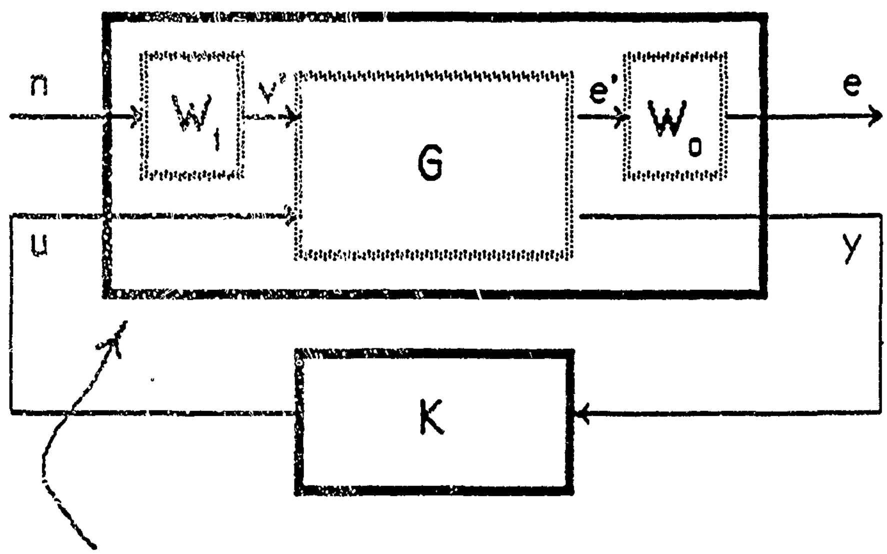
- Response Properties

Minimum Mean-Square-Error



$$\min E \{ e(t)^T e(t) \}$$

Resulting Design Model



Overall Plant:

$$\begin{bmatrix} e \\ y \end{bmatrix} = P(n,u) = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} n \\ u \end{bmatrix}$$

Closed Loop Responses:

$$\begin{aligned} e &= \left[P_{11} + P_{12}[I - KP_{22}]^{-1} KP_{21} \right] n \\ &= F(P,K)\, n \end{aligned}$$

Resulting "$\mathbb{H}_2$-Optimization Problem"

$$\boxed{\min_{\text{stabilizing } K} E\{ e(t)^T e(t) \}}$$

$$E\{e(t)^T e(t)\} = E\{ [F(P,K)n](t)^T [F(P,K)n](t) \}$$

$$= \mathrm{Tr}\{ \int_0^\infty F(t)\; F(t)^T\, dt \}$$

impulse response of operator F

$$= \mathrm{Tr}\{ \frac{1}{2\pi} \int_{-\infty}^\infty F(j\omega) F(j\omega)^* \, d\omega \}$$

frequency response of operator F

$$= \| F(s) \|_2^2$$

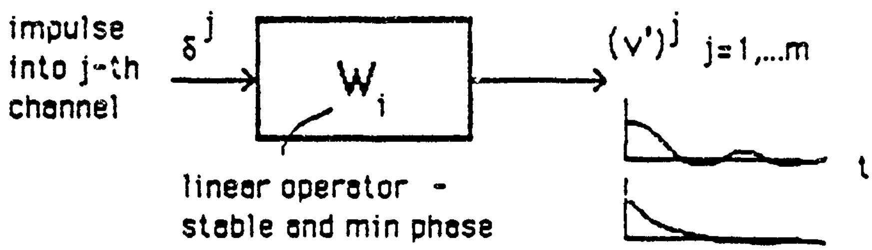
$$\| F(s) \|_2 \triangleq \sqrt{\mathrm{Tr}\, \frac{1}{2\pi} \int_{-\infty}^\infty F(j\omega) F(j\omega)^* \, d\omega}$$

= 2-NORM on the Hardy space of stable transfer functions

An Equivalent Deterministic Performance Objective

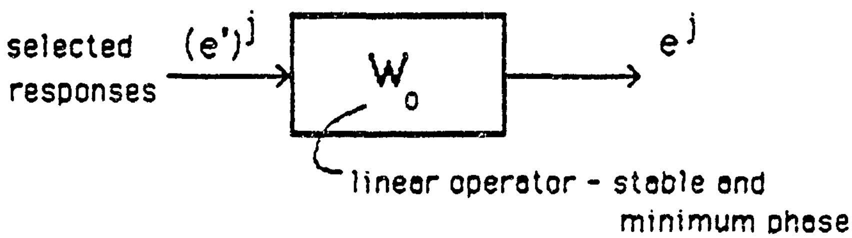
- **Input Class**

m specific time functions in $\mathbb{L}_2$ representable as impulse responses of linear systems



- **Response Properties**

Minimum $\mathbb{L}_2$-Norm of m corresponding weighted responses



$$\min \sum_{j=1}^{m} \| e^j(t) \|_2^2$$

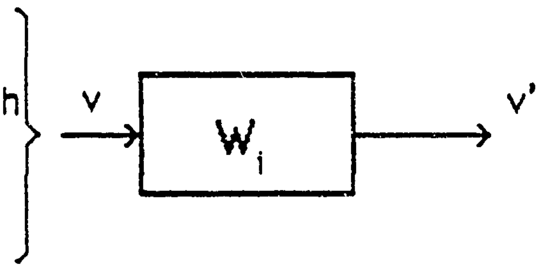
$$\| e(t) \|_2 \triangleq \sqrt{\int_0^\infty e(t)^T e(t)\, dt}$$

2-NORM on the space of time functions

Alternate Performance Objectives

- Input Class

All possible
$\mathbb{L}_2$-functions with
bounded norm
$\|v(t)\|_2 \leq 1$



- Response Properties

Minimum $\mathbb{L}_2$ Norm of the largest weighted response

selected
responses

$$\min \sup_{v} \|e(t)\|_2$$

Resulting " $\mathbb{H}_\infty$-Optimization Problem "

$$\boxed{\min_{\text{stabilizing } K} \; \sup_{v} \; \|e(t)\|_2}$$

$$\sup_{v} \|e(t)\|_2 = \sup_{v} \|F(P,K)\, v(t)\|_2$$

$$\triangleq \|F(P,K)\|_{i2}$$

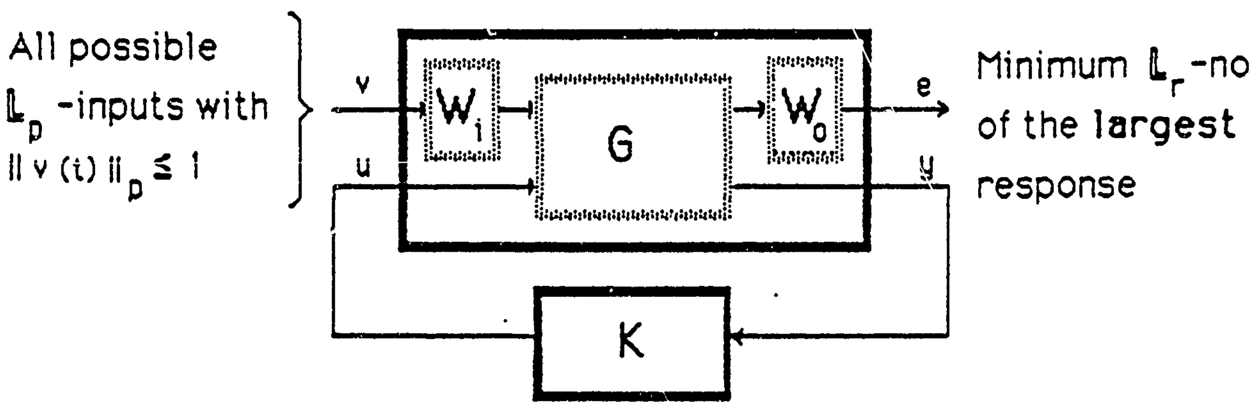
operator norm induced by the
2-norm on input/output time functions

$$= \sup_{\omega} \bar{\sigma}[\,F(j\omega)\,]$$

$$\triangleq \|F(s)\|_{\infty}$$

∞-Norm on the space of
stable transfer functions

More General Performance Objectives ?



- Trivial solutions (opt F=0) whenever $r > p$
- Little engineering precedent for

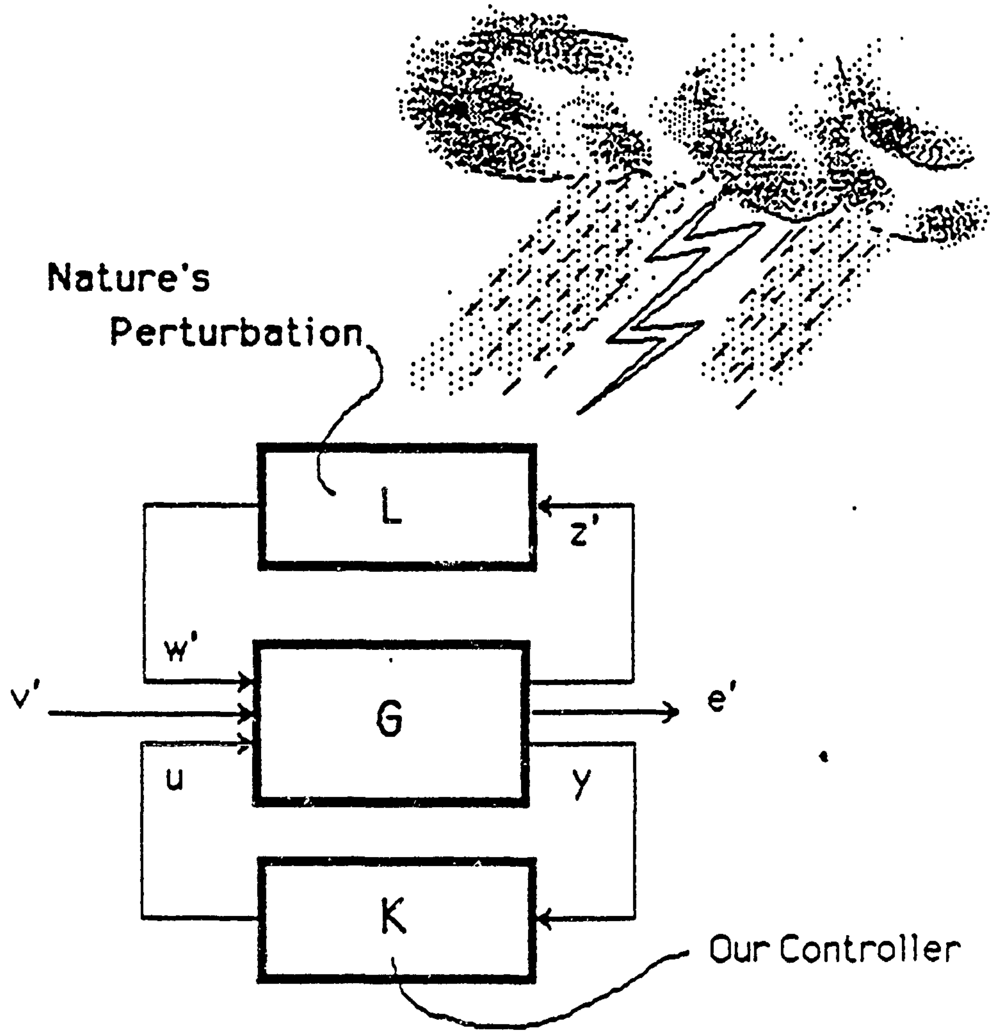
$$3 \leq r \leq p < \infty$$

- Bounded signal problems ($r=p=\infty$) indirectly covered by square-integrable problems

For SISO Operators (Doyle/Gahberg):

$$\|F\|_{i\infty} = \|F(t)\|_1 \leq \underbrace{2M}_{\text{Mc Millan Degree}} \|F\|_{i2}$$

Robustness



- Stability must be maintained in presence of L
- Some minimum performance level must be maintained in the presence of L

Some Philosophy About Perturbations

Nature's Perturbations

- are unknown,
- are potentially catastrophic for any control, and
- defy mathematical description .

Nevertheless, we must

- represent them by mathematical models ,
- specify maximum levels of severity within those models , and
- design controls to work successfully for this specification

The **"Leap of Faith"** that the selected representation will protect against the real thing remains the burden of engineers, not of mathematicians

A Model of "Unstructured" Perturbations

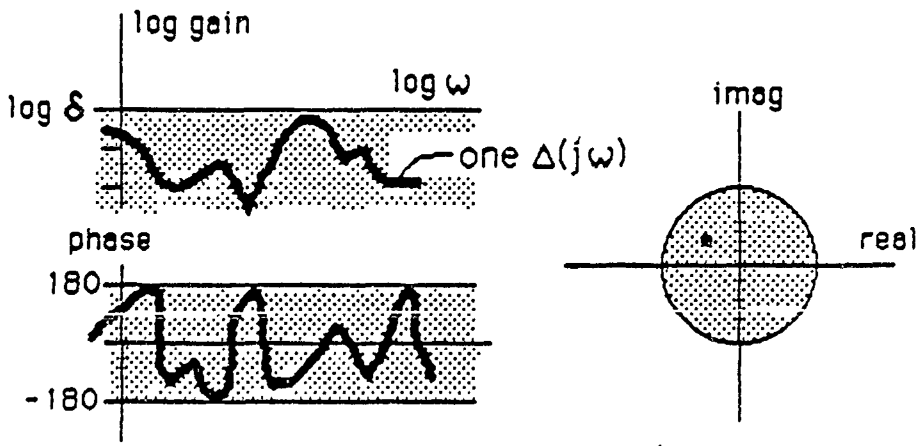
$$L(z') = W_o \, \Delta \, W_i \, (z')$$

stable minimum phase
LFDTI weighting operators

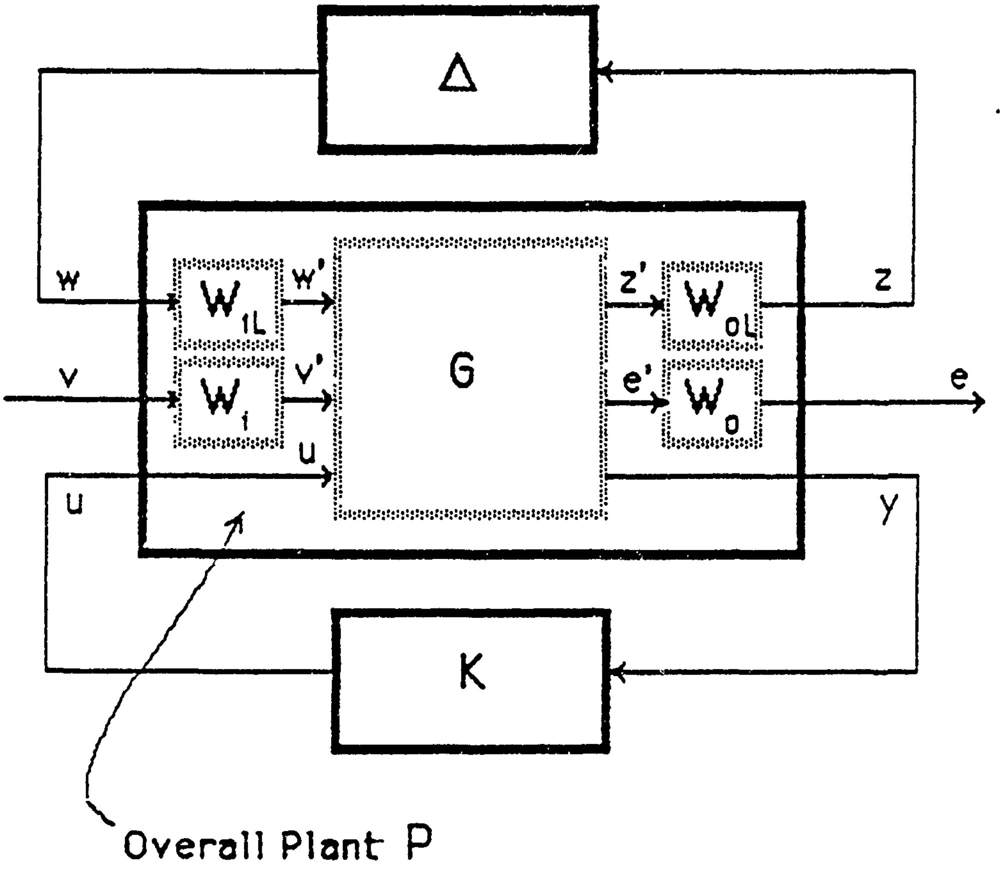
any LFDTI operator from a norm-bounded set

$$\Phi = \{ \Delta \mid \Delta \text{ stable}, \ \|\Delta\|_{i2} \le \delta \}$$

Transfer Functions in a Disk: $\bar{\sigma}[\Delta(j\omega)] \le \delta$

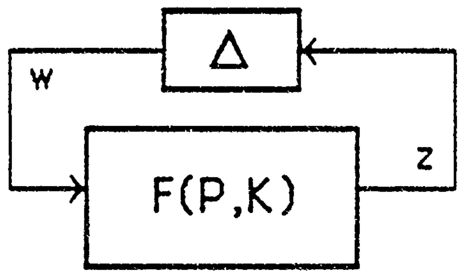


Design Model



Robust Stability

System S :



Theorem

Given : i. $F(P,K)$ stable

ii. $\Delta \in \mathcal{P}$ [stable with $\|\Delta\|_{i2} \leq \delta$]

Then system S is stable if and only if $\| F \|_{i2} < \frac{1}{\delta}$

Resulting Analysis Test for Robust Stability:

S remains stable for all $\Delta \in \mathcal{P}$

iff

$\bar{\sigma}[\, F(j\omega)\,] < \frac{1}{\delta}$ for all $\omega \leq \infty$

NOTE: This is a non-conservative test for stability with respect to the perturbation set $\mathcal{P}$

Other Stability Robustness Tests

Set of Perturbations	Corresponding Robustness Tes
• Individual scalar perturbations acting one at a time $\mathcal{P}_1 = \left\{ \Delta \;\middle|\; \begin{array}{l} \Delta = \mathrm{diag}(0,..0,\Delta_j,0,..0) \\ \Delta_j \text{ stable}, \;\; \|\Delta_j\| \le \delta \end{array} \right\}$	$\lvert F_{jj}(j\omega)\rvert < \frac{1}{\delta} \quad \omega \le$ for eac
• A single scalar perturbation acting simultaneously in each loop $\mathcal{P}_{1d} = \left\{ \Delta \;\middle|\; \begin{array}{l} \Delta = \mathrm{diag}(\Delta,\Delta,..\;\Delta) \\ \Delta \text{ stable}, \;\; \|\Delta\| \le \delta \end{array} \right\}$	$\rho\,[F(j\omega)] < \frac{1}{\delta} \quad \omega \le$
• A single multivariable perturbations $\mathcal{P} = \left\{ \Delta \;\middle|\; \Delta \text{ stable}, \;\; \|\Delta\| \le \delta \right\}$	$\bar{\sigma}\,[F(j\omega)] < \frac{1}{\delta} \quad \omega \le$
• m multivariable perturbations acting simultaneously $\mathcal{X} = \left\{ \Delta \;\middle|\; \begin{array}{l} \Delta = \underset{\text{diag}}{\text{block}}(\Delta_1, \Delta_2, ..\Delta_m) \\ \Delta_j \in \mathcal{P} \end{array} \right\}$	$\mu\,[F(j\omega)] < \frac{1}{\delta} \quad \omega \le$

Conservatism ?

