BOUNDED FEEDBACK AND STRUCTURAL ISSUES IN LINEAR MULTIVARIABLE SYSTEMS

BY

XU YI SHAN

THESIS SUBMITTED FOR THE

DEGREE OF

DOCTOR OF PHILOSOPHY

IN

CONTROL ENGINEERING CENTRE,
DEPARTMENT OF ELECTRICAL, ELECTRONIC AND
INFORMATION ENGINEERING,
CITY UNIVERSITY,
LONDON EC1V OHB

July 13, 1992

ACKNOWLEDGEMENTS

I would like to thank sincerely Dr. Nicos Karcanias, my supervisor, for his patience, constant encouragement and valuable advice throughout the past four years.

Also I would like to thank the research students in the Control Engineering Centre for many interesting discussions. In addition, thanks goes to Robert Oldaker, who spent much of his precious time reading through my thesis and making some worthwhile comments. I am also indebted to Mrs. J. Rivellini, secretary of the Control Engineering Centre, for her invaluable help.

The financial support from the Chinese government and the British Council is gratefully acknowledged, without which my study in the U.K. would have been impossible.

Finally I would like to thank my parents for their constant support in the past years.

DECLARATION

The author grants powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to him. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

ABSTRACT

In this thesis, two main issues are addressed: Bounded state feedback and evaluation of structural characteristics of large scale systems with ill-defined models.

The problems related to bounded state feedback are the closed-loop eigenvalue mobility and stabilisability of an unstable system when subject to bounded state feedback. Also related are the problems of developing measures for quantitative controllability and measures for the distance of an unstable polynomial from stability as well as the root distribution of summation of polynomials. The mobility of the closed-loop eigenvalues and the stabilisability of unstable systems are studied via the investigation of root distribution of bounded coefficient polynomials. In this thesis, direct and inverse root inclusion problems are defined and results are obtained for different class of polynomials. Then the bound on the state feedback gain is transformed into the bound on the coefficients of the closed-loop characteristic polynomials and then necessary and sufficient conditions for closed-loop eigenvalue mobility and stabilisability are derived.

The problems of evaluating the structural characteristics of large scale systems with ill-defined mathematical models are also studied. The working model characteristics and the desirable features of control theory concerning the design of large scale processes with ill-defined models have first been discussed. Next, the useful indicators for integral stabilisability and integral controllability based on the steady-state gain information are discussed. Large scale systems with structural models are then introduced and the concepts of structural McMillan degree, poles and zeros both finite and at infinity are defined. The evaluation of the structural McMillan degree, the zeros and poles both finite and at infinity are translated into finding the paths of minimum or maximum weight of integer matrices. Algorithms are also proposed and assessed.

MATHEMATICAL NOTATIONS AND SYMBOLS

— R, C: fields of real, complex numbers.

— $R(s)$: field of rational functions in the variable s with real coefficients.

— $R[s]$: ring of polynomials in s with real coefficients.

— $R_{pr}(s)$: ring of proper rational functions.

— $\mathcal{F}$: denotes a general field, or ring.

— $\mathcal{F}^{p\times k}$: set of matrices with $p \times k$ dimensions and elements over $\mathcal{F}$, thus $R^{p\times k}(s)$, $R^{p\times k}[s], \cdots$ denote the corresponding set of matrices with elements over $R(s), R[s], \cdots$

— $R_{\mathcal{P}}(s)$: ring of proper rational function with have no poles in a symmetric set of the complex plane Ω, which excludes at least one point of the real axis.

— $\mathcal{V}$: denotes a finite dimensional vector space over some field $\mathcal{F}$; usual cases are the real vector spaces (R-vector spaces), rational vector spaces ($R(s)$-vector spaces).

— $\dim\mathcal{V}$: denotes the dimension of a vector space.

— $\mathcal{F}^n$: set of all n-dimensional vectors (n-tuples) of elements of $\mathcal{F}$.

$R^n, C^n, R^n(s), \cdots$: n-dimensional vector spaces over $\mathcal{F}$.

— If $\mathcal{V}$ is a subspace of $R^n, (R^n(s))$, the $\underline{v} \in \mathcal{V}$ denotes a vector of $R^n(R^n(s))$ that belongs to $\mathcal{V}$. If $\dim\mathcal{V} = d$ and $\{\underline{v}_1, \cdots, \underline{v}_d\}$ is a basis of $\mathcal{V}$, then $V = [\underline{v}_1, \cdots, \underline{v}_d] \in R^{n\times d}$ denotes a basis matrix of $\mathcal{V}$.

— If $H \in \mathcal{F}^{p\times k}, \mathcal{F}$ a field, then $\rho_{\mathcal{F}}(H)$ denotes the rank of H over $\mathcal{F}$, $\mathcal{N}_r\{H\}$ the right null space and $\mathcal{N}_l(H)$ the left null space of H.

— $H \in \mathcal{F}^{p\times p}, |H|$ denotes the determinant of H.

— State space description:

$$S(A,B,C,D): \quad \begin{cases} \dot{\underline{x}} = A\underline{x} + B\underline{u}, & A \in R^{n\times n}, B \in R^{n\times l}, C \in R^{m\times n}, D \in R^{m\times l} \\ \underline{y} = C\underline{x} + D\underline{u} & \underline{x} \in R^n, \underline{u} \in R^l, \underline{y} \in R^m \end{cases}$$

Assumptions: $\rho(B) = l, \rho(C) = m$

N: left annihilator of B, $(NB = 0, \rho(N) = n - l, N \in R^{(n-l)\times n})$

$B^{\dagger}$: left inverse of B $(B^{\dagger}B = I_l, \rho(B^{+}) = l, B^{\dagger} \in R^{l\times n})$

M: right annihilator of C, $(CM = 0, \rho(M) = n - m, M \in R^{n\times(n-m)})$

$L \in R^{l\times n}$: state feedback.

$Q \in R^{m\times n}$: output injection.

$F \in R^{l\times m}$: output feedback, $K \in R^{l\times m}$: squaring down.

$T \in R^{n\times n}, |T| \neq 0$: state coordinate transformation.

$R \in R^{l \times l}, |R| \neq 0$: input coordinate transformation.
$P \in R^{m \times m} |P| \neq 0$: output coordinate transformation.
$\sigma(A)$: spectrum of A (eigenvalues of A including multiplicities).
λ: eigenvalue of A.
$\underline{u}_\lambda$: eigenvector of A for λ eigenvalue.
$J(A)$: Jordan canonical block of A.
$A = UJ(A)V$: Jordan decomposition of A.
$\mathcal{G}_{oc}$: Output controllability Grammian.
$\mathcal{G}_{sc}$: State controllability Grammian.
$\mathcal{G}_{so}$: State observability Grammian.
$P_c(s)$: Controllability pencil.
$P_o(s)$: Observability pencil.
I_c, I_o Controllability and observability indices.
Q_c: Controllability matrix.
Q_o: Observability matrix.
Q_{oc}: Output controllability matrix.
$\mathcal{R}_c(s)$: Restricted controllability pencil.
$\mathcal{R}_o(s)$: Restricted observability pencil.
c.m.i.: column minimal indices.
r.m.i.: row minimal indices.

— Transfer function description

$G(s) = C(sI - A)^{-1}B + D \in R^{m \times l}(s)$: transfer function matrix.
$r = \rho_{R(s)}\{G(s)\}$: normal rank of $G(s)$.
$t(s) \in R[s], \partial[t]$: degree of $t(s)$.

— $G(s) = N_r(s)D_r(s)^{-1} = D_l(s)^{-1}N_l(s)$.

$N_r(s) \in R^{m \times l}[s], D_r(s) \in R^{l \times l}[s]$: Right Matrix Fraction Description (R.M.F.D.).

$N_l(s) \in R^{m \times l}[s], D_l(s) \in R^{m \times m}[s]$: Left Matrix Fraction Description (L.M.F.D.).

— For $T(s) \in R^{p \times k}[s]$:

$$S(T) = \left[\begin{array}{ccc|c} f_1(s) & & & 0 \\ & \ddots & & \\ & & f_r(s) & \\ \hline & 0 & & 0 \end{array}\right] \qquad r = \rho_{R(s)}\{T(s)\}$$

$f_i(s)$: invariant polynomials of $T(s)$.
$S(T)$: Smith normal form of $T(s)$.

— For $T(s) \in R^{p\times k}(s)$:

$M(T)$: Smith-McMillan form of $T(s)$.

$$M(T) = \left[\begin{array}{ccc|c} \frac{\epsilon_1(s)}{\psi_1(s)} & & & 0 \\ & \ddots & & \\ & & \frac{\epsilon_r(s)}{\psi_r(s)} & \\ \hline & 0 & & 0 \end{array}\right] \qquad r = \rho_{R(s)}\{T(s)\}$$

$\epsilon_i(s)$: invariant zero polynomials, $\epsilon_1(s)/\epsilon_2(s)/\cdots/\epsilon_r(s)$.

$\psi_i(s)$: invariant pole polynomials, $\psi_r(s)/\psi_{r-1}(s)/\cdots/\psi_1(s)$.
(/) divides

$p(s) = \Pi_{i=1}^{r}\psi_i(s)$: pole polynomial of $P(s)$.

$z(s) = \Pi_{i=1}^{r}\epsilon_i(s)$: zero polynomial of $T(s)$.

$\delta_M(T) = \partial[P]$: McMillan degree of $T(s)$.

— $R_u^{m\times m}[s]$: set of $m\times m$ $R[s]$-unimodular matrices.

— $R_{bpr}^{m\times m}(s)$: Set of $m\times m$ $R_{pr}(s)$-unimodular matrices, biproper.
($U(s) \in R_{pr}^{m\times m}(s)$, then $U^{-1}(s) \in R_{pr}^{m\times m}(s)$).

— $T(s) \in R^{p\times k}[s]$: $T(s) = s^d T_d + \cdots + sT_1 + T_0, T_i \in R^{p\times k}, T_d \neq 0$,

$d =$deg$(T(s)) \equiv \partial_s[T]$: scalar degree of $T(s)$

$\delta = \partial_m[T]$: matrix degree of $T(s)$ (maximal degree amongst the maximal order minors of $T(s)$).

— $t(s) = n(s)/d(s) \in R(s)$

$\delta_\infty(t) = \partial[d] - \partial[n]$: valuation at infinity of $t(s)$.

— $T(s) \in R^{p\times k}(s)$,

$M_\infty(T)$: Smith-McMillan form at infinity of $T(s)$.

$$M_\infty(T) = \left[\begin{array}{ccc|c} s^{q_1} & & & 0 \\ & \ddots & & \\ & & s^{q_r} & \\ \hline & 0 & & 0 \end{array}\right] \qquad r = \rho_{R(s)}\{T(s)\}, q_1 \geq q_2 \geq \cdots \geq q_r$$

$q_i > 0$: orders of infinite poles, $q_i < 0$.

$|q_i|$: orders of infinite zeros.

$\delta_M^\infty(T) = \sum q_i$:

q_i: McMillan degree at infinity of $T(s)$.

$\nu(T) = \delta_M(T) + \delta_M^\infty(T)$: extended McMillan degree of $T(s)$.

— $G \in C^{m\times l}, G = Y\sum U^*$: singular value decomposition (SVD)
$Y \in C^{m\times m}, U \in C^{l\times l}$ unitary matrices.

$\sum =$ p-diag$\{\sigma_1, \cdots, \sigma_r\} \in R^{m\times l}, r = \min(m, l), \sigma_1 \geq \cdots \geq \sigma_r$.

p-diag: (pseudo-diagonal) $\left[\begin{array}{cccc} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \\ \hline & & 0 & \end{array}\right]$, or $\left[\begin{array}{cccc|c} \sigma_1 & & & & \\ & \sigma_2 & & & 0 \\ & & \ddots & & \\ & & & \sigma_r & \end{array}\right]$

$\sum(G)$: set of singular values of G.

$\bar{\sigma}(G)$: maximal singular value.

$\underline{\sigma}(G)$: minimal singular value.

Columns of Y, U: left, right singular vectors of G.

$A \in \mathcal{F}^{P\times K}$:

$$A = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_t \end{bmatrix}, A_i \in \mathcal{F}^{P_i\times K_i}, A = \text{b-diag}\{A_1, \cdots, A_t\}(\text{block-diagonal})$$

$A = \text{diag}\{A_1, \cdots, A_t\}$, if $A_i \in \mathcal{F}^{P_i\times P_i}$

— Polynomials

$P[s]$: Monic polynomials in $R[s]$ with fixed degree n.

$f(s) \in P[s]$: $f(s) = s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n$

$\underline{\alpha}_f$: coefficient vector of $f(s) \in P[s]$.

$\|\underline{\alpha}_f\|$: l_2 norm of the coefficient vector $\underline{\alpha}_f$.

$\|\underline{\alpha}_{[\bullet]}\|$: l_2 norm of the coefficient vector $\underline{\alpha}_{[\bullet]}$.

$P_r(s)$: Rational functions in s.

$P_{pr}(s)$: Proper rational functions in s.

$P^+[s]$: $f(s) \in P[s]$ with all the roots in the left half plane.

$P^-[s]$: $f(s) \in P[s]$ with all the roots in the right half plane.

$P^{+,\gamma}[s]$: $f(s) \in P^+[s]$ and the coefficients are bounded by $\|\underline{\alpha}_f\|_2 \leq \gamma$.

$P^{-,\gamma}[s]$: $f(s) \in P^-[s]$ and the coefficients are bounded by $\|\underline{\alpha}_f\|_2 \leq \gamma$.

$P^-[s]$: $f(s) \in P[s]$ with all the roots in the right half plane.

$P^\gamma[s]$: $f(s) \in P[s]$ and $|\underline{\alpha}_f| \leq \gamma$.

$\| \bullet \|_2$ l_2-norm

$\| \bullet \|_\infty$ l_∞-norm

$\gg$ greater greater than

$\tilde{n}$ Integer set $\{1, 2, \cdots, n\}$

$\sum$ Sum of

$\equiv$ equivalent

$\sim$ equal modulo

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

The edifice of control theory consists of two main parts, system analysis and system design. Modelling of systems, however, forms the pillar of the edifice. System analysis is to investigate the interaction between the system and the embedding environment, as well as the internal behaviour of systems, in relation to the system structure and the system parameters. The behaviour of systems is expressed in terms of system properties and the possession of the properties by a particular system or a class of systems is expressed through indicators. Another aspect of system analysis is to establish the interelations between the systems belonging to the same class. It was found that some of the system properties may remain unchanged while the systems are subject to generalised transformations which include state transformation, state feedback, output feedback or output injection, etc. The properties which remain unchanged are the invariants of the system under specific transformations. System properties, property indicators and system invariants facilitate the classification of systems into different categories. The results of system analysis also provide the tools for system design.

The most important system property is that of system stability. Closely related concepts are system controllability, stabilisability, observability and detectability. If a system is unstable, one naturally asks whether it is stabilisable. System controllability and stabilisability are equivalent if a system is minimal. It is a well-known result that a system can be stabilised by using state feedback and the closed-loop poles of a system can be arbitrarily assigned as long as the system is controllable [Won.,2] and that a system is stabilisable if all the controllable modes are controllable [Dav.,1]. However, this is true only with the important underlying assumption that the gains of the controller are by no means restricted. This important assumption unfortunately often breaks down in reality, since due to nonlinearity, or saturation reasons gains are usually restricted. In fact, the controller gains and power are always restricted. Because controllability of systems is treated almost exclusively as a qualitative property, by examining the corresponding indicators, one can only draw conclusions on whether it is possible to stabilise the system by using state feedback or not but one fails to give answers to further questions such as the degree of controllability of a system, or how far can one move the system eigenvalues when subject to bounded state feedback. In this context, further quantitative measures for system controllability have to be developed. Measures for quantitative controllability are also needed in the study of transferring the states of a system to a certain distance when the control signals are bounded by energy. In the last decade, quantitative controllability has been studied [Eis.,1] [Bol. & Lu,1] [Tar.,1] in relation to the controller gain, the distance to uncontrollability, etc. However, no explicit result has been obtained as far as pole mobility of systems under bounded state feedback is concerned.

In the vast literature of control theory, the majority of the results are concerned with linear time invariant systems whose mathematical descriptions either in the frequency domain or in the state-space domain are assumed to be known exactly. However, the need for the control theory to cope with systems with uncertain mathematical descriptions arises from real applications [Doy. & Ste.,1] [Lin,1] [Rei.,1] [Mor. & Ste.,1] [Hin. & Pri.,1] [Kar.,1] etc, not only because exact modelling of complex systems is impossible, but also due to the fact that the systems and the embedding environment are subject to constant change; thus, even if exact modelling were possible at one stage, the model can not represent the system all the time faithfully. Of course, errors may be introduced deliberately into the model in order to simplify the analysis by ignoring certain system dynamics such as the very high frequency response, or to linearise the problem, etc. Further, one would like to take a structural model in order to accommodate not just one system but a whole family of systems. For instance, this is particularly important in designing a controller for a system which is required to operate at different operating points. Due to the diverse sources of modelling errors, the errors of system models can be classified into two categories: structural or parametric and both have important physical interpretations. To work with systems whose models are not exact, provides control theory with fresh challenges.

In this work, the following two main issues have been addressed.

- Bounded state feedback;
- Structural evaluation of large scale (dimension) generic transfer functions.

With respect to the first issue, pole mobility and stabilisability of systems under bounded state feedback is examined. Based on the fact that a link between the bound on the state feedback gain and the bound on the coefficient of the characteristic polynomial can be established [Kar. & Shan,2] [Shan & Kar.,1], the problem of eigenvalue mobility and stabilisability of systems when subject to bounded state feedback can be studied via the root distribution of bounded coefficient polynomials. The problem of establishing relationships between the coefficients and the roots of a polynomial has been an important issue for the last two centuries and an excellent account of the classical results is given in [Mar.,1]. One of the key problems in the geometry of zeros of polynomials is the establishment of regions containing all roots of a given polynomial. The classical results provide different upper bounds for the region containing all roots of a bounded coefficient polynomial; however, such bounds are rather weak. Two important problems are formulated in this thesis, which have not been addressed before: the first is the definition of the least region in the complex plane which contains all roots of bounded norm polynomials, and the second is the definition of the maximal region of the complex plane such that, if a polynomial has its roots in it, then its coefficient vector is bounded. For the case of stable polynomials the first problem is solved, whereas the second is solved for any type of polynomials [Kar. & Shan,1].

The classical and the newly derived results on the properties of bounded norm polynomials allow the derivation of criteria characterising the mobility of closed-loop poles under bounded norm state feedback. This mobility is characterised by the size of the region the poles move. It is shown that the degree of controllability of the given system is an important parameter in the characterisation of the size of the pole mobility region; in fact, the closer to uncontrollability the system is, the smaller the size of the pole mobility region under bounded state feedback. The problem of pole mobility is closely related to stabilisability under bounded feedback. It is shown that controllability alone is not adequate for stabilisability under bounded state feedback. Different criteria for stabilisability are derived using the results on bounded coefficient polynomials. The problem of stabilisability under bounded feedback is in a sense dual to that of robustness under certainty [Hin. & Pri.,1] [Hin. & Pri.,2]; however, the tools for solving these two problems are different. The approach adopted in this thesis is based on the idea of finding estimates for the distance of an unstable polynomial from the set of stable, Hurwitz polynomials. An implicit assumption throughout all of the work for the bounded state feedback is that the state variables are physical variables and thus it makes sense to define performance constraints on them, as well as impose constraints on the norm of the state feedback which is used. In this sense, the degree of the controllability which affects the pole mobility is always related to the given physical variables coordinate system.

For the second main issue, the objective is to develop tools for diagnosis of generic system properties based on simple models of systems. It has been argued by many researchers [Lin,1] [Ros.,1] [Rei.,1] [Mor. & Ste.,1] that structural information of systems are vital in system design. It has been pointed out especially in [Kar.,1] that at the early design stages of a process, it would be highly desirable for the control theory to get involved in order to result in a design which will lead to quality controller design. However, control theory faces a challenge in working on ill-defined models. The evaluation of the potential from the control viewpoint of a given system depends on the nature, as well as the values of structural characteristics, known as systems invariants. System invariants affect the shape and values of control design indicators and also enter into the solvability conditions characterising the solvability of control synthesis problems. The need to introduce control theory tools into the "Early Process Design" [Kar.,2], that is to use control theory in evaluating alternative process flowsheets, determine procedures for selection of measurement and actuation schemes, for large dimension problems, implies that the control theory tools have to be adjusted for an uncertain models environment. The uncertain models in Early Process Design are characterised by uncertainty in the parameters, as well as dynamic complexity. For such large dimension uncertain models, an issue that arises is the evaluation of the generic values of certain structural characteristics. This problem is referred to as "structural identification" and aims at producing algorithms for evaluating the generic value of different types of invariants, without using exact methodologies, which due to large dimensions may not be suitable. One of the most important concepts that enters all generic solvability conditions of control problems (see [Byrn.,1], etc) is that of the McMillan degree. For large dimension transfer function matrices, we consider here the problem of evaluating the generic McMillan degree, as well as the structure of generic infinite zeros. These are two of the many problems which may be addressed within the general framework of "structural identification" issues. These tools would help in screening and selecting various designs at early stages. A topic that is closely related to the evaluation of "Early Process Design" is that of inferring properties characterising control quality, of a dynamic model, from simpler models, such as the steady state models. In fact, this issue has to do with determining the range of predictability of a property starting from simple and progressively going to more and more complex models. Issues related to the prediction of dynamic performance properties from simple type models and especially, steady state models, belong to the overall problem of assessing Early Process Schemes with control theory tools and are also examined in this thesis.

Control systems design is based around a system of design indicators and sets of system invariants. The link of system invariants to indicators is not always clear; however, there are strong indications that system invariants define, in a way, the limits of what can be achieved by compensation. An effort to establish the links between system invariants and indicators runs through the spirit of the present work, and an attempt to enlarge the set of performance indicators is also made. The main theme of this thesis is to investigate a number of system properties under bounded feedback and link such properties to design indicators on one hand and on the other hand to investigate issues related to evaluation of structural characteristics of large scale uncertain models arising in Early Process design studies. The thesis is structured as follows:

In Chapter 2, some of the fundamental concepts both in mathematics and systems theory are reviewed. These will provide the necessary concepts and tools used in the later developments. It is then followed by a description of the most important system properties, property indicators and system invariants under different transformations both in the frequency and state space domain.

In Chapter 3, the quantitative aspect of the fundamental properties, such as controllability, observability, output disturbability, will be examined. Quantitative measures will be provided for controllability and observability. The measures are based on the singular values of the appropriately defined Grammians. The quantitative controllability provide answers to problems such as how far can one move the state when subject to bounded energy control. Indeed the reachable states from zero initial state of a controllable system, when subject to bounded energy control, can be parametrised by using the singular values of the state controllability Grammian. The quantitative output controllability of the system then will be further developed to provide indicators for input-output interaction which gives vital information in selecting control structure. Finally, more recently developed quantitative measures of controllability in connection to pole mobility of bounded feedback control are surveyed.

In Chapter 4, classical results concerning the root distribution of polynomials in relation to the coefficients will be reviewed. The recent developments in this area initiated by Kharitonov's work together with the later results along the same line will also be examined.

In the first part of Chapter 5, the root distribution of l_2 norm bounded coefficient polynomials will be investigated. Two main problems concerning the root distribution of bounded coefficient polynomials will be looked into. Those are the root inclusion and inverse root inclusion problems. In examining these problems, the polynomials will be further classified into subfamilies. For stable and totally unstable polynomials, the minimum bound has been obtained for the root inclusion problem, while minimum rectangular regions have been established for the inverse root inclusion problem. For the general case, general upper bound for the root inclusion problem is given by employing the classic results reviewed in Chapter 4. In the second part of Chapter 5 the root distribution of the summation of polynomials will be investigated. Some interesting results have been obtained for special polynomials.

In Chapter 6, the problem of bounded state feedback will be investigated. It is first translated into the root distribution problem of bounded coefficient polynomials. The results developed in Chapter 4 and Chapter 5 will then be deployed to develop regions for the closed-loop poles. Effects of the controllability matrix on the closed-loop pole mobility will be discussed. Both SISO and MIMO systems will be investigated.

In Chapter 7, some issues arising in the evaluation of Early Process Models with structural criteria will first be discussed. The issues include the model environment of early design stages and the desirable properties of control theory for this type of problem will be outlined. Next, steady-state models will be deployed to develop informative indicators for closed-loop system properties such as closed-loop stabilisability, controllability as well as robustness. Then the concepts of generic structured transfer matrices, generic McMillan degree, finite and infinite poles and zeros will be defined and algorithms for evaluating the generic McMillan degree, finite and infinite poles and zeros will be given and assessed. These results aim at providing a fast assessment for screening designs at early stages when the models are ill-defined, by avoiding the existing algorithms which may be difficult to use for large scale type problems.

Chapter 2

SYSTEM PROPERTIES, PROPERTY INDICATORS AND INVARIANTS

2.1 Introduction

System analysis is the study of system behaviour together with the interactions between the systems and their embedding environments. System behaviour can be characterised by a set of system properties, the possession of which by the systems can be expressed in terms of the values and graphs of corresponding indicators. The properties of systems are determined by the structure and the values of the parameters of the systems. For instance, the stability of the system depends on the eigenvalues of the system which are purely affected by the system matrix A while the controllability of a system is determined by both the system matrix A and the input matrix B which represents the connection with its environment. The definition of properties and their dependence upon the system structure and the values of the parameters is the first task of system analysis.

The system properties can be classified into different categories depending upon the types of models employed, the nature of the system indicators, whether or not the indicator can be used for different properties, etc..

The study of the relationship between the systems in terms of structure, parameters and their properties is important. The results lay down the basis for system classification and parametrisation. System classification in terms of system structure or system properties is extremely important in the sense that a characteristic system from a group of systems can be studied in great detail and the results can be readily applied to the other systems in the same group. For instance, the pole assignability of the whole group of controllable systems via state feedback control demonstrates this clearly. In this chapter, some available results of system classification in terms of system properties and structures are presented.

System analysis aims at producing results which will set the guidelines for system design and compensation. In order to shape the system properties, different compensation schemes can be used. The system properties have been studied when the systems are subject to compensations. The structure and the parameters of the compensator affect some of the system properties while the others may remain intact. Those properties that remain unchanged under a specific compensation scheme will be called invariants of the particular compensation scheme. Because the system invariants remain unchanged under a certain type of compensation, they can be used to characterise not only a single system but a whole family of systems and in turn they characterise the limits in the shaping of the property indicators when subject to this particular compensation.

The canonical form of systems is a concept which is closely related to the system invariants. Because an invariant represents a group of systems, a canonical element from the group stands out as the representative of the whole group and therefore the study of the whole group can be carried out by using this element.

The study of system properties, property indicators and system invariants has been carried out for both linear and nonlinear system. Results for linear systems have been well established while such results for nonlinear systems are still in the early stages of development. Only linear systems are considered in this chapter.

The chapter is structured as follows: Essential background mathematics are presented in section (2.2). In section (2.3), definition and classification of system properties, property indicators and system invariants are given. In section (2.4), open loop system properties and property indicators are examined whereas the closed loop system properties and property indicators are discussed in section (2.5). System invariants are reviewed in section (2.6). Finally the conclusion is given in section (2.7).

2.2 Background mathematics

In this section, some essential background mathematics are presented. Mathematics is extensively deployed in control theory. From basic definitions to system representation, to system analysis and finally to system design, at each stage, the involvement of mathematics is inextricable. In fact, mathematics facilitate the development of control theory from the very beginning. It is also true that almost every branch of mathematics can find its application in control theory. The mathematics reviewed here does not serve to be a comprehensive survey, it covers those topics which are essential for the coming sections. These are: basic definitions, polynomial matrices, transformations and matrix pencils.

2.2.1 Rings and fields

A *ring* $(R, +, \cdot)$ is a set R, together with two binary operations $+$ and $\cdot$ on R satisfying the following axioms. For any element $a, b, c \in R$

- $(a + b) + c = a + (b + c)$
- $(a + b) = (b + a)$
- there exists $0 \in R$ called the zero, such that $a + 0 = a$
- there exists $(-a) \in R$ such that $a + (-a) = 0$
- $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
- there exists $1 \in R$ such that $1 \cdot a = a \cdot 1 = a$
- $a \cdot (b + c) = a \cdot b + a \cdot c$ and $(a + b) \cdot c = a \cdot c + b \cdot c$

  The ring $(R, +, \cdot)$ is called a *commutative ring* if
- $a \cdot b = b \cdot a$ for all $a, b \in R$

The *units* of a ring are the elements whose multiplicative inverses are also in the ring.

A *field* is a ring in which every nonzero element has a multiplicative inverse, i.e. a field is a nontrivial commutative ring R satisfying the following extra axiom

- for each nonzero element $a \in R$ there exists $a^{-1} \in R$ such that $a \cdot a^{-1} = 1$

The set of real numbers, R, and the set of complex numbers, C, are rings which are often used in control theory.

The set of all polynomials in s with coefficients from F ($F = R$ or C) is denoted as $F[s]$. That is

$$F[s] = \{a_0 s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n : a_i \in F\} \tag{2.1}$$

then the set of polynomials forms a ring $(F[s], +, \cdot)$.

The set of rational functions in s with coefficients from F ($F = R$ or C) is denoted as $F(s)$. That is

$$F(s) = \{n(s)/d(s) : n(s), d(s) \in F[s]\} \tag{2.2}$$

$(F(s), +, \cdot)$ is a field.

The set of proper rational functions is denoted by $F_{pr}(s)$ ($F = R$ or C), or

$$F_{pr}(s) = \{n(s)/d(s) : n(s), d(s) \in F[s], d(s) \not\equiv 0, \deg(d(s)) = \deg(n(s))\} \tag{2.3}$$

where $\deg(d(s))$ and $\deg(n(s))$ are the degrees of the polynomials $d(s)$ and $n(s)$. $(F_{pr}(s), +, \cdot)$ is a ring.

2.2.2 Polynomial matrices and Smith forms [Gan.,1]

A polynomial matrix $P(s) \in R^{m \times l}[s]$ is a rectangular matrix whose elements are polynomials is s

$$P(s) = \begin{bmatrix} p_{11}(s) & \cdots & p_{1l}(s) \\ \vdots & & \vdots \\ p_{m1}(s) & \cdots & p_{ml}(s) \end{bmatrix} \tag{2.4}$$

where $p_{ij}(s) \in R[s]$.

A square polynomial matrix is *nonsingular* if $|P(s)| \not\equiv 0$; otherwise is called *singular*.

The elementary row and column operations on $P(s)$ are

(1). Multiplication of any row (column) by a number $c \neq 0$;

(2). Interchange of any two rows (column);

(3). Addition to any row (column) of any other row (column) multiplied by any arbitrary polynomial $b(s)$.

These elementary row (column) operations are equivalent to pre-multiplying (post-multiplying) the matrix $P(s)$ with a nonsingular polynomial matrix whose determinant is a constant. These polynomial matrices will be called *elementary matrices.*

Let $P(s)$ be a polynomial matrix of rank r, i.e., the matrix has minors of order r not identically zero, but all the minors of order greater than r are identically equal to zero. Elementary operations do not alter the rank of a polynomial matrix. Denote by $\Delta_i(s)$ the greatest common divisor of all the minors of order i in $P(s), (i = 1, 2, ..., r)$. Further

$$\lambda_i(s) = \frac{\Delta_i(s)}{\Delta_{i-1}(s)}, \qquad \Delta_0(s) = 1 \tag{2.5}$$

The polynomials $\lambda_1(s), \lambda_2(s), ..., \lambda_r(s)$ are called *invariant polynomials* of $P(s)$.

Further if the invariant polynomials are decomposed into their irreducible factors

$$\begin{aligned}
\lambda_1(s) &= (s-\lambda_1)^{\eta_{11}}(s-\lambda_2)^{\eta_{12}}\cdots(s-\lambda_k)^{\eta_{1k}} \\
\lambda_2(s) &= (s-\lambda_1)^{\eta_{21}}(s-\lambda_2)^{\eta_{22}}\cdots(s-\lambda_k)^{\eta_{2k}} \\
&\vdots \\
\lambda_r(s) &= (s-\lambda_1)^{\eta_{r1}}(s-\lambda_2)^{\eta_{r2}}\cdots(s-\lambda_k)^{\eta_{rk}}
\end{aligned} \tag{2.6}$$

Then all the powers among $(s-\lambda_1)^{\eta_{11}}, \cdots, (s-\lambda_k)^{\eta_{rk}}$ are called the *elementary divisors* of the polynomial matrix $P(s)$.

The elementary row and column operations on $P(s)$ change neither the invariant polynomials, nor the elementary divisors. So by applying elementary row and column operations on $P(s)$, it can be brought to diagonal form

$$\left[\begin{array}{cccc|c}
\lambda_r(s) & & & 0 & \\
& \lambda_{r-1}(s) & & & 0_{r,l-r} \\
& & \ddots & & \\
0 & & & \lambda_1(s) & \\
\hline
& & 0_{m-r,r} & & 0_{m-r,l-r}
\end{array}\right] \tag{2.7}$$

and all the polynomial matrices which are attainable by applying elementary row and column operations can be transformed to the above form unique up to constant scale of the invariant polynomials; therefore it is referred to as the *Smith* canonical form.

Nonsingular polynomial matrices whose determinants are independent of the variable s are referred to as *unimodular.*

2.2.3 Matrix divisors, greatest common divisors and coprimeness [Kai.,1]

For two polynomial matrices $\{N(s), D(s)\}$ with the same number of columns, if there exist polynomial matrices $\bar{N}(s), \bar{D}(s)$ and $R(s)$ such that

$$N(s) = \bar{N}(s)R(s), \qquad D(s) = \bar{D}(s)R(s)$$

the $R(s)$ is called a *right common divisor* of $N(s)$ and $D(s)$. If for any right common divisor of $N(s)$ and $D(s)$, say, $R_1(s)$, there exists a polynomial matrix $W(s)$ such that $R(s) = W(s)R_1(s)$, then $R(s)$ is a *greatest common right divisor* (gcrd) of $N(s)$ and $D(s)$. Greatest common right divisors are not unique, but they can differ by only unimodular factors.

Two polynomial matrices are *right coprime* if their gcrds are unimodular.

Left common divisors, greatest common left divisors and left coprimeness can be defined accordingly as above.

2.2.4 Transformations of matrices and Jordan forms [Gan.,1]

Two matrices $A, B \in F^{n\times n}$ ($F = R$ or C) are similar if there exists a nonsingular matrix T such that

$$B = T^{-1}AT \tag{2.8}$$

Two matrices are similar if and only if their characteristic matrices defined as $sI - A$ and $sI - B$ have the same invariant polynomials and therefore the same elementary divisors.

Assume that the elementary divisors are $(s-\lambda_1)^{\eta_1}, (s-\lambda_2)^{\eta_2}, \ldots, (s-\lambda_r)^{\eta_r}$ where λ_i are not necessarily distinct. η_i will be referred to as *algebraic multiplicity* of λ_i. If some of the elementary divisors have the same value for λ_i, then the total number of the elementary divisors will be referred to as the *geometric multiplicity* of λ_i. By applying similarity transformations, all similar matrices can be transformed to the Jordan block diagonal form $F = \text{diag}\,\{F_1, F_2, \cdots, F_r\}$ where

$$F_i = \left.\begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix}\right\} \eta_i$$

2.2.5 Matrix pencil and Kronecker canonical form [Gan.,1]

Matrix polynomials of the form $F + sG$ are defined as *matrix pencils* where $F, G \in R^{l\times n}$ (or $C^{l\times n}$). The pencil $F+sG$ is *regular* if $F, G \in R^{n\times n}$ (or $C^{n\times n}$) and $|F+sG| \neq 0$; otherwise it is called *singular*. Two matrix pencils $F_1 + sG_1$ and $F_2 + sG_2$ are defined to be strictly equivalent if there exist constant matrices P and Q such that

$$P(F_1 + sG_1)Q = F_2 + sG_2 \tag{2.9}$$

For regular matrix pencils, define $\tilde{s}F + sG$. The invariant polynomials of this polynomial matrix can be obtained as

$$\lambda_1(s,\tilde{s}) = \frac{\Delta_n(s,\tilde{s})}{\Delta_{n-1}(s,\tilde{s})}, \lambda_2(s,\tilde{s}) = \frac{\Delta_{n-1}(s,\tilde{s})}{\Delta_{n-2}(s,\tilde{s})}, ..., \lambda_n(s,\tilde{s}) = \Delta_1(s,\tilde{s}) \tag{2.10}$$

where $\Delta_i(s,\tilde{s}), i = 1, 2, ..., n$ are the greatest common divisors of all the $i \times i$ minors of $\tilde{s}F + sG$, and $\lambda_i(s,\tilde{s})$ and $\Delta_i(s,\tilde{s})$ are homogeneous polynomials in s and $\tilde{s}$. Then the elementary divisors of the form $\tilde{s}^q$ will be referred to as the *infinite elementary divisors* while the rest are referred to as *finite elementary divisors*. Two regular matrix pencils are equivalent if and only if they both share the same infinite and finite elementary divisors.

For singular matrix pencils, there exist nontrivial solutions $\underline{x}, \underline{y}$ to the following equations

$$(F + sG)\underline{x} = 0 \text{ and } (F^T + sG^T)\underline{y} = 0 \tag{2.11}$$

In the solution spaces $\underline{x}^0$ and $\underline{y}^0$, there exist linearly independent bases with minimum degrees in s, say, $\underline{x}_1(s), \underline{x}_2(s), ..., \underline{x}_p(s)$ and $\underline{y}_1(s), \underline{y}_2(s), ..., \underline{y}_q(s)$ with degrees $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_p$ and $\nu_1 \leq \nu_2 \leq \cdots \leq \nu_q$, respectively. Then the set $\{\mu_1, \mu_2, \cdots, \mu_p\}$ and $\{\nu_1, \nu_2, \cdots, \nu_q\}$ will be referred to as the *right Kronecker indices* and *left Kronecker indices.*

Two matrix pencils are strictly equivalent if and only if they have the same finite and infinite elementary divisors as well as the same left and right Kronecker indices.

Strictly equivalent matrix pencils can be transformed by strictly equivalent transformations into canonical Kronecker form

$$P(F + sG)Q = \text{block.diag } \{L_{\mu_1}, ..., L_{\mu_p}, \tilde{L}_{\mu_1}, ..., \tilde{L}_{\nu_q}, sJ - I, sI - F\} \tag{2.12}$$

where $\{F, J, \{L_{\nu_i}\}, \{\tilde{L}_{\nu_i}\}\}$ are unique matrices such that

- F is in Jordan form;

- J is a nilpotent Jordan matrix, i.e. a matrix in Jordan form with all zero eigenvalues;

- L_μ is a $\mu \times (\mu + 1)$ matrix of the form

$$\begin{bmatrix} s & -1 & 0 & \cdots & 0 & 0 \\ 0 & s & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & 0 & \cdots & s & -1 \end{bmatrix}$$

- $\widetilde{L}_\nu$ is a $(\nu + 1) \times \nu$ matrix with s along the diagonal and -1 along the first subdiagonal.

2.2.6 Rational matrix, Smith-McMillan form, McMillan degree and matrix fraction description

A rational matrix $T(s) \in R^{m \times l}(s)$ is a matrix whose entries are rational functions in s. Let $d(s)$ be the monic least common multiple of the denominators of all the entries of $T(s)$; then $T(s)$ can be written as

$$T(s) = N(s)/d(s) \tag{2.13}$$

where $N(s)$ is a polynomial matrix. By applying elementary column and row operations on $N(s)$, it can be transformed into the Smith form, $N(s) = L(s)S(s)Q(s)$, where $S(s)$ is in Smith form, $L(s)$ and $Q(s)$ are unimodular matrices. The matrix $T(s)$ can be transformed into the *Smith-McMillan form*

$$M(s) = L(s) \left[\begin{array}{cccc|ccc} \frac{\epsilon_1(s)}{\psi_1(s)} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \frac{\epsilon_2(s)}{\psi_2(s)} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \frac{\epsilon_r(s)}{\psi_r(s)} & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{array} \right] Q(s) \tag{2.14}$$

where r is the normal rank of $T(s)$, $\frac{\epsilon_i(s)}{\psi_i(s)}, i = 1, 2, ..., r$ are coprime.

The *McMillan degree* of the rational matrix $T(s) \in R^{m \times l}(s)$ is the sum of the orders of $\psi_i(s)$, $\sum_{i=1}^{r} \deg \{\psi_i(s)\}$. The *poles* of $T(s)$ are the roots of the denominator polynomial $\psi_i(s), i = 1, 2, ..., r$ and the *zeros* of $T(s)$ are the roots of the numerator polynomial $\epsilon_i(s), i = 1, 2, ..., r$.

A rational matrix $G(s) \in R^{m\times l}(s)$, whose normal rank equals $\min\{m, l\}$, can be factorised factors as

$$G(s) = D_l^{-1}(s)N_l(s) = N_r(s)D_r^{-1}(s) \tag{2.15}$$

where $N_l(s), N_r(s) \in R^{m\times l}[s], D_l(s) \in R^{m\times m}[s], D_r(s) \in R^{l\times l}[s]$ with $\det D_l(s)$, $\det D_r(s) \neq 0$. Then $D_l^{-1}(s)N_l(s)$ and $N_r(s)D_r^{-1}(s)$ are called the *left matrix fraction description* and *right matrix fraction description* of the rational matrix $G(s)$, respectively.

2.3 Definition and classification of system properties and property indicators

System analysis is generally based on the mathematical models of the systems. The models can be given either in the frequency domain or in the time domain. In the time domain, they are usually a set of differential equations which have been derived from the physical laws underlying the systems, or from input-output identification and subsequent realisation. In general, the set of differential equations can be transformed into the state-space model description $S(A, B, C, D)$

$$S(A, B, C, D) : \begin{cases} \dot{\underline{x}}(t) = A\underline{x}(t) + B\underline{u}(t) \\ \underline{y}(t) = C\underline{x}(t) + D\underline{u}(t) \end{cases} \tag{2.16}$$

where A, B, C and D are constant matrices and $A \in R^{n\times n}$ is the state matrix, $B \in R^{n\times l}$ input matrix, $C \in R^{m\times n}$ output matrix, $D \in R^{m\times l}$ direct transfer matrix. $\underline{x} \in R^n$ is the state variable vector, $\underline{u} \in R^l$ the input variable vector and $\underline{y} \in R^m$ the output variable vector; it is also assumed that B and C are full rank matrices.

In the input-output sense, the system is represented by the frequency domain transfer function model

$$G(s) = C(sI - A)^{-1}B + D \tag{2.17}$$

which may also be represented by the right or left coprime MFD's [Cal. & Fra.,1] [Kai.,1]

$$G(s) = N_r(s)D_r^{-1}(s) = D_l^{-1}(s)N_l(s) \tag{2.18}$$

where $N_r(s), N_l(s) \in R^{m\times l}[s], D_r(s) \in R^{l\times l}[s]$, and $D_l(s) \in R^{m\times m}[s]$ and the coprimeness implies that the matrices

$$T_r(s) = \begin{bmatrix} N_r(s) \\ D_r(s) \end{bmatrix}, \qquad T_l(s) = \begin{bmatrix} N_l(s) & D_l(s) \end{bmatrix} \tag{2.19}$$

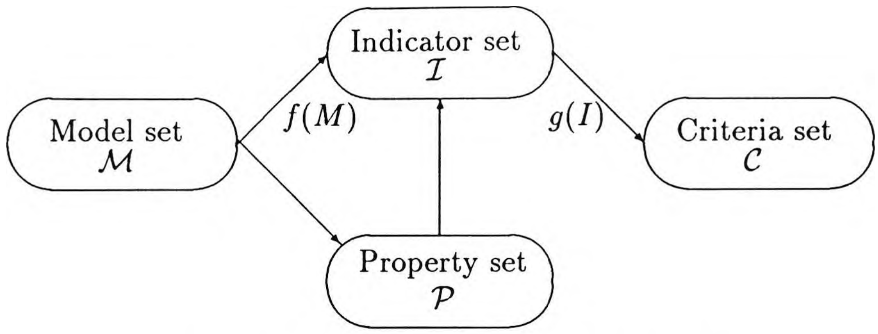


Figure 2.1: Model, property, indicator and criteria

have no zeros (full rank for all $s \in C$). For a given initial condition $\underline{x}(0)$ and an input $\underline{u}(t)$ to the system, the solution $\underline{x}(t)$ and $\underline{y}(t)$ to the equation (2.16) will be referred to as the state trajectory and the output trajectory, respectively.

A state-space model given in (2.16) provides information for both the input-output relation and the state variables while a transfer function model given in (2.17) describes input-output relations only. For this reason, the former will also be called an internal model while the latter will be called an external model.

The properties of the systems are studied in terms of the system structures and parameters. The definitions of property, property indicator and criteria are given below [Kar.,1].

Let $\mathcal{M}$ denote the set of system models, which are given either in transfer function or in state-space form. $M \in \mathcal{M}$ be a system model; denote $\mathcal{P}$ the set of all possible properties of interest and it should be referred to as the property set; denote $\mathcal{I}$ the set of property indicators. Associated with every property $P \in \mathcal{P}$, there may exist more than one indicators. Finally we denote by $\mathcal{C}$ a general set with elements, numbers, graphical statements, criteria, etc. and we shall call it the criteria set. Then we have the following definition.

Definition 2.1 : [Kar.,1] *A system property indicator is a function $f : \mathcal{M} \rightarrow \mathcal{I}$ defined on $M \in \mathcal{M}$ for each $P \in \mathcal{P}$. If $f(M)$ is the image of M under f, then a P-property test is a function g: $\mathcal{I} \rightarrow \mathcal{C}$.*

□

A diagrammatical illustration of the above definition is given in Figure (2.1).

Certainly, there may be more than one indicator for a system property. The corresponding criteria for the P-property test depends on the indicators used. Furthermore, a single indicator may be used for indicating different system properties.

An example of system models, system properties, property indicators and property criteria is given below.

Example (2.1):

models	property	property indicators	property criteria
$\dot{\underline{x}} = A\underline{x} + B\underline{u}$	controllability of the states	$[B, AB, \cdots, A^{n-1}B]$	full rank
		$[sI - A, -B]$	full rank for all $s \in C$
$\dot{\underline{x}} = A\underline{x}$	asymptotic stability of free motion	eigenvalues of A	negative real parts
$\dot{\underline{x}} = A\underline{x} + B\underline{u}$ $\underline{y} = C\underline{x} + D\underline{u}$	internal system poles		

Classification of the properties can be made depending on the models used, the nature of the criteria associated with the properties, etc.

The properties can be classified as *internal* and *external* depending on whether the model is internal or external.

If the criteria associated with a property are of binary nature, then the property is called *qualitative*; otherwise, if the criteria are defined in terms of a range of values, which expresses the "degree" of possession of the property by the model, then the property is called *quantitative.*

The properties can further be classified in terms of genericity [Won.,1] [MacL. & Bi If $\mathcal{M}$ is a family of models characterised by a fixed common structure but otherwise arbitrary parameters, then with every model $M \in \mathcal{M}$ we may associate a parameter vector $\underline{a}(M)$ in the parameter space R^n. A property is called *generic* if it holds true for almost all $M \in \mathcal{M}$. Otherwise, a property will be called *non-generic* if it is possessed only by a subset of systems, M', in the whole family of systems and the corresponding parameter vectors of M', $\underline{a}(M')$, form only a proper variety in the parameter space R^n. For instance, if we take the set of $n \times n$ real matrices, the property of having distinct eigenvalues is generic, whereas having repeated eigenvalues is a nongeneric property. A generic property may also be referred to as a structural property.

A property indicator which is used for assessing only a single property will be called *simple*; an indicator will be called a *multiple* indicator if it can assess more than one system property. Different indicators which show the same system property are called *equivalent.*

2.4 Open-loop system properties and property indicators

In this section, the open-loop system properties and their corresponding indicators are reviewed. The properties include system stability, controllability, observability and disturbability which are of fundamental interest in system analysis.

2.4.1 System stability

The most important property of any system is the stability property. It is related to the behaviour of all the possible trajectories which may be generated for different families of initial conditions and control inputs. The stability of a system is equivalent to the stability of the equilibrium points. For linear time invariant systems, stability can be defined in the following senses [Kai.,1] [Chen,1]. Note that the origin ($\underline{x} = 0$) is always an equilibrium point for $S(A, B, C, D)$ models.

Definition 2.2 : [Chen,1] *A system with a state-space model $S(A, B, C, D)$ and a transfer function $G(s)$ is:*

(i) Internally stable in the sense of Lyapunov: *If for any initial state $\underline{x}(0)$, the zero input response remains bounded for all $t > 0$.*

(ii) Asymptotically internally stable: *If for any initial state $\underline{x}(0)$, the zero input response remains bounded for all $t > 0$ and tends to zero as $t \to \infty$. This property is also referred to in short as internal stability.*

(iii) Bounded input bounded output stable: *If for any bounded input $\underline{u}(t)$, the zero initial state output response is bounded.*

(iv) Totally stable: *If for any initial state $\underline{x}(0)$ and any bounded input $\underline{u}(t)$, the output, as well as the state variables are bounded.*

□

The indicators corresponding to the above definitions are defined based on the eigenvalues and poles of the system [Rou.,1] [Chen,1].

Theorem 2.1 *Consider the system $S(A, B, C, D)$ with $G(s)$ as the transfer function, and let $\{\lambda_i = \sigma_i + j\omega_i, i \in \tilde{n}\}, \{p_j = \sigma_j + j\omega_j, j \in \tilde{v}\}$ be the sets of eigenvalues, poles respectively. The system is*

(i) *Lyapunov internally stable, if and only if* $\sigma_i \leq 0$ *for all* $i \in \tilde{n}$, *and those with* $\sigma_i = 0$ *have a simple structure, i.e. the algebraic multiplicity equals the geometric multiplicity for each of the distinct eigenvalues;*

(ii) *Asymptotically internally stable, if and only if* $\sigma_i < 0, \forall i \in \tilde{n}$;

(iii) *Bounded input bounded output stable, if and only if* $\sigma_j < 0, j \in \tilde{v}$;

(iv) *Totally stable, if it is Lyapunov internally stable and bounded input bounded output stable.*

□

Equivalent indicators for stability can also be defined on the characteristic polynomial and pole polynomial of a system. The Routh-Hurwitz test of a polynomial offers an alternative to calculating the exact values of the eigenvalues. If a characteristic polynomial is Hurwitz, then all the eigenvalues, which are the roots of the characteristic polynomial, will have negative real parts and thus imply internal stability. The same applies to the bounded input bounded output stability in relation to the pole polynomial.

If a system is internally stable, it implies that the system is also BIBO stable and thus in turn it also implies it is totally stable. However, BIBO stability implies neither internal stability nor total stability since the internal and external descriptions of a system are not always equivalent. If the external and the internal descriptions of a system are equivalent, i.e. the system described in state space is both controllable and observable, then BIBO stability implies internal stability and thus total stability.

The criteria for stability properties are of binary nature and thus stability is a qualitative property.

2.4.2 Controllability, observability properties and indicators

The controllability of a system deals with the interactions between either the inputs and the state variables or the inputs and the outputs of a system while the observability is concerned with the interaction between the state variables and the outputs.

Definition 2.3 : [Kal.,2] *The state space model* $S(A, B, C, D)$ *will be called:*

(i) State-controllable, *if there exists a finite time interval* T, $T > 0$, *such that for any initial state* $\underline{x}(0)$ *and an arbitrary state* $\underline{x}_1$ *in the state space, there always exists an input* $\underline{u}(t)$ *defined on* $[0, T]$ *such that the system can be driven from the initial state* $\underline{x}(0)$ *at* $t = 0$ *to state* $\underline{x}_1$, *at time* T $[\underline{x}(T) = \underline{x}_1]$. *Otherwise, the system is called uncontrollable.*

(ii) State observable, *if there exists a finite time interval* T, $T > 0$, *such that the knowledge of the input to the system,* $\underline{u}(t)$, *and the output of the system,* $\underline{y}(t)$, *over the time interval* $[0, T]$ *suffices to determine the initial state* $\underline{x}(0)$. *Otherwise, the system is called unobservable.*

(iii) Output controllable, *if for any output vector* $\underline{y}$, *there always exists a finite time interval* T, $T > 0$, *and an input* $\underline{u}(t)$ *defined over* $[0, T]$ *such that the output of the system can be steered from the initial output state* $\underline{y}(0) = \underline{0}$ *to* $\underline{y}(T) = \underline{y}_1$. *Furthermore, if the output of a system is not only controllable, but also can be steered on a preassigned curve over any period of time, then the system will be called* output function controllable [Ros.,1].

□

The controllability property can be further distinguished as controllability or reachability depending on whether one tries to transfer from an arbitrary initial state to the origin or from the zero initial state to an arbitrary final state, i.e., controllability refers to the transferring to the origin of an arbitrary initial state $\underline{x}(0)$ while reachability refers to the transferring of the zero initial state to an arbitrary final state in the state space. The difference in the definition is important when more general models are considered.

The corresponding indicators for the controllability and observability properties are given below [Kai.,1] [Chen,1] [Kar. & Mac.,1] [Ros.,1].

Theorem 2.2 *The state space model* $S(A, B, C, D)$ *with* n, l, m the number of states, *inputs and outputs, respectively, is*

(i) *State-controllable if and only if either of the following equivalent conditions hold true:*

- *All the rows of* $e^{At}B$ *are linearly independent on* $[0, \infty)$ *over* C.
- *All the rows of* $(sI - A)^{-1}B$ *are linearly independent over* C.
- *The state controllability Grammian* $G_{sc} = \int_0^T e^{A\tau}BB^T e^{A^T\tau} d\tau$ *is nonsingular for any* $T > 0$.

- *The $n \times nl$ controllability matrix defined as*

$$Q_c = [B, AB, A^2B, \cdots, A^{n-1}B] \tag{2.20}$$

has rank n.

- *The controllability pencil defined as $P_c(s) = [sI - A, -B]$ has full row rank for all $s \in C$, or equivalently it has no finite elementary divisors.*
- *The restricted controllability pencil $R_c(s) = sN - NA$, has rank n for all $s \in C$, where N is a left annihilator of B.*

(ii) *State-observable if and only if either of the following equivalent conditions hold true:*

- *All the rows of Ce^{At} are linearly independent on $[0, \infty)$ over C.*
- *All the rows of $C(sI - A)^{-1}$ are linearly independent over C.*
- *The observability Grammian $G_{so} = \int_0^T e^{A^T\tau}C^TCe^{A\tau}d\tau$ is nonsingular for any $T > 0$.*
- *The $nm \times n$ observability matrix defined as*

$$Q_o = [C^T, A^TC^T, (A^T)^2C^T, \cdots, (A^T)^{n-1}C^T] \tag{2.21}$$

has rank n.

- *The observability pencil defined as $P_o(s) = \begin{bmatrix} sI - A \\ -C \end{bmatrix}$ has full column rank for all $s \in C$.*
- *The restricted controllability pencil $R_o(s) = sM - AM$, has rank n for all $s \in C$, where M is the right annihilator of C.*

(iii) *Output controllable (for the strictly proper case, $D = 0$), if and only if either of the equivalent conditions hold true:*

- *All rows of $G(s) = C(sI - A)^{-1}B$ are linearly independent over C.*
- *The $m \times nl$ matrix $Q_{oc} = [CB, CAB, CA^2B, \cdots, CA^{n-1}B]$ has rank m.*

(iv) *Output function controllable if and only if the rank of $G(s)$ is equal to m over $R(s)$, where $R(s)$ denotes the field of rational functions with real coefficients.*

□

Remark 2.1: In terms of control quality, questions such as how close a system is to a controllable or uncontrollable one are often raised in the context of constrained control. The indicators for bounded energy controllability can be defined based on the singular values of the controllability Grammian which will be developed in the next chapter. These indicators will not only provide means of characterising the reachable states when a system is driven by a bounded energy control, but also quantifying input-output interaction and thus provide measures for selecting control structures.

□

The controllability and observability of systems are invariants under nonsingular coordinate transformations. For an appropriately chosen coordinate transformation, the system can be brought to a canonical form in which the state variables are partitioned into controllable-observable, controllable-unobservable, uncontrollable-observable and uncontrollable-unobservable sets.

Theorem 2.3 [Kal.,2] *For the state-space model $S(A, B, C, D)$, there exists a non-singular coordinate transformation P, $\underline{x}' = P\underline{x}$, such that the corresponding state space description $S'(A', B', C', D')$ has the form*

$$\begin{bmatrix} \dot{\underline{x}}'_{co} \\ \dot{\underline{x}}'_{c\bar{o}} \\ \dot{\underline{x}}'_{\bar{c}o} \\ \dot{\underline{x}}'_{\bar{c}\bar{o}} \end{bmatrix} = \begin{bmatrix} A'_{co} & 0 & A'_{13} & 0 \\ A'_{21} & A'_{c\bar{o}} & A'_{23} & A'_{24} \\ 0 & 0 & A'_{\bar{c}o} & 0 \\ 0 & 0 & 0 & A'_{\bar{c}\bar{o}} \end{bmatrix} \begin{bmatrix} \underline{x}'_{co} \\ \underline{x}'_{c\bar{o}} \\ \underline{x}'_{\bar{c}o} \\ \underline{x}'_{\bar{c}\bar{o}} \end{bmatrix} + \begin{bmatrix} B'_{c\bar{o}} \\ B'_{co} \\ 0 \\ 0 \end{bmatrix} \tag{2.22}$$

$$\underline{y} = \begin{bmatrix} C'_{co} & 0 & C'_{c\bar{o}} & 0 \end{bmatrix} \begin{bmatrix} \underline{x}'_{co} \\ \underline{x}'_{c\bar{o}} \\ \underline{x}'_{\bar{c}o} \\ \underline{x}'_{\bar{c}\bar{o}} \end{bmatrix}$$

where $\underline{x}'_{c\bar{o}}, \underline{x}'_{co}, \underline{x}'_{\bar{c}o}, \underline{x}'_{\bar{c}\bar{o}}$ are controllable-unobservable, controllable-observable, uncontrollable-observable and uncontrollable-unobservable state variable sets, respectively.

□

A diagrammatic representation of the decomposition is given in Fig. (2.2).

The transfer function of a system is invariant under nonsingular state coordination transformations, therefore all the systems $S(A, B, C, D)$, which can be transformed to $S'(A', B', C', D)$ by appropriate nonsingular coordination transformations, share the same transfer function, i.e.,

$$G(s) = C'_{co}(sI - A'_{co})^{-1}B'_{co} + D. \tag{2.23}$$

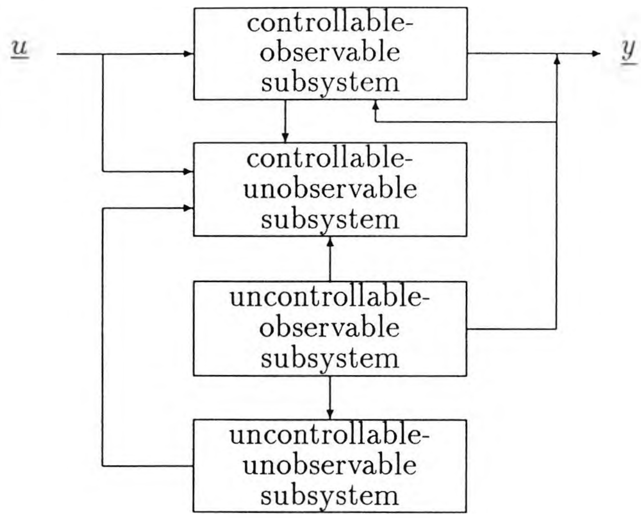


Figure 2.2: Canonical decomposition

This establishes the relationship between the descriptions of systems in the time and the frequency domains. It shows that the transfer function represents only the controllable and observable subsystem, not in general the overall system. The transfer function and the state space descriptions are equivalent if and only if the system is both controllable and observable, or when the system have the same system poles and external poles.

Closely related to the system controllability and observability properties are the system stabilisability and detectability.

Definition 2.4 : [Won.,1] *The state space model $S(A, B, C, D)$ is*

(i) Stabilisable, *if the unstable states of A are contained in the controllable subspace of the system.*

(ii) Detectable, *if the unobservable states of the system are contained in the stable eigenspace of A.*

□

Theorem 2.4 [Won.,1] *The system $S(A, B, C, D)$ is*

(i) *Stabilisable, if and only if its uncontrollable eigenvalues are stable.*

(ii) *Detectable, if and only if the unobservable eigenvalues are stable.*

□

These two more relaxed conditions have implications in the use of the transfer functions as the design models. If the system is both stabilisable and detectable, then the transfer functions may be used for feedback design, but not otherwise.

The uncontrollable, unobservable and the uncontrollable-unobservable eigenvalues are also referred to as input-, output- and input-output-decoupling zeros [Ros.,1] and the corresponding sets, including multiplicities, will be denoted by Z_{IZ}, Z_{OZ}, Z_{IOZ}, respectively. These sets can be computed in the following way [Ros.,1] [MacF. & Kar.,1] [Kar. & Mac.,1].

Theorem 2.5 (i) *Z_{IZ} is defined by the roots of the finite elementary divisors of $P_c(s) = [sI - A, -B]$, or equivalently $R_c(s) = sN - NA$, where N is the left annihilator of B;*

(ii) *Z_{OZ} is defined by the roots of the finite elementary divisors of $P_c(s) = \begin{bmatrix} sI - A \\ -C \end{bmatrix}$ or equivalently $R_o(s) = sM - AM$, where M is the right annihilator of C;*

(iii) $Z_{IOZ} = Z_{IZ} \cap Z_{OZ}$

□

The output function controllability is an important concept in process control [Ros.,1] where the output is sometimes required to follow certain trajectories. In this context, the output controllability is not sufficient to address this issue because in order to follow certain trajectories, impulse inputs may be required. Obviously, this is unrealistic in real applications. A dynamical system can only track a prescribed output up to $p - 1$ independent derivatives as stated in the following theorem.

Theorem 2.6 [Ske.,1] *For the system $S(A, B, C, D)$, we have that*

(i) *A system cannot track smoothly an output function $\underline{c}(t)$ with more than $p - 1$ independent derivatives.*

(ii) *A dynamical system can track smoothly the vector function $\underline{c}(t)$ up to its first $p < n - 1$ derivatives, if $\underline{c}(t)$ is sufficiently smooth to have r derivatives, and* *the matrix M_r defined below has rank $(r+1)m$,*

$$M_r = \begin{bmatrix} D & 0 & \cdots & 0 & 0 \\ CB & D & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ CA^{r-2}B & CA^{r-3}B & & D & 0 \\ CA^{r-1}B & CA^{r-2}B & \cdots & CB & D \end{bmatrix}. \qquad (2.24)$$

□

Strictly proper systems ($D = 0$) can not have rank$(M_r) = (r+1)l$ satisfied under any circumstance and therefore can only track an output $\underline{c}(t)$ with a constant offset, $\underline{y}(t) - \underline{c}(t)$ =constant, if rank$(M_r) = rl$. In other words, to have a smooth tracking of an output which has more than $(r+1)l$ derivatives requires necessarily a higher order dynamical system.

In real applications, cases may arise when the responses to some input signals are required to be suppressed if not at all eliminated. Such input signals are referred to as disturbances or noises. A typical internal type disturbance model is obtained by expanding $S(A, B, C, D)$ as $S(A, B, C, D, H)$ in the following form

$$\dot{\underline{x}}(t) = A\underline{x}(t) + B\underline{u}(t) + H\underline{\omega} \qquad (2.25)$$

$$\underline{y}(t) = C\underline{x}(t) + D\underline{u}(t) \qquad (2.26)$$

where $\underline{\omega} \in R^d, H \in R^{n\times d}$ are the disturbance-vector and the disturbance matrix, respectively.

Definition 2.5 *:* [Ske.,1] *The state space model $S(A, B, C, D, H)$ is called* output disturbable *if the system rejects the disturbance $\underline{\omega}$ in the output.*

□

For the above system property, an indicator for the output disturbability based on the Markov parameters of the systems can be defined.

Theorem 2.7 [Ske.,1] *For an arbitrary disturbance $\underline{\omega}(t)$, the system $S(A, B, C, D, H)$ completely rejects $\underline{\omega}(t)$ if and only if*

$$Q_d = \begin{bmatrix} CH, & CAH, & \cdots, & CA^{n-1}H \end{bmatrix} = 0 \qquad (2.27)$$

□

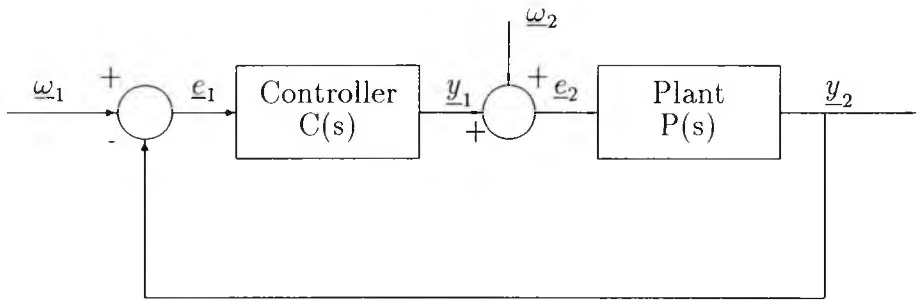


Figure 2.3: Complete feedback configuration

2.5 Closed-loop system properties and property indicators based on input-output models

In order to improve system performance, systems need to be controlled by using different control schemes, such as feedforward, feedback, etc. Feedback control is by far the most widely used scheme. The effects of feedback on the system properties which have been summarised in the previous section are discussed in this section.

2.5.1 General feedback configuration

Figure (2.3) shows the complete feedback configuration. Let $P(s) \in R^{m \times l}, C(s) \in R^{l \times m}(s)$ be the plant, controller transfer functions, $\underline{\omega}_1, \underline{\omega}_2$ the external control or disturbance signals, $\underline{e}_1, \underline{e}_2$ the controller and plant inputs and $\underline{y}_1, \underline{y}_2$ the controller and plant outputs. Such a configuration can represent either feedback or cascade compensation. Different design problems, such as tracking, disturbance rejection, can be accommodated. Therefore it is referred to as the complete feedback configuration.

The system equations are defined by

$$\begin{bmatrix} \underline{e}_1 \\ \underline{e}_2 \end{bmatrix} = \begin{bmatrix} \underline{\omega}_1 \\ \underline{\omega}_2 \end{bmatrix} - \begin{bmatrix} 0 & P \\ -C & 0 \end{bmatrix} \begin{bmatrix} \underline{e}_1 \\ \underline{e}_2 \end{bmatrix}, \quad \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} \underline{e}_1 \\ \underline{e}_2 \end{bmatrix} \tag{2.28}$$

or by

$$\underline{e} = \underline{\omega} - FG\underline{e}, \quad \underline{y} = G\underline{e} \tag{2.29}$$

where $\underline{e} = \begin{bmatrix} \underline{e}_1 \\ \underline{e}_2 \end{bmatrix}$, $\underline{\omega} = \begin{bmatrix} \underline{\omega}_1 \\ \underline{\omega}_2 \end{bmatrix}$, $\underline{y} = \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix}$, $G = \begin{bmatrix} C & 0 \\ 0 & P \end{bmatrix}$, $F = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$.

The feedback configuration is said to be well formed if

$$|I + FG| = |I + PC| = |I + CP| \neq 0 \tag{2.30}$$

Then we may define the transfer functions $H(P,C), W(P,C)$ by

$$\begin{bmatrix} \underline{e}_1 \\ \underline{e}_2 \end{bmatrix} = (I+FG)^{-1}\underline{\omega} = H(P,C)\underline{\omega} \tag{2.31}$$

$$\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix} = G(I+FG)^{-1}\underline{\omega} = W(P,C)\underline{\omega} \tag{2.32}$$

(2.33)

where

$$W(P,C) = \begin{bmatrix} C(I+PC)^{-1} & -CP(I+CP)^{-1} \\ PC(I+PC)^{-1} & P(I+CP)^{-1} \end{bmatrix} \tag{2.34}$$

$$H(P,C) = \begin{bmatrix} (I+PC)^{-1} & -P(I+CP)^{-1} \\ C(I+PC)^{-1} & (I+CP)^{-1} \end{bmatrix}$$

Further, we can also define the transfer functions

$$\begin{bmatrix} \underline{y}_1 \\ \underline{e}_1 \end{bmatrix} = T(P,C)\begin{bmatrix} \underline{\omega}_1 \\ \underline{\omega}_2 \end{bmatrix}, \quad \begin{bmatrix} \underline{y}_2 \\ \underline{e}_2 \end{bmatrix} = R(P,C)\begin{bmatrix} \underline{\omega}_2 \\ \underline{\omega}_1 \end{bmatrix} \tag{2.35}$$

where $T(P,C)$ and $R(P,C)$ take the following forms

$$T(P,C) = \begin{bmatrix} C(I+PC)^{-1} & -C(I+PC)^{-1}P \\ (I+PC)^{-1} & (I+PC)^{-1}P \end{bmatrix} \tag{2.36}$$

$$R(P,C) = \begin{bmatrix} P(I+CP)^{-1} & P(I+CP)^{-1}C \\ (I+CP)^{-1} & (I+CP)^{-1}C \end{bmatrix}$$

For system analysis we further define some matrices which are important in the study of feedback:

$$\begin{array}{ll} Q \equiv PC, & Q' \equiv CP \\ F \equiv I+PC, & F' \equiv I+CP \\ L \equiv I+(PC)^{-1} & \\ S \equiv (I+PC)^{-1}, & S' \equiv (I+CP)^{-1} \end{array} \tag{2.37}$$

where Q, Q' are referred to as the return ratio matrices, F, F' as the return difference matrices, L as the inverse-return difference matrix and S, S' as the sensitivity matrices.

Theorem 2.8 [Kuc.,1] *All transfer function matrices, $W(P,C), H(P,C), T(P,C)$ and $R(P,C)$ associated with the general feedback configuration have the same pole polynomial.*

□

The above theorem is important because it shows that it is possible to study the pole assignment and stabilisation problems using any of the above transfer function functions. Assume that the plant and the controller are in matrix fraction descriptions, then the pole polynomial of the system can be obtained as presented in the following theorem.

Theorem 2.9 [Kuc.,1] [Vid.,1] *If $P = A_l^{-1}B_l = B_r A_r^{-1}, C = D_l^{-1}N_l = N_r D_r^{-1}$ are left, right coprime MFDs, then*

(i) *A left, right coprime MFD for $H(P,C)$ is defined by*

$$H(P,C) = \begin{bmatrix} A_l & B_l \\ -N_l & D_l \end{bmatrix}^{-1} \begin{bmatrix} A_l & 0 \\ 0 & D_l \end{bmatrix} = \begin{bmatrix} D_r & 0 \\ 0 & A_r \end{bmatrix} \begin{bmatrix} D_r & B_r \\ -N_r & A_r \end{bmatrix}^{-1} \tag{2.38}$$

(ii) *The pole polynomial or pole function of $H(P,C)$ is*

$$f \sim |V_1| \sim |V_2|, \tag{2.39}$$

where $V_1 = A_l D_r + B_l N_r, V_2 = D_l A_r + D_l B_r$ and $\sim$ means equal modulo constants for $R[s]$ case, or modulo $R_{\mathcal{P}}(s)$ units.

□

An important issue for the feedback system is that of well-posedness, i.e., the properness of all the transfer functions associated with the feedback system.

Theorem 2.10 [Vid.,1] *H, C, P are proper transfer functions, then the general feedback configuration is well posed if and only if*

$$|I + C(\infty)P(\infty)| = |I + P(\infty)C(\infty)| \neq 0 \tag{2.40}$$

□

So if one of P, C is a strictly proper, then either $P(\infty) = 0$ or $C(\infty) = 0$, then the above inequality will always be satisfied and therefore the system is always well-posed.

The controllability and observability of the general feedback configuration are related to the controllability and observability of the plant and the controller in the following way.

