AD-A069 766 ILLINOIS UNIV AT URBANA-CHAMPAIGN COORDINATED SCIENCE LAB F/G 12/1

MODEL REDUCTION AND DECENTRALIZED CONTROL OF LARGE SCALE SYSTEM--ETC(U)

AUG 78 E C TSE DAAB07-72-C-0259

UNCLASSIFIED R-820 NL

1 OF 3

AD

A069766

LEVEL

REPORT DC-18 AUGUST, 1978

COORDINATED SCIENCE LABORATORY

DECISION AND CONTROL LABORATORY

AD A069766

MODEL REDUCTION AND DECENTRALIZED CONTROL OF LARGE SCALE SYSTEMS USING CHAINED AGGREGATION

DDC FILE COPY

UNIVERSITY OF ILLINOIS – URBANA, ILLINOIS

79 06 12 152

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 MODEL REDUCTION AND DECENTRALIZED CONTROL OF LARGE SCALE SYSTEMS USING CHAINED AGGREGATION,		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER 14 R-820, UILU-ENG 78-2213
7. AUTHOR(s) 10 Edwin Chi Yu Tse		8. CONTRACT OR GRANT NUMBER(s) 15 DAAB-07-72-C-0259, ERDA-EX-76-C-01-2088 NSF ENG 74-20091
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory University of Illinois at Urbana-Champaign Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 12 257 p.
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE 11 Aug[redacted]978
		13. NUMBER OF PAGES 244
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 9 Doctoral thesis,		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Large Scale Systems
Model Reduction
Decentralized and Hierarchical Control

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report proposes the model reduction of linear time invariant systems and the syntehsis of decentralized control with parallel computations by utilizing the given system information structure, e.g., structural and output information.

We first develop the chained aggregation procedure which leads to the generalized Hessenberg system representation (GHR). The structural features of the GHR are studied. The influences of the Hessenberg Blocks on trajectory and eigenvalue perturbation are examined. Secondly we develop the generalized

next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

mt

097 700

20. ABSTRACT (continued)

QL (GQL) and the restricted QL (RQL) algorithms to be used in conjunction with the GHR for construction of reduced order models. The GHR is shown to exhibit a class of plausible reduced order models suitable for further parameter adjustment if necessary. Finally, we exploit the use of GHR structure for the synthesis of decentralized control in the tracking of reference interconnection variables in an interconnected large scale system. A hierarchical structure is proposed which allows decentralized and parallel computations. Model reduction and control synthesis are illustrated using power system and rocket dynamics as examples.

UILU-ENG 78-2213

MODEL REDUCTION AND DECENTRALIZED CONTROL OF LARGE SCALE SYSTEMS USING CHAINED AGGREGATION

by

Edwin Chi Yu Tse

This work was supported in part by the Joint Services Electronics Program (U. S. Army, U. S. Navy and U. S. Air Force) under Contract DAAB-07-72-C-0259, in part by the Department of Energy under Contract ERDA-EX-76-C-01-2088 and in part by the National Science Foundation under Grant NSF INT 77-20969.

MODEL REDUCTION AND DECENTRALIZED CONTROL OF LARGE SCALE SYSTEMS USING CHAINED AGGREGATION

BY

EDWIN CHI YU TSE

B.S., University of Illinois, 1973
M.S., University of Illinois, 1975

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1978

Thesis Advisor: Professor William R. Perkins

Urbana, Illinois

MODEL REDUCTION AND DECENTRALIZED CONTROL OF LARGE SCALE SYSTEMS
USING CHAINED AGGREGATION

Edwin Chi Yu Tse, Ph.D.
Coordinated Science Laboratory and
Department of Electrical Engineering
University of Illinois at Urbana-Champaign, 1978

This thesis proposes the model reduction of linear time invariant systems and the synthesis of decentralized control with parallel computations by utilizing the given system information structure, e.g., structural and output information.

We first develop the chained aggregation procedure which leads to the generalized Hessenberg system representation (GHR). The structural features of the GHR are studied. The influences of the Hessenberg blocks on trajectory and eigenvalue perturbation are examined. Secondly we develop the generalized QL (GQL) and the restricted QL (RQL) algorithms to be used in conjunction with the GHR for construction of reduced order models. The GHR is shown to exhibit a class of plausible reduced order models suitable for further parameter adjustment if necessary. Finally, we exploit the use of GHR structure for the synthesis of decentralized control in the tracking of reference interconnection variables in an interconnected large scale system. A hierarchical structure is proposed which allows decentralized and parallel computations. Model reduction and control synthesis are illustrated using power system and rocket dynamics as examples.

ACKNOWLEDGMENT

The author wishes to thank his advisor, Professor W. R. Perkins, for his constant encouragement and guidance during the preparation of this thesis. Most sincere thanks also go to Professor J. V. Medanic for numerous hours of fruitful discussions and invaluable suggestions and guidance. He would also like to thank Professors J. B. Cruz, Jr., and P. V. Kokotovic for serving on his dissertation committee.

The author also wishes to express his appreciation to his wife, Helen, for her understanding, patience and encouragement.

Thanks are also given to Mrs. Rose Harris and Mrs. Wilma McNeal for their excellent typing.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

One purpose of this thesis is to develop methodology for reduced order modeling of a large scale system utilizing the given system information structure, e.g. structural and output informations. Then using insights gained from the reduced order modeling problem and the decentralized structure of a interconnected system with decentralized output information, control synthesis procedure is proposed that allows decentralized feedback control for tracking of interconnection variables. In addition, the proposed strategy allows parallel computations without excessive information transfer between subsystems.

Often high system order prevents a detail analysis and loss of structural insights. Moreover, often one is only interested in an input-output behavior of the system, without requiring a complete detail of the internal system variables. Consequently, a reduced order model that best characterizes the input-output behavior is desirable. The problem of obtaining reduced order models of dynamic systems has been considered from many points of view, such as curve fitting, minimization of some functional of the error between compatibly defined outputs of the system and the reduced order model, reduction by continued-fraction expansion, by spatial decomposition (ε-coupling), by temporal decomposition (singular perturbation), and by modal decomposition, [17 - 21, 31 - 35]. Other references on reduced order modeling may be found in literature overviews [36 - 38] and turtorial surveys on large scale systems in general [38 - 39]. However many existing methods produce reduced order models that bear no resemblance in structure to the full order system. In addition, it is difficult to relate the states of the reduced order model to those of the original system. Thus use of reduced order model to design control strategy to be used on the actual system can be unsatisfactory.

The methodology for reduced order modeling proposed in this thesis represents an attempt to eliminate some of the drawbacks in existing methods. The approach is based on the chained aggregation procedure, a generalization of the concept of dynamic aggregation [1]. It systematically transforms the system to a new representation called the generalized Hessenberg system representation (GHR). The GHR is induced by system output information structure. In general the GHR defines a collection of plausible reduced order models and defines in each case, the structure of the model. It is shown that the GHR may be decomposed into an aggregate subsystem S^i and a corresponding residual subsystem $\bar{S}^i$. An approximate aggregated model may be obtained by neglecting the coupling between S^i and $\bar{S}^i$ if weak coupling exists. The residual $\bar{S}^i$ then represents a weakly observable subsystem in the available output. On the other hand, if the weakly observable part cannot be simply neglected, it also defines a subset of model parameters appropriate for adjustment. Adjustment of these parameters is performed by the generalized QL (GQL) or the restriced QL (RQL) algorithms which are a generalization of the well known QR type algorithms [28]. A property of this generalization is that it orders the system eigenvalues according to their dominance. Thus it also provides a prescriptive method for the time scale decompositions and applications of singular perturbation theory [20] particularly when dominant modes are strongly observable in the outputs. The GQL algorithm is developed in detail together with its convergence property and implications on the GHR.

Decentralized control design has been studied by several researchers [38 - 39, 63 - 65]. Earlier works have concentrated on the existence and stabilizing property of decentralized control [64 - 65]. Applications to realistic large scale system using decentralized control were studied by Tripathy and Davison [66]. It is shown that for a large scale system with interconnected structures, desirable stability performance can be achieved by decentralized control without resort to the more complicated centralized control. However, existing proposed methods in general employs centralized computation in designing the decentralized control. This can be undesirable and often the computation savings is not immediately clear. To take advantage of the decentralized structure in a large scale system, often parallel computations are desirable at subsystem level. The actions of various local decision makers are coordinated by a coordinator who in general is mainly concerned with overall system level goals. In this thesis, we adopt an interconnected composite system representation where each subsystem is interconnected to other subsystems only through subsystem outputs. Thus if time evolution or dynamic equations governing the desired interconnection inputs are known, each set of subsystem equations becomes independent from others providing these desired interconnection inputs are agreed upon and achieved. It is shown that using the GHR structural insights from the first part of the thesis, a hierarchical decentralized control strategy can be developed.

If subsystem decision makers mutually agree to allocate the system interconnection variables to a coordinator who specifies the desired trajectory base on some system level goal, then a decoupling results. Parallel computations can be carried out by subsystems under coordination by the coordinator. It is also shown that a local control u_i may be decomposed into three basic components. One for achieving aggregation and decoupling; one for controlling the aggregate and thus specifies the reference system interconnections; and finally one non-interactive component, if exists, allows the control of the residual.

The thesis is organized as follow. In Chapter 2, the basic idea of dynamic aggregation by Aoki [1] is reviewed and re-interpreted. Extension of Aoki's results lead to the chained aggregation procedure and the GHR system representation. Properties of GHR representation and other topics related to GHR are studied. Chapter 3 developes the generalized QL theory to be used for model reduction in Chapter 4. Algebraic and geometric convergence properties are given. In Chapter 4, the model reduction problem via the GHR is studied. The implications of using existing model reduction techniques on the GHR are examined. Finally, the GQL theory is applied to define the RQL algorithm for construction of reduced order models by the GHR. The synthesis of decentralized control and the formulation of decentralized control strategy with parallel computation is given in Chapter 5. It is shown a decomposed optimization problem can be formulated as a coordinator level problem and a set of local subsystem level problems. The construction of reduced order models and the synthesis of decentralized hierarchical controls are illustrated in Chapter 6 using typical large scale system as bench-mark models. Performance of developed methodologies is supported by many simulation runs.

CHAPTER 2

CHAINED AGGREGATION AND THE GENERALIZED HESSENBERG REPRESENTATION OF A LINEAR SYSTEM

This chapter begins with a re-eximination of the basic notion of aggregation first intorduced by Aoki [1,2]. Aggregation is first viewed as a result of linear transformation on the system and re-interpreted in terms of observability of a linear time invariant system (LTI) with respect to the aggregation matrix. The definitions of complete aggregability and aggregability are introduced which facilitate the extension of the basic aggregational idea into a chained aggregation procedure. The end result is a new system representation denoted as the generalized Hessenberg representation (GHR). The chained aggregation algorithm which defines the GHR is given.

To study the implication and significance of the GHR, several properties such as invariant internal structure, observability and controllability of subsystems composing the GHR, and influence of the Hessenberg block $F_{i,i+1}$ on the aggregates are examined in the remaining sections of the chapter, several illustrative examples are also given.

2.1. Basic Aggregation

The concept of static aggregation is well known to economists. Dynamic aggregation was first introduced in 1968 by Aoki [1]. Since then, the subject has been studied by several researchers [3,4,5,6,7,10,11]. We first present a brief review and reinterpretation of basic results in dynamic aggregation [2,8]. Consider a linear time invariant system

$$\dot{x} = Ax + Bu \tag{2.1}$$

where $x \in R^n$, $u \in R^m$. A reduced r_1-th order aggregated state $z_1, r_1 < n$, is defined as $z_1 = Cx$ where without loss of generality it is assumed that the $r_1 \times n$ <u>aggregation matrix C</u> has full row rank. In the sense of Aoki, a linear time invariant model

$$\dot{z}_1 = F_{11}z_1 + G_1u \tag{2.2}$$

is desired to describe the dynamic variations of the aggregated state $z_1(t)$. By definition of z_1, we have

$$\dot{z}_1 = CAx + CBu \tag{2.3}$$

From (2.2) and (2.3), the aggregation conditions are

$$CA = F_{11}C \tag{2.4a}$$

$$CB = G_1 \tag{2.4b}$$

system (2.2) is termed the aggregated model of (2.1). It is worth pointing out that (2.2) is an exact dynamic aggregate of (2.1) if and only if F_{11} and G_1 are chosen such that (2.4) is satisfied. To see this, let us introduce the following definitions:

<u>Definition 2.1</u> [1,2]: System (2.1) is said to be <u>completely aggregable with respect to C</u> if there exists an $r_1 \times r_1$ matrix F_{11} such that $CA \equiv F_{11}C$.

Generalizing Definition 2.1, we have:

<u>Definition 2.2</u> [13]: System (2.1) is said to be <u>aggregable with respect to C</u> if there exists a matrix $\overline{C}$ of maximal rank such that (2.1) is completely

aggregable with respect to the expanded matrix

$$\tilde{C} = \begin{bmatrix} C \\ \bar{C} \end{bmatrix}$$

<u>Remark</u>: The generalization to Definition 2.2 is a consequence of the fact that given a triple (A,B,C), there may not exist a matrix F_{11} such that (2.4a) is identically satisfied. The possibility of enlarging the aggregated state z_1 which retains the selected C as a submatrix so that $\tilde{C}A \equiv F\tilde{C}$ for some matrix F motivates Definition 2.2.

To examine what is achieved by aggregation given (A,B,C), it is advantageous to view complete aggregation as an outcome of a linear transformation of (2.1). We partition the aggregation matrix C as

$$C = [C_{11} \quad \bar{C}_{11}]$$

where C_{11} is an $r_1 \times r_1$ nonsingular matrix (this can always be done since C is assumed to have maximal rank). Define the linear transformation on (2.1) as:

$$z = Tx, \quad T = \begin{bmatrix} C_{11} & \bar{C}_{11} \\ 0 & I_{n-r_1} \end{bmatrix}, \quad z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{2.5}$$

Application of (2.5) to (2.1) yields

$$\dot{z} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} z + \begin{bmatrix} G_1 \\ \bar{G}_1 \end{bmatrix} u. \tag{2.6}$$

where

$$F_{11} = C_{11}A_{11}C_{11}^{-1} + \overline{C}_{11}A_{21}C_{11}^{-1}$$

$$F_{12} = C_{11}A_{12} + \overline{C}_{11}A_{22} - (C_{11}A_{11} + \overline{C}_{11}A_{21})C_{11}^{-1}\overline{C}_{11}$$

$$F_{21} = A_{21}C_{11}^{-1} \tag{2.7}$$

$$F_{22} = A_{22} - A_{21}C_{11}^{-1}\overline{C}_{11}$$

$$G_1 = CB, \ \overline{G}_1 = B_2$$

and $A_{i,j}$, B_i, $i,j = 1,2$ are submatrices of A, B of (2.1) partitioned compatible with T. The resulting representation may be considered to be an interconnection of two subsystems, the <u>aggregate subsystem</u> S^1:

$$\dot{z}_1 = F_{11}z_1 + F_{12}z_2 + G_1u. \tag{2.8a}$$

and the <u>residual subsystem</u> $\overline{S}^1$:

$$\dot{z}_2 = F_{21}z_1 + F_{22}z_2 + G_1u. \tag{2.8b}$$

If (2.1) is completely aggregable, Definition 2.1 says a matrix $F_{11} \in R^{r_1 \times r_1}$ exists such that (2.4a) holds. The partitions of A and C in

$$[C_{11} \quad \overline{C}_{11}] \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \equiv F_{11}[C_{11} \quad \overline{C}_{11}]$$

implies

$$F_{11} \equiv C_{11}A_{11}C_{11}^{-1} + \overline{C}_{11}A_{21}C_{11}^{-1}$$

and

$$C_{11}A_{12} + \bar{C}_{11}A_{22} - (C_{11}A_{11} + \bar{C}_{11}A_{21})C_{11}^{-1}\bar{C}_{11} \equiv 0$$

Thus comparing with (2.7), we see $F_{12} \equiv 0$. Hence the composite system reduces to a tandem configuration with the aggregate subsystem driving the residual and no feedback from the residual into the aggregate. See Fig. 2.1.

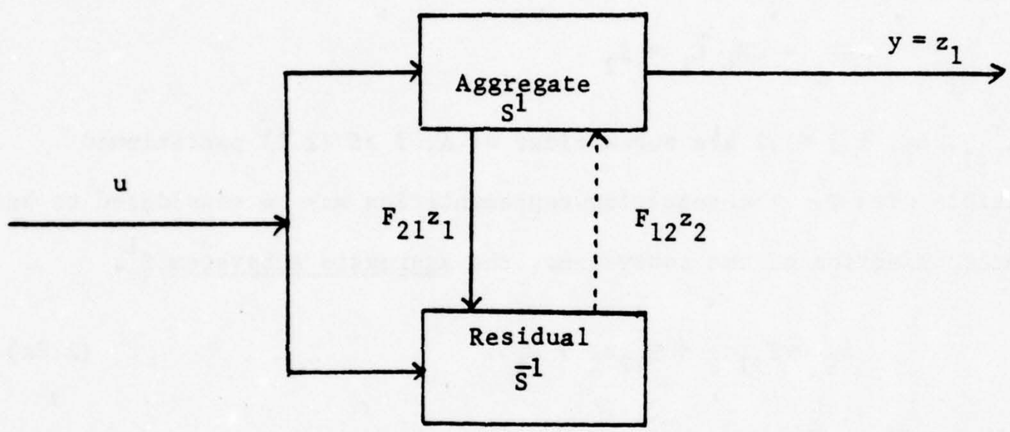


Figure 2.1. Tandem configuration of the aggregate and the residual subsystems under complete aggregation.

If we view the aggregated variable z_1 as the output of system (2.1), when complete aggregability occurs ($F_{12} \equiv 0$), the substate z_2 is unobservable through the aggregate. We realize (2.6) with $F_{12} \equiv 0$, and the output equation

$$y = z_1 = [I \quad 0]z$$

is an observable decomposition cannonical form [9]. Important and illuminating connections between aggregation and observability were summarized in [10, 11]. The basic conclusions may be stated as follows:

Theorem 2.1. Given a pair (A, C), the following statements are equivalent:

(a) The system (2.1) is completely aggregable with respect to C.

(b) $\mathcal{R}\{(CA)^T\} \subseteq \mathcal{R}\{C^T\}$, where $\mathcal{R}\{X\}$ denotes the range space of X.

(c) $A\,\mathrm{Ker}\{C\} \subseteq \mathrm{Ker}\{C\}$.

Moreover, if M is the observability matrix of (A,C), and rank $M = p$, then (A,C) is aggregable, i.e. there exists a matrix $\bar{C} \in R^{p \times r}$ with maximal rank such that $(A,\tilde{C})$ is completely aggregable, $\tilde{C} = [C^T \bar{C}^T]^T$.

Proof: Statements (a) - (c) are obtained trivially using Definition 2.1 and the observation that (2.6) with $F_{12} \equiv 0$ is the observable decomposition canonical structure. To prove the remaining part of the Theorem, notice if

$$\text{rank } M = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = p,$$

then

$$\text{rank} \begin{bmatrix} CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = \text{rank } M_o = p - r_1$$

Select any $p - r_1$ linearly independent rows of M_o and denote as $\bar{C}$, then $\tilde{C}$

induces a map such that the representation in appropiate basis has the observable decomposition cannonical structure [11]. Apply (a) - (c) and Definition 2.1 - 2.2, the conclusion follows.

2.2. The Chained Aggregation Algorithm and the Generalized Hessenberg System Representation

Given the LTI system (2.1), the aggregated state is defined as $z_1 = Cx$ where the aggregation matrix is in general selected according to some engineering judgement or requirement. For simplicity, hence forth we shall choose the system output matrix as the desired aggregation matrix. Thus consider the following LTI system:

$$\dot{x} = Ax + By,$$

$$y = Cx \tag{2.9}$$

where $y \in R^{r_1}$ is the available output, and consider the problem of constructing a reduced order model that reflects the input-output characteristics of the system. In the spirit of aggregation, if the above system is completely aggregable with respect to the output matrix C then the reduced order model will have the form (2.2) with $y = z_1$, and will exactly describe the dynamics of the output y.

However, by far the more interesting and realistic case occurs when the system (2.1) is not completely aggregable with respect to C; that is when (A, C) is a completely observable pair or the observable subspace is of dimension greater than r_1 but less than n. An extension of aggregation called <u>chained aggregation</u> is now proposed. Chained aggregation is in essence a sequence of aggregation steps performed on successively defined residual subsystems that either reduces (2.9) into a particular observable decomposition canonical form when (A, C) is aggregable or into a structured representation called the generalized Hessenburg representation when (A,C) is not aggregable. Let us first introduce the following definition:

<u>Definition 2.3</u>: A system representation (F,G,D) is said to be a Generalized Hessenberg Representation (GHR) if

$$\dot{z} = Fz + Gu$$

$$y = Dz \tag{2.10}$$

where

$$F = \begin{bmatrix} F_{11} & F_{12} & 0 \; \dots \; 0 & \dots & \dots & 0 \\ F_{21} & F_{22} & F_{23} & 0 \; \dots & \dots & 0 \\ \vdots & & & & & \vdots \\ F_{j1} & F_{j2} & \dots & F_{jj} \; F_{jj+1} & 0 \; \dots & 0 \\ \vdots & & & & & \\ F_{k-2,1} & \dots & & \dots & F_{k-2,k-1} & 0 \\ F_{k-1,1} & F_{k-1,2} & \dots & & \dots & F_{k-1,k} \\ F_{k,1} & F_{k,2} & \dots & & \dots & F_{k,k} \end{bmatrix}, G = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ \vdots \\ G_{k-1} \\ G_k \end{bmatrix}$$

$$D = [I_1 \;\; 0 \dots 0]^* \tag{2.11}$$

$^*I_i$ denotes a $r_i \times r_i$ dimensional identity matrix.

with $F_{ii} \in R^{r_i \times r_i}$, $r_i \geq r_{i+1}$*, $\sum_{i=1}^{k} r_i = n$.

<u>Remark</u>: The representation (2.10) is called the Generalized Hessenberg Representation because the block matrices $F_{ij} = 0$ for $i = 1,2,\ldots k-2$, $j = i+2,\ldots,k$ and thus F generalizes the concept of an n-th order Hessenberg matrix characterized by $f_{ij} = 0$ for $i = 1,2,\ldots k-2$, $j = i+2,\ldots,n$ well known in numerical literature [23].