- Tests are conservative (sufficient but not necessary) whenever

$$\left\{ \begin{array}{l} \text{set of true} \\ \text{perturbations} \end{array} \right\} \subset \left\{ \begin{array}{l} \text{set of modeled} \\ \text{perturbations} \end{array} \right\}$$

- Simple norm-bounded covering sets are often conservative

  Examples

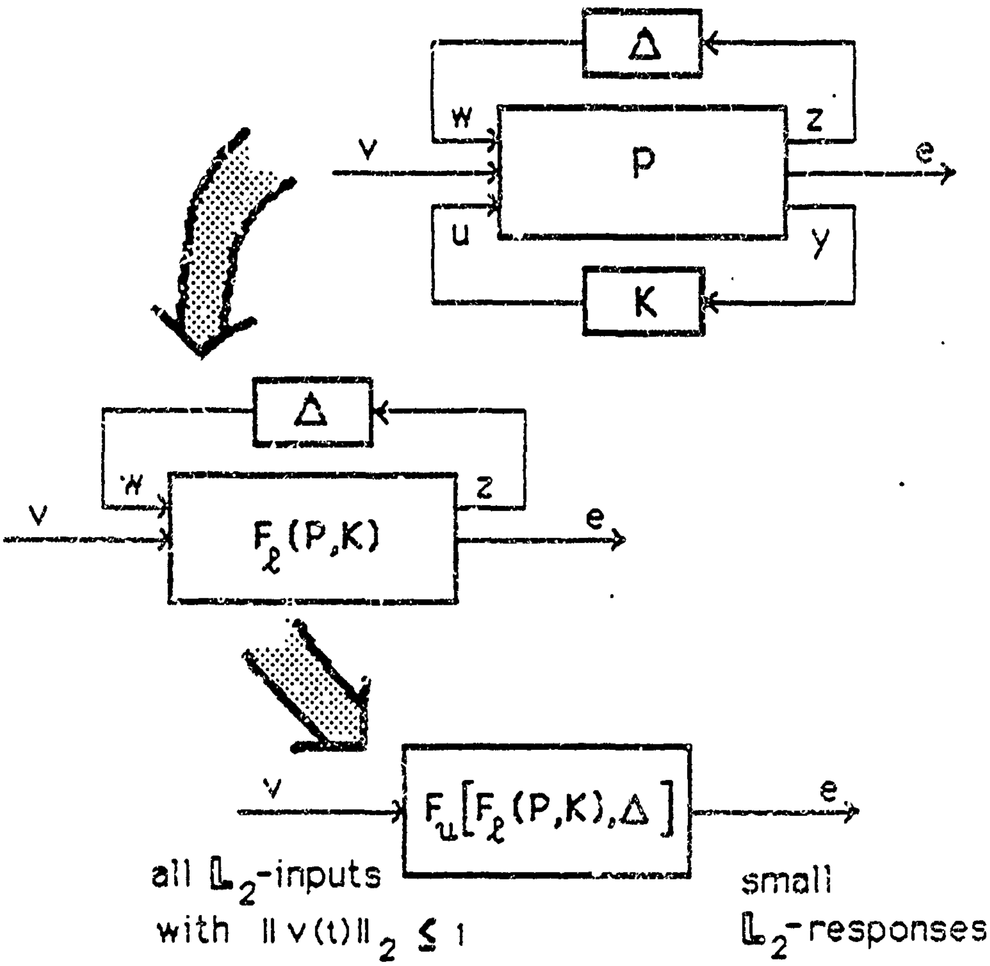
  - Real parameter variations
  - Deliberately neglected dynamics (time delays, high freq lags)
  - etc

- Of norm-bounded sets, induced 2-norm ($\mathbb{H}_\infty$) sets are least conservative

For SISO Operators : $\|M\|_{i2} \leq \|M\|_{ip}$

$$\therefore \quad \|F\|_{ip} < \frac{1}{\|\Delta\|_{ip}} \implies \|F\|_{i2} < \frac{1}{\|\Delta\|_{i2}}$$

Robust Performance



Robust Performance Specification:

$$\left\| F_u[F_\ell(P,K),\Delta] \right\|_{i,2} < \frac{1}{\delta} \quad \text{for all } \Delta \in \mathcal{P}$$

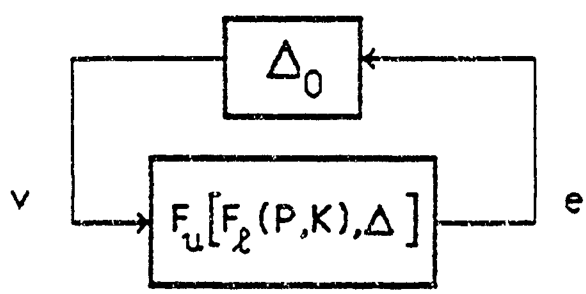
Robust Performance Specs Viewed as Stability Conditions

$\boxed{\text{Theorem}}$

$$\left\| F_u[F_\ell(P,K),\Delta] \right\|_{i2} < \frac{1}{\delta}$$

iff (1) $F_u[F_\ell(P,K),\Delta]$ is stable, and

(2) the following system is stable for all $\Delta_0 \in \mathcal{P}$

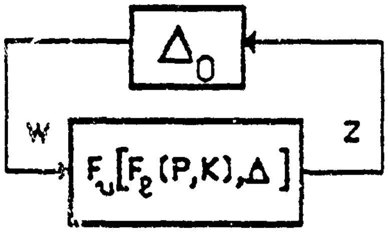


(stability robustness theorem used backwards)

Resulting Condition for Robust Stability and Robust Performance

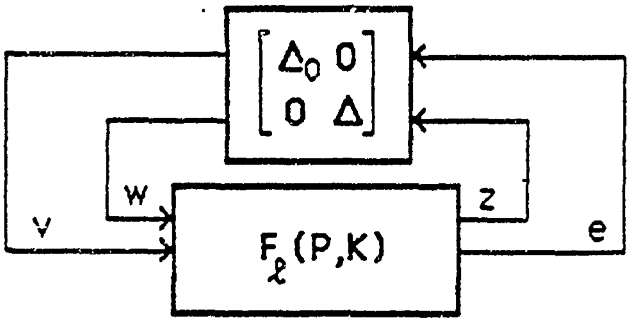
$F_u[F_\ell(P,K),\Delta]$ stable for all $\Delta \in \mathcal{P}$,

and



stable for all $\Delta_0 \in \mathcal{P}$

$$\Updownarrow$$

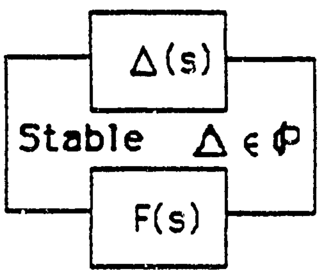


stable for all $\Delta_0, \Delta \in \mathcal{P}$

> Stability and performance robustness are achieved simultaneously if and only if our feedback system is stable for all perturbations with a particular 2x2 block-diagonal structure

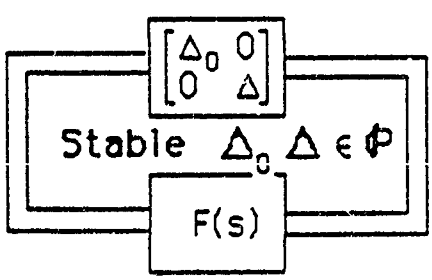
Once More for Emphasis . . . .

- Conditions for Robust Stability and Robust Performance are Equivalent in the induced 2-Norm ($\mathbb{H}_\infty$) .



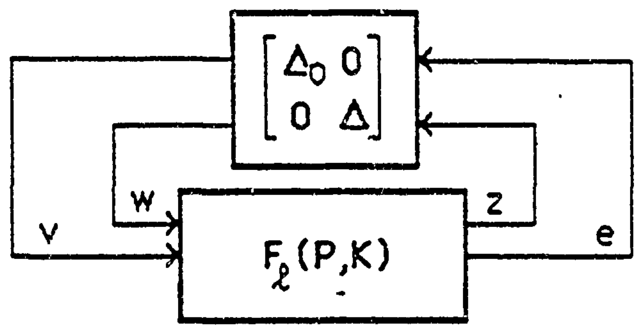
$$\Longleftrightarrow \quad \| F(s) \|_\infty < \frac{1}{\delta}$$

- Simultaneous Stability and Performance are guaranteed by a Stability Condition



A Conservative Analysis Test for Robust Stability and Performance

System S



remains stable for all $\Delta_0, \Delta \in \mathcal{D}$ if

$$\bar{\sigma}[\,F_\ell(j\omega)\,] < \frac{1}{\delta} \quad \text{for all } \omega \le \infty$$

(standard stability robustness theorem with structure ignored)

Arbitrarily Conservative

Example: $F_\ell = \begin{bmatrix} 0 & f \\ 0 & 0 \end{bmatrix}$ $\quad \bar{\sigma}[\,F_\ell(j\omega)\,] = |f(j\omega)|$

Test gives stability only for $\sup_\omega |f(j\omega)| < \frac{1}{\delta}$

Actually, stability is never lost

$$\det\left\{ I - F \begin{bmatrix} \Delta_0 & 0 \\ 0 & \Delta \end{bmatrix} \right\} = \det \begin{bmatrix} 1 & -f\Delta \\ 0 & 1 \end{bmatrix} \neq 0$$

A Tight Stability Test for Structured Perturbations

- **Set of Structured Perturbations**

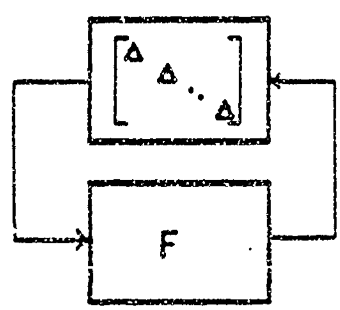
$$\mathcal{X} = \left\{ \Delta \;\middle|\; \begin{array}{l} \Delta = \underset{\text{diag}}{\text{block}}[\Delta_1, \Delta_2, \ldots, \Delta_m] \\ \Delta_j \in \mathcal{P} \end{array} \right\}$$

$$= \left\{ \Delta(s) \;\middle|\; \begin{array}{l} \Delta(s) = \underset{\text{diag}}{\text{block}}[\Delta_1(s), \ldots, \Delta_m(s)] \\ \bar{\sigma}[\Delta_j(j\omega)] \le \delta \cdot \text{ for all } \omega \text{ and all } j \end{array} \right\}$$

- **A New Function of a Matrix:**
  The "Structured Singular Value (SSV)"

$$\mu[F(j\omega)] \triangleq \frac{1}{\min\left\{ \delta \;\middle|\; \begin{array}{l} \det[I - F(j\omega)\Delta(j\omega)] = 0 \\ \text{for some } \omega \text{ and } \Delta \in \mathcal{X}(\delta) \end{array} \right\}}$$

Theorem



System S
is stable for all
$\Delta \in \mathcal{X}(\delta)$ iff $\mu[F(j\omega)] < \dfrac{1}{\delta}$ for all $\omega \le \infty$

Mu Makes a Difference !

Example $F = \begin{bmatrix} 0 & f \\ 0 & 0 \end{bmatrix}$ $\qquad \bar{\sigma}[\,F(j\omega)\,] = |f(j\omega)|$

$$\Delta = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix}$$

Singular Value Test gives stability for $\sup_{\omega} |f(j\omega)| < \frac{1}{\delta}$

Structured Singular Value Test :

$$\mu[F(j\omega)] \triangleq \frac{1}{\min\left\{ \delta \;\middle|\; \begin{array}{l} \det[\,I - F(j\omega)\Delta(j\omega)\,] = 0 \\ \text{for some } \omega \text{ and } \Delta \in \mathcal{X}(\delta) \end{array} \right\}}$$

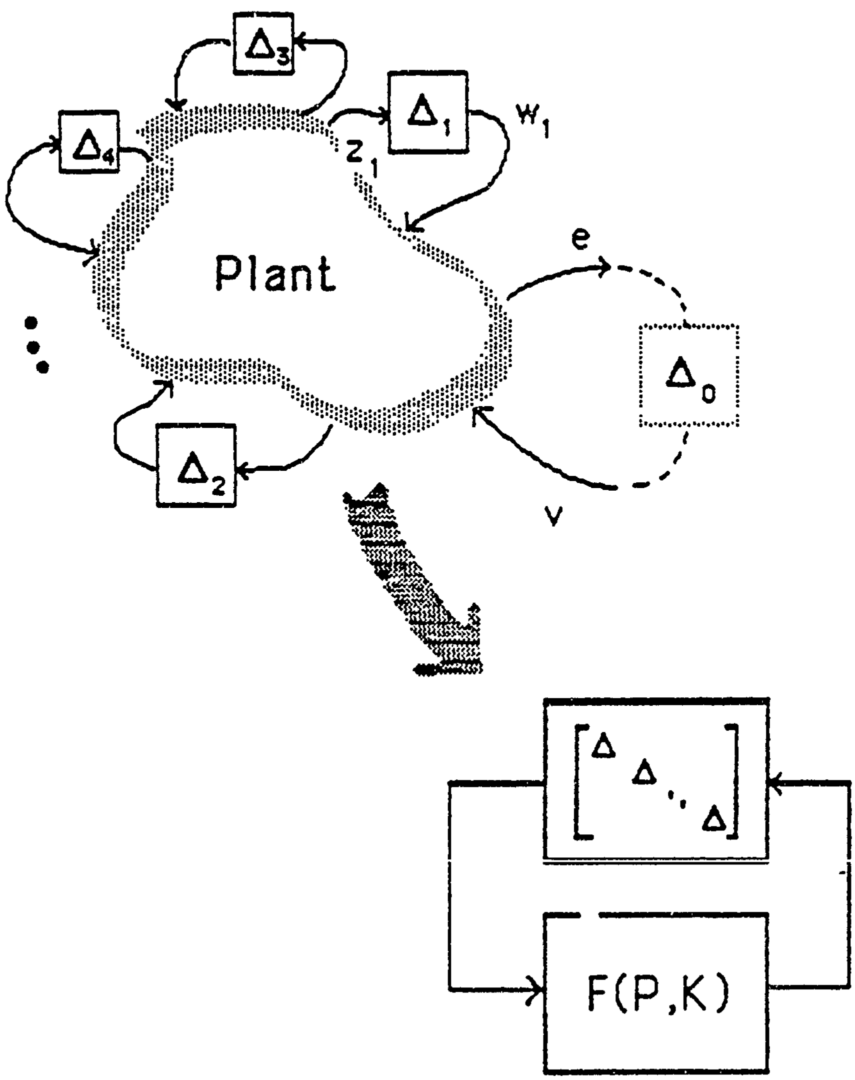
$$= \frac{1}{\min\left\{ \delta \;\middle|\; \det \begin{bmatrix} 1 & -f(j\omega)\,\Delta_2 \\ 0 & 1 \end{bmatrix},\ |\Delta_2| < \delta \right\}}$$

$$= \frac{1}{\infty} = 0$$

$\therefore$ Stability is maintained for all

$\Delta \in \mathcal{X}(\delta)$ with $\delta = \frac{1}{\mu} = \infty$

Mu is a Potent Tool !