Let S_P, S_C be the plant, controller state space model and S_f be the state space description of the general feedback configuration. Then

Theorem 2.11 [Vid.,1] *For the well-posed general feedback configuration the following properties hold true*

(i) *S_f is controllable and observable, if and only if both S_P, S_C are controllable and observable.*

(ii) *S_f is stabilisable and detectable, if and only if both S_P, S_C are stabilisable and detectable.*

□

The internal stability of $H(P,C)$ can be related to the BIBO stability of $H(P,C)$ and thus it is possible to establish means of studying the internal stability in terms of the BIBO stability of $H(P,C)$.

Theorem 2.12 [Vid.,1] *Consider the well-posed general feedback system with both S_P, S_C stabilisable and detectable. S_f is internally stable, if and only if $H(P,C)$ is BIBO stable.*

□

The internal stability is related to the MFD descriptions of $P(s)$ and $C(s)$ in the following manner.

Theorem 2.13 [Vid.,1] *Consider the well-posed feedback system S_f with S_P, S_C both stabilisable and detectable, then*

(i) *If $P = A_l^{-1}B_l = B_r A_r^{-1}, C = D_l^{-1}N_l = N_r D_r^{-1}$ are $R_{\mathcal{P}}(s)$ coprime MFDs, then S_f is internally stable if and only if either of the equivalent conditions hold true:*

$$A_l D_r + B_l N_r = V_1, \quad V_1 \in R_{\mathcal{P}}(s) \text{ — unimodular} \tag{2.41}$$

$$D_l A_r + N_l B_r = V_2, \quad V_2 \in R_{\mathcal{P}}(s) \text{ — unimodular} \tag{2.42}$$

(ii) *If S_P, S_C are controllable and observable and $P = A_l^{-1}B_l = B_r A_r^{-1}, C = D_l^{-1}N_l = N_r D_r^{-1}$ are $R[s]$-coprime MFDs, then the characteristic polynomial of S_f is defined by*

$$|sI - A_f| \sim |A_l B_r + B_l N_r| \sim |D_l A_r + N_l B_r| \tag{2.43}$$

Furthermore, if $F = I + PC, F' = I + CP$ are the return difference matrices, and A_P, A_C are the plant, controller state matrices, then

$$|I + PC| \sim |I + CP| \sim \frac{|sI - A_f|}{|sI - A_P||sI - A_C|}. \tag{2.44}$$

□

Substituting V_1, V_2 in equations (2.41) and (2.42) with the identity matrices, they become Diofantine equations [Kuc.,1] [Vid.,1]. And the solutions (A_l, B_l), (D_r, N_r) or (A_r, B_r), (C_l, N_l) are referred to as mutually stabilising pairs. The solutions for a fixed P give the whole family of controllers which stabilises the plant while the solution for a fixed controller gives the whole family of plants which will be stabilised by the controller. These results are important in the parametrisation of the controllers that stabilise a given plant or the plants for a given controller. Equation (2.44) establishes the important property of the determinant of the return difference as the ratio of closed loop pole polynomials of the feedback system, which underlines the multivariable Nyquist and root locus theories.

2.5.2 Closed-loop stability properties and stability indicators — Characteristic gain and characteristic frequency loci

For a square $(m = l)$ system $S(A, B, C, D)$ the conditions for transmitting a frequency s in the standard scalar k output feedback configuration are defined by [MacF. & Kar.,2]

$$\begin{bmatrix} sI_n - A & -B \\ -C & gI_m - D \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{u} \end{bmatrix} = \underline{0}, \text{ or } P(s,g) \begin{bmatrix} \underline{x} \\ \underline{u} \end{bmatrix} = \underline{0} \tag{2.45}$$

where $g = k^{-1}, \underline{x} \in C^n, \underline{u} \in C^l$. $P(s,g)$ is a two variable object and it is referred to as the *system net*. The above equality defines two eigenvalue-eigenvector problems

$$[sI_n - S(g)]\underline{x} = \underline{0}, \quad S(g) = A + B(gI_m - D)^{-1}C \tag{2.46}$$

$$[gI_m - G(s)]\underline{u} = \underline{0}, \quad G(s) = D + C(sI_n - A)^{-1}B \tag{2.47}$$

$S(g)$ is called a characteristic frequency function and $G(s)$ a characteristic gain function. The eigenvalue-eigenvector problem defined on $S(g)$ (parameter dependent on g) defines the closed-loop eigenvalues as a function of g and $\underline{x}$ is the corresponding eigenvector; if g takes real values, the solution $s(g)$ to equation (2.46) defines the multivariable root loci. Similarly, the problem defined on $G(s)$ (parameter dependent on s) defines the open-loop characteristic gains as functions of frequency and $\underline{u}$ will be the corresponding characteristic directions; if s takes values on the Nyquist contour, this problem defines the multivariable Nyquist diagrams and directions [Ros.,3][MacF.,1] [MacF. & Kou.,1] [MacF. & Pos.,1] [Pos. & MacF.,1]. The

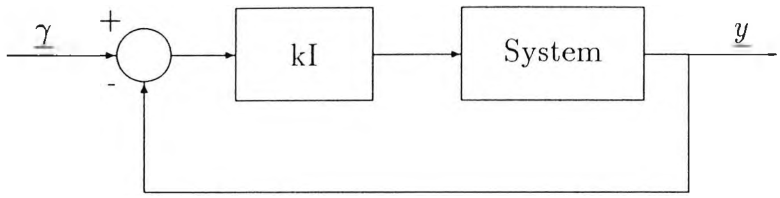


Figure 2.4: Standard scalar negative feedback configuration

characteristic equations,

$$\Delta(g,s) = |gI_m - G(s)| = g^m + a_1(s)g^{m-1} + \cdots + a_m(s) = 0 \qquad (2.48)$$

$$\nabla(s,g) = |sI_n - S(g)| = s^n + b_1(g)s^{n-1} + \cdots + b_n(g) = 0 \qquad (2.49)$$

are equivalent for all $s \in \sigma(P), g \in \sigma(D)$. The roots of the above equations are analytic functions of s, g, respectively.

Definition 2.6 *:* [Pos. & MacF.,1]

(i) *The Nyquist diagrams or the characteristic gain loci are defined to be the m branches $g_i(s)$ of $\Delta(g,s)$ when s takes values on the Nyquist contour, $s = j\omega$, and the eigenvalue-eigenvector decomposition (defined frequency by frequency)*

$$G(j\omega) = W(j\omega)\Lambda(j\omega)S(j\omega)^{-1}, \quad \Lambda(j\omega) = diag\{g_i(j\omega)\} \qquad (2.50)$$

as the characteristic gain decomposition of $G(j\omega)$, where the columns of $W(j\omega)$ are the eigenvectors of corresponding $g_i(j\omega)$ and they are called the Nyquist directions or characteristic frequency vectors.

(ii) *If g takes values on the real axis including ∞, then the n branches $s_i(g)$ of $\nabla(s,g)$ are defined as the characteristic frequency loci, or root loci, and the corresponding eigenvectors of (2.47) as the characteristic frequency vector.*

□

The Nyquist and root locus diagrams may be used to assess the closed-loop stability from open-loop information for the standard scalar negative feedback configuration in Figure 2.4.

The root loci are the trajectories of the closed-loop poles for different values of k, so the stability property of the closed-loop system can be inferred directly. The closed-loop system stability can also be inferred from the Nyquist diagrams by making use of the following Nyquist theorem [MacF. & Pos.,1].

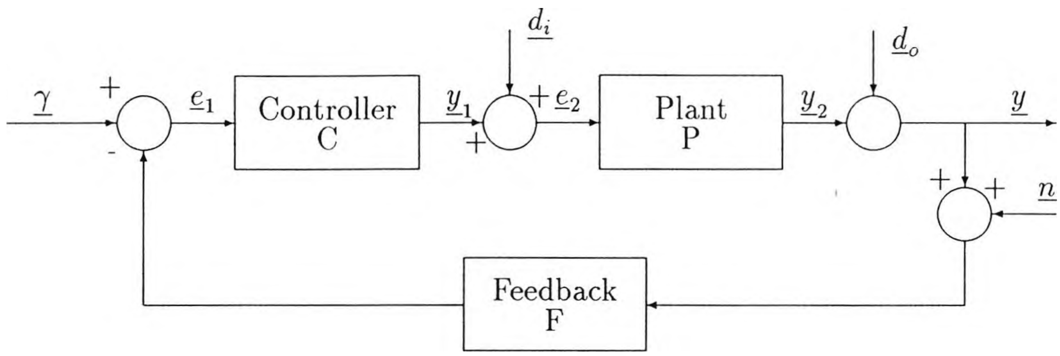


Figure 2.5: Design configuration

Theorem 2.14 [MacF. & Pos.,1] *If the open-loop system is stabilisable and detectable with $G(s)$ as its transfer function, then the closed-loop configuration given will be closed-loop stable, if and only if the net sum of anticlockwise encirclement of the critical point $(-\frac{1}{k}, j0)$ by the set of characteristic gain loci of $G(s)$ is equal to the total number of right-half plane poles of $G(s)$.*

□

Apart from being a stability indicator, the characteristic gain loci may also be used as closed-loop property indicators for such as tracking, noise rejection under the assumption of relative normality of $G(s)$ [MacF. & Kou.,1].

2.5.3 Closed-loop performance in terms of properties of return ratio, difference and sensitivity matrices

Consider the design configuration in Figure (2.4) which takes into account the sensor noise, input and output disturbances. P, C, F are proper transfer function matrices of the plant, precompensator, feedback compensator, respectively (with appropriate dimensions), and $\underline{r}$, $\underline{d}_i$, $\underline{d}_o$, $\underline{n}$ are the reference, plant input disturbance, sensor noise vector signals, respectively. Assume that P, C, F are stabilisable and detectable and that C, F stabilises the feedback configuration, and thus all transfer function matrices from any external signal to the output $\underline{y}$ are stable.

Define the transfer functions from all external signals to the output by: $H_r : \underline{r} \rightarrow \underline{y}, H_{od} : \underline{d}_o \rightarrow \underline{y}, H_{id} : \underline{d}_i \rightarrow \underline{y}, H_n : \underline{n} \rightarrow \underline{y}$, then

$$\begin{aligned} H_r &= PC(I+FPC)^{-1} = (I+PCF)^{-1}PC \\ H_{od} &= (I+PCF)^{-1} \\ H_{id} &= (I+PCF)^{-1}P \\ H_n &= -(I+PCF)^{-1}PCF \end{aligned}$$

By linearity, the output of the system is given by

$$\underline{y} = H_r \underline{r} + H_{od} \underline{d}_o + H_{id} \underline{d}_i + H_n \underline{n} \tag{2.51}$$

So for external signals with given bandwidth, the above equation describes the effect of each signal on the overall system response in terms of frequency domain transfer functions, $H_r(j\omega), H_{od}(j\omega), H_{id}(j\omega), H_n(j\omega)$ which are expressed in terms of the return ratio PCF, return difference $(I + PCF)$ and sensitivity matrix $(I + PCF)^{-1}$. The frequency domain study of tracking, disturbance or noise rejection, sensitivity to plant parameter variation, robustness, etc. involves the notion of gain of transfer function matrices, which is defined in terms of the singular values.

2.5.3.1 Singular value and polar decomposition of transfer function matrices

The singular values of a system are very important property indicators of system properties in the frequency domain [Pos. Edm. & MacF.,1] [Doy. & Ste.,1] [MacF. & Sco.,1]. The linear system $G(s) \in R^{m \times l}(s)$ is a matrix valued function of the complex variable s. If we evaluate $G(s)$ at each $s \in C$, especially on the Nyquist contour, then for $G(j\omega)$ we may define the Singular Value Decomposition (SVD) and Polar Decomposition.

Singular Value Decomposition: [Gan.,1] [Pos. Edm. & MacF.,1] Let $G = G(j\omega) \in C^{m \times l}$ and $\sigma_i, i \in \tilde{r}$ be the singular values of G which are ordered as $\sigma_r \geq \cdots \geq \sigma_1 > 0$ and define $\Sigma = \text{diag}(\sigma_r, ..., \sigma_1)$ $(r = \min\{m, l\})$ then G may be expressed as

$$G = Y \Sigma U^* \tag{2.52}$$

where

(i) if $m \geq l$, then $Y \in C^{m \times l}, \Sigma \in R^{l \times l}, U^* \in C^{l \times l}, Y^*Y = I_l, U^*U = I_l = UU^*$.

(ii) if $m < l$, then $Y \in C^{m \times m}, \Sigma \in R^{m \times m}, U^* \in C^{m \times l}, Y^*Y = I_m, U^*U = I_m = UU^*$.

which is called the singular value decomposition. Y and U are referred to as output-input singular vector frame matrices and Σ the principal gain matrix.

Closely related to the singular value decomposition is the polar decomposition.

Polar Decomposition: [Pos. Edm. & MacF.,1] Let $G = G(j\omega) \in C^{m \times l}$ and consider the SVD of G as in the above. The G may be expressed as

$$G = (Y \Sigma Y^*)(YU^*) = M_l \Phi \tag{2.53}$$

$$G = (YU^*)(U \Sigma U^*) = \Phi M_r \tag{2.54}$$

where Φ, M_l, M_r are referred to as phase, left, right modulus matrices. Equations (2.53) and (2.54) are referred to as the polar decompositions of G. If G is square, YU^* is unitary and its characteristic decomposition is expressed by

$$YU^* = \Phi = P\Theta P^*, \Theta = \text{diag}(e^{j\theta_i}) \tag{2.55}$$

where P is unitary and the set of angles θ_i are defined as the principal phases.

Since $G(s)$ is analytic, the plots of singular values are continuous functions of ω. In the case when the system is square, the singular values $\sigma_i(j\omega)$ will be called the principal gain functions. The principal phases $\theta_i(j\omega)$ are also analytic functions of ω. The plots of $\sigma_i(j\omega)$ define the multivarialbe amplitude Bode diagrams whereas the plots of $\theta_i(j\omega)$ give the multivariable phases Bode diagrams.

An important concept in analysis and design, which is related to the SVD of $G(s)$ is that of the vector gain. For a system

$$\underline{y}(s) = G(s)\underline{u}(s), \qquad s = jw \tag{2.56}$$

then one may define the *vector gain* of $G(s)$ for input $\underline{u}(s)$ as

$$\text{gain } G|_{\underline{u}} = \frac{\|\underline{y}(s)\|_2}{\|\underline{u}(s)\|_2} \tag{2.57}$$

where $\|\cdot\|_2$ denotes the Euclidean norm. If $\bar{\sigma}(jw), \underline{\sigma}(jw)$ denote the minimal and maximal singular values of $G(jw)$, the

$$\underline{\sigma}(jw) \leq \frac{\|\underline{y}(s)\|_2}{\|\underline{u}(s)\|_2} \leq \bar{\sigma}(jw), \text{ for all } \underline{u}(s) \tag{2.58}$$

The above property is known as the Min-Max theorem and has important implications in analysing signal tracking, noise rejection and sensitivity to parameter variations.

2.5.3.2 Closed-loop performance indicators

The singular values of the return ratio, return difference and sensitivity matrices are indicators of the system tracking, noise rejection and sensitivity to plant parameter variations [MacF. & Sco.,1] [Pos. Edm. & MacF.,1] [Doy. & Ste.,1].

Denote $\underline{\sigma}(G)$ and $\bar{\sigma}(G)$ the smallest and the largest singular values of $G(j\omega)$ and the range of frequencies of interest as Ω. Then

- The system has good tracking if $\underline{\sigma}(PCF) \gg 1$ for all $\omega \in \Omega$.

- The system has good plant disturbance rejection if $\underline{\sigma}(I + PCF) \gg 1$ for all $\omega \in \Omega$.

- The system has good sensor noise rejection if $\underline{\sigma}(I + (PCF)^{-1}) \gg 1$ over all $\omega \in \Omega$ or equivalently $\underline{\sigma}(PCF)$ is small for all $\omega \in \Omega$.

2.6 State-space invariants

Given the whole set of systems described in linear state space form, some of the systems share similar properties, such as stability, controllability, observability, transient response to step input, etc. By using different system properties, the systems can be classified into different categories. The reason for different systems sharing the same property is due to the fact that the systems share the same underlying structure, or even the same system parameters. In fact, the similarity in properties among systems lies in the equivalence relation among the systems. The equivalence relation among the systems can be defined mathematically as:

Definition 2.7 : [MacL. & Bir.,1] *If A is a set, then a relation R on A is a subset $A \times A$ [Cartesian product, set of ordered pairs $(x, y), (x, y) \in A$]. A relation R is called an equivalence relation if it is*

- *Reflexive, i.e. $(x, x) \in R, x \in R$;*

- *Symmetric, i.e. $(x, y) \in R$, then $(y, x) \in R$; and*

- *Transitive, i.e. $(x, y) \in R, (y, z) \in R$, then $(x, z) \in R$*

□

In the context of control theory, the equivalence relation can be interpreted as state, input and output coordinate transformations or in feedback sense as the state feedback, output feedback and output injection.

For a specifically defined equivalence relation, R^0, denote $R^0(x)$ as the subset of elements which are equivalent to $x \in A$, then the whole set of linear systems can be decomposed into a set of disjointed classes $A = U^0(x_i)$. And each of the classes can be represented by an element, say, $x_0^i \in R^0(x_i)$ and x_0^i will be called the representative of $R^0(x_i)$. The set of all representatives, T, of the equivalence classes will be called a system of distinct representatives.

The study of the properties of a system is one but not the final goal of control theory; control theory also aims at producing a system with desirable properties by employing different compensation schemes, such as feedforward, state feedback, output feedback, etc. So the study of the properties under different compensation schemes gives guidance in choosing the different compensation schemes. In fact, some of the properties will remain unchanged with a certain compensation and these properties will be called the invariants under this particular compensation scheme. The mathematical definition of system invariants is given below [MacL. & Bir.,1].

Let $\mathcal{M}$ be a family of linear models, E an equivalent relation defined on $\mathcal{M}$, $E(M)$ the equivalent class of $M \in \mathcal{M}$ and let $\mathcal{M}/E$ be the quotient set of orbit (set of all equivalence classes). We may define

Definition 2.8 : [MacL. & Bir.,1] *Let $\mathcal{M}$ be a family of models, I a set, E an equivalence relation defined on $\mathcal{M}$.*

(i) *A function $f : \mathcal{M} \to I$ is called invariant of E, when $M_1 E M_2$, implies $f(M_1) = f(M_2)$. Also, f is called a complete invariant for E when $f(M_1) = f(M_2)$ implies $M_1 E M_2$.*

(ii) *A set of invariants $f_i : \mathcal{M} \to I_i, i = 1, 2, ..., k$, is a complete invariant for E on $\mathcal{M}$, if the map defined by $f : \mathcal{M} \to \Pi_{i=1}^{k} I_i : M \to f(M) = \{f_1(M), ..., f_k(M)\}$ is a complete invariant for E on $\mathcal{M}$. The complete set of invariants is called independent.*

□

So a complete invariant defines a one to one correspondence between the R equivalence $R(x)$ and the image of f. By specialising the complete invariant f such that a canonical element $c \in R(x)$ uniquely characterises $R(x)$ and $c \in R(x)$ will be referred to as the canonical form.

The importance of the study of the equivalence relation, system invariants and canonical forms lies in the fact that the systems can be studied in groups. The result of the analysis of a particular system can be extended to every system which is equivalent to that particular system. And this particular system will usually be chosen to be the representative of the class. By employing special invariant of the system, canonical forms can be formed in such a way that it is simple, informative and thus reduce the amount of effort in analysis.

The invariants reveal the system capabilities and limitations under coordination transformations or feedback control because the equivalence between systems can be interpreted in the sense of coordinate transformations as well as feedback control. These invariants will give valuable information in system compensation by using these compensation schemes.

In the sequel, the system invariants in state space domain are reviewed.

2.6.1 Equivalence relations

The equivalence relations which can be defined on the systems described by state space models $S(A, B, C, D)$ are the system coordinate transformation, input coordinate transformation, output coordinate transformation and the state feedback, output feedback and output injection. So the most general types of transformations that may be applied on $S(A, B, C, D)$ systems involve all those defined by Q, T, R which are state, output, input coordinate transformations, state feedback L and output injection, F. Based on Q, T, R, L, F transformations, the following ordered set of transformations can be defined [Kar.,1]

$$\mathcal{H}_k \equiv \{H_k : H_k = (Q, T, R; L, F)\} \quad (2.59)$$
$$\mathcal{H}_B^r \equiv \{H_B^r : H_B^r = (Q, R; L)\} \quad (2.60)$$
$$\mathcal{H}_B^l \equiv \left\{H_B^l : H_B^l = (Q, T; F)\right\} \quad (2.61)$$
$$\mathcal{H}_c \equiv \{H_c : H_c = (Q, T, R; 0, 0) = (Q, T, R)\} \quad (2.62)$$
$$\mathcal{H}_c^{is} \equiv \left\{H_c^{is} : H_c^{is} = (Q, 0, R) = (Q, R)\right\} \quad (2.63)$$
$$\mathcal{H}_c^{os} \equiv \{H_c^{os} : H_c^{os} = (Q, T, 0) = (Q, T)\} \quad (2.64)$$
$$\mathcal{H}_c^{s} \equiv \{H_c^{s} : H_c^{s} = (Q, 0, 0) = (Q)\} \quad (2.65)$$

These transformations form groups, $\mathcal{H}_k, \mathcal{H}_B^r, \mathcal{H}_B^l$ will be referred to as the Kronecker, right- and left-Brunovsky groups and $\mathcal{H}_c, \mathcal{H}_c^{is}, \mathcal{H}_c^{os}, \mathcal{H}_c^{s}$ as general, input-state, state-output and state-coordinate groups, respectively. The study of the system equivalence can be carried out using the matrix pencils. The action of these different groups on the systems can be expressed as action on the pencils associated with the corresponding type of system pencils under consideration. So

(i) Action of $\mathcal{H}_k, \mathcal{H}_c$ on $S(A, B, C, D)$ is defined by:

$$\begin{bmatrix} sI - A' & -B' \\ -C' & -D' \end{bmatrix} = \begin{bmatrix} Q^{-1} & F \\ 0 & T \end{bmatrix} \begin{bmatrix} sI - A & -B \\ -C & -D \end{bmatrix} \begin{bmatrix} Q & 0 \\ L & R \end{bmatrix} \quad (2.66)$$

(ii) Action of $\mathcal{H}_B^r, \mathcal{H}_c^{is}$ on $S(A, B)$ is defined by

$$\begin{bmatrix} sI - A' & -B' \end{bmatrix} = Q^{-1} \begin{bmatrix} sI - A & -B \end{bmatrix} \begin{bmatrix} Q & 0 \\ L & R \end{bmatrix} \quad (2.67)$$

(iii) Action of $\mathcal{H}_B^l, \mathcal{H}_c^{os}$ on $S(A,C)$ is defined by:

$$\begin{bmatrix} sI - A' \\ -C' \end{bmatrix} = \begin{bmatrix} Q^{-1} & F \\ 0 & T \end{bmatrix} \begin{bmatrix} sI - A \\ -C \end{bmatrix} Q \tag{2.68}$$

and finally

(iv) Action of $\mathcal{H}_c^s$ on $S(A)$ is defined by

$$sI - A' = Q^{-1}(sI - A)Q \tag{2.69}$$

For the given groups, invariants and canonical forms may be defined. The groups of transformations fall into two categories, those which involve coordinate transformations only and those which also involve feedback. So we distinguish the invariants and canonical forms for these two different groups accordingly.

2.6.2 Invariants and canonical forms under coordinate transformation

2.6.2.1 State coordinate transformations $\mathcal{H}_c^s$ on $S(A)$

For systems $S(A), \dot{\underline{x}} = A\underline{x}$, coordinate transformations are equivalent to similarity transformations. The structure of the eigenvalues defines the invariants and canonical form [Gan.,1], [Kar. & Kal.,1].

Theorem 2.15 *If $\Lambda(A)$ is the root range of A, and $\mathcal{S}(A,\lambda) = \{\nu_1, ..., \nu_q\}$ is the Segre characteristic for every $\lambda \in \Lambda(A)$, then the set $\{\mathcal{S}(A,\lambda_i), \lambda_i \in \Lambda(A)\}$ is a complete invariant for similarity equivalence on $S(A)$. The corresponding canonical form is the Jordan canonical form:*

$$J(A) = diag\{\cdots, J(\lambda_i), \cdots\} \tag{2.70}$$

where $J(\lambda_i) = diag\{J_{\nu_1}(\lambda_i); ...; J_{\nu_q}(\lambda_i)\}$ *and* $J_k(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix} \in C^{k \times k}$

□

The invariant and canonical form may be computed algebraically by use of the Smith-form of $sI - A$, or by alternative means based on sequences of numbers

[Kar. & Kal.,1]. The maximum of the geometric multiplicities of eigenvalues is denoted by μ and referred to as the Segre index.

The similarity invariants define the nature of elementary motions of $S(A)$ and characterise stability properties. For eigenvalues on the imaginary axis, it is essential to compute the corresponding Segre characteristics, since it defines the difference between Lyapunov stability and instability. The Segre index μ (maximum of q for all eigenvalues) defines the minimum number of inputs, outputs which are needed for controllability, observability, when inputs and outputs are to be selected.

If $\nu(\lambda)$ is the maximal value in $\mathcal{S}(A,\lambda)$, then $\bar{n} = \sum \nu(\lambda)$ defines the degree of the minimal polynomial of A. Alternative canonical forms, such as those of the companion type may be found in [Gan.,1].

2.6.2.2 State and input coordinate transformations on S(A,B)

The invariants and canonical forms of $S(A,B)$ under the action of $\mathcal{H}_c^{is}$ are reviewed. The canonical forms have important implications in identification and state space design. Throughout this section we assume that $S(A,B)$ has n state, l inputs and $\text{rank}(B) = l$.

For the pair (A,B) we define the sequence of matrices

$$Q_{c,k} \equiv \begin{bmatrix} B, & AB, & \dots & A^k B \end{bmatrix}, k = 0,1,2,\dots \tag{2.71}$$

where $Q_{c,n-1} \equiv Q_c$ is the controllability matrix and $\text{rank}(Q_{c,k}) \leq \text{rank}(Q_{c,k+1})$.

Definition 2.9 : [Kai.,1] *The smallest integer μ for which $rank(Q_{c,\mu}) = rank(Q_{c,\mu+1})$ is defined as the controllability index of $S(A,B)$. If we assume that the linearly independent columns of Q_c in order from left to right have been found and rearrange these independent columns as*

$$\underline{b}_1, A\underline{b}_1, \dots, A_{\mu_1 - 1}\underline{b}_1; \cdots ; \underline{b}_l, A\underline{b}_l, \dots, A_{\mu_l - 1}\underline{b}_l \tag{2.72}$$

then the set of indices $\{\mu_i, i = 1,2,\dots,l\}$ are called the controllability indices of $S(A,B)$.

□

Some important properties of these indices are summarised by the following result [Kai.,1] [Chen,1] [Kar. & Mac.,1]. Note $\mu_i \geq 1$, for all $i = 1,2,\dots l$ and the zero value appears only when $\text{rank}(B) < l$.

Theorem 2.16 *For the set* $I_c = \{\mu_i, i = 1, 2, ..., l\}$ *of controllability indices of* $S(A, B)$ *the following hold true:*

- *The controllability index* $\mu = max\{\mu_1, \mu_2, ..., \mu_l\}$*;*
- *If* $\bar{n}$ *is the degree of the minimal polynomial of* A*, then*

$$\frac{n}{l} \leq \mu \leq \min(\bar{n}, n - l + 1) \leq n - l + 1; \tag{2.73}$$

- $\mu_1 + \mu_2 + \cdots + \mu_l \leq n$ *and equality holds if and only if the system is controllable. Furthermore,* $\sum_{i=1}^{l} \mu_i \equiv n_c$ *is the dimension of the controllable space of the system and* $n - n_c$ *defines the total number of uncontrollable modes;*
- *The controllability indices are invariant under state and input coordinate transformations and state feedback;*
- *The set* I_c *is the same as the set of column minimal indices of the pencil* $P_c(s) = [sI - A, -B]$*;*
- *The set* $I_c = \{\mu_i, i = 1, 2, ..., l\}$ *defines the set of column minimal indices* $\{\widetilde{\mu_i}\}$ *of the pencil* $R_c(s) = sN - NA$ *by the rule* $\widetilde{\mu_i} = \mu_i - 1, i = 1, 2, ..., l$ [Kar. & Mac.,1]*;*
- *If* $G(s) = N(s)D^{-1}(s)$ *is any* $R[s]$*-right coprime MFD with* $D(s)$ *column reduced and* $S(A, B, C)$ *is a minimal realisation of* $G(s)$ *(assume* $G(s)$ *strictly proper), then the column degrees of* $D(s)$ *define the controllability indices of* $S(A, B)$*.*

□

The set of controllability indices and the set of finite elementary divisors of $P_c(s)$ pencil are invariant under $\mathcal{H}_c^{is}$ group, but they are not complete, i.e. more invariants are needed to define a complete set.

Defining a complete set of invariants for $S(A, B)$ under $\mathcal{H}_c^{s}, \mathcal{H}_c^{is}$ groups is related to the theory of canonical form, which is extensively treated in [Kai.,1]. The Popov canonical form [Pop.,1], is a unique form under similarity of $S(A, B)$. Such a canonical form contains all additional information about the new invariants, which are now a set of real numbers. We may illustrate the structure of this canonical form in terms of an example. Thus consider a controllable system with $n = 5, \mu_1 = 2, \mu_2 = 3, l = 2$. The Popov canonical form has the following general structure, where $\times$s denote uniquely defined nonzero constants.

$$A_c = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ \times & \times & \times & \times & \times \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \times & \times & \times & \times & \times \end{bmatrix} \quad B_c = \begin{bmatrix} 0 & 0 \\ 1 & \times \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{2.74}$$

The subsystems defined by the diagonal blocks are intercoupled and this coupling is defined by the crate diagram [Kai.,1] and it is an invariant. This unique canonical form is a useful tool in system identification. If input coordinate transformations are also used, then the canonical form has an identical A matrix, A_c, but the $\times$s in B_c are eliminated. Different types of "pseudo canonical" forms (not unique) exist in the literature [Kai.,1], which once more demonstrates the minimal indices structure of the pair, but not all the nonzero elements are uniquely defined.

2.6.2.3 State output coordinate transformation on $S(A, B)$

The definitions and results presented for (A, B) pairs have their equivalents for the case of (A, C) pairs by using transposition duality arguments, that is (A^T, C^T) is first seen as a state, input pair and by transposition and use of the changes: controllability $\Longleftrightarrow$ observability, right MFD $\Longleftrightarrow$ left MFD, input $\Leftrightarrow$ output etc, all definitions and results may be stated for state, output pairs (A, C). The set of observability indices is denoted by $I_o = \{\theta_i, i = 1, 2, ..., m\}$, where m is the number of outputs and θ denotes the observability index which now satisfies the following inequality

$$\frac{n}{m} \leq \theta \leq min(\bar{n}, n - m + 1) \leq n - m + 1 \tag{2.75}$$

where $\bar{n}$ is the degree of the minimal polynomial.

2.6.2.4 State coordinate transformations on $S(A, B, C)$

Because $S(A, B)$ and $S(A, C)$ are subsystems of $S(A, B, C)$, so the sets of controllability, observability indices are invariants as well as the sets of input, output decoupling zeros and finite, infinite zeros. Also the additional invariants which will be defined under the Kronecker group $\mathcal{H}_k$ are also invariant under $\mathcal{H}_c^s$, since $\mathcal{H}_c^s$ is a subgroup of $\mathcal{H}_k$.

If Q is a transformation which brings (A, B) to Popov form (A_c, B_c) as defined earlier, then the output $C_c = CQ$ is uniquely defined and (A_c, B_c, C_c) is an input based canonical form. Similarly, if Q' is a transformation that brings (A, C)

to the corresponding Popov form (A_o, C_o), then $B_o = Q^{-1}B$ is uniquely defined and (A_o, B_o, C_o) is an output based canonical form. The Popov canonical forms (A_c, B_c, C_c) and (A_o, B_o, C_o) are related to the realisation of transfer functions based on canonical right, left MFDs, which are in an "echelon type form" [Kai.,1]. Alternative canonical forms based on balancing the controllability and observability Grammians have also been defined and give important insight in model reduction.

The canonical forms and invariants under coordinate transformations are important in system parametrisation, identification and model reduction. Some of them provide convenient forms for state space design.

2.6.3 Invariants and canonical forms under coordinate transformations and feedback

Possible feedback schemes are state feedback, output feedback and output injection. The invariants and canonical forms with state feedback and output injection are well established while those associated with output feedback are still in development. So the main emphasis is placed on the invariants and canonical forms under state feedback and output injection.

2.6.3.1 Coordinate transformations and state feedback on $S(A, B)$

Under the action of $\mathcal{H}_B^r$ group (input, state coordinate transformations and state feedback) on $S(A, B)$ systems, an equivalence class of systems $E_B(A, B)$, which will be referred to as the Brunovsky orbit of $S(A, B)$, can be obtained. If $I_c = \{\mu_i, i = 1, 2, ..., l\}$ are the set of controllability indices, or equivalently column minimal indices of $P_c(s)$, and $D_{ID} = \{(s - \lambda_i)^{\tau_i}, s \in C, i = 1, 2, ..., k\}$ is the set of finite elementary divisor of $P_c(s)$, then we may summarise the properties of $E_B(A, B)$ as follows [Bru.,1], [Kal.,2], [Kar. & Mac.,1].

Theorem 2.17 *For the Brunovsky orbit $E_B(A, B)$, the following hold true:*

1. *The sets I_c, D_{ID} are complete and independent invariants of $E_B(A, B)$;*

2. *There is a uniquely defined canonical form, the generalised Brunovsky form, $S(A_B, B_B)$, which in pencil form is described by*

$$P_c^B(s) = \begin{bmatrix} sI - A_B & -B_B \end{bmatrix} = \begin{bmatrix} sI - A_c & 0 & -B_c \\ 0 & sI - A_{ID} & 0 \end{bmatrix} \quad (2.76)$$

where $A_{ID} = diag\{J_{\tau_i}(\lambda_i), i = 1,2,...,k\}$, $J_{\tau_i}(\lambda_i)$ *is the Jordan Block associated with* $(s-\lambda_i)^{\tau_i}$, $A_c = diag\{\mathcal{H}_j, j = 1,2,...,\mu_1\}$, $\mathcal{H}_j$ *is the* $j\times j$ *standard nilpotent matrix and* $B_c = bk\text{-}diag\{\underline{w}_j, j = 1,...,\mu_1\}$ *with* $\underline{w}_j = \begin{bmatrix} 0, & 0, & ..., & 1 \end{bmatrix}^T \in R^j$.

□

$S(A_B, B_B)$ is the controllable subsystem. So if $S(A, B)$ is controllable, then $sI - A_{ID}$ will disappear in (2.76). The controllability indices and the structure and values of input decoupling zeros are the only invariants under $\mathcal{H}_B^r$.

Remark 2.2: Controllability indices are essential for system identification and study of control theory problems such as assignment of Jordan forms by state feedback [Ros.,1], structure and parametrisation of controllable subspaces [Won.,1], etc.

□

2.6.3.2 Coordinate transformation and output injection

The results in the previous subsection have their dual counterparts for Brunovsky orbit $E_B(A, C)$, obtained from $S(A, C)$ under $\mathcal{H}_B^l$. The duality is based on transposition. So elementary divisors of $P_o(s)$ and D_{op} which define the structure of the output decoupling zeros and observability indices are complete invariants and the corresponding canonical form can be obtained from the transposition of (A_B, B_B).

2.6.3.3 Coordinate transformations, state feedback and output injection on $S(A, B, C, D)$: Kronecker invariants and canonical form

With the action of the Kronecker group $\mathcal{H}_k$ on $S(A, B, C, D)$ which is referred to as the Kronecker orbit of $S(A, B, C, D)$ can be formed. The invariants and the canonical form is defined on the pencil of the system, which is

$$P(s) = \begin{bmatrix} sI - A & -B \\ -C & -D \end{bmatrix} \tag{2.77}$$

with the assumption that $P(s)$ is characterised by Kronecker invariants [Gan.,1], which are defined below.

Definition 2.10 : *For the system* $S(A, B, C, D)$ *described by* $P(s)$, *we define:*

- $D_z = \{(s - z_i)^{\tau_i}, i \in \Pi\}$ *the set of finite elementary divisors, which defines the finite zero structure of* $S(A, B, C, D)$; *the number* $n_f = \sum_{i=1}^{\pi} \tau_i$ *is called the finite zero order of the system.*

- $D_\infty = \{\hat{s}^{q_i} : 1 = q_1 = ... = q_\delta < q_{\delta+1} < q_\sigma\}$ *the set of infinite elementary divisors of the system. For those* $q_i = 1$, *they are called linear infinite zero divisors and those of* $\hat{s}^q, q > 1$ *are called nonlinear infinite zero divisors. The number* $n_\infty = \sum_{i=1}^{\sigma}(q_i - 1)$ *is defined as the infinite zero order of the system.*

- $I_r = \{\epsilon_i : 0 \leq \epsilon_1 \leq \epsilon_2 \leq \cdots \leq \epsilon_p\}, I_l = \{\eta_i, 0 \leq \eta_1 \leq \cdots \leq \eta_t\}$ *are the sets of column minimal indices, row minimal indices of* $P(s)$, *respectively and they are called the right, left indices of the system. The numbers* $n_r = \sum_{i=1}^{p} \epsilon_i, n_l = \sum_{i=1}^{t} \eta_i$ *are called the right-, left-orders of the system, respectively.*

□

The finite and infinite zero structure is characterised in physical terms by frequency transmission problems. The right and left indices are associated with the blocking of families of signals which are not necessarily of the simple exponential type [Kar. & Kou.,1].

The importance of the D_Z, D_∞, I_r and I_l sets defined on $S(A, B, C, D)$ is described below [Tho.,1] [Mor.,1] [Kar. & Mac.,1].

Theorem 2.18 *For the Kronecker orbit* $E_k(A, B, C, D)$, *the following hold true.*

- *The set* $\{D_z, D_\infty, I_r, I_l\}$ *defined on* $S(A, B, C, D)$ *is a complete and independent invariant.*

- *There is a uniquely defined canonical form, the Kronecker canonical form* $S(A_k, B_k, C_k, D_k)$ *which in pencil form is presented as*

$$P_k = \begin{bmatrix} sI - A_k & -B_k \\ -C_k & -D_k \end{bmatrix}$$
$$= \begin{bmatrix} sI - A_c & 0 & 0 & 0 & -B_c & 0 & 0 \\ 0 & sI - A_\eta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & sI - A_\infty & 0 & 0 & -B_\infty & 0 \\ 0 & 0 & 0 & sI - A_f & 0 & 0 & 0 \\ 0 & -C_\eta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -C_\infty & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -I_\delta \end{bmatrix} \quad (2.78)$$

where

$$A_c = diag\{A_j : j = \epsilon_1, ..., \epsilon_p\} \in R^{n_r \times n_r}$$
$$A_\eta = diag\{A_j : j = \eta_1, ..., \eta_t\} \in R^{n_l \times n_l}$$

$$A_\infty = diag\{A_j : j = f_1, ..., f_{\sigma-\delta}, f_i = q_i - 1, i = \delta+1, ..., \sigma\} \in R^{n_o \times n_o}$$
$$A_f = diag\{J_{\tau_i}(z_i) : i = 1, 2, ..., \pi\} \in R^{n_f \times n_f}$$

where $J_\tau(z_i)$ *are Jordan blocks characterising* $(s - z_i)^{\tau_i}$, $A_j = H_j$ *is the* $j \times j$ *standard nilpotent matrix and*

$$\begin{aligned} C_\eta &= bl.diag\{\underline{v}_j^T : j = \eta_1, ..., \eta_t\} \in R^{t \times n_l} \\ C_\infty &= bl.diag\{\underline{v}_j^T : j = f_1, ..., f_{\sigma-\delta}\} \in R^{t(\sigma-\delta) \times n_\infty} \\ B_\infty &= bl.diag\{\underline{w}_j : j = f_1, ..., f_{\sigma-\delta}\} \in R^{n_o \times (\sigma-\delta)} \\ B_c &= bl.diag\{\underline{w}_j : j = \epsilon_1, ..., \epsilon_p\} \in R^{n_r \times p} \end{aligned}$$

where $\underline{v}_j^T = [1, 0, ...,] \in R^{1 \times j}$ *and* $\underline{w}_j = [0, 0, ..., 0, 1]^T \in R^{j \times 1}$

- *If* r, ρ_G *are the ranks of* $P(s), G(s)$, *respectively, then the following relationships hold true among the numbers of the invariants*

  (a) $r = n + \rho_G, p = l - \rho_G, t = m - \rho_G, n = n_f + n_\infty + n_r + n_l$;

  (b) $\sigma = \rho_G$, *and* $\delta = \rho(D)$;

  (c) There are zero cmi and zero rmi, if and only if $[B^T D^T]^T, [C, D]$ *are rank deficient, respectively.*

  (d) The transfer function matrix of $S(A_k, B_k, C_k, D_k)$ *is*

$$\begin{aligned} G_k(s) &= C_k(sI - A_k)^{-1}B_k + D_k = \begin{bmatrix} M_\infty^*(s) & 0_{\sigma,p} \\ 0_{t,\sigma} & 0_{t,p} \end{bmatrix} \\ M_\infty^*(s) &= diag\{s^{1-q_1}, ..., s^{1-q_\mu}\} \end{aligned} \tag{2.79}$$

where $G_k(s)$ *is the Smith form at* $s = \infty$ *of* $G(s)$ [Var. Lim. & Kar.,1].

□

The above summarise the results demonstrating the structure of state space models under the most general types of transformations that may be applied on state space models. The results may be simplified in an obvious manner for strictly proper systems, where $D = 0$.

Remark 2.3: The number of divisors at infinity of $P(s)$ is equal to the rank of $G(s)$. There exists a number of linear divisors at infinity equal to the rank of D; for strictly proper systems, all divisors at infinity are nonlinear, i.e., $q_i \geq 2$. The orders of infinite zeros are defined by $f_i = q_i - 1$, when q_i are the degrees of nonlinear divisors at infinity. The f_i define the generic asymptotic locus pattern and terminal Nyquist phases. If rank(D) = rank$(G(s))$, then $G(s)$ has no infinite zeros, or equivalently all q_i's are equal to 1. For strictly proper, square systems with $m = l$, all orders of infinite zeros of $G(s)$ are $f_i = 1$ if and only if rank$(CB) = m = l$; higher order of infinite zeros emerge when rank$(CB) < m = l$.

□

Remark 2.4: The Kronecker form $S(A_k, B_k, C_k, D_k)$ is maximally uncontrollable and unobservable and the dimension of the minimal system is defined by the infinite zero order [Kar. & Mac.,1]. State feedback and output injection are equivalent to post- and pre-multiplication of transfer function by $R_{pr}(s)$-unimodular matrices [Hau. & Hey.,1]; the special element of $\mathcal{H}_k$, that reduces S to its Kronecker form, is equivalent to a pair of $R_{pr}(s)$-unimodular matrices which reduces $G(s)$ to its Smith form at $s = \infty$ of $G(s)$ [Var. Lim. & Kar.,1].

□

Remark 2.5: For right regular systems rank$(CB) = l, n_r = 0$ (no right indices) and for left regular systems rank$(CB) = m, n_l = 0$ (no left indices). For left-right regular systems rank$(CB) = m = l$ (square nondegenerate systems), $n_r = n_l = 0$ and

$$n_f + n_\infty = n \tag{2.80}$$

which shows that total number of finite and infinite zeros is equal to the dimension of the state space. For such systems, the total number of finite zeros satisfies the conditions:

1. $D \neq 0 : n_f < n$, if and only if rank$(D) = m = l$;
2. $D = 0 : n_f < n - m = n - l$ and equality holds if and only if rank$(CB) = m = l$.

For strictly proper square systems, the number $n - m = n - l$ defines an upper bound on the total number of finite zeros.

The right and left indices are related to synthesis problems such as squaring down, model matching etc.

□

Remark 2.6: The finite zeros of $P(s)$ and $Z(s)$ (zero pencil) are the same. If $q'_i, i = 1, 2, ..., \tau$ are the degrees of the divisors at $s = \infty$, with $q'_i \geq 3$ of $P(s)$, then the degrees of the restricted zero divisors of $Z(s)$ are $q'_i - 2$.

□

2.7 Summary

In this chapter, some of the fundamental linear system properties, property indicators and system invariants in both the time domain and the frequency domain have been surveyed. These properties, property indicators and system invariants provide some of the main tools for system analysis and system design. Some of the well-known qualitative system indicators, such as controllability and observability can be developed into quantitative indicators. Further they can be used as indicators for input-output interactions. These will be examined in the next chapter.

Chapter 3

FURTHER QUANTITATIVE PROPERTIES AND INDICATORS

3.1 Introduction

In Chapter 2, some of the basic system properties and the corresponding property indicators have been reviewed. Properties such as system controllability, observability and output disturbability are qualitative properties.

The qualitative properties of a system are important in the sense that they reveal the capabilities and limitations of the system. For instance, if a system is state controllable, then there always exists a certain input signal with which any given initial state of a system can be brought to zero in a finite time interval. Otherwise, if a system is not state controllable, then there exists a set of system initial conditions which cannot be brought to zero with any control signal in a finite time.

In reality, questions such as "how close is a system to being controllable or uncontrollable" are asked. The distance of a controllable from an uncontrollable one when subject to complex and real perturbations has been studied [Bol. & Lu,1].

In connection with the energy of the control signals, the quantitative side of the system properties can be studied. In Sections 3.2, the quantitative controllability, observability are characterised in terms of the singular values of the controllability Grammian and observability Grammian in relation to the minimal energy needed in control. So the singular values of the controllability Grammian are the indicators of quantitative controllability while the singular values of the observability Grammian are indicators of quantitative observability. Quantitative output tracking and output disturbability have also been defined and the singular values of correspondingly defined matrices are developed as indicators.

In Section 3.3, the quantitative state controllability indicators are employed to characterise the set of initial conditions which can be brought to the origin by a bounded energy control. This subset is proved to be a hyper-elipsoid. In fact, the initial conditions can be characterised by the singular values and the amount of energy for control.

In Section 3.4, the problem of selecting the inputs and outputs of a control system is discussed. An empirical criteria for coupling the outputs to the inputs are reviewed. Then the quantitative output controllability developed in Section 3.2 is further developed as a new interaction indicator and employed as a criterion for selecting the appropriate input-output pairing scheme in decentralised control. Finally in Section 3.5, existing indicators for eigenvalue mobility with respect to both state-feedback and output feedback gain will be reviewed.

3.2 Quantitative controllabilities and indicators

3.2.1 Output controllability

Let a system described by state space model (A, B, C, D, H, J) be as follows:

$$\begin{aligned} \dot{\underline{x}} &= A\underline{x} + B\underline{u} + H\underline{\omega} \\ \underline{y} &= C\underline{x} + D\underline{u} + J\underline{\omega} \end{aligned} \tag{3.1}$$

where $A \in R^{n\times n}, B \in R^{n\times l}, C \in R^{m\times n}, D \in R^{m\times l}, H \in R^{n\times p}, J \in R^{m\times p}, \underline{x} \in R^n, \underline{u} \in R^l, \underline{y} \in R^m$ and $\underline{\omega} \in R^p$.

As defined in Definition (2.3), if a system is output controllable, there always exists an input $\underline{u}(t)$ such that for an arbitrarily specified $\underline{y}(t_f)$ can be reached for an arbitrary $\underline{y}(t_0)$. In the following we try to find among all the possible inputs a particular input that needs the minimum energy and subsequently investigate the relationship between the minimum energy and the output controllability Grammian.

The output of the system (3.1) at t_f is given as

$$\begin{aligned} \underline{y}(t_f) &= C\Phi(t_f, t_0)\underline{x}(t_0) + \int_{t_0}^{t_f} C(t_f)\Phi(t_f, \sigma)\left[B(\sigma)\underline{u}(\sigma) + H(\sigma)\underline{\omega}(\sigma)\right] d\sigma + \\ &\quad + D\underline{u}(t_f) + J\underline{\omega}(t_f) \end{aligned} \tag{3.2}$$

where $\Phi(t_f, t_0)$ is the state transition matrix. In the case of linear time invariant systems, $\Phi(t_f, t_0) = e^{A(t_f - t_0)}$.

For the given system with the initial and final output values, the input signal to yield this output can be found in the sequel [Ske.,1]. First we find the set of feasible initial state conditions $\underline{x}(t_0)$. From equation (3.2), the initial output condition $\underline{y}(t_0)$ must satisfy the following condition:

$$\underline{y}(t_0) = C(t_0)\underline{x}(t_0) + D(t_0)\underline{u}(t_0) + J(t_0)\omega(t_0) \tag{3.3}$$

so for a feasible initial state, the above equation will have a solution for $\underline{x}(t_0)$ and so is the following equation

$$\begin{bmatrix} C(t_0) & D(t_0) \end{bmatrix} \begin{bmatrix} \underline{x}(t_0) \\ \underline{u}(t_0) \end{bmatrix} = \underline{y}(t_0) - J(t_0)\underline{\omega}(t_0) \tag{3.4}$$

The matrix $\begin{bmatrix} C(t_0) & D(t_0) \end{bmatrix}$ will have a right inverse. However, the solution to the initial values $\underline{x}(t_0)$ and $\underline{u}(t_0)$ might not be unique. If there are more than one solution, we can select one arbitrarily.

Having found the feasible conditions for the state variables, the input to transfer the output from $\underline{y}(t_0)$ to $\underline{y}(t_f)$ can be calculated as follows:

$$\begin{aligned} &\textstyle\int_{t_0}^{t_f} \left[C(t_f)\Phi(t_f,\sigma)B(\sigma) + D(t_f)\delta(t_f - \sigma)\right] \underline{u}(\sigma)d\sigma = \\ &\underline{y}(t_f) - C(t_f)\Phi(t_f,t_0)\underline{x}(t_0) - J(t_f)\underline{\omega}(t_f) - \textstyle\int_{t_0}^{t_f} C(t_f)\Phi(t_f,\sigma)H(\sigma)\underline{\omega}(\sigma)d\sigma \end{aligned} \tag{3.5}$$

Define

$$\begin{aligned} \underline{y}^o(t_f) &\equiv \underline{y}(t_f) - C(t_f)\Phi(t_f,t_0)\underline{x}(t_0) - J(t_f)\underline{\omega}(t_f) - \int_{t_0}^{t_f} C(t_f)\Phi(t_f,\sigma)H(\sigma)\underline{\omega}(\sigma)d\sigma \\ \mathcal{G}(\sigma) &\equiv C(t_f)\Phi(t_f,\sigma)B(\sigma) + D(t_f)\delta(t_f - \sigma) \end{aligned} \tag{3.6}$$

then

$$\underline{y}^o(t_f) = \int_{t_0}^{t_f} \mathcal{G}(\sigma)\underline{u}(\sigma)d\sigma \tag{3.7}$$

From the definition, $\underline{y}^o(t_f)$ is assumed to be an arbitrary vector. It can be verified by direct substitution that equation (3.7) is solved by

$$\underline{u}(\sigma) = \mathcal{G}^T(\sigma)\left[\ \textstyle\int_{t_0}^{t_f} \mathcal{G}(\epsilon)\mathcal{G}^T(\epsilon)d\epsilon\ \right]^{-1} \underline{y}^o(t_f) \tag{3.8}$$

From equation (3.8), in order to have a solution for $\underline{u}(\sigma)$

$$\mathcal{Y}(t_f, t_0) \equiv \int_{t_0}^{t_f} \mathcal{G}(\sigma)\mathcal{G}^T(\sigma)d\sigma \tag{3.9}$$

must be invertible.

The necessity for $\mathcal{Y}(t_f, t_0)$ to be invertible can be proved by showing that for any control which transfers $\underline{x}(t_0)$ to $\underline{x}(t_f)$, $\mathcal{Y}(t_f, t_0)$ will have to be invertible. Rewrite equation (3.7) as:

$$\underline{y}^o(t_f) = \int_{t_0}^{t_f} \begin{bmatrix} g_1^T(\sigma)\underline{u}(\sigma)d\sigma \\ \vdots \\ g_p^T(\sigma)\underline{u}(\sigma)d\sigma \end{bmatrix}, \quad \mathcal{G}^T(\sigma) = \begin{bmatrix} g_1(\sigma) & \cdots & g_p(\sigma) \end{bmatrix} \tag{3.10}$$

where p is the number of outputs.

Because $\underline{y}^o(t_f)$ is arbitrary, then each element $\underline{y}_i^o(t_f), i = 1, 2, ..., p$ is arbitrary. Thus a certain linear combination of the elements of the vector $g_i(\sigma)$ gives a different value $\underline{y}_i^o(t_f)$ for each i. This requires the vectors $g_i(\sigma), i = 1, 2, ..., p$ to be linearly independent on the interval $\sigma \in [t_0, t_f]$, which is equivalent to

$$\int_{t_0}^{t_f} \mathcal{G}(\sigma)\mathcal{G}^T(\sigma)\underline{u}(\sigma)d\sigma > 0 \tag{3.11}$$

Define the output controllability Grammian matrix as

$$\mathcal{G}_{oc}(t_0, t_f) \equiv \int_{t_0}^{t_f} C(t_f)\Phi(t_f, \sigma)B(\sigma)B^T(\sigma)\Phi^T(t_f, \sigma)C^T(t_f)d\sigma \tag{3.12}$$

Then equation (3.9) becomes

$$\mathcal{Y}(t_f, t_0) = D(t_f)D^T(t_f)\delta + D(t_f)B^T(t_f)C^T(t_f) + C(t_f)B(t_f)D^T(t_f) + \mathcal{G}_{oc}(t_0, t_f) \tag{3.13}$$

where δ is the Dirac delta function.

So we have the following result.

Theorem 3.1 [Ske.,1] *The system (3.1) is output controllable at t_0 if equation (3.13) has the property $\mathcal{Y}(t_f, t_0) > 0$ for some $t_f > t_0$. If the output is controllable, then one control that transfers $\underline{y}(t_0)$ to $\underline{y}(t_f)$ is given by equation (3.8).*

□

For strictly proper linear time invariant systems, the matrices are constants and the output controllability Grammian of the system becomes

$$\mathcal{G}_{oc}(t_0, t_f) = C \left[\int_{t_0}^{t_f} e^{A(t_f-\epsilon)}BB^Te^{A^T(t_f-\epsilon)}d\epsilon \right] C^T = C \left[\int_0^{t_f-t_0} e^{A\sigma}BB^Te^{A^T\sigma}d\sigma \right] C^T \tag{3.14}$$

and the control $\underline{u}(\sigma)$ is given as

$$\begin{aligned} \underline{u}(\sigma) &= B^T\Phi^T(t_f, \sigma)C^T \left[\int_{t_0}^{t_f} C\Phi(t_f, \sigma)BB^T\Phi^T(t_f, \sigma)C^Td\sigma \right]^{-1} \underline{y}^o(t_f) \\ &= B^Te^{A^T(t_f-\sigma)}C^T \left[C(\int_0^{t_f-t_0} e^{A\sigma}BB^Te^{A^T\sigma}d\sigma)C^T \right]^{-1} \underline{y}^o(t_f) \\ &= \mathcal{G}^T(\sigma)\mathcal{G}_{oc}^{-1}(t_f, t_0)\underline{y}^o(t_f) \end{aligned} \tag{3.15}$$

The energy of the control which drives the system from $\underline{y}(t_0)$ at $t = t_0$ to $\underline{y}(t_f)$ at $t = t_f$ is given in the following:

$$
\begin{aligned}
E &= \int_{t_0}^{t_f} \underline{u}^T(\sigma)\underline{u}(\sigma)d\sigma \\
&= \int_{t_0}^{t_f} \left\{ \underline{y}^{oT}(t_f) \left[\int_{t_0}^{t_f} \mathcal{G}(\epsilon)\mathcal{G}^T(\epsilon)d\epsilon \right]^{-1} \mathcal{G}(\sigma)\mathcal{G}^T(\sigma) \left[\int_{t_0}^{t_f} \mathcal{G}(\epsilon)\mathcal{G}^T(\epsilon)d\epsilon \right]^{-1} \underline{y}^o(t_f) \right\} d\sigma \\
&= (\underline{y}^o)^T(t_f)\mathcal{G}_{oc}^{-1}(t_f, t_0)\underline{y}^o(t_f) \qquad (3.16)
\end{aligned}
$$

Because the system is output controllable, the output controllability Grammian, $\mathcal{G}_{oc}(t_f, t_0)$, is a symmetric positive definite real matrix. For a symmetric positive definite real matrix, we have the following theorem.

Theorem 3.2 [Gan.,1] *Given a real, symmetric and positive definite matrix $\mathcal{G}$, there always exists a set of orthonormal eigenvectors $\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_n$ with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$. Set $Q = \left[\begin{array}{cccc} \underline{u}_1 & \underline{u}_2 & \ldots & \underline{u}_n \end{array} \right]$, then*

$$
Q^T\mathcal{G}Q = diag(\lambda_1, \lambda_2, \ldots, \lambda_n) \quad or \quad \mathcal{G} = Q\,diag(\lambda_1, \lambda_2, \ldots, \lambda_n)Q^T \qquad (3.17)
$$

where for the transformation matrix Q satisfies the following

$$
\begin{aligned}
QQ^T &= I \\
Q^{-1} &= Q^T \\
|Q| &= 1.
\end{aligned}
$$

□

Further, the quadratic form defined by the matrix A has the following property.

Theorem 3.3 [Gan.,1] *If matrix A is a real symmetric and positive definite, then the quadratic defined by the matrix satisfies*

$$
\underline{x}^T A \underline{x} \leq \lambda_1 \underline{x}^T \underline{x} \qquad (3.18)
$$

where λ_1 is the largest of the eigenvalues.

□

Assuming that the output controllability of the system has a set of singular values as $\sigma_1 \geq \sigma_2 \geq ..., \geq \sigma_n$, then there exists an orthonormal transformation U such that

$$
\mathcal{G}_{oc}(t_f, t_0) = U\text{diag}\left\{ \begin{array}{cccc} \sigma_1, & \sigma_2, & \cdots, & \sigma_n \end{array} \right\} U^T \qquad (3.19)
$$

and the inverse of which is

$$\mathcal{G}_{oc}^{-1}(t_f, t_0) = U\text{diag}\left\{ \sigma_1^{-1}, \quad \sigma_2^{-1}, \quad \cdots, \quad \sigma_n^{-1} \right\} U^T \tag{3.20}$$

so equation (3.16) satisfies

$$E = \int_{t_0}^{t_f} \underline{u}^T(\sigma)\underline{u}(\sigma)d\sigma \leq (1/\sigma_n)\underline{y}^{oT}(t_f)\underline{y}^o(t_f) \tag{3.21}$$

so the singular values of the output controllability Grammian of a system is a very important measure. When the smallest singular value is big, then the energy consumed in transferring the outputs from $\underline{y}(t_0)$ to $\underline{y}(t_f)$ will be small. It is shown that the shorter the time available for control action, the more the energy is needed to steer the output from $\underline{y}(t_0)$ to $\underline{y}(t_f)$ [Sei.,1]. This can also be demonstrated by looking at the singular values of the finite time output controllability Grammian. Indeed, it is shown in the following examples, the shorter the time available (smaller t_f in our case), the smaller the singular values of the output controllability Grammian. In summary, we have the following proposition.

Proposition 3.1 *The singular values of the output controllability Grammian of a linear time invariant system are important indicators for the energy needed to transform from one output state to another. In particular, the minimum energy needed to transform from one output state to another is reciprocal to the minimal singular value.*

□

On computing the singular values of the controllability Grammian, we first have to find the state transition matrix of the system, $\Phi(t, t_0)$. For a time varying system, commercial numerical packages should be used. However, in the case of linear time invariant system, by using the Jordan canonical form of the system we shall be able to find analytic solution to the state transition matrix [Gan.,1]. When the order of the system is low, the following equations can be employed to compute the Grammian at finite time t_f.

$$X(t) = \int_{t_0}^{t_f} e^{A\sigma}BB^Te^{A^T\sigma}d\sigma, \quad X(0) = 0 \tag{3.22}$$

set $e^{A\sigma} = Y(\sigma)$ so

$$\dot{Y}(\sigma) = Ae^{A\sigma} = AY \tag{3.23}$$

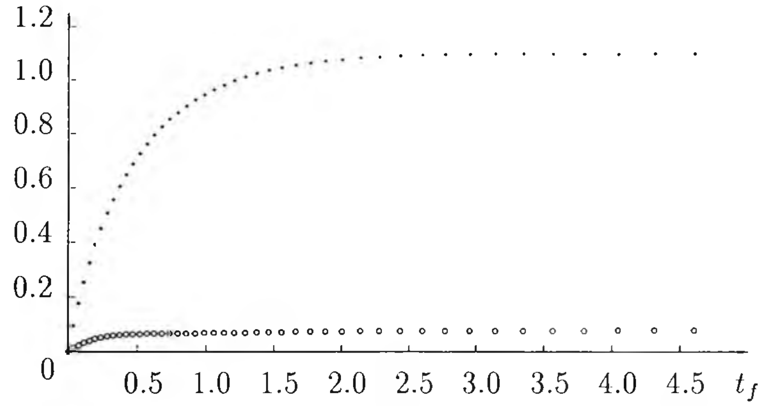


Figure 3.1: Singular Values of the Output Controllability Grammian of Example 3.1

with initial condition $Y(0) = e^{A0} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$ by combining the above equations together, we have

$$\begin{cases} \dot{X}(t) = Y(t)BB^TY^T(t) & , \quad X(0) = 0 \\ \dot{Y}(t) = AY(t), & \quad Y(0) = I \end{cases} \tag{3.24}$$

so the Grammian $X(t)$ can be computed by integrating equation (3.24) from t_0 to t_f. For the following examples, we use ode23.m or ode45.m subroutines in Matlab to perform the integration [Matlab]. For a detailed description of the algorithm, see Appendix I.

In the following we give some examples.

Example (3.1): Study the following linear time invariant system which is described by state-space model as:

$$\dot{\underline{x}} = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \underline{x} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{u} \tag{3.25}$$

$$\underline{y} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \underline{x} \tag{3.26}$$

The system is output controllable, because rank$[CB\ CAB\ CA^2B] = 2$. The singular values of the output controllability Grammian are plotted against t in Figure (3.1). Clearly, the singular values are functions of the final time t_f. When t_f is small, the smallest singular value is also small. So by using (3.21), the energy needed is large.

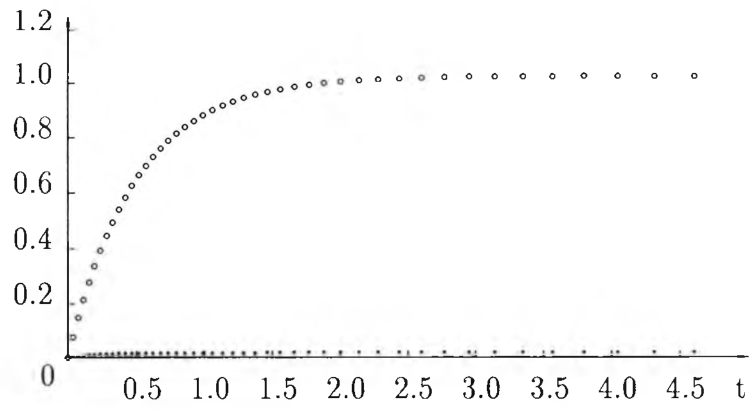


Figure 3.2: Singular Values of the Output Controllability Grammian of Example 3.2

Example (3.2): Study the following system with a slight different input matrix B.

$$\underline{\dot{x}} = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \underline{x} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 0.5 \end{bmatrix} \underline{u} \tag{3.27}$$

$$\underline{y} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \underline{x} \tag{3.28}$$

The singular values of the output controllability Grammian is shown in Figure (3.2). This system is also output controllable. However, the control will be more difficult than the first one in the sense that more energy is needed on transferring the same initial state. By examining the smallest singular value of the output controllability Grammian, it is smaller compared with the first system.

3.2.2 State controllability

From the definition (2.3), the system (3.1) is state controllable at t_0 if and only is $\mathcal{G}_{sc}(t_0, t_f) > 0$, for some $t_f > t_0$, where $\mathcal{G}_{sc}(t_0, t_f)$ is defined by equation (3.12) except $C = I$ [Ske.,1] [Kai.,1] or

$$\mathcal{G}_{sc}(t_f, t_0) \equiv \int_{t_0}^{t_f} \Phi(\sigma, t_0) B(t) B^T(t) \Phi^T(\sigma, t_0) d\sigma. \tag{3.29}$$

The derivation would be the same as the output controllability by setting the output matrix $C = I$ in Section 3.2.1.

For the case of linear time invariant systems, the state controllability is given by the following theorem:

Theorem 3.4 [Kai.,1] [Ske.,1] *The time invariant system $\dot{\underline{x}} = A\underline{x} + B\underline{u}$ is state controllable if and only if the state controllability Grammian $\mathcal{G}_{sc}$ is positive definite.*

$$\mathcal{G}_{sc}(+\infty, 0) \equiv \int_0^{+\infty} e^{At} B B^T e^{A^T t} dt, \quad \mathcal{G}_{sc} > 0, \tag{3.30}$$

or equivalently

$$rank[B, AB, ..., A^{n-1}B] = n \tag{3.31}$$

□

However, the state controllability Grammian at infinity is only a qualitative measure. In the following, the energy argument is used to develop the singular values of the state controllability Grammian at a finite t_f as a quantitative controllability measure of the control. For the case when $D = 0$, one control with minimum energy to transfer the initial state $\underline{x}(t_0)$ to $\underline{x}(t_f)$ is given as:

$$\underline{u}(t) = \mathcal{F}^T(t) \left[\int_{t_0}^{t_f} \mathcal{F}(\epsilon)\mathcal{F}^T(\epsilon) d\epsilon \right]^{-1} [\underline{x}(t_f) - \underline{x}(t_0)] \tag{3.32}$$

where $\mathcal{F}(\epsilon) \equiv \Phi(\epsilon, t_0)B(\epsilon)$. So the energy of the control to transfer the system from an initial state $\underline{x}(t_0)$ to a final state $\underline{x}(t_f)$ in a finite time interval $t_f - t_0$ is

$$\begin{aligned} E &= \int_{t_0}^{t_f} \underline{u}^T(t)\underline{u}(t) dt \\ &= \int_{t_0}^{t_f} \{[\underline{x}(t_f) - \underline{x}(t_0)]^T \left[\int_{t_0}^{t_f} \mathcal{F}(\epsilon)\mathcal{F}^T(\epsilon) d\epsilon \right]^{-1} \times \\ &\quad \mathcal{F}(\sigma)\mathcal{F}^T(\sigma) \left[\int_{t_0}^{t_f} \mathcal{F}(\epsilon)\mathcal{F}^T(\epsilon) d\epsilon \right]^{-1} [\underline{x}(t_f) - \underline{x}(t_0)]\} d\sigma \\ &= [\underline{x}(t_f) - \underline{x}(t_0)]^T \left[\int_{t_0}^{t_f} \mathcal{F}(\epsilon)\mathcal{F}^T(\epsilon) d\epsilon \right]^{-1} [\underline{x}(t_f) - \underline{x}(t_0)] \end{aligned} \tag{3.33}$$

In the linear time invariant case, the energy to transfer the state $\underline{x}(t_0)$ to the final state $\underline{x}(t_f)$ in a finite time interval $t_f - t_0$ is

$$E = [\underline{x}(t_f) - \underline{x}(t_0)]^T \left[\int_0^{t_f - t_0} e^{A\sigma} B B^T e^{A^T \sigma} d\sigma \right]^{-1} [\underline{x}(t_f) - \underline{x}(t_0)] \tag{3.34}$$

Because the involvement of the inverse of the state controllability Grammian in the energy, the energy is reciprocal to the singular values of the matrix, and therefore the set of singular values is a suitable quantitative measure for the control.

Further, the measure can be used when the control action is confined to a finite time interval $t_f - t_0 < \infty$.

3.2.3 State observability

The state observability of a system is concerned with the problem of when it is possible to know the internal properties of the system, given the output $\underline{y}(t)$ and the system inputs $\underline{u}(t)$ [Kai,1, etc.]. The outputs of the following system:

$$\begin{aligned} \dot{\underline{x}}(t) &= A(t)\underline{x}(t) + B(t)\underline{u}(t) \\ \underline{y}(t) &= C(t)\underline{x}(t) + D(t)\underline{u}(t) \end{aligned} \tag{3.35}$$

is given by

$$\underline{y}(t) = C(t)\Phi(t,t_0)\underline{x}(t_0) + \int_{t_0}^{t} [D(t)\Phi(t,\sigma)B(\sigma) + C(t)\delta(t-\sigma)]\,\underline{u}(\sigma)d\sigma \tag{3.36}$$

As defined in Definition (2.3), if the states of the system are observable at time t_0, then the states of the system at any other time $t \neq t_0$ can also be deduced from the expression $\underline{x} = \Phi(t,t_0)\underline{x}(t_0) + \int_{t_0}^{t} [C(t)\Phi(t,\sigma)B(\sigma) + D(t)\delta(t-\sigma)]\,\underline{u}(\sigma)d\sigma$. Hence, the observability of the system is equivalent to the observability of the states at $t = t_0$.

Define all the known terms in equation (3.36) as

$$\mathcal{Y}(t) \equiv \underline{y}(t) - \int_{t_0}^{t} [C(t)\Phi(t,\sigma)B(\sigma) + D(t)\delta(t-\sigma)]\,\underline{u}(\sigma)d\sigma \tag{3.37}$$

then we have

$$\mathcal{Y}(t) = C(t)\Phi(t,t_0)\underline{x}(t_0) \tag{3.38}$$

which is equivalent to the following unforced system

$$\begin{aligned} \dot{\underline{x}}(t) &= A(t)\underline{x}(t), \\ \underline{y}(t) &= C(t)\underline{x}(t) \end{aligned} \tag{3.39}$$

The solution $\underline{x}(t_0)$ to equation (3.39) can be found in the following way. Multiply equation (3.39) from the left by $[C(t)\Phi(t,t_0)]^T$ and integrate:

$$\int_{t_0}^{t_f} \left| \; \Phi^T(t,t_0)C^T(t)\underline{y}(t) \; \right| dt = \int_{t_0}^{t_f} \Phi^T(t,t_0)C^T(t)C(t)\Phi(t,t_0)\underline{x}(t_0)dt \tag{3.40}$$

or

$$\underline{x}(t_0) = \left| \; \int_{t_0}^{t_f} \Phi^T(t,t_0)C^T(t)C(t)\Phi(t,t_0)dt \; \right|^{-1} \int_{t_0}^{t_f} \Phi^T(t,t_0)C^T(t)\underline{y}(t)dt \tag{3.41}$$

Note that for equation (3.41) to exist, the matrix

$$\mathcal{G}_{so}(t_0,t_f) \equiv \int_{t_0}^{t_f} \Phi^T(t,t_0)C^T(t)C(t)\Phi(t,t_0)dt \tag{3.42}$$

must be invertible. Matrix $\mathcal{G}_{so}(t_0,t_f)$ is defined to be the observability Grammian of the system. Because $\mathcal{G}_{so}(t_0,t_f)$ satisfies

$$\frac{d}{dt}\mathcal{G}_{so}(t_0,t) = -\mathcal{G}_{so}(t_0,t)A(t) - A^T(t)\mathcal{G}_{so}(t_0,t) - C^T(t)C(t), \quad \mathcal{G}_{so}(t_0,t_0) = 0 \tag{3.43}$$

and also from the structure of $\mathcal{G}_{so}(t_0, t_f), \mathcal{G}_{so}(t, t_f)$ will at least be nonnegative. Hence, if $\mathcal{G}_{so}(t_0, t_f)$ is nonsingular, it is also positive definite. So it is concluded that the system (3.35) is observable if $\mathcal{G}_{so}(t_0, t_f) > 0$. The sufficiency is established by contradiction.

Suppose the system is observable at t_0, but there exists no $t_f > t_0$ such that the columns of $C(t)\Phi(t, t_0)$ are linearly independent over $[t_0, t_f]$. Then there exists an $n \times 1$ zero constant vector $\underline{\alpha}$ such that $C(t_0)\Phi(t, t_0)\underline{\alpha} = \underline{0}$ for all $t > t_0$.

