DECENTRALIZED CONTROL SYSTEM DESIGN:

DYNAMIC COUPLING FROM A GEOMETRIC VIEWPOINT

by

William H. Bennett

Thesis submitted to the Faculty of the Graduate School
of the University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
1984

ABSTRACT

Title of Dissertation: Decentralized Control System Design: Dynamic Weak Coupling From a Geometric Viewpoint

William Henry Bennett, Doctor of Philosophy, 1984

Thesis directed by: John S. Baras
Professor,
Electrical Engineering Department

This dissertation provides a new, frequency dependent, notion of dynamic weak coupling between subsystems which is useful for the design of decentralized control systems. An abstract geometric Nyquist criterion for multiloop systems is used to develop both a new measure of system stability margin and a new measure of subsystem weak coupling. The measure of stability margin developed has the advantage over standard measures of exposing certain additional internal stability properties of a feedback system. The weak coupling measure is useful for estimating stability properties (and therefore certain control system design objectives) of a decentralized control system and appears to be more generally applicable than other available measures of weak coupling.

The essential topological features of the abstract Nyquist criterion employed in this dissertation involve near intersection between a certain pair of linear subspaces (parameterized by the complex frequency variable s) of the direct sum of all the system inputs and outputs. The measures employed are derived from the idea of the gap between subspaces. Computational methods are provided based on the idea of principal angles between a pair of linear subspaces.

A review of some well known methods for design of decentralized control systems using other notions of subsystem weak coupling is provided. Some examples are included which serve to illustrate the ideas and compare with other well known techniques.

Dedicated to

Marie and Christine

with love

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude for the encouragement and guidance of Professor John S. Baras. His inspiration has helped me to formulate and achieve strong career objectives.

I would also like to thank Professor P.S. Krishnaprasad for his councel and numerous discussions about the geometric system theory. It is a pleasure to also thank the other members of my dissertation committee including: Professor W.S. Levine, Professor A. Tits, and Professor C. Berenstein.

Special thanks are due for Dr. K.T. Alfriend and Dr. R.L. Lindberg for their support of this research during my employment with the Naval Research Laboratory.

Last -- but not least -- I want to thank my wife, Marie, for her love and support throughout the last several years.

TABLE OF CONTENTS

LIST OF FIGURES

1. Introduction

In this dissertation we focus on certain design issues for control of large composite systems which can be modeled as interconnections of simpler dynamical systems. Our treatment is based on the available utility of frequency response models for linear, time-invariant dynamic systems. Design methods yielding decentralized control strategies for linear systems are considered which exploit various notions of dynamic weak coupling between subsystems.

The efforts described in this dissertation were originally motivated by a family of design methods for multivariable feedback control (MIMO) called Inverse Nyquist Array (INA) methods which were originally developed by Rosenbrock and his colleagues. Various extensions of these methods have been developed and are discussed in chapter 3.

This line of research began with [BE2] in which the present author extended the basic notion of weak coupling (diagonal dominance of transfer functions) used in INA methods to a more general setting appropriate for decentralized control. We sought in [BE1-2] (and others have also sought [LI1-2],[NW4]) to provide a methodology for design of decentralized control where certain partitions can be considered "natural" with respect to some notion of dynamic *weak coupling* between certain dominant subsystems based on the system frequency response. Such a methodology would allow the design process itself to be decentralized in the sense that different methods could be applied to design each local controller. This approach borrows the loop-at-a-time strategy of INA methods without employing the cumbersome requirement to modify system weak coupling by compensation (thus realizing a true decentralized control structure.) We focus on notions of dynamic weak coupling which are frequency dependent since we are interested in feedback control of complex systems using simple control structures with guaranteed stability margin properties.

Our original motivation for this research was a particular technical problem associated with extending INA type design methods to decentralized design. A particularly useful aspect of the INA method for design is its ability to provide local estimates for the contribution to the system stability margin properties. This allowed the designer to assess the sensitivity of the system performance to changes in each individual (scalar) gain. This is accomplished by recognizing certain inclusion regions for relevant Nyquist loci are centered on the Nyquist loci for each individual loop gain. To extend this to general decentralized feedback we need to consider some appropriate notion of a MIMO Nyquist contour.

Our approach is to exploit a certain analysis for closed loop systems suggested by algebraic geometry. Here we employ an abstract Nyquist contour living on a complex manifold (known as the Grassman manifold). Along the way, we develop a new geometric stability margin for feedback systems which is shown to have certain desirable advantages over the classical analysis even for the SISO case. Moreover, the geometric viewpoint which is based on the topology of the Grassman manifold permits the construction of appropriate estimates for the decentralized control problem.

We start by considering in chapter 2 the nature of feedback control based on frequency response models for linear systems. Certain limitations on the performance of feedback control are shown to be evident from the frequency response. Central to frequency domain design methods is the Nyquist stability criterion. We review the technical basis for this powerful result and several extensions. We review the concept of system *gain and phase margins* which will play a central role in the sequel.

In chapter 3 we review several results in design of decentralized control based on various notions of dynamic weak coupling between subsystems. Since several of these methods originate from the INA method of Rosenbrock we review this method. We focus on several recent extensions of the INA method and highlight their similarities and limitations. Finally, we conclude this chapter with a discussion of the fundamental technical problem which originally motivated this research.

In chapter 4 we discuss the development of a new *geometric* notion of a system stability margin based on frequency response data. We start with some background material on the geometric system theory popularized by Hermann, Martin, Brockett, Byrnes, and others. We consider the topology of the Grassman manifold (which is the natural setting for much of the geometric system theory) based on the *gap-metric*. Using this metric we construct a geometric stability margin for feedback control and consider its utility by way of some examples.

In chapter 5 we employ again the gap-metric and the geometric stability margin to develop a new notion of weak coupling between subsystems based on frequency response data. This notion of weak coupling involves the relative orientation of certain "curves" on the Grassman manifold which are the abstract representations for certain system frequency responses. In this setting we provide "local" estimates (at the subsystem level) for the contributions to the overall system stability margin with decentralized control. For comparison we conclude with an example originally considered in [BE2].

2. Design Requirements for Feedback Compensation: The Concept of Frequency Shaping

2.1. Feedback Compensation and Frequency Response Models

The concept of feedback compensation is fundamental in several engineering disciplines. From very early work on governors for rotating machinery to electronic circuit design the use of feedback for compensation has been essential [MA2]. Although recursive structures are apparent in natural processes, the introduction of such structures with the goal of altering dynamic behavior is a fundamental engineering design problem.

2.1.1. Benefits from Feedback Compensation

A typical introduction to control system design involves a demonstration of the benefits achievable with feedback compensation. In particular, sensitivity to errors in the dynamic models for components of a dynamical system configured in closed loop will be greatly reduced over a similar dynamical system configured as a cascade of components.

We can state rather broadly the various benefits from feedback compensation as follows:

(i) regulation or disturbance rejection,

(ii) servo-following,

(iii) insensitivity to incremental model parametric changes,

(iv) robustness with respect to large model changes and nonlinear effects.

Moreover, all these benefits can often be met to some extent simultaneously.

Intuitively, such desirable behavior is achieved using feedback only when the recursive dynamics are "stable"; i.e., signals convergence to steady state values under the influence of constant exogenous influences. To allow precise quantitative analysis of this behavior various theories have been developed; all of which - as might be expected - characterize quantitatively the dynamic convergence properties of feedback systems. The point is that the cost of using feedback is that the resulting system can oscillate or fail to converge due to improper design. This fundamental heuristic observation has lead to the development of a large body of frequency domain analysis for feedback systems with the primary goal of characterizing and predicting dynamic convergence of feedback systems.

2.1.2. Frequency Domain Models and Feedback

Due to the inherent robustness properties of feedback control practicing engineers can often employ simple linear models for system dynamics. In this dissertation we deal with systems which are linear, realizable (causal), and time-invariant. Thus the output time history $y(t)$ can be obtained from the input or forcing function $u(t)$ via a convolution operation,

$$y(t) = \int_0^{\infty} w(t-\tau)\, u(\tau) d\tau ,$$

whose kernel, $w(t)$, is the impulse response of the system which is assumed initially (at $t=0$) at rest. For such systems the unilateral Laplace transform of the impulse response,

$$h(s) = \oint_C w(\tau) e^{-s\tau} d\tau ,$$

defines the frequency domain model or transfer function of the system. The assumption of causality implies that $h(s)$ converges for s in some half plane which includes $\mathrm{Re}\,s > \alpha$, for some α and the function $h(s)$ is analytic there. The Laplace transform exits whenever the closed contour C is contained in this half plane.

2.1.3. Limitations on Frequency Response of Causal Systems and Feedback Performance

For the class of systems we are considering, many fundamental properties can be represented via path integrals in the complex plane. The significance of this for engineering is that such path integrals can be computed directly from experimental data on the system. Certain practical assumptions which further restrict the class of systems can be used to characterize some fundamental limitations in the use of feedback directly in terms of such path integrals.

The first example is how causality (a fundamental limitation on the time response of a system) is reflected in the frequency domain.

Theorem 2.1: (The Paley-Wiener Criterion)

Let $w(t) = 0$ for $t<0$; i.e. $w(t)$ represents the impulse response of a causal system. Given the Fourier transform (if it exists),

$$H(j\omega) = \int_0^{\infty} w(t) e^{-j\omega t} dt,$$

if $H(j\omega)$ satisfies

$$\int_{-\infty}^{+\infty} |H(j\omega)|^2 d\omega < \infty \qquad (2.1)$$

and $H(j\omega)$ is analytic $\omega \in \mathbf{R}$ then

$$\int_{-\infty}^{+\infty} \frac{|\ln|H(j\omega)||}{1+\omega^2} d\omega < \infty. \tag{2.2}$$

Conversely if (2.1) and (2.2) hold a phase response $\varphi(\omega)$ can be found such that $H(j\omega) = |H(j\omega)| e^{j\varphi(\omega)}$ is the transform of a causal system.

Proof: (cf. [ZA1, pp.423]).

This well known result places clear restrictions on the phase response of a causal systems. The relationship between magnitude and phase can be further clarified by considering a specific physically motivated class of transfer functions.

Definition : The class X of transfer functions T(s) satisfy:

(i) $T(s)$ is analytic in $\text{Re}s > 0$.

(ii) $\overline{T(s)} = T(\bar{s})$ for $\text{Re}s > 0$.

(iii) $T(s)$ has finitely many singularities on the $j\omega$ axis such that at each singularity $j\omega_0$

$$\lim_{s \to j\omega_0} (s - j\omega_0) T(s) = 0$$

(i.e., they are simple poles),

(iv) As $|s| \to \infty$, $\frac{T(s)}{s} \to 0$ uniformly in $|\arg s| \le \frac{\pi}{2}$.

For transfer functions in the class X it is possible to characterize the imaginary part directly in terms of a path integral involving the real part only.

Theorem 2.2: If $T(s)$ is in class X then for any ω_1

$$\operatorname{Im} T(j\omega_1) = \frac{2\omega_1}{\pi}\int_0^{\infty} \frac{\operatorname{Re} T(j\omega) - \operatorname{Re} T(j\omega_1)}{\omega^2 - \omega_1^2} d\omega \tag{2.3}$$

Proof: (cf. [ZA1, pp.430]).

The significance of this theorem follows by considering $T(s) = \ln H(s)$. Then one can state a definition for the notion of a minimum phase transfer function directly in terms of a path integral.

Definition : [ZA1, pp.435] If H(s) is in class X and satisfies Paley-Wiener criterion then we say H(s) is minimum phase if and only if its phase response is

$$\arg H(j\omega_1) = \frac{2\omega_1}{\pi}\int_0^{\infty} \frac{\ln|H(j\omega)| - \ln|H(j\omega_1)|}{\omega^2 - \omega_1^2} d\omega. \tag{2.4}$$

The well known algebraic characterization of this idea for rational transfer functions is that the zeros of $H(s)$ (the roots of the numerator polynomial in a coprime representation of $H(s)$) are contained in Re s<0. The above definition is clearly more general and also potentially more practical given experimental input/output data.

Finally, we will state a result whose practical significance will be postponed until section 2.2. Consider a transfer function from class X, $H(s) = R(s) + jI(s)$ with $R(s)$ and $I(s)$ both real valued. $H(s)$ has a series expansion about $s = \infty$

$$H(s) = R_\infty + \frac{I_\infty}{s} + \frac{R_2}{s^2} + \frac{I_2}{s^3} + \cdots \tag{2.5}$$

Theorem 2.3: [HO2,pp.306] Given $H(s)$ in class X and $H(s) = R(s) + jI(s)$, with $R(s)$, $I(s)$ both real, then,

$$\int_0^{\infty}[R(j\omega)-R_{\infty}]d\omega = \frac{\pi}{2}I_{\infty}. \qquad (2.6)$$

Proof: Because the integrand is holomorphic in the right half plane we can apply Cauchy's theorem. Thus $\oint_C[H(s)-H(\infty)]ds$ = $\oint_C[H(s)-R_{\infty}]ds$ = 0, where C is a closed contour consisting of a portion of $j\omega$ axis, $-jR<j\omega<jR$, and a semicircular contour in right half plane of radius R centered at s=0 and R is allowed to grow without bound. So the path integral consists of two parts,

$$-\lim_{R\to\infty}\int_{-R}^{R}(H(j\omega)-R_{\infty})j\,d\omega + \lim_{R\to\infty}\int_{-\pi/2}^{\pi/2}(H(s)-R_{\infty})j\,d(Re^{j\vartheta}) \qquad (2.7)$$

$$= -I_1 + I_2 = 0.$$

Since by assumption $H(.)$ is in class X, $R(s)=R(-s)$ and $I(s)=-I(-s)$ and we can write

$$I_1 = j2\int_0^{\infty}(R(j\omega)-R_{\infty})d\omega$$

and

$$I_2 = \lim_{R\to\infty}\int_{-\pi/2}^{\pi/2} I_{\infty}j\,d\vartheta = j\pi I_{\infty}.$$

And the result follows directly.

■

2.2. Broadband Objectives in Feedback Design

2.2.1. Sensitivity Reduction

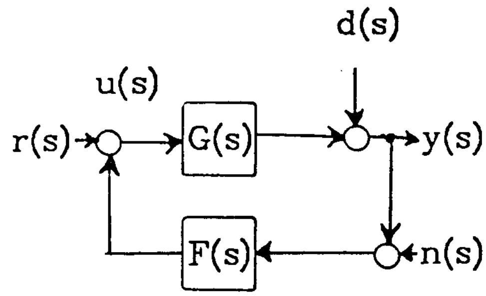


Figure 2.1: Configuration For Feedback

A basic model for feedback compensation is shown in figure 2.1. In the general case $G(s)$ will be a $p \times m$ rational transfer function matrix representing the (linear) plant model and $F(s)$ is an $m \times p$ rational transfer function matrix representing the controller or compensator. The m-vectors $y(s)$, $d(s)$ and $n(s)$ are respectively the p-vectors $r(s)$ and $u(s)$ are respectively the command inputs and control signals.

For the purposes of the following discussion on feedback properties we will assume (temporarily) $m = p$ and that $F(s)$ and $G(s)$ are both invertible over the field of rational functions. The basic issues of feedback design involve choosing $F(s)$ to achieve some desired system response, $y(s)$, to the commands, $r(s)$, subject to considerations of the effect of disturbances $d(s)$ and sensor noise, $n(s)$. Additionally restrictions on internal loop dynamics (e.g., $u(s)$) must be considered in any complete design methodology. The following relations can be derived from fig. 2.1:

$$y = [G^{-1} + F]^{-1} r - [F^{-1} G^{-1} + I]^{-1} n + [I + GF]^{-1} d \tag{2.8}$$

$$u = [I + FG]^{-1} r - [G + F^{-1}]^{-1} n - [G + F^{-1}]^{-1} d. \tag{2.9}$$

Where we recognize the *sensitivity operators* as follows:

$$S_1 \triangleq [I + GF]^{-1},$$

$$S_2 \triangleq [I + FG]^{-1}.$$

Additionally we have the relations

$$I - S_1 = [I + F^{-1}G^{-1}]^{-1} \tag{2.10}$$

$$I - S_2 = [I + G^{-1}F^{-1}]^{-1},$$

and

$$G S_2 F = I - S_1,$$

$$F S_1 G = I - S_2.$$

Thus the relation between the system output and the command input, sensor noise, and load disturbance (2.3) can be written

$$y = S_1 G r - (I - S_1) n + S_1 d. \qquad (2.8')$$

Clearly S_1 and $(I - S_1)$ cannot both be small (in any suitable norm) over identical frequency bands. Also the ability to "scale" the system response y to follow commands r is severely limited when r and d have the same spectral content. On consideration of internal loop dynamics we see from (2.4) that

$$u = S_2 r + F S_1 n + F S_1 d. \qquad (2.9')$$

Considering the relation (2.10) we see the potential for tradeoff in feedback design based on the spectral content of $r(s)$, $n(s)$, and $d(s)$. Design flexibility is greatest when $r(s)$, $n(s)$, and $d(s)$ are each band limited and have spectral content which is "separated". In classical feedback design for single input single output (SISO) (i.e., $p=m=1$) these frequency shaping ideas form the basis for standard engineering design (cf. [DO2]).

Returning to the SISO closed loop configuration ($p=m=1$) we refer to the sensitivity operator $S=S_1=S_2$. With this viewpoint for feedback compensation, it is clear that the design objectives can be stated in terms of frequency dependent constraints on the sensitivity operator representing desirable spectral content of the endogenous signals $u(t)$ given the spectral content of the exogenous signals $r(t)$, $n(t)$, and $d(t)$. This then is the concept of *frequency shaping objectives* in feedback design.

2.2.2. Physical Considerations in Frequency Shaping Design

In addition to the considerations discussed in Section 2.2.1 there are more subtle constraints arising from considerations of realizability and physical limitations on available components which in practice characterize the limitations in feedback design. As discussed in section 2.1.3 the assumption of causality impacts the achievable gain and phase responses of real systems. In this section we consider the impact of causality and some additional physical assumptions on feedback performance.

From section 2.2.1 we are concerned with the achievable balance between regions of the frequency axis where

(i) $|S(j\omega)|<1$ or $|1+fg(j\omega)|>1$ (corresponding to sensitivity reduction) and

(ii) $|S(j\omega)|>1$ or $|1+fg(j\omega)|<1$.

The following relevant conclusion is attributed to Horowitz [H02]. It is based on the observation that in practice any loop transmission $fg(s)$ is equivalent to a composite (series connection) of at least two or more separate physical systems (say $f(s)$ and

$g(s)$) each of which have the property that $|g(s)|$ (resp. $|f(s)|$) approaches zero as $s\to\infty$; i.e., they are systems with finite available power at the output. For instance if $f(s)$ and $g(s)$ are rational transfer functions then they are strictly proper (have relative degree of denominator strictly greater than numerator). Thus Horowitz observes that "almost every practical loop transmission has at least two more poles than zeros" [HO2, pp.306].

Corollary 2.4: For practical causal loop transmissions the following balance on feedback sensitivity holds.

$$\int_0^\infty \ln|1+fg(j\omega)|\,d\omega = 0. \qquad (2.11)$$

Proof: Let $H(s)=\ln[1+fg(s)]$ and apply theorem 2.3. Horowitz's observation allows that $R_\infty = I_\infty = 0$. ∎

In [HO3] Horowitz and Shaked provide a dissertation on the relative merits of design methods which appear to ignore this restriction and as a result claim excellent performance. In particular, the case of state variable feedback is examined.

This equation appears to focus attention near the boundary between the two regions described above (i.e., where $|1+fg(j\omega)|=1$) which is often referred to as "gain crossover". Furthermore, the balance given suggests possible tradeoffs between the "breadth" and "depth" of these regions. However, this is complicated (again due to causality) based on the following result due to Bode.

Corollary 2.5: [HO2, pp.313] For $fg(s)$ a minimum phase transfer function,

$$\arg[fg(j\omega_1)] = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{d\ln|fg(j\omega)|}{d\omega}\left(\ln\coth\frac{|\omega|}{2}\right)d\omega. \qquad (2.12)$$

Clearly increased rate of decay in $fg(s)$ is achieved only at the expense of additional phase lag in $fg(s)$. As we will see in the next section this may be crucial in the region where $|1+fg(j\omega)|<1$.

2.3. Stability Margins and Their Role in Classical Frequency Domain Design

The fundamental analysis which makes frequency domain design for feedback practical employs a collection of tools which permit the designer to predict the relative stability properties of the closed loop system based on physically measurable quantities. The historical evolution of technical contributions to the development of these tools has been surveyed recently in [MA2]. In this section, we provide the background necessary to understand the contributions contained in chapter 4. We review the technical limitations inherent in the natural extension of stability analysis for SISO systems to the MIMO case by exploiting certain algebraic structures. In particular the use of a matrix return difference will be examined. The essential technical aspects of the Nyquist criterion and its applicability to MIMO systems will next be examined. Finally, the notions of gain and phase margin employed in SISO design are examined along with various extensions to MIMO problems.

2.3.1. Limitations and Extensions of the Return Difference

We can consider two classes of feedback structures. Single-input, single-output (SISO) feedback obeys the relations

$$y(s)=g(s)e(s)\ ,\ e(s)=u(s)-f(s)y(s) \tag{2.13}$$

where $u(s)$, $y(s)$, and $e(s)$ are unilateral Laplace transforms of signals representing the input, output, and error respectively.

Multi-input, multi-output (MIMO) feedback obeys the relations:

$$y(s) = G(s)e(s) \quad , \quad e(s) = u(s) - F(s)y(s), \tag{2.14}$$

where $u(s), e(s) \in R^m(s)$† and $y(s) \in R^n(s)$. Thus $f(s), g(s) \in R(s)$ are transfer functions of the feedback compensator and plant, respectively. For MIMO $F(s) \in R^{m \times p}(s)$ and $G(s) \in R^{p \times m}(s)$ are the obvious matrix extensions. The resulting closed-loop transfer functions are

$$h(s) = \frac{g(s)}{1 + f(s)g(s)} \qquad \text{(SISO)}$$

and

$$H(s) = G(s)[I_m + F(s)G(s)]^{-1} = [I_p + G(s)F(s)]^{-1} G(s) \, \text{(MIMO)}$$

Since the fundamental issues in feedback design with which we are concerned center on the questions of stability we start by clarifying this point. We restrict the question initially to lumped (finite dimensional) models moving rational transfer functions. For the purposes of engineering we will be most concerned with *exponential stability.*

Definition : [DE1] A transfer function $H(s) \in R^{p \times m}(s)$ is exponentially stable if and only if (i) it is proper, and (ii) all its poles have negative real parts.

Fact: If a transfer function, $H(s)$, of a causal system is exponentially stable, then given

---

† We use R(s) to denote the field of rational functions in s with coefficients in $\mathbf{R}$ while R[s] denotes the ring of polynomials over $\mathbf{R}$. $R^m(s)$ is the appropriate m-dimensional vector space extension over this field.

$$W(t) = \frac{1}{2\pi j} \int_{\alpha - j\infty}^{\alpha + j\infty} H(s) e^{st} ds \,, \tag{2.15}$$

the impulse response of the system, there exists constants $k>0$ and $\alpha>0$ depending on $H(s)$ such that,

$$\|W(t)\| \leq k\, e^{-\alpha t} \tag{2.16}$$

for all $t>0$ [DE1].

At this point we make a further restriction which we will later relax.

Assumption : $F(s)=F$ is nondynamic constant feedback gain matrix. This is clearly a simplification which in the light of Horowitz criterion is a deviation from reality. Nevertheless it serves to clarify the following technical points.

Without any assumption as to $g(s)$ or $G(s)$ being proper we can state the following.

Theorem 2.6: [DE1, pp.60] Let $g(s) \in R(s)$ and $G(s) \in R^{p \times m}(s)$. Then with $h(s)$ and $H(s)$ as above we can state:

(i) $p_c \in \mathbf{C} \cup \{\infty\}$ is a pole of $h(s)$ if and only if p_c is a zero of $1 + fg(s)$.

(ii) If $p_c \in \mathbf{C} \cup \{\infty\}$ is a zero of $\det[I_m + FG(s)]$ then p_c is a pole of $H(s)$.

(iii) If $p_c \in \mathbf{C} \cup \{\infty\}$ is a pole of $H(s)$ then either p_c is a zero of $\det[I_m + FG(s)]$ or p_c is a pole of $G(s)$.

(iv) If $G(s)$ is exponentially stable, then $H(s)$ is exponentially stable if and only if $\det[I_m + FG(s)] \neq 0$ for $s \to \infty$ and the zeros of $\det[I_m + FG(s)]$ all have negative real parts.

Proof : (cf. [DE1,pp.61]).

Remark: We make implicit use of the fact that $\det[I_m + FG(s)] = \det[I_p + G(s)F]$ throughout. (cf. theorem 4.13 for further details.)

The above theorem suggests some limitations in the extension of SISO stability analysis to the MIMO case. The gap is however closed in the following case.

Theorem 2.7: [DE1,pp.67] Let $G(s) \in \mathbb{R}_p^{p \times m}(s)$ be proper and $F \in \mathbb{R}^{m \times p}$ constant. Given a factorization $G(s) = N(s)D^{-1}(s)$ where $N(s) \in \mathbb{R}^{p \times m}[s]$ and $D(s) \in \mathbb{R}^{m \times m}[s]$ with $N(s)$ and $D(s)$ coprime. Assume $\det[I_m + FG(\infty)] \neq 0$. Then $H(s) = G(s)[I_m + FG(s)]^{-1}$ is exponentially stable if and only if $\det[D(s) + FN(s)] \in \mathbb{R}[s]$ has all its zeros with negative real parts. Moreover, $H(s)$ is proper and p_c is a pole of $H(s)$ if and only if p_c is a zero of $\det[D(s) + FN(s)]$.

Proof : (cf.[DE1]).

As further clarification we state the following.

Corollary 2.8: With the assumptions of the above theorem,

$$\det[\, I_m + FG(s)\,] = \frac{\det G(s)}{\det H(s)}$$

$$= \det[I_m + FG(\infty)] \prod_{i=1}^{n} \frac{(s - p_{ci})}{(s - p_{oi})} \tag{2.17}$$

where $\{p_{oi},\ i=1, \cdots, n\}$ (resp. $\{p_{ci},\ i=1, \cdots, n\}$) are the open (resp. closed) loop poles of $G(s)$ (resp. $H(s)$), and n is the McMillan degree of $G(s)$.

This clearly shows how cancellations in the factors of $\det[I + FG(s)]$ can cause problems. However under the assumptions of the theorem if $G(s)$ is exponentially stable then $H(s)$ is exponentially stable if and only if all the zeros of $\det[I_m + FG(s)]$

have negative real parts.

We next consider the more realistic case by relaxing the assumption that $F(s)=F$. By allowing $F(s) \in R^{p \times m}(s)$ we can introduce further problems due to cancellations even for SISO feedback. In this case we will be concerned with two classes of poles. We refer to *internal* poles as the poles of the transfer function

$$H_e(s) = [I_m + F(s)G(s)]^{-1}$$

where

$$e(s) = H_e(s)u(s) \tag{2.18}$$

and *external* poles as the poles of the closed-loop transfer function

$$H_y(s) = G(s)[I_m + F(s)G(s)]^{-1}$$

where

$$y(s) = H_y(s)u(s). \tag{2.19}$$

Exponential stability follows if both the internal and external poles have negative real parts. Clearly the internal poles have been canceled in forming $H_y(s)$. To see how consider SISO case. Let $g(s)=n_o(s)/d_o(s)$ and $f(s)=n_f(s)/d_f(s)$. Where the pairs $(n_o(s), d_o(s))$ and $(n_f(s), d_f(s))$ are each relatively coprime. However with,

$$h_e = (1+fg)^{-1} = \frac{d_o\, d_f}{n_f\, n_o + d_f\, d_o} \tag{2.20}$$

$$h_y = \frac{g}{(1+fg)} = \frac{n_o\, d_f}{n_f\, n_o + d_f\, d_o} \tag{2.21}$$

it is clear that $d_o d_f$ (resp. $n_o d_f$) and $n_f n_o + d_f d_o$ are not necessarily coprime due to possible cancellations between n_o, d_f and

n_f, d_o.

Now consider the MIMO case. Given right (resp. left) coprime factors $G=N_o D_o^{-1}$ (resp. $F=D_f^{-1}N_f$) with $N_o \in R^{p\times m}[s]$, $N_f \in R^{m\times p}[s]$, $D_o, D_f \in R^{m\times m}[s]$ we can write the resulting transfer functions of interest as,

$$H_e = D_o[D_f D_o + N_f N_o]^{-1}D_f \tag{2.22}$$

$$H_y = N_o[D_f D_o + N_f N_o]^{-1}D_f . \tag{2.23}$$

Remark : Without loss of generality the coprime factors (N_o, D_o) and (N_f, D_f) can be chosen such that $\det D_o \neq 0$ and $\det D_f \neq 0$. We assume that $\det[I_m + F(\infty)G(\infty)] \neq 0$ which is consistent with assuming, for instance, that G, F, H_e, H_y are all proper. Under these conditions $\det[D_f D_o + N_f N_o] \neq 0$ since H_e is bounded and non-singular for $|s| \to \infty$.

The significance of $\det[D_f D_o + N_f N_o]$ is given by the following theorem.

Theorem 2.9: Let $\{p_{e,i}: i=1,\ldots,n_e\}$ (resp. $\{p_{y,i}: i=1,\ldots,n_y\}$) denote the set of poles, counting multiplicity, of the transfer function H_e (resp. H_y). Then the roots of $\det[D_f D_o + N_f N_o] = 0$ form the set $\{p_{e,i}: i=1,\ldots,n_e\} \cup \{p_{y,i}: i=1,\ldots,n_y\}$.

Proof : Using the factorizations given we rewrite (10) and (11) by introducing auxiliary (internal) variables

$$D_o v = e \,,\; y = N_o v \,,$$

$$w = N_f y \,,\; D_f u - D_f e = w \,.$$

This system of equations can be rewritten

$$(N_o N_f + D_f D_o) v = D_f u$$

$$y = N_o v$$

$$e = D_o v$$

$$w = N_f y = N_f N_o v .$$

Clearly $\det(N_f N_o + D_f D_o) = 0$ is the characteristic equation for the feedback system.