An " $\mathbb{H}_\mu$-Optimization Problem "

$$\boxed{\min_{\text{stabilizing } K} \ \sup_{\omega} \ \mu\left[F_\ell(P,K)(j\omega)\right]}$$

- Optimized Robust Stability and Performance

$$\sup_{\omega} \mu = \mu_{opt} \Longrightarrow$$

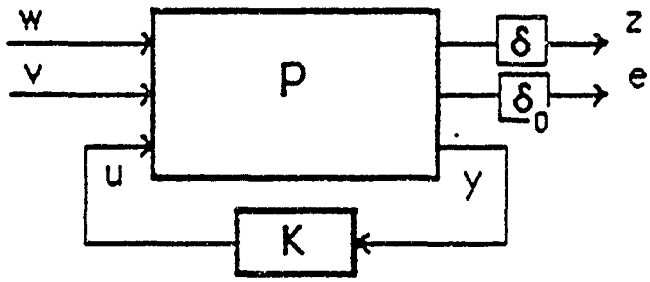
Stability for all

$$\Delta \in \mathcal{X}\left(\frac{1}{\mu_{opt}}\right)$$

Guaranteed Performance Level

$$\left\| F_u\left[F_\ell(P,K),\Delta\right] \right\|_{i2} < \mu_{opt}$$

Other "$\mathbb{H}_\mu$-Problems"



- Maximized Performance Subject to Stability Robustness Constraints

$$\max_{\text{stabilizing } K} \left\{ \delta_0 \;\middle|\; \sup_{\omega} \mu [F(P, K, \delta_0, \delta = 1] < 1 \right\}$$

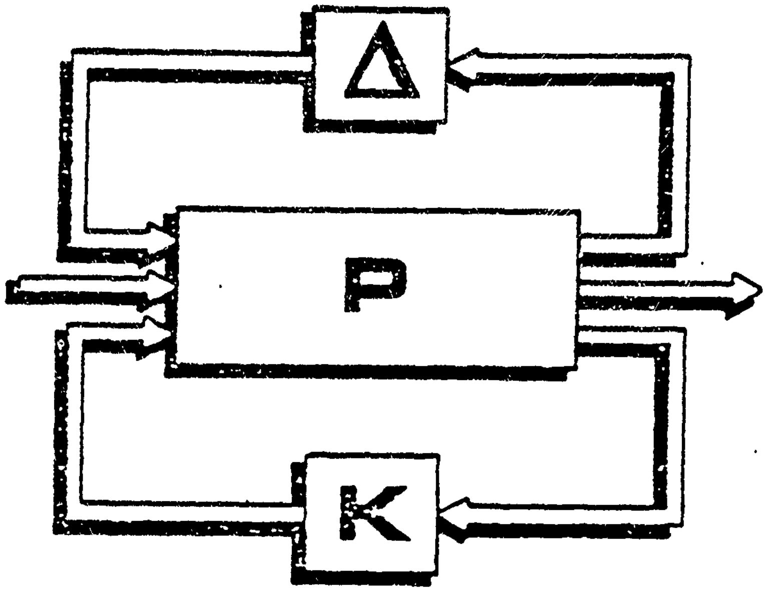
- Maximized Stability Robustness Subject to Performance Constraints

$$\max_{\text{stabilizing } K} \left\{ \delta \;\middle|\; \sup_{\omega} \mu [F(P, K, \delta_0 = 1, \delta] < 1 \right\}$$

Summary of Formal Synthesis Problems

Class of Inputs	Performance Objective	Set of Perturbations	Synthesis Problem
White Noise	$E(e^Te)$	$\Delta = 0$	Min $\| F(s) \|_2$
Impulse Responses	$\Sigma \| e^j(t) \|_2$	$\Delta = 0$	
$\mathbb{L}_2$-bounded Signals	$\sup \| e(t) \|_2$	$\Delta = 0$	Min $\| F(s) \|_\infty$
—	—	$\Delta \in \mathcal{P}$	$\{ \sup_\omega \bar{\sigma}[F(j\omega)] \}$
$\mathbb{L}_2$-bounded Signals	$\sup \| e(t) \|_2$	$\Delta \in \mathcal{X}$	Min $\| F(s) \|_\mu$
—	—	$\Delta \in \mathcal{X}$	$\{ \sup_\omega \mu[F(j\omega)] \}$

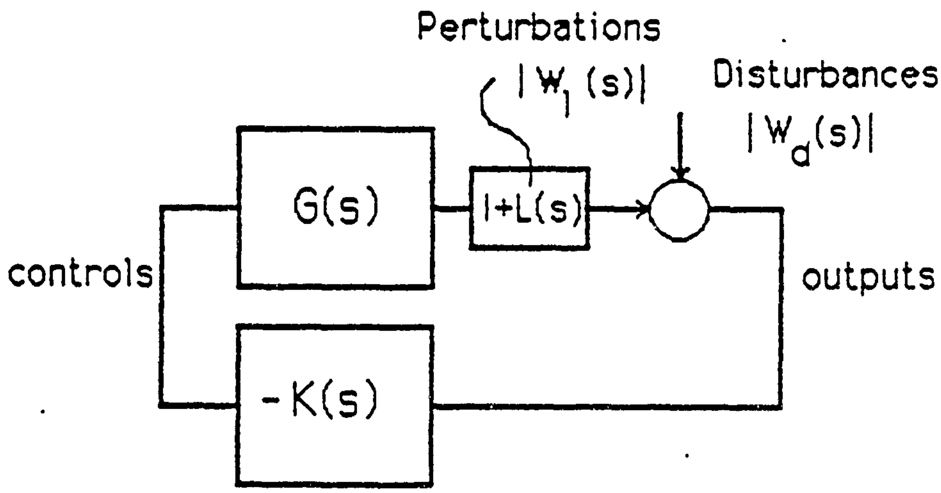
ONR/HONEYWELL WORKSHOP



Historical Perspective

A Generic Classical Control Problem

All SISO



Design Objectives Expressed in Terms of Desired $\boxed{\textbf{Loop Shapes}}$

- Small outputs in response to disturbances

$$\left| [1+GK(j\omega)]^{-1} W_d(j\omega) \right| < 1$$

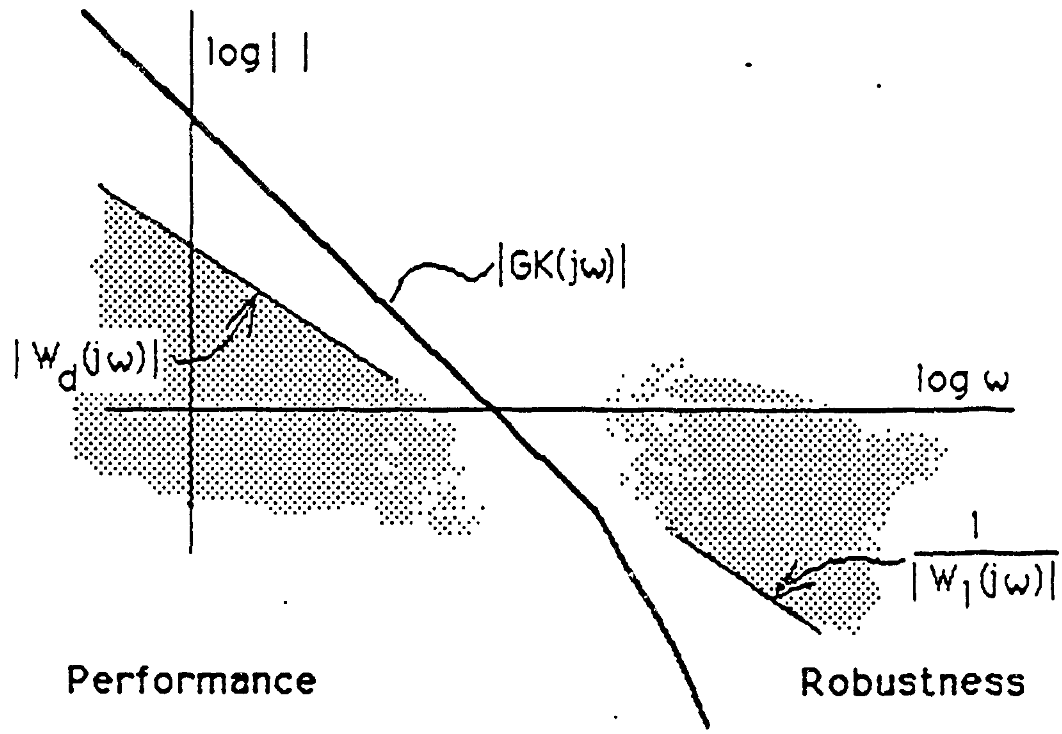
$$|GK(j\omega)| > |W_d(j\omega)| \qquad GK \gg 1$$

- Stability in the face of perturbations

$$\left| GK[1+GK(j\omega)]^{-1} W_1(j\omega) \right| < 1$$

$$|GK(j\omega)| < \frac{1}{|W_1(j\omega)|} \qquad GK \ll 1$$

Desired Loop Shapes



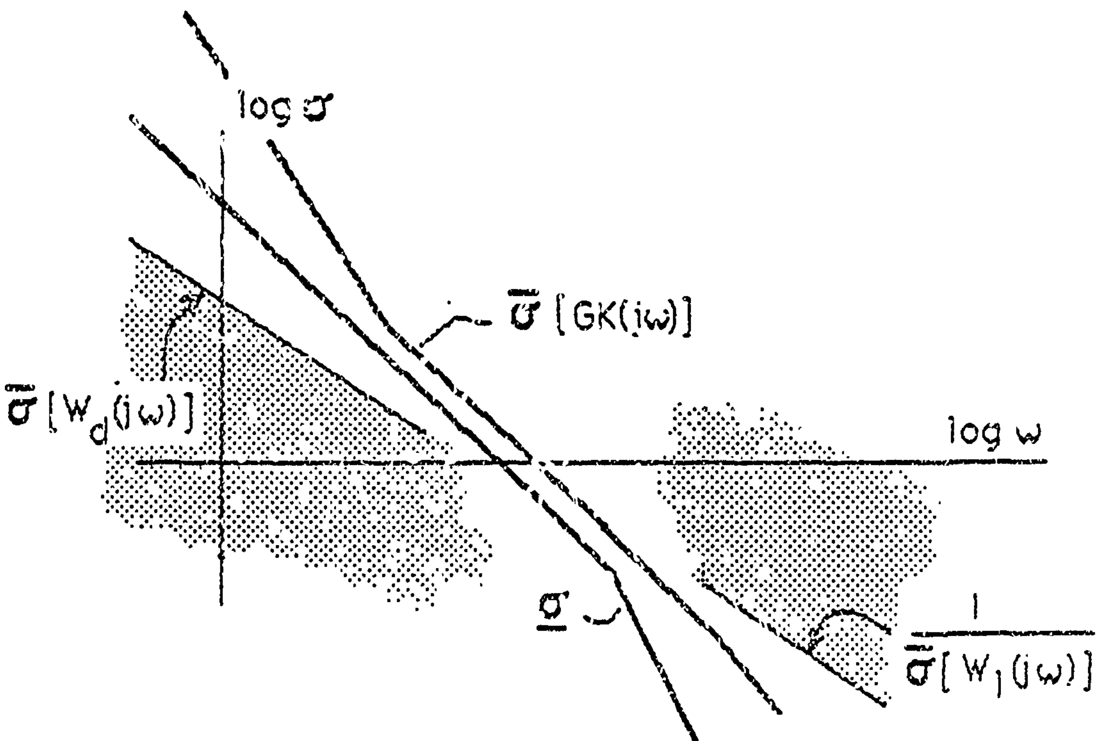
Performance

$$\left| [I+GK(j\omega)]^{-1} W_d(j\omega) \right| < 1$$

Robustness

$$\left| GK[I+GK(j\omega)]^{-1} W_1(j\omega) \right| < 1$$

Multivariable Generalization of Loop Shaping Ideas

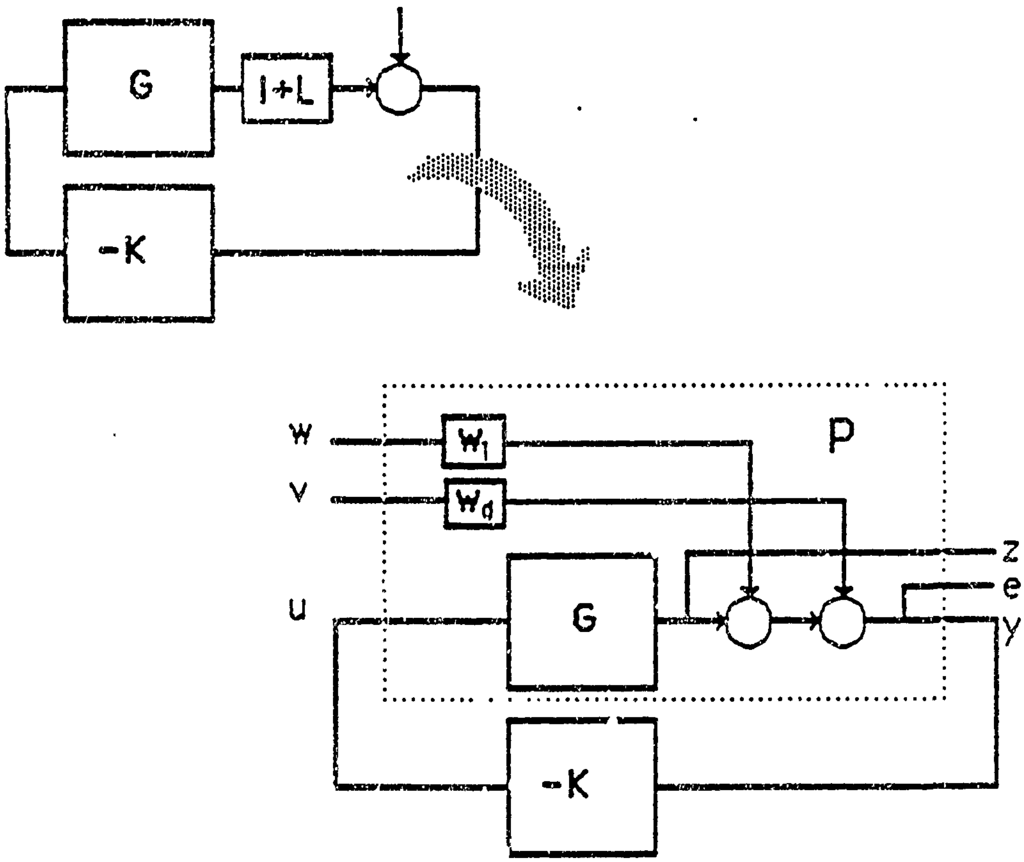


Performance

Robustness

$\bar{\sigma}\left[[I+GK(j\omega)]^{-1} W_d(j\omega)\right] < 1$ $\bar{\sigma}\left[GK[I+GK(j\omega)]^{-1} W_1(j\omega)\right] < 1$

Interpretation of Loop Shaping
In Terms of General Design Model



$$F_\ell(P,K) = \begin{bmatrix} GK(I+GK)^{-1}W_1 & GK(I+GK)^{-1}W_d \\ (I+GK)^{-1}W_1 & (I+GK)^{-1}W_d \end{bmatrix}$$

SSV for the Classical Loop

- **Definition**

$$\mu[F(j\omega)] \triangleq \frac{1}{\min\left\{ \delta \;\middle|\; \begin{array}{l} \det[I - F(j\omega)\Delta(j\omega)] = 0 \\ \text{for some } \omega \text{ and } \Delta \in \mathcal{X}(\delta) \end{array} \right\}}$$

with $\Delta = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_d \end{bmatrix}$

- **Some Calculations**

$$I - F_\ell \Delta = I - \begin{bmatrix} GK(I+GK)^{-1} W_1 \\ (I+GK)^{-1} W_d \end{bmatrix} \begin{bmatrix} \Delta_1 & \Delta_d \end{bmatrix}$$

is singular iff

$$I - \Delta_1 GK(I+GK)^{-1} W_1 - \Delta_d (I+GK)^{-1} W_d$$

is singular

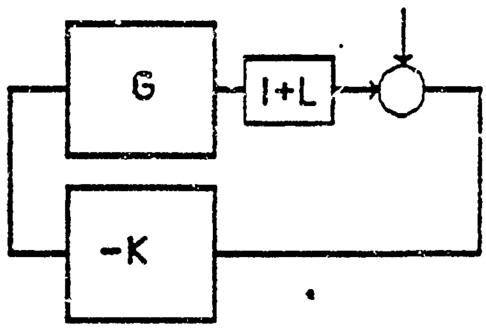
$$\therefore \quad \mu[F_\ell] \leq \bar{\sigma}[GK(I+GK)^{-1} W_1] + \bar{\sigma}[(I+GK)^{-1} W_d]$$

Loop Shapes which satisfy classical objectives tend to minimize μ

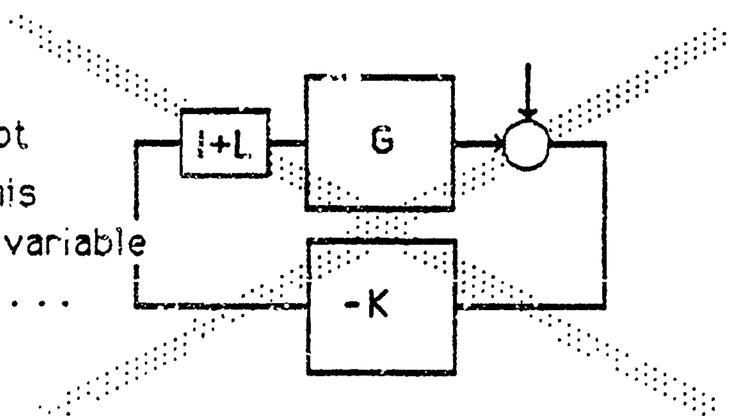
Limitations of Loop Shaping

All design objectives must be reflected to one point in the loop

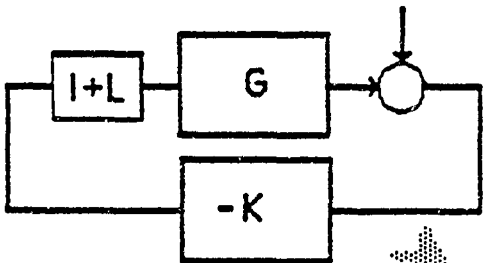
Loop Shaping works for this multivariable loop . . .

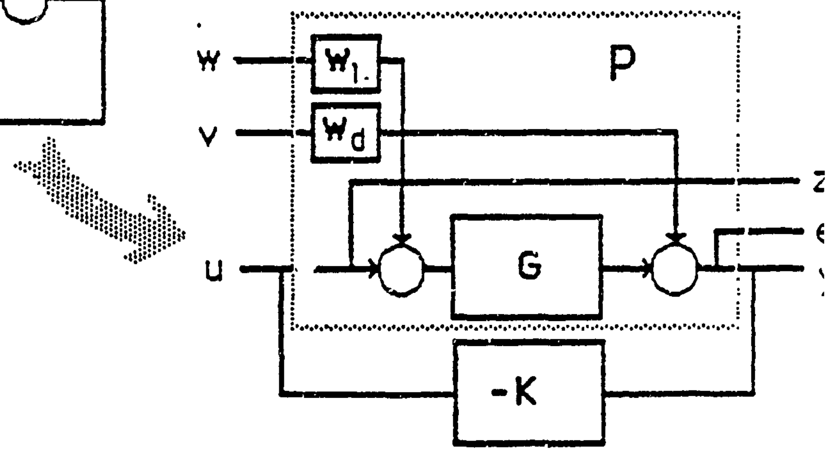


but not for this multivariable loop . . .



What Goes Wrong ?