If we choose $\underline{x}(t_0) = \underline{\alpha}$; then

$$\mathcal{Y}(t) = C(t)\Phi(t, t_0)\underline{\alpha} = 0 \tag{3.44}$$

for all $t > t_0$.

Hence, the initial state $\underline{x}(t_0) = \underline{\alpha}$ cannot be detected. This contradicts the assumption that the system is observable. Therefore, if the system is observable, there exists a finite t_f such that the columns of $C(t)\Phi(t, t_0)$ are linearly independent over $[t_0, t_f]$ and $\mathcal{G}_{so}(t_0, t_f) > 0$.

Theorem 3.5 [Ske.,1] *The linear system (3.35) is observable at time t_f, if and only if the solution $\mathcal{G}_{so}(t_0, t_f)$ is positive definite.*

□

For time invariant systems, the observability of the system can be simplified; from equation (3.42),

$$\mathcal{G}_{so}(t_0, t_f) = \int_{t_0}^{t_f} e^{A^T(\sigma - t_0)} C^T C e^{A(\sigma - t_0)} d\sigma = \int_0^{t_f - t_0} e^{A^T\sigma} C^T C e^{A\sigma} d\sigma \tag{3.45}$$

if we take $e^{A\sigma} = \sum_{i=0}^{n-1} \alpha_i(\sigma) A^i$, then

$$\begin{aligned} \mathcal{G}_{so}(t_0, t_f) &= \int_0^{t_f - t_0} \sum_{i=0}^{n-1} \alpha_i(\sigma)(A^i)^T C^T C \sum_{i=0}^{n-1} \alpha_i(\sigma) A^i d\sigma \qquad (3.46) \\ &= \begin{bmatrix} C^T, & A^T C^T, & \ldots, & (A^{n-1})^T C^T \end{bmatrix} \Omega(\sigma)(t_f, t_0) \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \\ &= \omega_0^T \Omega(t_f, t_0) \omega_0 \end{aligned}$$

where Ω is defined by

$$\Omega(t_f, t_0) \equiv \int_0^{t_f - t_0} \begin{bmatrix} \alpha_0(\sigma) I_p \\ \vdots \\ \alpha_{n-1}(\sigma) I_p \end{bmatrix} \begin{bmatrix} \alpha_0(\sigma) I_p, & \ldots, & \alpha_{n-1}(\sigma) I_p \end{bmatrix} d\sigma \tag{3.47}$$

Because $\Omega(t_0, t_f)$ is positive definite, the nonsingularity condition of $\mathcal{G}_{so}(t_0, t_f)$ is satisfied if and only if ω_0 is of full rank.

The rank of the observability matrix is a qualitative measure of the system. Further quantitative measure is required in order to assess the error of the observation of $\underline{x}(t_0)$. Also when the observer of the system is gain bounded, it is important to study the observability quantitatively against the change of the gain.

The singular values of the observability Grammian presents a good quantitative measure. From equation (3.41):

$$\underline{x}(t_0) = [\mathcal{G}_{so}(t_0, t_f)]^{-1} \int_{t_0}^{t_f} \Phi^T(\sigma, t_0) C^T(\sigma) \mathcal{Y}(\sigma) d\sigma \tag{3.48}$$

where

$$\mathcal{G}_{so}(t_0, t_f) = \int_{t_0}^{t_f} \Phi^t(\sigma, t_0) C^T C(\sigma) \Phi(\sigma, t_0) d\sigma$$

and $\mathcal{G}_{so}(t_0, t_f) = \int_0^{t_f - t_0} e^{A^T \sigma} C^T C e^{A\sigma} d\sigma$ for time invariant cases. Because of the involvement of the inverse of the observability Grammian, when it is near singular, then the error in the observation will be big. When the element of the observation matrix changes, the system observability Grammian will also change, and therefore the singular values.

On calculating the singular values of the observability Grammian, it can be taken as the dual problem of calculating the singular values of the state controllability Grammian by taking the A and B matrices as $A \Leftarrow A^T, B \Leftarrow C^T$.

In the following we present two examples.

Example (3.3): Study the same system as in example (3.1)

$$\dot{\underline{x}} = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \underline{x} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{u} \tag{3.49}$$

$$\underline{y} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \underline{x} \tag{3.50}$$

The singular values of the observability Grammian against time t_f are shown in Figure (3.3). It is shown clearly that when the observation time is short, the observation will be difficult and the error of the observation will be big.

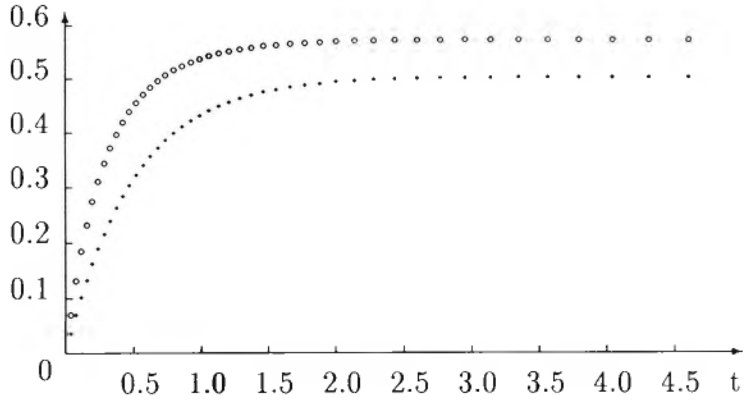


Figure 3.3: Singular values of the observability Grammian of Example 3.3

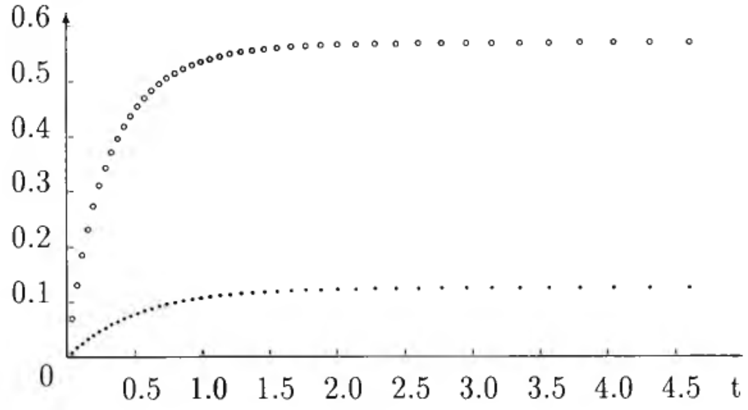


Figure 3.4: Singular values of the observability Grammian of Example 3.4

Example (3.4): Study the above system with a different output matrix $C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0.5 & 0 \end{bmatrix}$,

$$\underline{\dot{x}} = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \underline{x} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{u} \tag{3.51}$$

$$\underline{y} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0.5 & 0 \end{bmatrix} \underline{x} \tag{3.52}$$

The singular values of the observability Grammian are shown in Figure (3.4). Compared with Figure (3.3), the smallest singular value of the observability Grammian is smaller for the same given observation period, and therefore it is more difficult to observe the initial state.

3.2.4 Output tracking

In the definition (2.3) of output controllability, nothing has been said about the tracking of the a specified function $\underline{\mathrm{y}}(t)$. The definition is only concerned with getting to $\underline{y}(t_f)$. For $\underline{y}(t)$ to track an arbitrarily specified $\underline{\mathrm{y}}(t)$, the required $\underline{u}(t)$ might have to contain impulses. To synthesise a well-behaved function $\underline{u}(t)$, we shall restrict the class of $\underline{\mathrm{y}}(t)$ which we will have up to $\beta \leq n-1$ time derivatives. To have $\underline{y}(t) \equiv \underline{\mathrm{y}}(t)$ for all $t > t_0$, we must require all derivatives to be equal. So

$$\begin{aligned} \underline{y}(t) &= \underline{\mathrm{y}}(t), \\ \frac{d\underline{y}(t)}{dt} &= \frac{d\underline{\mathrm{y}}(t)}{dt}, \\ \frac{d^2\underline{y}(t)}{dt^2} &= \frac{d^2\underline{\mathrm{y}}(t)}{dt^2}, \\ &\vdots \\ \frac{d^\beta\underline{y}(t)}{dt^\beta} &= \frac{d^\beta\underline{\mathrm{y}}(t)}{dt^\beta}, \end{aligned} \tag{3.53}$$

where

$$\frac{d^i\underline{y}(t)}{dt^i} = CA^i\underline{x}(t) + CA^{i-1}B\underline{u} + CA^{i-2}B\frac{d\underline{u}}{dt} + \cdots + D\frac{d^i\underline{u}}{dt^i} \tag{3.54}$$

Equation (3.53) can be written in matrix form:

$$\mathcal{Y} \equiv \begin{bmatrix} \frac{d\underline{\mathrm{y}}(t)}{dt} \\ \frac{d^2\underline{\mathrm{y}}(t)}{dt^2} \\ \vdots \\ \frac{d^\beta\underline{\mathrm{y}}(t)}{dt^\beta} \end{bmatrix} - \begin{bmatrix} C \\ CA \\ \vdots \\ CA^\beta \end{bmatrix} \underline{x}(t) = \begin{bmatrix} D & 0 & 0 & \cdots \\ CB & D & 0 & \cdots \\ \vdots & & \ddots & \\ CA^{\beta-1}B & \cdots & CB & D \end{bmatrix} \begin{bmatrix} \underline{u} \\ \frac{d\underline{u}}{dt} \\ \vdots \\ \frac{d^{\beta-1}\underline{u}}{dt^{\beta-1}} \end{bmatrix} \tag{3.55}$$

where $\mathcal{Y}$ is taken as an arbitrary vector of dimension $(\beta-1)m$. From Cayley-Hamilton theorem that A^β is not independent of $A^{\beta-k}, k = 1, 2, ..., \beta - 1$, if $\beta \geq n$. Hence, we cannot track a function with more that $n-1$ independent derivatives. The solution for the vector of $\underline{u}$ derivatives is a linear algebra problem which leads to the following results:

Theorem 3.6 [Kai.,1] [Ske.,1] [Lue.,1] *The linear system (3.1) can track the vector function $\underline{\mathrm{y}}(t)$ up to its first $\beta \leq n-1$ derivatives if $\underline{\mathrm{y}}(t)$ is sufficiently smooth to* *have β derivatives and if the matrix*

$$\mathcal{U}_\beta \equiv \begin{bmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{\beta-1}B & CA^{\beta-2}B & CA^{\beta-3}B & \cdots & D \end{bmatrix} \tag{3.56}$$

has full row rank $(\beta - 1)m$. Conversely, if $\underline{y}(t)$ is a known output of the system (3.1), then an input $\underline{u}(t)$ which generates that output can be determined if $\mathcal{U}_\beta$ has rank $(\beta - 1)m$.

□

Note that the conditions of Theorem (3.6) can never be satisfied whenever $D = 0$. In this case an arbitrary $\underline{\mathbf{y}}(t)$ cannot be matched exactly, but if $\text{rank}(\mathcal{U}_\beta) = \beta m$, then all derivatives $\frac{d^i}{dt^i}\left\{\underline{\mathbf{y}}(t)\right\}, i = 1, 2, ..., \beta$, can be tracked (Note the omission of i=0). This means that "tracking" occurs with a constant offset $\underline{y}(t) - \underline{\mathbf{y}}(t) = \text{constant}$.

From equation (3.55) it is clear that to track an expected output $\underline{\mathbf{y}}(t)$, the input to the system is determined by the inverse of $\mathcal{U}_\beta$; so the singular values of this matrix defines the nature of the control and therefore can be used as the quantitative measure of the tracking.

3.2.5 Output disturbability

In real applications, a system is usually disturbed by some unavoidable sources of noise. In this case, it would be desirable to eliminate the response in the output to the undesirable system inputs. The undesirable disturbances are denoted in (3.1) as $\underline{\omega}(t)$, then the output disturbability is defined for output controllability as in Definition (2.3), except that the "input" $\underline{\omega}(t)$ is taken as the disturbances. The question posed here is whether the response to the disturbance can be totally eliminated.

Let $J = 0$ in (3.1). For $\underline{y}(t)$ to follow only its undisturbed response, it is required

$$\int_0^t Ce^{A(t-\sigma)}H\underline{\omega}(\sigma)d\sigma \equiv 0 \tag{3.57}$$

from $e^{A(t-\sigma)} = \sum_{i=0}^{n-1} A^i\alpha_i(t-\sigma)$, the above equation can be expanded to be

$$\int_0^t C\sum_{i=0}^{n-1} A^i\alpha_i(t-\sigma)H\underline{\omega}(\sigma)d\sigma \equiv 0 \tag{3.58}$$

or

$$\sum_{i=0}^{n-1} CA^iH \int_0^t \alpha_i(t-\sigma)\underline{\omega}(\sigma)d\sigma \equiv 0 \tag{3.59}$$

which leads to the following result.

Theorem 3.7 [Ske.,1] *Complete disturbance rejection is accomplished in system (3.1) with $D = 0$ for arbitrary disturbance $\underline{\omega}(t)$ if and only if the Markov parameters $M_i \equiv CA^iH$ are zero from $i = 0, ..., n-1$.*

□

However, it is almost impossible to reject disturbances totally. In order to measure how close the system is to total rejection of the disturbance, we need to know how small are the Markov parameters. Define a matrix

$$\mathcal{M} = \begin{bmatrix} CH & CAH & CA^2H & \cdots & CA^{n-1}H \end{bmatrix},$$

then the maximal singular value of the matrix can be used as a quantitative measure. The smaller the maximal singular value of the matrix, the better the quality of disturbance rejection of the system.

3.3 Characterisation of the relatively controllable initial state set

From the definition of the state controllability, if a system is controllable, then there always exists a control and a finite time, such that the final state can be reached. In this definition, no assumption is made about the nature of the control signal. For instance, the input signal might be required to contain impulses which is impossible in reality.

In real application, the energy of a control signal is always bounded. We define the relatively controllable initial state set for a bounded energy.

Definition 3.1 : *The* **relatively controllable initial state set** *is the set of initial states which can be steered by a control to the origin in a finite time T and the energy of the control signal is upper bounded by E.*

□

Assuming that a system described by state space model (A, B, C) as follows:

$$\begin{aligned} \dot{\underline{x}}(t) &= A\underline{x}(t) + B\underline{u}(t) \\ \underline{y}(t) &= C\underline{x}(t) \end{aligned} \tag{3.60}$$

For a given initial state $\underline{x}(t_0)$, the control to transfer the initial state to the origin is given in (3.32), which is

$$\underline{u}(t) = -\mathcal{F}^T(\sigma)\left[\ \int_{t_0}^{t_f} \mathcal{F}(\epsilon)\mathcal{F}^T(\epsilon)d\epsilon\ \right]^{-1} \underline{x}(t_0) \tag{3.61}$$

where $\mathcal{F}(\epsilon) \equiv \Phi(\epsilon, t_0)B(\epsilon)$ and the energy is given in (3.33), which is

$$E = \underline{x}^T(t_0)\left[\ \int_{t_0}^{t_f} \mathcal{F}(\epsilon)\mathcal{F}^T(\epsilon)d\epsilon\ \right]^{-1} \underline{x}(t_0) \tag{3.62}$$

Theorem 3.8 *The control given in (3.61) gives the minimum energy control compared with any other control signals which can bring the* $\underline{x}(t_0)$ *to the origin.*

□

Proof:

(a). Construction of the control: Given system (3.60), the initial state $\underline{x}(t_0)$ and the final state $\underline{x}(t_f) = 0$, the task is to find $u(\bullet)$ so that,

$$0 = \underline{x}(t_f) = \Phi(t_f, t_0)\underline{x}(t_0) + \int_{t_0}^{t_f} \Phi(t_f, \tau)B(\tau)\underline{u}(\tau)d\tau \tag{3.63}$$

or

$$-\Phi(t_f, t_0)\underline{x}(t_0) = \int_{t_0}^{t_f} \Phi(t_f, \tau)B(\tau)\underline{u}(\tau)d\tau \tag{3.64}$$

which can be rewritten as

$$-\underline{x}(t_0) = \int_{t_0}^{t_f} \mathcal{L}(\tau)\underline{u}(\tau)d\tau \tag{3.65}$$

where $\mathcal{L}(\tau) = \Phi(t_0, \tau)B(\tau)$.

The integration in (3.65) can be approximated by a finite sum where the infinite number of indeterminates $\underline{u}(\tau)$ are replaced by a finite number of indeterminates $\underline{u}(\tau_i), i = 0, 1, \ldots, N-1$ as

$$-\underline{x}(t_0) = \sum_{i=0}^{N-1} \mathcal{L}(\tau_i)\underline{u}(\tau_i)\triangle \tag{3.66}$$

where $\tau_i = t_0 + i\triangle; i = 0, 1, \ldots, N-1; \triangle = \frac{t_f - t_0}{N}; \mathcal{L}(\tau_i) = \Phi(t_0, \tau_i)B(\tau_i)$.

Let $\mathcal{U} = \left[\ \underline{u}(\tau_0) \ \ \ldots \ \ \underline{u}(\tau_{N-1})\ \right]^T, \mathcal{L} = \left[\ \mathcal{L}(\tau_0)\triangle \ \ \ldots \ \ \mathcal{L}(\tau_{N-1})\triangle\ \right]$, which is an $n \times N$ matrix.

Then (3.66) can be rewritten as

$$\mathcal{L}\mathcal{U} = -\underline{x}(t_0) \tag{3.67}$$

a set with more unknowns than equations.

Clearly, there can be a solution $\mathcal{U}$ if and only if $-\underline{x}(t_0) \in \mathcal{R}(\mathcal{L})$; i.e. $-\underline{x}(t_0)$ is a linear combination of the columns of $\mathcal{L}$. Since $\underline{x}(t_0)$ is an arbitrary n-vector, there can be a solution if and only if at least n columns of $\mathcal{L}$ are linearly independent. But the fact that row rank equals to the column rank means that there will be a solution if and only if

$$\text{the rows of } \mathcal{L}(\tau) \text{ are linearly independent;} \tag{3.68}$$

i.e. if and only if $\mathcal{L}$ has full rank. Clearly, condition (3.68) is plausible owing to the fact that (3.65) will have a solution if and only if the rows of $\mathcal{L}(\tau)$ are linearly independent.

Because there are more unknowns in the equations, there will be more than one solution of to the approximate equation (3.66) when (3.65) is met.

One particular solution is

$$\mathcal{U}^* = -\mathcal{L}^{-1}[\mathcal{L}\mathcal{L}^T]^{-1}\underline{x}(t_0) \tag{3.69}$$

$\mathcal{L}\mathcal{L}^T$ is invertible because the matrix $\mathcal{L}$ has full rank.

(b). Minimality of the control energy: [Kai.,1] In fact, the particular solution is a minimum energy solution among all the solutions which will bring the initial state to the final state. The minimality of the energy of the particular solution is given in the following.

The particular solution $\mathcal{U}^*$ to the equation (3.67) is not unique because we can clearly add to it any vector θ such that

$$\mathcal{C}\theta = 0 \tag{3.70}$$

There will always be such vectors because only n of the N columns of $\mathcal{C}$ are linearly independent. Therefore there are many solutions, but it turns out that the solution $\mathcal{U}^*$ has a minimum-length property

$$||\mathcal{U}^* + \theta|| \geq ||\mathcal{U}^*|| \tag{3.71}$$

for all θ satisfying (3.70). The proof follows by noting that

$$||\mathcal{U}^* + \theta||^2 \equiv (\mathcal{U}^* + \theta)^T(\mathcal{U}^* + \theta) = \mathcal{U}^{*T}\mathcal{U}^* + \theta^T\theta + 2\mathcal{U}^{*T}\theta \tag{3.72}$$

But by (3.69) and (3.70)

$$\mathcal{U}^{*T}\theta = \underline{x}^T(t_0)[\mathcal{L}\mathcal{L}^T]^{-1}\mathcal{L}\theta = 0 \tag{3.73}$$

so that, as claimed in (3.71),

$$||\mathcal{U}^* + \theta||^2 = ||\mathcal{U}^*||^2 + ||\theta||^2 \geq ||\mathcal{U}^*||^2 \tag{3.74}$$

□

First we study the state controllability Grammian of the system. It has been proved that if the system is controllable, then the controllability Grammian matrix of the system is always positive for a given time interval $t_f - t_0$.

Also from the definition of the controllability Grammian

$$\mathcal{G}_{sc} \equiv \int_{t_0}^{t_f} \Phi(\sigma, t_0)B(t)B^T(t)\Phi^T(\sigma, t_0)d\sigma \tag{3.75}$$

it is clear that the controllability Grammian of the system is also symmetric. So from theorem (3.2) we have

$$\mathcal{G}_{sc} = Q^T\Lambda Q \quad QQ^T = I \quad Q^{-1} = Q^T \tag{3.76}$$

where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n), \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$.

The inverse of the controllability Grammian is also a real symmetric and positive definite matrix, because

$$\begin{aligned}\mathcal{G}_{sc}^{-1} &= [Q^T\Lambda Q]^{-1} = Q^{-1}\Lambda^{-1}Q^{T^{-1}} = Q^T\Lambda^{-1}Q \\ &= Q^T \begin{bmatrix} \frac{1}{\lambda_1} & & & 0 \\ & \frac{1}{\lambda_2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\lambda_n} \end{bmatrix} Q \end{aligned} \tag{3.77}$$

The result concerning the set of relatively controllable initial states can be stated in the following:

Proposition 3.2 *The set of relatively controllable initial states which are possible to be transferred to zero with the given energy E in a finite time $t_f - t_0$ is determined in the Q coordinate by the condition:*

$$\frac{\epsilon_1^2}{\lambda_1} + \frac{\epsilon_2^2}{\lambda_2} + \cdots + \frac{\epsilon_n^2}{\lambda_n} \leq E \tag{3.78}$$

which is a hyper-elipsoid. The vector $[\epsilon_1, \epsilon_2, ..., \epsilon_n]^T$ is defined as $\underline{\epsilon} = Q^T\underline{x}$.

□

Remark 3.1: Because $|Q| = 1$, the set of initial conditions which are possible to be brought to zero in the original coordination is a rotated version of the hyper-elipsoid defined in (3.78).

□

Remark 3.2: Since $\mathcal{G}_{sc}$ is a function of $t_f - t_0$, the singular values of $\mathcal{G}_{sc}$ are also functions of $t_f - t_0$, so is the hyper-elipsiod.

□

From expression (3.78), it is clear that for a given amount of energy, the set of the initial states which can be brought to zero is defined by the eigenvalues of the controllability Grammian. Due to the symmetry and the positive-definiteness of the controllability Grammian, the eigenvalues are the square roots of the singular values.

Using linearity argument, it can be shown that from a given initial state and a given amount of energy, the accessible final states will also be a hyper-elipsoid set. The bigger the eigenvalues, the further final states can be reached.

Theorem 3.9 *For a given linear time invariant system with zero initial condition, when subject to bounded energy control, the maximal reachable states form a hyper-ellipsoid whose axes are defined by the singular values of the controllability Grammian and the energy bound.*

□

3.4 Input-output interaction indicators and simple control scheme selection

In multi-input multi-output systems, one output is affected by more than one input in general. In industrial control, simple SISO control schemes, in which one input is controlled only by one output, are favoured and therefore it is important to choose the input variable which is most closely related to the manipulated output variable. Here we discuss how to select an output from the output set to control an input from the input set with the assumption that the input and output sets have already been selected. The couplings are selected usually based on empirical interaction rules. The interaction among the various subsystems can be assessed by Relative Gain Array [Bri.,1], Relative Dynamic Array [Wit. & McA.,1] [Tun. & Edg.,1] [Gag. & Seb.,1] [Gro. Mor. & Hol.,1] [Man. Sav. & Ark.,1] and Block Relative Gain which will be reviewed first. However, these rules lack theoretical justifications. By employing quantitative output controllability we can establish an interaction measure between the inputs and the outputs on a solid basis. Having developed a new interaction measure, we finally put forward a novel criterion for input output coupling.

3.4.1 Relative Gain Array

Relative Gain Array [Bri.,1] is a steady-state interaction measure among the inputs and the outputs. For a square system with n inputs and n outputs, the Relative Gain Array is defined as a matrix

$$G_R = [\lambda_{ij}] \tag{3.79}$$

where $\lambda_{ij} = \frac{(\partial y_i/\partial u_j)_{u_k=0,k\neq j}}{(\partial y_i/\partial u_j)_{y_l=0,l\neq i}}$. $(\partial y_i/\partial u_j)_{u_k=0,k\neq j}$ is the steady-state gain between u_j and y_i when no control is applied to the system. $(\partial y_i/\partial u_j)_{y_l=0,l\neq i}$ is the steady-state gain between u_j and y_i when feedback control involving all other inputs $u_k, k = 1, ..., n, k \neq j$ and all other outputs $y_l, l = 1, ..., n, l \neq i$ is applied to the system such that at the steady state all $y_l, l = 1, ..., n, l \neq i$ are held at their nominal values. The interaction measure is only applicable when the controller contains integral action in each of the control loop.

The rows and columns of G_R in equation (3.79) satisfy

$$\sum_{i=1}^{n} \lambda_{ij} = \sum_{j=1}^{n} \lambda_{ij} = 1 \tag{3.80}$$

and the value λ_{ij} is an indicator of interaction between i-th input and j-th output. The bigger the value, the stronger the interaction. Relative Gain Array can also be used as an indicator for selecting the input-output pairing.

3.4.2 Relative dynamic array

Relative Dynamic Array [Gag. & Seb.,1] is an interaction measure based on the open-loop responses of the system. The Relative Dynamic Array is defined as

$$G_D = [\lambda_{ij}], \quad \lambda_{ij}(\theta) = \phi_{ij}(\theta)\tilde{\phi_{ji}}(\theta) \tag{3.81}$$

where $\phi_{ij}(\theta)$ is defined as the integral of the open-loop step response, $y_i(t)$, to a unit step change in u_j at time, $t = 0$,

$$\phi_{ij}(\theta) = \int_0^\theta y_i(t)dt \quad (i, j = 1, 2, ..., n) \tag{3.82}$$

and in (3.81) $\tilde{\Phi}(\theta) = [\tilde{\phi_{ji}}(\theta)] = [\Phi^T(\theta)]^{-1}$. The time period, θ, over which the integration is carried out is arbitrary depending on the signal frequencies under study; it was suggested θ be specified as 20% to 100% of the dominant time constant of the process. The elements in a row or a column satisfy

$$\sum_{i=1}^{n} \lambda_{ij} = \sum_{j=1}^{n} \lambda_{ij} = 1 \tag{3.83}$$

The values of the elements of the Relative Dynamic Array are indicators of dynamic interaction between the inputs and the outputs. If $\theta \to \infty$, then the Relative Dynamic Array approached Relative Gain Array asymptotically.

3.4.3 Block relative gain array [Man. Sav. & Ark.,1]

Block Relative Gain is a measure of interaction between subsystems. When each of the sub-systems contains only one input and one output, then it degenerates to Relative Gain Array. Block Relative Gain is defined for an $n \times n$ square system as follows. Assuming that the system consists of two square subsystems, one has a dimension $m \times m$ and the other is $(n - m) \times (n - m)$

$$\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix} = G \begin{bmatrix} \underline{u}_1 \\ \underline{u}_2 \end{bmatrix}, G = \begin{bmatrix} \overbrace{G_{11}}^{m} & \overbrace{G_{12}}^{n-m} \\ G_{21} & G_{22} \end{bmatrix} \begin{matrix} \}m \\ \}n-m \end{matrix} \tag{3.84}$$

the plant is to be controlled by a decentralised control structure in which the first m outputs, $\underline{y}_1$, are interconnected with the first m inputs, $\underline{u}_1$, and the last $n - m$ outputs, $\underline{y}_2$, to the last $n - m$ inputs, $\underline{u}_2$. The feedback configuration is shown in Figure (3.5). where the feedback F and the gain K have the form

$$K = \begin{bmatrix} \overbrace{K_1}^{m} & \overbrace{0}^{n-m} \\ 0 & K_2 \end{bmatrix} \begin{matrix} \}m \\ \}n-m \end{matrix} \qquad F = \begin{bmatrix} \overbrace{F_1}^{m} & \overbrace{0}^{n-m} \\ 0 & F_2 \end{bmatrix} \begin{matrix} \}m \\ \}n-m \end{matrix} \tag{3.85}$$

then

$$\left.\frac{\partial \underline{y}_1}{\partial \underline{u}_1}\right|_{\substack{\underline{y}_2 = 0 \\ F = 0}} = G_{11} \tag{3.86}$$

$$\left.\frac{\partial \underline{y}_1}{\partial \underline{u}_1}\right|_{\substack{\underline{y}_2 = 0 \\ F_1 = 0 \\ F_2 = I}} = ([G^{-1}]_{11})^{-1} = G_{11} - G_{12}G_{22}^{-1}G_{21} \tag{3.87}$$

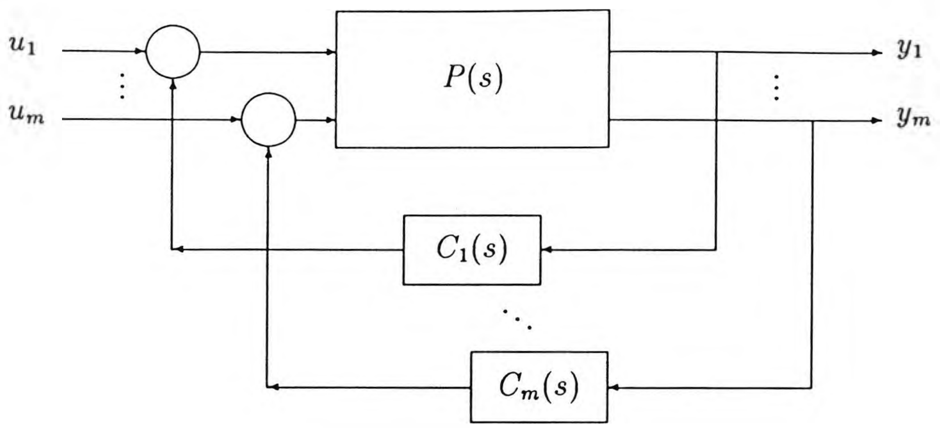


Figure 3.5: Simple Control Configuration

where $[G^{-1}]_{11}$ is the first $m \times m$ block of G^{-1}. In (3.87), G_{11} denotes the block gain between $\underline{y}_1$ and $\underline{u}_1$ when all the loops are open. Similarly in (3.87), $([G^{-1}]_{11})^{-1}$ denotes the block gain between $\underline{y}_1$ and $\underline{u}_1$ when the first m loops are open and the last $n-m$ loops are closed and under perfect control, i.e. $\underline{y}_2 = 0$. Then the left and right Block Relative Gain can be defined as

$$\mathrm{BRG}_l = \left[\frac{\partial \underline{y}_1}{\partial \underline{u}_1} \bigg|_{\substack{\underline{y}_2 = 0 \\ F = 0}} \right] \bullet \left[\frac{\partial \underline{y}_1}{\partial \underline{u}_1} \bigg|_{\substack{\underline{y}_2 = 0 \\ F_1 = 0 \\ F_2 = I}} \right]^{-1} = G_{11}[G^{-1}]_{11} \quad (3.88)$$

$$\mathrm{BRG}_r = \left[\frac{\partial \underline{y}_1}{\partial \underline{u}_1} \bigg|_{\substack{\underline{y}_2 = 0 \\ F_1 = 0 \\ F_2 = I}} \right]^{-1} \bullet \left[\frac{\partial \underline{y}_1}{\partial \underline{u}_1} \bigg|_{\substack{\underline{y}_2 = 0 \\ F = 0}} \right] = [G^{-1}]_{11}G_{11} \quad (3.89)$$

The left (or right) Block Relative Gain can be served as an indicator for the interaction between the two subsystems. When BRG_l is close to identity, the interaction is weak and otherwise, the interaction is strong. So the both RGA and BRG can be used as an indicators for selecting the input-output pairing.

3.4.4 Energy criterion for input-output interaction

The above methods suffer from the drawback that the system must be square. Further, in each of the control loop, there must be an integral unit. However, in real application, systems need not to be square and not all the control loops have to have an integral unit. In order to overcome these drawbacks, we can develop the quantitative output controllability into an interaction indicator between the inputs and the outputs. As having been pointed earlier the singular values of the output controllability are related to the energy needed to transfer the output from an initial zero state to a particular final state. If an input is not connected to an output, then the desired state of that output cannot be achieved by exercising control from that

	u_{11}	u_{12}	$\cdots$	u_{1p_1}	u_{21}	$\cdots$	u_{2p_2}	$\cdots$	u_{s1}	$\cdots$	u_{sp_s}
y_{11}	0.01			0.54			0.2				
y_{12}		0.31							0.33		0.02
$\vdots$											
y_{1m_1}		0.02			0.6		0.07				0.34
y_{21}	0.12	0.22									0.89
$\vdots$											
y_{sm_s}				0.35			0.71				0.6

Table 3.1: Input-output interaction table.

particular input. If the input is weakly connected to an output then the energy needed will be large. Further, the energy needed to achieve a certain transformation of the output is proportional to the inverse of the singular values, therefore, the greater the values of the singular values, the stronger the interaction.

In the following we propose the procedure for fine control scheme selection.

Step 1. Calculate the output controllability from any of the input $u_i, u_i \in \mathcal{S}_j$ in the input set which is associated with the output j.

Step 2. Select the input u_i^T which gives the maximal value for the indicator.

□

Continue this procedure for every output in the output set, then finally one input for each of the output can be found and the controller design can be carried out based on the selected inputs and the outputs. The control input for the output will be chosen in such a way that the interaction indicator based on the output controllability Grammian gives the largest among all the input sets. Clearly, the system need not to be square. Table 3.1 shows an example of the values of the interaction indicators for a fixed time T.

Remark 3.3: The selection procedure works on systems which are both square and non-square.

□

Remark 3.4: The selection is basically an open loop criterion. The effect of the feedback loop will be difficult to account for, because the procedure for calculating the singular values of the output controllability Grammian needs the exact values of the parameters of the feedback loop.

□

Remark 3.5: The algorithm is based on the computation of e^{At} of the system and works well for small dimension systems. When the dimension of the system is large, some other fast algorithm for calculating the output controllability Grammian should be investigated and implemented.

□

Remark 3.6: The selection procedure can be applied to the remaining subsystem when some of the inputs and outputs have already been fixed for particular control purposes.

□

3.4.5 Example

In the following we study a 2×2 system which is described both in the state space and transfer model as:

$$G(s) = \begin{bmatrix} \frac{1}{s+1} & \frac{1}{(s+1)(s+2)} \\ \frac{1}{(s+1)(s+3)} & \frac{1}{s+3} \end{bmatrix} \tag{3.90}$$

and the corresponding irreducible state-space realisation of the system is:

$$A = \begin{bmatrix} -1 & & & \\ & -1 & & \\ & & -2 & \\ & & & -3 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & -1 \\ -\frac{1}{2} & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 0 & 1 \end{bmatrix}.$$

The RGA of the system can be calculated as

$$G_R = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \tag{3.91}$$

So by using the selection criteria, 1-1 and 2-2 pairing is favoured.

Now, if we look at the quantitative output controllability indicators which are displayed in Fig. (3.6). For t=4 sec., the singular values of the output controllability Grammian from inputs to the outputs are as shown in Table 3.

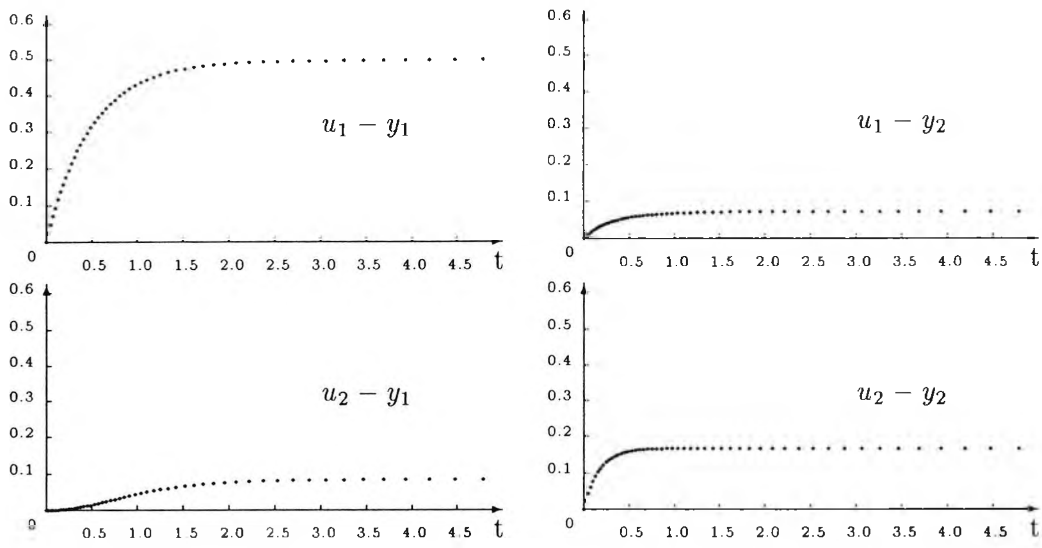


Figure 3.6: Singular values of the output controllability Grammian from u_i to y_j

Table 3. Singular values of the output controllability Grammian from u_i to y_j (i,j=1,2) (t=4sec)

	y_1	y_2
u_1	0.500	0.042
u_2	0.084	0.164

So the pairing $1-1$ and $2-2$ is also favoured.

3.5 Measures for eigenvalue mobility

Since the set of system eigenvalues is a very important indicator for various system properties, the reallocation of the eigenvalues via state feedback as well as output feedback has important implications in improving system stability, dynamic responses, robustness, etc. It has been proved that the set of closed-loop eigenvalues can be arbitrarily assigned by state feedback if and only if the system under consideration is controllable [Won.,2]. However, controllability of a system does not give any information on how far can the eigenvalues be moved when the system is subject to constrained feedback. Developing measures for near uncontrollability of a controllable system is important in answering questions such as how far can the states of a system be reached when subject to bounded energy controls, which has been studied in Section 3.3, or how far can the closed-loop eigenvalues be shifted when the feedback gains are constrained. The mobility of the closed-loop eigenvalues was found to be related to the distance of the controllable system from the set of uncontrollable ones [Eis.,1] [Bol. & Lu,1] as well as to the angle between the input direction and the set of eigenvectors [Tar.,1].

3.5.1 Distance between a controllable and the nearest uncontrollable system

Basically all systems are generically controllable, or put it in another way, the set of controllable systems is open and dense in the finite dimensional parameter space. In the neighbourhood of a controllable (A, B), there may not exist any uncontrollable pair. Therefore defining the distance between a controllable pair and a nearest uncontrollable pair makes sense [Bol. & Lu,1].

Definition 3.2 *For a given pair (A, B) define the distance between (A, B) and a nearest uncontrollable pair by*

$$\begin{aligned} \mu(A, B) &= \min_{\Delta A, \Delta B} \|\Delta A, \Delta B\|_2 \\ \textit{subject to} & \quad (A + \Delta A, B + \Delta B) \textit{ being uncontrollable} \end{aligned}$$

where $\Delta A \in C^{n \times n}, \Delta B \in C^{n \times l}$. When the perturbations are real, the distance is defined as

$$\begin{aligned} \mu_r(A, B) &= \min_{\Delta A, \Delta B \text{ real}} \|\Delta A, \Delta B\|_2 \\ \textit{subject to} & \quad (A + \Delta A, B + \Delta B) \textit{ being uncontrollable} \end{aligned}$$

where $\Delta A \in R^{n \times n}, \Delta B \in R^{n \times l}$.

□

By using singular value decomposition, the distances are equivalent to the minimum of the smallest singular value $[sI - A, B]$ for all the complex or real values of s, i.e.

$$\begin{aligned} \mu(A, B) &= \min_{s \in C} \sigma_n(sI - A, B) \\ \mu_r(A, B) &= \min_{s \in R} \sigma_n(sI - A, B) \end{aligned}$$

where $\sigma_n(sI - A)$ is the smallest singular value of $[sI - A, B]$.

Computing $\mu_r(A, B)$ numerically can be an involved process. However, an upper bound on $\mu_r(A, B)$ can be obtained based on the singular values of the controllability matrix $Q_c = \left[B, \quad BA, \quad \cdots, \quad BA^{n-1} \right]$.

Applying singular value decomposition on Q_c, we have

$$Q_c = Y \left[\sum \quad 0 \right] U^* \tag{3.92}$$

where $\sum =$ diag $\{\sigma_1, \cdots, \sigma_n\}, Y \in R^{nxn}$ and $U^* \in R^{nl \times nl}$ are orthogonal matrices. Then an estimate for the distance is given as [Bol. & Lu,1]

$$\mu(A, B) \leq \mu_r(A, B) \leq \left(1 + \frac{\|A_F\|_2}{\sigma_{n-1}}\right) \sigma_n \tag{3.93}$$

where A_F is the companion form of A.

Based on the distance $\mu(A, B)$, the mobility of eigenvalues when subject to state feedback control can be stated as:

Theorem 3.10 [Bol. & Lu,1] *Assume that (A, B) is a controllable pair and λ_n is a simple eigenvalue of A. Then for any sufficiently small $\gamma > 0$, there exists a feedback matrix K with norm bounded by γ such that all the closed-loop eigenvalues of the system differ from λ_n by at least $\gamma\mu(A, B)$.*

□

Remark 3.7: The result presented above is only applicable to the case when the norm bound γ on the gain is sufficiently small and matrix A is simple.

□

3.5.2 A direct measure of eigenvalue mobility

A direct measure of eigenvalue mobility, when subject to output feedback can be developed [Tar.,1].

Consider a linear multivariable system

$$\begin{cases} \dot{\underline{x}}(t) = A\underline{x}(t) + B\underline{u}(t) \\ \underline{y}(t) = C\underline{x}(t) \end{cases} \tag{3.94}$$

with A having distinct eigenvalues $\lambda_i, i = 1, 2, ..., n$. An output feedback law is applied as:

$$\underline{u} = -\delta K \underline{y} \tag{3.95}$$

where δK is the incremental $m \times l$ feedback matrix. The set of eigenvalues of the closed-loop system is denoted as $\hat{\lambda}_i, \hat{\lambda}_i = \lambda_i + \delta\lambda_i$. Then we have the following definition of eigenvalue mobility.

Definition 3.3 [Tar.,1] *The mobility of the eigenvalues λ_i of the system (A, B, C) under output feedback is defined to be*

$$\mu_i = \|\frac{\delta\lambda_i}{\delta K}\|_2 \tag{3.96}$$

where $\frac{\delta\lambda_i}{\delta K} = \begin{bmatrix} \frac{\delta\lambda_i}{\delta K_{11}} & \cdots & \frac{\delta\lambda_i}{\delta K_{1l}} \\ \vdots & & \vdots \\ \frac{\delta\lambda_i}{\delta K_{m1}} & \cdots & \frac{\delta\lambda_i}{\delta K_{ml}} \end{bmatrix}$

□

Let $\underline{e}_i$ and $\underline{f}_i$ denote the normalised right and left eigenvectors of A, then it has been shown that the mobility of λ_i may be expressed by:

$$\mu_i = \|C\underline{e}_i\underline{f}_i^T B\|_2 \tag{3.97}$$

Remark 3.8: It is clear from the above result that if an eigenvalue of a system is uncontrollable, then $\underline{f}_i B = 0$ and therefore, the eigenvalue mobility is zero.

□

The eigenvalue mobility can further be related to the distance between the eigenvalue and the zeros of the system. In fact, let $G(s)$ be the transfer function. Then it can be expressed as:

$$G(s) = C(sI - A)^{-1}B = \sum_{i=1}^{n} \frac{R_i}{(s - \lambda_i)} \tag{3.98}$$

where R_i is the residue matrix

$$R_i = \left[(s - \lambda_i)C(sI - A)^{-1}B\right]_{s=\lambda_i} \tag{3.99}$$

Set $h_i(s) = \Pi_{l=1,l\neq i}^{n}(s - \lambda_l)$ and $W(s) = C$ adj $(sI - A)B$, then

$$R_i = \frac{W(\lambda_i)}{h_i(\lambda_i)} \tag{3.100}$$

and

$$\mu_i = \frac{1}{\alpha_i}\|R_i\|_2 = \frac{\|W(\lambda_i)\|_2}{\alpha_i|h_i(\lambda_i)|} \tag{3.101}$$

where α_i is a positive constant relating the eigenvectors to the normalised eigenvectors. Take an element $w_{ij}(s), i = 1, ..., m, j = 1, ..., l$ from $W(s)$ and write it in the following form

$$w_{ij}(s) = a_{ij}(s - z_{ij1})(s - z_{ij2}) \cdots (s - z_{ijp}) \tag{3.102}$$

where a_{ij} is a constant and z_{ijk} are the roots of $w_{ij}(s) = 0$. Then the eigenvalue mobility is related to the distance between the eigenvalue and the zeros of the system in the following way:

Theorem 3.11 [Tar.,1] *The eigenvalue mobility of μ_i satisfies*

$$\mu_i = \frac{1}{\alpha_i |h_i(\lambda_i)|} \left[\sum_{i=1}^{m} \sum_{j=1}^{l} a_{ij}^2 (\lambda_i - z_{ij1})^2 \cdots (\lambda_i - z_{ijp})^2 \right]^{1/2} \tag{3.103}$$

□

A eigenvalue mobility measure for systems with repeated eigenvalues has also been developed, though it is much more complicated [Tar.,1].

Remark 3.9: If λ_i is a zero of $W(s)$, then the mobility of this eigenvalue is zero. This complies with the existing theory because in this case there is a pole-zero cancellation and therefore the system is uncontrollable.

□

3.6 Summary

In this chapter, further quantitative measures have been developed for output controllability, state controllability and observability. These quantitative measures are based on the output controllability Grammian, state controllability Grammian and observability Grammian. By using the quantitative state controllability, the set of initial states which can be brought to the origin with bounded energy control signals has been parameterized.

The quantitative output controllability then is further developed to be an interaction measure between the inputs and the outputs. Other interaction measures based on the transfer functions of systems have been reviewed. A scheme for the selection of control structure has been proposed and the input-output interaction indicator, obtained earlier, is used as an criterion for the selection of simple control structures.

Though the problem of closed-loop pole assignment of controllable systems via state feedback has been solved, some related issues, such as relative pole mobility with respect to the controller gain, or how far can one move the poles when the controller gain is bounded, are still open. Some measures for relative pole mobility have recently been developed [Bol. & Lu,1] [Tar.,1], but the problem of pole mobility of a system when subject to bounded state feedback has yet to be addressed. This will be the main concern of the next three chapters.

Chapter 4

ROOT LOCATION OF POLYNOMIALS — BACKGROUND RESULTS

4.1 Introduction

As having been reviewed in Chapter 2, the poles of a system are important stability indicators. When the parameters of a system are subject to uncertainties, system stability might be at stake. Stability robustness of systems under parameter uncertainties has been extensively investigated [Lun.,1] [Doy. & Ste.,1] [Kai.,1] [Chen,1] [Ack.,2] etc. Here we are interested in the effect of controllability or near uncontrollability on the pole mobility under state feedback and what can be achieved under bounded feedback such as stabilisability. The effect of state feedback on the mobility of the closed-loop poles is manifested through the coefficients of the closed-loop characteristic polynomials [Kar. & Shan,2]. Thus, this chapter provides some results on the root distribution of polynomials in relation to their coefficients. In Section 4.2, important classical results concerning the root distributions of polynomials are presented [Mar.,1] [Xu,1]. Results on the root distribution with respect to the coefficient variations of polynomials are reviewed in Section 4.3 [Ost.,1] [Gan.,1] [Arg.,1] [Arg.,2].

The relation between the coefficients of the characteristic polynomial and the physical parameters is usually complicated. But if we assume that the variations of the physical parameters are confined within upper and lower bounds, so will be the coefficients of the characteristic polynomial. Polynomials whose coefficients are within interval regions will be termed as interval polynomials. The root distributions of interval polynomials have been studied by Kharitonov [Kha.,1] and it has been proved that in the complex coefficient case, the whole set of interval polynomials will have all the roots in the left half of the complex plane if and only if eight specially chosen polynomials all have their roots in the left half of the complex plane. The number of the special polynomials reduces to four for polynomials with real coefficients. Since the introduction of this seminal result, many applications and extensions based on this have been developed. The robust control theory especially has taken a new impetus by the study of properties of interval polynomials [Barm.,1] [Bia. & Gar.,1] [Bie. Hwa. & Bha.,1] [Hol. Looz. & Bar.,1] [Cha. & Bha.,1]. In Section 4.4, some of the Kharitonov type of results are reviewed.

In connection to stability robustness, problems such as how far a stable system is away from instability, are very important. Having found the distance of a stable system from instability, maximal allowable variations in the parameter space can be studied. Results obtained by D. Hinrichsen and A. J. Pritchard [Hin. & Pri.,1] [Hin. & Pri.,2] are reviewed in Section 4.5. A dual problem concerning the minimum distance of an unstable polynomial from stable polynomials has been put forward and methods for solving the problem have also been suggested.

4.2 Survey of previous classical results on polynomials

The relationship between the root distribution and the coefficients of the polynomials has been studied for a long time. Some results are summarised below from Marden [Mar.,1].

Theorem 4.1 *All the roots of $f(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n \in P[s]$ lie in the circle $|s| \leq r$, where r is the positive root of the equation:*

$$|\alpha_1|s^{n-1} + \ldots + |\alpha_{n-1}|s + |\alpha_n| - s^n = 0 \tag{4.1}$$

□

The result given in [Mori,1] is basically the same.

Theorem 4.2 *All the roots of $f(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n \in P[s]$, lie in the circle*

$$|r| \leq 1 + \max |\alpha_i|, \tag{4.2}$$

where $i = 1, 2, ..., n$.

□

Theorem 4.3 *The root r_1 of largest modulus of $f(s) = s^n + \alpha_1 s^{n-1} + \ldots + \alpha_n \in P[s]$, satisfies the inequality*

$$(2^{1/n} - 1)r \leq |r_1| \leq r \tag{4.3}$$

where r is the positive root of the equation (4.1).

□

Theorem 4.4 *For any p and q such that $p > 1, q > 1$,*

$$\tfrac{1}{p} + \tfrac{1}{q} = 1 \tag{4.4}$$

the polynomial $f(s) = \alpha_0 s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n$, has all its roots in the circle

$$|r| < \left[1 + \textstyle\sum_{i=0}^{n-1} |\alpha_i|^p\right]^{q/p} < (1 + n^{q/p} M^q)^{1/q} \tag{4.5}$$

where $M = \max |\alpha_i/\alpha_0|, i = 1, 2, ..., n$.

□

Corollary 4.1 *From the above result for $p = q = 2$ we have the following*

$$|s| < \{1 + \sum_{j=1}^{n} |\alpha_j|^2\}^{\frac{1}{2}} < \sqrt{1 + nM^2} \tag{4.6}$$

□

Some further useful bounds are given in Marden [Mar.,1] and they are summarised below:

Corollary 4.2 *Let $f(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n \in P[s]$. Discs which contain all roots of $f(s)$ are defined below by their radii:*

$$|s| \leq \left[1 + |\alpha_n|^2 + |\alpha_{n-1} - \alpha_n|^2 + \ldots + |\alpha_1 - \alpha_2|^2 + |1 - \alpha_1|^2\right]^{1/2} \tag{4.7}$$

$$|s| \leq \sum_{j=1}^{n} |\alpha_i|^{1/i} = (|\alpha_1| + |\alpha_2|^{1/2} + |\alpha_3|^{1/3} + \ldots + |\alpha_n|^{1/n}) \tag{4.8}$$

$$|s| \leq \gamma = \max(|\alpha_n/\alpha_{n-1}|, 2|\alpha_{n-1}/\alpha_{n-2}|, \ldots, 2|\alpha_1|) \tag{4.9}$$

□

For special subfamilies polynomials the following results are also stated from Marden [Mar.,1].

Corollary 4.3 *If $f(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n \in P[s]$ and $\alpha_n \geq \alpha_{n-1} \geq \ldots \geq \alpha_2 \geq \alpha_1 \geq 1$, then $f(s)$ has no root in the disc $|s| < 1$.*

□

Corollary 4.4 *If $f(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n \in P[s], \alpha_i > 0$, then all its roots lie in the circle $|s| \leq \rho$, where*

$$\rho = \max\{\alpha_n/\alpha_{n-1}, \alpha_{n-1}/\alpha_{n-2}, \ldots, \alpha_2/\alpha_1, \alpha_1\} \tag{4.10}$$

□

Corollary 4.5 *If $f(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_p s^{n-p} + \alpha_{p+1} s^{n-p-1} + \ldots + \alpha_{n-1} s + \alpha_n \in P[s]$ and the following condition holds true:*

$$|\alpha_p| > 1 + \sum_{i=1, i\neq p}^{n} |\alpha_i| \tag{4.11}$$

then $f(s)$ has exactly p roots in the unit circle.

□

Corollary 4.6 *If $f(s) = s^n + \alpha_1 s^{n-1} + \ldots + \alpha_{n-1} + \alpha_n \in P[s]$. All roots of $f(s)$ lie in the circle $|s| \leq \max(L, L^{1/(n+1)})$ where L is the length of the polygonal line joining in succession the points $0, \alpha_n, \alpha_{n-1}, \ldots, \alpha_2, \alpha_1, 1$.*

□

Note that the length of the polygonal line joining succession the points $0, \alpha_n, \alpha_{n-1}, \ldots, \alpha_2, \alpha_1, 1$ is

$$L = |\alpha_n| + |\alpha_{n-1} - \alpha_n| + \ldots + |\alpha_1 - \alpha_2| + |1 - \alpha_1| \tag{4.12}$$

and thus the bound in Corollary (4.6) is expressed in terms of the successive differences of the coefficients.

Some improved results are given below.

Theorem 4.5 [Xu,1] *For the polynomials* $f(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n \in P[s]$, *if there exists a constant* r *and a positive integer* m *such that*

$$\sum_{i=m}^{mn} \left| \sum_{j_{m-1}=m-1}^{i-1} \sum_{j_{m-2}=m-2}^{j_{m-1}-1} \cdots \sum_{j_1=1}^{j_2-1} \prod_{k=1}^{m} \alpha_{j_k - j_{k-1}} \right| r^{-i} < 1 \tag{4.13}$$

where $j_0 = 0, j_m = i$ *and* $a_i = 0$ *if* $i \notin N \equiv \{0, 1, ..., n\}$, *then the roots* r_k $(k \in N)$ *of the polynomial lie within a circle of radius* r *centered on the origin of the complex plane, i.e.,*

$$|r_k| < r, \quad i \in N. \tag{4.14}$$

If

$$\sum_{i=0}^{m(n-1)} \left| \sum_{j_{m-1}=0}^{i} \sum_{j_{m-2}=0}^{j_m-1} \cdots \sum_{j_1}^{j_2} \prod_{k=1}^{m} \alpha_{j_k - j_{k-1}} \right| r^{mn-i} < |\alpha_n|^m \tag{4.15}$$

where $j_0 = 0, j_m = i, \alpha_i = 0$ *as* $i \geq n$ *on the left side of the inequality (4.15), then we have*

$$|r_k| > r, \quad i \in N. \tag{4.16}$$

□

Theorem 4.6 [Xu,1] *If there are positive integers* p, m *and a real number* s $(|s| < r)$ *such that*

$$\sum_{i=m}^{m(n+p)} \left| \sum_{j_{m-1}=m-1}^{i-1} \sum_{j_{m-2}=m-2}^{j_{m-1}-1} \cdots \sum_{j_1=1}^{j_2-1} \prod_{k=1}^{m} (\alpha_{j_k - j_{k-1}} - s\alpha_{j_k - j_{k-1} - p}) \right| r^{-i} < 1 \tag{4.17}$$

where $j_0 = 0, j_m = i$ *and* $\alpha_i = 0$ *as* $i \notin N$, *then the roots of the polynomial satisfy*

$$|r_k| < r, \quad i \in N. \tag{4.18}$$

□

4.3 Perturbation results

When the coefficients of a polynomial are subject to uncertainties, so are its roots. In the literature, perturbation results are abundant. The results can mainly be classified into two categories. Firstly, the root distribution of the perturbed polynomials with respect to the roots of the nominal polynomials, and secondly, the roots of the perturbed polynomials with respect to a region which, for instance, can be a half plane or a circular region of particular interest. In the stability analysis of control systems, the second class of results is more important.

4.3.1 Relation between the roots of the nominal and perturbed polynomials

Consider two polynomials

$$f(s) = s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n \tag{4.19}$$

$$g(s) = s^n + b_1 s^{n-1} + \cdots + b_{n-1}s + b_n \tag{4.20}$$

Let the n roots of $f(s)$ be $x_1, ..., x_n$, those of $g(s), y_1, ..., y_n$. The estimates for the differences between x_i and y_i in terms of the expressions $|b_i - a_i|$ is given below.

Set

$$\gamma = \max_i(1, |x_i|, |y_i|), \qquad \Gamma = \max_{i>0}(|a_i|^{1/i}, |b_i|^{1/i}), \tag{4.21}$$

and it is well known that $\gamma \leq 2\Gamma$.

Introduce the expression

$$\epsilon = \sqrt[n]{\sum_{i=1}^{n} |b_i - a_i|\gamma^{n-i}} \tag{4.22}$$

Then we have the following results.

Theorem 4.7 [Ost.,1] *In the ϵ neighbourhood of any root of $f(s)$, x_0, there is always one root of $g(s)$, y_0.*

□

Theorem 4.8 [Ost.,1] *Let the n roots of $f(s)$ be $x_1, ..., x_n$ and those of $g(s)$ be $y_1, ..., y_n$. Then the roots of $f(s), g(s)$ can be ordered in such a way that we have*

$$|x_i - y_i| < 2n\epsilon \quad (i = 1, 2, ..., n). \tag{4.23}$$

□

The value ϵ which depends on the roots of $g(s)$ can be estimated in the following way. Set

$$\delta = \max_i |b_i - a_i|, \tag{4.24}$$

which is a known number. Then $\epsilon^n \leq \delta \sum_{i=1}^{n-1} (2\Gamma)^i$.

On the other hand, we have for any $\mu \geq 0$ the relation

$$\sum_{i=0}^{n-1} \mu^i \leq \max(1, \mu^n) \min\left(n, \frac{1}{|1-\mu|}\right) \tag{4.25}$$

So we obtain

$$\epsilon \leq \delta^{1/n} \max(1, 2\Gamma)^n \sqrt[n]{\min\left(n, \frac{1}{|1+2\Gamma|}\right)}. \qquad (4.26)$$

Another estimate is obtained by assuming

$$|b_i - a_i| \leq \sigma \Gamma^i, \quad (i = 1, 2, ..., n); \qquad (4.27)$$

then

$$\epsilon^n \leq \sigma \sum_{i=1}^{n} \Gamma^n (2\Gamma)^{n-i} = \sigma \Gamma^n (1 + 2 + \cdots + 2^{n-1}) \qquad (4.28)$$

$$\epsilon \leq 2\Gamma \sigma^{1/n}. \qquad (4.29)$$

A later result without involving the roots of the perturbed polynomial is presented below.

Theorem 4.9 [Ost.,1] *Let* $\Gamma_1 = \max\{1, |a_i|, |b_i|\}$, $d = \sum_{i=1}^{n} |a_i - b_i|$, *then the roots of the perturbed polynomial are related to the original polynomial as*

$$|y_i - x_i| \leq (n+2)\Gamma_1 d^{1/n}. \qquad (4.30)$$

□

Remark 4.1: The results presented here are useful later on in studying the closed-loop pole mobility and system stabilisability when a system is subject to bounded norm feedback, where the perturbation expresses the feedback term.

□

The absolute deviation of the roots of the perturbed polynomial from the nominal ones may not be the main interest. Instead, very often it is desirable to know the sensitivity of the roots when subject to perturbations. The relative sensitivity of the roots with respect to the perturbations is given as

Theorem 4.10 [Ost.,1] *Consider two polynomials*

$$f(s) = s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n \qquad (4.31)$$

$$g(s) = s^n + b_1 s^{n-1} + \cdots + b_{n-1}s + b_n \qquad (4.32)$$

and assume that $a_n \neq 0$ *i.e. that* $f(s)$ *has* n *finite* x_i $(i = 1, 2, ..., n)$ *and* $x_i \neq 0$. *Assume further that for a certain positive* τ *with*

$$4n\tau^{1/n} \leq 1 \qquad (4.33)$$

we have

$$|b_i - a_i| \leq \tau |a_i|, \quad (i = 0, 1, ..., n) \tag{4.34}$$

Then the n roots $y_1, ..., y_n$ of $g(s)$ can be ordered in such a way that

$$\left| \frac{y_i}{x_i} - 1 \right| < 8n\tau^{1/n}, \quad (i = 1, ..., n) \tag{4.35}$$

□

4.3.2 Perturbation with respect to the left half of the complex plane

The eigenvalues of a system are the indicators of the system stability. If all the eigenvalues of the system have negative real parts, the system will be stable. Calculating the eigenvalues of the system is equivalent to finding all the roots of the characteristic polynomial. However, without actually computing the roots of a polynomial, it can be decided indirectly whether or not all the roots of a polynomial are in the left half of the complex plane by using Routh-Hurwitz Theorem or Hermite-Biehler Theorem. If a polynomial has all its roots in the left half of the complex plane, the polynomial is also referred to as *Hurwitz*.

Theorem 4.11 (Routh-Hurwitz Theorem [Gan.,1]) *All the roots of a real polynomial $f(s) = \alpha_0 s^n + b_0 s^{n-1} + \alpha_1 s^{n-2} + b_1 s^{n-3} + \cdots + \alpha_m s + b_m \in R[s]$, $(\alpha_0 \neq 0)$ have negative real parts if and only if the following inequalities hold true,*

$$\alpha_0\Delta_1 > 0, \Delta_2 > 0, \alpha_0\Delta_3 > 0, \Delta_4 > 0, ..., \begin{matrix} \alpha_0\Delta_n > 0 \\ \Delta_n > 0 \end{matrix} \tag{4.36}$$

where $\Delta_1, ..., \Delta_n$ are defined as

$$\Delta_1 = H\begin{pmatrix} 1 \\ 1 \end{pmatrix} = b_0, \Delta_2 = H\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = \begin{vmatrix} b_0 & b_1 \\ \alpha_0 & \alpha_1 \end{vmatrix}, \cdots,$$

$$\Delta_n = H\begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix} = \begin{vmatrix} b_0 & b_1 & \cdots & b_{n-1} \\ \alpha_0 & \alpha_1 & \cdots & \alpha_{n-1} \\ 0 & b_0 & \cdots & b_{n-2} \\ 0 & \alpha_0 & \cdots & \alpha_{n-2} \\ . & . & . & . \end{vmatrix}. \tag{4.37}$$

and H is defined to be the Hurwitz matrix as

$$H = \begin{bmatrix} b_0 & b_1 & \cdots & b_{n-1} \\ \alpha_0 & \alpha_1 & \cdots & \alpha_{n-1} \\ 0 & b_0 & \cdots & b_{n-2} \\ 0 & \alpha_0 & \cdots & \alpha_{n-2} \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \left(\begin{array}{lll} \alpha_k = 0 & for & k > \frac{n}{2}, \\ b_k = 0 & for & k > \frac{n-1}{2}, \end{array} \right). \tag{4.38}$$

□

and Hermite-Biehler Theorem is given as:

Theorem 4.12 (Hermite-Biehler Theorem [Gan.,1]) *Let $p(s) = h(s^2) + sg(s^2)$, then $p(s)$ is Hurwitz if and only if there exist positive real λ_i, μ_i and c such that*

$$h(-\omega^2) = (\lambda_1 - \omega^2)(\lambda_2 - \omega^2)\cdots(\lambda_{l_1} - \omega^2) \tag{4.39}$$

$$g(-\omega^2) = c(\mu_1 - \omega^2)(\mu_2 - \omega^2)\cdots(\mu_{l_2} - \omega^2) \tag{4.40}$$

$$0 < \lambda_1 < \mu_1 < \lambda_2 < \mu_2 < \cdots$$

where $\begin{cases} l_1 = \frac{n}{2}, & l_2 = \frac{n}{2} - 1 \quad for\ n\ even \\ l_1 = \frac{n+1}{2}, & l_2 = \frac{n-1}{2} \quad for\ n\ odd \end{cases}$

□

When the polynomial coefficients are subject to perturbations, the Hurwitzness of all possible polynomials with the given perturbations can be checked in the frequency domain by making use of the Hermite-Biehler theorem and continuity argument ([Arg.,1], [Arg.,2]). Furthermore, by solving a set of inequality equations iteratively, the maximal allowable perturbations can be decided. The main results are outlined as follows:

Consider the nominal polynomial

$$p(s) = a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0 \in R[s] \tag{4.41}$$

which is assumed to be Hurwitz. Let the coefficients a_i, $i = 0, 1, ..., n$ be subject to perturbations by an amount δa_i, $i = 0, 1, ..., n$. The first problem is to check whether the polynomials are Hurwitz when $|\delta a_i| \leq \triangle a_i$, $i = 0, 1, ..., n$, while the second is to calculate the maximal $\triangle a_i$ which guarantees the Hurwitzness of the perturbed polynomials. Denote the set of all possible polynomials as $P(s, \delta a)$.

If only the odd (even) coefficients were to be perturbed, the result is given as:

Theorem 4.13 [Arg.,2] *The perturbed polynomials $P(s, \delta a)$ remain Hurwitz if the odd coefficients satisfy*

$$\max_{\triangle a_i,\ i=1,3,5,\ldots} \triangle g(\omega_k) \leq |g(j\omega_k)|, \quad \forall \omega_k \tag{4.42}$$

where ω_k *are the frequencies at which the polynomial* $h(s)$ *satisfies* $h(j\omega_k) = 0$ *and*

$$\triangle g(\omega_k) = \triangle a_1 \omega_k + \triangle a_3 \omega_k^3 + \triangle a_5 \omega_k^5 + \cdots \tag{4.43}$$

□

Define

$$\begin{aligned} g_1(j\omega) &= g(j\omega) - \triangle g(j\omega) \\ g_2(j\omega) &= g(j\omega) + \triangle g(j\omega) \end{aligned}$$

where $\triangle g(\omega)$ is an odd-coefficient polynomial which satisfies equation (4.42).

Since $\triangle g(\omega) > 0, \ \forall\omega, \ \omega \neq 0$, it is obvious that all odd-coefficient polynomials $g(j\omega)$ with perturbed coefficients satisfying equations (4.42) also satisfy

$$|g_1(j\omega)| \leq |g(j\omega)| \leq |g_2(j\omega)|, \quad \forall\omega. \tag{4.44}$$

Let $\omega_{l_0}, \omega_{l_1}$ and $\omega_{l_2}, l = 1, 2, ...$, be the intersection frequencies of the polynomials $g(j\omega), g_1(j\omega)$ and $g_2(j\omega)$ with the real axis, respectively, and let the frequency band

$$\triangle\omega_l = \{\omega : \omega_{l_{\min}} \leq \omega \leq \omega_{l_{max}}\} \tag{4.45}$$

be a band of frequencies centred around ω_{l_0} and bounded by $\omega_{l_{\max}}$ and $\omega_{l_{\min}}$, where $\omega_{l_{\max}}$ and $\omega_{l_{\min}}$ are the larger and smaller frequencies of $\{\omega_{l_1}, \omega_{l_2}\}$, respectively. Therefore the frequencies ω_l at which any perturbed polynomial $p(j\omega)$ with $g(j\omega)$ satisfying equation (4.42) intersects the real axis will lie inside the band $\triangle\omega_l$, i.e.

$$\omega_{l_{\min}} \leq \omega_l \leq \omega_{l_{\max}}. \tag{4.46}$$

The result can be stated as:

Theorem 4.14 [Arg.,2] *Let the odd-coefficient perturbations be chosen such that they satisfy equation (4.42). The perturbed polynomial will then be Hurwitz, if the complementary set of even-coefficient perturbations satisfies*

$$\max_{\omega_l} |\triangle h(j\omega_l)| \leq \min |h(j\omega_l)|, \tag{4.47}$$

$$\omega_{l_{\min}} < \omega_l < \omega_{l_{\max}}. \tag{4.48}$$

□

4.4 Kharitonov type results

In the last decade, perturbation results on polynomials have made great advance due to Kharitonov's theorem [Kha.,1]. In this section, the Kharitonov's result will be reviewed followed by later results developed along the same line.

Theorem 4.15 [Kharitonov's Theorem] [Kha.,1] *The polynomials*

$$P(s) = \alpha_0 s^n + \alpha_1 s^{n-1} + \cdots + \alpha_{n-1} s + \alpha_n, \quad \alpha_k \in [u_k, v_k], \quad u_k \le v_k. \tag{4.49}$$

where the real coefficients α_k take arbitrary values in the closed intervals $[u_k, v_k]$, are strictly Hurwitz if and only if the following four polynomials are strictly Hurwitz:

$$P_1(s) = v_0 s^n + u_1 s^{n-1} + u_2 s^{n-2} + v_3 s^{n-3} + \cdots \tag{4.50}$$

$$P_2(s) = v_0 s^n + v_1 s^{n-1} + u_2 s^{n-2} + u_3 s^{n-3} + \cdots \tag{4.51}$$

$$P_3(s) = u_0 s^n + u_1 s^{n-1} + v_2 s^{n-2} + v_3 s^{n-3} + \cdots \tag{4.52}$$

$$P_4(s) = u_0 s^n + v_1 s^{n-1} + v_2 s^{n-2} + u_3 s^{n-3} + \cdots \tag{4.53}$$

□

It has been proved that the number of polynomials to be checked can be reduced when the order of the polynomials is less than 6 [And. Jur. & Man.,1]. In the literature, many simpler proofs were provided, Bose [Bos. & Shi,1], Yeung [Yeu. & Wan.,1] are among them. Extensions to Kharitonov's results are abundant and are mainly along two directions. In Kharitonov's theorem the coefficients are assumed to be perturbed independently. When perturbations on the coefficients are dependent, the results yielded by Kharitonov's theorem are very conservative. So the first category of extensions is to introduce dependent perturbations in the coefficients. The result obtained in [Bar.,1] which can accommodate linear dependent perturbations is perhaps the best along this line of development. The second category of extension is to replace the left half complex plane by an arbitrary region, especially, the unit circle which is important for the stability test of discrete-time systems. Work such as those in [Soh. Ber. & Dab.,1], [Barm.,2], [Cie.,1], [Zeh.,1], [Vic.,1], [Tes. & Vic.,1], etc. is along this line. In the following, we examine the stability of the polynomials when the perturbations are linearly dependent.

Study the polynomials $P(s)$,

$$P(s) = s^n + (\alpha_1 + \delta_1)s^{n-1} + (\alpha_2 + \delta_2)s^{n-2} + \cdots + (\alpha_{n-1} + \delta_{n-1})s + (\alpha_n + \delta_n) \tag{4.54}$$

where the perturbations δ_i are linearly dependent on a set of independent variables, $q_1, q_2, \ldots, q_m$ and the set of independent variables take values from the intervals $q_i \in$

$[\underline{q}_i, \bar{q}_i]$. In this case, the polynomial set $P(s)$ can be expressed as a *polytope of polynomials*, i.e. the convex hull of m^2 vertex polynomials, or

$$P(s) = \sum_{i=1}^{m^2} \lambda_i p_i(s) \tag{4.55}$$

where $\sum_{i=1}^{m^2} \lambda_i = 1$ and the m^2 vertex polynomials are defined by the relation between the coefficient perturbations δ_i and the values of $\underline{q}_i$ and $\bar{q}_i$. So the stability of the set of perturbed polynomials is equivalent to the stability of the newly formed polytope.

