INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of “sectioning” the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.
Wang, Zheng-Zhi

SYNTHESIS OF LINEAR MULTIVARIABLE FEEDBACK SYSTEMS IN INFINITE INDEX NORM

Rice University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

PH.D. 1985
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark √.

1. Glossy photographs or pages ______
2. Colored illustrations, paper or print ______
3. Photographs with dark background ______
4. Illustrations are poor copy √
5. Pages with black marks, not original copy ______
6. Print shows through as there is text on both sides of page ______
7. Indistinct, broken or small print on several pages √
8. Print exceeds margin requirements ______
9. Tightly bound copy with print lost in spine ______
10. Computer printout pages with indistinct print ______
11. Page(s) ______ 110 lacking when material received, and not available from school or author.
12. Page(s) ______ seem to be missing in numbering only as text follows.
13. Two pages numbered ______. Text follows.
14. Curling and wrinkled pages ______
15. Dissertation contains pages with print at a slant, filmed as received ______
16. Other ________________________________
    ___________________________________
    ___________________________________
    ___________________________________

University
Microfilms
International
RICE UNIVERSITY

SYNTHESIS OF LINEAR MULTIVARIABLE FEEDBACK SYSTEMS IN INFINITE INDEX NORM

by

ZHENG-ZHI WANG

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

APPROVED, THESIS COMMITTEE:

Dr. J. B. Pearson
Professor of Electrical Engineering
Chairman

Dr. Rui J. P. de Figueiredo
Professor of Electrical Engineering and Mathematical Sciences

Dr. Richard A. Tapia
Professor of Mathematical Sciences

HOUSTON, TEXAS

AUGUST 1984
SYNTHESIS OF LINEAR MULTIVARIABLE FEEDBACK
SYSTEMS IN INFINITE INDEX NORM

by

ZHENG-ZHI WANG

ABSTRACT

The deficiency of the widely used LQG method is that depends heavily on the precision of plant parameters and the noise spectrum. The robustness problem can be formalized in singular value analysis. With the help of operator theory a new method for the synthesis of linear multivariable feedback systems in $H_\infty$ norm is developed from the singular value analysis. The positive feature of $H_\infty$ norm synthesis is the transparency for robustness conditions, the weighting functions are directly related to the specifications of design requirements.

In this dissertation the LQG problem is restated as an interpolation problem in $H_2$ space. The interpolation problem in simplest case can be solved by an explicit formula. The $H_\infty$ optimal norm can be obtained from the consideration of the ratio of two $H_2$ norms. The close relations and similarities between $H_\infty$ and $H_2$ are brought out. The total $H_\infty$ optimal solutions can be constructed by the unitary dilation from the interpolation space. The explicit formulas in $s$ domain for these purposes are given, including the repeated zeros case and the degenerate case. The optimal solutions must belong to the degenerate case, in this case the problem can be solved by separating the singular part of Pick matrix from
the regular part by a Cholesky decomposition. These results are also developed in a recursive version for repeated zeros.

The zeros and interpolation condition vectors of a system can be determined numerically by an algorithm to solve eigenvalues and eigenvectors of a pencil. To convert the two-sided problem to a one-sided problem and to convert the nonsquare problem to a square problem are related to the spectral factorization which is discussed in detail.

The optimal solutions of the nonsquare problem need not be all-pass, which is related to the existence of a critical point.

The theory applied to the sensitivity design problem can be considered as an extension of the classical lead-lag design method from SISO to MIMO with more profound mathematical background. The robust stability problem can also be formalized and solved in the framework. The robust sensitivity design introduces a new type of mathematical problem, which can be approximated in our framework in certain situations. The regulation, tracking, filtering and optimal controller design problem under the inexact knowledge of noise spectrum can be solved in the general model in $H_\infty$ space by introducing proper weighting functions.
ACKNOWLEDGEMENTS

I sincerely thank Professor J. B. Pearson, my advisor, for the chance he gave me to study the multivariable control theory under his guidance, for his continual encouragement, enlightenment and advice in my research work.

I am grateful to professor R. J. Defigueiredo, professor R. A. Tapia for serving on my thesis committee and their helpful advice.

I would like to thank my schoolmates Dr. Chang, Mr. Ganesh, Mr. Datta for their many discussions and helps.

I want to express my gratitude to Rice University where my study abroad is so happy and meaningful.
# Table of Contents

1. Introduction ............................................ 1

2. Hardy Space and Krein Space
   2.1 Introduction ........................................ 9
   2.2 Hardy Space ........................................ 9
   2.3 Krein Space ......................................... 14

3. $H_2$ Optimal Norm Problem
   3.1 Introduction ....................................... 17
   3.2 Interpolation in $H_2$ Space ...................... 17
   3.3 Inner and Outer Factorization .................... 20
   3.4 Two-sided Problem ................................ 24
   3.5 The Case of Repeated Zeros ....................... 26

4. $H_\infty$ Optimal Norm
   4.1 Introduction ....................................... 32
   4.2 The Case of Distinct Zeros ....................... 32
   4.3 The Case of Repeated Zeros ....................... 37
   4.4 Another Derivation ................................ 39

5. Construction of All $\Phi$ in Unit Ball
   5.1 Introduction ....................................... 48
   5.2 The Case of Nondegenerate and Distinct Zeros .... 49
   5.3 The Case of Degenerate and Distinct Zeros ....... 61
   5.4 The Case of Degenerate and Repeated Zeros ....... 75
   5.5 Two-sided Problem in Square Case ................. 77

6. Recursive Algorithm
   6.1 Introduction ....................................... 83
6.2 Interpolation Pair and Recursive Method
6.3 Treatment of Degeneration
6.4 Interpolation on Imaginary Axis or at Infinity

7. NONSQUARE PROBLEM

7.1 Introduction
7.2 Spectral Factrization
7.3 Recursive Algorithm
7.4 Optimal Infinite Index Norm Problem in Nonsquare Case

8. SYNTHESIS OF CONTROL SYSTEMS

8.1 Introduction
8.2 Design of Sensitivity
8.3 Robust Stability of Plant Uncertainty
8.4 Robust Sensitivity Design under Plant Uncertainty
8.5 General Model

9. CONCLUSIONS

10. BIBLIOGRAPHY
CHAPTER 1 INTRODUCTION

Since N. Wiener proposed the Wiener-Hopf method in his book "Extrapolation, Interpolation and Smoothing of Stationary Time Series" in 1947 and R. E. Bellman invented dynamic programming in 1953, scientists and engineers in control area soon broke through the obstacles laid in their advanced road. The elegant LQR and LQG method thus became very popular tools of system theory, the quadratic formulation was used because anything else would result in a nonlinear optimization problem. At the beginning of 60's, Kalman considered the problem in state space description and separated it into a recursive estimation filter and a feedback controller based on the information from the estimation filter. This significant discovery of Kalman was generalized to a more sophisticated stage by the innovation and martingale theory introduced by Kalilath into the control area. Stimulated by the triumph of state space description, the multivariable theory in frequency domain is also developing, and the LQR problem in multivariable case in frequency domain was solved by Youla, Jabr, Bonginorno [14] in 1976 by the introducing of the now well known Y-J-B parameter matrix.

As everything has its contrary, LQR needs perfect measurement of all states which is impossible in reality, while LQG suffers from its too close dependence on precision of plant parameters and spectrum of the noise. Once the Kalman filter is used, the forgetting factor is always introduced to prevent the blockage of innovation channel. LQR has a nice property of infinite gain margin and 60° phase margin, but LQG is without
such property since the noise exists in the measurement. For minimal phase plants Doyle and Stein [3] suggested a recovery procedure to increase the gain margin and phase margin of the solution of LQG when they are not enough, but it has already destroyed the original optimal property. The recovery method cannot be used for nonminimal phase plants. Willems [3] viewed the plant uncertainty as Gaussian random vector noise inputs and tried to make a compromise between the optimal criterion and plant uncertainty through the Kolmogorov and Chapman equation. Moore, Mingori and Anderson [52] recently suggested considering LQG with frequency shaped weighting, in this case the plant uncertainty can be approximated through a LQG design for a minimal phase plant by adding fictitious plant noise.

The situation is just as two designers Chandler and Potts [2] of Flight Dynamics Laboratory complained to the control society at the IEEE Conference on Decision and Control in 1983:

"The road to applying modern control theory to fighter control design has been a very rocky one. LQR in its purest form has repeatedly been found unsuccessful. When it has worked, it has been so modified as to be unrecognizable." and they also presented "the basic requirements for a flight control synthesis theory, plus reasons and examples are given revealing modern control theory to be highly deficient --- in particular LQR, LQG, singular value theory, and eigenvalue / eigenstructure assignment."

Most of what they complained are correct. In fact, the classical lead-lag method is still widely used in many practical control system
To lay down the profound mathematical background for the classical lead-lag design method and to describe the robustness under plant uncertainty, it was Zames [1] who first introduced the infinite index Hardy space into multivariable system theory in 1981, and he immediately came to the conclusion that the sensitivity function can always be reduced to zero for minimal phase plants. The answer for nonminimal phase plants is not quite so simple. It was also Zames and Francis [4] who first realized that the sensitivity reduction problem of nonminimal phase plants is an interpolation problem. By the end of 1981 they solved the problem in the single input single output case from that the fact the optimal sensitivity function should be an all pass function. Another more significant method was introduced by them -- what is called the Nevanlinna-Schur transformation, which maps the unit disc onto the unit disc and a given point on the disc to the centre of the disc. In the spring of 1982, Francis [5] again solved the single input and single output problem by Sarason's theorem. Inspired by the success of Francis, Professor Pearson's group [10] realized that Sarason's theorem can also be used to solve the multi-input and multi-output problem and then constructed all the solutions for MIMO problem by the matrix version Nevanlinna-Schur algorithm which was proposed by Delsarte, Genin, Kamp [11] in 1978. At the same time Francis, Helton and Zames [8] solved the MIMO problem by a new developed Ball-Helton theory [9]. Safonov and Verma [12] also solved the MIMO problem by reforming the problem into a model reduction problem which was recently solved by Kung and Lin [28] via Adamjan, Arov and
Krein's method [18].

It is now clear that the mathematical background of our control problem is originated from Nevanlinna-Pick problem:

For given interpolation conditions

$$F(z_i) = w_i, \quad i = 1, \cdots, n,$$

(1.1)

to find all positive definite analytical functions in the unit disc. Pick had shown the necessary and sufficient condition for the existence of solution is

$$\begin{bmatrix} 1-w_i\bar{w}_i \\ -1/z_i z_j \end{bmatrix} \succeq 0. \quad (1.2)$$

This criterion matrix is called Pick matrix.

Nevanlinna developed the structure of the solutions in operator version, and then Akhiezer and Krein [26] developed many interesting results in the title or moment problem which is now one of the most important resources of operator theory. The Szego orthogonal polynomial [33] now widely used in information theory is also closely related to above problem.

On the other hand the Hardy space theory was developed prosperously since Beurling [30] and Lax [31] laid the cornerstone theorem in their names:

Let shift operator $\sigma$ be defined in $H^2$ by $\sigma f = e^{i\theta} f$, then a subspace $M$ of $H^2$ is invariant under $\sigma$ if and only if $M = UH^2$ for some inner function $U$. The inner function is determined up to a constant or absolute
value one [22].

Taking inner function or inner matrix as the denominator of the transformation function matrix, observing the behavior of the transformation function matrix under the action of shift operator \( \sigma \) in the modulus \( M' = H^2 \otimes UH^2 \), Fuhrmann proposed his famous realization for linear system [32] [22].

\( H^2 \otimes UH^2 \) can be considered a defection space of \( H^2 \). \( UH^2 \) is the deletion part of \( H^2 \). By viewing contraction operator \( T \) as a function of the projection of shift operator \( \sigma \) on the defection space \( H^2 \otimes UH^2 \), Nagy and Foias et al [21] developed the contraction operator theory which now is the main topic of operator theory. As they pointed out, for a long time definite results in the theory of operator on Hilbert space only had been known for self-adjoint, unitary and normal operators; but now the situation has been completely changed.

The merit of Nagy-Foias theory to the control system area is that any control system can be considered a contraction operator \( T \). Let inner function or inner matrix \( U \) contain the zeros of the system (here different from Fuhrmann's idea, now \( U \) is considered as the numerator of the system), by regulating the controller the system will run over whole space \( H^2 \otimes UH^2 \). the norm of contraction operator \( T \) can be found by shrinking the operator \( T \) as small as possible in the space \( H^2 \otimes UH^2 \). Therefore the optimal norm of the optimal control system is determined by the interpolation values at zeros of the system. To find the optimal controller or to find the optimal system can be solved by trying to find the deletion part \( UH^2 \) under the condition of keeping the optimal norm, which
is also called unitary dilation.

Mathematicians Nagy and Foias developed their theory in a very abstract language for general operators, therefore their results are hardly to be used for control engineering. In this dissertation, many explicit formulas for rational matrix version contraction operator are derived which are suitable for our multivariable control system design.

Even in a theoretical sense the problems proposed in control design are more complicated than what appeared in mathematical literature. For example, control problem in most cases is two-sided problem as $H^2\Theta U H^2 V$, but it is not difficult to overcome. A more difficult problem is that $U$ and $V$ are not square in most control problems, $U$ is tall and $V$ is flat matrix. This problem has not been discussed in mathematical literature according to author's knowledge. Several researchers in the control area tackled the above problem by different methods. Kwakernaak [41] had listed a set of equations for the simplest case (SISO) by assumption of the maximal singular value of optimal system being allpass. Unfortunately from my experience working on this problem this assumption is not always true, therefore Kwakernaak's method can only get a special solution even in the SISO case. Furthermore, his method seems difficult to be extended into a general routine performed by computer. If it is extended to MINO case, a set of complex equations is obtained and is hopeless to find the answer. Another attempt -- iteration process -- was proposed independently by Francis [56] and Doyle [40].

In this dissertation, the iteration method is also considered a main tool to tackle the nonsquare problem, but it is improved by introducing
two spectral factorization methods which are easier than what Francis and Doyle used, and two inequalities are suggested to estimate the range of the optimal norm and accelerate the convergence rate of iteration process.

The dissertation is organized as follows:

The first chapter is to emphasize the importance of introducing new synthesis methods to control theory, the historical background and the related mathematical background of a new synthesis method by infinite norm now developing rapidly are briefly mentioned. The second chapter gives the basic preliminary terminologies of Hardy space and Krein space. With the help of spectra and interpolation vectors an explicit expression will be given for LQG problem in the simplest case from the view of interpolation which is simpler than Youla’s method [14]. The inner matrix plays a very important role in our theory, it is discussed for the simplest case in chapter 3. The $H_{\infty}$ optimal norm can be derived from the expression for $H_2$ optimal solution, which will appear in chapter 4. Another method to find $H_{\infty}$ optimal norm for repeated zeros case based on Sarason’s theory will be also discussed in chapter 4. Chapter 5 is devoted to the problem of constructing all the matrices in the unit ball by unitary dilation, some explicit formulas especially for degenerate case are shown in this chapter. A recursive method is offered in chapter 6, which can be considered an extension of the Nevanlinna algorithm to include the repeated zeros case, the consideration of interpolation pair as functions of $s$ reduces the complexity of the problem. Chapter 6 will also consider the special problem of some zeros on the imaginary axis or
at infinity which is important to control problems. Chapter 7 deals with
the nonsquare problem which is of great importance to practical control
problems. Chapter 8 is devoted to apply all these results to synthesize
and design control systems. Chapter 9 gives the conclusions and some
further considerations.
CHAPTER 2 HARDY SPACE AND KREIN SPACE

2.1 Introduction

In this chapter some basic concepts and terminology in mathematics are introduced. Hardy space consists of all matrix-valued complex functions which are analytic on the open unit disc or on the open right half plane. It is the best setting for studying stability of control systems. Krein space is an indefinite norm space, the norm of a vector in Krein space can be positive, negative or neutral. Imbedding the graph of operator into Krein space, the graph should be negative when the operator norm is less than 1, therefore the graph in Krein space gives a geometrical description of operator.

2.2 Hardy space

The set of all complex matrix sequences \( \{ F_k \} \in \mathbb{C}^{m \times r} \), \( -\infty < k < +\infty \) satisfying the square-summability condition

\[
\sum_k \text{tr} F_k^* F_k < \infty
\]  

(2.1)

is denoted by \( \ell^2_{m \times r} \). Where \( \text{tr} \) means trace, and * means conjugate transpose.

After introducing the inner product

\[
\langle [F_k], [G_k] \rangle = \sum_k \text{tr} F_k^* G_k,
\]

\( \ell^2_{m \times r} \) is a Hilbert space. The norm of sequence \( \{ F_k \} \) is
\[ \| F_k \|_2 = \left( \sum_k \text{tr} F_k^* F_k \right)^{1/2}. \quad (2.3) \]

For each $L^2_{\text{mxr}}$ sequence \([F_k]\), define a matrix-valued function $F(e^{i\theta})$ as

\[ F(e^{i\theta}) = \sum_k F_k e^{ik\theta}. \quad (2.4) \]

The set of all these matrix-valued complex functions defined on the unit circle is called $L^2_{\text{mxr}}$. $L^2_{\text{mxr}}$ is a Hilbert space under the inner product

\[ \langle F, G \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{tr} F(e^{i\theta})^* G(e^{i\theta}) \, d\theta \quad (2.5) \]

The subset of $L^2_{\text{mxr}}$ consisting of all sequences which are zero for $k < 0$ is denoted by $L^2_{\text{mxr}}^+$. $L^2_{\text{mxr}}^+$ is a subspace of $L^2_{\text{mxr}}$, and is also a Hilbert space under inner product (2.2).

Each $L^2_{\text{mxr}}^+$-sequence \([F_k]\) provides a matrix-valued function.

\[ F(z) = \sum_{k=0}^{\infty} F_k z^k. \quad (2.6) \]

Every entry of $F(z)$ is analytic in the open unit disc $|z|<1$. The set of all such matrix-valued functions is called the Hardy space $H^2_{\text{mxr}}$.

For each $F$ in $H^2_{\text{mxr}}$, the limit

\[ \lim_{r \to 1^-} F(re^{i\theta}) \quad (2.7) \]

exists for almost all $\theta$, the limit yielding a unique matrix-valued function.
\[ F(e^{i\theta}) = \sum_{k=0}^{\infty} F_k e^{i k \theta} \]  
(2.8)

which is in \( L^2_{m\mathbf{x}r} \). Therefore \( H^2_{m\mathbf{x}r} \) is isomorphic to the subspace of \( L^2_{m\mathbf{x}r} \) whose Fourier coefficients are zero for negative indices. After imbedding \( H^2_{m\mathbf{x}r} \) into \( L^2_{m\mathbf{x}r} \) according to (2.2), (2.3), (2.4), we can say \( H^2_{m\mathbf{x}r} \) is a subspace of \( L^2_{m\mathbf{x}r} \).

The space of all essentially bounded \( m\mathbf{x}r \) matrix-valued functions on the unit circle is denoted by \( L^\infty_{m\mathbf{x}r} \). The \( L^\infty_{m\mathbf{x}r} \) norm of \( F \) is

\[ \| F \|_\infty = \text{ess sup}_\theta \sigma[ F(e^{i\theta}) ] \]  
(2.9)

where \( \sigma[A] \) is the maximal singular value of \( A \).

The Hardy space \( H^\infty_{m\mathbf{x}r} \) consists of all matrix-valued functions which are analytic and of bounded modulus on the open unit disc

\[ \| F \|_\infty = \sup_{|z| < 1} \sigma[ F(z) ] . \]  
(2.10)

For any element \( F(z) \) in \( H^2_{m\mathbf{x}r} \), define a shift operator \( \sigma \) as following

\[ \sigma F = z \cdot F(z) . \]  
(2.11)

From (2.6) and (2.11), we find

\[ H^2_{m\mathbf{x}r} = C_{m\mathbf{x}r} \oplus \sigma C_{m\mathbf{x}r} \oplus \sigma^2 C_{m\mathbf{x}r} \oplus \cdots \]  
(2.12)

\( L^2_{m\mathbf{x}r}, H^2_{m\mathbf{x}r}, L^\infty_{m\mathbf{x}r} \) and \( H^\infty_{m\mathbf{x}r} \) defined on the unit disc as above can be used to describe discrete time control systems. For continuous control systems, we need to define the corresponding space on right half plane instead of on unit disc.
$L^mxr_2$ is the space of $mxr$ matrix-valued function $F(j\omega)$ with the norm

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \text{tr} F(j\omega)^* F(j\omega) \, d\omega < \infty. \tag{2.13}$$

It is a Hilbert space under inner product

$$\langle F, G \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{tr} F(j\omega)^* G(j\omega) \, d\omega. \tag{2.14}$$

Hardy space $H^mxr_2$ is the set of $mxr$ matrix-valued complex functions $F(s)$ which are analytic in the open right half plane and which satisfy the condition

$$\sup_{\sigma>0} \int_{-\infty}^{\infty} \text{tr} F(\sigma+j\omega)^* F(\sigma+j\omega) \, d\omega < \infty \tag{2.15}$$

Sometimes we need $H^mxr_2$ space which is the set of $mxr$ matrix-valued complex functions $F(s)$ which are analytic in the open left half plane and which satisfy the condition (2.15) for $\sigma < 0$.

The inner product between two functions $F$ and $G$ in $H^mxr_2$ is defined by

$$\langle F, G \rangle = \frac{1}{2\pi j} \int_{j\omega}^{j\omega} \text{tr} F(s)^* G(s) \, ds, \tag{2.16}$$

where

$$F(s)^* = F^T(-s), \tag{2.17}$$

"$-"$ means that all the coefficients in $F(s)$ are replaced by their complex conjugate, "$T$" means matrix transpose, and $s$ becomes $-s$. 
Hardy space $H^m_{\infty}$ is the set of matrix-valued complex functions $\Psi(s)$ which are analytic and of bounded modulus in the open right half plane.

Each $\Psi$ in $H^m_{\infty}$ can be considered as a multiplication operator from $H^m_{\infty}$ to $H^m_{\infty}$, and the $H^m_{\infty}$ norm of $\Psi$ is defined by the norm of multiplication operator

$$
||\Psi||_m = \sup \{ ||\Psi d||_2, \quad ||d||_2 \leq 1, \quad d \in H^m_{\infty} \} 
$$

$$
= \sup_{d \in H^m_{\infty}} \frac{||\Psi d||_2}{||d||_2}. 
$$

(2.18)

We can find the relation between the two Hardy spaces on unit disc and right half plane.

(2.5) can be written as follows,

$$
\langle f, g \rangle = \frac{1}{2\pi j} \int_{|z|=1} \text{tr} f(z)^* g(z) \frac{dz}{z},
$$

(2.19)

where

$$
f(z)^* = \bar{f}(\frac{1}{z}).
$$

(2.20)

The transformation from right half plane to unit disc is

$$
z = \frac{s-1}{s+1}
$$

(2.21)

Putting (2.21) into (2.19), we have

$$
\langle f, g \rangle = \frac{1}{2\pi j} \int_{-j=\infty}^{+j=\infty} \text{tr} \left[ \frac{1}{s+1} f(s-1) \right]^* \frac{1}{s+1} g(s-1) ds.
$$

(2.22)
Compare (2.16) and (2.22), we have the map

\[ f(z) \to F(s) = \sqrt{\frac{1}{2}} \frac{f(s-1)}{s+1} \]  \hspace{1cm} (2.23)

which gives a one to one correspondence between the element in \( H^2_{\text{mix}} \) and in \( H^2_{\text{mix}} \).

On the other hand, the map

\[ f(z) \to F(s) = f(s) \]  \hspace{1cm} (2.24)

gives the one to one correspondence between the element in \( H^\infty_{\text{mix}} \) and the element in \( H^\infty_{\text{mix}} \).

From (2.11) and (2.21), we can define the multiplication operator

\[ \sigma F = \frac{s-1}{s+1} F(s) \]  \hspace{1cm} (2.25)

as a shift operator on \( H^\text{mix}_2 \).

\( B^\text{mix}_{2,1} \) and \( B^\text{mix}_{\infty,1} \) can be defined in similar way as \( B^\text{mix}_2 \) and \( B^\text{mix}_\infty \) except allowing at most \( 1 \) poles in right half plane.

### 2.3 Krein Space

Krein space \( K^m.r \) is a combination of \( H^m_2 \) and \( H^r_2 \), \( H^m_2 \oplus H^r_2 \), under the indefinite inner product

\[ [f, g] = \langle f_1, g_1 \rangle - \langle f_2, g_2 \rangle \]  \hspace{1cm} (2.26)

where \( f \) and \( g \) are in \( K^m.r \), \( f_1 \) and \( g_1 \) are in \( H^m_2 \), \( f_2 \) and \( g_2 \) are in \( H^r_2 \).

Denote
\[ f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, \quad (2.27) \]

\[ g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}, \quad (2.28) \]

\[ J = \begin{bmatrix} I_m & 0 \\ 0 & -I_r \end{bmatrix}, \quad (2.29) \]

then (2.26) can be written as

\[ [f, g] = \langle f, Jg \rangle, \quad (2.30) \]

the last inner product is defined in \( H_{2}^{m+r} \).

Since Krein space \( K_{m,r} \) is an indefinite norm space, it contains positive, negative and neutral subspace.

A subspace is called positive, negative, neutral, if all its vectors \( f \) are \( [f, f] \geq 0 \), \( [f, f] \leq 0 \), \( [f, f] = 0 \) respectively.

A subspace is called \( M \)-maximal negative subspace if it can not be properly contained in a negative subspace of \( M \).

The companion space of \( M \), denoted by \( M' \) is the subspace of vectors orthogonal to all vectors of \( M \) in \( K_{m,r} \).

In this chapter, we introduced several Hardy spaces. \( H_{2}^{m,r} \) and \( H_{m,r}^{m} \) are defined on the unit disc, \( H_{2}^{m,r} \) and \( H_{m,r}^{m} \) are defined on the right half plane. The relations between them are shown by (2.23) and (2.24). The shift operator on \( H_{2}^{m,r} \) is shown in (2.25). Krein space is a combination of two Hardy spaces under indefinite inner product. In chapter 5 we will
construct the contraction operator with the help of its graph in Krein space. Before the main topic of the optimal $H_\infty$ norm problem, the optimal $H_2$ norm problem will be discussed in the next chapter.
CHAPTER 3  \( H_2 \) OPTIMAL NORM PROBLEM

3.1 Introduction

The \( H_2 \) optimal norm problem is discussed in this chapter. The Linear Quadratic Gaussian problem can be reduced to an optimal problem in \( H_2 \) space. There exist simple expressions for optimal norm and solution when the right half plane zeros are known by viewing it an interpolation problem in \( H_2 \) space. This method is different from the Youla's method. The inner factor is the most important concept in our whole problem, a simple formula for inner factor is given in this chapter. To get at the core of the matter, we first discuss the distinct zeros case, then point out the general expressions in repeated zeros case. A problem left is how to find all (right half plane) zeros of a matrix-valued rational function and the corresponding interpolation condition vectors, which will be discussed in chapter 7.

3.2 Interpolation in \( H_2 \) Space

Let us consider an optimal problem in \( H_2 \) space.

\[
\min_{H \in H^{mxr}_2} \| W - U H \|_2^2,
\]

(3.1)

where \( W, U \) are in \( H^{mxr}_2, H^{mxm}_n \) respectively. It can be considered an interpolation problem provided all right half plane zeros of \( U \) are known. For simplicity of statement, we first suppose \( U(s) \) has only a finite number of distinct zeros in the right half plane, say \( s_1, \ldots, s_n \), and has no zeros on the imaginary axis. In this case we can find \( \alpha_i, \beta_i \) in \( \mathbb{C}^m \)
and \( G \) respectively such that

\[
\alpha_i^* U(s_i) = 0 \\
\beta_i^* = \alpha_i^* W(s_i) \quad i=1, \ldots, n.
\]  
(3.2)

Define

\[
e_i = \frac{1}{s+s_i} \alpha_i
\]  
(3.3)

\[
Q = \begin{bmatrix}
\alpha_i^* & \alpha_j^* \\
- & - \\
\end{bmatrix} \\
\begin{bmatrix}
s_i & s_j \\
+ & + \\
\end{bmatrix} \\
\begin{bmatrix}
i, j=1, \ldots, n
\end{bmatrix}
\]  
(3.4)

\[
a = [\alpha_1, \ldots, \alpha_n]
\]  
(3.5)

\[
\beta = [\beta_1, \ldots, \beta_n]
\]  
(3.6)

**THEOREM 3.1**

Problem (3.1) has a unique solution

\[
\Psi_1 = [e_1, \ldots, e_n] Q^{-1} \beta^*
\]  
(3.7)

and the optimal \( H_2 \) norm is

\[
\|\Psi_1\|_2 = \sqrt{\text{tr} \ \beta Q^{-1} \beta^*}
\]  
(3.8)

**(Proof)**

If

\[
\Psi = W - U H
\]  
(3.9)
for some \( \mathcal{H}_2^{mxr} \), then \( \mathcal{Y} \) should satisfy the interpolation conditions

\[
\beta_1^* = a_i^* \mathcal{Y}(s_i), \quad i=1,\ldots,n. \tag{3.10}
\]

On the other hand, if \( \mathcal{Y} \) satisfies (3.10), then we can find a \( \mathcal{H} \) in \( \mathcal{H}_2^{mxr} \) such that (3.9) exists. It's easy to check \( \mathcal{Y}_1 \) in form (3.7) satisfies the interpolation condition (3.10). Now we are going to show \( \mathcal{Y}_1 \) takes the minimal \( \mathcal{H}_2 \) norm among all the elements \( \mathcal{Y} \) in form (3.9) for all \( \mathcal{H} \) in \( \mathcal{H}_2^{mxr} \). For any \( \mathcal{Y} \) in (3.9), we can write

\[
\mathcal{Y} = \mathcal{Y}_1 - \mathcal{Y}_0 \tag{3.11}
\]

where \( \mathcal{Y}_0 \in \mathcal{H}_2 \) and

\[
\mathcal{Y}_0 = \bigcup \mathcal{H}_0 \tag{3.12}
\]

for some \( \mathcal{H}_0 \in \mathcal{H}_2^{mxr} \), therefore

\[
a_i^* \mathcal{Y}_0(s_i) = 0, \quad i=1,\ldots,n. \tag{3.13}
\]

From (3.7), (3.13) and (3.4), we have

\[
\langle \mathcal{Y}_1, \mathcal{Y}_0 \rangle = 0. \tag{3.14}
\]

\[
\| \mathcal{Y}_1 \|_2^2 = \| \mathcal{Y}_1 + \mathcal{Y}_0 \|_2^2
\]

\[
= \| \mathcal{Y}_1 \|_2^2 + \| \mathcal{Y}_0 \|_2^2
\]

\[
\geq \| \mathcal{Y}_1 \|_2^2 \tag{3.15}
\]

(3.15) means that \( \mathcal{Y}_1 \) takes the minimal value of problem (3.1). From (3.7) we have (3.8) after a little manipulation. Theorem 3.1 is proved.
3.3 Inner and Outer Factorization

Inner and outer factorization is a most important procedure for our problem. In this section we only consider a special case, that is $U$ in $H^m_{\infty}$ with null rank $m$ on the imaginary axis and at infinity, we want to factorize it as follows

$$U = U_i U_o$$  \hspace{1cm} (3.16)

where $U_i$ is in $H^m_{\infty}$, contains all right half plane zeros of $U$, and satisfies

$$U_i^* U_i = I_{m \times m}.$$  \hspace{1cm} (3.17)

$U_o$ is in $H^m_{\infty}$ and contains only left half plane zeros. $U_i$ is called an inner factor of $U$. $U_o$ is called an outer factor of $U$.

In general case, $U$ has zeros on the imaginary axis or at infinity, then we have following factorization:

$$U = U_i U_s U_o$$  \hspace{1cm} (3.18)

where $U_s$ is a singular factor of $U$, and it contains all zeros on the imaginary axis and at the infinity.

The next theorem gives an expression for the inner factor when all the right half plane zeros of $U$ are distinct.

**Theorem 3.2**

For $U$ in $H^m_{\infty}$ with rank $m$, and all right half plane zeros $s_1, \ldots, s_n$ of $U$ are distinct,
\[ a_j U(s_j) = 0, \quad j=1, \ldots, n, \quad (3.19) \]

The expression

\[ U_i = I_m \dot \det U \cdot \begin{bmatrix} e_1, \ldots, e_n \end{bmatrix} Q^{-1} a^* \quad (3.20) \]

is an inner factor of \( U \). In (3.20), \( e_i, Q, a \) are from (3.3), (3.4), (3.5).

(Proof)

From expression (3.20), it is easy to check (3.19).

Again from (3.20), \( \det U_i \) is nth order in \( s \), therefore \( U_i \) contains and only contains the right half plane zeros of \( U \). The next step is to check (3.17). From (3.3) and (3.4), we have following identity:

\[
\begin{bmatrix}
  e_1, \ldots, e_n \\
  \\
  \\
  \\
\end{bmatrix}^* \begin{bmatrix}
  e_1, \ldots, e_n \\
  \\
  \\
  \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
  -s+s_1 \\
  \\
  \\
  \\
\end{bmatrix} \begin{bmatrix}
  -s+s_1 \\
  \\
  \\
  \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
  1 \\
  \\
  \\
  \\
\end{bmatrix} \begin{bmatrix}
  1 \\
  \\
  \\
  \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
  -s+s_1 \\
  \\
  \\
  \\
\end{bmatrix} \begin{bmatrix}
  -s+s_1 \\
  \\
  \\
  \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
  1 \\
  \\
  \\
  \\
\end{bmatrix} \begin{bmatrix}
  1 \\
  \\
  \\
  \\
\end{bmatrix}
\]

Now, from (3.20) we have

\[ U_i^* U_i \]

\[ = (I - aQ^{-1}[e_1, \ldots, e_n]^*) (I - [e_1, \ldots, e_n]Q^{-1} a^*) \]

\[ = I - aQ^{-1}[e_1, \ldots, e_n]^* - [e_1, \ldots, e_n]Q^{-1} a^* \]
\[ + a Q^{-1} [e_1, \ldots, e_n]^* [e_1, \ldots, e_n] Q^{-1} a^*. \]  \hspace{1cm} (3.22)

Put \((3.21)\) into the last term in \((3.22)\), we can find that it is split into two parts and they can be canceled by the second and third term respectively, and thus \((3.17)\) is proved.