$F_\ell(P,K) =$

$$\begin{bmatrix} KG(I+KG)^{-1}W_1 & K(I+GK)^{-1}W_d \\ G(I+KG)^{-1}W_1 & (I+GK)^{-1}W_d \end{bmatrix} = \begin{bmatrix} K \\ I \end{bmatrix} \begin{bmatrix} G(I+KG)^{-1}W_1 & (I+GK)^{-1}W_d \end{bmatrix}$$

$I - F_\ell \Delta$ is nonsingular iff

$I - G(I+KG)^{-1}W_1\Delta_1 K - (I+GK)^{-1}W_d\Delta_d$ is nonsingular

$K^{-1}K$

$$\boxed{\mu[F_\ell] \le \left[\frac{\bar{\sigma}K}{\underline{\sigma}K}\right] \bar{\sigma}[KG(I+KG)^{-1}W_1] + \bar{\sigma}[(I+GK)^{-1}W_d]}$$

Classical objectives do not minimize μ unless condition of K remains small

An Aside: The Scaling Implicit in μ

$$F_\ell(P,K) = \begin{bmatrix} KG(I+KG)^{-1}W_1 & K(I+GK)^{-1}W_d \\ G(I+KG)^{-1}W_1 & (I+GK)^{-1}W_d \end{bmatrix}$$

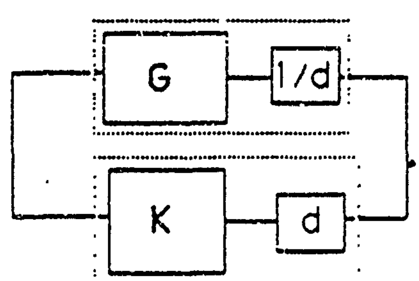
$$= \begin{bmatrix} K \\ I \end{bmatrix} (I+GK)^{-1} \begin{bmatrix} GW_1 & W_d \end{bmatrix}$$

- To compute μ, find the minimizing D for $\bar{\sigma}[DF_\ell D^{-1}]$

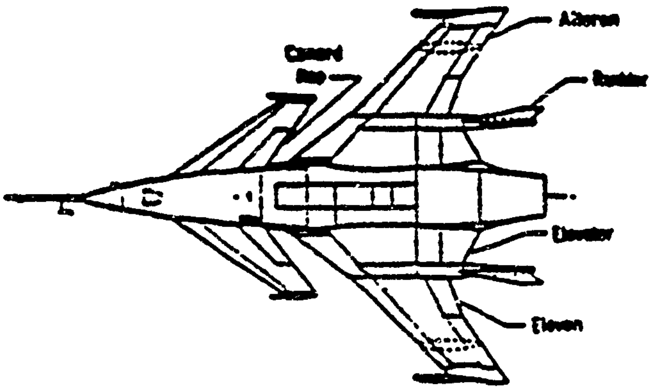
$$DF_\ell D^{-1} = \begin{bmatrix} dK \\ I \end{bmatrix} (I+\frac{1}{d}G\,dK)^{-1} \begin{bmatrix} \frac{1}{d}GW_1 & W_d \end{bmatrix}$$

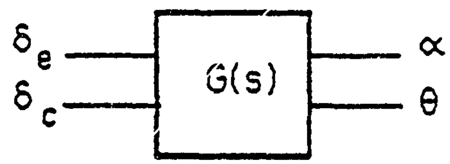
$$D = \begin{bmatrix} dI & 0 \\ 0 & I \end{bmatrix}$$

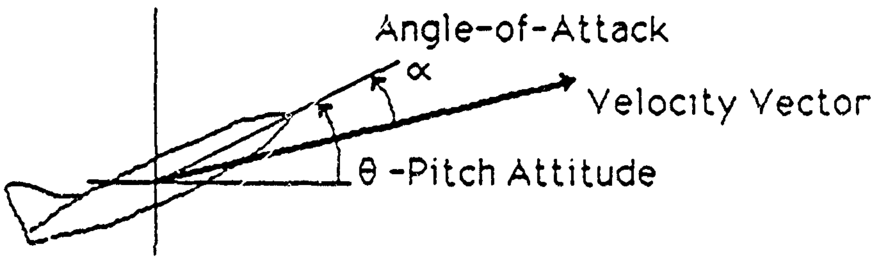
- Equivalent to a change of "units" in the problem



Decoupled Flight Path / Attitude Control

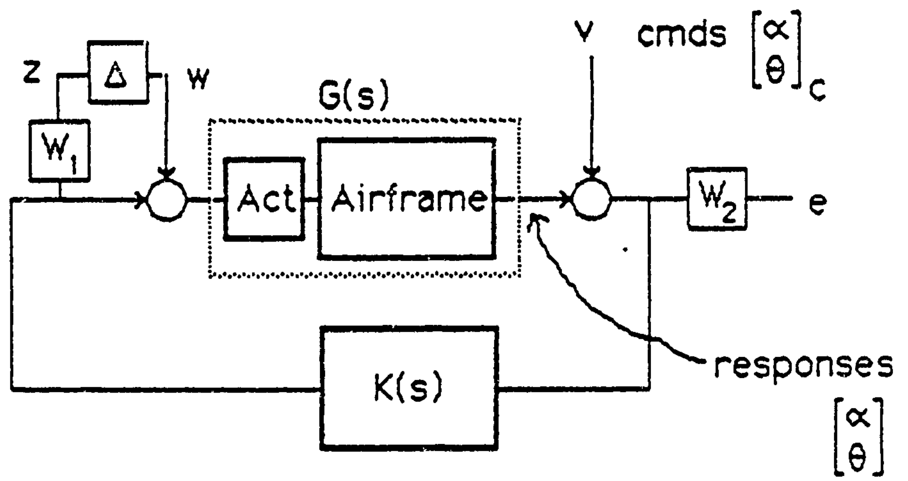






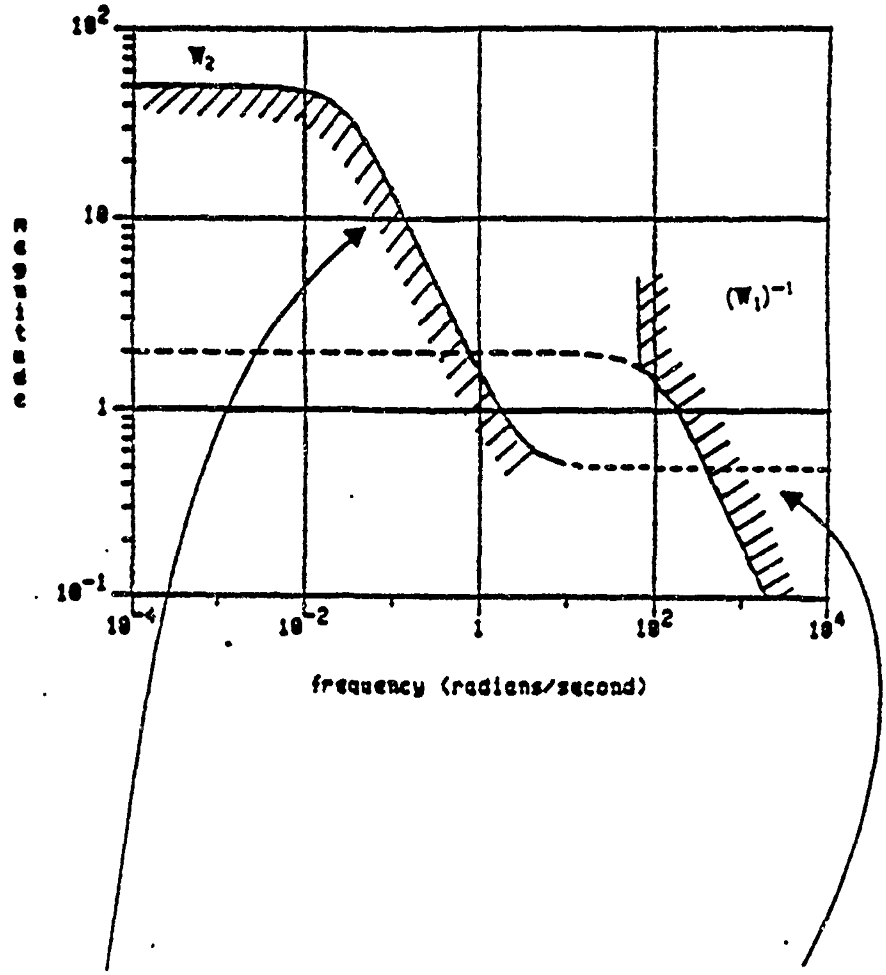
Decoupled Flight Path / Attitude Control

In Our General Framework



$$F_{\ell}(P,K) = \begin{bmatrix} W_1 KG(I+KG)^{-1} & W_1 K(I+GK)^{-1} \\ W_2 G(I+KG)^{-1} & W_2 (I+GK)^{-1} \end{bmatrix}$$