We close this section by discussing the physical interpretation of the quantity $I_m + F(s)G(s)$ (or $I_p + G(s)F(s)$) which are generalizations in terms of matrices of the return difference. The role of these terms in stability analysis is readily apparent from the preceding discussion. Also, (and this will be important later) the potential for hidden modes (cancellations) in forming the return difference is apparent even in the SISO case and further complicated by MIMO feedback.

Definition : Consider the feedback system shown in figure 2.2 which

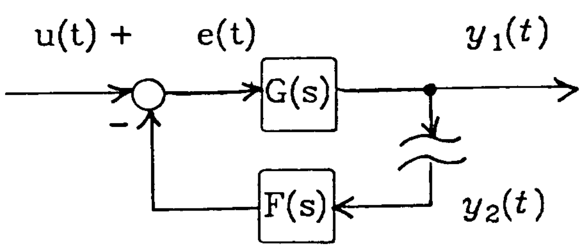


Figure 2.2: The General Feedback Configuration

represents the feedback configuration of (2.14). Associated with the point 1 we define the return difference as

$$R(s) = I_p + G(s)F(s) . \tag{2.24}$$

The physical significance of this terminology follows by the following experimental procedure for measuring the response $R(s)$. First cut the feedback loop at point 1. Set $u(t)=0$. Inject the signal $y_2(t)$ and measure the response $y_1(t)$. Then,

$$\begin{aligned} y_2(t) - y_1(t) &= y_2(t) + G(s)F(s)y_2(t) \\ &= R(s)y_2(t). \end{aligned} \tag{2.25}$$

2.3.2. Stability Analysis and the Nyquist Criterion

The utility of the Nyquist criterion and associated stability tests of Bode and Nichols for engineering stems from two practical observations:

1) the test is based on data which can be obtained directly from measurements on the actual system in contrast to for instance Routh-Hurwitz tests which are based on data which must be derived by interpolation from measurable quantities
2) The test provides additional information about the relative stability of the feedback system which has proven useful in designing compensation.

The second observation is the subject of section 2.3.3. In this section we review the technical basis for the Nyquist criterion. The utility of the method has motivated various research in providing extensions which enlarge the class of systems for which the test is valid. We include some background on more relevant extensions here for completeness.

The fundamental technical approach most commonly used to prove the Nyquist criterion for SISO case involves the principle of the argument from complex analysis. This well known result gives the difference between the number of zeros and poles of a rational function contained in a given closed region in terms of a winding number. The winding number is obtained by evaluating the function (e.g., the return difference) along a closed contour which is the boundary for the region. With respect to the usual concerns for stability we construct a closed elementary contour, D, in the complex plane consisting of a relatively large portion of the $j\omega$ axis, $-j\omega_0<j\omega<j\omega_0$, and a semi-circle centered at the origin of radius ω_0. If poles or zeros of the open loop transfer function exist on the imaginary axis the contour D is modified by indenting into the left half plane around these points. The traditional (SISO) Nyquist locus Γ is then the image of D under the map $fg(s)$.

Theorem 2.10 (Nyquist Stability Criterion): Given $g(s)\in R(s)$ and $f(s)=f$ construct the appropriate contour D as above. Let $fg(s):D\rightarrow\Gamma$. Let p_o (resp. p_c) be the number of open-loop (resp. closed loop) poles contained in the closed right half plane. Then the closed loop system is exponentially stable if and only if

$$N(\Gamma;-1) = -p_o.^{\dagger} \qquad (2.26)$$

Remark: A natural assumption for physical systems which is implicit in the use of Nyquist stability tests is that the loop transfer function is a *proper* rational function. The technical requirement for this assumption become obvious as one considers the construction of the closed contour D. This contour is required to enclose a finite area in the right half plane within which all closed loop poles relating to stability must exist. Clearly, for nonproper transfer

† $N(C;z)$ denotes the number of clockwise encirclements of the closed contour C about the point z in $\mathbb{C}$.

functions there will be some number of system poles at $s = \infty$.

Proof : Note that Γ is a translated version of Γ' where $1+fg(s):D \to \Gamma'$ so that $N(-\Gamma;-1) = N(\Gamma';0)$. Then by principle of argument applied to the return difference with eqn. (2.17).

$$N(\Gamma';0) = p_c - p_o .$$

∎

Remark: Even at this fundamental level the Nyquist test may not be conclusive for $f(s)$ dynamic because as discussed above cancellations can occur in forming $1+f(s)g(s)$. However such cancellations can often be treated as special cases in various design methods where an intimate knowledge of the pole/zero structure of the plant and compensation is assumed available. Such methods however work well only for low order systems (and those which can be approximated by low order models).

It is important to recognize that the practical significance of Nyquist criterion has motivated a massive amount of research focusing on extending the applicability to (1) systems with distributed parameter effects, (2) systems with non-linear effects, and (3) systems using MIMO feedback. Although much of this work has been highly successful various technical limitations remain. We hope to make some contribution to this general thrust in this dissertation.

The reference [DE1] gives a comprehensive discussion of these extensions. In considering stability of systems with nonlinearities a different notion of stability is employed. In this case a system is thought of as a mapping from a function space of input waveforms to another function space of output waveforms. The system is called bounded-input/bounded-output (BIBO) stable if for all bounded inputs only bounded outputs are produced. It is well known that for linear, lumped-parameter, systems, BIBO stability is equivalent to exponential stability and we will not discuss this further.

Distributed parameter effects have been included in Nyquist type tests by considering a class of systems whose impulse responses are generalized functions in the sense that they can contain a sequence of Dirac delta (impulse) functions , $\delta(t)$, [DE1]. For instance

$$w(t) = \begin{cases} w_a(t) + \sum_{i=0}^{\infty} w_i \delta(t - t_i), & t \geq 0 \\ 0, & t < 0 \end{cases} \tag{2.27}$$

with $w_a(t) \in L_1(\mathrm{R}_+)$, $w_i \in \mathrm{R}$, $\sum_{i=1}^{\infty} |w_i| < \infty$ is such a possible impulse response. Desoer and Vidyasagar [DE1] provide a Nyquist type stability test for such systems based on constructing a convolution algebra of impulse response functions (where convolution is the product operation). Their result, stated simply, provides necessary and sufficient conditions for stability (in BIBO sense) of a feedback loop with $w(t)$ as above in terms of the condition

$$\inf_{\mathrm{Re}\, s \geq 0} |1 + f g(s)| > 0. \tag{2.28}$$

where $g(s)$ is the Laplace transform of $w(t)$ [DE1,pp.91]. This can be checked from a Nyquist plot.

Of fundamental importance in this dissertation are the various extensions of the Nyquist criterion for MIMO feedback. Rosenbrock was first to provide a complete statement of a Nyquist type criterion generalized to cover MIMO feedback. His approach focused on providing a rational function $\varphi(s)$ whose zeros (resp. poles) are the required closed loop poles (resp. open loop poles). One way to do this (not exactly the approach used by Rosenbrock) is to first obtain coprime factorizations for $G(s)=N_G(s)D_G^{-1}(s)$, the plant, and $F(s)=D_F^{-1}(s)N_F(s)$, the compensator. Then an appropriate choice for $\varphi(s)$ is

$$\varphi_1(s) = \det[I_m + F(s)G(s)]\det D_G(s)\det D_F(s). \qquad (2.29)$$

Then one can construct a Nyquist type stability test (cf.[RO1,pp.140]) by constructing the contour Γ_1 as $\varphi_1(s): D \to \Gamma_1$. Then a necessary and sufficient condition for exponential stability is that $N(\Gamma_1;0) = 0$. However this approach requires a more detailed internal model of both the plant and compensator (removing some of the practical benefits of Nyquist test being based on measurable data). The alternative suggested by Rosenbrock is to obtain the integer p_o which is the number zeros of $\det D_G(s)\det D_F(s)$ (which are the open loop poles) contained in the closed right half plane. Then the Nyquist contour based on Γ_2 with $\varphi_2(s)=\det[I_m + F(s)G(s)]$, $\varphi_2(s): D \to \Gamma_2$ can be used to provide necessary and sufficient condition for stability as

$$N(\Gamma_2; 0) = -p_o\,. \qquad (2.30)$$

Assuming p_o can be determined the Nyquist test can be performed based on measurable quantities. However the resulting Nyquist contour provides very little insight as to how to modify the compensator elements of $F(s)$ to achieve (or enhance) stability. Here Rosenbrock proposed focusing on a specific class of transfer functions which are diagonally dominant on the contour D. (We will consider the technical aspects of this in more detail in chapter 3.) For such transfer functions he showed that one could stabilize the system with decentralized compensation (i.e., $F(s)$ are diagonal matrices). The focus was to provide insight for the design of the required diagonal elements to enhance overall stability. The approach allowed the designer to focus attention on those loops which were most critical in achieving enhanced stability. For those systems which were naturally diagonally dominant the approach was quite successful. In order to extend the approach to the class of MIMO transfer functions Rosenbrock proposed, as a first step, designing series multiloop dynamic compensation to achieve diagonal dominance. This is however not always easy to do in practice.

A major concern in engineering design of feedback systems is sensitivity of closed loop performance to variation in parameters. In [PO3] Postlethwaite and MacFarlane develop a generalization of the Nyquist test for MIMO feedback which focuses on sensitivity of stability properties due to variation of a single parameter $k \in \mathbb{C}$. They choose to focus on a particular loop-breaking configuration which identifies an $m \times m$ matrix return difference $I_m + kG(s)$ where $G(s)$ is the loop transmission transfer function containing dynamics of the plant and compensator as well as any sensors and/or actuator dynamics. The physical significance of the "gain" parameter k which appears equally in all m loops is not discussed (cf. [PO3-4]). From a practical point of view this setup remains somewhat less than general. However, this allows concentration on the variation of a "critical point" depending on k with respect to a set of fixed Nyquist loci.

Theorem 2.11: Let G(s) $\in R^{m \times m}(s)$ be proper. Let $\lambda_i(s)$ for $i=1,\ldots,m$ be eigenvalues of the matrix $G(s)$. Then define m Nyquist contours Γ_i via $\lambda_i(s): D \to \Gamma_i$. Let p_o be the number of open loop poles of G(s) in the closed right half plane. Then the closed loop system is exponentially stable if and only if,

$$\sum_{i=1}^{m} N(-1/k;\Gamma_i) = -p_o \,. \tag{2.31}$$

Proof : (cf.[PO3]).

The proof employed by Postlethwaite and MacFarlane centers on the analysis of the function $\mu(g,s) = \det[gI_m - G(s)]$ (where $g = \frac{-1}{k}$) using algebraic function theory. Then $\mu(g,s) = 0$ defines a relation $s(g)$ called the characteristic gain loci. In this setup the Nyquist contours are the branches of $s(g)$ and are constructed (via analytic continuation) on a multi-sheeted Riemann surface.

Alternately Desoer and Wang [DE2] provide a proof for this same theorem based on techniques from analytic function theory. This permits extension of the result to cover the class of distributed parameter systems whose impulse response belongs to the convolution algebra described above (cf.[DE1]).

Despite limitations due to practical interpretation of the scalar parameter k, the isolation of a critical point depending on k allows the designer to portray graphically the set of stabilizing k. The graphical appeal is also considered important for computer-aided design since one could readily plot these eigen loci for rather complex systems.

However, Doyle and Stein [DO3] point out that this test may be misleading in practice since the eigen loci can be very sensitive to more general perturbations than just changes in scalar k. By way of rebuttal, Postlethwaite et al [PO4] show how singular value analysis can be coupled with this approach to provide a certain level of robustness.

A geometric viewpoint is taken in Brockett and Byrnes [BO1] in describing a generalized Nyquist criterion. Here for the first time the general case of G(s) $\in R^{p\times m}(s)$ with $p \neq m$ is treated, although somewhat abstractly. Significantly, their approach avoids formulation of the return difference matrix and as a result allows the separate characterization of an abstract critical point (resulting from F) and a Nyquist locus (resulting from $G(j\omega)$). This formulation preserves most nearly the practical aspects of Nyquist criterion exploited in SISO system design [BO1].

Briefly,the setup is as follows †. The feedback equations 2.14 are written in matrix form as

$$\begin{bmatrix} G(s) & -I_p \\ I_m & F \end{bmatrix} \begin{Bmatrix} u(s) \\ y(s) \end{Bmatrix} = 0. \qquad (2.14')$$

where it is clear that a complex scalar s is a closed loop pole if and only if the $\ker[G(s), -I_p]$ intersects $\ker[I_m, F]$ in some nontrivial way. In the case that $p < m$ we can construct an abstract Nyquist locus Γ_G for the $p \times m$ transfer function matrix $G(s)$ by thinking of Γ_G as the image of the imaginary axis under the map $\mathbf{G}(s) = \ker[G(s), -I_p]$. For each $s = j\omega$ the object $\mathbf{G}(s) = \ker[G(s), -I_p]$ is a p-dimensional subspace of a $p+m$ dimensional complex space. Thus Γ_G can be thought of as a "curve" in the complex Grassmanian ($\Gamma_G \subseteq$Grass$(p, m+p)$) of p-dimensional planes in $m+p$ space. The question of whether $\mathbf{G}(s) = \ker[G(s), -I_p]$ intersects $\mathbf{F} = \ker[I_m, F]$ can be ascertained by utilizing the dual structure of the Grassmanian in the following way. The Schubert hypersurface associated with $\mathbf{F}$, an m-dimensional subspace of $\mathbf{C}^{m+p}$, is defined as

$$\sigma(\mathbf{F}) \triangleq \{\mathbf{M} \in \text{Grass}(p, m+p) : \dim(\mathbf{M} \cap \mathbf{F}) > 0\}, \qquad (2.32)$$

† The geometric theory of linear systems and feedback is reviewed in some detail in section 4.1.

a hypersurface in Grass$(p, m+p)$ representing the point $\mathbf{F} \in$Grass$(m, m+p)$. Here Grass$(m, p+m)$ is referred to as the *dual* of the space Grass$(p, p+m)$ [BR1]. Then $\mathbf{F}$ intersects $\mathbf{G}(s)$ if $\mathbf{G}(s)$ intersects $\sigma(\mathbf{F})$ in Grass$(p, m+p)$.

Using this dual structure (cf. [BR1] and [BY1] for details) the following theorem is provided.

Theorem 2.12: (Generalized Nyquist Theorem)

Suppose $G(s)$ is a proper rational $p \times m$ transfer function matrix with no poles on $\mathrm{Re}\, s=0$. Suppose the abstract Nyquist locus Γ_G does not intersect the Schubert hypersurface $\sigma(FB)$ defined by the feedback matrix F. Let p_o be the number of open loop poles of $G(s)$ in the closed right half plane (CRHP) and p_c be the number of closed loop poles in CRHP. Then

$$N(\Gamma_G; \sigma(\mathbf{F})) = p_c - p_o \tag{2.33}$$

where $N(\bullet;\bullet)$ is the number of encirclements of the abstract Nyquist locus Γ_G about the Schubert hypersurface $\sigma(\mathbf{F})$ taken in a positive direction on the Grassman manifold.

Proof: (cf. [BR4]).

Clearly theorem 2.12 does not admit any readily obvious graphical representation that would permit the determination of the winding number N (except in trivial cases). However the theorem does permit us to ascertain the stability of a MIMO feedback system involving a plant $G_1(s)$ with feedback F_1 by testing for homotopical equivalence with some other feedback system $G_2(s)$ with F_2 (of appropriate dimensions) which is known to be stable. To show such equivalence we will need a measure of how close a point $\mathbf{G}(s) \in$Grass$(p, m+p)$ is to some Schubert hypersurface

$\sigma(F) \in \text{Grass}(p, m+p)$.

2.3.3. Stability Margins From Return Difference

The primary benefit offered by frequency domain analysis and especially the Nyquist criterion for engineering design of SISO feedback is its ability to display immediately not only whether (or not) a feedback configuration is stable but how stable (or unstable) the loop is. Moreover, the graphical display of the Nyquist contour allows the designer to assess how certain critical design parameters can be chosen to enhance the system stability margin.

The notion of stability margin can be quantified in several (non-equivalent) ways. Take $g(s) \in R(s)$ to be the Laplace transform of a causal system with impulse response $w(t)$, where

$$w(t) = \frac{1}{2\pi j} \oint_C g(s)\, e^{st}\, ds \tag{2.34}$$

(the inverse Laplace transform) converges for C a closed contour in the region of convergence for the transform (cf. section 2.1.2). With a simple constant feedback f, the closed loop response has transfer function

$$h(s) = \frac{g(s)}{1 + f g(s)}. \tag{2.35}$$

Then $z(t)$, the inverse Laplace transform of $h(s)$, is of *exponential order* if there exists $\alpha, M \in \mathbb{R}$ with $M>0$ such that $|z(t)| < Me^{\alpha t}$ for $t>0$. Clearly z(t) is exponentially stable if and only if one can find such $\alpha<0$. Thus the ***abscissa of convergence***, $\alpha_0 \in \mathbb{R}$, which is the infimum of all such α, is often referred to as a stability margin.

However this notion of stability margin may not be very useful for practical feedback design since it offers no direct characterization of possible oscillatory responses. Indeed, the utility of the Nyquist criterion is to establish directly a stability margin concept of this type. Furthermore, by including a physically motivated design parameter (e.g. loop gain) direct insight is obtained from a Nyquist plot as to the choice of feedback gains to enhance stability margin.

Consider again a causal transfer function, $g(s)$, and a nominal scalar gain, f_0. For the purposes of design the feedback gain is parametrized as $f = kf_0$. Assume the nominal feedback system is exponentially stable. Then we ask whether there exists a finite loop gain variation $k \in \mathrm{R}$ for which the system supports a nondecreasing oscillation? This question is answered classically by computing the system *gain margin*, i.e.

$$g_m \triangleq \inf_{k>1} \{k \in \mathrm{R} \mid fg(s) = -1, \text{for } \mathrm{Re}\, s \geq 0\}. \qquad (2.36)$$

The usual way to compute g_m is to find a frequency ω_1 such that $\measuredangle fg(j\omega_1) = -\pi$†. Then

$$g_m = \frac{1}{|fg(j\omega_1)|}.$$

Similarly the *phase margin*, is a real number ϑ, $2\pi \geq \vartheta \geq 0$ such that

$$\vartheta_m \triangleq \inf_{2\pi \geq \vartheta \geq 0} \{\vartheta \mid fg(s) = -e^{j\vartheta}, \text{ for } \mathrm{Re}\, s \geq 0\}, \qquad (2.37)$$

which is computed by finding a frequency ω_2 such that $|fg(j\omega_2)| = 1$, then

$$\vartheta_m = -\measuredangle fg(j\omega_2). \qquad (2.38)$$

† The notation $\measuredangle$ is used to indicate the argument (phase) of a complex number.

In some cases it has been observed that a reduction in gain (i.e., $k<1$) may also lead to instability. This corresponds to case where there is more than one frequency at which "gain crossover" occurs (i.e., $|1+fg(j\omega)|=1$). This case is classically referred to as *conditional stability.*

Despite the popularity of these classical notions we will take the viewpoint in this dissertation (following [HE1]) that a slightly more conservative measure of stability margin is appropriate.

Definition : The combined *gain-phase* margin, $g_{\vartheta m} \in \mathrm{R}_+$, is given by

$$g_{\vartheta m} \triangleq \inf_{s=j\omega} |1+fg(s)| . \tag{2.39}$$

Remark 1 : Clearly $g_{\vartheta m}$ is just the euclidean distance between the Nyquist locus resulting from some loop breaking configuration and the "critical point" at -1.

Remark 2 : For $g(s) \in \mathrm{R}(s)$ the Nyquist contour Γ is a closed contour and therefore a compact set in $\mathbf{C}$. We can therefore replace infimum with minimum. Now conditional stability means that $|1+fg(j\omega)|$ has several local minima so that the set of local minimizers of the function

$$|1+fg(j\omega)| \tag{2.40}$$

may *not* be a compact set in R.

Remark 3 : As a measure of stability margin $g_{\vartheta m}$ is more conservative than using g_m and ϑ_m since it accounts for possible simultaneous gain and phase variations. This can be seen in figure 2.3 which gives a Nichols plot (i.e. gain $|fg(j\omega)|$ in dB vs. phase $\measuredangle fg(j\omega)$). The combined g_m gives a rectangular region which includes the contour, Γ. However, $g_{\vartheta m}$ defines a circle tangent to the contour, Γ.

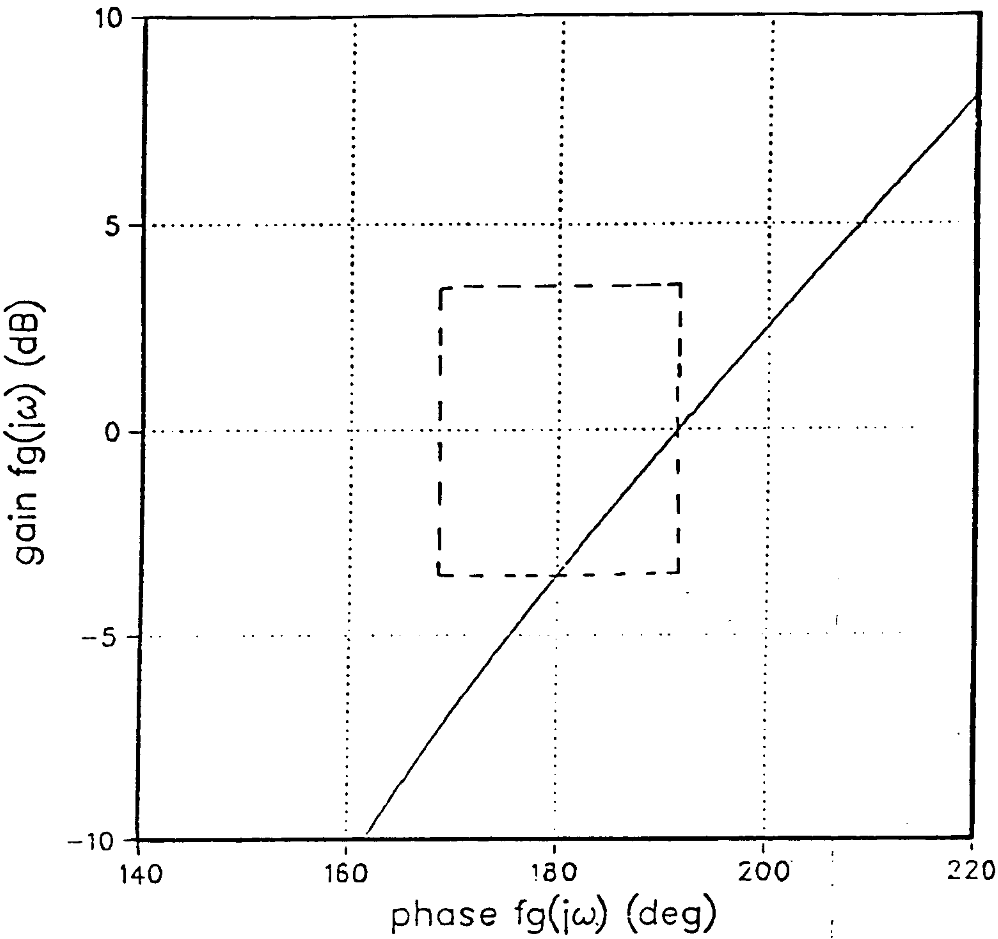


Figure 2.3: Phase and Gain Margin Specifications on a Nichols Plot

Remark 4 : The practical extension of these ideas to MIMO systems is probably most meaningful by generalization of $g_{\vartheta m}$. This is at least in part due to lack of general significance of any other notion of multivariable phase which can take into account the individual phase of each scalar transfer function $g_{ij}(s)$ appearing in $G(s)$.

Some recent work on extending these notions of stability margins to MIMO feedback [DO3,LE1] has focused on the characterization of specific classes of allowable (non-destabilizing) perturbations in terms of a measure of stability employing the minimum singular value of a matrix return difference (say $I_m + FG(s)$) depending on loop breaking location. That singular value analysis is an appropriate tool in studying perturbations has been well known by numerical analysts (cf. [BJ1]). However for this analysis it leads to the definition of at least four different measures of stability margin; e.g.,

$$
\begin{aligned}
g_i &\triangleq \inf_{s=j\omega} \sigma_{\min}[I_m + FG(s)], \dagger \\
g_o &\triangleq \inf_{s=j\omega} \sigma_{\min}[I_p + G(s)F], \\
\hat{g}_i &\triangleq \inf_{s=j\omega} \sigma_{\min}[I_m + \{FG(s)\}^{-1}], \\
\hat{g}_o &\triangleq \inf_{s=j\omega} \sigma_{\min}[I_p + \{G(s)F\}^{-1}],
\end{aligned}
\tag{2.41}
$$

each of which may lead to different characterization of allowable perturbations since in general these measures all achieve different values. These measures are however not completely independent as discussed by Safanov et al [SA1].

As an example of the type of perturbation result which one can obtain using singular analysis consider perturbation $\Delta(s) \in R^{m \times m}(s)$ appearing as

$$FG(s) = (I_m + \Delta(s))\, FG_m(s) \tag{2.42}$$

where $FG(s)$ is the true $m \times m$ loop gain transfer function, $FG_m(s)$ is the transfer function of the nominal model, and $\Delta(s)$ is some unknown perturbation. Assume all transfer functions are strictly proper.

Theorem 2.13: [DO3] If $G_m(s)[I_m + FG_m(s)]^{-1}$ and $\Delta(s)$ are both exponentially stable then $G(s)[I_m + FG(s)]^{-1}$ is also exponentially stable for all $\Delta(s)$ satisfying

† We use the notation $\sigma_{\min}(A)$ to denote the minimum singular value of the matrix A.

$$\sigma_{\max}(\Delta(s)) = \|\Delta(s)\|_2 \leq l(\omega) \tag{2.43}$$

if

$$\sigma_{\min}[I_m + (FG_m(j\omega))^{-1}] > l(\omega) \tag{2.44}$$

for all $\omega \in \mathbf{R}$.

Proof: (cf. [DO3]).

Results such as this give somewhat conservative characterization of the robustness properties of MIMO feedback in that there can exist many perturbations $\Delta(s)$ which violate the bound given but do not cause instability. The problem is that the nature of the perturbation is not physically motivated in contrast to the SISO case where $\Delta(s)$ can be related to simple gain changes in the loop. However, these results suggest guidelines for good MIMO feedback design in terms of the stability margins given.

3. Results in Design of Decentralized Control Systems Based on Weak Dynamic Coupling

For the purpose of control system design and analysis the term "large scale system" is most often applied to a dynamical system which can be modeled as an interconnection of low-order dynamic systems. It is often desirable to build feedback control for such systems based on constrained measurements. If the set of system measurements (outputs) and the set of system commands (inputs) are each partitioned into disjoint sets with a "local" feedback control corresponding between each subset of inputs and outputs, the control strategy is called *decentralized.*

Associated with the partitioning of the inputs and outputs one can partition conformally a model for the dynamic system. This identifies certain "local" dynamic systems and possibly dynamic interactions. This is one way in which a system model can be viewed as a large-scale system. For the class of linear, time-invariant models considered in this dissertation, a natural way to describe such model partitions is in terms of partitions of a transfer function matrix into submatrices. Those blocks appearing on the diagonal will be called *local subsystems;* otherwise called *interaction dynamics*.

An essential issue in developing design methods for decentralized control was first considered by Wang and Davison [WA1]. They focused on the question of when a linear system which is both completely observable and controllable can be stabilized by decentralized control. Their analysis highlights conditions for the existence of "fixed modes" with respect to decentralized feedback. An algebraic characterization of the existence of fixed modes was first given by Anderson and Clements [AN2] and later by Vidyasagar and

Viswanadham [VI1]. In [VI1] the authors exploit the algebraic characterization of fixed modes to develop a method for pole relocation using decentralized feedback. Their method does not explicitly employ the partitioned model structure discussed above.

Several closely related analytical devices for determining stability of such partitioned models (often called interconnected models) with decentralized feedback have appeared [AR1, CO1, LA1, MI1], which are based on a sufficient condition which is loosely interpreted as *dynamic weak coupling*. The idea of dynamic weak coupling suggests that in some sense the local subsystem dynamics dominate those of the interactions. An alternate technical approach to describe dynamic weak coupling provides a frequency dependent measure of weak coupling appropriate for design of decentralized control [BE1-3, HU2, LI1, NW4]. These methods all have in common that they are motivated (at least technically) as generalizations of a well known design method for MIMO systems developed by Rosenbrock called Inverse Nyquist Array (INA) [RO1-2].

Since this dissertation provides yet another generalization of INA methods which is focused on a particular practical aspect of the design problem for decentralized control we dedicate this chapter to a review of this line of research. We start by discussing the INA method which is based on the notion of a *diagonally dominant* transfer function matrix. This method was not initially suggested as a design tool for decentralized control. Next we discuss a procedure originally proposed by this author for decentralized control which employs the notion of a block diagonally dominant transfer function matrix [BE1-2]. Efforts by Limebeer and Nwokah to generalize the notion of diagonal dominance have also been extended to the case of partitioned matrices [LI1, NW4]. These results will be reviewed. Finally, in this chapter we describe the difficulty encountered in attempting to generalize a particularly important practical aspect of INA design to the case of a general (non-trivial) transfer matrix partition.

It is important to recognize that, for systems where dynamic weak coupling is evident, the susceptibility (sometimes called "integrity") of a decentralized feedback to failures of local controllers can be readily evaluated. This remark is in contrast with design methods such as [VI1, LE2] which can potentially deal with a much larger class of systems by avoiding the need for a weak coupling assumption. As a result such general methods require that feedback robustness properties be evaluated on a global basis [LE1, chapt.V].

The weak coupling methods discussed in this chapter are based in an essential way on frequency domain models and the Nyquist stability test. Such tests have practical utility since they are based on frequency response data which can be obtained from direct measurements on the system. In practice the availability of such data is limited to finite frequency bands. We will find it convenient to restrict consideration to systems where the loop transmission transfer functions are all *strictly proper*. This assumption is used for instance by Rosenbrock [RO1,pp.153] to provide sufficient conditions for a Nyquist stability test where diagonal dominance is required only on the finite portion of the contour D along the $j\omega$ axis (with possible small indentations around finite singularities.) In [BE1,pp.42] the present author obtained similar results for the case of general partitioned transfer functions which are *block diagonally dominant* along the corresponding finite frequency range. Further discussion of the ramifications of this assumption will be postponed until chapter 5.