The dual outer–inner factorization is mentioned in theorem 3.2'.

**THEOREM 3.2’**

For \(V\) in \(H^r_{\infty}\) having rank \(r\), and all right half plane zeros \(s_1, \ldots, s_n\) of \(V\) are distinct,

\[ V(s_j) a_j = 0, \quad j = 1, \ldots, n. \] \hspace{1cm} (3.25)

the expression

\[
\begin{bmatrix}
 a^* \\
 s_j \\
 s + s_1 \\
 0 \\
 a_n \\
 s + s_n \\
\end{bmatrix}
\]

\[ V_i = I_{r \times r} - a Q^{-1} \] \hspace{1cm} (3.24)

is an inner factor of \(V\), and \(V = V O' V_i\), where

\[
\tilde{Q} = \begin{bmatrix}
 a^* \\
 s_i \\
 s_j \\
 a_n \\
 s + s_n \\
\end{bmatrix}
\] \hspace{1cm} (3.25)
Example 3.1 (Blaschke product)

Consider the scalar case. $U$ has $n$ right half plane zeros $s_1, \ldots, s_n$.

$$U = \frac{(s-s_1) \ldots (s-s_n)}{(s+a_1) \ldots (s+a_n)}$$  \hspace{1cm} (3.26)

Let

$$a_i = 1, \quad i = 1, \ldots, n,$$  \hspace{1cm} (3.27)

(3.19) becomes

$$U_i = 1 - \left[ \begin{array}{c} \frac{1}{s+s_1} \\ \vdots \\ \frac{1}{s+s_n} \end{array} \right] Q^{-1} \left[ \begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right],$$  \hspace{1cm} (3.28)

Notice

$$\det \left[ I + AB \right] = \det \left[ I + BA \right],$$  \hspace{1cm} (3.29)

we have

$$U_i = \det \left[ I_n - \left[ \begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right] \left[ \begin{array}{c} \frac{1}{s+s_1} \\ \vdots \\ \frac{1}{s+s_n} \end{array} \right] Q^{-1} \right]$$

$$= \det Q^{-1} \det \left[ Q - \left[ \begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right] \left[ \begin{array}{c} \frac{1}{s+s_1} \\ \vdots \\ \frac{1}{s+s_n} \end{array} \right] \right]$$

$$= \det Q^{-1} \det \left[ \frac{s-s_j}{(s_j+s_k)(s+s_k)} \right], j, k = 1, \ldots, n.$$
\[
\begin{align*}
\det Q^{-1} \det \left[ \text{diag}(s-s_1, \ldots, s-s_n) \right] \det Q \det \left[ \text{diag} \left( \frac{1}{s+s_1}, \ldots, \frac{1}{s+s_n} \right) \right] \\
= \frac{(s-s_1) \cdots (s-s_n)}{(s+s_1) \cdots (s+s_n)},
\end{align*}
\]

which is a Blaschke product and plays a basic role in interpolation theory.

### 3.4 Two-sided Problem

In control theory we need to solve a two-sided optimal problem:

\[
\min_{H \in \mathbb{H}_2^{mx}} \| W - U H V \|_2
\]

(3.31)

In this section we only consider the case of $U$ and $V$ being square rational matrices, namely, $U$, $V$, $W$ are in $\mathbb{H}_\infty^{mx}$, $\mathbb{H}_\infty^{rx}$, $\mathbb{H}_2^{mx}$ respectively. $U$ and $V$ have full rank on the imaginary axis and at infinity.

From the last section, we have outer-inner factorization:

\[
V = V_0 V_i.
\]

(3.32)

Let $\Psi$ be the Blaschke product such that $\Psi V_i^*$ is in $\mathbb{H}_\infty^{rx}$, we have

\[
\begin{align*}
\| W - U H V \|_2 \\
= \| W - U H V_0 V_i \|_2 \\
= \| \Psi V_i^* - \Psi U H V_0 \|_2 \\
= \| W_1 - U_1 H_1 \|_2.
\end{align*}
\]

(3.33)

where $W_1 = \Psi V_i^*$ is in $\mathbb{H}_2^{mx}$, $U_1 = \Psi U$ is in $\mathbb{H}_\infty^{mx}$, $H_1 = HV_0$ can vary over
whole $H^m_{\infty}$ as $H$ varies over whole $H^m_{\infty}$}. Problem (3.31) becomes

$$\min_{H_1 \in H^m_{\infty}} || W_1 - U_1 H_1 ||_2.$$  \hfill (3.34)

This problem has been discussed in section 3.1.

**Example 3.2**

To solve problem (3.31) for

$$W = \begin{bmatrix} -20(2s-1) / (s+1)(s+10) & 10 / (s+10) \\ 10 / (s+1)(s+10) & s+10 \end{bmatrix}$$

$$U = \begin{bmatrix} s-1 \\ s+3 \end{bmatrix}$$

$$V = \begin{bmatrix} 0 & -1 \\ s-2 & 3 \\ 2(s+1) & 3 \end{bmatrix}.$$

From (3.24) we can get $V_i$,

$$V_i = \begin{bmatrix} s-2 \\ s+2 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Take $\phi = \frac{s-2}{s+2}$, we have

$$W_1 = \begin{bmatrix} -20(2s-1) / (s+1)(s+10) & 10(s-2) / (s+2)(s+10) \\ (s+1)(s+10) & (s+2)(s+10) \end{bmatrix}$$

$$U_1 = \begin{bmatrix} (s-1)(s-2) \\ (s+3)(s+2) \end{bmatrix}.$$  

the problem becomes in the form of (3.1), which can be solved by theorem 3.1.
Let
\[ s_1 = 1, \quad a_1 = 1, \]
\[ s_2 = 2, \quad a_2 = 1, \]
the optimal solution is
\[ \Phi_1 = \begin{bmatrix} \frac{20(s-5)}{11(s+1)(s+2)} & \frac{20(s-2)}{11(s+1)(s+2)} \end{bmatrix}. \]

It's optimal $H_2$ norm is
\[ \|\Phi_1\|_2 = 6.618282. \]

The original one is
\[ \Phi = \begin{bmatrix} \frac{20(s-5)}{11(s+1)(s+2)} & \frac{20}{11(s+1)} \end{bmatrix}, \]
and the optimal $H$ is
\[ H = \begin{bmatrix} \frac{10(s+3)(143s+134)}{11(s+1)(s+2)(s+10)} & \frac{560(s+3)}{11(s+2)(s+10)} \end{bmatrix} \]
which is in $H_2^{1 \times 2}$.

3.5 The Case of Repeated Zeros

Suppose $U(s)$ has repeated right half plane zeros $s_1, \ldots, s_n$, and the multiplicity of $s_i$ is $l_i$.

Expanding $U(s)$ at $s_i$ into Taylor series:
\[ U(s) = U_i^0 + \frac{1}{1!} U_i^1 (s-s_i) + \frac{1}{2!} U_i^2 (s-s_i)^2 + \ldots \quad (3.35) \]
Since $s_i$ is a $i$th order zero of $U$, we can find an analytic function $\alpha^*(s)$ in $H^m_\infty$, ($\gamma$ is determined by $s_i$), such that
\[ \alpha^*(s) U(s) = O(s-s_i)^{i_i}, \quad i=1,\ldots,n, \] (3.36)
where $O(x^n)$ means the same order as $x^n$ when $x \to 0$. Expanding $\alpha^*(s)$ at $s_i$:
\[ \alpha^*(s) = a_i^0 + \frac{1}{1!} a_i^1 (s-s_i) + \frac{1}{2!} a_i^2 (s-s_i)^2 + \ldots \] (3.37)
where $U_i^l$ and $a_i^l$ are $l$th derivatives of $U(s)$ and $\alpha^*(s)$ at $s_i$. From (3.35), (3.36), (3.37), comparing the coefficients until $(s-s_i)^{i_i-1}$, we have
\[
\begin{bmatrix}
0^0 & 0^1 & \cdots & 0^{i_i-1} \\
0^0 & 0^1 & \cdots & 0^{i_i-1} \\
\vdots & \vdots & \ddots & \vdots \\
0^0 & 0^1 & \cdots & 0^{i_i-1} \\
\end{bmatrix}
\begin{bmatrix}
a_i^0 \\
a_i^1 \\
\vdots \\
a_i^{i_i-1} \\
\end{bmatrix}
= 0. \] (3.38)

The next step is to expand $W(s)$ into a Taylor series at $s_i$,
\[ W(s) = \frac{w^0}{i!} \left( s-s_i \right) + \frac{1}{2!} \frac{w^1}{i!} \left( s-s_i \right)^2 + \ldots \] (3.39)

Define
\[ \beta^*(s) = \alpha^*(s) W(s), \] (3.40)

$\beta^*(s)$ is in $H^m_\infty$. 
We can also expand $\beta^*(s)$ into Taylor series at $s_i$,

$$
\beta^*(s) = \beta_{i}^{0*} + \frac{1}{1!} \beta_{i}^{1*} (s-s_i) + \frac{1}{2!} \beta_{i}^{2*} (s-s_i)^2 + \ldots .
$$

(3.41)

Truncate (3.40) $(s-s_i)^j$ and higher order, we have

$$
[ \beta_{i}^{0*}, \beta_{i}^{1*}, \ldots, \beta_{i}^{j_i-1*} ] = [ a_{i}^{0*}, a_{i}^{1*}, \ldots, a_{i}^{j_i-1*} ].
$$

(3.42)

Now define

$$
V = [ v_{ij} ]_{i,j=1,\ldots,n}.
$$

(3.43)

where

$$
V_{ij} =
\begin{bmatrix}
\beta_{i}^{0,0} & \beta_{i}^{0,1} & \ldots & \beta_{i}^{0,j_i-1} \\
\beta_{i}^{1,0} & \beta_{i}^{1,1} & \ldots & \beta_{i}^{1,j_i-1} \\
\vdots & \vdots & \ddots & \vdots \\
\beta_{i}^{j_i-1,0} & \beta_{i}^{j_i-1,1} & \ldots & \beta_{i}^{j_i-1,j_i-1} \\
\end{bmatrix}
$$

(3.44)
\[ v_{i,j} = \frac{(-1)^{l_1^i + l_2^j} \cdot (l_1 + l_2)!}{(s_i + s_j) \cdot (s_1 + s_2 + 1)} \cdot I_{m \times m} \]  
(3.44)

\[ a_i = \begin{bmatrix} 0 \\ C_0^0 a_i \\ C_1^1 a_i \\ \vdots \\ C_{l_i - 1}^{l_i - 1} a_i \\ \end{bmatrix} \]  
(3.45)

\[ \beta_i = \begin{bmatrix} 0 \\ C_0^0 \beta_i \\ C_1^1 \beta_i \\ \vdots \\ C_{l_i - 1}^{l_i - 1} \beta_i \\ \end{bmatrix} \]  
(3.46)

\[ Q = [a_i^* v_{ij} a_j]_{i,j=1 \ldots n}, \]  
(3.47)

\[ a = [a_1, \ldots, a_n], \]  
(3.48)

\[ \beta = [\beta_1, \ldots, \beta_n], \]  
(3.49)
\[ g_i = \begin{bmatrix} 1 \\ \frac{-1}{(s+s_i)} \\ \frac{-1}{(s+s_i)^2} \\ \vdots \\ \frac{(-1)^{i-1}}{(s+s_i)^{i-1}} \\ \frac{(-1)^{i-1}}{(s+s_i)^{i}} \end{bmatrix}, \quad (3.50) \]

\[ A_i = \begin{bmatrix} 0 & a_1 & \cdots & a_{i-1} \\ a_i & 1 & \cdots & a_{i-1} \end{bmatrix}, \quad (3.51) \]

\[ B_i = \begin{bmatrix} 0 & \beta_1 & \cdots & \beta_{i-1} \\ \beta_i & 1 & \cdots & 1 \end{bmatrix}. \quad (3.52) \]

With these notations we can write down Theorem 3.3:

**THEOREM 3.3**

In the case of repeated zeros of \( U(s) \) as shown in (3.38), problem (3.1) has the unique solution:

\[ \hat{\Phi}_i = \begin{bmatrix} g_1 a_1 & \cdots & g_n a_n \end{bmatrix} Q^{-1} \begin{bmatrix} B_1 & \cdots & B_n \end{bmatrix}^* \quad (3.53) \]

and the optimal \( H_2 \)-norm is

\[ \| \hat{\Phi}_i \|_2 = \sqrt{\text{tr} \left[ B_1, \ldots, B_n \right] Q^{-1} \left[ B_1, \ldots, B_n \right]^*}. \quad (3.54) \]

**Proof**

This theorem is an extension of theorem 3.1 from distinct zeros to repeated zeros. The proof is the same as above and needs more manipulations, and is omitted.

With these notations we also can extend theorem 3.2 and theorem 3.2' of inner factor to the case of repeated zeros.
THEOREM 3.4

For \( U(s) \) in \( H_{\infty}^{m \times m} \), suppose \( U(s) \) has repeated zeros \( s_1, \ldots, s_n \) as shown in (3.38). The expression

\[
U_i = I_{m \times m} - \begin{bmatrix} s_1 a_1 & \ldots & s_n a_n \end{bmatrix} Q^{-1} \begin{bmatrix} A_1 & \ldots & A_n \end{bmatrix}^* \tag{3.55}
\]

is an inner factor of \( U \), and \( U = U_1 U_0 \).

THEOREM 3.4'

For \( V \) in \( H_{\infty}^{r \times r} \), suppose \( V \) has repeated zeros \( s_1, \ldots, s_n \) and

\[
\begin{bmatrix}
C^0 V^0 \\
C^0 V^1 \\
\vdots \\
C^0 V^{i-1} \\
C^0 V_i
\end{bmatrix} - \begin{bmatrix}
C^{i-1} V_i \\
C^{i-1} V^0 \\
\vdots \\
C^{i-1} V^{i-2} \\
C^{i-1} V_i
\end{bmatrix} = 0 
\tag{3.56}
\]

for \( i = 1, \ldots, n \). The expression

\[
V_i = I_{r \times r} - \begin{bmatrix} A_1 & \ldots & A_n \end{bmatrix} Q^{-1} \begin{bmatrix} A_1 & \ldots & A_n \end{bmatrix}^* \tag{3.57}
\]

is an inner factor of \( V \), and \( V = V_0 V_i \).

In this chapter the explicit solution for \( H_2 \) optimal problem is given. In chapter 8 we will show a simple example to solve LQG problem by the recipe. The explicit expression for \( H_2 \) is also used to find the \( H_{\infty} \)-optimal norm in the next chapter.
CHAPTER 4 $H_\infty$-OPTIMAL NORM

4.1 Introduction

In this chapter we will consider a $H_\infty$-optimal problem (4.1):

$$\min_{H \in H_\infty^{\text{mrx}}} \| W - UH \|_\infty$$  \hspace{1cm} (4.1)

where $W$ is in $H_\infty^{\text{mrx}}$, and $U$ is in $H_\infty^{\text{mxm}}$.

$H_\infty$ problem (4.1) and $H_2$ problem (3.1) are closely related. First we will derive a formula for $H_\infty$-optimal norm from the expression for the $H_2$ problem, then give another derivation for the same problem by an extension of Sarason theorem. We will construct all interpolation solutions in $B_{H_\infty^{\text{mrx}}}$ in the next chapter, it's formula is similar to the expression for the inner factor given in the last chapter.

4.2 The Case of Distinct Zeros

In this section only the case of distinct zeros is considered. As in section 3.2, right half plane zeros are $s_1, \ldots, s_n$, and the interpolation vectors are satisfied (3.2). Notice (2.18), (4.1) can be rewritten as following:

$$\min_{H \in H_\infty^{\text{mrx}}} \sup_{d \neq 0} \frac{\|WD - UEd\|_2}{\|d\|_2} \hspace{1cm} (4.2)$$

Since $H_2$ norm is only defined on the imaginary axis, we can consider $d$ as being minimal phase in the right half plane. First we calculate
by the method mentioned in chapter 3. Keeping in mind that \( W_d \) in (4.3) is in the same position as \( W \) in (3.1), define \( e_1, Q, \) and \( a \) according to (3.3), (3.4), (3.5):

\[
\beta^d = \begin{bmatrix} \beta^d_1, \ldots, \beta^d_n \end{bmatrix}
\]

(4.4)

where

\[
\beta^d_i = a^*_i W(s_i)d(s_i) = \beta^*_i d(s_i).
\]

(4.5)

(4.5) gives the relation between \( \beta^d_i \) and \( \beta^*_i \), where \( \beta^*_i \) is from (3.2).

Similar to (3.7), define

\[
\Phi^d = \begin{bmatrix} e_1, \ldots, e_n \end{bmatrix} Q^{-1} \beta^d
\]

(4.6)

then from the argument in theorem 3.1, we can write

\[
W_d - UH_d = \Phi^d_1 + \Phi^d_o
\]

(4.7)

where \( \Phi^d_1 \) is from (4.6), \( \Phi^d_o \in \mathbb{H}^2 \) and

\[
a^*_i \Phi^d_0(s_i) = 0, \quad i = 1, \ldots, n.
\]

(4.8)

From (4.6), (4.8), we have

\[
\langle \Phi^d_1, \Phi^d_0 \rangle = 0.
\]

(4.9)

Therefore

\[
\| W_d - UH_d \|_2^2 = \| \Phi^d_1 \|_2^2 + \| \Phi^d_0 \|_2^2
\]
\[ = \beta_d Q^{-1} \beta_d^* + \| d \|_2^2. \]  
(4.10)

The next step is to calculate \( \| d \|_2^2 \). Define the Blaschke product

\[ \varphi = \prod_{i=1}^n \frac{s-s_i}{s+s_i}, \]  
(4.11)

d can be projected into two orthogonal spaces: \( R^r_2 \otimes \mathcal{V}^r_2 \) and \( \mathcal{V}^r_2 \).

\[ d = d_\varphi + \mathcal{V}^r h \]  
(4.12)

where \( d_\varphi \) is in \( R^r_2 \otimes \mathcal{V}^r_2 \), and \( h \) is in \( R^r_2 \). Again we can write an expression for \( d_\varphi \) by the argument in theorem 3.1. Let

\[ a_i = I_r \]  
(4.13)

\[ b_{i}^{\varphi} = d(s_i) \]  
(4.14)

\[ c_{i}^{\varphi} = \frac{1}{s+s_i} I_r \]  
(4.15)

\[ V = \begin{bmatrix} I_r \\ s_i + s_j \\ s_i + s_j \end{bmatrix} \]  
(4.16)

\[ \beta^{\varphi} = \begin{bmatrix} d(s_1) \cdots \cdots d(s_n) \end{bmatrix} \]  
(4.17)

\[ d_\varphi = \begin{bmatrix} c_1^{\varphi} \cdots \cdots c_n^{\varphi} \end{bmatrix} V^{-1} \beta_{i}^{\varphi}. \]  
(4.18)

We have

\[ \langle d_\varphi , \varphi h \rangle = 0 , \quad V h \in R^r_2 \]  
(4.19)
\[ \| \mathbf{d} \|_2^2 = \| \mathbf{d}_\varphi \|_2^2 + \| \varphi \mathbf{n} \|_2^2 = \beta^\varphi \mathbf{v}^{-1} \beta \mathbf{d}^\star + \| \mathbf{h} \|_2^2 \]  

(4.20)

From (4.10) and (4.20), (4.2) can be written as follows:

\[ \min_{\varphi_0 \in \mathbb{H}_2^n} \sup_{\mathbf{b} \in \mathbb{H}_2^r} \frac{\beta^d \mathbf{Q}^{-1} \beta \mathbf{d}^\star + \| \varphi_0 \|_2^2}{\beta'^\varphi \mathbf{v}^{-1} \beta \mathbf{d}^\star + \| \mathbf{h} \|_2^2} \]

\[ d(s_1), \ldots, d(s_n) \]  

(4.21)

For any given \( d(s) \), \( \mathbf{h} \) and \( d(s_1), \ldots, d(s_n) \) are determined. For the minimal value of (4.21) we need

\[ \varphi_0^d = 0 \]  

(4.22)

From (4.12) we can think that \( d \) is determined by \( d_\varphi \) and \( \mathbf{h} \), and \( d_\varphi \) and \( \mathbf{h} \) can be chosen independently, therefore \( d(s_1), \ldots, d(s_n) \) and \( \mathbf{h} \) are independent. For the supremum of (4.21), we must choose

\[ \mathbf{h} = 0 \]  

(4.23)

From (4.22) and (4.23), (4.21) becomes

\[ \sup_{d(s_1), \ldots, d(s_n)} \frac{\beta^d \mathbf{Q}^{-1} \beta \mathbf{d}^\star}{\beta'^\varphi \mathbf{v}^{-1} \beta \mathbf{d}^\star} \]  

(4.24)

From (4.4), (4.5), (4.17), we can write (4.24) as follows:
\[
\begin{bmatrix}
\beta_1^* d(s_1) \\
\vdots \\
\beta_n^* d(s_n)
\end{bmatrix}^* \begin{bmatrix}
\beta_1^* d(s_1) \\
\vdots \\
\beta_n^* d(s_n)
\end{bmatrix} = \sup_{d(s_1), \ldots, d(s_n)} \begin{bmatrix}
[d(s_1)]^* \\
\vdots \\
[d(s_n)]^*
\end{bmatrix} Q^{-1} \begin{bmatrix}
[d(s_1)] \\
\vdots \\
[d(s_n)]
\end{bmatrix} \begin{bmatrix}
\beta_1^* \\
\vdots \\
\beta_n^*
\end{bmatrix} \begin{bmatrix}
[d(s_1)] \\
\vdots \\
[d(s_n)]
\end{bmatrix} \begin{bmatrix}
\lambda \mathbf{V}^{-1}
\end{bmatrix} = 0 (4.25)
\]

which can be easily solved by Lagrange indefinite multiplier method. The supremum of (4.25) should be the maximal generalized eigenvalue of

\[
\det \left\{ \text{diag} \left[ \beta_1^*, \ldots, \beta_n^* \right] Q^{-1} \text{diag} \left[ \beta_1^*, \ldots, \beta_n^* \right] - \lambda \mathbf{V}^{-1} \right\} = 0 \quad (4.26)
\]

It is also

\[
\det \left\{ \mathbf{R} - \lambda \mathbf{Q} \right\} = 0 . \quad (4.27)
\]

where \( \mathbf{R} \) is in the form of (4.32).

To sum up, we have proved the next theorem:

**Theorem 4.1**

For \( \mathbf{W} \in \mathbb{H}_\infty^{m \times r} \), \( \mathbf{U} \in \mathbb{H}_\infty^{m \times m} \), and all right half plane zeros \( s_1, \ldots, s_n \) of \( \mathbf{U} \) are distinct. Suppose \( \alpha_i, \beta_i \) (\( i = 1, \ldots, n \)) are

\[
\begin{align*}
\alpha_i^* \mathbf{W}(s_i) &= 0 , \quad (4.38) \\
\beta_i^* &= \alpha_i^* \mathbf{W}(s_i) , \quad (4.39)
\end{align*}
\]

the \( \mathbb{H}_\infty \)-optimal norm of
\[
\min_{H \in H_{\infty}} \| W - UH \|_\infty
\]  
(4.1)

is the positive square root of the maximal generalized eigenvalue of

\[
\det \{ R - \lambda Q \} = 0.
\]  
(4.30)

where

\[
Q = \begin{bmatrix}
    a_i^* & a_j^* \\
    -s_i & -s_j \\
\end{bmatrix}
\]  
(4.31)

\[
R = \begin{bmatrix}
    \beta_i^* & \beta_j^* \\
    -s_i & -s_j \\
\end{bmatrix}
\]  
(4.32)

4.3 The Case of Repeated Zeros

In this section we will extend the $H_{\infty}$ -optimal problem to the case of repeated zeros. Suppose $s_i$ is $i_{th}$ zero of $U$ according to (3.38) and (3.42). Let

\[
d(s) = d_0^i + \frac{1}{1!} d_1^i (s-s_i) + \frac{1}{2!} d_2^i (s-s_i)^2 + \ldots
\]  
(4.33)

\[
\beta^d(s)^* = \alpha^* (s) W(s) d(s) = \beta^* (s) d(s)
\]  
(4.34)

and expand

\[
\beta^d(s)^* = \beta_i^d0^* + \frac{1}{1!} \beta_i^d1^* (s-s_i) + \frac{1}{2!} \beta_i^d2^* (s-s_i)^2 + \ldots
\]  
(4.35)

then

\[
B_i^d = \beta_i^* d_i
\]  
(4.36)
where

\[
\begin{bmatrix}
  d_i^0 \\
  \vdots \\
  d_i^{i-1} \\
\end{bmatrix}, \quad B_i^{d*} = \begin{bmatrix}
  \beta_i \\
  \vdots \\
  \beta_i^{i-1*} \\
\end{bmatrix}
\]

(4.37)

Denote

\[
B_i^{d*} = \begin{bmatrix}
  d_1 \\
  \vdots \\
  d_n \\
\end{bmatrix}
\]

(4.38)

\[
B_i^{d*} = \begin{bmatrix}
  \beta_1^* d_1 \\
  \vdots \\
  \beta_n^* d_n \\
\end{bmatrix}
\]

(4.39)

and define \( V, a_i, \beta_i, Q, g_i \) as (3.43) to (3.50), then

\[
\| W - U H \|_2^2 = \text{tr} B_i^{d*} Q^{-1} B_i^{d*} + \tilde{\psi}_0
\]

(4.40)

Taking

\[
\varphi = \prod_{i=1}^{n} \frac{(s-s_i)}{(s+s_i)}
\]

(4.41)

and the decomposition of \( d \) is

\[
d = d_\varphi + \psi \ h
\]

(4.42)

\[
d_\varphi = \begin{bmatrix}
  g_1, g_2, \ldots, g_n \\
\end{bmatrix} V^{-1} B_i^{d*}
\]

(4.43)
Therefore

\[ \|d\|_2^2 = \text{tr} B^\psi V^{-1} B^\psi + \|h\|_2^2. \]  

(4.44)

From (4.40) and (4.44), using the same argument in theorem 4.1, we have

**Theorem 4.2**

Suppose \( W \) and \( U \) are in \( H_{\infty}^{mxr} \) and \( H_{\infty}^{mym} \) respectively, \( U \) has no imaginary axis zero. For all right half plane zero \( s_i \) of \( U \), the multiplicity of \( s_i \) is \( \ell_i \), and (3.38) (3.42) exist. The optimal \( H_{\infty} \) norm of

\[ \min_{H \in H_{\infty}^{mxr}} \| W - U H \|_{\infty} \]  

(4.1)

is the positive square root of the maximal eigenvalue of

\[ \det \{ R - \lambda Q \} = 0, \]  

(4.45)

where \( Q \) is from (3.47), and

\[ R = \begin{bmatrix} \beta_1^* & V_{ij} & \beta_j \end{bmatrix}_{i,j=1,\ldots,n} \]  

(4.46)

Now each \( a_i, \beta_i \) is a big matrix (see (3.45) (3.46)).

**4.4 Another Derivation**

In this section another derivation of \( H_{\infty} \) optimal norm is discussed, which is an extension of Sarason's theorem.
Sarason proved that for $\Phi$ in $H_\infty$, there is norm preserving between two spaces $H_\infty / \psi H_\infty$ and $K_\psi$:

$$\|\Phi\|_{H_\infty / \psi H_\infty} = \|\Phi\|_{K_\psi}, \tag{4.47}$$

where $K_\psi$ is the interpolation space $H_2 \ominus \psi H_2$.

Instead of scalar Blaschke product $\Phi$, we use full row rank rational matrix $U$. Without loss of generality, $U$ is considered an inner functional matrix, otherwise use the inner factor of $U$ instead of $U$. Define $K^\text{mxr}_U$ as the subspace $H_2^\text{mxr} \ominus U H_2^\text{mxr}$. $P$ is a projection operator, $P^\text{mxr}_{H_2^\text{mxr}}$ is the projection from $L_2^\text{mxr}$ to $K^\text{mxr}_U$, and $P^\text{mxr}_{H_2^\text{mxr}}$ is the projection from $L_2^\text{mxr}$ to $H^\text{mxr}_2$.

From the definition of $H_2$ and $H_\infty$ norm, we have:

$$\|\Phi\|_{H^\text{mxr}_\infty} = \sup_{d \in L_2^\text{mxr}} \frac{\|d\|_2}{\|d_1\|_2}$$

$$= \sup_{d_1 \in L_2^\text{mxr}, d_2 \in L_2^\text{mxr}} \frac{\langle d_1, \Phi d_2 \rangle}{\|d_1\|_2 \cdot \|d_2\|_2}. \tag{4.48}$$

With $\Phi = W + UH$, we recognize that $\Phi$ is an element of coset of $W$ in the factor space $H^\text{mxr}_\infty / U H^\text{mxr}_\infty$. The coset norm will be written as

$$\|\Phi\|_{H^\text{mxr}_\infty / U H^\text{mxr}_\infty} = \inf_{W \neq U H} \|W + U H\|_{H^\text{mxr}_\infty}, \tag{4.49}$$

On the other hand, we define the operator norm of $P^\text{mxr}_{H_2^\text{mxr}} \Phi^*$ on $K^\text{mxr}_U$:
\[ \| P_{rxl} \|_{k_U}^{m_{mxl}} = \sup_{g \in k_U^{m_{mxl}}} \| g \|_2 \]

\[ = \sup_{g_1 \in k_U^{m_{mxl}}} \| g_1 \|_2 \cdot \| h_2 \|_2 \]  
\[ = \sup_{g_1 \in k_U^{m_{mxl}}} \| g_1 \|_2 \cdot \| h_2 \|_2 \]  
\[ = \sup_{g_1 \in k_U^{m_{mxl}}} \| g_1 \|_2 \cdot \| h_2 \|_2 \]  

(4.50)

**Theorem 4.3**

For \( \Phi \) in \( H_\infty^{m_{mxr}} \),

\[ \| \Phi \|_{H_\infty^{m_{mxr}}/UH_\infty^{m_{mxr}}} = \| P_{rxl} \|_{k_U}^{m_{mxl}} \cdot \]  

(4.51)

(Proof)

There are three extensions of this theorem from Sarason's theorem:

(i) Extend scalar Blaschke product to any matrix-valued inner factor.

(ii) Sarason proved only the case of scalar \( \Phi \), and then extended it to the case of square matrix \( \Phi \). We consider \( \Phi \) in the cases of \( m \times r \) matrix no matter \( m \times r \), \( m \times r \) or \( m \times r \).

(iii) We use the operator norm of \( P_{rxl} \) instead of \( \Phi \).

Suppose for a \( \Phi \) in \( H_\infty^{m_{mxr}} \),

\[ \| \Phi \|_{H_\infty^{m_{mxr}}/UH_\infty^{m_{mxr}}} = \alpha \ . \]  

(4.52)
From (4.48) and (4.49) this means that

$$\inf_{h \in H^{mxr}_\infty} \sup_{d_1 \in L_2^{mxl}} \sup_{d_2 \in L_2^{rxl}} \frac{\langle d_1, (W+Uh)d_2 \rangle}{\|d_1\|_2 \cdot \|d_2\|_2} = \alpha.$$  \hspace{1cm} (4.53)

From (4.53), we conclude that

$$\langle d_1, Uh_d \rangle = 0,$$ \hspace{1cm} (4.54)

otherwise, when $\langle d_1, Uh_d \rangle > 0$, we use $-\alpha$ instead of $h$; and when $\langle d_1, Uh_d \rangle < 0$, we use $2\alpha$ instead of $h$, all will further decrease the norm. This means that for calculating

$$\sup_{d_1 \in L_2^{mxl}} \sup_{d_2 \in L_2^{rxl}} \frac{\langle d_1, (W+Uh)d_2 \rangle}{\|d_1\|_2 \cdot \|d_2\|_2}$$ \hspace{1cm} (4.55)

we only need consider these $d_1, d_2$ which satisfy (4.54) for all $h \in H^{mxr}_\infty$ and so $d_1 \in \tilde{U}_{\infty}^{mxl}$, $d_2 \in \tilde{h}_2^{rxl}$. Suppose there exists a $h$ in $H^{mxr}_\infty$, a $\tilde{h}_1$ in $\tilde{h}_2^{mxl}$ and a $h_2$ in $H^{rxl}_2$ such that

$$\frac{\langle \tilde{U}_{\infty} d_1, (W+Uh_2) \rangle}{\|\tilde{U}_{\infty} d_1\|_2 \cdot \|h_2\|} = \alpha,$$ \hspace{1cm} (4.56)

or

$$\frac{\langle \tilde{U}_{\infty} d_1, Wh_2 \rangle}{\|\tilde{U}_{\infty} d_1\|_2 \cdot \|h_2\|} = \alpha.$$ \hspace{1cm} (4.57)

Take the projection of $\tilde{U}_{\infty} d_1$ to $K^{mxl}_U$. 
\[ s_1 = P_{\tilde{U}^*_{KU}} U_{\tilde{U}_{KU}} \]  

(4.58)

since \( \tilde{U}_{KU} \) is in

\[ L_{\tilde{U}^*_{KU} H_{KU}^*} = H_{KU}^* H_{KU}^* \epsilon U_{\tilde{U}_{KU}} \]  

so write \( U_{\tilde{U}_{KU}} = \tilde{h}_3 + g_1 \) for some \( \tilde{h}_3 \) in \( H_{KU}^* \), we have

\[ \langle U_{\tilde{U}_{KU}}, Wh_2 \rangle = \langle \tilde{h}_3, Wh_2 \rangle = \langle W^* g_1, h_2 \rangle = \langle P_{H_{KU}^*} W^* g_1, h_2 \rangle \]  

(4.59)

From (4.58),

\[ \|g_1\|_2 \leq \|\tilde{h}_3\|_2 \]  

(4.60)

Combine (4.57), (4.59) and (4.60), we have

\[ \langle P_{H_{KU}^*} W^* g_1, h_2 \rangle \]  

\[ \frac{\|g_1\|_2 \cdot \|h_2\|_2}{\|g_1\|_2 \cdot \|h_2\|_2} \geq \alpha \]  

(4.61)

It is also

\[ \langle P_{H_{KU}^*} W^* g_1, h_2 \rangle \]  

\[ \frac{\|g_1\|_2 \cdot \|h_2\|_2}{\|g_1\|_2 \cdot \|h_2\|_2} \geq \alpha \]  

(4.62)

Since later we will prove that \( U_{\tilde{U}_{KU}}^* \) is in \( H_{KU}^* \) and therefore

\[ P_{H_{KU}^*} U_{\tilde{U}_{KU}}^* g_1 = 0 \]

From (4.62) we have
\[ \| P_{H_2}^{rxl} \|_{K_{mxl}}^* \|_{K_{U}}^{mxl} \geq \alpha. \]  

(4.63)

On the other hand, suppose

\[ \| P_{H_2}^{rxl} \|_{K_{U}}^* \|_{K_{mxl}} = \beta, \]  

(4.64)

from (4.50), we can find \( g_1 \in K_U^{mxl} \) and \( h_2 \in H_2^{rxl} \), such that

\[ \langle P_{H_2}^{rxl} \|_{H_2}^* g_1, h_2 \rangle \]

\[ \frac{\| g_1 \|_{H_2} \cdot \| h_2 \|_{H_2}}{\| g_1 \|_{H_2} \cdot \| h_2 \|_{H_2}} = \beta. \]  

(4.65)

This is

\[ \langle g_1, h_2 \rangle \]

\[ \frac{\| g_1 \|_{H_2} \cdot \| h_2 \|_{H_2}}{\| g_1 \|_{H_2} \cdot \| h_2 \|_{H_2}} = \beta. \]  

(4.66)

Write (4.66) in the form

\[ \langle UU^* g_1, Wh_2 \rangle \]

\[ \frac{\| UU^* g_1 \|_{H_2} \cdot \| Wh_2 \|_{H_2}}{\| UU^* g_1 \|_{H_2} \cdot \| Wh_2 \|_{H_2}} = \beta. \]  

(4.67)

We claim that \( \tilde{h}_1 = U^* g_1 \) is in \( H_2^{mxl} \). It is from

\[ \langle U^* g_1, h \rangle = \langle g_1, Uh \rangle = 0 \]

for all \( h \in H_2^{mxl} \) (notice \( g_1 \in K_U^{mxl} \) which is \( H_2^{mxl} \cap UH_2^{mxl} \)).