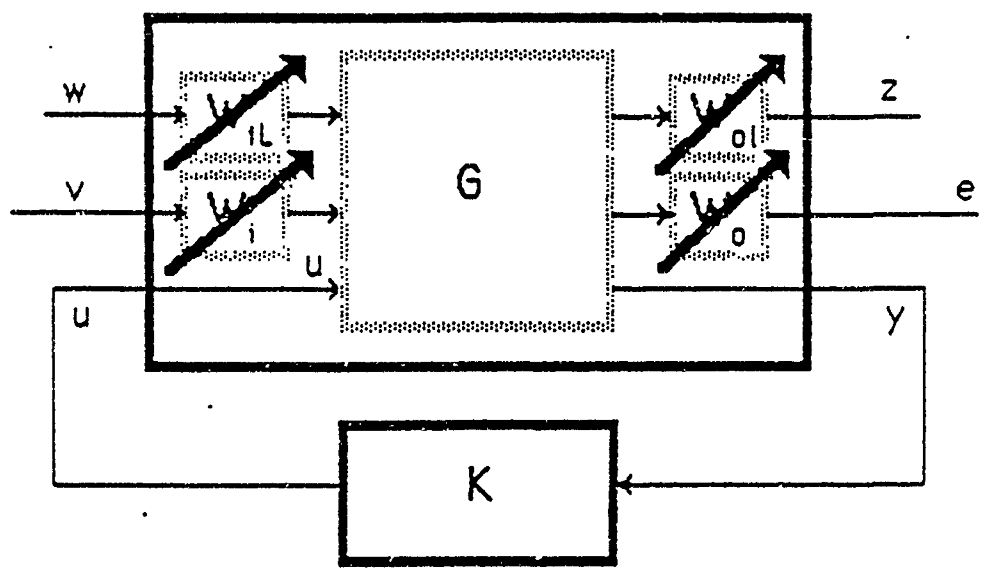
Selected Weighting Functions



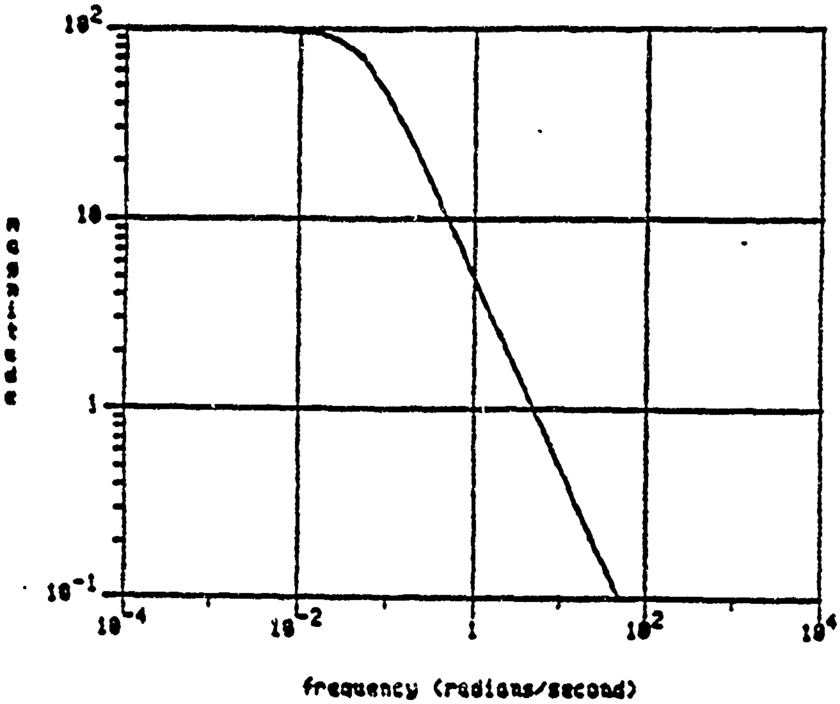
$|| W_2(I+GK)^{-1} ||_\infty \leq 1$

$|| W_1K(I+GK)^{-1}G ||_\infty \leq 1$

An Aside : Weights are the "Knobs" of the Formal Synthesis Problem

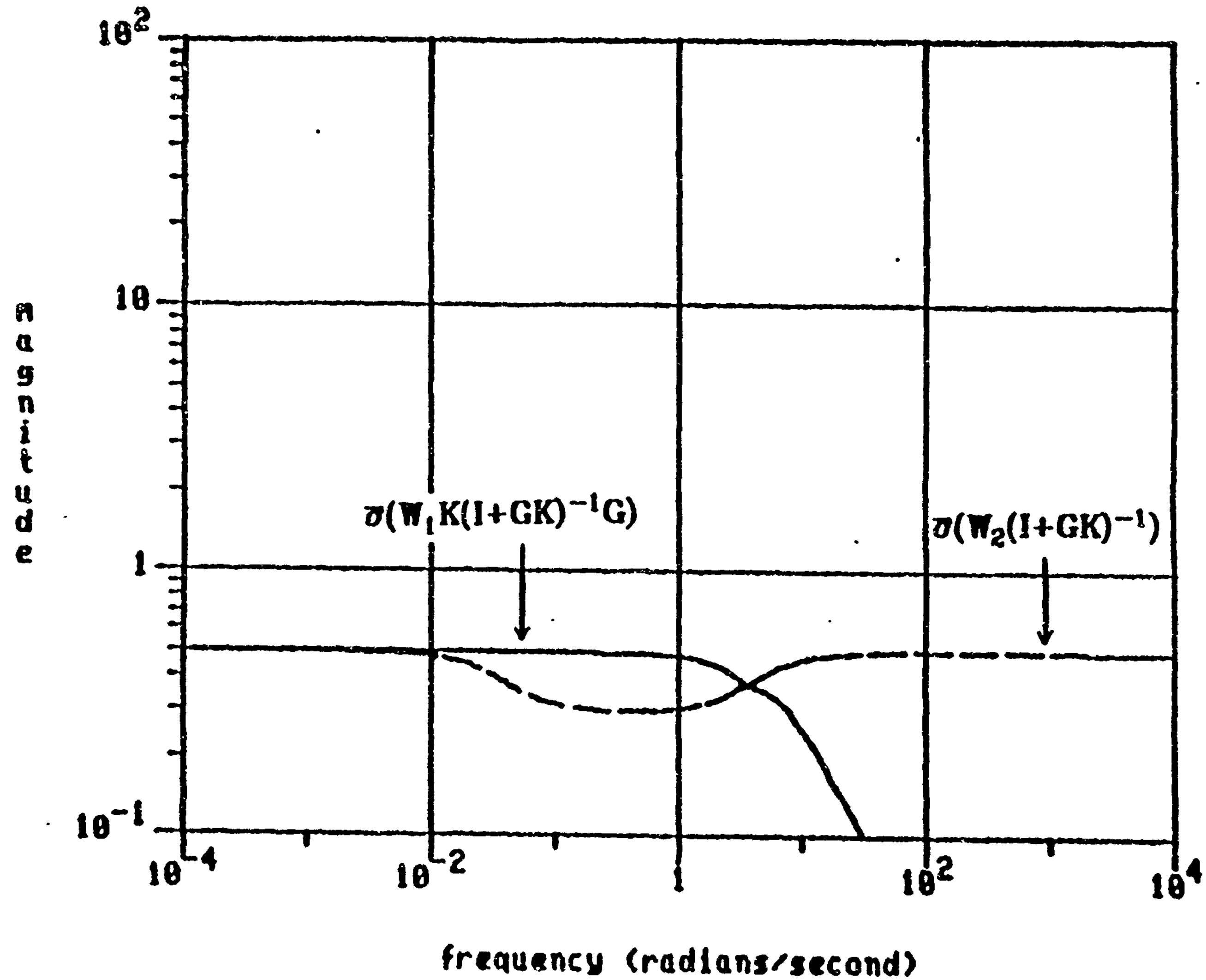


Selected Desired Loop Shape

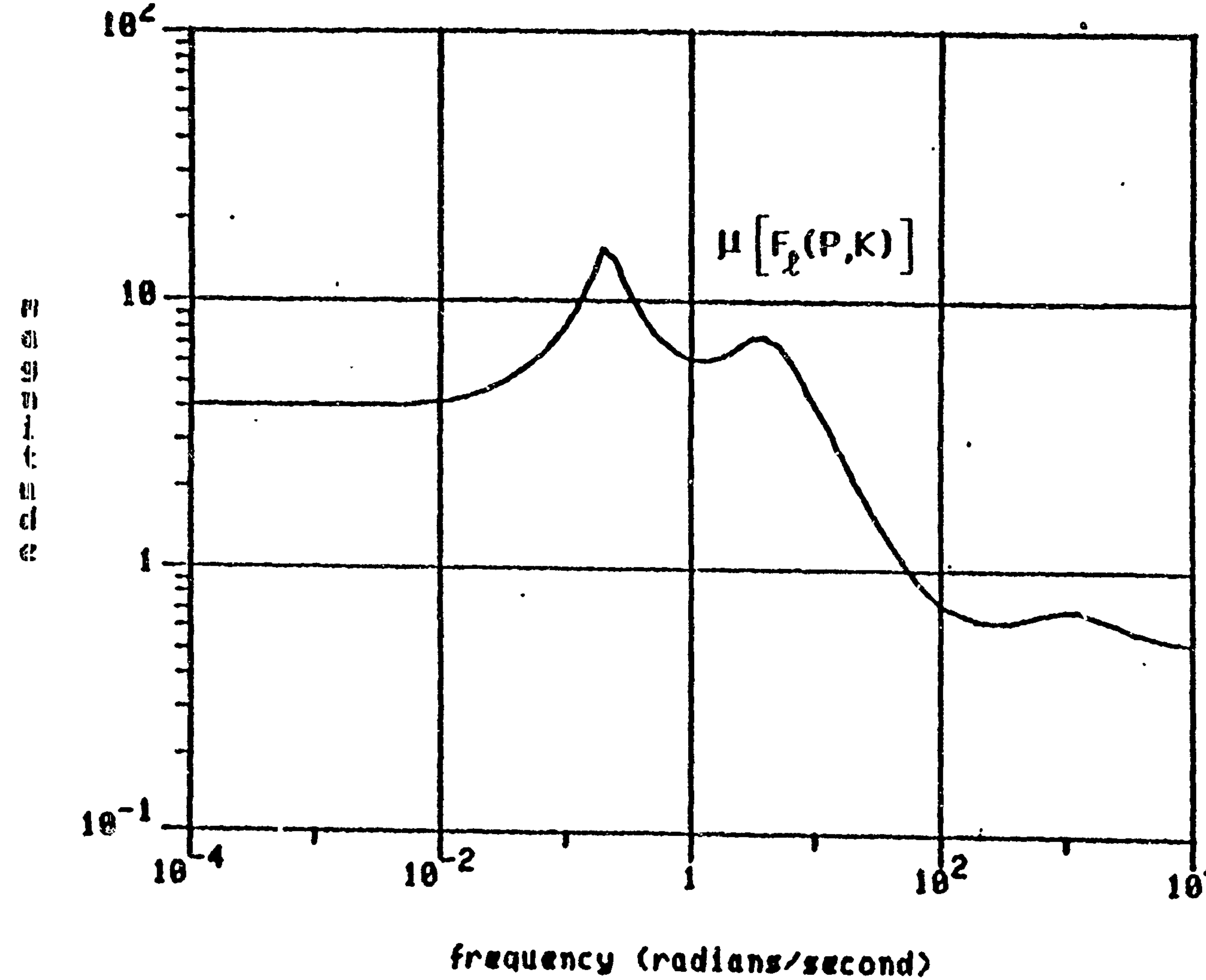


$\bar{\sigma}_i(KG)\ \forall i$

Resulting Loop-Shaping Design



Structured Singular Value of the Design



Implications

- The Loop-Shaping Design is Robustly Stable

$$\| W_1 KG (I+KG)^{-1} \|_\infty < 1$$

- Performance is Satisfactory for the Nominal Plant

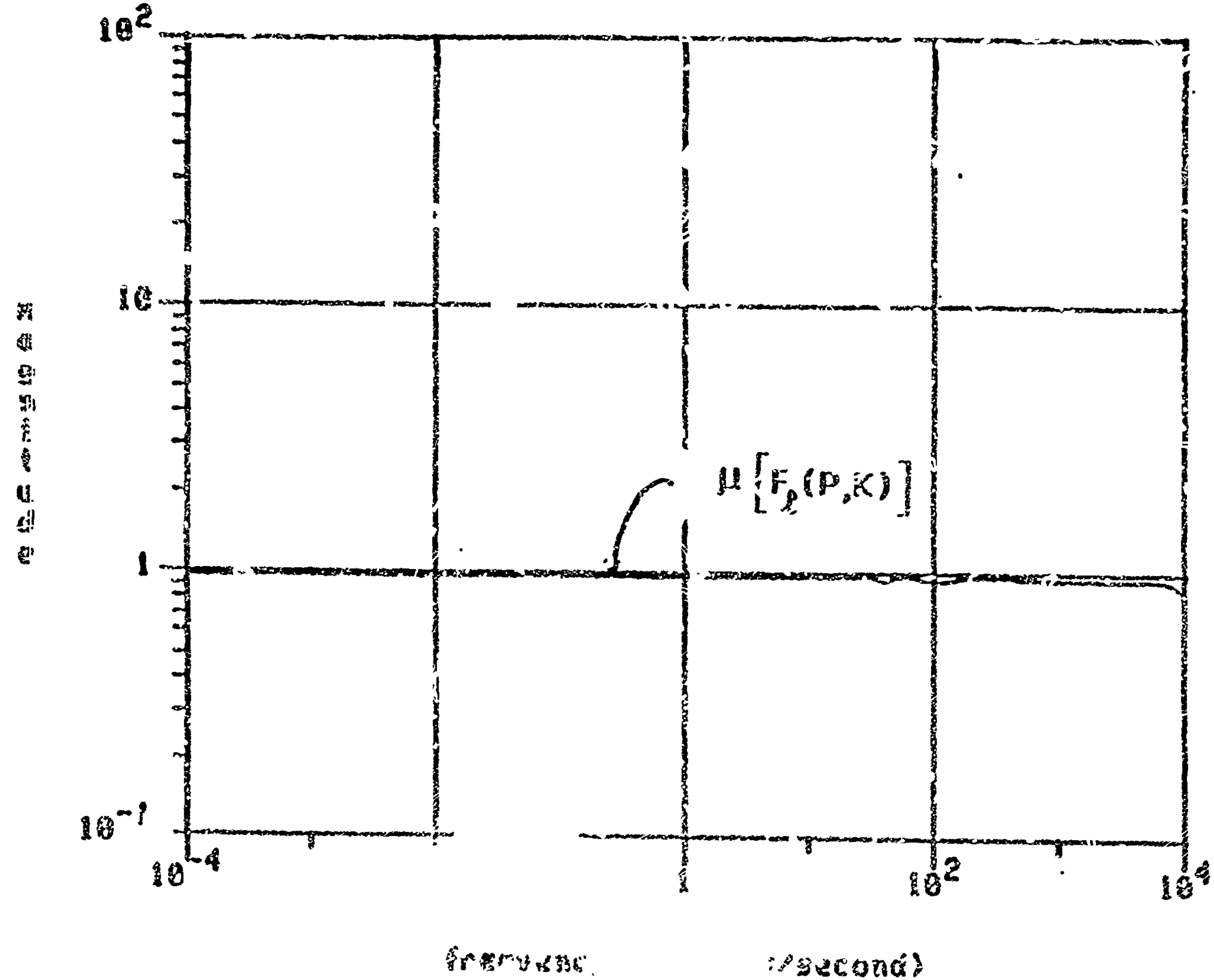
$$\| W_1 (I+GK)^{-1} \|_\infty < 1$$

- **But** there exists a Δ_1 with $\|\Delta_1\|_\infty \leq 1/15$ such that

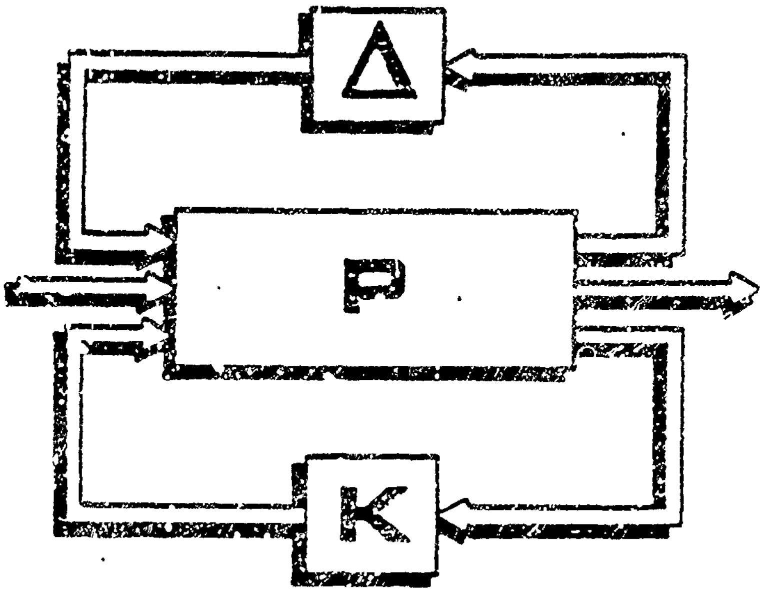
$$\| W_2 F_u [F_\ell (P,K), \Delta_1] \|_\infty \geq 15$$

$\therefore$ The design does not have robust performance !

Solution Achieved with $\mathcal{H}_\mu$-Optimization

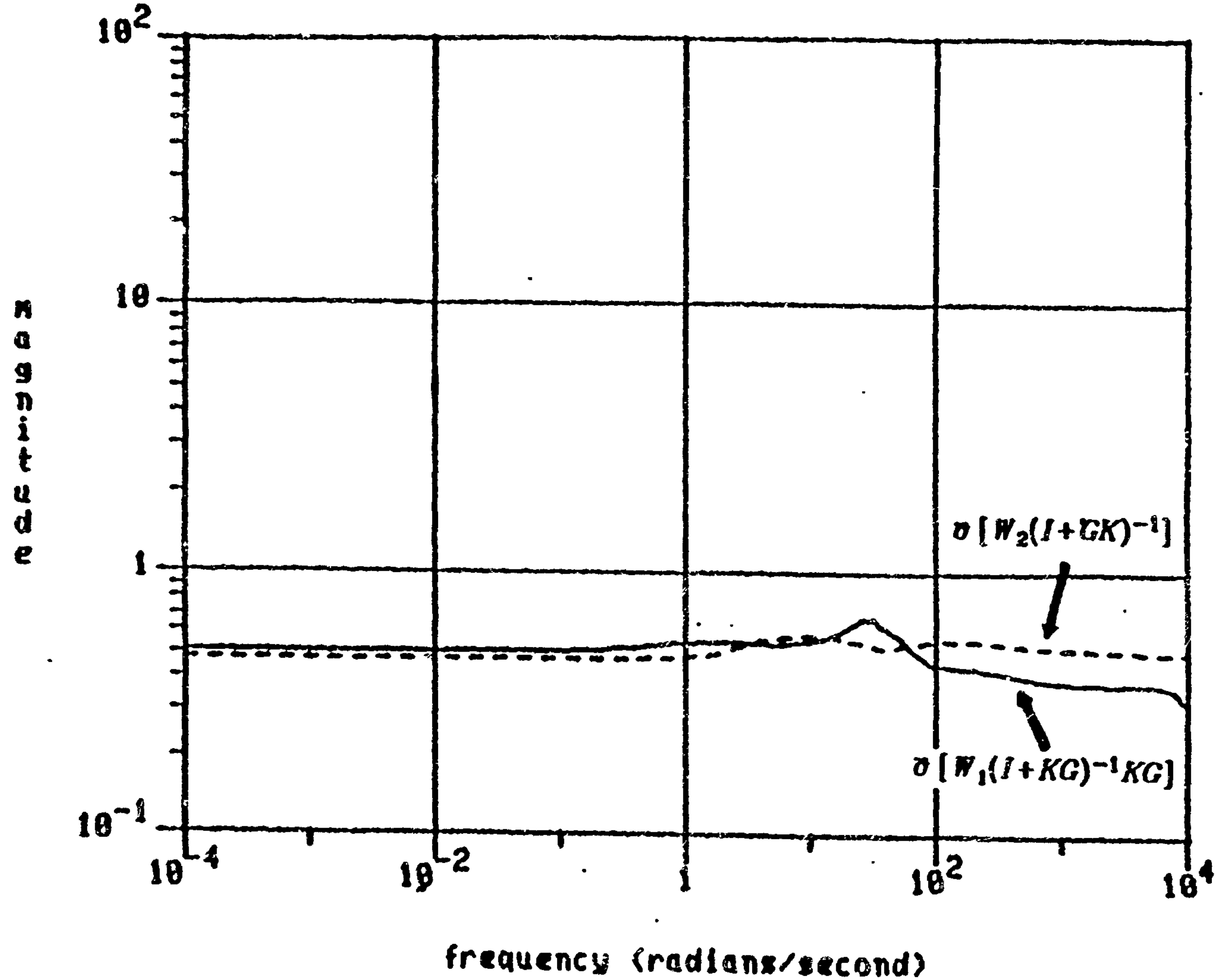


ONR/HONEYWELL WORKSHOP



Overview of Synthesis Solutions

Solution Achieved with $\mathbb{H}_{\mu}$-Optimization



Summary of Formal Synthesis Problems

Class of Inputs	Performance Objective	Set of Perturbations	Synthesis Problem
White Noise	$E(e^T e)$	$\Delta = 0$	Min $\| F(s) \|_2$
Impulse Responses	$\sum \| e^j(t) \|_2$	$\Delta = 0$	$\left\{ \int Tr[F^* F] \, d\omega \right\}^{1/2}$
$\mathbb{L}_2$-bounded Signals	$\sup \| e(t) \|_2$	$\Delta = 0$	Min $\| F(s) \|_\infty$
—	—	$\Delta \in \mathbb{P}$	$\left\{ \sup_\omega \bar{\sigma}[F(j\omega)] \right\}$
$\mathbb{L}_2$-bounded Signals	$\sup \| e(t) \|_2$	$\Delta \in \mathbb{X}$	Min $\| F(s) \|_\mu$
—	—	$\Delta \in \mathbb{X}$	$\left\{ \sup_\omega \mu[F(j\omega)] \right\}$

Status of Synthesis Solutions

- The $\mathbb{H}_2$-Problem

  Solutions completely known
  (Parallel Wiener-Hopf / LQG Solution)

- The $\mathbb{H}_\infty$-Problem

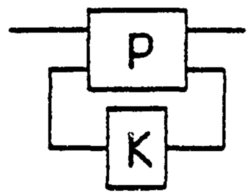
  Solutions available arbitrarily close
  to optimal

- The $\mathbb{H}_\mu$-Problem

  Approximate Solutions available through
  Iteration

$$\min_{D} \left\{ \min_{K} \left\| D(s)F(s)D(s)^{-1} \right\|_\infty \right\}$$

The $\mathbb{H}_\alpha$-Solution Process $\alpha = 2, \infty$



$$\min_{\text{Stabilizing } K} \left\| F_\ell(P,K) \right\|_\alpha$$

$$F_\ell = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21}$$

Parametrization

$$\min_{Q \in R\mathbb{H}_\alpha} \left\| T_{11} + T_{12} Q T_{21} \right\|_\alpha \qquad T_{12}^* T_{12} = I \quad T_{21} T_{21}^* = I$$

Unitary invariance

$$\min_{Q \in R\mathbb{H}_\alpha} \left\| \begin{matrix} R_{11} + Q & R_{12} \\ R_{21} & R_{22} \end{matrix} \right\|_\alpha$$

Projection | Dilation

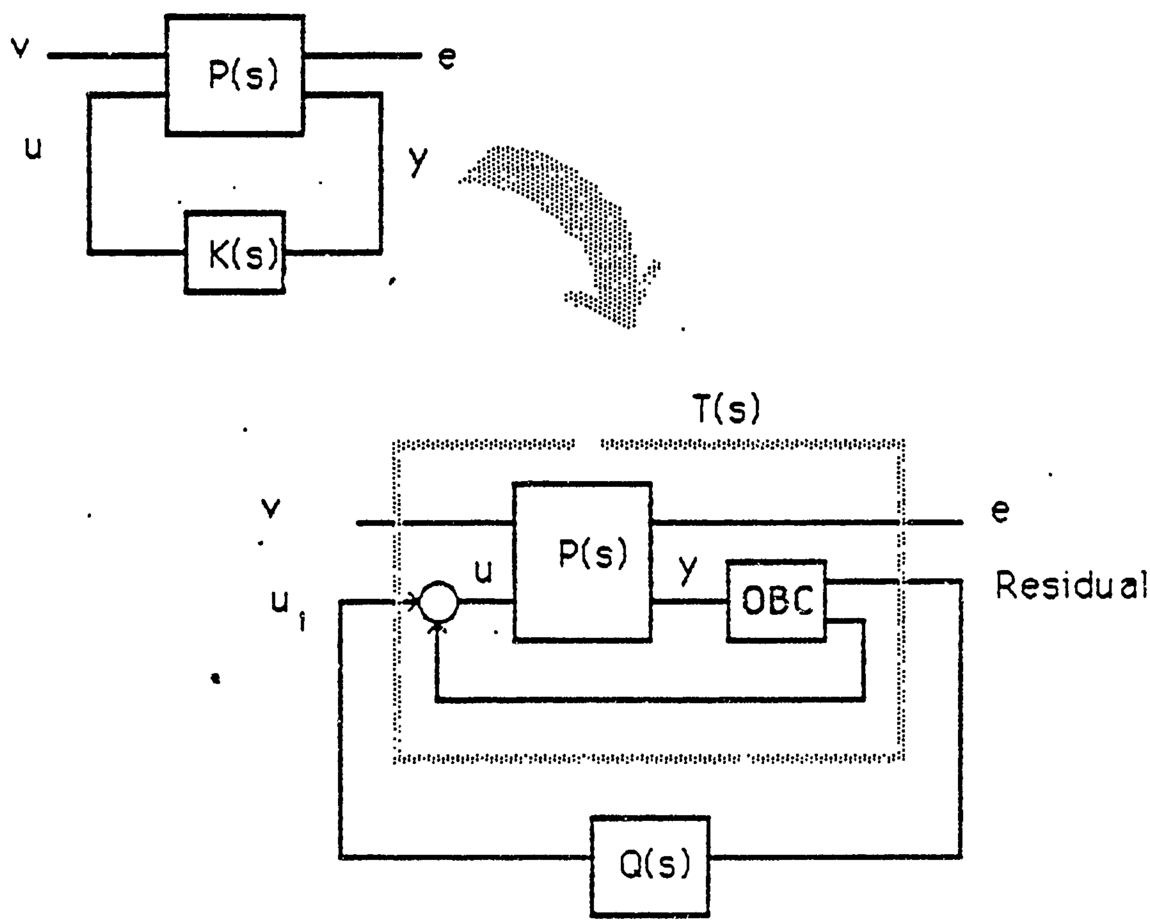
$$\min_{Q \in R\mathbb{H}_2} \left\| R_{11} + Q \right\|_2 \qquad \min_{Q \in R\mathbb{H}_\infty} \left\| G + \hat{Q} \right\|_\infty$$

$$Q = -P_{\mathbb{H}_2}(R_{11}) \qquad \hat{Q} = \text{Best } \mathbb{H}_\infty \text{ Approx of } \mathbb{L}_\infty \; G$$

Re-Parametrization

K_{opt}

Parametrization Highlights . . .