3.1. The Inverse Nyquist Array Method

The INA method [RO1] seeks to determine sufficient conditions for the choice of feedback compensators, based on a frequency dependent measure of loop weak coupling, in an independent loop-by-loop approach. The specific notion of weak coupling used is based on the idea of diagonal dominance of the system transfer function matrix for values of the Laplace variable s contained on a closed elementary contour in the complex plane. The term, diagonal dominance, is borrowed from a result due to Gershgorin providing sufficient conditions for the invertibility (or regularity) of a square matrix.

Definition : An $m \times m$ matrix, $Z(s)$, rational in s is said to be *diagonally dominant* on a closed contour D in the complex plane if for all s on D one of the following conditions is satisfied:

$$|z_{ii}(s)| > \sum_{j \neq i}^{m} |z_{ij}(s)| \qquad i=1, \cdots, m, \tag{3.1}$$

(row dominance) or

$$|z_{ii}(s)| > \sum_{j \neq i}^{m} |z_{ji}(s)| \qquad i=1, \cdots, m,$$

(column dominance).

This criterion can be easily evaluated graphically. Plot for each s on D a circle centered at $z_{ii}(s)$ with radius

$$\min\{ \sum_{j \neq i}^{m} |z_{ij}(s)|, \ \sum_{j \neq i}^{m} |z_{ji}(s)| \}$$

then (3.1) is satisfied if and only if the envelope swept out by these so called Gershgorin circles (often called the Gershgorin band), avoids the origin for each of m plots, $i=1,\ldots,m$.

The significance of these Gershgorin bands when applied to the inverse of the system transfer function matrix over the usual Nyquist contour is taken from another classic result in numerical linear algebra due to Ostrowski [RO1]. This result provides a bound on the distance between the inverse of the diagonal elements of a complex-valued matrix which is diagonally dominant and the diagonal elements of the matrix inverse (which are, of course, usually different). This involves some notation. Let $\hat{Z}(s) = Z^{-1}(s)$. The inverse of the diagonal elements of Z (or $\hat{Z}$) are z_{ij}^{-1} (resp. $\hat{z}_{ij}^{-1}$) where in general $z_{ii}^{-1} \neq \hat{z}_{ii}^{-1}$. Then the Ostrowski result states that if Z is diagonally dominant then

$$|\hat{z}_{ii}^{-1} - z_{ii}| < \varphi_i d_i < d_i \qquad (3.2)$$

for $i=1,\ldots,m$ where $d_i = \sum_{j \neq i}^{m} |z_{ij}|$ is the radius of ith Gerschgorin circle and

$$\varphi_i = \max_{j \neq i} \frac{d_j}{|z_{jj}|} \qquad (3.3)$$

is a "shrinking factor" associated with the dominance of the other, $j \neq i$, rows of Z. Notice that the φ_i for $i=1,\ldots,m$ are each $\varphi_i < 1$ if Z is diagonally dominant.

For the design of multivariable feedback (cf. fig. 2.1). We have the relations

$$y(s) = G(s)[I + F(s)G(s)]^{-1} r(s) \qquad (3.4)$$

and

$$r(s) = [F(s) + \hat{G}(s)] y(s) \qquad (3.5)$$

which can be used to study equivalently the closed-loop system response.

Remark : The inverse system representation of equation (3.5) is the relation favored by Rosenbrock in INA method. It is often suggested that (3.5) is preferable because of its "simpler form". However use of (3.5) is predicated on plants with square transfer function matrix; i.e., $p=m$ and requires the explicit calculation of the inverse matrix $\hat{G}(s)=G^{-1}(s)$ which can often be numerically ill-conditioned in practice. The real utility of the inverse representation results from the application of a result due to Ostrowski.

Let $Z(s)=F+\hat{G}(s)$ in (3.2) with F, a diagonal matrix, F = diag$\{f_1,f_2,\ldots,f_m\}$. We see that

$$|h_{ii}^{-1}(s)-(f_i+\hat{g}_{ii}(s))| < \varphi_i(s)\hat{d}_i(s) < \hat{d}_i(s), \tag{3.6}$$

for $i=1,\ldots,m$ and where

$$\hat{d}_i(s) = \sum_{j\neq i}^{m} |\hat{g}_{ij}(s)|, \tag{3.7}$$

$$\varphi_i(s) = \max_{j\neq i} \frac{\hat{g}_j(s)}{|f_j+\hat{g}_{jj}(s)|}, \tag{3.8}$$

and $h_{ii}(s)$ is the ith diagonal element of $H(s) = G(s)[I_m+FG(s)]^{-1}$.

Now consider what happens if the ith loop is opened; i.e., $f_i=0$ while other $f_j, j\neq i$ are fixed. Then $h_{ii}^{-1}(s)$, the inverse of the diagonal element of the transfer function matrix of the system with all other loops closed, is contained in the Gershgorin band about $\hat{g}_{ii}(s)$. The shrinking factor, $\varphi_i(s)$, remains unchanged so that the effect of the other local feedbacks must be to strictly reduce the width of the Gershgorin band about $\hat{g}_{ii}(s)$ (so long as $F+\hat{G}$ is diagonally dominant). Since $h_{ii}^{-1}(s)$ is the actual inverse transfer function

relating the ith input-output pair (with f_i=0) classical single-input/single-output techniques can be applied to choose the scalar compensator gain, f_i. The shrinking factor $\varphi_i(s)$ leads to estimates for the local stability margins associated with the ith loop resulting from the effect of the other loop compensators. Thus the Gershgorin bands define a broad or fuzzy Nyquist locus for the ith subsystem.

The major problem with the INA method for design of general multivariable control systems is the case when $\hat{G}(s)$ is not diagonally dominant. In this case, the method as developed by Rosenbrock, requires the construction of pre- (or post-) series compensators such that $\hat{Q}(s)=\hat{L}(s)\hat{G}(s)\hat{K}(s)$ can be made diagonally dominant. Available techniques for the synthesis of $K(s)$ and $L(s)$ which represent relatively low order, realizable, and stable multivariable plants are ad hoc at best. The search for techniques for synthesizing these compensators has been the subject of much research [RO1], [LE2]. Various algebraic results are now available for certain classes of problems. Also several gradient search procedures have been developed and tested. However the collection of results still presents a rather ad hoc approach.

We take the viewpoint in this research that synthesis of $K(s)$ and $L(s)$ may not be of fundamental interest for two reasons. First, there are other measures of diagonal dominance which may provide the necessary estimates when the plant is not diagonally dominant in the usual sense [BE1-3,LI1-2,NW1-4]. Second, for large scale systems, we are interested in identifying the structural aspects of the plant which admit a decentralized control solution embodying the partitioning of the information pattern imposed on the controller. Thus in this dissertation we will be interested in associating attractive partitions for the control problem with appropriate measures of subsystem interaction.

3.2. Block Diagonal Dominance and Decentralized Feedback Control

In references [BR1-3,FE1,FI1-2,JO1-2,VA1-2] results are provided which generalize the Gershgorin theories to the case of partitioned matrices in several ways. In [BE1-2] these ideas are applied to the problem of decentralized feedback control by appropriately generalizing the Nyquist array ideas of Rosenbrock. The result provides a framework for rationalizing the choice of local feedback compensators for the subsystems of the partitioned plant model in terms of a measure of subsystem interaction. Complete freedom is available as to the choice of the design method employed of each for the subsystem compensators so long as certain bounds on the compensator response can be guaranteed. These bounds can be stated in terms of several different (not necessarily equivalent) measures of subsystem interaction. Unlike in Rosenbrock's INA method, the individual subsystem designs require a multivariable design method. In this section we will summarize the salient aspects of these preliminary results. We will discuss the major limitations and open questions of the method.

3.2.1. Preliminary Definitions and Notation

Following [FE1] let A be an $n \times n$ complex matrix partitioned into $m \times m$ submatrices, A_{ij}, $i,j=1,\ldots,m$, where A_{ij} is $k_i \times k_j$; and $\sum_{j=1}^{m} k_i = n$. Let $|\bullet|$ denote a vector norm on the subspace X_i, for $i=1,\ldots,m$ of C^n implied by the partition above. Then the usual induced or subordinate matrix norm is

$$\|A_{ij}\| = \sup_{\substack{x \in \mathbf{X}_i \\ x \neq 0}} \frac{|A_{ij}\, x_i|}{|x_i|}$$

and its infimum or reciprocal is

$$[\![A_{ij}]\!] = \inf_{\substack{x \in \mathbf{X}_i \\ x \neq 0}} \frac{|A_{ij}\, x_i|}{|x_i|}$$

Clearly if A_{ij} is nonsingular then

$$[\![A_{ij}]\!] = \|A_{ij}^{-1}\|^{-1}.$$

Definition : Let $A(s)$ be an $n \times n$ matrix of rational functions in s which is partitioned into $m \times m$ submatrices, $A_{ij}(s)$. Let D be a closed elementary contour in the complex plane. Then $A(s)$ is said to be ***block diagonally dominant*** (BDD) on D if:

(i) $A_{ii}(s)$ has no pole on D for $i=1,\ldots,m$ and

(ii) for all s on D either

$$[\![A_{ii}]\!] > \sum_{j \neq i}^{m} \|A_{ij}\| \qquad i=1,\cdots,m, \tag{3.9}$$

(block row dominance) or

$$[\![A_{ii}]\!] > \sum_{j \neq i}^{m} \|A_{ji}\| \qquad i=1,\cdots,m,$$

(block column dominance).

Then we can state the main result of [BE2].

Theorem 3.1: Let $A(s)$ be block diagonally dominant on D, a closed elementary contour in the complex plane. Let s trace once around D clockwise. Let $\det A(s): D \to \Gamma_A$. Let $\det A_{ii}(s): D \to \Gamma_i$ for each $i=1,\ldots,m$. Then

$$N(\Gamma_A;0) = \sum_{i=1}^{m} N(\Gamma_i;0). \tag{3.10}$$

Proof: (cf. [BE2]).

3.2.2. Stability Tests for Decentralized Control and the Return Difference

To test for closed loop stability we choose a contour D as discussed in chapter 2. We take p_o to be the number of open-loop poles of $F(s)G(s)$ in the closed right half plane (CRH); i.e., $p_o = p_{o,G} + p_{o,F}$ where $p_{o,G}$ (resp. $p_{o,F}$) are the number of system poles of G (resp. F) in CRH.

Remark: Several types of cancellations can occur in forming $I+GF$ as discussed in chapter 2. In the sequel, when we discuss decentralized control, we restrict F to have a particular sparse structure (i.e. block diagonal). In this case from [WA1] other types of fixed modes can exist.

Now the fundamental utility of block diagonal dominance for the case of decentralized feedback control can be seen from the following result [BE2].

Theorem 3.2: Let D be the Nyquist contour as defined above. Take F=block diag$\{F_i, i=1, \cdots, m\}$ and partition $G(s)$ conformally with

F.

Let $\det(F_i^{-1} + G_{ii}(s)) : D \to \Gamma_i$. Assume that the compensator $F(s)$ is open-loop asymptotically stable and non-minimum phase. If $F^{-1}(s) + G(s)$ is block diagonally dominant for s on D then the closed-loop system is asymptotically stable if and only if:

$$-\sum_{i=1}^{m} N(\Gamma_i; 0) = p_{o,G} = p_o \,. \tag{3.11}$$

Proof: (cf. [BE2,thm.3-3,pp.38]).

Remark: Several other results are available in [BE2] when other forms of the return difference for the inverse system transfer function are block diagonally dominant.

In [BE1] the present author derives several conditions on the choice of the local compensators, $F_i(s)$ $i=1,\ldots,m$, which if satisfied guarantee that stability of each of the local subsystems with interaction dynamics ignored; i.e.,

$$H_{ii}(s) = [I + G_{ii}(s)\,F_i(s)]^{-1} G_{ii}(s) \tag{3.12}$$

for $i=1,\ldots,m$, implies stability of the complete system. One such condition is as follows (slightly restated from [BE1,pp.40]).

Corollary 3.3: If the local feedback compensators $F_i(s)$ $i=1,\ldots,m$ are each chosen so that:

$$G_{ii}(s)[I + F_i(s)G_{ii}(s)]^{-1} \tag{3.13}$$

are asymptotically stable and for all s on D

$$[\![F_i(s)]\!]^{-1} < \max \left\{ [\![G_{ii}(s)]\!] - \sum_{j \neq i}^{m} \|G_{ij}(s)\|, \right.$$

$$\left. \| G_{ii}(s) \| - \sum_{j \neq i}^{m} \| G_{ji}(s) \| \right\}, \tag{3.14}$$

for $i=1,\ldots,m$. Then the composite closed-loop system, $G(s)[I+F(s)G(s)]^{-1}$, is asymptotically stable whenever all interconnections are stable; i.e., all contributions to p_o come from open-loop poles in CRH found only in $G_{ii}(s)$ for $i=1,\ldots,m$.

3.3. Generalizations of Block Diagonal Dominance for Rational Transfer Functions

Gerschgorin's theorem is but one (albeit well known) example of a class of localization theorems available in classical linear algebra [FA1, chapt.3]. Such theorems, which provide various characterizations for inclusion (or exclusion) regions for the spectrum of a matrix, are often used in numerical analysis to provide perturbation results for the spectrum of a matrix in terms of parameters. A rather extensive literature exists on extensions of these results to the case of partitioned matrices [BR1-3, FE1, FI1-2, JO1-2]. (A survey of some of the more well known results for partitioned matrices was included in [BE2]). In many applications these results are used merely to provide regularity conditions for matrices (e.g. the sufficiency argument used to provide the condition on encirclements of the Nyquist locus as $N(\Gamma;0) = \sum_{i=1}^{n} N(\Gamma_i;0)$ is based on regularity results). In such cases, complex characterization of inclusion region for the spectrum is entirely superficial since we are more concerned with determining a maximal exclusion region which contains the origin.

In 1980 Nwokah [NW4] developed a generalization of block diagonal dominance for transfer function matrices based on the notion of a composite H-matrix. Subsequently in 1982, Limebeer [LI1] employed a similar generalization but focusing on graphical design methods based on the eigenloci characterization of the generalized Nyquist criterion. The differing goals lead to different viewpoints but resulted in essentially the same generalization. To be specific Nwokah was interested in employing regularity results while Limebeer sought to characterize inclusion regions.

This dissertation focuses on characterizing an entirely new notion of block dominance from an abstract geometric perspective. We focus on determining inclusion regions for an abstract characterization of the Nyquist locus itself. Thus in this section we provide a critical review of the salient aspects of the work on BDD generalizations of Nwokah and Limebeer. We will show their commonality and highlight the technical differences. Finally, we will motivate the fundamental technical problem addressed in chapter 5 of this dissertation.

3.3.1. Generalized BDD by Using a Composite H-matrix Test

The generalized notion of BDD for rational transfer function matrices developed by Nwokah [NW4] is based on a generalization of Hadamard matrices (H-matrix) for partitioned matrices.

Definition: Given $C \in \mathbb{C}^{m \times m}$ construct two $m \times m$ real matrices B and $W = \text{diag}\{w_1,\ldots,w_m\}$ such that for $i,j=1,\ldots,m$,

$$b_{ij} = \begin{cases} 0\,, & \text{for } i=j \\ |c_{ij}|\,, & \text{for } i \neq j \end{cases} \tag{3.15}$$

and for $i=1,\ldots,m$ with $w_i = |c_{ii}|$. Then we say C is an H-matrix (Hadamard matrix) if $W-B$ is on M-matrix [FI1]. (Some results on M-matrices are summarized in appendix B.). The $m \times m$ matrix $W-B$ is an M-matrix if and only if all its principal minors are positive.

This is generalized as follows.

Definition (composite H-matrix): Let $Z \in \mathbb{C}^{n \times n}$ be partitioned into $m \times m$ submatrices $Z=(Z_{ij})$ where $Z_{ij} \in \mathbb{C}^{n_i \times n_j}$ with $\sum_{j=1}^{m} n_j = n$. Construct two $m \times m$ real matrices B and W = diag $\{w_1,\ldots,w_m\}$ such that for $i,j=1,\ldots,m$,

$$b_{ij} = \begin{cases} 0, & \text{for } i=j \\ \|Z_{ij}\|, & \text{for } i \neq j \end{cases} \tag{3.16}$$

and for $i=1,\ldots,m$

$$w_i = [\!]Z_{ii}[\!]. \tag{3.17}$$

Then call Z a *composite H-matrix* with respect to this partitioning if $W-B$ is an M-matrix.

The significance of this definition is that regularity of Z follows.

Theorem 3.4: Let $Z=(Z_{ij})$ be $n \times n$ composite H-matrix. Then Z is non-singular and furthermore $|\det Z| \geq \det(W-B) > 0$.

Proof: (cf.[NW4] also see Appendix B).

In applying this concept to develop a stability test Nwokah makes a crucial observation that if $G(s) \in \mathbb{R}^{n \times n}(s)$ is a composite H-matrix for s on D and if $F(s) \in \mathbb{R}^{n \times n}(s)$ is block diagonal and non-singular for s on D then both of the following matrices are also composite H-matrices:

(i) $G(s)F(s)$

(ii) $I_n + G(s)F(s)$.

Thus Nwokah is lead to consider the special structure of $G(s)$ a square transfer function matrix representing the plant. In providing a stability test for decentralized feedback he employs the eigenloci Nyquist test [DE2] "locally"; i.e., he considers the eigenloci of the diagonal blocks $G_{ii}(s)F_i(s)$ for each $i=1,\ldots,m$.

Theorem 3.5: Let $G(s)\in\mathbb{R}^{n\times n}(s)$ be partitioned into $m\times m$ submatrices. Let the decentralized feedback be

$$F(s) = \text{block diag}\{F_1(s),\cdots,F_m(s)\}. \tag{3.18}$$

The family of sets $\Lambda_i(s) = \{\lambda_1^i(s),\ldots,\lambda_{n_i}^i(s)\}$ for $i=1,\ldots,m$ are the eigenvalues for each s and for each block i of $G_{ii}(s)F_i(s)$. Let $\gamma_j^i(s)$ for $j=1,\cdots,r_i\le n_i$ be an indexed family of circuits (loci) formed from the set $\Lambda_i(s)$. If $G(s)$ is a composite H-matrix for all s on D then the closed loop system is asymptotically stable if:

(i) $\gamma_j^i(s)$ does not intersect the point -1 for $j=1,\ldots,r_i$ and $i=1,\ldots,m$ and

(ii)

$$\sum_{i=1}^{m}\sum_{j=1}^{r_i} N(\gamma_j^i(s);-1) = -p_o. \tag{3.19}$$

where p_o is the number of open loop poles contained in closed right half plane.

Proof: (cf. [NW4]). The usual homotopy argument is applied with the regularity condition supplied by assumption that $G(s)F(s)$ is a composite H-matrix. (Note r_i is the number of closed circuits (loci) constructed from the eigenloci (cf.[DE2])).

The fact that composite H-matrix is a generalization of a block diagonally dominant matrix is recognized by Nwokah [NW4] but perhaps not fully exploited since he did seek to construct inclusion regions for the eigenloci. As discussed in section 3.1 (cf. eqn. (3.6)) the trivial partitioning employed by Rosenbrock leads to strong engineering interpretation for the Gershgorin circles.

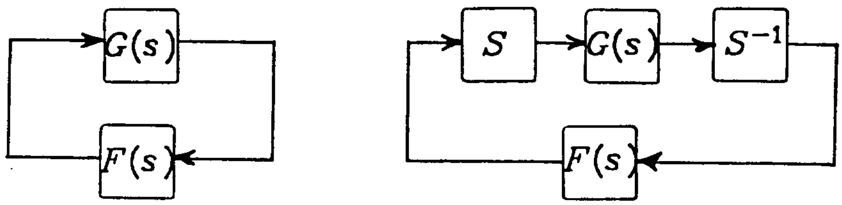
3.3.2. Generalized BDD and Inclusion Regions: A Graphical Test

In 1982, Limebeer reported the development of a generalized notion of a BDD matrix test based on an optimally "scaled" Gershgorin test. His method of proof (based on some recent work of Mees [ME1]) employed Perron-Frobenius theory of non-negative matrices. His result is not fundamentally different from the composite H-matrix test described above. The scaling approach, however, allows him to define generalized Gershgorin inclusion regions.

The scaling idea is as follows. Consider a transfer function $G(s) \in \mathbb{R}^{n \times n}(s)$ partitioned into $m \times m$ submatrices and a block diagonal (decentralized) feedback compensator F=block diag $\{F_1,...,F_m\}$ where $G_{ij} \in \mathbb{R}^{n_i \times n_j}(s)$ and $F_i \in \mathbb{R}^{n_i \times n_i}(s)$. Let S be an $n \times n$ diagonal scaling matrix

$$S = \text{block diag } \{s_1 I_{n_1}, \ . \ . \ . \ , s_m I_{n_m}\} \qquad (3.20)$$

with $s_i \neq 0$ for $i=1,...,m$. Then clearly stability of the loop pictured in fig. 3.1a is unchanged by scaling as shown in fig. 3.1b. However $G = S^{-1}GS$ can have drastically altered dominance properties. Indeed if one considers a block triangular matrix



(a) (b)

Figure 3.1: Scaling in Gershgorin Estimates

$$G = \begin{bmatrix} G_{11} & 0 \\ G_{21} & G_{22} \end{bmatrix}$$

which is not BDD it is clear that by appropriate choice of scaling s_1 and s_2,

$$G' = \begin{bmatrix} G_{11} & 0 \\ \frac{s_1}{s_2} G_{21} & G_{22} \end{bmatrix}$$

can be made BDD.

Definition: A matrix $Z \in \mathbb{C}^{n \times n}$ is *generalized block diagonally dominant* (GBDD) with respect to a given $m \times m$ partitioning if there exists an $x \in \mathbb{R}^m$, $x > 0$, such that either

$$\|Z_{ii}\| x_i > \sum_{j \neq i}^{m} \|Z_{ij}\| x_j \qquad \text{for } i = 1, \cdots, m, \tag{3.21}$$

or

$$\|Z_{ii}\| x_i > \sum_{j \neq i}^{m} \|Z_{ji}\| x_j \qquad \text{for } i = 1, \cdots, m.$$

The significance of this more general definition lies in the fact that inclusion regions for the eigenvalues of Z can be characterized as follows.

Theorem 3.6: If Z is partitioned into $m \times m$ submatrices then for any $x>0$, $x \in \mathbf{R}^m$ any eigenvalue of Z is contained in one of the sets

$$\left\{ \lambda \in \mathbf{C} : [\![\lambda I - Z_{ii}]\!] \le \sum_{i \ne j}^{m} \left(\frac{x_j}{x_i} \right) \| Z_{ij} \| \right\} \qquad (3.22)$$

for $i=1,\ldots,m$.

Proof: Obvious extension of basic result (e.g. [FE1]).

∎

By employing the Perron-Frobenius theory of non-negative matrices Limebeer demonstrates conditions under which a scaling x exists which can optimally reduce the inclusion regions provided above. This is accomplished by defining two non-negative test matrices $B, W \in \mathbf{R}^{m \times m}$ such that the elements of B satisfy

$$b_{ij} = \begin{cases} [\![Z_{ii}]\!], & \text{for } i=j \\ \| Z_{ij} \|, & \text{for } i \ne j \end{cases} \qquad (3.23)$$

and

$$W = \mathrm{diag}\{ [\![Z_{ii}]\!], i=1, \cdots, m \}. \qquad (3.24)$$

Very briefly, the Perron-Frobenius theory deals with nonnegative (element wise) matrices. Such matrices are classified according to the following notion of reducibility. We say that a nonnegative matrix A is ***reducible*** if it can be put into the form

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$$

by permutation of its rows and columns. (Thus a nonnegative matrix is *irreducible* if this cannot be done.) Associated with any irreducible nonnegative matrix is (according to the theory) a unique eigenvalue of A, λ_{PF}, of maximum modulus called the *Perron root* for which one can find an eigenvector which is strictly positive (element wise).

Then the result is stated as a theorem [LI1].

Theorem 3.7: Let $Z \in \mathbb{C}^{n \times n}$ be partitioned into $m \times m$ submatrices with the test matrices defined above. If B is irreducible then the following are equivalent

(a) $W^{-1}B$ has Perron-Frobenius eigenvalue

$$\lambda_{PF}(W^{-1}B) < 2, \tag{3.25}$$

(b) Z is generalized row block diagonally dominant

(c) there exists a scaling matrix $S>0$ (cf. Fig. 3.1)

$$S = \text{diag}\{x_1 I_{k_1}, x_2 I_{k_2}, \ldots, x_m I_{k_m}\} \tag{3.26}$$

such that $S^{-1}ZS$ is BDD by rows

(d) Z is generalized column block diagonally dominant

(e) there exists a scaling matrix $\bar{S}>0$ (cf. Fig. 3.1)

$$\bar{S} = \text{diag}\{\bar{x}_1 I_{k_1}, \bar{x}_2 I_{k_2}, \ldots, \bar{x}_m I_{k_m}\} \tag{3.27}$$

And furthermore if any of the above holds then Z is nonsingular.

Proof: (cf. [LI1]) It is sufficient to recognize that the Perron-Frobenius theory applied to the non-negative irreducible matrix $W^{-1}B$ gives

$$\begin{bmatrix} 1 & \bullet & \frac{\|Z_{1m}\|}{\|Z_{11}\|} \\ \bullet & 1 & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \frac{\|Z_{m1}\|}{\|Z_{mm}\|} & \bullet & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ \bullet \\ \bullet \\ \bullet \\ x_m \end{pmatrix} = \lambda_{PF} \begin{pmatrix} x_1 \\ \bullet \\ \bullet \\ \bullet \\ x_m \end{pmatrix}$$

where $x > 0$. From each row $i = 1, \ldots, m$ we get

$$\sum_{j \neq i}^{m} \|Z_{ij}\| x_j = \|Z_{ii}\| x_i \, (\lambda_{PF} - 1) . \quad \blacksquare \tag{3.28}$$

Then Limebeer provides a stability result for decentralized feedback focusing (as does Nwokah) on the case of $G(s) \in \mathbb{R}^{n \times n}(s)$ a square transfer function matrix and non-dynamic feedback of the form

$$F = \text{diag}\{f_1 I_{k_1}, \ldots, f_m I_{k_m}\} \, . \tag{3.29}$$

Theorem 3.8: With $G(s)$, F as above partitioned conformally into $m \times m$ partitions. Let $F^{-1} + G(s)$ be GBDD for all s on the Nyquist contour D. Let $\Lambda_i(s) = \{\lambda_1^i(s), \cdots, \lambda_{k_i}^i(s)\}$ $i = 1, \ldots, m$ be the eigenloci (generalized Nyquist contour) for the i^{th} diagonal block $G_{ii}(s) \in R^{k_i \times k_i}(s)$ for $i = 1, \ldots, m$ of image of D. Form circuits $\gamma_j^i(s)$ for $j = 1, \ldots, r_i \leq k_i$ which are continuous and analytic curves (loci) in $\mathbf{C}$. Then (stated somewhat differently from [LI1, Thm. 7]) the closed loop system is asymptotically (exponentially) stable if and only if:

(i) $\gamma_j^i(s)$ does not intersect the point $\frac{-1}{f_i}$ for each $i, j = 1, \ldots, m$

(ii)

$$\sum_{i=1}^{m} \sum_{j=1}^{r_i} N(\gamma_j^i(s); \frac{-1}{f_i}) = -p_o . \tag{3.30}$$

where as usual p_o is number of open loop poles in closed right half plane.

Proof: The proof differs from those above only in the technical basis for the regularity of $F^{-1} + G(s)$ on D - in this case being based on GBDD.)

It is acknowledged in [LI1] that GBDD is "not fundamentally different from" the H-matrix test of Nwokah [NW4]. Indeed, these definitions are equivalent as can be seen in [FI1, sect. 4] (cf. Appendix B). Limebeer's construction, however, focuses on the inclusion regions which he seeks to characterize graphically. The goal seems to be to provide, more completely, a generalization of the graphical design methods of INA to more general partitions by exploiting Nyquist eigenloci plots with inclusion regions super imposed. However, this approach (it seems to me) is doomed to failure for a number of technical reasons which we next discuss.

3.4. Problem of Estimating Local "Stability Margins"

The strength of the INA method for MIMO design comes from its graphical interpretation. As discussed in section 3.1 this follows from the bounds (3.6) which provide a new interpretation to the Gershgorin circles. Thus the envelope of the Gershgorin circles for each input-output pair centered on the corresponding Nyquist plot of the SISO response between that pair can be interpreted as an inclusion region for the true Nyquist locus for the pair with interaction dynamics included. Thus stability margins and gain/phase margins ***for each individual loop*** can be estimated directly from plots of the resulting brood or fuzzy Nyquist loci.

A natural extension of this idea to the more general case of decentralized feedback with arbitrary partitions would suggest the direct interpretation of inclusion regions for the eigenloci (via the generalized block Gershgorin results previously discussed). Thus the natural attempt would be to employ the generalized Nyquist eigenloci concept for MIMO stability analysis. However in practice this approach suffers from several deficiencies.

(1) Generally the inclusion regions

$$G_i = \left\{\lambda \in \mathbf{C} : [\![\lambda I - Z_{ii}]\!] < \sum_{j \neq i}^{m} \|Z_{ij}\|\right\} \qquad (3.31)$$

$$G_i' = \left\{\lambda \in \mathbf{C} : [\![\lambda I - Z_{ii}]\!] < \sum_{j \neq i}^{m} \|Z_{ji}\|\right\}$$

are *not* discs (except in the rather special case when Z is normal and the norm is "axis oriented"). The more general shape of G_i, G_i' may be difficult to determine.

(2) The sets G_i and G_i' (which of course coalesce in the case of GBDD) may be covered by discs for the purposes of graphical presentation as is proposed by Limebeer [LI1] by exploiting the eigenvectors of the diagonal blocks Z_{ii} $i=1,\ldots,m$. However, such bounds are not at all sharp and in fact can be totally useless when there are confluences in the eigenloci of the $G_{ii}(s) \in \mathbf{R}^{k_i \times k_i}(s)$ as s varies over D.

(3) Moreover, as discussed by Doyle and Stein [DO4] and by Postlethwaite and MacFarlane [PO4] it is not possible to provide a useful notion of MIMO gain/phase margin directly from the Nyquist-eigenloci plots without additional information. (Typically the minimum and maximum singular values are also required.) Thus the robustness question must be evaluated separately for MIMO systems.