Put \( \tilde{h}_1 \) into (4.67), and recognize (4.67) being
we have $\beta \leq \alpha$. Compare it with (4.63), we have $\alpha = \beta$, therefore (4.51) is proved.

Now we are in the position to give another proof of theorem 4.2

(Another proof of theorem 4.2)

Let

$$g_i^j = \sum_{t=0}^{i} \frac{(-1)^{i-t} \cdot (i-t)! \cdot c_i^t \cdot a_i^t}{(s+s_i^j)^{i+1-t}}, \quad (4.69)$$

$i=1, \ldots, n, \quad j=0, \ldots, i-1$.

It is easy to check

$$\langle g_i^j, U_h \rangle = 0$$

for all $h \in H^\alpha_2$, therefore these $g_i^j$ are the basis of $K^{\alpha\beta}_U$.

Any element $g$ in $K^{\alpha\beta}_U$ can be written as

$$g = \sum_{i=0}^{n} \sum_{j=0}^{i-1} g_i^j x_i^j. \quad (4.70)$$

then

$$P_{H^\alpha_2} W(s) g$$
\[ E \equiv \sum_{i=0}^{\infty} \frac{1}{i^t} \sum_{i=0}^{t} \frac{(-1)^{i-t}(i-t)!}{(s+s_i)^{i+1-t}} a_1^t x_i \]

\[ = \sum_{t=0}^{\infty} \sum_{i=0}^{t} \frac{(-1)^{i-t}(i-t)!}{(s+s_i)^{i+1-t}} a_1^t x_i \]

Where \( t' = t + i' \).

Write

\[ X = (x_1, \ldots, x_1, \ldots, x_n) \]

and after manipulation we have

\[ \langle g, g \rangle = X^T Q X \]  \hspace{1cm} (4.72)

and

\[ \langle P_{H_2}^r W^* g, P_{H_2}^r W^* g \rangle = X^T RX \]  \hspace{1cm} (4.73)

Where Q and R are from (3.47) and (4.46) respectively.
Therefore $\|P_{xx}W^*\|_{\text{mxx}}$ is

\[
\sup_{x, x' \in Q} \frac{x^*Rx'}{x^*Qx},
\]

(4.74)

now the theorem is obvious.

In this chapter we find that the optimal $H_\infty$-norm is the positive square root of the maximal eigenvalue of (4.45) by two different considerations. The first method only needs a little elementary mathematical background and shows the close relation between $H_2$ and $H_\infty$ problems. In next chapter we will try to organize all the $H_\infty$ optimal solutions.
CHAPTER 5 CONSTRUCTION OF ALL $\Psi$ IN UNIT BALL

5.1 Introduction

In the last chapter we have found the optimal norm of (4.1). After dividing by the optimal norm the optimal solutions should be in the unit ball. In this chapter we will consider a more general problem to construct all

$$\Psi = (W-UH_1) \cap BH_{m,i}^{mxr}$$

(5.1)

where $W$ is in $H_{m,1}^{mxr}$, $U$ is in $H_{m,1}^{r}$, $H_1$ can be any element in $H_{m,1}^{r}$, $BH_{m,i}^{mxr}$ means the unit ball in $H_{m,1}^{mxr}$.

The main tool to solve the problem is to imbed the space

$$M = \begin{bmatrix} W & U \\ I & 0 \end{bmatrix} E_{2}^{m+rx}$$

(5.2)

into an indefinite inner product space which is called Krein space. The interpolation conditions (3.2) becomes that $M'$ is spanned by $b_i$ ($i=1,\ldots,n$), where $b_i$ is defined by (5.5). The graph of $\Psi$ which is in the unit ball is corresponding to a maximal negative space in the Krein space. To find all solutions $\Psi$ of (5.1) is to find all the maximal negative spaces in $M$ which are perpendicular to $b_i$ ($i=1,\ldots,n$).

In the case of $l=0$, from theorem 4.1 the existence of solutions of (5.1) depends on the maximal eigenvalue $\lambda_{\text{max}}$ in (4.45):

(i) if $\lambda_{\text{max}} > 1$, there is no solution for (5.1).
(ii) if $\lambda_{\text{max}} \leq 1$, there exist solutions.

In this chapter we will construct all the solutions for case (ii). We even can consider a more general problem of $H_{\alpha}eH_{\beta}^{\text{max}}$ instead of $HeH_{\alpha}^{\text{max}}$ without creating any difficulties. Later we can find that the criterion for the existence of solutions of (5.1) is the $2+1$ th largest eigenvalue of (4.45) not greater than 1:

$$\lambda_{1+1} \leq 1$$  \hspace{1cm} (5.3)

Let

$$P = Q - R = \left[ a_i^* V_{ij} a_j - \beta_i^* V_{ij} \beta_j \right]_{i, j=1, \ldots, n},$$  \hspace{1cm} (5.4)

$P$ is called a Pick matrix. The case of $\det P \neq 0$ is called nondegenerate case, and the case of $\det P = 0$ is called degenerate case.

From (4.45) and (5.4), the case of $\lambda_{\text{max}} < 1$ is nondegenerate case, and the case of $\lambda_{\text{max}} = 1$ is degenerate case.

5.2 The Case of Nondegenerate and Distinct Zeros

In this section we only consider the simplest case:

(i) All right half plane zeros $s_1, \ldots, s_n$ of $U$ are distinct, and we have (3.2).

(ii) Nondegenerate case.

Define
\[ b_i = \frac{1}{s+s_i} \begin{bmatrix} a_i \\ \beta_i \end{bmatrix}, \quad i=1,\ldots,n, \quad (5.5) \]

the companion \( M' \) of \( M \) in \( \mathbb{K}^{m,r} \) is spanned by \( b_1,\ldots,b_n \). Since \( P \) is Hermitian, there exist a unitary matrix

\[ U_0 = [ u_1, \ldots, u_n ] \quad (5.6) \]

such that

\[ P = U_0 \cdot \text{diag} [ \gamma_1, \ldots, \gamma_n ] \cdot U_0^* \quad (5.7) \]

Suppose

\[ \gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_{n-1} > 0 > \gamma_{n-1+1} \geq \cdots \geq \gamma_n, \quad (5.8) \]

which means that the \( l+1 \)th largest eigenvalue of \((4.2')\) is less than \( 1 \).

To see this, suppose \( Q = NN^* \), the eigenvalues and eigenvectors of \((4.2')\) are \( \lambda_i \) and \( x_i \). Let \( y_i = N^* x_i \), then we have \( N^{-1}RN^{-1} y_i = \lambda_i y_i \), therefore we can choose \( Y = [ y_1, \ldots, y_n ] \) as a unitary matrix, and

\[ R = NY \cdot \text{diag} [ \lambda_1, \ldots, \lambda_n ] \cdot Y^* \]

then

\[ P = Q - R = NY \cdot [ 1-\lambda_1, \ldots, 1-\lambda_n ] \cdot Y^* \]

compare it with \((5.7)\) and use inertial law of quadratic form in \( U\)-space\([58, \text{pp.296}]\), we arrive at the above conclusion.
Lemma 5.1

Under condition (5.8),

\[ B_i = [ b_1, \ldots, b_n ] \cdot u_i \cdot |\gamma_i|^2, \quad i=1,\ldots,n \]  \hspace{1cm} (5.9)

are the normalized orthogonal basis of \( M' \) in \( k^{m,r} \).

(Proof)

\[ [ B_i, B_j ] = |\gamma_i|^2 u_i^* P u_j |\gamma_j|^2 \]

\[ = |\gamma_i|^2 u_i^* u_j |\gamma_j|^2 \]

\[ \begin{bmatrix} 0 & i\neq j \\ 1 & i=j \leq n-1 \\ -1 & i=j > n-1 \end{bmatrix} \]

Lemma 5.2

\[ [ b_1, \ldots, b_n ]^* \cdot J \cdot [ b_1, \ldots, b_n ] = \]

\[ \text{diag } [ \frac{1}{-s+s_1}, \ldots, \frac{1}{-s+s_n} ] \cdot P + P \cdot \text{diag } [ \frac{1}{s+s_1}, \ldots, \frac{1}{s+s_n} ]. \]  \hspace{1cm} (5.10)

and

\[ \begin{bmatrix} a_1, \ldots, a_n \\ b_1, \ldots, b_n \end{bmatrix}^* \cdot J \cdot \begin{bmatrix} a_1, \ldots, a_n \\ b_1, \ldots, b_n \end{bmatrix} = \]

\[ \text{diag } [ (s+s_1), \ldots, (s+s_n) ] \cdot P + P \cdot \text{diag } [(s-s_1), \ldots, (s-s_n)]. \]  \hspace{1cm} (5.11)
(Proof).

It is a direct calculation.

The left side of (5.10)

\[
\begin{bmatrix}
\frac{a_i^*a_j - \beta_i^* \beta_j}{(-s+s_i)(s+s_j)}
\end{bmatrix}_{i,j=1,\ldots,n}
\]

\[
= \begin{bmatrix}
\frac{a_i^*a_j - \beta_i^* \beta_j}{(-s+s_i)(s+s_j)} + \frac{a_i^*a_j - \beta_i^* \beta_j}{(s+s_i)(s+s_j)}
\end{bmatrix}_{i,j=1,\ldots,n}
\]

= right side of (5.10).

(5.11) can be proved in a similar way.

Lemma 5.3

Under the condition (5.8), the shift invariant subspace \( M \) of (5.2) can be expressed as

\[
M = L \cdot E_2^{m+r},
\]

where

\[
L = I_{m+r} - [ b_1, \ldots, b_n ] \cdot P^{-1} [ \alpha^*, -\beta^* ].
\]

\( L \) is \( J \)-unitary:

\[
L^* J L = J.
\]
(Proof)

Under the condition (5.8), det P \neq 0 and P^{-1} exists. From (5.13) L is in $H_2^{(m+r) \times (m+r)}$, therefore $L \cdot H_2^{m+r}$ is in $H_2^{m+r}$ and $K^{m+r}$.

From (5.13) and (5.10), we have

$$[ b_1, \ldots, b_n ]^* J L = -P \cdot \text{diag} \left[ \frac{1}{s+s_1}, \ldots, \frac{1}{s+s_n} \right] \cdot P^{-1} \cdot [ \alpha^*, -\beta^* ] \quad (5.15)$$

Therefore

$$[ b_i, Lh ] = 0, \quad \text{for all } h \in H_2^{m+r}, \quad i = 1, \ldots, n. \quad (5.16)$$

Since $b_i$ (i = 1, ..., n) are all the basis of $M'$ and $L \cdot H_2^{m+r}$ is in $K^{m+r}$, (5.16) tells us that

$$L \cdot H_2^{m+r} \subseteq M \quad (5.17)$$

We will prove the inverse direction of (5.17):

$$M \subseteq L \cdot H_2^{m+r}. \quad (5.18)$$

For this purpose we first prove (5.14). From (5.13),

$$L^* J L = J - [ \alpha ] [ P^{-1} [ b_1, \ldots, b_n ]^* - [ b_1, \ldots, b_n ] P^{-1} [ \alpha^*, -\beta^* ] + [ \beta ] P^{-1} [ b_1, \ldots, b_n ] J [ b_1, \ldots, b_n ] P^{-1} [ \alpha^*, -\beta^* ], \quad (5.19)$$

notice in (5.19), $[ b_1, \ldots, b_n ]^* J [ b_1, \ldots, b_n ]$ can be expressed in the form of right side of (5.10), the last term of (5.19) can exactly be canceled by the second and the third term of (5.19). Therefore (5.14) is correct.
Multiplying $J$ on the both sides of (5.14), we have

$$(JL^*)^* J = L = I_{m+r}. \quad (5.20)$$

From (5.20), we find

$$L^{-1} = JL^*J$$

$$= I_{m+r} - \begin{bmatrix} a \end{bmatrix} p^{-1} \begin{bmatrix} b \end{bmatrix} \begin{bmatrix} J \\ \vdots \\ J \end{bmatrix} \begin{bmatrix} a \\ \vdots \\ a \end{bmatrix}$$

$$= I_{m+r} - \begin{bmatrix} a \end{bmatrix} p^{-1} \begin{bmatrix} b \end{bmatrix} \begin{bmatrix} J \\ \vdots \\ J \end{bmatrix} \begin{bmatrix} a \\ \vdots \\ a \end{bmatrix} \quad (5.21)$$

Obviously $L^{-1}$ is not in $H^m_\infty$, but

$$L^{-1} M = \begin{bmatrix} a \\ \vdots \\ a \\ -s+s_1 \end{bmatrix} \begin{bmatrix} a \\ \vdots \\ a \\ -s+s_1 \end{bmatrix} \begin{bmatrix} W & U \\ I_r & 0 \end{bmatrix} H^m_2 \quad (5.22)$$

is still in $H^m_2$ since (3.2) exists.

Furthermore, for any $h \in H^m_2$, we have

$$\langle L^{-1}Mh, L^{-1}Mh \rangle = \langle Mh, Mh \rangle \leq \|M\|_\infty^2 \|h\|_2^2 < \infty, \quad (5.23)$$

where

$$M = \begin{bmatrix} W & U \\ I_r & 0 \end{bmatrix} \quad (5.24)$$

From (5.22) and (5.23) we have
therefore (5.18) is true. From (5.17) and (5.18) we have (5.12).

Lemma 5.4

Under condition (5.8), we have the orthogonal decomposition of the Krein space $K^{m,r}$:

$$K^{m,r} = B_1 \mathcal{C} \oplus B_2 \mathcal{C} \oplus \ldots \oplus B_n \mathcal{C} \oplus L_1 H_2 \oplus \ldots \oplus L_{m+r} H_2,$$  \hspace{1cm} (5.26)

and

$$B_1 \mathcal{C}, \ldots, B_{n-1} \mathcal{C}, L_1 H_2, \ldots, L_{m+r} H_2$$

are all positive spaces;

$$B_{n-1+1} \mathcal{C}, \ldots, B_n \mathcal{C}, L_{m+1} H_2, \ldots, L_{m+r} H_2$$

are all negative spaces.

$L_i (i=1, \ldots, m+r)$ is the column of $L$.

(Proof)

It's a conclusion from lemma 5.1 and lemma 5.3.

Lemma 5.5

Any shift invariant maximal negative subspace $\mathfrak{g}$ of $K^{m,r}$ has the form

$$\mathfrak{g} = \left[ \begin{array}{c} \mathfrak{H} \\ [I_{m} \mathcal{C} \oplus \mathcal{L} \oplus \mathcal{I}_{m+h_2} H_2] \end{array} \right]$$  \hspace{1cm} (5.21)

where $\mathfrak{H} \in \mathcal{B}(H_{m+h_2})$. 
Lemma 5.5' 

Any shift invariant negative subspace \( \mathcal{Q} \) of \( K^{m,r} \) has the form

\[
\mathcal{Q} = \begin{bmatrix}
\mathcal{Q}_1 \\
\mathcal{Q}_2 \\
I_r \\
\end{bmatrix}
\]

(5.28)

where \( \mathcal{Q}_1 \) is an inner function matrix with \( \lambda \) right half plane zeros, \( \mathcal{Q}_2 \in \mathbb{B}_{m,r} \), \( \lambda \) is the codimension of \( \mathcal{Q} \) in the maximal negative subspace of \( K^{m,r} \) which contains \( \mathcal{Q} \).

(Proof)

The proof of lemma 5.5 and lemma 5.5' is based on the concept of maximal negative space and negative space, the same as [8] and [9].

Lemma 5.6

Under condition (5.8), a shift invariant space \( \mathcal{Q} \) is a maximal negative subspace of \( M \) iff it is in the form

\[
\mathcal{Q} = L \cdot \begin{bmatrix}
X \\
I_r \\
\end{bmatrix} \cdot H_{2,r}^r
\]

(5.29)

where \( X \) is in \( \mathbb{B}_{m,r} \).

(Proof)

From (5.14)

\[
[\text{Lh}, \text{Lh}] = [h, h], \quad h \in K^{m,r}
\]

(5.30)

From (5.30) we conclude that
\[
L : K^{m,r} \rightarrow M
\]

is an isomorphism. From lemma 5.5, the maximal negative invariant space in \( K^{m,r} \) is in the form

\[
\begin{bmatrix}
X \\
I \\
I
\end{bmatrix}
\cdot H^r_2,
\]

therefore

\[
\phi = L \cdot \begin{bmatrix}
X \\
I \\
I
\end{bmatrix}
\cdot H^r_2
\]

is the maximal negative invariant subspace of \( M \), and vice versa.

**Theorem 5.1**

Under condition (5.8), we have

\[
[ W + U H_1 ] \cap BH^\text{mxf}_{\omega,1} = L(X) \tag{5.31}
\]

for some \( H_1 \in H^\text{mxf}_{\omega,1} \) and \( X \in BH^\text{mxf}_{\omega} \). where

\[
L(X) = ( L_{11}X + L_{12} ) \cdot ( L_{21}X + L_{22} )^{-1} \tag{5.32}
\]

\( L \) is from (5.13), and is divided into four blocks:

\[
L = \begin{bmatrix}
L_{11} & L_{12} \\
L_{21} & L_{22}
\end{bmatrix}
\]

(\( m \times r \))

(\( \begin{array}{c}
m \\
r
\end{array} \))

\[\text{Consider subspaces}\]
\[
\phi_1 = \begin{bmatrix}
L(X) \\
I_r
\end{bmatrix} \cdot Q_1 H_2^r
\]

(5.34)

\[
\phi_2 = L \cdot \begin{bmatrix}
X \\
I_r
\end{bmatrix} \cdot H_2^r
\]

(5.35)

\[
\phi_3 = \begin{bmatrix}
W + U H_1 \\
I_r
\end{bmatrix} \cdot Q_1 H_2^r
\]

(5.36)

\[
M = \begin{bmatrix}
W & U \\
I & 0
\end{bmatrix} \cdot H_2^{m+r}
\]

(5.37)

From lemma 5.6, \( \phi_2 \) is a maximal negative subspace of \( M \), and from lemma 5.4 there is a negative subspace of \( \ell \) dimension in \( M' \), therefore from lemma 5.5', \( \phi_2 \) should be in the form

\[
\phi_2 = \begin{bmatrix}
\phi_t \\
I_r
\end{bmatrix} \cdot Q_1 H_2^r
\]

(5.38)

Compare (5.38) with (5.35), \( (L_{21} X + L_{22}) \) should have only \( \alpha \) zeros in right half plane, and therefore \( L(X) \) has only \( \ell \) poles in right half plane. We also find

\[
\phi_t = L(X),
\]

(5.39)

therefore

\[
\phi_1 = \phi_2.
\]

(5.40)

Notice \( \phi_2 \in M \) we have

\[
\phi_1 \in M
\]

(5.41)
(5.41) means that for any $h_r \in H_2^r$ we can find a $h_m \in H_2^m$ such that
\[ L(X)Q_r h_r = WQ_r h_r + Uh_m. \]  
(5.42)

Let the correspondence from $h_r$ to $h_m$ be
\[ h_m = H^{-1}h_r \]  
(5.43)

$H$ is in $H_{\omega,1}^{mxr}$, the multiplicative property is from both sides of (5.42) being shift invariant. Put (5.43) into (5.42), we have
\[ L(X)Q_t = WQ_t + Uh \]
or
\[ L(X) = W + Uh_t \]  
(5.31)

and $H_t = HQ_t^{-1}$ is in $H_{\omega,1}^{mxr}$. The argument above tells us that for any $X$ in $BH_{\omega,1}^{mxr}$, we can find some $H_t$ in $H_{\omega,1}^{mxr}$ such that $W + Uh_t$ in $BH_{\omega,1}^{mxr}$, and (5.31) exists.

The next step is to prove the inverse direction. Suppose there exist a $H_t \in H_{\omega,1}^{mxr}$ such that
\[ W + UH_t \in BH_{\omega,1}^{mxr}, \]
and suppose $H_t = HQ_t^{-1}$, $H$ is in $H_{\omega,1}^{mxr}$ and $Q_t$ is inner function matrix with $\lambda$ zeros in right half plane. Then from lemma 5', $\phi_3$ is a negative subspace of $K_{\omega,r}^{mx}$. Obviously
\[ \phi_3 \in \mathcal{M} \]  
(5.44)
From lemma 5.4, under condition (5.8) \( M \) and \( l \) other negative vectors \( (B_{n-1+1}, . . . , B_n) \) span a maximal negative space of \( \mathbb{K}^{m,r} \). From the expression (5.36) \( \phi_3 \) and some \( l \) negative vectors can span a maximal negative subspace of \( \mathbb{K}^{m,r} \). Therefore \( \phi_3 \) is a maximal negative subspace of \( M \). \( \phi_3 \) is shift invariant, from lemma 5.6, it should be in the form

\[
\phi_3 = L \begin{bmatrix} X \end{bmatrix} \cdot H_2 \begin{bmatrix} 1 \end{bmatrix}
\]

(5.45)

for some \( X \in \mathbb{B}^{m,r}_{\infty} \). This is

\[
\phi_3 = \phi_2.
\]

(5.46)

From (5.36), (5.35) and (5.46), we have that \( (L_{21}X + L_{22}) \) has \( l \) zeros in right half plane, therefore \( L(X) \) in \( \mathbb{B}^{m,r}_{\infty,1} \) and

\[
\phi_2 = \phi_1.
\]

(5.47)

From (5.46) and (5.47) we have

\[
\phi_3 = \phi_1,
\]

(5.48)

which means, for any \( H_1 \in \mathbb{H}^{m,r}_{\infty,1} \) such that

\[ W + U H_1 \in \mathbb{B}^{m,r}_{\infty,1} \]

we can find a \( X \) in \( \mathbb{B}^{m,r}_{\infty} \) such that

\[ W + U H_1 = L(X). \]

(5.31) is proved.
5.3 The case of Degenerate and Distinct Zeros

This section we consider the degenerate case. In this case we can find a unitary matrix $U_0$ such that

$$ P = U_0 \cdot \text{diag} \left[ \gamma_1, \ldots, \gamma_{n-r_0}, 0, \ldots, 0 \right] \cdot U_0^* \quad (5.49) $$

from Cholesky decomposition, write

$$ U_0 = \begin{bmatrix} u_1, \ldots, u_n \end{bmatrix}, \quad (5.50) $$

and choose unitary matrices $V_1$ and $V_2$ such that

$$ \begin{bmatrix} * \\ u_{n-r_0+1} \\ \vdots \\ * \\ u_n \end{bmatrix} \cdot a^* \cdot V_1 = 0_{r_0 \times (m-r_0)}, \quad (5.51) $$

$$ \begin{bmatrix} * \\ u_{n-r_0+1} \\ \vdots \\ * \\ u_n \end{bmatrix} \cdot \beta^* \cdot V_2 = 0_{r_0 \times (m-r_0)}, \quad (5.52) $$

where

$V_1^L$ is the $1$st $\rightarrow (m-r_0)$ th columns of $V_1$, $V_2^L$ is the $1$st $\rightarrow (r-r_0)$ th columns of $V_2$.

Define

$$ [ L_+ \ L_- ] = [ I_{m+r_0} - (b_1, \ldots, b_n)(u_1, \ldots, u_{n-r_0}) \cdot \text{diag}(\gamma_1^{-1}, \ldots, \gamma_{n-r_0}^{-1}) $$
\[(u_1, \ldots, u_{n-r_0})^* (a^*, -b^*) \begin{bmatrix} v_1^L & 0 \\ 0 & v_2^L \end{bmatrix}, \quad (5.53)\]

and

\[L_0 = (b_1, \ldots, b_n)^* (u_{n-r_0+1}, \ldots, u_n), \quad (5.54)\]

\[N = L_0 \cdot H_2^o. \quad (5.55)\]

**Lemma 5.7**

\[
\begin{bmatrix}
I_{m-r_0} \\
\end{bmatrix}
\begin{bmatrix}
L_+ L_0 \\
L_- L_0 \\
\end{bmatrix}^* 
\begin{bmatrix}
I_{r-r_0} \\
0 \\
\end{bmatrix}
\begin{bmatrix}
L_+ L_0 \\
L_- L_0 \\
\end{bmatrix} 
= 
\begin{bmatrix}
I_{r-r_0} \\
0 \\
\end{bmatrix} 
\begin{bmatrix}
L_+ L_0 \\
L_- L_0 \\
\end{bmatrix}. 
\quad (5.56)
\]

(Proof)

The proof is a calculation similar to the proof in lemma 5.3 and mainly dependent on the result of lemma 5.2, and is omitted.

**Lemma 5.8**

\[
M \cap N' = (L_+ L_- L_0) \cdot H_2^o. \quad (5.57)
\]
(Proof)

A wandering subspace of shift invariant space $M$ in $K^{m,r}$ is defined by

$$\mathcal{L} = M \cap (\sigma M)'$$ (5.58)

where

$$\sigma = \frac{s-1}{s+1}$$ (5.59)

is shift operator.

According to Ball and Helton [9],

$$M \cap N' = \bigcup_{k>0} \sigma^k \mathcal{L}$$ (5.60)

We will use (5.60) as a start to prove (5.57). From (5.58),

$$\mathcal{L} = M \cap [(\sigma K^{m,r})' \cup \sigma M']$$

$$= M \cap \left[ \frac{\sqrt{2}}{s+1} C^{m,r} \cup \sigma M' \right]$$ (5.61)

Notice $M'$ is spanned by $b_1, \ldots, b_n$, and (5.59), (5.61), we have

$$(\sigma M)' = \frac{\sqrt{2}}{s+1} C + \frac{s-1}{s+1} [b_1, \ldots, b_n] \cdot |2D$$ (5.62)

where $C \in C^{m,r}$, $D \in C^n$.

Next step is to find $C$ and $D$ such that what expressed in (5.62) is also in $M$ and so that it represents. It requires (5.62) perpendicular to $b_1, \ldots, b_n$.
\[
[b_i, \frac{1}{s+1} + (b_1, \cdots, b_n) \cdot \frac{s-1}{s+1}, [2 \cdot D] = 0, \quad i=1, \cdots, n. \quad (5.63)
\]

This is

\[
(\alpha^*, -\beta^*) \cdot C + \text{diag}(s_1-1, \cdots, s_n-1) \cdot P \cdot D = 0, \quad (5.64)
\]

where \( P \) is Pick matrix of (5.4). In degenerate case, \( P \) is singular and has a decomposition (5.49). Put (5.49) into (5.64), we have

\[
\begin{bmatrix}
\ast \\
u_1 \\
\vdots \\
u_n
\end{bmatrix}
\begin{bmatrix}
\frac{1}{s_1-1} \\
\vdots \\
\frac{1}{s_n-1}
\end{bmatrix}
\begin{bmatrix}
\ast \\
u_1 \\
\vdots \\
u_n
\end{bmatrix}
\begin{bmatrix}
(\alpha^*, -\beta^*) \cdot C \\
0
\end{bmatrix}
\]

From (5.65) we have

\[
\begin{bmatrix}
\ast \\
u_{n-r_o} \\
\vdots \\
u_n
\end{bmatrix}
\begin{bmatrix}
\frac{1}{s_1-1} \\
\vdots \\
\frac{1}{s_n-1}
\end{bmatrix}
\begin{bmatrix}
\ast \\
u_{n-r_o} \\
\vdots \\
u_n
\end{bmatrix}
\begin{bmatrix}
(\alpha^*, -\beta^*) \cdot C = 0, \quad (5.66)
\end{bmatrix}
\]

and
\[ D = -(u_1, \ldots, u_{n-r_0})^T \begin{bmatrix} -1 \\ \vdots \\ \gamma_{n-r_0} \\ \vdots \\ \gamma_{n-r_0} \\ \vdots \\ -1 \end{bmatrix} \begin{bmatrix} u_1 \\ \vdots \\ u_{n-r_0} \\ \vdots \\ u_{n-r_0} \\ \vdots \\ \frac{1}{s-1} \end{bmatrix} \cdot (a^*,-\beta^*) \cdot C + (u_{n-r_0+1}, \ldots, u_n) \cdot E_1, \]  

(5.67)

where \( E_1 \in C^r \).

From (5.61), (5.62) and (5.67), we have

\[ \mathcal{L} = (I_{m+r} - (s-1) \cdot (b_1, \ldots, b_n) \cdot (u_1, \ldots, u_{n-r_0})^T \begin{bmatrix} -1 \\ \vdots \\ \gamma_{n-r_0} \\ \vdots \\ \gamma_{n-r_0} \\ \vdots \\ -1 \end{bmatrix}) \begin{bmatrix} \frac{1}{s-1} \\ \vdots \\ \frac{1}{s-1} \end{bmatrix} \cdot (a^*,-\beta^*) \cdot \sqrt{\frac{1}{s}} \cdot C \]

\[ + (b_1, \ldots, b_n) \cdot (u_{n-r_0+1}, \ldots, u_n) \cdot \frac{s-1}{s+1} \cdot E_1 \]

\[ = \mathcal{L}_1 + \mathcal{L}_{01} \]  

(5.68)

where we define

\[ \mathcal{L}_{01} = (b_1, \ldots, b_n) \cdot (u_{n-r_0+1}, \ldots, u_n) \cdot \frac{s-1}{s+1} \cdot E_1 \]  

(5.69)

which belongs to the neutral space \( N \).
Notice

\[(s-1) - (b_1, \ldots, b_n) = \begin{bmatrix} a \end{bmatrix} - \begin{bmatrix} b_1, \ldots, b_n \end{bmatrix} \begin{bmatrix} \alpha \end{bmatrix} \begin{bmatrix} \beta \end{bmatrix} \begin{bmatrix} -1 \end{bmatrix} \begin{bmatrix} s_1 + 1 \\ \vdots \\ - \frac{1}{s_n + 1} \end{bmatrix} \]

we can write \( L_1 \), the first part in (5.68), as follows:

\[ L_1 = \begin{bmatrix} y_{1}^{-1} & u_1 \\ \vdots & \vdots \\ y_{n-r_0}^{-1} & u_{n-r_0} \end{bmatrix} \begin{bmatrix} \gamma_{1}^{-1} \\ \vdots \\ \gamma_{n-r_0}^{-1} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_{n-r_0} \end{bmatrix} + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} \alpha, -\beta \end{bmatrix} \begin{bmatrix} b_1, \ldots, b_n \end{bmatrix} \begin{bmatrix} y_{1}^{-1} & u_1 \\ \vdots & \vdots \\ y_{n-r_0}^{-1} & u_{n-r_0} \end{bmatrix} \begin{bmatrix} \alpha, -\beta \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} s_1 + 1 \\ s_1^{-1} \\ \vdots \\ s_n^{-1} \end{bmatrix} \]

Notice (5.10) and (5.11), after a long manipulation, (5.71) becomes:

\[ L_1 = \begin{bmatrix} y_{1}^{-1} & u_1 \\ \vdots & \vdots \\ y_{n-r_0}^{-1} & u_{n-r_0} \end{bmatrix} \begin{bmatrix} \gamma_{1}^{-1} \\ \vdots \\ \gamma_{n-r_0}^{-1} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_{n-r_0} \end{bmatrix} + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} \alpha, -\beta \end{bmatrix} \begin{bmatrix} b_1, \ldots, b_n \end{bmatrix} \begin{bmatrix} y_{1}^{-1} & u_1 \\ \vdots & \vdots \\ y_{n-r_0}^{-1} & u_{n-r_0} \end{bmatrix} \begin{bmatrix} \alpha, -\beta \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} s_1 + 1 \\ s_1^{-1} \\ \vdots \\ s_n^{-1} \end{bmatrix} \]
\[
\begin{aligned}
&\cdot \begin{bmatrix} a \\ \gamma \end{bmatrix} \cdot \begin{bmatrix} \gamma_1^{-1} & u_1^* \\ \gamma_{n-r_o}^{-1} & u_{n-r_o}^* \end{bmatrix} \cdot \begin{bmatrix} 1 \\ s_{n-1} \end{bmatrix} \\
&\cdot \begin{bmatrix} \gamma_n^{-1} & \gamma_{n-r_o}^{-1} \\ \gamma_{n-r_o} & u_{n-r_o} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ s_{n-1} \end{bmatrix}
\end{aligned}
\]
\[
\cdot C \cdot \sqrt{\frac{1}{2}} \cdot \begin{bmatrix} b_1, \cdots, b_n \end{bmatrix} \cdot \begin{bmatrix} u_{n-r_o+1}, \cdots, u_n \end{bmatrix} \cdot E_2 \cdot \sqrt{\frac{1}{2}}
\]
\[
= \mathcal{L}_2 + \mathcal{L}_{02}
\]
\[(5.72)\]

where
\[
E_2 = \begin{bmatrix}
\sqrt{\frac{1}{2}} \\
\gamma_1 \\
\gamma_{n-r_o} \\
\gamma_{n-r_o} \\
\gamma_{n-r_o} \\
\gamma_{n-r_o}
\end{bmatrix}
\]
\[
\begin{bmatrix}
\gamma_1^{-1} & u_1^* \\
\gamma_{n-r_o}^{-1} & u_{n-r_o}^* \\
\gamma_n^{-1} & \gamma_{n-r_o}^{-1} \\
\gamma_{n-r_o} & u_{n-r_o} \\
\gamma_{n-r_o} & u_{n-r_o} \\
\gamma_{n-r_o} & u_{n-r_o}
\end{bmatrix}
\]
\[
\begin{bmatrix}
\gamma_1^{-1} & u_1^* \\
\gamma_{n-r_o}^{-1} & u_{n-r_o}^* \\
\gamma_n^{-1} & \gamma_{n-r_o}^{-1} \\
\gamma_{n-r_o} & u_{n-r_o} \\
\gamma_{n-r_o} & u_{n-r_o} \\
\gamma_{n-r_o} & u_{n-r_o}
\end{bmatrix}
\]
\[
\cdot \begin{bmatrix} \sqrt{\frac{1}{2}} \\ \gamma \end{bmatrix} \cdot \begin{bmatrix} \gamma_1^{-1} & u_1^* \\ \gamma_{n-r_o}^{-1} & u_{n-r_o}^* \end{bmatrix} \cdot \begin{bmatrix} 1 \\ s_{n-1} \end{bmatrix} \\
\cdot \begin{bmatrix} \gamma_n^{-1} & \gamma_{n-r_o}^{-1} \\ \gamma_{n-r_o} & u_{n-r_o} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ s_{n-1} \end{bmatrix}
\end{aligned}
\]
\[(5.73)\]

Again the last part in (5.72)
\[
\mathcal{L}_{02} = \begin{bmatrix} b_1, \cdots, b_n \end{bmatrix} \cdot \begin{bmatrix} u_{n-r_o+1}, \cdots, u_n \end{bmatrix} \cdot E_2 \cdot \sqrt{\frac{1}{2}}
\]
\[(5.74)\]

belongs to the neutral space \( N \).