$T_{22}(s)$ = (residual) / u = 0 (separation theorem)

$$F_{\ell}(T,Q) = T_{11} + T_{12}\,Q\,(I - T_{22}Q)^{-1}\,T_{21}$$
$$= T_{11} + T_{12}\,Q\,T_{21}$$

Parametrization Highlights

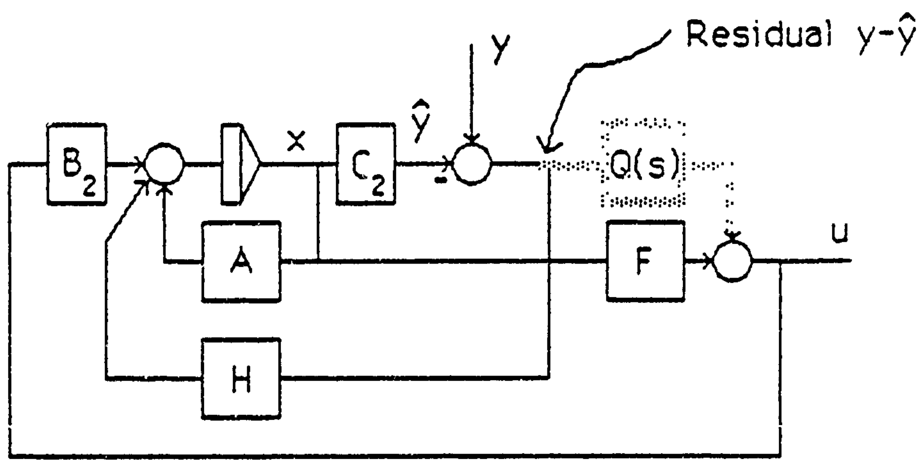
- **Youla Parametrization of all Stabilizing Controllers**

$$K(s) = (U_0 + MQ)(V_0 + NQ)^{-1}$$

where $P_{22} = NM^{-1}$, $K_0 = U_0 V_0^{-1}$ stabilizes, Q stable

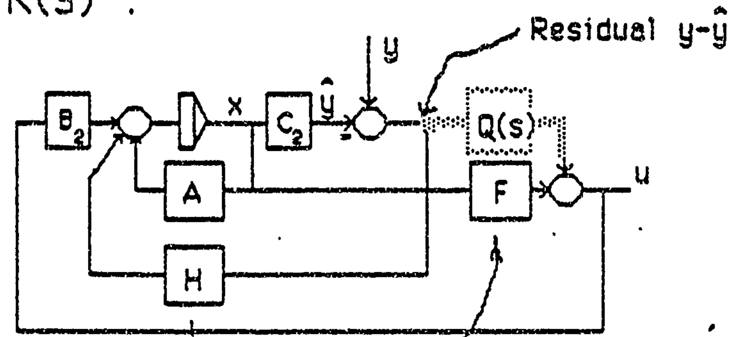
$$\Updownarrow$$

- **Observer-Based Compensator (OBC) with added stable dynamics**



Exploiting Unitary Invariance

$K(s)$:



Let H and F be Riccati Gains

Then $\begin{bmatrix} T_{12} & T_{12}^{\perp} \end{bmatrix}$ and $\begin{bmatrix} T_{21} \\ T_{21}^{\perp} \end{bmatrix}$

can be made unitary

$$\begin{bmatrix} T_{12} & T_{12}^{\perp} \end{bmatrix}^* \begin{bmatrix} T_{12} & T_{12}^{\perp} \end{bmatrix} = I$$

$$\begin{bmatrix} T_{21} \\ T_{21}^{\perp} \end{bmatrix} \begin{bmatrix} T_{21} \\ T_{21}^{\perp} \end{bmatrix}^* = I$$

$$\| T_{11} + T_{12} Q T_{21} \|_{\alpha} = \left\| \begin{bmatrix} T_{12} & T_{12}^{\perp} \end{bmatrix}^* \begin{bmatrix} T_{11} + T_{12} Q T_{21} \end{bmatrix} \begin{bmatrix} T_{21} \\ T_{21}^{\perp} \end{bmatrix}^* \right\|_{\alpha}$$

$$= \left\| \begin{matrix} R_{11} + Q & R_{12} \\ R_{21} & R_{22} \end{matrix} \right\|_{\alpha}$$

THE PRICE : unstable functions

Projections

$$\left\| \begin{matrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{matrix} \right\|_2 =$$

$$\left[\int_{-\infty}^{\infty} Tr(R_{11}+Q)^*(R_{11}+Q) + Tr(R_{12}^* R_{12}) + Tr(R_{21}^* R_{21}) + Tr(R_{22}^* R_{22})\, d\omega \right]^{1/2}$$

$\Downarrow$

It is sufficient to minimize $\| R_{11}+Q \|_2$

turns out to
be entirely unstable

must be
stable

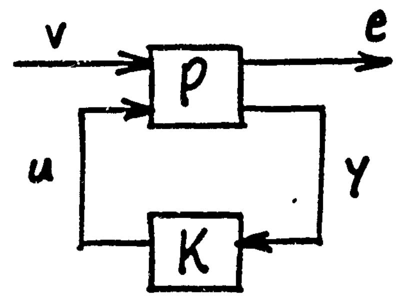
$$\therefore \quad Q_{opt} = 0$$

PART THREE:

FRANCIS'S NOTES

STABILIZATION

system



interconnection matrix

controller

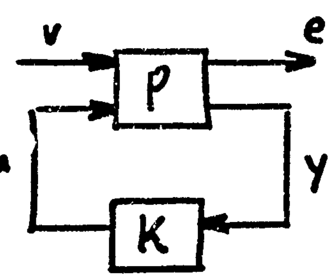
topics

1. well-posedness
2. internal stability
3. parametrization of K
4. state-space realization
5. closed-loop transfer matrix

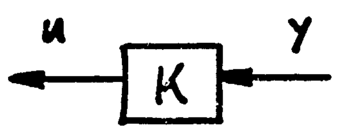
key references

1. Youla, Jabr, Bongiorno '76
2. Desoer, Liu, Murray, Saeks '80
3. Nett, Jacobson, Balas '84

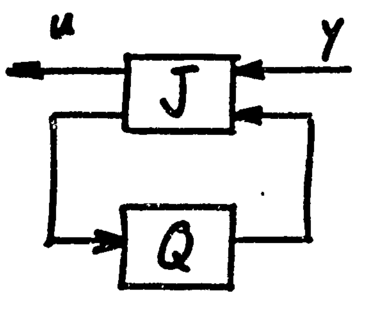
preview



parametrization of K

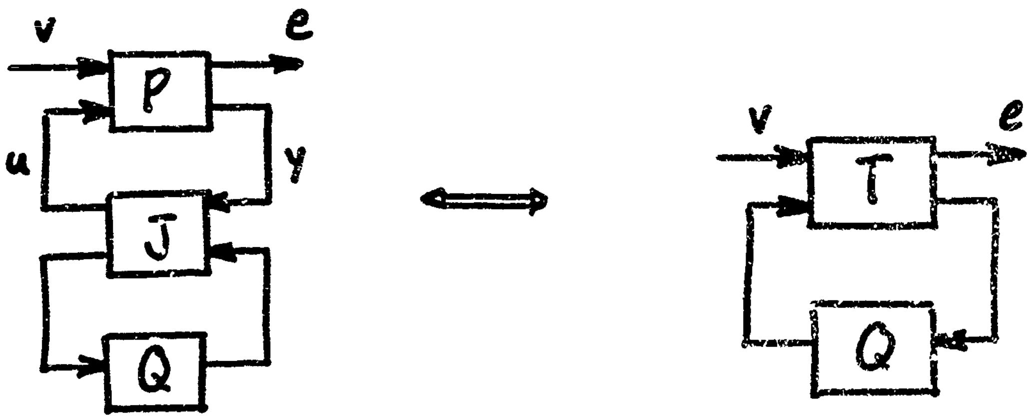


$\Longleftrightarrow$

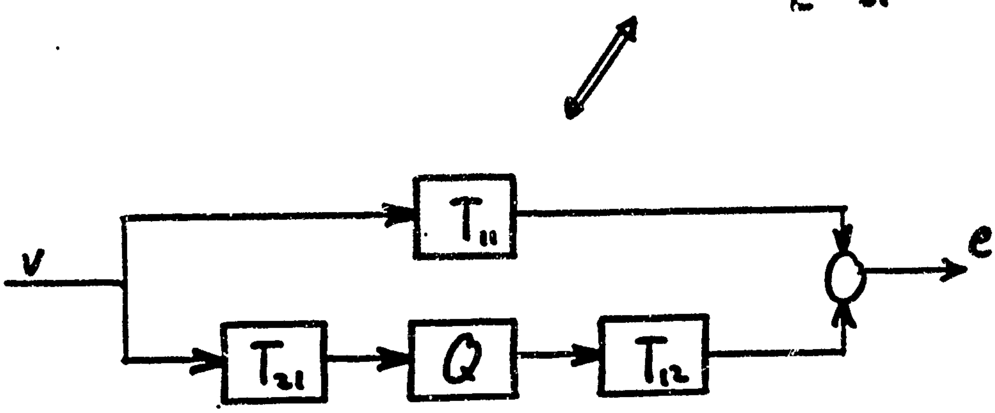


Q = 'free parameter'

closed-loop transfer matrix



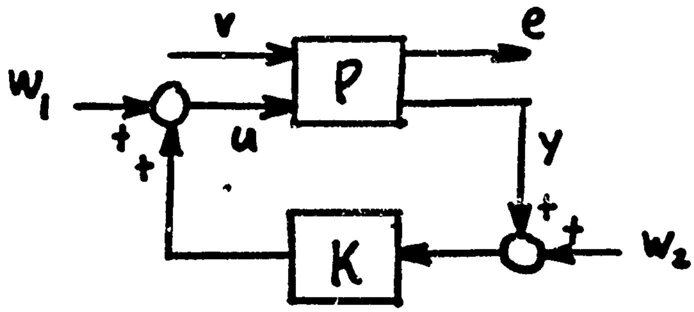
$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & 0 \end{bmatrix}$$



$$e = \underbrace{(T_{11} + T_{12}\, Q\, T_{21})}_{\text{affine in } Q}\, v$$

1. well-posedness
2. internal stability
3. parametrization of K
4. state-space realization
5. closed-loop transfer matrix

assume P, K proper



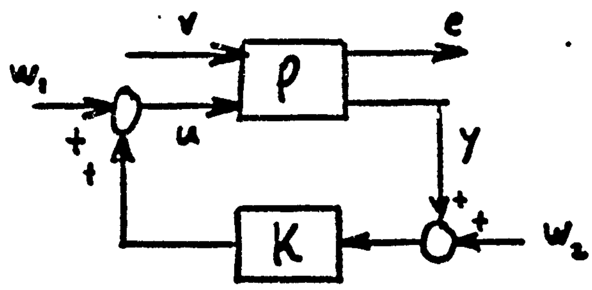
w_1, w_2 fictitious exogenous inputs

def'n system is well-posed if

transfer matrix $\begin{bmatrix} v \\ w_1 \\ w_2 \end{bmatrix} \longmapsto \begin{bmatrix} e \\ y \\ u \end{bmatrix}$

exists & is proper

suffices to consider $\begin{bmatrix} v \\ w_1 \\ w_2 \end{bmatrix} \longmapsto u$



$$\begin{bmatrix} e \\ y \end{bmatrix} = P \begin{bmatrix} v \\ u \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} v \\ u \end{bmatrix}$$

$$\begin{cases} y = P_{21} v + P_{22} u \\ u = w_1 + K(y + w_2) \end{cases}$$

$$u = w_1 + K\left[(P_{21} v + P_{22} u) + w_2 \right]$$

$$(I - KP_{22})u = w_1 + KP_{21} v + K w_2$$

therefore

well-posed $\Longleftrightarrow$ $(I - KP_{22})^{-1}$ exists

& is proper

conclusion

well-posedness

$$\iff I - K(\infty)P_{22}(\infty) \text{ invertible}$$

$$\iff I - P_{22}(\infty)K(\infty) \text{ invertible}$$

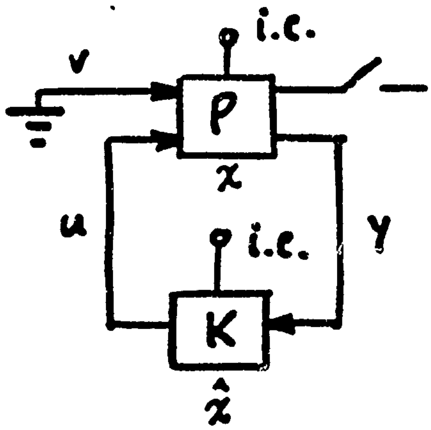
$$\iff \begin{bmatrix} I & -K(\infty) \\ -P_{22}(\infty) & I \end{bmatrix} \text{ invertible}$$

well-posedness assumed hereafter

1. well-posedness
2. internal stability
3. parametrization of K
4. state-space realization
5. closed-loop transfer matrix

notation

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] := D + C(sI - A)^{-1}B$$



def'n internal stability means

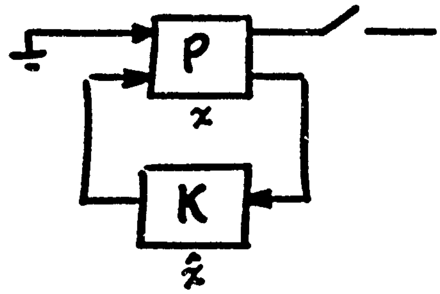
$$x(t),\ \hat{x}(t) \longrightarrow 0 \quad \forall \text{ i.c.}$$

characterization

minimal realizations

$$P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

$$K = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] \qquad \Longrightarrow \qquad P_{ij} = \left[\begin{array}{c|c} A & B_j \\ \hline C_i & D_{ij} \end{array}\right]$$



$$P \quad \begin{cases} \dot{x} = Ax + B_2 u \\ y = C_2 x + D_{22} u \end{cases} \qquad P_{22}(\infty) = D_{22}$$

$$K \quad \begin{cases} \dot{\hat{x}} = \hat{A}\hat{x} + \hat{B} y \\ u = \hat{C}\hat{x} + \hat{D} y \end{cases} \qquad K(\infty) = \hat{D}$$

eliminate u, y

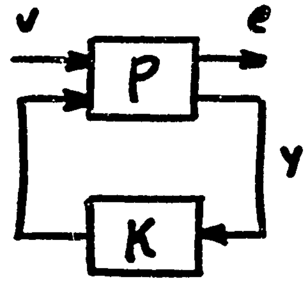
exists by well-posedness

$$\begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} I & -\hat{D} \\ -D_{22} & I \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0 & \hat{C} \\ C_2 & 0 \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$$

$$\frac{d}{dt}\begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \begin{bmatrix} \quad & \quad \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$$

internal stability $\Longleftrightarrow$ eigenvalues in $\operatorname{Re} s <$

existence



$$P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

lemma $\exists$ proper K achieving internal stability $\Longleftrightarrow$

(A, B_2) stabilizable

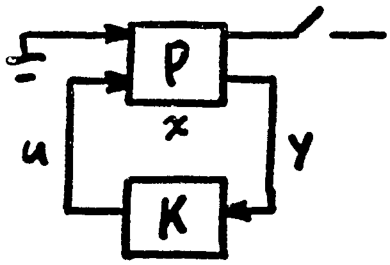
(unstable modes of P controllable from u)

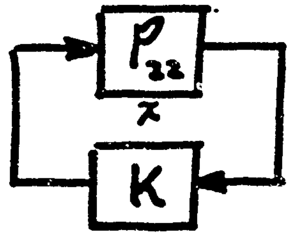
& (C_2, A) detectable

(unstable modes of P observable at y)

... assumed hereafter

reduction





$$\begin{cases}\dot{x} = Ax + B_2 u\\ y = C_2 x + D_{22} u\end{cases}$$

$$P_{22} = \left[\begin{array}{c|c} A & B_2\\ \hline C_2 & D_{22}\end{array}\right]$$

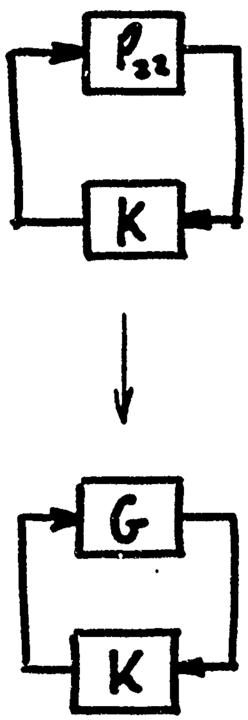
stabilizable &
detectable

lemma left system internally stable

$\Longleftrightarrow$ right " " "

... suffices to stabilize P_{22}

simplification of notation



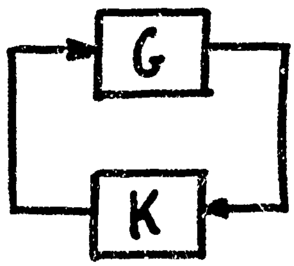
$$P_{22} = \left[\begin{array}{c|c} A & B_2 \\ \hline C_2 & D_{22} \end{array}\right]$$

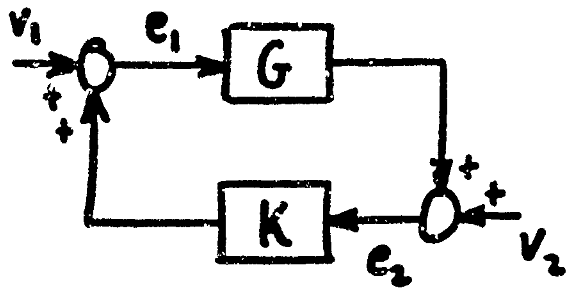
$$\downarrow$$

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

stabilizable
& detectable

internal stability vs input-output stability





$$\begin{bmatrix} I & -K \\ -G & I \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

internal stability $\Longleftrightarrow$ $\begin{bmatrix} I & -K \\ -G & I \end{bmatrix}^{-1} \in H_\infty$

1. well-posedness
2. internal stability
3. parametrization of K
4. state-space realization
5. closed-loop transfer matrix

factorizations over RH_∞

$f \in RH_\infty \iff f(s)$ real-rational,
proper,
analytic in $\mathrm{Re}\, s \geq 0$

... extend to vectors & matrices

def'n $M, N \in RH_\infty$ are right-coprime if

$\begin{bmatrix} M \\ N \end{bmatrix}$ is left-invertible in RH_∞

equivalently

1. $(\exists\, X, Y \in RH_\infty)\ XM + YN = I$
(Bezout eq'n)

2. every common right divisor of M, N is invertible in RH_∞

right-coprime factorization (rcf)

G real-rational $\Longrightarrow$

$(\exists M, N)$ M, N right-coprime

$$G = NM^{-1}$$

e.g.

$$\frac{1}{s-1} = \frac{\left(\frac{1}{s+1}\right)}{\left(\frac{s-1}{s+1}\right)}$$

left-coprime factorization (lcf)

...