There have been several attempts to apply the classical notions of robustness in terms of stability margins for multivariable systems [CR1,DO4,LE1,PO4,SA1,SA4]. Essentially, these methods express a measure of stability for a multivariable system in terms of a degree of regularity applied to the matrix return-difference. The degree of regularity frequently employed is subordinate to the Euclidean vector norm and can be shown to be the minimum singular value of $I + FG(j\omega)$ over the effective system bandwidth. The association of a bound

$$[\![F^{-1} + G(j\omega)]\!]_2 \geq l(\omega) > 0 \qquad (3.32)$$

with a measure of robustness with respect to plant uncertainty follows by showing that (3.32) guarantees that $\det[F^{-1} + G(j\omega)]$ and $\det[F^{-1} + G(j\omega) + \Delta G(j\omega)]$ are homotopically equivalent with respect to the origin for s on D whenever

$$l(\omega) \geq ||\Delta G(j\omega)||_2 \qquad (3.33)$$

for $\omega \in [-\omega_0, \omega_0]$ (recall from sect. 2.3.2 the construction of D).

Using the concept of block diagonally dominant (BDD) matrices this notion of robust design is applied (cf.[BE1]) to the case of decentralized feedback control as follows.

Definition: Consider the $n \times n$ plant transfer function matrix $G(s)$ partitioned into $m \times m$ submatrices. The quantities

$$d_i(s) = [\![G_{ii}]\!] - \sum_{j \neq i}^{m} \| G_{ij} \|, \tag{3.34}$$

$$d_i'(s) = [\![G_{ii}]\!] - \sum_{j \neq i}^{m} \| G_{ji} \|,$$

are called respectively the ***row margin of dominance***, and ***column margin of dominance***.

Theorem 3.9: (robust decentralized compensation):

The regularity requirement (3.32) is guaranteed for F = block diag$\{F_i, i=1, \ldots, m\}$ whenever $G(s)$ is BDD for $s = j\omega$ for $\omega \in [-\omega_0, \omega_0]$ with dominance margin (see (3.34)) sufficiently large so that compensators F_i $i=1,\ldots,m$ can be chosen which satisfy:

(i)

$$[\![F]\!]^{-1} < \min_{i \in [1,m]} \left[\max \{ d_i(j\omega), d'_i(j\omega) \} \right] - l(\omega) \tag{3.35}$$

for $\omega \in [-\omega_0, \omega_0]$ and

(ii)

$$G_{ii}(s)[I + F_i G_{ii}(s)]^{-1} \tag{3.36}$$

is asymptotically stable. Overall plant stability is guaranteed for all plant perturbations, $\Delta G(s)$, satisfying (3.33) whenever

$$p_o = \sum_{i=1}^{m} p_{o,i} \tag{3.37}$$

where p_o (resp. $p_{o,i}$) is the total number of open-loop unstable poles of $G(s)$ (resp. $G_{ii}(s)$).

Subsequently, Limebeer and Hung [HU1] developed robustness characterization for decentralized control exploiting GBDD and given in terms of bounds on the minimum singular value of $I + FG$. They study specific classes of perturbations significant for large-scale interconnected systems. Rather than dwell on the technical details of these arguments which employ the common tools of homotopical equivalence of some Nyquist loci (typically eigenloci) based on a regularity condition on $I + FG(s)$ for s on D we will provide a new result which we believe summarizes the thrust of this line of reasoning.

Consider that from the discussion in chapter 2 it is apparent that one realistic notion of stability margin for MIMO systems is based on the following measure (called by Helton gain-phase margin [HE1]).

Definition: Gain-phase margin of the MIMO feedback configuration of Fig. 2.1 is.

$$\delta_{pg} = \inf_{s \in D} \sigma_{\min}[I + FG(s)] . \tag{3.38}$$

From the standpoint of decentralized control and in the light of the utility of INA methods we wish to:

(i) provide estimates for δ_{pg} of the overall system in terms of some measure of relative dominance of subsystem to interaction dynamics,

(ii) identify dominant contributions to such estimates locally; i.e. identify those local controllers where interactions are strong.

A result which contributes to this goal in the spirit of the analysis used up to this point is based on the following block generalized Gershgorin result for singular values.

Theorem 3.10: Let $A \in \mathbf{C}^{m \times n}$ be partitioned into $l \times k$ submatrices so that $A_{ij} \in \mathbf{C}^{m_i \times n_i}$ with $\sum_{i=1}^{l} m_i = m$ and $\sum_{i=1}^{k} n_i = n$. Assume that $m_i = n_i$ for each $i = 1, \ldots, m$ (i.e., A_{ii} are square submatrices.) Let

$$d_i = \sum_{j \neq i}^{k} \|A_{ij}\|, \tag{3.39}$$

$$d_i{}' = \sum_{j \neq i}^{l} \|A_{ji}\|,$$

$$s_i = \max(d_i, d_i{}') \tag{3.40}$$

$$a_i = \|A_{ii}\| \tag{3.41}$$

for each $i = 1, \cdots, \min(l, k)$.

Without loss of generality assume $l < k$ and take

$$s = \max_{k+1 \leq i \leq l} \left\{ \sum_{j=1}^{k} \|A_{ij}\| \right\}. \tag{3.42}$$

Then each singular value of A belongs to one of the closed intervals

$$[\max(a_i - s_i, 0), a_i + s_i]$$

for $i = 1, \ldots, m$ and $[0, s]$.

Proof: Suppose σ is a singular value of A. Then there exists two vectors $x \in \mathbf{C}^n$ and $y \in \mathbf{C}^m$, both non-zero, such that

$$\sigma x = A^* y, \quad \sigma y = Ax. \tag{3.43}$$

Take $x^T = (x_1^T, \cdots, x_k^T)$, $y^T = (y_1^T, \cdots, y_k^T)$, partitioned conformally with A. The vectors x and y can be chosen such that for some subvector and some vector norm

$$\|y_i\| = \max\{\|x_1\|, \|x_2\|, \ldots, \|x_k\|, \|y_1\|, \ldots, \|y_l\|\},$$

and $\|y_i\|=1$. (For simplicity assume $\|.\|$ is the euclidean vector norm.)

Consider two cases as follows:

(i) Take i such that $k<i\leq l$ then write

$$\sigma y_i = \sum_{j=1}^{k} A_{ij}\, x_j \,. \tag{3.44}$$

Therefore,

$$|\sigma|\ \|y_i\| \leq \sum_{j=1}^{k} \|A_{ij}\|\ \|x_j\|$$

=>

$$|\sigma| \leq \sum_{j=1}^{k} \|A_{ij}\| \leq s\ ,$$

where the matrix norm applied to A_{ij} is induced by the (euclidean) vector norm. Thus $\sigma\in[0,s]$.

(ii) Consider $i<k$. Then we can write from (3.43) for some $i<k$

$$\sigma x_i - A^*_{ii}\, y_i = \sum_{j\neq i}^{l} A^*_{ji}\, y_j \tag{3.45}$$

and

$$\sigma y_i - A_{ii}\, x_i = \sum_{j\neq i}^{l} A^*_{ji}\, x_j \,. \tag{3.46}$$

From (3.45) and (3.46) we get

$$|\ \sigma\|x_i\| - \|A^*_{ii}\, y_i\|\ | \leq \sum_{j\neq i}^{l} \|A_{ji}\, y_j\| = d_i\,', \tag{3.47}$$

and

$$|\ \sigma\|y_i\| - \|A_{ii}\, x_i\|\ | \leq \sum_{j\neq i}^{l} \|A_{ij}\, x_j\| = d_i\,. \tag{3.48}$$

Clearly if $\sigma \geq a_i$ then by employing the definition of $[\!]\cdot[\!]$ and the assumption $\|x_i\|<1$,

$$|\sigma - a_i| \leq |\ \sigma\|x_i\| - [\!]A_{ii}[\!]\ | \leq |\ \sigma\|x_i\| - \|A^*_{ii} y_i\|\ | .$$

On the other hand if $\sigma < a_i$ then (since $\|y_i\|=1$ by construction),

$$|\sigma - a_i| \leq |\ \sigma - \|A_{ii} x_i\|\ | = |\ \sigma\|y_i\| - \|A_{ii} x_i\|\ | .$$

Thus we conclude using (3.47) and (3.48) that

$$|\ \sigma - [\!]A_{ii}[\!]\ | \leq \max(d_i', d_i) = s_i .$$

(The above result appropriately generalizes the result in [QI1].)

■

The application of this to the above problem then follows immediately.

Theorem 3.11: For the decentralized feedback system of figure 3.1 which is assumed exponentially stable let

$$d_i(s) = \sum_{j \neq i}^{m} \|F_i G_{ij}(s)\|$$

$$d_i'(s) = \sum_{j \neq i}^{m} \|F_j G_{ji}(s)\|$$

$$c_i(s) = \max(d_i(s), d_i'(s)) .$$

Then

$$\delta_{pg} \geq \inf_{s \in D} \min_{i \in [1,m]} \max\left([\!]I_{k_i} + F_i G_{ii}(s)[\!] - c_i(s),\ 0\right) . \qquad (3.49)$$

Proof: Immediate by substitution of $A = I_n + FG(s)$ which is therefore $n \times n$ complex valued.

In the remainder of this dissertation we will consider a new viewpoint towards answering this basic question of estimating locally the stability margins of a decentralized control system based on a geometric Nyquist criterion for MIMO systems described by Brockett and Byrnes [BO1]. The goal here is to develop a fresh and more generally applicable measure of weak coupling (like BDD or GBDD) which provides the required estimates.

4. Generalized Stability Margins from a Geometric Viewpoint

4.1. A Geometric View of MIMO Feedback

In this section we provide background on a particularly useful geometric construction of an abstract Nyquist contour for MIMO systems. We focus on some particularly salient properties of the abstract Nyquist contour for MIMO systems as discussed in [BR4]. We consider the relationship to more standard notions of the Nyquist contour and how this geometric notion can lead to a natural decomposition of the inputs and outputs (in special cases) relevant to the design of decentralized control.

We start with the general MIMO feedback equations with $u \in \mathrm{U}=\mathrm{C}^m$, $y \in \mathrm{Y}=\mathrm{C}^p$

$$G(s)u(s) = y(s)\,,\quad u(s) = -Fy(s) \tag{4.1}$$

with $G(s) \in \mathrm{R}^{p \times m}(s)$, $F \in \mathrm{R}^{m \times p}$ which we write suggestively as

$$\begin{bmatrix} I_m & F \\ G(s) & -I_p \end{bmatrix} \begin{bmatrix} u(s) \\ y(s) \end{bmatrix} = 0. \tag{4.1'}$$

The geometric theory of linear systems and feedback centers on the relative orientation of the two objects,

$$\mathbf{G}(s) = \ker\left[G(s)\,,\, -I_p\right] \tag{4.2}$$

and

$$\mathbf{F} = \ker\left[I_m\,,\, F\right] \tag{4.3}$$

which for any $s \in \mathbf{C}$ are a pair of subspaces in $\mathrm{U} \oplus \mathrm{Y} = \mathrm{C}^{p+m}$. Following Hermann and Martin [HE2] we can state the following.

Theorem 4.1: [HE2]. A complex number s_0 is a closed loop pole of the feedback equations (4.1) if and only if

$$\dim \left[G(s_0) \cap F \right] > 0. \qquad (4.4)$$

Proof: In the matrix form (4.1') this result is obvious since for rational functions the only singularities are poles.

∎

More significantly Hermann and Martin [HE2] suggested the following definition of an abstract Nyquist locus.

Definition: The (abstract) *Nyquist locus*, Γ_G, of a $p \times m$ transfer function $G(s)$ is an algebraic "curve" given by the map

$$s \mapsto \ker \left[G(s), -I_p \right]$$

as the image of the closed contour D. Γ_G is contained in the complex Grassman space consisting of all m-dimensional subspaces in $\mathbb{C}^{p+m}$. We consider a curve in a more general sense as an analytic map from a Riemann surface to a complex analytic manifold, viz., the complex Grassmanian. In this sense a curve has complex dimension one or real dimension two.

The complex Grassmanian is the set of all p-dimensional complex subspaces of $\mathbb{C}^n$, which we denote as Grass(p,n). Grass(p,n) admits the structure of an analytic manifold in this case of dimension $np - p^2$. A fundamental property of Grass(p,n), which was successfully exploited in [BR4] toward the construction of a generalized Nyquist test, is the duality between Grass(p,n) and Grass$(n-p,n)$. In particular, a canonical representation of a point $X \in$ Grass(p,n) is as a hypersurface $\sigma(X) \subset$ Grass$(n-p,n)$. This so called *Schubert hypersurface* is given by

$$\sigma(X) = \left\{ Y \in \text{Grass}(n-p,n) \colon \dim(X \cap Y) \geq 1 \right\}; \qquad (4.5)$$

i.e., all $Y \in \text{Grass}(n-p,n)$ which intersect $X \in \text{Grass}(p,n)$ nontrivially.

The most significant aspect of the abstract Nyquist locus, Γ_G, as constructed above is that it is fixed with respect to choice of feedback compensation, F, in contrast to methods which focus on eigenloci or determinants of a matrix return difference.

However, the construction is quite general and allows connections with more standard analyses. For instance, by change of basis in the space of inputs and outputs, U⊕Y, one can generate a new "rotated" Nyquist contour, Γ_2. In the particular case

$$\left[G(s), -I_p \right] \begin{bmatrix} I_m & F \\ G(s) & -I_p \end{bmatrix} = \left[0, \; G(s)F + I_p \right] \qquad (4.6)$$

reveals a "rotated" curve, Γ_2 as

$$\ker \left[0, \; G(s)F + I_p \right] \colon D \to \Gamma_2 .$$

However, the transformation

$$\begin{bmatrix} I_m & F \\ G(s) & -I_p \end{bmatrix}$$

represents a valid change of basis in $\mathbb{C}^{p+m}$ only for s not a closed loop pole. To say this another way Γ_2 is a valid curve contained in $\text{Grass}(m,p+m)$ whenever s is *not* a closed loop pole for s on D. (Of course the standard construction of D does not guarantee this.) Thus, from this geometric point of view one can see the advantage of working in the "natural" basis (given by (4.1')) in defining a "legitimate" Nyquist contour.

In the sequel we will need some background on certain natural algebraic structures in the study of linear systems and their effect on the abstract Nyquist locus. We will show, using these structures, the existence of certain classes of transfer functions with special properties (in terms of their Nyquist loci) of significance to decentralized control. In Chapter 5 we will exploit these ideas further.

We first consider transfer functions $G(s) \in \mathbb{R}^{p \times m}(s)$ (not necessarily proper) and the action of constant (nondynamic) change of basis in the composite space $\mathbf{U} \oplus \mathbf{Y} = \mathbb{C}^{p+m}$ in the form

$$\Sigma = \begin{bmatrix} S & K \\ H & -T \end{bmatrix}$$

with Σ: $\mathbb{C}^{p+m} \rightarrow \mathbb{C}^{p+m}$ nonsingular.

Lemma 4.2: The set of transfer functions $G(s) \in \mathbb{R}^{p \times m}(s)$ together with the Σ-action above provides the algebraic structure of a group. (We will call this the Σ-group). The action generates new transfer functions in $\mathbb{R}^{p \times m}(s)$ as

$$G(s) \underset{\Sigma}{\longmapsto} G_\Sigma(s)$$

where

$$\begin{aligned} G_\Sigma(s) &= (\, G(s)K + T)^{-1} \,(\, G(s)S - H\,) \qquad (4.7) \\ &= (\, TG(s) + H)\,(\, KG(s) + S)^{-1}. \end{aligned}$$

Proof: By construction Σ is in the general linear group of transformations on $\mathbb{C}^{p+m}$ so it remains to show only that $G_\Sigma(s) \in \mathbb{R}^{p \times m}(s)$. Now (4.7) comes from observing the Σ-action on the subspace $\mathbf{G}(s)$ as

$$\begin{bmatrix} T & H \\ K & S \end{bmatrix} \begin{bmatrix} G(s) \\ I_m \end{bmatrix} = \begin{bmatrix} TG(s) + H \\ KG(s) + S \end{bmatrix}$$

and

$$\left[G(s),\, -I_p\right] \begin{bmatrix} S & -K \\ H & T \end{bmatrix} = \left[G(s)S - H,\, -(G(s)K + T)\right]. \qquad \blacksquare$$

Clearly the Σ-action involves quadruples (S,T,K,H) which are identified as follows:

(i) S: $\mathbf{C}^m \to \mathbf{C}^m$, change of basis in $\mathbf{U}$ (space of inputs)

(ii) T: $\mathbf{C}^p \to \mathbf{C}^p$, change of basis in $\mathbf{Y}$ (space of outputs)

(iii) K: $\mathbf{C}^p \to \mathbf{C}^m$, with $u \mapsto u - Ky$ as the *feedback* operation

(iv) H: $\mathbf{C}^m \to \mathbf{C}^p$, with $y \mapsto y + Hu$ as the *feedforward* operation.

Next we consider a subgroup of the Σ-group on $\mathbf{R}_{sp}^{p\times m}(s) \subset \mathbf{R}^{p\times m}(s)$, the subset of ***strictly proper*** transfer functions in $\mathbf{R}^{p\times m}(s)$. This subgroup is commonly referred to as the ***feedback group*** [BY1] and involves transformations Σ_0 of the form

$$\Sigma_0 = \begin{bmatrix} S & -K \\ 0 & T \end{bmatrix}$$

on $\mathbf{U}\oplus\mathbf{Y}$.

Corollary 4.3: The set of strictly proper transfer functions $\mathbf{R}^{p\times m}(s)$ together with triples (S,K,T) form a group (called the feedback group) with respect to the action

$$G(s) \underset{\Sigma_0}{\longmapsto} G_{\Sigma_0}(s)$$

where

$$G_{\Sigma_0}(s) = (\,G(s)K + T\,)^{-1}\, G(s)S \tag{4.8}$$

$$= TG(s)\,(\,KG(s) + S)^{-1}.$$

This is clearly a subgroup of the Σ-group of rational transfer functions over $\mathbf{R}_{sp}^{p\times m}(s)$.

Proof: Let $H=0$ in Σ-group above.

Remark: Clearly the Σ group (resp. Σ_0-group) acts on the Nyquist locus Γ_G to yield a new $\Gamma_\Sigma \subset \text{Grass}(p, m+p)$ and equally on the Schubert hypersurface $\sigma(\mathbf{F})$ representing the fixed (with respect to s) point,

$$\mathbf{F} = \ker[\, I_m \,,\, F \,],$$

in $\text{Grass}(m, p+m)$. For instance

$$\begin{bmatrix} I_m \,, F \end{bmatrix} \begin{bmatrix} I_m & -F \\ 0 & I_p \end{bmatrix} = \begin{bmatrix} I_m \,, 0 \end{bmatrix}.$$

We next consider a special class of transfer functions whose Nyquist loci have special properties. This special class will play a central role in this dissertation so we will elaborate on their properties.

Definition: A transfer function $G(s) \in \mathbb{R}^{p \times m}(s)$ is said to be *degenerate* if and only if there exists some Schubert hypersurface $\sigma(\mathbf{X}) \subseteq \text{Grass}(m, p+m)$ associated with some $\mathbf{X} \in \text{Grass}(p, m+p)$ which contains the abstract Nyquist locus, Γ_G.

Clearly the class of degenerate transfer functions is invariant with respect to Σ-group (resp. Σ_0-group) actions on $\text{Grass}(m, p+m)$ (and therefore, on the transfer functions).

Several alternate conditions are available [BR4] to describe degeneracy. One alternate condition for degeneracy is that there exists a $m \times (p+m)$ matrix $[K_1, K_2]$ of rank $\min(m, p)$ such that

$$\det \begin{bmatrix} K_1 & K_2 \\ G(s) & -I_p \end{bmatrix} \equiv 0. \tag{4.9}$$

A third alternate condition is given by the following construction (stated somewhat differently from [BR4]). A subspace $\mathbf{X} \in$ Grass$(p, m+p)$ can be represented as the kernel of a $m \times (p+m)$ complex matrix X. Then $\Gamma_G \subset \sigma(\mathbf{X})$ in Grass$(m, p+m)$ if and only if for any s on D there exists a nontrivial linear combination of the columns of

$$\begin{bmatrix} G(s) \\ I_m \end{bmatrix} = \begin{bmatrix} g_1(s), \cdots, g_m(s) \end{bmatrix};$$

viz.,

$$\sum_{i=1}^{m} a_i \, g_i(s),$$

which lies in $\ker X$; i.e.,

$$X \left(\sum_{i=1}^{m} a_i \, g_i(s) \right) = 0. \tag{4.10}$$

Let

$$X = \begin{bmatrix} x_1^T \\ \bullet \\ \bullet \\ \bullet \\ x_m^T \end{bmatrix}.$$

Then (4.10) hold for a_i $i=1,\ldots,m$ not all zero if and only if

$$\det \Big[< x_j, \, g_i(s) > \Big]_{ij} \equiv 0. \tag{4.11}$$

Alternately, if we consider X partitioned as X = $[X_1, X_2]$ then the condition (4.11) can be written as

$$\det\left\{\begin{bmatrix}X_1, X_2\end{bmatrix}\begin{bmatrix}G(s)\\ I_m\end{bmatrix}\right\} \tag{4.12}$$

$$= \det\begin{bmatrix}X_1 G(s) + X_2\end{bmatrix} \equiv 0 .$$

To illustrate how degeneracy can occur, we consider two examples.

Example 4.1: Let $G(s) \in \mathbb{R}^{2\times 2}(s)$ be

$$G(s) = \begin{bmatrix} g_{11}(s) & \alpha g_{11}(s) \\ g_{12}(s) & \alpha g_{12}(s) \end{bmatrix}.$$

Clearly $G(s)$ is rank deficient over $\mathbb{R}(s)$. Consider

$$X = \begin{bmatrix} I_2, F \end{bmatrix}$$

$$F = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}.$$

Then by condition (4.12), in this case, since

$$\begin{bmatrix} I_2, F \end{bmatrix}\begin{pmatrix} G(s) \\ I_2 \end{pmatrix} = \begin{bmatrix} g_{11}(s)+1 & g_{12}(s)+2 \\ \alpha g_{11}(s)+2 & \alpha g_{12}(s)+4 \end{bmatrix}$$

is singular for all s, $G(s)$ is said to be degenerate.

Example 4.2: In this case take

$$G(s) = \begin{bmatrix} g_{11}(s) & g_{12}(s) \\ 0 & g_{22}(s) \end{bmatrix}.$$

Choose $\mathbf{X} = \ker X$ with

$$X = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then we get

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} g_{11}(s) & g_{12}(s) \\ 0 & g_{22}(s) \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & g_{22}(s) \\ 0 & 1 \end{bmatrix}$$

which is clearly singular for all $s \in \mathbf{C}$.

The significance of these two examples is now discussed. We consider a partitioning of the class of degenerate transfer functions into two mutually exclusive subclasses as follows:

(c1) Subclass 1 contains those degenerate transfer functions for which $\Gamma_G \subseteq$ Grass$(m,p+m)$ is contained in some Schubert hypersurface

$$\sigma(\mathbf{X}) \subseteq \text{Grass}(m,p+m)$$

where XB can be represented as the kernel of a $m \times (p+m)$ matrix of the form

$$X = \left[I_m \, , \, F\right]$$

with F a constant $m \times p$ matrix.

(c2) Subclass 2 contains those degenerate transfer functions where **X** cannot be represented as above.

Clearly example 1 is in subclass 1 and example 2 is in subclass 2.

The subclass 2 of degenerate transfer functions can be defined by the following alternate condition:

(c2') For $G(s)$ in subclass 2, Γ_G is contained in $\sigma(\mathbf{X})$ for some $\mathbf{X} \in$ Grass$(p,m+p)$ for which

$$\dim(\mathbf{X} \cap \mathbf{U}) > 0. \tag{4.13}$$

We remark that such a subspace $\mathbf{X}$ implies a direct sum decomposition of the space $\mathbf{U} \oplus \mathbf{Y}$ as

$$\mathbf{U} \oplus \mathbf{Y} = (\mathbf{U}_1 \oplus \mathbf{Y}_1) \oplus (\mathbf{U}_2 \oplus \mathbf{Y}_2) \tag{4.14}$$

according to constraints of the form

$$X_1\, u(s) = -X_2\, y(s) \tag{4.15}$$

where $\mathbf{X} = \ker[X_1, X_2]$ with $X_1 \colon \mathbf{C}^m \to \mathbf{C}^m$ singular; i.e., $u_2(s) \in \mathbf{U}_2 = \ker X_1$ (u_2 is underdetermined) and $y_2(s) \in \mathbf{Y}_2$ satisfies $y_2(s) \equiv 0$ from (4.15).

Next we consider the question as to whether this classification of degenerate transfer functions is invariant under Σ-group action.

Theorem 4.4: The classification of degenerate transfer functions $G(s)$ according to whether (or not) the Nyquist locus $\Gamma_G \subset \sigma(\mathbf{X})$ for some $\mathbf{X} = \ker[I_m, F]$ with F finite (or not) is invariant under the action of the feedback group (Σ_0-group). However, this classification is ***not*** invariant under the more general action of the Σ-group.

Proof: Given $\Gamma_G \subset \sigma(\mathbf{X})$ in Grass$(m, p+m)$ it is clear that the action of Σ applied to Γ_G (resp. $\sigma(\mathbf{X})$) yields Γ_G^Σ (resp. $\sigma(\mathbf{X}^\Sigma)$) such that $\Gamma_G^\Sigma \subset \sigma(\mathbf{X}^\Sigma)$. So the question amounts to whether or not we can go by Σ-group action from

$$\mathbf{X} = \ker \left[I_m \,,\, F\right]$$

to

$$\mathbf{X}^\Sigma = \ker \left[X_1 \,,\, X_2\right]$$

with $X_1 \in \mathbf{C}^{m \times m}$ singular and vice versa. The general Σ-group action on **X** takes the form (with $\mathbf{X} = \ker\,[I_m, F]$),

$$\left[I_m\,,\,F\right]\begin{bmatrix} S & -K \\ H & T \end{bmatrix} = \left[S + FH\,,\,FT - K\right].$$

Thus given $F \in \mathbb{C}^{p\times m}$ we seek an $S \in \mathbf{C}^{m\times m}$ nonsingular, and $H \in \mathbb{C}^{p\times m}$ such that for some $x \in \mathbf{C}^m$

$$\left[S + FH\right] x = 0 .$$

We give a constructive procedure. Assume $m>p$ and choose any $S \in \mathbf{C}^{m\times m}$ nonsingular. For instance, let $S = I_m$. Now choose an $H \in \mathbb{C}^{p\times m}$ such that there exists an H_l where $H_l H = I_m$. Thus, it is enough to choose $H_l \in \mathbb{C}^{p\times m}$ such that for some $w \in \mathbb{C}^p$

$$\left[H_l + F\right] w = 0 .$$

Then $x = Hw$ satisfies above condition. This proves the second claim.

The first claim is proved by considering that for the subgroup action of Σ_0 we get

$$\left[I_m\,,\,F\right]\begin{bmatrix} S & -K \\ 0 & T \end{bmatrix} = \left[S\,,\,FT - K\right],$$

where $S \in \mathbf{C}^{m\times m}$ is a (nonsingular) change of basis in **U**. Similarly, starting with $\mathbf{X} = \ker[X_1, X_2]$ with X_1 singular we get

$$\left[X_1, X_2\right]\begin{bmatrix} S & -K \\ 0 & T \end{bmatrix} = \left[X_1 S\,,\,X_2 T - X_1 K\right],$$

where $X_1 S \in \mathbf{C}^{m\times m}$, is singular.

∎

From the above theorem it is clear that the classification is meaningful in this geometric setting only for strictly proper transfer functions. It is well known that strictly proper transfer functions can be realized as

$$G(s) = C\left[sI - A\right]^{-1}B$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{p \times n}$ with minimal dimension n called the McMillan degree of $G(s)$. From the above examples it is clear that degenerate transfer functions are rather special. We have the following results of Brockett and Byrnes.

Theorem 4.5: [BR4] Let $G(s)$ be strictly proper with McMillan degree n. If $mp \leq n$ then nondegeneracy is generic in the space of strictly proper transfer functions. If $mp > n$, then every $G(s)$ is degenerate †.

Proof: (cf. [BR4], Thm. 4.2).

The notion of stability margins as discussed in Chapter 2 for SISO systems involves a measure of how close the Nyquist contour Γ approaches a "critical point", $\frac{-1}{f}$ in the complex plane. In this geometric setting we will be concerned with how nearly Γ_G intersects the appropriate fixed hypersurface $\sigma(\mathbf{F})$ for $\mathbf{F} = \ker[I_m, F]$. In chapter 5 we will exploit the decomposition of certain degenerate transfer functions to design decentralized control systems. Considering the last theorem we will need to approximate nondegenerate transfer functions by degenerate ones. In preparation for these results we discuss in the next section the topology of the Grassman manifold.

† In the context of this discussion, the statement that nondegeneracy is generic means that the set of all degenerate transfer functions can be defined by algebraic equations.

4.2. Angle Topology of the Grassman Space

4.2.1. Plücker Metric

As discussed above the Grassman space, Grass(p,n), is the space of all p-dimensional subspaces of $\mathbf{C}^n$. Clearly for any $\mathbf{N} \in$Grass(p,n) we can express p basis vectors for $\mathbf{N}$ in terms of some coordinate systems by writing an $n \times p$ matrix say B_1. In terms of some other coordinate system we can write a new matrix B_2. In this case there exists a nonsingular $p \times p$ matrix A such that $B_1 = B_2 A$ and conversely. Thus $\mathbf{N} = \text{image}(B_1) = \text{image}(B_2)$.

Following Byrnes, et al [BY1] we state

Definition: The Plücker coordinates of a $p \times n$ matrix B is the $\binom{p}{n}$ dimensional vector of determinants of all $p \times p$ submatrices of B.

It is easy to show that for any B_1 and B_2 as above their respective Plücker coordinates will differ by a scalar multiple of each other; i.e., their Plücker coordinates are aligned. One way to provide a notion of distance on Grass(p,n) is then to think of any two subspaces of dimension p say $\mathbf{N},\mathbf{M} \in$Grass($p,n$) in terms of their Plücker coordinates (say $u_N, v_M \in \mathbf{C}^{\binom{p}{n}}$). Then the function

$$d(\mathbf{N}, \mathbf{M}) = \frac{(1 - u_N^* v_M)}{\|u_N\|_2 \|v_M\|_2} \tag{4.16}$$

is an "angle" metric and obeys the property

$$-1 < d(\mathbf{N}, \mathbf{M}) < 1 .$$

4.2.2. Gap Metric

Another sort of angle metric can be described by thinking of points in Grass(p,n) in terms of the p-dimensional subspaces of $\mathbf{C}^n$ that they represent. Here we take an abstract "basis-free" viewpoint in describing the subspaces.