Associate the vector $f \equiv \left| \begin{array}{cccc} a_1 & a_2 & \cdots & a_n \end{array} \right|$ with the monic polynomial

$$f(s) = s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n \in P[s] \tag{4.56}$$

and let $\Omega \subset R^n$ be the set of vectors which is the entire polytope of m polynomials. Define $\Omega \cap H$ as the exposed sets where H is a nontrivial supporting hyperplane. One dimensional exposed sets will be called the exposed edges while the two dimensional exposed sets are the exposed faces. Then the result concerning the root distribution of the entire polytope is given as:

Theorem 4.16 [Bart. Hol. & Lin,1] *Let $\Omega \subset R^n$ be a polytope of polynomials. Then the boundary of the root set of $f \in \Omega$ is contained in the roots set of all the exposed edges of Ω.*

□

So the test of Hurwitzness of the entire polytope of polynomials reduces to the test of the Hurwitzness of the finite number of exposed edges. Further, the Hurwitzness of the exposed edges can be implemented using the one variable root locus method.

4.4.1 Maximal allowable perturbations under independent perturbations

In real applications, when a system is subject to perturbations, it is often important to know how much a system can be perturbed before reaching instability. Various methods have been presented to calculate the maximal allowable perturbations based on Kharitonov's theorem. Under the assumption that the coefficients can be perturbed independently and the relative bounds for the coefficients are a priori, then the maximal allowable perturbation is obtained by Barmish [Barm.,1], Bialas et. al. [Bia. & Gar.,1] with the counterpart results in the discrete time systems given by Soh et. al. [Soh. Ber. & Dab.,1]. However, the coefficients of the characteristic polynomial are usually perturbed dependently on a set of physical variables. So maximisation of the stability can be expressed in terms of the maximisation of the parametric physical variables. Results given in [Tes. & Vic.,1] are along this line.

Assume the polynomial

$$f(s) = s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n \in P[s] \tag{4.57}$$

to be strictly Hurwitz. Assume the coefficients are perturbed as

$$a_i - v_i\epsilon < b_i < a_i + w_i\epsilon, \quad i = 1, 2, ..., n. \tag{4.58}$$

with known nonnegative weightings (v_i, w_i). Then it is to decide the maximal allowable perturbation ϵ for the whole set of polynomials defined as

$$P(s, \epsilon) = s^n + b_1 s^{n-1} + \cdots + b_{n-1}s + b_n. \tag{4.59}$$

Applying the Routh-Hurwitz stability test, the whole set of polynomials is stable if and only if the Hurwitz testing matrix (for the case when n is even):

$$H(b_1, b_2, ..., b_n) = \begin{bmatrix} b_n & b_{n-2} & b_{n-4} & \cdots & b_2 & 0 & 0 & \cdots & 0 \\ 1 & b_{n-1} & b_{n-3} & \cdots & b_1 & 0 & 0 & \cdots & 0 \\ 0 & b_n & b_{n-2} & \cdots & b_4 & b_2 & 0 & \cdots & 0 \\ 0 & 1 & b_{n-1} & \cdots & b_3 & b_1 & 0 & \cdots & 0 \\ 0 & 0 & b_n & \cdots & b_6 & b_4 & b_2 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & b_5 & b_3 & b_1 & \cdots & 0 \\ & & . & . & . & . & . & & \end{bmatrix} \tag{4.60}$$

has positive leading principal minors. Then the following theorem holds:

Theorem 4.17 [Barm.,1] *Define four $n \times n$ matrices,*

$$\begin{aligned} Q_1(\epsilon) &= H(a_n + w_n\epsilon, a_{n-1} + w_{n-1}\epsilon, a_{n-2} - v_{n-2}\epsilon, a_{n-3} - v_{n-3}\epsilon, \\ &\quad a_{n-4} + w_{n-4}\epsilon, a_{n-5} + w_{n-5}\epsilon, a_{n-6} - v_{n-6}\epsilon, a_{n-7} - v_{n-7}\epsilon, ...) \\ Q_2(\epsilon) &= H(a_n - v_n\epsilon, a_{n-1} - v_{n-1}\epsilon, a_{n-2} + w_{n-2}\epsilon, a_{n-3} + w_{n-3}\epsilon, \\ &\quad a_{n-4} - v_{n-4}\epsilon, a_{n-5} - v_{n-5}\epsilon, a_{n-6} + w_{n-6}\epsilon, a_{n-7} + w_{n-7}\epsilon, ...) \\ Q_3(\epsilon) &= H(a_n - v_n\epsilon, a_{n-1} + w_{n-1}\epsilon, a_{n-2} + w_{n-2}\epsilon, a_{n-3} - v_{n-3}\epsilon, \\ &\quad a_{n-4} - v_{n-4}\epsilon, a_{n-5} + w_{n-5}\epsilon, a_{n-6} + w_{n-6}\epsilon, a_{n-7} - v_{n-7}\epsilon, ...) \\ Q_4(\epsilon) &= H(a_n + w_n\epsilon, a_{n-1} - v_{n-1}\epsilon, a_{n-2} - v_{n-2}\epsilon, a_{n-3} + w_{n-3}\epsilon, \\ &\quad a_{n-4} + w_{n-4}\epsilon, a_{n-5} - v_{n-5}\epsilon, a_{n-6} - v_{n-6}\epsilon, a_{n-7} + w_{n-7}\epsilon, ...) \end{aligned}$$

Let $\Delta_{ij}(\epsilon)$ *denote the* j^{th} *leading principal minor of* $Q_i(\epsilon)$ *and define*

$$\epsilon_i^* \equiv \min\{\epsilon \geq 0, \textit{there exists a } j \leq n \textit{ such that } \Delta_{ij}(\epsilon) \leq 0\}, \quad i = 1,2,3,4. \quad (4.61)$$

Then it follows that

$$\epsilon_{\max} = \min\{\epsilon_i^*\}, \quad i = 1,2,3,4. \quad (4.62)$$

□

4.4.2 Maximal allowable perturbations under linearly dependent perturbations

The coefficients of the characteristic polynomial are in general functions of some underlying physical variables. When the set of physical variables are under uncertainty, so are the coefficients of the characteristic polynomial. Let $\underline{p} = [p_1, p_2, ..., p_p]^T$ be the vector of parameters on which the polynomial coefficients depend. Further assume that this is a subset of the parameters which are subject to uncertainty.

A class of uncertain polynomials of n^{th} order is defined as $P(s,\underline{p}) = s^n + a_1(\underline{p})s^{n-1} + a_2(\underline{p})s^{n-2} + \cdots + a_{n-1}(\underline{p})s + a_n(\underline{p})$, where $\underline{p} \in \Omega_p$ is a compact subset in the parameter space $\Omega_p \subset R^n$ and $a_i(\underline{p}) : \Omega_p \rightarrow R, \; i = 1,2,..,n$ are real continuous functions of parameter vector $\underline{p}$.

Denote the nominal parameter vector as $\underline{p}^o \in \Omega_p$ corresponding to the nominal polynomial $P^o(s,\underline{p}^o) \in R^n$. Further assume that the functions $a_i(\underline{p}), \; i = 1,2,...,n$ are affine in $\underline{p}$, and Ω_p is assumed to be a convex polytope in R^p. Thus the mapping between the coefficients and the parameter space can be expressed as

$$\underline{a} = H\underline{p} + \underline{h} \quad (4.63)$$

where $\underline{a}(\underline{p}) = \begin{bmatrix} a_1, & a_2, & ..., & a_n \end{bmatrix}^T \in R^n, H \in R^{n\times p}$, and $\underline{h} \in R^n$.

The uncertainty set Ω_p of the parameter is assumed to satisfy

$$\Omega_p = \Omega(\rho) = \{\underline{p} \in R^p : \|T(\underline{p} - \underline{p}^o)\|_\infty^w \leq \rho\} \quad (4.64)$$

where ρ is a positive value, $T \in R^{n\times p}$ is a full column rank operator and $\|\cdot\|_\infty^w$ represents a weighted l_∞ norm.

A more general type of admissible root region Λ can be defined as a finite number of connected regions. A polynomial is called Λ-stable if all the roots of the polynomial are in the region Λ. The admissible region is defined according to the interest of the analysis. When the stability of a continuous time system is under investigation, the region would be the left hand side of the complex plane. If the system is further required to have a maximal damping ratio, then it could be a circular sector, etc. Assume that the boundary of $\Lambda, \partial\Lambda$, can either be defined in analytical or numerical form, which is assumed as

$$\partial\Lambda = \{s \in C : s = F(\gamma), \gamma \in R\}. \tag{4.65}$$

For practical cases, the admissible region will always be bounded. Then the set of Λ-stable polynomials $P(\Lambda)$ in the n dimensional space is bounded by surfaces defined as

$$P(\partial\Lambda) = \{\underline{a} \in R^n : s^n + a_1 s^{n-1} + a_2 s^{n-2} + \cdots + a_n = 0, s \in \partial\Lambda\} \tag{4.66}$$

which is a linear manifold depending on the one degree freedom movement s on $\partial\Lambda$. Because only real coefficient polynomials are considered, the movement of s along $\partial\Lambda$ in the upper or lower complex plane will be sufficient.

The corresponding parameter subset relating to $P(\partial\Lambda)$ is given by

$$\Omega(\partial\Lambda) = \{\underline{p} \in R^p, P(s, \underline{p}) = 0, s \in \partial\Lambda\}. \tag{4.67}$$

which defines the set that bounds the domain $\Omega(\Lambda)$ of parameters generating Λ-stable polynomials.

For a given $T, \underline{p}^o$ and weightings w_i, in order to find the maximal ρ for which $\Omega(\rho)$ contains only the parameters which generate Λ-stable polynomials, the following optimisation problem needs to be solved:

$$\rho^o = \inf_{\underline{p}, \gamma, \rho} \{\rho\} \tag{4.68}$$

subject to

$$\begin{cases} \rho \geq 0 \\ -\rho w_i \leq t_i'(\underline{p} - \underline{p}^o) \leq \rho w_i, \quad i = 1, 2, ..., m \\ P(s, \underline{p}) = 0, \quad s = F(\gamma) \end{cases} \tag{4.69}$$

where $t_i', \ i = 1, 2, ..., m$ are the row vectors of the matrix T. For a given $z \in \partial\Lambda$, define the following value:

$$\rho_s(z) = \inf_{\underline{p}, \rho} \{\rho\} \tag{4.70}$$

subject to

$$\begin{cases} \rho \geq 0 \\ -\rho w_i \leq t_i'(\underline{p} - \underline{p}^o) \leq \rho w_i, \quad i = 1, 2, ..., m \\ P(z, \underline{p}) = 0. \end{cases} \tag{4.71}$$

When $P(z,\underline{p}) = 0$ is empty, then set $\rho_s(z) = \infty$. So the optimisation is equivalent to

$$\rho^o = \inf_{z : z = F(\gamma)} \{\rho_s(z)\} \tag{4.72}$$

which is a one-parameter optimisation problem for each $z \in \partial\Lambda$, a linear programming problem in $p+1$ parameters needs to be solved with $2m+3$ constraints. The above result is stated as the following theorem:

Theorem 4.18 [Barm.,2] *Let Λ be a given admissible region of the complex plane. Let $\underline{p}^o$ be a given nominal parameter vector generating a Λ-stable polynomial and let ρ^o given by (4.72). Then the polytope $\Lambda(\rho^o)$ is maximal in the class of admissible polytopes generating Λ-stable polynomials.*

□

4.5 Stability radii of matrices

In this section, we review the results on the stability radii of a matrix when perturbed either by structured or unstructured disturbances. Given a nominal stable system described by

$$\dot{\underline{x}} = A\underline{x} \tag{4.73}$$

where $A \in K^{n \times n}$. It is assumed that A can be perturbed by either real or complex perturbations, therefore K is either the real field $K = R$ or the complex field $K = C$. The system matrix is assumed to be perturbed to

$$A \rightarrow A + D\Delta E \tag{4.74}$$

where $D \in K^{n \times l}$, $E \in K^{q \times n}$, which are known and fixed, define the structure of the perturbation while $\Delta \in K^{l \times q}$ is the unknown disturbance. The matrices D, E are determined by the nature of the disturbance. Note that output feedback effects on the matrix A are also of the same type. If they are identity matrices, then the perturbation is called unstructured and otherwise structured.

If $\| \bullet \|_{K^l}$ and $\| \bullet \|_{K^q}$ are given norms on K^l and K^q respectively, the measure of the perturbation matrix $\Delta \in K^{l \times q}$ is given by the corresponding induced norm

$$\|\Delta\| = \max_{\|y\|_{K^q}=1} \{\|\Delta y\|_{K^l}\} \tag{4.75}$$

Partition the complex plane into two complementary regions

$$C = C_s \cup C_u \tag{4.76}$$

where C_s is an open set and denotes the stable region whereas C_u unstable region. So for continuous time systems

$$C_s = C^- = \{s \in C : \text{Re } s < 0\} \text{ and } C_u = C^+ = \{s \in C : \text{ Re } s \geq 0\}$$

From the assumption, the nominal system satisfies

$$A \in K^{n \times n}, \quad \sigma(A) \subset C_s \tag{4.77}$$

The stability radius of A with respect to perturbations is defined as

Definition 4.1 : *Given a partition (4.76) and the perturbation norm (4.75), the* Stability Radius *of $A \in K^{n\times n}$ with respect to perturbations of the structure $(D, E) \in K^{n\times l} \times K^{q\times n}$ is defined by*

$$\gamma_K = \gamma_K(A; D, E; C_u) = inf\Big\{\|\Delta\|; \Delta \in K^{l\times q}, \sigma(A + D\Delta E) \cap C_u \neq 0\Big\} \tag{4.78}$$

If both the structure matrices D, E are the identity matrix I_n we obtain the unstructured stability radius

$$d_K = \gamma_K(A; I, I; C_u) = min\Big\{\|\Delta\|; \Delta \in K^{l\times q}, \sigma(A + \Delta) \cap C_u \neq 0\Big\} \tag{4.79}$$

$d_K(A, C_u)$ is the distance, within the normed space $(K^{n\times n}, \| \bullet \|)$, between A and the set of unstable matrices in $K^{n\times n}$

$$\mathcal{U}(K; C_u) = \Big\{X \in K^{n\times n}; \sigma(X) \cap C_u \neq 0\Big\} \tag{4.80}$$

□

If A, D, E are real, according to whether $\Delta \in C^{q\times l}$ or $\Delta \in R^{q\times l}$, two stability radii, γ_R or γ_C can be derived. They are called complex and real stability radii of A, respectively. Clearly

$$\gamma_R(A; D, E; C_u) \geq \gamma_C(A; D, E; C_u) \geq 0 \tag{4.81}$$

For mathematical concreteness, $\gamma_K = \gamma_k(A; D, E; C_u) = \infty$ if and only if there does not exist $\Delta \in K^{l\times q}$ with $\sigma(a + D\Delta E) \cap C_u \neq 0$.

Remark 4.2: The stability radius γ_K in invariant under similarity transformations

$$\gamma_K(A; D, E; C_u) = \gamma_K(TAT^{-1}; TD, E^{-1}; C_u) \tag{4.82}$$

Moreover, if $\sigma(A + D\Delta E) \cap \text{int}(C_u) \neq 0$ where $\text{int}(C_u)$ denotes the set of interior points of C_u, then by continuity argument the same is true for all matrices in a small neighbourhood of $\Delta \in K^{l \times q}$. Therefore, if A is C_s-stable

$$\gamma_K(A; D, E; C_u) = \gamma_K(A; D, E; \partial C_u) \tag{4.83}$$

where ∂C_u is the boundary of C_u.

□

Define $G(s)$ and the transfer matrix associated with the triple (A, D, E)

$$G(s) = E(sI - A)^{-1} D \tag{4.84}$$

then the stability radius can be characterised by $\max_{s \in \partial C_b} \|G(s)\|$:

Proposition 4.1 [Hin. & Pri.,1] *Given $A \in K^{n \times n}, \sigma(A) \subset C_s$, and it is perturbed to $A + D\Delta E$, then the stability radius as defined in (4.78) satisfies*

$$\gamma_K(A; D, E; C_u) \geq \left[\max_{s \in \partial C_u} \|G(s)\|\right]^{-1}. \tag{4.85}$$

□

The stability radius is related to the solution of the following optimisation problem

$$J_\rho(\underline{x}_0, \underline{v}) = \int_0^\infty \left[\|\underline{v}\|^2 - \rho^2 \|\underline{z}\|^2\right] dt \tag{4.86}$$

where

$$\begin{aligned} \dot{\underline{x}}(t) &= A\underline{x}(t) + D\underline{v}(t), \quad t \geq 0, \quad \underline{x}(0) = \underline{x}_0 \\ \underline{z}(t) &= E\underline{x}(t) \end{aligned}$$

and the associated parametrised algebraic Riccati equation

$$A^* X + XA - XDD^* X - \rho^2 E^* E = 0 \tag{4.87}$$

Let $\mathcal{H}_n \subset C^{n \times n}$ denote the real vector space of all Hermitian $n \times n$ matrices and $\mathcal{H}_n^+$ (resp. $\mathcal{H}_n^-$) the convex cones of positive semi-definite (resp. negative semi-definite) matrices in $\mathcal{H}_n$. Then the solution of the optimisation problem is related to the stability radius by the following theorem.

Theorem 4.19 [Hin. & Pri.,1] *Suppose $\sigma(A) \subset C^-, \rho > 0$. Then*

(i). *There exists a solution $P_\rho \in \mathcal{H}_n^-$ of (4.87) such that $\sigma(A - DD^* P_\rho) \subset C^-$ and P_ρ is unique among all Hermitian solutions with this property if $\rho < \gamma_C$;*

(ii). *If* $\rho = \gamma_C$ *there exists a solution* $P_\rho \in \mathcal{H}_n^-$ *satisfying* $\sigma(A - DD^* P_\rho) \subset C^-$ *and* $\sigma(A - DD^* P_\rho) \cap jR \neq 0$;

(iii). *If* $\rho > \gamma_C$ *there does not exist any Hermitian solution to (4.87).*

If (A, E) *is observable then* P_ρ *is negative definite for all* $\rho \in (0, \gamma_C]$.

□

The parametrised Hamiltonian matrix associated with (4.87) is

$$H_\rho = \begin{bmatrix} A & -DD^* \\ \rho^2 E^* E & -A^* \end{bmatrix} \tag{4.88}$$

An algorithm for computing γ_c is based on the following characterisation of γ_c in terms of H_ρ.

Proposition 4.2 [Hin. Kel. & Lin.,1] *If* H_ρ *is defined by (4.88), then*

$$\rho < \gamma_C(A; D, E) \text{ iff } \sigma(H_\rho) \cap iR = 0 \tag{4.89}$$

Moreover,

$$iw_0 \in \sigma(H_{\gamma_C}) \text{ iff } \|G(iw_o)\| = \max_{w \in R} \|G(iw)\|. \tag{4.90}$$

□

4.6 Minimal norm stabilisation

4.6.1 Stabilisation of a polynomial with minimal perturbation

An equally important problem to that of the stability robustness of a stable system, is the minimum norm stabilisation problem of an unstable system. This is to find a perturbation which stabilises the system while the norm of the perturbation in a certain sense is minimised [Kou.,1]. This will provide very useful information in bounded norm stabilisation. The results presented in the previous section can not be applied directly because of the fundamental assumption that the nominal system has to be stable. Here we present a method which can calculate both the minimum distance of an unstable system away from stable ones and the minimum distance of a stable system away from unstable ones.

We study the nominal characteristic polynomial of a system

$$f_o(s) = s^n + \alpha_1^o s^{n-1} + \cdots + \alpha_{n-1}^o s + \alpha_n^o \tag{4.91}$$

which is unstable. The nominal system is assumed to be perturbed by

$$\Delta f(s) = \delta_1 s^{n-1} + \delta_2 s^{n-2} + \cdots + \delta_n \tag{4.92}$$

where $\delta_1, \delta_2, ..., \delta_n$ denote the effect of parametric uncertainties or feedback as perturbations on the coefficients. If we define

$$Q(\Delta f) = \delta_1^2 + \delta_2^2 + \cdots + \delta_n^2$$

as the performance index to be minimised. Let $\Lambda(f)$ be the root set of $f(s)$; then the problem under consideration is:

- For unstable nominal system $\Lambda(f_0) \cap C^- \neq 0$, find a perturbation with minimum norm γ_{f^+} which stabilises the nominal system:

$$\gamma_{f^+} = \inf\left\{Q(\Delta f) : \Lambda\left(f_o(s) + \Delta f(s)\right) \cap C^- = 0\right\} \tag{4.93}$$

  where C^-, C^+ denote the left and right half of the complex plane.

□

The following stability theorem concerning polynomials provides the basis of this investigation.

Theorem 4.20 [Gan.,1] *Define for a given polynomial $f(s)$*

$$f(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_{n-1} s + \alpha_n \tag{4.94}$$

a rational function $r(s)$

$$r(s) = \begin{cases} \dfrac{Ev\{f(s)\}}{Od\{f(s)\}} & \text{if } n \text{ is even} \\ \\ \dfrac{Od\{f(s)\}}{Ev\{f(s)\}} & \text{if } n \text{ is odd} \end{cases} \tag{4.95}$$

where $Ev\{f(s)\}$ is the even part and $Od\{f(s)\}$ the odd part of $f(s)$. This rational function may be expanded as

$$r(s) = a_1 s + \cfrac{1}{a_2 s + \cfrac{1}{a_3 s + \cdots \cfrac{1}{a_{n-1}s + \cfrac{1}{a_n s}}}} \tag{4.96}$$

where a_i's are the ratios of two successive parameters in the first column of the Routh table. Then the number of negative a_i is equal to the number of the roots of $f(s)$ in C^+.

□

From equations (4.94–4.96), the relations between α_i and $a_i, i = 1, 2, ..., n$ can be established. Indeed, the coefficients, α_i, of the polynomial can be expressed in terms of a_i. For instance when $n = 4$, we have

$$\begin{aligned} r(s) &= a_1 s + \cfrac{1}{a_2 s + \cfrac{1}{a_3 s + \frac{1}{a_4 s}}} \\ &= \frac{a_1 a_2 a_3 a_4 s^4 + (a_1 a_2 + a_1 a_4 + a_3 a_4) s^2 + 1}{a_2 a_3 a_4 s^3 + (a_2 + a_4) s} \end{aligned} \tag{4.97}$$

so the fourth order polynomial can be expressed in terms of $a_i, i = 1, 2, 3, 4$ as

$$f(s) = s^4 + \frac{1}{a_1} s^3 + \frac{a_1 a_2 + a_1 a_4 + a_3 a_4}{a_1 a_2 a_3 a_4} s^2 + \frac{a_2 + a_4}{a_1 a_2 a_3 a_4} s + \frac{1}{a_1 a_2 a_3 a_4} \tag{4.98}$$

This facilitates a representation of the stable polynomials in terms of the coefficients a_i. For the fourth order case, for instance, the set of all stable polynomials $P^+[s]$ can be represented as

$$\begin{aligned} P^+[s] &\equiv \left\{ f(s) : \Lambda(f) \subset C^+ \right\} \\ &= \left\{ f(s) = s^4 + \frac{1}{a_1} s^3 + \frac{a_1 a_2 + a_1 a_4 + a_3 a_4}{a_1 a_2 a_3 a_4} s^2 + \frac{a_2 + a_4}{a_1 a_2 a_3 a_4} s + \frac{1}{a_1 a_2 a_3 a_4}; a_i > 0 \right\} \end{aligned}$$

The minimisation problem as defined above can now be rephrased for the fourth order case as

$$\gamma_{f^+} = \inf \left\{ Q(\Delta(f)) : a_i > 0, i = 1, 2, 3, 4 \right\} \tag{4.99}$$

where $Q(\Delta(f))$ has a concrete form:

$$Q(\Delta(f)) = (\frac{1}{a_1} - \alpha_1^o)^2 + (\frac{a_1 a_2 + a_1 a_4 + a_3 a_4}{a_1 a_2 a_3 a_4} - \alpha_2^o)^2 + (\frac{a_2 + a_4}{a_1 a_2 a_3 a_4} - \alpha_3^o)^2 + (\frac{1}{a_1 a_2 a_3 a_4} - \alpha_4^o)^2 \tag{4.100}$$

A procedure for calculating the distance that an unstable polynomial away from stable polynomials, γ_{f^+}, can be defined as shown below:

Minimum distance from stability:

1. For a given unstable polynomial $f_o(s)$, first construct the performance index $Q(\Delta(f))$ as in equation (4.100) where a_i are considered as the variables;

2. Choosing an initial vector $(a_1^{(0)}, a_2^{(0)}, ..., a_n^{(0)}) > 0$ and minimise $Q(\Delta(f))$ subject to $a_i > 0, i = 1, 2, 3, 4$.

Different constrained minimisation techniques can be used here. The procedure terminates at a local or a global minimum.

4.6.2 Destabilisation of a polynomial with minimal perturbations

The problem of destabilising a polynomial with minimal perturbations has been considered within the Kharitonov's framework [Barm.,1] etc. An alternative is to adopt an algorithmic approach similar to what has been presented in the previous section. We study the nominal characteristic polynomial of a system

$$f_o(s) = s^n + \alpha_1^o s^{n-1} + \cdots + \alpha_{n-1}^o s + \alpha_n^o \tag{4.101}$$

which is stable. The nominal system is assumed to be perturbed by

$$\Delta f(s) = \delta_1 s^{n-1} + \delta_2 s^{n-2} + \cdots + \delta_n \tag{4.102}$$

where $\delta_1, \delta_2, ..., \delta_n$ denote the effect of parametric uncertainties or feedback as perturbations on the coefficients. If we define $Q(\Delta f) = \delta_1^2 + \delta_2^2 + \cdots + \delta_n^2$ as the performance index to be minimised. Let $\Lambda(f)$ be the root set of $f(s)$; then the problem under consideration can be formulated as:

- For stable nominal system $\Lambda(f_0) \cap C^- = 0$, find a perturbation with minimum norm γ_{f^-} which destabilises the nominal system:

$$\gamma_{f^-} = \inf\left\{Q(\Delta f) : \Lambda\left(f_o(s) + \Delta f(s)\right) \cap C^- \neq 0\right\} \tag{4.103}$$

□

By deploying Theorem (4.20), for the fourth order case, the set of all unstable polynomials $P^-[s]$ can be represented as

$$\begin{aligned} P^-[s] &\equiv \left\{f(s) : \Lambda(f) \cap C^- \neq 0\right\} \\ &= \{f(s) = s^4 + \frac{1}{a_1}s^3 + \frac{a_1a_2 + a_1a_4 + a_3a_4}{a_1a_2a_3a_4}s^2 + \frac{a_2 + a_4}{a_1a_2a_3a_4}s + \frac{1}{a_1a_2a_3a_4}; \\ &\qquad \exists a_i < 0, i = 1,2,3,4\} \end{aligned}$$

The minimisation problem as defined above can now be rephrased for the fourth order case as

$$\gamma_{f^-} = \inf\ \{Q(\Delta(f)) : \exists a_i < 0, i = 1,2,3,4\} \tag{4.104}$$

where $Q(\Delta(f))$ has a concrete form:

$$Q(\Delta(f)) = (\frac{1}{a_1} - \alpha_1^o)^2 + (\frac{a_1a_2 + a_1a_4 + a_3a_4}{a_1a_2a_3a_4} - \alpha_2^o)^2 + (\frac{a_2 + a_4}{a_1a_2a_3a_4} - \alpha_3^o)^2 + (\frac{1}{a_1a_2a_3a_4} - \alpha_4^o)^2 \tag{4.105}$$

A procedure for calculating the distance that a stable polynomial is from unstable polynomials, γ_{f^-}, can be defined as shown below:

Minimum distance from stability:

1. For a given unstable polynomial $f_o(s)$, first construct the performance index $Q(\Delta(f))$ as in equation (4.105) where a_i are considered as the variables;
2. Choosing an initial vector $(a_1^{(0)}, a_2^{(0)}, ..., a_n^{(0)}) > 0$, which corresponds to a stable polynomial, and minimise $Q(\Delta(f))$ subject to the following different conditions

   (i) One of the coefficients a_i is negative.

   (ii) Two of the coefficients a_i are negative.

   (iii) Three of the coefficients a_i are negative.

   (iv) All the coefficients are negative.

Different constrained minimisation techniques can be used here. The procedure terminates at a local or a global minimum.

4.6.3 Examples

In this subsection, two examples are presented.

Example (4.1): For the first example, we consider an unstable polynomial which is given as

$$f^o(s) = s^4 + 4s^3 - 6s^2 + 5s + 2 \tag{4.106}$$

and its roots are $-0.28518, -5.29736$, and $0.79127 \pm j0.83534$. We try to find the stable polynomial which is of a minimum distance from it. The performance index is defined in the following form:

$$Q(\Delta(f)) = (\frac{1}{a_1} - 4)^2 + (\frac{a_1a_2 + a_1a_4 + a_3a_4}{a_1a_2a_3a_4} + 6)^2 + (\frac{a_2 + a_4}{a_1a_2a_3a_4} - 5)^2 + (\frac{1}{a_1a_2a_3a_4} - 2)^2 \tag{4.107}$$

and by employing the procedures outlined in the previous section, the optimisation problem can be solved with $Q(\Delta f) = 52.1166$. The perturbed polynomial is

$$f^o(s) + \Delta f(s) = s^4 + 4.98773s^3 + 0.0002s^2 + 0.0004s + 1.005^{-8} \tag{4.108}$$

which has its roots at $-4.9877, -0.000027$, and $-1.929^{-8} \pm j8.611^{-3}$.

Example (4.2): The second example is to find a destabilising perturbation with minimum perturbation. The stable polynomial is given as

$$f^o(s) = s^4 + 4s^3 + 6s^2 + 5s + 2 \tag{4.109}$$

whose roots are $-1, -2$, and $-0.5 \pm j0.866$.

The optimisation problem is to solve the following

$$\begin{aligned}\text{minimise} \quad Q(\Delta(f)) &= (\frac{1}{a_1} - 4)^2 + (\frac{a_1 a_2 + a_1 a_4 + a_3 a_4}{a_1 a_2 a_3 a_4} - 6)^2 \\ &+ (\frac{a_2 + a_4}{a_1 a_2 a_3 a_4} - 5)^2 + (\frac{1}{a_1 a_2 a_3 a_4} - 2)^2\end{aligned}$$

subject to

$$\text{any or all } a_i \leq 0, i = 1, 2, 3, 4$$

The minimum perturbation is found to be

$$\Delta f(s) = -0.00015s^3 - .000073s^2 + 0.000159s - 2.000053 \tag{4.110}$$

with $Q(\Delta(f)) = 4.00021$ and the perturbed polynomial has a set of roots at $0, -2.353$, and $-0.823388 \pm j1.202912$.

In both of the above examples, NAG library routines E04HBF and E04JBF have been used to carry out the minimisation and E04JBQ is used for finding the roots of the polynomials. The programmes can be found in Appendix 2 A & B.

4.7 Summary

The study of the relation between the coefficients and the root distributions of polynomials has been a subject of both immense theoretical and applications interest. Since the connection between the system stability and the root distribution of the corresponding characteristic polynomial was established, a tremendous amount of effort has been devoted to the further exploration of the subject. Because the system structure and system parameters are subject to uncertainties, so are the coefficients and the roots of the characteristic polynomial of the systems. The main effort is devoted to the study of the coefficients and in turn the roots of the characteristic polynomial due to parametric changes. Recent development motivated by Kharitonov's result in the subject is fast and still foregoing with a tremendous momentum. In this chapter, some classical results concerning the relation between the roots and the coefficients of polynomials have been reviewed first, together with some well-known perturbation results. These results are very general and have a wide range of applications. Then we surveyed results in the recently developed area due originally to Kharitonov's result.

In the study of stability robustness against parameter variations, it is very important to know how far a stable system is away from instability when the system is subject to either structured or unstructured perturbations. Some results due to Hinrichsen and Pritchard in this direction have also been reviewed. Further we studied the dual problem of finding the distance of an unstable polynomial from the set of stable ones. A unified procedure has been designed to calculate the minimal distance of an unstable polynomial from the stability domain as well as the minimal distance of a stable polynomial from the set of unstable polynomials. The generality of the classical results has important implications and yields a reduced sharpness. When we restrict ourselves to the study of some special groups of polynomials with special type of constraints on the coefficients, better results can be established. These will be further explored in the next chapter.

Chapter 5

ROOT DISTRIBUTIONS OF BOUNDED COEFFICIENT POLYNOMIALS AND SUM OF TWO POLYNOMIALS

5.1 Root Region of Bounded Coefficient Polynomials

5.1.1 Introduction

The classical problem of establishing relationships between coefficients of polynomials and roots [Mar.,1], has been an integral part of the study of stability and performance of linear control systems [Bar.,1], [Barm.,2]. An important problem arising in the bounded norm feedback design [Kar. & Shan,1] of linear systems, as well as in the robust design [Lun.,1] is the study of relationships between the norm of the coefficient vector of the polynomial and its location of roots in the complex plane. The two fundamental questions which are related to the root distribution and have to be studied are:

Problem 1: Define the infimal region Γ_γ^1 of the complex plane such that all polynomials of degree n and maximal norm γ have all their roots in it.

Problem 2: Define the maximal region Γ_γ^2 of the complex plane such that every polynomial having its roots there has norm less or equal to γ.

The solution of Problem 1 provides necessary conditions for the location of the roots, based on the value of the norm of the coefficient vector, but by no means sufficient. In fact, in Γ^1_γ we may find polynomials having all their roots there, but having norm larger than γ. The solution of Problem 2 aims at defining a subset of Γ^1_γ, say Γ^2_γ for which the norm constraint is automatically satisfied, as long as the polynomial has its roots there. Thus Problem 2 is linked to sufficiency conditions. Problem 1 is a classical problem addressed in the theory of polynomials with complex coefficients [Mar.,1]. In this chapter, as far as the first problem is concerned, we are to define the region Γ^1_γ and not just upper bounds as those defined in the literature for the case of stable or totally unstable polynomials. The second problem has not been addressed before in the classical literature. Problem 1 will be referred to as the *Direct Problem* of root enclosure, whereas Problem 2 as the *Inverse Problem* of root enclosure. These two problems are studied and solved in the case of stable and totally unstable polynomials. The boundaries of the Γ^1_γ and Γ^2_γ regions are defined as branches of algebraic functions parametrised by the norm value. The general case of polynomials with roots both in the left and right half of the complex plane is finally addressed; the direct problem is solved in the third degree case, whereas the inverse problem is solved in the general n-th order case. The results are presented for the case of the l_2-norm, but may readily be extended to other norms, such as the l_1, or l_∞-norm.

In Section 5.1.2, the problems to be solved are more clearly defined and it is shown how the classical results presented in the previous chapter can be used to establish the existence of upper bounds of the *Direct Problem.* In Section 5.1.3, the *Direct Problem* and the *Inverse Problem* are considered for the polynomials with all their roots in the left half of the complex plane as well as for the polynomials which have all their roots in the right half complex plane. In Section 5.1.4, general third polynomials are studied. Tighter solutions to the *Direct Problem* are obtained and for the *Inverse Problem*, maximal rectangular regions are established.

5.1.2 Definition of Problems

The problems studied here can be summarised as follows:

Problem 1: Define the region Γ^1_γ of the C-plane such that:

(i) $\forall f(s) \in P^\gamma[s], \Lambda_f \in \Gamma^1_\gamma$;

(ii) $\forall f(s) \in P[s]$ with $\Lambda_f \cap \Gamma^{1c}_\gamma \neq O$ (Γ^{1c}_γ is the C-complement of Γ^1_γ), $\|\underline{\alpha}_f\|_2 > \gamma$;

(iii) Γ^1_γ is the smallest region satisfying the above properties.

□

The classical results summarised in Chapter 4, define upper bounds (in terms of disks centred at the origin) with the property that they contain all zeros of $f(s) \in P^{\gamma}[s]$. For instance, from Corollary (4.1), a necessary region can be defined as the disc centred at the origin with a radius $|\rho| < \sqrt{1+\gamma^2}$. It is not clear, however, whether they have the second of the properties stipulated above and they are not the smallest of the regions characterised by properties (i) and (ii).

Defining the region Γ_{γ}^{1} does not necessarily imply that if for some $f(s) \in P[s]$, $\Lambda_f \in \Gamma_{\gamma}^{1}$, then $\|\underline{\alpha}_f\| \leq \gamma$, or that all polynomials with $\Lambda_f \in \Gamma_{\gamma}^{1}$ are elements of $P^{\gamma}[s]$. This leads to the study of the second problem.

Problem 2: Define a region Γ_{γ}^{2} of the C-plane such that:

(i) For $\forall f(s) \in P[s]$ with $\Lambda_f \in \Gamma_{\gamma}^{2}$, then $f(s) \in P^{\gamma}[s]$ (i.e. $\|\underline{\alpha}_f\| \leq \gamma$);

(ii) The region Γ_{γ}^{2} is the maximal region for which the above property holds true.

□

This problem has not been addressed before in the literature and it is motivated by the needs of the bounded norm feedback design. Problem (1) and (2) will be also referred to as the *Direct* and *Inverse problems of root enclosure.* Note that the existence of regions Γ_{γ}^{2} that satisfy the first of the properties has to be established before the second part is addressed.

5.1.3 Case of Stable Polynomials

In this section, Problem 1 and Problem 2 are studied in the case of stable polynomials $f(s) \in P^{+}[s]$. The existence and boundaries of the $\Gamma_{\gamma}^{1}, \Gamma_{\gamma}^{2}$ are established for this family of polynomials. The results naturally extend to the totally unstable polynomials.

5.1.3.1 Minimal necessary Root Region for Stable Polynomials with Norm Bounded Coefficients

In this subsection, the relation between the root region and the subset of the stable polynomials $f(s) \in \Gamma^{+,\gamma}[s]$ whose coefficient norms are upper bounded by γ is established. This is the smallest of the regions which contain all the roots of the norm bounded polynomials and it will be termed as the Γ-Prime Region of all

$f(s) \in P^+[s]$. This Γ-Prime Region provides the solution to Problem1 in the case of stable polynomials.

First we prove the following two Lemmas, which establish the relation between the norms of two polynomials $f(s), g(s)$ which are polynomials whose coefficients are all positive.

Lemma 5.1 *If $f(s)$ is a monic n-th degree polynomial whose coefficients are positive and $g(s)$ is a polynomial of degree less than or equal to $n-1$ having all its coefficients non-negative and they satisfy*

$$f(s) = g(s) + h(s) \tag{5.1}$$

then $\|\underline{\alpha}_f\| \geq \|\underline{\alpha}_g\|$ with equality holding only when $h(s) = 0$ or $f(s) = g(s)$.

Proof:

Without loss of generality, we assume $h(s)$ to be of order $n-1$, so

$$f(s) = s^n + \alpha_1 s^{n-1} + \ldots + \alpha_n \tag{5.2}$$

$$g(s) = s^n + \beta_1 s^{n-1} + \ldots + \beta_n \tag{5.3}$$

$$h(s) = t_1 s^{n-1} + \ldots + t_n \tag{5.4}$$

From the assumptions of $f(s), g(s)$ and $h(s), \alpha_i \geq 0, \beta_i \geq 0$ and $t_i \geq 0 (i = 1, 2, ..., n)$. So

$$\begin{aligned} f(s) &= g(s) + h(s) \\ &= s^n + (\beta_1 + t_1)s^{n-1} + \ldots + (\beta_n + t_n) \end{aligned} \tag{5.5}$$

$$\begin{aligned} \|\underline{\alpha}_f\| &= \sqrt{a_1^2 + a_2^2 + \ldots + a_n^2} \\ &= \sqrt{(\beta_1 + t_1)^2 + (\beta_2 + t_2)^2 + \ldots + (\beta_n + t_n)^2} \end{aligned} \tag{5.6}$$

So $\|\underline{\alpha}_f\| = \|\underline{\alpha}_g\|$ holds only when $t_i = 0$ for all $i = 1, 2, ..., n$.

□

Remark 5.1: The above Lemma holds true for the case when the polynomials $f(s)$ and $g(s)$ are both stable.

□

Let $f(s) \in P^+[s]$. If the polynomial has k real roots, then the root set is $\Lambda = \{-\sigma_1, ..., -\sigma_k, -\sigma_{k+1} \pm j\omega_{k+1}, ..., -\sigma_{(n-k)/2} \pm j\omega_{(n-k)/2}\}$ and define

$$\mu(f) = \{\max|\sigma_i|, i = 1, 2, ..., (n-k)/2\} \tag{5.7}$$

$$\nu(f) = \{\max|\omega_i|, i = k+1, k+2, (n-k)/2\} \tag{5.8}$$

Then we can prove the following Lemma.

Lemma 5.2 *Let $f(s) \in P^+[s]$ and define $\mu(f), \nu(f)$ correspondingly. Let $f_\mu(s) = s^{n-1}[s - \mu(f)], f_\nu(s) = s^{n-2}[s^2 + \nu^2(f)]$, and $\underline{\alpha}_{f_\mu}, \underline{\alpha}_{f_\nu}$ be the coefficient vectors of $f_\mu(s), f_\nu(s)$. Then the following are true:*

$$(i) \quad \|\underline{\alpha}_{f_\mu}\| \leq \|\underline{\alpha}_f\|$$
$$(ii) \quad \|\underline{\alpha}_{f_\nu}\| \leq \|\underline{\alpha}_f\|.$$

Proof:

(i) First we assume $\mu(f) = \sigma_l$ with $l < k$. Then

$$\begin{aligned} f(s) &= \Pi_{i=1,i\neq l}^{k}(s+\sigma_i)\Pi_{j=k+1}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2](s+\sigma_l) \\ &= s^{n-1}(s+\sigma_l)+\triangle f(s) \end{aligned} \tag{5.9}$$

By direct calculation, $f(s)$ is of order at most $n-1$ and all its coefficients non-negative; therefore by Lemma (5.1) $\|\underline{\alpha}_f\| \geq \|\underline{\alpha}_{f_\mu}\|$ and with equality holding only when $h(s) = 0$.

Next we consider the case $\mu(\sigma_l) = \sigma_l$ with $l > k$. Then

$$\begin{aligned} f(s) &= \Pi_{i=1}^{k}(s+\sigma_i)\Pi_{j=k+1,i\neq l}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2][(s+\sigma_l)^2+\omega_l^2] \\ &= (s+\sigma_l)s^{n-1}+\triangle_1 f(s)+\sigma_l s^{n-1}+\sigma_l^2 s^{n-2} \\ &= f_q(s)+\triangle_1 f(s)+\sigma_l s^{n-1}+\sigma_l^2 s^{n-2} \end{aligned} \tag{5.10}$$

Also from a direct calculation $\triangle_1 f(s)+\sigma_l s^{n-1}+\sigma_l^2 s^{n-2}$ is of order at most $n-1$ and all the coefficients non-negative. Following Lemma (5.1), $\|\underline{\alpha}_f\| \geq \|\underline{\alpha}_{f_q}\|$.

(ii) Let $\nu(f) = \omega_l$ then

$$\begin{aligned} f(s) &= \Pi_{i=1}^{k}(s+\sigma_i)\Pi_{j=k+1,i\neq l}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2][(s+\sigma_l)^2+\omega_l^2] \\ &= (s^2+\omega_l^2)s^{n-2}+\triangle f(s) \\ &= f_q(s)+\triangle f(s) \end{aligned} \tag{5.11}$$

where $\triangle f(s)$ is of order at most $n-1$ and all the coefficients are non-negative. Again from Lemma (5.1) $\|\underline{\alpha}_f\| \geq \|\underline{\alpha}_{f_q}\|$.

□

In what follows, we are going to consider the case of $R_0^{+,\gamma}[s], I_0^{+,\gamma}[s]$ sets of polynomials, that is those polynomials of $P^+[s]$ which have all their roots on the negative real axis and the imaginary axis respectively and have norm less than or equal to $\gamma > 0$. The sets $R_0^{+,\gamma}[s], I_0^{+,\gamma}[s]$ are defined as: $R_0^{+,\gamma}[s]$ is the set of polynomials $f(s) \in P^{+,\gamma}[s]$ with all the roots on the real axis, while $I_0^{+,\gamma}[s]$ is defined to be the polynomials $f(s) \in P^{+,\gamma}[s]$ with all the roots on the imaginary axis.

Proposition 5.1 *Let $\gamma > 0$ and $\psi(s) = s^{n-1}(s + z_h) \in R_0^{+,\gamma}[s], \phi(s) = s^{n-2}(s^2 + \omega_h^2) \in I_0^{+,\gamma}[s]$ with $\|\underline{\alpha}_\psi\| = \gamma, \|\underline{\alpha}_\phi\| = \gamma$. The following properties hold true:*
(i) If $f(s) \in R_0^{+,\gamma}[s]$, then $\mu(f) \leq z_h$;
(ii) If $f(s) \in I_0^{+,\gamma}[s]$, then $\nu(f) \leq \omega_h$.

Proof:

(i) If we assume that there is a polynomial $f(s)$ in $R_0^{+,\gamma}[s]$ which satisfies $\mu(f) > z_h$, then $\triangle(f) = \mu(f) - z_h = \lambda_l - z_h$ with the assumption that $\lambda_l = \mu(f)$, and

$$\begin{aligned} f(s) &= \Pi_{i=1}^n (s + \lambda_i) = (s + \lambda_l)\Pi_{i=1, i\neq l}^n (s + \lambda_i) \\ &= s^{n-1}(s + z_h) + \triangle f(s) \\ &= \psi(s) + \triangle f(s) \end{aligned} \tag{5.12}$$

and with direct calculation, $\triangle f(s)$ will be of order at most $n - 1$ and the coefficients nonnegative. So following Lemma (5.1), $\|\underline{\alpha}_f\| \geq \|\underline{\alpha}_\psi\|$.

(ii) The proof follows along similar lines.

□

So z_h and ω_h define the maximal real roots and the maximal imaginary roots for the whole set of polynomials which have all roots either on the real axis or on the imaginary axis. In the general case, we prove that a necessary region for the root distribution is given by the rectangular Φ_γ defined below.

Definition 5.1 *: For a given $\gamma > 0$ we define the following regions of C:*

$$\Phi_\gamma \equiv \left\{ s = \sigma \pm j\omega \in C^+ : -z_h \leq \sigma \leq 0, |\omega| \leq \omega_h \right\} \tag{5.13}$$

$$\Psi_1^+ \equiv \left\{ s = \sigma \pm j\omega \in C^+ : \sigma < -z_h \right\} \tag{5.14}$$

$$\Psi_2^+ \equiv \{ s = \sigma \pm j\omega \in C^+ : -z_h \leq \sigma \leq 0, |\omega| > \omega_h \} \tag{5.15}$$

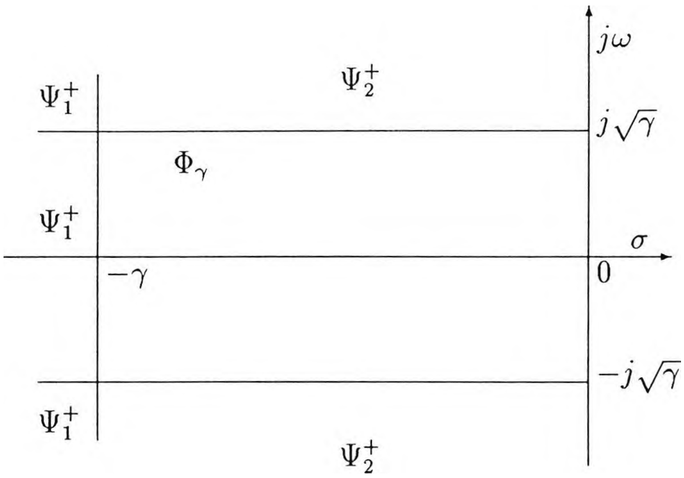


Figure 5.1: Necessary Region for the γ-Norm Bounded Stable Polynomials

where z_h, ω_h are defined by the polynomials:

$$\psi_\gamma(s) = s^{n-1}(s + z_h) \in P^{+,\gamma}[s], \quad z_h = \gamma \tag{5.16}$$

$$\phi_\gamma(s) = s^{n-2}(s + \omega_h^2) \in P^{+,\gamma}[s], \quad \omega_h = \sqrt{\gamma} \tag{5.17}$$

and the regions are shown is Figure (5.1).

□

Then we have the following result.

Theorem 5.1 *Let $\gamma > 0$ be given and $z_h = \gamma, \omega_h = \sqrt{\gamma}$, define a region in the left hand side of the complex plane as in Figure (5.1), then for all $f(s) \in P^{+,\gamma}[s]$ all roots of $f(s)$ are contained in the region Φ_γ.*

Proof:

There are two cases to be considered, either $f(s)$ has real roots (or complex conjugate roots) in Ψ_1^+, or complex conjugate roots in Ψ_2^+.

(1) Assume $f(s) \in P^{+,\gamma}[s]$, and there are k real roots and the rest are complex conjugate root pairs of the type $s = -\sigma \pm j\omega$, or

$$\Lambda_f = \{-\sigma_1, ..., -\sigma_k, -\sigma_{k+1} \pm j\omega_{k+1}, ..., -\sigma_{(n-k)/2} \pm j\omega_{(n-k)/2}\}$$

and with at least one root either (a) $s = -\sigma_l$ or (b) $s = -\sigma_l \pm j\omega_l$ contained in Ψ_2^+, where Ψ_2^+ is as defined in Figure (5.1).

(a) If $s = -\sigma_l$, then

$$\begin{aligned} f(s) &= \Pi_{i=1}^{k}(s+\sigma_i)\Pi_{j=k+1}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2] \quad (5.18) \\ &= (s+\sigma_l)s^{n-1}+\triangle f(s) \\ &= f_{\sigma_l}(s)+\triangle f(s) \end{aligned}$$

So by direct calculation $\triangle f(s)$ is of order at most $n-1$ and all its coefficients non-negative. So by Lemma (5.1), $\|\underline{\alpha}_f\| \geq \|\underline{\alpha}_{f_{\sigma_l}}\|$. As $s = -\sigma \in \Psi_1^+$, so $\triangle\sigma_l = \sigma_l - z_h > 0, \sigma_l = z_h + \triangle\sigma_l$ and $f_\sigma(s) = s^{n-1}(s+\sigma_l) = s^{n-1}(s+z_h+\triangle\sigma_l)$

$$\|\underline{\alpha}_{f_\sigma}\| = \sqrt{(z_h+\triangle\sigma_1)^2} > \sqrt{z_h^2} \quad (5.19)$$

so $\|\underline{\alpha}_f\| > \gamma$, which is a contradiction to the assumption.

(b) $s = -\sigma_l \pm j\omega_l \in \Psi_1^+$ then:

$$f(s) = \Pi_{i=1}^{k}(s+\sigma_i)\{\Pi_{j=k+1,i\neq l}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2]\}[(s+\sigma_l)^2+\omega_l^2] \quad (5.20)$$

Then by a similar argument as in (a), it follows that $\|\underline{\alpha}_f\| > \gamma$, which is also a contradiction to the $f(s) \in P^{+,\gamma}[s]$ assumption.

(2) Assume $f(s) \in P^{+,\gamma}[s], \Lambda = \{-\sigma_1, \ldots, -\sigma_k, -\sigma_{k+1} \pm j\omega_{k+1}, \ldots, -\sigma_{(n-k)/2} \pm j\omega_{(n-k)/2}\}$, there exists at least one pair of roots $\sigma_l \pm j\omega_l \in \Psi_2^+$, then

$$f(s) = \Pi_{i=1}^{k}(s+\sigma_i)\Pi_{j=k+1,i\neq l}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2][(s+\sigma_l)^2+\omega_l^2] \quad (5.21)$$

By Lemma (5.2), we have $\|\underline{\alpha}_f\| \geq \|\underline{\alpha}_{f_\omega}\|$, where $f_\omega(s) = s^{n-2}(s^2+\omega_l^2)$. From the assumption that $s = \sigma_l \pm j\omega_l \in \Psi_2^+$, so we have $\omega_l - \omega_h > 0$. Set $\triangle\omega_l = \omega_l - \omega_h$, and $\omega_l = \omega_h + \triangle\omega_l$, $f_\omega(s) = [s^2 + (\omega_h+\triangle\omega_l)^2]s^{n-2}$. The norm of the coefficient vector of the polynomial $f_\omega(s)$ is $\|\underline{\alpha}_\omega\| = \sqrt{(\omega_h+\triangle\omega_l)^4} = (\omega_h+\triangle\omega_l)^2 > \omega_h^2 = \gamma$. So the strict inequality follows i.e. $\|\underline{\alpha}_f\| > \gamma$ which is a contradiction to the assumption. Thus in both cases, if $f(s) \in P^{+,\gamma}[s]$, then $\Lambda_f \in \Phi_\gamma$. If $f(s)$ has roots outside Φ_γ, then the norm of the polynomial will always be greater than γ.

□

The region Φ_γ defined in Figure (5.1) gives a necessary condition for the location of roots of all bounded by γ stable polynomials. However, this region does not give the minimal necessary region. So we proceed to establish the minimal necessary region for the norm bounded polynomials $f(s) \in P^{+,\gamma}[s]$. Within the region Φ_γ, we define subsets $\Gamma_\gamma^+, \Phi_\gamma^+$ and Θ_γ^+:

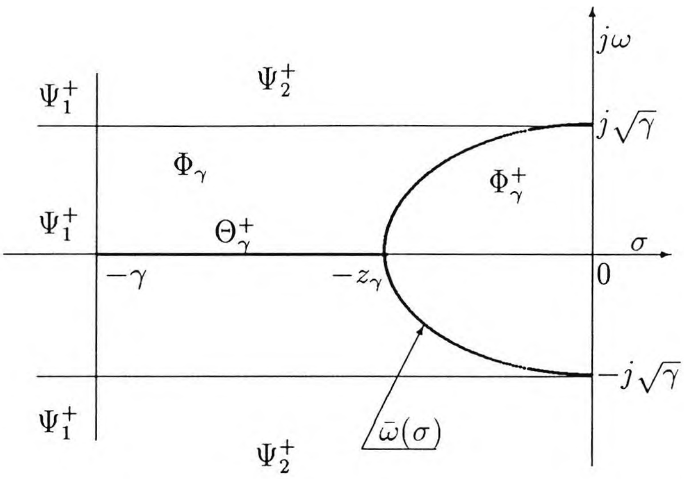


Figure 5.2: Γ-Prime Region of all Polynomials $f(s) \in P^+[s]$

Definition 5.2 : *Let $\gamma > 0$, then*

$$\Phi_\gamma^+ \equiv \left\{ s = \sigma \pm j\omega \in C^+ : -z_\gamma \le \sigma \le 0, |\omega| \le \bar{\omega}(\sigma) = \sqrt{-\sigma^2 + \sqrt{\gamma^2 - 4\sigma^2}} \right\} \quad (5.22)$$

$$\Theta_\gamma^+ \equiv \{ s = \sigma \pm j\omega \in C^+ : -z_h \le \sigma < -z_\gamma, \omega = 0 \} \quad (5.23)$$

$$\Gamma_\gamma^+ \equiv \Phi_\gamma^+ \cup \Theta_\gamma^+ \quad (5.24)$$

where $z_\gamma = \sqrt{-2 + \sqrt{4 + \gamma^2}}$, and the regions are shown in Figure (5.2). Γ_γ^+ will be referred to as the Γ-Prime Region of the polynomials $f(s) \in P^+[s]$ and $\delta\Gamma_\gamma^+$ denote the boundary of Γ_γ^+.

□

The importance of the above regions is described by the following result.

Theorem 5.2 *Let $\gamma > 0$ be the given norm bound. Γ_γ^+ be the Γ-Prime region of C^+ as defined in Definition (5.1). For all $f(s) \in P^{+,\gamma}[s]$ with $\Lambda_f = \{-\lambda_i, i = 1, 2, ..., n\}$ root set, the following properties hold true:*

(i) *For any $-\lambda_i \in \Lambda_f$, $-\lambda_i \in \Gamma_\gamma^+$.*

(ii) *If $f(s) \in P^{+,\gamma}[s]$ has roots on $\delta\Gamma_\gamma^+$, then either has one simple real root in Θ_γ^+, or a pair of roots on $\delta\Phi_\gamma^+$.*

(iii) *There exists no proper subset $p\Gamma_\gamma^+ \in \Gamma_\gamma^+$ with at least one point of $\delta\Gamma_\gamma^+$ excluded from $p\Gamma_\gamma^+$, that contains all zeros of the polynomials in $P^{+,\gamma}[s]$.*

Proof:

(i) By Theorem 5.4 for any $f(s) \in P^{+,\gamma}[s]$ we have that the real part and the imaginary part of the roots satisfy $0 \leq \sigma_i \leq \gamma$ and $0 \leq |\omega_i| \leq \sqrt{\gamma}$, for every $i = 1, 2, ..., (n-k)/2$.

(a) We assume that a pair of roots $-\sigma_l \pm j\omega_l$ with $z_\gamma < \sigma_l \leq \gamma, \omega_l \neq 0$, then $f(s)$ can be written as:

$$\begin{aligned} f(s) &= \Pi_{i=1}^{k}(s+\sigma_i)\Pi_{j=k+1,j\neq l}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2][(s+\sigma_l)^2+\omega_l^2] \\ &= s^{n-2}[(s+\sigma_l)^2+\omega_l^2]+\triangle f(s) \\ &= f_{\sigma_l,\omega_l}(s)+\triangle f(s) \end{aligned} \quad (5.25)$$

where $f_{\sigma_l,\omega_l} = s^{n-2}[(s+\sigma_l)^2+\omega_l^2]$ and $\triangle f(s)$ is of order at most $n-1$ and all its coefficients are non-negative. Then by Lemma (5.1)

$$\|\underline{\alpha}_f\| \geq \|\underline{\alpha}_{f_{\sigma_l,\omega_l}}\| \quad (5.26)$$

Furthermore,

$$\|\underline{\alpha}_{f_{\sigma,\omega}}\| = \sqrt{(2\sigma_l)^2+(\sigma_l^2+\omega_l^2)^2} > \sqrt{4\sigma_l^2+\sigma_l^4} > \sqrt{4z_\gamma^4} = \gamma \quad (5.27)$$

So $\|\underline{\alpha}_f\| > \gamma$ and this leads to a contradiction. Thus, there exists no polynomial of $f(s) \in P^{+,\gamma}[s]$ with a pair of complex conjugate roots having real part in the interval $[-\gamma, -z_\gamma)$.

(b) Next we assume $f(s) \in P^{+,\gamma}[s]$ and there is at least a pair of complex conjugate roots $-\sigma_l \pm j\omega_l$ such that $0 \leq \sigma_l \leq z_\gamma$ with

$$|\omega_l| > \bar{\omega}(\sigma_l) = \sqrt{-\sigma_l^2+\sqrt{\gamma^2-4\sigma_l}}$$

then

$$\begin{aligned} f(s) &= \Pi_{i=1}^{k}(s+\sigma_i)\Pi_{j=k+1,j\neq l}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2][(s+\sigma_l)^2+\omega_l^2] \\ &= s^{n-2}[(s+\sigma_l)^2+\omega_l^2]+\triangle f(s) = f_{\sigma_l,\omega_l}(s)+\triangle f(s) \end{aligned} \quad (5.28)$$

where $f_{\sigma_l,\omega_l} = s^{n-2}[(s+\sigma_l)^2+\omega_l^2]$ and $\triangle f(s)$ is of order at most $n-1$ and all its coefficients non-negative, so by Lemma (5.1) $\|\underline{\alpha}_f\| \geq \|\underline{\alpha}_{f_{\sigma_l,\omega_l}}\|$. As from the assumption that

$$|\omega_l| > \bar{\omega}(\sigma_l) = \sqrt{-\sigma_l^2+\sqrt{\gamma^2-4\sigma_l}}.$$

Denote $\triangle\omega_l = |\omega_l| - \bar{\omega}(\sigma_l) \geq 0$ and $|\omega_l| = \triangle\omega_l + \bar{\omega}(\sigma_l)$ then

$$\begin{aligned} f(s) &= s^{n-2}[(s+\sigma_l)^2+\bar{\omega}^2(\sigma_l)+2\bar{\omega}(\sigma_l)\triangle\omega_l+\triangle\omega_l^2] \\ &= s^{n-2}[(s+\sigma_l)^2+\bar{\omega}^2(\sigma_l)]+(2\bar{\omega}\triangle\omega_l+\triangle\omega_l^2)s^{n-2} \end{aligned} \quad (5.29)$$

Again from Lemma (5.1), $\|\underline{\alpha}_f\| > \gamma$ because $2\bar{\omega}\Delta\omega_l + \Delta\omega_l^2 > 0$, and this leads to a contradiction. So for any $f(s) \in P^{+,\gamma}[s]$, it is necessary to have all its roots contained in Γ_γ^+.

(ii) Assume that $f(s) \in P^{+,\gamma}[s]$ has at least two roots $-\sigma_1, -\sigma_2 \in \Theta^{+,\gamma}$, and $z_\gamma < \sigma_1, \sigma_2 \leq \gamma$. Then

$$\begin{aligned} f(s) &= \Pi_{i=1}^k (s+\sigma_i)\Pi_{j=k+1}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2] \\ &= s^{n-2}(s+\sigma_1)(s+\sigma_2)+\Delta f(s) \\ &= f_{\sigma_1,\sigma_2}(s)+\Delta f(s) \end{aligned} \tag{5.30}$$

where $f_{\sigma_1,\sigma_2}(s) = s^{n-2}(s+\sigma_1)(s+\sigma_2)$ and $\Delta f(s)$ is of order at most $n-1$ and having non-negative coefficients. As $\sigma_1, \sigma_2 > z_\gamma, \sigma_1 = \Delta\sigma_1 + z_\gamma, \sigma_2 = \Delta\sigma_2 + z_\gamma$, with $\Delta\sigma_1 > 0, \Delta\sigma_2 > 0, f_{\sigma_1,\sigma_2}(s)$ can be rewritten as

$$\begin{aligned} f_{\sigma_1,\sigma_2}(s) &= s^{n-2}(s+\Delta\sigma_1+z_\gamma)(s+\Delta\sigma_2+z_\gamma) \\ &= s^{n-2}(s+z_\gamma)^2+(\Delta\sigma_1+\Delta\sigma_2)s^{n-1}+\Delta\sigma_1\Delta\sigma_2 s^{n-2} \end{aligned} \tag{5.31}$$

Because the polynomial $s^{n-2}(s+z_\gamma)^2$ has norm γ it follows that $\|\underline{\alpha}_f\| > \gamma$ which is a contradiction.

Now assume that $f(s) \in P^{+,\gamma}[s]$ and apart from one pair of complex conjugate roots $\sigma_l \pm j\omega_l$ on $\delta\Gamma_\gamma^+$ of $0 \leq \sigma_l \leq z_\gamma, \omega_l = \sqrt{-\sigma_l^2 + \sqrt{\gamma^2 - 4\sigma_l^2}}$, there are at least (a) either one non-zero real root σ', or (b) a pair of non-zero complex conjugate roots $-\sigma' \pm j\omega'$. Then

(a)

$$\begin{aligned} f(s) &= \Pi_{i=1}^k (s+\sigma_i)\Pi_{j=k+1}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2] \\ &= s^{n-1}(s+\sigma')+\Delta_1 f(s) \\ &= f_{\sigma'}(s)+\Delta_1 f(s); \end{aligned} \tag{5.32}$$

(b)

$$\begin{aligned} f(s) &= \Pi_{i=1}^k (s+\sigma_i)\Pi_{j=k+1}^{(n-k)/2}[(s+\sigma_j)^2+\omega_j^2] \\ &= s^{n-2}[(s+\sigma')^2+\omega_l'^2]+\Delta_2 f(s) \\ &= f_{\sigma',\omega'}+\Delta f_2(s) \end{aligned} \tag{5.33}$$

By direct calculation, both $\Delta_1(s)$ and $\Delta_2(s)$ will have nonnegative coefficients. Because $\|\underline{\alpha}_{f_{\sigma',\omega'}}\| = \|\underline{\alpha}_{f_{\sigma'}}\| = \gamma$, then by Lemma (5.1) we have $\|\underline{\alpha}_f\| > \gamma$, which is a contradiction to the assumption. Thus (ii) follows.

(iii) If there exists a region $p\Gamma_\gamma^+$ which excludes a point $-\sigma \pm j\omega$ on the boundary $\delta\phi_\gamma^+$, then the polynomial $\psi(s) = s^{n-2}[(s+\sigma)^2 + \omega^2]$, has a root outside $p\Gamma_\gamma^+$; however, $\phi(s) \in P^{+,\gamma}[s]$ and this leads to contradiction. Let us now assume that the point $-\sigma \in \Theta_\gamma^+$ is excluded from $p\Gamma_\gamma^+$. Consider now the polynomial

$$f(s) = s^{n-2}(s+\sigma)(s+x), x > 0 \tag{5.34}$$

If we show that there exists such an x that $||\underline{\alpha}_f|| = \gamma$, then this will prove the result for this case since then $f(s) \in P^{+,\gamma}[s]$. Note that $\underline{\alpha}_f = [\sigma + x, x, 0, ..., 0]'$ and thus $||\underline{\alpha}_f||$ implies that x in defined as the solution of the equation

$$(1+\sigma^2)x^2 + (2\sigma)x + (\sigma^2 - \gamma^2) = 0 \tag{5.35}$$

We prove next that this equation has a positive solution. Consider the function defined by

$$T(s) = (1+\sigma^2)x^2 + (2\sigma)x + \sigma^2 - \gamma^2 \tag{5.36}$$

then $T(0) = -\gamma^2 < 0, t(+\infty) = +\infty > 0$, also

$$T'(s) = 2(1+\sigma^2)x + 2\sigma > 0, x \in [0, +\infty) \tag{5.37}$$

thus the equation will always have a positive solution and polynomials $f(s) \in P^{+,\gamma}[s]$ of this type exist.

□

The results so far have established existence of the properties of the Γ-Prime region Γ_γ^+ that contains the zeros of all bounded norm polynomials of $P^{+,\gamma}[s]$. This region is the minimal one that may be defined in the sense that any proper subset of Γ_γ^+ does not contain the zeros of all polynomials of $P^{+,\gamma}[s]$. This property is established below.

Corollary 5.1 *The Γ-Prime region Γ_γ^+ is the smallest region of C^+ that contains all zeros of all polynomials of $P^{+,\gamma}[s]$.*

Proof:

What we have to show is that if $p\Gamma_\gamma^+$ is a proper subset of Γ_γ^+ that excludes at least one point of Γ_γ^+, then $p\Gamma_\gamma^+$ does not contain all zeros of all polynomials of $P^{+,\gamma}[s]$. For proper subset $p\Gamma_\gamma^+$ that excludes at least one point of the effective boundary $\delta\Gamma_\gamma^+$, the result has been proved. We shall prove the result for regions that exclude at least one internal point.

We note first that since in $p\Gamma_{\gamma}^{+}$ the effective boundary has to be included, and the point at the origin $(0,0)$ has also to be included, otherwise the polynomials $s^{n-2}[(s+\sigma)^2+\omega^2]$ where $\omega=\omega_h(\sigma)$, which have norm γ are also excluded. Consider now a region that excludes at least a pair of complex conjugate internal points, say, $-\sigma \pm j\omega$. Because $-\sigma \pm j\omega$ is a pair of internal points, on the boundary, there exists a pair of complex conjugate roots $-\sigma \pm j\Omega$, where $\Omega=\omega_h(\sigma), \Omega>\omega$. Consider now the polynomials of $P^{+}[s]$

$$f_{\omega}(s) = s^{n-2}[(s+\sigma)^2+\omega^2] \tag{5.38}$$

$$f_{\Omega}(s) = s^{n-2}[(s+\sigma)^2+\Omega^2] \tag{5.39}$$

clearly, by construction $\|\underline{\alpha}_{f_{\Omega}}\| = \gamma > \|\underline{\alpha}_{f_{\omega}}\|$. The polynomial $f_{\omega}(s)$ thus $\|\underline{\alpha}_{f_{\omega}}\| < \gamma$ and belongs to $P^{+,\gamma}[s]$. Thus, by excluding the point $-\sigma \pm j\omega$, we also exclude the $f_{\Omega}(s) \in P^{+,\gamma}[s]$ and this completes the proof.

□

The Γ_{γ}^{+} region is thus the smallest region that contains all zeros of all polynomials of $P^{+,\gamma}[s]$. Note, however, that there exist polynomials having all their roots in the Γ_{γ}^{+} region which have $\|\underline{\alpha}_f\| > 0$. Defining a subset of Γ_{γ}^{+} with the property that all polynomials with their roots inside it have a norm less than or equal to γ, is the second problem we have set to solve, and it is considered next. In fact, we are interested in the maximal subset with this property.

5.1.3.2 Maximal Sufficient Root Region of C^{+} Which Guarantees Norm Bounded Coefficient Polynomials

In this part the sufficient condition for polynomials of $P^{+,\gamma}[s]$ having all their roots inside a region to have their norm bounded by γ is investigated in the case of stable polynomials. The resulted region provides solution to Problem 2.

First we investigate the case when all the roots of the polynomials are either on the real axis or on the imaginary axis. As defined in section , $R_0^{+,\gamma}[s], I_0^{+,\gamma}[s]$ are the sets of polynomial of $P^{+}[s]$ having all their roots either on the negative real axis or on the imaginary axis and their norms equal to γ. Further, we define

$$\phi(s) = (s+z_l)^n \in R_0^{+,\gamma}[s] \tag{5.40}$$

$$\psi(s) = s^{\delta}(s^2+\omega_l^2)^{(n-\delta)/2} \in I_0^{+,\gamma}[s] \tag{5.41}$$

where $\delta=0$ when n is even $\delta=1$ when n is odd. Set $\nu=(n-\delta)/2$ and z_l, ω_l are defined as the positive roots of the following equations:

$$(C_1^n)^2 z_l^2 + (C_2^n) z_l^4 + \ldots + (C_n^n)^2 z_l^{2n} = \gamma^2 \tag{5.42}$$

$$(C_1^{\nu})^2\omega_l^4 + (C_2^{\nu})^2\omega_l^8 + \ldots + (C_{\nu}^{\nu})^2\omega_l^{2\nu} = \gamma^2 \tag{5.43}$$

where $C_i^n = n(n-1)(n-2)\cdots(n-i)/i!$ and $i!$ is the factorial of i.

There is a unique positive solution to each of the above equations. The uniqueness of the solution is established below.

Corollary 5.2 *Both equations (5.42) and (5.43) always have a unique positive solution.*

Proof:

First we prove that equation (5.42) has a unique positive solution, then the uniqueness of the solution of equation (5.43) follows in a similar fashion. Consider

$$E(z) = (C_1^n)^2 z^2 + (C_2^n)^2 z^4 + \ldots + (C_n^n)^2 z^{2n} - \gamma^2 \tag{5.44}$$

then

$$E'(z) = 2(C_1^n)^2 z + 4(C_2^n)^2 z^3 + \ldots + 2n(C_n^n)^2 z^{2n-1} \tag{5.45}$$

where $E'(s)$ is the derivative of $E(s)$ with respect to s. So when $z > 0$,$E'(s) > 0$ which means that the function $E'(s)$ is monotonically increasing in $[0, +\infty)$. So a real positive z_l always exists and is uniquely determined by equation (5.42).

□

The sufficient condition when all the roots of the polynomials are on the real axis or on the imaginary axis is derived below.

Corollary 5.3 *Let $\gamma > 0$ and z_l, ω_l defined by equations (5.42) and (5.43), where $\delta = 0$ when n is even and $\delta = 1$ when n is odd. On both the axes, we define the two sections*

$$\Xi_\gamma^R \equiv \{s = \sigma \pm j\omega : 0 \geq \sigma \geq -z_l, \omega = 0\}; \Xi_\gamma^I \equiv \{s = \sigma \pm j\omega : |\omega| \leq \omega_l\}. \tag{5.46}$$

Then Ξ_γ^R and Ξ_γ^I define the maximal sufficient regions on either the real axis or the imaginary axis, i.e.