The first part in (5.72) can be written as follows:
\[ I_2 = (I_{m+r} - [b_1, \ldots, b_n] \cdot [u_1, \ldots, u_{n-r_0}]) \cdot \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_{n-r_0} \end{bmatrix} \]

\[
\begin{bmatrix}
\vdots \\
\gamma_1 \\
\vdots \\
\gamma_{n-r_0}
\end{bmatrix}
\]

\[
\begin{bmatrix}
u_1 \\
y_1 \\
\vdots \\
\gamma_{n-r_0}
\end{bmatrix}
\]

\[
\vdots
\]

\[
\begin{bmatrix}
[a^*, -\gamma^*] \cdot R, \sqrt{s} \\
[a^*, -\gamma^*] \\
\vdots \\
[a^*, -\gamma^*] \\
\end{bmatrix}
\]

where

\[
R = (I_{m+r} - [\alpha] \cdot [u_1, \ldots, u_{n-r_0}]) \cdot \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_{n-r_0} \end{bmatrix}
\]

\[
\begin{bmatrix}
\vdots \\
\gamma_1 \\
\vdots \\
\gamma_{n-r_0}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\gamma_1 \\
\vdots \\
\gamma_{n-r_0}
\end{bmatrix}
\]

\[
\begin{bmatrix}
[a^*, -\gamma^*] \cdot C
\end{bmatrix}
\]

It is easy to verify that (5.66) is equivalent to (5.77):

\[
\begin{bmatrix}
u_{n-r_0+1} \\
\vdots \\
u_n
\end{bmatrix}
\]

\[
\begin{bmatrix}
[a^*, -\gamma^*] \cdot R = 0
\end{bmatrix}
\]

Notice \( C^m \), and it must satisfy (5.66), therefore it has only \( m+r-r_0 \) degrees of freedom. Furthermore, if we give the degrees of freedom
to $E_2 e^{C^o}$ in (5.73), then $C$ will lose another $r_0$ degrees of freedom. So we find that $C$ has only $m + r - 2r_0$ degrees of freedom. In this case $R$ has also $m + r - 2r_0$ degrees of freedom from (5.76).

Let

$$R = \begin{bmatrix} v_1^L & 0 \\ 0 & v_2^L \\ \end{bmatrix} R_1,$$

where

$$R_1 \in C^{m + r - 2r_0},$$

$v_1^L$ is 1 th to $(m - r_1)$ th column of $m \times m$ unitary matrix $V_1$,

$v_2^L$ is 1 th to $(r - r_0)$ th column of $r \times r$ unitary matrix $V_2$, and $v_1^L, v_2^L$ satisfies (5.51), (5.52) respectively, then (5.77) is satisfied automatically.

Put (5.78) into (5.75), we have

$$\mathcal{L}_2 = [ L_+ \ L_- ] \cdot R_1 \cdot \sqrt{\frac{1}{s+1}}.$$

From the expressions of (5.69) and (5.74), we have

$$\mathcal{L}_{01} = L_0 \cdot E_1 \cdot \sqrt{\frac{1}{s+1}},$$

$$\mathcal{L}_{02} = L_0 \cdot E_2 \cdot \sqrt{\frac{1}{s+1}},$$

where $L_+, L_-, L_0$ come from (5.53), (5.54), and

$$R_1 \in C^{m + r - 2r_0}, E_1 \in C^{r}, E_2 \in C^{r}.$$
From (5.68), (5.72), we have

\[ L = L_2 + L_{01} + L_{02}. \]  

(5.82)

Put (5.82), (5.79), (5.80), (5.81) into (5.60), we have proved (5.57).

**lemma 5.9**

A shift invariant space \( \phi \) is a maximal negative subspace of \( M \cap N' \) iff it is in the form

\[
\phi = L \cdot I_{x-r_0} \cdot H_2^r.
\]

(5.83)

where

\[
M \cap N' = L \cdot H_2^{m+r-r_0},
\]

(5.84)

\[
L = \begin{bmatrix}
L_+ & L_{-} & L_0
\end{bmatrix},
\]

(5.85)

\( L_+ \), \( L_- \), \( L_0 \) are from (5.53), (5.54), \( X \) is in \( BH_\infty \).

(Proof)

From (5.57) we can write

\[
M \cap N' = [L_+ L_-] \cdot H_2^{m+r-2r_0} + L_0 \cdot H_2^r.
\]

(5.86)
The shift invariant maximal negative subspace in \([ L_+ L_- ] \cdot H_2^{m+r-2r_0} \) should be in the form

\[
\begin{bmatrix}
  X \\
  I_{r-r_0} \\
  I_{r-r_0}
\end{bmatrix} \cdot H_2^{r-r_0} \tag{5.87}
\]

according to the result of lemma 5.6.

Notice that \( L_0 \cdot H_2^r \) is neutral space and orthogonal to \([ L_+ L_- ] \cdot H_2^{m+r-2r_0} \), we have proved the lemma.

**Theorem 5.2**

Suppose \( P \) has \( i \) negative and \( r_0 \) zero eigenvalues. then

\[
(W + U H_1) \cap H_{\omega,1}^m = L(X) \tag{5.88}
\]

for \( H_1 \in H_{\omega,1}^m \) and \( X \in H_{\omega,1}^{m+r_0} \), where

\[
L(X) = (L_{11}L_{12} + L_{10}, L_{10})^{-1} \tag{5.89}
\]

\[
L = [L_+ L_-] = \begin{bmatrix}
  L_{11} & L_{12} & L_{01} \\
  L_{21} & L_{22} & L_{02}
\end{bmatrix}^m_{r} \tag{5.90}
\]

\[
L_+ L_- = \begin{bmatrix}
  r_{-r_0} \\
  r_{-r_0}
\end{bmatrix}^r_{r_0}
\]

(Proof)

Consider spaces

\[
\phi_1 = \begin{bmatrix}
  L(X) \\
  I_r \cdot H_2^r
\end{bmatrix} \tag{5.91}
\]
\[ \varphi_2 = L \cdot \begin{bmatrix} X & \cdot H_2^r \\ I_{r-r_0} \end{bmatrix}, \quad (5.92) \]

\[ \varphi_3 = \begin{bmatrix} W + U H_1 \\ I_r \end{bmatrix} \cdot Q_r H_2^r, \quad (5.93) \]

\[ M \cap N' = L \cdot R_2^{m+r-r_0} \]

\[ M = \begin{bmatrix} W & U \\ I_r & 0 \end{bmatrix} \cdot H_2^{m+r}. \quad (5.95) \]

We start from \( \varphi_2 \). From lemma 5.9, \( \varphi_2 \) is a shift invariant maximal negative subspace of \( M \cap N' \), it must contain

\[ (M \cap N') \cap (M \cap N')' = N. \]

Obviously,

\[ \varphi_2 \in N, \]

we are going to show that \( \varphi_2 \) is also a maximal negative subspace of \( M \).

Suppose there exist a negative space \( \mathcal{T} \) such that

\[ \varphi_2 \in \mathcal{T} \in M, \]

since

\[ N \in \varphi_2, \]

therefore
Since \( N \in \mathcal{T} \).

Since \( N \) is neutral space, \( \mathcal{T} \) is negative, they must be orthogonal to each other:

\[ [N, \mathcal{T}] = 0, \]

therefore,

\[ \mathcal{T} \in N'. \]

Now \( \mathcal{T} \) should be a negative subspace of \( M \cap N' \). Since \( \mathcal{Q}_2 \) is a maximal negative subspace of \( M \cap N' \), therefore \( \mathcal{Q}_2 = \mathcal{T} \), and \( \mathcal{Q}_2 \) must also be a maximal negative subspace of \( M \).

Since \( M' \) only contains a 1 dimension strict negative subspace and a neutral subspace which is also contained in \( N \) and \( \mathcal{Q}_2 \), from lemma 5', \( \mathcal{Q}_2 \) must be in the form:

\[ \mathcal{Q}_2 = \begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix}. \]

This means that \((L_{21}, L_{22}, L_{02})\) has only 1 zeros in the right half plane, and \( \mathcal{Q}_2 \) can be written in the form \( \mathcal{Q}_1 \). We have

\[ \mathcal{Q}_1 = \mathcal{Q}_2 \subset M \cap N' \subset M, \]

this is

\[ \mathcal{Q}_1 \in M. \]
From (5.91), (5.95), (5.97), using the same argument as in the proof of theorem 5.1, we get following result:

\[(m-r_0)(x-r_0)\]

For any \(X \in BH_{\omega,0}\), we can find a \(H_1 \in H_{0,1}^{mxr}\) such that

\[(W+UH_1) = L(X)\quad (5.88)\]

and \(W+UH_1\) or \(L(X)\) is in \(BH_{\omega,1}^{mxr}\).

Now we prove the inverse direction. Suppose there exist a \(H_1 \in H_{0,1}^{mxr}\)
such that

\[W+UH_1 \in BH_{\omega,1}^{mxr}\].

Let \(H_1 = HQ_1^{-1}\), \(H\) is in \(H_{\omega,0}^{mxr}\) and \(Q_1\) is inner and with \(\lambda\) right half plane zeros, then \(\hat{\phi}_3\) is a negative subspace of \(K_{\omega}^{mxr}\). Obviously, \(\hat{\phi}_3 \subset M\).

From (5.93) \(\hat{\phi}_3\) and \(\lambda\) other negative vectors span a maximal negative subspace of \(K_{\omega}^{mxr}\), but we already know that \(M\) and \(\lambda\) negative vectors in \(M'\) span a maximal negative subspace of \(K_{\omega}^{mxr}\), therefore \(\hat{\phi}_3\) must be a maximal negative subspace of \(M\). So it contains \(M \cap M'\). Since \(\hat{\phi}_3\) is shift invariant, \(\hat{\phi}_3\) contains \(N\). Notice \(N\) is neutral, and \(\hat{\phi}_3\) is negative space.

\[[N, \hat{\phi}_3] = 0\],

therefore \(\hat{\phi}_3 \subset N'\), and

\[\hat{\phi}_3 \subset M \cap N'\quad (5.98)\]

\(\hat{\phi}_3\) should be a maximal negative shift invariant subspace of \(M \cap N'\),

from lemma 5.9, there exist a \(X \in BH_{\omega,0}^{mxr}\), such that
\[ \phi_3 = \phi_2 . \quad (5.99) \]

From (5.92), (5.93), (5.99), \((L_{21} x + L_{22} y, L_{02})\) has \(i\) zeros in right half plane, therefore (5.92) can be written in the form (5.91):

\[ \phi_3 = \phi_1 . \quad (5.100) \]

From (5.100) we conclude that for any \(X_i\) in \(H^{mxr}_{\infty, 1}\) such that \((m-r-o)(r-r-o)\)
\(W + U X_i \in H^{mxr}_{\infty, 1}\), we can find a \(X\) in \(H^{mxr}_{\infty}\) such that \(W + U X = L(X)\).

Theorem 5.2 is proved.

5.4 The Case of Degenerate and Repeated Zeros

To complete our results we need to consider the repeated zeros case which is a bit cumbersome. Since the treatment and proof are the same as the distinct zeros case, we only write down the last results for practical use.

Theorem 5.3

When \(W\) and \(U\) are in \(H^{mxr}_{\infty}\) and \(H^{mxm}_{\infty}\) respectively, and \(U\) has repeated zeros \(s_1, \ldots, s_n\), and the multiple degree of \(s_i\) is \(i\) according to (3.38), (3.42). Let \(a_i, \beta_i, V_{ij}, g_i, Q, R\) in forms (3.45), (3.46), (3.44), (3.50), (3.47) and (4.46) respectively.

\[ A_i = [a_i^0, a_i^1, \ldots, a_i^{i-1}] , \quad (5.101) \]

\[ A = [A_1, A_2, \ldots, A_n] , \quad (5.102) \]
\[ B_i = \begin{bmatrix} \beta_i^0 & \beta_i^1 & \cdots & \beta_i^{i-1} \end{bmatrix}, \quad (5.103) \]
\[ B = \begin{bmatrix} B_1 & B_2 & \cdots & B_n \end{bmatrix}, \quad (5.104) \]
\[ P = Q - R, \quad (5.105) \]

find the Cholesky decomposition:
\[
P = U \cdot \begin{bmatrix} \gamma_1 & \cdots & \gamma_{\xi i - r_0} & \cdots & \gamma_{\xi i - r_0} \end{bmatrix} \cdot U^* \quad (5.106)
\]
and there are \( \gamma \) negative in \( \gamma_1, \ldots, \gamma_{\xi i - r_0} \).

Define
\[
[L_+ \, L_-] = [I_{m+r} - \{ g_1 a_1, g_2 a_2, \ldots, g_n a_n \}]
\]
\[
\cdot \{ u_1, \ldots, u_{\xi i - r_0} \} \cdot \text{diag}\{ \gamma_1^{-1}, \ldots, \gamma_{\xi i - r_0}^{-1} \} \cdot \{ u_1, \ldots, u_{\xi i - r_0} \}^*
\]
\[
\begin{bmatrix}
\begin{bmatrix} L_+^L & 0 \\
V_1 & 0
\end{bmatrix} \\
[A^* - B^*] \cdot L_2
\end{bmatrix} \quad (5.107)
\]
\[
L_o = \{ g_1 a_1, g_2 a_2, \ldots, g_n a_n \} \cdot \{ u_{\xi i - r_0 + 1}, \ldots, u_{\xi i} \} \quad (5.108)
\]

and divide
L = \begin{bmatrix} L_{11} & L_{12} & L_{01} \\ L_{21} & L_{22} & L_{02} \end{bmatrix}
\begin{array}{c}
\begin{bmatrix} m \\ r \end{bmatrix} \\
\begin{bmatrix} m-r_0 \\ r-r_0 \end{bmatrix} \\
r_0
\end{array}
\tag{5.109}

then

\phi = L(X) = (L_{11}X + L_{12}, L_{01}) (L_{21}X + L_{22}, L_{02})^{-1}
\tag{5.110}

represents total solutions of

\((W + U^T) \cap B_{\infty}^{max} \cap B_{\infty}^{max} \cap B_{\infty}^{max} \cap B_{\infty}^{max} \cap B_{\infty}^{max}
\tag{5.111}

for any X in \( B_{\infty}^{max} \).

5.5 Two-sided Problem in Square Case

A more general problem is

\[
\min_{H_{\infty}^{max}} \| W - UHV \|_{\infty}
\]
\tag{5.112}

for W, U, V in \( H_{\infty}^{max}, H_{\infty}^{max}, H_{\infty}^{max} \) respectively. Similar to \( H_2 \) problem in section 3.4, we can convert the two-sided problem into a one-sided problem. The \( H_\infty \) optimal norm of a one-sided problem was solved in chapter 4; and after normalizing by the optimal norm, all the solutions should be in the unit ball and can be organized by the method provided in this chapter. Generally speaking, there are many solutions for the \( H_\infty \) problem -- unlike the \( H_2 \) problem, which has a unique solution. When m or r is 1, the \( H_\infty \) optimal problem has a unique solution.
Theorem 5.4

The maximal singular values of any optimal solutions of two-sided problem (5.112) are all-pass on all frequencies when $U$ and $V$ are square rational matrices of full rank on the imaginary axis and at infinity.

(Proof)

We only need to prove

$$\bar{\sigma}[\hat{\Psi}(j\omega)] = 1, \quad \text{for all real } \omega$$  \hspace{1cm} (5.113)

for the normalized problem. Recall that for fixed $\omega$ we have

$$\bar{\sigma}^2[\hat{\Psi}(j\omega)] = \sup \frac{(\hat{\Psi}d)^* \hat{\Psi}d}{d(j\omega)}$$  \hspace{1cm} (5.114)

Since all optimal solutions are given in the form of (5.110), we can take a particular $\hat{d}$ as the last column of $L_{02}$, then $\hat{\Psi}$ becomes the last column of $L_{01}$. Notice

\[
\begin{bmatrix}
    L_{01} \\
    L_{02}
\end{bmatrix}
\]

is in the neutral shift invariant space of Krein space, therefore we have

\[
(\hat{\Psi}d)^* (\hat{\Psi}d) = \hat{d}^* \hat{d} \hspace{1cm} (5.115)
\]

Put (5.115) into (5.114), we have

$$\bar{\sigma}[\hat{\Psi}(j\omega)] \geq 1 \quad \text{for all real } \omega$$  \hspace{1cm} (5.116)

On the other hand, for this normalized problem the maximal singular value on $j\omega$ cannot be greater than 1. So we have (5.113).
Example 5.1

As an example, to solve (5.112) when

\[
W = \begin{bmatrix}
\frac{20(2s-1)}{(s+1)(s+10)} & -\frac{10}{(s+10)} \\
\end{bmatrix}
\]

\[
U = \frac{s-1}{s+3}
\]

\[
V = \begin{bmatrix}
0 & -1 \\
\frac{s-2}{2(s+1)} & 3 \\
\end{bmatrix}
\]

Just as in section 3.4 it can be transformed into a one-sided problem with

\[
W_1 = \begin{bmatrix}
\frac{20(2s-1)}{(s+1)(s+10)} & -\frac{10(s-2)}{(s+2)(s+10)} \\
\end{bmatrix}
\]

\[
U_1 = \frac{(s-1)(s-2)}{(s+3)(s+2)}
\]

The right half plane zeros are \( s_1 = 1, s_2 = 2 \). Taking \( a_1 = a_2 = 1 \) then

\[
\beta_1^* = \begin{bmatrix}
\frac{10}{11} & -\frac{10}{33} \\
\end{bmatrix}
\]

\[
\beta_2^* = \begin{bmatrix}
\frac{5}{3} & 0 \\
\end{bmatrix}
\]

solve

\[
\det ( R - \lambda Q ) = 0
\]

we get
\[ \lambda = 8.480644796 \]

The optimal \( H_\infty \) norm is

\[ U_0 = 2.912154666 \]

To find the optimal \( \hat{f} \), we normalize the problem into the unit ball by dividing by \( U_0 \). For simplicity we still use the same notation for the normalized problem. Now

\[ a_1 = a_2 = 1 \]

\[ \beta_1^* = (0.312171231, -0.104057077) \]

\[ \beta_2^* = (0.572313924, 0) \]

\[ b_1 = \frac{1}{s+1} \begin{bmatrix} 0.312171231 \\ -0.104057077 \end{bmatrix}, \quad b_2 = \frac{1}{s+2} \begin{bmatrix} 0.572313924 \\ 0 \end{bmatrix} \]

\[ P = \begin{bmatrix} 0.445860623 & 0.273780019 \\ 0.273780019 & 0.168114193 \end{bmatrix} = U_0^* \Gamma U_0^* \]

\[ U_0 = \begin{bmatrix} 0.852166151 & -0.523271294 \\ 0.523271294 & 0.852166151 \end{bmatrix} \]

\[ \Gamma = \begin{bmatrix} 0.613974816 & 0 \\ 0 & 0 \end{bmatrix} \]

Since \( m = r_0 = 1 \), \( V^L_1 \) does not exist. (5.52) becomes

\[ (0.324356309, 0.054450081) \cdot V^L_2 = 0 \]

solving it, we have
\[
V_2^L = \begin{bmatrix}
0.165554674 \\
-0.98640061
\end{bmatrix}
\]

Since \( V_1^+ \) does not exist therefore there is no \( L_+ \), from (5.53)

\[
L_- = \frac{1}{(s+1)(s+2)} \begin{bmatrix}
0.405638466s + 0.656955872 \\
0.165554674s^2 + 0.663438177s + 0.576337567 \\
-0.98620061s^2 - 2.984753184s - 2.024703929
\end{bmatrix}
\]

From (5.54), we have

\[
L_0 = \frac{1}{(s+1)(s+2)} \begin{bmatrix}
0.32894857s - 0.194376437 \\
0.324356309s + 0.161006065 \\
0.054450081s + 0.108900162
\end{bmatrix}
\]

Divide

\[
\begin{bmatrix}
L_{12} & L_{01} \\
L_{22} & L_{02}
\end{bmatrix}
\]

then

\[
\hat{H} = [L_{12}, L_{01}][L_{22}, L_{02}]^{-1}
\]

\[
= \begin{bmatrix}
0.986200167(s-0.496386418) & 0.165554673(s^2) \\
(0.590998717) & (s+0.590998717)
\end{bmatrix}
\]

The optimal \( \hat{H} \) is

\[
H = \begin{bmatrix}
(-2.871967s+1.218059)(s+2)(s+3) & (-0.48212081s+3.7342919)(s+3) \\
(s+1)(s+10)(s+0.590998717) & (s+10)(s+0.590998717)
\end{bmatrix}
\]

An explicit expression for all solutions in the unit ball and satisfying interpolation conditions is given in this chapter. Just as Francis, Helton and Zames, the geometrical description of Krein space is used for the consideration, but they haven't got the explicit expression.
The degenerate case can be treated by a Cholesky decomposition and by separating the singular part of Pick matrix from the regular part.

In the next chapter we will convert this method into a recursive process.
CHAPTER 6 RECURSIVE ALGORITHM

6.1 Introduction

A complete solution set is given in the last chapter which heavily depends on the computation of the inverse or Cholesky decomposition of matrix P. For a large dimension matrix P these computations are a heavy burden. To avoid large matrix computation a recursive algorithm is discussed in this chapter. This algorithm is a generalization of Nevanlinna and Schur algorithm.

We have found from chapter 5 that \( L(X) \) maps unit ball to unit ball if and only if \( L \) is \( J \)-unitary. Let \( P_i \) be a projection operator with the property \( P_i^J = JP_i \), then

\[
L = I - P_i + \frac{1}{s - i^*} P_i
\]

will be \( J \)-unitary for \( s \) on the imaginary axis. In this chapter we will extend this basic result to repeated zeros case.

6.2 Interpolation Pair and Recursive Method

Suppose

\[
\hat{\Psi}(s) = W(s) + U(s)H(s) \tag{6.1}
\]

has been normalized by dividing both sides by optimal norm \( \| \cdot \|_0 \). Our objective is to describe a recursive algorithm for obtaining a \( \hat{\Psi}(s) \) of unity norm satisfying the interpolation conditions (3.38) and (3.42).
For simplicity of expression we use $\xi$ and $\eta$ instead of $\alpha$ and $\beta$, $\xi(s) = \alpha^*(s)$, $\eta(s) = -\beta^*(s)$ and suppose

$$\eta(s) = -\xi(s) \gamma(s) = -\xi(s) W(s) \quad (6.2)$$

where """ means that all derivatives at $s_i$ ($i = 1, \ldots, n$) up to $(i-1)$th are equal on both sides. This is a restatement of the interpolation conditions (3.38) and (3.42).

We will give $W$, $\xi$, $\eta$ two subscripts: $i,j$ and two superscripts $i,t$. By $i$, we mean that we expand them at the point $s_i$, $j$ represents the $j$th derivative. By $j$ we mean that we interpolate at $s_j$, $t$ means the $t$th interpolation operation at point $s_j$.

We will use $\xi(s)$ and $\eta(s)$ to transmit the interpolation conditions in the recursive process. Therefore $\xi(s)$ and $\eta(s)$ are called the interpolation pair. Suppose the recursive process is at the $t$th interpolation operation at point $s_j$. $(6.2)$ becomes

$$\eta^t_j(s) = -\xi^t_j(s) W^t_j(s) \quad (6.3)$$

Expanding $\xi^t_j(s)$, $\eta^t_j(s)$ and $W^t_j(s)$ at $s_i$, we have

$$\xi^t_j(s) = \xi_{1j} + \frac{1}{1!} \xi_{2j} \cdot (s-s_i) + \frac{1}{2!} \xi_{3j} \cdot (s-s_i)^2 + \ldots \quad (6.4)$$

$$\eta^t_j(s) = \eta_{1j} + \frac{1}{1!} \eta_{2j} \cdot (s-s_i) + \frac{1}{2!} \eta_{3j} \cdot (s-s_i)^2 + \ldots \quad (6.5)$$

$$W^t_j(s) = W_{1j} + \frac{1}{1!} W_{2j} \cdot (s-s_i) + \frac{1}{2!} W_{3j} \cdot (s-s_i)^2 + \ldots \quad (6.6)$$

Since $(6.3)$ is correct for $0$th to $(i-1)$th coefficients at $s_i$ ($i = 1, 2, \ldots, n$), we have
\[
\begin{bmatrix}
0_t & 1_t & \ldots & 1_{l_i-1,t} \\
\eta_{ij} & \ldots & \ldots & \eta_{ij}
\end{bmatrix}
= \\
\begin{bmatrix}
\begin{bmatrix}
0_{0t} & \ldots & 1_{0t} \\
0_{0t} & \ldots & \ldots & 1_{0t}
\end{bmatrix} & \ldots & \begin{bmatrix}
\ldots & 1_{i-2,t} \\
\ldots & \ldots & 1_{i-1,t}
\end{bmatrix} \\
\begin{bmatrix}
\ldots & 1_{i-2,t} \\
\ldots & \ldots & \ldots & 1_{i-2,t}
\end{bmatrix} & \ldots & \begin{bmatrix}
\ldots & 1_{i-1,t} \\
\ldots & \ldots & \ldots & 1_{i-1,t}
\end{bmatrix}
\end{bmatrix}
\]

(6.3) can also be written as

\[
\begin{bmatrix}
\xi_j^t(s) \\
\eta_j^t(s)
\end{bmatrix}
= \\
\begin{bmatrix}
0_t \\
I
\end{bmatrix}
\]

(6.8)

Let

\[
L_j^t(s) = \begin{bmatrix}
A_j^t(s) & B_j^t(s) \\
C_j^t(s) & D_j^t(s)
\end{bmatrix}
\]

(6.9)

and

\[
W_j^t(s) = L_j^t(s) W_j^{t+1}(s)
\]

(6.10)

then (6.8) becomes
\[ [\xi_j^t(s), \eta_j^t(s)] \cdot L_j^t(s) \cdot \begin{bmatrix} W_j^{t+1}(s) \\ I \end{bmatrix} = 0. \] (6.11)

Denote

\[ [\xi_j^{t+1}(s), \eta_j^{t+1}(s)] = [\xi_j^t(s), \eta_j^t(s)] \cdot L_j^t(s) \] (6.12)

(6.11) becomes

\[ [\xi_j^{t+1}(s), \eta_j^{t+1}(s)] \cdot \begin{bmatrix} W_j^{t+1}(s) \\ I \end{bmatrix} = 0 \] (6.13)

This means that the form of (6.10) is valid for the whole recursive process. For simplicity we will write

\[ \gamma(s) = [\xi(s), \eta(s)] \] (6.14)

Then (6.12) can be rewritten as

\[ \gamma_j^{t+1}(s) = \gamma_j^t(s) \cdot L_j^t(s) \] (6.15)

Expanding \( L_j^t(s) \) at \( s_i \),

\[ L_j^t(s) = L_{ij}^0 + \frac{1}{1!} L_{ij}^1 (s-s_i) + \frac{1}{2!} L_{ij}^2 (s-s_i)^2 + \ldots \] (6.16)

and expanding \( \gamma_j^t(s) \) and \( \gamma_j^{t+1}(s) \) at \( s_i \) in the similar way, we have

\[
\begin{bmatrix}
0, t+1 & 1, t+1 & \cdots & i_i-1, t+1 \\
\gamma_{ij}^t, \gamma_{ij}^t, \cdots, \gamma_{ij}^t
\end{bmatrix} = \]
We call (6.10) the $t$ th interpolation operation at $s_j$ when

$$L_j^t(s) = (I - p_j^t) + \frac{s-s_j}{s+s_j} p_j^t$$  (6.18)

where

$$p_j^t = \begin{bmatrix}
\hat{\xi}^t \\
\hat{\xi}^t (\hat{\eta}^t)^{-1} \\
-\hat{\eta}^t (\hat{\xi}^t)^{-1}
\end{bmatrix}$$  (6.19)

To explain why (6.10) is an interpolation operation, let us consider the interpolation process at $s_j$ from $t=0$ to $t= i_j - 1$.

In (6.18), at $s = s_j$, we have

$$L_j^{00} = L_j^0(s_j) = I - p_j^0$$

Notice

$$[\xi_{jj}^{00}, \eta_{jj}^{00}] p_j^0 = [\xi_{jj}^{00}, \eta_{jj}^{00}]$$

or
\[
\gamma_{jj}^0 p_j^0 = \gamma_{jj}^0
\]
\[(6.20)\]

Consider the first block in (6.17) at \(t=0, i=j\). We have

\[
\begin{align*}
\gamma_{jj}^0 &= \gamma_{jj}^0 L_j^0 \\
&= \gamma_{jj}^0 (I - p_j^0) \\
&= 0
\end{align*}
\]
\[(6.21)\]

This is \(\xi_{jj}^0 = \eta_{jj}^0 = 0\). So (6.7) is true for the first block at \(t=1\) and \(i=j\), no matter what \(W_{jj}^{01}\) is. Go back to \(t=0\) and \(i=j\), the first block in (6.7) also holds which means that

\[
W_{jj}^0(s) = L_j^0(s) [W_{jj}^1(s)]
\]

satisfying the 0th interpolation condition of (6.7) no matter what \(W_{jj}^1(s)\) is.

Now consider interpolation for \(W_{jj}^1(s)\). In (6.17) let \(t=1, i=j\). Consider the first two blocks:

\[
\begin{bmatrix}
\gamma_{jj}^{02} & \gamma_{jj}^{12}
\end{bmatrix} = \begin{bmatrix}
\gamma_{jj}^{01} & \gamma_{jj}^{11}
\end{bmatrix} \cdot \begin{bmatrix}
C_{0j}^0 & C_{1j}^1
\end{bmatrix}
\]
\[(6.22)\]

Notice \(\gamma_{jj}^{01} = 0\) [from (6.21)], and

\[
\gamma_{jj}^{11} p_j^0 = \gamma_{jj}^{11} (I - p_j^0) = 0
\]

Therefore

\[
\begin{bmatrix}
\gamma_{jj}^{02} & \gamma_{jj}^{12}
\end{bmatrix} = \begin{bmatrix}
0 & 0
\end{bmatrix}
\]
\[(6.23)\]

so (6.7) is true for the first two blocks at \(t=2\) and \(i=j\), no matter what \(W_{jj}^{02}\) and \(W_{jj}^{12}\) are. Go back to \(t=0\) and \(i=j\); the first two blocks in
(6.7) also hold which means that

\[ W_j^0(s) = L_j^0(s) \{ L_j^1(s)(W_j^2(s)) \} \]

satisfying the 0th and 1st interpolation condition of (6.7) no matter what \( W_j^2(s) \) is. The process goes on until all blocks in (6.17) and (6.7) are satisfied at \( i=j \).

\[ W_j^0(s) = L_j^0(s) \{ L_j^1(s) \{ L_j^{j-1} \{ \cdots \{ L_j^1(s) \{ W_j^2(s) \} \} \} \} \} \]

will satisfy all the interpolation condition at \( s_j \) no matter what \( W_j^2(s) \) is.

Denote

\[ W_{j+1}^0(s) = W_j^j(s) \]

Just as at \( s_j \), we can do the same procedure at \( s_{j+1} \) and so on, until all \( s_i \) (\( i = 1, \ldots, n \)) are considered.

One important thing should be noted. (6.17) is not only used to calculate \( \gamma_{jj}^{1, t+1} \), but also to calculate all \( \gamma_{ij}^{i, t+1} \) (\( i > j \)) for later use.

**Recursive Algorithm**

The recursive algorithm is started from \( j=1, t=0 \) by putting

\[ \xi_{i1}^{10} = \alpha_i^* \] \[ \lambda = 1, \ldots, \lambda_i = 1, \ldots, n \] \[ (6.24) \]

and calculating all \( \eta_{i1}^{10} \) from (6.7).
From \( j=1 \) to \( j=n \), and at each \( j \) from \( t=0 \) to \( t=j-1 \) step by step, according to (6.19), (6.18), calculate \( L_j^t(s) \) from \( \gamma_{jj}^{tt} \) and calculate \( \gamma_{jj}^{t+1,t+1} \) from (6.17) for the next step. At the same time, calculate all \( \gamma_{ij}^{t+1} \) from (6.17) for later use.