(mini-lemma M, N right-coprime

$$NM^{-1} \in RH_\infty \Longleftrightarrow M^{-1} \in RH_\infty$$

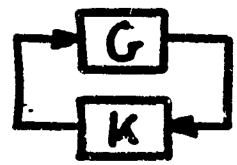
proof

$$XM + YN = I$$

$$X + YNM^{-1} = M^{-1} \quad)$$

characterization of internal stability

$$G = NM^{-1} = \tilde{M}^{-1}\tilde{N}$$
$$K = UV^{-1} = \tilde{V}^{-1}\tilde{U}$$



$$\begin{bmatrix} I & K \\ G & I \end{bmatrix}^{-1} = \begin{bmatrix} I & UV^{-1} \\ NM^{-1} & I \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} M & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} M & U \\ N & V \end{bmatrix}^{-1}$$

right-coprime

internal stability

$$\Longleftrightarrow \begin{bmatrix} I & -K \\ -G & I \end{bmatrix}^{-1} \in RH_\infty$$

$$\Longleftrightarrow \begin{bmatrix} I & K \\ G & I \end{bmatrix}^{-1} \in RH_\infty$$

$$\Longleftrightarrow \begin{bmatrix} M & U \\ N & V \end{bmatrix} \text{ invertible in } RH_\infty$$

$$\Longleftrightarrow \begin{bmatrix} \tilde{V} & -\tilde{U} \\ -\tilde{N} & \tilde{M} \end{bmatrix} \quad \text{" \quad " \quad "}$$

one stabilizer

choose (via next section) $M, N, \ldots$ s.t.

$$G = NM^{-1} = \tilde{M}^{-1}\tilde{N}$$

$$\begin{bmatrix} \tilde{V}_0 & -\tilde{U}_0 \\ -\tilde{N} & \tilde{M} \end{bmatrix}\begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix} = I$$

set $K_0 = U_0 V_0^{-1} = \tilde{V}_0^{-1}\tilde{U}_0$

all stabilizers

$$\begin{bmatrix} I & -Q \\ 0 & I \end{bmatrix}\begin{bmatrix} \tilde{V}_0 & -\tilde{U}_0 \\ -\tilde{N} & \tilde{M} \end{bmatrix}\underbrace{\begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix}\begin{bmatrix} I & Q \\ 0 & I \end{bmatrix}}_{\begin{bmatrix} M & U_0 + MQ \\ N & V_0 + NQ \end{bmatrix}} = I$$

$$K = (U_0 + MQ)(V_0 + NQ)^{-1}$$

$$= (\tilde{V}_0 + Q\tilde{N})^{-1}(\tilde{U}_0 + Q\tilde{M})$$

theorem formula for all stabilizers

$$K = (U_0 + MQ)(V_0 + NQ)^{-1}$$

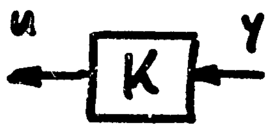
$$= (\tilde{V}_0 + Q\tilde{N})^{-1}(\tilde{U}_0 + Q\tilde{M})$$

$$= K_0 + \tilde{V}_0^{-1} Q (I + V_0^{-1} N Q)^{-1} V_0^{-1}$$

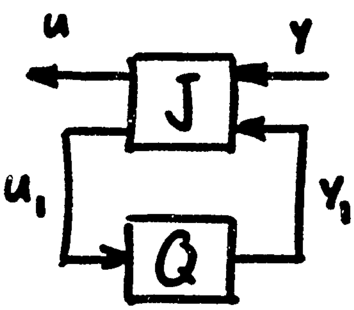
'free parameter' $Q \in RH_\infty$

$(V_0 + NQ)(\infty)$ invertible

block diagram



$$K = K_0 + \tilde{V}_0^{-1} Q (I + V_0^{-1} N Q)^{-1}$$



$$J = \begin{bmatrix} K_0 & \hat{V}_0^{-1} \\ V_0^{-1} & -V_0^{-1} N \end{bmatrix}$$

recap

get $M, N, \ldots$ s.t.

$$G = NM^{-1} = \tilde{M}^{-1}\tilde{N}$$

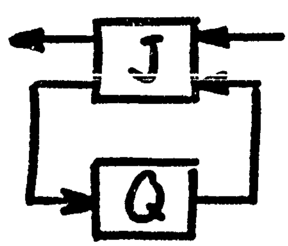
$$\begin{bmatrix} \tilde{V}_0 & -\tilde{U}_0 \\ -\tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix} = I$$

set

$$K_0 = U_0 V_0^{-1}$$

$$J = \begin{bmatrix} K_0 & \tilde{V}_0^{-1} \\ V_0^{-1} & -V_0^{-1}N \end{bmatrix}$$

then K is parametrized via



1. well-posedness
2. internal stability
3. parametrization of K
4. state-space realization
5. closed-loop transfer matrix

state-space realization of rcf

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

choose F s.t. $A+BF$ stable

$$y = Gu$$

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \end{cases}$$

$$\begin{cases} \dot{x} = (A+BF)x + B(u - Fx) \\ y = (C+DF)x + D(u-Fx) \end{cases}$$

$$\begin{cases} \dot{x} = (A+BF)x + Bw \\ u = Fx + w \\ y = (C+DF)x + Dw \end{cases} \qquad (w := u - Fx)$$

$$\begin{cases} u = \{F[sI-(A+BF)]^{-1}B + I\}w \\ \quad =: Mw \\ y = \{(C+DF)[sI-(A+BF)]^{-1}B + D\}w \\ \quad =: Nw \end{cases}$$

$$y = NM^{-1}u$$

conclusion

$$G = N M^{-1}$$

$$\begin{bmatrix} M \\ N \end{bmatrix} = \left[\begin{array}{c|c} A+BF & B \\ \hline F & I \\ C+DF & D \end{array}\right]$$

state-space realization of lcf

choose H s.t. $A+HC$ stable

$$G = \tilde{M}^{-1} \tilde{N}$$

$$[\,\tilde{M} \;\; \tilde{N}\,] = \left[\begin{array}{c|cc} A+HC & H & B+HD \\ \hline C & I & D \end{array}\right]$$

recall

we want

$$\begin{bmatrix} \tilde{V}_0 & -\tilde{U}_0 \\ -\tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix} = I$$

method

1. get a stabilizer K_0 via observer
2. do state-space realizations of rcf and lcf of K_0

step 1

$$\begin{cases} \dot{\hat{x}} = A\hat{x} + Bu + H(C\hat{x} + Du - y) \\ u = F\hat{x} \end{cases}$$

$$K_0 = \left[\begin{array}{c|c} A+BF+HC+HDF & -H \\ \hline F & 0 \end{array}\right]$$

step 2

$$K_0 = U_0 V_0^{-1} = \tilde{V}_0^{-1} \tilde{U}_0$$

$$\begin{bmatrix} V_0 \\ U_0 \end{bmatrix} = \left[\begin{array}{c|c} A+BF & -H \\ \hline C+DF & I \\ F & 0 \end{array}\right]$$

$$[\tilde{V}_0 \quad \tilde{U}_0] = \left[\begin{array}{c|cc} A+HC & -(B+HD) & -H \\ \hline F & I & 0 \end{array}\right]$$

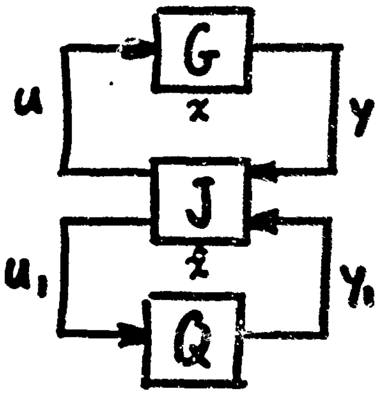
$$V_0 = \left[\begin{array}{c|c} A+BF & -H \\ \hline C+DF & I \end{array}\right] \qquad \tilde{V}_0 = \left[\begin{array}{c|c} A+HC & -(B+HD) \\ \hline F & I \end{array}\right]$$

$$U_0 = \left[\begin{array}{c|c} A+BF & -H \\ \hline F & 0 \end{array}\right] \qquad N = \left[\begin{array}{c|c} A+BF & B \\ \hline C+DF & D \end{array}\right]$$

$$J = \begin{bmatrix} U_0 V_0^{-1} & \tilde{V}_0^{-1} \\ V_0^{-1} & -V_0^{-1} N \end{bmatrix}$$

$$= \left[\begin{array}{c|cc} A+BF+HC+HDF & -H & B+HD \\ \hline F & 0 & I \\ -(C+DF) & I & -D \end{array}\right]$$

block diagrams



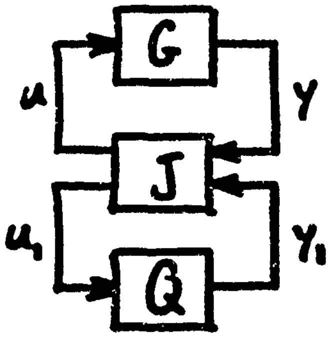
$$G = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$$ for simplicity

$$J = \left[\begin{array}{c|cc} A+BF+HC & -H & B \\ \hline F & 0 & I \\ -C & I & 0 \end{array}\right]$$

$$G \begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases}$$

$$J \begin{cases} \dot{\hat{x}} = (A+BF+HC)\hat{x} - Hy + By_1 \\ u = F\hat{x} + y_1 \\ u_1 = -C\hat{x} + y \end{cases}$$

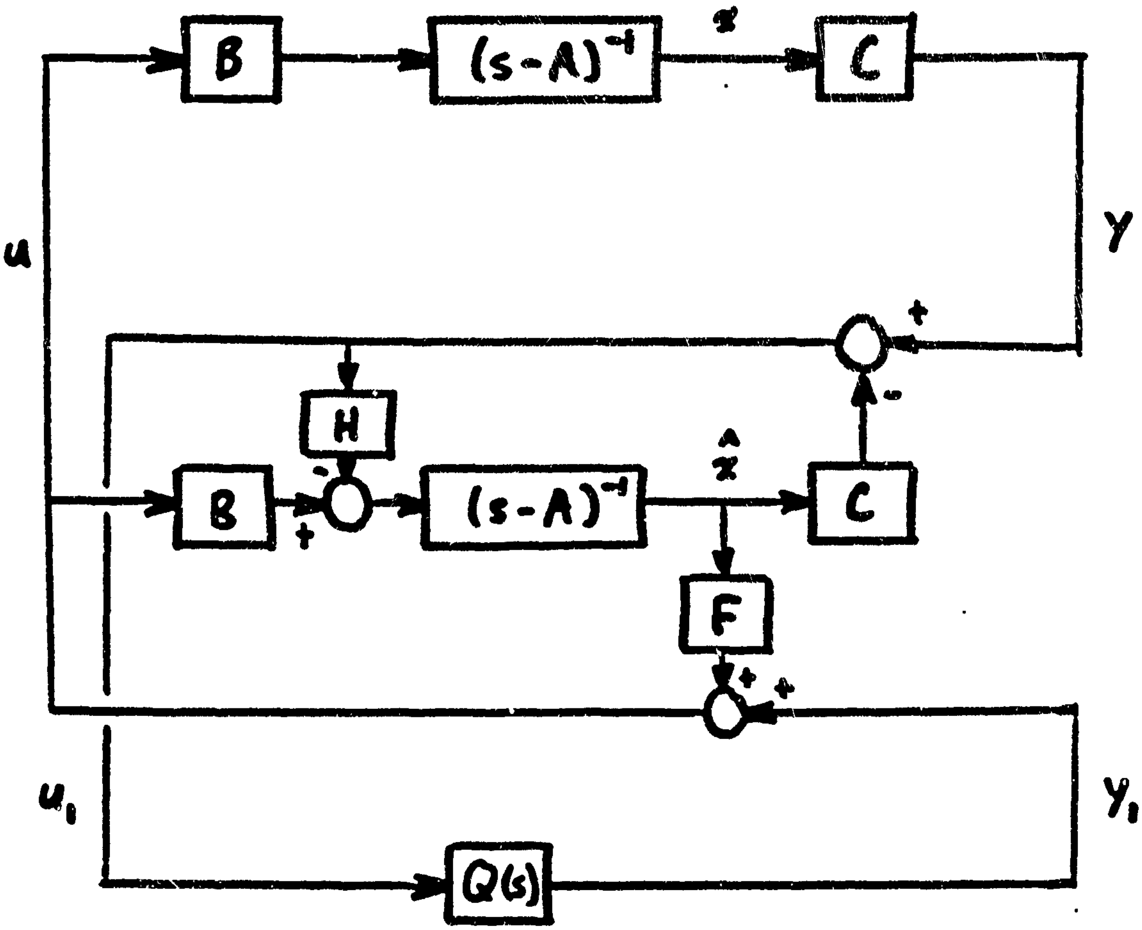
$$\begin{cases} \dot{\hat{x}} = A\hat{x} + Bu - Hu_1 \\ u = F\hat{x} + y_1 \\ u_1 = -C\hat{x} + y \end{cases}$$



$$u_1 = y - \hat{y}$$
$$= Cx - C\hat{x}$$
$$= Ce\,, \qquad \dot{e} = (A+HC)e$$

$\therefore$ transfer function

$y_1 \longrightarrow u_1$ is zero!



$$Q = \left[\begin{array}{c|c} A_a & B_a \\ \hline C_a & 0 \end{array}\right] \qquad A_a \text{ stable}$$

representation theorem

G, K strictly proper

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$$

if K is a stabilizer, then

$$G \longleftrightarrow G_e = \left[\begin{array}{cc|c} A & 0 & B \\ 0 & A_a & 0 \\ \hline C & 0 & 0 \end{array}\right]$$

$$=: \left[\begin{array}{c|c} A_e & B_e \\ \hline C_e & 0 \end{array}\right]$$

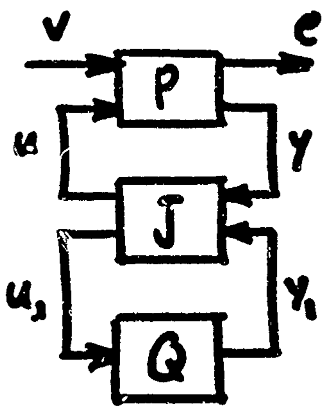
A_a stable

$(\exists F_e, H_e)$

$$K = \left[\begin{array}{c|c} A_e + B_e F_e + H_e C_e & -H_e \\ \hline F_e & 0 \end{array}\right]$$

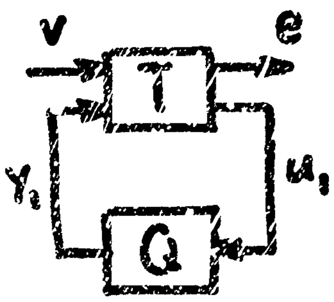
conclusion stabilization $\Longleftrightarrow$
augment with stable dynamics,
use observer-based-controller

1. well-posedness
2. internal stability
3. parametrization of K
4. state-space realization
5. closed-loop transfer matrix



$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

$$J = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}$$



$$T_{22} = 0 \qquad T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$$

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} = \begin{bmatrix} P_{11} & 0 \\ 0 & J_{22} \end{bmatrix} + \begin{bmatrix} P_{12} & 0 \\ 0 & J_{21} \end{bmatrix} \begin{bmatrix} I & -J_{11} \\ -P_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & J_{12} \\ P_{21} & 0 \end{bmatrix}$$

Conclusion

parametrization of closed-loop transfer matrix

$$e = (T_{11} + T_{12} Q T_{21}) v$$

$$T_{11} = \left[\begin{array}{cc|c} A+B_2F & -B_2F & B_1 \\ 0 & A+HC_2 & B_1+HD_{21} \\ \hline C_1+D_{12}F & -D_{12}F & D_{11} \end{array}\right]$$

$$T_{12} = \left[\begin{array}{c|c} A+B_2F & B_2 \\ \hline C_1+D_{12}F & D_{12} \end{array}\right]$$

$$T_{21} = \left[\begin{array}{c|c} A+HC_2 & B_1+HD_{21} \\ \hline C_2 & D_{21} \end{array}\right]$$

$Q \in RH_\infty$

$I + D_{22} Q(\infty)$ invertible

H_∞ Approximation

problem

given $R(s)$ real-rational
strictly proper
analytic in $\operatorname{Re} s \le 0$
(anti-stable)

find $Q(s)$ real-rational
proper
analytic in $\operatorname{Re} s \ge 0$
(stable)

such that

$$\| R - Q \|_\infty = \text{minimum}$$

where

$$\| F \|_\infty := \sup_\omega \bar{\sigma} [F(j\omega)]$$

preview

1. $\min \| R - Q \|_\infty$

=: dist (R, stable matrices)

= ‖ Hankel operator of R ‖

2. an optimal Q can be computed by state-space methods

key reference

Glover '84

topics

1. function spaces
2. Hankel operators
3. the distance formula
4. some history
5. state-space sol'n
6. general distance formula

time domain

L_2 : $f(t) \in \mathbb{C}^n$, $-\infty < t < \infty$

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(t)^* g(t)\, dt$$

$$\| f \|_2 = \left[\int_{-\infty}^{\infty} f(t)^* f(t)\, dt \right]^{1/2}$$

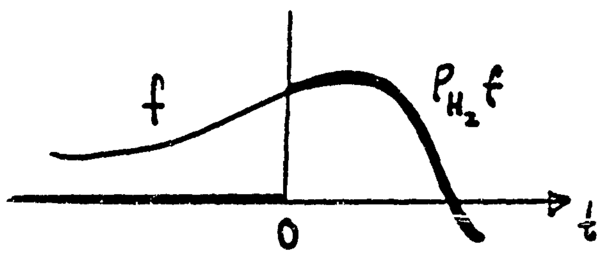
H_2 : $f(t) = 0$, $t < 0$

(causal part of L_2)

$H_2^\perp$: $f(t) = 0$, $t > 0$

(anti-causal part of L_2)

$P_{H_2} : L_2 \longrightarrow H_2$ truncation projection



$P_{H_2^\perp} : L_2 \longrightarrow H_2^\perp$

frequency domain

L_2 : $f(j\omega) \in \mathbb{C}^n$, $-\infty < \omega < \infty$

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(j\omega)^* g(j\omega)\, d\omega$$

$$\| f \|_2 = \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} f(j\omega)^* f(j\omega)\, d\omega \right]^{1/2}$$

H_2 : $f(s)$ analytic in $\mathrm{Re}\, s > 0$

(stable part of L_2) $\quad \sup_{\sigma > 0} \int_{-\infty}^{\infty} f(\sigma + j\omega)^* f(\sigma + j\omega)\, d\omega < \infty$

$H_2^\perp$: $f(s)$ analytic in $\mathrm{Re}\, s < 0$

(anti-stable part of L_2) $\quad \sup_{\sigma < 0} \int_{-\infty}^{\infty} f(\sigma + j\omega)^* f(\sigma + j\omega)\, d\omega < \infty$

$P_{H_2} : L_2 \longrightarrow H_2 \qquad P_{H_2^\perp} : L_2 \longrightarrow H_2^\perp$

L_∞: $F(j\omega) \in \mathbb{C}^{m \times n}$

$$\|F\|_\infty = \sup_\omega \bar{\sigma}\left[F(j\omega)\right]$$

H_∞: $F(s)$ analytic & bounded in $\mathrm{Re}\, s > 0$

(stable part of L_∞)

rational functions

$f(s)$ rational.