<u>Remark</u>: Given a triple (A,B,C), the GHR displays a block structure with dimensions r_i, $1 \leq i \leq k$. It will be shown subsequently (Theorem 2.3) that this block structure is fixed once (A,B,C) is given. In particular, the numbers r_i is only dependent on the pair (A,C), and is therefore said to be induced by the system information structure characterized by the output matrix C. Furthermore, the system may be considered to be the interconnection of k subsystems, characterized by the subsystem matrices F_{ii} and mutually coupled by the interactions F_{ij}, $j = 1,\ldots i+1$. Hence the GHR is a sequence of subsystems coupled so that each subsystem receives feedforward signals from <u>all</u> subsystems preceding it and a feedback signal <u>only</u> from the subsystem immediately following it in the interconnection. The interconnection pattern has a particular form and exhibits the essential feedback links within the system. This is shown in Fig. 2.2.

---

*$r_i \geq r_{i+1}$ always holds if (A,C) is non-aggregable. If (A,C) is aggregable, then $F_{j,j+1} = 0$, $1 < j \leq k-1$. In this case, the residual which represents the unobservable subspace of dimension $n - \sum_{i=1}^{j} r_i$ can be represented by a scalar Hessenberg form with $r_i \geq r_{i+1}$ still hold for $\forall i$.

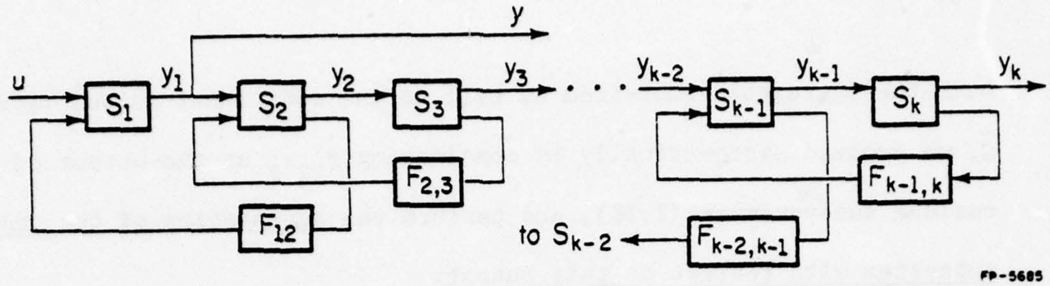


Figure 2.2. Structural diagram of the GHR (most feedforward paths have been eliminated to simplify this structural diagram).

Theorem 2.2: Any LTI system (2.9) can be transformed into the Generalized Hessenberg Representation (2.10)-(2.11).

This theorem will be established by a constructive proof which links the GHR to aggregation [1]. The procedure will be explained in a series of steps, the steps forming the chained aggregation procedure.

Proof:

Step 1: Recalling that now $y = z_1$ and using (2.5) we have

$$\begin{bmatrix} \dot{y} \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} F_{11} & F'_{12} \\ F'_{21} & F'_{22} \end{bmatrix} \begin{bmatrix} y \\ x_2 \end{bmatrix} + \begin{bmatrix} G_1 \\ G'_2 \end{bmatrix} u \tag{2.12}$$

as the equivalent representation of (2.9). If the system (2.9) is completely aggregable with respect to C, $F'_{12} \equiv 0$ and (2.12) is the GHR. If not, proceed with Step 2.

Step 2: If the system (2.9) is not completely aggregable with respect to C, then $F'_{12} \neq 0$. It may still be possible to obtain aggregation by enlarging the output vector as a consequence of Definition 2.2. To obtain both the aggregable subsystem as well as the additional output matrix $\bar{C}$, we proceed systematically by considering $F_{12}'x_2$ as the output of the residue subsystem in (2.12), and perform the <u>aggregation of the residue subsystem with respect to this output</u>.

Denote the new output $F_{12}'x_2$ by w_2. Since $F_{12}' \in R^{r_1 \times n_2}$, where $n_2 = n - r_1$, in general F_{12}' does not have full rank. To extract linearly independent components of w_2, note that there exists a nonsingular matrix E_2 such that

$$E_2 w_2 = E_2 F_{12}' x_2 = \begin{bmatrix} C_2 \\ 0 \end{bmatrix} x_2 \tag{2.13}$$

where $C_2 \in R^{r_2 \times n_2}$, $r_2 \le r_1$, E_2 is the product of elementary Gaussian elimination matrices [12] each differing from the unit matrix in only certain rows. Define

$$y_2 = C_2 x_2.$$

Then

$$w_2 = E_2^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} y_2 = F_{12} y_2$$

Hence the residual subsystem of (2.12) is now described by

$$\dot{x}_2 = F_{21}' y_1 + F_{22}' x_2 + G_2' u \tag{2.14a}$$

$$y_2 = C_2 x_2. \tag{2.14b}$$

Now proceed with chained aggregation. Partition C_2 as

$$C_2 = [C_{22} \quad \bar{C}_{22}] \tag{2.15}$$

where C_2 is assumed nonsingular,* and define the $n_2 \times n_2$ nonsingular matrix

$$D_2 = \left[\begin{array}{c:c} C_{22} & \bar{C}_{22} \\ \hdashline 0 & I_{n_3} \end{array}\right] \tag{2.16}$$

where $n_3 = n_2 - r_2$. Applying (2.16) to (2.14) we obtain after two chained aggregation steps:

$$\begin{aligned} \dot{y}_1 &= F_{11}y_1 + F_{12}y_2 + \qquad\quad G_1u \\ \dot{y}_2 &= F_{21}y_1 + F_{22}y_2 + F'_{23}x_3 + G_2u \\ \dot{x}_3 &= F'_{31}y_1 + F'_{32}y_2 + F'_{33}x_3 + G'_3u. \end{aligned} \tag{2.17}$$

with $y_1 \equiv y$. If now $F'_{23} = 0$, the chained aggregation procedure terminates. If not, the procedure is repeated using the residual in (2.17) namely (2.17c) and $w_3 = F'_{23}x_3$ as the new output equation. Clearly, the process of chained aggregation terminates in a finite number of steps, i.e. k, where min $\{k\} = 1$ and max $\{k\} = n - r_1$. This completes the proof of Theorem 2.2.

It is possible to define the process of chained aggregation as a sequence of linear transformations on the full system (2.9). Define

---

*If C_{22} is singular then a further permutation of the components of x_2 is necessary to transform (2.14) into an equivalent form in which the resulting C_{22} would be nonsingular. Such a permutation always exists.

$$T_i = \left[\begin{array}{ccc:c} I_1 & & & \\ & \ddots \quad 0 & & 0 \\ 0 & & I_{i-1} & \\ \hdashline & & & \\ & 0 & & D_i \end{array}\right], \quad i = 1,2,\ldots k-1 \tag{2.18}$$

where the matrix D_i is the transformation matrix on the residual similarly defined as in (2.5) or as in (2.16) except for the i-th step of the chained aggregation. Then the transformation H that brings the original system (2.9) into the GHR (2.10)-(2.11) is

$$H = T_{k-1}\, T_{k-2} \cdots T_1 . \tag{2.19}$$

Note that H is lower block triangular and partitioned compatibility with the GHR. If, in general, it becomes necessary to permute the components of the residue subvector x_i during intermediate steps of chained aggregation to assure that C_{ii} is nonsingular, then H is given by

$$H = T_{k-1}P_{k-1}T_{k-2}P_{k-2} \cdots T_1P_1 \tag{2.20}$$

where P_i is the permutation matrix at i-th step of chained aggregation required to make C_{ii} non-singular (C_i defined similar to (2.15)). As a consequence H is no longer upper block triangular.

Remark: If the system (2.9) is of the form

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Nu \end{aligned} \tag{2.9'}$$

The chained aggregation procedure can still be applied. Let H be the GHR transformation for (2.9) and applied to (2.9'), then

$$z = \begin{bmatrix} C \\ \bar{H} \end{bmatrix} x = \begin{bmatrix} Cx \\ \bar{H}x \end{bmatrix} = \begin{bmatrix} z_1 \\ \bar{z} \end{bmatrix}$$

Clearly

$$y = z_1 + Nu$$

Thus (2.10) is modified to be

$$\begin{aligned} \dot{z} &= Fz + Gu \\ y &= Dz + Nu \end{aligned} \tag{2.10'}$$

<u>Remark</u>: Because of the strong relation between aggregability and observability, the generalized Hessenberg transformation matrix H, (2.20) is closely related to the observability matrix M of the pair (A,C). In particular if $H^T = [\bar{H}^T \;\; \tilde{H}^T]$, then $\bar{H} \in \text{span}\,\{m_1,\ldots,m_p\}$, $p \leq n$, where $m_1,\ldots,m_p$ is the set of linearly independent row vectors of M. Also, since chained aggregation will identify the observable subspace when the system is aggregable, it is noteworthy to point out the relation with results in [14,15] concerning the unobservable subspace. The procedures are complementary in the sense that chained aggregation identifies the observable subspace by gradually extending $\mathcal{R}(C)$ while the procedure in [14] and related results in [15] identify the unobservable subspace by gradually decreasing $\mathcal{N}(C)$.

Based on the structure of the GHR we now extend the definition of the aggregate and residual subsystems as follows: Given (2.11) define $F^i, H^i, G^i, E^i, \bar{F}^i$, and $\bar{G}^i$ as follows:

$$F^i = \begin{bmatrix} F_{11} & F_{12} & 0 & \cdots & 0 \\ F_{21} & F_{22} & F_{23} & \cdots & 0 \\ \vdots & & & & \vdots \\ F_{i1} & F_{i2} & F_{i3} & \cdots & F_{ii} \end{bmatrix} \quad E^i = \begin{bmatrix} 0 & \cdots\cdots & & 0 \\ 0 & \cdots\cdots & & 0 \\ \vdots & & & \\ 0 & & 0\cdots & 0 \\ F_{i,i+1} & & 0\cdots & 0 \end{bmatrix} \quad G^i = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_i \end{bmatrix} \tag{2.21}$$

$$H^i = \begin{bmatrix} F_{i+1,1} & \cdots & F_{i+1,i} \\ \vdots & & \\ F_{k1} & \cdots & F_{k,i} \end{bmatrix} \quad \bar{F}^i = \begin{bmatrix} F_{i+1,i+1} & \cdots & F_{i+1,k} \\ \vdots & & \vdots \\ F_{k,i+1} & \cdots & F_{k,k} \end{bmatrix}$$

$$\bar{G}^i = \begin{bmatrix} G_{i+1} \\ \vdots \\ G_k \end{bmatrix}$$

Then for $i = 1,\ldots,k-1$ and $p_i = \sum_{j=1}^{i} r_j$, the p_i-th dimensional subsystem

$$\dot{z}^i = H^i z^i + E^i \bar{z}^i + G^i u \tag{2.22}$$

is called the <u>aggregate subsystem</u> S^i and the $(n-p_i)$-th dimensional subsystem

$$\dot{\bar{z}}^i = H^i z^i + \bar{F}^i \bar{z}^i + \bar{G}^i u \tag{2.23}$$

is called the residual subsystem $\bar{S}^i$.

<u>Remark</u>: In representing a LTI system (2.9) in its GHR, three distinct

cases exist: (a) (A,C) is completely aggregable. In this case $F_{12} \equiv 0$ and the subsystem S^1 is decoupled from the residual; (b) (A,C) aggregable. In this case, there exists some $i \in \{1,2,\ldots k-1\}$ such that $F_{i,i+1}$ 0 and the subsystem S^i is decoupled from its corresponding residual subsystem; (c) (A,C) non-aggregable. In Chapter 4, it will be seen how S^i may be used as a reduced order model of (2.9) that approximates the input-output behavior of the full system.

2.3. Properties of the Generalized Hessenberg Representation

In this section, we shall examine several properties of the GHR and some problems of immediate interest that relate to the GHR.

2.3.1. Invariant Internal Structures of the GHR

Theorem 2.3. The internal structure of the GHR characterized by the indices $r_1,\ldots,r_k$ is unique.

Proof: It is sufficient to show that if an arbitrary linear transformation is initially applied to the system (2.9), rank F_{12} with F_{12} defined as in (2.6), is invariant with respect to this transformation. The theorem then follows by induction on the step i of the chained aggregation procedure. Given the system (2.9) let the transformation (2.5) bring the system into the representation (2.6). The dimension of the next layer in the GHR is then uniquely defined by rank F_{12}, with F_{12} given by

$$F_{12} = C_{11}(A_{12}-A_{11}C_{11}^{-1}C_{12}) + C_{12}(A_{22}-A_{21}C_{11}^{-1}C_{12}) \qquad (2.24)$$

where A_{ij}, $i,j=1,2$ comes from a decomposition of A compatible with that of T in (2.5). Note we have rewritten $C=[C_{11} \quad C_{12}]$.

Suppose an initial similarity transformation P is applied to (2.9) and the system transformed into

$$\dot{z} = \tilde{A}z + \tilde{B}u$$
$$y = \tilde{C}z \tag{2.25}$$

with $z = Px$, $\tilde{A} = PAP^{-1}$ and $\tilde{B} = PB$. The only restriction on P is that the leading square submatrix of $\tilde{C}$ is nonsingular. Apply now the first step of chained aggregation to (2.25). Defining, analogously to (2.5), the transformation

$$\tilde{T} = \begin{bmatrix} \tilde{C}_{11} & \tilde{C}_{12} \\ \cdots & \cdots \\ 0 & I \end{bmatrix} = \begin{bmatrix} CP^{-1} \\ \cdots \\ 0 \quad I \end{bmatrix} \tag{2.26}$$

$$\tilde{F} = \tilde{T}\tilde{A}\tilde{T}^{-1} = \begin{bmatrix} CP^{-1} \\ \cdots \\ 0 \quad I \end{bmatrix} (PAP^{-1}) \begin{bmatrix} CP^{-1} \\ \cdots \\ 0 \quad I \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} C_{11} & C_{12} \\ P_{21} & P_{22} \end{bmatrix} A \begin{bmatrix} C_{11} & C_{12} \\ P_{21} & P_{22} \end{bmatrix}^{-1} = \hat{T}A\hat{T}^{-1}. \tag{2.27}$$

Decomposte P and $\hat{T}^{-1}$ compatibly with $\hat{T}$ so that

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}, \hat{T} = \begin{bmatrix} C_{11} & C_{12} \\ P_{21} & P_{22} \end{bmatrix}, \hat{T}^{-1} = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}. \tag{2.28}$$

From (2.27) and in analogy with (2.24) it is found that

$$\tilde{F}_{12} = C_{11}(A_{11}W_{12} + A_{12}W_{22}) + C_{12}(A_{21}W_{12} + A_{22}W_{22}). \quad (2.29)$$

But from $\hat{T}\,(\hat{T}^{-1}) = I$, it follows that $C_{11}W_{12} + C_{12}W_{22} = 0$, $P_{21}W_{12} + P_{22}W_{22} = I$ and thus

$$W_{12} = -C_{11}^{-1}C_{12}W_{22}$$
$$W_{22} = (P_{22} - P_{21}C_{11}^{-1}C_{12})^{-1} \quad (2.30)$$

and

$$\tilde{F}_{12} = F_{12}W_{22}$$

Thus, rank F_{12} is invariant if W_{22} is nonsingular. We show finally that under condition that C_{11} and $\tilde{C}_{11}$ are both nonsingular, W_{22} as given by (2.30) is always nonsingular. To see this note that from (2.27) $\hat{T} = \tilde{T}P$ and therefore

$$\det \hat{T} = \det \begin{bmatrix} C_{11} & C_{12} \\ P_{21} & P_{22} \end{bmatrix} \neq 0.$$

However, by assumption $\det C_{11} \neq 0$, and thus

$$\det \hat{T} = \det C_{11} . \det\,(P_{22} - P_{21}C_{11}^{-1}C_{12}) \neq 0$$

Thus W_{22} is nonsingular, and the proof is completed.

The next theorem, Theorem 2.4, indicates that the indices r_i

are really a property of the pair (A,C), i.e. $r_i \forall i$ are completely determined once (A,C) is specified. It compliments Theorem 2.3 in the sense Theorem 2.3 suggests r_i are invariant with respect to linear transformations and system representation, but does not identify what determines r_i. Assume H is given by (2.19). In this case H is block upper triangular. Let us introduce the following notation:

$$H = [H_{ij}] = \begin{bmatrix} H_{11} & H_{12} & \cdots\cdots & H_{1k} \\ 0 & H_{22} & H_{23}\cdots & H_{2k} \\ \vdots & & & \\ 0 & \cdots 0 & H_{k-1,k-1} & H_{k-1,k} \\ 0 & \cdots & 0 & I \end{bmatrix} = \begin{bmatrix} H_1 \\ \vdots \\ H_k \end{bmatrix}$$

$$H_{ij} \in R^{r_i \times r_j}, \; i,j = 1,2,\ldots k$$

$$C_i = [H_{i,i} \;\cdots\; H_{ik}]$$

$$A = [A_1 \;\; A_2 \;\; \cdots \;\; A_k], \;\; A_i \in R^{n \times r_i}$$

$$A^i = [A_1 \;\cdots\; A_i], \;\; \bar{A}^i = [A_{i+1} \;\cdots\; A_k] \tag{2.31}$$

Note that the form of H is a consequence of (2.18) and can be shown from the steps of the chained aggregation algorithm. Furthermore,

$[H_{11} \;\; H_{12} \;\; \cdots \;\; H_{1k}] = C$ and H_{11} is nonsingular by assumption.

<u>Theorem 2.4</u>. Given the triple (A,B,C), a GHR may be obtained by direct computation from the relation

$$HA = FH \tag{2.32a}$$

$$G = HB \tag{2.32b}$$

Furthermore, the indices $r_1, \ldots r_k$ of the GHR is completely determined by the pair (A,C).

Proof: Using (2.31) in (2.32a) and equating left and right hand side layer by layer, we have

Layer 1

$$H_1 A_1 = F_{11} H_{11} \tag{2.33a}$$

$$H_1 \bar{A}^1 = F_{11}[H_{12} \cdots H_{1k}] + F_{12} C_2 \tag{2.33b}$$

Layer 2

$$H_2 A_1 = F_{21} H_{11} \tag{2.34a}$$

$$H_2 A_2 = F_{21} H_{12} + F_{22} H_{22} \tag{2.34b}$$

$$H_2 \bar{A}^2 = F_{21}[H_{13} \cdots H_{1k}] + F_{22}[H_{23} \cdots H_{2k}] + F_{23} C_3 \tag{2.34c}$$

Layer i

$$H_i A_1 = F_{i1} H_{11} \tag{2.35a}$$

$$H_i A_2 = F_{i1} H_{12} + F_{i2} H_{22} \tag{2.35b}$$

$$H_i A_3 = F_{i1} H_{13} + F_{i2} H_{23} + F_{i3} H_{33} \tag{2.35c}$$

$$\vdots$$

$$H_i A_i = F_{i1} H_{1i} + F_{i2} H_{2i} + \cdots + F_{ii} H_{ii} \tag{2.36}$$

$$H_i \bar{A}^i = F_{i1}[H_{1,i+1} \cdots H_{1k}] + F_{i2}[H_{2,i+1} \cdots H_{2,k}] + \cdots$$
$$+ F_{ii}[H_{i,i+1} \cdots H_{i,k}] + F_{i,i+1} C_{i+1} \tag{2.37}$$

Notice in (2.32a), although we can partition H and F according to their general structure, we know only H_1, A and r_1. The matrices H_i, F_{ij}, and indices r_i $2 \leq i \leq k$ remain to be determined. From (2.33a), knowing $H_{11} = C_{11}$ is nonsingular, we immediately have

$$F_{11} = H_1 A_1 H_{11}^{-1} \tag{2.38}$$

Thus

$$F_{12}C_2 = H_1\bar{A}^{-1} - F_{11}[H_{12} \;\cdots\; H_{1k}] = X_1 \tag{2.39}$$

From the construction of the GHR through chained aggregation, we know rank F_{12} = rank $C_2 = r_2$. Since $F_{12} \in R^{r_1 \times r_2}$ and $C_2 \in R^{r_2 \times \bar{n}_1}$, $\bar{n}_1 = n - r_1$, it follows that, rank $F_{12}C_2 \equiv r_2$. By (2.39), X_1 is completely determined. Thus

$$r_2 \equiv \text{rank } X_1 \tag{2.40}$$

Furthermore, given a matrix X_1 with rank r_2, we can always find a factorization F_{12} and C_2 of maximal rank. Thus r_2, F_{12} and C_2 are completely determined. The non-uniqueness of the GHR comes from the fact that the factorization $X_1 = F_{12}C_2$ is not unique.

Knowing F_{11}, F_{12}, C_2 and r_2, we may move on to the equations of the next layer to determine F_{21}, F_{22}, F_{23}, C_3 and r_3. To prove the theorem for the i-th layer, let us assume that $F_{j,k}$, H_j, r_j for $j = 1,2$ $j = 1,2,\ldots i-1$ are known. In addition, H_i, r_i is also known from calculations of $i-1^{th}$ layer. Since H is assumed to be given by (2.19), H_{ii} is nonsingular for all i. From (2.35a)

$$F_{i1} = H_i A_1 H_{11}^{-1} \tag{2.41}$$

Similarily

$$F_{i2} = (H_i A_2 - F_{i1} H_{12}) H_{22}^{-1} \tag{2.42}$$

In general for a fixed j, $j \leq i$

$$F_{ij} = (H_i A_j - \sum_{k=1}^{j-1} F_{ik} H_{kj}) H_{jj}^{-1} \tag{2.43}$$

To obtain $F_{i,i+1}$, we use (2.37)

$$F_{i,i+1} C_{i+1} = H_i \bar{A}^i - \sum_{j=1}^{i} F_{ij} [H_{j,i+1} \cdots H_{j,k}] = X_i \tag{2.44}$$

By same argument as before, $F_{i,i+1} \in R^{r_i \times r_{i+1}}$. $C_{i+1} \in R^{r_{i+1} \times \bar{n}_i}$, $\bar{n}_i = n - \sum_{j=1}^{i} r_j$. Futhermore, rank $F_{i,i+1} = r_{i+1}$, rank $C_{i+1} = r_{i+1}$ implies rank $F_{i,i+1} C_{i+1} =$ rank $X_i = r_{i+1}$. Knowing r_{i+1}, we may factor X_i into $F_{i,i+1}$ and C_{i+1} each having maximal rank. Thus $F_{i,k}$, $k = 1,2,\ldots i+1$, H_{i+1} have been determined. By induction, all submatrices of H and F, r_i, $i = 1,\ldots,k$ can be solved for. In addition, the matrices H, F and indices r_i are induced by only $H_1 = C$ and A, thus is an invariant property of the pair (A,C). Having obtained H, G is trivial from (2.32b). This completes the proof.