Multiply the \( L_j^t(s) \) \((j=1, \ldots, n, t=0, \ldots, j-1)\) and denote the product as

\[
L(s) = \prod_{t=0}^{j-1} \prod_{j=1}^{n} L_j^t(s)
\]

(6.25)

\( L(s) \) is a \((m+r)\times(m+r)\) matrix which can be partitioned into four blocks,

\[
L(s) = \begin{bmatrix}
A(s) & B(s) \\
C(s) & D(s)
\end{bmatrix}
\]

(6.26)

where \( A(s) \), \( B(s) \), \( C(s) \) and \( D(s) \) are \( mxm \), \( mxr \), \( rxm \), \( rxr \) matrices.

Then

\[
\Psi(s) = L(s)X(s)
\]

\[
= [A(s)X(s) + B(s)] \cdot [C(s)X(s) + D(s)]^{-1}
\]

(6.27)

will satisfy all the interpolation conditions for any \( X(s) \) in \( BH_m^{mxr} \).

### 6.3 Treatment of Degeneration

Since we have normalized \( \eta(s) \), for some step of interpolation process at \( j,t \), we will prove later the following case must occur,

\[
\det \left[ \begin{array}{cc}
\xi_{jj}^{tt} & \xi_{jj}^{tt}\ast \\
\eta_{jj}^{tt} & \eta_{jj}^{tt}\ast
\end{array} \right] = 0
\]

(6.28)
In this case the inverse in (6.19) does not exist. This case is called degeneration.

To deal with this case, come back to (6.7). Let \( i = j \), it follows from the interpolation process that

\[
\xi_{jj}^{tt} = 0 \quad \text{for} \quad i = 0, 1, \ldots, t-1
\]

we have

\[
\eta_{jj}^{tt} = -\xi_{jj}^{tt} w_{jj}^{tt}
\]

(6.29)

(6.28) is

\[
\det \left[ \xi_{jj}^{tt} (I - w_{jj}^{tt} w_{jj}^{tt*}) \xi_{jj}^{tt*} \right] = 0
\]

write

\[
I - w_{jj}^{tt} w_{jj}^{tt*} = YY^* \]

then

\[
\det \left[ \xi_{jj}^{tt} Y(\xi_{jj}^{tt} Y)^* \right] = 0
\]

which means that some linear combination of rows of \( \xi_{jj}^{tt} \) is zero. Therefore \( Y \) has zero eigenvalues and

\[
YY^* = V_1 \cdot \text{diag} \left\{ 1 - \lambda_1, \cdots, 1 - \lambda_{n-r_o}, 0_{r_o} \right\} \cdot V_1^*
\]

(6.31)

where \( V_1 \) is a unitary matrix. In this case we can find another unitary matrix \( V_2 \) such that

\[
w_{jj}^{tt} = V_1 \cdot \text{diag} \left\{ \sqrt{\lambda_1}, \cdots, \sqrt{\lambda_{n-r_o}}, I_{r_o} \right\} \cdot V_2^*
\]

(6.32)
Let

\[ W_{jj}^{tt} = V_1^* W_{jj}^{tt} V_2 \]  

(6.33)

From (6.32) we have

\[ W_{jj}^{tt} \begin{bmatrix} W_{tt}^{tt} \\ I_r \end{bmatrix} = \begin{bmatrix} W_{tt}^{jj} \\ 0 \end{bmatrix} \begin{bmatrix} I_r \\ 0 \end{bmatrix} \]  

(6.34)

Using \( V_1 \) and \( V_2 \), take the same transformation on whole function matrix \( W_{jj}^t(s) \),

\[ W_{jj}^t(s) = V_1^* W_{jj}^t(s) V_2 \]  

(6.35)

Since \( W_{jj}^t(s) \) has the maximal norm 1, and \( W_{jj}^{tt} \) has the form (6.34), \( W_{ij}^t \) must have the same form for \( i > j \), all \( i \), and \( i = j, \ l \geq t \),

\[ W_{ij}^t = \begin{bmatrix} I_r \\ 0 \end{bmatrix} W_{jj}^t \begin{bmatrix} I_r \\ 0 \end{bmatrix} \]  

(6.36)

After discarding the degenerate part, we can use

\[ \tilde{\eta}^t_j(s) = - \tilde{\xi}^t_j(s) \tilde{W}_{jj}^t(s) \]  

(6.37)

for the further interpolation process. Since the inverse of \( \tilde{\eta}^{tt}_{jj} \) and \( \tilde{\eta}^{*tt}_{jj} \) now exists, one gets

\[ \tilde{L}^t_j(s) \ldots \tilde{L}^t_n(s)(\tilde{X}) = \sqrt{\tilde{L}^t_j(s)}(\tilde{X}) \]  

(6.38)
\[ (m-r_0)x(r-r_0) \]

where \( \tilde{x}(s) \) is in \( \mathcal{B} \).

Denote the product of all previous interpolation matrices as \( L_j^{t-1} \),

\[ L_j^{t-1} = L_1^{t-1} \ldots L_j^{t-1} \tag{6.39} \]

In our degenerate case, (6.27) should be modified as follows

\[ \Phi(s) = L_j^{t-1} \quad \begin{bmatrix} \text{1} \\ \text{0} \end{bmatrix} \quad V_1 \quad \begin{bmatrix} \text{1} \\ \text{0} \\ \text{I} \end{bmatrix} \quad V_2 \]

\[ = \left[ A \tilde{x} + B \right] \cdot \left[ C \tilde{x} + D \right]^{-1} \tag{6.40} \]

where

\[ \begin{bmatrix} [A \ B] \\ [C \ D] \end{bmatrix} = L_j^{t-1} \cdot \begin{bmatrix} \text{1} \\ \text{0} \end{bmatrix} \quad \tilde{x}^t \quad L_j^{-1} \]

\[ = \begin{bmatrix} [E] \\ [F] \end{bmatrix} = L_j^{t-1} \cdot \begin{bmatrix} \text{1} \\ \text{0} \end{bmatrix} \quad \begin{bmatrix} \text{1} \\ \text{0} \end{bmatrix} \tag{6.41} \]

and

\[ V_1^L \] contains the \( i^{th} \) to \( (m-r_o)^{th} \) columns of \( V_1 \),

\[ V_1^R \] contains the \( (m-r_o+1)^{th} \) to \( m^{th} \) columns of \( V_1 \),

\[ V_2^L \] contains the \( i^{th} \) to \( (r-r_o)^{th} \) columns of \( V_2 \),
\( V_2^R \) contains the \((r-r_o+1)^{th}\) to \(r^{th}\) columns of \( V_2 \).

This is from

\[
\begin{bmatrix}
\hat{\xi} \\
[\hat{A}_j^t \hat{X} + \hat{B}_j^t \ V_1 \ 0 \ I] V_2^* + \hat{E}_j^t-1 \\
\end{bmatrix}
= \begin{bmatrix}
\hat{A}_j^t \hat{X} + \hat{B}_j^t & 0 \\
[\hat{A}_j^t \hat{X} + \hat{B}_j^t & 0] \\
\end{bmatrix}
\begin{bmatrix}
\hat{C}_j^t \hat{X} + \hat{D}_j^t \ V_1 \ 0 \ I \\
\end{bmatrix}
+ \begin{bmatrix}
\hat{C}_j^t \hat{X} + \hat{D}_j^t & 0 \\
[\hat{C}_j^t \hat{X} + \hat{D}_j^t & 0] \\
\end{bmatrix}
\begin{bmatrix}
\hat{E}_j^t-1 \ V_1 \ 0 \ I \\
\end{bmatrix}
+ \begin{bmatrix}
\hat{E}_j^t-1 \ V_1 \ 0 \ I \\
\end{bmatrix}
\]

\[ 6.4 \textbf{ Interpolation on Imaginary Axis or at Infinity } \]

Until now the interpolation on imaginary axis or at infinity is deliberately avoided, but it is of great significance for control problems. For example in order to track a constant signal an internal model of \( \frac{1}{s} \) should be introduced into the loop which will cause an interpolation at \( s = 0 \). Another example is that the sensitivity function should be 1 at
infinity when plant is strictly proper and compensator is proper.

Suppose \( s_1 = j\omega_1 \), and we want to interpolate at this point:

\[
\beta_1^* = a_1^* \hat{\theta}(s_1)
\]

(6.43)

and keep \( \hat{\theta}(s) \) in \( BH^\infty_\omega \). Obviously there exist solutions for this problem iff

\[
a_1^* a_1 - \beta_1^* \beta_1 \geq 0
\]

(6.44)

Notice that the Blaschke factor

\[
\frac{s-s_1}{s+s_1}
\]

becomes constant 1 for all \( s \) when \( s_1 \) is on imaginary axis, therefore (6.18) also becomes constant identity matrix which now cannot be used for the interpolation operation. Recall Blaschke factor is an inner factor, and from operator theory [21] the singular inner factor for \( s_1 = j\omega_1 \) is

\[
\exp \left( \frac{1+s_1}{s-s_1} \right)
\]

(6.45)

So we can put it into the position of the Blaschke factor in (6.18), this new \( L(s) \) is a \( J \)-unitary matrix which maps \( BH^\infty_\omega \) to \( BH^\infty_\omega \). But it is actually useless since it only can be used to realize the interpolation condition (6.43) in a generalized (i.e. singular) sense. Besides, this \( L(s) \) is not a rational function and difficult to be realized by a practical compensator.
In the special case:

\[ a_1^*a_1 - \beta_1^*\beta_1 = 0, \]  
\[ L_1 = I_{m+r} - \frac{1+s_1s}{s-s_1} \begin{bmatrix} a_1^*a_1 & -a_1^*\beta_1 \\ \beta_1^*a_1 & -\beta_1^*\beta_1 \end{bmatrix} \]  

is J-unitary. It can be used for interpolation at \( s_1 \) in a perfect sense.

Suppose interpolation points are \( s_1, s_2, \ldots, s_n \); where \( s_1 \) is on imaginary axis or at infinity, other points are on the open right half plane. For this problem, we can first solve the normal interpolation problem at \( s_2, \ldots, s_n \) and get

\[ \hat{g}(s) = L(\tilde{X}) \]  

where

\[ L = \begin{bmatrix} A & B & E \\ C & D & F \end{bmatrix} \]  

\[ (m-r)x(r-r_0) \]  

and \( \tilde{X} \) is in \( \mathcal{B}H^\infty \).

To satisfy the interpolation condition (6.43) at \( s_1 \), we have

\[ \begin{bmatrix} a_1^* & -\beta_1^* \end{bmatrix} \begin{bmatrix} A & B & E \\ C & D & F \end{bmatrix} \begin{bmatrix} \tilde{X} & 0 \\ 0 & I_{r-r_0} \end{bmatrix} = 0 \]  

Define
\[
\begin{bmatrix}
\alpha_x^* & -\beta_x^*
\end{bmatrix} = \begin{bmatrix}
\alpha_1^* & -\beta_1^*
\end{bmatrix} \begin{bmatrix}
A(s_1) & B(s_1) \\
C(s_1) & D(s_1)
\end{bmatrix}
\]  \hspace{1cm} (6.51)

then for (6.43) iff
\[
\beta_x^* = \alpha_x^* \tilde{x}(s_1)
\]  \hspace{1cm} (6.52)

and
\[
\alpha_1^* E(s_1) = \beta_1^* F(s_1)
\]  \hspace{1cm} (6.53)

Generally speaking, (6.53) only can be satisfied by chance. Therefore, (6.44) cannot guarantee the solvability of our problem when the right half plane interpolation is in the degenerate case.

When the right half plane interpolation is in the nondegenerate case, \(E, F\) will not exist, so (6.53) will not exist. We can always find \(X\) in \(BH^{\text{mfr}}\) which satisfies (6.52) under the condition (6.44), since we have the following relation:
\[
\alpha_x^* \alpha_x - \beta_x^* \beta_x = \alpha_1^* \alpha_1 - \beta_1^* \beta_1 \geq 0
\]  \hspace{1cm} (6.54)

Therefore we conclude that in the case of right half plane interpolation problem being nondegenerate, the total interpolation problem becomes to find \(X\) in \(BH^{\text{mfr}}\) to satisfy (6.52) which is always solvable when (6.44) exists.

Especially, when \(s_1\) is at infinity, from (5.13) we have
\[
L(\infty) = I_{m^{\text{fr}}}
\]

Therefore
\[ \Phi(\omega) = X(\omega) \]

In this case the interpolation condition on \( X(\omega) \) can be found without any calculation.

In the case of right half plane interpolation being degenerate, the requirement \( 6.53 \) is hardly to be satisfied. In this case we can use Zames-Francis sequence method [6] to overcome the difficulty.

For example, in many control problems plants are strictly proper, so in the corresponding interpolation problem

\[ \Phi = W - UH, \quad (6.55) \]

\( U \) has a blocking zero at infinity. For simplicity we suppose that \( U \) is square, and

\[ U = \frac{1}{s+1} U_1(s) \]

where \( U_1(s) \) has only right half plane zeros, \( \det U_1(\omega) \neq 0 \), and \( U_1(\omega) \) is finite.

Suppose after solving the interpolation problem for \( U_1 \) we have

\[ \Phi(s) = L(\widetilde{X}) \quad (6.56) \]

which is in the degenerate case.

Since \( U(\omega) = 0 \) we need consider the interpolation problem at infinity:

\[ \Phi(\omega) = W(\omega) = L(\widetilde{X}(\omega)) \quad (6.57) \]
We already mentioned that it is difficult to meet the requirement in degenerate case. In this case we can first drop the interpolation condition at infinity, put (6.56) into (6.55) and find

$$R = (s+1)U_{\frac{1}{4}}^{-1} \cdot (W-L[X])$$  (6.58)

In this case, the "solution" $R$ must be improper, and it is not in $H_\infty$. To find proper solution, choose

$$\hat{R} = \frac{n}{s+n} H$$  (6.59)

as our solution. Put $\hat{R}$ into (6.55) we have

$$\hat{\phi} = \frac{s}{s+n} W + \frac{n}{s+n} L[\bar{X}]$$  (6.60)

which satisfies the interpolation condition at infinity:

$$\hat{\phi}(\infty) = W(\infty)$$  (6.61)

we also can find, $\hat{\phi}$ satisfies all other interpolation conditions.

When $W$ is strictly proper ($W(\infty) = 0$), we can prove that $\hat{\phi}$ is in $BH^\infty_{\infty}$ as $n \to \infty$.

When $W(s)$ is not strictly proper, under the condition

$$\sigma(W(\infty)) \leq 1,$$

Zames and Francis proposed a doubly indexed sequence to get a proper solution sequence in the SISO case.

In our experience the singly indexed sequence (6.60) can also be used for practical system design with a little increase of $H_\infty$ norm in the case of $W(\infty) \neq 0$. This parameter $n$ is also of benefit to designer for other
requirements of specifications.

Another similar method is to consider suboptimal solutions instead of the optimal one. In this case Pick matrix is always positive definite, so the problem now comes back to the nondegenerate case. We have pointed that in the nondegenerate case, there are enough degrees of freedom for \( X \) to meet the interpolation conditions on imaginary axis or at infinity when (6.44) exists.

In this chapter a recursive method is introduced, which is numerically better than the normal Nevanlinna algorithm which must calculate square root of matrix for Halmos matrix [11]. This recursive method is also extended to repeated zeros case. The interpolation problem on imaginary axis and at infinity is also discussed in this chapter.

In the next chapter we will discuss a more general problem -- the nonsquare problem (\( U \) and \( V \) are not square).
CHAPTER 7 NONSQUARE PROBLEM

7.1 Introduction

In most control problems we meet a general problem as follows

\[
\min_{\mathbf{H} \in \mathbb{H}_m^{\times r}} \| \mathbf{W} - \mathbf{UH} \mathbf{V} \|_\infty
\]

(7.1)

where \( \mathbf{W}, \mathbf{U}, \mathbf{V} \in \mathbb{H}_m^{\times r'}, \mathbb{H}_m^{\times m}, \mathbb{H}_r^{\times r'} \) respectively, and \( m' > m, r < r' \). We call it the nonsquare problem.

Many authors attacked this problem by different methods. Kwakernaak [41] proposed a method which seems to have completely solved the one-sided nonsquare problem in the one dimensional case \( (m=r=1) \). But it is actually difficult to be used for practical problems since a very complicated equation is introduced even in the one dimensional case. A set of complicated associated equations which cannot be expected to be solved by routine computer program would be introduced when it is extended to the multidimensional case.

Francis [56] first proposed an iteration process to solve the nonsquare problem by reforming it into a square problem. Recently, Doyle [40] proposed a nice method to find the inner-outer factorization of a nonsquare rational matrix by its state space expression and gave a regular procedure for the iteration method.

In this chapter, we will also apply the iteration method to solve the nonsquare problem. Two different spectral factorization methods for inner-outer factorization of a nonsquare matrix are introduced. The first
one uses the transfer function matrix in s-domain directly and is simpler computationally than Doyle's method which needs transformation into state space as a preparation. The second one is attractive from its simplicity and efficiency, but has a shortcoming or only being effective for spectral factorization of inside-outside of the unit disc (so, effective in z-domain).

We will also derive two simple formulas for the estimation of optimal norm from the result at each step. An example shows that high precision can be reached in a few steps of iteration with the help of these two formulas.

One thing that should be emphasized is that the maximal singular value of the optimal solution need not to be constant on all frequencies for the nonsquare problem, which is completely different from the square case. A simple example can be used to expose the new property. For example,

$$\min_{H \in H_\infty} \| \begin{bmatrix} H \\ \frac{1}{s+1} \end{bmatrix} \|_{\infty}$$

(7.2)

obviously

$$H = 0$$

is a solution, and

$$\| \begin{bmatrix} 0 \\ \frac{1}{s+1} \end{bmatrix} \|_{\infty} = 1$$

but now
\[
\begin{bmatrix}
0 \\
\frac{1}{s+1}
\end{bmatrix}_j \omega = \frac{1}{\omega^2 + 1}
\]

is not constant for all frequencies.

Another interesting thing is that the optimal solution is not unique even in \( m = r = 1 \) case, for example,

\[
H = \frac{ss}{s+1}, \quad 0 \leq s \leq 1
\]

they all cause

\[
\| \begin{bmatrix}
H \\
\frac{1}{s+1}
\end{bmatrix}\|_{\infty} = 1
\]

and they all are the optimal solutions of (7.2). Moreover,

\[
\begin{bmatrix}
\frac{ss}{s+1} \\
\frac{1}{s+1}
\end{bmatrix}_j \omega = \sqrt{\frac{2}{\omega^2 + 1}}
\]

Let

\[\varepsilon = 1\]

the maximal singular value in this case is constant on all frequencies and which is also the optimal norm. Therefore we conclude that Kwakernaak's method can only be used to find some special solution for the nonsquare problem, since in his method the optimal solution is searched within the set of function matrices whose maximal singular value is constant.
A new phenomenon accompanying the maximal singular value of the optimal solution not being constant is the existence of a critical point which we will discuss in more detail in section 7.4 (see case B(ii) and example 7.2).

To deal with the nonsquare problem we first need to extend the concept of inner-outer factorization to a nonsquare matrix:

\[
U = \begin{bmatrix}
M_1 \\
0
\end{bmatrix}
\]  
\[V = \begin{bmatrix}
M_2 & 0
\end{bmatrix}
\]  
\[
\Theta \ast \Theta = I_{m' \times m'}
\]  
\[
\Phi \ast \Phi = I_{r' \times r'}
\]

where \( \Theta \) is in \( H_{\infty}^{m' \times m'} \), \( \Phi \) is in \( H_{\infty}^{r' \times r'} \), \( M_1 \) is in \( H_{\infty}^{m \times m} \), and outer. \( M_2 \) is in \( H_{\infty}^{r \times r} \), and outer.

We only show how to get (7.3). Rational matrix \( U(s) \) can be written as a ratio of two polynomial matrices \( N(s) \) and \( D(s) \):

\[
U(s) = N(s) D(s)^{-1}
\]  

Suppose

\[
N(s) = N_0 + N_1 s + \cdots + N_n s^n
\]

where \( N_i \in C^{m' \times m} \), \( i = 0, 1, \ldots, n \). Then we can find \( Q_i \in C^{(m' - m) \times m'} \)
(i = 0, ..., t) such that

\[
\begin{bmatrix}
N_0 & N_1 & \ldots & N_n \\
N_0 & N_1 & \ldots & N_n \\
\vdots & \vdots & \ddots & \vdots \\
N_0 & N_1 & \ldots & N_n \\
\end{bmatrix}
\begin{bmatrix}
Q_0 \\
Q_1 \\
\vdots \\
Q_t \\
\end{bmatrix} = 0.
\tag{7.9}
\]

We can form a \(m' \times (m'-m)\) polynomial matrix as follows:

\[
Q(s)^* = Q_0^T + Q_1^T(-s) + \cdots + Q_t^T(-s)^t
\tag{7.10}
\]

From (7.9) we have

\[
Q(s)^*N(s) = 0.
\tag{7.11}
\]

Denote

\[
\Omega(s) = N(s)^*N(s)
\]

\[
= a_0 + a_1 s + \cdots + a_{2n} s^{2n},
\tag{7.12}
\]

if we can find the following spectral factorization:

\[
\Omega(s) = Z(s)^*Z(s)
\tag{7.13}
\]

where \(Z(s)\) is a \(m \times m\) polynomial matrix and only contains left half plane zeros, then

\[
\Theta_1 = N(s)Z(s)^{-1}
\tag{7.14}
\]

is in \(H_{m' \times m}^\infty\), and

\[
\Theta_1^* \Theta_1 = I_{m \times m},
\tag{7.15}
\]

and
\[ M_1(s) = Z(s)D(s)^{-1} \]  

(7.16)

is minimal phase, in \( U_{m \times m}^{\infty} \). Now we have

\[ U = \Theta_1 \cdot M_1 \]  

(7.17)

Similarly, we can find

\[ Q(s) \cdot Q(s)^* = Z_1(s) \cdot Z_1(s)^* \]  

(7.18)

and \( Z_1(s) \) is a \((m'-m) \times (m'-m)\) polynomial matrix and only contains left half plane zeros, then

\[ \Theta_2 = Q(s)^* \cdot Z_1(s)^{-1} \]  

(7.19)

has all its zeros in the right half plane, is in \( U_{m' \times (m'-m)}^{\infty} \), and

\[ \Theta_2^* \Theta_2 = I_{(m'-m) \times (m'-m)} \]  

(7.20)

From (7.11), (7.15) and (7.20) we have

\[ \Theta = [ \Theta_1, \Theta_2 ] \]  

(7.21)

is an inner matrix:

\[ \Theta^* \Theta = I_{m' \times m'} \]  

(7.22)

and (7.3) exists.

7.2 Spectral Factorization

In this section we will discuss the spectral factorization of \( Q(s) \).

Denote \( N = 2n \), (7.12) becomes:
\[ \Omega(s) = a_0 + a_1 s + \cdots + a_N s^N. \quad (7.23) \]

Suppose \( N_n \) in (7.8) is full column rank, then \( a_N \) is full rank (\( a_N \) is square matrix). We also suppose that \( s_i \) is a \( \hat{i} \)th zero of \( \Omega(s) \), then we can find

\[ \xi(s) = \xi^0_i + \frac{1}{1!}\xi^{1}_i (s-s_i) + \frac{1}{2!}\xi^{2}_i (s-s_i)^2 + \cdots \quad (7.24) \]

such that

\[ \Omega(s) \xi(s) = 0(s-s_i)^{\hat{i}}_i. \quad (7.25) \]

Expanding

\[ \Omega(s) = \Omega^0_i + \frac{1}{1!}\Omega^{1}_i (s-s_i) + \frac{1}{2!}\Omega^{2}_i (s-s_i)^2 + \cdots \quad (7.26) \]

then (7.25) is

\[
\begin{bmatrix}
    c_i^0 \Omega_i^0 \\
    c_i^1 \Omega_i^0 \\
    \vdots \\
    c_i^{i-1} \Omega_i^0 \\
\end{bmatrix}
\begin{bmatrix}
    \xi_i^0 \\
    \xi_i^1 \\
    \vdots \\
    \xi_i^{i-1} \\
\end{bmatrix}
= 0
\quad (7.27)
\]

On the other hand, \( \Omega(s) \) is corresponding to a pencil

\[ \Omega_i(s) = A s + B \quad (7.28) \]

where
\[ A = \begin{bmatrix}
  a_N \\
  & I \\
  & \ddots \\
  & & & I
\end{bmatrix} \quad (7.29) \]

\[ B = \begin{bmatrix}
  a_{N-1} & a_{N-2} & \cdots & \cdots & a_0 \\
  -I & & & & \\
  & -I & & & \\
  & & \ddots & & \\
  & & & -I & 0
\end{bmatrix} \quad (7.30) \]

Define

\[ R(s) = \begin{bmatrix}
  s^{N-1} & I \\
  s^N & I \\
  \vdots & \\
  s & I
\end{bmatrix} \quad (7.31) \]

and define

\[ \eta(s) = R(s)\xi(s) \quad (7.32) \]

then

\[ \Omega_1(s)\eta(s) = \begin{bmatrix}
  \Omega(s)\xi(s) \\
  0 \\
  \vdots \\
  0
\end{bmatrix} = 0((s-s_i)^l_i) \quad (7.33) \]

Expanding \( R(s) \) and \( \eta(s) \) at \( s_i \):

\[ R(s) = R_i^0 + \frac{1}{1!} R_i^1(s-s_i) + \frac{1}{2!} R_i^2(s-s_i)^2 + \cdots \quad (7.34) \]
\[
\eta(s) = \eta_i^0 + \frac{1}{i!} \eta_i^1 (s-s_i) + \frac{1}{2!} \eta_i^2 (s-s_i)^2 + \cdots 
\]  
\hspace{1cm} (7.35)

Truncate (7.32) until \( \eta_i \) th order of \((s-s_i)\):

\[
\begin{bmatrix}
0 \\
\eta_i \\
1 \\
\vdots \\
\eta_{i-1} \\
\end{bmatrix} = 
\begin{bmatrix}
C_{R_i}^0 & 0 \\
C_{R_i}^1 & C_{R_i}^0 \\
\vdots & \vdots \\
C_{R_i}^{i-1} & C_{R_i}^{i-2} & \cdots & 0 \\
\end{bmatrix}
\begin{bmatrix}
\xi^0 \\
\xi^1 \\
\vdots \\
\xi_{i-1} \\
\end{bmatrix} 
\]  
\hspace{1cm} (7.36)

We also can truncate (7.33) as follows:

\[
\begin{bmatrix}
A \eta_i + B \\
A & A \eta_i + B \\
2A & A \eta_i + B \\
\vdots & \vdots \\
(\eta_i - 1)A & \eta_i \\
\end{bmatrix} 
\begin{bmatrix}
0 \\
\eta_i \\
1 \\
\vdots \\
\eta_{i-1} \\
\end{bmatrix} = 0. 
\]  
\hspace{1cm} (7.37)

We can rewrite (7.37) into following important formula:

\[
B \cdot \begin{bmatrix}
1 & 0 & 1 & 1 & \cdots & \frac{1}{(i_i-1)!} \\
0 & 1 & 1 & \cdots & \frac{1}{i_i-1} \\
\end{bmatrix} 
\]  
\[
= -A \cdot \begin{bmatrix}
1 & 0 & 1 & 1 & \cdots & \frac{1}{(i_i-1)!} \\
0 & 1 & 1 & \cdots & \frac{1}{i_i-1} \\
\end{bmatrix} \cdot J_i 
\]  
\hspace{1cm} (7.38)

where \( J_i \) is Jordan block:
PLEASE NOTE:

This page not included with original material. Filmed as received.

University Microfilms International
B = -ATJT^{-1} \tag{7.44}

Put (7.44) into (7.28), we have proved the fundamental result stated in this lemma.

Let us pick some eigenvalues, say \(\lambda_k, \ldots, \lambda_k\), and define
\[
\mathbf{R}(s) = \begin{bmatrix}
\lambda_k & \cdots & \lambda_k \\
\vdots & \ddots & \vdots \\
\lambda_k & \cdots & \lambda_k
\end{bmatrix}
\tag{7.45}
\]

\(\eta(s) = \mathbf{R}(s) \cdot \xi(s). \tag{7.46}\)

Truncate it until \((i_1 - 1)\) th order of \((s-s_i)\):
\[
\begin{bmatrix}
0 \\
C_i^0 \\
C_i^1 \\
\vdots \\
C_i^{i_1 - 1}
\end{bmatrix}
\begin{bmatrix}
C_i^0 \\
C_i^1 \\
C_i^2 \\
\vdots \\
C_i^{i_1 - 1}
\end{bmatrix}
\begin{bmatrix}
\xi_i \\
\xi_i \\
\xi_i \\
\vdots \\
\xi_i
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
0 \\
\vdots \\
0
\end{bmatrix}
\tag{7.47}
\]

Denote
\[
\mathbf{T} = \begin{bmatrix}
1 & \frac{1}{i_1 - 1} & \cdots & \frac{1}{i_1 - 1}
\end{bmatrix}
\tag{7.48}
\]

It is pointed out that \(\mathbf{T}\) is actually the low-right block \(T_{22}\) of \(T\) with compatible dimension:
\[
\begin{bmatrix}
T_{11} & T_{12} \\
T_{21} & T_{22}
\end{bmatrix}
\]

(7.49)

Denote

\[
I = \begin{bmatrix}
J_{k_1} \\
\vdots \\
J_{k}
\end{bmatrix}
\]

(7.50)

**Theorem 7.1**

Under the condition of \( I \) of full rank, define

\[
X_I(s) = T(sI-I)^{-1}
\]

(7.51)

then \( X_I(s) \) must be in the form

\[
X_I(s) = \begin{bmatrix}
I & \dotsc & I \\
I & \dotsc & I \\
\vdots & \ddots & \vdots \\
I & \dotsc & I
\end{bmatrix}
\begin{bmatrix}
x_{n-1} & x_{n-2} & \cdots & x_1 & x_0 \\
-I & & & & \\
& -I & & & \\
& & \ddots & & \\
& & & -I & 0
\end{bmatrix}
\]

(7.52)

Furthermore, let

\[
X(s) = x_0 + x_1s + \cdots + x_{n-1}s^{n-1} + Is^n
\]

(7.53)

then \( X(s) \) is with spectra of \( z_{k_1}, \dotsc, z_k \) to corresponding degree, and

\[
\Omega(s) = Y(s)X(s)
\]

(7.54)

for some polynomial matrix \( Y(s) \).
(Proof)

To prove (7.52) we only need check if

\[
\begin{bmatrix}
    x_{n-1} & x_{n-2} & \cdots & x_1 & x_0 \\
    -I \\
    \cdots \\
    -I & 0
\end{bmatrix}
\begin{bmatrix}
    I \ 
\end{bmatrix}
\begin{bmatrix}
    I \\
    -I \\
    \cdots \\
    -I & 0
\end{bmatrix}
\begin{bmatrix}
    I \\
\end{bmatrix}
\]

(7.55)

It is a long manipulation from (7.48), and is omitted.

Obviously \(X_I(s)\) and \(X(s)\) are with the spectra of \(s_{k_1}, \cdots, s_k\) to corresponding degree.

What is left to prove is (7.54). For this purpose let us consider a sequence

\[
\{ u_M, \ M = 0, 1, \cdots \}
\]

which satisfies

\[
x_0 u_M + x_1 u_{M+1} + \cdots + x_{n-1} u_{M+n-1} + u_{M+n} = 0
\]

(7.56)

It can be rewritten into the following form:

\[
\begin{bmatrix}
    u_{M+n} \\
    u_{M+n-1} \\
    \vdots \\
    u_{M+1} \\
\end{bmatrix}
\begin{bmatrix}
    x_{n-1} & x_{n-2} & \cdots & x_1 & x_0 \\
    -I \\
    \cdots \\
    -I & 0
\end{bmatrix}
\begin{bmatrix}
    u_{M+n-1} \\
    u_{M+n-2} \\
    \vdots \\
    u_M
\end{bmatrix}
\]

(7.57)

We have already mentioned that
\[
\begin{bmatrix}
x_{n-1} & x_{n-2} & \cdots & x_1 & x_0 \\
-1 & & & & \\
& -1 & & & \\
& & \ddots & & \\
& & & -1 & 0
\end{bmatrix} = -T J T^{-1}
\] (7.58)

therefore (7.57) becomes

\[
\begin{bmatrix}
u_{N+n} \\
u_{N+n-1} \\
 \vdots \\
u_{M+1} \\
u_{M}
\end{bmatrix} = T J T^{-1} \begin{bmatrix}
u_{N+n-1} \\
\vdots \\
u_{M} \\
u_0 
\end{bmatrix}
\] (7.59)

Repeatedly using (7.59) we get

\[
\begin{bmatrix}
u_{M+n-1} \\
u_{M+n-2} \\
 \vdots \\
u_{M} \\
u_0
\end{bmatrix} = T J^M T^{-1} \begin{bmatrix}
u_{N-1} \\
\vdots \\
u_0 
\end{bmatrix}
\] (7.60)

Denote

\[
\begin{bmatrix}
u_{N-1} \\
u_0
\end{bmatrix} = D
\]

notice (7.48), the last row of (7.60) is

\[
u_M = \begin{bmatrix}
\frac{1}{1! k_1} & \cdots & \frac{1}{i_1 k_1^{-1}} \\
\frac{1}{(i_1 - 1)! k_1} & \cdots & \frac{1}{i_k k_1^{-1}} \\
\frac{1}{(i_k - 1)! k_1} & \cdots & \frac{1}{i_k k_1^{-1}}
\end{bmatrix}J^M D
\] (7.61)

Put (7.61) into

\[
a_0 u_M + a_1 u_{N+1} + \cdots + a_N u_{N+N}
\]

and notice (7.27), after long manipulation we get a nice result:
\[ a_0 u_M + a_1 u_{M+1} + \cdots + a_N u_{M+N} = 0 \]  
(7.62)

which means that any sequence satisfying (7.56) must satisfy (7.62).

Now suppose

\[ \Omega(s) = Y(s) \cdot X(s) + Z(s) \]  
(7.63)

where \( Y(s) \) and \( Z(s) \) are polynomial matrices and \( Z(s) \) is a remainder,

\[ Z(s) = z_0 + z_1 s + \cdots + z_t s^t \]  
(7.64)

and \( z_t \neq 0, \quad 0 \leq t < n \).