$f \in L_2 \iff$ strictly proper, no poles on $\mathrm{Re}\, s = 0$

$f \in H_2 \iff$ " " " " in $\mathrm{Re}\, s \geq 0$

$f \in H_2^\perp \iff$ " " " " " $\mathrm{Re}\, s \leq 0$

$f \in L_\infty \iff$ proper, no poles on $\mathrm{Re}\, s = 0$

$f \in H_\infty \iff$ " " " in $\mathrm{Re}\, s \geq 0$

prefix $\mathcal{R}$: real-rational

connection

time domain freq domain

$$f(t) \xleftrightarrow[\text{transform}]{\text{Fourier}} f(j\omega)$$

$$L_2 \xleftrightarrow{\text{isomorphism}} L_2$$

$$H_2 \xleftrightarrow{\text{isomorphism}} H_2$$

1. function spaces
2. Hankel operators
3. the distance formula
4. some history
5. state-space sol'n
6. general distance formula

frequency domain

$R(s)$ real-rational
strictly proper
analytic in $\mathrm{Re}\, s \le 0$
(anti-stable)

$\therefore\ R \in H_2^{\perp}$

time domain

$r(t)$ = inverse Fourier trans of $R(j\omega)$

$r(t) = 0$ for $t > 0$
(anti-causal)

$\therefore\ r \in H_2^{\perp}$

def'n of Hankel operator

frequency domain

$$\Gamma_R : H_2 \longrightarrow H_2^{\perp}$$

$$\Gamma_R f = P_{H_2^{\perp}} R f$$

$$\begin{tikzcd}
H_2 \arrow[r, "\text{multiply by } R"] \arrow[dr, "\Gamma_R"'] & L_2 \arrow[d, "\text{take anti-stable part}"] \\
& H_2^{\perp}
\end{tikzcd}$$

time domain

$$\Gamma_r : H_2 \longrightarrow H_2^{\perp}$$

$$\Gamma_r f = P_{H_2^{\perp}} r * f$$

$$\begin{tikzcd}
H_2 \arrow[r, "\text{convolve with } r"] \arrow[dr, "\Gamma_r"'] & L_2 \arrow[d, "\text{take anti-causal pa}"] \\
& H_2^{\perp}
\end{tikzcd}$$

isomorphisms $\Longrightarrow \| \Gamma_R \| = \| \Gamma_r \|$

adjoint operator

time domain

$$\check{r}(t) := r(-t)'$$

$$\Gamma_r^* : H_2^\perp \longrightarrow H_2$$

$$\Gamma_r^* f = P_{H_2} \check{r} * f$$

$$\begin{tikzcd}
H_2^\perp \arrow[r, "\substack{\text{convolve} \\ \text{with } \check{r}}"] \arrow[dr, "\Gamma_r^*"'] & L_2 \arrow[d, "\substack{\text{take} \\ \text{causal par}}"] \\
& H_2
\end{tikzcd}$$

state-space representations

$$R = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \quad \text{minimal}$$

eigenvalues of A in $\operatorname{Re} z > 0$

controllability gramian

$$Y = \int_{-\infty}^{0} e^{At} B B' e^{A't} \, dt$$

observability gramian

$$X = \int_{-\infty}^{0} e^{A't} C' C e^{At} \, dt$$

lemma $\Gamma_r^* \Gamma_r$ and YX have the same nonzero eigenvalues.

$$\therefore \quad \| \Gamma_r \| = \left[\lambda_{max} (YX) \right]^{1/2}$$

proof

$$\Gamma \Longleftarrow \Gamma_r$$

let

$\sigma^2 \neq 0$ an eigenvalue of $\Gamma^*\Gamma$

$f \in H_2$ a corr eigenvector

then $\Gamma^*\Gamma f = \sigma^2 f$

define

$$h := \frac{1}{\sigma}\Gamma f$$

then

$$\begin{cases} \Gamma f = \sigma h \\ \Gamma^* h = \sigma f \end{cases}$$ (Schmidt pair)

where

$$\begin{cases} \Gamma f = P_{H_2^\perp} \, r * f \\ \Gamma^* h = P_{H_2} \, \check{r} * h \end{cases}$$

$$\begin{cases} r(t) = C e^{At} B \,, & t < 0 \\ \check{r}(t) = B' e^{-A't} C' \,, & t > 0 \end{cases}$$

$$\begin{cases} \Gamma f = \sigma h \\ \Gamma^* h = \sigma f \end{cases}$$

$$\begin{cases} \displaystyle\int_0^\infty C e^{A(t-\tau)} B f(\tau)\,d\tau = \sigma h(t), & t<0 \\ \displaystyle\int_{-\infty}^0 B' e^{A'(\tau-t)} C' h(\tau)\,d\tau = \sigma f(t), & t>0 \end{cases}$$

$$\begin{cases} C e^{At}\Big[\underbrace{\displaystyle\int_0^\infty e^{-A\tau} B f(\tau)\,d\tau}_{=:\,v}\Big] = \sigma h(t), & t<0 \\ B' e^{-A't}\Big[\underbrace{\displaystyle\int_{-\infty}^0 e^{A'\tau} C' h(\tau)\,d\tau}_{=:\,w}\Big] = \sigma f(t), & t \end{cases}$$

$$\begin{cases} e^{A't} C' C e^{At} v = \sigma e^{A't} C' h(t), & t<0 \\ e^{-At} B B' e^{-A't} w = \sigma e^{-At} B f(t), & t>0 \end{cases}$$

integrate

$$\begin{cases} X v = \sigma w \\ Y w = \sigma v \end{cases}$$

$$Y X v = \sigma^2 v \qquad //$$

conclusion

to compute $\|\Gamma_r\|$

do a realization $R = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$

solve Lyapunov eq'ns

$$AY + YA' + BB' = 0$$
$$A'X + XA + C'C = 0$$

then

$$\|\Gamma_r\| = \left[\lambda_{max}(YX)\right]^{1/2}$$

1. function spaces
2. Hankel operators
3. the distance formula
4. some history
5. state-space sol'n
6. general distance formula

theorem

$$\operatorname{dist}(R, \text{stable matrices})$$
$$= \operatorname{dist}(R, H_\infty)$$
$$= \| \Gamma_r \|$$

easy inequality ($\geq$)

$$F \in L_\infty$$

$$M_F : H_2 \xrightarrow[\text{by } F]{\text{multiply}} L_2$$

$$\| F \|_\infty = \| M_F \|$$

$$= \sup_{\substack{g \in H_2 \\ \|g\|_2 = 1}} \| Fg \|_2$$

let $Q \in H_\infty$

$$\|R - Q\|_\infty = \| M_R - M_Q \|$$

$$\geq \| P_{H_2^\perp} (M_R - M_Q) \|$$

$$= \| P_{H_2^\perp} M_R \|$$

$$\begin{tikzcd} H_2 \arrow[r, "M_Q"] \arrow[dr, "0"'] & L_2 \arrow[d] \\ & H_2^\perp \end{tikzcd}$$

$$= \| \Gamma_R \|$$

$\therefore$ dist $(R, H_\infty) \geq \| \Gamma_R \|$

... remains to show $\exists Q$ s.t.

$$\| R - Q \|_\infty = \| \Gamma_R \|$$

function spaces : unit disk

L_2 : $f(e^{j\theta}) \in \mathbb{C}^n$, $0 \le \theta < 2\pi$

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(e^{j\theta})^* g(e^{j\theta}) d$$

$$f(e^{j\theta}) = \sum_{-\infty}^{\infty} f_k e^{jk\theta}$$

H_2 : $f(e^{j\theta}) \in L_2$

(stable part of L_2) $f(z)$ analytic in $|z| < 1$

$$f(z) = \sum_{0}^{\infty} f_k z^k$$

$H_2^\perp$: $f(e^{j\theta}) \in L_2$

(anti-stable part of L_2) $f(z)$ analytic in $|z| > 1$

$$f(z) = \sum_{-\infty}^{-1} f_k z^k$$

L_∞: $F(e^{j\theta}) \in \mathbb{C}^{m \times n}$

$$\|F\|_\infty = \sup_\theta \bar{\sigma}\left[F(e^{j\theta})\right]$$

H_∞: $F(e^{j\theta}) \in L_\infty$

(stable part of L_∞) $F(z)$ analytic in $|z| < 1$

rational functions

$f(z)$ rational

$f \in L_2 \Longleftrightarrow f \in L_\infty \Longleftrightarrow$ no poles on $|z| = 1$

$f \in H_2 \Longleftrightarrow f \in H_\infty \Longleftrightarrow$ " " in $|z| \leq 1$

$f \in H_2^\perp \Longleftrightarrow$ strictly proper, no poles in $|z| \geq$

approximation problem

$R(z)$ real-rational
strictly proper
analytic in $|z| \geq 1$
$\therefore R \in H_2^{\perp}$

Hankel operator

$$\Gamma_R : H_2 \longrightarrow H_2^{\perp}$$

$$\Gamma_R f = P_{H_2^{\perp}} R f$$

matrix rep of Γ_R

$$f(z) = \sum_0^{\infty} f_k z^k \quad , \quad R(z) = \sum_{-\infty}^{-1} R_k z^k$$

$$h := \Gamma_R f \, , \quad h(z) = \sum_{-\infty}^{-1} h_k z^k$$

$$\begin{bmatrix} h_{-1} \\ h_{-2} \\ h_{-3} \\ \vdots \\ \\ \end{bmatrix} = \begin{bmatrix} R_{-1} & R_{-2} & R_{-3} & \cdots \\ R_{-2} & R_{-3} & R_{-4} & \cdots \\ R_{-3} & R_{-4} & & \\ R_{-4} & & & \\ \vdots & & & \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ \vdots \\ \\ \end{bmatrix}$$

find $Q \in H_\infty$ s.t.

$$\| R + Q \|_\infty = \| \Gamma_R \|$$

$$\| R + Q \|_\infty = \| M_{R+Q} \|$$

$$H_2 \xrightarrow{M_{R+Q}} L_2$$

matrix rep of M_{R+Q}

$$\begin{bmatrix} \vdots & & \\ Q_1 & Q_0 & \\ Q_0 & R_{-1} & \\ R_{-1} & R_{-2} & \\ R_{-2} & R_{-3} & \\ R_{-3} & & \\ & & \end{bmatrix}$$

$$\left[\begin{array}{c|cc}
\vdots & \vdots & \\
Q_1 & Q_2 & \cdots \\
\hline
Q_0 & R_{-1} & \cdots \\
\hline
R_{-1} & R_{-2} & \cdots \\
R_{-2} & R_{-3} & \cdots \\
\vdots & \vdots &
\end{array}\right]$$

bottom

choose Q_0 to minimize $\|$ bottom $\|$

min $\|$ bottom $\|$ $= \| \Gamma_R \|$

(Parrott / Davis - Kahan - Weinberger)

$$
\left[\begin{array}{c|cc}
\vdots & \vdots & \\
Q_2 & Q_1 & \cdots \\
\hline
Q_1 & Q_0 & \cdots \\
\hline
Q_0 & R_{-1} & \cdots \\
R_{-1} & R_{-2} & \cdots \\
R_{-2} & R_{-3} & \cdots \\
\vdots & \vdots &
\end{array}\right]
\quad \Big\downarrow \text{ bottom}
$$

choose Q_1 to minimize $\| \text{bottom} \|$

$$\min \| \text{bottom} \| = \| \Gamma_R \|$$

etc //

... end of proof of distance formula

1. function spaces
2. Hankel operators
3. the distance formula
4. some history
5. state-space sol'n
6. general distance formula

1. Nevanlinna - Pick theory

Pick '16
Nevanlinna '19 } scalar case

Delsarte, Genin, Kamp '79 } matrix case

define

$$\{p_i\} = \text{poles of } R$$

$$b(s) = \prod_i (s-p_i)/(s+p_i)$$

then

$$|b(j\omega)| = 1 \quad \text{(all-pass)}$$

$$R_1 := bR \in H_\infty$$

$$\|R - Q\|_\infty = \|R_1 - bQ\|_\infty$$

let

$$Q_1 := R_1 - bQ$$

then

$$Q \in H_\infty \iff Q_1 \in H_\infty \text{ and } Q_1(p_i) = R_1(p_i)$$

so problem becomes

$$\min \|Q_1\|_\infty \; : \; Q_1 \in H_\infty, \; Q_1 = R_1 \text{ at } p_i$$

2. Sarason theory '67

3. Adamjan - Arov - Krein theory '71, '78

eg scalar case

let $\sigma = \| \Gamma_R \|$

(f, h) a Schmidt pair

$$\begin{cases} \Gamma_R f = \sigma h & f \in H_2 ,\ h \in H_2^{\perp} \\ \Gamma_R^{*} h = \sigma f & \| f \|_2 = \| h \|_2 = 1 \end{cases}$$

then

$$R - Q_{opt} = \sigma \, h / f$$

4. Ball - Helton theory '83

1. function spaces
2. Hankel operators
3. the distance formula
4. some history
5. state-space sol'n
6. general distance formula

key tools

Lyapunov lemma

$$\begin{cases} (A,B) \text{ stabilizable} \\ (\exists P)\ P \geq 0,\ AP + PA' + BB' = 0 \end{cases}$$

$\Longrightarrow$ A stable (eigenvalues in $\operatorname{Re} s < 0$)

little lemma

$$\begin{cases} G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \\ (\exists P) \quad AP + PA' + BB' = 0 \\ \qquad\qquad CP + DB' = 0 \end{cases}$$

$$\Longrightarrow \quad G(s)\,G(-s)' = DD'$$

proof of little lemma

$$G(s) = D + \underbrace{C(s-A)^{-1}B}_{=:\,G_0(s)}$$

$$AP + PA' + BB' = 0$$

$$BB' = (s-A)P + P(-s-A')$$

$$C(s-A)^{-1}\Big[\qquad\qquad\qquad\Big](-s-A')^{-1}C'$$

$$G_0(s)G_0(-s)' = \underbrace{CP}_{-DB'}(-s-A')^{-1}C' + C(s-A)^{-1}\underbrace{PC'}_{-BD'}$$

$$= -DG_0(-s)' - G_0(s)D'$$

$$G(s)G(-s)' = \big[D + G_0(s)\big]\big[D' + G_0(-s)'\big]$$

$$= DD' + \big[DG_0(-s)'$$

$$+ G_0(s)D' + G_0(s)G_0(-$$

$$= DD' \quad //$$

... back to approximation problem

realization

$$R = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} \quad \text{minimal}$$

define

$$\sigma := \| \Gamma_r \|$$

to be found :

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

error system

$$G_e := R - Q$$

$$= \left[\begin{array}{cc|c} A & 0 & B \\ 0 & \hat{A} & \hat{B} \\ \hline C & -\hat{C} & -\hat{D} \end{array}\right] =: \left[\begin{array}{c|c} A_e & B_e \\ \hline C_e & D_e \end{array}\right]$$

conclusion Q is optimal if

$\hat{A}$ is stable , $\| G_e \|_\infty = \sigma$

lemma

suppose

1. $(\hat{A}, \hat{B})$ stabilizable

2. $(\exists P_e)$ $A_e P_e + P_e A_e' + B_e B_e' = 0$

$$C_e P_e + D_e B_e' = 0$$

$$D_e D_e' = \sigma^2 I$$

$$\hat{P} \geq 0 \quad \text{where} \quad P_e = \begin{bmatrix} \cdot & \cdot \\ \cdot & \hat{P} \end{bmatrix}$$

then Q is optimal

proof

little lemma $\Longrightarrow G_e(s)\, G_e(-s)' = \sigma^2 I$

$$\Longrightarrow \| G_e \|_\infty = \sigma$$

also $\hat{A}\hat{P} + \hat{P}\hat{A}' + \hat{B}\hat{B}' = 0$

Lyapunov lemma $\Longrightarrow$ $\hat{A}$ stable //

summary

find (A, B, C), $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$, P_e s.t.

$$R = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \quad \text{minimal}$$

$(\hat{A}, \hat{B})$ stabilizable

$$A_e P_e + P_e A_e' + B_e B_e' = 0$$

$$C_e P_e + D_e B_e' = 0$$

$$D_e D_e' = \sigma^2 I$$

$$\hat{P} \geq 0$$

then

$$Q = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] \quad \text{is optimal}$$

construction

step 1

find singular values of Γ_r

$$\underbrace{\sigma_1 = \cdots = \sigma_k}_{=:\,\sigma} > \sigma_{k+1} \geq \cdots \qquad \Sigma = \begin{bmatrix} \sigma_{k+1} & & \\ & \ddots & \\ & & \ddots \end{bmatrix}$$

find a balanced realization of R

$$R = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \quad \text{minimal}$$

controllability gramian

= observability gramian

$$= - \begin{bmatrix} \sigma I & 0 \\ 0 & \Sigma \end{bmatrix}$$

step 2 partition

$$A = \underbrace{\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}}_{k} \updownarrow k \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \quad C = [\, C_1 \;\; C_2 \,]$$

step 3

choose $\hat{D}$ s.t.

$$\hat{D} B_1' + \sigma C_1 = 0$$

$$\hat{D} \hat{D}' = \sigma^2 I$$

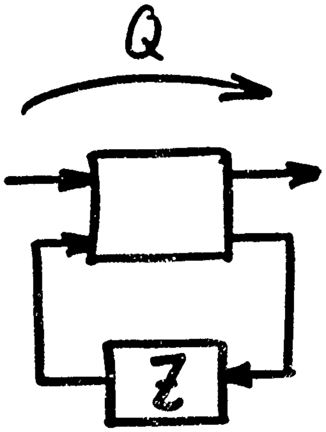
step 4

$$\hat{B} := -(\sigma^2 - \Sigma^2)^{-1} (\Sigma B_2 + \sigma C_2' \hat{D})$$

$$\hat{A} := (-A_{22} + B_2 \hat{B}')'$$

$$\hat{C} := C_2 \Sigma + \hat{D} B_2'$$

parametrization of all optimal Q's



'free parameter'

$Z \in H_\infty$, $\|Z\|_\infty \leq 1$

Q = linear fractional transformation of Z

1. function spaces
2. Hankel operators
3. the distance formula
4. some history
5. state-space sol'n
6. general distance formula

general approximation problem

given

$R(s)$ real-rational
strictly proper
analytic in $Re\ s \leq 0$

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

find

$Q(s)$ real-rational
proper
analytic in $Re\ s \geq 0$

such that

$$\left\| \begin{bmatrix} R_{11} - Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_{\infty} = \text{minimum}$$

$$\min = \operatorname{dist}\left(R , \begin{bmatrix} H_\infty & 0 \\ 0 & 0 \end{bmatrix} \right)$$

$$\begin{matrix} H_2 \\ \oplus \\ H_2 \end{matrix} \xrightarrow{\quad M_R \quad} \begin{matrix} L_2 \\ \oplus \\ L_2 \end{matrix}$$

$$M_R \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

$$\left\| R - \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \right\|_\infty = \| M_R - M_{\left[\begin{smallmatrix} Q & 0 \\ 0 & 0 \end{smallmatrix}\right]} \|$$

$$\geq \left\| \begin{bmatrix} P_{H_2^\perp} & 0 \\ 0 & I \end{bmatrix} \left(M_R - M_{\left[\begin{smallmatrix} Q & 0 \\ 0 & 0 \end{smallmatrix}\right]} \right) \right\|$$

$$= \underbrace{\left\| \begin{bmatrix} P_{H_2^\perp} & 0 \\ 0 & I \end{bmatrix} M_R \right\|}_{=:\, \Gamma}$$

theorem

$$dist\left(R, \begin{bmatrix} H_\infty & 0 \\ 0 & 0 \end{bmatrix} \right) = \| \Gamma \|$$