(1) $\forall f(s) \in P^+[s]$, *if* $\Lambda_f \in \Xi_\gamma^R$, *then* $f(s) \in R_0^{+,\gamma}[s]$;

(2) $\forall f(s) \in P^+[s]$, *if* $\Lambda_f \in \Xi_\gamma^I$, *then* $f(s) \in I_0^{+,\gamma}[s]$

Proof:

(i) First we prove the real roots case. Let $\Lambda_f = \{-\lambda_i, i = 1, 2, ..., n\} \in \Xi_\gamma^R$ be the root set. As from the assumption $\triangle\lambda_i = z_l - \lambda_i$. So $z_l = \lambda_i + \triangle\lambda_i$ and

$$\begin{aligned} \phi(s) &= (s + \lambda_1 + \triangle\lambda_1)...(s + \lambda_n + \triangle\lambda_n) \\ &= f(s) + \triangle f(s) \end{aligned} \tag{5.47}$$

where $\triangle f(s)$ is of order at most $n-1$ and all its coefficients non-negative. So following Lemma (5.1) we have $\|\underline{\alpha}_f\| < \|\underline{\alpha}_\phi\| = \gamma$. This establishes the result that all the polynomials satisfying the condition will have the norm bounded by γ.

Next we establish the maximality of the region. Assume that there exists $z_m > z_l$ such that the region $[-z_m, 0]$ also gives the sufficient condition. But if this were the case then we can put all the zeros at $z = -z_m$ and construct a polynomial as

$$f_m(s) = (s + z_m)^n = (s + z_l + \triangle z_l)^n \tag{5.48}$$

where $\triangle z_l = z_m - z_l > 0$. So

$$f_m(s) = (s + z_l)^n + \triangle f(s) \tag{5.49}$$

where $\triangle f(s)$ is of order at most $n-1$ and all its coefficients non-negative. By Lemma (5.1) $\|\underline{\alpha}_{f_m}\| > \|\underline{\alpha}_\phi\|$. So the maximality follows.

(ii) The case when all the roots of the polynomials are on the imaginary axis can be established in much a similar fashion.

□

When the roots of a polynomial are not confined to the axes, the sufficient root region can also be established and is illustrated in Figure (5.3). We consider the case when the order of the polynomial is even. A polynomial $f(s) \in P^+[s]$ which has $n/2$ pairs of roots at $s = -\sigma \pm j\omega$ can be expressed as:

$$\begin{aligned} f(s) = [(s &+ \sigma)^2 + \omega^2]^{n/2} \\ = s^n &+ 2\sigma C_1^{n/2} s^{n-1} \\ &+ [(\sigma^2 + \omega^2) C_1^{n/2} + 4\sigma^2 C_2^{n/2}] s^{n-2} \\ &+ [4(\sigma^3 + \sigma\omega^2) C_2^{n/2} + 8\sigma^3 C_3^{n/2}] s^{n-3} \\ &+ [(\sigma^2 + \omega^2)^2 C_2^{n/2} + 12\sigma^2(\sigma^2 + \omega^2) C_3^{n/2} + 16\sigma^4 C_4^{n/2}] s^{n-4} + \dots \\ &+ (\sigma^n + \omega^2 C_1^{n/2} \sigma^{n-2} + \omega^4 C_2^{n/2} \sigma^{n-4} + \dots + \omega^n) \end{aligned} \tag{5.50}$$

Let $\|\underline{\alpha}_f\| = \gamma$. Then

$$\begin{aligned}\gamma^2 &= (2\sigma C_1^{n/2})^2 + [(\sigma^2+\omega^2)C_1^{n/2} + 4\sigma^2 C_2^{n/2}]^2 + \\ &+[4(\sigma^3+\sigma\omega^2)C_2^{n/2} + 8\sigma^3 C_3^{n/2}]^2 + \ldots \\ &+(\sigma^n + \omega^2 C_1^{n/2}\sigma^{n-2} + \omega^4 C_2^{n/2}\sigma^{n-4} + \ldots + \omega^n)^2\end{aligned} \quad (5.51)$$

Define $D(\sigma,\omega) = (2\sigma C_1^{n/2})^2 + [(\sigma^2+\omega^2)C_1^{n/2} + 4\sigma^2 C_2^{n/2}]^2 + [4(\sigma^3+\sigma\omega^2)C_2^{n/2} + 8\sigma^3 C_3^{n/2}]^2 + \cdots + (\sigma^n + \omega^2 C_1^{n/2}\sigma^{n-2} + \omega^4 C_2^{n/2}\sigma^{n-4} + \cdots + \omega^n)^2 - \gamma^2$. When $\omega = 0$, there exists a unique real positive solution $\underline{\sigma}$ to the equation $D(\sigma, 0) = 0$. Further, for every $0 \leq \sigma_0 \leq \underline{\sigma}$, there exists a unique real positive solution $\underline{\omega}(\sigma_0)$ to $D(\sigma_0,\omega) = 0$. The uniqueness of the solution to equation $D(\sigma_0,\omega) = 0$ is established by the following two inequalities:

$$\begin{aligned}D(\sigma_0, 0) &= (2\sigma_0 C_1^{n/2})^2 + (\sigma_0^2 C_1^{n/2} + 4\sigma_0 C_2^{n/2})^2 + \ldots + \sigma_0^n - \gamma^2 \quad (5.52)\\ &\leq (2\underline{\sigma} C_1^{n/2})^2 + (\underline{\sigma}^2 C_1^{n/2} + 4\underline{\sigma} C_2^{n/2})^2 + \ldots + \underline{\sigma}^n - \gamma^2 \\ &= 0 \\ \frac{dD(\sigma_0,\omega)}{d\omega} &= 2(2C_1^{n/2}\omega)[(\sigma_0^2+\omega^2)C_1^{n/2} + 4\sigma_0^2 C_2^{n/2}] \quad (5.53)\\ &+2(8\sigma C_2^{n/2}\omega)[4(\sigma^3+\sigma\omega^2)C_2^{n/2} + 8\sigma^3 C_3^{n/2}] + \ldots \\ &+2(2\sigma^{n-2}C_1^{n/2}\omega + 4\sigma^{n-4}C_2^{n/2}\omega^3 + \ldots + n\omega^{n-1}) \times \\ &(\sigma^n + \omega^2 C_1^{n/2}\sigma^{n-2} + \omega^4 C_2^{n/2}\sigma^{n-4} + \ldots + \omega^n) \\ &> 0\end{aligned}$$

for all $\omega \in [0, +\infty)$.

□

As shown in Figure (5.3), for $s \in \Gamma_\gamma^+$, we can define a subregion Ξ_γ^+ as:

Definition 5.3 : *For all $\gamma > 0$ we define:*

$\Xi_\gamma^+ \equiv \{s = \sigma \pm j\omega \in C^+ : |\omega| \leq \underline{\omega}(\sigma), -\underline{\sigma} \leq \sigma \leq 0\}.$

□

In the case when the order of the polynomials $f(s) \in P^+[s]$ is odd, we study the following polynomial which has $(n-1)/2$ pairs of complex conjugate roots at $s = -\sigma \pm j\omega$ and a single real root at $s = -\sigma$, or $f(s) = (s+\sigma)[(s+\sigma)^2 + \omega^2]^{(n-1)/2}$. Following a similar argument, we can also establish the region $\Xi_\gamma^+ \equiv \{s = \sigma \pm j\omega \in C^+ : |\omega| \leq \underline{\omega}(\sigma), -\underline{\sigma} \leq \sigma \leq 0\}$.

In what follows, we prove that every rectangular region as shown in Figure (5.3) inside Ξ_γ^+ gives a sufficient region for polynomials to be norm bounded by γ.

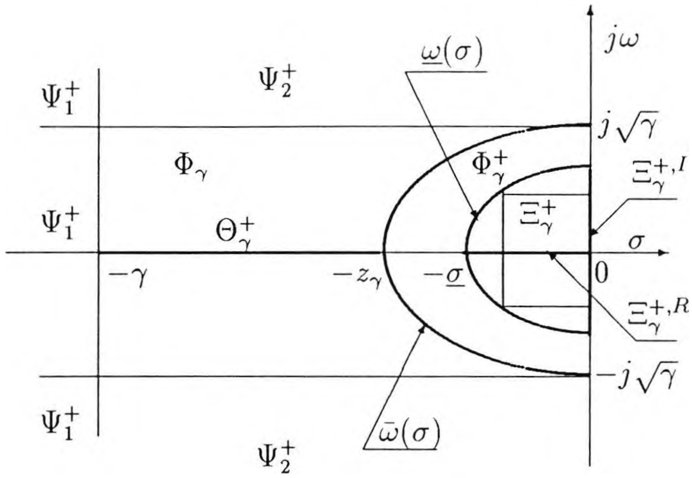


Figure 5.3: Sufficient Root Regions, Ξ_γ^+, for Stable Polynomials

Proposition 5.2 *For any $f(s) \in P^+[s]$ and a given norm $\gamma > 0$, we can obtain a boundary given by $(\sigma, \underline{\omega}(\sigma))$ as defined in equation (5.52). Then for any polynomial $f(s) \in P^+[s]$ with its roots satisfying $\mu(f) \leq \sigma$ and $\nu(f) \leq \omega_\sigma$, $\|\underline{\alpha}_f\| \leq \gamma$.*

Proof:

Assume that a polynomial $f(s) \in P^{+,\gamma}[s]$ with all its roots satisfying $\mu(f) \leq \sigma$ and $\nu(f) \leq \underline{\omega}(\sigma)$. For every $s = \sigma_i \pm j\omega_i$ set $\triangle\sigma_i = \sigma - \sigma_i, \triangle\omega_i = \underline{\omega}(\sigma) - \omega_i$, then

$$f'(s) = [(s+\sigma)^2 + \omega^2]^{n/2} = f(s) + \triangle f(s) \tag{5.54}$$

By direct calculation, $\triangle f(s)$ is of order at most $n-1$ and all its coefficients are non-negative. So by Lemma (5.1), $\|\underline{\alpha}_f\| < \|\underline{\alpha}_{f'}\| = \gamma$.

□

Remark 5.2: Having proved the above proposition, we can generate a sequence of rectangulars, each of which gives a sufficient condition for the polynomials to have their norm bounded by γ as shown in Figure (5.3). Each of the rectangulars gives the maximal rectangular region which satisfies the sufficiency requirement.

5.1.3.3 Case of Completely Unstable Polynomials

The set of completely unstable polynomials $P^-[s]$ is defined to be the set of polynomials which are in $P[s]$ and all the roots are in the right hand side of the complex plane. $P^{-,\gamma}[s]$ is a subset of $P^-[s]$ within which all the polynomials have their coefficient norm bounded by γ.

Given any polynomial $f(s) = s^n + \alpha_1 s^{n-1} + \ldots + \alpha_{n-1}s + \alpha_n \in P^+[s], \Lambda_f = \{\lambda_i, i = 1, 2, \ldots, n\}$ it corresponds to a completely unstable polynomial $f^-(s) = s^n + \alpha_1^- s^{n-1} + \ldots + \alpha_{n-1}^- s + \alpha_n^- \in P^+[s], \Lambda_f = \{-\lambda_i, i = 1, 2, \ldots, n\}$. So $f^-(s)$ is constructed from $f(s)$ by setting s to be $-s$. The two polynomials have the same coefficient vector norm as shown below. Because the coefficients are related to the roots of the polynomials as:

$$\alpha_1 = \sum_{i=1}^{n} \lambda_i, \quad \alpha_2 = \sum_{(i_1,i_2)\in Q_{2,n}} \lambda_{i_1} \cdot \lambda_{i_2}, \quad \ldots,$$

$$\alpha_k = \sum_{(i_1,\ldots i_k)\in Q_{k,n}} \lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_k}, \quad \ldots, \quad \alpha_n = \lambda_1\lambda_2\ldots\lambda_n$$

where $Q_{k,n}$ denotes the set of lexicographically ordered, strictly increasing sequences of k integers from $\{1, 2, ..., n\}$. So we have $|\alpha_i| = |\alpha_i^-|, i = 1, 2, \ldots, n$, and thus $\|\underline{\alpha}_f\| = \|\underline{\alpha}_{f^-}\|$.

Following a dual argument, the necessary and sufficient regions can be established. In fact, the minimal necessary and the maximal sufficient region are the mirror images of those defined for stable polynomials'.

Remark 5.3: For the necessary region, we define the Γ^--Prime region for all polynomials $f(s) \in P^-[s]$ as the mirror image of Γ-Prime region with respect to the imaginary axis as in Definition (5.1), i.e.

Definition 5.4 : *The Γ^--Prime region of all polynomials $f(s) \in P^-[s]$.*

Let $\gamma > 0$, then

$$\Phi_\gamma^- \equiv \left\{s = \sigma \pm j\omega \in C^- : 0 \le \sigma \le z_\gamma, |\omega| \le \bar{\omega}(\sigma) = \sqrt{-\sigma^2 + \sqrt{\gamma^2 - 4\sigma^2}}\right\} \quad (5.55)$$

$$\Theta_\gamma^- \equiv \{s = \sigma \pm j\omega \in C^- : z_{h\gamma} < \sigma \le z_h, \omega = 0\} \quad (5.56)$$

$$\Gamma_\gamma^- \equiv \Phi_\gamma^- \cup \Theta_\gamma^- \quad (5.57)$$

where $z_\gamma = \sqrt{-2 + \sqrt{4 + \gamma^2}}$.

□

For the defined Γ^--Prime region, we have

Theorem 5.3 *Let $\gamma > 0$ be the given norm bound. Γ_γ^- be the Γ^--Prime region of C^- as defined in Definition (5.3). For all $f(s) \in P^{-,\gamma}[s]$ with $\Lambda_f = \{\lambda_i, i = 1, 2, ..., n\}$ root set, the following properties hold true:*

(i) *For any* $\lambda_i \in \Lambda_f, \lambda_i \in \Gamma_\gamma^-$.

(ii) *If* $f(s) \in P^{-,\gamma}[s]$ *has roots on the edge of the* Γ^-*-Prime region,* $\delta\Gamma_\gamma^-$*, then either has one simple real root in* Θ_γ^-*, or a pair of roots on* $\delta\Phi_\gamma^-$.

(iii) *There exists no proper subset* $p\Gamma_\gamma^- \in \Gamma_\gamma^-$ *with at least one point of* $\delta\Gamma_\gamma^-$ *excluded from* $p\Gamma_\gamma^-$*, that contains all zeros of the polynomials in* $P^{-,\gamma}[s]$.

□

Remark 5.4: For the sufficient region, we define the Ξ_γ^- region for all polynomials $f(s) \in P^-[s]$ as the mirror image of Ξ_γ^+ region with respect to the imaginary axis as in Definition (5.3), i.e,

Definition 5.5 : *The* Ξ_γ^- *region of all polynomials* $f(s) \in P^-[s]$,

$$\Xi_\gamma^- \equiv \left\{ s = \sigma \pm j\omega \in C^- : |\omega| \leq \underline{\omega}(\sigma), 0 \leq \sigma \leq \underline{\sigma} \right\}$$

and for this region we have

Proposition 5.3 *For any* $f(s) \in P^-[s]$ *and a given norm* $\gamma > 0$*, we can obtain a boundary given by* $(\sigma, \underline{\omega}(\sigma))$ *as defined in equation (5.52). Then for any polynomial* $f(s) \in P^-[s]$ *with its roots satisfying* $\mu(f) \leq \sigma$ *and* $\nu(f) \leq \omega_\sigma$*, we have* $\|\underline{\alpha}_f\| \leq \gamma$.

□

5.1.4 Case of General Polynomials

In the general case, the roots of the polynomials may lie in both the left and the right hand side of the complex plane. Results concerning the distribution of the roots when the coefficients are bounded by l_2 norm can be adapted from Theorem (4.2). Because for $f(s) = s^n + \alpha_1 s^{n-1} + \alpha_2 s^{n-2} + \ldots + \alpha_{n-1} s + \alpha_n$, $\sqrt{\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_n^2} \leq \gamma$, $\alpha_i \leq \gamma$, and by Theorem (4.2), we have the following proposition.

Proposition 5.4 *If* $f(s) \in P^\gamma$*, then all the roots of the polynomials are in the circle centred at the origin with a radius* σ,

$$\sigma = 1 + \gamma.$$

□

The proof follows directly from Theorem (4.2). This result establishes an upper bound and not the minimal bound. However, in some special cases, much tighter necessary regions can be obtained. In the following we establish the minimal disk-type approximation of the necessary region for the case of third order monic polynomials; by using an analytic method, the exact boundary can also be obtained for this case.

5.1.4.1 Necessary Root Region for Third Order Monic Polynomials with Norm Bounded Coefficients

A general monic third order polynomial is of the following form $f^3(s) = s^3 + \alpha_1 s^2 + \alpha_2 s + \alpha_0$. For bounded polynomials of this type we have:

Proposition 5.5 *All the complex roots of the γ-bounded polynomials $f^3(s)$ are within the disk centred at the origin with a radius $\rho < \sqrt{\sqrt{\gamma^2+4}}$.*

Proof:

We prove the above proposition in the following way: assume that $f(s)$ has a pair of complex conjugate roots, $z = \sigma \pm j\omega$, which are outside the disk, then there exists no real root $s = a$ such that $\|\underline{\alpha}_{f^3}\| \leq \gamma$ stands.

Let $\Lambda_{f^3} = \{-a, \sigma \pm j\omega\}$ be the root set, where the complex roots are outside the disk centred at the origin with a radius $\rho = \sqrt{\sqrt{\gamma^2+4}}$, or

$$\sqrt{\sigma^2+\omega^2} > \rho = \sqrt{\sqrt{\gamma^2+4}} \tag{5.58}$$

then

$$\begin{aligned} f^3(s) &= (s+a)[(s+\sigma)^2+\omega^2] \\ &= s^3 + (a+2\sigma)s^2 + (\sigma^2+2a\sigma+\omega^2)s + a(\sigma^2+\omega^2) \end{aligned} \tag{5.59}$$

The coefficient vector norm should satisfy

$$\begin{aligned} \|\underline{\alpha}_{f^3}\| &= (a+2\sigma)^2 + (\sigma^2+2a\sigma+\omega^2))^2 + a^2(\sigma^2+\omega^2)^2 \\ &= [1+4\sigma^2+(\sigma^2+\omega^2)^2]a^2 + 4(\sigma+\sigma\omega^2+\sigma^3)a + [4\sigma^2+(\sigma^2+\omega^2)^2] \\ &\leq \gamma^2 \end{aligned} \tag{5.60}$$

so we have the following inequality

$$[1+4\sigma^2+(\sigma^2+\omega^2)^2]a^2 + 4(\sigma+\sigma\omega^2+\sigma^3)a + [4\sigma^2+(\sigma^2+\omega^2)^2-\gamma^2] \leq 0 \tag{5.61}$$

where $1+4\sigma^2+(\sigma^2+\omega^2)^2>0$. Because $\sqrt{\sigma^2+\omega^2}>\rho=\sqrt{\sqrt{\gamma^2+4}}$,

$$\begin{aligned}\triangle_a &= [4(\sigma+\sigma\omega^2+\sigma^3)]^2-4[1+4\sigma^2+(\sigma^2+\omega^2)^2][4\sigma^2+(\sigma^2+\omega^2)^2-\gamma^2] \\ &\leq 4[-12\sigma^4-16\sigma^2-4-4\omega^4] \\ &< 0 \end{aligned} \quad (5.62)$$

So we have proved that no complex roots are possible to be outside the disk defined above.

□

The previous result of a necessary region has established the existence for the case of general third order polynomials. Yet a tighter region can be established.

Proposition 5.6 *For a given coefficient vector norm bound* $\gamma>0$, *a necessary complex root region for third order polynomials is given by the disk centred at the origin with a radius* $\rho=\sqrt{\sqrt{\gamma^2+\delta}}$, *where* $\delta=\gamma\times\frac{\sqrt{\gamma^2+4}-\gamma}{2}\leq 1.0$.

Proof:

Assume that a norm bounded third order polynomial has one real root $a+j0$ and a pair of complex conjugate roots $\sigma\pm j\omega$. Further assume that the pair of complex conjugate roots are outside the disk defined which is centred at the origin with a radius $\rho=\sqrt{\sqrt{\gamma^2+\delta}}$. Then

$$\begin{aligned}f^3(s) &= (s+a)[(s+\sigma)^2+\omega^2] \quad (5.63)\\ &= s^3+(a+2\sigma)s^2+(\sigma^2+2a\sigma+\omega^2)s+a(\sigma^2+\omega^2)\end{aligned}$$

The coefficient vector norm should satisfy

$$\begin{aligned}\|\underline{\alpha}_{f^3}\| &= (a+2\sigma)^2+(\sigma^2+2a\sigma+\omega^2)^2+a^2(\sigma^2+\omega^2)^2\leq\gamma^2 \quad (5.64)\\ &= [1+4\sigma^2+(\sigma^2+\omega^2)^2]a^2+4(\sigma+\sigma\omega^2+\sigma^3)a+[4\sigma^2+(\sigma^2+\omega^2)^2] \\ &\leq \gamma^2\end{aligned}$$

For inequality (5.64) to hold, there should exist a real solution to "a". So the following inequality has to be satisfied:

$$\begin{aligned}\Delta_a &= (4\sigma+4\sigma^3+4\sigma\omega^2)^2+4(1+4\sigma^2+\sigma^4+2\sigma^2\omega^2+\omega^4)\times \\ &\quad (\gamma^2-4\sigma^2-\sigma^4-\omega^4-2\sigma^2\omega^2)\geq 0\end{aligned} \quad (5.65)$$

However, when $\sigma^2+\omega^2>\rho^2=\sqrt{\gamma^2+\delta}$, $\Delta_a/4<-\{\delta^2+(1+\gamma^2)\delta-\gamma^2+[4\delta\sigma^2+(\sqrt{\gamma^2+\delta}-4\sigma^2)^2]\}$. So if the pair of complex conjugate roots are outside the defined region, inequality (5.65) will not be true, neither will be inequality (5.64). So for every third order polynomial, all the complex conjugate roots will fall inside the defined disk if their coefficient vectors are bounded by $\gamma > 0$.

□

Remark 5.5: The results established above are much better when compared with the result stated in Proposition (5.5). For instance when $\gamma = 0.5$, then the present proposition gives a result $\sqrt{\sqrt{\gamma^2 + \delta}} \doteq 0.895$, while Proposition (5.5) gives $1 + \gamma = 1.5$. When $\gamma = 100$, then the present proposition gives a result $\sqrt{\sqrt{\gamma^2 + 4}} \doteq 10, \sqrt{\sqrt{\gamma^2 + \delta}} \doteq 10$, while Proposition (5.5) gives $1 + \gamma = 101$, which is more than 10 times bigger. The larger the given bound, the bigger the difference.

□

Also, the results are better when compared with those obtained by Boses and Luther [Boe. & Lut.,1]. The result adapted from their Theorem 1 is as follows:

Theorem 5.4 *For a given* $\gamma, \|\underline{\alpha}_f\| \leq \gamma$, *then all zeros of* $f(s) = s^n + a_1 s^{n-1} + \ldots + a_{n-1}s + a_n$ *lie in the unit disk* $|z| < R$ *where*

$$R := \begin{cases} \left\{\frac{\gamma[1-n\gamma]}{[1-(n\gamma)^{1/n}]}\right\}^{1/n}, \gamma \leq 1/n \\ \min\left\{(1+\gamma)\left(1 - \frac{\gamma}{[(1+\gamma)^{1/n} - n\gamma]}\right), 1 + 2n(n\gamma - 1)(n+1)\right\}, \quad \gamma \geq 1/n \end{cases}$$

□

So when $n = 3, \gamma = 0.2$, the present result gives $\sqrt{\sqrt{\gamma^2 + \delta}} \doteq 0.68$ while Theorem (5.4) gives $R = 0.7995$. For $n = 3, \gamma = 0.5, \sqrt{\sqrt{\gamma^2 + \delta}} \doteq 0.895$ while Theorem (5.4) gives $R = 1.75$. For both $\gamma \leq 1/n$ and $\gamma \geq 1/n$, the result obtained here are much better. If we compare this result with Proposition (5.4), we find that Theorem (5.4) usually gives a tighter region than when γ is small and the opposite is true when γ is big for higher order polynomials. For instance when $n = 4$ and $\gamma = 0.2$, Theorem (5.4) gives $R = 0.9266$ while Proposition (5.4) gives $1 + \gamma = 1.2$. However, when $n = 4$ and $\gamma = 5$, Theorem (5.4) gives $R = 7.627$ while Proposition (5.4) gives $1 + \gamma = 6$.

5.1.4.2 Sufficient Polynomial Root Region

The sufficient root region for the general polynomial case is shown as in Figure (5.4). The region consists of the sufficient region when all the polynomials are stable and the sufficient region when all the polynomials are completely unstable. The sufficiency of the region can be established as follows:

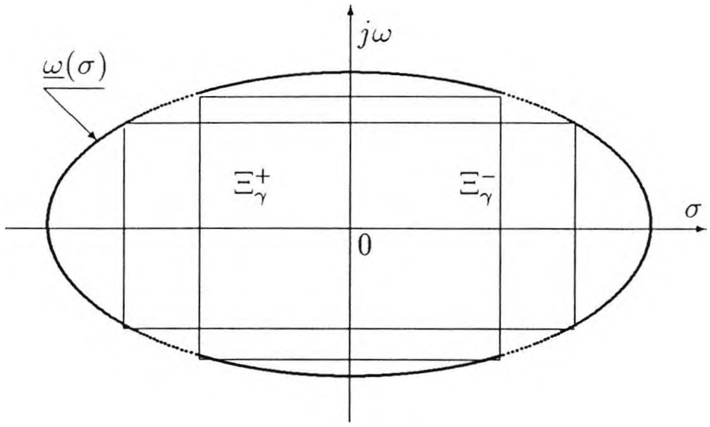


Figure 5.4: Sufficient Root Regions for General Polynomials

Proposition 5.7 *Let $\Xi = \Xi_\gamma^+ \cup \Xi_\gamma^-$ be a rectangular with its edges on $\underline{\omega}(\sigma)$ as shown in Figure (5.1). For all $f(s) \in P[s]$ such that $\Lambda_f \in \Xi$ we have that $\|\underline{\alpha}_f\| \leq \gamma$.*

Proof:

First we define another polynomial $f^+(s)$ which takes all the roots of $f(s)$ which are in the left hand side of the complex plane. For the roots in the right hand side of the complex plane, we take the mirror images. From the assumption, all the roots of $f^+(s)$ are in the sufficient region. Then the sufficiency can be established by showing $\gamma = \|\underline{\alpha}_{f^+}\| \geq \|\underline{\alpha}_f\|$.

From the direct relation between the roots and the coefficients of the polynomials,

$$\alpha_1 = \sum_{i=1}^{n} \lambda_i, \quad \alpha_2 = \sum_{(i_1,i_2)\in Q_{2,n}} \lambda_{i_1} \cdot \lambda_{i_2}, \quad \ldots,$$

$$\alpha_k = \sum_{(i_1,\ldots i_k)\in Q_{k,n}} \lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_k}, \quad \ldots, \quad \alpha_n = \lambda_1\lambda_2\ldots\lambda_n$$

where $Q_{k,n}$ denotes the set of lexicographically ordered, strictly increasing sequences of k integers from $\{1, 2, ..., n\}$.

If we set $f^+(s) = s^n + \alpha_1^+ s^{n-1} + \ldots + \alpha_{n-1}^+ s + \alpha_n^+$, then $|\alpha_i^+| \geq |\alpha_i|$, for $i = 1, 2, ..., n$. So $\gamma \geq \|\underline{\alpha}_{f^+}\| \geq \|\underline{\alpha}_f\|$.

□

So every rectangular as shown in Figure (5.4) provides a maximal sufficient root region for the general polynomial case. The region defined by $\underline{\omega}(\sigma)$ thus describes the locus of the edges of all rectangulars Ξ for which the bounded norm property holds true.

5.2 Root distribution of the sum of two polynomials

5.2.1 Introduction

In this section, the root distribution of the sum of two polynomials are studied. The importance of the study can be highlighted by the control problems such as stability robustness and state feedback. In the study of control robustness, assume that a system has a nominal characteristic polynomial

$$f_0(s) = s^n + \alpha_1 s^{n-1} + \alpha_2 s^{n-2} + \cdots + \alpha_{n-1} s + \alpha_n$$

and the system is perturbed by parametric disturbances, then the actual system characteristic polynomial will be

$$f(s) = f_0(s) + \Delta f(s) = s^n + (\alpha_1 + \beta_1)s^{n-1} + \cdots + (\alpha_{n-1} + \beta_{n-1})s + (\alpha_n + \beta_n) \quad (5.66)$$

where $\Delta f(s) = \beta_1 s^{n-1} + \beta_2 s^{n-2} + \cdots + \beta_{n-1} s + \beta_n$ is the additive disturbance. The knowledge about the disturbance will usually be limited. Here we assume that the disturbances satisfy the following condition

$$\sqrt{\beta_1^2 + \beta_2^2 + \cdots + \beta_n^2} \leq r.$$

It is important to study the root distribution of the polynomial in equation (5.66).

The other important problem is the study of bounded gain state feedback. Given a controllable single-input single-output system

$$\dot{\underline{x}} = A\underline{x} + \underline{b}u(t) \quad (5.67)$$

and assume that it has a characteristic polynomial

$$f_o(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_{n-1} s + \alpha_n \quad (5.68)$$

When the system is under bounded state feedback with the feedback gain bounded as

$$\sqrt{k_1^2 + k_2^2 + \cdots + k_n^2} \leq \gamma_f \quad (5.69)$$

then, the closed-loop characteristic polynomial

$$f_c(s) = s^n + \beta_1 s^{n-1} + \cdots + \beta_{n-1} s + \beta_n \quad (5.70)$$

can be considered as the open loop characteristic polynomial being disturbed, i.e.

$$f_c(s) = f_o(s) + \Delta f_s = s^n + (\alpha_1 + \Delta\alpha_1)s^{n-1} + \cdots + (\alpha_{n-1} + \Delta\alpha_{n-1})s + (\alpha_n + \Delta\alpha_n) \quad (5.71)$$

where the bound on the disturbance is proportional to the bound on the feedback gain.

Equation (5.71) takes the same form as equation (5.66). In light of the results presented in the previous section, the root distributions of the polynomials with bounded coefficients can be studied. The polynomial in equation (5.66) can be considered to be the sum of two polynomials each with a known root distribution, i.e.

$$f(s) = \frac{1}{2}[f_1(s) + f_2(s)] \quad (5.72)$$

$$f_1(s) = s^n + 2\alpha_1 s^{n-1} + 2\alpha_2 s^{n-2} + \cdots + 2\alpha_{n-1} s + 2\alpha_n \quad (5.73)$$

$$f_2(s) = s^n + 2\beta_1 s^{n-1} + 2\beta_2 s^{n-2} + \cdots + 2\beta_{n-1} s + 2\beta_n \quad (5.74)$$

where $f_1(s)$ is a fixed polynomial while $f_2(s)$ has varying coefficients. So the root distribution of the original polynomial $f(s)$ is equivalent to the root distribution of the sum of the two polynomials $f_1(s)$ and $f_2(s)$.

In the literature, root distribution of the composite polynomials has been studied. Results are summarised in Marden [Mar.,1]. The most fundamental result is the continuity theorem relating the roots of a polynomial and the coefficients.

The roots of a polynomial are continuous functions of the coefficients of the polynomial.

Theorem 5.5 [Mar.,1] *Let*

$$f_o(s) = s^n + \alpha_1 s^{n-1} + \alpha_2 s^{n-2} + \cdots + \alpha_n = \Pi_{j=1}^{p}(s - s_j)^{m_j} \quad (5.75)$$

$$f(s) = s^n + (\alpha_1 + \Delta\alpha_1)s^{n-1} + (\alpha_2 + \Delta\alpha_2)s^{n-2} + \cdots + (\alpha_n + \Delta\alpha_n) \quad (5.76)$$

and let

$$0 < r_k < \min|s_k - s_j|, \qquad j = 1, 2, ..., k-1, k+1, ..., p. \quad (5.77)$$

There exists a positive number ϵ such that, if

$$|\Delta\alpha_i| \leq \epsilon \text{ for } i = 1, 2, ..., n, \quad (5.78)$$

then $f(s)$ has precisely m_k roots in the circle with centre at s_k and radius r_k.

□

A result on the root distribution of the linear combination of polynomials given in [Mar.,1] is stated below:

Theorem 5.6 [Mar.,1] *The zeros of the linear combination*

$$F(s) = \lambda_1 f_1(s) + \lambda_2 f_2(s) + \cdots + \lambda_p f_p(s) \tag{5.79}$$

where $\lambda_j \neq 0, j = 1, 2, ..., p$, *lie in the locus* Γ *of the roots of the equation,*

$$\lambda_1(s-\alpha_1)^{n_1} + \lambda_2(s-\alpha_2)^{n_2} + \cdots + \lambda_p(s-\alpha_p)^{n_p} = 0 \tag{5.80}$$

where the $\alpha_1, \alpha_2, ..., \alpha_p$ *vary independently over the regions* $C_1, C_2, ..., C_p$, *respectively, and* $f_i(s)$ *and* C_i *are defined as*

$$f_i(s) = s^{n_i} + \alpha_{i,1}s^{n_i-1} + \cdots + \alpha_{i,n_i}, \qquad i = 1, 2, ..., p, \tag{5.81}$$

The zeros of $f_i(s)$ *are assumed to lie in a circular region* C_k. *Unless otherwise specified, the region will be bounded by a circle* C_k *with centre* c_k *and a radius of* r_k.

□

The particular case $p = 2$ and $n_1 = n_2 = n$ is one in which we can readily determine Γ.

Corollary 5.4 *For the case when* $p = 2$ *and* $n_1 = n_2 = n$, *we write* $\lambda_2/\lambda_1 = -\lambda$ *and denoted by* $\omega_1, \omega_2, ..., \omega_n$ *the* n^{th} *root of* λ *with* $\omega_1 = 1$ *when* $\lambda = 1$. *The roots of the previous equation are*

$$s_k = \frac{\alpha_1 - \omega_k \alpha_2}{1 - \omega_k} \tag{5.82}$$

where $k = 1, 2, ..., n$ *when* $\lambda \neq 1$ *and* $k = 2, 3, ..., n$ *when* $\lambda = 1$. *The locus* Γ *will then consist of the ensemble of loci* Γ_k *of the* s_k *when* α_1 *and* α_2 *vary over their circular* C_1 *and* C_2 *respectively.*

□

By making use of the above theorem, we study the root distribution of the sum of the following two polynomials. The two polynomials are third order and have three roots as $\sigma, \sigma \pm j\omega$ and $-\sigma, -\sigma \pm j\omega$. So

$$\begin{aligned} f_1(s) &= (s-\sigma)[(s-\sigma)^2 + \omega^2] \\ f_2(s) &= (s+\sigma)[(s+\sigma)^2 + \omega^2] \end{aligned} \tag{5.83}$$

We can draw two circles C_1 and C_2 which encloses the roots of $f_1(s)$ and $f_2(s)$ respectively. Then the roots of the sum of the two polynomials should be contained

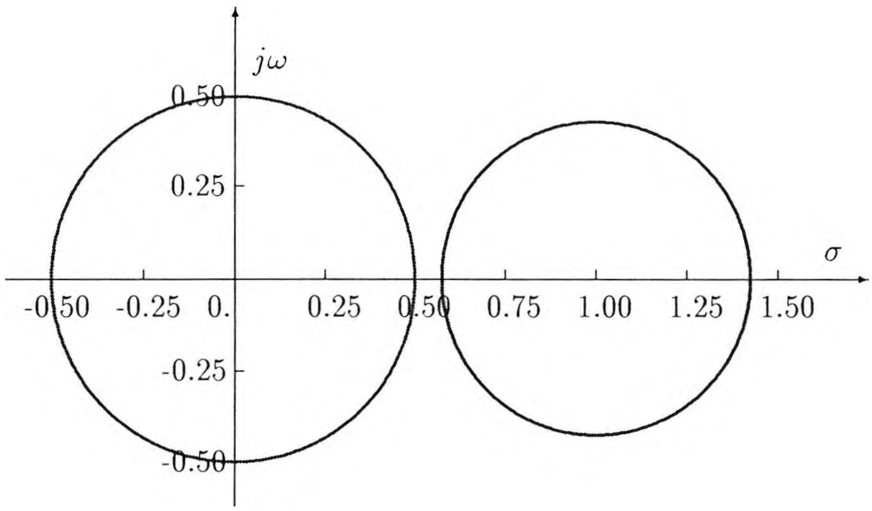


Figure 5.5: Root Distribution Region of the Sum Polynomial.

within the contours given by

$$s_1 = \frac{\alpha_1 + \alpha_2}{2} \tag{5.84}$$

$$s_2 = \frac{\alpha_1 - (\frac{1}{2} + j\frac{\sqrt{3}}{2})\alpha_2}{1 - (\frac{1}{2} + j\frac{\sqrt{3}}{2})\alpha_2} \tag{5.85}$$

$$s_3 = \frac{\alpha_1 - (\frac{1}{2} - j\frac{\sqrt{3}}{2})\alpha_2}{1 - (\frac{1}{2} - j\frac{\sqrt{3}}{2})\alpha_2} \tag{5.86}$$

where α_1 and α_2 vary over circles C_1 and C_2 respectively. The region obtained is shown in Figure (5.5). Compare this region with the actual roots of the polynomial, which are located all along the imaginary axis, it is clear that the result given by the theorem is far too conservative. The conservativeness is perhaps introduced by the necessity to have circular region C_i to enclose all the roots of $f_i(s)$.

The root distribution of a linear combination of polynomials with respect to the common roots of the all the polynomials has also been studied [Kar. Gia. & Hub.,1]. The concepts of exact zeros and almost zeros have been defined. The exact zeros will be the zeros of any linear combination of polynomials while the almost zeros act as poles attracting the roots of the linear combination of polynomials. It was shown that disk centred at the almost zeros with finite radius can be defined and the disk will include at least one root of the linear combination of polynomials. Estimates for the radius have also been given [Kar. Gia. & Hub.,1].

In this section, the root distribution of the summation of some special class of polynomials are studied. In Section 5.2.2, second order polynomials are studied and tight root distribution regions are obtained. In Section 5.2.3, the results are generalised to third order polynomials and to general polynomials in Section 5.2.4.

5.2.2 Second Order Polynomial Case

In this subsection we prove that if two second order polynomials have roots which form mirror image with respect to a line that is parallel to the imaginary axis then the sum polynomial of the two will have roots all on that line. Further we prove that for two polynomials both with real roots, then the root region of the sum polynomial can be defined by the smallest and the largest roots of the polynomials. First we study the following problem:

(a). **Two Summand polynomials having roots as mirror image with respect to a line parallel to the imaginary axis**

We consider first the following special cases:

1. Summand polynomials with equal real roots.

Proposition 5.8 *Assume that the two summand polynomials each has two equal real roots, $f_1(s) = (s-\sigma_1)^2$ and $f_2(s) = (s-\sigma_2)^2$. Then the roots of the summation polynomial will have no root except on the $\{s \in C : s = \sigma \pm j\omega, \sigma = \frac{\sigma_1+\sigma_2}{2}, \omega \geq 0\}$ line.*

Proof:

The proof is straightforward. Because

$$f_1(s) = s^2 + 2\sigma_1 s + \sigma_1^2, \qquad f_2(s) = s^2 + 2\sigma_2 s + \sigma_2^2 \tag{5.87}$$

and $f_1(s) + f_2(s) = 2[s^2 + (\sigma_1+\sigma_2)s + \frac{\sigma_1^2+\sigma_2^2}{2}]$
so $f_1(s) + f_2(s)$ has two roots at

$$s = \frac{-(\sigma_1+\sigma_2) \pm \sqrt{-(\sigma_1^2 - 2\sigma_1\sigma_2 + \sigma_2^2)}}{2} = \frac{-(\sigma_1+\sigma_2) \pm j|\sigma_1 - \sigma_2|}{2} \tag{5.88}$$

which are on the line defined by $\{s \in C : s = \sigma \pm j\omega, \sigma = \frac{\sigma_1+\sigma_2}{2}, \omega \geq 0\}$.

□

2. Summand polynomials with different roots.

Proposition 5.9 *Assume one of the polynomials has roots at $s_1 = \sigma - \gamma_1, s_2 = \sigma - \gamma_2$ and the other at $s_1 = \sigma + \gamma_1, s_2 = \sigma + \gamma_2, \gamma_1 \geq 0, \gamma_2 \geq 0$. So*

$$f(s) = f_1(s) + f_2(s) = 2s^2 - 4\sigma s + 2\gamma_1\gamma_2 \tag{5.89}$$

so the roots which are $\sigma \pm j\sqrt{\gamma_1\gamma_2}$ are on the mid-line $\sigma \pm j\omega$.

□

3. Case of summand polynomials having equal imaginary parts.

Proposition 5.10 *Assume one of the summand polynomials has roots at* $s = -\sigma_1 \pm j\omega_1$ *and the other at* $s = -\sigma_2 \pm j\omega_1$ *then*

$$f_1(s) + f_2(s) = 2s^2 + 2(\sigma_1 + \sigma_2)s + \sigma_1^2 + \sigma_2^2 + 2\omega_1^2 \tag{5.90}$$

which has roots at

$$s = \frac{-(\sigma_1 + \sigma_2) \pm \sqrt{-(\sigma_1 - \sigma_2)^2 - 4\omega_1^2}}{2} \tag{5.91}$$

So the roots are located on the mid-line drawn between the two pairs of imaginary roots.

□

So from the above result, we have the following:

Proposition 5.11 *In the case of second order polynomials, if two summand polynomials have roots on either side of a line which is parallel to the imaginary axis and the roots form mirror image with respect to this line, then the summation polynomial will have all its roots on this line.*

□

Employing the above proposition, we study the root distribution for the summation polynomial $f(s) = f_1(s) + f_2(s)$ when $f_1(s)$ has fixed roots at $\gamma_1, \gamma_2, \gamma_1 \leq \gamma_2$ while the roots of $f_2(s)$ are varying but confined in the region defined as $\delta_2 = [\gamma_3, \gamma_4], \gamma_3 \leq \gamma_4$. We have the following Corollary.

Corollary 5.5 *The summation polynomial* $f(s)$ *will have all its roots in the region defined as* $\delta = \{s \in C : s = \sigma \pm j\omega, \frac{\gamma_1+\gamma_3}{2} \leq \sigma \leq \frac{\gamma_2+\gamma_4}{2}\}$.

Proof:

The sum of the two polynomials is

$$f_1(s) + f_2(s) = 2[s^2 + \frac{(\gamma_1 + \gamma_2) + (\gamma_3 + \gamma_4)}{2}s + \frac{\gamma_1\gamma_2 + \gamma_3\gamma_4}{2}] \tag{5.92}$$

substituting s with $= \sigma + j\omega$, the sum polynomial can be expressed in terms of its real and imaginary parts:

$$\begin{aligned} f_1(s) + f_2(s) &= [(\sigma^2 - \omega^2) + \frac{(\gamma_1 + \gamma_2) + (\gamma_3 + \gamma_4)}{2}\sigma + \frac{\gamma_1\gamma_2 + \gamma_3\gamma_4}{2}] + \\ &+2j\omega[2\sigma - \frac{\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4}{2}] \end{aligned} \tag{5.93}$$

The roots of the polynomial are the variables which make both the real and imaginary parts zero. Taking the imaginary part of the sum polynomial

$$2j\omega[2\sigma - \frac{\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4}{2}] = 0 \tag{5.94}$$

which leads to $\sigma = -\frac{\gamma_1+\gamma_2+\gamma_3+\gamma_4}{4}$. From the assumption, we have $\frac{\gamma_1+\gamma_3}{2} \leq \sigma \leq \frac{\gamma_2+\gamma_4}{2}$ and thus the above is proved.

□

(b). **Two summand polynomials of which one has a pair of complex conjugate roots while the other has two real roots**

Proposition 5.12 *If one of the polynomials has two equal real roots γ while the other has a pair of complex conjugate roots $\sigma \pm j\omega$. In this case, we prove that the roots of the sum polynomial will be distributed along the line $\frac{\sigma+\gamma}{2} + j\omega, \omega \in R$.*

Proof:

Because

$$f_1(s) = (s - \gamma)^2, f_2(s) = (s - \sigma)^2 + \omega^2 \tag{5.95}$$

and the sum polynomial is

$$f_1(s) + f_2(s) = 2s^2 - 2(\gamma + \sigma)s + (\gamma^2 + \sigma^2 + \omega^2) \tag{5.96}$$

so the roots of the summation polynomial are

$$s_{1,2} = \frac{\gamma + \sigma \pm \sqrt{-(\gamma - \sigma)^2 - 2\omega^2}}{2} \tag{5.97}$$

□

From the above proposition we immediately have the following corollary.

Corollary 5.6 *For a fixed polynomial $f_1(s)$ with a pair of complex conjugate roots at $\sigma_1 \pm j\omega_1$ and a polynomial $f_2(s)$ with real roots in the region defined as $\delta_1 = [\gamma_1, \gamma_2]$, then the roots of the summation polynomial of the two are confined in the region $\delta = \left\{s : s = \sigma \pm j\omega : \frac{\gamma_1+\sigma_1}{2} \leq \sigma \leq \frac{\gamma_2+\sigma_1}{2}\right\}$.*

Proof:

From proposition 5.12, the sum polynomial $[(s-\sigma_1)^2+\omega_1]+(s+\gamma_1)^2$ has all roots distributed along the line $\frac{\sigma_1+\gamma_1}{2}$ and the sum polynomial $[(s-\sigma_1)^2+\omega_1]+(s+\gamma_2)^2$ has all roots distributed along the line $\frac{\sigma_1+\gamma_2}{2}$. By using continuity argument, the roots of the polynomial $f_1(s)+f_2(s)$ will have all its roots in the region δ.

□

(c). **Both the summand polynomial having complex conjugate roots**

Proposition 5.13 *If the two summand polynomials are $f_1(s)=(s-\sigma_1)^2+\omega_1^2$ and $f_2(s)=(s-\sigma_2)^2+\omega_2^2$, i.e. both having complex conjugate roots. Then the sum polynomial will have all its roots on the line at $s=\frac{\sigma_1+\sigma_2}{2}$.*

Proof:

$$f_1(s)=s^2-2\sigma_1 s+\sigma_1^2+\omega_1^2, f_2(s)=s^2-2\sigma_2 s+\sigma_2^2+\omega_2^2 \tag{5.98}$$

and

$$f_1(s)+f_2(s)=2[s^2+(\sigma_1+\sigma_2)s+\frac{\sigma_1^2+\sigma_2^2}{2}] \tag{5.99}$$

so $f_1(s)+f_2(s)$ has two roots as

$$s_{1,2}=\frac{(\sigma_1+\sigma_2)\pm\sqrt{-(\sigma_1-\sigma_2)^2-(\omega_1^2+\omega_2^2)}}{2} \tag{5.100}$$

which are on the mid-line of $\frac{\sigma_1-\sigma_2}{2}$.

□

From the above proposition 5.13, we can have the following corollary.

Corollary 5.7 *If the summand polynomials $f_1(s), f_2(s)$ have roots distributed in the stripes defined as $\phi_1=\{s=\sigma\pm\omega:\gamma_1\leq\sigma\leq\gamma_2\}$ and $\phi_2=\{s=\sigma\pm\omega:\gamma_3\leq\sigma\leq\gamma_4\}$, then the roots of the summation polynomial will always fall in the stripe defined as $\Phi=\{s=\sigma\pm\omega:\frac{\gamma_1+\gamma_3}{2}\leq\sigma\leq\frac{\gamma_2+\gamma_4}{2}\}$.*

Proof:

From proposition 5.13, the summation polynomial $[(s-\gamma_1)^2+\omega^2]+[(s-\gamma_3)^2+\omega^2]$ has all its roots along the line $\frac{\gamma_1+\gamma_3}{2}$ and the polynomial $[(s-\gamma_2)^2+\omega^2]+[(s-\gamma_4)^2+\omega^2]$ has all its roots along the line $\frac{\gamma_2+\gamma_4}{2}$. So from continuity argument, all the roots of $f_1(s)+f_2(s)$ are in the stripe defined as Φ.

□

5.2.3 Case of third order polynomials

(a). Roots of the summand polynomials are on either side of a line and form mirror images with respect to a line which is parallel to the imaginary axis

Proposition 5.14 *If the roots of the summand polynomials $f_1(s), f_2(s)$ form mirror images with respect to a line which is parallel to the imaginary axis, then the summation polynomial of the two third order polynomials will have all its root distributed along this line.*

Proof:

(i) Summand polynomials with equal real roots.

In the case when the polynomials have three equal roots, $f_1(s)$ and $f_2(s)$ are

$$f_1(s) = s^3 + 3\gamma_1 s^2 + 3\gamma_1^2 s + \gamma_1^3 \quad (5.101)$$

$$f_2(s) = s^3 + 3\gamma_2 s^2 + 3\gamma_2^2 s + \gamma_2^3 \quad (5.102)$$

the sum of these two polynomials is

$$f_1(s) + f_2(s) = 2[s^3 + \frac{3\gamma_1 + 3\gamma_2}{2}s^2 + \frac{3\gamma_1^2 + 3\gamma_2^2}{2}s + \frac{\gamma_1^3 + \gamma_2^3}{2}] \quad (5.103)$$

set $s = \sigma + j\omega$, then

$$\begin{aligned} & f_1(\sigma + j\omega) + f_2(\sigma + j\omega) \\ = \; & 2[\sigma^3 + j3\sigma^2\omega - 3\sigma\omega^2 - j\omega^3 + \frac{3\gamma_1 + 3\gamma_2}{2}(\sigma^2 - \omega^2 + 2j\sigma\omega) \\ & + \frac{3\gamma_1^2 + 3\gamma_2^2}{2}(\sigma + j\omega) + \frac{\gamma_1^3 + 3\gamma_2^3}{2}j\omega] \\ = \; & 2[\sigma^3 - 3\sigma\omega^2 + \frac{(3\gamma_1 + 3\gamma_2)(\sigma^2 - \omega^2)}{2} + \frac{3\gamma_1^2 + 3\gamma_2^2}{2}\sigma + \frac{\gamma_1^3 + \gamma_2^3}{2} \\ & + j3\sigma^2\omega - j\omega^3 + (3\gamma_1 + 3\gamma_2)j\sigma\omega + \frac{3\gamma_1^2 + 3\gamma_2^2}{2}j\omega] \end{aligned} \quad (5.104)$$

We first take the imaginary part, which is

$$3j\omega[\sigma^2 + (\gamma_1 + \gamma_2)\sigma + (\frac{\gamma_1^2 + \gamma_2^2}{2} - \frac{\omega^2}{3})] = 0 \quad (5.105)$$

this leads to either $\omega = 0$ or $\sigma^2 + (\gamma_1 + \gamma_2)\sigma + (\frac{\gamma_1^2+\gamma_2^2}{2} - \frac{\omega^2}{3}) = 0$.

Next, we take the real part, which is

$$\begin{aligned}
& \sigma^3 - 3\sigma\omega^2 + \frac{3\gamma_1+3\gamma_2}{2}(\sigma^2-\omega^2) + \frac{3\gamma_1^2+3\gamma_2^2}{2}\sigma + \frac{\gamma_1^3+\gamma_2^3}{2} \\
= & \sigma^3 + \frac{3\gamma_1+3\gamma_2}{2}\sigma^2 + \frac{3\gamma_1^2+3\gamma_2^2}{2}\sigma + \frac{\gamma_1^3+\gamma_2^3}{2} - (3\sigma\omega^2 + \frac{3\gamma_1+3\gamma_2}{2}\omega^2) \\
= & \frac{1}{2}[(\sigma^3+3\gamma_1\sigma^2+3\gamma_1^2\sigma+\gamma_1^3)+(\sigma^3+3\gamma_2\sigma^2+3\gamma_2^2\sigma+\gamma_2^3)] \\
& -3\omega^2[\sigma+\frac{\gamma_1+\gamma_2}{2}] \\
= & \frac{1}{2}[(\sigma+\gamma_1)^3+(\sigma+\gamma_2)^3] - 3\omega^2[\sigma+\frac{\gamma_1+\gamma_2}{2}] = 0 \qquad (5.106)
\end{aligned}$$

<u>Case 1</u>: When $\omega = 0$, from the real part equation (5.106) we have

$$\frac{1}{2}[(\sigma+\gamma_1)^3+(\sigma+\gamma_2)^3] = 0 \qquad (5.107)$$

where σ, γ_1 and γ_2 are real. And therefore $f(\sigma) = \frac{1}{2}[(\sigma+\gamma_1)^3+(\sigma+\gamma_2)^3$ attains zero only when $\sigma = \frac{\gamma_1+\gamma_2}{2}$.

So when $\omega = 0, \sigma = \frac{\gamma_1+\gamma_2}{2}$ is one of the roots of the summation polynomial.

Second, we consider the case when $\omega \neq 0$.

<u>Case 2</u>: When $\omega \neq 0$, then from equation (5.105) and (5.106) we have

$$\begin{cases} \frac{1}{2}[(\sigma+\gamma_1)^2+(\sigma+\gamma_2)^2] - \frac{\omega^2}{3} = 0 \\ \frac{1}{2}[(\sigma+\gamma_1)^3+(\sigma+\gamma_2)^3] - 3\omega^2[\sigma+\frac{\gamma_1+\gamma_2}{2}] = 0 \end{cases} \qquad (5.108)$$

From equation (5.108) we first show that $\sigma = -\frac{\gamma_1+\gamma_2}{2}, \omega = \frac{\sqrt{3}}{2}(\gamma_2-\gamma_1)$ is a pair of solutions to the equation. From the uniqueness argument of the zeros of the third order polynomials, we have proved that the roots of the summation of the two polynomials will have zeros only on the line defined by $-\frac{\gamma_1+\gamma_2}{2}$.

(ii) Case when the roots of the polynomials form mirror images with respect to a line parallel to the imaginary axis.

Given any two monic third order polynomials, if their roots form mirror images with respect to a line parallel to the imaginary axis, then the summation polynomial of the two will have all its roots on this line.

We prove that on the line there exist 3 roots of the polynomial; therefore the polynomial will have no root other than on this line.

Assume that the variable takes values from the line. For $s = \frac{\sigma_2+\sigma_3}{2} + j\omega$ where ω takes values from $-\infty$ to $+\infty$, then the sum polynomial can be presented as

$$\begin{aligned}
f_1(s) + f_2(s) &= \gamma_1 e^{i\alpha_1}\gamma_2 e^{i\alpha_2}\gamma_3 e^{i\alpha_3} + \gamma_1 e^{i(\pi-\alpha_1)}\gamma_2 e^{i(\pi-\alpha_2)}\gamma_3 e^{i(\pi-\alpha_3)} \\
&= \gamma_1\gamma_2\gamma_3 e^{i(\alpha_1+\alpha_2+\alpha_3)} + \gamma_1\gamma_2\gamma_3 e^{i[3\pi-(\alpha_1+\alpha_2+\alpha_3)]}
\end{aligned}$$

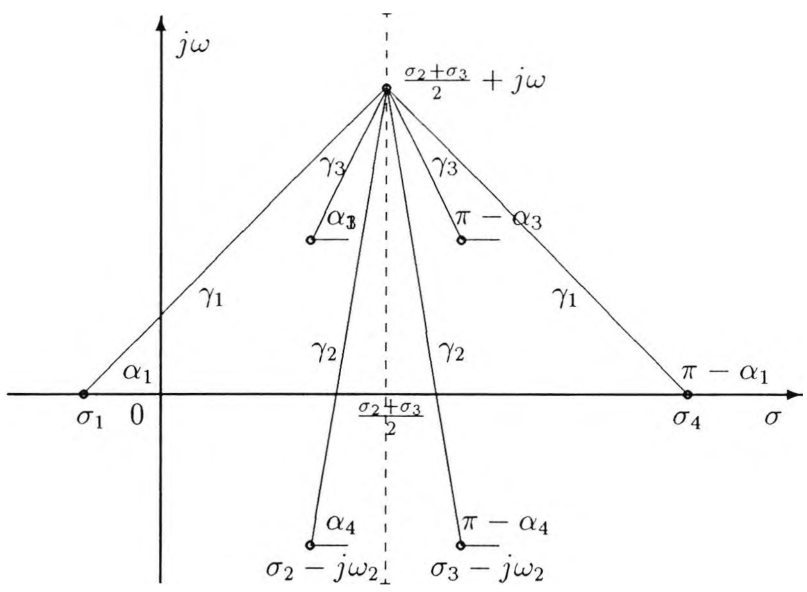


Figure 5.6: Root distributions of the sum of two symmetric third order polynomials

So the polynomial attains zero if and only if

$$\alpha_1+\alpha_2+\alpha_3+(2k+1)\pi = 3\pi-(\alpha_1+\alpha_2+\alpha_3) \tag{5.109}$$

where k takes integer numbers and $\alpha_1+\alpha_2+\alpha_3$ takes values from $-\frac{3\pi}{2}$ to $\frac{3\pi}{2}$. From (5.109) we have

$$\begin{aligned}
2(\alpha_1+\alpha_2+\alpha_3)+(2k+1)\pi &= 3\pi \\
\frac{-3\pi\times 2}{2} \leq 2(\alpha_1+\alpha_2+\alpha_3) &= 3\pi-(2k+1)\pi \leq \frac{3\pi\times 2}{2} \\
6\pi \geq (2k &+ 1)\pi \geq 0 \\
-\frac{1}{2} \leq & \; k \leq 2.5
\end{aligned}$$

Thus there exist three roots of the polynomial on the line.

□

(b). **Summand polynomials with real roots**

Corollary 5.8 *If we assume that the summand polynomials $f(s)$ and $g(s)$ have real roots and all their real roots are distributed in the regions*

$$\delta_1 = \{s\in C: s=\sigma\pm j\omega, \omega=0, \gamma_1\leq\sigma\leq\gamma_3\} \tag{5.110}$$
$$\delta_2 = \{s\in C: s=\sigma\pm j\omega, \omega=0, \gamma_4\leq\sigma\leq\gamma_6\}. \tag{5.111}$$

then its roots are all confined within the region defined as

$$\Gamma = \{s\in C: s=\sigma\pm j\omega, \omega=0, \frac{\gamma_1+\gamma_4}{2}\leq\sigma\leq\frac{\gamma_3+\gamma_6}{2}\}. \tag{5.112}$$

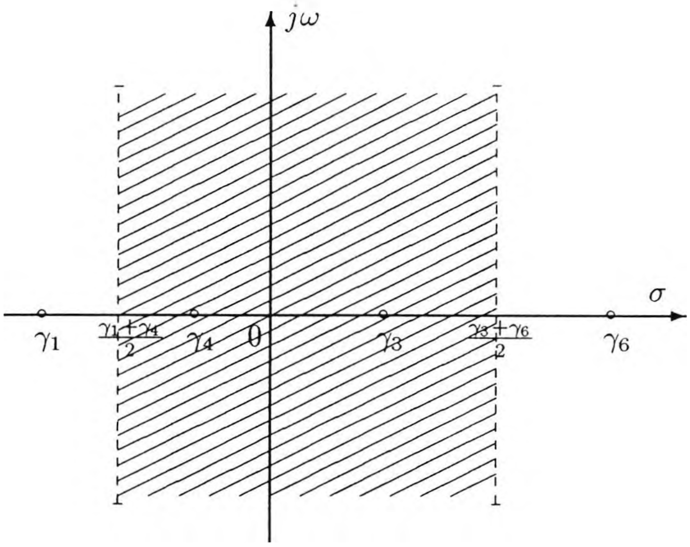


Figure 5.7: Root distributions of the sum of two third order polynomials

Proof:

From the assumption, each of the two summand polynomials has three real roots. Let $-\mu_1, -\mu_2, -\mu_3$ be the three real roots of $f_1(s)$ and $-\nu_1, -\nu_2, -\nu_3$ be the three real roots of $f_2(s)$. Further assume $\mu_1 \leq \mu_2 \leq \mu_3$ and $\nu_1 \leq \nu_2 \leq \nu_3$.

We construct two polynomials using the minimal and the maximal roots of the summand polynomials as in the following:

(1) For polynomial $f(s)$, we construct two polynomials $f_1(s)$ and $f_2(s)$

$$\begin{aligned} f_1(s) &= (s+\mu_3)^3 \\ f_2(s) &= (s+\mu_1)^3 \end{aligned} \tag{5.113}$$

(2) For polynomial $g(s)$, we construct two polynomials $g_1(s)$ and $g_2(s)$

$$\begin{aligned} g_1(s) &= (s+\nu_3)^3 \\ g_2(s) &= (s+\nu_1)^3 \end{aligned} \tag{5.114}$$

From proposition 5.14, the summation polynomial of $f_1(s)$ and $g_1(s)$will have all its roots distributed along line l_1 and the roots of the summation polynomial of $f_1(s)$ and $g_2(s)$ are distributed along the line l_2. If $g(s)$ takes three real roots in the region δ_2 as defined earlier, then using the continuity argument, the roots of the summation polynomials $f_1(s) + g(s)$ will have all its roots distributed in between the lines l_1 and l_2. Similarly, it can be proved that the roots of the summation polynomials $f(s) + g(s)$ are distributed in the region defined by Γ.

5.2.4 High order polynomials

For the general case, the root distribution of the sum polynomial is more difficult. General results as given by Marden may be applied. However, they are very conservative. In this section, we shall generalise some of the results developed in earlier sections for high order polynomials.

(a). **Summand polynomials of which each has n equal real roots**

Proposition 5.15 *If the two summand polynomials each of which has n equal real roots, then the summation polynomial will have all its roots on the line* $-\frac{\gamma_1+\gamma_2}{2}$, *where* γ_1 *and* γ_2 *are the roots of the summand polynomial* $f_1(s)$ *and* $f_2(s)$, *respectively.*

Proof:

We distinguish two different cases, i.e. when the order of the polynomial is even or odd. We will prove for the case when the order of the polynomial is even, that, on the line $\{s = \sigma + j\omega, \sigma = -\frac{\gamma_1+\gamma_2}{2}, \omega \geq 0\}$ we will have $n/2$ roots of the summation polynomial.

Assume $s = \sigma + \omega, \sigma = -\frac{\gamma_1+\gamma_2}{2}, \omega \geq 0$. Then the sum polynomial can be expressed as a function of ω when the variable is in the half line

$$\begin{aligned} f(s) &= (s+\gamma_1)^n + (s+\gamma_2)^n \\ &= (\frac{\gamma_1-\gamma_2}{2} + j\omega)^n + (-\frac{\gamma_1-\gamma_2}{2} + j\omega)^n \end{aligned} \tag{5.115}$$

set $a = \frac{\gamma_1-\gamma_2}{2}$, then

$$\begin{aligned} f(s) &= (a+j\omega)^n + (-a+j\omega)^n \\ &= \gamma^n e^{in\alpha} + \gamma^n e^{in(\pi-\alpha)} \end{aligned} \tag{5.116}$$

where $\alpha = cos\frac{a}{\sqrt{a^2+\omega^2}}, \gamma = \sqrt{a^2+\omega^2}$.

When ω varies from 0 to $+\infty, \alpha$ changes from 0 to $\pi/2$. From equation (5.117), it is clear that $f(s)$ attains zero when

$$n(\pi-\alpha) = n\alpha + (2k+1)\pi \tag{5.117}$$

Because n is even, it can be expressed as $n = 2n_1, n_1 = n/2$. Then the equality (5.117) becomes as

$$2n_1(\pi-\alpha) = an_1\alpha + (2k+1)\pi \tag{5.118}$$

$$4n_1\alpha = (2n_1-2k)\pi + \pi \tag{5.119}$$

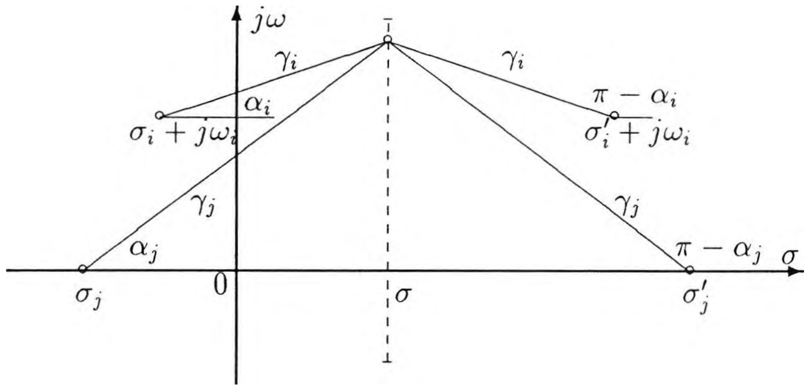


Figure 5.8: Root distribution of the sum of two n-th order polynomials

when α varies from 0 to $\pi/2$, we have the following inequality

$$0 \leq (2n_1 - 2k)\pi + \pi < 2n_1\pi \tag{5.120}$$

which leads to the following inequality

$$1/2 < k \leq \frac{2n_1 + 1}{2} \tag{5.121}$$

for $k = 1, 2, ..., n$, it satisfies the equality (5.117). Therefore, there are $2n_1 = n$ zeros of the summation polynomial on the $-\frac{\gamma_1+\gamma_2}{2}$ line. And from the uniqueness argument, all the roots of the summation polynomial are located on the line.

The case when n is odd can be proved similarly, except that on the half line $(s = \sigma + j\omega, \sigma = -\frac{\gamma_1+\gamma_2}{2}, \omega \geq 0)$, there are $\frac{n+1}{2}$ roots.

□

(b). Roots of the summand polynomials form mirror images with respect to a line parallel to the imaginary axis

Theorem 5.7 *If the roots of the two summand polynomials form mirror images with respect to a line which is parallel to the imaginary axis, then the roots of the summation polynomial will be distributed along this line.*

□

Proof:

The roots of $f_1(s)$ and $f_2(s)$ form mirror images with respect to line l as shown in Figure (5.8). We prove that on line l, there exist n roots of the summation polynomial.

Assuming s moves along line l, i.e. $s = \sigma_0 + j\omega$, where σ_0 is fixed while ω varies from $-\infty$ to ∞, then the summation polynomial can be expressed as

$$\begin{aligned} f_1(s) + f_2(s) &= \gamma_1 e^{i\alpha_1}\gamma_2 e^{i\alpha_2}\cdots\gamma_n e^{i\alpha_n} + \gamma_1 e^{i(\pi-\alpha_1)}\gamma_2 e^{i(\pi-\alpha_2)}\cdots\gamma_n e^{i(\pi-\alpha_n)} \\ &= \gamma_1\gamma_2\cdots\gamma_n e^{i(\alpha_1+\alpha_2+\cdots+\alpha_n)} + \gamma_1\gamma_2\cdots\gamma_n e^{i[n\pi-(\alpha_1+\alpha_2+\cdots+\alpha_n)]} \end{aligned}$$

So the polynomial attains zero if and only if

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n + (2k+1)\pi = n\pi - (\alpha_1 + \alpha_2 + \cdots + \alpha_n) \tag{5.122}$$

where k takes integer numbers and $\alpha_1 + \alpha_2 + \cdots + \alpha_n$ takes values from $-\frac{n\pi}{2}$ to $\frac{n\pi}{2}$. From (5.122) we have

$$\begin{aligned} 2(\alpha_1 + \alpha_2 + \cdots + \alpha_n) + (2k+1)\pi &= n\pi \\ \frac{-n\pi \times 2}{2} \le 2(\alpha_1 + \alpha_2 + \cdots + \alpha_n) &= n\pi - (2k+1)\pi \le \frac{n\pi \times 2}{2} \\ 2n\pi \ge (2k &+ 1)\pi \ge 0 \\ -\frac{1}{2} \le \quad k &\le \frac{2n-1}{2} \end{aligned}$$

there indeed exist n roots of the polynomial on the line.

□

(c). Summand polynomials each having n different real roots

Denote the roots of the summand polynomial $f(s)$ and $g(s)$ as $a_i, b_i, i = 1, 2, ..., n$, $a_1 \le a_2 \le \cdots \le a_n, b_1 \le b_2 \le \cdots \le b_n$. We construct two extreme polynomials for each of the summand polynomial.

(1) For polynomial $f(s)$, we construct two polynomials $f_1(s)$ and $f_2(s)$

$$\begin{aligned} f_1(s) &= (s + a_n)^n \\ f_2(s) &= (s + a_1)^n \end{aligned} \tag{5.123}$$

(2) For polynomial $g(s)$, we construct two polynomials $f_1(s)$ and $g_2(s)$

$$\begin{aligned} g_1(s) &= (s + b_n)^n \\ g_2(s) &= (s + b_1)^n \end{aligned} \tag{5.124}$$

Corollary 5.9 *If both the summand polynomials have real roots, and their roots are distributed in the region defined as in the regions δ_1 and δ_2*

$$\delta_1 = \{s \in C : s = \sigma \pm j\omega, \omega = 0, \gamma_1 \le \sigma \le \gamma_2\} \tag{5.125}$$
$$\delta_2 = \{s \in C : s = \sigma \pm j\omega, \omega = 0, \gamma_3 \le \sigma \le \gamma_4\}. \tag{5.126}$$

then its roots are all confined within the region defined as

$$\Gamma = \{s \in C : s = \sigma \pm j\omega, \omega = 0, \frac{\gamma_1 + \gamma_3}{2} \leq \sigma \leq \frac{\gamma_2 + \gamma_4}{2}\}. \tag{5.127}$$

Proof:

The proof follows along similar line to those of the proof of Corollary 5.8.

□

5.3 Summary

In this chapter, the root distribution regions of bounded coefficient polynomials as well as the sum of two polynomials have been investigated. In the first section, bounded coefficient polynomials have been considered. For stable and completely unstable polynomials minimal necessary regions and maximal rectangular sufficient regions have been established. For the general case, much tighter upper bounds of the necessary region for third order polynomials have also been established, whereas the sufficiency condition based on rectangulars has been shown to extend to the case of general polynomials.

In the second section, the problem of root distribution of the sum of two polynomials with known root distributions has been addressed. For the case when the roots of the summand polynomials form mirror images with respect to a line which is parallel to the imaginary axis, it has been proved that all the roots of the summation polynomial of the two are distributed along this line. Stripe regions have been obtained for the roots of the summation polynomial if the summand polynomials have only real roots but they are confined to an interval. These results are improvements over the previous results available for the roots of the summation of two polynomials.

Chapter 6

POLE MOBILITY AND STABILISABILITY OF LINEAR SYSTEMS UNDER BOUNDED STATE FEEDBACK

6.1 Introduction

Closed-loop control schemes have many advantages over open-loop control schemes; tolerating modelling inaccuracy, parameter variation, operating point drifting, just to name a few. The most important feature, above all, is to shape the dynamic response of the system, i.e. to choose and construct a feedback scheme in such a way that the system performance satisfies the user given specifications.

In the time domain, both state feedback and output feedback schemes can be used. In general, an output feedback control scheme is more practical than a state feedback scheme because the outputs of a system are accessible while the state variables of a system are, in general, not accessible. However, the state feedback scheme has been studied because either the states may be accessible or can be reconstructed using observers; furthermore, the additional significance of the state feedback scheme is due to that output feedback is a special case of state feedback and thus what can be achieved under state feedback provides an upper limit of what can be achieved by output feedback.

An important use of state variable feedback is to reallocate the set of system eigenvalues. The eigenvalues, which are equivalent to the system poles when the state space representation is minimal, are important system property indicators. The most important property shown by the set of system eigenvalues is the system stability. If a system is stable, then it is necessary and sufficient that the eigenvalues do not have positive real parts. The reallocation of the system eigenvalues by using constant gain state feedback has been studied by many researchers, among them are Wonham [Won.,2] and Davison [Dav.,1] and an excellent account of the recent status of the problem may be found in [Mun.,1]. They have proved that the closed-loop eigenvalues can be arbitrarily assigned using constant gain state variable feedback if and only if the system is controllable. Further, they concluded that if the unstable modes of the system are all controllable, then it is always possible to find a state feedback control scheme such that the resulting system is stabilised. An implicit assumption of the standard theory so far has been that the feedback gains are not constrained; this implies that under certain conditions the gains required may be very high and thus the ideal solution predicted by the theory may not be possible to implement. It is the aim of this work here to investigate the effect of the bounded norm assumption of the feedback on system properties such as pole mobility, stabilisability etc.