2.3.2. Controllability and Observability Properties of GHR

Given the LTI system (2.9) specified by the triple (A,B,C), if the system is aggregable, (therefore not observable by Theorem 2.1) then we know the pair (A,C) induces the aggregate which represents (2.9) in its observable subspace. On the other hand, if (2.9) is uncontrollable, by working with the dual, we can apply the chained aggregation procedure to extract the controllable subspace. In this sense, the chained aggregation may be used to obtain minimal realization from a given system representation. This is summarized by the following theorem.

Theorem 2.5. Let M and N be the observability and controllability matrices of (A,C) and (A,B) respectively. Given LTI (2.9), if rank $MN = \ell < n$, then (2.9) is equivalent to the representation

$$\begin{bmatrix} \dot{z}_a \\ \dot{z}_b \\ \dot{z}_c \end{bmatrix} = \left[\begin{array}{cc:c} F_{aa} & F_{ab} & 0 \\ \hdashline 0 & F_{bb} & 0 \\ \hdashline & F_{abc} & F_{cc} \end{array}\right] \begin{bmatrix} z_a \\ z_b \\ z_c \end{bmatrix} + \begin{bmatrix} I_m \\ 0 \\ 0 \\ G_c \end{bmatrix} u \tag{2.45a}$$

$$y = [D_a \quad D_b \quad 0] \begin{bmatrix} z_a \\ z_b \\ z_c \end{bmatrix} \tag{2.45b}$$

Futhermore, the ℓ-dimensional system

$$\dot{z}_a = F_{aa} z_a + \begin{bmatrix} I_m \\ 0 \end{bmatrix} u \tag{2.46a}$$

$$y = D_a z_a \tag{2.46b}$$

is the minimal realization of (2.9).

Proof: Given (2.9), since (A,C) is unobservable, by Theorem 2.1. there is a GHR transformation H_1 that transforms (2.9) into

$$\dot{\hat{x}} = \begin{bmatrix} \bar{A}^T & 0 \\ F_{abc} & F_{cc} \end{bmatrix} \hat{x} + \begin{bmatrix} \bar{C}^T \\ G_c \end{bmatrix} u \quad , \hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} . \tag{2.47a}$$

$$y = [I_{r_1} \quad 0]\hat{x} \tag{2.47b}$$

The aggregate

$$\dot{\hat{x}}_1 = \bar{A}^T \hat{x}_1 + \bar{C}^T u \tag{2.48a}$$

$$y = [I_{r_1} \quad 0]\hat{x}_1 = \bar{B}^T \hat{x}_1 \tag{2.48b}$$

is observable and uncontrollable in general. Hence its dual

$$\dot{w} = \bar{A}w + \bar{B}v \tag{2.49a}$$

$$\eta = \bar{C}w, \quad \eta \in R^m, \ w \in R^p \tag{2.49b}$$

is controllable and unobservable. Applying a second GHR transformation H_2 on (2.49), we get

$$\dot{\hat{w}} = \begin{bmatrix} F_{aa}^T & 0 \\ F_{ab}^T & F_{bb}^T \end{bmatrix} \hat{w} + \begin{bmatrix} D_a^T \\ D_b^T \end{bmatrix} v \tag{2.50a}$$

$$\eta = [I_m \quad 0]\hat{w} \tag{2.50b}$$

where $F_{aa}^T \in R^{\ell \times \ell}$ and

$$\dot{\hat{w}}_a = F_{aa}^T \hat{w}_a + D_a^T v \tag{2.51a}$$

$$\eta = [I_m \quad 0]\hat{w}_a \tag{2.51b}$$

is both controllable and observable since (2.50) is controllable. Hence (2.48) is linearly equivalent* to the dual of (2.50), i.e.

$$\begin{bmatrix} \dot{z}_a \\ \cdots \\ \dot{z}_b \end{bmatrix} = \left[\begin{array}{c:c} F_{aa} & F_{ab} \\ \hdashline 0 & F_{bb} \end{array}\right] + \begin{bmatrix} z_a \\ \cdots \\ z_b \end{bmatrix} + \begin{bmatrix} I_m \\ 0 \\ \cdots \\ 0 \end{bmatrix} \tag{2.52a}$$

$$y = [D_a \quad D_b] \begin{bmatrix} z_a \\ z_b \end{bmatrix} \tag{2.52b}$$

Augmenting the residual of (2.47) and (2.52), (2.45) results. The observability and controllability of (2.46) follows.

Next we examine the observability and controllability properties of the aggregate subsystem S^i when the GHR is arbitrarily decomposed into an aggregated subsystem S^i and a residual subsystem $\bar{S}^i$ as defined in (2.21) - (2.23).

<u>Lemma 2.1</u>. Consider a linear time-invariant system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \tag{2.53}$$

$$y = [C_1 \quad C_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

*[42] Two representation are linearly equivalent if the states are related by a linear transformation.

where (A, B) and (C, A) are controllable and observable respectively. Then $\{A_{11},[A_{12}\ \ B_1]\}$ and $\{A_{22},[A_{21}\ \ B_2]\}$ are completely controllable and $\left\{\begin{bmatrix} A_{21} \\ C \end{bmatrix}, A_{11}\right\}$ and $\left\{\begin{bmatrix} A_{12} \\ C \end{bmatrix}, A_{22}\right\}$ are completely observable.

Proof: (A,B) is completely controllable if and only if

$$\text{rank}\begin{bmatrix} A_{11}-\lambda I & A_{12} & B_1 \\ & & \\ A_{21} & A_{22}-\lambda I & B_2 \end{bmatrix} = n_1 + n_2 = n \quad \forall\lambda \in C$$

which implies

$$\left.\begin{aligned} &\text{rank}\,[A_{11}-\lambda I \ \vdots\ A_{12}\ \ B_1] = n_1 \\ &\text{rank}\,[A_{22}-\lambda I \ \vdots\ A_{21}\ \ B_2] = n_2 \end{aligned}\right\} \quad \forall\lambda \in C$$

Observability is established similarly.

If the system (2.53) is viewed as a composite system consisting of two interacting subsystems, Lemma 2.1 states that if the composite is completely controllable, then the individual subsystems are controllable by the joint influence of both interaction and external control u. However, we cannot claim the controllability of each subsystem by either the respective interaction or controls alone. Recalling the special structure of the GHR, it follows that the aggregated subsystem S^i is always controllable through both external control and interactions from the residual subsystem while the residual $\bar{S}^i$ is always observable through both the available outputs and the interaction from the residual system that enters the aggregate. It is expected that the influence of the residual $\bar{S}^i$ on the aggregate S^i in a given decomposition of the GHR is not solely dependent on the magnitude of the interconnection, i.e. $\|E^i\|$, but depends significantly on the controllability of the aggregate through the interconnection $F_{i,i+1}z_{i+1}$ as well as on the observability of the residual through this interconnection. These considerations prompt us to examine the controllability and observability properties of subsystems of the GHR.

<u>Definition 2.4</u>: The subsystem $\bar{S}^i$ is said to be <u>interaction observable</u> by S^i if $(E^i,\bar{F}^i)$ is an observable pair.

<u>Theorem 2.6</u>: Consider the Generalized Hessenberg Representation. $\bar{S}^i$ is interaction observable by S^i if and only if the blocks $F_{j,j+1} \neq 0 \ \forall j \geq i$.

Proof. Form

$$N = \begin{bmatrix} F_{11}-\lambda I & F_{12} & 0 & \cdots\cdots & 0 & 0 \\ F_{21} & F_{22}-\lambda I & F_{23} & 0 \cdots\cdots & & 0 \\ F_{31} & F_{32} & F_{33}-\lambda I & & & \\ \vdots & & & \ddots & & \\ \vdots & & & & F_{k-1,k-1}-\lambda I & F_{k-1,k} \\ F_{k1} & & & & F_{k,k-1} & F_{kk}-\lambda I \\ I_1 & 0 & \cdots\cdots & 0 & 0 & 0 \end{bmatrix}$$

For $\bar{S}^{k-1}$ to be observable by S^{k-1} through $F_{k-1,k}z_k$, we must have

$$\operatorname{rank}\begin{bmatrix} F_{kk}-\lambda I \\ F_{k-1,k} \end{bmatrix} = r_k = \operatorname{rank}\begin{bmatrix} F_{k-1,k} \\ F_{kk}-\lambda I \end{bmatrix} \quad \forall \lambda \in C.$$

But this is exactly the last r_k columns of N in (2.54). By construction, $F_{k-1,k}$ has full rank r_k. Hence z_k is interaction observable by S^{k-1}. Next consider the interaction observability of $\bar{S}^{k-2}$ by S^{k-2}, we have

$$\text{rank}\begin{bmatrix} F_{k-1,k-1}-\lambda I & F_{k-1,k} \\ F_{k,k-1} & F_{kk}-\lambda I \\ F_{k-2,k-1} & 0 \end{bmatrix} = \text{rank}\begin{bmatrix} F_{k-2,k-1} & 0 \\ F_{k-1,k-1}-\lambda I & F_{k-1,k} \\ F_{k,k-1} & F_{kk}-\lambda I \end{bmatrix} = r_{k-1}+r_k$$

$$\forall \lambda \in C$$

By construction, because each of $F_{k-2,k-1}$ and $F_{k-1,k}$ are of rank r_{k-1} and r_k respectively, then $\bar{S}^{k-2}$ is interaction observable by S^{k-2}. Moving from lower right hand corner, and proceeding upwards, we establish that all residual subsystem $\bar{S}^i$ are interaction observable by S^i for any partition $i = 1,2,...k-1$ if $F_{j,j+1} \neq 0$ and $F_{j,j+1}$ has maximal rank, $j \geq i$. But this later condition is guaranteed by the chained aggregation procedure. Furthermore, note as we proceed from the lower right hand corner of (2.54), the blocks proceeding i for the S^i under consideration do not enter as conditions for interaction observability. Thus only the blocks $F_{j,j+1}$ for $j \geq i$ need to be considered.

Convesely if $\bar{S}^i$ is interaction observable by S^i then the matrix N_i has full column rank where

$$N_i = \begin{bmatrix} F_{i,i+1} & 0 & \cdots\cdots & 0 \\ F_{i+1,i+1}-\lambda I & F_{i+1,i+2} & 0 \cdots\cdots & 0 \\ F_{i+2,i+1} & F_{i+2,i+2}-\lambda I & \cdots\cdots\cdots & 0 \\ \vdots & \vdots & F_{k-1,k-1}-\lambda I & F_{k-1,k} \\ F_{k,i+1} & F_{k,i+2} & \cdots\cdots & F_{k,k}-\lambda I \end{bmatrix}$$

$$= [\Gamma_{i+1}\ \Gamma_{i+2} \cdots \Gamma_k]$$

i.e. rank $N_i = \sum_{j=i+1}^{k} r_j \ \ \forall \lambda \in C$ which implies Γ_j, $j = i+1,\ldots k$ must have maximal column rank. For $j = k$, we have

$$\text{rank} \begin{bmatrix} F_{k-1,k} \\ \\ F_{k,k}-\lambda I \end{bmatrix} = r_k \quad \forall \lambda \in C.$$

Since for $\lambda \in \sigma\,(F_{k,k})$,

$$\text{rank} \begin{bmatrix} F_{k-1,k} \\ \\ F_{k,k}-\lambda I \end{bmatrix} = \text{rank} \begin{bmatrix} F_{k-1,k} \\ \\ 0 \end{bmatrix} = r_k$$

Thus $F_{k-1,k} \neq 0$. Proceeding inductively, we see that $F_{j,j+1} \neq 0 \ \ \forall j=i,\ldots.k$. This completes the proof.

<u>Remark</u>: Theorem 2.6 is applicable whether $F_{i,i+1}$ is scalar or is a block matrix. This means the dynamical behavior $\bar{S}^i$ is <u>always</u> observable through interaction $F_{i,i+1}z_{i+1}$ by S^i and consequently $\bar{S}^i$ will always affect the

dynamics of the aggregate subsystem S^i.

Next we consider the interaction controllability.

<u>Definition 2.5</u>: A subsystem S^i is said to be <u>interaction controllable</u> by by $\bar{S}^i$ if (F^i, E^i) is a controllable pair.

<u>Theorem 2.7</u>. Given the <u>scalar</u> Hessenberg system representation, then for any partition into S^i and $\bar{S}^i$, S^i is interaction controllable by $\bar{S}^i$ if and only if $f_{j,j+1} \neq 0$ for all $j \leq i$.

Proof: Follows similarily the proof of Theorem 2.6, and is omitted.

The extension to the matrix case of interaction controllability results of the scalar Hessenburg Representation is not possible. In general, we cannot guarantee the interaction controllability of S^i by $\bar{S}^i$ even with the assumption of complete composite controllability by control. It is possible, however, to characterize the increase in the dimension of the subspace controlled by the interaction when the aggregated subsystem is extended from S^{i-1} to S^i. Defining rank $[F^i - \lambda I \;\; E^i] = \sigma_i$, $\forall \lambda \in C$, we have:

<u>Theorem 2.8</u>: The increase in the dimension of the subspace controlled by the interaction when the aggregate subsystem is expanded from S^{i-1} to S^i is bounded above r_i and below by r_{i+1}, i.e., $r_{i+1} \leq \sigma_i - \sigma_{i-1} \leq r_i$, $i=1,2,\ldots,k$.

Proof: Consider the rank condition

$$\operatorname{rank}\,[F_{11} - \lambda I \;\; F_{12}] = \sigma_1 \quad \forall \lambda \in C \tag{2.55}$$

and

$$\operatorname{rank}\begin{bmatrix} F_{11}-\lambda I & F_{12} & 0 \\ F_{21} & F_{22}-\lambda I & F_{23} \end{bmatrix} = \sigma_2 \quad \forall \lambda \in C .$$

Since rank $[F_{23}] = r_3$, combining with (2.55) it is obvious

$$\sigma_1 + r_3 \leq \sigma_2 \leq \sigma_1 + r_2.$$

Proceeding in a similar manner we have

$$\text{rank} \begin{bmatrix} F^{i-1}-\lambda I & E^{i-1} & 0 \\ & & \\ F_{i1} \cdots F_{ii-1} & F_{ii}-\lambda I & F_{ii+1} \end{bmatrix} = \sigma_i, \quad \forall \lambda \in C$$

Then it is obvious

$$\sigma_{i-1} + r_{i+1} \leq \sigma_i \leq \sigma_{i-1} + r_i$$

$$r_{i+1} \leq \sigma_i - \sigma_{i-1} \leq r_i.$$

2.3.3. The GHR and Its Hessenberg Blocks - The Effect on Output Trajectories

For a GHR arbitrarily decomposed into S^i and $\bar{S}^i$, S^i is coupled to $\bar{S}^i$ only by the Hessenberg block $F_{i,i+1}$. Thus it is of interest to study the influence of the Hessenberg blocks on S^i and $\bar{S}^i$. The present section will characterize the trajectory deviations when a particular interconnection $F_{i,i+1}z_{i+1}$ is neglected from S^i. In the next section, the influence of the Hessenberg blocks will be assessed from an eigenvalue perturbation point of view if $F_{i,i+1}z_{i+1}$ is severed from S^i.

We recall that when (A,C) of (2.9) is observable (and thus non-aggregable), none of the $F_{i,i+1}$ in the GHR will be zero and thus no decoupling of the aggregate from the residual exists. On the other hand, when (A,C) is unobservable, this decoupling exists. Lying between these two extremes is the case when the residual is only weakly observable at the aggregate through $F_{i,i+1}z_{i+1}$. By weakly observable we mean that the dynamic of the residual has a weak effect on the dynamics of the aggregate. If the influence of $F_{i,i+1}z_{i+1}$ on S^i is sufficiently small, then neglecting such interconnection results the aggregate (2.56a) and (2.56c) which may be viewed as a reduced order model of the full order system (2.9). We now wish to establish some bounds on the effect of coupling term $F_{i,i+1}z_{i+1}$. Specifically, we ask how well does:

$$\dot{\hat{z}}^i = F^i\hat{z}^i + G^i u \tag{2.56a}$$

$$\dot{\hat{\bar{z}}}^i = H^i\hat{z}^i + \bar{F}^i\hat{\bar{z}}^i + \bar{G}^i u \tag{2.56b}$$

$$\hat{y} = [I_1 \quad 0 \ \vdots \ 0] \begin{bmatrix} \hat{z}^i \\ \hat{\bar{z}}^i \end{bmatrix} \tag{2.56c}$$

approximate the actual GHR

$$\dot{z}^i = F^i z^i + E^i\bar{z}^i + G^i u \tag{2.57a}$$

$$\dot{\bar{z}}^i = H^i z^i + \bar{F}^i\bar{z}^i + \bar{G}^i u \tag{2.57b}$$

$$y = [I_1 \quad 0 \ \vdots \ 0] \begin{bmatrix} z^i \\ \bar{z}^i \end{bmatrix} \tag{2.57c}$$

where $i \in \{1,2,\ldots k-1\}$.

Since k-1 possible approximate GHR (2.56) exist by neglecting any one of the coupling $F_{i,i+1}$ in (2.57). It becomes necessary to rank each approximation. Since y and $\hat{y}$ are subvectors of z^i and $\hat{z}^i$ respectively, in order to best approximate $y(t)$ by $\hat{y}(t)\,\forall t$, it is sufficient to keep the growth of $e_i(t) = z^i(t) - \hat{z}^i(t)$ small. We ask specifically:

- How does $e_i(t)$ evolve in general?
- What is the asymtotic behavior of $e_i(t)$?
- Which interconnection $F_{i,i+1}z_{i+1}, i=1,\dots k-1$ results in minimum growth of $\|e_i(t)\| \forall t$ where $\|e_i(t)\| \equiv \{\mathrm{tr}\ e_i^T(t)e_i(t)\}^{1/2}$?

Instead of asking how well (2.56) approximates (2.57) when $u = 0$, let us assume controls are non-zero, but bounded, i.e.

$$u \in U = \{u(t) \,|\, \|u(t)\| \le K < \infty \ \forall t\} \tag{2.58}$$

Define

$$e_i(t) = z^i(t) - \hat{z}^i(t) \tag{2.59}$$

Theorem 2.9 and Corollary 2.9 characterize the evolution of and asymtotic behavior of $\|e_i(t)\|$. The proofs are given in Appendix A.

<u>Theorem 2.9</u>: Consider the approximation of (2.57) by (2.56). If there exists constants α_i, ρ_i, $i = 1,2$, then $e_i(t)$ satisfies:

$$\|e_i(t)\| \le \frac{k_3\alpha}{\alpha+\mu} + (k_1' + K_2)e^{\alpha t} + \left[\frac{(k_1' + k_2 + k_3)\mu^2 + (k_1'\alpha + k_2\alpha + k_1'')\mu + k_1''\alpha}{\mu(\alpha+\mu)}\right] e^{(\alpha+\mu)t} \tag{2.60}$$

where α_i, ρ_i, k_i, α and μ are defined in Appendix A.

<u>Corollary 2.9</u>: Assume F^i and $\bar{F}^i$ are stable matrices under the hypothesis of Theorem 2.9, then

$$0 \le \lim_{t \to \infty} \|e_i(t)\| \le \frac{\alpha k_3}{\alpha+\mu} < \infty \tag{2.61}$$

if and only if $\alpha+\mu<0$. (α,μ and k_3 are defined in Appendix A).

<u>Remark</u>: From the definition of k_3, we see if u belongs to the class of controls that exponentially approaches zero, then $\|e(t)\|$ also approaches zero as $t\to\infty$.

Now for a given decomposition of the system into S^i and $\bar{S}^i$, any one of $F_{i,i+1}$ may be neglected in order to approximate (2.57) by (2.56), resulting the approximate GHR and its associated instantaneous error norm $\|e_i(t)\|$. Because of the fact that y and $\hat{y}$ are subvectors of z^i and $\hat{z}^i$ small growth in $e_i(t)$ warrants the conclusion that the growth of $y(t)-\hat{y}(t)$ is also small. An essential exponential in (2.60) that indicates the growth of $e_i(t)$ is $e^{(\alpha^i+\mu^i)t}$ where the quantities α_1, α_2, α and μ are appropiately indexed as α_1^i, α_2^i, α^i and μ^i, $i\in\{1,2,...k-1\}$ according to partitions in the GHR. We observe three possibilities:

<u>Possiblity (a)</u>:

F^i, $\bar{F}^i$ stable	=>	$\alpha^i<0$ $(\alpha^i+\mu^i)\leq 0$ possible	=>	for minimum instantaneous norm $\|e_i(t)\|$ or small growth in $\|e_i(t)\|$, we desire a small $(\alpha^i+\mu^i)$.

<u>Possibility (b)</u>:

F^i, $\bar{F}^i$ unstable	=>	$\alpha^i>0$ $(\alpha^i+\mu_i)\geq 0$ always	=>	for minimum growth on $\|e_i(t)\|$, we desire a small $(\alpha^i+\mu^i)$.

<u>Possibility (c)</u>:

F^i stable, $\bar{F}^i$ unstable or vice versa	=>	$\alpha^i>0$ $(\alpha^i+\mu^i)\geq 0$ always	=>	for minimum growth in $\|e_i(t)\|$, we desire a small $(\alpha^i+\mu^i)$.

Thus the requirement to keep the growth of $\|e_i(t)\|$ small can be related as keeping the value of $(\alpha^i + \mu^i)$ as small as possible. Such conclusion can therefore be interpreted as a criteria for choosing $i \in \{1,\ldots k-1\} = \mathcal{K}$ such that the influence of Hessenberg blocks $F_{i,i+1}$ on introducing small error between (2.56) and (2.57) may be ranked. Define

$$\beta_i = \alpha^i + \mu^i, \quad i = 1,\ldots k-1 \tag{2.62}$$

Then of the k-1 approximation (2.56), the interconnection $F_{j,j+1}z_{j+1}$, if neglected, where j is defined

$$\beta_j = \min_{i \in \mathcal{K}} \{\beta_i = \alpha^i + \mu^i\}$$

may be accepted as resulting in smallest growth in the error $\|e_i(t)\|$. The rest of the interconnections may be ranked accordingly.