Let $\mathbf{M},\mathbf{N} \subseteq \mathbf{C}^n$ be subspaces. Let dim $\mathbf{M} = p$ and dim $\mathbf{N} = m$. The following function called the gap (or aperture) between $\mathbf{M}$ and $\mathbf{N}$ is defined in Kato [KA1,pp.197].

Definition: The gap between $\mathbf{M}$ and $\mathbf{N}$ is

$$\delta(\mathbf{M},\mathbf{N}) = \max \left\{ \sup_{\substack{\|x\|=1 \\ x\in\mathbf{M}}} \inf_{y\in\mathbf{N}} \|x-y\|, \sup_{\substack{\|y\|=1 \\ y\in\mathbf{N}}} \inf_{x\in\mathbf{M}} \|y-x\| \right\} \tag{4.17}$$

The gap function obeys the following properties which follow immediately from the definition.

(P1) $\delta(\mathbf{M},\mathbf{N}) = 0$ if and only if $\mathbf{M} = \mathbf{N}$.

(P2) $\delta(\mathbf{M},\mathbf{N}) = \delta(\mathbf{N},\mathbf{M})$.

(P3) $0 \le \delta(\mathbf{M},\mathbf{N}) \le 1$.

Furthermore, the gap function obeys the following property which will be significant for us:

(P4) $\delta(\mathbf{M},\mathbf{N}) < 1$ if and only if dim($\mathbf{M}$) = dim($\mathbf{N}$) (cf. [KA1, pg. 200]).

In general, the gap, $\delta(\mathbf{M},\mathbf{N})$, is not a metric.† However, after modifying the definition slightly by taking infimums over the appropriate unit ball in each subspace, the resulting modified gap, $\hat{\delta}(\mathbf{M},\mathbf{N})$, obeys the triangle inequality and thus becomes a metric.

† In the sequel we will consider the gap based on the euclidean vector norm. In this case the gap is a metric.

Property (P4) clearly indicates that using this modified "gap-metric" one can actually form a basis for a topology of the Grassman space via neighborhoods of the form

$$B_\varepsilon(\mathbf{M}) = \left\{\mathbf{N} \in \text{Grass}(p,n) \colon \delta(\mathbf{M},\mathbf{N}) \leq \varepsilon < 1\right\}.$$

4.2.3. Orthogonal Projections in Unitary Space and the Gap Metric

In a unitary space, E^n, we can employ the notion of an orthogonal projector to represent a subspace $\mathbf{M} \subseteq E^n$; e.g., take $E^n = \mathbf{C}^n$ and the natural inner product $\langle x,y \rangle = \bar{x}^T y$. If $P_{\mathbf{M}}$(resp. $P_{\mathbf{N}}$) is an orthogonal projector whose range is the subspace $\mathbf{M} \subseteq E^n$ (resp. $\mathbf{N} \subseteq E^n$) then using the natural Euclidean norm we can state the following:

Theorem 4.6:

$$\delta(\mathbf{M},\mathbf{N}) = \|P_{\mathbf{M}} - P_{\mathbf{N}}\|_2 .$$

Proof: (cf. Kato [KA1]).

Property (P4) of the gap is then related to the following fact.

Theorem 4.7: Any two orthogonal projectors $P_{\mathbf{M}}$, $P_{\mathbf{N}}$ which satisfy $\|P_{\mathbf{M}} - P_{\mathbf{N}}\| < 1$ are unitarily equivalent. That is, there exists a unitary transformation U such that $UP_{\mathbf{M}}U^* = P_{\mathbf{N}}$. ($U$ is unitary if $U^*U = I$).

Proof: (cf. Kato [KA1]).

Unitary transformations have an intuitive geometric appeal because they represent orthogonal rotations of the given vector space coordinate system. Thus we see that with the structure of a unitary space the gap-metric (here the gap function $\delta(\bullet,\bullet)$ becomes naturally a metric) takes on a particularly natural geometric appeal. Indeed, in finite dimensional unitary spaces, for which we have interest, the transformation U of theorem 4.7 can be represented by an easily computable matrix. In Kato [KA1] these results (and others) are used to study perturbations of linear operations on infinite dimensional spaces. In Stewart [ST1] similar ideas are applied to certain numerical problems in the computation of invariant subspaces for matrix (finite dimensional) operators. As we discuss in the subsequent sections our concern is slightly different but will follow along the same line of reasoning.

4.2.4. Near Intersection Between Subspaces and the Minimum Gap Function

From the statement of the generalized Nyquist criterion above it is clear that we will be interested in characterizing the "near" intersection between certain pairs of subspaces. On the Grassmanian manifold this is characterized by near intersection between a point $\mathbf{M}\in\text{Grass}(p,n)$ and a Schubert hypersurface $\sigma(\mathbf{N})\subseteq\text{Grass}(p,n)$ associated with the subspace $\mathbf{N}\in\text{Grass}(n-p,n)$.

Towards this end we provide the following:

Definition: Let $\mathbf{M}\subseteq\mathbf{C}^n$ be a p-dimensional subspace and $\mathbf{N}\subseteq\mathbf{C}^n$ an m-dimensional subspace. Then the *minimum gap* (or *min-gap*) between $\mathbf{M}$ and $\mathbf{N}$ in $\mathbf{C}^n$ is given by,

$$\gamma(\mathbf{M},\mathbf{N}) = \min\left\{ \inf_{\substack{\|x\|=1 \\ x\in\mathbf{M}}} \inf_{y\in\mathbf{N}} \|x-y\|,\ \inf_{\substack{\|y\|=1 \\ y\in\mathbf{N}}} \inf_{x\in\mathbf{M}} \|y-x\| \right\} \quad (4.18)$$

Obviously, the minimum gap satisfies the properties:

(P1) $0 \le \gamma(\mathbf{M},\mathbf{N}) \le \delta(\mathbf{M},\mathbf{N})$

(P2) $\gamma(\mathbf{M},\mathbf{N}) = 0$ if and only if $\dim(\mathbf{M} \cap \mathbf{N}) > 0$.

Based on (P2) it is clear that for the abstract Nyquist criterion described in section 2.3.2 that the min-gap can provide a measure of distance between the abstract Nyquist contour Γ_G and the abstract critical point $\sigma(\mathbf{F})$ as

$$\min_{\mathrm{Re}s=0} \gamma(\mathbf{G}(s), \mathbf{F}), \tag{4.19}$$

where $\mathbf{G}(s) \in \mathrm{Grass}(m, p+m)$ and $\mathbf{F} \in \mathrm{Grass}(p, n)$. In section 4.3 we consider this further.

Following the line of reasoning of section 3.3 we make the following claim.

Corollary 4.8: If $P_{\mathbf{M}}$ and $P_{\mathbf{N}}$ are both orthogonal projectors in $\mathbf{C}^n$ with image$(P_{\mathbf{M}}) = \mathbf{M}$, image$(P_{\mathbf{N}}) = \mathbf{N}$ then

$$\gamma(\mathbf{M}, \mathbf{N}) = \| P_{\mathbf{M}} - P_{\mathbf{N}} \|_2 .^{\ddagger} \tag{4.20}$$

4.2.5. Canonical Angles Between Subspaces

There is a natural notion of angles between pairs of subspaces in a unitary space. In finite dimensional spaces these angles can be computed from singular values of a particular matrix. If we let $\mathbf{M},\mathbf{N}$ be a pair of subspaces of $\mathbf{C}^n$ with $\dim \mathbf{M}=p$, $\dim \mathbf{N}=m$. Assume $m>p$. Then we say the smallest angle between $\mathbf{M}$ and $\mathbf{N}$ (cf. [BJ1]), $\vartheta_1(\mathbf{M},\mathbf{N}) = \vartheta_1 \in [0, \frac{\pi}{2}]$, is given by

‡ Note that in finite dimensional unitary spaces that the right hand side of (4.20) is just the minimum singular value of the matrix difference.

$$\cos \vartheta_1 = \max_{\substack{u \in \mathbf{M} \\ \|u\|_2 = 1}} \max_{\substack{v \in \mathbf{N} \\ \|v\|_2 = 1}} u^* v . \tag{4.21}$$

Following Björck and Golub [BJ1] we define recursively the *principal angles*, ϑ_k , $k=1,...,p$ as follows.

Definition: The principal angles $\vartheta_k \in [0, \frac{\pi}{2}]$ between **M** and **N** are given recursively for $k=1,2, \cdots ,p$ by

$$\cos \vartheta_k = \max_{\substack{u \in \mathbf{M} \\ \|u\|_2 = 1}} \max_{\substack{v \in \mathbf{N} \\ \|v\|_2 = 1}} u^* v = u_k{}^* v_k \tag{4.22}$$

subject to the constraints

$$u_j{}^* u = 0 , \; v_j{}^* v = 0 \tag{4.23}$$

for $j=1, \cdots ,k-1$. We call the set of vectors $\{u_1, \cdots ,u_p, v_1, \cdots , v_p\}$ the *principal vectors* for the pair of subspaces.

In this section we review how the principal angles can be computed for a pair of subspaces. The relation between certain principal angles and the gap will be clarified using orthogonal projectors. The result will be a computational procedure for determining the gap, $\delta(\bullet,\bullet)$, and the min-gap, $\dot{\gamma}(\bullet,\bullet)$, between a pair of subspaces **M**,**N**. Moreover using the principal vectors we can compute a basis for the intersection, $\mathbf{M} \cap \mathbf{N}$. For the problem of multivariable feedback such a basis can be used to describe how certain modal behavior of the system is reflected from an input-output view point.

The main computational result which we exploit requires that we have a unitary basis for each of the subspaces **M** and **N**. Since this can be obtained conceptually using a Gram-Schmidt procedure (and in practice using Householder reflections) we assume that we have a pair of matrices $Q_M \in \mathbf{C}^{n \times m}, Q_N \in \mathbf{C}^{n \times p}$ with $Q_N{}^* Q_N = I_m$ and $Q_M{}^* Q_M = I_p$ such that image$(Q_M) = \mathbf{M}$ and image$(Q_N) = \mathbf{N}$.

Theorem 4.9: Given Q_M and Q_N such that image$(Q_M) = \mathbf{M}$ and image$(Q_N) = \mathbf{N}$ each a subspace of $\mathbf{C}^n$. Compute the singular value decomposition (SVD) of

$$Q_M{}^* Q_N = Y_M \, C \, Y_N{}^* \tag{4.24}$$

where

$$Y_M{}^* Y_M = Y_M Y_M{}^* = Y_N{}^* Y_N = I_p \, , \tag{4.25}$$

$$C = \cos\Theta = \operatorname{diag}\left[\sigma_1, \ldots, \sigma_p\right]$$

with singular values $\sigma_1 \geq \cdots \geq \sigma_p$ and

$$\Theta = \operatorname{diag}\left[\vartheta_1, \ldots, \vartheta_p\right].$$

Then $\vartheta_1 \leq \cdots \leq \vartheta_p$ are the principal angles between $\mathbf{M}$ and $\mathbf{N}$. The columns of and $U = Q_M Y_M$, $V = Q_N Y_N$ are the principal vectors.

Proof: (cf. [BJ1,thm.1]).

■

Corollary 4.10: Let $P_{\mathbf{M}} = Q_M Q_M{}^*$ be an orthogonal projector on $\mathbf{M}$. Then compute the SVD of

$$\left(I_n - P_{\mathbf{M}}\right) Q_N = W_M \, S \, Y_N{}^* \tag{4.26}$$

where $S = \sin\Theta$. Here W_M gives the principal vectors in the orthogonal complement, $\mathbf{M}^{per}$, associated with the pair of subspaces $\mathbf{M}, \mathbf{N}$.

Theorem 4.11: As above, let $P_{\mathbf{M}}$ and $P_{\mathbf{N}}$ be orthogonal projectors on **M** and **N** respectively. Then the nonzero eigenvalues of $P_{\mathbf{M}} - P_{\mathbf{N}}$ are $\pm\sin\vartheta_i$ for $i=1, \cdots, p$.

Proof: (cf. [ST1,thm.2.5]).

Finally we can state as a corollary to theorem 4.11:

Corollary 4.12: With the above notation

$$\delta(\mathbf{M},\mathbf{N}) = \|P_{\mathbf{M}} - P_{\mathbf{N}}\|_2 = \sin\vartheta_p\,. \tag{4.27}$$

$$\gamma(\mathbf{M},\mathbf{N}) = [\![P_{\mathbf{M}} - P_{\mathbf{N}}]\!]_2 = \sin\vartheta_1\,. \tag{4.28}$$

Proof: See Theorem 4.6 and corollary 4.8.

4.2.6. Computational Procedures for Obtaining the Gap and Min-Gap

Following the procedure suggested by corollaries 4.12 and 4.10 we can provide a straightforward, numerically stable, procedure for computing the gap or min-gap between a pair of finite dimensional subspaces $\mathbf{M},\mathbf{N}\in\mathbf{C}^n$ in terms of some matrix representations. In this most general form, the procedure is computationally intensive.

Let **M**,**N** be represented as $\mathbf{M}=\text{image}(M)$, $\mathbf{N}=\text{image}(N)$ where $M\in\mathbf{C}^{n\times p}$, $N\in\mathbf{C}^{n\times m}$. The following procedure can be coded directly for computer solution using, for instance, LINPACK routines [DO1].

Procedure for Computing the Gap or Min-Gap

Given: M, N, a pair of $n \times p$ (resp. $n \times m$) matrices

Step_1: Obtain a unitary basis for **M**. Conceptually, this is done by obtaining a QR factorization of M

$$M = \left[Q_M \, , \, Z_M\right] \begin{bmatrix} R_M \\ 0 \end{bmatrix}$$

where R_M is right triangular matrix and $[Q_M \, , Z_M]$ is unitary. Then Q_M is the required $n \times p$ matrix of unitary bases for **M**.

Step_2: Obtain a unitary basis for **N**. Again, employ a QR factorization as

$$N = \left[Q_N \, , \, Z_N\right] \begin{bmatrix} R_N \\ 0 \end{bmatrix}$$

with R_N right triangular. Now, Z_N is the $n \times (n-p)$ matrix of bases for $\mathbf{M}^{\mathrm{per}}$.

Step_3: To compute the gap, $\delta(\mathbf{M}, \mathbf{N})$, (resp. min-gap, $\gamma(\mathbf{M}, \mathbf{N})$) obtain the maximum (resp. minimum) singular value of the $(n-p) \times m$ matrix

$$Z_N^* \, Q_M \, .$$

Remark: The QR factorization outlined here can be performed using a numerically stable algorithm involving the use of Householder reflections to compute the transformations. This has been implemented efficiently in LINPACK routine CQRDC [DO1].

Remark: The product of unitary matrices can be obtained in a numerically stable way. Then application of a standard algorithm can provide the required singular value. The routine CSVDC is available in LINPACK for computing these quantities [DO1].

Computations for examples in the later sections of this dissertation were completed using this procedure codified in MATLAB [MO1]. The MATLAB code is included in appendix C.

4.3. A Geometric Stability Margin Based on the Minimum Gap

In section 2.3.3 we discussed some background on classical notions of stability margins for SISO feedback and extensions of these ideas to the case of MIMO feedback. We suggested that a combined gain-phase margin, $g_{\vartheta m}$, given in (2.39) could be thought of as the euclidean distance in the complex plane between the Nyquist locus, Γ_{fg}, where $fg(s)$: $D \rightarrow \Gamma_{fg}$ and a critical point at $s=-1$. Alternately, $g_{\vartheta m}$ can be thought of as the euclidean distance between a translated form of Γ_{fg} (which is the image of the contour D under the map $1+fg(s)$, the return difference) and a critical point at the origin. This viewpoint has been appropriately generalized in several ways (cf. eqn.(2.41)) for MIMO case by focusing on a measure of regularity (typically the minimum singular value) of a matrix return difference.

For the purposes of designing feedback compensation, it is usually more convenient to consider a slightly different stability margin; viz.,

$$g'_{\vartheta m} = \inf_{s \in D} | f^{-1} + g(s) | ,$$

which is the euclidean distance between the Nyquist contour, Γ_g (fixed with respect to $g(s)$) and a critical point, $s=-\frac{1}{f}$, depending on the choice of feedback. Furthermore, insight can be gained for the design of dynamic (lead/lag) feedback compensation in this setup by considering a locus of critical points, Γ_f where $-f^{-1}(s)$: $D \to \Gamma_f$ [R01, pp. 56-59].

Clearly any generalization of $g'_{\vartheta m}$ to the case of MIMO feedback based on regularity of the matrix function $F^{-1} + G(s)$ on D can be meaningful only in special cases (e.g. $p=m$ and F nonsingular). However, the geometric Nyquist criterion discussed in section 2.3.2 suggests applying the topology of the Grassman space to construct a measure of how nearly the subspaces $\mathbf{G}(s) = \ker[G(s), -I_p]$ and $\mathbf{F} = \ker[I_m, F(s)]$ intersect in $\mathbf{U} \oplus \mathbf{Y}$.

4.3.1. Definition and Properties of a Geometric Stability Margin

In this dissertation we employ the minimum gap function, $\gamma(\mathbf{N},\mathbf{M})$, between a pair of subspace $\mathbf{N}$ and $\mathbf{M}$ of a unitary space to measure the distance between the abstract Nyquist locus, $\Gamma_G \subseteq \mathrm{Grass}(m, p+m)$ and the Schubert hypersurface, $\sigma(\mathbf{F})$, representing the fixed critical point $\mathbf{F} = \ker[I_m, F] \in \mathrm{Grass}(p, m+p)$.

Definition (Geometric Stability Margin): Given $G(s) \in \mathbf{R}^{p \times m}(s)$ and $F \in \mathbf{R}^{m \times p}$ the *geometric stability margin*, g_{sm}, is a real number $0 \le g_{sm} \le 1$ given by,

$$g_{sm} \triangleq \inf_{s \in D} \gamma(\mathbf{G}(s), \mathbf{F}). \qquad (4.29)$$

In this section we focus on some properties of the geometric stability margin which clarify its relation to more traditional stability margins.

Theorem 4.13: A point $s_0 \in \mathbf{C}$ is a closed loop pole of the feedback equations (4.1) if and only if either of the following holds

(i)

$$\det[I_p + G(s_0)F] = \det[I_m + FG(s_0)] = 0 \qquad (4.30)$$

(ii)

$$\gamma\left(\ker[G(s_0), -I_p], \ker[I_m, F]\right) = 0 \qquad (4.31)$$

Proof:

(i) From (2.17) we see that

$$\det[I_m + FG(s)] = \det[I_m + FG(\infty)]\prod_{i=1}^{n}\frac{(s - p_{ci})}{(s - p_{co})}$$

where n is the McMillan degree of $G(s)$ so that s_0 is a closed loop pole if and only if it is a root of $\det[I_m + FG(s)]$.

Now

$$\begin{bmatrix} I_m & F \\ G(s) & -I_p \end{bmatrix} = \begin{bmatrix} I_m + FG(s) & F \\ 0 & -I_p \end{bmatrix}\begin{bmatrix} I_m & 0 \\ -G(s) & I_p \end{bmatrix},$$

implies that

$$\det\begin{bmatrix} I_m & F \\ G(s) & -I_p \end{bmatrix} = -\det[I_m + FG(s)].$$

Also,

$$\begin{bmatrix} I_m & F \\ G(s) & -I_p \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ G(s) & I_p + G(s)F \end{bmatrix}\begin{bmatrix} I_m & F \\ 0 & -I_p \end{bmatrix}.$$

implies that

$$\det\begin{bmatrix} I_m & F \\ G(s) & -I_p \end{bmatrix} = -\det[I_p + G(s)F].$$

(ii) From section 4.2 we see by definition of the minimum gap function that

$$\gamma\Big[G(s_0),\, F\Big] = 0.$$

if and only if

$$\dim(G(s_0) \cap F) > 0. \qquad \blacksquare$$

Next we would like to show that for,

(1)

$$\varphi_1(s) = \sigma_{\min}[\, I_m + FG(s)],$$

(2)

$$\varphi_2(s) = \sigma_{\min}[I_p + G(s)F],$$

(3)

$$\varphi_3(s) = \gamma(G(s),\, F),$$

which map the closed contour $D \subseteq \mathbf{C}$ into $\mathbf{R}_+$ (resp. $[0,1] \in \mathbf{R}$ for (3)), that, if for some s_1, $\varphi_1(s_1)$ achieves its minimum on D then $\varphi_2(s_1)$ and $\varphi_3(s_1)$ also achieve their respective minimums at s_1. To do this we consider some further aspects of the topology of Grassman manifolds.

Lemma 4.14: With $Y \in \text{Grass}(p, m+p)$ the set

$$B_\varepsilon(Y) = \Big\{X \in \text{Grass}(p, m+p):\ \delta(XB,\, YB) < \varepsilon\Big\}, \qquad (4.32)$$

is a convex, neighborhood of $\text{Grass}(p, m+p)$ for $\varepsilon < 1$.

Proof: We must show that the set $B_\varepsilon(Y)$ is convex. Normally, a convex set is one for which a straight line segment between any two points in the set is itself contained in the set. Here the "straight line segment" consists of a curve in Grass($p,p+m$). We employ a simple parametrization suggested by Stewart [ST1].

To make the argument concrete we represent points in Grass($p,p+m$) in terms of unitary matrices. Thus let $X \in$ Grass($p,p+m$) be represented as

$$X = \text{image}\, X_1$$

where $X_1 \in C^{(p+m)\times p}$ satisfies $X_1^* X_1 = I_p$; i.e., X_1 has orthonormal columns which span X. Let $X_2 \in C^{(p+m)\times m}$ have columns which form an orthonormal basis for the perpendicular complement of X in C^{p+m}. Thus $[X_1, X_2]$ is a unitary matrix. Also, given any other $Y \in$ Grass ($p,p+m$) there exists a unitary basis for Y of the form

$$Y_1 = (X_1 + X_2 T)(I_p + T^* T)^{-1/2}$$

where $Y_1 \in C^{(p+m)\times p}$ and $T \in C^{m\times p}$ and the square root denotes the unique positive definite square root of the positive definite matrix $I_p + T^* T$.

We parametrize a "path" between X and Y confined to Grass($p,p+m$) in terms of matrices by parametrizing

$$T_\alpha = \alpha T$$

where $\alpha \in R$, $0 \le \alpha \le 1$.

To show that this "path" is confined to the set $B_\varepsilon(Y)$ if $X \in B_\varepsilon(Y)$ we focus on the relationship between the gap δ and the maximum principal angle ϑ_p given in (4.27).

Let

$$Y_\alpha = (X_1 + X_2 T_\alpha)(I_p + T_\alpha^* T_\alpha)^{-1/2}.$$

To compute the principal angles between $\mathbf{Y}_\alpha$ and $\mathbf{X}$ we obtain (following theorem 4.11) the singular values of the $p \times p$ matrix

$$Y_\alpha^* X_1 = (I + T_\alpha^* T_\alpha)^{-1/2}.$$

Let τ_i for $i=1, \cdots, \min(p,m)$ be the singular values of T. Then we see that the principal angles satisfy

$$\cos \vartheta_i = (1 + \tau_i^2)^{-1/2}.$$

Thus the singular values of T are

$$\tau_i = \tan \vartheta_i .$$

So the question of convexity is answered by recognizing that as α goes from 0 to 1

$$\|T_\alpha\| = \tau_{\max}(\alpha) < \tan \arcsin \varepsilon$$

since obviously

$$\|T_\alpha\| = \|\alpha T\| = \alpha \|T\|.$$

■

Lemma 4.15: Let $\Gamma_G \subseteq \mathrm{Grass}(m, p+m)$ be the abstract Nyquist contour associated with the $G(s) \in R^{p \times m}(s)$ and $\mathbf{F} = \ker[I_m, F]$. For any $\mathbf{X} \in \mathbf{B}_{g_{sm}}(\mathbf{F}) \subseteq \mathrm{Grass}(p, m+p)$

$$N(\Gamma_G; \sigma(\mathbf{F})) = N(\Gamma_G; \sigma(\mathbf{X})). \tag{4.33}$$

Proof:

Assume, for simplicity, that $\varphi_3(s)$ has a unique global minimum occurring at $s_1 \in D$. Then there exists some $X_0 \in \mathrm{Grass}(p, p+m)$ on the boundary of $B_{g_{sm}}(F)$, so that $\delta(X_0, F) = g_{sm}$ and

$$\gamma(\Gamma_G; \sigma(X_0)) = 0.$$

To show the result we must construct a homotopy of subspaces, say $X_\alpha = \ker X_\alpha$, relating F to X_0 which does not leave the manifold. This is done (as in the proof of Lemma 4.14) in terms of matrices as

$$X_\alpha = [Y_1 + Y_2(\alpha T)][I_p + (\alpha T)^*(\alpha T)]^{-1/2}, \qquad (4.34)$$

for some $T \in \mathbb{C}^{m \times p}$ where for instance $X_0 = \mathrm{image}(Y_1)$ and $X_0^{per} = \mathrm{image}(Y_2)$. T is chosen so that $F = \mathrm{image}(X_1)$. Now consider that the condition

$$N(\Gamma_G; \sigma(F)) \neq N(\Gamma_G; \sigma(X_0))$$

holds on the manifold if and only if for some $0 \leq \alpha < 1$ the contour Γ_G intersects the critical point $\sigma(X_\alpha)$. Thus we will be done if we show that no such α exists under the assumptions given.

By the convexity of $B_{g_{sm}}(F)$ and (4.29) we get

$$\delta(X_\alpha, X) < \inf_{s \in D} \gamma(G(s), F)$$

for $0 \leq \alpha < 1$ which completes the proof.

∎

Finally, to clarify the extent to which g_{sm} provides similar information as $g_{\vartheta m}$ with respect to gain variations in F we provide theorem 4.16. With respect to an appropriately constructed contour D (which avoids poles of $G(s)$ on $j\omega$ axis) the values $\gamma(G(s), F)$ on D form a proper subset of $[0,1] \subset \mathbb{R}$ which is both closed and bounded. Therefore, we can replace the definition (4.29) with

$$g_{sm} = \min_{s \in D} \gamma(G(s), F). \tag{4.29'}$$

In the following theorem we will consider the case (which is most typical in practice) when the set

$$\arg \min_{s \in D} \gamma(G(s), F)$$

consists of a single point $s^* \in D$. More generally, this set will consist of a countable number of points on D.

Theorem 4.16: (Main result of chapter 4) If

$$g_{sm} = \min_{s \in D} \gamma(G(s), F) > 0, \tag{4.35}$$

and

$$s^* = \arg \min_{s \in D} \gamma(G(s), F), \tag{4.36}$$

then there exits a "gain" $K_1 \in \mathbb{C}^{p \times p}$ such that

$$\min_{s \in D} \gamma(G(s), X) = \gamma(G(s^*), X) = 0, \tag{4.37}$$

for some $X = \ker[I_m, FK_1]$ if and only if there exits a $K_2 \in \mathbb{C}^{p \times p}$ such that

$$\inf_{s \in D} \det[I_p + G(s)FK_2] = \det[I_p + G(s^*)FK_2] = 0. \tag{4.38}$$

Moreover, K_1 and K_2 both satisfy

$$\|K_i\| \geq \frac{1}{\|G(s^*)F\|}. \tag{4.39}$$

Proof: With conditions (4.35) and (4.36) assume we can find a K_1

such that (4.37) holds (the proof of the last theorem suggests a constructive way to do this). Now (4.37) holds if and only if

$$\det \begin{bmatrix} I_m & FK \\ G(s^*) & -I_p \end{bmatrix} = 0. \tag{4.40}$$

By theorem 4.13, (4.40) holds if and only if

$$\det [I_m + G(s^*)FK] = 0.$$

This can be stated equivalently using the matrix infimum as

$$[\![I_m + G(s^*)FK]\!] = 0.$$

By construction of D we have that $G(s)F$ is nonsingular for s on D; viz., at s^*. Therefore, (4.40) holds if and only if

$$[\![(G(s^*)F)^{-1} + K]\!] = 0.$$

Recall that for X nonsingular $[\![X]\!] = \|X^{-1}\|^{-1}$. Then by the triangle inequality for matrix norms we get that

$$[\![I_m + G(s^*)FK]\!] \geq [\![(G(s)F)^{-1}]\!] - \|K\|.$$

At s^* this implies that any K for which (4.39) holds also satisfies

$$\|K\| > [\![(G(s^*)F)^{-1}]\!].$$

And (4.39) follows immediately from the definition of the matrix infimum.

■

4.3.2. Application of the Geometric Stability Margin and Some Examples

In this section we seek to demonstrate some salient features of the geometric analysis of stability margins for feedback systems. We indicate, by way of illustration, that the geometric stability margin proposed in the previous section has some peculiar properties which can extend its application to more general settings than the classical case. Indeed, even for SISO analysis, the geometric analysis can provide additional useful information which can be lost using the classical approach. This section can be considered a sequel to section 2.3.1 (which should be reviewed at this time) in which some of the limitations of that analysis are relaxed.

As in the previous section, our discussion focuses on the properties of the maps

$$\varphi_1(s) = \sigma_{\min}[\, I_m + FG(s)],$$

$$\varphi_2(s) = \sigma_{\min}[I_p + G(s)F],$$

$$\varphi_3(s) = \gamma(\mathrm{G}(s), \mathrm{F});$$

viz., their respective local minima on D,

$$g_{\vartheta m\,1} = \inf_{s \in D} \varphi_1(s),$$

$$g_{\vartheta m\,2} = \inf_{s \in D} \varphi_2(s),$$

and

$$g_{\vartheta m\,3} = \inf_{s \in D} \varphi_3(s).$$

In the SISO (classical) case where $p = m = 1$ we get

$$\varphi_1(s) = 1 + fg(s) = \varphi_2(s)$$

regardless of where the single loop is broken (cf. Fig. 2.2). However, $\varphi_3(s)$ is fundamentally different from $\varphi_1(s) = \varphi_2(s)$ even in this case.