The standard state feedback theory of linear systems relies extensively on the assumption that there are no constraints on the elements of the controller. In practice, however, very high gains are unrealistic and undesirable either due to saturation and nonlinearity issues, or due to the excessive energy demands associated with the control action. Studying certain issues of the norm restricted state feedback theory is the aim of this chapter; the particular emphasis is on the issues related to eigenvalue mobility under bounded norm feedback. When the gain of the controller is restricted, stabilisability and thus also pole assignability may not be guaranteed under the controllability assumption. The mobility of poles (eigenvalues) under bounded gain feedback is central in the extension of the standard theory to the case of bounded norm state feedback design. For simplicity of analysis, we assume that the feedback gains are constrained as

$$\|\underline{l}\|_n \leq \gamma_f \tag{6.1}$$

where $\underline{l}$ is a vector in the case of SISO systems and a matrix for MIMO systems and $\|\cdot\|_n$ denotes l_n norm. It should be stressed here that the norm of a given feedback matrix may vary, if we carry out a coordinate transformation on the system. It will be assumed that the system states are physical variables on which constraints are naturally imposed and thus coordinate transformations changing a given physical state description to another description of a general type make no sense here as far as preserving constraints on high gain.

The mobility of the closed-loop eigenvalues of a system is related to the quantitative controllability of the system as well as the constraints on the controller. Different measures have been used to define the degree of controllability and the relative pole mobility of the closed-loop systems has been studied [Paig.,1] [Eis.,1] [Bol. & Lu,1] [Tar.,1] etc. The mobility of the closed-loop eigenvalues with respect to the constraints on the controller gain can be investigated in a number of ways depending on how the constraints on the gain are defined. In some of the previous work [Ack.,2] [Bie. Hwa. & Bha.,1] [Kee. Bha. & How.,1] the constraints on the controller gain are defined parametrically, i.e. the controller gain vector takes values from some real intervals and the closed-loop poles are equivalent to the roots of corresponding interval polynomials. So the Kharitonov's theorem can be applied [Kha.,1]. The problem of mobility under output feedback has been considered in [Kar. & Gia.,1], where discs trapping the closed-loop poles, for certain families of systems, are defined. The results of the previous approaches, however, do not apply to the case of norm constrained state feedback. In this chapter the constraints on the controller gain is defined in terms of the l_2 norm. Here we first translate the bound on the controller gain into the bound on the coefficients of the closed-loop characteristic polynomial and then establish the regions containing the closed-loop eigenvalues.

In Section 6.2 the SISO system are investigated and in Section 6.3, we study the MIMO systems. For SISO systems, regions have been established for the closed-loop eigenvalues to be assignable for the given bounded norm feedback. If the system contains unstable open-loop poles, the stabilisability of the systems has also been studied. Similar results have also been established for MIMO systems.

6.2 Bounded norm feedback and the pole mobility, stabilisability of SISO systems

In this section, the problem of closed-loop pole mobility with bounded state feedback controller is considered. A controllable single-input single-output system is assumed with a state-space description given below:

$$\begin{aligned} \dot{\underline{x}} &= A\underline{x} + \underline{b}u(t) \\ y &= \underline{c}^T\underline{x} \end{aligned} \tag{6.2}$$

where $\underline{x} \in R^n, A \in R^{n\times n}, \underline{b} \in R^{n\times 1}, u \in R, y \in R,$ and $\underline{c} \in R^{n\times 1}$.

Assume that the system is under state feedback with feedback gain vector $\underline{k} \in R^n$. So the closed-loop system becomes:

$$\begin{aligned} \dot{\underline{x}} &= (A + \underline{b}\underline{k}^T)\underline{x} + \underline{b}u(t) \\ y &= \underline{c}^T \underline{x}. \end{aligned} \tag{6.3}$$

Here we study the system pole mobility when the feedback gain is constrained. The closed-loop poles will always form a symmetric set with respect to the real axis because the system parameters will always assume real values.

First we define the constraint on the feedback gain vector $\underline{k}$. For mathematical simplicity, we use l_2 norm, $||\underline{k}||_2$ and the bound on the gain will be denoted by γ_f, $\gamma_f > 0$. So the gain vector will satisfy $||\underline{k}||_2 \leq \gamma_f$, $\gamma_f > 0$. So the gain vector will satisfy $||\underline{k}||_2 \leq \gamma_f$, or

$$\sqrt{k_1^2 + k_2^2 + \cdots + k_n^2} \leq \gamma_f \tag{6.4}$$

The coefficient vector of the closed-loop characteristic polynomial is related to the open-loop characteristic polynomial coefficients and the feedback gain in the following way [Kai.,1]: assume that the desired closed-loop characteristic polynomial is

$$f_c(s) = s^n + \beta_1 s^{n-1} + \cdots + \beta_{n-1}s + \beta_n \tag{6.5}$$

while the open-loop characteristic polynomial is:

$$f_o(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_{n-1}s + \alpha_n \tag{6.6}$$

then

$$\underline{k}^T = (\underline{\alpha}^T - \underline{\beta}^T)T^{-1}(\underline{\alpha})Q^{-1}(A, \underline{b}) \tag{6.7}$$

where $\underline{\alpha}^T = \left[\begin{array}{cccc} \alpha_1, & \alpha_2, & \cdots, & \alpha_n \end{array} \right]^T$, $\underline{\beta}^T = \left[\begin{array}{cccc} \beta_1, & \beta_2, & \cdots, & \beta_n \end{array} \right]^T$, $T(\underline{\alpha})$ is the characteristic Toeplitz matrix of A , or

$$T(\underline{\alpha}) = \begin{bmatrix} \alpha_{n-1} & \alpha_{n-2} & \alpha_{n-3} & \cdots & \alpha_1 & 1 \\ \alpha_{n-2} & \alpha_{n-3} & \alpha_{n-4} & \cdots & 1 & 0 \\ \cdot & \cdot & \cdot & \cdots & 0 & 0 \\ \alpha_1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{6.8}$$

clearly, $|T(\underline{\alpha})| \neq 0$. $Q(A, \underline{b})$ is the controllability matrix of the open-loop system $(A, \underline{b})$, i.e.

$$Q(A, \underline{b}) = \left[\begin{array}{cccc} \underline{b}, & A\underline{b}, & \cdots, & A^{n-1}\underline{b} \end{array} \right]. \tag{6.9}$$

From (6.7), the coefficient vector of the closed-loop characteristic polynomial satisfies

$$\underline{\beta}^T = \underline{\alpha}^T - \underline{k}^T R(A, \underline{b}) \tag{6.10}$$

where $R(A, \underline{b}) = Q(A, \underline{b})T(\underline{\alpha})$ and will be referred to as the *scaled controllability matrix.* So the norm of the coefficient vector of the closed-loop characteristic polynomial satisfies the following inequalities:

$$\begin{aligned} \|\underline{\beta}^T\|_2 &= \|\underline{\alpha}^T - \underline{k}^T R(A, \underline{b})\|_2 \\ &\leq \|\underline{\alpha}^T\|_2 + \|\underline{k}^T R(A, \underline{b})\|_2 \\ &\leq \|\underline{\alpha}^T\|_2 + \|\underline{k}^T\|_2 \cdot \|R(A, \underline{b})\|_2 \\ &\leq \|\underline{\alpha}^T\|_2 + \gamma_f \cdot \|R(A, \underline{b})\|_2 \end{aligned} \tag{6.11}$$

Set

$$\gamma_p \equiv \|\underline{\alpha}^T\|_2 + \gamma_f \cdot \|R(A, \underline{b})\|_2, \tag{6.12}$$

then

$$\|\underline{\beta}^T\|_2 \leq \gamma_p. \tag{6.13}$$

Equation (6.13) has translated the norm bound on the feedback gain into the norm bound on the coefficients of the closed-loop characteristic polynomial. It is shown clearly that the coefficients of the closed-loop characteristic polynomial are bounded when the controller gain is bounded; therefore, the closed-loop poles cannot be arbitrarily allocated.

Remark 6.1: From equations (6.12) and (6.13) it is shown clearly that the bound on the coefficients of the closed-loop characteristic polynomial is affected both by the bound on the feedback as well as the l_2 norm of the scaled controllability matrix $R(A, \underline{b})$.

□

Remark 6.2: The l_2 norm of the scaled controllability matrix $R(A, \underline{b})$ is affected by the coordinate transformation on the system. Let Q be a nonsingular coordinate transformation, then the l_2 norm of the new scaled controllability matrix becomes

$$\|R(A', \underline{b}')\|_2 = \|Q[\underline{b}, A\underline{b}, A^2\underline{b}, ..., A^{n-1}\underline{b}]Q^{-1}T(\underline{\alpha})\|_2 = \gamma_Q \cdot \|R(A, \underline{b})\|_2 \tag{6.14}$$

where γ_Q is the condition number of the transformation matrix Q. So the study of the pole mobility of the closed-loop system should always be carried out with respect to the natural coordinates and without using coordinate transformations which may affect the meaning of the norm conditions. It is also clear from the above that the bounded feedback pole mobility study makes sense only when the states are natural variables for which imposing constraints makes sense.

□

The mobility of the closed-loop poles depends on the scaled controllability matrix $R(A, \underline{b})$. In the following, it can be demonstrated that if a system is nearly uncontrollable, i.e. the smallest singular value of the scaled controllability matrix is very small, then excessively large controller gain is needed to move the poles away from their original location. In fact, let us rewrite equation (6.10) as

$$\underline{\alpha}^T - \underline{\beta}^T = \underline{k}^T R(A, \underline{b}) \tag{6.15}$$

and also write the scaled controllability matrix in SVD form

$$R(A, \underline{b}) = U \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \end{bmatrix} V \tag{6.16}$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ are the singular values of the scaled controllability matrix. Then equation (6.15) is equivalent to

$$(\underline{\alpha}^T - \underline{\beta}^T)[\underline{w}_1, \underline{w}_2, ..., \underline{w}_n] = \underline{k}^T [\underline{u}_1, \underline{u}_2, ..., \underline{u}_n] \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \end{bmatrix} \tag{6.17}$$

where $[\underline{w}_1, \underline{w}_2, ..., \underline{w}_n] = V^{-1}$, $[\underline{u}_1, \underline{u}_2, ..., \underline{u}_n] = U$. From equation (6.17), the relation between the controller gain and the singular values of the scaled controllability matrix can be presented as:

$$(\underline{\alpha}^T - \underline{\beta}^T) \times \underline{w}_i = \sigma_i \underline{k}^T \underline{u}_i, \quad i = 1, 2, .., n \tag{6.18}$$

and thus the l_2 norm of the controller gain satisfies

$$||\underline{k}||_2 \leq \frac{1}{\sigma_i} \times \frac{||\underline{\alpha}^T - \underline{\beta}^T||_2 \cdot ||\underline{w}_i||_2}{||\underline{u}_i||_2} = \frac{||\underline{\alpha}^T - \underline{\beta}||_2}{\sigma_i}, \quad i = 1, 2, ..., n \tag{6.19}$$

therefore nearly uncontrollable systems will have smaller singular values and will need larger gains to control.

Remark 6.3: The mobility of the closed-loop poles depends not only on the bound on the controller gain, but also on the singular values of the scaled controllability matrix. The bigger the singular values, the easier the poles are able to be moved away from their original location.

□

The problems studied in this section are:

(i) Establishing upper bounds for the pole mobility region which contains all the closed-loop eigenvalues under bounded norm feedback.

(ii) Establishing upper bounds for the pole mobility region which contains all the eigenvalues of the stable closed-loop system under bounded norm feedback.

(iii) If some of the open-loop poles are unstable, examining the stabilisability under bounded state feedback.

The results presented in Chapter 5 on the root distributions of norm bounded polynomials can be applied on the root distributions of the closed-loop characteristic polynomials. The necessary root inclusion regions for the bounded coefficients closed-loop characteristic polynomial define the necessary assignable closed-loop pole regions for the given bounded state feedback system. In the following we study the necessary assignable pole region for the general case by using Theorem (4.9), Proposition (5.5) and then the necessary assignable pole region employing Theorem (5.2) when the closed-loop poles are all assumed to be stable.

6.2.1 Pole Mobility Regions which contain all the closed-loop eigenvalues under bounded norm feedback

The results stated in the previous section may now be used for the study of pole mobility of SISO systems under bounded norm feedback.

(1). <u>The Ostrowski's pole mobility region.</u>

The result presented in Theorem (4.9) provides the links between the distance of the roots and the coefficients of two polynomials. Assume the same notations as in Theorem (4.9). Now $f(s)$ and $g(s)$ represent the open- and closed-loop characteristic polynomials, respectively, i.e.

$$f(s) = f_o(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_{n-1}s + \alpha_n$$
$$g(s) = f_c(s) = s^n + \beta_1 s^{n-1} + \cdots + \beta_{n-1}s + \beta$$

By employing equation (6.15), the values of $\Gamma_1 = \max\{1, |\alpha_i|, |\beta_i|\}$ and $d = \sum_{i=1}^{n} |\alpha_i - \beta_i|$ can be calculated as

$$\begin{aligned}\Gamma_1 &= \max\{1, |\alpha_i|, |\beta_i|\} \\ &\leq \max\{1, |\alpha_i| + \gamma_f \cdot \|R(A, \underline{b})\|_2\} \end{aligned} \quad (6.20)$$

$$\begin{aligned} d &= \sum_{i=1}^{n} |\alpha_i - \beta_i| \\ &\leq n \cdot \gamma_f \cdot \|R(A, \underline{b})\|_2 \end{aligned} \tag{6.21}$$

then by applying Theorem (4.9), we can compute the maximal distance, δ, between the open-loop and closed-loop poles by

$$\delta = \max_{i,j} |\lambda_i - \mu_j| \leq (n+2)\Gamma_1 d^{1/n}. \tag{6.22}$$

where $\lambda_i, \mu_j, i, j = 1, 2, ..., n$ are the open- and closed-loop poles of the system. A result on the closed-loop pole mobility under bounded feedback is described below:

Theorem 6.1 *The poles of the SISO closed-loop system under a bounded state feedback $\underline{k}$, such that $\|\underline{k}\|_2 \leq \gamma_f$, will be in the circles which are centred at the open-loop poles with a radius δ where*

$$\delta = (n+2)\Gamma_1 d^{1/n} \tag{6.23}$$

and Γ_1, d are defined as above.

□

Remark 6.4: The bound obtained for the closed-loop poles with respect to the open-loop poles is shown to be proportional to $(\gamma_f \cdot \|R(A, \underline{b})\|_2)^{(1+\frac{1}{n})}$. Thus a system with larger values of γ_f and $\|R(A, \underline{b})\|_2$ will have greater closed-loop pole mobility.

□

(2). Single disc type pole mobility region.

By using Theorem (5.4), we can establish another upper bound for pole mobility region for the closed-loop system. Let us define a circle, D_f, in the complex plane with the centre at the origin such as:

$$D_f \equiv \{s \in \mathrm{C}; |s| \leq 1 + \gamma_p\}. \tag{6.24}$$

where γ_p is defined by equation (6.12), i.e. $\gamma_p \equiv \|\underline{\alpha}^T\|_2 + \gamma_f \cdot \|R(A, \underline{b})\|_2$. From Theorem (5.4), every eigenvalue of the closed-loop system is contained in a circle defined by the following result.

Theorem 6.2 *For a controllable system $(A, \underline{b})$, if the state feedback gain is constrained by l_2-norm as $\|\underline{k}\|_2 \leq \gamma_f$, then all the eigenvalues of any closed-loop system will have all its roots in the circle D_f.*

□

Remark 6.5: This result demonstrates that the bound on the closed-loop poles is proportional to the value $\gamma_f \cdot \|R(A, \underline{b})\|_2$. Compared with the result in Theorem (6.1), Theorem (6.2) is less conservative when the value $\gamma_f \cdot \|R(A, \underline{b})\|_2$ is large.

□

(3). Pole mobility of third order systems:

For lower dimension systems ($n \leq 3$) a much tighter necessary closed-loop pole region can be established by using Theorem (5.6).

Theorem 6.3 *If the system is of order* 3, *then all the complex conjugate closed-loop poles lie within the circle centred at the origin with radius*

$$\rho = \sqrt{\sqrt{\gamma_p^2} + \triangle},$$

where $\triangle = \gamma_p \times \frac{\sqrt{\gamma_p^2+4}-\gamma_p}{2} \leq 1.0$ *and all the real roots satisfy* $|\mu_i| \leq 1 + \gamma_p$.

□

Remark 6.6: Compared with Theorems (6.1) and (6.2), the result given in Theorem (6.3) is the least conservative for third order systems because the bound on the closed-loop poles is proportional to $\sqrt{\gamma_f \cdot \|R(A, \underline{b})\|_2}$.

□

A comparison of the different results will be made at the end by an example.

6.2.2 Pole Mobility Region of the stable closed-loop eigenvalues under bounded norm feedback

When the closed-loop system is stable under the constrained feedback, then a tighter region for all the roots of the closed-loop characteristic polynomial can be obtained.

For $\gamma_p \geq 0$, define as in Definition (5.2) the Γ_p-Prime region as:

$$\Phi^+_{\gamma_p} \equiv \{s = \sigma \pm j\omega \in C^+ : 0 \geq \sigma \geq -z_{\gamma_p}, |\omega| \leq \bar{\omega}(\sigma)\} \quad (6.25)$$

$$\Theta^+_{\gamma_p} \equiv \{s = \sigma \pm j\omega \in C^+ : -z_h \leq \sigma < -z_{\gamma_p}, \omega = 0\} \quad (6.26)$$

$$\Gamma^+_{\gamma_p} \equiv \Phi^+_{\gamma_p} \cup \Theta^+_{\gamma_p} \quad (6.27)$$

where $z_h = \gamma_p, z_{\gamma_p} = \sqrt{-2 + \sqrt{4 + \gamma_p^2}}$ and $\bar{\omega}(\sigma) = \sqrt{-\sigma^2 + \sqrt{\gamma_p^2 - 4\sigma^2}}$. Then all the stable closed-loop system will have all its characteristic polynomial roots in Γ_p-

Prime region. Thus, if γ_p is defined as indicated by equation (6.12), we have the following theorem.

Theorem 6.4 *For a controllable system* $(A, \underline{b})$, *if the state feedback gain is constrained by* l_2*-norm as* $\|\underline{k}\|_2 \leq \gamma_f$, *then all the eigenvalues of any stable closed-loop system will have all its roots in the* Γ_p*-Prime region.*

Proof:

For a controllable system $(A, \underline{b})$, if the state feedback gain is constrained by l_2-norm as $\|\underline{k}\|_2 \leq \gamma_f$, then from equations (6.10-6.12), it is shown that the coefficient vector of the closed-loop polynomial is bounded by γ_p, or

$$\|\underline{\beta}^T\|_2 \leq \gamma_p \equiv \|\underline{\alpha}^T\|_2 + \gamma_f \cdot \|R(A, \underline{b})\|_2. \tag{6.28}$$

So from Theorem (5.2), it follows that all the eigenvalues of any stable closed-loop system under this bounded controller will be in the Γ_p-Prime region.

□

Remark 6.7: The circular region D_f contains $\Gamma^+_{\gamma_p}$, or $\Gamma^+_{\gamma_p} \subset D_f$. So the conservativeness of the maximal region is reduced in the case when all the closed-loop systems are stable by using $\Gamma^+_{\gamma_p}$ instead of D_f. The $\Gamma^+_{\gamma_p}$ region is the smallest uncertainty region for the mobility of the stable closed-loop eigenvalues under the bounded state feedback.

□

Remark 6.8: For the necessary pole mobility regions obtained here there exist points which will never be closed-loop eigenvalues of the state feedback system with the given bound constraint.

□

A comparison of the different results, which also demonstrates the conservativeness of them, is given below.

Example (6.1): A helicopter near hover can be described by the equations

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -0.002 & -1.4 & 9.8 \\ -0.01 & -0.4 & 0 \\ 0 & 1.0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 9.8 \\ 6.3 \\ 0 \end{bmatrix} u \tag{6.29}$$

where x_1 = horizontal velocity, x_2 = pitch rate, x_3 = pitch angle, u= rotor tilt angle.

The open-loop characteristic polynomial is

$$f_o(s) = s^3 + 0.42s^2 - 0.006s + 0.098$$

with open-loop poles at

$$\begin{aligned} \lambda_1 &= -0.6565 \\ \lambda_{2,3} &= 0.1183 \pm j0.3678 \end{aligned}$$

therefore the controllability matrix and the Toeplitz matrix of $\underline{\alpha}$ are

$$Q(A,\underline{b}) = \begin{bmatrix} 9.8 & -9.016 & 65.58552 \\ 6.3 & -2.618 & 1.13736 \\ 0 & 6.3 & -2.618 \end{bmatrix}, T(\underline{\alpha}) = \begin{bmatrix} -0.006 & 0.42 & 1 \\ 0.42 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (6.30)$$

so

$$R(A,\underline{b}) = Q(A,\underline{b})T(\underline{\alpha}) = \begin{bmatrix} 62.7142 & -4.9 & 9.8 \\ 0 & 0.028 & 6.3 \\ 0.028 & 6.3 & 0 \end{bmatrix} \quad (6.31)$$

The l_2 norm of the matrix $R(A,\underline{b})$ is

$$\|R(A,\underline{b})\|_2 = 62.1742 \quad (6.32)$$

Assume the feedback gain vector is bounded by $\gamma_f = \|\underline{k}\|_2 \leq 1$, then the norm bound on the coefficients of the closed-loop characteristic polynomial in equation (6.12) is

$$\begin{aligned} \gamma_p &= \|\underline{\alpha}^T\|_2 + \gamma_f \cdot \|R(A,\underline{b})\|_2 \\ &= 63.1455 \end{aligned}$$

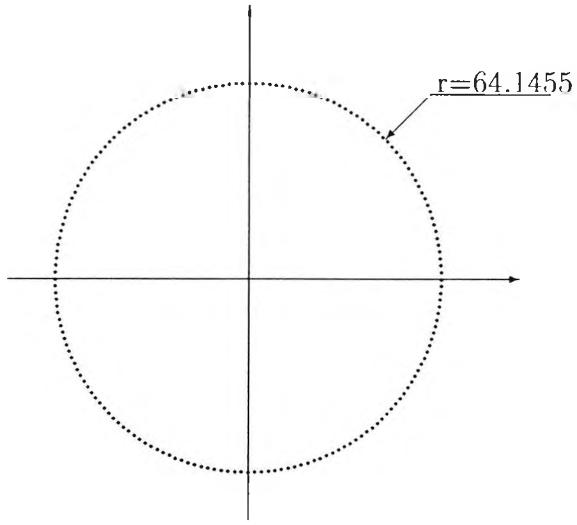
So the necessary closed-loop pole regions for the general case can be established.

(1). From Theorem (6.1), the closed-loop poles are located in the circles which are centred at the open-loop poles $\lambda_i, i = 1,2,3$ with a radius δ, where

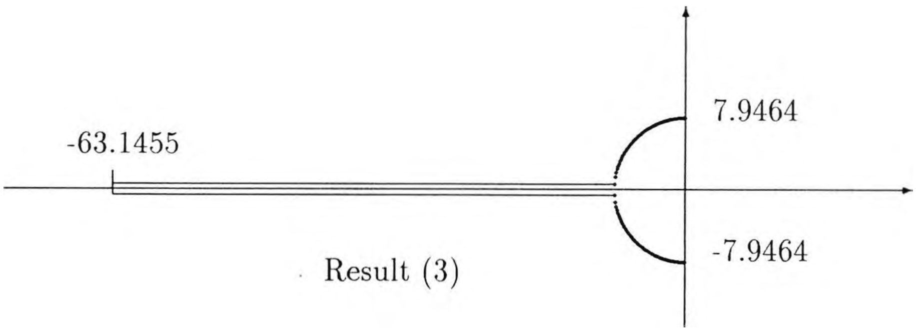
$$\delta = (n+2)\Gamma_1 d^{1/n} = 1506.2 \quad (6.33)$$

(2). From Theorem (6.2), we can obtain a circle D_f centred at the origin with a radius $1 + \gamma_p$, or

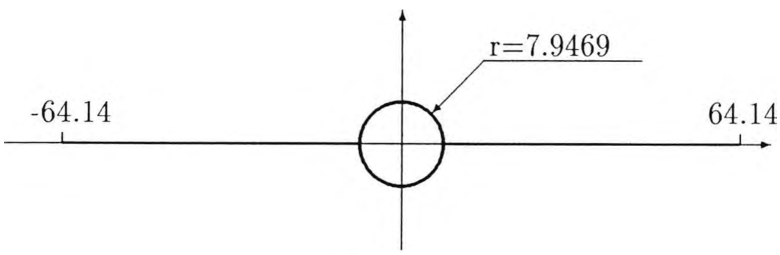
$$D_f = \{s \in C; |s| \leq 1 + \gamma_p = 64.1455\} \quad (6.34)$$



Result (2)



Result (3)



Result (4)

Figure 6.1: Upper Bounds for Closed-loop Pole Mobility Regions

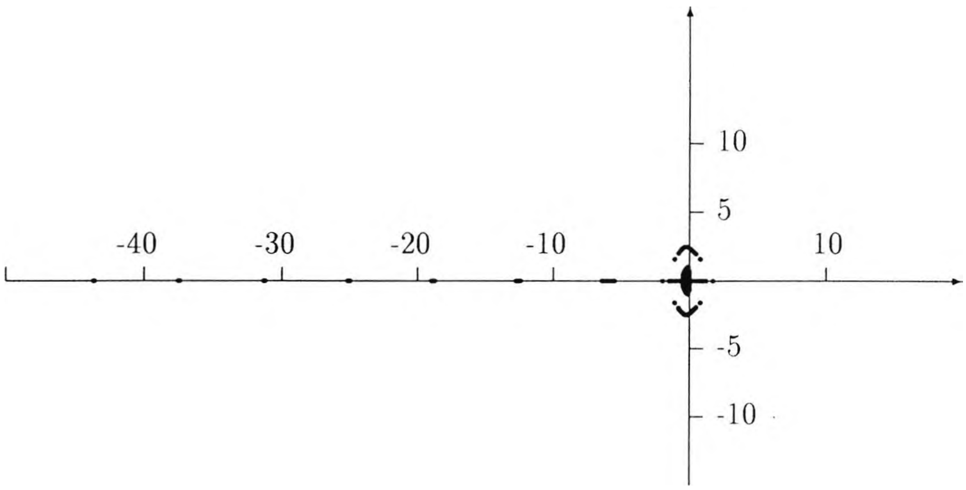


Figure 6.2: Actual Closed-loop Pole Distribution Under Bounded Feedback

(3). Yet, we can establish a much tighter necessary closed-loop pole region by using Theorem (6.3). From Theorem (6.3), the complex conjugate closed-loop poles must lie within the circle

$$\left\{ s \in C; |s| \leq \rho = \sqrt{\sqrt{\gamma_p^2 + \gamma_p \times \frac{\sqrt{\gamma_p^2 + 4} - \gamma_p}{2}}} = 7.9469 \right\} \tag{6.35}$$

while the real closed-loop poles μ_j satisfying $|\mu_j| \leq 1 + \gamma_p = 64.1455$.

(4). If we require all the closed-loop poles to be stable, then from Theorem (6.4), all the closed-loop poles are in the Γ_p-Prime region as defined in (6.27)

$$\begin{aligned} \Phi_{\gamma_p}^+ &\equiv \left\{ s = \sigma \pm j\omega \in C^+ : -7.822 \leq \sigma \leq 0, |\omega| \leq \sqrt{-\sigma^2 + \sqrt{63.1455^2 - 4\sigma^2}} \right\} \\ \Theta_{\gamma_p}^+ &\equiv \{ s = \sigma \pm j\omega \in C^+ : -63.1455 \leq \sigma < -7.822, \omega = 0 \} \\ \Gamma_{\gamma_p}^+ &\equiv \Phi_{\gamma_p}^+ \cup \Theta_{\gamma_p}^+ \end{aligned}$$

The results are presented in Figure (6.1), while the actual pole distribution of the closed-loop system is shown in Figure (6.2). It is shown clearly that the result obtained by using Theorem (6.1) is the most conservative and the results yielded by applying Theorem (6.3) and Theorem (6.4) are the least conservative.

Remark 6.9: The norm of the scaled controllability matrix, $R(A, \underline{b})$ affects the upper bounds for the closed-loop pole mobility in such a way that the bigger the norm, the larger the mobility of the system.

□

6.2.3 Stabilisability of unstable SISO systems

In this subsection, the stabilisability of open-loop unstable SISO systems is considered. Given an open-loop unstable system, then we say that the system is γ_f-stabilisable if there exists a feedback $\underline{k}$ which satisfies $\|\underline{k}\|_2 \leq \gamma_f$ such that the resulted closed-loop system has all the eigenvalues in the left half of the complex plane; otherwise, the unstable system is not γ_f-stabilisable under any γ_f-bounded feedback.

In the following we define a region, D_n, with the property that if an unstable system has a root outside this region, then the system is not stabilisable by the bounded state feedback; this region will be referred to as the necessary stabilisable region. This region is defined in terms of the bound of the state feedback gain, γ_f and the scaled controllability matrix, $R(A, \underline{b})$.

Necessary stabilisable region

Assume that the open-loop characteristic polynomial and $f_o(s)$ and the desirable closed-loop polynomial $f_c(s)$ are:

$$\begin{aligned} f_o(s) &= s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n, \\ f_c(s) &= s^n + \beta_1 s^{n-1} + \cdots + \beta_n \end{aligned}$$

and that the bound on the state feedback gain is $\|\underline{k}\|_2 \leq \gamma_f$. From equation (6.7), it should be remembered that

$$\|\underline{\alpha}\|_2 \leq \|\underline{\beta}\|_2 + \gamma_f \cdot \|R(A, \underline{b})\|_2 \equiv \gamma_n \tag{6.36}$$

Using the results for pole mobility we have:

Theorem 6.5 *For the system $(A, \underline{b})$, if A has an unstable eigenvalue outside the circle*

$$D_n \equiv \{s \in \mathrm{C}; |s| \leq 1 + \gamma_n\} \tag{6.37}$$

then the system is not γ_f-stabilisable.

Proof:

If A has an unstable eigenvalue outside the circle D_n, it follows from Theorem (5.4) that the coefficient vector of the open-loop characteristic polynomial, $\underline{\alpha}$, satisfies

$$\|\underline{\alpha}\|_2 \geq \gamma_n \tag{6.38}$$

which is contradictory to equation (6.36). Therefore, the open-loop poles must lie inside the circle D_n.

□

Therefore, the part of D_n disc which lies in the right half plane defines a necessary region within which all unstable eigenvalues must lie for the system to be γ_f-stabilisable.

Another necessary condition for system stabilisation can be developed by using Theorem (3.1). As stated in Theorem (3.1), the closed-loop poles of a system when subject to a bounded state feedback $\underline{k}$, such that $\|\underline{k}\|_2 \leq \gamma_f$ can be moved at most to a distance from the open-loop poles by $\delta = (n+2)\Gamma_1 d^{1/n}$, where Γ_1 and d are defined in equations (6.20) and (6.21). In order to stabilise a system, this distance should be larger than the distance of any unstable pole from the imaginary axis. If we define Δ as

$$\Delta = \max\{|real(\lambda_i)|, \lambda_i \in \Lambda_{ou}\} \tag{6.39}$$

where Λ_{ou} is the set of open-loop unstable eigenvalues, then we have the following necessary condition for system stabilisability.

Theorem 6.6 *Let $(A, \underline{b})$ be a stabilisable pair. A necessary condition for this pair to be γ_f-stabilisable is that*

$$\delta \geq \Delta \tag{6.40}$$

where $\delta = (n+2)\Gamma_1 d^{1/n}$ with Γ_1, d defined by (6.20), (6.21) and Δ is as defined by (6.39) above.

□

The above result defines an alternative type of necessary condition for γ_f-stabilisability in terms of the distance of the unstable eigenvalues from the imaginary axis. These two results may be used in a negative way to infer non-stabilisability. Sufficient conditions are considered next.

6.2.4 Sufficient conditions for stabilisability

A sufficient condition for γ_f-stabilisability can be obtained by employing equation (6.7).

Corollary 6.1 *Given a system $(A, \underline{b})$ and the set of desired closed-loop poles as Λ_c, then the system can be γ_f-stabilised, if the controller $\underline{k}$ satisfies*

$$\|\underline{k}\|_2 \geq \|(\underline{\alpha} - \underline{\beta}) \cdot R^{-1}(A, \underline{b})\|_2. \tag{6.41}$$

□

This is an obvious result that provides the means for working out sufficient conditions. If we set $\Lambda_c = \{0, 0, ..., 0\}$, a sufficient condition for γ_f-stabilisability is given below:

Remark 6.10: A sufficient condition for γ_f-stabilisability is that

$$\|\underline{k}\|_2 \geq \|\underline{\alpha} \cdot R^{-1}(A, \underline{b})\|_2. \tag{6.42}$$

□

Remark 6.11: From (6.41), it follows that $\|\underline{k}\|_2$ satisfies the inequality

$$\bar{\gamma}^{-1} \leq \frac{\|\underline{k}\|_2}{\|\underline{\alpha} - \underline{\beta}\|_2} \leq \underline{\gamma}^{-1} \tag{6.43}$$

where $\underline{\gamma}, \bar{\gamma}$ are the maximum and minimum singular values of $R(A, \underline{b})$.

□

Another sufficient condition for stabilisability can also be obtained if we first find a stable characteristic polynomial, $f_c(s)$, which has a minimum distance from the unstable characteristic polynomial. Instead of using $\Lambda_c = \{0, 0, ..., \}$ in Remark 6.10, the set of roots of $f_c(s)$ is used. Then the controller gain to achieve the assignment for this set of closed-loop poles defines a sufficient condition for stabilisability. In order to find the closest stable characteristic polynomial from the unstable open-loop characteristic polynomial, the procedure described in Section 4.6 can be used. However, the performance index to be minimised is a function of $\underline{k}$ and the scaled-controllability matrix $R(A, \underline{b})$. Indeed,

$$Q(\Delta f) = \|\underline{k}^T R^{-1}(A, \underline{b})\|_2^2 \tag{6.44}$$

Having found the closest stable characteristic polynomial, say $f_c^o(s)$, the corresponding controller gain, $\underline{k}^o$, can be computed

$$\underline{k}_o^T = (\underline{\beta}_o^T - \underline{\alpha}^T) R^{-1}(A, \underline{b}) \tag{6.45}$$

where $\underline{\beta}_o$ is the coefficient vector of $f_c^o(s)$. Therefore, by continuity argument, another sufficient condition for stabilisability is that

$$\|\underline{k}\|_2 \geq \|\underline{k}_o\|_2 \tag{6.46}$$

Example (6.2):

Given an unstable system $(A, \underline{b})$,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & -5 & 6 & -4 \end{bmatrix}, \underline{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

and the open-loop characteristic polynomial is

$$f^o(s) = s^4 + 4s^3 - 6s^2 + 5s + 2$$

with roots $-0.28518, -5.29736$, and $0.79127 \pm j0.83534$.

From Remark 6.10, a sufficient condition is given by

$$\|\underline{k}\|_2 \geq \|\underline{\alpha} \cdot R^{-1}(A, \underline{b})\|_2 = 9.0$$

As shown in Example (4.2), a closest stable polynomial is found to be

$$f_c^o(s) = s^4 + 4.98773s^3 + 0.0002s^2 + 0.0004s + 1.005^{-8}$$

with roots at $-4.9877, -0.000027$, and $-1.929^{-8} \pm j8.611^{-3}$. So the controller gain $\underline{k}_o$ is

$$\begin{aligned} \underline{k}_o^T &= (\underline{\beta}_o^T - \underline{\alpha}^T)R^{-1}(A, \underline{b}) \\ &= \begin{bmatrix} 0.98773 & 6.0002 & -4.9996 & -2 \end{bmatrix} \end{aligned}$$

So another sufficient condition can be presented as

$$\|\underline{k}\|_2 \geq \|\underline{k}_o\|_2 = 8.1224$$

6.3 Bounded norm feedback of MIMO systems

6.3.1 Introduction

An important study in the closed-loop design of the MIMO systems is the closed-loop eigenvalue placement via state feedback. It was proved by Wonham [Won.,2] that the closed-loop eigenvalues of a MIMO system can be arbitrarily assigned via state feedback if and only if the system is controllable. Different pole assignment methods, as well as formulae for computing the controller gain matrices have been presented [Won.,2] [Bas. & Gur.,1] [Ack.,2] [May. & Mur.,1] [Ros.,1] [Mun.,1] [Moo.,1]

[Verg. & Kai.,1] [Kar. & Gia.,2] [Pat.,1]. Here we first study the methods of assigning the set of desired closed-loop eigenvalues to the system. We present two ways of closed-loop eigenvalue assignment. The first method will be based on the controller form of the system while the second will be using dyadic feedback. Then the eigenvalue mobility of the system, when subject to bounded state feedback will be investigated for different pole placement methods. In Section 6.3.2, the methods of pole assignment are presented while in section 6.3.3, closed-loop pole mobility of the MIMO systems under bounded norm state feedback are presented.

6.3.2 Methods of pole assignment

Given the following MIMO system,

$$\begin{aligned} \dot{\underline{x}} &= A\underline{x}(t) + B\underline{u}(t) \\ \underline{y} &= C^T\underline{x}(t) \end{aligned} \tag{6.47}$$

where $\underline{x} \in R^n, A \in R^{n\times n}, B \in R^{n\times l}, \underline{u} \in R^l, \underline{y} \in R^m$, and $C \in R^{n\times m}$.

The system is subject to state feedback

$$\underline{u} = K\underline{x}(t) + G\underline{v}(t) \tag{6.48}$$

where $K \in R^{l\times n}, G \in R^{l\times l}$ and $\underline{v} \in R^l$ is the real system input.

So the closed-loop system is equivalent to

$$\begin{aligned} \dot{\underline{x}} &= (A + BK)\underline{x}(t) + BG\underline{u}(t) \\ \underline{y} &= C^T\underline{x}(t) \end{aligned} \tag{6.49}$$

If the system (A, B) is controllable, or equivalently, $Q = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$ is of full rank, then it is proved that the closed-loop eigenvalues of the system can be assigned arbitrarily when the desired closed-loop eigenvalues form a symmetric set [Won.,1].

Theorem 6.7 [Won.,2] *For a MIMO system as given in (6.47), the desired closed-loop eigenvalues can always be assigned if and only if the system is controllable.*

□

For controllable MIMO systems, the pole assignment can be carried out either based on the canonical controller form of the system or by a direct method if the matrix A is simple. First, the pole assignment based on the canonical controller form of the system is investigated.

(a). **Pole Assignment Based on Controller Form**

In order to obtain a canonical controller form of a system, we first have to select a new coordinate system. Because the system is assumed to be controllable, so the controllability matrix, $Q = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$ of the system is of full rank. From this $n \times nl$ matrix, we can select a set of n vectors which are linearly independent. Different from the case of SISO systems, the selection of the n linearly independent vectors can be done in different ways. Depending on the way the set of linearly independent vectors is selected, the controller form will in general be different. In this sense, the controller form cannot be defined as canonical. However, if the selection procedure of the linearly independent vectors is fixed, then for similar matrices, we will always obtain the same controller form.

For a given controllability matrix, $Q = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$, of the system, we select from left to right until n linearly independent vectors are found. Rearrange then in the form

$$Q_o = \left\{ \begin{array}{cccccccc} b_1 & Ab_1 & \cdots & A^{k_1-1}b_1 & b_2 & \cdots & A^{k_l-1}b_l \end{array} \right\} \tag{6.50}$$

and define

$$\sigma_1 = k_1, \quad \sigma_2 = k_1 + k_2, ..., \sigma_l = \sum_{i=1}^{l} k_i = n \tag{6.51}$$

Now let

$$q_i = \text{ the } \sigma_i\text{th row of } Q_o^{-1} \tag{6.52}$$

and form

$$T^{-1} = \begin{bmatrix} (A^T)^{k_1-1}q_1^T & ... & q_1^T & (A^T)^{k_2-1}q_2^T & ... & q_l^T \end{bmatrix}^T \tag{6.53}$$

Then the newly formed transformation matrix T will transform the system pair (A, B) into controller form, i.e.,

$$T^{-1}AT = A_c, \qquad T^{-1}B = B_c \tag{6.54}$$

where A_c and B_c are in the forms

$$
A_c = \left[\begin{array}{ccccccccccccccc}
\times & \times & \cdots & \times & \times & \times & \times & \cdots & \times & \times & \times & \times & \cdots & \times & \times \\
1 & 0 & \cdots & 0 & 0 & & & & & & & & & & \\
0 & 1 & \cdots & 0 & 0 & & & & & & & & & & \\
 & & \vdots & & & & & & & & & & & & \\
0 & 0 & \cdots & 1 & 0 & & & & & & & & & & \\
\times & \times & \cdots & \times & \times & \times & \times & \cdots & \times & \times & \times & \times & \cdots & \times & \times \\
 & & & & & 1 & 0 & \cdots & 0 & 0 & & & & & \\
 & & & & & 0 & 1 & \cdots & 0 & 0 & & & & & \\
 & & & & & & & \vdots & & & & & & & \\
 & & & & & 0 & 0 & \cdots & 1 & 0 & & & & & \\
\times & \times & \cdots & \times & \times & \times & \times & \cdots & \times & \times & \times & \times & \cdots & \times & \times \\
 & & & & & & & & & & 1 & 0 & \cdots & 0 & 0 \\
 & & & & & & & & & & 0 & 1 & \cdots & 0 & 0 \\
 & & & & & & & & & & & & \vdots & & \\
 & & & & & & & & & & 0 & 0 & \cdots & 1 & 0
\end{array}\right]
\qquad
B_c = \left[\begin{array}{ccccc}
1 & \times & \times & \cdots & \times \\
0 & 0 & 0 & \cdots & 0 \\
 & & \vdots & & \\
 & & \vdots & & \\
 & & \vdots & & \\
0 & 1 & \times & \cdots & \times \\
0 & 0 & 0 & \cdots & 0 \\
 & & \vdots & & \\
 & & \vdots & & \\
 & & \vdots & & \\
0 & 0 & 0 & \cdots & 1 \\
0 & 0 & 0 & \cdots & 0 \\
 & & \vdots & & \\
 & & \vdots & & \\
 & & \vdots & &
\end{array}\right]
\tag{6.55}
$$

Under the transformation

$$\underline{x} = T\underline{x}_c$$

and the state equation of (6.47) becomes

$$
\begin{aligned}
\dot{\underline{x}}_c &= (A_c + B_c K_c)\underline{x}_c(t) + B_c G \underline{v}(t) \\
\underline{y} &= C\underline{x}_c(t)
\end{aligned}
\tag{6.56}
$$

where $\{A_c, B_c\}$ are as in (6.55) and

$$K_c \equiv KT \tag{6.57}$$

Because there are different ways of choosing the set of linearly independent vectors from the controllability matrix, there are more than one controllability canonical forms. However, when the canonical form is fixed, then the matrix to transform the system into the specific controllability canonical form is uniquely defined.

The uniqueness of the transformation matrix is proved in the following:

Proposition 6.1 *The nonsingular matrix which transforms a system into its canonical controllability form is uniquely defined.*

Proof:

Assume that there are two matrices T_1, T_2 which both are nonsingular and can carry out the following:

$$T_1 = T_2 + \triangle T \tag{6.58}$$

and

$$
\begin{aligned}
T_1^{-1} A T_1 &= A_c, \qquad B_c = T_1^{-1} B \\
T_2^{-1} A T_2 &= A_c, \qquad B_c = T_2^{-1} B
\end{aligned}
\tag{6.59}
$$

Because in both of the cases, (A_c, B_c) are the same and therefore the controllability matrices are the same. After transformation, the controllability matrix of a system becomes

$$Q = T_1 Q_0, \qquad Q = T_2 Q_0 \tag{6.60}$$

where Q_0 is the original controllability matrix.

So we result in $T_1 Q_0 = T_2 Q_0$ or $\triangle T Q_0 = 0$.

This shows that $\triangle T$ must be orthogonal to Q_0 which is of full rank, and henceforth we conclude that $\triangle T = 0$.

□

(b). **Procedure of Pole Assignment Based on Controller Form**

The procedure to assign the desired closed-loop eigenvalues to the system can be carried out on the controller form in (6.55) as follows.

- Choose an appropriate precompensator G to set all the entries marked x in the $\{1st, , (k_1+1)th, (k_1+k_2+1)th, ...\}$ rows of B_c zero. This can always be done by elementary transformations because of the appropriately located 1s in these rows. Let a nonsingular matrix G be chosen to represent these elementary transformations, or

$$B_c G = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ & & \cdots & & \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ & & \cdots & & \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \\ & & \cdots & & \end{bmatrix} \tag{6.61}$$

Define $B_c G = B_o^o$, $B_c = B_o^o G^{-1}$ and also define $\tilde{K}_c = G^{-1} K_c$, $\;K_c = G\tilde{K}_c$ so

$$B_c K_c = B_c G \tilde{K}_c \;=\; B_o^o \tilde{K}_c \tag{6.62}$$

$$= \begin{bmatrix} \tilde{k}_{11} & \tilde{k}_{12} & \cdots & \tilde{k}_{1n} \\ 0 & 0 & \cdots & 0 \\ & \vdots & & \\ \tilde{k}_{21} & \tilde{k}_{22} & \cdots & \tilde{k}_{2n} \\ 0 & 0 & \cdots & 0 \\ & \vdots & & \\ \tilde{k}_{m1} & \tilde{k}_{m2} & \cdots & \tilde{k}_{mn} \\ 0 & 0 & \cdots & 0 \\ & \vdots & & \end{bmatrix}$$

- A method of getting a desired characteristic polynomial

$$f_c(s) = s^n + \beta_1 s^{n-1} + \cdots + \beta_n \tag{6.63}$$

is the following:

1. Choose the $2^{nd}, 3^{rd}, \cdots$ rows of K_c to make all elements in the $\{(k_1 + 1), (k_1 + k_2 + 1), \cdots\}$ rows zero except for the $\{\{k_1 + 1, k_1\}, \{k_1 + k_2 + 1, k_1 + k_2\} \cdots\}$ elements, which should be made equal to 1.
2. Now choose the first row of K_c to make the first row of $A_c - B_o^o K_c$ equal to $\begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{bmatrix}$.

Then clearly we shall have

$$\begin{aligned} p(s) &= \det(sI - A_c - B_o^o \tilde{K}_c) = \det(sI - A_c - B_c G \tilde{K}_c) \qquad (6.64) \\ &= \det(sI - A_c - B_c K_c) \\ &= \det(sI - A_c - T^{-1} BKT) \\ &= \det(sI - A - BK) \end{aligned}$$

Remark 6.12: This gives only the sufficient condition for K to achieve the pole assignment. If, by using this method, the pole assignment cannot be solved with the bounded norm condition, then this does not imply that the desired closed-loop eigenvalues assignment is not possible.

□

(c). Direct Method of Pole Assignment

By first transforming the system into its controller canonical form and then carrying out the pole assignment is one way of achieving the desired pole assignment. Another way of achieving pole assignment is by using general linearisation methods

[Kar. & Gia.,2]. A special case of the general linearisation methods is the Dyadic Feedback Controller method. If the system matrix A is cyclic, and the system (A, B) is controllable, then there exists at least one vector $\underline{q}$ such that $\{A, B\underline{q}\}$ is controllable.

Proposition 6.2 [Won.,1] *If (A, B) is controllable and A is cyclic, there exists at least one vector $\underline{q}$ such that $(A, B\underline{q})$ is controllable. In fact, it holds that "almost any" $m \times 1$ vector $\underline{q}$ will suffice.*

□

For these systems, the pole assignment of the MIMO system can be reduced to a SISO pole assignment problem. Therefore, the simple relation between the coefficients of the open-loop characteristic polynomial and the desired closed-loop characteristic polynomial can be obtained as shown in (6.7). Because the closed-loop system characteristic polynomial $f_c(s)$ is:

$$f_c(s) = \det(sI - A + BK) \tag{6.65}$$

then clearly the coefficients of the closed-loop characteristic polynomial depend on $k_{i,j}$ in a multilinear way as long as $l \geq 2$ [Kar. & Gia.,2]. If we write K in the form

$$K = UL = \begin{bmatrix} \underline{u} & \bar{U} \end{bmatrix} \begin{bmatrix} \underline{l}^T \\ \bar{L} \end{bmatrix} = \underline{u}\underline{l}^T + \bar{U}\bar{L} \tag{6.66}$$

where $U \in R^{l \times q}, K \in R^{q \times n}$ and $\bar{U}, \bar{L}$ are arbitrarily fixed matrices of dimensions $l \times (q-1)$ and $(q-1) \times n$. If $\bar{U}\bar{L} = 0$ then $K = \underline{u}\,\underline{l}^T$ which is a dyadic feedback. Now the closed-loop characteristic polynomial becomes

$$\begin{aligned} f_c(s) &= \det(sI - A - BK) \\ &= \det[sI - (A + B\bar{U}\bar{L}) - B\underline{u}\underline{l}^T] \\ &= \det(sI - \tilde{A} - \underline{b}_u \underline{l}^T) \end{aligned} \tag{6.67}$$

where $\tilde{A} = A + B\bar{U}\bar{L}$ and $\underline{b}_u = B\underline{u}$. So the nonlinear problem is reduced to a linear one, which is equivalent to a SISO system problem, i.e.:

$$\begin{aligned} \dot{\underline{x}} &= \tilde{A}\underline{x} + \underline{b}_u u(t) \\ y &= \underline{c}\underline{x} \end{aligned} \tag{6.68}$$

with state feedback

$$u(t) = \underline{l}^T \underline{x} + v(t) \tag{6.69}$$

If the system $(\tilde{A}, \underline{b}_q)$ is controllable, the eigenvalues of the matrix $(\tilde{A} + \underline{b}_q \underline{k}^T)$ can be assigned arbitrarily and so are the eigenvalues of the matrix $(A+BK)$. Assuming the open-loop characteristic polynomial as

$$f_o(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n \tag{6.70}$$

and the desired closed-loop characteristic polynomial is

$$f_c(s) = s^n + \beta_1 s^{n-1} + \cdots + \beta_n \tag{6.71}$$

then

$$\underline{\beta}^T = \underline{\alpha}^T - \underline{k}^T Q(A, \underline{b}_q) T(\underline{\alpha}) \tag{6.72}$$

where the controllability matrix $Q(A, \underline{b}_q)$ is defined as

$$\begin{aligned} Q(A, \underline{b}) &= \begin{bmatrix} \underline{b}_q & A\underline{b}_q & \cdots & A^{n-1}\underline{b}_q \end{bmatrix} \\ &= \begin{bmatrix} B\underline{q} & AB\underline{q} & \cdots & A^{n-1}B\underline{q} \end{bmatrix} \\ &= \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \left[\begin{array}{c|c|c|c} \begin{matrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{matrix} & & & \\ \hline & \begin{matrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{matrix} & & \\ \hline & & \ddots & \\ \hline & & & \begin{matrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{matrix} \end{array}\right] \end{aligned} \tag{6.73}$$

Denote

$$\begin{bmatrix} \begin{array}{c|c|c|c} q_1 & & & \\ q_2 & & & \\ \vdots & & & \\ q_m & & & \\ \hline & q_1 & & \\ & q_2 & & \\ & \vdots & & \\ & q_m & & \\ \hline & & \ddots & \\ \hline & & & q_1 \\ & & & q_2 \\ & & & \vdots \\ & & & q_m \end{array} \end{bmatrix} \equiv T(\underline{q}). \tag{6.74}$$

Based on the procedures for carrying out closed-loop pole assignment, the pole mobility of the closed-loop eigenvalues with respect to the bounded gain state feedback is studied in the next section.

6.3.3 Pole mobility of MIMO system under bounded state feedback

In the SISO case, the norm bound on the state feedback gain is defined to be the l_2 norm. A natural extension of the l_2 norm for the vector case to the matrix case is the Frobenius norm. So in MIMO systems, we assume that the norm of the state feedback gain matrix K is bounded by the Frobenius norm. The definition of the Frobenius norm of a matrix $A = [\alpha_{ij}]$ is:

$$||A||_f = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij}^2} \tag{6.75}$$

In the following we study the closed-loop pole mobility of the MIMO system under the Frobenius norm bounded state feedback, $||K||_f \leq \gamma$. Because there are different ways of pole assignment, the pole mobility of the closed-loop system depends on the specific method employed. In the sequel, a sufficient closed-loop pole region is first obtained based on the pole assignment using the controller canonical form. A necessary region is then obtained by using the dyadic feedback for systems when A is cyclic.

(a). **Sufficient Assignable Closed Loop Pole Region**

From section 6.3.2, the feedback gain matrix can be obtained for the desired closed-loop characteristic polynomial in such a way that the nonzero rows $\tilde{\underline{k}}_i^T$, $i = 2, 3, ..., m$ in

$$B_c K T = B_o^o \tilde{K}_c = \begin{bmatrix} \tilde{k}_{11} & \tilde{k}_{12} & \cdots & \tilde{k}_{1n} \\ 0 & 0 & \cdots & 0 \\ & & \cdots & \\ \tilde{k}_{21} & \tilde{k}_{22} & \cdots & \tilde{k}_{2n} \\ 0 & 0 & \cdots & 0 \\ & & \cdots & \\ \tilde{k}_{m1} & \tilde{k}_{m2} & \cdots & \tilde{k}_{mn} \\ 0 & 0 & \cdots & 0 \\ & & \cdots & \end{bmatrix} = \begin{bmatrix} \tilde{\underline{k}}_1^T \\ 0 \\ \vdots \\ \tilde{\underline{k}}_2^T \\ 0 \\ \vdots \\ \tilde{\underline{k}}_m^T \\ 0 \\ \vdots \end{bmatrix} \tag{6.76}$$

are chosen to make the closed-loop system matrix as

$$A_c - B_o^o \tilde{K}_c = \tag{6.77}$$

$$\left[\begin{array}{ccccccccccccccccc} \times - \tilde{k}_{11} & \times - \tilde{k}_{12} & \cdots & \times - \tilde{k}_{1,i-1} & \times - \tilde{k}_{1i} & & & & & & & & & & & & \times - \tilde{k}_{1n} \\ 1 & 0 & \cdots & 0 & 0 & & & & & & & & & & & & \\ 0 & 1 & \cdots & 0 & 0 & & & & & & & & & & & & \\ & & \cdots & & & & & & & & & & & & & & \\ 0 & 0 & \cdots & 1 & 0 & & & & & & & & & & & & \\ \hline 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & & 0 & 0 & \cdots & 0 & 0 \\ \hline & & & & & 1 & 0 & \cdots & 0 & 0 & & & & & & & \\ & & & & & 0 & 1 & \cdots & 0 & 0 & & & & & & & \\ & & & & & & & \cdots & & & & & & & & & \\ & & & & & 0 & 0 & \cdots & 1 & 0 & & & & & & & \\ & & & & & & & & & & \ddots & & & & & & \\ \hline 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & & 1 & 0 & 0 & \cdots & 0 & 0 \\ \hline & & & & & & & & & & & & 1 & 0 & \cdots & 0 & 0 \\ & & & & & & & & & & & & 0 & 1 & \cdots & 0 & 0 \\ & & & & & & & & & & & & & & \cdots & & \\ & & & & & & & & & & & & 0 & 0 & \cdots & 1 & 0 \end{array}\right]$$

or they satisfy

$$\tilde{k}_{ij} = -\times_{ij}, \quad j \neq i+1, i = 2, 3, ..., m, \quad j = 1, 2, ..., m \tag{6.78}$$

$$\tilde{k}_{ij} = -\times_{ij} + 1, \quad j = i+1, i = 2, 3, ..., m \tag{6.79}$$

and in terms of Frobenius norm, the matrix $B_c K T$ should satisfy

$$\|B_c K T\|_f = \sqrt{\|\tilde{\underline{k}}_1^T\|_2^2 + \sum_{i=2}^{m} \sum_{j=1, j\neq i+1}^{n} \times_{ij}^2 + \sum_{i=2, j=i+1}^{m} (\times_{ij} - 1)^2} \tag{6.80}$$

where $\times_{ij}, i = 2, 3, ..., m, j = 1, 2, ..., n$ of matrix A_c. So the vector $\tilde{\underline{k}}_1$ satisfies

$$\begin{aligned} \|\tilde{\underline{k}}_1\|_2 &= \sqrt{(\|B_c K T\|_f)^2 - \sum_{i=2}^{m} \sum_{j=1, j\neq i+1}^{n} \times_{ij}^2 - \sum_{i=2, j=i+1}^{m} (\times_{ij} - 1)^2} \\ &\leq \sqrt{\gamma^2 \cdot \|B_c\|_f^2 \cdot \|T\|_f^2 - \sum_{i=2}^{m} \sum_{j=1, j\neq i+1}^{n} \times_{ij}^2 - \sum_{i=2, j=i+1}^{m} (\times_{ij} - 1)^2} \\ &\equiv \gamma_k \end{aligned} \tag{6.81}$$

From (6.77), the norm relation between the closed-loop characteristic polynomial vector, the first row of matrix A_c and the gain vector $\underline{\tilde{k}}_1$ stands as

$$\underline{\beta} = \underline{\tilde{k}}_1 - \underline{\times}_1 \tag{6.82}$$

and therefore

$$\|\underline{\beta}\|_2 = \|\underline{\tilde{k}}_1\|_2 + \|\underline{\times}_1\|_2 \leq \gamma_k + \sqrt{\sum_{i=1}^{n} \times_{1i}^2} \equiv \gamma_\beta \tag{6.83}$$

Remark 6.13: From (6.81) and (6.83), it is clear that the transformation T introduced in transforming the system into controller form affects the norm bound on the coefficients of the closed-loop characteristic polynomial. When $\|T\|_f$ is big, the bound γ_β on the coefficient of the closed-loop characteristic polynomial will be relaxed and the results subsequently obtained for the root distribution will also be conservative.

□

Having obtained the norm bound on the coefficient vector of the closed-loop characteristic polynomial, the maximal reachable closed-loop poles of the system can be obtained.

Theorem 6.8 *For a controllable system (A, B), if the state feedback gain is constrained by Frobinus-norm as $\|K\|_f \leq \gamma$, then all the eigenvalues of any closed-loop system will have all its roots in the circle D*

$$D = \{s \in C : |s| \leq 1 + \gamma_\beta\} \tag{6.84}$$

where $\gamma_\beta = \gamma_k + \sqrt{\sum_{i=1}^{n} \times_{1i}^2}$.

□

If the closed-loop poles are all stable, then they are necessarily within the γ_β-Prime region which is defined in Definition (5.2).

Theorem 6.9 *For a controllable system (A, B), if the state feedback gain is constrained by Frobinus-norm as $\|K\|_f \leq \gamma$, then all the eigenvalues of any stable closed-loop system will have all its roots in the γ_β-Prime region.*

□

(b). **Maximal Assignable Pole Region**

If a system has A cyclic, then as shown in section (6.3.2), the MIMO pole assignment problem can be transformed to a scalar case by dyadic state feedback.

The relation between the coefficients of the open-loop characteristic polynomial and the closed-loop characteristic polynomial and the gain matrix of the resulting scalar pole assignment problem is presented as in (6.72), or

$$\underline{\beta}^T = \underline{\alpha}^T - \underline{k}^T Q(A, \underline{b}) T(\underline{\alpha}) \tag{6.85}$$

where $\underline{b} = B\underline{q}$, $K = \underline{q}\underline{k}^T$, and the controllability matrix Q is defined as

$$\begin{aligned} Q(A, \underline{b}) &= \begin{bmatrix} \underline{b} & A\underline{b} & \cdots & A^{n-1}\underline{b} \end{bmatrix} \\ &= \begin{bmatrix} B\underline{q} & AB\underline{q} & \cdots & A^{n-1}B\underline{q} \end{bmatrix} \\ &= \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} T(\underline{q}) \end{aligned} \tag{6.86}$$

and $T(\underline{\alpha})$ is the characteristic Toeplitz matrix of A.

$$T(\underline{\alpha}) = \begin{bmatrix} \alpha_{n-1} & \alpha_{n-2} & \alpha_{n-3} & \cdots & \alpha_1 & 1 \\ \alpha_{n-2} & \alpha_{n-3} & \alpha_{n-4} & \cdots & 1 & 0 \\ & & \cdots & & & \\ \alpha_1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}. \tag{6.87}$$

From condition (6.85), the following inequalities hold true

$$\begin{aligned} \|\underline{\beta}^T\|_2 &= \|\underline{\alpha}^T - \underline{k}^T Q(A, \underline{b}) T(\underline{\alpha})\|_2 \\ &\leq \|\underline{\alpha}^T\|_2 + \|\underline{k}\|_2 \cdot \|[BAB \cdots A^{n-1}B]\|_f \cdot \left\|T(\underline{q})\right\|_f \cdot \|T(\underline{\alpha})\|_f \end{aligned} \tag{6.88}$$

Because $K = \underline{q}\underline{k}^T$, so the Frobenius norm of the matrix satisfies

$$\|K\|_f = \|\underline{q}\underline{k}^T\|_f = \|\underline{q}\|_2 \cdot \|\underline{k}\|_2 \tag{6.89}$$

and the Frobenius norm of the matrix

$$\left\|T(\underline{q})\right\|_f = m \times \|\underline{q}\|_2 \tag{6.90}$$

so inequality in (6.88) becomes

$$\begin{aligned} \|\underline{\beta}^T\|_2 &\leq \|\underline{\alpha}^T\|_2 + m \cdot \|\underline{k}\| \cdot \|[BAB \cdots A^{n-1}B]\|_f \cdot \|\underline{q}\|_2 \cdot \|T(\underline{\alpha})\|_f \\ &= \|\underline{\alpha}^T\|_2 + m \cdot \|K\|_f \cdot \|Q(A, B)\|_f \cdot \|T(\underline{\alpha})\|_f \\ &\leq \|\underline{\alpha}^T\|_2 + m \cdot \gamma \cdot \|Q(A, B)\|_f \cdot \|T(\underline{\alpha})\|_f. \end{aligned} \tag{6.91}$$

Define the norm bound on the coefficients of the closed-loop characteristic polynomial as γ_c, or

$$\gamma_c \equiv \|\underline{\alpha}^T\|_2 + m \cdot \gamma \cdot \|Q(A, B)\|_f \cdot \|T(\underline{\alpha})\|_f \tag{6.92}$$

Based on the norm bound on the coefficients of the closed-loop characteristic polynomial, the maximal necessary region for the poles to be assignable can be established as in section 5.1.3.

Theorem 6.10 *Let (A, B) be a controllable pair and let the matrix A be cyclic; further assume that the state feedback gain K is bounded by the Frobenius norm, $\|K\|_f \leq \gamma$. A symmetric set of eigenvalues is not assignable as closed-loop eigenvalues of $A+BK$, if they are outside the γ_c-Prime region as defined in Definition (5.2), where $\gamma_c = \|\underline{\alpha}^T\|_2 + m \cdot \gamma \cdot \|Q(A, B)\|_f \cdot \|T(\underline{\alpha})\|_f$.*

□

Remark 6.14: The inequality presented in equation (6.88) is the most general result for all the possible choices of $\underline{q}^T = [\ q_1 \quad q_2 \quad \dots \quad q_m\ ]^T$. A different choice of $\underline{q}$ will give rise to different controllability matrix $Q(A, B\underline{q})$ and therefore a different norm. It is possible to choose some $\underline{q}$ such that the exact maximal norm of the matrix $Q(A, B)T(\underline{q})T(\underline{\alpha})$ can be calculated. However, this will inevitably involve a numerical optimisation approach and it is not pursued here.

□

A comparison of the results obtained with these two different approaches will be given in the example below.

6.3.4 Example

In the following we give an example.

We study the following system

$$\underline{\dot{x}} = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix} \underline{x} + \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{u} \tag{6.93}$$

which has an open-loop characteristic polynomial

$$f_o(s) = s^3 + 6s^2 + 11s + 6$$

and the open-loop poles are $s = -1, -2, -3$.

We first investigate the closed-loop pole assignment via controller-form and then we use dyadic feedback.

(a) <u>Controller-form method</u>: The controllability matrix of the system is

$$Q(A, B) = \begin{bmatrix} 1 & 0 & -1 & -1 & 1 & 4 \\ 1 & 0 & -2 & 0 & 4 & 0 \\ 0 & 1 & 0 & -3 & 0 & 9 \end{bmatrix} \tag{6.94}$$

searching the controllability matrix $Q(A, B)$ from left to right to find 3 independent vectors and rearrange then in the order $\begin{bmatrix} \underline{b}_1 & A\underline{b}_1 & A^2\underline{b}_1 & \dots & A^2\underline{b}_2 \end{bmatrix}$ which gives P,

$$P = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6.95}$$

and a transformation matrix T, which transforms the original system into the controller-form, can be obtained as

$$T = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad T^{-1} = \begin{bmatrix} -1 & 2 & -1 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6.96}$$

by which the system can be transformed into the following controller-form:

$$A_c = T^{-1}AT = \left[\begin{array}{cc|c} -3 & -2 & 1 \\ 1 & 0 & 0 \\ \hline 0 & 0 & -3 \end{array}\right]$$

$$B_c = T^{-1}B = \left[\begin{array}{cc} 1 & -1 \\ 0 & 0 \\ \hline 0 & 1 \end{array}\right]$$

so we can choose $G = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and

$$B_o^o = B_c G = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Set

$$K_c = KT = \begin{bmatrix} \tilde{k}_{11} & \tilde{k}_{12} & \tilde{k}_{13} \\ \tilde{k}_{21} & \tilde{k}_{22} & \tilde{k}_{23} \end{bmatrix}$$

then

$$B_o^o K_c = \begin{bmatrix} \tilde{k}_{11} & \tilde{k}_{12} & \tilde{k}_{13} \\ 0 & 0 & 0 \\ \tilde{k}_{21} & \tilde{k}_{22} & \tilde{k}_{23} \end{bmatrix}$$

The gains $\tilde{k}_{2j}, j = 1, 2, 3$ are selected to make the matrix $A_c + B_o^o K_c$ satisfy the following condition

$$A_c + B_o^o K_c = \begin{bmatrix} -3 + \tilde{k}_{11} & -2 + \tilde{k}_{12} & \tilde{k}_{13} \\ 1 & 0 & 0 \\ \tilde{k}_{21} & \tilde{k}_{22} & -3 + \tilde{k}_{23} \end{bmatrix} = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{6.97}$$

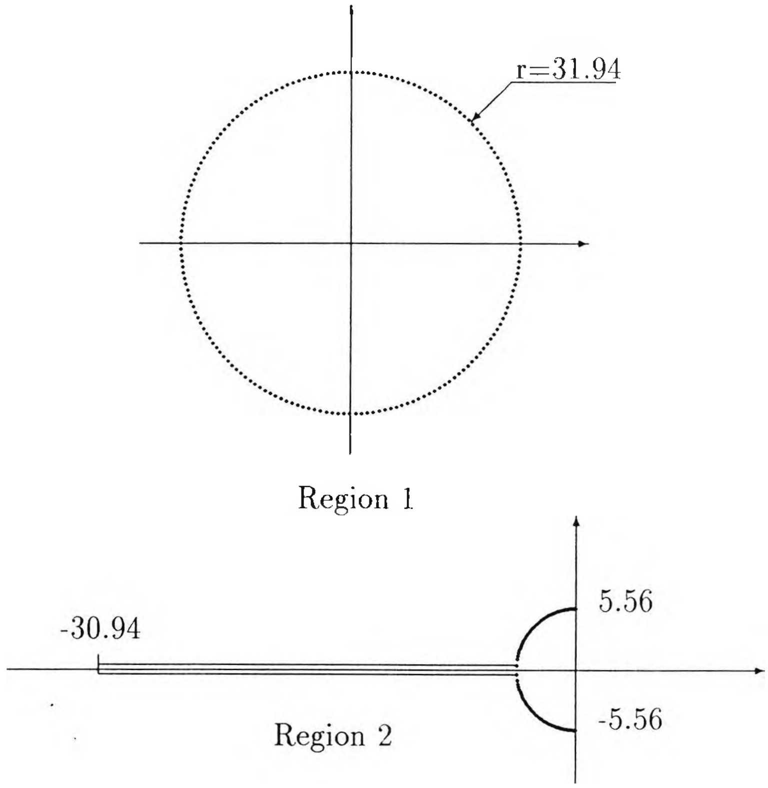


Figure 6.3: Necessary closed-loop pole regions via controller-form

where $\begin{vmatrix} \beta_1 & \beta_2 & \beta_3 \end{vmatrix}$ are the coefficients of the desired closed-loop characteristic polynomial.

So following equation (6.81), we can obtain the norm of $\tilde{\underline{k}}_1$ as

$$\|\tilde{\underline{k}}_1\|_2 \leq \gamma_{\tilde{k}} = \sqrt{30 \cdot \gamma_{\tilde{k}}^2 - 10} \tag{6.98}$$

which demonstrates clearly that the norm bound of the gain matrix must satisfy $\gamma_k \geq .333$ for this particular example.

The coefficients of the possible assignable closed-loop characteristic polynomial must satisfy

$$\|\underline{\beta}\|_2 \leq \|\tilde{\underline{k}}_1\|_2 + \sqrt{\sum_{j=1}^{n} \times_{1j}^2} = \sqrt{14} + \sqrt{30\gamma_{\tilde{k}}^2 - 10} \equiv \gamma_\beta \tag{6.99}$$

Deploying Theorems (6.8), (6.9) we can establish the following necessary closed-loop pole regions.

Region 1: The poles of the closed-loop system with the controller norm bounded by $\|K\|_f \leq \gamma_k$ lie in the circle D which has its centre at the origin and a radius $r = 1 + \gamma_\beta$ and $\gamma_\beta = \sqrt{14} + \sqrt{30\gamma_{\tilde{k}}^2 - 10}$.

Region 2: If all the closed-loop poles are stable then they are necessary confined by the γ_β-Prime region as defined in Definition (5.2).

If we take $\gamma_k = 5$, $\gamma_\beta = 30.94$, then the regions are plotted in Figure (6.3).

(b) Dyadic feedback method: Set $K = \underline{q}\underline{l}^T = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 & l_3 \end{bmatrix}$. The vector $\underline{q}$ should be chosen to make the controllability matrix $Q(A, B\underline{q})$ to be of full rank, i.e.

$$\text{rank}\left(Q(A, B\underline{q})\right) = \text{rank}\left(\begin{bmatrix} q_1 & -q_1 - q_2 & q_1 + 4q_2 \\ q_1 & -2q_1 & 4q_1 \\ q_2 & -3q_2 & 9q_2 \end{bmatrix}\right) = 3 \qquad (6.100)$$

which gives

$$q_1 \neq 0, q_2 \neq 0 \text{ and } q_2 \neq 2q_1.$$

By applying inequality (6.91) we have

$$\begin{aligned} \|\underline{\beta}\|_2 &\leq \|\underline{\alpha}^T\|_2 + m \cdot \gamma \cdot \|Q(A, B)\|_f \cdot \|T(\underline{\alpha})\|_f \\ &= 13.63 + 220.06\gamma_k \end{aligned}$$

If we take $\gamma_k = 5$, then $\|\underline{\beta}\|_2 \leq 1320.35$. Two similar necessary regions **Region** 1^o and **Region** 2^o have been obtained for the closed-loop poles as in figure (6.4). However, they are substantially larger compared with Region 1 and Region 2.

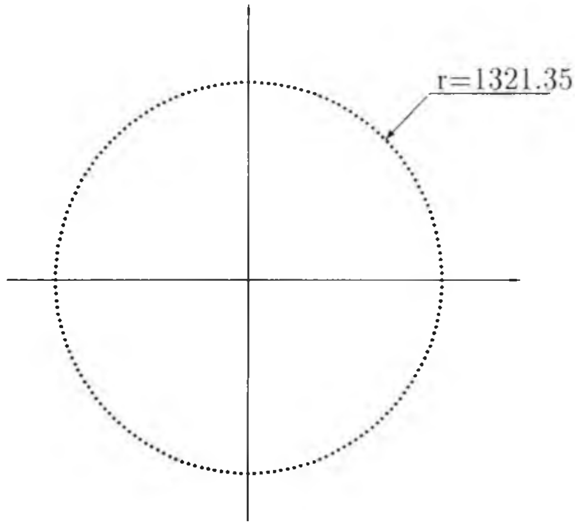
Using the general result of root distribution for third order bounded coefficient polynomials presented in Chapter 6, a much less conservative region can be obtained for the closed-loop poles. The results is shown in **Region** 3^o in figure (6.4).