Since the sequence \( \{ u_M \} \) satisfying (7.56) and (7.62), it must also satisfy following:

\[ z_0 u_M + z_1 u_{M+1} + \cdots + z_t u_{M+t} = 0 \]  
(7.65)

for all \( M \geq 0 \). Let \( M = 0 \), we have

\[ z_0 u_0 + z_1 u_1 + \cdots + z_t u_t = 0 \, . \]

From (7.60) \( u_0, \cdots, u_{n-1} \) can be chosen arbitrary, the only possibility is

\[ z_0 = z_1 = \cdots = z_t = 0 \]

(7.54) is thus proved.

Theorem 7.1 is a fundamental result for spectral factorization. Once the eigenvalues and eigenvectors of \( \Omega(s) \) are found, we can organize any polynomial factors of \( \Omega(s) \) by the recipe of theorem 7.1.
As an application of theorem 7.1 is an algorithm to find the left half plane and right half plane spectral factorization of \( \Omega(s) \). Suppose \( \Omega(s) \) has no imaginary axis zeros, and from the structure of (7.12), it's zeros are symmetric to imaginary axis and \( k \) must be even. Suppose \( s_{k+1}^{-}, \ldots, s_k^{-} \) are left half plane zeros of \( \Omega(s) \), and let \( n = N/2 \), then \( \Sigma \)
in (7.48) must be of full rank. Therefore we can apply theorem (7.1) and find \( x_0^{-}, \ldots, x_{n-1}^{-} \) from (7.51), and organize \( \chi(s) \) according to (7.54). Obviously \( \chi(s) \) contains and only contains all the left half plane zeros of \( \Omega(s) \).

One thing that should be mentioned is that \( \chi(s) \) can be determined by pure real computation when \( \Omega(s) \) has only real coefficients. In this case all eigenvalues and eigenvectors must be appear conjugate in pairs in (7.40) and (7.42). We can give a transformation

\[
R_i = \begin{bmatrix}
1 & 1 \\
2 & 2j
\end{bmatrix}
\]

(7.66)

to every conjugate pair of eigenvectors, and every Jordan block becomes

\[
J_{i,i+1}^{\chi} = R_i^{-1} \begin{bmatrix}
J_i & 0 \\
0 & J_{i+1}
\end{bmatrix} R_i = \begin{bmatrix}
\text{Re} J_i & \text{Im} J_i \\
-\text{Im} J_i & \text{Re} J_i
\end{bmatrix}
\]

(7.67)

where
\[
J_{i}^{\text{Re}} = \begin{bmatrix}
\text{Re}(s_i) & 1 \\
& \text{Re}(s_i) & \ddots \\
& & \ddots & 1 \\
& & & \text{Re}(s_i)
\end{bmatrix}
\]

(7.68)

\[
J_{i}^{\text{Im}} = \text{diag} \left[ \text{Im}(s_i), \cdots, \text{Im}(s_i) \right]
\]

(7.69)

Now $\mathbf{J}$ should be revised into the form:

\[
\mathbf{J}^R = \text{diag} \left[ J_{n+1,n+2}^{R}, J_{n+3,n+4}^{R}, \cdots \right]
\]

(7.70)

and $\mathbf{I}$ should be revised into the form:

\[
\mathbf{I}^R = \left\{ \frac{1}{0!} \text{Re}(n+1), \frac{1}{0!} \text{Im}(n+1), \cdots, \\
\frac{1}{(n+1)!} \text{Re}(n+1), \frac{1}{(n+1)!} \text{Im}(n+1), \cdots \right\}
\]

(7.71)

Now

\[
X_I(s) = s\mathbf{I} - \mathbf{I}^R \mathbf{R} \mathbf{R}^{-1}
\]

(7.72)

can be calculated with pure real computation.

**Algorithm for Left-Right Factorization**

(i) Suppose $N_n$ in (7.8) is full column rank, calculate

\[
\Omega(s) = N(s)^* N(s)
\]

\[
= a_0 + a_1 s + \cdots + a_{2n} s^{2n}
\]

(7.73)

from given $N(s)$. Therefore $a_N$ is square matrix of full rank.
(ii) Denote

\[
A = \begin{bmatrix}
  a_{2n} & \cdots & a_2 & a_1 & a_0 \\
  & I_m & & & \\
  & & I_m & & \\
  & & & I_m & \\
\end{bmatrix}
\]

\[
B = \begin{bmatrix}
  a_{2n-1} & a_{2n-2} & \cdots & a_1 & a_0 \\
  & -I_m & & & \\
  & & -I_m & & \\
  & & & -I_m & \\
  & & & & 0
\end{bmatrix}
\]

Since \(A\) is square matrix of full rank, \([A, -B]\) is a regular pencil, we can find all generalized eigenvalues \(s_i\) \((i = 1, \cdots, k)\) and generalized eigenvectors \(v_i^0, \cdots, v_i^{l-1}\) of the pencil \([A, -B]\) by EISPACK. Suppose all eigenvalues with negative real part are:

\[s_i \quad i = k', k'+1, \cdots, k\]

and \(s_{k'}, s_{k'+1}\) are conjugate pair, \(\cdots\). Recall \(v_i\) is 2nx1 vector, we take the bottommost \(nm\) rows of \(v_i\) as \(v_i'\). Denote

\[
\mathbf{I}^R = \begin{bmatrix}
  \text{Re}(v^0_{k'}), \text{Im}(v^0_{k'}), \cdots, \text{Re}(v^1_{k'}), \text{Im}(v^1_{k'}), \cdots
\end{bmatrix}
\]

\[
J^R = \text{diag} \left[ J^R_{k', k'+1}, J^R_{k'+2, k'+3}, \cdots \right].
\]

(iii) Calculate
\[
\begin{bmatrix}
x_{n-1} & x_{n-2} & \cdots & x_1 & x_0 \\
-I_m & & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
\end{bmatrix} \begin{bmatrix}
X_1 \\
X_2 \\
\vdots \\
X_m \\
X_{m+1} \\
\vdots \\
X_{n-1} \\
X_n \\
\end{bmatrix} = \begin{bmatrix}
-I_m & & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
& & & & \\
\end{bmatrix}
\] (7.78)

(iv) Write

\[
X(s) = x_0 + x_1 s + \cdots + x_{n-1} s^{n-1} + I_m s^n
\] (7.79)

and

\[
Z(s) = (N N_n^n)^{1/2} X(s).
\] (7.80)

then

\[
\Omega(s) = Z(s)^* Z(s)
\] (7.81)

is a left-right spectral factorization of \( \Omega(s) \).

Our whole theory is built on the knowledge of the interpolation points and the corresponding interpolation vectors. Now we are in the position to determine them. From (3.36) and (7.7) we have

\[
\alpha^* (s) N(s) = 0(s-s_i)^{\frac{i}{i}}
\] (7.82)

which is also:

\[
N(-s)^* \alpha(-s) = 0(s-s_i)^{\frac{i}{i}}
\]

Compare it with (7.25) we can substitute \( N(s)^T \) and \( \alpha(-s) \) for \( \Omega(s) \) and \( \xi(s) \) respectively, and get following results.
When \( U(s) \) or \( N(s) \) are square matrices given by (7.7) and (7.8), suppose \( N_n \) is of null rank. Denote

\[
E = \begin{bmatrix}
N_n^T \\
\vdots \\
I_m \\
\vdots \\
I_m \\
\end{bmatrix}
\]  
(7.83)

\[
F = \begin{bmatrix}
N_{n-1}^T & N_{n-2}^T & \cdots & N_0^T \\
-I_m \\
-I_m \\
\ddots \\
-I_m & 0
\end{bmatrix}
\]  
(7.84)

The interpolation point \( s_i \) is the generalized eigenvalue (with positive real part) pencil \( \{ E, -F \} \), it can be found by a routine algorithm. We also get the corresponding eigenvectors by the algorithm, pick up the bottommost \( m \) rows, thus we get a set of "shortened eigenvectors". The transpose and conjugate of each shortened eigenvector multiplied by \((i-1)!\) according to its \( i \)th position in the Jordan block will be the interpolation vector \( d_i^* \).

7.3 Recursive Algorithm

The results in section 7.1 and 7.2 can also be used for \( z \)-domain. Furthermore there is a simple recursive algorithm for inside-outside unit disc which cannot be expected for left-right factorization in \( s \)-domain. This recursive algorithm is useful and efficient since the dimension of
operated matrices is the same as the dimension of $\Omega(z)$. The algorithm in the last section needs the computations on the pencils of large dimension, which is not efficient and sometimes is beyond the capability of the computer.

Suppose

$$\Omega(z) = z^n N(z)^* N(z)$$

$$= a_0 + a_1 z + \cdots + a_n z^n$$  \hspace{1cm} (7.85)

and suppose there is no zero on the unit circle. We want to factorize

$$\Omega(z) = Y(z)X(z)$$

$$= (a_0 + y_1 z + \cdots + y_n z^n) (I + x_1 z + \cdots + x_n z^n)$$  \hspace{1cm} (7.86)

such that $X(z)$ has only zeros in the unit disc and $Y(z)$ has only zeros in the outside of the unit disc. We suppose $a_i, x_i, y_i \in \mathbb{C}^{m \times m}$.

Consider the difference equation

$$a_0 e^{M+N} + a_1 e^{M+N-1} + \cdots + a_N e^{M} = 0 \hspace{1cm} (7.87)$$

it can be written in the following form:

$$\begin{bmatrix}
    a_N \\
    0 \\
    \vdots \\
    0 \\
\end{bmatrix}
\begin{bmatrix}
    e^{M} \\
    e^{M+1} \\
    \vdots \\
    e^{M+N-1} \\
\end{bmatrix}
= 
\begin{bmatrix}
    I_m \\
    I_m \\
    \vdots \\
    I_m \\
\end{bmatrix}
\begin{bmatrix}
    e^{M+N} \\
    e^{M+N-1} \\
    \vdots \\
    e^{M+2} \\
\end{bmatrix}$$
\[
\begin{bmatrix}
 a_{N-1} & a_{N-2} & \cdots & a_1 & a_0 \\
 -I_m & & & & \\
 & -I_m & & & \\
 & & \ddots & & \\
 & & & -I_m & 0 \\
\end{bmatrix}
\begin{bmatrix}
 e_1 \\
 e_2 \\
 \vdots \\
 e_{M+N-1} \\
 e_{M+N} \\
\end{bmatrix} = 0. 
\] (7.88)

Denote

\[
E_M = \begin{bmatrix}
 e_1 \\
 \vdots \\
 e_{M+N-1} \\
\end{bmatrix} 
\] (7.89)

(7.88) becomes

\[
AE_M + BE_{M+1} = 0 
\] (7.90)

Define $m \times mN$ matrix as follows:

\[
G_o = \begin{bmatrix}
 I & x_{o,N-2}^1 & \cdots & \cdots & \cdots & \cdots & x_{o,o}^1 \\
 I & x_{o,N-3}^2 & \cdots & \cdots & \cdots & \cdots & x_{o,o}^2 \\
 \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 I & x_{o,N-n}^n & \cdots & \cdots & \cdots & \cdots & x_{o,o}^n \\
\end{bmatrix} 
\] (7.91)

The elements $x'_i$ of $G_o$ can be chosen in an arbitrary way and only need to satisfy a loose (generic) condition that will be mentioned later.

Suppose we have found some
\[
E_0 = \begin{bmatrix}
  e_0 \\
  \vdots \\
  e_{N-1}
\end{bmatrix}
\]  
(7.92)

which satisfies

\[
G_0 E_0 = 0,
\]  
(7.93)

then from (7.90) we have

\[
E_0 = -A^{-1}BE_1
\]  
(7.94)

Putting (7.94) into (7.93), we have

\[
G_0 A^{-1}BE_1 = 0
\]  
(7.95)

Let

\[
G_1 = R_1 G_0 A^{-1}B
\]  
(7.96)

where \(R_1\) is a constant matrix to transform \(G_1\) in the same version of \(G_0\) as shown in (7.91), then we have

\[
G_1 E_1 = 0
\]  
(7.97)

which can be considered as one step ahead of (7.93).

The recursive process can be continued as follows:

At \(N\) th step we have

\[
G_N E_N = 0
\]  
(7.98)

let
\[ G_{M+1} = R_{M+1} G_M A^{-1} B \]  

then

\[ G_{M+1} E_{M+1} = 0 \]  

The key is to find \( R_{M+1} \) to keep \( G_{M+1} \) in the version of (7.91).

Denote

\[
G_M = \begin{bmatrix}
I & x_{M,N-2} & \cdots & \cdots & \cdots & 1 \\
I & x_{M,N-3} & \cdots & \cdots & 2 & x_{M,o} \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
I & x_{M,N-n} & \cdots & \cdots & \cdots & x_{M,o}
\end{bmatrix}
\]  

Write the \( j \)th row of \( G_M \) as a column:

\[
f_M^j = \begin{bmatrix}
I \\
x_{M,N-j-1} \\
\vdots \\
x_{M,o}
\end{bmatrix} \quad j = 1, \cdots, n. \]  

denote

\[
A_{M}^0 = \begin{bmatrix}
-a_{N-1} & a_{N} \\
a_{N-2} & \ddots \\
\vdots & \ddots & \ddots \\
a_{o} & 0
\end{bmatrix}
\]  

(7.103)
\[
A^j_M = \begin{bmatrix}
0 & 1 & & \\
& 0 & \ddots & \\
& & \ddots & 1 \\
-x_{M,N-j-1}^j & -x_{M,N-j-2}^j & \cdots & -x_{M,N-o}^j \\
-x_{M,N-j}^j & -x_{M,N-j-2}^j & \cdots & -x_{M,N-o}^j \\
\vdots & \vdots & \ddots & \vdots \\
-x_{M,N-j}^j & -x_{M,N-j-2}^j & \cdots & -x_{M,N-o}^j \\
-x_{M,N-j}^j & -x_{M,N-j-2}^j & \cdots & -x_{M,N-o}^j \\
\end{bmatrix}
\] (7.104)

for \( j = 1, \cdots, n-1 \).

The recursive process is

\[
f_{M+1}^1 = (x_{N-1}^1 - a_{N} x_{M,N-2}^1)^{-1} \otimes I \cdot A^0_M \cdot f_{M}^1
\] (7.105)

\[
f_{M+1}^j = (x_{M,N-j}^{j-1} - x_{M,N-j-1}^j)^{-1} \otimes I \cdot A^{j-1} M \cdot f_{M}^j
\] (7.106)

for \( j = 2, \cdots, n \). Where \( I \) is identity matrix \( I_m \).

**Recursive Algorithm**

(i) Start from \( G_0 \) by arbitrary choice of \( x^* \) in (7.91).

(ii) Calculate \( f_1^1, \cdots, f_n^1 \) with the help of (7.105) and (7.106).

(iii) Write down \( G_1 \) according to (7.101) and \( f_1^1, \cdots, f_n^1 \), and then find \( A^1_1(j = 1, \cdots, n-1) \) according to (7.104).

(iv) Calculate \( f_2^1, \cdots, f_n^1 \) and so on.

(v) The recursive process gives us \( n \) sequences:

\[
\{f_M^j, \quad M = 1, 2, \cdots \} \quad j = 1, \cdots, n
\]

We will prove that
and the limit offers a polynomial matrix

$$X(z) = x_0 + x_1 z + \cdots + x_{n-1} z^{n-1} + Iz^n$$

which contains and only contains all the zeros of $\Omega(z)$ in the unit disc.

(vi) Sometimes the following cases will happen:

$$\det(a_{N-1}^{-1} a_{N} x_{N, N-2}) = 0$$

or

$$\det(x_{M, N-j}^{-1} x_{j, M, N-j-1}) = 0$$

In these cases, we should stop the recursive process and choose a new arbitrary $x_j$, and again start a new recursive process.

**Theorem 7.2**

In the recursive algorithm above we have

$$f_M^n \rightarrow \begin{bmatrix} I_m \\ \vdots \\ x_0 \end{bmatrix} \quad \text{as } M \rightarrow \infty$$

generically and
\begin{equation}
X(z) = x_o + x_1z + \cdots + x_{n-1}z^{n-1} + I_m z^n
\end{equation}

contains and only contains all the zeros of \( \Omega(z) \) in the unit disc.

(Proof)

Put (7.47) into (7.90) we have

\begin{equation}
T^{-1}E_M = JT^{-1}E_{M+1}
\end{equation}

Denote

\begin{equation}
T^{-1}E_M = \begin{bmatrix}
- \bar{F}_M \\
\bar{F}_M \\
\bar{E}_M
\end{bmatrix}
\end{equation}

then we have

\begin{equation}
\begin{bmatrix}
- \bar{F}_M \\
\bar{F}_M \\
\bar{E}_M
\end{bmatrix} = \begin{bmatrix}
\bar{J} & 0 \\
0 & \bar{J}
\end{bmatrix} \begin{bmatrix}
- \bar{F}_{M+1} \\
\bar{F}_{M+1} \\
\bar{E}_{M+1}
\end{bmatrix}
\end{equation}

therefore

\begin{equation}
\bar{F}_M = \bar{J} \cdot \bar{F}_{M+1}
\end{equation}

\begin{equation}
\bar{E}_M = \bar{J} \cdot \bar{E}_{M+1}
\end{equation}

If \( \bar{J} \) is the diagonal block matrix of all Jordan blocks corresponding to the eigenvalues of \( \Omega(z) \) outside the unit disc and \( \bar{J} \) is the diagonal block matrix of all Jordan blocks corresponding to the eigenvalues of \( \Omega(z) \) in the unit disc, then
\[ \| J^{-M} \|_\infty \to 0 \quad \text{as} \ M \to \infty \]  \hspace{1cm} (7.118)

\[ \| J^{-M} \|_\infty \to 0 \quad \text{as} \ M \to 0 \]  \hspace{1cm} (7.119)

From (7.116) and (7.117) we have

\[ F_M = \frac{J^{-M} F_0}{r_{\infty}} \]  \hspace{1cm} (7.120)

\[ F_M = \frac{J^{-M} F_0}{r_{\infty}} \]  \hspace{1cm} (7.121)

Denote the ratio

\[ r_M = \frac{F_M}{F_{\infty - M}} \]  \hspace{1cm} (7.122)

we have

\[ r_M = J^{-M} r_0 J^M \to 0 \quad \text{as} \ M \to \infty \]  \hspace{1cm} (7.123)

From (7.89), (7.101), the bottommost row of (7.98) is

\[ I_M e_M + \kappa_{M,n-1} e_{M+1} + \cdots + \kappa_{M,0} e_{M+n} = 0 \]  \hspace{1cm} (7.124)

which means that

\[ \frac{F_M}{F_{\infty - M}} + \kappa_{M,n} \frac{F_{M+1}}{F_{\infty - M+1}} = 0 \]  \hspace{1cm} (7.125)

where
\[
X_M = \begin{bmatrix}
  x_{M,n-1}^n & x_{M,n-2}^n & \cdots & x_{M,1}^n & x_{M,0}^n \\
  -I & \ddots & \ddots & \ddots & \ddots \\
  \vdots & \ddots & \ddots & \ddots & \ddots \\
  -I & \ddots & \ddots & \ddots & \ddots \\
  -I & 0 & \ddots & \ddots & \ddots \\
\end{bmatrix}
\]  
(7.126)

and

\[
\overline{E}_M = \begin{bmatrix}
  e_M^1 \\
  \vdots \\
  e_{M+n-1} \\
\end{bmatrix}
\]  
(7.127)

is the top half part of \(E_M\) (notice \(N = 2n\), we also define the bottom half part of \(E_M\) as

\[
E_M = \begin{bmatrix}
  e_{M+n}^1 \\
  \vdots \\
  e_{N+N-1} \\
\end{bmatrix}
\]  
(7.128)

Obviously

\[
\overline{E}_{M+n} = E_M. 
\]  
(7.129)

(7.125) can be rewritten as follows:

\[
\overline{E}_{M+n} + X_{M+n}\overline{E}_{M+n+1} = 0
\]  
(7.130)

or

\[
E_M + X_{M+n}E_{M+1} = 0.
\]  
(7.131)

Coming back to (7.114) we have
\[
E_M = T \cdot \begin{bmatrix}
-F_M \\
T_{21}' & T_{22}' & E_M \\
F_M & &
\end{bmatrix}
\] (7.132)

Partitioning \(T\) as in (7.52) and noticing that \(E_M\) is the bottom half part of \(E_M'\), we have

\[
E_M = \begin{bmatrix}
-T_{21}' & T_{22}' & E_M \\
F_M & &
\end{bmatrix}
\]

\[
= \begin{bmatrix}
T_{21}' & T_{22}' & \begin{bmatrix}
F_M \\ r_M \cdot E_M \\
\end{bmatrix}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
T_{21}' & T_{22}' & \begin{bmatrix}
T_{21}' & T_{22}' & F_M \\
F_M & &
\end{bmatrix}
\end{bmatrix}
\]

\[
(7.133)
\]

and

\[
E_{M+1} = \begin{bmatrix}
T_{21}' & T_{22}' & \begin{bmatrix}
F_M \\ r_{M+1} \cdot E_{M+1} \\
\end{bmatrix}
\end{bmatrix}
\]

(7.134)

Put (7.133) and (7.134) into (7.131) we have

\[
\begin{bmatrix}
T_{21}' & T_{22}' & \begin{bmatrix}
F_M \\ r_M \cdot E_{M+1} \\
\end{bmatrix}
\end{bmatrix} + \begin{bmatrix}
F_M \\ r_{M+1} \cdot E_{M+1} \\
\end{bmatrix} = 0
\]

(7.135)

Since \(r_M \to 0\) as \(M \to \infty\), we have

\[
( T_{22}' + X_{M+n} T_{22} ) \cdot E_{M+1} \to 0, \quad \text{as } M \to \infty
\]

(7.136)
But from (7.121) we have

$$F_{M+1} \rightarrow \infty \quad \text{as } M \rightarrow \infty$$  \hspace{1cm} (7.137)

therefore

$$T_{22} I + X_{M+n} T_{22} \rightarrow 0 \quad \text{as } M \rightarrow \infty$$  \hspace{1cm} (7.138)

or

$$X_M \rightarrow -T_{22} I T_{22}^{-1} \quad \text{as } M \rightarrow \infty.$$  \hspace{1cm} (7.139)

Recall that (7.58) is also true for z domain, we have

$$\begin{bmatrix}
  x_{n-1} & x_{n-2} & \cdots & x_1 & x_0 \\
  -I \\
  \vdots \\
  -I \\
  -I & 0
\end{bmatrix}$$

and thus (7.111) is proved and

$$X(z) = x_0 + x_1 z + \cdots + x_{n-1} z^{n-1} + I z^n$$  \hspace{1cm} (7.141)

contains and only contains all the zeros of \( \Omega(z) \) in the unit disc (notice 7.54).

7.4 Optimal Infinite Index Norm Problem in Nonsquare Case

In this section we will discuss the optimal \( H_\infty \) norm problem in the most general case: that is, to solve (7.1) when \( U \) and \( V \) are nonsquare rational matrices.
Lemma 7.2

\[ \| \begin{bmatrix} x \\ y \end{bmatrix} \|_\infty \leq \sigma \]  \quad (7.142)

iff

\[ \| x((\sigma^2 - Y^*) Y)^{-1/2} \|_\infty \leq 1 \]  \quad (7.143)

(Proof)

(7.142) is actually

\[ d^* \begin{bmatrix} x^* \\ y^* \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} d \]
\[ \leq \sigma^2, \quad \text{for all } d \neq 0 \]

or

\[ d^* x x d \leq d^* ((\sigma^2 I - Y^*) Y) d \]

Denote

\[ \tilde{d} = ((\sigma^2 I - Y^*) Y)^{1/2} d \]

then

\[ \tilde{d}^* ((\sigma^2 I - Y^*) Y)^{-1/2} x x ((\sigma^2 I - Y^*) Y)^{-1/2} \tilde{d} \]
\[ \leq 1 \]
\[ \tilde{d}^* \tilde{d} \]

which is equivalent to (7.143). All the arguments are also true in the inverse direction.
Lemma 7.3

\[ \| [X, Y] \|_{\infty} \leq U \]  
(7.144)

iff

\[ \| (I^{2} - YY^{*})^{-1/2} X \|_{\infty} \leq 1 \]  
(7.145)

(Proof)

First we prove a useful relation:

\[ \| A \|_{\infty} = \| A^{*} \|_{\infty} \]  
(7.146)

which is from

\[ \| A \|_{\infty} = \sup_{\omega} \sigma(AA^{*})_{j\omega} \]

\[ = \sup_{\omega} \sigma(AA^{*})_{j\omega} = \| A^{*} \|_{\infty} \]

By using (7.146) we conclude that (7.144) iff

\[ \| \begin{bmatrix} X^{*} \\ Y^{*} \end{bmatrix} \|_{\infty} \leq U \]  
(7.147)

iff (from lemma 7.2)

\[ \| X^{*}((I^{2} - YY^{*})^{-1/2} \|_{\infty} \leq 1 \]

iff (7.145) is true. (using (7.146) again)

Now we are in a position to solve (7.1). By the results of section 7.1, 7.2, 7.3 we can find the inner-outer factorization of U:
\[ U = \begin{bmatrix} M_1 \\ 0 \end{bmatrix} \] \hspace{1cm} (7.148)

For a given positive real \( \ell \) suppose we find some \( h \in \mathbb{H}_n^{mxr} \) such that
\[ \| W - UhV \|_\infty \leq \ell \] \hspace{1cm} (7.149)

This is equivalent to
\[ \| \Theta^* W - \begin{bmatrix} M_1 hv \\ 0 \end{bmatrix} \|_\infty \leq \ell \] \hspace{1cm} (7.150)

Write
\[ \Theta^* W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \begin{bmatrix} m & m' - m \end{bmatrix} \] \hspace{1cm} (7.151)

(7.150) can be rewritten as follows:
\[ \| \begin{bmatrix} W_1 - M_1 hv \\ W_2 \end{bmatrix} \|_\infty \leq \ell \] \hspace{1cm} (7.152)

Using lemma 7.2, (7.152) iff
\[ \| (W_1 - M_1 hv)(T^{1/2}I_r - W_2^* W_2)^{-1/2} \|_\infty \leq 1 \] \hspace{1cm} (7.153)

Again, we can find an outer-inner factorization:
\[ V(T^{1/2}I_r - W_2^* W_2)^{-1/2} = [ M_2, 0 ] \] \hspace{1cm} (7.154)

(7.153) becomes
\[ \| W_1 (T^{1/2}I_r - W_2^* W_2)^{1/2} \|_\infty \leq \| \begin{bmatrix} M_2 & 0 \end{bmatrix} \|_\infty \leq 1 \] \hspace{1cm} (7.155)
Write
\[ W_1 \left( (I_m - W_2^* W_2)^{-1/2} \right)^* = \left[ W_3, W_4 \right] \]
(7.156)

Then (7.155) becomes
\[ \| W_3 - M_1 h M_2 \|_m \leq 1 \]
(7.157)

By using lemma 7.3, (7.157) iff
\[ \| \left( I_m - W_4^* \right)^{-1/2} (W_3 - M_1 h M_2) \|_m \leq 1 \]
(7.158)

(7.158) is a typical $L_\infty$/$H_\infty$ problem which can be transformed $H_\infty$ problem and solved by the results from chapter 4 to chapter 6.

Suppose
\[ \min_{\lambda \in H_\infty^{ex}} \| \left( I_m - W_4^* \right)^{-1/2} (W_3 - M_1 h M_2) \|_m = \| \lambda \|
(7.159)

For a given positive real $\lambda$ there are two possible cases:

(A) the process from (7.149) to (7.158) can go through, which means that the spectral factors
\[ \frac{1}{(I_m - W_2^* W_2)^{1/2}} \text{ and } \frac{1}{(I_m - W_4^* W_4)^{1/2}} \]
exist and are in $H_\infty^{ex}$.

(B) The process from (7.149) to (7.158) cannot go through, which means that at least one of the spectral factors does not exist or has a zero on the imaginary axis.
For case (A) there are still three possible subcases:

A(i) \( \lambda = 1 \), we have \( \min\) = \( \lambda \).

A(ii) \( \lambda > 1 \), then \( \min\) > \( \lambda \). Furthermore (7.159) means that there is \( h \in \mathbb{R}^n \) such that

\[
(W_3 - M_1 h M_2)(W_3 - M_1 h M_2)^* \leq \lambda (I - W_4 W_4^*) \leq \lambda I - W_4 W_4^* \tag{7.160}
\]

which is

\[
\begin{bmatrix}
(W_3 - M_1 h M_2)^* \\
(W_3 - M_1 h M_2)
\end{bmatrix}
\begin{bmatrix}
W_4 \\
W_4^*
\end{bmatrix} \leq \lambda I \tag{7.161}
\]

or

\[
\begin{bmatrix}
(W_3 - M_1 h M_2)^* \\
(W_3 - M_1 h M_2)
\end{bmatrix} \leq \lambda I \tag{7.162}
\]

From (7.156) and (7.154), (7.162) becomes

\[
(W_3 - M_1 h V)^*(W_3 - M_1 h V) \leq \lambda (I - W_2 W_2^*) \leq \lambda I - W_2 W_2^* \tag{7.163}
\]

or

\[
\begin{bmatrix}
W_3 - M_1 h V \\
W_2
\end{bmatrix} \leq \lambda I \tag{7.164}
\]

or

\[
\| W - UhV \|_\infty \leq \bar{\lambda} \| U \|
\]

which means that

\[
\| U \| \leq \| U \|_{\text{opt}} \leq \bar{\lambda} \| U \| \tag{7.165}
\]
A(iii) \( \lambda < 1 \), then \( \mu_{\text{opt}} < \mu \). Furthermore, (7.159) means that

\[
(W_3 - M_1 hM_2)(W_3 - M_1 hM_2)\leq \lambda (I_m - W_4 W_4)^* \tag{7.167}
\]

or

\[
\begin{bmatrix}
(W_3 - M_1 hM_2)^* \\
W_4
\end{bmatrix} \leq (\lambda + (1-\lambda) ||W_4||_\infty^2) I_m \tag{7.168}
\]

or

\[
\begin{bmatrix}
(W_3 - M_1 hM_2), W_4
\end{bmatrix}^* \begin{bmatrix}
(W_3 - M_1 hM_2), W_4
\end{bmatrix} \leq (\lambda + (1-\lambda) ||W_4||_\infty^2) I_r \tag{7.169}
\]

or

\[
(W_1 - M_1 hV)^*(W_1 - M_1 hV) \leq (\lambda + (1-\lambda) ||W_4||_\infty^2)(I_r - W_2^* W_2) \tag{7.170}
\]

or

\[
\begin{bmatrix}
W_1 - M_1 hV \\
W_2
\end{bmatrix} \leq \begin{bmatrix}
W_1 - M_1 hV \\
W_2
\end{bmatrix} \tag{7.171}
\]

\[
|| W - U_b V ||_\infty \leq \frac{1}{\sqrt{\left(\lambda^2 + (1-\lambda) (1-||W_4||_\infty^2) \cdot ||W_2||_\infty^2\right)}} \tag{7.172}
\]

therefore we conclude that

\[
\mu_{\text{opt}} \leq \frac{1}{\sqrt{\left(\lambda^2 + (1-\lambda) (1-||W_4||_\infty^2) \cdot ||W_2||_\infty^2\right)}} \tag{7.173}
\]
For case (B) there are two subcases:

B(i) After increasing /, it will come back to case (A); and then the search process can iterate in A(ii) and A(iii) and at last arrive at A(i). We can consider case B(i) is the same as case (A) inherently.

B(ii) After increasing /, the search process cannot arrive at A(i) at all. In this case, we can find a critical point /_c such that for /_c + ε (ε is any positive real) the process from (7.149) to (7.158) can go through, but the optimal λ of (7.159) is less than 1 and cannot approach 1 even ε approaches 0. That means /_c + ε (ε > 0) belongs to A(iii) but does not belong to A(i), and /_c belongs to case (B). In this case we can prove that the critical point /_c is the optimal norm of problem (7.1), and there is a set of optimal solutions whose maximal singular value need not to be constant at all frequencies.

All these are summed up into following algorithm:

**Algorithm For Solving (7.1)**

(i) Pick any h ∈ \( H^m^{max} \) (for example let h = 0) and compute

\[ || W - UhV ||_\infty \]

and take it or some positive real less than it as an initial estimate of \( /_{opt} \), denote as \( /_1 \).

(ii) According to \( /_1 \) we can compute (7.149) to (7.158) successively and solve (7.159) to find \( \lambda_1 \).

(iii) If \( \lambda_1 = 1 \), ...
\[ U_{\text{opt}} = U_1. \] (7.174)

(iv) If \( \lambda_1 > 1 \), then

\[ U_1 < U_{\text{opt}} \leq \sqrt{\lambda_1} U_1 \] (7.175)

and take some \( U_2 \) in \((U_1, \sqrt{\lambda_1} U_1)\) as the next estimate of \( U_{\text{opt}} \).

(v) If \( \lambda_1 < 1 \), then

\[ U_{\text{opt}} \leq \sqrt{\frac{U_1^2 (1 - \lambda_1) \cdot \|W_1\|^2 + (1 - \lambda_1) (1 - \|W_1\|^2) \cdot \|W_2\|^2}{\|W_1\|_\infty}} \] (7.176)

and take some positive \( U_2 \) satisfying (7.176) as the next estimate of \( U_{\text{opt}} \).

(vi) According to \( U_2 \) repeat (ii) to find \( \lambda_2 \), and according to \( \lambda_2 \) repeat (iii) or (iv) or (v), and so on until the optimal norm of (7.159) approaches 1

\[ \lambda_n \to 1. \] (7.177)

In this case we get an optimal solution of (7.1) whose maximal singular value is constant at all frequencies.

(vii) As \( U \) is shrunk, before \( \lambda_n \) increases to 1, case B(ii) happens; then we can find a critical point \( U_c \) by iteration. \( U_c \) is the optimal norm of problem (7.1), and the maximal singular value of the optimal solution need not be constant at all frequencies in this case.