Corollary 4.17: In the SISO case where $p = m = 1$ the geometric stability margin becomes,

$$g_{sm} = \inf_{s \in D} \varphi_3(s) \tag{4.41}$$

where

$$\varphi_3(s) = \frac{|g(s)f + 1|}{\sqrt{(1 + |g(s)|^2)(1 + |f|^2)}}. \tag{4.42}$$

Proof: We consider a constructive approach based on the computational procedure given in section 4.2.6. We can represent the two subspaces alternately as,

$$\mathbf{G}(s) = \ker[g(s), -1] = \text{image}\begin{bmatrix} 1 \\ g(s) \end{bmatrix},$$

and

$$\mathbf{F} = \ker[1, f] = \text{image}\begin{bmatrix} -f \\ 1 \end{bmatrix}.$$

Thus we obtain normalized basis vectors for the pair of 1-dimensional subspaces as,

$$\mathbf{F} = \text{image}\ \frac{\begin{bmatrix} -f \\ 1 \end{bmatrix}}{\sqrt{1 + |f|^2}}.$$

and for the orthogonal compliment,

$$\mathbf{G}^{per} = \text{image}\ \frac{\begin{bmatrix} -\bar{g}(s) \\ 1 \end{bmatrix}}{\sqrt{1 + |g(s)|^2}}.$$

$$G(s)F + I_p = R_G^*(s)Q_G^*(s)Q_F R_F. \tag{4.43}$$

The relation (4.43) amounts to nothing more than an algebraic statement of the computational procedure used.

Example 4.3: To illustrate these ideas we consider a simple SISO example. For some loop breaking let the loop transmission be

$$fg(s) = \frac{1}{s(s+1)(2s+1)}.$$

The resulting Nyquist contour for $s = j\omega$ with $\omega \in [0.25, 8.0]$ is displayed in Figure 4.1. The relevant euclidean distance between this curve and the critical point at $s = -1$ is given by $\phi_2(j\omega)$ which is displayed in figure 4.2 giving $g_{\vartheta m} = 0.7$ occurring at $\omega^* = 1$ rad/sec.

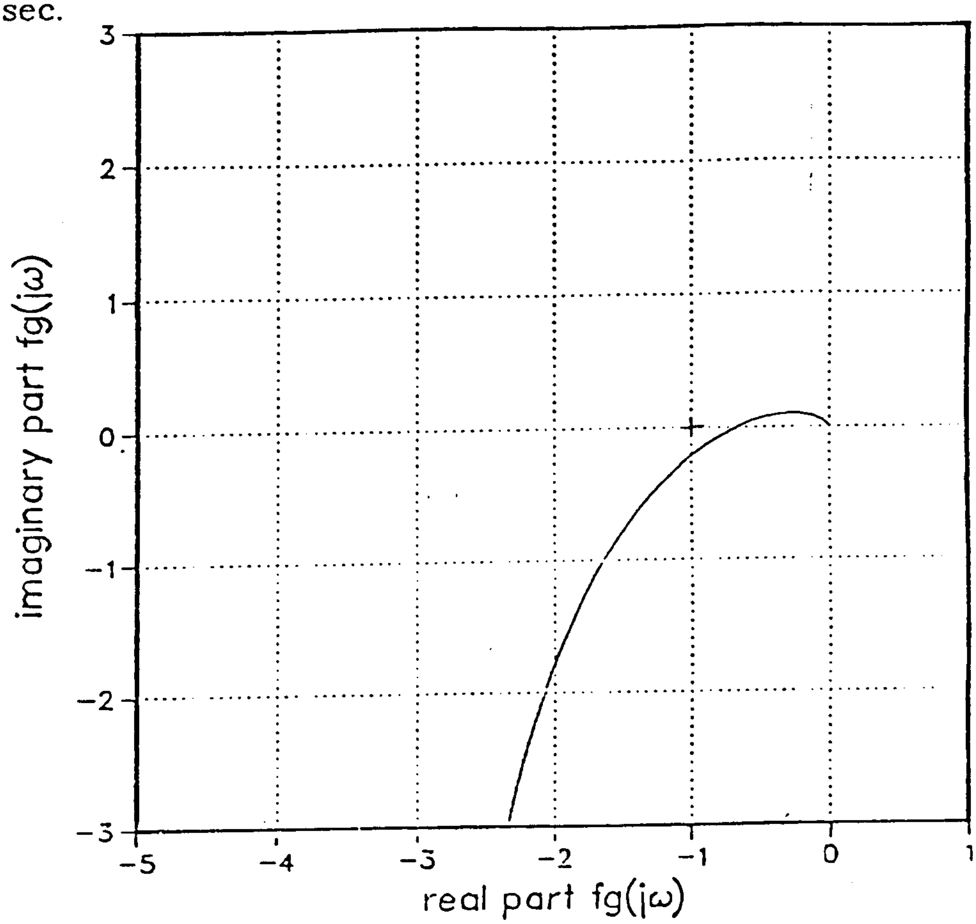


Figure 4.1: The Nyquist plot for Example 4.3.

$$G(s)F + I_p = R_G^*(s)Q_G^*(s)Q_F R_F . \tag{4.43}$$

The relation (4.43) amounts to nothing more than an algebraic statement of the computational procedure used.

Example 4.3: To illustrate these ideas we consider a simple SISO example. For some loop breaking let the loop transmission be

$$fg(s) = \frac{1}{s(s+1)(2s+1)} .$$

The resulting Nyquist contour for $s = j\omega$ with $\omega \in [0.25, 8.0]$ is displayed in Figure 4.1. The relevant euclidean distance between this curve and the critical point at $s = -1$ is given by $\varphi_2(j\omega)$ which is displayed in figure 4.2 giving $g_{\vartheta m} = 0.7$ occurring at $\omega^* = 1$ rad/sec.

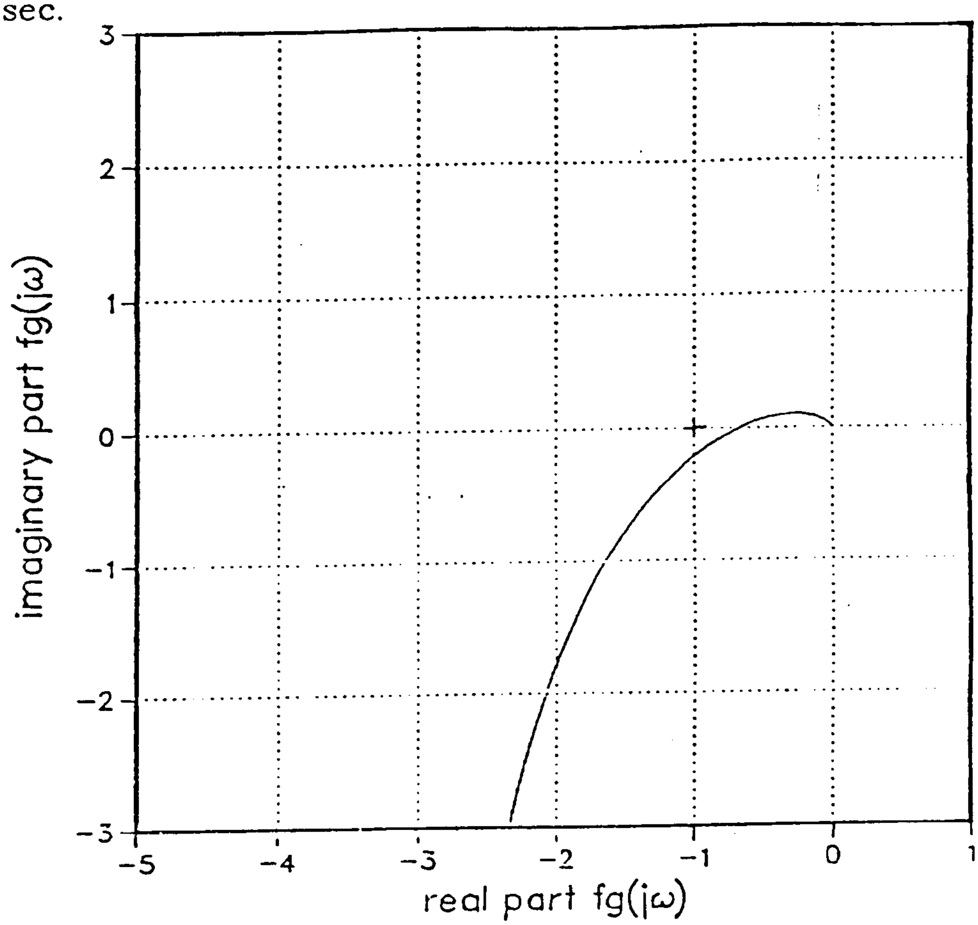


Figure 4.1: The Nyquist plot for Example 4.3.

The curve $\varphi_3(j\omega)$ is displayed in figure 4.3 giving $g_{sm} \doteq 0.41$ at $\omega^* =$ 1 rad./sec. The value $g_{sm} = 0.41$ suggests a minimum principle angle between

$$\ker\left[1\,,\,1\right]$$

and

$$\ker\left[\frac{1}{s(s+1)(2s+1)}\,,\,-1\right]$$

for $s=j\omega$ of ϑ=24.2 degrees. The significance of the asymptotic value $\varphi_3(0) = 0.707$ (or ϑ = 45 degrees) comes from the following observation. With reference to (4.1') for $g(s) = 1/s(s+1)(2s+1)$ we see that $g(s)$ has a pole at the origin. (Of course D would be appropriately indented to avoid this pole.) Thus $g(j\omega)\rightarrow\infty$ as $\omega\rightarrow 0$ which in the geometric picture of (4.1') implies that

$$\ker[g(j\omega)\,,\,-1] \rightarrow \mathbf{Y}$$

in terms of the angle (gap) metric. Thus 45 degrees is just the angle between $\mathbf{U} = \ker[1,0]$ and $\mathbf{F} = \ker[1,1]$.

It is important to recognize that the geometric stability margin analysis we are discussing appropriately generalizes, from classical frequency domain analysis, the notion of "distance" between a Nyquist contour and a fixed point *without* employing the return difference. Motivated by a discussion of some physical considerations in feedback design (cf. corollary 2.4) we considered in section 2.3.1 some limitations of the analysis of closed loop stability based on the return difference. In particular, the Horowitz criterion (2.11) is inferred from the observation that any physical feedback control system will involve *dynamic* compensation with a *dynamic* plant (i.e., $g(s), f(s) \in \mathbf{R}_{sp}(s)$). The complication for stability

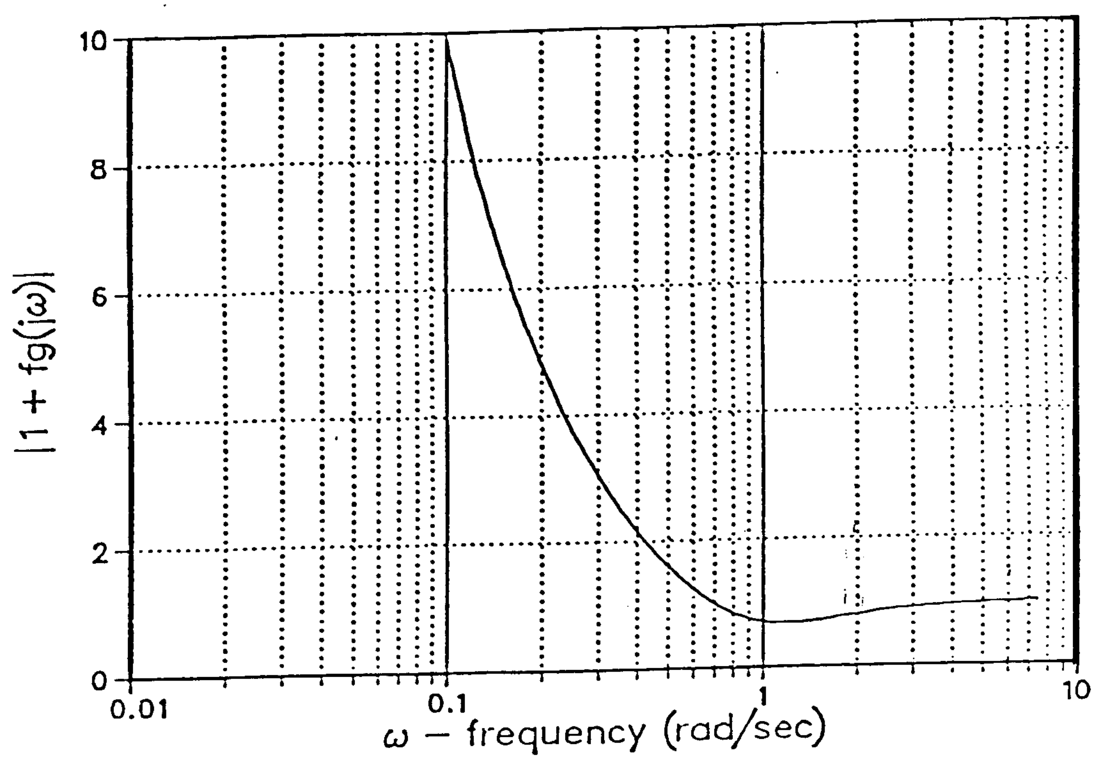


Figure 4.2: $\varphi_2(j\omega)$ vs. ω for example 4.3.

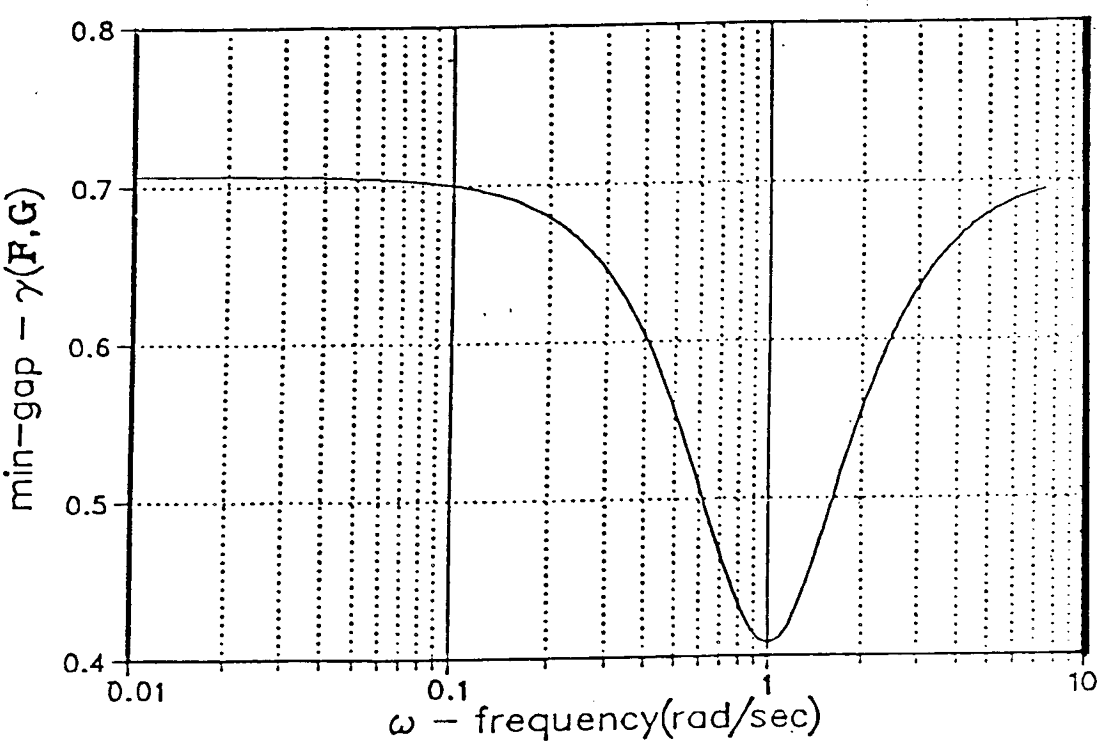


Figure 4.3: $\varphi_3(j\omega)$ vs. ω for example 4.3.

analysis involves the possible existence of *internal poles* (cf. (2.18)) of the resulting closed loop transfer function. Such internal poles exist due to possible cancellations in forming the loop transmission $F(s)G(s)$ when the McMillan degree of $F(s)G(s)$ is strictly less than the sum of the respective McMillan degrees of $F(s)$ and $G(s)$.

In terms of the geometric picture of feedback,

$$\begin{bmatrix} I_m & F(s) \\ G(s) & -I_p \end{bmatrix} \begin{pmatrix} u(s) \\ y(s) \end{pmatrix} = 0, \tag{4.44}$$

we see that we may have difficulty in estimating the regularity, $[\![\cdot]\!]$, for all s on D of

$$\Sigma = \begin{bmatrix} I_m & F(s) \\ G(s) & -I_p \end{bmatrix}$$

as a map on $\mathbf{U} \oplus \mathbf{Y}$ in terms of a transformed basis,

$$\begin{bmatrix} I_m & F(s) \\ G(s) & -I_p \end{bmatrix} \begin{bmatrix} I_m & F(s) \\ G(s) & -I_p \end{bmatrix} = \begin{bmatrix} I_m + F(s)G(s) & 0 \\ 0 & I_p + G(s)F(s) \end{bmatrix} \tag{4.45}$$

The basis given by the right hand side of (4.45) suggests we can estimate $[\![\Sigma]\!]$ in terms of the regularity of an operator on $\mathbf{U}$ *or* $\mathbf{Y}$. However, for s_1 in the neighborhood of an internal pole the transformation

$$\begin{bmatrix} I_m & F(s_1) \\ G(s_1) & -I_p \end{bmatrix}$$

will be poorly conditioned despite the fact that $I_m + F(s_1)G(s_1)$ and $I_p + G(s_1)F(s_1)$ may be relatively well conditioned. Moreover, this observation holds despite the fact that,

$$-\det \begin{bmatrix} I_m & F(s) \\ G(s) & -I_p \end{bmatrix} = \det[I_m + F(s)G(s)] = \det[I_p + G(s)F(s)].$$

We remark that $|det(\bullet)|$ (in contrast to the matrix infimum, $[\![\bullet]\!]$) is *not* a measure of regularity which admits any useful perturbation analysis [DA1,DO3-4,HO1].

Before we consider the question of dynamic compensation further we discuss some aspects of the geometric theory of linear systems with respect to a characterization of open loop poles. First, we recognize that the map

$$s \quad \ker[G(s), -I_p]$$

is well defined on the domain $\mathbf{C}\cup\{\infty\}$ - $\{p_{oi}\}$, where $\{p_{oi}\}$ is the set of n open loop poles of $G(s)\in\mathbf{R}^{p\times m}(s)$. More generally, we can consider the equation

$$y(s) = G(s)u(s)$$

for each $s\in\mathbf{C}\cup\{\infty\}$ in terms of the set of all possible ordered pairs in $\mathbf{U}\times\mathbf{Y}$; i.e.,

$$\text{graph}\{G(s)\} = \left\{\Big[u(s),\ G(s)u(s)\Big]\in \mathbf{U}\times\mathbf{Y}\right\}.$$

So the geometric viewpoint focuses on identifying the equivalence between graph$\{G(s)\}$ and $\ker[G(s), -I_p]$ for $s\in\mathbf{C}\cup\{\infty\}$.

Next we remark that given a (left) coprime factorization

$$G(s) = D^{-1}(s)\, N(s)$$

then clearly

$$\mathbf{G}(s) = \ker[G(s), -I_p] = \ker[N(s), -D(s)]$$

has rank m over the field of rational functions. This means that rank $\mathbf{G}(s) = m$ for $s\in\mathbf{C}\cup\{\infty\}$ - $\{p_{oi}\}$. Thus $\mathbf{G}(s)$ can be thought of as an element of Grass(m,$p+m$) for $s\in\mathbf{C}\cup\{\infty\}$ - $\{p_{oi}\}$; but for some sequence $s_n \to p_{oi}$ the sequence $\mathbf{G}(s_n)$ does not have a limit on Grass$(m, p+m)$. Nevertheless, if we think in terms of graph$(G(s))$ it is clear that

$$\mathbf{X}_{oi} = \lim_{n \to \infty} \text{graph}\{G(s_n)\} \subseteq \mathbf{U} \oplus \mathbf{Y}$$

is a subspace of $\mathbf{U} \oplus \mathbf{Y}$ such that

$$\dim(\mathbf{X}_{oi} \cap \mathbf{Y}) \geq 1 .$$

Thus we make the following observation based on the angle topology of the gap.

Theorem 4.18: Let $p_{o1} \in \{p_{oi}\}$ be an internal closed loop pole of (4.44) (cf. (2.18)). Consider any sequence $s_n \in \mathbf{C}$ which approachs p_{o1}. Then

$$\lim_{n \to \infty} \gamma(\mathbf{G}(s_n), \mathbf{F}(s_n)) = 0 .$$

Proof: As usual,

$$\mathbf{G}(s) = \ker[G(s), -I_p]$$

and

$$\mathbf{F}(s) = \ker[I_m, F(s)]$$

are two subspaces of $\mathbf{U} \oplus \mathbf{Y} = \mathbf{C}^{p+m}$. Now p_{o1} is and internal closed loop pole of (4.44) if there exists some $A(s) \in \mathbf{R}^{p \times p}(s)$ such that

$$G(s) = A(s)\, G'(s),$$

$$F'(s) = A(s)\, F(s),$$

and p_{o1} is a pole of $A(s)$. Thus

$$I_m + F(s)G(s) = I_m + F'(s)G'(s).$$

Now as $s_n \to p_{o1}$

$$\ker[A(s)G'(s), -I_p] \to \mathbf{Y}$$

in the sense of the min-gap; i.e., there exists a sequence of unit vectors,

$$v_n \in \ker[G(s_n), -I_p]$$

such that for some $u \in \mathbf{U}$

$$\lim_{n \to \infty} v_n^* u = 1.$$

However,

$$\ker[I_m, F'(s)A^{-1}(s)] \to \mathbf{Y}$$

also in the same sense (since p_{o1} is therefore a zero of $A^{-1}(s)$).

Finally, if s_1 cannot be described as above then it can be described as a pole of some $B(s) \in \mathbf{R}^{m \times m}(s)$ with

$$G'(s) = G(s)B(s)$$

$$F(s) = F'(s)B(s)$$

and the result follows similarly. ∎

Example 4.4: To illustrate the efficacy of the above theorem we consider again example 4.3. Consider a new loop breaking strategy which reveals a "hidden" mode. In Figure 4.4 we have redrawn the closed loop configuration of Example 4.3 to illustrate the new loop breaking where,

$$g(s) = g_1(s)g_2(s) = \frac{1}{s(s+1)(2s+1)}$$

$$g_1(s) = \frac{1}{(2s + 1)[(s + 0.25)^2 + 0.1]},$$

$$g_2(s) = \frac{[(s + 0.25)^2 + 0.1]}{s(s + 1)}.$$

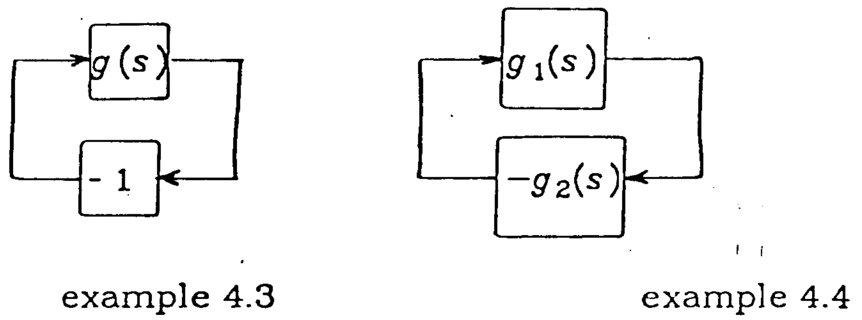


Figure 4.4: The Loop Breaking Configuration for Example 4.4.

In figure 4.5 we plot $\varphi_3(j\omega)$ for $\omega \in [0.1, 8.]$. The significance of the curve in figure 4.5 in the region near ω=0.1 can be explained from figure 4.6 where we plot $g_1(j\omega)$ and $-1/g_2(j\omega)$. Clearly, $\varphi_3(j\omega)$ is a frequency dependent measure of the distance between $g_1(j\omega)$ and $-1/g_2(j\omega)$ in terms of the min-gap.

We remark that the same information can be obtained by examining a plot of $\left| \frac{1}{g_2(j\omega)} + g_1(j\omega) \right|$. The point here is merely that $\varphi_3(j\omega)$ appropriately generalizes this kind of analysis to the general MIMO ($p \neq m$) case as we next demonstrate.

Example 4.5: Consider again example 4.3 where we reveal still another possible internal loop breaking configuration as illustrated in figure 4.7. Here H(s) is the 2×1 transfer function

$$H(s) = \begin{bmatrix} \dfrac{2s + 1}{s[(s + .025)^2 + .01]} \\ \dfrac{1}{2s + 1} \end{bmatrix}$$

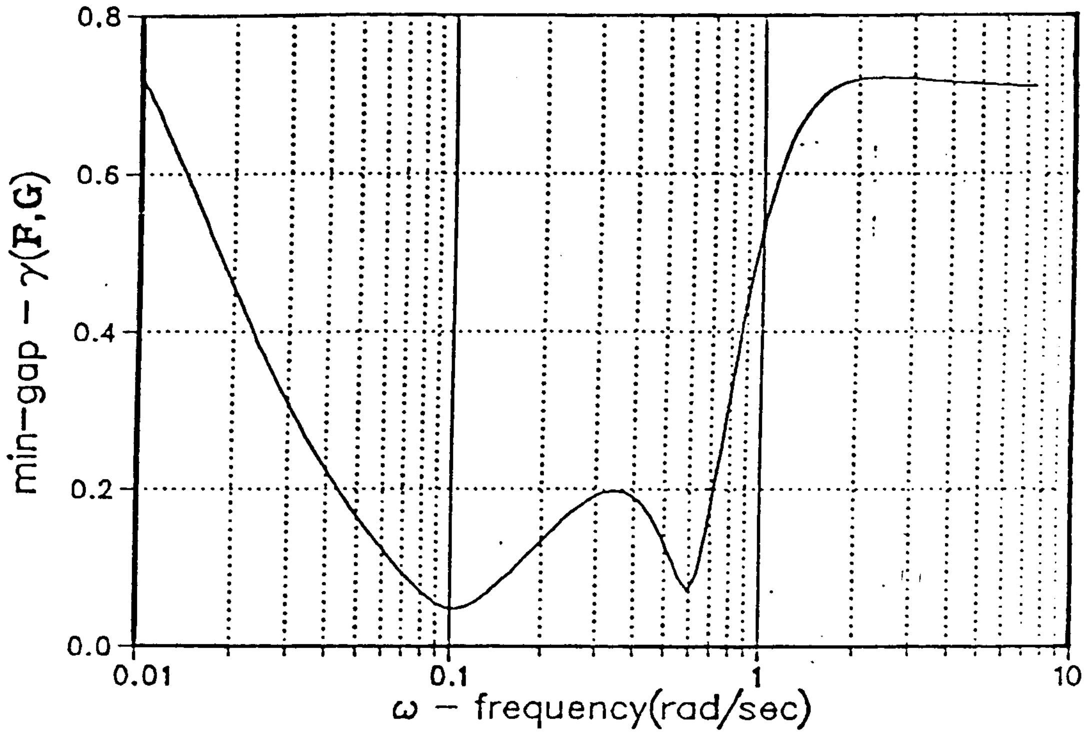


Figure 4.5: $\varphi_3(j\omega)$ vs. ω for example 4.4.

and

$$F(s) = \left[\frac{[(s + .025)^2 + .01]}{s + 1}, -4\right].$$

Again, a hidden mode is revealed. We plot in figure 4.8 the curve $\gamma(\mathrm{H}(s), \mathrm{F}(s))$ for $s = j\omega$ with $\omega \in [0.1, 8.]$.

Finally, we state a caveat. In the case that a hidden mode exists for some dynamic feedback configuration; e.g., $f(s) = f'(s)/a(s)$ and $g(s) = a(s)g'(s)$, then the function,

$$\left|\frac{1}{f(s)} + g(s)\right| = |a(s)| \left|\frac{1}{f'(s)} + g'(s)\right|,$$

amounts to a scaling which can degrade the numerical conditioning of the computational problem. Use of the min-gap function *does*

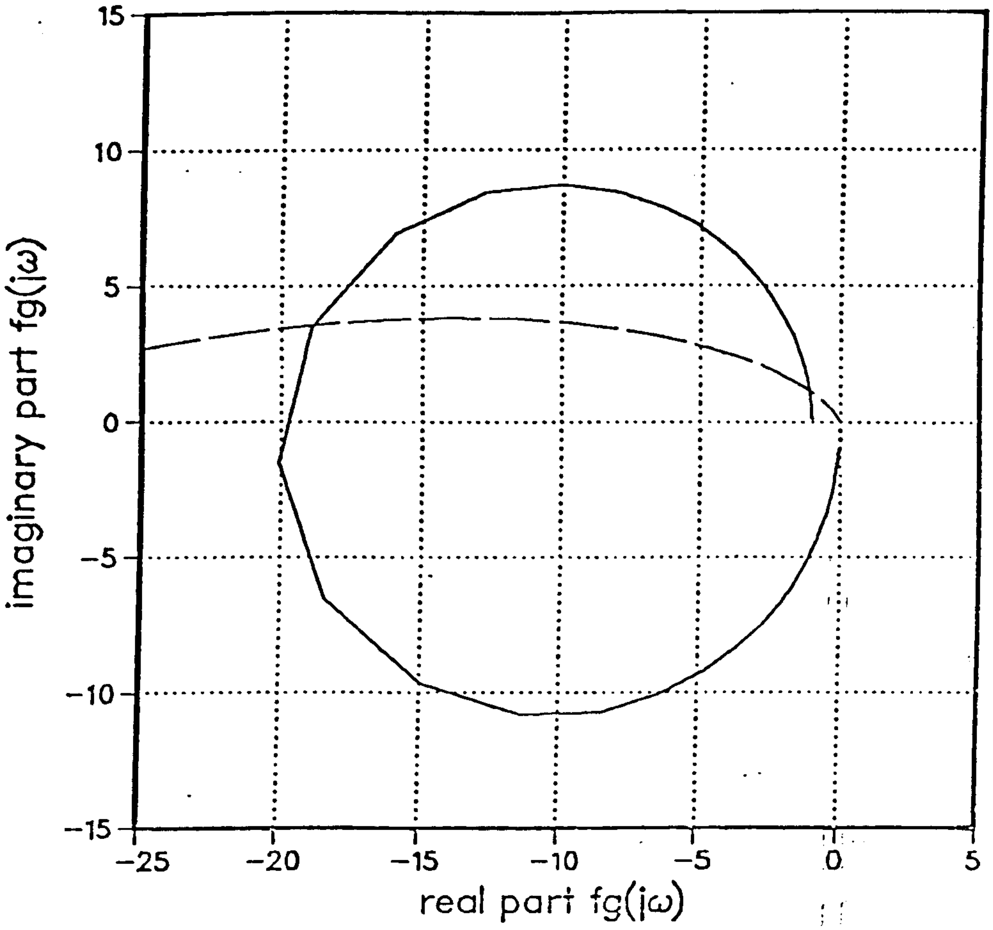


Figure 4.6: A Nyquist Plot for example 4.4.

not mitigate this problem. Indeed,

$$\begin{bmatrix} 1/a(s) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & f'(s) \\ g'(s) & -1 \end{bmatrix} \begin{bmatrix} a(s) & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & f(s) \\ g(s) & -1 \end{bmatrix},$$

is again a scaling of the feedback equations which can degrade the numerical conditioning of the problem of computing the principle angles via singular value analysis in the neighborhood of a root of $a(s)$.