In Figure (6.5), numerical tests of the closed-loop poles are displayed. It is shown that the closed-loop poles are mainly distributed on the real axis and the theoretical result given is quite conservative.

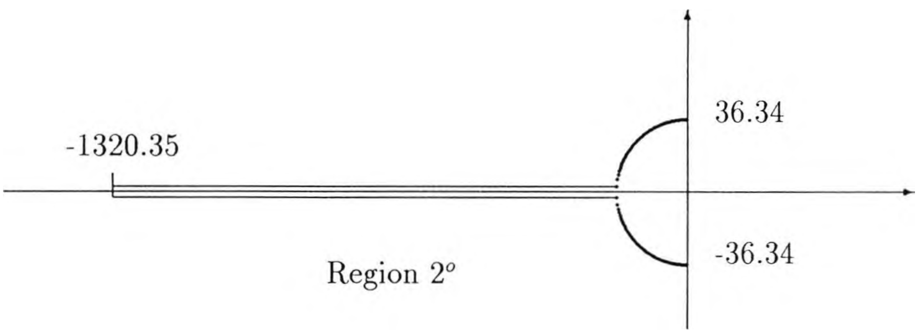
6.3.5 Stabilisability of unstable MIMO systems with bounded norm state feedback

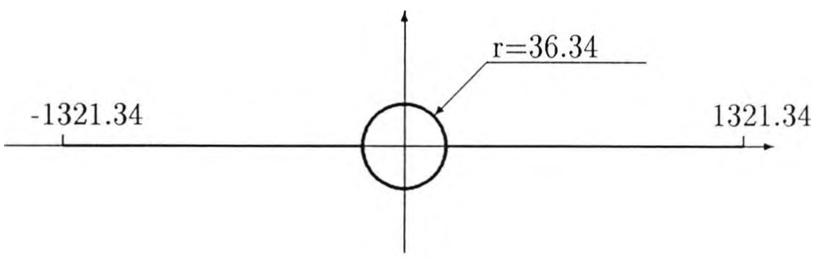
In this subsection, the stabilisability of unstable MIMO systems are studied when the state feedback gain matrix is norm bounded. By a system to be stabilisable we mean that there exists a state feedback satisfying the given bound such that the open-loop unstable MIMO system can be stabilised. Here only the controllable systems (A, B) with A being cyclic are studied.

The assumption that the system matrix A is cyclic is not too restrictive in the sense that for a controllable pair (A, B), almost all $A + BK$ is cyclic with K arbitrarily chosen.



Region 1°





Region 3°

Figure 6.4: Necessary closed-loop pole regions via dyadic feedback

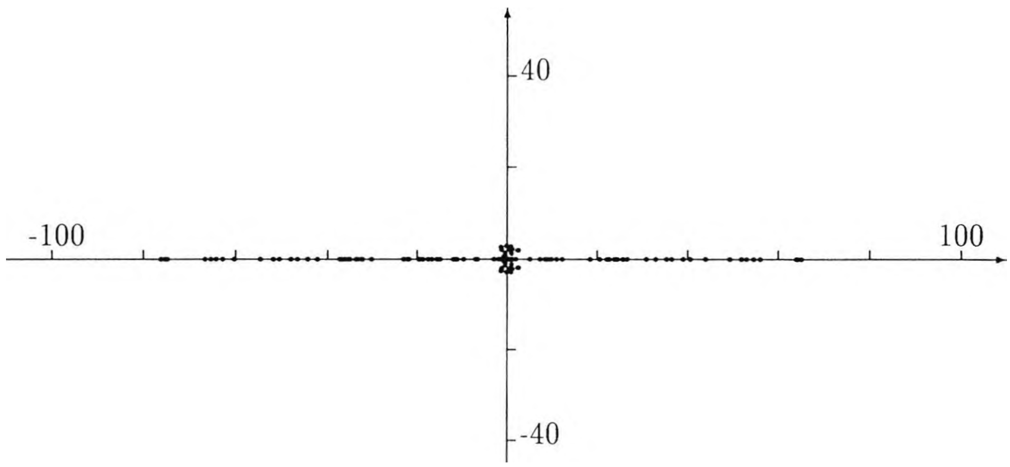


Figure 6.5: Numerical test of the closed-loop distribution

Lemma 6.1 [Kai.,1] *For a controllable pair, any closed-loop system with almost any feedback K, $A + BK$ is cyclic. Obviously, this is true when K is norm bounded.*

□

Because the system (A, B) is controllable and matrix A is cyclic, the stabilisation of the unstable MIMO system can be transformed into a SISO stabilisation problem as described in Section 6.3.2. Given the system

$$\begin{aligned} \dot{\underline{x}} &= A\underline{x}(t) + B\underline{u}(t) \\ \underline{y} &= C\underline{x}(t) \end{aligned} \tag{6.101}$$

and with a state feedback law

$$\underline{u} = K\underline{x}(t) + G\underline{v}(t) \tag{6.102}$$

where the state feedback gain satisfies the condition

$$\|K\|_f \leq \gamma \tag{6.103}$$

By using dyadic state feedback, the relation between the coefficients of the open-loop characteristic polynomial and the closed-loop characteristic polynomial is given by

$$\underline{\beta}^T = \underline{\alpha}^T - \underline{k}^T Q(A, \underline{b}) T(\underline{\alpha}) \tag{6.104}$$

where the controllability matrix $Q(A, \underline{b})$ is defined as

$$\begin{aligned} Q(A, \underline{b}) &= \begin{bmatrix} \underline{b} & A\underline{b} & \cdots & A^{n-1}\underline{b} \end{bmatrix} \\ &= \begin{bmatrix} B\underline{q} & AB\underline{q} & \cdots & A^{n-1}B\underline{q} \end{bmatrix} \\ &= \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} T(\underline{q}) \end{aligned} \tag{6.105}$$

and $T(\underline{q})$ is defined in equation (6.74). If we now set

$$\underline{\delta} = \underline{k}^T Q(A, \underline{b}) T(\underline{\alpha}) \tag{6.106}$$

then we have

$$\|\underline{\delta}\|_2 = m \cdot \gamma \cdot \|Q(A, B)\|_f \cdot \|T(\underline{\alpha})\|_f \tag{6.107}$$

(a). Necessary stabilisable region

Let the the set of desired stable eigenvalues of the closed-loop system be denoted by $\Lambda_c = \{\lambda_i, i = 1, n\}$. In the following we look at the necessary condition for an unstable system to be stabilisable and further has Λ_c as its closed-loop eigenvalues. From equation (6.72), we have

$$\|\underline{\alpha}\|_2 \leq \|\underline{\beta}\|_2 + m \cdot \gamma \cdot \|Q(A, B)\|_f \cdot \|T(\underline{\alpha})\|_f \equiv \gamma_n \tag{6.108}$$

Then a disk can be obtained such that any system with open-loop eignevalues outside this disk cannot be stabilised by the bounded state feedback.

Theorem 6.11 *For a given system $(A, \underline{b})$, if the open-loop system has an unstable eigenvalue outside the circle*

$$D_n \equiv \{s \in C; |s| \leq 1 + \gamma_n\} \tag{6.109}$$

where γ_n is defined in (6.108). Then the system is not Λ_c-stabilisable under the bounded state feedback.

Proof:

From equation (6.108), if the closed-loop poles were to be moved to Λ_c with the bounded norm feedback, it is necessary that $\|\underline{\alpha}\|_2 \leq \gamma_n$. However, if the open-loop system has an eigenvalue outside D_n, following Proposition 5.4 the norm of $\underline{\alpha}$ will necessarily be greater than γ_n which is impossible. This completes the proof.

□

(b). Sufficient condition for the existence of stabilising controllers

In this subsection, we study the problem of finding a controller such that the closed-loop can be stabilised. The norm of this controller gives a sufficient bound for the system to be stabilised. A controller can be found by putting all the closed-loop poles at the origin. In the SISO case, the solution is trivial because a unique solution can be found for the controller by solving equation (6.7). In the MIMO case, however, the controller cannot be uniquely defined.

If we place all the roots of a system at the origin, then the system can be considered to be marginally stabilised. Further, we assume that dyadic feedback is used, i.e. $K = \underline{q} \times \underline{l}^T$. For this particular case, the closed-loop characteristic of the system is simply

$$f_c(s) = s^n$$

and correspondingly the coefficient vector is $\underline{\beta} = \underline{0}$. From equation (6.66-6.73), we have

$$\underline{\alpha}^T = \underline{l}^T Q(A, \underline{b}) T(\underline{\alpha}) \tag{6.110}$$

and the gain vector $\underline{l}^T$ can be found. However, it is necessary that the input direction of the dyadic feedback $\underline{q}$ has to be fixed beforehand because it affects $Q(A, \underline{b})$. The controller norm is in turn affected by the input direction since $\|K\| = \|\underline{q}\underline{l}^T\|$. Therefore, choosing different dyadic feedback direction $\underline{q}$ will result in different controller gain $\underline{l}$ and different controller K as well.

6.4 Summary

In this chapter, the problem of closed-loop pole mobility has been addressed, when the state feedback controller is norm bounded: in this case the closed-loop poles of the feedback system cannot be arbitrarily assigned even though the system is assumed to be controllable. By deploying the classical and recent results on the root distribution of bounded coefficient polynomials in Chapter 6, we have established

for SISO systems:

- A region which contains all the poles of the closed-loop system when under bounded norm state feedback. This region is defined in terms of the system matrix $(A, \underline{b})$, the norm bound on the controller gain vector, γ_k. If the desired closed-loop characteristic polynomial has a zero outside this region, then there exists no controller which will give rise to the expected closed-loop characteristic polynomial and at the same time the norm of the controller is bounded by the given value. For systems with low order ($n \leq 3$) a much tighter region has been established.

- A region that contains all the possible stable closed-loop poles with the bounded norm controller. If any desired closed-loop pole is outside this region, the closed-loop system is not attainable using any feedback with the norm bounded by the given value.

- A necessary region has been obtained which contains all the open-loop poles which are to be stabilised by the given bounded norm controller.

for MIMO systems:

- When assigning the closed-loop poles via controller form, a region which contains all the poles of the closed-loop system when under bounded norm state feedback has been established. The region is defined in terms of the system matrix (A, B), and the norm bound on the controller gain vector, γ_k. If the desired closed-loop characteristic polynomial has a zero outside this region, then there exists no controller which will give rise to the expected closed-loop characteristic polynomial and at the same time the norm of the controller is bounded by the given value. Also we have obtained a necessary region which contains the all the stable closed-loop poles with the norm bounded controller.

- When assigning the closed-loop poles using dyadic feedback scheme, a region is established which contains all the possible closed-loop poles with the bounded norm controller. If any desired closed-loop pole is outside this region, the closed-loop system is not attainable using any feedback with the norm bounded by the given value. A tighter region is obtained for low order systems. Further, if we have established a region which contains all the stable closed-loop poles with the bounded controller.

- A necessary region is obtained for the open-loop poles which are possibly stabilisable by the bounded controller if only the dyadic feedback control scheme is allowed.

The regions established here are quite conservative. The conservativeness is introduced mostly by the inequalities used to convert the norm bound on the controller onto the bound on the coefficients of the closed-loop characteristic polynomial.

Chapter 7

ISSUES ON STRUCTURAL SYNTHESIS OF LARGE SCALE SYSTEMS

7.1 Introduction

In all the previous chapters, the properties and property indicators are studied for the systems whose mathematical models are known exactly. In practice, however, working on systems which can be described by exact mathematical models is more than a luxury. The nature of the models on which the analysis is carried out is determined by the problems one poses, the data available and the environment within which one works. For instance, as argued in [Kar.,2], if the control theory is to intervene in process design at early stages, which is believed to be beneficial, then the control theory has to work on systems whose models are ill-defined. By an ill-defined model we mean that a model captures only the main characteristics of a system and the values of the parameters may not be exact. For instance, a steady-state model features only the steady state information while the dynamics of the system is ignored. In general, errors are introduced when systems are represented by ill-defined models. The errors can be classified into two categories: unstructured or parametric. If the model of a system features only the structural information, then the error will be unstructured. Otherwise, the error will be parametric.

A lot of research effort has been devoted to the study of system properties based on ill-defined models. The main interest is to predict the true system properties in exploring the ill-defined models. The problems are tackled in both the frequency domain and the state-space domain. The frequency domain analysis caters for the unstructured errors naturally [Doy. & Ste.,1] [Kar.,2] while state-space domain analysis caters for parametric errors [Rei.,1] [Lin,1] [Mor. & Ste.,1]. The properties which are dependent on the structure of a system will be referred to be the structural or generic properties. For instance the controllability, observability of a system may be determined by the underlying graph structure of a system and thus they have a structural version [Lin,1] known as the structural controllability and structural observability. The structural properties are important in that these properties are generically possessed by all the systems which may have different parametric values, but share the same underlying graph structure; therefore the study of the structural properties is relevant not only to one particular system but to a class of systems. This analysis, on one hand, provides means to detect certain structural characteristics, and on the other hand enables the prediction of solvability of certain control problems and determines lower bounds for the required numbers of inputs, outputs needed for the presence of certain system properties.

In this chapter, we will be dealing mainly with ill-defined models. In Section 7.2, we will first discuss the general issues arising with the modelling of large interconnected systems. We will look into the model characteristics in early process design and outline the desirable features of the control theory in early process design aiming at providing a general framework for structural analysis. In Section 7.3, we concentrate on issues related to predicting the system properties based on the steady-state models. Necessary conditions for the system controllability, stabilisability and system integrity related to the steady-state models are reviewed. In Section 7.4, the notions of generic structured transfer function and finite generic McMillan degree will be introduced. Then methods for evaluating the finite generic McMillan degree will be discussed. Finally in Section 7.5 generic properties of structured transfer function at infinity will be investigated.

7.2 General issues arising with the modelling of large interconnected systems

7.2.1 Model characteristics in early process design

The specific problems arising with the attempt to intervene in early process design with control theory design tools are due to the nature of the models available. At the early design stages, the models available to the control engineers are often very simple, in the sense that the dynamics are not known exactly and the numerical values of the parameters are also inexact. Further, the models of the processes will be of very large dimension. These characteristics of the models available contribute considerably to the difficulties faced in carrying out the expected meaningful intervention in the early process design with control theory and design tools.

Most of the existing control theories apply readily to systems with well defined models either described in the transfer function or state space but fail to work on systems with inexact models. First, it should be mentioned the general models the control theory will have to work on at an early process design stage. Considering the nature of the models, it is strongly argued in [Kar.,2] that the system be described by *External Structural Dynamical Models* for early process design. By External Structural Dynamical Models, we mean the models which are of the transfer function type with the elements being dynamical SISO models whose parameters are not known exactly. They are more appropriate than the exact state space models, the exact transfer function models and the structural state space models used so far, due to the following reasons:

- A state space model requires the knowledge of the exact number of states. In an early process design environment, where the subprocesses may be modelled with varying accuracy models, this implies an updating of the overall state space model every time the subprocess models are modified.
- The models of subprocesses may contains delay units or more general distributive parameter characteristics; such features may be more easily handled in the frequency domain, rather than in the time domain, where the corresponding state space model may become infinite dimensional.
- A transfer function model indicates clearly the structure of interconnections and the possible reasons for interactions. A general state space model obtained by minimal realisation does not reflect the structure of interconnections, since its states are not necessarily associated with the subprocesses. Models with physical variables, which reflect the structure of interconnections may be obtained, but they are not necessarily minimal.

So the tools developed later should work on the External Structural Dynamical Models.

The next main characteristic of early process design models is that the models are under constant modification in terms of both the dynamical complexity and parameter accuracy. This characteristic presents the second challenge to the development of useful tools in control theory for the intervention in early process design.

It is a common practice that the design of a process starts off with a process flowsheet which represents the main components, interconnections, etc. Associated with the flowsheet is the data file with the detailed description of the components, the values of the parameters, interconnections, etc. The mathematical models describing the overall system may have varying complexity in dynamical terms and accuracy in terms of parameter values. For instance, we can associate a Boolean graph representing the existence of the components and the interconnections among the components qualitatively or a quantitative steady-state model representing the system steady-state response when subject to step input.

Let $\mathcal{M}$ denote the whole set of models associated with a given flowsheet. The models of large dimension, structured (and possibly sparse), transfer function matrices with fixed dynamic complexity and variable parameter uncertainty will be denoted as $\{M_i\}$, $M_i \in \mathcal{M}$. All such linear models stem from the same flowsheet and thus they have the same graph structure, which is defined by a Boolean matrix. The family $\{M_i\}$ has two parametric dimensions, the dynamic complexity and the parameter accuracy. By *dynamic complexity* we refer to the order, type of the rational approximation, which is used for each of the functions representing process elements; this complexity may be measured by the ordered set of McMillan degrees of approximations of the process elements. For models with the same dynamic complexity, the term parameter accuracy is used in the standard way. The dynamic complexity and parameter accuracy, introduce some ordering for the set $\mathcal{M}$.

The first and simplest model in $\mathcal{M}$ is the graph model M_g, represented by a Boolean matrix; the only information indicated by such a model is the generic existence of the components and the couplings between the inputs and outputs for almost all frequencies. The next subfamily of models are the steady-state models $\{M_0\}$, and the elements of this subfamily are characterised by their accuracy; such models indicate the steady-state coupling of inputs and outputs in a quantitative way for step inputs. The steady state models express the value of transfer functions at $s = 0$ and their associated graph may differ from that of M_g; conceptually such models differ from the rest in the family, since they express certain aspects of the overall behaviour predicted by the dynamic models.

The rest of the families of models, i.e. $\{M_1\}, \{M_2\}$, etc. represent the families of dominant lag, multi-lag models etc, and $\{M_k\}$ represents a family of well defined dynamic models, used in the final design stages. The evolution of the models is shown in Figure (7.1).

The dynamic complexity of the models should be measured by the McMillan degree which indicates the total number of internal variables associated with a minimal

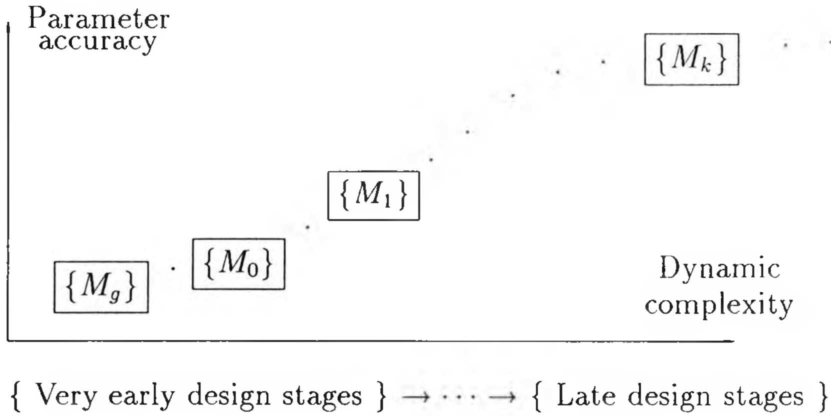


Figure 7.1: Evolutionary development of Early Process Design models

representation of $G(s)$.

The objective is to evaluate design alternatives and make choices using low complexity and accuracy models and some set of comparison criteria. These criteria are based on the properties of the ill-defined models and should have a prediction capability, as far as the properties of higher complexity and accuracy models, which are not available at this stage. Since we deal with families of linear systems, the standard concepts and tools of control theory, such as system properties, property-, design-indicators, invariants etc., are well defined for every model in $\mathcal{M}$. From the large set of properties and invariants of linear systems theory, we have to select those which are most relevant in the context of rough dynamic models. In deciding about the relevance of properties, invariants the following characteristics should be taken into consideration:

(i). *Universality,* that is, if they are defined on M_i, then they are also defined on all $M_{i+j}, j = 1, 2, ...$

(ii). *Computability,* that is there exists a robust algorithm for checking the presence, or computing the values.

(iii). *Parameter insensitivity,* that is their presence is independent of parameter variations which are within the system graph structure.

The important control problems associated with early process design are addressed in the next subsection.

7.2.2 Desirable features of control theory in process design

Conventionally, the design of a chemical process is carried out solely by the chemical engineers according to the chemical-physical laws of the process. The control engineers will be involved at a very late stage of the design. It is argued that an early intervention by the control engineers in the design of a process will be beneficial [Kar.,2] [Mor. & Ste.,1]. This role of control theory differs substantially from their previous roles in the design of a system. Early decisions on

(i). Identification of control quality characteristics of large scale models;

(ii). Identification and sorting out of all possible sets of inputs (and potential actuators) and output (and potential sensors);

(iii). Decomposability, partitioning of large scale control design problems into smaller dimension problems; and finally

(vi). Selection and design of structured, simple control schemes

will not only be desirable, but also be very beneficial from the control point of view in getting a good control structure and finally achieving quality control of the process.

The possible contribution of the control theory listed above has to be complementary to the practical rules, experience and techniques, which are currently available. A short description of the concrete problems involved in the problems and their relevance to the early design of a process is given below.

(i) For a given system with specified inputs and outputs, the problem is to identify the model characteristics, which precondition the existence of a simple control scheme, or exclude undesirable system response characteristics for a given system with specified inputs and outputs. It is of particular importance in choosing alternative systems resulted from selecting different input and output sets. It is required that the selection algorithm can work on large dimension and ill-defined models efficiently. The difficulty lies in the definition of desirable and undesirable system properties and characteristics.

(ii) Although the whole sets of possible inputs and outputs can be listed, not all of the potential inputs and outputs will be used for actuation and measurement purposes due to economical or physical limits. Process considerations and experience may define subsets of input and output variables and a further selection of them has to be done based on the properties of the resulting model. Avoiding decisions which may lead to system models with undesirable properties, is the most important task here. The most important problems are:

- Define some estimates of lower bounds on the number of process inputs and outputs which are essential for an easy control problem;
- Evaluate alternative choices of input and output variables with operability and control quality criteria;
- Define the actuators and sensors scheme such that the resulting model has certain desirable features from the control design viewpoint.

(iii) Most of the process applications are of large dimension in nature and the design techniques are difficult to apply. Reducing the overall design problem to smaller dimension problems by rearrangement of the inputs and output is an important problem. Specific problems are:

- Derive interaction measures among subsystems when the overall system is partitioned into subsystems.
- Examine compensation transformations which may reduce the inherent process coupling, and thus lead to subsystems with small dimensions.

(vi) In order to design a process, the coupling of measurement and actuation variables should be as simple as possible due to reliability considerations. So the structure of the control schemes and their suitability for the control quality specifications have to be defined and assessed. Important issues which arise in the design of such structured control schemes are:

- Derive tests and criteria for determining the range of possible control structures;
- Examine methods for designing the dynamics of the already structured controllers;
- Derive tests for measuring the robustness (under model uncertainties) and integrity (under actuator, sensor failures) of the structured control system and suggest methods for designing robust, high integrity schemes;
- Examine methods which will allow the design of control schemes that may guarantee good performance not only for one, but for many process operating points.

7.3 System properties based on the steady state gain information

As discussed in Chapter 3, useful interaction measure between the inputs and the outputs can be obtained from the steady state gain information. As a matter of fact, some more system properties can be inferred from the steady state gain information, in particular, the relative gain array (RGA), which has been defined in subsection (3.4.1).

Based upon the steady state gain information, the system stabilisability, system controllability, system integrity as well as property robustness can be studied. In process control, a system is usually operating at steady state and the steady state error is required to be zero; therefore in each of the control loops, there should be an integrator. Under this assumption, the stabilisability and controllability of the closed-loop system can be inferred from the steady state gain and/or the RGA matrix [San. & Ath., 1] [Nie.,1] [Gro. Mor. & Hol.,1]. System integrity, which is defined to be the stability of the system when one or more of the system sensors and/or system actuators fail, can also be studied from these information [Mor.,1] [Gro. Mor. & Hol.,1]. Furthermore, useful information concerning system sensitivity when subject to model-reality errors in modelling can be extracted from the steady state RGA information [Gro. Mor. & Hol.,1].

7.3.1 Integral stabilisability

In the following, closed-loop control systems with a specific class of controllers will be studied. The special feature of the class of controllers is that in each of the controller channels, there exists an integrator. Such controllers are termed as integral controllers and the stabilisability of the system is referred to as *integral stabilisability*. Because every loop of the integral controller contains an integrator, it can be decomposed into a matrix of integrators $\frac{k}{s}I$ and a compensator matrix $C(s)$; where k is a positive constant and I is the identity matrix of appropriate dimension. The system has a general configuration as shown in Figure (7.2). Let $H(s) = G(s)C(s)$. Observe that for a proper $G(s)$ the newly defined transfer function $H(s)$ could well be non-proper. In the following we study the closed-loop stability with integral controllers. First the following definition is given.

Definition 7.1 : [Gro. Mor. & Hol.,1] *A system $H(s) = G(s)C(s)$ is called integrally stabilisable if there exists a $k > 0$ such that the closed-loop system in Figure (7.2) is stable and has zero tracking error for all asymptotically constant inputs.*

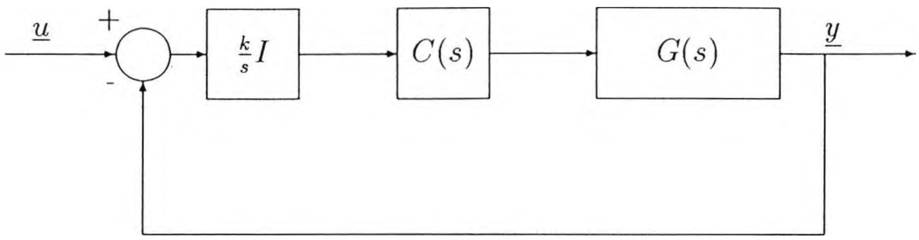


Figure 7.2: General Configuration with Integral Controller

□

A necessary condition for a system to be integrally stabilisable is given by the following result:

Theorem 7.1 [Gro. Mor. & Hol.,1] *Assume that $H(s)$ is a proper rational transfer matrix. $H(s)$ is integrally stabilisable only if det($H(0)$) > 0.*

□

For SISO systems, the above theorem becomes necessary and sufficient.

Corollary 7.1 *Assume that $h(s)$ is a proper rational transfer matrix. $h(s)$ is integrally stabilisable if and only if $h(0) > 0$.*

□

Specifically when the system is controlled in a decentralised manner, i.e. the compensator $C(s)$ has a diagonal form, then concerning the stability of the individual control loops and the overall stabilisability of the system, we have the following result.

Theorem 7.2 [Nie.,1] [Gro. Mor. & Hol.,1] *Consider the control system of the configuration shown in Figure (7.2) with plant transfer matrix $G(s)$ and diagonal compensator $C(s)$. The following conditions hold true:*

- *$G(s)$ is open-loop stable;*
- *$H(s) = G(s)C(s)$ is rational and proper;*
- *Any one-variable control system, obtained from the multivariable system by opening any $n-1$ feedback loops, is stable.*

Then there exists no k with $k > 0$, such that the overall system with all the closed-loops operating is stable if

$$\frac{detG(0)}{\Pi_{i=1}^{n} g_{ii}(0)} < 0 \tag{7.1}$$

Furthermore, the above condition becomes necessary and sufficient for SISO or 2×2 system.

□

Remark 7.1: The above result provides an important guideline for designing a system in a decentralised manner. It is a common practice that on designing a large system, the overall design is first broken it into smaller subsystems, usually SISO systems, and then stabilising controllers are designed for each of the subsystems. As it is pointed out in the above theorem, if the decentralised system were to be integrally stabilisable, it is necessary that both the individual small systems be stable and

$$\frac{\det G(0)}{\Pi_{i=1}^{n} g_{ii}(0)} \geq 0.$$

Otherwise, one is put into the dilemma of having either unstable individually designed loops or an unstable overall system. So the above theorem can be employed to be a necessary test for system stability.

□

From Theorem 7.2, the system integrity can also be studied. By integrity, we mean that a system will remain stable even though some of the system sensors and actuators fail in operation.

Definition 7.2 : *The system in Figure (7.2) is jth-sensor failure sensitive if the overall system is integrally stabilisable but the reduced system with the failed jth sensor removed ($k_j = 0$) is not.*

□

From the above definition, if a system is not jth-sensor failure sensitive, then the system can be suitably set up in such a way that even if the jth sensor fails, the stability of the remaining system can still be maintained with an appropriate choice of the controller gain i.e. the system integrity is ensured. Otherwise, the system will go unstable as soon as the jth-sensor fails regardless how the controller is set up.

Similarly, jth-actuator failure sensitive can be defined though more consideration is needed here. When an actuator fails, the corresponding manipulated input will lose control. Worse still, the control signal will assume a wrong value. For example, if a phneumatic valve gets stuck and remains fully open all the time. Generally, this will affect all the loops of the system. For simplicity, it is assumed that only the jth output is left uncontrolled when the control over u_i is lost.

Definition 7.3 : *The system shown in Figure (7.2) is jth-actuator failure sensitive if the overall system is integrally stabilisable while the reduced system with the jth actuator and the jth sensor removed and with the controller appropriately reduced is not integrally stabilisable.*

□

The conditions for system jth-sensor failure sensitive and jth-actuator failure sensitive follows from theorem 7.2 directly.

Theorem 7.3 [Gro. Mor. & Hol.,1] *Assume that $H(s)$ is rational, proper and integrally stabilisable. The system is jth-sensor failure sensitive if $det(G^{jj}(0)C^{jj}(0)) < 0$, where $G^{jj}(0)$ denotes the remaining matrix with jth column and jth row removed.*

□

Theorem 7.4 [Gro. Mor. & Hol.,1] *Assume that $H(s)$ is rational, proper and integrally stabilisable. The system is jth-actuator failure sensitive if $det(G^{jj}(0)C^{jj}(0)) < 0$.*

□

With respect to decentralised control scheme, the overall stabilisability is related to the stabilisability of the individual control loop in the following way through the relative gain array (RGA) which was defined in subsection (3.4.1) as

$$\text{RGA} = [\lambda_{ij}](G), \qquad i, j = 1, ..., n \tag{7.2}$$

Theorem 7.5 [Gro. Mor. & Hol.,1] *If $\lambda_{jj}(G) < 0$, then for any compensator $C(s)$ with the properties*

- *$G(s)C(s)$ is proper;*
- *$C_{jl} = C_{lj} = 0, \forall l \neq j,$*

the closed-loop system shown if Figure (7.2) has at least one of the following properties

- *The closed-loop system is unstable;*
- *loop j is unstable with all the other loops opened;*
- *the closed-loop system is unstable as loop j is removed.*

□

The above theorem has important practical implications. It has been a common practice in process design that each loop of a system is firstly individually designed and the stability of the individual loops are ensured, then the system is integrated together. However, unless the diagonal elements of the steady state gain matrix, $\lambda_{jj}(G) > 0$, the stability of the so designed overall system can not be achieved. Further, for systems with $\lambda_{jj}(G) > 0$ and an appropriately designed diagonal controller, an overall stabilisable system will be ith-sensor failure sensitive or ith-actuator failure sensitive.

Remark 7.2: The conditions given in the above theorem is only a necessary condition. Further, for some systems with $H(s)$ improper, even the condition does not hold, there still may exist $k > 0$ such that the overall system is stabilisable as demonstrated by the following example.

□

Example 7.1: Given a SISO system $g(s) = \frac{-5(s+1)}{2s+1}$ which is controlled by an PI controller, $kc(s)/s = k(s+1)/s$. From the definition $h(s) = \frac{-5(s+1)(s+1)}{2s+1}$ which is improper.

The characteristic equation is

$$(2 - 5k)s^2 + (1 - 10k)s - 5k = 0$$

and the system is stable for $0.4 < k < \infty$, despite $h(0) < 0$. Such system will be termed as conditionally stable.

In practice, systems which are only conditionally stable are difficult to tune because the range of the parameters for the system to be conditionally stable is not only difficult to decide but also subject to constant change as the operating point, system environment, parameters are changing. To remedy the above weakness, a stronger definition is given on stability and for which sufficient condition can be obtained from the steady state information.

7.3.2 Integral controllability and steady state information

If a system $H(s)$ as defined in Figure (7.2) is improper, the integral stabilisability of the system cannot be determined from the steady state gain information. Further concepts are needed to remedy the weakness of integral stabilisability.

Definition 7.4 : [Gro. Mor. & Hol.,1] *The open-loop stable system $H(s)$ is called integrally controllable if there exists a $k^* > 0$ such that the closed-loop system shown in Figure (7.2) is stable for all values of k varying in the interval $0 < k \leq k^*$ and has zero tracking error for asymptotically constant disturbances.*

□

Remark 7.3: The difference of this definition from that of system stabilisability is that a range of positive gains starting from zero will make the system stable rather than any exact value. A practical consequence of this definition is that integrally controllable systems can be tuned on-line starting with a very low gain for which stability is guaranteed, and then increasing the gain until acceptable performance is achieved.

□

Conditions for a system to be integrally controllable are given below.

Theorem 7.6 [Gro. Mor. & Hol.,1] *The stable system $H(s)$ is integrally controllable if all the eigenvalues of $H(0)$ lie in the open right half of the complex plane. The system $H(s)$ is not integrally controllable, if any of the eigenvalues of $H(0)$ lie in the open left half of the complex plane.*

□

A system which is only conditionally integrally stabilisable will not be integrally controllable as stated in the next corollary.

Corollary 7.2 [Gro. Mor. & Hol.,1] *If $det(H(0)) < 0$, then $H(s)$ is not integrally controllable.*

□

The relationship between the integral controllability and integral stabilisability is that all integrally controllable systems are integrally stabilisable, but not vice versa. The case that integrally stabilisable systems are not integrally controllable happens only when an even number of eigenvalues of $H(0)$ are in the left half plane because a necessary condition for integrally stabilisability is if $\det(H(0)) > 0$.

However, for SISO systems, the two concepts are equivalent.

Theorem 7.7 [Gro. Mor. & Hol.,1] *Any proper rational system $h(s)$ which is integrally stabilisable is also integrally controllable.*

□

Integrally controllable systems have an advantage over integrally stabilisable systems in that the loops can be tuned starting from very small gains. On the other hand for integrally stabilisable systems, it is difficult to compute the stable range for the gain.

For an integrally controllable system, the stability of the remaining system resulting from taking out the loops with failed sensors or failed actuators can be studied as for integrally stabilisable systems. Some further definitions on system sensor failure tolerant and system actuator failure tolerant are given.

Definition 7.5 *: The system $H(s)$ shown in Figure (7.2) is j sensor failure tolerant if both the overall system and the reduced system with the j-th sensor removed are integrally controllable, and the system is actuator failure tolerant if both the complete system and the reduced system with the jth actuator and the jth sensor removed are integrally controllable.*

□

Conditions for system to be sensor failure tolerant or to be actuator failure tolerant are direct consequences of the definition and theorem 7.6.

Corollary 7.3 [Gro. Mor. & Hol.,1] *The system with $H(s)$ rational is jth sensor failure tolerant if all the eigenvalues of $H(0)$ and $H^{jj}(0)$ are in the open right half complex plane. It is not jth sensor failure tolerant if any of the eigenvalues of $H(0)$ of $H^{jj}(0)$ are in open left half complex plane.*

□

Corollary 7.4 [Gro. Mor. & Hol.,1] *The system with $H(s)$ rational is jth actuator failure tolerant if all the eigenvalues of $H(0)$ and $G^{jj}(0), C^{jj}(0)$ are in the open right* *half complex plane. It is not jth actuator failure tolerant if any of the eigenvalues of* $H(0)$ *of* $G^{jj}(0), C^{jj}(0)$ *are in open left half complex plane.*

□

For small dimensional systems, the RGA of the system also provides some information on the system sensor and/or actuator failure tolerance.

Theorem 7.8 [Gro. Mor. & Hol.,1] *Let* $G(s)$ *be a* 2×2 *system. If* $\lambda_{jj}(G) > 0$, *then there exists a diagonal compensator* $C(s)$ *such that* $H(s)$ *is 1st-sensor and/or actuator failure tolerant and 2nd-sensor and/or actuator failure tolerant.*

□

For 3×3 systems, a weaker result is given by:

Corollary 7.5 [Gro. Mor. & Hol.,1] *Let* $G(s)$ *be a* 3×3 *system. If* $\lambda_{jj}(G) > 0, j = 1, 2, 3$. *If a diagonal compensator* $C(s)$ *can be found such that:*

- $H(s) = G(s)C(s)$ *is integrally controllable and*
- $h_{jj}(0) > 0, j = 1, 2, 3,$

then the closed-loop system shown in Figure (7.2) is jth-sensor and/or actuator failure tolerant, for $j = 1, 2, 3$.

□

So the eigenvalues of the steady state gain matrix $H(0)$ provides information not only on the integral controllability of a system, but also on the integrity of a system when some of the sensors and/or actuators fail.

7.3.3 Robustness and the RGA

One of the most important properties a system should possess is robustness against model-reality differences. A controller designed for a given model, when applied to the real system, may result in performance which is quite different from specifications. So it is important to analyse the sensitivity of system stability with respect to the model-reality difference.

From the previous subsections, it is shown that the integral stabilisability of a system is closely related to the diagonal elements of the RGA of the system. So the sensitivity of the elements of the RGA when the elements of the steady state gain matrix are subject to uncertainties is an important measure of the system sensitivity. The relative sensitivity of the elements of the RGA is related to the elements of the steady state matrix in the following way.

Theorem 7.9 [Gro. Mor. & Hol.,1] *Consider the $n \times n$ transfer function matrix G with its inverse $\tilde{G} = G^{-1}$ and its associated RGA. Relative changes in the g_{ij}'s and $\tilde{g}_{ij}$'s are related as*

$$\frac{d\lambda_{ij}}{\lambda_{ij}} = (1 - \lambda_{ij})\frac{dg_{ij}}{g_{ij}}$$

and

$$\frac{d\lambda_{ij}}{\lambda_{ij}} = \frac{\lambda_{ij} - 1}{\lambda_{ij}}\frac{d\tilde{g}_{ij}}{\tilde{g}_{ij}}$$

□

So it is clearly shown that the sensitivity to error in the modelling of g_{ij} increases when the value λ_{ij} is far away from unity.

System design when subject to model reality difference has been studied quite extensively [Doy. & Ste.,1]. Depending on the domain the analysis is carried out, the models used and the nature of the model-reality difference, different methods can be used to represent the model-reality difference. One of the most popular ways is to use the l_1, l_2 or l_∞ norm. The l_1, l_2 or l_∞ norms are defined as follows. Given a vector $\underline{1}$

$$\begin{aligned} \|\underline{1}\|_1 &\equiv \max_i l_i \\ \|\underline{1}\|_2 &\equiv \sqrt{\sum_{i=1}^{n} l_i^2} \\ \|\underline{1}\|_\infty &\equiv \sum_{i=1}^{n} l_i \end{aligned}$$

The induced norm of a matrix is defined as

$$\|A\| \equiv \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

and the condition number of a matrix as

$$\gamma(A) \equiv \|A\|\|A^{-1}\|.$$

Then for the l_2 norm, the following stands

$$\begin{aligned} \|A\|_2 &= \bar{\sigma}(A) \\ \|A^{-1}\|_2 &= \underline{\sigma}(A) \\ \gamma(A) &= \bar{\sigma}(A)/\underline{\sigma}(A). \end{aligned}$$

where $\bar{\sigma}(A)$ and $\underline{\sigma}(A)$ are the maximum and the minimum of the singular values of matrix A.

A matrix with $\gamma(A)$ close to 1 is called well-conditioned while a matrix with $\gamma(A) \gg 1$ is called ill-conditioned. From the analysis by [Wil.,1] [Gol. & van Loan,1] etc. the eigenvalues of a matrix that is well-conditioned are less sensitive to perturbations than the eigenvalues of an ill-conditioned one.

Because the eigenvalues of the steady state gain matrix are indicators of integral controllability, so the sensitivity of the eigenvalues of the steady-state matrix to model reality difference is also the stability robustness. The maximal allowable difference between the model and reality to retain integral controllability is defined by:

Theorem 7.10 [Gro. Mor. & Hol.,1] *Assume the model $\tilde{G}(s)$ with controller $C(s)$ is integrally controllable. Then the plant $G(s)$ with the same controller $C(s)$ is integrally controllable if the difference is bounded as*

$$\|G - \tilde{G}\|_2 < \frac{1}{\gamma(\tilde{G})}\|\tilde{G}\|_2. \tag{7.3}$$

where $\gamma(\tilde{G})$ is the condition number of $\tilde{G}$ and $\| \bullet \|_2$ denotes the Euclidean norm.

□

Clearly, the system is more error tolerant when $\gamma(\tilde{G})$ is close to 1.

The sensitivity of a system is invariant with respect to input output scaling. The condition number of a matrix, however, is not and so neither is the relation presented in (7.3). Realising that the RGA of a system is independent of the input-output scaling, efforts have been made to explore the relationship between the sensitivity and the RGA. In view of equation (7.3), the system is least sensitive to model-reality error when the condition number attains its minimum when subject to optimal scaling. Denote the minimum condition number as $\gamma^*(G)$, then it is related to the l_1 norm of the RGA for 2×2 systems as:

Theorem 7.11 [Gro. Mor. & Hol.,1] *For a 2×2 transfer matrix G, the minimum condition number γ^* is given by*

$$\gamma^* = \|RGA\|_1 + \sqrt{\|RGA\|_1 - 1} \tag{7.4}$$

□

Therefore, from equation (7.4), γ^* is bounded above by $2 \times \|RGA\|_1$ and approached $\|RGA\|_1$ as it becomes large. So large elements in the RGA result in large

$\|RGA\|_1$ and in turn implies large γ^*. So when $|\lambda_{jj}|$ is large, the system will be very sensitive to model-reality error.

However, there is no proved simple relationship between the minimum condition number and the l_1 norm of RGA.

7.4 Generic McMillan degree of structured rational matrices

7.4.1 Introduction

In this section, the generic McMillan degree of a rational matrix is investigated. From the definition, the McMillan degree of a rational matrix can be calculated from the orders of the denominators of the matrix in Smith-McMillan form. So algorithms can be designed to first transform the rational matrix into Smith-McMillan form by using unimodular transformations and then find the sum of the orders of the denominators. As pointed out by MacFarlane, Karcanias, etc., this method is impractical in terms of computations to obtain the Smith-McMillan form. An alternative has been suggested by MacFarlane & Karcanias [MacF. & Kar.,1], that is to obtain the pole polynomial as the least common multiple of the minors of all orders. The order of this least common multiple gives the McMillan degree. This method may also be used for computation of the generic form of the Smith-McMillan form or the given structure system, as well as the unstable McMillan degree. This method does not require the transformation of the rational matrix into Smith-McMillan form and computationally is more practical.

It is a fact that, in the early design stages, the exact values of the parameters in the elements of the transfer function are not known exactly. Yet, it is desirable to have some knowledge on the McMillan degree since it indicates the complexity of the system. Given the structure of the transfer function matrix and the type of the non-zero entries of the matrix, the evaluation of the McMillan degree of such systems will be termed as generic evaluation of the McMillan degree and the McMillan degree will be termed as generic McMillan degree of the element structured uncertain model. The poles and zeros of a system provide important information in the study of the multivariable root-locus design, system properties at infinity, etc. When the parameters of the system are not given exactly, we also wish to calculate the generic degree of the poles and zeros at infinity.

In this section we first define the genericity of transfer function matrices, the

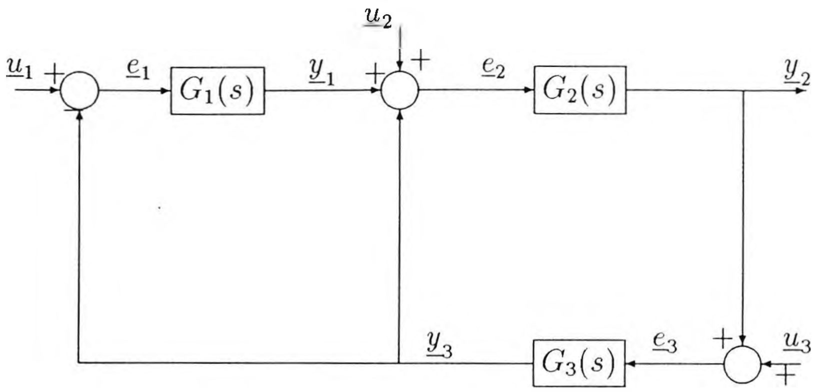


Figure 7.3: A general system

McMillan degree and poles, zeros both finite and at infinity. Then, methods for evaluating the generic McMillan degree both finite and infinite are presented. Finally we assess the methods which are proposed here.

7.4.2 Generic structured transfer functions and generic McMillan degree

As pointed in section (7.2.1), the models at early design stage provide only some structural information about the system. The structural information includes the fixed poles in the transfer function matrix, the orders of the nonzero entries of the matrix and some repeated patterns due to specific dynamic units which are modelled with certain known dynamic complexity. But otherwise, the values of the parameters of the transfer functions are not known exactly. Assume that all the entries of the transfer function are proper rational fractions. For a system of dimension $m \times l$, a structured overall transfer function matrix $H(s)$ is $H(s) \in R_{pr}^{m \times l}[s]$. The notion of structured transfer function matrices of variable complexity is demonstrated by the following example.

Example 7.2: Consider the system of Figure (7.3) which can be represented by the aggregated model as in Figure (7.4), where the three subsystems are described by:

$$\underline{y}_1 = G_1 \underline{u}_1$$
$$\underline{y}_2 = G_2 \underline{u}_2$$
$$\underline{y}_3 = G_3 \underline{u}_3$$

and for the aggregated model, we have

$$\underline{y} = G(I + FG)^{-1}\underline{u}$$

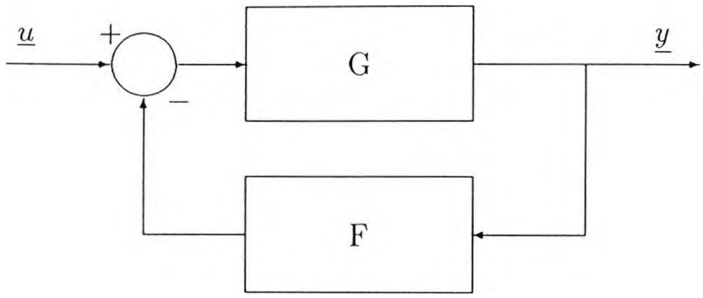


Figure 7.4: Aggregated model

where $\underline{u} = [\underline{u}_1^T \underline{u}_2^T \underline{u}_3^T]^T, \underline{y} = [\underline{y}_1^T \underline{y}_2^T \underline{y}_3^T]^T$ and

$$G = \begin{bmatrix} G_1 & & 0 \\ & G_2 & \\ 0 & & G_3 \end{bmatrix}$$

$$F = \begin{bmatrix} 0 & 0 & -I \\ 0 & 0 & I \\ -I & 0 & 0 \end{bmatrix}$$

The matrix F represents the interconnections while the matrix G represents the dynamics of the system. Assume the subsystems are all of dimension 2×2 and

$$G_1 = \begin{bmatrix} \frac{1}{(s+1)(s+5)(s+10)} & \frac{6}{(s+1)(s+2)(s+5)(s+10)} \\ \frac{1}{(s+2)(s+5)(s+10)} & \frac{1}{(s+1)(s+2)(s+5)(s+10)} \end{bmatrix}$$

$$G_2 = \begin{bmatrix} \frac{2.5}{s+1} & 0 \\ 0 & \frac{0.75}{s+2} \end{bmatrix}$$

$$G_3 = \begin{bmatrix} \frac{0.12(s-1)}{(s+1)(s+2)} & \frac{s}{(s+1)(s+2)} \\ \frac{6.0}{(s+1)(s+2)} & \frac{s-2}{(s+1)(s+2)} \end{bmatrix}$$

Following the procedure in the previous section, the structure of the system can be modelled by the Boolean model as

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

and by the steady-state model as

$$G_0 = \begin{bmatrix} 0.02 & 0.01 & 0 & 0 & 0 & 0 \\ 0.001 & 0.001 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.375 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.06 & 0 \\ 0 & 0 & 0 & 0 & 3.0 & -1 \end{bmatrix}$$

If we model the system by taking only the dominant pole, then we have

$$G_1 = \begin{bmatrix} \frac{0.02}{s+1} & \frac{0.01}{s+1} & 0 & 0 & 0 & 0 \\ \frac{0.002}{s+2} & \frac{0.001}{s+1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{2.5}{s+1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{0.75}{s+2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-0.06(s-1)}{s+1} & \frac{s}{s+1} \\ 0 & 0 & 0 & 0 & \frac{3.0}{s+1} & \frac{s-2}{s+1} \end{bmatrix}$$

and so on.

□

The above example demonstrates that the composite system is modelled in terms of the subsystem models $\{G_1, G_2, G_3\}$ and the interconnection structure matrix F. Such models are referred to as *Internal Progenitor Models* (IPM) [Kar.,3] and always lead to "structured" transfer function matrices for the composite system (when the subsystem models are fixed). A structured transfer function matrix is a transfer function with certain elements fixed to zero, some elements being constant and other elements expressing the simple dynamics of the subsystems. The transfer function of a composite system for which the underlined interconnection matrix F is not known or not explicitly stated, is called an *External Progenitor Model* [Kar.,3] and the only evidence in the interconnection structure is that provided by the structured nature of the overall transfer function. Structured transfer function matrices frequently arise as models in the Early Process Design Stages [EPIC] and some of their basic problems associated with their structural characteristics will be considered here. It will be assumed throughout the following, that the structured transfer function is given but the underlined matrix F is not known. Some of the issues discussed next are of particular importance for large dimension matrices.

An example of a 3×3 structured proper rational transfer function matrix, $H(s)$ is given below

$$H(s) = \begin{bmatrix} A_1^2 A_2 & A_1 & A_3 \\ A_3 & A_2^2 & A_1 A_2 A_3 \\ A_1 & A_4 & A_1^2 \end{bmatrix} \tag{7.5}$$

where the elements $A_i, i = 1, ..., 4$ are repeated patterns representing, for instance, constant terms, the first or second order dynamics, $A_1 = \frac{c_1}{s+a_1}, A_2 = \frac{c_2}{s^2+b_1 s+b_2}, \ldots,$ where the a_1, b_1, b_2 etc are fixed and the c_i are constants which take generic values. By using partial fraction expansion we can decompose the transfer function matrix in the following manner:

$$\begin{aligned} H(s) &= \begin{bmatrix} A_1^2 A_2 & A_1 & A_3 \\ A_3 & A_2^2 & A_1 A_2 A_3 \\ A_1 & A_4 & A_1^2 \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} A_1^2 & A_1 & 0 \\ 0 & 0 & A_1 \\ A_1 & 0 & A_1^2 \end{bmatrix}}_{H_1(s)} + \underbrace{\begin{bmatrix} A_2 & 0 & 0 \\ 0 & A_2^2 & A_2 \\ 0 & 0 & 0 \end{bmatrix}}_{H_2(s)} + \underbrace{\begin{bmatrix} 0 & 0 & A_3 \\ A_3 & 0 & A_3 \\ 0 & 0 & 0 \end{bmatrix}}_{H_3(s)} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & A_4 & 0 \end{bmatrix}}_{H_4(s)} \end{aligned} \tag{7.6}$$

and the matrices $H_i(s)$ will be called simple structured matrices.

In general, we define the structured and simple structured matrices as:

Definition 7.6 : *The* structured transfer function matrix *of a system is a transfer function matrix whose entries consist of elementary dynamical terms. The elementary dynamical terms represent the basic dynamics of the subsystems which may appear in more than one entries depending on the structure of the system. If the structured matrix consists of entries with the same elementary dynamical term, then it is called* simple.

□

A structured transfer function matrix can always be decomposed into a set of simple structured transfer function matrices by use of partial fraction expansion method for each of the dynamic terms.

Remark 7.4: If $\{\lambda_1, \lambda_2, ..., \lambda_p\}$ are the fixed pole locations of a structured transfer function, then $H(s)$ may always be expressed as

$$H(s) = H_1(s) + H_2(s) + \cdots + H_p(s)$$

where $H_i(s)$ are simple structured transfer functions, corresponding to λ_i fixed pole.

□

For the so-defined structured transfer function matrices, we define the generic McMillan degree as

Definition 7.7 : *The generic McMillan degree of the structured transfer function $H(s) \in R_{pr}^{m\times l}[s]$ is the McMillan degree when the gain parameters of the entries take generic values.*

□

Remark 7.5: In the computation of any minor of a generic rational matrix, there is no pole zero cancellation occurring.

□

By the above definition, Remark (7.4) and the definition of McMillan degree based on the minors [MacF. & Kar.,1], we have the result.

Proposition 7.1 *The generic McMillan degree of the structured transfer function matrix $H(s) \in R_{pr}^{m\times l}[s]$ is equal to the sum of the generic McMillan degree of the matrices $H_1(s), H_2(s), \cdots$. That is, if $\delta_{gm}(H_i)$ denotes the generic McMillan degree of a structured transfer function matrix H_i, then*

$$\delta_{gm}(H) = \delta_{gm}(H_1) + \delta_{gm}(H_2) + \cdots. \tag{7.7}$$

Proof: Without loss of generality, we assume that $H_i(s)$ and $H_j(s)$ are two simple structured transfer function matrices of $H(s)$ which are associated with the fundamental dynamics $A_i = \frac{1}{s+\lambda_i}$ and $A_j = \frac{1}{s+\lambda_j}$, $\lambda_i \neq \lambda_j$. We prove that if the generic McMillan degrees of the matrices $H_i(s), H_j(s)$ are $\delta_{gm}(H_i), \delta_{gm}(H_j)$, then the contributions of the terms $A_i = \frac{1}{s+\lambda_i}$ and $A_j = \frac{1}{s+\lambda_j}$ towards the generic McMillan degree of the structured transfer function matrix $H(s)$ are exactly $\delta_{gm}(H_i)$ and $\delta_{gm}(H_j)$.

Because the generic McMillan degrees of the matrices $H_i(s)$ and $H_j(s)$ are $\delta_{gm}(H_i)$ and $\delta_{gm}(H_j)$, the least common multiples of all minors of the matrices $H_i(s)$ and $H_j(s)$ are $\frac{\alpha}{(s+\lambda_i)^{\delta_{gm}(H_i)}}$ and $\frac{\beta}{(s+\lambda_j)^{\delta_{gm}(H_j)}}$, respectively. Under the genericity assumption, there will be no pole-zero cancellations among the constituent parts in the determinants of the minors; therefore the terms in the least common multiple of all minors of the matrix $H(s)$ due to terms $A_i = \frac{1}{s+\lambda_i}$ and $A_j = \frac{1}{s+\lambda_j}$ are of the forms $\frac{\alpha'}{(s+\lambda_i)^{\delta_{gm}(H_i)}}$ and $\frac{\beta'}{(s+\lambda_j)^{\delta_{gm}(H_j)}}$, respectively, i.e., the contributions towards the generic McMillan degree of the structured transfer function matrix due to the fundamental dynamics $A_i = \frac{1}{s+\lambda_i}$ and $A_j = \frac{1}{s+\lambda_j}$ are $\delta_{gm}(H_i)$ and $\delta_{gm}(H_j)$, respectively. This proves the Proposition.

□

Remark 7.6: The evaluation of the generic McMillan degree of a structured transfer function matrix is reduced to finding the generic McMillan degrees of the simple structured matrices $H_i(s)$.

□

In the following, we look into the methods of computing the generic McMillan degree of the simple structured transfer function matrices. First we define the concepts of order, path and weight.

Definition 7.8 : *Given a simple structured matrix* $H_i(s) \in R_{pr}^{m\times l}, m \leq l$, *the* order *of an entry in the matrix is the power of the fundamental dynamics, a* path *is a sequence of m elements selected from the matrix with no two elements from the same column or from the same row. The* length of a path *is the number of non-zero elements in the path. The* weight of a path *is defined to be the sum of the orders of the elements in the path. The maximal weight of all the independent paths of the matrix is denoted as* $\Gamma(H_i)$.

□

Remark 7.7: The constant terms of the structured transfer function matrix do not contribute to the weight. From the definition of generic McMillan degree, the constant terms do not contribute to the generic McMillan degree. So the constant terms are equivalent to fixed zero elements.

□

Remark 7.8: The zero and constant entries do not contribute to the weight of a path.

□

Remark 7.9: A path with the maximal weight does not necessarily have to be the longest path, as displayed by the following simple structured matrix

$$H(s) = \begin{bmatrix} A^4 & A \\ A^2 & 0 \end{bmatrix}.$$

For this matrix there are two paths,

$$\begin{cases} h_{11} \rightarrow h_{22} \text{ with length 1 and weight 4} \\ h_{12} \rightarrow h_{21} \text{ with length 2 and weight 3} \end{cases}$$

□

In the following we study the simple structured matrices $H_i(s)$. Because only the non-zero dynamic elements need to be considered, and the non-zero entries represent the same dynamic unit with different orders, for the simplicity of notation, we use the orders of the dynamic of the entries only. For example, in (7.6) the matrix $H_1(s)$ is simplified to be

$$H_1(s) = \begin{bmatrix} A_1^2 & A_1 & 0 \\ 0 & 0 & A_1 \\ A_1 & 0 & A_1^2 \end{bmatrix} \longrightarrow I_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 2 \end{bmatrix}.$$

In general, a map can be defined between a simple structured matrix $H_i(s)$ and an integer matrix I_i such that the entries of the integer matrix correspond to the orders of the entries in $H_i(s)$.

Concerning the relationship between the generic McMillan degree and the weight of paths, we have the following result.

Proposition 7.2 *The generic McMillan degree of the simple structured matrix $H_i(s)$ is equal to the maximal weight,*

$$\delta_{gm}(H_i) = \Gamma(H_i).$$

Proof: For a given simple structured matrix $H_i(s)$ with a maximal weight $\Gamma(H_i(s))$, we prove that the order of the least common multiple of all the minors of all orders of the matrix $H_i(s)$ is generically $\Gamma(H_i(s))$ and so is the generic McMillan degree. Without loss of generality, we assume $A_i = \frac{1}{s+\lambda_i}$. We first prove that $\delta_{gm}(H_i) \geq \Gamma(H_i)$. Assume that the length of a maximum weight path is l_Γ, then from the definition of path, it is clear that there exists a minor of order $l_\Gamma \times l_\Gamma$ which contains this path. The denominator of the determinant of the minor is generically $(s+\lambda_i)^{\Gamma(H_i)}$ because there is no cancellation involved under the genericity assumption, therefore, the order of the least common multiple of all minors is greater or equal to $\Gamma(H_i)$. Second we prove that if $\delta_{gm}(H_i) > \Gamma(H_i)$, then there exists an independent path whose weight w^* satisfies $w^* > \Gamma(H_i)$. If $\delta_{gm}(H_i) > \Gamma(H_i)$, then there must exists at least a minor whose denominator has an order exactly as $\delta_{gm}(H_i)$. Because there is no cancellation among the terms of the constituent parts to the minor, there exists a set of elements selected from different rows and columns whose product is in the form $\frac{\alpha}{(s+\lambda_i)^{\delta_{gm}(H_i)}}$. In other words, there exists a path which has a weight w^* and $w^* = \delta_{gm}(H_i) > \Gamma(H_i)$. This contradicts to the assumption that $\Gamma(H_i)$ is the maximal weight. So we have $\delta_{gm}(H_i) = \Gamma(H_i)$.

□

Having defined the generic McMillan degree we investigate the methods of calculating the generic McMillan degree in the next section.

7.4.3 Methods for evaluating the generic McMillan degree

In this subsection, we present three ways of evaluating the generic McMillan degree of a structured rational transfer function matrix and the merits and disadvantages of these methods are assessed.

(a): Direct method

The first method is to find all the possible independent paths for the $m \times l$ matrix and the corresponding weight is calculated for each of the path. The maximum among all the weights gives the generic McMillan degree.

This method involves very heavy computations when the order of the system is high. For a matrix of dimension $m \times m$, the number of paths equals to the factorial of m, which is $m!$. The fixed zeros in the structured matrix are not taken into account. So even when the matrix is of sparse nature, the number of calculations involved will be the same as in the case of non-sparse matrices.

(b): Leading order coefficient matrix method

For a given integer matrix $A \in N^{m \times l}$, where N represents integer matrices, define the *weight of a column* k_i to be

$$k_i \equiv \text{ the maximum of the } i\text{th column of } A \tag{7.8}$$

and the *complexity* of A as $\delta(A)$ to be

$$\delta(A) = \sum_{i=1}^{l} k_i \tag{7.9}$$

We can define row weight and complexity in a similar manner. In general, we can always write

$$A = A_{lc} S + L \tag{7.10}$$

where

$$\begin{aligned} S &\equiv \operatorname{diag}\{k_i, i = 1, 2, ..., l\} \\ A_{lc} &= \text{the maximum-column-weight coefficient matrix} \end{aligned}$$

or the leading (column) weight matrix, of A

L = the smaller remaining elements

The matrix A_{lc} is a Boolean matrix and by rank of A_{lc} we mean the usual structural rank [Rei.,1]. The next Proposition follows.

Proposition 7.3 *The generic McMillan degree of an $m \times l$ simple matrix with $m \geq l$ is equal to the sum of the elements in S, i.e. $\delta(A) = \sum_{i=1}^{l} k_i$ if the rank of the leading weight matrix is l; otherwise it is always less than $\sum_{i=1}^{l} k_i$.*

Proof: We show that the maximal weight is less or equal to the complexity. If matrix A_{lc} is of full column rank, then they are equal. From the definition of the complexity and the maximal weight of an integer matrix, it is clear that

$$\Gamma(A) \leq \delta(A) \tag{7.11}$$

Inequality may hold because the entries with values equal to the weights of columns may appear in the same row, therefore they will not appear in the same path.

If the matrix A_{lc} has rank l (nonsingular), then a path can be selected from different rows and columns with length l and maximal weight $\sum_{i=1}^{l} k_i$. So it follows from Proposition (7.1) that the generic McMillan degree of A is $\sum_{i=1}^{l} k_i$.

□

Applying a dual argument, a dual statement can be made if the leading row coefficient matrix of A, A_{mr} is of rank m. However, when the leading row or column coefficient matrix is rank deficient, nothing can be said on the generic McMillan degree. Of course, it is possible to select a column from matrix L to replace the columns in A_{lc} which are linearly dependent on the other columns. The criterion to select the columns from L is complicated and it is believed that the criterion is also difficult to implement.

(c): A more effective search algorithm

Since most of the computation time will be spent on summation operations, in the following we present an algorithm aiming at reducing the summation operations. The algorithm is first presented for lower order matrix cases, then it will be extended to the general cases.

(i). **3rd order case**

For the third order matrix case, we look at the following matrix

$$M_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

There are two independent paths which include a_{33} entry as an element, namely,

$$a_{11} \to a_{22} \to a_{33} \text{ and } a_{12} \to a_{21} \to a_{33}$$

To calculate the weights of the two independent paths requires $2 \times 2 = 4$ number of summation operations, those are $a_{11}+a_{22}+a_{33}$ and $a_{12}+a_{21}+a_{33}$. Because both paths contain a_{33} as an element, there is no need to carry out all the summation operations before we can decide which path gives the larger weight. Indeed, we first carry out the summation operations $a_{11}+a_{22}$ and $a_{12}+a_{21}$. By then we know which of the two independent paths should be selected, i.e., we can take either $a_{11} \to a_{22}$ or $a_{12} \to a_{21}$ depending which gives the larger weight. Having chosen the particular route, the maximal weight containing a_{33} entry can be obtained by just adding the value a_{33}. So instead of doing 4 summation operations, only three summation operations are needed, i.e., $a_{11} + a_{22}$, $a_{12} + a_{21}$ and plus one more summation operation either for $(a_{11}+a_{22})+a_{33}$ or for $(a_{12}+a_{21})+a_{33}$. For the purpose of generalisation, we express the number of summation operations as $(2! + 1) = 3$.

The above analysis can be carried over to the cases when the independent paths contain either a_{13} or a_{23}. The number of summation operations needed for both of the cases are also $(2!+1) = 3$. So the total amount of summation operations needed for the 3rd order matrix case is $3 \times (2! + 1) = 9$.

Compare with the original method, the number of summation operations needed accounts only $9/12 = 75\%$, a reduction of 25%. The percentage of summation operations needed will be substantially less when the order of the matrices increases. As a matter of fact, the percentage approaches 0 when the order approaches ∞.

(ii). **4th order case**

Now, we look at a 4th order matrix

$$M_4 = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

We first consider the independent paths which include a_{44} entry as an element. From the definition of independent paths, the rest of the elements of an independent path have to be selected from the submatrix

$$M_{41} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Instead of carrying out all the summation operations of the independent paths selected from this submatrix with the element a_{44}, we select an independent path from all the $3! = 6$ independent paths which gives the maximal weight. Only the weight of this maximum weight path need to be added up with the value of a_{44}. Further, this weight gives the maximal weight for all the independent paths containing the a_{44} entry as an element.

As having been shown earlier, the number of summation operations which is needed for selecting the maximal weight for the submatrix M_{41} is $3 \times (2! + 1) = 9$, the number of summation operations needed to obtain the maximal weight for all the independent paths which contain a_{44} as an element is just $3 \times (2! + 1) + 1 = 10$.

The above analysis can be implemented directly on the cases when the independent paths containing a_{14}, a_{24} or a_{34} as an element. So the number of summation operations needed to give the maximal weight for the 4×4 matrix is $4 \times (3 \times (2! + 1) + 1) = 40$.

The reduction in the number of summation operations needed is $3 \times 4! - 40 = 32$, which constitutes $32/72 = 44.4\%$.

(iii). **General order case**

The analysis in the previous two sections can be directly generalised to the general case. Given an $n \times n$ matrix,

$$M = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

If we fix a_{nn} as one element in the independent paths, the rest of the elements have to be selected from the submatrix

$$M_{n1} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} \\ a_{21} & a_{22} & \cdots & a_{2n-1} \\ & & \ddots & \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n-1} \end{bmatrix}$$

The number of summation operations needed to obtain the maximum weight path from all the independent paths is $(n-1)(1+(n-2)(1+(n-3)(\cdots + 3(1+2!)\cdots)$, and the number of summation operations needed to give the maximal weight among all the independent paths containing the a_{nn} entry as an element requires exactly 1 more summation operation, i.e., $1+(n-1)(1+(n-2)(1+(n-3)(\cdots+3(1+2!)\cdots)$. So the amount of summation operations needed to give the maximal weight for an $n \times n$ matrix is $n(1+(n-1)(1+(n-2)(1+(n-3)(\cdots+3(1+2!)\cdots)$. The expression can be rewritten as

$$\begin{aligned} \Delta' &= n(1+(n-1)(1+(n-2)(1+(n-3)(\cdots+3(1+2!)\cdots) \\ &= n+n(n-1)+n(n-1)(n-2)+n(n-1)(n-2)(n-3)+\cdots+n! \end{aligned}$$

So the percentage of summation operations needed compared with the original method for the general case is

$$\frac{\Delta'}{\Delta} = \frac{n+n(n-1)+n(n-1)(n-2)+n(n-1)(n-2)(n-3)+\cdots+n!}{(n-1)n!} \simeq \frac{1}{n-1}$$

and

$$\frac{\Delta'}{\Delta} \to 0, \text{ as } n \to \infty$$

(iv). **Comparison of the results**

In this subsection, we compare the new algorithm with the original method. The following table gives the number of summation operations needed, the percentage of operations for the new method compared with the original one for different orders of matrices.

order	3	4	5	10	15	20
original method	12	72	480	32659200	1.830×10^{13}	4.622×10^{19}
new method	9	40	205	6235300	2.247×10^{12}	4.180×10^{18}
percentage	75%	55.56%	42.71%	19.09%	12.27%	9.04%

So this method is much better compared with the method proposed in (a).

Remark 7.10: The algorithm can be implemented for finding the minimum weight path with obvious changes.

□

(d). **Diagonalisation procedure**

From the definitions of the path and the weight of a path, it is clear that by exchanging rows and columns, the path and weight will not be changed. So diagonalisation procedures involving only row and column permutations can be used. The diagonalisation procedure can be very beneficial for large dimensional and sparse matrices. Assuming that the system can be diagonalised into two diagonal blocks of dimensions n_1 and n_2, $n_1 + n_2 = n$, and we carry out the search of the maximal weight for each of the matrix using method (a), the total number of independent paths to be searched is $n_1! + n_2!$. Compared with the total number of independent paths in the original matrix $n!$, it will be much smaller. Of course, different methods can be used for searching the maximal weight path of the submatrices.

7.5 Generic properties of generic transfer functions at infinity

The poles and zeros of a system at $s = \infty$ are important in the sense that they provide the information on the system behaviour at $s = \infty$ [Kai.,1] [Var. Lim. & Kar.,1]. There are several problems in which it is important to keep track of the behaviour at $s = \infty$. Poles at $s = \infty$ characterise nonproper systems (or systems with differentiators), as may arise in constructing inverse systems, while the zeros at ∞ are important, for example, in studying the asymptotic behaviour of multivariable root loci, decoupling, etc. For scalar systems with n poles and m zeros, $m < n$, m of the closed-loop poles will converge toward the m finite poles will the remaining $n - m$ poles will converge to the $n - m$ zeros at infinity. A similar conclusion can be made for multivariable systems.

However the definitions given for the poles and the zeros in Section (2.2.6) do not extend to $s = \infty$ as it can be shown that the $R[s]$-unimodular matrices used to transform $N(s)$ into the Smith-form can have both poles and zeros at ∞. So by unimodular transformations, the information at $s = \infty$ will be destroyed. It is important to observe that the pole-zero information of the system is preserved for any finite frequency when under $R[s]-$unimodular transformations but not always the structure at infinity. If we make the bilinear transformation

$$s = \frac{a\lambda + b}{c\lambda + d} \tag{7.12}$$

where $c \neq 0$ and $ad - bc \neq 0$, which will merely transform the complex s plane into itself, then this transformation will move the point at $s = \infty$ to the point $\lambda = -\frac{d}{c}$. The Smith-McMillan form for $H(\lambda)$ will accurately reflect the behaviour of $H(\lambda)$ for all point except those at $s = \infty$. In particular, the Smith-McMillan structure at $\lambda = -\frac{d}{c}$ will accurately reflect that of $H(s)$ at $s = \infty$. If the constants a, b, c, d are chosen as $a = 0 = d, b = c = 1$, then $s = \frac{1}{\lambda}$ and with the substitution $s = \frac{1}{\lambda}$, the pole-zero information at $s = \infty$ can be studied by $\lambda = 0$.

For instance, a simple unimodular matrix

$$U(s) = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} \tag{7.13}$$

with substitution $s = \frac{1}{\lambda}$

$$U(\lambda^{-1}) = \begin{bmatrix} 1 & \lambda^{-1} \\ 0 & 1 \end{bmatrix} \tag{7.14}$$

whose Smith-McMillan for can be calculated as

$$\begin{bmatrix} \lambda^{-1} & 0 \\ 0 & \lambda \end{bmatrix} \tag{7.15}$$

which shows that $U(s)$ has a pole and zero at $s = \infty$.

An alternative way of calculating the pole-zero structure of a system both at finite and infinite frequencies is to use the valuation method [Kai.,1] [Var. Lim. & Kar.,1] which characterizes the Smith-McMillan form directly. Define for a scalar rational function $g(s)$ the discrete valuation at $s = \infty$ by:

$$\begin{aligned} v_\infty(g) &\equiv \text{the } \infty \text{ valuation of } g(s) \\ &= \text{denominator degree - numerator degree} \end{aligned} \tag{7.16}$$

for example,

$$v_\infty\left[\frac{s+1}{(s-1)^2(s+2)}\right] = 2$$

and for matrix case, the definition is extended as,

$$\begin{aligned} v_\infty^{(i)}(g) \equiv {} & \text{the algebraically smallest } \infty \text{ valuation of all} \\ & \text{the } i \times i \text{ minors of } H(s). \end{aligned} \tag{7.17}$$

Then the Smith-McMillan form at $s = \infty$ is defined by

$$M_\infty(s) = \text{diag}\left\{s^{\sigma_1(\infty)}, ..., s^{\sigma_r(\infty)}\right\} \tag{7.18}$$

where

$$\sigma_1(\infty) = v_\infty^{(1)}, \sigma_2(\infty) = v_\infty^{(2)} - v_\infty^{(1)}, \ldots \tag{7.19}$$

The family of unstructured generic models in the frequency domain is defined by:

Definition 7.9 : *For transfer function models $G(s)$, where $G(s)$ is proper, it is assumed that the number of inputs (l), outputs (m) are fixed, but the $g_{ij}(s)$ elements of $G(s)$ are generic proper rational functions. The family of systems is denoted by $\sum_{pr}(l,m)$.*

□

Then for generic systems in the family $\sum_{pr}(l,m)$, we have the following result.