Example 7.1

Case A(i) will happen.
We only take a one-sided problem as an illustrative example. In this case \( W_3, W_4 \) will not exist.

\[
\begin{bmatrix}
  s+10 \\
  s+1 \\
  0
\end{bmatrix}
\leq
\begin{bmatrix}
  s+10 \\
  (s+1)^2 \\
  0.1
\end{bmatrix}
\|H\|_\infty
\]

(7.178)

Now

\[
U = \begin{bmatrix}
  s+10 \\
  (s+1)^2 \\
  0.1
\end{bmatrix}
\]

\[
U^* U = \frac{0.01s^4 - 1.02s^2 + 100.01}{(-s+1)^2(s+1)^2} - M_1^* M_1
\]

where

\[
M_1 = \frac{0.1(s^2 + 17.378435s + 100.005)}{(s+1)^2}
\]

is an outer function.

\[
\begin{bmatrix}
  10(s+10) \\
  s^2 + 17.38435s + 100.005 \\
  (s+1)^2 \\
  s^2 + 17.38435s + 100.005
\end{bmatrix}
\]

\[
\Theta_1 = UM^{-1}
\]

\[
\begin{bmatrix}
  10(-s+10) \\
  s^2 + 17.378435s + 100.005 \\
  -(-s+1)^2 \\
  s^2 + 17.378435s + 100.005
\end{bmatrix}
\]

\[
\Theta_2 = \]

\[
\begin{bmatrix}
  \frac{-(-s+1)^2}{s^2 + 17.378435s + 100.005} \\
  \frac{10(-s+10)}{s^2 + 17.378435s + 100.005} \\
  \frac{s^2 + 17.378435s + 100.005}{s^2 + 17.378435s + 100.005}
\end{bmatrix}
\]
\[ \theta = [ \theta_1, \theta_2 ] \]
\[ U = \theta^* \begin{bmatrix} M_1 \\ 0 \end{bmatrix} \]
\[ \theta^* W = \begin{bmatrix} 10(-s+10)(s+10) \\ (s+1)(s^2 - 17.378435s + 100.005) \\ -(s+1)(s+10) \\ s^2 - 17.378435s + 100.005 \end{bmatrix} \]

Let

\[ \hat{H} = M_1 H \]

(7.178) becomes

\[ \min_{\hat{H} \in \mathcal{H}_\infty} \| \frac{\begin{array}{c} 10(-s+10)(s+10) \\ (s+1)(s^2 + 17.378435s + 100.005) \\ -(s+1)(s+10) \\ s^2 + 17.378435s + 100.005 \end{array}}{s^2 - 17.378435s + 100.005} \|_\infty \]

now

\[ W_2 = \frac{-(s+1)(s+10)}{s^2 + 17.378435s + 100.005} \]

and

\[ \| W_2 \|_\infty = 1 \]

(0) 0th step is to solve

\[ \min_{\hat{H} \in \mathcal{H}_\infty} \| \frac{\begin{array}{c} 10(-s+10)(s+10) \\ (s+1)(s^2 + 17.378435s + 100.005) \\ -(s+1)(s+10) \\ s^2 + 17.378435s + 100.005 \end{array}}{s^2 - 17.378435s + 100.005} \|_\infty \]
The optimal norm is

\[ \sqrt{\lambda_o} = 0.8949914 \]

therefore

\[ \gamma_{opt} \leq \sqrt{\lambda_o + \|w_2\|^2} = 1.376586 \]

(1) 1st step. Suppose we choose

\[ \gamma_1 = 1.269126 \]

By lemma 7.2 problem (7.178) becomes

\[
\min_{\hat{\theta} \in \mathbb{H}} \left\| \frac{10(-s+10)(s+10)}{s^2-17.37844s+100.005} - \frac{s^2-17.37844s+100.005}{s^2+20.67493s+161.9075} \hat{\theta} \right\|_{\omega}
\]

\[
= \frac{0.781461(s^2+20.67493s+161.9075)}{s^2+20.67493s+161.9075} \hat{\theta} \omega
\]

Solving it we get

\[ \lambda_1 = 0.883754 < 1 \]

and estimate

\[ \gamma_{opt} \leq \sqrt{\lambda_1 \gamma_1^2 + (1-\lambda_1) \|w_2\|^2} \]

\[
= \sqrt{0.8837654 \times (1.269126)^2 + (1-0.8837654) \times 1^2} = 1.2408459
\]

(2) Suppose we want to enlarge the estimation step and guess

\[ \gamma_2 = 1.1313708 \]

By lemma 7.2 the problem becomes
\[
\min_{\tilde{u}_n(\tilde{s})} \left\| \frac{10(-s+10)(s+10)}{0.529150(s+1)(s^2+23.05509s+212.983)} - \frac{s^2-17.37844s+100.005}{s^2+23.05509s+212.983} \hat{\mu}_{\infty} \right\|_{\infty}
\]

Solving it we get

\[\lambda_2 = 1.469069 > 1\]

which means that \(\lambda_2\) should be in the range of

\[\frac{1}{\lambda_2} \leq \frac{1}{\lambda_2} = (1.1313708, 1.3715774)\]

This step tells us

\[U_{\text{opt}} > 1.1313708\]

and the upper-bound contains nothing new since we already knew lower upper-bound from step (1).

(3) From (1) and (2) we guess

\[U_3 = 1.2303658\]

By lemma 7.2 the problem becomes

\[
\min_{\tilde{u}_n(\tilde{s})} \left\| \frac{10(-s+10)(s+10)}{0.716798(s^2+21.12154s+171.0881)} - \frac{s^2-17.37844s+100.005}{s^2+21.12154s+171.0881} \hat{\mu}_{\infty} \right\|_{\infty}
\]

Solving it we get

\[\lambda_3 = 0.9970056 < 1\]

Therefore

\[U_{\text{opt}} < \frac{1}{\lambda_3} \left( \frac{\lambda_3^2}{2} + (1-\lambda_3) \|w\|^2_{\infty} \right) = 1.2297404\]
and so on. For this simple example we find that after only 6 iterations

$$U_6 = 1.22956$$

with a precision of $10^{-5}$ for $U_{opt}$.

**Example 7.2**

Case B(ii) will happen.

$$\min_{H \in \mathcal{H}_m} \| \begin{bmatrix} \frac{1}{s+1} \\ \frac{1}{s+1} \end{bmatrix} 0 + \begin{bmatrix} 1 \\ \frac{s+1}{\sqrt{s+1/2}} \end{bmatrix} \cdot H \|_m$$  \hfill (7.179)

Suppose the norm is not larger than $U$, by using inner-outer factorization

$$\begin{bmatrix} \frac{1}{s+1} \\ \frac{1}{s+1} \end{bmatrix} = \begin{bmatrix} -\frac{s-1}{s+1} \\ \frac{s-1}{s+1} \end{bmatrix} \begin{bmatrix} s+1/2 \\ s+1/2 \end{bmatrix} \begin{bmatrix} \frac{1}{s+1} \\ \frac{1}{s+1} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{s+1}{s+1/2} \end{bmatrix}$$

(7.179) becomes

$$\| \begin{bmatrix} \frac{1}{(s+1)(-s+\sqrt{2})} + \frac{s+1/2}{s+1} H \end{bmatrix} \|_m \leq U \hfill (7.180)$$

or

$$\| \begin{bmatrix} \frac{-1}{-s+\sqrt{2}} \\ \frac{-1}{-s+\sqrt{2}} \end{bmatrix} \|_m \leq 1 \hfill (7.181)$$

when
\[ U > \frac{1}{\sqrt{2}} \] (7.182)

otherwise the spectral factorization will be meaningless and

\[ \frac{1}{\sqrt{d+1}} \frac{1}{\sqrt{2d^2-1}} \] (7.183)

will not be in \( H^\infty \).

When \( U > \frac{1}{\sqrt{2}} \), we have

\[ |W|_{s=\sqrt{2}} = \frac{1}{(\sqrt{2}+1)(\sqrt{2}/2\sqrt{2^2-1})} < \frac{1}{\sqrt{2}+1} < 1 \]

therefore we cannot expect

\[ \lambda \rightarrow 1 \]

The critical point is obviously

\[ U_c = \frac{1}{\sqrt{2}} \] (7.184)

which is also the optimal norm of problem (7.179). For critical point

\[ U_c = \frac{1}{\sqrt{2}} \], the solutions of (7.180) are

\[ H = \frac{(s-1)X + \sqrt{2}}{2X - \sqrt{2}(s+2)} \] (7.185)

where \( X \) be any element in \( BH^\infty \). Put (7.185) into the left side of (7.180), it becomes
\[ G = \begin{bmatrix} \frac{s((s-2)x+2)}{\sqrt{2}(s-\sqrt{2})(\sqrt{2}x-(s+2))} \\ \frac{1}{s-\sqrt{2}} \end{bmatrix} \]  

(7.186)

and it easy to check

\[ \sigma^2[G(j\omega)] = \frac{\omega^2}{2(\omega^2+2)} |X(j\omega)|^2 + \frac{1}{\omega^2+2} \]  

(7.187)

which is not larger than \( \frac{1}{2} \) when \( |X(j\omega)| < 1 \) and is obviously a function of \( \omega \). Only when

\[ |X(j\omega)| = 1 \quad \text{for all } \omega \]  

(7.188)

we have a special solution with

\[ \sigma[G(j\omega)] = \frac{1}{\sqrt{2}} = \text{constant} \]  

(7.189)

But in general, optimal solutions have the property that the maximal singular value is not constant on all frequencies as shown in (7.187).

In this chapter, we solved the nonsquare problem by an iterative method. The key point is to find an efficient procedure for spectral factorization. Two algorithms are suggested for spectral factorization, which are performed satisfactorily on the computer. In this chapter we also find that the interpolation points and the interpolation vectors are the generalized eigenvalues with positive real part and the shortened generalized eigenvectors of a pencil, when \( U \) in (3.1) or (4.1) is square matrix and the coefficient matrix of the highest order of the numerator of \( U \) is full rank.
In the next chapter, we will try to apply all these results to control systems.
CHAPTER 8 SYNTHESIS OF CONTROL SYSTEMS

8.1 Introduction

Feedback control systems must be internally stable, that is, the transfer function between any two points of the loop must be stable (i.e., analytic for Re s \( \geq 0 \)), otherwise the system could not bear any disturbance. This requirement can be translated into the Hardy space language. The sensitivity design requirements, the robust stability requirements, the robust sensitivity design, the robust regulation and tracking, the control system design given lack of the precise knowledge of the spectrum of the disturbance and noise, all these can be discussed in terms of \( H_\infty \) norm. Therefore the \( H_\infty \) synthesis method would offer a new road for the control system design. On the other hand the widely used LQG method is actually a synthesis method in \( H_2 \) norm. They are also closely related.

Let us start from the simplest configuration of a feedback system as follows:

\[
y = P u + d
\]

\[
u = -C y
\]

where \( P \) is a plant and \( C \) is a compensator. Suppose \( P \) is a \( r \times m \) rational matrix and \( C \) is a \( m \times r \) rational matrix. \( P \) and \( C \) can be written as ratios of two proper stable matrices which are also called generalized polynomial matrices and are obviously elements of \( H_\infty \) \( \rightarrow X \) spaces.
To do this let

\[ s = \frac{1}{\lambda} - a, \quad a \text{ is any positive real.} \quad (8.3) \]

Now \( P \) and \( C \) becomes rational matrices of \( \lambda \), we can find \( \hat{A}(\lambda), \hat{B}(\lambda), \hat{A}_1(\lambda), \hat{B}_1(\lambda), \hat{P}_c(\lambda), \hat{Q}_c(\lambda) \) of polynomial matrices in \( \lambda \) such that

\[ P = \hat{A}(\lambda)^{-1}\hat{B}(\lambda) = \hat{B}_1(\lambda)\hat{A}_1(\lambda)^{-1} \quad (8.4) \]

\[ C = \hat{P}_c(\lambda)\hat{Q}_c(\lambda) \quad (8.5) \]

Putting

\[ \lambda = \frac{1}{s+a} \quad (8.6) \]

back into (8.4), (8.5), they become

\[ P = A(s)^{-1}B(s) = B_1(s)A_1(s)^{-1} \quad (8.7) \]

\[ C = P_c(s)Q_c(s)^{-1} \quad (8.8) \]

now \( A, B, A_1, B_1, P_c, Q_c \) are elements in \( H^r_{\text{rf}}, H^r_{\text{rm}}, H^m_{\text{mx}}, H^r_{\text{rm}}, H^m_{\text{mr}} \), \( H^r_{\text{mf}} \) respectively.

The matrix

\[ H = \begin{bmatrix} (I+PC)^{-1} & (I+PC)^{-1}P \\ C(I+PC)^{-1} & C(I+PC)^{-1}P \end{bmatrix} \quad (8.9) \]

determines internal stability of the system.

For the generalized polynomial matrices \( A, B, A_1, B_1 \), when \( A, B \) coprime and \( A_1, B_1 \) coprime there exist compatible generalized polynomial
matrices $X, Y, X_1, Y_1$ such that Bezout identity exists:

$$
\begin{bmatrix}
  A & B \\
  -Y_1 & X_1
\end{bmatrix}
\begin{bmatrix}
  X & -B_1 \\
  Y & A_1
\end{bmatrix}
= \begin{bmatrix}
  I & 0 \\
  0 & I
\end{bmatrix}
$$

(8.10)

Let

$$C = (Y+A_1K)(X-B_1K)^{-1}$$

(8.11)

or

$$
\begin{bmatrix}
  Q_c \\
  P_c
\end{bmatrix}
= \begin{bmatrix}
  X & -B_1 \\
  Y & A_1
\end{bmatrix}
\begin{bmatrix}
  I \\
  K
\end{bmatrix}
$$

(8.12)

where $K$ is Y-J-B parameter matrix. Then (8.9) takes the following form:

$$H = \begin{bmatrix}
  (X-B_1K)A & (X-B_1K)B \\
  (Y+A_1K)A & (Y+A_1K)B
\end{bmatrix}
$$

(8.13)

From (8.13) internal stability is equivalent to $K$ stable or $K$ being in $H_\infty^{mxr}$.

8.2 Design of Sensitivity

In the configuration of (8.1) and (8.2), the relationship between output $y(s)$ and disturbance $d(s)$ is given by

$$y(s) = [I+P(s)C(s)]^{-1}d(s)$$

define

$$S(s) = [I+P(s)C(s)]^{-1}$$

(8.14)

as the sensitivity matrix function of the feedback system and
\[ T(s) = P(s)C(s)[I+P(s)C(s)]^{-1} \] (8.15)

as the complementary sensitivity matrix function. Obviously,

\[ S(s) + T(s) = I. \] (8.16)

The sensitivity function \( S \) is a measure of the ability of the feedback system to reduce the disturbance and also of the capability to maintain the output under the parameter changes of the plant.

In the practical design of control systems, one of the tasks of the control engineer is to find a stabilizing compensator to meet the requirement of

\[ \sigma[S(j\omega)] \leq |M_s(j\omega)|, \quad \text{for all } \omega \] (8.17)

where \( |M_s(j\omega)| \) comes from the following consideration:

(i) Deep disturbance rejection should exist for a given frequency range according to practical requirements (mostly in the low frequency range).

(ii) When \( P \) is strictly proper and compensator is proper,

\[ S(j\omega) = I \] (8.18)

which suggests that \( |M_s(j\omega)| \) should be near 1 in the high frequency range.

(iii) Disturbance response should not be increased at any frequency, that is

\[ |M_s(j\omega)| \leq 1, \quad \text{for all } \omega. \] (8.19)
For nonminimal phase plant, as first pointed out by Zames, (8.19) can never be satisfied, and it should be relaxed so that

$$|M_s(j\omega)| \geq 1, \text{ for some } \omega.$$  

Once the curve $|M_s(j\omega)|$ is given from practical requirements, the analytic explicit expression can be found from the shape of the curve, and then can be extended from $j\omega$ to $s$ plane analytically, say $M_s(s)$. $M_s(s)$ can be chosen stable and minimal phase.

(8.17) can be rewritten as follows:

$$-\sigma (M_s^{-1}(j\omega)S(j\omega)) \leq 1, \text{ for all } \omega.$$  \hspace{1cm} (8.20)

Since $M_s^{-1}(s)$ and $S(s)$ are analytic in the right half plane we have

$$-\sigma [M_s^{-1}(s)S(s)] \leq 1, \text{ Re } s \geq 0.$$  \hspace{1cm} (8.21)

Recall

$$S = (X-B_1K)A$$

$$||\psi||_\infty = \sup \sigma [\psi(j\omega)]$$

(8.20) is

$$||M_s^{-1}(X-B_1K)A||_\infty \leq 1.$$  \hspace{1cm} (8.22)

Factorize $A$ into outer-inner parts

$$A = A_o A_i$$

and denote
W = M_s^{-1}X_0

U = M_s^{-1}B_1

H = KA_0

(8.22) becomes

\[ \|W-UB\|_\infty \leq 1 \]  

which was discussed in chapter 5 and chapter 7.

U is a r \times m matrix. When the degree of control freedom is equal to or more than the degree of output freedom (m \geq r), (8.23) can be solved by the results of chapter 4 and chapter 5; when the degree of control freedom is less than the degree of output freedom (m < r), U is now a tall matrix and the problem (8.23) can be solved by the results of chapter 7.

**Example 8.1**

Suppose plant

\[
P = \begin{bmatrix}
\frac{1}{s+1} & \frac{s-1}{(s+1)^2} \\
\frac{s-2}{(s+1)^2} & 0
\end{bmatrix} = A^{-1}B = B_1A_1^{-1}
\]

where

\[
A = A_1 = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]

\[
B = B_1 = P.
\]
Take

\[
X = X_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

\[
Y = Y_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
\]

then the Bezout identity exists.

Suppose \(|M_s(j\omega)|\) is as shown in figure 8.1, its analytical expression in s domain is

\[
M_s(s) = \frac{(s+0.5)(s+0.01)^2}{(s+0.4119602)(s^2+0.05080398s+0.0061899)}
\]

The sensitivity design problem becomes to solve

\[\|W-UB\|_\infty \leq 1\]

where

\[
W = \frac{s^3+0.47s^2+0.0301s+0.00233}{s^3+0.52s^2+0.0101s+0.00005}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

\[
U = \begin{bmatrix}
\frac{1}{s+1} & \frac{s-1}{(s+1)^2} \\
\frac{s-2}{(s+1)^2} & 0
\end{bmatrix}
\]

U has two zeros: \(s_1 = 1, s_2 = 2\).

For \(s_1 = 1\),

\[
a_1^* = (1, 2)
\]
\[ \beta_1^* = (0.98202791, 1.96405582) \]

For \( s_2 = 2 \),

\[ a_2^* = (0, 1) \]

\[ \beta_2^* = (0, 0.98440633) \]

Solve the generalized eigen-equation

\[
\begin{vmatrix}
5\lambda - 4.821894 & 2\lambda - 1.9334289 \\
2 & 3 \\
2\lambda - 1.9334289 & \lambda - 0.9690558 \\
3 & 4
\end{vmatrix} = 0
\]

we get

\[ \lambda_{\max} = 0.9710774 < 1 \]

Therefore the sensitivity design problem has solutions.

The Pick matrix is

\[ P = \begin{bmatrix}
0.08905296 & 0.022190339 \\
0.022190339 & 0.07736044
\end{bmatrix} \]

\[ a^* = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix} \]

\[ \beta = \begin{bmatrix}
0.98202791 & 1.96405582 \\
0 & 0.98440633
\end{bmatrix} \]
\[
\begin{bmatrix}
1 \\
2 \\
\frac{1}{s+1} 0.98202791 \\
1.96405582
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 \\
1 \\
\frac{1}{s+2} 0 \\
0.98440633
\end{bmatrix}
\]

from (5.13) we have

\[
L = \frac{1}{(s+1)(s+2)} \begin{bmatrix}
-36.367879s-76.735788 \\
34.188432s+68.376864 \\
-38.660356s-77.320712 \\
33.842576s-43.478134
\end{bmatrix}
\begin{bmatrix}
38.660356s+77.320712 \\
-33.57998s+43.746714 \\
s^2+40.969874s+77.931097 \\
-33.234357s+42.696744
\end{bmatrix}
\begin{bmatrix}
34.188432s+68.376864 \\
-155.955488s-88.578612 \\
33.573994s+67.147989 \\
-156.6394s-89.491412
\end{bmatrix}
\]

One important thing which should be mentioned is that

\[
L(\infty) W(\infty) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

\[
U \text{ strictly proper and } K \text{ in } H^s_{\infty} \text{ requires}
\]

\[
X(\infty) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

which is an interpolation problem at infinity. From chapter 6, for the existence of solution is that the $H_{\infty}$ norm of the interpolation value
matrix at infinity is not larger than 1 in the nondegenerate case.

Let us take

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

then

$$\Phi = \frac{1}{s^2 + 0.023591s + 0.3868105} \begin{bmatrix} s^2 + 0.0108777s + 0.3868627 & 0.00621576(s+2) \\ 0.00621576(s-2) & s^2 - 0.0108777s + 0.3868627 \end{bmatrix}$$

and then

$$G = \frac{1}{\delta} \begin{bmatrix} -0.00621576(s+1)^2 & -0.0155312(s+1)^2(s+0.099143) \\ (s+0.5)(s+0.01)^2 & (s^2 - 0.477079s + 0.313477) \end{bmatrix}$$

where

$$\delta = (s+0.4119602)(s^2 + 0.0580398s + 0.0061899)(s^2 + 0.023591s + 0.3868015)$$

The compensator should be

$$C = \frac{0.001(s+1)^2}{s^3 + 0.52s^2 + 0.0101s + 0.0005}(s^4 + 0.7729683s^2 + 0.1495851)$$

$$ \begin{bmatrix} -6.2158s^5 - 3.06806s^4 & -15.5312s^5 + 5.7395s^4 \\ -26.5998s^3 - 1.15042s^2 & -9.97663s^3 + 1.78196s^2 \\ -0.0726572s - 0.00612194 & -1.60214s - 0.186588 \\ -31.0708s^5 + 4.27032s^4 & 9.31544s^5 + 3.577278s^4 \\ -20.758s^3 + 0.745436s^2 & -0.992606s^3 + 0.736173s^2 \\ -3.45017s - 0.378888 & -1.81912s - 0.174968 \end{bmatrix}$$
8.3 Robust Stability of Plant Uncertainty

The importance of obtaining robustly stable feedback control systems has been recognized by engineers for a long time. First of all, any model is an approximation of reality and one cannot give an exact model for a practical plant. Secondly, our theory only deals with linear time-invariant systems — slow changes of plant parameters, nonlinearities and dynamic effects must be neglected in the theory. Thirdly, for lightening the burden of computation lower order approximation models are often used.

The problem of robust stability under plant uncertainty was discussed by many authors. In this section the main results of Doyle, Stein [37], Lehtomaki, Sandell, Athans [38], Safonov [43], Vidyasagar [44] and Kimura [44] are briefly summarized. Their results depend heavily on the requirement that the number of unstable poles of the nominal plant is the same as the number of unstable poles of the actual plant. To avoid this constraint, an extension of these results to a more general case is discussed in this section.

Denote the nominal plant by $P$ and the actual plant by $P'$, the difference between the true plant and the model is

$$\Delta P = P' - P$$

We suppose that controller $C$ is fixed and precisely known.

The right half plane pole is called an unstable pole, the closed right half plane pole (including the imaginary axis) is called a generalized unstable pole.
Since asymptotic stability is considered in most control problems, all generalized unstable poles of the closed loop must be excluded.

Suppose the numbers of generalized unstable poles of $P$, $P'$, $C$, $PC$, $P'C$ are $n^+_p, n^+_p', n^+_c, n^+_pc, n^+_p'c$, respectively. Obviously, we have

$$n^+_p + n^+_c \geq n^+_pc$$
$$n^+_p' + n^+_c \geq n^+_p'c$$

the equalities only take place when there are no cancellations of generalized unstable zeros and poles.

From now on, "The closed loop of $P$ and $C$ is internally stable" is said as "$C$ stabilizes $P$" for short. In this case there is no generalized unstable pole-zero cancellation between $P$ and $C$. So we have

$$n^+_pc = n^+_p + n^+_c$$

Suppose the numbers of generalized unstable zeros of $\det [I+PC]$ and $\det [I+P'C]$ are $z$ and $z'$ respectively. Then for the internal stability of the closed loop of $P$ and $C$ iff

(i) $z = 0$;

(ii) $n^+_pc = n^+_p + n^+_c$.

Consider the Nyquist contour $D_\delta$ which is the limitation of $D^\delta_R$ as $\delta \to 0^-$ and $R \to +\infty$, where $D^\delta_R$ is composed of a line

$$\text{Re } s = \delta, \quad \delta < 0$$

and a right half circle

$$|s| = R, \quad R > 0.$$
Let

\[ N(0, \det[I+PC], D_+^e) \]

represent the number of encirclements of the origin of the Nyquist curve of \( \det [I+PC] \) when \( s \) varies over the Nyquist contour \( D_+^e \).

**Lemma 8.1**

\[ N(0, \det [I+PC], D_+^e) = n_{pc}^+ - z \]  \hspace{1cm} (8.24)

\[ N(0, \det [I+P'C], D_+^e) = n_{p'C}^+ - z' \]  \hspace{1cm} (8.24)

**Lemma 8.2**

If

(i) \( P \) and \( P' \) have same number of generalized unstable poles.

(ii) \( C \) stabilizes \( P \).

(iii)

\[ \det \left[ I + (1-\varepsilon)P(s)C(s) + \varepsilon P'(s)C(s) \right] \neq 0 \quad \text{for } s \text{ on } D_+^e \quad 0 < \varepsilon < 1 \]  \hspace{1cm} (8.23)

then \( C \) also stabilizes \( P' \).

(Proof)

From (ii), the closed loop of \( P \) and \( C \) is internally stable, we have

\[ n_{pc}^+ = n_p^+ + n_c^+ , \quad \text{and} \quad z = 0 . \]
From (iii), \( \det [I + PC] \) and \( \det [I + P'C] \) are homotopic on \( D_+ \), therefore we have

\[
N(0, \det [I+PC], D_+) = N(0, \det [I+P'C], D_+) ,
\]
or

\[
n_{pc}^+ - z = n_{p'c}^+ - z' .
\]

Since \( z = 0 \), and \( z' \geq 0 \), we have

\[
n_{p'c}^+ \geq n_{pc}^+ .
\]

Therefore we have

\[
n_{p'}^+ + n_{c}^+ \geq n_{p'c}^+ \geq n_{pc}^+ \geq n_{p}^+ + n_{c}^+ .
\]

From (i), \( n_{p'}^+ = n_{p'}^+ \) we conclude that

\[
n_{p'}^+ + n_{c}^+ = n_{p'c}^+ = n_{pc}^+ = n_{p}^+ + n_{c}^+ .
\]

Therefore there is no cancellation between \( P' \) and \( C \), we also find \( z' = 0 \). These two conditions guarantee the internal stability of the closed loop of \( P' \) and \( C \).

**Theorem 8.1**

Let

\[
P' = [I + L\Theta R]P
\]

(8.21)

where \( L, R \) are compatible minimal phase stable matrices, \( \Theta \) runs \( BL_{\infty}^X \) and keeps \( P' \) and \( P \) having the same number of generalized unstable poles. If \( C \)
stabilizes \( P \), then \( C \) also stabilizes all \( P' \) iff

\[
\| RPC(I+PC)^{-1} L \|_{\infty} < 1. \tag{8.28}
\]

(Proof)

(i) if part,

\[
det [ I + (1-\varepsilon)PC + \varepsilon P' C ]
\]

\[
= det [I+PC] \cdot det [ I + \varepsilon \Theta RPC(I+PC)^{-1} L ]
\]

\#0 for \( s \in D^+ \), \( 0 \leq \varepsilon \leq 1 \)

by lemma 8.2, \( C \) also stabilizes all \( P' \).

(ii) only if part. Suppose

\[
\| RPC(I+PC)^{-1} L \|_{\infty} \geq 1 \tag{8.29}
\]

then there exists some point on the imaginary axis, say \( j \omega_o \), such that

\[
\sigma [ RPC(I+PC)^{-1} L ] = \sigma_o \geq 1, \quad \text{at } j \omega_o \tag{8.30}
\]

therefore we can find a vector \( x \) and a constant unitary matrix \( U_o \) such that

\[
RPC(I+PC)^{-1} L \cdot x = \sigma_o U_o x \quad \text{at } j \omega_o
\]

Let

\[
y = (I+PC)^{-1} L(j \omega_o) \cdot x
\]

then
\[ [I+(I+L(-\sigma_o^{-1}U_o^*)R)PC] y = 0 \quad \text{at } j\omega_o \]  \hspace{1cm} (8.31)

Let

\[ P' = (I+L(-\sigma_o^{-1}U_o^*)R)P \]  \hspace{1cm} (8.32)

which is in the cone (8.27), and (8.31) becomes

\[ [I+P'C] y = 0 \ , \quad \text{at } j\omega_o \]  \hspace{1cm} (8.33)

which is

\[ \det [I+P'C] = 0 \ , \text{at } j\omega_o \]  \hspace{1cm} (8.34)

or it has a generalized pole at \( j\omega_o \), that is contradictory to \( C \) stabilizing all \( P' \) in the cone (8.27).

Similarly we can write down

**Theorem 8.1'**

Let

\[ P' = P[I+L0R] \]  \hspace{1cm} (8.35)

where \( L, R \) are compatible minimal phase and stable matrices, \( \Theta \) runs \( BI_{\infty} \times \)
and keeps \( P' \) and \( P \) having the same number of generalized unstable poles.

If \( C \) stabilizes \( P \), then \( C \) stabilizes all \( P' \) in cone (8.35) iff

\[ \|RCP(I+CP)^{-1}L\|_{\infty} \leq 1 \]  \hspace{1cm} (8.36)

**Remark 8.1**

The condition \( P \) and \( P' \) have the same number of generalized unstable poles cannot be omitted, otherwise theorem 8.1 and 8.1' will be false.
Now we are going to consider the general case of dropping the requirement of $m = m'$, that means we want to give a criterion for robust stabilizability design even though $P$ and $P'$ have a different number of generalized unstable poles.

**Theorem 8.2**

For nominal plant $P = A^{-1}B = B_1A_1^{-1}$, and let $P' = A'^{-1}B'$, $A' = A + \Delta A$, $B' = B + \Delta B$; $A$, $B$, $A_1$, $B_1$, $A'$, $B'$, $\Delta A$, $\Delta B$ are all in $H^\infty$ with compatible dimension. When the perturbations of $\Delta A$ and $\Delta B$ are in the ball of

$$|| \begin{bmatrix} L^{-1}A^{-1}\Delta A & L^{-1}A^{-1}\Delta B \\ L^{-1}A^{-1}\Delta A & L^{-1}A^{-1}\Delta B \end{bmatrix} ||_{\infty} \leq 1,$$

(8.37)

where $L$, $L_1$, $L_2$ are compatible minimal phase stable rational matrices; $C$ which stabilizes $P$ also stabilizes all $P'$ in (8.37) if and only if

$$\frac{R_1 (X-B_1K)AL}{R_2 (Y+A_1K)AL} ||_{\infty} < 1.$$ 

(8.38)

(Proof)

The loop matrix for actual plant $P'$ and $C$ is

$$H' = \begin{bmatrix} (X-B_1K)D^{-1} - A' & (X-B_1K)D^{-1} - B' \\ (Y+A_1K)D^{-1} - A' & (Y+A_1K)D^{-1} - B' \end{bmatrix}$$

(8.39)

where

$$D = I + \Delta A(X-B_1K) + \Delta B(Y+A_1K)$$

(8.40)

(i) if part.
\[
\det \left[ (1-\varepsilon)I + \varepsilon D \right] = \det \left[ I + \varepsilon \left[ \Delta A(X-B_1 K) + \Delta B(Y+A_1 K) \right] \right]
\]
\[
= \det \begin{bmatrix}
I + \varepsilon [A \Delta B_1^{-1} L^{-1} A^{-1} \Delta A_1^{-1} L^{-1} A^{-1} \Delta BR_1^{-1}] & R_1(X-B_1 K) AL \\
R_2(Y+A_1 K) AL & R_2(Y+A_1 K) AL
\end{bmatrix}
\]
\[
\neq 0 \quad \text{on } D_+ \quad (8.41)
\]

Therefore \( \det I \) and \( \det D \) are homotopic to origin point for \( D_+ \). Notice \( \det D \) contains no generalized unstable poles, therefore it also contains no unstable zeros, so \( U' \) is stable for all \( P' \) in (8.37).

(ii) only if part.

Write
\[
\begin{bmatrix}
R_1(X-B_1 K) AL \\
R_2(Y+A_1 K) AL
\end{bmatrix} = \phi \begin{bmatrix} M \\ 0 \end{bmatrix} \quad (8.42)
\]
as an inner-outer factorization. If (8.38) does not exist then
\[
\|M\|_\infty = \sigma_0 \geq 1
\]

We can find some vector \( x \) and a constant unitary matrix \( U_0 \) such that
\[
M \cdot x = \sigma_0 U_0 \cdot x, \quad \text{at } j\omega_0 \quad (8.43)
\]

Define
\[
y = AL \cdot x \quad (8.44)
\]
and
\[
\begin{bmatrix}
\Delta A_o & \Delta B_o
\end{bmatrix} = \mathcal{A}\begin{bmatrix}
-\sigma_o^{-1}U_o^* & 0
\end{bmatrix} \varphi_o^* \begin{bmatrix}
R_1 & 0
\end{bmatrix} \begin{bmatrix}
0 & R_2
\end{bmatrix}
\]
(8.45)

where

\[
\varphi_o^* = \varphi^* |_{j\omega_o}
\]

then we have

\[
\begin{bmatrix}
L^{-1}A^{-1}\Delta A_o R_1^{-1} & L^{-1}A^{-1}\Delta B_o R_2^{-1}
\end{bmatrix} = \begin{bmatrix}
-\sigma_o^{-1}U_o^* & 0
\end{bmatrix} \varphi_o^*
\]

which is in (8.37) and we also have

\[
\begin{bmatrix}
X-B_1K
\end{bmatrix} \begin{bmatrix}
I + [\Delta A_o & \Delta B_o]
\end{bmatrix} \begin{bmatrix}
Y-A_1K
\end{bmatrix} \cdot y = 0, \quad \text{at } j\omega_o
\]

or

\[
D\cdot y = 0, \quad \text{at } j\omega_o
\]
(8.46)

which means that \( D \) has a generalized unstable zero \( j\omega_o \) for a special plant \( P' \) in (8.37).