2.3.4. The GHR and Its Hessenberg Blocks - An Eigenvalue Perturbation Analysis

Section 2.3.3 examines the influence of the Hessenberg blocks $F_{i,i+1}$ on the aggregate subsystem S^i in terms of error introduced between the exact and approximate trajectories. In this section, we shall examine the influence of neglecting $F_{i,i+1}$ in terms of perturbation of open-loop eigenvalues. The motivation to study eigenvalue perturbation lies in the fact that (2.56a) - (2.56c) results from a perturbation of the full system (2.9) if the coupling $F_{i,i+1}z_{i+1}$ is neglected. Open loop behavior of a system trajectory is in general related to the eigenvalues that are retained or approximated in the reduced order model. Since the input-output behavior is our major concern we note that, if the Hessenberg block $F_{i,i+1}$ is neglected, the input-output behavior will be solely governed by the aggregate S^i (see section 2.3.2). Denote the system matrices of (2.56) and (2.57) with index i suppressed:

$$A_c = \begin{bmatrix} F & E \\ H & \bar{F} \end{bmatrix} \qquad \hat{A}_c = \begin{bmatrix} F & 0 \\ H & \bar{F} \end{bmatrix} \tag{2.63}$$

and let $\sigma(A)$ denote the set of eigenvalues of A. If $\sigma(\hat{A}_c)$ and $\sigma(A_c)$ have similar eigenvalue pattern or eigenvalues that are close in numerical values, then the dynamics of $\hat{y}(t)$ can be expected to be close to that of $y(t)$. In fact, due to the structure of the GHR (S^i is coupled to $\bar{S}^i$ through only $F_{i,i+1}z_{i+1}$), the closeness of $\hat{y}$ to y may be explained by the idea of weak coupling when eigenvalue separation exists between $\sigma(F)$ and $\sigma(\bar{F})$[22]. Since this need not be the case, we shall study how the *eigenvalues of* $\hat{A}_c$ *differ from those of* A_c and to bound the eigenvalue of A_c in terms of a Gorschgorin disc [23] if this is possible.

Let λ_i, $\hat{\lambda}_i$ denote the i-th eigenvalue of A_c and $\hat{A}_c$ respectively. Also let x_i, y_i be the normalized eigenvector and adjoint eigenvector of $\hat{A}_c$ corresponding to $\hat{\lambda}_i$. X is defined as the modal matrix of $\hat{A}_c$ and $Y = X^{-1}$. Furthermore, $\Delta A_c = A_c - \hat{A}_c$ where

$$\Delta A_c = \begin{bmatrix} 0 & E \\ 0 & 0 \end{bmatrix} \tag{2.64}$$

For simplicity, assume A_c and $\hat{A}_c$ each have distinct eigenvalues. We wish to bound $|\lambda_i - \hat{\lambda}_i|$ in terms of the perturbation ΔA_c. The following lemmas are necessary.

Lemma 2.2. Given $n \times n$ matrix $A = [a_{i,j}]$, $i,j = 1,\ldots n$. Every eigenvalue of A lies in at least one of the discs (Gorschgorin discs) with center a_{ii} and radius $\sum_{j \neq i} |a_{ij}|$.

Lemma 2.3. [23] If s of the Gorschgorin discs form a connected set which is isolated from other discs, then there are precisely s eigenvalues of A within this connected set.

Base on Lemma 2.3, the objective is to isolate the Gorschgorin disc in which λ_i lies from the rest by using linear transformation on A_c, thus reducing the radius of each disc. Using the definitions of various quantities given in Appendix B, we have the following theorem.

Theorem 2.10: Consider the matrices A_c and $\hat{A}_c$ defined by (2.64), then the i-th Gorschgorin disc of A_c can be isolated from the rest so that

$$\hat{\lambda}_i + \beta_{ii} - r_i \leq \lambda_i \leq \hat{\lambda}_i + \beta_{ii} + r_i \tag{2.65}$$

where $\beta_{ij} \equiv y_i^T \Delta A_c x_i$, $i, j = 1,\ldots n$ if and only if the following conditions hold:

$$\text{(I)} \quad \underline{\rho}_i = \max_{\substack{j \\ j \neq i}} \frac{\sum_{\substack{k \neq j \\ k \neq i}} |\beta_{jk}|}{|(\hat{\lambda}_i + \beta_{ii}) - (\hat{\lambda}_j + \beta_{jj})|} < 1$$

$$\text{(II)} \quad \underline{\gamma}_i \leq \overline{\gamma}_i$$

where β_{ij}, $\underline{\gamma}_i$ and $\overline{\gamma}_i$ are defined in Appendix B.

The proof of Theorem 2.10 is given in Appendix B. An example will be given at the end of the section to illustrate the theorem.

Finally, by treating ΔA_c as a perturbation on A_c, we note one may study the sensitivity of eigenvalues of A_c with respect to ΔA_c.

Definition 2.6: Let the first order sensitivity matrix of λ_k with respect to ΔA_c be defined as

$$ {}_xS_k^{(1)} \equiv \frac{\partial \lambda_k}{\partial E} = \left[\frac{\partial \lambda_k}{\partial e_{ij}} \right] , \quad E = [e_{i,j}] \tag{2.66} $$

and the total first order change of $\lambda_k(A_c)$ with respect to ΔA_c be defined as

$$ \Delta \lambda_k^{(1)} \equiv \operatorname{tr} E^T \, {}_xS_k^{(1)} \tag{2.67} $$

To see the validity of (2.67), let $E = [e_{i,j}]$ be a 2×2 matrix, then

$$ \Delta \lambda_k^{(1)} = \operatorname{tr} E^T \, {}_xS_k^{(1)} $$

$$ = e_{11} \frac{\partial \lambda_k}{\partial e_{11}} + e_{12} \frac{\partial \lambda_k}{\partial e_{12}} + e_{21} \frac{\partial \lambda_k}{\partial e_{21}} + e_{22} \frac{\partial \lambda_k}{\partial e_{22}} . $$

Remark: The left subscript x in ${}_xS_k^{(1)}$ denotes the sensitivity given a particular matrix representation. This is important because sensitivities are not invariant to linear transformations on matrix representations.

Applying the results on eigenvalue sensitivity with respect to a parameter [24], we have

$$ A_c x_k = \lambda_k x_k $$

Denote $E = [E_{ij}]$, then

$$ A_c \frac{\partial x_k}{\partial e_{ij}} + \frac{\partial A_c}{\partial e_{ij}} x_k = \lambda_k \frac{\partial x_k}{\partial e_{ij}} + \frac{\partial \lambda_k}{\partial e_{ij}} x_k $$

which gives

$$\frac{\partial \lambda_k}{\partial e_{ij}} \equiv \frac{< \frac{\partial A_c}{\partial e_{ij}} x_k, y_k >}{< x_k, y_k >} \tag{2.68}$$

where the right hand side of (2.68) is evaluated according to $\hat{A}_c$. Hence to first order:

$$\lambda_k = \hat{\lambda}_k + \Delta \lambda_k^{(1)}$$

We note the computation of the sensitivity matrix is relatively simple because the eigenvector and adjoint eigenvector x_i, y_i of $\hat{A}_c$ may be obtained from lower order computations due to the lower block triangular structure of $\hat{A}_c$.

The following simple example illustrates Theorem 2.10 and the eigenvalue sensitivity with respect to ΔA_c.

Example. Consider a scalar Hessenberg matrix A_c and an approximate $\hat{A}_c$ where,

$$A_c = \begin{bmatrix} -2.5 & 1.2 & 0 & 0 \\ 1.0 & 1.5 & 0.1 & 0 \\ 1.4 & 2.2 & 0.62 & -3.43 \\ 0.6 & 1.6 & 0.96 & -4.3 \end{bmatrix}, \quad \hat{A}_c = \begin{bmatrix} -2.5 & 1.2 & 0 & 0 \\ 1.0 & 1.5 & 0 & 0 \\ 1.4 & 2.2 & 0.62 & -3.43 \\ 0.6 & 1.6 & 0.96 & -4.3 \end{bmatrix}.$$

First let us consider the bounding of λ_i using the Gorschgorin disc. The matrix β is computed to be

$$\beta = \begin{bmatrix} 0.008009 & -0.0180304 & 0.0180445 & -0.02747 \\ -0.039071 & 0.08796 & -0.088028 & 0.13401 \\ -0.00915 & 0.020603 & -0.020619 & 0.031389 \\ 0.021968 & -0.049455 & 0.049494 & -0.075348 \end{bmatrix}$$

According to (B8) of Appendix B, the lower bounds for ρ_i are $\rho_1 = 0.069$ $\rho_2 = 0.033$, $\rho_3 = 0.061$ and $\rho_4 = 0.059$. Choose $\rho_i = 0.5\ \forall i$, it is trivial to show ϵ_i exists for $\forall i$ and $r_1 = 0.00263$, $r_2 = 0.01306$, $r_3 = 0.00335$ and $r_4 = 0.01734$. The results are tabulated in Table 2.1.

Table 2.1. Estimates of $\lambda_i(A_c)$ based on $\hat{\lambda}_i(\hat{A}_c)$ and the bounds of Gorschgorin discs.

Eigenvalues of $A_c : \lambda_i$	Eigenvalues of $\hat{A}_c : \hat{\lambda}_i$	Estimated Bound on λ_i
-2.7725	-2.7804	$-2.77502 \leq \lambda_1 \leq -2.7697$
1.865	1.7804	$1.855 \leq \lambda_2 \leq 1.881$
-3.5215	-3.501	$-3.5251 \leq \lambda_3 \leq -3.5183$
-0.25104	-0.17904	$-0.271 \leq \lambda_4 \leq -0.237$

Next consider the first order eigenvalue sensitivities with respect to $f_{23} = 0.1$ in A_c. The sensitivity matrices of λ_i, $i = 1,2,3,4$ are defined according to (2.66) and (2.68) while the total first order variation is computed according to (2.67).

$$\frac{\partial \lambda_1}{\partial E} = \begin{bmatrix} -0.285681 & -0.040774 \\ 0.080091 & 0.011431 \end{bmatrix}, \quad \frac{\partial \lambda_2}{\partial E} = \begin{bmatrix} 0.20549 & 0.096208 \\ 0.879594 & 0.411818 \end{bmatrix}$$

$$\frac{\partial \lambda_3}{\partial E} = \begin{bmatrix} -0.076532 & -0.091946 \\ -0.206189 & -0.247716 \end{bmatrix} \quad \frac{\partial \lambda_4}{\partial E} = \begin{bmatrix} 0.156713 & 0.036507 \\ -0.753484 & -0.175529 \end{bmatrix}$$

and

$$\Delta \lambda_1^{(1)} = 0.0080091 \quad , \quad \Delta \lambda_2^{(1)} = 0.0879594$$

$$\Delta \lambda_3^{(1)} = -0.0206189 \ , \quad \Delta \lambda_4^{(1)} = -0.0753484$$

The computed results and comparisons are tabulated in Table 2.2

Table 2.2. Eigenvalue estimate of A_c based on eigenvalues of $\hat{A}_c$ using 1st order sensitivities.

Eigenvalues of of A_c : λ_i	Eigenvalues of $\hat{A}_c$: $\hat{\lambda}_i$	Estimated λ_i	% error $= \left\lvert \frac{\lambda_i - \hat{\lambda}_i}{\lambda_i} \right\rvert$
-2.7725	-2.7804	-2.77239	0.003%
1.865	1.7804	1.8684	0.18%
-3.5215	-3.501	-3.5216	0.003%
-0.25104	-0.17904	-0.2544	1.33%

CHAPTER 3

THE GENERALIZED QL ALGORITHM

The scalar Hessenberg matrix has been used extensively in numerical analysis. It is useful in eigenvalue analysis [23,25-27], particularly in conjunction with the powerful QR algorithms developed by Francis [28]. Extensive treatment is also given by Wilkinson [23]. Working with the upper scalar Hessenberg matrix F, the QR algorithm factors F into an unitary matrix and a upper triangular matrix, and iteratively transform F into a upper triangular matrix, thus displaying the eigenvalues on the diagonal. Taking advantage of the block structure of the generalized Hessenberg matrix, we envision a parallel development for the purpose of exhibiting weak coupling between some aggregate S^i and associated residual $\bar{S}^i$, if such weak coupling indeed exists in the given system. The result is the generalized QL (GQL) algorithm developed in this chapter. We point out that weak coupling, if it exists, is a function of system representation, and is only exhibited in certain particular system representations. The GQL algorithm does not introduce weak coupling, it merely transforms systems into a more suitable representation to bring out weak coupling. The motivation for using the particular generalized QL algorithm is found in the fact that its successful application to the GHR would have the effect of reducing <u>in the limit</u>, the plant matrix in the GHR from generalized Hessenberg form to block triangular form. In that process, some Hessenberg blocks will be converging to zero matrices. Depending on the rate of convergence, the algorithm is expected to place in evidence after a finite number of iterations those Hessenberg blocks

$F_{i,i+1}$ that may be neglected with acceptable approximation error. The main objective is to understand under what conditions convergence occurs and how the algorithm may be applied to bring out any weak coupling within the system structure.

The generalized QL algorithm to be developed is a particular variant of the standard QR algorithm [23,28] and is constructed for the block partitioned Hessenberg matrices. At each iteration, the system matrix in the GHR is factored into the product of a block unitary matrix Q and a lower block triangular matrix L, each compatibly partitioned as the generalized Hessenberg system matrix. The QL factorizations stem from the needs of reduced order modeling to be described later while the block structure is consistent with the GHR structure.

The chapter is organized as follows. In Section 3.1, the fundamental QR algorithm is modified to obtain the basic QL algorithm. The extension to block Hessenberg matrix results in the generalized QL (GQL) algorithm which is presented in Section 3.2. The extension to block matrices leads to a whole class of problems which must be investigated. In particular, the convergence and computational aspects of the generalized QL algorithm are presented in Sections 3.3 and 3.4 respectively. Finally, the application of the generalized QL algorithm to GHR is briefly discussed in Section 3.5, leaving the model reduction problem via the GHR and GQL in the next chapter.

3.1. The Basic QL Algorithm and its Properties

Let A be an arbitrary but real $n \times n$ matrix, Q a $n \times n$ unitary matrix, and L a $n \times n$ lower triangular matrix. Then:

Result 1: The factorization $A = QL$ is always possible.

Result 2: The factorization $A = QL$ is unique if A is nonsingular and the diagonal elements of L are restricted to be positive and real.

Let $A_1 = A$, then the QL algorithm is defined by the following ierative procedure:

$$A_s = Q_s L_s \qquad s = 1,2,3,\ldots \tag{3.1}$$

$$A_{s+1} = L_s Q_s \qquad s = 1,2,3,\ldots \tag{3.2}$$

where the factor matrices Q_s and L_s are unitary and lower triangular respectively *and subscript s denotes iteration number. Then*

Result 3: All A_s are similar to each other, and, in particular, are similar to A_1.

Result 4: Let $P_s \equiv Q_1 Q_2 \cdots Q_s$, $U_s \equiv L_s L_{s-1} \cdots L_1$. Then

(i) $$A_{s+1} = Q_s^T \cdots Q_1^T A_1 Q_1 \cdots Q_s = P_s^T A_1 P_s \tag{3.3}$$

(ii) $$P_s U_s = A_1^s. \tag{3.4}$$

Thus $P_s U_s$ achieves a QL decomposition of A_1^s.

Result 5: Given a $n \times n$ matrix A_1 with real eigenvalues. Then the QL algorithm converges to a lower triangular matrix $\bar{A}$ with eigenvalues of A_1 on the diagonal in ascending magnitude, that is,

$$\operatorname{diag} \bar{A} = (\lambda_1, \ldots, \lambda_n)$$

with

$$|\lambda_1| < |\lambda_2| < \cdots |\lambda_n|.$$

<u>Result 6</u>: Given A with distinct real or complex-conjugate pairs of eigenvalues, the eigenvalues in the limiting matrix $\bar{A}$ are again ordered with respect to their absolute value as in Result 5, with the distinction that 2×2 diagonal block entries correspond to complex-conjugate pairs of eigenvalues.

3.2. The Generalized QL Algorithm and its Properties

As in the case of QR algorithm, the QL algorithm presented in Section 3.1 is applicable to any arbitrary initial matrix. It has been noted in [28] that the amount of computation is significantly reduced if the initial matrix is Hessenberg. Instead of concentrating the development on arbitrary block matrices, a generalization of the QL algorithm suitable for application to the generalized Hessenberg matrix of the GHR is now developed. Consider the system matrix F of the GHR (2.10). For notational purposes denote F by F_1. The factor matrices Q_i, L_i, $i=1,2,\ldots$ are said to be compatible with F_1 if they have the same block partitions as F_1.

The generalized QL algorithm states

$$F_s = Q_s L_s \qquad s=1,2,3,\ldots \tag{3.5}$$

$$F_{s+1} = L_s Q_s \qquad s=1,2,3,\ldots \tag{3.6}$$

where

Q_s = $n\times n$ block unitary matrix compatible with F_1

L_s = $n\times n$ lower block triangular matrix, compatible with F_1.

The following results then hold:

<u>Theorem 3.1</u>: Given a matrix F in generalized Hessenberg form, the factorization $F = QL$ is always possible where Q is unitary and L is lower block triangular and each compatible with F.

<u>Proof</u>: The theorem will be proved by a constructive procedure referred to as the <u>successive orthogonalization method</u>. Since $F = QL$ and using the notation

$$[F_1 \ \cdots \ F_k] = [Q_1 \ \cdots \ Q_k] \begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & \vdots \\ \vdots & & \ddots \\ L_{k1} & \cdots & L_{kk} \end{bmatrix} \tag{3.7}$$

where F_i, Q_i are $n \times r_i$ matrices and L_{ij} are $r_i \times r_j$ matrices. We have

$$\begin{aligned} F_k &= Q_k L_{kk} \\ F_{k-1} &= Q_{k-1} L_{k-1,k-1} + Q_k L_{k,k-1} \\ &\vdots \\ F_i &= Q_i L_{ii} + Q_{i+1} L_{i+1,i} + \cdots + Q_k L_{ki} \\ &\vdots \\ F_1 &= Q_1 L_{11} + Q_2 L_{21} + \cdots + Q_k L_{k1} \end{aligned} \tag{3.8}$$

The successive orthogonalization method starts by choosing Q_k such that $Q_k^T Q_k = I_k$.* This is always possible if columns of Q_k are orthonormal vectors. From (3.8a), we immediately have $L_{kk} = Q_k^T F_k$. Next we choose Q_{k-1} such that $Q_{k-1}^T Q_{k-1} = I_{k-1}$ and $Q_{k-1}^T Q_k = 0$. This is always possible if columns of Q_{k-1} are chosen such that each column of Q_{k-1} is orthogonal to each column of Q_k and all columns of Q_{k-1} are mutually orthonormal. From the

---

*Here I_i denotes the $r_i \times r_i$ dimensional identity matrix.

second equation of (3.8) we get $L_{k-1,k-1} = Q_{k-1}^T F_{k-1}$ and $L_{k,k-1} = Q_k^T F_{k-1}$. In general, Q_i, $i = k,k-1,\ldots,1$ are chosen in succession starting from Q_k such that: (a) columns of Q_i are mutually orthonormal and (b) each column of Q_i is orthogonal to each column of Q_j, $i < j \le k$, i.e., $Q_i^T Q_i = I_i$, $Q_i^T Q_j = 0 \;\; \forall j > i$. This is always possible. Once Q_i is selected and since Q_j, $j > i$ are known from previous steps, we have

$$\begin{aligned} L_{ii} &= Q_i^T F_i \\ L_{i+1,i} &= Q_{i+1}^T F_i \\ &\vdots \\ L_{k,i} &= Q_k^T F_i . \end{aligned} \tag{3.9}$$

Thus the matrix L is obtained block column-wise starting from the kth block column. Clearly each L_{ij}, $\forall i,j = 1,2,\ldots,k$ is well defined and with Q_i, $i = 1,2,\ldots,k$ appropriately selected, the generalized QL factorization of an arbitrary generalized Hessenberg matrix always exists.

Analogous to Fact 2, the generalization to block Hessenberg starting matrix is

Theorem 3.2: The decomposition $F = QL$ is unique if F is nonsingular and L is restricted to have positive definite (negative definite) and symmetric diagonal blocks L_{ii} $\forall i$.

Proof: By hypothesis F is nonsingular, thus L is nonsingular with positive definite diagonal blocks $L^{(i,i)} \triangleq L_{ii}$, $i = 1,2,\ldots,k$. Assume that there exists two decompositions such that $F = Q_1 L_1 = Q_2 L_2$. Defining

$$V = Q_2^T Q_1 = L_2 L_1^{-1} \tag{3.10}$$

Q_i being unitary gives

$$V^T V = Q_1^T Q_2 Q_2^T Q_1 = I. \tag{3.11}$$

Thus equating both sides of (3.11) and using (3.10) we have $V = \mathrm{diag}(V_{ii})$, $V_{ii} \in R^{r_i \times r_i}$, with $V_i^T V_{ii} = I_i$, $i = 1,2,\ldots,k$. From (3.10) $V_{ii} = L_2^{(i,i)}[L_1^{(i,i)}]^{-1}$. Hence $L_2^{(i,i)T} L_2^{(i,i)} = L_1^{(i,i)T} L_1^{(i,i)}$. If $L^{(i,i)}$ is restricted to be symmetric and positive definite (or all negative definite), then $L_1^{(i,i)} = L_2^{(i,i)}$ which implies $V_{ii} = I_i$ and it follows that $L_1 = L_2$ and $Q_1 = Q_2$. This proves that if L is restricted to have positive definite and symmetric diagonal blocks, then the QL decomposition is unique. It should be noted that a unique decomposition $F = QL$ can be attained if F is nonsingular and L is restricted to have <u>negative definite</u> and <u>symmetric</u> diagonal blocks $L^{(i,i)}$ for all i. However, there is no loss in generality in taking the $L^{(i,i)}$ to be positive definite.

<u>Theorem 3.3</u> (i): All F_s are similar and thus similar to the initial generalized Hessenberg matrix F_1; (ii) define $P_s = Q_1 \ldots Q_s$ and $U_s = L_s \ldots L_1$. Then $P_s U_s = F_1^s$, i.e., $P_s U_s$ achieves a QL decomposition of F_1^s.

<u>Proof</u>: (i) by definition, $F_s = Q_s L_s$, then

$$Q_s^T F_s Q_s = Q_s^T (Q_s L_s) Q_s = L_s Q_s = F_{s+1}.$$

In particular,

$$F_{s+1} = Q_s^T \cdots Q_1^T F_1 Q_1 \cdots Q_s = P_s^T F_1 P_s.$$

(ii)
$$\begin{aligned} P_s U_s &= Q_1 \cdots Q_{s-1}(Q_s L_s) L_{s-1} \cdots L_1 \\ &= Q_1 \cdots Q_{s-2} Q_{s-1} F_s L_{s-1} L_{s-2} \cdots L_1 \\ &= F_1 Q_1 \cdots Q_{s-2}(Q_{s-1} L_{s-1}) L_{s-2} \cdots L_1 \\ &= F_1^s. \end{aligned}$$

<u>Theorem 3.4</u>: Given an initial generalized Hessenberg matrix F, all QL iterates are structurally invariant, i.e., all F_s are generalized Hessenberg matrices compatibly partitioned as in F.