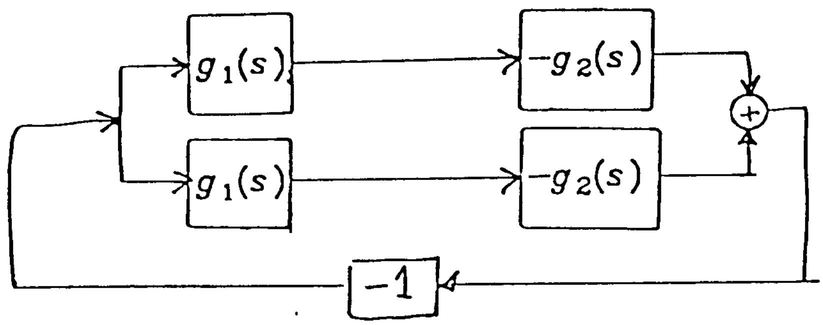


Figure 4.7: The Loop Breaking Configuration for Example 4.5.

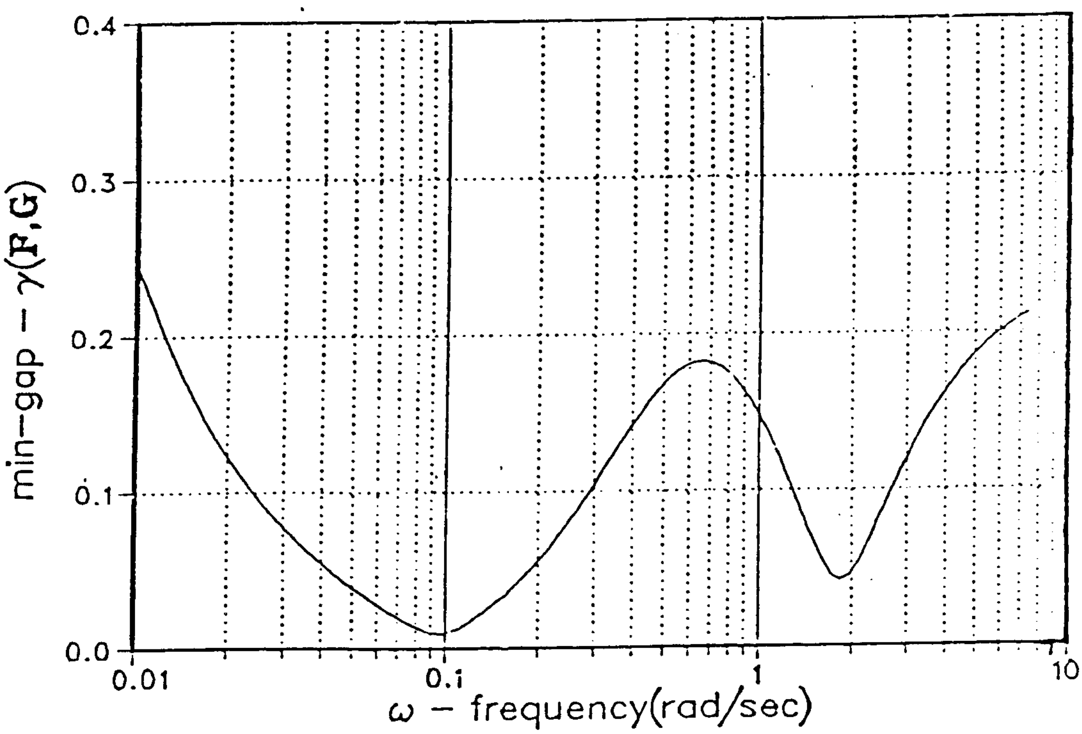


Figure 4.8: $\varphi_3(j\omega)$ vs. ω for example 4.5.

5. Dynamic Weak Coupling from a Geometric Viewpoint

In this section we develop a new notion of dynamic weak coupling useful for the design of decentralized control. The method is presented as an alternative to methods based on block diagonal dominance (BDD) (and its various extensions) as discussed in chapter 3. Our approach is based on the viewpoint (obtained from algebraic geometry) of an abstract Nyquist criterion for MIMO feedback [BR4] which was discussed in detail in sections 2.3.2 and 4.1. This provides a more general setting for testing for weak coupling than is provided by BDD methods by permitting more general partitions to be considered.

As discussed in chapter 3 and in the introduction to this dissertation this research was motivated as an extension of the well known Inverse Nyquist Array (INA) methods for MIMO system design popularized by Rosenbrock [RO1-2]. These methods (which were not originally proposed as tools for design of decentralized control) can result in a decentralized control structure when the open loop plant has transfer function which is diagonally dominant. Moreover, in this case, the graphical procedure provides directly a method for estimating the contribution of each local controller to the overall stability properties (and thus performance) of the decentralized scheme. In the second section of this chapter we discuss the utility of the geometric stability margin and the abstract Nyquist criterion for providing an approach to this problem for more general decentralized schemes (cf. section 3.4) which are weakly coupled in a technical sense which we define in the sequel.

5.1. A Geometric Measure of Dynamic Weak Coupling

The interpretation of diagonal dominance as an indication of weak coupling was first used implicitly by Rosenbrock and his coworkers [RO1-2]. The notion of weak coupling here is mostly intuitive since clearly a diagonally dominant transfer function matrix is "close" to a strictly diagonal (or decoupled) transfer function.

Another notion of weak coupling (originally proposed by Aplevich [AP1]) was shown to have some weak connection with diagonal dominance by Hutcheson [HU3]. A measure of dynamic system interaction was constructed using a geometric argument applied to a time domain system model. Essentially, the *interaction index* [AP1] is the angle between the system impulse response, $G(t)$, and an orthogonal projection of $G(t)$ in the space of square matrix valued functions with square integrable elements onto the subspace of completely decoupled systems; viz., $G^D(t)$, where,

$$\left[G^D(t)\right]_{ij} = \begin{cases} g_{ii}(t), & \text{for } i \neq j \\ 0, & \text{for } i \neq j \end{cases} .$$

It is shown in [HU3] that this angle can be computed as

$$\sin^2\vartheta = \frac{\|G(t) - G^D(t)\|^2}{\|G(t)\|^2} = \frac{\sum_{i,j \neq i} \|g_{ij}(t)\|^2}{\sum_{i,j} \|g_{ij}(t)\|^2},$$

where the norms are the natural L_2 norms of the impulse response functions. Then Hutcheson shows that given $G(t)$ is a stable system with proper transfer function matrix which is diagonally dominant for all $s = j\omega$ with $\omega \in [0, \infty)$ then the interaction index will have an angle $\vartheta < \frac{\pi}{4}$.

For our purposes this result will not be very important for several reasons. First, the connection with diagonal dominance (the extension to BDD is straightforward) is rather weak since the converse relation does not hold in general. The definition of the interaction index is rather special requiring stable system responses for all transfer functions. Most importantly, the interaction index does not support any readily apparent analysis of the relative stability properties of the system subject to compensation.

In section 4.1 we considered a generalized notion of a Nyquist contour suggested by algebraic geometry as a curve on the Grassman manifold. We considered a special class of transfer functions which are *degenerate* with respect to this picture in the sense that there exists some abstract critical point (a Schubert hypersurface in Grass(m,p+m)) which strictly contains the relevant Nyquist contour. Furthermore, we demonstrated by example (cf. Example 4.2) that a subclass of transfer functions with the property that the Schubert hypersurface which contains Γ_G intersects the space of inputs, $\mathbf{U}$, nontrivially, provides a natural decomposition of the space of inputs and outputs as in (4.14). In this section the significance of this (rather special) class of transfer functions for decentralized control will be discussed. Then by employing the topology of the gap metric we provide a mechanism for approximating a broader class of transfer functions by ones which have this special property. We show how this can be interpreted as *weak coupling* in the spirit of BDD methods.

Again, as discussed in the preamble to chapter 3 and section 2.3.2 it will be necessary for limitations of practical application based on finite bandwidth data to limit ourselves to transfer functions which are *strictly proper*. This guarantees that the weak coupling criterion (formerly BDD) need be satisfied only on the closed interval of the $j\omega$ axis, say $\omega \in [0, \omega_0]$, rather than on all of the contour D. Thus we start with the following assumption.

Assumption: Let $G(s) \in \mathbf{R}_{sp}^{p \times m}(s)$; i.e., $G(s)$ is *strictly proper*.

We consider initially, the special class of degenerate transfer functions $G(s)$ with the following property.

Property 5.1: $G(s)$ is a degenerate transfer function with the property that its abstract Nyquist contour Γ_G is contained in some Schubert hypersurface $\sigma(\mathbf{X}) \in \text{Grass}(m, p+m)$ where $\mathbf{X} \in \text{Grass}(p, p+m)$ intersects the m dimensional subspace $\mathbf{U}$ nontrivially; i.e.,

$$\dim\left[\mathbf{X} \cap \mathbf{U}\right] > 0 .$$

We note from the discussion leading to theorem 4.4 that the other class of degenerate transfer functions are rank deficient over the field of rational functions. Such models are ill-posed and can be reformulated (by dropping redundancy in the inputs and outputs) without consequence to the question of model decomposition as we next consider.

Theorem 5.1: Let $G(s) \in \mathbf{R}^{p \times m}(s)$ have property 5.1. Then there exists a change of basis in the space of inputs $\mathbf{U}$, $u = Su$, and in the space of outputs, $\mathbf{Y}$, $y = Ty$, such that in the new basis for $\mathbf{U} \oplus \mathbf{Y}$ the resulting transfer function $G^{\Delta}(s) = TG(s)S^{-1}$ has the form

$$G^{\Delta}(s) = \begin{bmatrix} G_1(s) & * \\ 0 & G_2(s) \end{bmatrix}$$

(or the form

$$G^{\Delta}(s) = \begin{bmatrix} G_1(s) & 0 \\ * & G_2(s) \end{bmatrix}), \dagger$$

where $G_1(s)$ and $G_2(s)$ are both full rank over the field of rational functions.

This implies a partitioning of the space

$$\mathbf{U} \oplus \mathbf{Y} = (\mathbf{U}_1 \oplus \mathbf{Y}_1) \oplus (\mathbf{U}_2 \oplus \mathbf{Y}_2)$$

as suggested in (4.14).

Proof: By property 5.1 we see that the Nyquist contour Γ_G is contained in a Schubert hypersurface $\sigma(\mathbf{X})$ in Grass(m,p+m) which means that for all $s = j\omega$

$$\ker[\mathrm{G}(s), -I_p] \subseteq \sigma(\mathbf{X}).$$

We represent $\mathbf{X}$ as $\ker[X_1, X_2]$ and by the discussion in section 4.1 the $m \times m$ matrix X_1 is rank deficient. Thus solutions to

$$X_1 u(s) = -X_2 y(s) \tag{5.1}$$

are of interest. Corresponding to solutions $u(s)$ in $\ker X_1$ which is of dimension say $m_2 < m$ we have solutions $y(s)$ in $\ker X_2$ which has dimension say $p_2 < p$. Then choose a change of basis, S, in $\mathbf{U}$ such that $\tilde{u} = Su$ has the form

$$\tilde{u} = \begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix}$$

where

$$\begin{bmatrix} 0 \\ \tilde{u}_2 \end{bmatrix} \in \ker X_1 S$$

† Here * stands for "don't care".

and $\tilde{y} = Ty$ has the form

$$\tilde{y} = \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix}$$

where

$$\begin{bmatrix} 0 \\ \tilde{y}_2 \end{bmatrix} \in \ker X_2 S .$$

■

Of course it is possible that $G^{\Delta}(s)$, as constructed above, may have a diagonal submatrix $G_1^{\Delta}(s)$ whose Nyquist contour (now constructed in a lower dimensional Grassman manifold, $\mathrm{Grass}(m_1, m_1+p_1)$) is also degenerate with property 5.1. In such a case we can proceed to apply the theorem again to further decompose $\mathrm{U} \oplus \mathrm{Y}$.

Next, we consider the significance of such a decomposition for decentralized feedback. We start by assuming this procedure of decomposition has already been carried out so that we start with $G(s)$ in a block upper (or lower) triangular form. We consider *decentralized feedback*

$$F = \text{block diag}\{F_1, \ldots, F_k\}$$

where F is partitioned conformally with $G(s)$ and k is consequently the number of submatrices of $G(s)$ appearing on the diagonal.

Theorem 5.2: With $G(s)$ block upper (resp. lower) triangular and F block diagonal as above we construct the abstract Nyquist contour Γ_G for $G(s)$ as usual and the critical point $\sigma(\mathrm{F})$ in $\mathrm{Grass}(m, p+m)$. Let $G_{ii} \in \mathrm{R}^{p_i \times m_i}(s)$ and $F_i \in \mathrm{R}^{m_i \times p_i}$ and construct for each $i=1,\ldots,k$

$$\ker\left[G_{ii}(s),\ -I_{p_i}\right]: D \to \Gamma_i$$

and

$$\mathbf{F}_i = \ker\left[I_{m_i},\ F_i\right]$$

with appropriate Schubert hypersurface $\sigma(\mathbf{F}_i)$ on the manifold Grass(m_i, p_i+m_i). Then

$$N(\Gamma_G;\ \sigma(\mathbf{F})) = \sum_{i=1}^{k} N(\Gamma_i;\ \sigma(\mathbf{F}_i)) \tag{5.2}$$

where the encirclements on the left hand side are counted on Grass$(m, p+m)$ and those on the right hand side of (5.2) are counted respectively on Grass(m_i, p_i+m_i).

Proof: Assume, the Nyquist contour is not contained in $\sigma(\mathbf{F})$ on Grass$(m, p+m)$ so we can determine $N(\Gamma_G;\ \sigma(\mathbf{F}))$. For the purposes of this proof we recognize that $N(\Gamma_G;\ \sigma(\mathbf{F}))$ on Grass$(m, p+m)$ is the same as $N(\beta(s);\ 0)$ in the complex plane where $\beta(s)$ is the rational function

$$\beta(s) = \det\begin{bmatrix} I_m & F \\ G(s) & -I_p \end{bmatrix}. \tag{5.3}$$

Then the result follows by showing that under the conditions of the theorem

$$\det\begin{bmatrix} I_m & F \\ G(s) & -I_p \end{bmatrix} = \prod_{i=1}^{k} \det\begin{bmatrix} I_{m_i} & F_i \\ G_{ii}(s) & -I_{p_i} \end{bmatrix}. \tag{5.4}$$

We show this by applying a series of similarity transformations to

$$\det\begin{bmatrix} I_m & F \\ G(s) & -I_p \end{bmatrix}$$

involving permutations which essentially reorder the basis of $U \oplus Y$ to the form

$$\begin{bmatrix} u_1 \\ y_1 \\ u_2 \\ y_2 \\ \bullet \\ \bullet \\ \bullet \\ u_k \\ y_k \end{bmatrix}$$

where the pairs $\begin{bmatrix} u_i \\ y_i \end{bmatrix}$ are the inputs and outputs for the diagonal blocks of $G(s)$.

■

Of course, transfer functions with property 5.1 are indeed special. From theorem 4.5 we recall that degeneracy is non-generic in the space of all proper transfer functions. Although such obvious decompositions are rather special there are other, less restrictive, conditions which are also sufficient for equation (5.2) to hold. For example, if $F = \text{block diag}\{F_1, \ldots, F_k\}$ is chosen such that $I_m + FG(s)$ is BDD on D then (5.2) will hold.

Condition (5.2) is a *weak coupling* condition in the spirit of Rosenbrock's INA methods which allows decentralization of the design process. We next construct a *new* sufficient condition based in an essential way on the topology of the Grassman manifold. Here we exploit the gap metric and minimum gap function as discussed in section 4.2.

From the point of view of stability, represented by the abstract Nyquist criterion, it is natural to consider a transfer function G(s) as being well approximated by some $G^{\Delta}(s)$ which satisfies property 5.1 if their respective Nyquist contours are "close" in the sense of the gap metric; i.e.,

$$\delta\{\Gamma_G\,,\,\Gamma_{G^{\Delta}}\} < \varepsilon(\omega)\,. \tag{5.5}$$

With respect to decentralized control the required bound, $\varepsilon(\omega)$, for the weak coupling condition (5.2) to hold is given by the following.

Theorem 5.3 (Main weak coupling result): Let $G(s)\in\mathbf{R}_{sp}^{p\times m}(s)$ with associated abstract Nyquist contour Γ_G. Let the feedback $F\in\mathbf{R}^{m\times p}$ be a block diagonal matrix

$$F \;=\; \text{block diag}\,\{F_1, \ldots, F_k\}$$

where $F_i\in\mathbf{R}^{m_i\times p_i}$, $\sum_{i=1}^{k} m_i = m$, and $\sum_{i=1}^{k} p_i = p$. Let $G^{\Delta}(s)$ be the block upper (resp. lower) triangular part of $G(s)$ with respect to a partitioning which is conformal with F. For each $i=1,\ldots,k$ define abstract Nyquist contours for the diagonal blocks of $G(s)$

$$\mathbf{G}_i(s) \;=\; \ker\left[G_{ii}(s)\,,\,-I_{p_i}\right] : D \to \Gamma_i\,,$$

each living in its appropriate submanifold $\Gamma_i\subseteq\text{Grass}\,(m_i, p_i+m_i)$. And for each local feedback (diagonal block of F) define an appropriate abstract critical point on each submanifold as $\sigma(\mathbf{F}_i)\subseteq\text{Grass}\,(m_i, p_i+m_i)$ where $\mathbf{F}_i \;=\; \ker[I_{m_i}, F_i]$. Then if the condition

$$\min_{i\in[1,k]} \gamma(\mathbf{G}_i(s)\,,\,\mathbf{F}_i) > \delta(\mathbf{G}(s),\,\mathbf{G}^{\Delta}(s)) \tag{5.6}$$

is satisfied for all s on D then,

$$N(\Gamma_G\,;\,\sigma(\mathbf{F})) \;=\; \sum_{i=1}^{k} N(\Gamma_i\,;\,\sigma(\mathbf{F}_i))\,. \tag{5.7}$$

Proof: First consider the following lemma.

Lemma 5.4: Let $\mathbf{N}_1, \mathbf{N}_2 \subseteq \mathbf{C}^n$ be subspaces of dimension p, and $\mathbf{N}_3 \subseteq \mathbf{C}^n$ be of dimension m. So that $\mathbf{N}_1, \mathbf{N}_2 \in$ Grass(p,n), and $\mathbf{N}_3 \in$ Grass$(n-p,n)$. If

$$\gamma(\mathbf{N}_1, \mathbf{N}_3) > \delta(\mathbf{N}_1, \mathbf{N}_2)$$

then

$$\mathbf{N}_2 \cap \mathbf{N}_3 = \{0\}.$$

Proof of Lemma: From section 4.2 $\gamma(\mathbf{N}_1, \mathbf{N}_3)$ gives a measure of the distance (in the gap metric) between $\mathbf{N}_1 \in$Grass(p,n) and the Schubert hypersurface $\sigma(\mathbf{N}_3) \subseteq$Grass$(p,n)$. The result of the lemma is then obvious by the gap-metric topology of Grass(p,n).

∎

Next define the matrix

$$G_\varepsilon(s) = G^\Delta(s) + \varepsilon[G(s) - G^\Delta(s)]$$

for $0 \le \varepsilon \le 1$. The associated subspaces

$$\mathbf{G}_\varepsilon(s) = \ker[G_\varepsilon(s), -I_p]$$

will satisfy

$$\delta(\mathbf{G}(s), \mathbf{G}^\Delta(s)) > \delta(\mathbf{G}_\varepsilon(s), \mathbf{G}^\Delta(s)) \tag{5.8}$$

for $0 \le \varepsilon \le 1$ by the gap-metric topology of Grass$(p,m+p)$. Also note that

$$\min_{i=1,\ldots,k} \gamma(\mathbf{G}_i(s), \mathbf{F}_i) = \gamma(\mathbf{G}^\Delta(s), \mathbf{F})$$

where $\mathbf{G}^\Delta(s) = \ker[G^\Delta(s), -I_p]$. Thus clearly (5.6) guarantees under the assumptions given that if $G^{\Delta}(s) \cap F = \{0\}$ for s on D then $G(s) \cap F = \{0\}$ for all s on D.

Let

$$A(\varepsilon,s) = \begin{bmatrix} G_{\varepsilon}(s) & -I_p \\ I_m & F \end{bmatrix}$$

and

$$A_i(s) = \begin{bmatrix} G_{ii}(s) & -I_{p_i} \\ I_{m_i} & F_i \end{bmatrix}.$$

Then let

$$\beta(\varepsilon, s) = \frac{\det A(\varepsilon, s)}{\prod_{i=1}^{k} \det A_i(s)}$$

map D into Γ_β, a closed curve in the complex plane. Now Γ_β does not encircle the origin since otherwise there must exist some s on D and $0 \le \varepsilon \le 1$ with $\beta(\varepsilon,s) = 0$. However this means that $G_\varepsilon(s)$ intersects F in some nontrivial way -- a situation which is precluded by (5.6) using the lemma and (5.8). Finally application of the principle of the argument to $\beta(\varepsilon,s)$ gives the result.

∎

We remark that theorem 5.3 provides a new notion of dynamic weak coupling in the spirit of Rosenbrock's INA method. The method can effectively deal with more general partitions for transfer functions than BDD methods obtained in [BE2]. Recall from the definition of BDD (cf. section 3.2.1) that no partitions for which any diagonal block of $G(s)$ is ***not*** a square matrix can satisfy BDD. Of course, it is always possible to apply a BDD test to a matrix return difference, say $I_m + FG$, representing some loop breaking at the inputs. Thus any partitioning of the input space **U** will lead to a partitioning of $I_m + FG$ with square submatrices on the main diagonal. This approach can, however, be disadvantageous for design of the individual "local" feedback compensators. Indeed, several researchers [NW1-4] have found the alternate form $F^{-1} + G(s)$ more convenient for design (Rosenbrock prefers the inverse formulation, $F + G^{-1}(s)$). For example, Nwokah [NW4] shows that if $G(s)$ is a composite H-matrix and F is *any* block diagonal (decentralized) feedback then $F^{-1} + G(s)$ is also a composite H-matrix. Such formulation places an artificial technical limitation on the number of inputs and outputs ($p = m$) which we avoid in the formulation of theorem 5.3.

The condition for dynamic weak coupling (5.6) is clearly quite different from any BDD condition. It does not seem possible to suggest that weak coupling in the sense of (5.6) can imply or be implied by weak coupling in the sense of BDD. Consider the following example:

Example 5.1: We reconsider the design example of [BE2, chapter 4]. Here

$$G(s) = \left[\begin{array}{cc|c|cc} \frac{15}{s-5} & \frac{-5.25}{s+3.5} & \bullet & 0 & \frac{2.5}{s+1} \\ \frac{5.25}{s+3} & \frac{21}{s-6} & \bullet & \frac{2.5}{s+1} & 0 \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ 0 & \frac{5}{s+2} & \bullet & \frac{18}{s-6} & \frac{-4.5}{s+3} \\ \frac{5}{s+2} & 0 & \bullet & \frac{7}{s+4} & \frac{17.5}{s-5} \end{array}\right]$$

was shown to be BDD for $s = j\omega$ with $\omega \in [0, 25 \text{ rad./sec.}]$ with respect to the 2x2 partitioning shown above with each $G_{ij}(s) \in \mathbb{R}^{2\times2}(s)$.

To test for weak coupling in the sense of (5.6) we plot in figure 5.1 two curves as follows:

(i) $\gamma(G^{\Delta}(j\omega), Y)$

(ii) $\gamma(G^{\Delta}(j\omega), Y) - \delta(G(j\omega), G^{\Delta}(j\omega))$.

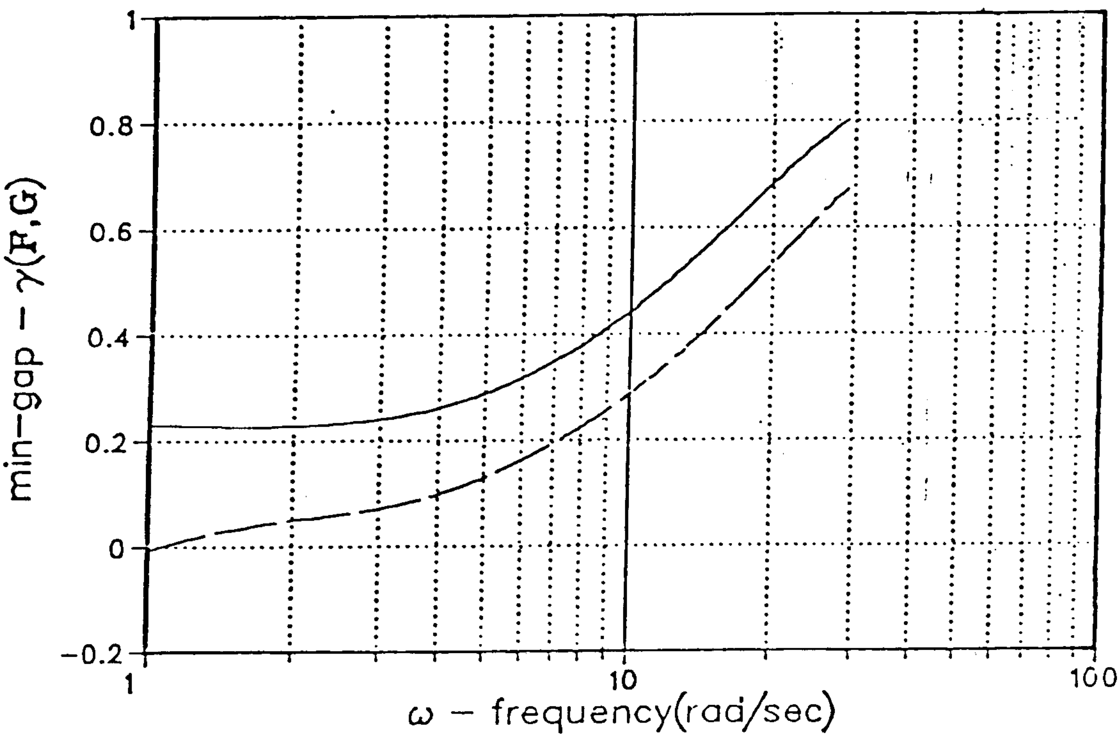


Figure 5.1: Illustrating weak coupling in Example 5.1

Here

$$G^{\Delta}(s) = \begin{bmatrix} G_{11}(s) & 0 \\ 0 & G_{22}(s) \end{bmatrix}$$

with

$$G_{11}(s) = \begin{bmatrix} \frac{15}{s-5} & \frac{-5.25}{s+3.5} \\ \frac{5.25}{s+3} & \frac{21}{s-6} \end{bmatrix}$$

and

$$G_{22}(s) = \begin{bmatrix} \frac{18}{s-6} & \frac{-4.5}{s+3} \\ \frac{7}{s+4} & \frac{17.5}{s-5} \end{bmatrix}.$$

It is seen from Fig. 5.1 that $G(s)$ is open-loop weakly coupled in the interval $1 \le \omega \le 25$. Note that to test for "open-loop" weak coupling using (5.6) we have taken $F = 0$ so that

$$\ker[I_m\,,\,0] = \mathbf{Y}.$$

5.2. Local Stability Margin Analysis for Decentralized Systems with Weak Coupling

The significant aspect of theorem 5.3 for our purposes is that it generalizes (although somewhat abstractly) the notion of a "broad" or "fuzzy" Nyquist locus for the individual, local feedback loops (obtained in [RO1] by using a result of Ostrowski) to the case of partitioned G(s). This leads to a sequential design approach for the local feedback compensators.

In particular, condition (5.6) can be tested (at least conceptually) by appending for each $i=1,...,k$ and for each s on D, a neighborhood of radius $\delta(\mathbf{G}(s), \mathbf{G}^{\Delta}(s))$ about the corresponding point on each abstract Nyquist contour, Γ_i in its appropriate Grassman space. These neighborhoods then sweep out a subset of each Grassman manifold which contains the Γ_i. Then if each subset avoids its corresponding abstract critical point $\sigma(\mathbf{F}_i) \subseteq \mathrm{Grass}(m_i, p_i+m_i)$ then (5.6) is satisfied and conversely. We provide, by example, a simple illustration of how the local contribution to the system stability margin (in the sense of chapter 4) can be estimated.

Example 5.2: Consider the problem of example 5.1 again. In [BE2, chapter 4] we showed that the 4×4 transfer function matrix is BDD with respect to the given partition and has 2×2 diagonal blocks which are each diagonally dominant. Thus we applied Nyquist array methods to choose a decentralized compensator of the form

$$F = \text{block diag}\left\{3,3,3,3\right\}$$

which stabilizes the system. However, the procedure in [BE2] does not suggest a method for estimating the stability margins.