Theorem 7.12 [Kar. & Kou.,1] *For generic systems of the $\sum_{pr}(l,m)$ family, the following properties hold true:*

- *The generic element of $\sum_{pr}(l,m)$ has no infinite zeros. If the system is strictly proper, then the generic system of $\sum_{pr}(l,m)$ has $\min\{m,l\}$ number of first order infinite zeros.*
- *If $m \neq l$, the generic system has no finite zeros. If $m = l$, the generic proper system has n finite zeros and the generic strictly proper system has $n-m$ finite zeros, where n is the McMillan degree of $G(s)$.*

□

In the following we consider a family of systems described in the frequency domain with a fixed structure. For instance, a system $H(s)$ is structurally given as

$$H(s) = \begin{bmatrix} A_1^2 A_2 & A_1 & A_3 \\ A_3 & A_2^2 & A_1 A_2 A_3 \\ 0 & A_4 & A_1^2 \end{bmatrix} \tag{7.20}$$

where $A_i, i = 1,2,3,4$ are dynamic models in the frequency domain whose orders are fixed while the coefficients in both the denominator and the numerator may take arbitrary values. We define the valuation element-wisely at infinity as:

$$u_\infty(H) \equiv \text{The element-wise valuation of } H(s) \tag{7.21}$$
$$= [u_\infty^{ij}]; \tag{7.22}$$

where $u_\infty^{ij} = v_\infty(h_{ij}), i = 1,...,m, j = 1,...,l$.

First we state the following result.

Proposition 7.4 [Var. Lim. & Kar.,1] *Given two generic rational functions $f(s)$ and $g(s)$ with valuations $\alpha = v_\infty(f)$ and $\beta = v_\infty(g)$, respectively. Then we have*

$$v_\infty(f \cdot g) = v_\infty(f) + v_\infty(g) \tag{7.23}$$
$$v_\infty(f \pm g) = \min\{v_\infty(f), v_\infty(g)\} \tag{7.24}$$

□

Remark 7.11: A structured transfer function matrix $H \in R_{pr}^{l\times l}$ can be transformed into an integer matrix $u_\infty(H)$ of the same dimension whose entries are the valuation of the corresponding entries in the structured transfer function matrix.

□

The path, weight, etc can be defined for the integer matrix $u_\infty(H)$ as in Definition (7.8). We can now establish the following result concerning the $i \times i$ valuation of $H(s)$ at infinity as:

Proposition 7.5 *The $i \times i$th order valuation of $H(s)$ at infinity is equivalent to the minimal weight of length i in matrix $u_\infty(H)$.*

Proof:

If the minimal weight of any independent paths with length i is t_i, we have to prove that there exists an $i \times i$ order minor in $H(s)$ such that the valuation of this minor at infinity is t_i. Construct an $i \times i$ order minor by taking the minor just containing the minimum weight path. Then the order of the minor will be $i \times i$ and by employing Proposition (7.4), the result is established.

□

For the system given in equation (7.20), if we assume that the components take the following forms

$$\begin{aligned} A_1 &= \frac{a_{11}}{s^2 + b_{11}s + b_{12}} \\ A_2 &= \frac{a_{21}s^2 + s_{22}s + a_{23}}{b_{21}s^4 + b_{22}s^3 + b_{23}s + b_{24}} \\ A_3 &= \frac{a_{31}s + a_{32}}{b_{31}s + b_{32}} \\ A_4 &= \frac{a_{41}}{s^2 + b_{41}s + b_{42}} \end{aligned}$$

then the element-wise valuation of the matrix $H(s)$ is given as

$$u_\infty(H) = \begin{bmatrix} 6 & 2 & 0 \\ 0 & 4 & 4 \\ \infty & 2 & 4 \end{bmatrix} \tag{7.25}$$

The generic valuation of 1×1 minors of the matrix $H(s)$ at infinity is the minimum among all the entries, or $v_\infty^{(1)}(H) = 0$. The minimal weight paths of the

2×2 minors

$$\begin{bmatrix} 6 & 2 \\ 0 & 4 \end{bmatrix}, \begin{bmatrix} 6 & 0 \\ 0 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 4 & 4 \end{bmatrix},$$
$$\begin{bmatrix} 0 & 4 \\ \infty & 2 \end{bmatrix}, \begin{bmatrix} 4 & 4 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 4 \\ \infty & 4 \end{bmatrix},$$
$$\begin{bmatrix} 6 & 2 \\ \infty & 2 \end{bmatrix}, \begin{bmatrix} 6 & 0 \\ \infty & 4 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 2 & 4 \end{bmatrix}$$

are $2, 0, 4, 2, 2, 6, 4, 8, 10, 2$. So the generic valuation of the 2×2 minors at infinity is 0, i.e., $v_\infty^{(2)}(H) = 0$. Finally the generic valuation of the 3×3 at infinity is the minimum weight path of $u_\infty(H)$ which is 2, so $v_\infty^{(3)}(H) = 2$. Finally we have

$$\begin{aligned} v_\infty^{(1)}(H) &= 0 \\ v_\infty^{(2)}(H) &= 0 \\ v_\infty^{(3)}(H) &= 2 \end{aligned}$$

and therefore the rational matrix generically has one zero of order 2 at infinity and no generic poles.

Remark 7.12: Assuming that the normal rank of the matrix $H(s)$ is r, because we consider only the proper system, the difference between the total number of zeros and the total number of poles at infinity is given by the minimal weight of all the independent paths with a length r.

□

Remark 7.13: From the above definition and analysis, it is clear that the number of generic poles and zeros at infinity remains unchanged as long as the element-wise valuation of the rational matrix is the same, because then all the minors of all orders will be the same. In fact, this happens when the valuations of the elements A_i are fixed.

□

Remark 7.14: In order to find the minimal weight of all the independent paths, the methods proposed in Section 7.4.3 for finding the maximal weight of all the independent paths can be adapted. The only change needed is to mark the minimum weight path instead of the maximum weight path.

□

7.6 Summary

In this chapter, we first discussed some of the general issues arising from the modelling of interconnected complex systems from the control engineer's view point. The attention is focused on process modelling at different design stages exploring the possibilities of applying control theories and tools in the design process. Two equally important aspects are highlighted: model environment and features of control theory. On one hand, the model environment in which the control theories wish to work is generally poor. Under these circumstances, it is argued [Kar.,2] that External Structural Dynamic Models are better suited for the life circle of design. On the other hand, features of control theory such as defining the minimal number of inputs, eliminating bad designs at early stages will be desirable.

Next, we looked at the properties which can be inferred from the steady-state models of systems. The properties include integral stabilisability, integral controllability, system sensor failure sensitivity, system actuator failure sensitivity and system robustness. The corresponding property indicators are based simply on the steady-state models and are vary useful in screening bad designs at an early stage.

Then we defined the notions of generic structured transfer functions, generic McMillan degree and generic poles and zeros both finite and at infinity. For the generic structured transfer functions, we transformed the problem of evaluating the generic McMillan degree into a problem of searching the maximal or minimal weight paths of the correspondingly defined integer matrices. Methods for searching the maximal or minimal weight path are then proposed and assessed.

Chapter 8

CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK

8.1 Conclusions

The two main issues addressed in this thesis are bounded state feedback and evaluation of structural characteristics. The topics covered in relation to bounded state feedback are quantitative controllability, the distance a stable polynomial from instability or of an unstable polynomial from stability and the root distribution of bounded coefficient polynomials. These are investigated in Chapters 4, 5 and 6, respectively. The problems of evaluating structural characteristics are discussed in Chapter 7.

In Chapter 3, quantitative state controllability, output controllability, quantitative observability of systems based on the state-space singular values of the Controllability Grammian and Observability Grammian have been developed. Next, the singular values of the state Controllability Grammian are developed to be indicators for parametrising the initial conditions in the state-space which can be brought to the origin with bounded energy controls. Indeed, the subset of initial conditions form a hyper-elipsoid whose axes are defined by the singular values of the state Controllability Grammian. Then, the quantitative output controllability is also developed to be an interaction measure based on energy between the inputs and outputs; this indicator can give important information in selecting input-output pairing schemes when the problem of selection of control structure is considered. In addition to the theoretical development of the property indicators, implementation of the property indicators based on MATLAB have also been produced. Finally, other quantitative controllability measures based on the distance that a controllable system away from uncontrollability and eigenvalues and eigenvectors are reviewed.

Because a link between the bound on the state feedback gain and the bound on the coefficients of the characteristic polynomial of a system can be established, the pole mobility of a system under bounded state feedback can be examined through the study of root distribution of bounded coefficient polynomials. So in Chapter 4, various classical results concerning root distribution in relation to the coefficients have been reviewed. Also included are the recent most well-known result of Kharitonov's theorem and the variations developed thereafter.

If one can find the minimum distance of an unstable polynomial from stability and the corresponding closest stable polynomial, then one could similarly find the minimum distance of an unstable system from stability and the corresponding closest stable system. Then by making use of this minimum distance a necessary condition imposed on the feedback gain for a system to be stabilisable could be produced in the feedback context. This may provide an alternative approach in bounded gain stabilisation. Along this line, the problem of finding the distance of a stable polynomial from the set of unstable polynomials, as well as the distance of an unstable polynomial from stable polynomials has been discussed. A novel method has been proposed for calculating either the distance a stable polynomial away from instability or the distance of an unstable polynomial from stability. Some of the results were later employed to establish an upper bound on the root inclusion problem of bounded coefficient polynomials in Chapter 6.

In studying the root distribution of bounded coefficient polynomials, the problems of inverse and direct root inclusion of polynomials have been formulated in Chapter 5 for bounded coefficient polynomials. The bound on the coefficient is given in terms of the l_2 norm on the coefficient vector. For different subfamilies of polynomials, different root distribution regions have been established, that is:

- For the direct root inclusion problem, minimum regions with algebraically defined boundaries have been obtained for the set of l_2 norm bounded stable polynomials $P^{+,\gamma}[s]$ as well as the set of l_2 norm bounded totally unstable polynomials $P^{-,\gamma}[s]$.

- Maximum rectangular regions have been obtained for the inverse root inclusion problem for both the stable and totally unstable polynomial sets $P^{+,\gamma}[s], P^{-,\gamma}[s]$.

- For the general polynomials $P[s]$, upper bound for the direct root inclusion problem is obtained. In addition, for the lower order polynomials, a tighter upper bound for the direct root inclusion problem has also been produced. With respect to the inverse root inclusion problem, the region is given by combining the maximum rectangular regions for both the stable and totally unstable polynomials.

The direct root inclusion problem for the set of norm bounded general polynomial needs further investigation. The aim is to find a minimum region which contains all the coefficient bounded polynomials $P[s]$. The general results derived so far are still by no means minimal. If the bound on the coefficients is given in other forms, for instance in terms of l_1, l_3 or l_∞, etc. the impact on corresponding direct and inverse root inclusion problems should also be investigated. Also addressed in Chapter 5 is the problem of root distribution for the sum of two monic polynomials. Results show that the roots of the sum of two monic polynomials whose roots are symmetrically distributed with respect to a line parallel to the imaginary axis are distributed along the line solely. And applications based on this result have been made for the case of root distribution of the sum of two polynomials each of which has a root distribution in a subset of the real axis.

In Chapter 6, the closed-loop pole mobility for SISO systems under bounded state feedback has been studied. The bound on the controller gain is given in terms of the l_2 norm. First the norm bound on the controller gain is transformed into the bound on the coefficients of the characteristic polynomial. Then, results developed in the previous chapters are deployed to establish the upper bound for the closed-loop poles. If an open-loop system is unstable, then sufficient conditions have been obtained for the system to be stabilisable by the bounded feedback.

Closed-loop pole mobility for MIMO systems under bounded state feedback has also been studied. However, the norm bound on the controller gain is given in terms of Frobenius norm instead of l_2 norm. The closed-loop pole mobility regions are established for the cases when special pole assignment methods are used, namely dyadic feedback scheme and controller form pole assignment scheme. If an open-loop system is unstable, then sufficient conditions are given for the system to be stabilisable under bounded state feedback. In dyadic feedback, or indeed for the general linearisation feedback scheme, the selection of the input direction affects the coefficient bound on the closed-loop characteristic polynomials and hence the final mobility regions. The problem of how to select the best input direction to render maximal pole mobility under bounded state feedback has still to be answered.

Evaluation of structural characteristics are addressed in Chapter 7, where the working model characteristics and the desirable features of control theory concerning the design of large scale processes with ill-defined models have first been discussed. It is argued that external structural dynamic models should be used in the design and a family of models with increasing complexity should be developed in order to suit the need at different design stages. However, the models should have the characteristics such as Universality, Computability and Parameter insensitivity. Then, some existing results concerning the development of useful tools based on the steady-state gain information has been reviewed. Not only integral stabilisability and integral controllability can be inferred from the steady-state gain of a system, but also system integrity and robustness. Next, we have defined the concepts of generic structured transfer functions, generic McMillan degree at both finite and infinite frequencies. It has been proved that the evaluation of the generic McMillan degree, generic poles and zeros both finite and at infinity can be transformed into searching independent paths with either maximal or minimal weight of integer matrices. Different searching algorithms have been proposed and assessed.

8.2 Suggestions for further research work

Further research work is needed in both the areas of bounded state feedback design and evaluation of structural characteristics of large scale systems with ill-defined mathematical models.

The mechanism how the system matrix and the input matrix affect the quantitative controllability, i.e., the singular values of the Controllability Grammians is not very well understood; nor is the mechanism how the system and input matrices affect the mobility of system eigenvalues under feedback control. In addition, the effects of feedback on the singular values of either Controllability Grammian, or controllability matrix of the closed-loop systems need to be further explored. It is believed that these will provide vital information in feedback design.

In this work, the bounded state feedback problem is tackled via the study of the root distribution of bounded coefficient polynomials. The results derived on the mobility of closed-loop eigenvalues under bounded state feedback control are still quite conservative. The main cause of the conservativeness is due to the inequalities used in transforming the bound on the feedback onto the bound on the coefficients of the closed-loop characteristic polynomials. In order to reduce the conservativeness, one could possibly establish tighter regions for the direct root inclusion problem of bounded coefficient polynomials for the general case. However, the alternative approach by finding the minimal distance an unstable system from stability numerically is believed to be the key to this answer. Along this line, the nature of the numerical problem such as the existence of global minimum should be properly investigated.

As to the problem of evaluating the structural characteristics of large scale systems with ill-defined models, only a general framework is outlined in this work. In order to provide a set of comprehensive and systematic tools in assisting process design at early stages, much more further work is needed in this direction. For instance, the selection of control structures such as defining the lower bound on the number of inputs and outputs and selecting inputs and outputs for actuation and measurement purposes, the evaluation of system properties like stability robustness against structure of parameter variations, etc., should be examined further. Computationally economical tools need to be developed in order to work on large scale systems. Considerations should also be given to systems which are of sparse nature.

Bibliography

[Ack.,1] Ackermann, J.,*Abtastregulung*. Springer, Berlin. 1972.

[Ack.,2] Ackermann, J., "Parameter space design of robust control systems". IEEE Trans. on Automatic Control, vol. AC-25, pp. 1058-1072. 1980.

[And. Jur. & Man.,1] Anderson, B.D.O., Jury, E.I. and Mansour, M., "On robust Hurwitz polynomials". IEEE Trans. on Automa. Cont., vol. AC-32, pp.909-913. 1987.

[Arg.,1] Argoun, M.B., "Frequency domain conditions for the stability of perturbed polynomials". IEEE Trans. on Automa. Cont., vol. AC-32, pp.913-916. 1987.

[Arg.,2] Argoun, M.B., "Stability of a Hurwitz polynomial under coefficient perturbations: Necessary and sufficient conditions". Int. J. Control, vol.45, No.2, 739-744. 1987.

[Barm.,1] Barmish, B.R., "Invariance of the strict Hurwitz property for polynomials with perturbed coefficients". IEEE Tran. on Automatic Control, vol.AC-29, pp.935-936. 1984.

[Barm.,2] Barmish, B.R., "A generalisation of Kharitonov's four-polynomial concept for robust stability problems with linearly dependent coefficient perturbations". IEEE Trans. Automat. Contr., vol. AC-34, No.2, pp.157-165, 1989.

[Bar.,1] Barnett, S., *Polynomials and Linear Control Systems*, Marcel Dekker Inc. New York. 1983.

[Bas. & Gur.,1] Bass, R.W. and Gura, I., "High order design via state space considerations". Proceedings of the 1965 Joint Automatic Control Conference, Troy, N.Y. 1965.

[Bart. Hol. & Lin,1] Bartlett, A. C., Hollot, C. V. and Lin, H., "Root locations of an entire polytope of polynomials: It suffices to check the edges". Mathematics of Control, Signals and Systems, vol. 1, Number 1, pp.61-71. 1988.

[Bia. & Gar.,1] Bialas, S. and Garloff, J., "Stability of polynomials under coefficient perturbations". IEEE Trans. on Automatic Control, vol. AC-30, No.3. 1985.

[Bie. Hwa. & Bha.,1] Biernacki, R. M., Hwang, H. and Bhattacharyya, S. P.,"Robust stability with structured real parameter perturbations". IEEE Trans. on Automatic Control, vol. AC-32, pp. 495-506. 1987.

[Boe. & Lut.,1] Boese, F.G. and Luther, W.J., "A note on a classical bound for the moduli of all zeros of a polynomial". IEEE Trans. on Automatic Control, vol. AC-34, No.9. 1989.

[Bol. & Lu,1] Boley, D.L. and Lu, W. S., "Measuring how far a controllable system is from an uncontrollable one". IEEE Trans. on automatic Control, AC-31, pp.249-251. 1986.

[Bos. & Shi,1] Bose, N.K. and Shi, Y.Q., "A simple general proof of Kharitonov's generalised stability criterion". IEEE Trans. on Circuits and Systems, vol. CAS-34, pp. 1233-1248. 1987.

[Bra. & Pea.,1] Brasch, F. M. and Pearson, J. B., "Pole placement using dynamic compensators". IEEE Trans. Auto. Control, vol. AC-15, pp. 34-43. 1970.

[Bri.,1] Bristol, E.H., "On a new measure of interaction for multivariable process control". IEEE Trans. Auto. Control, vol. AC-11, p133. 1966.

[Bru.,1] Brunovsky, "A classification of linear controllable systems". Kybernetika, 6 (3), pp. 173-188. 1970.

[Byrn.,1] Byrnes, C. I., "Algebraic and geometric aspects of the analysis of feedback systems". *Geometric Methods in Linear Systems Theory.* C.I. Byrnes and C. F. Martins (Ed.). Dordrecht: Reidel. 1980.

[Cal. & Fra.,1] Callier, Frank, M., *Multivariable Feedback Systems*, Dowden & Culver, Inc. 1982.

[Cha. & Bha.,1] Chapellat, H. and Bhattacharyya, S.P., "A generalisation of Kharitonov's theorem: Robust stability of interval plants". IEEE Trans. on Auto. Control, vol.34, pp.306-311. 1989.

[Chen,1] Chen, C.T., *Introduction to Linear systems*, Holt-Rinehart and Winston, New York. 1984.

[Cie.,1] Cieslik, J., "On possibilities of the extension of Kharitonov's stability test for interval polynomials to discrete systems". IEEE TRans. Automat. Contr., vol. AC-32. pp237-238. 1987.

[Dav.,1] Davison, E.J., "On pole assignment in multivariable linear systems". IEEE Transactions on Automatic Control, vol. AC-13, No. 6. Dec. pp.747-748. 1968.

[Dav.,2] Davison, E.J., "On pole assignment in linear systems with incomplete state feedback". IEEE Transactions on Automatic Control, vol. AC-15, No. 6. June. pp.348-351. 1970.

[Dem.,1] Demmel, J. W., "A counterexample for two conjectures about stability". IEEE Trans. on Automatic Control, vol. AC-32, pp.340-342. 1987.

[Dic.,1] Dickman, A., "On the robustness of multivariable linear feedback systems in state-space representation". IEEE Trans. on Automatic Control, vol. AC-32, pp. 407-410. 1987.

[Doy. & Ste.,1] Doyle, J. C. and Stein, G., "Multivariable feedback design:Concepts for a classical/modern synthesis". IEEE Trans. on Automatic Control, vol. AC-26, pp. 4-16. 1981.

[Eis.,1] Eising, R., "Between controllable and uncontrollable". Systems and Control Letters, vol. 4, pp. 263-264, 1984.

[EPIC] EPIC Project: ESPRIT II Project 2090: Early Process Design Integrated with Control.

[Gag. & Seb.,1] Gagnepain, J.P. and Seborg, D.E., "An analysis of process interactions with applications to multiloop control system design". Ind. Eng. Chem. Process Des. Dev., 21, pp5-11. 1982.

[Gah.,1] Gahinet, P. M., "Distance to the nearest uncontrollable pair and algebraic Riccati equation". Proceedings of the 29 Conference on Decision and Control. Honolulu, Hawaii. December, 1990. pp. 273-278.

[Gan.,1] Gantmacher, F.R., *The Theory of Matrices.* Volume I and II. Chelsea Publishing Company. New York, N.Y. 1959.

[Gia. Kar. & Kal.,1] Giannakopoulos, C., Karcanias, N. and Kalogeropoulos, G., *Multivariable Control.* S.G. Tzafestas (ed.). Reidel Publishing Company. Chapter 2. "The theory of polynomial combinants in the context of linear systems". pp. 27-41. 1984.

[Gol. & van Loan,1] Golub, G. and van Loan, C., *Matrix Computations.* Baltimore, MD: Johns Hopkins Press. 1983.

[Gro. Mor. & Hol.,1] Grosdidier, P., Morari, M., and Holt, B.R., "Closed-loop properties from steady-state gain information". Ind. Eng. Chem. Fund., 24, p221. 1985.

[Hau. & Hey.,1] Hautus, M. L. J. and Heymann, M, "Linear feedback — An algebraic approach". SIAM J. Control and Optimisation. vol. 16, pp. 83-105. 1978.

[Hin. & Pri.,1] Hinrichsen, D., and Pritchard, A.J., "Stability radius for structured perturbations and the algebraic Riccati equation". Systems & Control Letters 8. pp. 105-113. 1986.

[Hin. & Pri.,2] Hinrichsen, D., and Pritchard, A.J., "Real and complex stability radii: A survey". 1991.

[Hin. Kel. & Lin.,1] Hinrichsen, D., Kelb, B. and Linnemann, A., "An algorithm for the computation of the structured complex stability radius". Automatica. vol. 25, No. 5. pp. 771-775. 1989.

[Hol. Looz. & Bar.,1] Hollot, C.V. Looze, D.P. and Bartlett, A.C., "Unmodeled dynamics: Performance and stability via parameter space methods". Proceedings of the 26th Conference on Decision and Control. Los Angeles, CA. 1987.

[Kai.,1] Kailath, T., *Linear Systems*, Prentice-Hall, Englewood Cliffs, 1983.

[Kal.,1] Kalman, R.E., "Irreducible realizations and the degree of a rational matrix". SIAM J., Vol.13, pp.520-544. 1965.

[Kal.,2] Kalman, R.E., "Kronecker invariants and feedback". Proceeding of Conference on Ordinary Differential Equation (Ed. L.Weiss). NLR Math. Res. Cent., pp.459-471. 1965.

[Kar.,1] Karcanias, N., "System properties: Property, design indicators and system invariants". EPIC reports. Deliverable 1.1. Section 3. 1989.

[Kar.,2] Karcanias, N., "Early process design integrated with control". EPIC report No. 2090. 1991.

[Kar.,3] Karcanias, N., "The model projection problems in the global instrumentation of a process". Submitted for publication.

[Kar. & Gia.,1] Karcanias, N. and Giannakopoulos, C., "Grassmann invariants, almost zeros and the determinantal zero, pole assignment problems of linear multivariable systems". International Journal of Control, vol.40. pp.673-698. 1984.

[Kar. & Gia.,2] Karcanias, N. and Giannakopoulos, C., "Frequency assignment problems in linear multivariable systems: exterior algebra and algebraic geometry methods". *Multivariable Control: New Concepts and Tools*, edited by S. G. Tzafestas (Dordrecht: D. Reidel), pp. 211-232. 1984.

[Kar. Gia. & Hub.,1] Karcanias, N. Giannakopoulos, C. and Hubbard, M., "Almost zeros of a set of polynomials of $R[s]$". International Journal of Control, Vol. 38, pp. 1213-1238. 1983.

[Kar. & Kal.,1] Karcanias, N. and Kalogeropoulos, "On the Segre', Weyr characteristics of right (left) regular pencils". International Journal of Control, Vol. 44, pp. 991-1015. 1986.

[Kar. & Kou.,1] Karcanias, N. and Kouvaritakis, B., "The output zeroing problems and its relationship to the invariant zero structure: a matrix pencil approach". International Journal of Control, Vol.30, pp.395-415. 1979.

[Kar. Lai. & Gia.,1] Karcanias, N., Laios, B. and Giannakopoulos, C., "Decentralized determinantal assignment problem: fixed and almost fixed modes and zeros". International Journal of Control, vol.48, pp.129-147. 1988.

[Kar. & Mac.,1] Karcanias, N. and MacBean, P., "Structural invariants and canonical forms of linear multivariable systems". Third IMA Conference on Control Theory, Academic Press, pp.257-282. 1981.

[Kar. & Shan,1] Karcanias, N. and Shan, X.Y., "Root distribution of bounded coefficients polynomials". International Journal of Control, vol. 53, pp. 929-949. 1991.

[Kar. & Shan,2] Karcanias, N. and Shan, X.Y., "Pole mobility and stabilisability of linear systems under bounded state feedback: The SISO case". Submitted to CDC. 1992.

[Kee. Bha. & How.,1] Keel, L. H., Bhattacharyya, S. P. and Howze, J. W., "Robust control with structured perturbations". IEEE Trans. on Automatic Control, vol. 33, pp. 68-78. 1988.

[Kha.,1] Kharitonov, V.L., "Asymptotic stability of an equilibrium position of a family of systems of linear differential equations". Deffertial. Uravnen., vol.14, No.11, pp.2086-2088, 1978.

[Kou.,1] Koussiouris, T., Private communication. 1991.

[Kuc.,1] Kucera, V., *Discrete Linear Control: The polynomial Equation Approach*, John Wiley & Sons, New York, 1979.

[Leh. Gas. Lev. Ste. San. & Ath.,1] Lehtomaki, N. A., Gastanon, D. A., Levy, B. C., Stein, G., Sandell JR., N. R. and Athans, M, "Robustness and modelling error characterization". IEEE Trans. on Automatic Control, vol. AC-29, pp. 212-220. 1985.

[Lin,1] Lin, C. T., "Structural controllability". IEEE Trans. on Automatic Control, vol. AC-19, pp. 201-208. 1974.

[Lue.,1] Luenberger, D. G., "An introduction to observers". IEEE Trans on Automatic Control, vol. AC-11, pp. 596-602. 1971.

[Lun.,1] Lunze, J.A.N., *Robust Multivariable Feedback Control*, Prentice Hall, London. 1989.

[MacF.,1] MacFarlane, A.G.J., "Return-difference and return-ratio matrices and their use in analysis and design of multivariable feedback control systems". Proc. IEE, 177, pp. 2037-2049. 1970.

[MacF. & Kar.,1] MacFarlane, A.G.J. and Karcanias, N., "Poles and zeros of linear multivariable systems: A survey of the algebraic, geometric and complex variable theory". International Journal of Control, Vol.24, pp.33-74. 1976.

[MacF. & Kar.,2] MacFarlane, A.G.J. and Karcanias, N., "Relations between state space and frequency response concepts". Proc. 7th Triennial World IFAC Congr. Pergamon, Oxford. 1978.

[MacF. & Kou.,1] MacFarlane, A.G.J. and Kouvaritakis, B., "A design technique for linear multivariable feedback systems". International Journal of control, Vol.25, pp.837-874. 1977.

[MacF. & Pos.,1] MacFarlane, A.G.J. and Postlethwaite, I., "The generalised Nyquist stability criterion and multivariable root loci". International Journal of Control, Vol.25, pp.81-127. 1977.

[MacF. & Sco.,1] MacFarlane, A.G.J. and Scott-Jones, D. F. A., "Vector gain". International Journal of Control, vol.29, pp. 65-91. 1979.

[MacL. & Bir.,1] MacLane, S. and Birkhoff, G., *Algebra*, second edition, McMillan. New York. 1979.

[Man. Sav. & Ark.,1] Manousiouthakis, V., Savage, R., and Arkun, Y., "Synthesis of decentralised process control structures using the concept of block relative gain". AICHE Journal, vol.32, No.6, pp991-1003. 1986.

[Mar.,1] Marden, M., *The Geometry of the Zeros of a Polynomial in a Complex Variable.* 1949.

[Mart.,1] Martin, J. M., "State space measures for stability robustness". IEEE Trans. on Automatic Control, vol. AC-32, pp. 509-512. 1987.

[Matlab] *Matlab Manual*, Cambridge Control. 1990.

[May. & Mur.,1] Mayne D. and Murdoch, P., "Modal control of linear-invariant systems". Int. J. Control, 11, pp.223-227. 1970.

[Moo.,1] Moore, B. C., "On the flexibility offered by state feedback in multivariable systems beyond closed loop eigenvalue assignment". IEEE Trans. on Automatic Control, vol. AC-21, pp.689-692. 1976.

[Mor.,1] Morari, M., "Robust stability of systems with integral control". 22nd IEEE Conference on Decision and Control, San Antonio, TX. 1983.

[Mor. & Ste.,1] Morari, M., and Stephanopoulous, G., "Studies in the synthesis of control structures for chemical processes. II: Structural aspects and the synthesis of alternative feasible control schemes". AICHE, Journal, vol.26, No.2, p232. 1980.

[Mori,1] Mori, T., "Note on the absolute value of the roots of a polynomial". IEEE Trans. on Automatic Control, vol. AC-29, No. 1. 1984.

[Mor.,1] Morse, A.S., "Structural invariants of linear multivariable systems". SIAM J. Control, vol. 11, pp. 446-465. 1973.

[Mun.,1] Munro, N., "Pole assignment: A review of methods". Systems & Control Encyclopedia (Editor M.G. Singh), Pergamon Press. pp. 3710-3717. 1986.

[Mun.,2] Munro, N., (Editor) *Modern Approaches to Control System Design.* IEE Control Engineering Series, vol. 9.

[Nie.,1] Niederlinski, A., "A heuristic approach to the design of linear multivariable interacting control systems". Automatica, pp. 691-701. 1971.

[Ost.,1] Ostrowski, A.M., *Solution of Equations and Systems of Equations.* 2nd edition. Academic Press. 1966.

[Paig.,1] Paige, C. C., "Properties of numerical algorithms related to computing controllability". IEEE Trans. on Automatic Control, vol. AC-26, pp. 130-138. 1981.

[Pat.,1] Patel, R.V. and Munro, N., *Multivariable System Theory and Design*. Pergamon Press. 1982.

[Pop.,1] Popov, V.M., *Some Properties of Control systems with Matrix Transfer Functions*, in Lecture Notes in Mathematics, Vol.144, Springer-Verlog, pp.169-180,1969.

[Pos. & MacF.,1] Postlethwaite, I. and MacFarlane, A. G. J., *A Complex Variable Approach to the Analysis of Linear Multivariable Feedback Systems*. Springer-Verlag. 1979.

[Pos. Edm. & MacF.,1] Postlethwaite, I., Edmunds, J. M. and MacFarlane, A. G. J., "Principal gains and principal phases in the analysis of linear multivariable feedback systems". IEEE Trans. Autom. Control, AC-26, pp. 4-16. 1981.

[Rei.,1] Reinschke, K.J., *Multivariable Control: A Graph-Theoretical Approach*. Springer-Verlag. 1988.

[Ros.,1] Rosenbrock, H.H., *State-space and Multivariable Theory*, Nelson, London. 1970.

[Ros.,2] Rosenbrock, H.H., "Structural properties of linear dynamical systems". International Journal of Control, Vol.20, pp.191-202, 1974.

[Ros.,3] Rosenbrock, H. H., "On the design of linear multivariable control systems". Proc. Third IFAC Congress, London. 1, pp.1-16. 1966.

[Rou.,1] Routh, E.J., *Stability of Motion*. London: Taylor and Francis. 1975.

[San. & Ath., 1] Sandell, H. and Athans, M., "On "type L" multivariable linear systems". Automatica, pp.131-136. 1973.

[Sar. & Ker.,1] Saridereli, M.K. and Kern, F.J., "The stability of polynomials under correlated coefficient perturbations". in Proc. IEEE Conf. Decision Contr., Los Angeles, CA. 1987.

[Sei.,1] Seidman, T.I., "How violent are fast controls?" Mathematics of Control, Signals ans Systems, vol. No.1. 1988.

[Shan & Kar.,1] Shan, X.Y. and Karcanias, N., "Pole mobility and stabilisability of linear systems under bounded state feedback: The MIMO case". Submitted to IMA Control Conference. Manchester. September, 1992.

[Ske.,1] Skelton, R.E., *Dynamic Systems Control.* John Wiley & Sons. 1988.

[Soh. Ber. & Dab.,1] Soh, C.B., Berger, C.S. and Dabke, K.P., " On the stability properties of polynomials with perturbed coefficients'. IEEE Trans. Automat. Contr., vol. AC-30, pp.1033-1036, 1985.

[Soh. Ber. & Dab.,2] Soh, C.B., Berger, C.S. and Dabke, K.P., "Addendum to "On the stability properties of polynomials with perturbed coefficients" ". IEEE Trans. on Automatic Control, vol. AC-32. No.3. pp.239-240. 1987.

[Soh. & Eva.,1] Soh, Y.C. and Evans, R.J., "Characterisation of allowable perturbations of a closed-loop system with pole placement controller". IEEE Trans. on Automatic Control, vol. AC-33. No.5. pp.466-469. 1988.

[Soh,1] Soh, Y.C., "Strict Hurwitz property of polynomials under coefficient perturbation". IEEE Trans. on Automatic Control, vol. 34. pp.629-632. 1989.

[Tar.,1] Tarokh, M., "Measures for controllability and eigenvalue mobility". Proceedings of the 29th Conference on Decision and Control. Honolulu, Hawaii, Dec. 1990. pp.1143-1148.

[Tes. & Vic.,1] Tesi, A. and Vicino, A., "Robustness analysis for linear dynamical systems with linearly correlated parametric uncertainties". IEEE Trans. on Automatic Control, vol. 35. No. 2. pp.186-191. 1990.

[Tho.,1] Thorp, J.P., "The singular pencil of linear dynamical systems". Int. J. Control, vol. 18, pp. 577-596. 1973.

[Tun. & Edg.,1] Tung, and Edgar, AICHE J. vol. 27, p.690. 1981.

[Var. Lim. & Kar.,1] Vardulakis, A.I.G., Limebeer, D.J., and Karcanias, N., "Structure and Smith-McMillan form of rational matrix at infinity". International Journal of Control, Vol.35, pp.701-725. 1982.

[Verg. & Kai.,1] Verghese, G. and Kailath, T., "Fixing the state feedback gain by choice of closed-loop eigensystem". Proceedings of the 1977 IEEE Conference on Decision and Control, New Orleans, Dec. 1977. pp. 1245-1248.

[Vic.,1] Vicino, A., "Some results on robust stability of discrete time systems". IEEE Trans. Automat. Contr., vol. 33. pp. 844-847. 1988.

[Vid.,1] Vidyasagar, M., *Control System Synthesis: A Factorization Approach*, MIT Press, Cambridge, Mass. 1985.

[Wei. & Yed.,1] Wei, K.H. and Yedavalli, R.K., "Invariance of strict Hurwitz property for uncertain polynomials with dependent coefficients". IEEE Trans. on Automa. Cont., vol. AC-32, pp.907-909. 1987.

[Wil.,1] Wilkinson, J.H., *The Algebraic Eigenvalue Problem*, Clarenton Press, Oxford. 1965.

[Wit. & McA.,1] Witcher, M.F., and McAvoy, T.F., "Interacting control systems: Steady state and dynamic measurement of interaction". ISA Trans., 16(3), p35. 1977.

[Won.,1] Wonham, W.M., *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, New York. Second Edition. 1979.

[Won.,2] Wonham, W.M., "On pole assignment in multi-input controllable linear systems". IEEE Trans. Automat. Control, vol. AC-12, No. 6, pp. 660-665. 1967.

[Xu,1] Xu, D.Y., "On the bounds for the roots of a polynomial". IEEE Trans. on Automatic Control, vol. AC-32. No.4. pp336-337. 1987.

[Yeu. & Wan.,1] Yeung, K.S. and Wang, S.S., "A simple proof of Kharitonov's theorem". IEEE Trans. Automatic Control, vol. AC-32, No.9, pp.822-823. 1987.

[Zeh.,1] Zeheb, E.,"Necessary and sufficient conditions for root clustering in simple connected domains". IEEE Trans. Automat. Contr., vol. 33. pp. 984-986. 1988.

Appendix: I

I.A. MATLAB code for the computation of controllability Grammian computation.

```
% This programme calculates the controllability Grammian as a function
% of time. The singular values of the controllability Grammian shows
% the controllability of the eigenvalues in the sense that the bigger
% the singular values, the less the energy needed to transfer a system
% from an initial state to a desired final state.
% These are the initial condition assignment
disp('please specify the data source:either by m-file or from keyboard
sou=input('input matrices A, B, and time t from the keyboard,  ''y'' or
''n''?','s');
if sou=='y'
a=input('The system matrix A is:')
b=input('The input matrix B is:')
tff=input('The final time tff is:')
else
disp('please specify the m-file which contains the data')
inn=input( ' ','s');
inn=['load ',inn];
eval(inn)
a
b
tff
end
[n,m]=size(a);
[nn,mm]=size(b);
m=mm;
t0=0.;
stem1=0;
stem2=0;
    for i=1:2*n*n
        x0(i)=0.;
    end
    for i=1:n
        x0((i-1)*n+i)=1.;
    end
x0=x0';
[t,x]=sode23('st',n,m,a,b,t0,tff,x0);
i=length(t);
for i1=1:i;
    for j=1:n;
        for k=1:n;
            stem1(j,k)=x(i1,n*n+(j-1)*n+k);
            txout((j-1)*n+k)=x(i1,n*n+(j-1)*n+k);
        end
    end
```

```
    xxout=[xxout;txout];
    [v,s,u]=svd(stem1);
    temcond=rcond(s);
    for i2=1:m;
        stem2(i2)=s(i2,i2);
    end
    singular=[singular;stem2];
    condition=[condition;temcond];
end
plot(t,singular)
title('singular values of controllability Grammian')
xlabel('t')
ylabel('singular values')
disp('strike any key to continue'),pause
plot(t, condition)
title('condition number of controllability Grammian')
xlabel('t')
ylabel('condition number')
```

I.B. MATLAB code for the computation of output controllability Grammian computation.

```
% outgra.m calculates the output controllability Grammian of a
% system. The singular values of the output controllability
% Grammian of the system are indicators of the output assignability
% of the system.
% These are the initial condition assignment
stem1=0;
stem2=0;
disp('please specify the data source:either by m-file or from keyboard
sou=input('input matrices A, B, and time t from the keyboard,  ''y'' or
 ''n''?','s');
if sou=='y'
   a=input('The system matrix A is:')
   b=input('The input matrix B is:')
   c=input('The output matrix C is:')
   tff=input('The final time tff is:')
else
   disp('please specify the m-file which contains the data')
   inn=input('','s');
   inn=['load ',inn];
   eval(inn)
   a
   b
   c
   tff
end
[n,m]=size(a);
```

```
[nn,mm]=size(b);
m=mm;
t0=0.;
for i=1:2*n*n
    x0(i)=0.;
end
for i=1:n
    x0((i-1)*n+i)=1.;
end
x0=x0';
[t,x]=sode23('st',n,m,a,b,t0,tff,x0);
i=1;
while tff-t(i)>0,i=i+1;end
      for i1=1:i
          for j=1:n;
              for k=1:n;
                  stem1(j,k)=x(i1,n*n+(j-1)*n+k);
                  txout((j-1)*n+k)=x(i1,n*n+(j-1)*n+k);
              end
          end
          xxout=[xxout;txout];
          [v,s,u]=svd(c*stem1*c');
          temcond=rcond(s);
          for i2=1:m
              stem2(i2)=s(i2,i2);
          end
          singular=[singular;stem2];
          condition=[condition;temcond];
      end
plot(t,singular)
title('singular values of output controllability Grammian')
xlabel('t')
ylabel('singular values')
disp('strike any key to continue'),pause
plot(t, condition)
title('condition number of output controllability Grammian')
xlabel('t')
ylabel('condition number')
```

I.C. MATLAB code for the computation of observability Grammian computation.

```
% This programme calculates the singular values of the finite time
% observability Grammian. The singular values of the observability
% are the indicators of the accurcy of the observation of the
% states.

% These are the initial condition assignment
disp('please specify the data source:either by m-file or from keyboard
```

```
sou=input('input matrices A, B, and time t from the keyboard,  ''y'' or  n'
if sou=='y'
   a=input('The system matrix A is:')
   c=input('The output matrix C is:')
   tff=input('The final time tff is:')
else
   disp('please specify the m-file which contains the data')
   inn=input('','s');
   inn=['load ',inn];
   eval(inn)
   a
   c
   tff
end
t0=0;
stem1=0;
stem2=0;
a=a';
b=c';
[n,m]=size(a);
[nn,mm]=size(b);
   for i=1:2*n*n
       x0(i)=0.;
   end
   for i=1:n
       x0((i-1)*n+i)=1.;
   end
x0=x0';
[t,x]=sode23('st',n,m,a,b,t0,tff,x0);
i=1;
while tff-t(i)>0,i=i+1;end
for i1=1:i
    for j=1:n;
        for k=1:n;
            stem1(j,k)=x(i1,n*n+(j-1)*n+k);
            txout((j-1)*n+k)=x(i1,n*n+(j-1)*n+k);
        end
    end
    xxout=[xxout;txout];
    [v,s,u]=svd(stem1);
    temcond=rcond(s);
    for i2=1:m;
        stem2(i2)=s(i2,i2);
    end
    singular=[singular;stem2];
    condition=[condition;temcond];
end
plot(t,singular)
```

```
title('singular values of observability Grammian')
xlabel('t')
ylabel('singular values')
disp('strike any key to continue'),pause
plot(t, condition)
title('condition number of obsevability Grammian')
xlabel('t')
ylabel('condition number')
```

I.D. Affilated programme for intergration which is adopted from ode23.m is MATLAB.

```
function [tout, yout] =sode23(F,n,m,a,b,t0, tfinal, y0, tol, trace)
%SODE23  Integrate a system of ordinary differential equations using
%
% [tout, yout] = sode23(F, n,m,a,b,t0, tfinal, y0, tol, trace)
%
% INPUT:
% F     - String containing name of user-supplied problem description.
%         Call: yprime = fun(t,y) where F = 'fun'.
%         t      - Time (scalar).
%         y      - Solution column-vector.
% yprime - Returned derivative column-vector; yprime(i) = dy(i)/dt.
% n - dimension of matrix a.
% t0     - Initial value of t.
% tfinal- Final value of t.
% y0     - Initial value column-vector.
% tol    - The desired accuracy. (Default: tol = 1.e-3).
% trace - If nonzero, each step is printed. (Default: trace = 0).
%
% OUTPUT:
% tout   - Returned integration time points (row-vector).
% yout   - Returned solution, one solution column-vector per tout-value
%
% The result can be displayed by: plot(tout, yout).
% Initialization
pow = 1/3;
if nargin < 9, tol = 0.001; end
if nargin < 10, trace = 0; end
 % Initialization
t = t0;
hmax = (tfinal - t)/5;
hmin = (tfinal - t)/20000;
h = (tfinal - t)/100;
y = y0(:);
tout = t;
yout = y.';
```

```
tau = tol * max(norm(y, 'inf'), 1);
if trace
   cla, t, h, y
end
% The main loop
   while (t < tfinal) \& (h >= hmin)
         if t + h > tfinal, h = tfinal - t;
   end
% Compute the slopes
      s1 = feval(F, t, y,n,m,a,b);
      s2 = feval(F, t+h, y+h*s1',n,m,a,b);
      s3 = feval(F, t+h/2, y+h*(s1'+s2')/4,n,m,a,b);
% Estimate the error and the acceptable error
      delta = norm(h*(s1' - 2*s3' + s2')/3,'inf');
      tau = tol*max(norm(y,'inf'),1.0);
% Update the solution only if the error is acceptable
      if delta <= tau
         t = t + h;
         y = y + h*(s1' + 4*s3' + s2')/6;
         tout = [tout; t];
         yout = [yout; y.'];
      end
      if trace
         home, t, h, y
      end
% Update the step size
      if delta = 0.0
         h = min(hmax, 0.9*h*(tau/delta) ^ pow);
      end
   end;
   if (t < tfinal)
      disp('SINGULARITY LIKELY.')
      t
   end
```

I.E. Function st.m

```
function xdot=st(t,x,n,m,a,c)
b=c*c';
for i=1:n;
    for j=1:n;
        tem=0;
        for h=1:n;
            tem=tem+a(i,h).*x((h-1)*n+j);
        end
        xdot((i-1)*n+j)=tem;
    end
end
```

```
for i=1:n;
    for j=1:n;
        tem2=0;
        for k=1:n;
            tem1=0;
            for h=1:n;
                tem1=tem1+x((i-1)*n+h).*b(h,k);
            end
            tem2=tem2+tem1.*x((j-1)*n+k);
         end
         xdot(n*n+(i-1)*n+j)=tem2;
    end
end
```

Appendix: II

II.A. FORTRAN code for computing the minimum distance of a stable polynomial from instability (Fourth order case)

```
      INTEGER  N,I
      DOUBLE   PRECISION BL(4), BU(4),X(4),B(5),BB(5),PL,PU,NL,NU,
     *                   PXI,NXI,XX(15,4),FF(15),ROT(2,15,4),F,RT(2,4
      DATA BB/1.,4.,6.,5.,2./
      COMMON//BB(5)
      EXTERNAL MINROOT,OROOT,MINFUN
c
c
      N=4
      PRINT *,'COEFFICIENTS OF THE ORIGINAL POLYNOMIAL ARE'
      DO 15 I=1,5
 15   B(I)=BB(I)
      PRINT *,(B(I),I=1,5)
      CALL OROOT(N,B)
      PL=1.e-20
      PU=1.E20
      NL=-1.E20
      NU=-1.E-20
      PXI=10.
      NXI=-10.
C
C
C
      DO 300 IOVAL=1,N
      IF (IOVAL.EQ.1) THEN
              ILOCAL=N
              DO 99 I=1,ILOCAL
              BU(I)=PU
              BL(I)=PL
              X(I)=PXI
 99           CONTINUE
              DO 98 J=1,ILOCAL
              BU(J)=NU
              BL(J)=NL
              X(J)=NXI
              CALL MINFUN(N,X,BU,BL,F)
                      DO 299 IX=1,4
 299                  XX(J,IX)=X(IX)
              FF(J)=F
              BU(J)=PU
              BL(J)=PL
              X(J)=PXI
              CALL MINROOT(N,X,RT)
```

```
          DO 449 IX=1,4
          ROT(1,J,IX)=RT(1,IX)
449       ROT(2,J,IX)=RT(2,IX)
98        CONTINUE
          ELSE IF (IOVAL.EQ.2) THEN
          BU(1)=NU
          BL(1)=NL
          X(1)=NXI
          BU(2)=NU
          BL(2)=NL
          X(2)=NXI
          BU(3)=PU
          BL(3)=PL
          X(3)=PXI
          BU(4)=PU
          BU(4)=PL
          X(4)=PXI
          CALL MINFUN(N,X,BU,BL,F)
                DO 298 IX=1,4
298             XX(5,IX)=X(IX)
                FF(5)=F
          CALL MINROOT(N,X,RT)
          DO 448 IX=1,4
          ROT(1,5,IX)=RT(1,IX)
448       ROT(2,5,IX)=RT(2,IX)
          BU(1)=NU
          BL(1)=NL
          X(1)=NXI
          BU(3)=NU
          BL(3)=NL
          X(3)=NXI
          BU(2)=PU
          BL(2)=PL
          X(2)=PXI
          BU(4)=PU
          BU(4)=PL
          X(4)=PXI
          CALL MINFUN(N,X,BU,BL,F)
                DO 297 IX=1,4
297             XX(6,IX)=X(IX)
                FF(6)=F
          CALL MINROOT(N,X,RT)
          DO 447 IX=1,4
          ROT(1,6,IX)=RT(1,IX)
447       ROT(2,6,IX)=RT(2,IX)
          BU(1)=NU
          BL(1)=NL
          X(1)=NXI
```

```
      BU(4)=NU
      BL(4)=NL
      X(4)=NXI
      BU(3)=PU
      BL(3)=PL
      X(3)=PXI
      BU(1)=PU
      BU(1)=PL
      X(1)=PXI
      CALL MINFUN(N,X,BU,BL,F)
              DO 296 IX=1,4
296           XX(7,IX)=X(IX)
              FF(7)=F
      CALL MINROOT(N,X,RT)
      DO 446 IX=1,4
      ROT(1,7,IX)=RT(1,IX)
446   ROT(2,7,IX)=RT(2,IX)
      BU(3)=NU
      BL(3)=NL
      X(3)=NXI
      BU(2)=NU
      BL(2)=NL
      X(2)=NXI
      BU(1)=PU
      BL(1)=PL
      X(1)=PXI
      BU(4)=PU
      BU(4)=PL
      X(4)=PXI
      CALL MINFUN(N,X,BU,BL,F)
              DO 295 IX=1,4
295           XX(8,IX)=X(IX)
              FF(8)=F
      CALL MINROOT(N,X,RT)
      DO 445 IX=1,4
      ROT(1,8,IX)=RT(1,IX)
445   ROT(2,8,IX)=RT(2,IX)
      BU(4)=NU
      BL(4)=NL
      X(4)=NXI
      BU(2)=NU
      BL(2)=NL
      X(2)=NXI
      BU(1)=PU
      BL(1)=PL
      X(1)=PXI
      BU(3)=PU
      BU(3)=PL
```

```
              X(3)=PXI
              CALL MINFUN(N,X,BU,BL,F)
                      DO 294 IX=1,4
294                   XX(9,IX)=X(IX)
                      FF(9)=F
              CALL MINROOT(N,X,RT)
              DO 444 IX=1,4
              ROT(1,9,IX)=RT(1,IX)
444           ROT(2,9,IX)=RT(2,IX)
              BU(3)=NU
              BL(3)=NL
              X(3)=NXI
              BU(4)=NU
              BL(4)=NL
              X(4)=NXI
              BU(1)=PU
              BL(1)=PL
              X(1)=PXI
              BU(2)=PU
              BU(2)=PL
              X(2)=PXI
              CALL MINFUN(N,X,BU,BL,F)
                      DO 293 IX=1,4
293                   XX(10,IX)=X(IX)
                      FF(10)=F
              CALL MINROOT(N,X,RT)
              DO 443 IX=1,4
              ROT(1,10,IX)=RT(1,IX)
443           ROT(2,10,IX)=RT(2,IX)
       ELSE IF (IOVAL.EQ.3) THEN
              ILOCAL=N
              DO 96 I=1,ILOCAL
              BU(I)=NU
              BL(I)=NL
              X(I)=NXI
96            CONTINUE
              DO 95 J=1,ILOCAL
              BU(J)=PU
              BL(J)=PL
              X(J)=PXI
              CALL MINFUN(N,X,BU,BL,F)
                      DO 292 IX=1,4
292                   XX(10+J,IX)=X(IX)
                      FF(10+J)=F
              CALL MINROOT(N,X,RT)
              DO 442 IX=1,4
              ROT(1,10+J,IX)=RT(1,IX)
442           ROT(2,10+J,IX)=RT(2,IX)
```

```
                BU(J)=NU
                BL(J)=NL
                X(J)=NXI
 95             CONTINUE
      ELSE IF (IOVAL.EQ.4) THEN
                ILOCAL=N
                DO 93 I=1,ILOCAL
                BU(I)=NU
                BL(I)=NL
                X(I)=NXI
 93             CONTINUE
                CALL MINFUN(N,X,BU,BL,F)
                        DO 291 IX=1,4
 291                    XX(15,IX)=X(IX)
                        FF(15)=F
                CALL MINROOT(N,X,RT)
                DO 441 IX=1,4
                ROT(1,15,IX)=RT(1,IX)
 441            ROT(2,15,IX)=RT(2,IX)
      END IF
 300  CONTINUE
      DO 259 I=1,15
                PRINT *,'F= ',FF(I)
                DO 258 J=1,4
                       PRINT *,'AT X(I) =',XX(I,J)
 258            CONTINUE
                DO 257 J=1,4
                       PRINT *,'ROOTS=',ROT(1,I,J),'+i',ROT(2,I,J)
 257            CONTINUE
 259  CONTINUE

      STOP
      END
C
C
      SUBROUTINE MINFUN(N,X,BU,BL,F)
      INTEGER N,LH,LIW,LW,ISTATE(4), IW(2),NOUT
      DOUBLE PRECISION  ETA, F, FEST, STEPMX, XTOL
      INTEGER  IBOUND, IFAIL, INTYPE, IPRINT, J, MAXCAL
      LOGICAL  LOCSCH
      DOUBLE PRECISION BL(4), BU(4), DELTA(4), G(4), HESD(4)
      DOUBLE PRECISION HESL(6),W(36),X(4)
      EXTERNAL E04HBF, E04JBF, E04JBQ, FUNCT,
     *                MONIT
      PRINT *,'*********************************************'
      LH=6
      LIW=2
      LW=36
```

```
      IFAIL=1
      CALL E04HBF(N,FUNCT,X,NF,DELTA,HESL,LH,
     *             HESD,F,G,IW,LIW,W,LW,IFAIL)
      IPRINT=0
c     PRINT *,'**********************************************'
      LOCSCH=.TRUE.
      INTYPE=0
      MAXCAL=40*N*(N+5)
      ETA=.5e0
      XTOL=0.0e0
      STEPMX=4.0e4
      FEST=4.0
      IBOUND=0
      IFAIL=1
      CALL E04JBF(N,FUNCT,MONIT,IPRINT,LOCSCH,INTYPE,E04JBQ,
     * MAXCAL,ETA,XTOL,STEPMX,FEST,DELTA,IBOUND,BL,BU,X,HESL,
     * LH,HESD,ISTATE,F,G,IW,LIW,W,LW,IFAIL)
      IF (IFAIL.NE.0) WRITE(NOUT,FMT=996) IFAIL
      IF (IFAIL.NE.1) THEN
      PRINT *,'THE FUNCTION AT EXIT IS F=', F
      PRINT *,'at the point ', (X(J),J=1,N)
      END IF
               IF (IFAIL.EQ.2) THEN
                      WRITE (NOUT,FMT=993) (ISTATE(J),J=1,N)
                      WRITE (NOUT,FMT=992) (HESL(J),J=1,LH)
                      WRITE (NOUT,FMT=991) (HESD(J),J=1,N)
      END IF
c     PRINT *, 'For minimization ifail=', IFAIL
996   FORMAT(///' ERROR EXIT TYPE',I3, 'SEE ROUTINE DOCUMENT')
993   FORMAT(' WHERE ISTATE CONTAINS', 4I5, '.')
992   FORMAT(' HESL CONTAINS',/' ', 1P, 6e20.4)
991   FORMAT(' AND HESD CONTAINS', 1P,4e20.4)
      RETURN
      END
C
      SUBROUTINE FUNCT(IFLAG,N,XC,FC,GC,IW,LIW,W,LW)
      DOUBLE PRECISION    FC,XJ
      INTEGER IFLAG,LIW,LW,N
      DOUBLE PRECISION   GC(N),W(LW),XC(N)
      INTEGER   IW(LIW)
      DOUBLE PRECISION DE1,DE3,BB(5)
      COMMON//BB(5)
c     PRINT *, 'IFLAG=',IFLAG
      IF (IFLAG.NE.3) THEN
      DE1=XC(1)*XC(2)*XC(3)*XC(4)
      DE3=XC(1)*XC(2)+XC(1)*XC(4)+XC(3)*XC(4)
c     if (XC(1).eq.0)  PRINT *, '!!!!!!!',DE1,(XC(i),i=1,4)
      FC=(1./XC(1)-BB(2))**2+(DE3/DE1-BB(3))**2
```

```
     *    +((XC(2)+XC(4))/DE1-BB(4))**2+(1/DE1-BB(5))**2
c         PRINT *,(XC(i),i=1,4),FC
          ELSE
          XJ=XC(1)+GC(1)
          DE1=XJ*XC(2)*XC(3)*XC(4)
          DE3=XJ*XC(2)+XJ*XC(4)+XC(3)*XC(4)
          GC(1)=(1./XJ-BB(2))**2+(DE3/DE1-BB(3))**2
     *    +((XC(2)+XC(4))/DE1-BB(4))**2+(1/DE1-BB(5))**2
C
          XJ=XC(2)+GC(2)
          DE1=XC(1)*XJ*XC(3)*XC(4)
          DE3=XC(1)*XJ+XC(1)*XC(4)+XC(3)*XC(4)
          GC(2)=(1./XC(1)-BB(2))**2+(DE3/DE1-BB(3))**2
     *    +((XJ+XC(4))/DE1-BB(4))**2+(1/DE1-BB(5))**2
C
          XJ=XC(3)+GC(3)
          DE1=XC(1)*XC(2)*XJ*XC(4)
          DE3=XC(1)*XC(2)+XC(1)*XC(4)+XJ*XC(4)
          GC(3)=(1./XC(1)-BB(2))**2+(DE3/DE1-BB(3))**2
     *    +((XC(2)+XC(4))/DE1-BB(4))**2+(1./DE1-BB(5))**2
C
          XJ=XC(4)+GC(4)
          DE1=XC(1)*XC(2)*XC(3)*XJ
          DE3=XC(1)*XC(2)+XC(1)*XJ+XC(3)*XJ
          GC(4)=(1./XC(1)-BB(2))**2+(DE3/DE1-BB(3))**2
     *    +((XC(2)+XJ)/DE1-BB(4))**2+(1./DE1-BB(5))**2
          END IF
c         PRINT *,FC,GC(1),GC(2)
c         PRINT *, 'BB(*****)=',(BB(i),i=1,5)
          RETURN
          END
C
C
          SUBROUTINE MINROOT(N,X,RT)
          DOUBLE PRECISION X(N),C(5),TE,RE(5),IM(5),TOL,RT(2,4)
          INTEGER N,M,IFAIL,I
          EXTERNAL C02AEF
          M=N+1
          TE=X(1)*X(2)*X(3)*X(4)
          C(1)=1.
          C(2)=1/X(1)
          C(3)=(X(1)*X(2)+X(1)*X(4)+X(3)*X(4))/TE
          C(4)=(X(2)+X(4))/TE
          C(5)=1/TE
          DO 25 I=1,M
          PRINT *,'C(',I,')=',C(I)
 25       CONTINUE
          TOL=1.E-20
```

```
      IFAIL=1
      CALL C02AEF(C,M,RE,IM,TOL,IFAIL)
      IF (IFAIL.NE.0) THEN
      PRINT *,'THE PROCEDURE HAS FAILED'
      ELSE
             DO 30 I=1,N
              RT(1,I)=RE(I)
              RT(2,I)=IM(I)
C             PRINT *,'ROOT ',I,'=',RE(I),' +i ',IM(I)
 30   CONTINUE
C     PRINT *,'IFAIL FOR FINDING THE ROOTS IS ',IFAIL
      END IF
      RETURN
      END
C
      SUBROUTINE OROOT(N,B)
      DOUBLE PRECISION B(5),RE(5),IM(5),TOL
      INTEGER N,M,IFAIL,I
      EXTERNAL C02AEF
      M=N+1
      IFAIL=1
      TOL=1.E-20
      CALL C02AEF(B,M,RE,IM,TOL,IFAIL)
      IF (IFAIL.NE.0) THEN
              PRINT *,'THE PORCEDURE HAS FAILED'
      ELSE
              DO 35 I=1,N
 35           PRINT *,'OROOT',I,'= ',RE(I),'+i ',IM(I)
C     PRINT *,'IFAIL FOR FINDING THE ROOTS IS',IFAIL
      END IF
      RETURN
      END
C
      SUBROUTINE MONIT(N,XC,FC,GC,ISTATE,GPJNRM,COND,
     *                 POSDEF,NITER,NF,IW,LIW,W,LW)
      INTEGER          NOUT
      DOUBLE  PRECISION COND,FC,GPJNRM
      INTEGER          LIW,LW,N,NF,NITER
      LOGICAL POSDEF
      DOUBLE  PRECISION GC(4),W(36),XC(4)
      INTEGER          ISTATE(4),IW(2)
      INTEGER          ISJ,J
      NOUT=6
c     PRINT *, NITER,NF,FC,GPJNRM
c     PRINT *, '    J          X(J)              G(J)             STATUS
      DO 20 J=1,N
      ISJ=ISTATE(J)
      IF (ISJ.GT.0) THEN
```

```
            PRINT *,'  ', J,XC(J),GC(J),'      FREE'
        ELSE IF (ISJ.EQ.-1) THEN
            PRINT *,'  ', J,XC(J),GC(J),'    UPPER BOUND'
        ELSE IF (ISJ.EQ.-2) THEN
            PRINT *,'  ', J,XC(J),GC(J),'      LOWER BOUND'
        ELSE IF (ISJ.EQ.-3) THEN
            PRINT *,'  ', J,XC(J),GC(J),'     CONSTANT'
        END IF
 20     CONTINUE
        END
```

II.B. FORTRAN code for computing the minimum distance of an unstable polynomial from stability (Fourth order case)

```
        INTEGER   N,I
        DOUBLE PRECISION BL(4), BU(4),X(4),B(5),BB(5),PL,PU,NL,NU,
     *                       PXI,NXI,XX(15,4),FF(15),ROT(2,15,4),F,RT(2,4
        DATA BB/1.,4.,-6.,5.,2./
        COMMON//BB(5)
        EXTERNAL      MINROOT,OROOT,MINFUN
c
c
        N=4
        PRINT *,'COEFFICIENTS OF THE ORIGINAL POLYNOMIAL ARE'
        DO 15 I=1,5
 15     B(I)=BB(I)
        PRINT *,(B(I),I=1,5)
        CALL OROOT(N,B)
        PL=1.e-20
        PU=1.E20
        PXI=10.
C
C
C
        ILOCAL=N
        DO 99 I=1,ILOCAL
            BU(I)=PU
            BL(I)=PL
            X(I)=PXI
 99         CONTINUE
            CALL MINFUN(N,X,BU,BL,F)
            PRINT *, '****&&&****',(X(i),i=1,4)
c           DO 299 IX=1,4
c 299               XX(J,IX)=X(IX)
c                   FF(J)=F
            CALL MINROOT(N,X,RT)
            DO 449 IX=1,4
            ROT(1,J,IX)=RT(1,IX)
```

```
 449            ROT(2,J,IX)=RT(2,IX)
        STOP
        END
C
C
        SUBROUTINE MINFUN(N,X,BU,BL,F)
        INTEGER N,LH,LIW,LW,ISTATE(4), IW(2),NOUT
        DOUBLE PRECISION  ETA, F, FEST, STEPMX, XTOL
        INTEGER  IBOUND, IFAIL, INTYPE, IPRINT, J, MAXCAL
        LOGICAL  LOCSCH
        DOUBLE PRECISION BL(4), BU(4), DELTA(4), G(4), HESD(4)
        DOUBLE PRECISION HESL(6),W(36),X(4)
        EXTERNAL E04HBF, E04JBF, E04JBQ, FUNCT,
     *                MONIT
        PRINT *,'********************************************'
        LH=6
        LIW=2
        LW=36
        IFAIL=1
        CALL E04HBF(N,FUNCT,X,NF,DELTA,HESL,LH,
     *             HESD,F,G,IW,LIW,W,LW,IFAIL)
        IPRINT=0
c       PRINT *,'********************************************'
        LOCSCH=.TRUE.
        INTYPE=0
        MAXCAL=40*N*(N+5)
        ETA=.5e0
        XTOL=0.1e-15
        STEPMX=4.0e4
        FEST=4.0
        IBOUND=0
        IFAIL=1
        CALL E04JBF(N,FUNCT,MONIT,IPRINT,LOCSCH,INTYPE,E04JBQ,
     * MAXCAL,ETA,XTOL,STEPMX,FEST,DELTA,IBOUND,BL,BU,X,HESL,
     * LH,HESD,ISTATE,F,G,IW,LIW,W,LW,IFAIL)
        IF (IFAIL.NE.0) WRITE(NOUT,FMT=996) IFAIL
        IF (IFAIL.NE.1) THEN
        PRINT *,'THE FUNCTION AT EXIT IS F=', F
        PRINT *,'at the point ', (X(J),J=1,N)
        END IF
              IF (IFAIL.EQ.2) THEN
                      WRITE (NOUT,FMT=993) (ISTATE(J),J=1,N)
                      WRITE (NOUT,FMT=992) (HESL(J),J=1,LH)
                      WRITE (NOUT,FMT=991) (HESD(J),J=1,N)
        END IF
c       PRINT *, 'For minimization ifail=', IFAIL
996     FORMAT(///' ERROR EXIT TYPE',I3, 'SEE ROUTINE DOCUMENT')
993     FORMAT(' WHERE ISTATE CONTAINS', 4I5, '.')
```

```
992     FORMAT(' HESL CONTAINS',/' ', 1P, 6e20.4)
991     FORMAT(' AND HESD CONTAINS', 1P,4e20.4)
        RETURN
        END
C
        SUBROUTINE FUNCT(IFLAG,N,XC,FC,GC,IW,LIW,W,LW)
        DOUBLE PRECISION  FC,XJ
        INTEGER IFLAG,LIW,LW,N
        DOUBLE PRECISION GC(N),W(LW),XC(N)
        INTEGER  IW(LIW)
        DOUBLE PRECISION DE1,DE3,BB(5)
        COMMON//BB(5)
c       PRINT *, 'IFLAG=',IFLAG
        IF (IFLAG.NE.3) THEN
        DE1=XC(1)*XC(2)*XC(3)*XC(4)
        DE3=XC(1)*XC(2)+XC(1)*XC(4)+XC(3)*XC(4)
c       if (XC(1).eq.0)  PRINT *, '!!!!!!!',DE1,(XC(i),i=1,4)
        FC=(1./XC(1)-BB(2))**2+(DE3/DE1-BB(3))**2
     *     +((XC(2)+XC(4))/DE1-BB(4))**2+(1/DE1-BB(5))**2
c       PRINT *,(XC(i),i=1,4),FC
        ELSE
        XJ=XC(1)+GC(1)
        DE1=XJ*XC(2)*XC(3)*XC(4)
        DE3=XJ*XC(2)+XJ*XC(4)+XC(3)*XC(4)
        GC(1)=(1./XJ-BB(2))**2+(DE3/DE1-BB(3))**2
     *     +((XC(2)+XC(4))/DE1-BB(4))**2+(1/DE1-BB(5))**2
C
        XJ=XC(2)+GC(2)
        DE1=XC(1)*XJ*XC(3)*XC(4)
        DE3=XC(1)*XJ+XC(1)*XC(4)+XC(3)*XC(4)
        GC(2)=(1./XC(1)-BB(2))**2+(DE3/DE1-BB(3))**2
     *     +((XJ+XC(4))/DE1-BB(4))**2+(1/DE1-BB(5))**2
C
        XJ=XC(3)+GC(3)
        DE1=XC(1)*XC(2)*XJ*XC(4)
        DE3=XC(1)*XC(2)+XC(1)*XC(4)+XJ*XC(4)
        GC(3)=(1./XC(1)-BB(2))**2+(DE3/DE1-BB(3))**2
     *     +((XC(2)+XC(4))/DE1-BB(4))**2+(1./DE1-BB(5))**2
C
        XJ=XC(4)+GC(4)
        DE1=XC(1)*XC(2)*XC(3)*XJ
        DE3=XC(1)*XC(2)+XC(1)*XJ+XC(3)*XJ
        GC(4)=(1./XC(1)-BB(2))**2+(DE3/DE1-BB(3))**2
     *     +((XC(2)+XJ)/DE1-BB(4))**2+(1./DE1-BB(5))**2
        END IF
c       PRINT *,FC,GC(1),GC(2)
c       PRINT *, 'BB(*****)=',(BB(i),i=1,5)
        RETURN
```

```
      END
C
C
      SUBROUTINE MINROOT(N,X,RT)
      DOUBLE PRECISION X(N),C(5),TE,RE(5),IM(5),TOL,RT(2,4)
      INTEGER N,M,IFAIL,I
      EXTERNAL C02AEF
      M=N+1
      TE=X(1)*X(2)*X(3)*X(4)
      C(1)=1.
      C(2)=1/X(1)
      C(3)=(X(1)*X(2)+X(1)*X(4)+X(3)*X(4))/TE
      C(4)=(X(2)+X(4))/TE
      C(5)=1/TE
      DO 25 I=1,M
      PRINT *,'C(',I,')=',C(I)
 25   CONTINUE
      PRINT *,(X(i),i=1,4)
      PRINT *,'&&&&&&&&&&&'
      TOL=1.E-20
      IFAIL=1
      CALL C02AEF(C,M,RE,IM,TOL,IFAIL)
      IF (IFAIL.NE.0) THEN
      PRINT *,'THE PROCEDURE HAS FAILED'
      ELSE
              DO 30 I=1,N
              RT(1,I)=RE(I)
              RT(2,I)=IM(I)
              PRINT *,'ROOT ',I,'=',RE(I),' +i ',IM(I)
 30   CONTINUE
C     PRINT *,'IFAIL FOR FINDING THE ROOTS IS ',IFAIL
      END IF
      RETURN
      END
C
      SUBROUTINE OROOT(N,B)
      DOUBLE PRECISION B(5),RE(5),IM(5),TOL
      INTEGER N,M,IFAIL,I
      EXTERNAL C02AEF
      M=N+1
      IFAIL=1
      TOL=1.E-20
      CALL C02AEF(B,M,RE,IM,TOL,IFAIL)
      IF (IFAIL.NE.0) THEN
              PRINT *,'THE PORCEDURE HAS FAILED'
      ELSE
              DO 35 I=1,N
 35           PRINT *,'OROOT',I,'= ',RE(I),'+i ',IM(I)
```

```
C       PRINT *,'IFAIL FOR FINDING THE ROOTS IS',IFAIL
        END IF
        RETURN
        END
C
        SUBROUTINE MONIT(N,XC,FC,GC,ISTATE,GPJNRM,COND,
     *                    POSDEF,NITER,NF,IW,LIW,W,LW)
        INTEGER           NOUT
        DOUBLE  PRECISION COND,FC,GPJNRM
        INTEGER           LIW,LW,N,NF,NITER
        LOGICAL           POSDEF
        DOUBLE  PRECISION GC(4),W(36),XC(4)
        INTEGER           ISTATE(4),IW(2)
        INTEGER           ISJ,J
        NOUT=6
c       PRINT *, NITER,NF,FC,GPJNRM
c       PRINT *, '    J          X(J)             G(J)            STATUS
        DO 20 J=1,N
        ISJ=ISTATE(J)
        IF (ISJ.GT.0) THEN
                PRINT *,'  ', J,XC(J),GC(J),'      FREE'
        ELSE IF (ISJ.EQ.-1) THEN
                PRINT *,'  ', J,XC(J),GC(J),'      UPPER BOUND'
        ELSE IF (ISJ.EQ.-2) THEN
        PRINT *,'  ', J,XC(J),GC(J),'      LOWER BOUND'
        ELSE IF (ISJ.EQ.-3) THEN
                PRINT *,'  ', J,XC(J),GC(J),'     CONSTANT'
        END IF
  20    CONTINUE
        END
```