<u>Proof</u>: If F is nonsingular, the assertion is obvious because L_s will be nonsingular and

$$\begin{aligned} F_{s+1} &= L_s L_{s-1} \cdots L_1 F L_1^{-1} \cdots L_s^{-1} \\ &= U_s F U_s^{-1}. \end{aligned}$$

Since U_s is lower block triangular and compatibly partitioned then F_{s+1}, $\forall s = 1,2,\ldots$ will be generalized Hessenberg and compatibly partitioned. If F_1 is singular the result can be shown to hold by recalling the possibilities of structuring the block matrices in Q via successive orthogonalization. From (3.8) we have $L_{kk} = Q_k^T F_k$ and thus

$$L_{kk} = [Q_{1k}^T \cdots Q_{k-1,k}^T Q_{k,k}^T] \begin{bmatrix} 0 \\ \vdots \\ 0 \\ F_{k-1,k} \\ F_{k,k} \end{bmatrix} = Q_{k-1,k}^T F_{k-1,k} + Q_{k,k}^T F_{k,k}. \tag{3.12}$$

By selecting $Q_{k-1,k}$ and $Q_{k,k}$, L_{kk} can be made to take on arbitrary value. From the submatrices of the block matrix QL and using the structure of F, we have $Q_{ik}L_{kk} = 0$ $\forall i = 1,2,\ldots,k-2$ and arbitrary L_{kk} implies $Q_{ik} = 0$ $\forall i = 1,2,\ldots,k-2$. In a similar manner, $L_{k-1,k-1} = Q_{k-1}^T F_{k-1}$ shows that $L_{k-1,k-1}$ is independent of $Q_{i,k-1}$ for $i = 1,2,\ldots,k-3$. Thus by the same argument, $Q_{i,k-1} = 0$ for $i = 1,2,\ldots,k-3$. The rest is by induction. Thus the unitary matrix Q is also generalized Hessenberg. Consequently $F_2 = LQ$ is also generalized Hessenberg.

3.3. Convergence of the Generalized QL Algorithm

Having established the existence and uniqueness of the generalized QL factorization and defined the GQL algorithm, consider next its convergence properties. Both a geometric and an algebraic approach are used. From the geometric point of view, we shall first show the relationship between the generalized QL algorithm and the power method used in numerical procedures for eigenvalue extraction [29]. Based on the properties of the power method, conclusions on the convergence of the QL algorithm will be drawn. In addition, it will be shown that if the generalized QL algorithm converges, a partial ordering or eigenvalues will be induced so that the leading submatrix of the limiting matrix $\overline{F}$ contains the dominant eigenvalues, i.e., eigenvalues with smallest moduli. On the other hand, the convergence process and the structure of the limiting matrix $\overline{F}$ at convergence is best illustrated using an algebraic approach. Furthermore, partial convergence is defined and its implication on generalized Hessenberg strating matrix is studied.

From the proof of Theorem 3.3, we note

$$\overline{F} = \lim_{s\to\infty} F_{s+1} = \lim_{s\to\infty} P_s^T F_1 P_s . \tag{3.13}$$

Thus, F_s converges to a constant $\overline{F}$ if and only if $P_s \to P$ as $s \to \infty$. On the other hand, from Theorem 3.2, the factorization $F_s = Q_s L_s$ is not unique if F_1 (thus F_s) is singular. Since $P_s = Q_1 \cdots Q_s$, in order to prove convergence of the GQL algorithm, it is necessary to assume unique factorization at each iteration. Henceforth, except noted to the contrary, it is assumed that F_1 is nonsingular. Such an assumption may not be unreasonable because many systems possess nonzero eigenvalues.

From Theorem 3.3,

$$P_s = F_1^s U_s^{-1}. \tag{3.14}$$

Because U_s is lower block triangular, (3.14) implies that the last q <u>block columns</u> of P_s are linear functions of the last q <u>block columns</u> of the matrix F_1^s. Let the $n \times n$ identity matrix I be partitioned compatibly with F_1, i.e.,

$$I = \begin{bmatrix} I_1 & 0 & & 0 & & 0 \\ 0 & I_2 & & 0 & & 0 \\ \vdots & 0 & \cdots & I_{k-q+1} & \cdots & \\ \vdots & \vdots & & \vdots & & 0 \\ 0 & 0 & & 0 & & I_k \end{bmatrix}.$$

The following notations and definitions will be used in the sequel. We define $\tilde{\mathcal{E}}_q$ as the space spanned by the last q block columns of I, i.e.,

$$\tilde{\mathcal{E}}_q = \mathcal{R}\{I^q\}$$

where

$$I^q = \begin{bmatrix} 0 & & 0 \\ \vdots & & \\ 0 & & \vdots \\ I_{k-q+1} & \cdots & \\ 0 & & 0 \\ \vdots & & \\ 0 & & I_k \end{bmatrix}.$$

Let $\mathcal{F}_1$ be the linear operator corresponding to the matrix F_1 and let $\mathcal{X}_o$ be an arbitrary subspace of $\mathcal{R}^n$ spanned by p linearly independent n-vectors $x_1, x_2, \ldots, x_p$. An iterative sequence of subspaces $\{\mathcal{X}_s\}$ will be denoted by

$$\mathcal{X}_s \equiv \mathfrak{F}_1 \ldots \mathfrak{F}_1 \mathcal{X}_o = \mathfrak{F}_1^s \mathcal{X}_o, \qquad s = 1,2,\ldots \tag{3.15}$$

and $\mathcal{X}_o$ will be called the starting subspace of $\mathcal{X}_s$. Let the eigenvalues of F_1 and their respective eigenvectors be ordered by their modulus:

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$$

$$v_1, v_2, \ldots, v_n.$$

Then any p dimensional subspace spanned by the leading p eigenvectors is called dominant in modulus. If $|\lambda_p| = |\lambda_{p+1}|$, then there will be more than one such dominant in modulus invariant subspaces. Let $\mathcal{D}(p,\mathcal{R}^n)$ denote the set of all such dominant in modulus spaces. By cardinality of a set, we refer to the number of elements in that set. When card $\mathcal{D}(p,\mathcal{R}^n) = 1$, we shall denote this unique dominant in modulus invariant subspace by $D(p,\mathcal{R}^n)$. In general, for an arbitrary $n \times n$ matrix A compatible with a given generalized Hessenberg matrix F, let $(A)_q$ denote the submatrix of A consisting of the last q <u>block columns</u> of A and $(\mathcal{A})_q$ be the subspace of $\mathcal{R}^n$ spanned by the linearly independent columns of $(A)_q$. Furthermore, we introduce the following definitions:

<u>Definition 3.1</u>: An operator $\mathfrak{F}_1$ in $\mathcal{R}^n$ is said to be defective if there are less than n linearly independent eigenvectors for F_1, where F_1 is the matrix representation of $\mathfrak{F}_1$ in $\mathcal{R}^n$.

<u>Definition 3.2</u>: Let $\mathcal{X}_p$ be an arbitrary p dimensional subspace of $\mathcal{R}^n$ and $\mathcal{N}_p$ be a p dimensional invariant subspace of $\mathfrak{F}_1$. $\mathcal{X}_p$ is said to be deficient in $\mathcal{N}_p$ if $\mathcal{X}_p \cap \mathcal{N}_p^c \neq 0$.

<u>Remark</u>: If $\mathcal{X}_p$ is not deficient in $\mathcal{N}_p$, then $\mathcal{X}_p$ is a complement of $\mathcal{N}_q^c$, i.e., $\mathcal{R}^n = \mathcal{X}_p \oplus \mathcal{N}_p^c$. In this case it is equivalent to say $\mathcal{X}_p$ and $\mathcal{N}_p^c$ must not

contain common elements. However, in general we cannot claim $\mathcal{X}_p$ contains $\mathcal{N}_p$ or vice versa. A simple example shows this. Let $\mathcal{N}_p = \text{span}\{x_1, x_2\}$, $\mathcal{N}_p^c = \text{span}\ \{x_3\}$ where x_i are linearly independent and $x_i \in \mathbb{R}^3$. Clearly $R^3 = \text{span}\{x_1, x_2, x_3\}$. Let $\mathcal{X}_p = \text{span}\{x_1 + x_3, x_2\} = \text{span}\{v_1, v_2\}$. Clearly $\mathcal{X}_p \oplus \mathcal{N}_p^c = \mathbb{R}^3$ but neither $\mathcal{X}_p$ contains $\mathcal{N}_p$ nor $\mathcal{N}_p$ contains $\mathcal{X}_p$.

From (3.14), we note

$$(\mathcal{P}_s)_q = \mathcal{F}_1^s (U_s^{-1})_q \equiv \mathcal{X}_s. \tag{3.16}$$

Since $\mathcal{X}_s$ must converge $\forall q$ in order for F_s to converge, the convergence of $\mathcal{X}_s$ must be guaranteed. $\mathcal{X}_s$ is formed by $\mathcal{F}_1$ repeatedly operating on columns of U_s^{-1}. Thus the following facts are necessary:

<u>Fact 1</u> [29]: For an arbitrary vector $x \in \mathbb{R}^n$, let λ_M and ν_M be the eigenvalues of A with maximum modulus and its associated eigenvectors. Under appropriate conditions, the power sequence $\{A^s s,\ s = 0,1,2,\ldots\}$ converges to ν_M.

For the simplest case of distinct eigenvalues, this may be seen as follows. Let u_o be an arbitrary n-vector and define two sequences u_s and t_s by

$$u_{s+1} = t_{s+1}/\max(t_{s+1}), \qquad s = 0,1,2,\ldots$$
$$t_{s+1} = Au_s, \qquad s = 0,1,2,\ldots$$

where max(x) equals the element of maximum modulus of x. Then

$$u_s = \frac{A^s u_o}{\max(A^s u_o)}, \qquad s = 0,1,2,\ldots$$

Next express u_o in terms of the eigenvectors

$$u_o = \sum_{i=1}^{n} \alpha_i \nu_i .$$

Hence

$$A^s u_o = \sum_{i=1}^{n} \alpha_i \lambda_i^s V_i = \lambda_M^s [\alpha_M V_M + \sum_{i \neq M} (\lambda_i/\lambda_M)^s \alpha_i V_i].$$

Since $|\lambda_i| < |\lambda_M|$ $\forall i$, then as $s \to \infty$, if $\alpha_M \neq 0$, we have

$$A^s u_o \sim \lambda_M^s \alpha_M V_M.$$

Also $\max(A^s u_o) \sim \lambda_M^s \alpha_M \max(V_M)$, giving us

$$u_s = \frac{A^s u_o}{\max(A^s u_o)} \sim \frac{V_M}{\max(V_M)} .$$

The more general case of repeated eigenvalues may be found in [23].

Fact 2 [29]: For arbitrary n-vectors x, y, and z, the subspace spanned by x, y, z and its successive iterates obtained by the power method:

$$[x^s, y^s, z^s] = A^s[x,y,z]$$

under appropriate conditions, converges to the invariant subspace spanned by eigenvectors associated with eigenvalues of largest modulus.

The following lemma from [29] then allows us to relate the convergence of the generalized QL algorithm to that of the power method.

Lemma 3.1 [29]: If $\mathcal{J}$ is a nondefective operator in $\mathcal{R}^n$ and card $\mathcal{D}(p,\mathcal{R}^n) = 1$, then the power sequence $\{\mathcal{J}^s \mathcal{Z}\}$ where $\mathcal{Z} = \text{span}\{x_1, \ldots, x_p\}$ converges to $D(p,\mathcal{R}^n)$ linearly as $s \to \infty$ if and only if $\mathcal{Z}$ is not deficient in $D(p,\mathcal{R}^n)$.

Lemma 3.1 may be interpreted as follows. In order for the initial space $\mathcal{Z}$ to converge under the power method, $\mathcal{Z}$ must not contain common elements with $D^c(p,\mathcal{R}^n)$ which is the invariant subspace formed by the n-p eigenvectors other than the p eigenvectors associated with the p eigenvalues of maximum modulus.

A relationship between the generalized QL algorithm and power method may be established as follows. Since U_s^{-1} is lower block triangular, it is obvious that

$$(U_s^{-1})_q \in \tilde{\mathcal{E}}_q .$$

From (3.16), we have

$$\mathcal{X}_s \sim \mathcal{F}_1^s \tilde{\mathcal{E}}_q . \tag{3.17}$$

where $\sim$ denotes equivalency. Comparing (3.16) and (3.17), we conclude that the sequence $\{\mathcal{X}_s\}$ is generated from the starting subspace $\tilde{\mathcal{E}}_q$. In particular, the generalized QL algorithm is seen to be equivalent to a nested sequence of k power sequences with starting spaces $\tilde{\mathcal{E}}_1, \tilde{\mathcal{E}}_2, \ldots, \tilde{\mathcal{E}}_k$ For the convergence of the generalized QL algorithm, we immediately have the following theorem.

<u>Theorem 3.5</u>: The generalized QL algorithm converges if and only if the power sequence $\{\mathcal{X}_s\} = \{\mathcal{F}^s \tilde{\mathcal{E}}_q\}$ converges for all $q = 1,2,\ldots,k$.

Furthermore, in view of Fact 2 of the power method, if the generalized QL algorithm converges, $(\mathcal{P}_s)_q$, $\forall q = 1,\ldots,k$ converges to the $\sum_{i=k-q+1}^{k} r_i = m_q$ dimensional subspace of $\mathcal{R}^n$ spanned by the m_q eigenvectors corresponding to the m_q eigenvalues of largest magnitude. In other words, $P = \lim_{s\to\infty} P_s$ is ordered such that columns of $(P)_q$ spans $D(m_q, \mathcal{R}^n)$. Because

$$\lim_{s\to\infty} P_s^T F_1 P_s = P^T F_1 P = \overline{F}$$

then the leading $\sum_{i=1}^{\ell} r_i \times \sum_{i=1}^{\ell} r_i$ submatrix of $\overline{F}$ will represent the restriction of $\overline{F}$ to the eigenspace spanned by the $\sum_{i=1}^{\ell} r_i$ eigenvectors corresponding to smallest modulus. Hence we can conclude that the leading submatrix of $\overline{F}$

always contains the dominant (eigenvalues with smallest modulus) eigenvalues of F_1.

Theorem 3.6: If the generalized QL algorithm converges, then the eigenvalues of the leading submatrix of the limiting matrix $\bar{F}$ are the dominant eigenvalues of F_1.

Using Lemma 3.1 and Theorem 3.5, the convergence of the generalized QL algorithm may be stated as follows:

Theorem 3.7: If

(i) $\mathfrak{F}_1$ nondefective, and

(ii) $\mathfrak{F}_1$ has a unique eigenspace $D(m_q, \mathcal{R}^n)$, $m_q = \sum_{i=k-q+1}^{k} r_i$ corresponds to m_q eigenvalues of largest magnitude, and

(iii) F_1 is nonsingular and generalized Hessenberg,

then the generalized QL algorithm converges if and only if $(\mathcal{L}_1^{-1})_q$ is not deficient in $D(m_q, \mathcal{R}^n)$ $\forall q = k, k-1, \ldots, 1$, where $(\mathcal{L}_1^{-1})_q$ denotes the space spanned by the last q block columns of L_1^{-1} from the decomposition $F_1 = Q_1 L_1$.

Proof: Rewrite (3.14)

$$P_s = F_1^s U_s^{-1} = F_1 L_1^{-1}(L_2^{-1}, \ldots, L_s^{-1}).$$

Consider an alternate sequence

$${}_1P_s = F_1^s L_1^{-1}.$$

Hence

$$({}_1\mathcal{P}_s)_q = \mathfrak{F}_1^s (\mathcal{L}_1^{-1})_q.$$

Now ${}_1P_s \to {}_1\bar{P}_s$ if and only if $\forall q$, $(\mathcal{L}_1^{-1})_q$ is not deficient in $D(m_q, \mathcal{R}^n)$. Since $(L_2^{-1} \ldots L_s^{-1})$ is also lower block triangular of the same form as L_1^{-1}, thus

$$(U_s^{-1})_q \in (\mathcal{L}_1^{-1})_q.$$

Hence $\mathcal{F}_1^s$ in effect operated on $(\mathcal{L}_1^{-1})_q$ in (3.16). Thus P_s converges if and only if ${}_1P_s$ converges. This proves the assertion.

In order to understand the causes of convergence in general and what is the structure of the limiting matrix $\overline{F}$, let X be the modal matrix of F_1, then

$$F_1 = XDX^{-1} = XDY \tag{3.18}$$

where D is block diagonal and compatible with F_1, i.e., D = block diag D = block diag $(\Lambda_1, \Lambda_2, \ldots, \Lambda_k)$, $\Lambda_i \in R^{r_i \times r_i}$. If $S(\Lambda_i)$ denotes the spectrum of Λ_i, then min $S(\Lambda_i)$, max $S(\Lambda_i)$ denotes the eigenvalues in $S(\Lambda_i)$ with minimal and maximal modulus respectively. Assume further than

$$\max S(\Lambda_i) < \min S(\Lambda_{i+1}), \quad i = 1, \ldots, k-1. \tag{3.19}$$

Note that cases with eigenvalues of equal moduli including repeated eigenvalues have not been ruled out. The only assumption is that eigenvalues with equal moduli may be grouped into a particular Λ_i and thus satisfy (3.19).* Factorize Z and Y as

$$X = QL \tag{3.20a}$$

$$Y = U\tilde{L} \tag{3.20b}$$

where Q is unitary, L is lower block triangular, U is unit upper block triangular (i.e., unit matrices on the main block diagonal) and $\tilde{L}$ is lower block triangular. Moreover, Q, L, U, and $\tilde{L}$ are all compatible with F_1. Using (3.20) in $P_s U_s = F_1^s$,

---

*This depends on the number of eigenvalues of equal modulus and on r_i which is determined by the structure of the generalized Hessenberg matrix F_1.

$$
\begin{aligned}
P_s U_s = F_1^s &= XD^s Y \\
&= QLD^s U\tilde{L} \\
&= QL(D^s UD^{-s})D^s\tilde{L}.
\end{aligned} \tag{3.21}
$$

Since U is <u>unit</u> upper block triangular,

$$
[D^s UD^{-s}]_{i,j} = \begin{cases} [D_{i,i}]^s U_{i,j}[D_{j,j}]^{-s} & i \leq j \\ 0 & \text{otherwise} \end{cases}. \tag{3.22}
$$

Hence we can write

$$
D^s UD^{-s} = I + E_s \tag{3.23}
$$

where E_s is upper block diagonal with zero main diagonal blocks and for $j > i$, the (i,j)th block is given by (3.22). Due to (3.19), D_{jj} has eigenvalues with larger moduli than those of D_{ii}. Then $\lim_{s\to\infty} E_s = 0$. Using (3.23) in (3.21), we have

$$
\begin{aligned}
P_s U_s = F_1^s &= QL(I + E_s)D^s\tilde{L} \\
&= Q(I + LE_s L^{-1})LD^s\tilde{L} \\
&= Q\tilde{Q}_s \cdot \tilde{L}_s LD^s\tilde{L}
\end{aligned} \tag{3.24}
$$

where the product $\tilde{Q}_s\tilde{L}_s$ results from a generalized QL decomposition of $I + LE_s L^{-1}$. Because Q and $\tilde{Q}_s$ are unitary and $\tilde{L}_s$, L, D^s, and L are lower block triangular, it follows from the uniqueness of the QL decomposition

$$
P_s = Q\tilde{Q}_s \tag{3.25a}
$$

$$
U_s = \tilde{L}_s LD^s\tilde{L}. \tag{3.25b}
$$

Moreover, from the fact that $\lim_{s\to\infty} E_s = 0$, it must be that

$$
\lim_{s\to\infty} \tilde{L}_s = I
$$

$$
\lim_{s\to\infty} \tilde{Q}_s = I.
$$

Hence

$$\lim_{s\to\infty} P_s \equiv Q = \text{constant matrix} \tag{3.26a}$$

and

$$\lim_{s\to\infty} U_s = \lim_{s\to\infty} LD^s\tilde{L}. \tag{3.26b}$$

Furthermore,

$$\lim_{s\to\infty} F_{s+1} = \lim_{s\to\infty} P_s^T F_1 P_s = Q^T F_1 Q = \bar{F} \tag{3.27}$$

and using (3.18)

$$\bar{F} = Q^T(XDX^{-1})Q = LDL^{-1}. \tag{3.28}$$

Equations (3.26)-(3.28) show F_s indeed converges and the limit $\bar{F}$ will be lower block triangular and compatible with F_1. In other words, when the GQL algorithm is applied to a generalized Hessenberg matrix F, the Hessenberg block $(F_s)_{i,i+1}$ will tend to a zero matrix when convergence is achieved. Thus convergence is a consequence of (3.22) and the assumption that D_{jj} has eigenvalues with larger moduli than those of D_{ii}.

Next consider the relationship of geometric and algebraic approaches and the implications of (3.19) on the convergence of the GQL algorithm. Using (3.25b) in (3.14), we get

$$P_s = F_1^s \cdot \tilde{L}^{-1} D^{-s} L^{-1} \tilde{L}_s^{-1}.$$

In addition, (3.25b) indicates

$$(\tilde{\mathcal{L}}^{-1}\mathcal{D}^{-s}\mathcal{L}^{-1}\tilde{\mathcal{L}}_s^{-1})_q \sim (\mathcal{U}_s^{-1})_q.$$

Since $U_s^{-1} = L_s^{-1}L_{s-1}^{-1} \ldots L_1^{-1}$, $(\mathcal{U}_s^{-1})_q \sim (\mathcal{L}_1)_q^{-1}$. Thus

$$(\tilde{\mathcal{L}}^{-1}D^{-s}\mathcal{L}^{-1}\tilde{\mathcal{L}}_s^{-1})_q \sim (\mathcal{L}_1^{-1})_q.$$

In view of Theorem 3.7, we are interested in the relationship of $(\tilde{\mathcal{L}}^{-1}D^{-s}\mathcal{L}^{-1}\tilde{\mathcal{L}}_s^{-1})_q$ and $D(m_q,R^n)$. To this end, using (3.20) and (3.25), we find

$$\tilde{L}^{-1}D^{-s}L^{-1}\tilde{L}_s^{-1} = Q\cdot LD^{-s}L^{-1}\tilde{Q}_s. \tag{3.29}$$

Because $\tilde{Q}_s \to I$ as $s \to \infty$, $LD^{-s}L^{-1}\tilde{Q}_s$ $LD^{-s}L^{-1}$. From (3.29), we conclude

$$(\tilde{\mathcal{L}}^{-1}\mathcal{D}^{-s}\mathcal{L}^{-1}\tilde{\mathcal{L}}_s^{-1})_q \sim (\mathcal{Q})_q. \tag{3.30}$$

Recall $Q = XL^{-1}$ where X is the modal matrix of F_1 with columns ordered according to (3.19). Since rank $(Q)_q$ = rank$(X)_q$, $(\mathcal{Q})_q \sim D(m_q,\mathcal{R}^n)$. From (3.27) we conclude $(\tilde{\mathcal{L}}^{-1}\mathcal{D}^{-s}\mathcal{L}^{-1}\tilde{\mathcal{L}}_s^{-1})_q \sim D(m_q,\mathcal{R}^n)$. Hence $(\mathcal{L}_1^{-1})_q$ is not deficient in $D(m_q,\mathcal{R}^n)$. Thus Theorem 3.7 guarantees the convergence of the generalized QL algorithm as long as F_1 has nonzero eigenvalues and the eigenvalues of F_I may be grouped according to (3.19).