We will prove that this \( j\omega_o \) is a generalized unstable pole of \( H' \). If it is not so, then \( H' \) is analytic at \( j\omega_o \). Write

\[
H' = \begin{bmatrix}
X-B_1K \\
Y-A_1K
\end{bmatrix} D^{-1} \begin{bmatrix}
A' & B'
\end{bmatrix}
\]
(8.47)

Since \( A' \) and \( B' \) are coprime, we can find \( X' \) and \( Y' \) in \( R_{n+1} \) with compatible dimension such that
\[
[ A', B' ] \cdot \begin{bmatrix} X' \\ Y' \end{bmatrix} = I
\] (8.48)

Multiplying

\[
[ A, B ] \quad \text{and} \quad \begin{bmatrix} X' \\ Y' \end{bmatrix}
\]

to (8.47) from left and right respectively, we have

\[
[ A, B ] \cdot H' \cdot \begin{bmatrix} X' \\ Y' \end{bmatrix} = D^{-1}
\] (8.49)

which is a contradiction since the left side of (8.49) is analytic at \( j\omega_0 \), but the right side of (8.49) has a pole at \( j\omega_0 \).

The theorem is proved.

Problem (8.38) is also in the form which was discussed in chapter 7.

8.4 Robust Sensitivity Design Under Plant Uncertainty

A more practical problem is to find compensator \( C \) such that all the sensitivity function matrices satisfying

\[
\| R_1 S' L_1 \|_\infty < 1
\] (8.50)

for all plant \( P' \) in uncertainty cone

\[
P' = P(I+L_2 \Theta_2 R_2)
\] (8.51)

and

\[
S' = (I+P'C)^{-1}
\] (8.52)

where \( R_1, R_2, L_1, L_2 \) are compatible stable minimal phase weighting func-
tion matrices, $\Theta_2$ runs $BL^{X}_\infty$ and keeps in $P$ and $P'$ the same number of
generalized unstable poles.

Put (8.52) into (8.50) and using an argument similar to that in
section 8.3, we have that (8.50) exists if and only if

$$\det(I+L_1\Theta_1 R_1 + P' C) \neq 0$$ (8.53)

in closed right half plane for all $\Theta_1 e^{BH^{X}_\infty}$ and the loop of $P'$ and $C$ is
internally stable (i.e. $\det[I+P'C] \neq 0$ in the closed rhp). In this case
$\Theta_1$ represents a fictitious perturbation used to satisfy (8.50). This type
of perturbation was first introduced by Doyle-Wall-Stein [39].

Put (8.51) into (8.53), we have

$$\det[I+L_1\Theta_1 R_1 + P(I+L_2\Theta_2 C)] \neq 0$$ (8.54)

for all $\Theta_1 e^{BH^{X}_\infty}$ and $\Theta_2 e^{BL^{X}_\infty}$ with $P$ and $P'$ having the same number of gen-
eralized unstable poles. From the homotopic conception (8.54) only needs
to exist on $D_+$ when $P$ and $P'$ have the same number of generalized
unstable poles. (8.54) can be rewritten into

$$\det[I+PC] \cdot \det[I+(I+PC)^{-1}L_1\Theta_1 R_1 + (I+PC)^{-1}PL_2\Theta_2 C] \neq 0$$ (8.55)

on $D_+$. The nominal closed loop should be internally stable,

$$\det[I+PC] \neq 0$$ (8.55b)

on $D_+$ (and in closed right half plane), therefore problem becomes

$$\det \left[ I + [(I+PC)^{-1}L_1 \quad (I+PC)^{-1}PL_2] \cdot \begin{bmatrix} \Theta_1 R_1 \\ \Theta_2 R_2 C \end{bmatrix} \right] \neq 0$$ (8.57)
on $D_+$. It can be rewritten as
\[
\det \begin{bmatrix}
    \Theta_1 & 0 \\
    I + \begin{bmatrix}
    0 & \Theta_2 \\
    \end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
    R_1(I+PC)^{-1}L_1 & R_1(I+PC)^{-1}PL_2 \\
    R_2C(I+PC)^{-1}L_1 & R_2C(I+PC)^{-1}PL_2
\end{bmatrix}
\neq 0 \quad (8.58)
\]

on $D_+^*$. Define
\[
G = \begin{bmatrix}
    R_1(I+PC)^{-1}L_1 & R_1(I+PC)^{-1}PL_2 \\
    R_2C(I+PC)^{-1}L_1 & R_2C(I+PC)^{-1}PL_2
\end{bmatrix}
\quad (8.59)
\]

which is the weighted matrix of $R$. And since now the problem is only considered on the imaginary axis, $\Theta_2$ can be considered only in $BH_{\omega}^{X}$. instead of from $BH_{\omega}^{I}$. Problem is simplified as
\[
\det \begin{bmatrix}
    \Theta_1 & 0 \\
    I + \begin{bmatrix}
    0 & \Theta_2 \\
    \end{bmatrix}
\end{bmatrix}
G \neq 0
\quad (8.60)
\]
on the imaginary axis for all $\Theta_1, \Theta_2 \in BH_{\omega}^{X}$.

Definition

Define $\mu(\mathbf{G})$ as
\[
\sup \{ \rho \mid \det \begin{bmatrix}
    \Theta_1 & 0 \\
    \rho^{-1} & \Theta_2 \\
\end{bmatrix}
G_j \omega = 0, \text{ some } \omega; \Theta_1, \Theta_2 \in BH_{\omega}^{X} \}
\quad (8.61)
\]
which is a special case of Doyle's $\mu$ function [57].

We summarize the foregoing into the following theorem:
Theorem 8.3

To solve the sensitivity design problem (8.50) under plant uncertainty (8.51) is equivalent to solving

\[ \min_{K \in \mathbb{H}_\infty^{m \times r}} \mu(G(K)) < 1 \]  

(8.62)

where

\[ G(K) = \begin{bmatrix} R_1(X-B_1K)A & R_1(X-B_1K)B L_2 \\ R_2(Y+A_1K)A & R_2(Y+A_1K)B L_2 \end{bmatrix} \]  

(8.63)

(8.62) is a new type of mathematical problem which we still don't know how to solve. To put it into our framework, the first attempt is to use \( \|G\|_\infty \) instead of \( \mu(G) \), obviously

\[ \mu(G) \leq \|G\|_\infty \]  

(8.64)

so it is perhaps too conservative. A modification is to use \( \|D^{-1}GD\|_\infty \) instead of \( \mu(G) \), where

\[ D = \begin{bmatrix} d_1I_1 & 0 \\ 0 & d_2I_2 \end{bmatrix} \]  

(8.65)

and \( d_1 \) and \( d_2 \) are any scalar rational functions of \( s \), \( I_1 \), \( I_2 \) are identity matrices with compatible dimensions, since we also have

\[ \mu(G) = \mu(D^{-1}GD) \]  

(8.66)

and

\[ \mu(G) \leq \|D^{-1}GD\|_\infty. \]  

(8.67)
The solution of

$$\min \min_{D,K} \| D^{-1}GD \|_\infty < 1$$  \hspace{1cm} (8.68)

must be also a solution of problem (8.62). It is a moderately conservative treatment.

**Example 8.2**

For plant

$$p = \frac{1}{s+1}$$  \hspace{1cm} (8.69)

with $43\%$ uncertainty, to find a compensator $C$ such that the sensitivity reduction is always less than

$$\left| \frac{j \omega + 1}{0.9 j \omega + 10} \right|$$  \hspace{1cm} (8.70)

for all possible plants.

**Solution**

For plant (8.69), $A = A_1 = 1$, $B = B_1 = \frac{1}{s+1}$, $X = X_1 = 1$, $Y = Y_1 = 0$.

To represent the plant uncertainty of $43\%$, in (8.51) take

$$L_2 = 1, \quad R_2 = 0.43$$

For the requirement of the sensitivity reduction, take

$$L_1 = 1, \quad R_1 = \frac{0.9 s + 10}{s + 1}$$

in (8.50). Now (8.63) becomes
\[
G = \begin{bmatrix}
\frac{0.9s+10(1-K)}{s+1} & \frac{0.9s+10(1-K)-1}{s+1} \\
0.43K & \frac{0.43K}{s+1}
\end{bmatrix}
\]

Our problem is to solve

\[
\mu(G) < 1
\]

Taking

\[
D = \begin{bmatrix}
1 & 0 \\
0 & 0.48(s+1)
\end{bmatrix}
\]

the problem

\[
\| D^{-1}GD \|_\infty < 1
\]

is

\[
\left\|
\begin{bmatrix}
\frac{0.9s+10}{s+1} \\
0
\end{bmatrix}
- \begin{bmatrix}
\frac{0.9s+10}{s+1} \\
0.895833
\end{bmatrix}
\cdot \frac{K}{s+1} \cdot \{1 \pm 0.48\}
\right\|_\infty < 1
\]

or

\[
\left\|
\begin{bmatrix}
(-0.9s+10)(0.9+10) \\
-1.612517 s + 100.80257
\end{bmatrix}
\right\|_\infty < \frac{1}{1.2304}
\]

or

\[
\left\|
\begin{bmatrix}
0.895833(0.9s+10) \\
-1.612517 s + 100.80257
\end{bmatrix}
\right\|_\infty < \frac{1}{1.2304}
\]

or

\[
\left\|
\begin{bmatrix}
(-0.9s+10)(0.9s+10) \\
-1.612517 s^2+100.80257
\end{bmatrix}
\right\|_\infty < 1
\]

and the solutions are
\[
\hat{Y} = \frac{(s-19.307028)X+17.61388}{-17.61388X+(s+19.307028)}
\]

the interpolation condition at \( s = \infty \) is

\[
\hat{Y}(\infty) = -0.8984986
\]

Taking

\[ X = -0.8984986 \]

a special solution is

\[
\hat{Y} = \frac{-0.8984986s+34.961218}{s+35.133075}
\]

and is corresponding to

\[
K = \frac{35.891467(s+1)(s+7.5681219)}{(s+35.133075)(s+7.9064879)}
\]

and

\[
C = \frac{35.891467(s+7.5681219)}{s+6.1482323}
\]

The sensitivity function for the nominal plant is

\[
S = \frac{(s+1)(s+6.1492323)}{(s+35.133251)(s+7.906448)}
\]

and for the perturbed plant

\[
p' = \frac{0.57}{s+1}
\]

is

\[
S' = \frac{(s+1)(s+6.1492323)}{(s+19.23174)(s+8.3671935)}
\]

they all satisfy the sensitivity specification (8.70).
Sometimes the solution for (8.50) causes too large controller output \( u \). To avoid it a modification of (8.50) is to design a compensator such that

\[
\begin{bmatrix}
R_1 S' L_1 \\
R_2 Q' L_2
\end{bmatrix}
\]

\[
\| \cdot \|_\infty < 1
\]

(8.73)

for all plant \( P' \) in the uncertainty cone

\[
P' = P(I + L_0 R O R_0)
\]

(8.74)

where

\[
Q' = (I + P'C)^{-1}
\]

(8.75)

is the Zames function matrix, and \( R_2 Q' L_2 \) is the controller cost.

By a similar argument, the problem is equivalent to solving

\[
\begin{bmatrix}
R_1 (X-B_1K)A_1 & R_1 (X-B_1K)BL_0 & 0 \\
R_2 (Y+A_1K)A_2 & 0 & R_2 (Y+A_1K)BL_0 \\
R_0 (Y+A_1K)A_1 & R_0 (Y+A_1K)BL_0 & 0 \\
R_0 (Y+A_1K)A_2 & 0 & R_0 (Y+A_1K)BL_0
\end{bmatrix}
\]

\[
\mu < 1
\]

(8.76)

subject to the structure of

\[
\begin{bmatrix}
\theta \\
\theta_0 \\
\theta_0
\end{bmatrix}
\]

(8.77)

which is still an open problem even after introducing the substitution of \( D^{-1}GD \) since it cannot be reduced into the form
\[ \| W - UKV \|_\infty < 1 \]  
(8.78)

Another similar practical problem is to design a compensator such that

\[ \| R_1 S' L_1 \|_\infty < 1 \]  
(8.79)

and

\[ \| R_2 Q' L_2 \|_\infty < 1 \]  
(8.80)

for all \( p' \) in (8.51), which is equivalent to

\[
\begin{bmatrix}
R_1 (X-B_1 K) A L_1 & R_1 (X-B_1 K) B L_0 \\
R_0 (Y+A_1 K) A L_1 & R_0 (Y+A_1 K) B L_0
\end{bmatrix}
\]

\[
\min \mu \mu \mu
K \mu \mu
\]

\[
R_2 (Y+A_1 K) A L_2 & R_2 (Y+A_1 K) B L_0 \\
R_0 (Y+A_1 K) A L_2 & R_0 (Y+A_1 K) B L_0
\]

\[ < 1 \]  
(8.81)

subject to the structure of

\[
\begin{bmatrix}
\theta_1 \\
\theta_0 \\
\theta_2 \\
\theta_0
\end{bmatrix}
\]

(8.82)

8.5 General Model

In this section a general model is considered:

\[ y = -H u + G_1 d \]  
(8.83)
\[ z = P u + G_2 d \]  

(8.84)

where \( y \) is the measured output, \( z \) is the output of the plant, and \( d \) contains disturbance and measurement noise. For meaningful control research we suppose that \( H \) contains all unstable patterns of \( P \). This model can cover most control problems such as regulation, tracking, filtering, and optimal controller design etc.

Let

\[ H = A^{-1}B = B_1 A_1^{-1} \]

and the Bezout identity as above. Define the feedback controller

\[ F = (Y + A_1 K)(X - B_1 K)^{-1} \]  

(8.85)

then

\[ Z = (PYAG_1 + G_2 + PA_1 KAG_1) \cdot d \]  

(8.86)

When the spectrum of stochastic disturbance \( d \) is given the design problem becomes

\[ \min_{K \text{ stable proper}} \| (PYAG_1 + G_2 + PA_1 KAG_1) d \|_2 \]  

(8.87)

which is the well known Weiner-Hopf-Kalman problem.

When the spectrum of stochastic disturbance \( d \) is unclear, we can design the system according to

\[ \min_{K \text{ stable proper}} \| R(PYAG_1 + G_2 + PA_1 KAG_1) L \| = \]  

(8.88)

where \( R \) and \( L \) are weighting function matrices which come from the practical requirements of control system design.
Consider the state space model

\[ \dot{x} = Ax + Bu + D_0d_0 \quad (8.89) \]

\[ y = Cx + Na \quad (8.90) \]

then

\[ y = C(sI-A)^{-1}Bu + C(sI-A)^{-1}D_0d_0 \quad (8.91) \]

and consider

\[ z = \begin{bmatrix} R_1x \\ R_2u \end{bmatrix} \quad (8.92) \]

the output of the system, then the state space system can be written in model (8.83) and (8.84) with

\[ \Pi = -C(sI-A)^{-1}B \quad (8.93) \]

\[ P = \begin{bmatrix} R_1(sI-A)^{-1}B \\ R_2 \end{bmatrix} \quad (8.94) \]

\[ G_1 = [C(sI-A)^{-1}D_0 N] \quad (8.95) \]

\[ G_2 = \begin{bmatrix} R_1(sI-A)^{-1}D_0 & 0 \\ 0 & 0 \end{bmatrix} \quad (8.96) \]

\[ d = \begin{bmatrix} d_0 \\ n \end{bmatrix} \quad (8.97) \]
Suppose a system is
\[ \dot{x} = -x + u + d_0 \]  
\[ y = 2x + n \]  
\[ z = \begin{bmatrix} x \\ \frac{1}{2}u \end{bmatrix} \]  
\[ (8.98) \quad (8.99) \quad (8.100) \]

Let \( A = A_1 = 1, \ B = B_1 = \frac{2}{s+1}, \ \dot{X} = \dot{X}_1 = 1, \ Y = Y_1 = 0 \). When \( d_0, n \) are independent unit white noise the \( H_2 \) problem (8.87) is
\[
\min_{K \text{ stable proper}} \| \begin{bmatrix} 1 & 0 \\
0 & 0 \end{bmatrix} + \begin{bmatrix} 1 \\
\frac{1}{2} \end{bmatrix} K \begin{bmatrix} 2 & 1 \\
\frac{2}{s+1} & 1 \end{bmatrix} \|_2 \]  
\[ (8.101) \]

After finding the inner-outer factorization, (8.101) becomes
\[
\min_k \| \begin{bmatrix} -2 & \frac{2}{s+1} \\
\frac{2}{s+1} & (s-15)^2 \\
\frac{4}{(s+1)(s-15)^2} + \frac{(s+15)^2}{2(s+1)^2} & \end{bmatrix} \|_2 \]  
\[ (8.102) \]

We only need to solve
\[
\min_k \| 4 - \frac{(s+15)^2}{(s+1)(s+15)^2} - \frac{2}{2(s+1)^2} \|_2 \]  
\[ (8.103) \]

Which is a repeated zero problem can be reached by (3.52) and (3.53), the optimal solution is
\[
\Phi = -0.381966 s + 2.09017 \\
\frac{1}{(s+15)^2} \]  
\[ (8.103) \]
therefore

\[ K = \frac{-0.763932(s+1)}{(s+1.5)^2} \]  \hspace{1cm} (8.104)

\[ F = \frac{-0.76392}{s+3.472136} \]  \hspace{1cm} (8.105)

Recently Doyle [40] proved the well known separation principle by
Zames parameter matrix Q. We will use the separation principle to check
our calculation in pure frequency domain with Y-J-B parameter K.

The Kalman filter is

\[ K = XC'(NN')^{-1} \]

and \( X \) is the semipositive solution of Riccati equation

\[ AX + XA' - XC'(NN')^{-1}CX + D'_{0}D'_{0} = 0 \]

which for our problem is

\[-2X - 4X^2 + 1 = 0 \]

\[ X = 0.6180339 \]

The LQR optimal controller is

\[ F = (R'_{2}R_{2})^{-1}B'Y \]

where \( Y \) is the semipositive solution of Riccati equation

\[ A'Y + YA - YB(R'_{2}R_{2})^{-1}B'Y + R'_{1}R_{1} = 0 \]

which for our problem is
\[-2Y - 4Y^2 + 1 = 0\]

\[F = 1.236068\]

So the whole compensator is

\[F = -F(sI - A + BE + KC)^{-1}K\]

\[= \frac{-0.763932}{s + 3.472136}\]

which is the same as (8.105)

When \(R_1, R_2\) in (8.94) and \(d\) in (8.87) are functions of frequency then the problem will be much more complicated in the framework of Kalman state space, even in the recently developed Moore, Mingorri and Anderson's theory [52] which can only deal with minimal phase plants. But it can be easily solved by the interpolation theory of \(H_2\) in a uniform method.

When the spectra of \(d\) and \(n\) are unknown, the design becomes to solve (8.88) by properly choosing weighting functions, which is an extension of classical design method. As an example we only compute a simple case:

\[R = L = 1\]

that is to solve

\[
\min_K || \begin{bmatrix} \frac{1}{s+1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{s+1} \\ \frac{1}{2} \end{bmatrix} K \left[ \begin{bmatrix} 2 \\ s+1 \end{bmatrix} \right] ||_{\infty} \tag{8.106}
\]
Suppose the infinite norm is not larger than $\mathcal{U}$, and using the inner-outer factorization we have

$$
\| \begin{bmatrix}
\frac{2}{(s-1)(s+1)\sqrt{5}} & 0 \\
\frac{1}{(s+1)\sqrt{5}} & 0
\end{bmatrix} + \begin{bmatrix}
\frac{s+1}{2(s+1)} & \frac{2}{s+1} \\
\frac{1}{2(s+1)} & \frac{1}{s+1}
\end{bmatrix} \|_{\infty} \leq \mathcal{U}
$$

or

$$
\| \begin{bmatrix}
\frac{2(s+1)}{(s+1)(s+1)\sqrt{5}} & 0 \\
\frac{1}{(s+1)\sqrt{5}} & 0
\end{bmatrix} + \begin{bmatrix}
\frac{s+1}{2(s+1)} & \frac{2(s+1)}{s+1} \\
\frac{1}{2(s+1)} & \frac{1}{s+1}
\end{bmatrix}, (\mathcal{U}^{-1}) \|_{\infty} \leq 1
$$

when $\mathcal{U} > \frac{1}{\sqrt{5}}$.

Let

$$s_1 = \sqrt{\frac{10-\mathcal{U}^2}{2} + j\sqrt{\frac{16\mathcal{U}^2-\mathcal{U}^4}{2}}}, \quad s_2 = s_1^{-1},$$

when $\mathcal{U} \geq \frac{1}{4}$, then (8.108) becomes

$$
\| \begin{bmatrix}
\frac{4(s+1)}{(s+1)(s+1)\sqrt{5}} & 0.5(s+1)(s+s_1)(s+s_2) \\
(s+1)(s+1)(\frac{1}{s+1})^2 & (s+1)^2(\frac{1}{s+1})^2
\end{bmatrix}
\|_{\infty} \leq 1
$$

or

$$
\| \begin{bmatrix}
\frac{4\mathcal{U}}{(s+1)(s+1)\sqrt{5}^2} & 0.5(\frac{1}{s-s_1})(s-s_2)(s+s_1)(s+s_2) \\
(s+1)(s+1)(\frac{1}{s+1})^2 & (s+1)^2(\frac{1}{s+1})^2
\end{bmatrix}
\|_{\infty} \leq 1
$$

After iteration process we can find
\( / = 0.5 \)  

which causes the optimal norm in (8.110) to be 1, and optimal function is

\[ \dot{\phi} = \frac{3-s}{3+s} \]  

(8.112)

Therefore

\[ K = -\frac{s+1}{s+3} \]

and the optimal compensator is

\[ F = -1 \]  

(8.113)

Now changing the problem a little, the phenomenon of "critical point" which was discussed in the last chapter will occur.

\[ \dot{x} = -x + u + d_0 \]  

(8.114)

\[ y = 2x + u \]  

(8.115)

\[ z = \begin{bmatrix} x \\ u \end{bmatrix} \]  

(8.116)

the only change is that we pay more attention to the control cost.

Problem will rise in \( H_\infty \) optimization:

\[
\begin{array}{c}
\begin{bmatrix}
\frac{1}{s+1} & 0 \\
0 & 0
\end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\frac{1}{s+1} \\
\frac{2}{s+1}
\end{bmatrix}
K \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\end{array}
\]

By the same argument, suppose the norm is not larger than \( /, \) we have
\[
\| \begin{bmatrix}
1 & 0 \\
-\frac{s+\sqrt{2}}{s+1} & 1 \\
0 & -\frac{s+1}{(s+\sqrt{2})}
\end{bmatrix} \| \leq U \leq 1 \]  
\text{when}

\[
\| \begin{bmatrix}
1 & 0 \\
\frac{s+\sqrt{2}}{s+1} & \frac{2(s+\sqrt{2})}{(s+1)(s+\sqrt{2}/2-1)}
\end{bmatrix} \| \leq 1
\]

Denote

\[
s_1 = \frac{\sqrt{7-u^2}}{2} + j\frac{\sqrt{10u^2-9-u^4}}{2}, \quad s_2 = \frac{1}{s_1},
\]

then it becomes

\[
\| \begin{bmatrix}
\frac{2U}{(s-s_1)(s-s_2)(s+s_1)(s+s_2)} \\
\frac{U(s-s_1)(s-s_2)(s+s_1)(s+s_2)}{(s+1)(s+\sqrt{2}^2-1)(s+\sqrt{5}^2-1)(s+1)^2(s+\sqrt{2}^2-1)(s+\sqrt{5}^2-1)}/s_1
\end{bmatrix} \|_\infty \leq 1
\]

By iteration we can find

\[
U = 0.6732
\]

for optimal norm of the left hand of (8.120) being 1. But this \(U\) violates the condition (8.119). In this case an imaginary pole appears in (8.120) which comes from \(U\) less than

\[
\| \begin{bmatrix}
\frac{-s+1}{(s+1)(-s+\sqrt{2})}
\end{bmatrix} \|_\infty = \frac{1}{\sqrt{2}}
\]

It is impossible since \(U\) is the whole norm of (8.118). The critical
point is

$$U = \frac{1}{|12|}$$

which is the optimal infinite norm of our problem (8.117), but in this case the optimal infinite norm of left side of (8.120) is only

$$|\hat{\lambda}| = 0.6460179 < 1$$

The general model (8.83) and (8.84) can deal with the regulation and tracking problems. In general d will consist of two types of inputs: unstable ones, e.g. steps, ramps, sinusoids, for which asymptotic regulation is desired and stable ones which affect only the error transient response.

We first consider asymptotic regulation for the unstable inputs and then error reduction for the stable ones as measured by the $H_\infty$ norm of the error transfer function. Let

$$d = M r$$

(8.123)

where $M$ is a rational matrix that is used to describe the asymptotic regulation objectives and $r$ is a proper rational function analytic in the closed right-half-plane.

From (8.86) and (8.123) we will achieve asymptotic regulation with internal stability if and only if

$$[PA_1KAG_1 + PYAG_1 + G_2] M$$

and $K$ are analytic in the closed right-half-plane.
From Cheng and Pearson [48] the necessary and sufficient conditions for this to take place are that there exist stable, proper matrices $N$, $K$ and $W$ such that

$$Q_2^{-1}Q_1 = N$$  \hspace{1cm} (8.124)

$$PA_1K + W\tilde{Q}_1 = P_2N1$$  \hspace{1cm} (8.125)

where

$$\text{AG}_1M = P_1Q_1^{-1} = \tilde{Q}_1^{-1}P_1$$  \hspace{1cm} (8.126)

$$\text{PTAG}_1M + G_2M_+ = P_2Q_2^{-1}$$  \hspace{1cm} (8.127)

$$\Pi_1P_1 + \Pi_2Q_1 = I$$  \hspace{1cm} (8.128)

and $X_+$ represents the unstable part of partial fraction expansion of $X$.

When (8.124) and (8.125) are satisfied and $PA_1$ and $\tilde{Q}_1$ have no common zeros, the general solution to (8.125) is

$$K = K_0 + R\tilde{Q}_1$$

$$W = W_0 + PA_1R$$  \hspace{1cm} (8.129)

where $R$ is an arbitrary stable matrix, and $K_0$ and $W_0$ represent any particular solution to (8.125). This allows the parameterization of all stabilizing compensators that are also asymptotic regulators. We can therefore rewrite (8.86) as follows:

$$z = [PA_1RT_1 + T_2]r$$  \hspace{1cm} (8.130)
where

\[ T_1 = \tilde{Q}_1 A G_1 M \]

\[ T_2 = P A_1 K_0 AG_1 M + (P Y A G_1 + G_2) M \]

Clearly \( T_2 \) is stable and proper since \( K_0 \) is a solution of the asymptotic regulator problem. It is also clear that \( T_1 \) is stable and proper.

We are now ready to treat the stable inputs represented by \( r \) in (8.130).

To include all possibilities, we will modify the transfer function of (8.130) on the left and on the right by stable rational matrices \( W_1 \) and \( W_2 \) so that (8.130) becomes

\[ \dot{z} = W_1 z = (W + U R V) \dot{r} \]  

(8.131)

where

\[ U = W_1 P A_1 \]

\[ V = T_1 W_2 \]

\[ W = W_1 T_2 W_2 \]

\[ r = W_2 \dot{r} \]

Roughly, the matrix \( W_1 \) can serve to describe the range of frequencies over which the regulated variables are to be considered and \( W_2 \) the range of frequencies over which the expected input will occur.

To summarize the above results, we have parameterized a general regulator problem with internal stability by means of a stable proper.
rational matrix \( R \) as given in (8.131). \( R \) determines a compensator that asymptotically regulates with internal stability through (8.129) and (8.85). The only assumptions made in this development were that (i) system (8.83) and (8.84) is admissible and (ii) \( PA_1 \) and \( \tilde{G}_1 \) have no common zeros. Assumption (ii) is necessary in order for (8.129) to represent a general solution to (8.125). If this is not true, then stable \( K \) and \( W \) can be produced by an unstable \( R \) in (8.129), and the optimization problem to be discussed becomes more technically involved.

The design problem becomes to solve

\[
\begin{align*}
\text{min} & \quad \| W + U R V \|_\infty \\
R \text{ proper, stable} & \tag{8.132}
\end{align*}
\]

which is the main topic of the dissertation.
CHAPTER 9 CONCLUSIONS

This dissertation deals with a very important topic of control theory, namely the infinite norm synthesis of linear multivariable system. Once the plant is not precise or the noise spectra are not exactly known as is normally the case, the LQG solution will no longer be optimal in practice and sometimes cannot even guarantee a stable feedback system. The robust design of feedback system can be performed in infinite norm space. At first glance the infinite norm synthesis method may seems very conservative, but it is not so because weighting functions can be used in the design. Kwakernaak noted that $H_\infty$ and $H_2$ optimal problem can be made equivalent in SISO case by proper choice of weighting functions. The advantage of infinite norm synthesis is the transparency for robustness, the weighting functions are directly related to the design specifications. This is reason behind the increasing interest in the $H_\infty$ problem in recently three years.

With a view to bring theory as close as possible to practical system design, most of formulas in the dissertation are explicitly given in the s-domain in their simplest form, while those for the repeated zeros case are also included in detail.

The well-known LQG problem was restated in the interpolation language of $H_2$ in chapter 3. This method seems much simpler than the variation method used by Youla et al in 1976. It is also a little more general than Kalman version because this method can consider frequency shaped weighting without creating any complexity. Instead of solving
Riccati equation the key turns out to be finding the inner-outer factorization, and the later was solved in a natural setting in chapter 8 by a routine computer algorithm which looks like a generalization of the algorithm used for solving the Riccati equation.

$H_2$ and $H_\infty$ optimal problems are closely related. In fact $H_\infty$ norm can be considered a ratio of two $H_2$ norms. Hence once the expression for optimal $H_2$ norm solution has been found the optimal $H_\infty$ norm can easily be found by the Lagrange method. There are many similarities between (3.20) and (5.13). In fact, $L$ can be considered an inner function matrix in Krein space. Compare (3.7) and (5.13) there are many similarities between them as well. But the optimal solution in $H_2$ space is unique and determined by (3.7), the solutions in $H_\infty$ are the linear fractional combination of four (sometimes six) blocks of $L$ and are enormous in most cases, therefore there are many differences between two problems also.

It is shown that the $H_\infty$ optimal problem must be a degenerate case and the degenerate problem can be easily solved without increasing any computational burden. Chapter 5 can be considered a development of the method of Francis, Helton and Zames; explicit expressions for both nondegenerate case and degenerate case are given in this chapter.

A recursive method was developed in chapter 6, which is an extension of the Nevanlinna and Schur algorithm. The main idea is to introduce an interpolation pair and expand the pair into Taylor series. It is interesting to find all derivatives can be embedded by the same procedure in this recursive method.
In the recursive method the basic formula is a variety of the general formula (5.13) when only one interpolation is taking place (n=1). Therefore the recursive method in chapter 6 can also be extended to batch version when (6.18) is substituted by (5.13) corresponding to the batch. If all the interpolation points are taken in one batch the recursive method is reduced to what was shown in chapter 5.

The nonsquare problem is of the most importance in control theory. Some new properties appear in the nonsquare problem. For example, the maximal singular value of the optimal solution of a nonsquare problem may vary with frequency, it is different from the all-pass property of the maximal singular value of the optimal solution of the square problem. To solve the nonsquare problem the first step is to find all free terms by inner-outer factorization. The second key step is to find the optimal norm, once it is found by some recipe then it will be reduced to a square problem. An attempt is to consider these free terms in a fictitious detection space, then the problem can be solved by the original method. To remove the fictitious detection space the number of interpolation points must be increased to infinity, therefore to find the optimal norm of a nonsquare problem one must solve an eigenequation with infinite dimension. The maximal eigenvalue of the truncated eigenequation converges slowly when some poles of the free terms are near imaginary axis. Therefore we prefer iteration to truncation. With the help of two inequalities one can determine the range of the optimal norm at each step, the convergence rate is very fast.
Recall that in chapter 1 the criticism of Chandler and Potts cited is not only on LQG but also on singular value theory: "Singular value theory appears to be modern control theorist answer to the uncertainty problem. But it is not a synthesis technique, it is purely analysis." But now the situation is changing since the singular value theory has been developed into a systematic synthesis method by introducing operator and Hardy space theory into the control area. Several attempts are made in chapter 8 to show how feedback control systems can be synthesized naturally in the $H_\infty$ framework.

Many design problems of feedback control systems can be related to an associated problem: To solve

$$\| W_1 - U_1 H \| < 1 \quad (9.1)$$

and

$$\| W_2 - U_2 H \| < 1 \quad (9.2)$$

or more inequalities with the same element $H$ of Hardy space in all inequalities. It is also to solve

$$\| \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix} - \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \begin{bmatrix} H & 0 \\ 0 & H \end{bmatrix} \|_{H_{\infty}} < 1 \quad (9.3)$$

Many control system design problems will become tractable once this problem is solved.

An attempt is to solve
\[
\begin{bmatrix}
  t_1 w_1 \\
  t_2 w_2
\end{bmatrix}
- \begin{bmatrix}
  t_1 u_1 \\
  t_2 u_2
\end{bmatrix} H \leq 1
\]  
(9.4)

where \( t_1, t_2 \) satisfying the following relation

\[
t_1(s) * t_1(s) + t_2(s) * t_2(s) = 1
\]  
(9.5)

Obviously any solutions of (9.3) must satisfy (9.4), but solutions of (9.4) perhaps are not solutions of (9.3). If there is no solution for (9.4) then there is no solution for the associated problem (9.3). Pick one solution of (9.4) and check if it satisfies the two inequalities of (9.1) and (9.2), if it does then we have found a solution of the associated problem (9.3), otherwise try another solution of (9.4).

We can also try to find the optimal solution of (9.4) then check if it satisfies the two inequalities of (9.1) and (9.2). If it can not fit the two inequalities, observe the frequency range where the misfit occurs, and revise \( t_1(s) \) and \( t_2(s) \) according to the observed information, then try to find the optimal solution of (9.4) again with new \( t_1(s) \) and \( t_2(s) \). This method resembles that of Makila, Westerlad, and Toivonen [55] to solve the similar problem for constrained LQJ problem.
BIBLIOGRAPHY


19. V.M. Adamjan, D.Z. Arov, and M.G. Krein, "Infinite Hankel Block Matrices and Related Extension problems", *AMS Transal.* (2) 111, 1978, 133-156


42. L. Silverman and M. Bettayeb, "Optimal Approximation of Linear Systems". **FACC.** 1980. FA8-A.


51. D. C. Youla and N. N. Kazanjan, "Bauer-Type Factorization of Positive Matrices and the theory of Matrix Polynomial Orthogonal on the


58. F. R. Gantmacher, "The Theory of Matrices", *vol I.*