Here we examine the local controller associated with the first two inputs and outputs. In figure 5.2 we plot two curves as follows:

(i) $\gamma(\Gamma_1, \sigma(F_1))$

(ii) $\gamma(\Gamma_1, \sigma(F_1)) - \delta(G, G^{\Delta})$

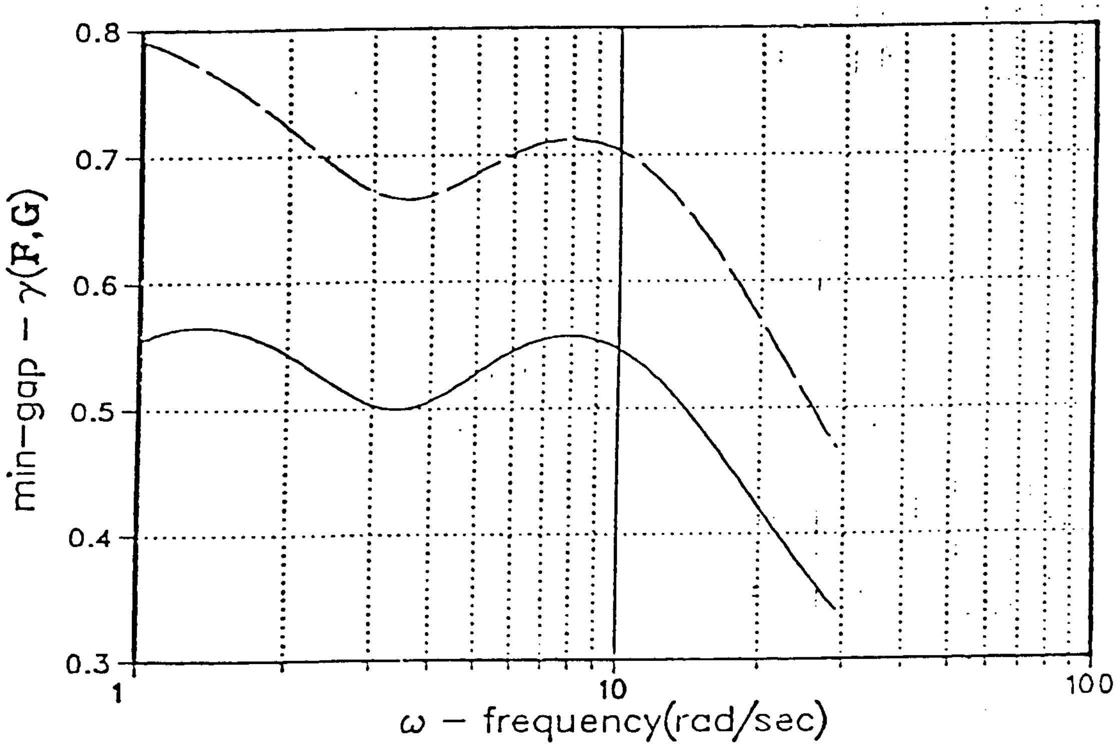


Figure 5.2: Illustrating stability margin estimation for Example 5.2

where Γ_1 and G^{Δ} are given in example 5.1 and $F_1 = 3I_2$. From (4.29) and Fig. 5.2 we get as an estimate of the geometric stability margin for these loops $g_{sm} \geq 0.33$.

6. CONCLUSIONS AND DIRECTIONS

In this dissertation we have considered the development of a fresh line of analysis for the general problem of defining stability margins for feedback control based on the frequency response of the loop components. We have employed a viewpoint, suggested by algebraic geometry, of an abstract Nyquist contour living on a complex Grassman manifold. By exploiting the natural topology of this manifold we have developed a way to compare relative stability properties for various systems and subsystems in feedback configuration which applies equally to SISO and MIMO loops.

Significantly, this approach avoids formulation of a (possibly matrix-valued) return-difference for the feedback system. Instead, we compute the distance between an abstract critical point and an appropriate Nyquist contour directly in terms of a gap-metric. The construction is completely general allowing for non-square transfer functions for individual loop components.

The advantage of the gap-metric approach (in contrast to the Plücker metric) for this application is that it allows computationally efficient algorithms to be developed. Such algorithms have excellent numerical stability properties and can be codified using standard modules such as are available in LINPACK [DO1]. The computational procedure is based on the principal angles between a pair of subspaces of a finite dimensional vector space. Using standard singular value analysis one can readily compute the principle vectors. We believe that the use of this principle vector analysis will play an important role in developing iterative design methods based on the method of feasible directions. Such methods are becoming increasingly feasible for practical engineering design with the development of powerful interactive software systems such as DELIGHT [NY1] and supported by fast minicomputers and high quality graphic displays.

REFERENCES

[AN1]B.D.O. Anderson and J.B. Moore, *Linear Optimal Control* , Prentice Hall, 1971.

[AN2]B.D.O.Anderson and D.Clements, "Algebraic Characterization of Fixed Modes in Decentralized Control", *Automatica* , Vol.17, No.5, 1981, pp. 703-712.

[AP1]J. D. Aplevich, "A Geometrically- Motivated Scalar Measure of System Interaction," *Proc. IFAC Symposium on Large Scale Systems Theory and Applications* , Udine, Italy, 1976.

[AR1]M. Araki, "Input-Output Stability of Componsite Feedback Systems," *IEEE Trans. on Auto Control.* Vol. AC-21, April 1976.

[BE1]W. H. Bennett and J. S. Baras, "Block Diagonal Dominance and the Design of Decentralized Compensators", Univ. of Md. Tech. Res. Rep. SRR-83-2, 1983.

[BE2]W.H. Bennett, *Block Diagonal Dominance and Design of Decentralized Feedback Compensators* , M.S. Thesis, Univ. of Maryland, 1979.

[BE3]W. H. Bennett and J. S. Baras, "Decentralized Stabilization of Large Scale Systems by Frequency Domain Methods", *Proc. 17th Conf. on Info. Sci. & Systems* , Johns Hopkins Univ., Baltimore, Md., 1983.

[BJ1]A. Björck and G. H. Golub, "Numerical Methods for Computing Angles Between Linear Subspaces," *Math of Comp.* , Vol. 27, No. 123, July 1973, pp. 579-594.

[BR1]J. L. Brenner, "Gerschgorin Theorems, Regularity Theorems, and Bounds for Determinants of Partitioned Matrices", *SIAM J. Appl. Math.* , Vol. 19, No. 2, September 1970, pp. 443-450.

[BR2]J. L. Brenner, "Gerschgorin Theorems, Regularity Theorems, and Bounds for Determinants of Partitioned Matrices - Part II", *Pac. J. Math.* , Vol. 39, No. 1, 1971.

[BR3]J. L. Brenner, "New Root-Location Theorems for Partitioned Matrices", *SIAM J. Appl. Math.* , Vol. 16, No. 4, Sept. 1968.

[BR4]R.W.Brockett and C.I.Byrnes, "Multivariable Nyquist Criteria, Root Loci, and Pole Placement: A Geometric Viewpoint", *IEEE Trans. on Auto. Control* , Vol. AC-26, No.1, Feb.1981, pp. 271-283.

[BY1]C.Byrnes, M.Hazewinkel, C.Martin,& Y.Rouchaleau, "Introduction to Geometrical Methods for the Theory of Linear Systems", in *Geometric Methods for the Theory of Linear Systems* , D.Reichel, Dordrect, 1980.

[CO1]P. A. Cook, "On the Stability of Interconnected Systems," *Int. J. Control* , Vol. 20, No. 3, 1974, pp.407-415.

[CO2]J.P.Corfmat and A.S.Morse, "Decentralized Control of Linear Multivariable Systems", *Automatica* , Vol.12, 1976, pp.479-495.

[CR1]J.B.Cruz,Jr., J.S.Freudenberg, & D.P.Looze, "A Relationship Between Sensitivity and Stability of Multivariable Feedback Systems", *IEEE Trans. on Auto. Control , Vol.AC-26, No. 1, Feb. 1981, pp. 66-74.*

[DA1]G. Dahlquist and A. Bjorck, *Numerical Methods* , Prentice Hall, Inc., 1974.

[DE1]C. A. Desoer and M. Vidyassagar, *Feedback Systems: Input-Output Properties* , Academic Press, 1975.

[DE2]C.A. Desoer and Y.T. Wang, "On the Generalized Nyquist Criterion", *IEEE Trans. on Auto. Control* , Vol. AC-25, No.2, April 1980, pp.187-196.

[DO1]J.J.Dongarra, C.B.Moler, J.R.Bunch, & G.W.Stewart, *LINPACK Users Guide* , SIAM Press, 1979.

[DO2]R. C. Dorf, *Modern Control Systems* , Addison-Wesley, 1974.

[DO3]J.C. Doyle, "Robustness of Multiloop Linear Feedback Systems", *IEEE Conf. on Dec. & Control* , Jan. 1979, San Diego, CA., pp. 12-18.

[DO4]J.C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis", *IEEE Trans. on Auto. Control* , Vol. AC-26, No. 1, pp. 4-16.

[FA1]V. N. Faddeeva, *Computational Methods of Linear Algebra* , Dover, Pub., 1959.

[FE1]D. G. Feingold and R. S. Varga,"Block Diagonally Dominant Matrices and Generalizations of the Gerschgorin Circle Theorems", *Pac. J. Math.* , Vol. 12, 1962, pp. 1241-1250.

[FI1]M. Fiedler and V. Pták, "On Matrices with Non-positive Off-diagonal Elements And Positive Principal Minors", *Czech. Math. J.* , Vol. 87, No. 12, 1962, pp. 382-400.

[FI2]M. Fiedler and V. Pták, "Generalized Norms Of Matrices And The Location Of The Spectrum", *Czech. Math. J.* , Vol. 87, No. 12, 1962, pp. 558-570.

[GA1]F. R. Gantmacher, *The Theory of Matrices, Vol. I & II* , Chelsea Pub., 1959.

[GO1]G. Golub and W. Kahan, "Calculations the Singular Values and Pseudo-Inverse of a Matrix", *J. SIAM Numer Anal.* , *Vol. 2, No. 2, 1965, pp. 205-224.*

[GO2]G. Golub and C. Reinsch, "Singular Value Decomposition and Least Squares Solutions," *Numer. Math.* , Vol. 4, 1970, pp. 403-420.

[HE1]J. W. Helton, "An H^∞ Approach to Control", *Proc. IEEE Control Decision Conf.* , Dec. 1983, San Antonio, Tx., pp. 607-611.

[HE2]R. Hermann and C.F. Martin, "Application of Algebraic Geometry to Linear Sytems Theory", *IEEE Trans. on Auto. Control, Vol. AC-22, 1977, pp. 19-25.*

[HO1]A. S. Householder, *The Theory of Matrices in Numerical Analysis* , Dover Pub., 1975.

[HO2]I. M. Horowitz, *Synthesis of Feedback Systems* , Academic Press, 1963.

[HO3]I. M. Horowitz and U. Shaked, " Superiority of Transfer Function Over State-Variable Methods in Linear Time-Invariant Feedback System Design", *IEEE Trans. on Auto. Control* , Vol. AC-20, No. 1, Feb. 1975, pp. 84-97.

[HU1]Y. S. Hung and D. J. N. Limebeer, "Robust Stability of Additively Perturbed Interconnected Systems", Rep. CUED/F-CAMS/TR-220, Cambridge Univ., 1982.

[HU2]Y. S. Hung and D. J. N. Limebeer, "A Frequency Response Design Technique for Decentralized Control Systems",

[HU3]W.J.Hutcheson, "Decoupling of Multivariable Systems by Function Optimization", *Int. J. Control* , 1979, Vol.30, No.6, pp. 1001-1012.

[JO1]R. L. Johnson, "Gerschgorin Theorems for Partitioned Matrices", *Linear Algebra and Its Applications* , Vol. 4, (1971), pp. 205-220.

[JO2]R. L. Johnson and B. T. Smith,"Calculation of Isolated Gerschgorin Discs", *Numer. Math.* ,Vol. 16, 1970, pp. 22-31.

[KA1]T. Kato, *Perturbation Theory for Linear Operators* , Springer-Verlag, 1976.

[KA2]T. Kailath, *Linear Systems* , Prentice-Hall, Inc., 1980.

[KW1]H.Kwakernaak and R.Sivan, *Linear Optimal Control Systems* , New York:Wiley-Interscience, 1972.

[LA1]E. L. Lasley and A.N. Michel, "Input-Output Stability of Interconnected Systems," *IEEE Trans. on Auto. Control* , Vol. AC-21, Feb. 1976,pp. 84-89.

[LA2]A. Laub, "Efficient Multivariable Frequency Response Computations", *IEEE Control Dec. Conf.* , December 1980.

[LE1]N.A.Lehtomaki, N.R.Sandell, & M.Athans, "Robustness Results in Linear-Quadratic Gaussian Based Multivariable Control Designs", *IEEE Trans. on Auto. Control* , Vol.AC-26, No.1, Feb.1981, pp.75-92.

[LE2]G. G. Leininger, "New Dominance Characteristics for the Multivariable Nyquist Array Method", *Int. J. Control* , 1979, Vol. 30, No. 3, pp. 459-475.

[LE3]W.K. Leung, *Block Sequential Return Difference Method for Large Scale Control Systems* , Ph.D. Dissertation, Univ. of Md., 1983.

[LI1]D. J. N. Limebeer, "The Application of Generalized Diagonal Dominance to Linear System Stability Theory", *Int. J. Control* , Vol. 36, No. 2, 1982, pp. 185-212.

[LI2]D. J. N. Limebeer, "An Inclusion Region for the Eigenvalues of Partitioned Matrices", *Int. J. Control* , Vol. 37, No. 2, 1983, pp. 429-436.

[LI3]D. J. N. Limebeer, "Robust Stability of Interconnected Systems", rep. CUED/F-CAMS/TR-227, Cambridge Univ., 1982.

[MA1]A.G.J. MacFarlane and J.J. Belletrutti,"The Characteristic Locus Design Method", *Automatica* , Vol. 9, No. 5, September 1973, pp. 575-588.

[MA2]A.G.J. MacFarlane, *Frequency-Response Methods in Control Systems* , IEEE Press, 1979.

[MA3]A.G.J. MacFarlane, "A Survey of Some Recent Results in Linear Multivariable Feedback Theory", *Automatica* , Vol. 8, No. 4, July 1972, pp. 455-492.

[ME1]A.I. Mees, *System and Control Letters* , Vol.1, 1981, pp.155.

[MI1]A. N. Michel and R. K. Miller, *Qualitative Analysis of Large Scale Systems* , Academic Press, 1977.

[MO1]Cleve Moler, *MATLAB Users' Guide, CS81-1, University of New Mexico, Albuquerque, NM., 1982.*

[MU1]N. Munro & V. Zakian, "Inversion of Rational Polynomial Matrices", *Electronic Letters* , Sept. 1970, Vol. 6, No. 19, pp. 628.

[NW1]O. D. I. Nwokah, "Estimates for the Inverse of a Matrix and Bounds for Eigenvalues", *Linear Alg. & Appl.* , Vol. 22, 1978, pp. 283-291.

[NW2]O. D. I. Nwokah, "The Convergence and Local Minimality of Bounds for Transfer Functions", *Int. J. Control* , Vol. 30, No. 2, 1979, pp. 195-202.

[NW3]O. D. I. Nwokah, "A Note on the Stability of Multivariable Systems", *Int. J. Control*, Vol. 31, No. 3, 1980, pp. 587-592.

[NW4]O. D. I. Nwokah, "Generalized Nyquist Stability Criterion for Composite Feedback Control Systems", *Large Scale Systems*, No. 2, North-Holland Pub., 1981, pp. 185-190.

[NY1]W.T. Nye, "DELIGHT: An Interactive System for Optimization-Based Engineering Design", Memo. No. UCB/ERL M83/33, University of Calif., Berkeley, May 1983

[PO1]E.Polak and D.Q.Mayne, *On the Solution of Singular Value Inequalities Over a Continuum of Frequencies*, Memorandum Report No. UCB/ERL M80/8, Electronics Research Lab, Univ. of Calif., Berkeley, Jan. 1980.

[PO2]E.Polak and D.Q.Mayne, "An Algorithm for Optimization Problems with Functional Inequality Constraints", *IEEE Trans. on Auto. Control*, Vol. AC-21, No.2, April 1976, pp. 184-193.

[PO3]I. Postlethwaite, and A. G. J. MacFarlane, *A Complex Variable Approach to the Analysis of Linear Multivariable Feedback Systems*, Springer-Verlag, No. 12, 1979.

[PO4]I. Postlethwaite, J. M. Edmonds, and A. G. J. MacFarlane, "Principle Gains and Principal Phases in the Analysis of Linear Multivariable Feedback Systems", *IEEE Trans. on Automatic Control* : Special Issue on Linear Multivariable Control Systems, Vol. AC-20. No. 1, pp. 32-46.

[QI1]L. Qi, "Some Simple Estimates for Singular Values of a Matrix", *Linear Algebra & Applic.* , Vol. 56, 1984, pp. 105-119.

[RO1]H. H. Rosenbrock, *Computer-Aided Control System Design*, Academic Press, Inc., London, 1974.

[RO2]H. H. Rosenbrock, "Design of Multivariable Control Systems Using the Inverse Nyquist Array", *Proc. IEEE* ,Vol. 116, No. 11, Nov. 1969, pp. 1929-1936.

[SA1]M. G. Safanov, A. J. Laub, and G. L. Hartman, "Feedback Properties of Multivariable Systems: The Role and Use of the Return Difference Matrix", *IEEE Trans. on Automatic Control* : Special Issue on Linear Multivariable Control Systems, Vol. AC-26, No.1, 1980.

[SA4]N.R. Sandell, "Robust Stability of Systems with Application to Singular Perturbations", *Automatica* , Vol.15,1981, pp.467-470.

[ST1]G.W. Stewart, "Error and Perturbation Bounds for Subspaces Associated with Certain Eigenvalue Problems," *SIAM Review* , Vol. 15, No. 4, Oct. 1973.

[ST2]G.W. Stewart, *Introduction to Matrix Computations*, Academic Press, 1973.

[VA1]R.S. Varga, "Minimal Gerschgorin Sets", *Pac. J. Math.* , Vol. 15, No. 2, 1965, pp. 719-729.

[VA2]R.S. Varga, "Minimal Gerschgorin Sets for Partitioned Matrices", *SIAM J. Numer. Anal.* , Vol. 7, No. 4, December 1970, pp. 493-507.

[VI1]M. Vidyasgar and N. Viswanadham, "Algebraic Characterization of Decentralized Fixed Modes and Pole Placement", *Proc. IEEE Conf. on Dec. and Control* , Dec. 1982, Orlando, Fld., pp. 501-505.

[WA1]S.H. Wang and E.J. Davison, "On the Stabilization of Decentralized Control Systems", *IEEE Trans. on Auto. Control* , Vol.AC-18, No.5, 1973, pp. 473-478.

[WO1]W.A. Wolovich, *Linear Multivariable Systems* , Springer-Verlag, 1974.

[YO1]D.C. Youla, H.A. Jabr, and J.J. Bongiorno, Jr., "Modern-Wiener-Hopf Design of Optimal Controller - Part II: The Multivariable Case," *IEEE Trans onAuto. Control* , June 1976.

[ZA1]L.A. Zadeh and C.A. Desoer, *Linear System Theory: The State Space Approach* , McGraw-Hill, 1963.

[ZA2]G. Zames, "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses", *IEEE Trans. on Auto. Control* , Vol.AC-26, No.2, April 1981, pp. 301-320.

[ZA3]G. Zames and D. Bensoussan, "Multivariable Sensitivity Reduction and Decentralized Control", *Proc. JACC*, 1983.

APPENDIX A

Geometric Interpretation of the Singular Value Decomposition

In [BJ1] Björck and Golub describe a method for computing the principal angles between a pair of subspaces of a linear, finite dimensional vector space which exploits the properties of the singular value decomposition (SVD) for matrices in an essential way. In this appendix we discuss the geometric interpretation of the SVD which is central to the computational results in [BJ1]. These results are used in section 4.2.6 to provide a computational procedure (based on numerically stable algorithms for computing the SVD [DO1]) for determining the geometric stability margin based on the minimum gap as the sine of the minimum principal angle.

Consider a finite dimensional vector space **X** and two vectors $x, y \in \mathbf{X}$. We usually compute the angle, ϑ, between these vectors in terms of the natural (euclidean) inner product as

$$\cos\vartheta = \frac{<x, y>}{||x||\,||y||}$$

where the norms are

$$||x|| = \sqrt{<x, x>}$$

and for **X** a unitary space

$$<x, y> = x^{*} y$$

which is the natural extension of euclidean vector space over the field of complex numbers.

We next generalize this idea to a pair of subspaces of **X**

Consider a pair of subspaces of $\mathbf{X}$ and to make the ideas concrete we take $\mathbf{X} = \mathbf{C}^n$. Let $\mathbf{F} \subseteq \mathbf{C}^n$ have dimension $p < n$ and $\mathbf{G} \subseteq \mathbf{C}^n$ have dimension $q < n$. Following [BJ1] we define the principal angles between a pair of subspaces.

Definition: The *principal angles*, $\vartheta_k \in [0, \pi/2]$ between $\mathbf{F}$ and $\mathbf{G}$ are given recursively for $k = 1, 2, \ldots, \min(p, q)$ by

$$\cos\vartheta_k = \max_{\substack{x \in \mathbf{F} \\ \|x\|_2 = 1}} \max_{\substack{y \in \mathbf{G} \\ \|y\|_2 = 1}} \langle x, y \rangle = x_k^* y_k \tag{A.1}$$

subject to the constraints

$$x_j^* x = 0 \text{ and } y_j^* y = 0 \tag{A.2}$$

for $j = 1, \ldots, k-1$.

Now let the columns of an $n \times p$ matrix, A, form an orthonormal (unitary) basis for the subspace $\mathbf{F}$ and the columns of an $n \times q$ matrix B form a unitary basis for $\mathbf{G}$. (Thus $A^* A = I_p$ and $B^* B = I_q$.)

Of course, this means that the subspace $\mathbf{F}$ can be characterized as the image (or range) of A taken as a linear map from $\mathbf{C}^p$ to $\mathbf{C}^n$; i.e.,

$$\mathbf{F} = \left\{ Af : \text{for all } f \in \mathbf{C}^p \right\}.$$

And similarly,

$$\mathbf{G} = \left\{ Bg : \text{for all } g \in \mathbf{C}^q \right\}.$$

To compute the minimum principal angle, ϑ_1, we take from (A.1)

$$\cos\vartheta_1 = \max_{\substack{x\in\mathbf{F}\\ \|x\|_2=1}} \max_{\substack{y\in\mathbf{G}\\ \|y\|_2=1}} x^* y \tag{A.3}$$

$$= \max\left\{(Af)^*(Bg) : f \in \mathbf{C}^p,\ g \in \mathbf{C}^q,\ \|Af\| = \|Bg\| = 1\right\}.$$

But since the columns of A form a unitary basis

$$\|Af\|^2 = f^*A^*Af = f^*f = \|f\|^2.$$

Therefore, (A.3) implies

$$\cos\vartheta_1 = \max\left\{f^*A^*Bg : f \in \mathbf{C}^p,\ g \in \mathbf{C}^q,\ \|f\| = \|g\| = 1\right\}.$$

Now the Rayleigh quotient characterization of the singular values [ST2, pp.321] is

$$\sigma_k = \max_{\|w\|=\|z\|=1} w^* C z$$

subject to the constraints

$$w_j^* w = 0 \text{ and } z_j^* z = 0$$

for $j=1, \cdots, k$.

Thus it follows that the *principal angles*, ϑ_k, can be computed from the singular values of the matrix $A^* B$ via

$$\cos\vartheta_k = \sigma_k ,$$

for each $k=1, \ldots, \min(p,q)$.

APPENDIX B

Results on M-Matrices

This appendix includes a summary of the basic properties of M-matrices. These properties were fully illuminated in the seminal paper of Fiedler and Pták [FI1]. We include this material to clarify some subtle technical aspects of some recent extensions of diagonal dominance and INA design methods as discussed in section 3.3 of this dissertation.

The theory of M-matrices is intimately related to the theory of nonnegative matrices. An $n \times n$ matrix A is nonnegative (resp. positive), denoted by $A \geq 0$ (resp. $A > 0$), if for every $i,j = 1, \ldots, n$ $a_{ij} \geq 0$ (resp. $a_{ij} > 0$). This provides a partial ordering on the set of such matrices in the sense that $B \geq A$ means $B - A \geq 0$.

The following results of Perron and Frobenius will be needed.

Definition: A matrix $A \geq 0$ is said to be *reducible* if there exists a permutation, P, such that

$$P\,A\,P^{-1} = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix}$$

otherwise it is called *irreducible*.

Theorem B.1: Let $A \geq 0$ be an $n \times n$ matrix. Then there exists an eigenvalue of A, λ_{PF}, called the *Perron root of A* with the properties:

(i) $\lambda_{PF}(A) \geq 0$

(ii) $\lambda_{PF}(A) = \rho(A)$

where $\rho(A) \triangleq \max |\lambda(A)|$ is the *spectral radius* of A. Moreover, if A is *irreducible* then $\lambda_{PF}(A)$ is distinct, strictly positive, and has a corresponding eigenvector which is strictly positive.

The following useful property of diagonally dominant matrices is related to the spectral properties of such matrices and will be instructive to consider.

Property B.2: If A is an $n \times n$ matrix which is row diagonally dominant then

$$\rho(I_n - A_D^{-1} A) < 1,$$

where A_D is the diagonal of A.

Proof: Choose $\lambda \in \lambda(I - A_D^{-1} A)$. Then there exists a vector $x \neq 0$ such that

$$\lambda x = x - A_D^{-1} A x.$$

Choose any index i such that

$$|x_i| = \max_j |x_j| > 0.$$

Then

$$\lambda x_i = \sum_{j \neq i} \left(\frac{a_{ij}}{a_{ii}} \right) x_j$$

$$|\lambda| \; |x_i| \leq \left(\sum_{j \neq i} \frac{|a_{ij}|}{|a_{ii}|} \right) \max_j |x_j| < |x_i|.$$

Therefore any $\lambda \in \lambda(I - A_D^{-1} A)$ satisfies $|\lambda| < 1$.

The condition which an $n \times n$ matrix A must satisfy to be an M-matrix which is most often quoted involves the following two requirements:

(P1)A has off-diagonal elements which are non-positive

(P2)all principal minors of A are positive.

It will be sufficient for our discussion to consider only a few equivalent requirements for property (P2). (Fiedler and Pták [FI1] discuss several additional alternate requirements and properties of M-matrices.)

Theorem B.3: Let an $n \times n$ matrix A have off-diagonal elements which are non-positive. Then we say A is an M-matrix if any of the following equivalent conditions are met:

(C1)all principal minors of A are positive

(C2)there exists a vector $x \geq 0$;such that $A x > 0$

(C3)there exists a diagonal matrix D with positive diagonal elements such that $W = AD$ is *row* diagonally dominant.

Proof: (cf. [FI1]).

The significance of M-matrices for many problems involving perturbations is that they can be ordered in a way which is much stronger than for nonnegative matrices.

Theorem B.4: Let A and B both have off-diagonal elements which are non-positive. Assume that A is an M-matrix and $B \geq A$. Then B is also an M-matrix and satisfies the following:

(P1) $0 \le B^{-1} \le A^{-1}$

(P2) $\det B \ge \det A > 0$

(P3) $A^{-1}B \ge I$ and $B^{-1}A \ge I$

(P4) $A^{-1}B$ and $B^{-1}A$ are also M-matrices

(P5) $\rho(I - B^{-1}A) < 1$ and $\rho(I - A^{-1}B) < 1$

Proof: (cf. [FI1]).

Finally, we provide the following claim.

Claim B.5: Let Z be an $n \times n$ complex matrix. Then Z is generalized block diagonally dominant (in the sense of Limebeer, cf. theorem 3.7 of section 3.3.2) if and only if Z is a composite H-matrix (in the sense of Nwokah, cf. section 3.3.1) with respect to the same partition of Z.

Proof: Assume that Z is a composite H-matrix. From eqns. (3.16) and (3.17) this means that the test matrix

$$W - B = \begin{bmatrix} [\![Z_{11}]\!] & \cdots & -\|Z_{n1}\| \\ \vdots & \ddots & \vdots \\ -\|Z_{n1}\| & \cdots & [\![Z_{11}]\!] \end{bmatrix}$$

is an M-matrix. From property (C3) of theorem B.3 we see that this implies that there exists a diagonal matrix $D>0$ such that $(W - B)D$ is row diagonally dominant. Restating this as

$$d_i \, [\![Z_{ii}]\!] > \sum_{j \ne i} d_j \|Z_{ij}\|$$

for each index i of the partition we get the definition of generalized block diagonal dominance eqn. (3.21).

The proof is completed by essentially reversing the argument. Let Z be generalized block diagonally dominant. Then a diagonal scaling matrix D is obtained as in (3.26).

APPENDIX C

MATLAB Code for Computation of the Gap and Min-gap Functions

The following MATLAB code was used to compute the examples in chapters 4 and 5. The computational approach employed was described in section 4.2.6 of this dissertation.

```
// MATLAB SOURCE - MGAP.SRC - Wm. Bennett - 4/84
// PURPOSE: TO COMPUTE THE GEOMETRIC STABILITY MARGIN (PHI3(W))
// ASSOCIATED WITH A FEEDBACK PAIR (F,G)
// APPROACH: 1) COMPUTE THE PRINCIPAL ANGLES BETWEEN A PAIR OF SUBSPACES
//        2) MIN-GAP IS MINIMUM SIN(PRINCIPAL ANGLE)
// ASSUMES: G IS PXM, F IS MXP, AND M>P
AA = <EYE(M);G>;
BB = <F;-EYE(P)>;
<QA;RA> = QR(AA);
<QB;RB> = QR(BB);
QA1 = QA(:,1:M);
ZB1 = QB(:,P+1:P+M);
MGAP = SVD(ZB1'*QA1);// RETURNS AN ORDERED VECTOR OF THE PRINCIPAL SINES
```

```
// MATLAB SOURCE - GAP.SRC - Wm. Bennett - 4/84
// PURPOSE: TO COMPUTE THE GAP-METRIC BETWEEN A PAIR OF ABSTRACT NYQUIST
// LOCI ASSOCIATED WITH A PAIR OF FREQUENCY RESPONSES FOR G AND GD
// APPROACH: 1) COMPUTE THE PRINCIPAL ANGLES BETWEEN A PAIR OF SUBSPACES
//        2) GAP IS MAXIMUM SIN(PRINCIPAL ANGLE)
// ASSUMES: G AND GD ARE BOTH PXM WITH M>P
GG = <EYE(M);G>;
GGD = <EYE(M);GD>;
<QG,RG> = QR(GG);
<QD,RD> = QR(GGD);
QG1 = QG(:,1:M);
ZD1 = QD(:,M+1:P+M);
GAP = SVD(ZD1'*QG1);// RETURNS AN ORDERED VECTOR OF THE PRINCIPAL SINES
```