In the remaining part of the chapter, we shall introduce the notion of partial convergence. The motivation lies in the fact the dimensions of the submatrices $F^{(i,i)} = F_{i,i}$ of the GHR are defined apriori once the output matrix C is specified in the original system representation. It is possible that eigenvalues with equal moduli appear in successive blocks Λ_i of (3.19). The simplest example occurs when the output is one dimensional and there exist complex conjugate pairs of eigenvalues so that due to the dimension of blocks $F^{(i,i)}$ (which is 1 by 1 if output is scalar), one complex eigenvalue is associated with Λ_i and its complex conjugate is associated Λ_{i+1}. In view of (3.22), $\lim_{s\to\infty}[D^sUD^{-s}]^{(i,i+1)} \neq 0$. The nonzero elements will have constant modulus but different in argument for each value of s. To this end, let us introduce the following definition.

<u>Definition 3.3</u>: The generalized QL algorithm is said to have partial convergence if there exists a set $\mathcal{K}_o \subset \mathcal{K}$ with $\mathcal{K}_o = \{q_{r_1}, q_{r_2}, \ldots, q_{r_t}\}$, $\mathcal{K} = \{1,2,\ldots,k-1\}$ such that

$$\lim_{s\to\infty} F_s^{(q_i,q_i+1)} = 0 \qquad \forall q_i \in \mathcal{K}_o.$$

To see that F_s still converges to a lower block triangular matrix as $s\to\infty$ but not all $F_s^{(i,i+1)}$, $i\in\mathcal{K}$ converges to a zero matrix, let us, for simplicity, assume that Λ_t, Λ_{t+1} have eigenvalues of equal modulus. The extension to more general case where there are isolated groups of Λ_i's with eigenvalues of equal modulus is obvious. Equation (3.19) now becomes

$$\max S(\Lambda_1) < \min S(\Lambda_2) < \cdots < \max S(\Lambda_t) \le \min S(\Lambda_{t+1}) < \cdots < \max S(\Lambda_{k-1}) < \min S(\Lambda_k). \tag{3.31}$$

Then instead of (3.23), we may write

$$D^s U D^{-s} = \overline{U}_s + E_s \tag{3.32}$$

where $\lim_{s\to\infty} E_s = 0$, with $\overline{U}_s$ of the form

$$\overline{U}_s = \begin{bmatrix} I_1 & 0 & \cdots & & & & & & 0 \\ 0 & I_2 & 0 & \cdots & & & & & 0 \\ \vdots & & \ddots & & & & & & \vdots \\ \vdots & & & I_{t-1} & 0 & & & & \vdots \\ \vdots & & & & I_t & X & 0 & \cdots & \vdots \\ \vdots & & & & & I_{t+1} & 0 & \cdots & \vdots \\ \vdots & & & & & & I_{t+2} & 0 \cdots & 0 \\ \vdots & & & & & & & \ddots & \vdots \\ 0 & \cdots & & & & & 0 & \cdots & I_k \end{bmatrix}. \tag{3.33}$$

Recall from (3.21),

$$\begin{aligned} P_s U_s &= QL(D^s U D^{-s}) D^s \tilde{L} \\ &= X\bar{U}_s (I + \bar{U}_s^{-1} E_s) D^s \tilde{L} \\ &= X\bar{U}_s (I + G_s) D^s \tilde{L} \end{aligned}$$

where $\lim_{s\to\infty} G_s = 0$. Writing $X\bar{U}_s = \tilde{Q}_s \tilde{L}_s$ then

$$P_s U_s = \tilde{Q}_s \tilde{L}_s (I + G_s) D^s \tilde{L}$$

or as $s \to \infty$

$$\tilde{Q}_s \sim P_s.$$

Hence

$$\begin{aligned} F_{s+1} &= P_s^T F_1 P_s \\ &= P_s^T X D X^{-1} P_s \\ &= P_s^T X \bar{U}_s (\bar{U}_s^{-1} D \bar{U}_s) \bar{U}_s^{-1} X^{-1} P_s \\ &\sim \tilde{L}_s (\bar{U}_s^{-1} D \bar{U}_s) \tilde{L}_s^{-1}. \end{aligned} \tag{3.34}$$

Now

$$X\bar{U}_s = QL\bar{U}_s = \tilde{Q}_s \tilde{L}_s. \tag{3.35}$$

If we define

$$\tilde{Q}_s = Q \begin{bmatrix} I & 0 & 0 \\ 0 & W_s^{-1} & 0 \\ 0 & 0 & I \end{bmatrix}, \quad W_s \in \mathbb{R}^{p\times p}, \quad p = r_t + r_{t+1}$$

and

$$\tilde{L}_s = \begin{bmatrix} I & 0 & 0 \\ 0 & W_s & 0 \\ 0 & 0 & I \end{bmatrix} L\bar{U}_s.$$

Clearly they satisfy (3.35). Denote

$$L = \begin{bmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} D_1 & & & \\ & \Lambda_t & & 0 \\ & & \Lambda_{t+1} & \\ 0 & & & D_2 \end{bmatrix}$$

from (3.34) we have

$$F_{s+1} \sim \begin{bmatrix} I & 0 & 0 \\ 0 & W_s & 0 \\ 0 & 0 & I \end{bmatrix} LDL^{-1} \begin{bmatrix} I & 0 & 0 \\ 0 & W_s^{-1} & 0 \\ 0 & 0 & I \end{bmatrix}$$

$$\sim \begin{bmatrix} \hat{L}_{11} & 0 & 0 \\ W_s\hat{L}_{21} & W_s\hat{L}_{22}W_s^{-1} & 0 \\ \hat{L}_{31} & \hat{L}_{32}W_s^{-1} & \hat{L}_{33} \end{bmatrix}$$

where $\hat{L}_{ij}$ $\forall i,j$ are independent of iteration number s, and

$$W_s\hat{L}_{22}W_s^{-1} = W_sL_{22}\begin{bmatrix} \Lambda_t & 0 \\ 0 & \Lambda_{t+1} \end{bmatrix} L_{22}^{-1}W_s^{-1}. \tag{3.36}$$

Thus as $s \to \infty$, F_{s+1} eventually becomes a fixed-structure lower block triangular matrix except for block rows and block columns corresponding to Λ_t and Λ_{t+1}. Even though the diagonal blocks correspond to Λ_t and Λ_{t+1} do not converge, their eigenvalues are those of Λ_t and Λ_{t+1} as a consequence of (3.36).

3.4. Computational Aspects of the GQL Algorithm

Successive iterates of the generalized QL algorithm requires the factorization of F_s into block matrices Q_s and L_s and obtain F_{s+1} as in (3.6). In this case, both factor matrices Q_s and L_s must be solved for. Alternatively, the successive GQL iterates may be obtained as a result of linear transformation using the relationship

$$F_{s+1} = L_s F_s L_s^{-1} \qquad , \; s = 1,2,\ldots \tag{3.37}$$

While there are many computational procedures for the standard QR algorithm that may be adapted to the generalized QL algorithm, the presence of block structure in F_s, Q_s and L_s and with reasonably small dimensions associated with each block, new alternative approaches may be developed. In view of (3.37), it is sufficient to obtain L_s only in order to obtain the next iterate F_{s+1}. We have from (3.5)

$$F_s^T F_s = L_s^T L_s . \tag{3.38}$$

When F and L are written in terms of their block entries

$$F = [F_1, \ldots, F_k], \qquad L = \begin{bmatrix} L_{11} & 0 \;\ldots\; 0 \\ L_{21} & L_{22} \; 0 \;..0 \\ \vdots & \ddots \\ \vdots & \qquad \ddots \\ L_{k1} & \qquad L_{kk} \end{bmatrix} \tag{3.39}$$

then equating both sides of (3.38), we find that

$$L_{ii}^T L_{ii} = F_i^T F_i - \sum_{m=i+1}^{k} L_{mi}^T L_{mi}, \qquad i = k, k-1, \ldots, 1 \tag{3.40}$$

$$L_{ij} = (L_{ii}^T)^{-1}[F_i^T F_j - \sum_{m=i+1}^{k} L_{mi}^T L_{mj}], \qquad j = 1,2,\ldots,i-1. \tag{3.41}$$

(3.40)-(3.41) constitute a set of recursive equations for L, of course knowing L, $Q = FL^{-1}$. While (3.38) always has a nonsingular L_s as a solution, it is not evident that this solution may be constructed to be lower block triangular and have symmetric, posiitive definite diagonal blocks as required for uniqueness of factorization. An examination of (3.40) indicates that L_{ii} may be solved for only if the right hand side of (3.40) is at least positive semi-definite and L_{ij} may then be obtained using (3.41) only if the right hand side of (3.40) is positive definite. We now proceed to show that under the assumption of nonsingularity of F_1, the right hand quantity of (3.40) is indeed positive definite. From (3.8),

$$L_{ii}^T L_{ii} = [F_i - \sum_{m=i+1}^{k} Q_m L_{mi}]^T [F_i - \sum_{m=i+1}^{k} Q_m L_{mi}]. \quad (3.42)$$

Expanding the right hand side and using (3.9), we have

$$\begin{aligned} L_{ii}^T L_{ii} &= F_i^T F_i + \sum_{m=i+1}^{k} L_{mi}^T L_{mi} - \sum_{m=i+1}^{k} L_{mi}^T L_{mi} - \sum_{m=i+1}^{k} L_{mi}^T L_{mi} \\ &= F_i^T F_i - \sum_{m=i+1}^{k} L_{mi}^T L_{mi}. \end{aligned} \quad (3.43)$$

Thus (3.42) is equivalent to (3.40). We can conclude that the right hand side of (3.40) is at least positive semi-definite. Due to the nonsingularity of L, the right hand side then must be positive definite. It is obviously symmetric. Note during step i, all quantities on the right hand side have already been obtained. From [30], a unique positive definite and symmetric L_{ii} exists using (3.40). Thus when Q is not explicitly required, (3.40)-(3.41) provide a simple set of equations for L.

Next consider the solution of (3.40)-(3.41). Denote the right hand side of (3.40) by $\tilde{Q}_i$, $i = k, k-1, \ldots, 1$. The problem therefore reduces to solving a quadratic equation for L_{ii}, inverting the L_{ii} and then solving for L_{ij}. For a fixed i, the right hand side of (3.40) is completely defined from previous calculations. Since $\tilde{Q}_i$ is symmetric and positive definite, then the matrix of eigenvectors of $\tilde{Q}_i$ is unitary. Denote the matrix of eigenvectors of $\tilde{Q}_i$ by V_i, then

$$V_i^T L_{ii}^T L_{ii} V_i = V_i^T \tilde{Q}_i V_i = \mathrm{diag}(\lambda_1^i, \lambda_2^i, \ldots, \lambda_{r_i}^i) = \Lambda_i \tag{3.44}$$

where $\{\lambda_1^i, \ldots, \lambda_{n_i}^i\} = \sigma\{\tilde{Q}_i\}$. $\sigma\{\tilde{Q}_i\}$ denotes the set of eigenvalues of $\tilde{Q}_i$. Thus

$$L_{ii}^T L_{ii} = (V_i \Lambda_i^{\frac{1}{2}} V_i^T)(V_i \Lambda_i^{\frac{1}{2}} V_i^T)$$

giving us

$$L_{ii} = V_i \Lambda_i^{\frac{1}{2}} V_i^T \tag{3.45}$$

which is symmetric and positive definite if $\lambda_i^{\frac{1}{2}}$ is taken to be positive. Once L_{ii} is obtained, L_{ij}, $j = 1,2,\ldots,i-1$ may be obtained by using (3.41). Note block submatrices of L are obtained successively starting from the kth block column. See Figure 3.1 for a summary of the algorithm.

3.5. Application of the Generalized QL Algorithm to the GHR

While the generalized QL algorithm may be used solely for numerical purposes, we recall the primary goal to develop the GQL theory is to apply it to the GHR so that any weak interconnection between S^i and $\bar{S}^i$, if it exists, may be brought into evidence. However, we stress the GQL algorithm does not create weak interconnections, it merely reorganizes system representations, and in that process, block diagonalizes the system matrix according to eigenvalue magnitudes. We discuss two particular aspects. The first is the use of the factor matrix L_s in $F_s = Q_s L_s$ as a linear transformation of the GHR and show the general effect of GQL algorithm on GHR. The second is how the property of GQL algorithm may be used to advantage when applied to the GHR in order to indicate a two-time scale decomposition of the system into a fast and a slow subsystem. While the second aspect is

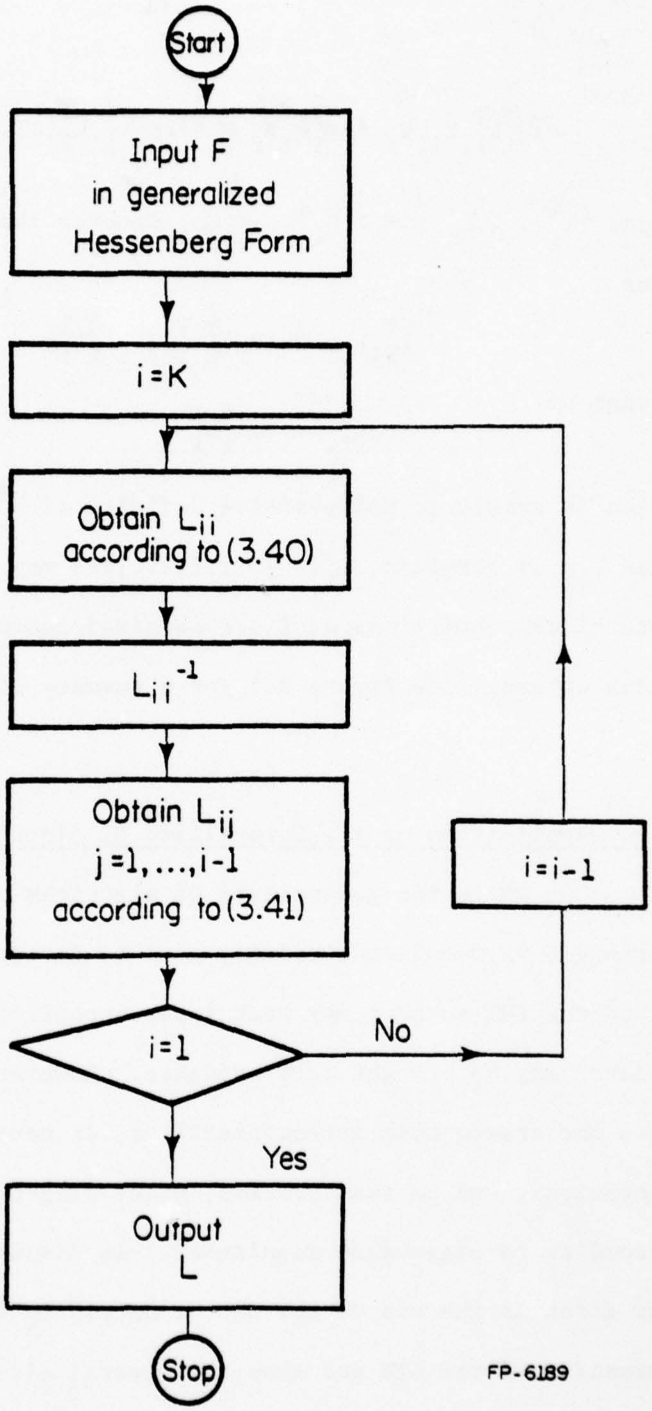


Figure 3.1. Flow chart of subroutine for obtaining L for the generalized QL factorization F=QL.

in fact a particular consequence of the first aspect, its repercussions are of separate significance because of applications within the theory of singular perturbation [20]. The generalized QL algorithm provides a prescriptive procedure for identifying slow and fast subsystems and thus represents a convenient tool for achieving thisccrucial step for the application of singular perturbation results.

In view of (3.28) which gives the structure of the limit $\overline{F}$, we observe that under conditions of convergence some or all of the super diagonal blocks in F_s will gradually reduce towards the zero matrix and, thus would tend to become weak links in the system structure. If indeed some of these links may be neglected with acceptable approximation error this would result in the residual being unobservable to the aggregate, achieving an approximate aggregation. Thus the aggregate may become an acceptable reduced order model of (2.9). We now make the above discussion more concrete by applying the sequence of similarity transformations characterized by the matrices L_s, $s = 1,2,3,\ldots$ to the GHR. Denote the GHR from (2.10) as

$$\dot{z}_1 = F_1 z_1 + G_1 u, \qquad z_1 \in R^n \tag{3.46}$$

$$y_m = Dz_1, \qquad y_m \in R^{r_1}$$

with $z = z_1$, $F = F_1$, and $G = G_1$ employed to simplify the notation in this section. Since F_1 is nonsingular, it has a unique block QL decomposition (Theorem 3.2) $F_1 = Q_1 L_1$. Premultiply (3.46) by L_1. Then

$$L_1 \dot{z}_1 = L_1 F_1 z_1 + L_1 G_1 u = L_1 Q_1 L_1 z_1 + L_1 G_1 u.$$

Define $z_2 = L_1 z_1$, $G_2 = L_1 G_1$. Then $\dot{z}_2 = F_2 z_2 + G_2 u$ and $y_m = DL_1^{-1} z_2$. A sequence of transformations yields the system representation:

$$\dot{z}_{s+1} = F_{s+1} z_{s+1} + G_{s+1} u \tag{3.47}$$

$$y_m = DL_1^{-1} \ldots L_s^{-1} z_{s+1} = DU_s^{-1} z_{s+1} \tag{3.48}$$

where

$$z_{s+1} = L_s z_s = L_s \ldots L_1 z_1 = U_s z_1 \tag{3.49}$$

$$G_{s+1} = L_s G_s = L_s \ldots L_1 G_1 = U_s G_1 .$$

Since $U_s = \prod_{i=1}^{s} L_i$ is lower block triangular, the r_1-dimensional leading subvector of z_s is the transformed measured output y_m. This is precisely why we require the decomposition $F = QL$ instead of $F = QR$. Otherwise, the transformed output y_m will be a linear combination of all system states z_{s+1} instead of measured output only. This is crucial because we want to "preserve" relationship between the measured outputs and state variables of any reduced order model constructed from the GHR.

The usefulness of the GQL algorithm in singular perturbation theory stems from Theorem 3.6. Traditionally singular perturbed system assumed the structure:

$$S_s: \; \dot{x}_1 = A_{11} x_1 + A_{12} x_2 + B_1 u \tag{3.50a}$$

$$S_f: \; \dot{x}_2 = 1/\mu (A_{21} x_1 + A_{22} x_2 + B_2 u) \tag{3.50b}$$

where μ is a small parameter and the fast eigenvalues (nondominant) of the composite system is associated with the fast subsystem S_f. Unfortunately, many systems possessing a two-time scale property are not directly expressed in the form of (3.50). Thus a systematic procedure to be applied to a general system representation so that (3.50) results will be useful.

The effect of applying a sequence of QL iterations on the system structure, as depicted in Figure 2.2 of Chapter 2, is that the subsystems will eventually be ordered so that S_1, as an independent subsystem, contains the dominant dynamics, S_2 is the system with which the next group of eigenvalues, in order of dominance, is associated, and so on to the last subsystem S_k which is the fastest subsystem as it is associated with eigenvalues with largest moduli. It is however, important to note that as the iterations are allowed to go to infinity, while the essential feedback links would tend to become weaker and weaker, the effect of the iterations on some other parameters is to make them larger and larger, e.g., allow them to grow unboundedly. To see this recall that from (3.25b), lim $U_s = LD^s\tilde{L}$, and when F_1 contains eigenvalues with modulus greater than one, U_s will grow unbounded. Thus if the QL algorithm is applied to the final goal of reducing all superdiagonal blocks to arbitrary neighborhood of the zero matrix, the inevitable results is the appearance of large gain factors in the control matrix and some feedforward links, accompanied with negligible entries in the output matrix. On the other hand when there exist some eigenvalues of moduli smaller than one this will have the inevitable result of the appearance of high gains in the output matrix of the resulting GHR. Thus, the primary aim of applying the generalized QL algorithm is to detect after a relatively small number of iterations whether some of the superdiagonal blocks have a high rate of convergence towards a zero matrix such that by neglecting them,an acceptable approximation error between the outputs of the full order model and the outputs of the approximate model in which such terms have been neglected is acceptable.

Assume after s iterations, we have a new GQL processed GHR representation:

$$\dot{z}_s = F_s z_s + G_s u \tag{3.51a}$$

$$y_m = [U_{s-1}^{(1,1)}]^{-1}[I_1 \; 0 \; \ldots \; 0] z_s \tag{3.51b}$$

with

$$F_s = \begin{bmatrix} F_s^j & E_s^j \\ H_s^j & \bar{F}_s^j \end{bmatrix} \qquad G_s = \begin{bmatrix} G_s^j \\ \bar{G}_s^j \end{bmatrix}$$

$$z_s = \begin{bmatrix} z_s^j \\ \bar{z}_s^j \end{bmatrix}, \quad z_s^j \in R^p, \quad p = \sum_{i=1}^{j} r_i .$$

If superdiagonal block $F_{j,j+1}^s$ in E_s^j is appreciably reduced, a candidate reduced order model is obtained by placing $E_s^j = 0$ and thus has the following form:

$$\begin{aligned} \dot{z}_r &= F_s^j z_r + G_s^j u \\ \hat{y}_m &= [U_{s-1}^{(1,1)}]^{-1}[I_1 \; 0 \; \ldots \; 0] z_r \end{aligned} \tag{3.52}$$

which results from retaining only the strongly observable part of the resulting system. (3.52) may be considered as an <u>approximate aggregate model</u> of (2.9). A desirable feature of (3.52) is that (a) approximate output $\hat{y}_m$ is only a linear combination of measured output y_m; (b) the aggregated reduced order model state z_r is linearly related to the original system state x of (2.9). This is so because z_s of (3.51) is related to x by a linear transformation $z_s = U_s Hx$. Such a relation is particularly useful if (3.52) is used to design feedback control laws which is subsequently applied to (2.9).

CHAPTER 4

GENERALIZED HESSENBERG REPRESENTATION AND REDUCED ORDER MODELING

Given a large scale dynamic system with many variables, one is frequently interested in the behavior and control of only a specific subset of variables. Thus a gross or aggregated input-output relation that does not involve details of unconcerned internal variables is of interest. A reduced order model that models the dynamics of these variables with acceptable approximation error is frequently sufficient. In this chapter, we shall study how GHR may be used for the construction of such a reduced order model induced by a set of variables (outputs are assumed) and show how the GQL algorithm can be used to combine structural information and weak coupling in order to achieve approximate aggregation. In the most general case, crucial steps in the construction of a reduced order model via GHR involves the definition of the model structure, the selection of adjustable parameters within the defined structure and the adjustment of these parameters to some appropriate values. We shall show that the GHR is a useful representation of a linear system for the purpose of constructing a reduced order model because it determines a collection of candidate structures, and, in each case, singles out parameters most appropriate for further adjustment if this is necessary. We establish these assertions by showing how some existing model reduction methods when applied to a system in GHR, only adjust a subset of GHR parameters within the selected reduced order model structure.

In this chapter we concentrate on the case where (2.9) is observable, thus not aggregable. In this case $F_{i,i+1} \neq 0$, $\forall i$. Recall the definition of aggregate and residual subsystem introduced in Chapter 2, i.e., equations (2.21)-(2.23) where

$$F^i = \begin{bmatrix} F_{11} & F_{12} & 0 & \cdots & 0 \\ F_{21} & F_{22} & F_{23} & \cdots & 0 \\ \vdots & & & & \\ F_{i1} & F_{i2} & F_{i3} & \cdots & F_{ii} \end{bmatrix} \quad E^i = \begin{bmatrix} 0 & & & & 0 \\ 0 & & & & 0 \\ \vdots & & & & \\ 0 & & 0 & \cdots & 0 \\ F_{i,i+1} & & 0 & \cdots & 0 \end{bmatrix} \quad G^i = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_i \end{bmatrix} \tag{4.1}$$

$$H^i = \begin{bmatrix} F_{i+1,1} & \cdots\cdots & F_{i+1,i} \\ \vdots & & \\ F_{k1} & \cdots\cdots & F_{k,i} \end{bmatrix} \quad \bar{F}^i = \begin{bmatrix} F_{i+1,i+1} & \cdots & F_{i+1,k} \\ \vdots & & \vdots \\ F_{k,i+1} & & F_{k,k} \end{bmatrix} \quad \bar{G}^i = \begin{bmatrix} G_{i+1} \\ \vdots \\ G_k \end{bmatrix}$$

and

$$\text{Aggregate } S^i: \quad \dot{z}^i = F^i z^i + E^i \bar{z}^i + G^i u \tag{4.2a}$$

$$\text{Residual } \bar{S}^i: \quad \dot{\bar{z}}^i = H^i z^i + \bar{F}^i \bar{z}^i + \bar{G}^i u. \tag{4.2b}$$

In general if for some i, the norm $\|H^i\|$ is much smaller than the norms of $\|F^i\|$, $\|E^i\|$, $\|\bar{F}^i\|$, $\|G^i\|$, and $\|\bar{G}^i\|$, then $F_{i,i+1}$ may be neglected to achieve approximate aggregation induced by (A,C). The resulting reduced order model is

$$\dot{z} = F^i z + G^i u \qquad , z \in R^p, \quad p = \sum_{j=1}^{i} r_j \tag{4.3a}$$

$$y = [I_1 \;\; 0 \ldots 0] z \;, \; y \in R^{r_1}. \tag{4.3b}$$

On the other hand, if none of the Hessenberg block $F_{i,i+1}$ can be neglected after some analysis such as trajectory error growth or eigenvalue perturbation, we propose that an approximate reduced order model be constructed in the form:

$$\dot{\tilde{z}} = \tilde{F}^i \tilde{z} + \tilde{G}^i u \tag{4.4}$$

$$\hat{y} = [I_1 \;\; 0 \ldots 0] \tilde{z},$$

for some $i = 1,\ldots,k-1$. Here, $\tilde{F}^i$ and $\tilde{G}^i$ have the same structure as F^i and G^i in (4.1) except that some parameters in $\tilde{F}^i$ and $\tilde{G}^i$ are to be adjusted to more appropriate values to compensate for the effect of the neglected term $F_{i,i+1}z_{i+1}$.

A natural choice is to adjust only the subset of the model parameters in $F_{i,j}$, $j = 1,\ldots,i$, and in G_i to achieve

$$\sum_{j=1}^{i} \tilde{F}_{i,j}\tilde{z}_j + \tilde{G}_i u \approx \sum_{j=1}^{i+1} F_{i,j}z_j + G_i u \tag{4.5}$$

and effecitvely compensate for the effect of the neglected term $F_{i,i+1}z_{i+1}$. This is indeed what results when singular perturbation theory [20] is applicable for the decomposition of slow and fast phenomena; when the procedure of norm-minimization proposed in [2] is generalized in conjunction with chained aggregation; and when an approach based on transforming the reduced order modeling problem to a constrained regulator problem [43] is developed. We first show when each of the above methods is used on a system in GHR, what and how parameters are adjusted so that (4.5) is approximated. The drawbacks of these methods then motivates our major proposal, the restricted QL algorithm. The remaining sections are organized as follows. The use of GHR with singular perturbation method is presented in Section 4.1, model reduction by norm minimization is presented in Section 4.2, while model reduction using GHR formulated as an output regulator problem is presented in Section 4.3. Finally, we proceed to develop the RQL algorithm which effectively modifies only a subset of model parameters while retaining the general structure of an appropriate reduced order model based on the properties of the generalized QL algorithm.

4.1. GHR and Model Reduction by Singular Perturbation

We now analyze what is the meaning of applying singular perturbation to a GHR in terms of parameter adjustment in the slow subsystem. Assume the existence of wide separation in eigenvalue modulus in (2.9). It has been pointed out that the system may not be in the suitable form for singular perturbation analysis and how GHR and GQL methods may be used to bring a general system into the suitable form. Assume a GQL processed GHR already exhibits an eigenvalue separation, i.e., the dominant eigenvalues are approximately specified by F^i and the fast eigenvalues are approximately those of $\bar{F}^i$ for some index i. Then to obtain the singular perturbation system (3.50), we normalize the residual subsystem by dividing the equation of the residual in (4.2) by $\|\bar{F}^i\|$ and introducing the notation $\tilde{F}^i = \dfrac{\bar{F}^i}{\|\bar{F}^i\|}$, $\tilde{H}^i = \dfrac{H^i}{\|\bar{F}^i\|}$, $\tilde{G}^i = \dfrac{\bar{G}^i}{\|\bar{F}^i\|}$ and $\mu = \dfrac{1}{\|\bar{F}^i\|}$, the composite system is brought to the form:

$$
\begin{aligned}
z^i &= F^i z^i + E_i \bar{z}^i + G^i u \\
\mu \dot{\bar{z}}^i &= \tilde{H}_i z^i + \tilde{F}_i \bar{z}^i + \tilde{G}^i u.
\end{aligned}
\tag{4.6}
$$

Under appropriate conditions, the reduced order model is obtained by placing $\mu = 0$ and solving the fast subsystem for $\bar{z}^i$ and substituting $\bar{z}^i$ into the aggregate. The reduced order model then is

$$\dot{\tilde{z}}^i = (F^i - E^i(\tilde{F}^i)^{-1}\tilde{H}^i)\tilde{z}^i + (G^i - E^i(\tilde{F}^i)^{-1}\tilde{G}^i)u \tag{4.7a}$$

$$\hat{y} = [I_1 \quad 0 \quad \dots \quad 0]\tilde{z}^i. \tag{4.7b}$$

Considering the expression for E^i in (4.1) we find that the correction terms $E^i(\tilde{F}^i)^{-1}\tilde{H}^i$ and $E^i(\tilde{F}^i)^{-1}\tilde{G}^i$ have the structure

$$E^i(\tilde{F}^i)^{-1}\tilde{H}^i = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 0 \\ & & & \\ K_{i,1} & K_{i,2} & \cdots & K_{i,i} \end{bmatrix}, \quad E^i(\tilde{F}^i)^{-1}\tilde{G}^i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \\ M_i \end{bmatrix}. \tag{4.8}$$

Thus, the effect of applying singular perturbation to a system in the GHR is equivalent to modifying only the parameters in $F_{i,j}$, $j = 1,\ldots,i$, and G_i in the GHR according to (4.7a).

4.2. GHR and Model Reduction by Norm-Minimization

It was suggested in [1,10,13] that when the pair (A,C) is not completely aggregable, the dynamics of the aggregated state variables can be approximated with a low order system in which F is selected to minimize the norm of CA-FC. The norm-minimizing F is given by

$$F = CAC^T(CC^T)^{-1} \tag{4.9}$$

and the reduced order model is

$$\dot{y} = Fy + CBu. \tag{4.10}$$

A basic shortcoming of this procedure is that the reduced order model is restricted to be of the same dimension as the measured output if the output matrix C is used as an aggregation matrix. However, with the introduction of chained aggregation, this approach may be generalized to produce reduced order models of dimension p_i, $i = 1,\ldots,k-1$, where $p_i = \sum_{j=1}^{i} r_j$. Suppose a model of dimension p_i is desired. Consider that the system has been transformed after (i-1) step of chained aggregation into the form where S^{i-1} is in GHR while $\bar{S}^{i-1}$ as yet is not. Continuing with chained aggregation would eventually yield the GHR as in (2.10). However, stopping at step (i-1) results in the system representation

$$\begin{bmatrix} \dot{z}_1 \\ \vdots \\ \dot{z}_{i-2} \\ \dot{z}_{i-1} \\ \cdots \\ \dot{x}_r \end{bmatrix} = \left[\begin{array}{cccc|ccc} F_{1,1} & F_{1,2} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots & & \vdots \\ F_{i-2,1} & \cdots & & F_{i-2,i-1} & 0 & \cdots & 0 \\ F_{i-1,1} & \cdots & & F_{i-1,i-1} & & \bar{C} & \\ \hline & & \bar{E} & & & \bar{A} & \end{array}\right] \begin{bmatrix} z_1 \\ \vdots \\ z_{i-2} \\ z_{i-1} \\ \cdots \\ x_r \end{bmatrix} + \begin{bmatrix} G_1 \\ \vdots \\ G_{i-2} \\ G_{i-1} \\ \cdots \\ \bar{B} \end{bmatrix} u . \tag{4.11}$$

Instead of continuing with step i of the chained aggregation procedure on the residual subsystem

$$\dot{x}_r = \bar{A}x_r + \bar{E}z^{i-1} + \bar{B}u \tag{4.12a}$$

$$z_i = \bar{C}x_r \tag{4.12b}$$

we shall choose a matrix $\bar{F}$ such that $\|\bar{C}\bar{A} - \bar{F}\bar{C}\|$ is minimized in accordance with (4.9). The approximately aggregated residual subsystem (4.12) thus becomes

$$\dot{z}_i = \bar{F}z_i + \bar{C}\bar{E}z^{i-1} + \bar{C}\bar{B}u \tag{4.13}$$

where $\bar{F} = \bar{C}\bar{A}\bar{C}^T(\bar{C}\bar{C}^T)^{-1}$ and z^{i-1}, consistent with the notations given in (4.1), is the composition of the subvectors z_j, $j = 1,2,\ldots,i-1$. Subsystem (4.13) coupled with the existing aggregate S^{i-1} at step i-1 of chained aggregation comprises a corrected reduced order model. Comparing the effects of norm-minimization characterized by (4.12)-(4.13) with the model S^i obtained directly from GHR by neglecting E^i, it follows immediately that the resulting model will have the form (4.4) with $F_{i,j}$, $j = 1,\ldots,i-1$ and G_i the same as those in the complete GHR, $\forall i$, while $F_{i,i}$ is adjusted to a new value given by

$$\tilde{F}_{i,i} = \bar{C}\bar{A}\bar{C}^T(\bar{C}\bar{C}^T)^{-1}. \tag{4.14}$$

Thus, the effect of using norm-minimizing at step i is equivalent to adjusting only the parameters $F_{i,i}$ in the complete GHR.

4.3. GHR and Model Reduction by Optimal Output Regulator Design

In this section, the model parameter adjustment will be carried out using the optimization procedure. Unlike [19] where all reduced order model parameters are obtained through optimization, the proposed approach adjusts only a subset of parameters within the chosen GHR structure, namely blocks F_{ij}, $j = 1,\ldots,i$. Assume a model of the form (4.4) is sought. Let us consider only open-loop behavior of the system and its reduced order model. In terms of (4.5), the adjustment on blocks F_{ij}, $j = 1,2,\ldots,i$ is to find δF_{ij}, $j = 1,\ldots,i$ such that $\|F_{i,i+1}z_{i+1} - \sum_{j=1}^{i}\delta F_{ij}z_j\|$ is minimized, i.e.

$$F_{i,i+1}z_{i+1} \approx \sum_{j=1}^{i}\delta F_{ij}z_j. \tag{4.15}$$

Clearly complete compensation cannot be achieved, i.e. equality cannot hold for (4.15). In the sequel, we shall see how the problem may be reformulated as an optimal output regulator problem and that (4.15) be solved by applying additional feedback such that $\tilde{F}$ of (4.4) is obtained.

Let

$$D = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ F_{i,i+1} \end{bmatrix}$$

AD-A069 766 ILLINOIS UNIV AT URBANA-CHAMPAIGN COORDINATED SCIENCE LAB F/G 12/1

MODEL REDUCTION AND DECENTRALIZED CONTROL OF LARGE SCALE SYSTEM--ETC(U)

AUG 78 E C TSE DAAB07-72-C-0259

UNCLASSIFIED R-820 NL

2 OF 3

AD

A069766

the reduced order modeling problem by parameter optimization may be reformulated in the following way: Given the augmented system

$$\dot{w} = F^i w + Dv$$
$$\dot{z}^i = F^i z^i + E^i \bar{z}^i \qquad (4.17)$$
$$\dot{\bar{z}} = H^i z^i + \bar{F}^i \bar{z}^i$$

determine the "control" v of the form

$$v = -Kw$$

so that the following quadratic criterion is minimized:

$$J(w_o, z_o^i, \bar{z}_o^i, K) = \tfrac{1}{2}\int_0^\infty [(w-z^i)^T(w-z^i) + v^T R v]\,dt. \qquad (4.18)$$

Introducing the notation

$$A = \begin{bmatrix} F^i & 0 & 0 \\ 0 & F^i & E^i \\ 0 & H^i & \bar{F}^i \end{bmatrix},\ B = \begin{bmatrix} D \\ 0 \\ 0 \end{bmatrix},\quad Q = \begin{bmatrix} I & -I & 0 \\ -I & I & 0 \\ 0 & 0 & 0 \end{bmatrix},\ X = \begin{bmatrix} w \\ z^i \\ \bar{z}^i \end{bmatrix} \qquad (4.19)$$

$$C = [I \quad 0 \quad 0]$$

the problem reduces to the following:

$$\dot{z} = Ax + Bv$$
$$y = Cx \qquad (4.20)$$
$$J_o = \tfrac{1}{2}\int_0^\infty (x^T Q x + v^T R v)\,dt.$$

Because v = -Kw = -Ky. (4.20) becomes a constrained output regulator problem [43] which has an optimal linear feedback solution dependent on the initial condition of state x. The optimal feedback gain K is obtained by reformulating (4.20) as a parameter optimization problem where

$$v^* = -K^*y = -K^*Cx \tag{4.21a}$$

$$\tilde{J}(K^*) = \min_K \{ \underset{x_o}{E} [\tfrac{1}{2} \int_o^\infty x^T (Q + C^T K^T RKC) x dt \} \tag{4.21b}$$

subject to

$$\dot{x} = (A - BK^*C)x. \tag{4.21c}$$

The formulation of the problem in terms of average value of the criterion is reasonable in the reduced order modeling problem since the model should be acceptable for all, or at least for a wide range of initial conditions of the system.

The solution K^* is then known to satisfy the following set of necessary conditions [43]:

$$\begin{aligned} &(A-BKC)^T M + M(A-BKC) + C^T K^T RKC + Q = 0 \\ &N(A-BKC)^T + (A-BKC)N + E(x_o x_o^T) = 0 \\ &K = R^{-1} B^T MNC^T (CNC^T)^{-1} \end{aligned} \tag{4.22}$$

with

$$\tilde{J} = \tfrac{1}{2} \operatorname{tr} ME(x_o x_o^T). \tag{4.23}$$

Unlike [43] where the initial condition is assumed to be uniformly distributed over a unit sphere, because w is meant to approximate z in (4.17), the appropriate initial condition on x_o should be

$$x_o = \begin{bmatrix} z_o \\ z_o \\ \bar{z}_o \end{bmatrix} = \begin{bmatrix} I & 0 \\ I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} z_o \\ \bar{z}_o \end{bmatrix} = \pi \begin{bmatrix} z_o \\ \bar{z}_o \end{bmatrix}. \tag{4.24}$$

The correct assumption on initial condition should be $[z_o^T \;\; \bar{z}_o^T]^T$, is uniformly distributed over a unit sphere, i.e.

$$E\left\{\begin{bmatrix} z_o \\ \bar{z}_o \end{bmatrix} [z_o^T \;\; \bar{z}_o^T]\right\} = \frac{1}{n} .$$

Then

$$E(x_o \, x_o^T) = \frac{1}{n}\pi\,\pi^T. \qquad (4.25)$$

Consequently the necessary conditions (4.22) for optimal K becomes:

$$(A-BKC)^T M + M(A-BKC) + C^T K^T RKC + Q = 0 \qquad (4.26a)$$

$$N(A-BKC)^T + (A-BKC)N + \frac{1}{n}\pi\,\pi^T = 0 \qquad (4.26b)$$

$$K = R^{-1}B^T MNC^T (CNC^T)^{-1}. \qquad (4.26c)$$

Thus, the original parameter optimization problem has been reduced to a problem of determining the triplet (M,N,K) as the solution of a set of coupled algebraic matrix equations. While the conditions under which convergence to a solution with different iteration schemes are known [43,44], it is difficult to satisfy them. In particular, if for some $K = K_i$, (4.26a) is solved for M_i and given K_i, M_i, (4.26b) is solved for N_i and the pair M_i, N_i is used in (4.26c) to obtain the successive K_{i+1}, it has been shown that the scheme will converge to a solution if at every step $A-BK_iC$ is a stable matrix. However, there is no guarantee that K_i generated through this procedure will be stabilizing. Nevertheless this as well as modified schemes, have been found applicable in practice.

To solve the reduced order modeling problem by transforming it into an equivalent output regulator problem, it is therefore necessary to solve two Lyapunov equations of order $\bar{n}$, $\bar{n} = n+p_i$, and compute the $p_i \times p_i$ mmatrix K. However, from the structure of the equivalent system and the defined criterion, (4.19)-(4.20), it may be shown that if the resulting average value of $\tilde{J}$ is not of interest (which it might be for possible comparison of two different reduced order models) it is actually only necessary to solve two, non-square linear matrix equations of order $p_i \times (n+p_i)$. To see this decompose M and N as

$$M = [M_1 \;\; M_2 \;\; M_3] = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{12}^T & M_{22} & M_{23} \\ M_{13}^T & M_{23}^T & M_{33} \end{bmatrix}, \; N = [N_1 \; N_2 \; N_3] = \begin{bmatrix} N_{11} & N_{12} & N_{13} \\ N_{12}^T & N_{22} & N_{23} \\ N_{13}^T & N_{23}^T & N_{33} \end{bmatrix}. \tag{4.27}$$

Equation (4.26c) then takes the form

$$K = R^{-1}D^T M_1 N_1^T. \tag{4.28}$$

Equations (4.26a) in expanded form become:

$$\begin{aligned}
&(F^i-DK)^T M_{11} + M_{11}(F^i-DK) + I + K^T RK = 0 \\
&(F^i-DK)^T M_{12} + M_{12}F^i + M_{13}E^i - I = 0 \\
&(F^i-DK)^T M_{13} + M_{12}H^i + M_{13}\bar{F}^i = 0 \\
&F^{i^T} M_{22} + E^{i^T} M_{23}^T + M_{22}F^i + M_{23}E^i + I = 0 \\
&F^{i^T} M_{23} + E^{i^T} M_{33} + M_{22}H^i + M_{23}\bar{F}^i = 0 \\
&H^{i^T} M_{23} + F^{i^T} M_{33} + M_{23}^T H^i + M_{33}\bar{F}^i = 0
\end{aligned} \tag{4.29}$$

while (4.26b) becomes:

$$\begin{aligned}
&N_{11}(F^i-DK)^T + (F^i-DK)N_{11} + \frac{1}{n} I = 0 \\
&N_{12}F^{i^T} + N_{13}E^{i^T} + (F^i-DK)N_{12} + \frac{1}{n} I = 0 \\
&N_{12}H^{i^T} + N_{13}\bar{F}^{i^T} + (F^i-DK)N_{13} = 0 && (4.30) \\
&N_{22}F^{i^T} + N_{23}E^{i^T} + F^i N_{22} + E^i N_{23}^T + \frac{1}{n} I = 0 && (4.30) \\
&N_{22}^T H^{i^T} + N_{23}\bar{F}^{i^T} + F^i N_{23} + E^i N_{33} = 0 \\
&N_{23}^T H^{i^T} + N_{33}\bar{F}^{i^T} + H^i N_{23} + \bar{F}^i N_{33} + \frac{1}{n} I = 0.
\end{aligned}$$

Thus, we find that (4.28), (4.29a)-(4.29c), and (4.30a)-(4.30c) are decoupled from the remaining equations and thus that the desired gain matrix K may be obtained by solving only this restricted set of coupled equations:

$$K = R^{-1}D^TM_1M_1^T$$
$$N_1A^T + (F^i-DK)N_1 + \frac{1}{n}P = 0, \quad P = [I \quad 0 \quad 0] \tag{4.31}$$
$$(F^i-DK)^TM_1 + M_1A + P = 0.$$

When K^* is obtained, the reduced order model of the system is characterized by

$$\tilde{F}^i = F^i-DK^* = F^i-DR^{-1}D^TM_1^*N_1^* \tag{4.32}$$

and the expected values of J is

$$\tilde{J} = \frac{1}{2n}\,\mathrm{Trace}\,[\sum_{i=1}^{2}\sum_{j=1}^{2}M_{ij} + M_{33}], \quad M_{ji} = M_{ij}^T. \tag{4.33}$$

Of course, in the actual solution of (4.31), the fact that (4.29a) is decoupled from (4.29b), (4.29c), and that (4.30a) is decoupled from (4.30b), (4.30c) can be utilized.

4.4. GHR and Model Reduction by the Restricted QL Algorithm

From previous sections we have demonstrated how existing model reduction techniques may be adapted and applied for constructing reduced order models. Two particular features of the modeling procedure should be noted: (a) the utilization of the GHR structure and (b) the adjustment of only a subset of parameters through a well-defined procedure. Nevertheless, each approach possesses certain drawbacks. Singular perturbation is only applicable when two-time scale property exists; norm minimization is not invariant with respect to system representations and finally output regulator formulation cannot guarantee existence and uniqueness of solutions to (4.22).

We now present a general method based on the GQL theory developed in Chapter 3. The approach taken here is to: (I) transform (2.9) into its GHR by chained aggregation; (II) choose a particular decomposition into S^i and $\bar{S}^i$ based on structural considerations*; (III) apply the GQL algorithm to the residual $\bar{S}^{i-1}$; (IV) construction of reduced order model by combining S^{i-1} and the reduced equivalent of $\bar{S}^{i-1}$ that exhibits the dominant dynamics in $\bar{S}^{i-1}$.

Consider a GHR of a LTI system and assume that based on the internal structure of the GHR, plausible model may have dimension $p_i = r_1 + \cdots + r_i$, where r_j, $j = 1,\ldots,i$ is the dimension of S_1 through S_i in the GHR. Then for a model of dimension p_i, the plant matrix F^i and the control matrix G^i are then seen as rational initial guesses in the construction of a reduced order model. We propose here that the interaction term should be compensated and that this generally be accomplished by the adjustment of some parameters in F^i and G^i when the interaction term is dropped. We now describe how the selection of parameters to be adjusted and the subsequent adjustment of these parameters can be achieved methodologically by applying the restricted QL algorithm.

---

*By structural considerations, we mean tests for controllability and observability; trajectory and eigenvalue perturbation analysis, etc. selection of reduced order model order may also be carried out as described in [45-46].

Decompose the plant matrix F as

$$F = \begin{bmatrix} F^{i-1} & E^{i-1} \\ H^{i-1} & \bar{F}^{i-1} \end{bmatrix} \triangleq \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \equiv F_1. \tag{4.34}$$

Note that from the introduced notation the following relation holds between F^i and F^{i-1}:

$$F^i = \left[\begin{array}{c|c} & 0 \\ & \vdots \\ F^{i-1} & 0 \\ & F_{i-1,i} \\ \hline F_{i,1} \cdots F_{i,i-1} & F_{i,i} \end{array}\right]. \tag{4.35}$$

It also follows that:

$$B_1 = \begin{bmatrix} 0 & 0 \cdots\cdots\cdots 0 \\ 0 & 0 \cdots\cdots\cdots 0 \\ \vdots & \\ 0 & 0 \cdots\cdots\cdots 0 \\ F_{i-1,i} & 0 \cdots\cdots\cdots 0 \end{bmatrix} \tag{4.36}$$

and that $F_{i,1},\ldots,F_{i,i-1}$ in (4.35) are the top layer blocks in C_1, while $F_{i,i}$ is the upper-left corner block matrix of D_1.

Define now the similarity transformation

$$T = \begin{bmatrix} I & 0 \\ 0 & L \end{bmatrix}$$

the block-partition of T being compatible with (4.34) and define L such that it achieves the generalized QL decomposition of D_1, i.e.

$$D_1 = QL$$

for some unitary matrix Q. It is immediately clear from the properties of the generalized QL algorithm that this transformation achieves the following:

(i) T defines a similarity transformation and thus <u>all eigenvalues of F are preserved</u>

(ii) L defines a QL decomposition of D_1 and thus <u>all eigenvalues of D_1 are preserved.</u>

Consider now a sequence of iterations such that in the k-th iterate we have

$$F_k = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix}, \quad T_k = \begin{bmatrix} I & 0 \\ 0 & L_k \end{bmatrix}, \quad D_k = Q_k L_k. \qquad (4.37)$$

Then, since $F_{k+1} = T_k F_k T_k^{-1}$ we have

$$F_{k+1} = \begin{bmatrix} A_k & B_k L_k^{-1} \\ L_k C_k & L_k D_k L_k^{-1} \end{bmatrix}, \quad G_{k+1} = \begin{bmatrix} G_k^{i-1} \\ L_k \bar{G}_k^{i-1} \end{bmatrix}. \qquad (4.38)$$

Thus after a number of iterations we observe that letting $L = L_k L_{k-1} \cdots L_1$, the following can be concluded:

(a) $A_{k+1} = A_1 = F^{i-1}$ and thus F^{i-1} has remained unaffected;

(b) $B_{k+1} = B_1 L^{-1}$. Because of (4.36) $F_{i-1,i}$ has been adjusted;

(c) $C_{k+1} = LC_1$ and thus the blocks $F_{i,j}$, $j = 1,\ldots,i-1$ have been adjusted;

(d) $D_k = LD_1 L^{-1}$ and thus $F_{i,i}$, the left upper corner block of D_1, has been adjusted;

(e) The block $F_{i,i+1}$ next to $F_{i,i}$ in D_1 has been reduced in norm;

(f) The eigenvalues of D_1 have been ordered in dominance and the dominant eigenvalues of D_1 are now contained in the adjusted F_{ii};

(g) $G_{k+1}^{i-1} = G_1^{i-1}$ and thus $G_1,\ldots,G_{i-1}$ is unaffected by the transformation;

(h) $\overline{G}_{k+1}^{i-1} = L\overline{G}_{1}^{i-1}$ and thus G_i has been adjusted.

Therefore, if based on the structural properties of the GHR and the physical insight concerning the system being considered, it is decided that a reduced order model containing the first i layers of the GHR should be used as the basis for constructing a reduced order model of dimension p_i, then the restricted QL algorithm provides a methodological way for modifying those parameters of the model contained in the blocks $F_{i-1,i}$, $F_{i,j}$, $j = 1,\ldots,i$ and G_i. The final result obtained by neglecting $F_{i,i+1}$ in the resulting RQL processed GHR is a model characterized by the following plant and control matrices:

$$\tilde{F}^i = \begin{bmatrix} F_{1,1} & F_{1,2} & \cdots & 0 & 0 \\ F_{2,1} & F_{2,2} & \cdots & 0 & 0 \\ \vdots & \vdots & & & \vdots \\ F_{i-1,1} & F_{i-1,2} & \cdots & F_{i-1,i-1} & \tilde{F}_{i-1,i} \\ \tilde{F}_{i,i} & \tilde{F}_{i,2} & & \tilde{F}_{i,i-1} & \tilde{F}_{i,i} \end{bmatrix}, \quad G = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_{i-1} \\ \tilde{G}_i \end{bmatrix} \tag{4.39}$$

where the result may be interpreted as if the coupling term $F_{i,i+1}z_{i+1}$ has been approximated via a linear combination

$$F_{i,i+1}z_{i+1} \approx \sum_{j=1}^{i} (\tilde{F}_{i,j} - F_{i,j})z_j + (\tilde{G}_i - G_i)u$$

with additional adjustment of the parameters in the block $F_{i-1,i}$.

The effect of applying the described method to construct a reduced order model can best be interpreted in the concept of dynamic equivalents of external system often used in power systems [47] and model reduction philosophy in general. From (4.37) and the form of T_k, we note the RQL algorithm in effect applies only the generalized QL transformations to the residual subsystem $\bar{S}^{i-1}$. By virtue of Theorem 3.6, the net effect is the shifting of the dominant dynamics in the residual towards leading submatrices in D_k. Thus the leading subsystem in the residual represents a reduced equivalent of the residual. Specifically, suppose that a decomposition has been made and (F^{i-1}, G^{i-1}) defines a desired but fixed internal subsystem while $(\bar{F}^{i-1}, \bar{G}^{i-1})$ characterizes the external subsystem with $F_{i-1,i}z_i$ serving as a coupling between the two. This is illustrated in Fig. 4.1. Applying the RQL iterations to system of Fig. 4.1, this amounts to obtaining an equivalent of the external subsystem $\bar{S}^{i-1}$ and replacing $\bar{S}^{i-1}$ by its dominant approximate (order $r_i < n-p_{i-1}$) after a finite number of RQL iterations. This is shown in Fig. 4.2. The reduced order model is then the composition of the untouched internal subsystem plus the newly obtained dynamic equivalent of the residual, Fig. 4.3. As seen from (4.39), the internal model of the system remains unaffected by the equivalencing procedure, while the interconnections and the model of the equivalent have been adjusted and constructed, respectively.

The procedure is therefore intuitively appealing and provides a result that is desirable, and in fact, is often pursued in practical applications. The procedure however, seems to provide more than is usually obtainable by existing methods. First, there is a rational choice available concerning the dimension of the model, and in fact a collection of reduced order models of various system order. Second, most of the model parameters do not need to be adjusted (at least in principle) and while the adjustment of those that do have to be modified is done in a straightforward and well-defined manner. Moreover, the parameter adjustment procedure has an appealing interpretation in terms of interconnections, equivalents and

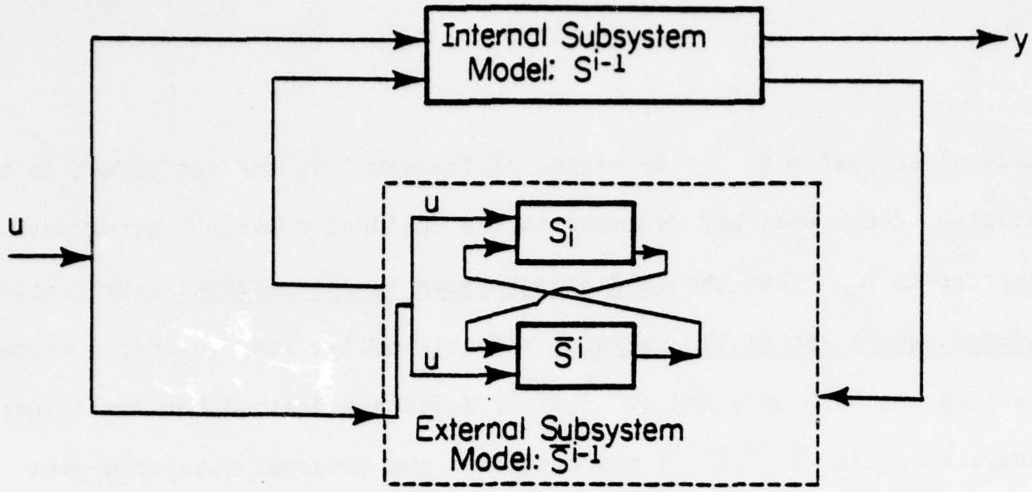


Figure 4.1. Decomposition of GHR for the application of RQL algorithm.

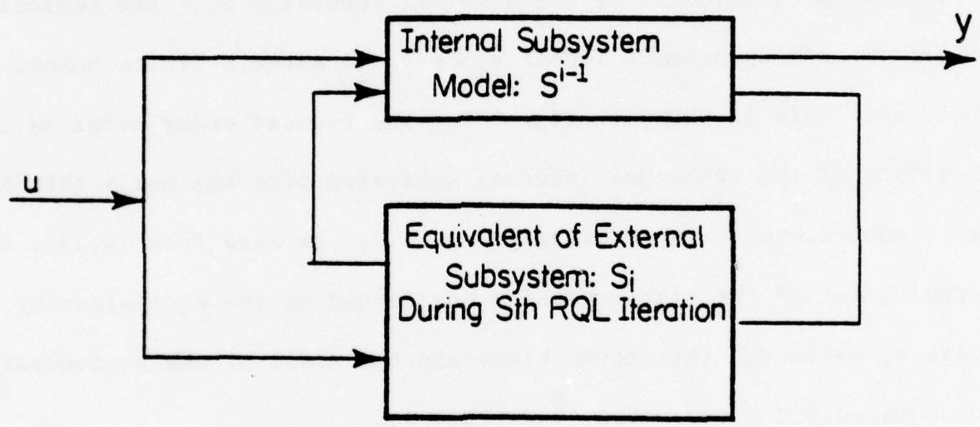


Figure 4.2. Construction of reduced order model by RQL algorithm.

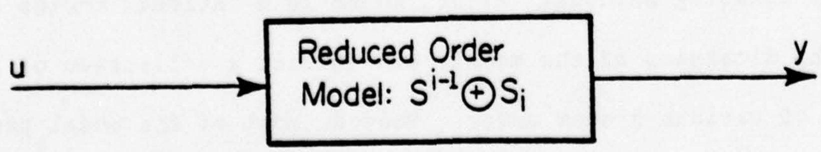


Figure 4.3. Effective reduced order model of the large scale system via RQL algorithm.

reduced order model. And finally in this reduced order modeling procedure not only do we define precisely the equivalent of a system interconnected to another system, but also in the process of obtaining it we are not forced to make any assumptions concerning the control efforts appearing in the external system. The control variables are not lost in the equivalent as is the case in all other equivalencing procedures known to us, but remain present in the reduced order model. The effect of applying RQL algorithm to a GHR for the construction of reduced order model will be studied via examples in Chapter 6 while the application of other methods will be subject to further investigation and will be reported elsewhere. As a remark, we believe there is a close connection between order reduction and weak observability. Recently Moore has used singular value decomposition to quantify weak observability [68-69]. However such results appeared late in the research stage and are not studied in this thesis.

4.5. Robustness of Reduced Order Model for Control Design - A Stability Analysis

One of the fundamental purposes in constructing reduced order models arises from the need for such models in various synthesis or control problems. Although a control design based on a reduced order model may be satisfactory when applied to the reduced model, e.g. stabilizes the closed-loop reduced order model, it is possible that when such a control is applied to the full order actual system, the performance is not satisfactory, e.g. the system becomes unstable. Thus a basic test of validity of the reduced order model for control design can be assessed in terms of stability properties. Other criteria may be related to suboptimality of cost functions if optimization is important. We will only consider the stability question in this thesis. Sufficient conditions for a control strategy selected using the reduced order model to stabilize the actual system are derived. Two different approaches are taken. In the first, stability is guaranteed using Lyapunov functions and bounds on system matrix perturbations. In the second, a frequency domain approach is adopted.

4.5.1. Problem Statement

Consider the GHR of (2.9) arbitrarily decomposed into S^i and $\bar{S}^i$:

$$\begin{bmatrix} \dot{z}^i \\ \dot{\bar{z}}^i \end{bmatrix} = \begin{bmatrix} F^i & E^i \\ H^i & \bar{F}^i \end{bmatrix} \begin{bmatrix} z^i \\ \bar{z}^i \end{bmatrix} + \begin{bmatrix} G^i \\ \bar{G}^i \end{bmatrix} u \tag{4.40a}$$

$$y = [D^i \quad 0] \begin{bmatrix} z^i \\ \bar{z}^i \end{bmatrix} \tag{4.40b}$$

where $z = [T_r^T \quad \bar{T}_r^T]^T x$. Assume $\bar{F}^i$ is stable. Let the approximate aggregated model of (2.9) be

$$\dot{z}_r = F^i z_r + G^i u \tag{4.41a}$$

$$\hat{y} = D^i z_r . \tag{4.41b}$$

Suppose the control law

$$u = K z_r \tag{4.42}$$

is designed based on (4.41) and stabilizes (4.41). Furthermore, suppose in controlling (2.9) (equivalently (4.40)), the control (4.42) is implemented as

$$u = K z^i = K T_r x . \tag{4.43}$$

Then the problem is to obtain conditions under which stability is preserved for the closed loop system

$$\dot{x} = (A + BKT_r)x . \tag{4.44}$$

4.5.2. A Stability Condition Via Perturbation Analysis

Since stability is invariant to linear transformations, then stability of (4.44) is equivalent to the stability of

$$\begin{bmatrix} \dot{z}^i \\ \dot{\bar{z}}^i \end{bmatrix} = \begin{bmatrix} F^i+G^iK & E^i \\ H^i+\bar{G}^iK & \bar{F}^i \end{bmatrix} \begin{bmatrix} z^i \\ \bar{z}^i \end{bmatrix} . \tag{4.45}$$

Introduce the notation

$$F_c = \begin{bmatrix} F^i+G^iK & 0 \\ H^i+\bar{G}^iK & \bar{F}^i \end{bmatrix} \qquad \Delta F = \begin{bmatrix} 0 & E^i \\ 0 & 0 \end{bmatrix} \tag{4.46}$$

By assumption, (F^i+G^iK) is stable because (4.42) stabilizes (4.41). Since $\bar{F}^i$ is also stable, then there exists a positive definite and symmetric matrix P such that

$$PF_c + F_c^TP = -Q \tag{4.47}$$

where Q is any positive definite and symmetric matrix.

Let

$$v = z^TPz \tag{4.48}$$

be a possible Lyapunov function for (4.45). The objective is to bound ΔF such that $v > 0$ and $\dot{v} \leq 0$ when evaluated along (4.45).

Definition 4.1: The control law (4.43) is said to be robust* if stability of $(F^i + G^iK)$ implies (4.44) is stable.

Let $\|\cdot\|$ denote the Euclidean norm† and $\lambda_{min}(A)$, $\lambda_{max}(A)$ denote the minimum and maximum eigenvalues of A respectively. Furthermore, for a given

---

*The term robustness has been defined elsewhere [61,62] to mean stability or system structural properties under parameter variations.

†Given a n×m matrix K, the Euclidean norm, $\|K\|$ is defined as $\|K\| = \{\sum_{i=1}^{n} \sum_{j=1}^{m} |k_{ij}|^2\}^{\frac{1}{2}}$ while the spectral norm $\|K\|_s \equiv \{\lambda_{max}(KK^T)\}^{\frac{1}{2}}$.

$Q > 0$, let P be defined by (4.47). Then Theorem 4.1 gives sufficient conditions for control (4.43) to be robust in terms of the norm of perturbation ΔF. The proof is given in Appendix C.

Theorem 4.1: The control law (4.43) is robust if the neglected Hessenberg block $F_{i,i+1}$ is such that

$$\|\Delta F\| \leq \tfrac{1}{2}\,\frac{\lambda_{min}(Q)}{\lambda_{max}(P)} \,. \tag{4.49}$$

We close this section by noting the sufficient condition (4.49) is very easy to check. Due to the block triangular structure of F_c, P may be solved for using lower order calculations.

4.5.3. A Stability Condition Via Frequency Domain Analysis

In this section, an alternative approach is taken to examine stability of (4.44). The method developed in Section 4.5.2 can be conservative due to the use of matrix norms. The frequency domain method developed here offers additional insights without introducing such matrix boundings.

Since (4.42) stabilizes the reduced order model (4.41), then for any $Q > 0$ (> 0 means positive definite and symmetric) there exists a $P > 0$ such that

$$P(F^i + G^iK) + (F^i + G^iK)^TP = -Q. \tag{4.50}$$

Furthermore (4.45) still holds. Define a Lyapunov function for (4.45) as

$$v = \tfrac{1}{2} z^T\tilde{K}z \tag{4.51}$$

where

$$\tilde{K} = \begin{bmatrix} P & 0 \\ 0 & M \end{bmatrix} \tag{4.52}$$

and M is some positive definite and symmetric matrix. Then differentiating

(4.51) along (4.45), we have from (4.46)

$$\dot{v} = \frac{1}{2} z^T[\Delta F^T\tilde{K} + F_c^T\tilde{K} + \tilde{K}F_c + \tilde{K}\Delta F]z \tag{4.53}$$

since

$$\tilde{K} = \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & M \end{bmatrix}$$

and denote

$$F = [F_{oi,j}] \qquad F_c = [F_{cij}] \qquad i,j = 1,2$$

After some algebraic manipulation, (4.53) becomes

$$\dot{v} = \tfrac{1}{2}[z^{iT} \quad \bar{z}^{iT}] \begin{bmatrix} -Q & PF_{012} + F_{c21}^T M \\ F_{012}^T P + MF_{c21} & MF_{022} + F_{022}^T M \end{bmatrix} \begin{bmatrix} z^i \\ \bar{z}^i \end{bmatrix}$$

$$= \tfrac{1}{2} z^T \Lambda(M) z \tag{4.54}$$

where

$$\Lambda(M) = \begin{bmatrix} -Q & PF_{012} + F_{c21}^T M \\ F_{012}^T P + MF_{c21} & MF_{022} + F_{022}^T M \end{bmatrix}. \tag{4.55}$$

Thus (4.45) is stable if and only if there exists a $M > 0$ such that $\Lambda(M) < 0$. The subsequent development is aimed to obtain conditions under which <u>M exists and $\Lambda(M) < 0$</u>. Before proceeding, let us note two identities. Define

$$\pi(M) \equiv -Q - (PF_{012} + F_{c21}^T M)(MF_{022} + F_{022}^T M)^{-1}(F_{012}^T P + MF_{c21}) \tag{4.56}$$

$$R(M) \equiv MF_{022} + F_{022}^T M + (MF_{c21} + F_{012}^T P)Q^{-1}(F_{c21}^T M + PF_{012}) \tag{4.57}$$

$$L_1 \equiv (MF_{022} + F_{022}^T M)^{-1}(F_{012}^T P + MF_{c21}) \tag{4.58}$$

$$L_2 \equiv -(MF_{c21} + F_{012}^T P)Q^{-1}. \tag{4.59}$$

Clearly Q^{-1} and $(MF_{022} + F_{022}^T M)^{-1}$ exists. Then

Identity I:

$$\Lambda(M) = \begin{bmatrix} I & 0 \\ L_2 & I \end{bmatrix} \begin{bmatrix} -Q & 0 \\ 0 & R(M) \end{bmatrix} \begin{bmatrix} I & 0 \\ L_2 & I \end{bmatrix}^T \tag{4.60}$$

Identity II:

$$\Lambda(M) = \begin{bmatrix} I & 0 \\ L_1 & I \end{bmatrix}^T \begin{bmatrix} \pi(M) & 0 \\ 0 & MF_{022} + F_{022}^T M \end{bmatrix} \begin{bmatrix} I & 0 \\ L_1 & I \end{bmatrix} \tag{4.61}$$

The two identities may be easily verified using the definitions (4.56)-(4.59). Due to the identities, we have the following theorem:

<u>Theorem 4.2</u>: The following conditions are equivalent

(i) $\Lambda(M) < 0$

(ii) $R(M) < 0$ and $-Q < 0$

(iii) $\pi(M) < 0$ and $(MF_{022} + F_{022}^T M) < 0$.

<u>Remark</u>: An interesting observation follows from Theorem 4.2. If indeed a $M > 0$ exists such that $\Lambda(M) < 0$, (iii) implies the existence of a $\overline{Q} > 0$ with

$$MF_{022} + F_{022}^T M = -\overline{Q} \tag{4.62}$$

Since (4.62) is a Lyapunov equation, $F_{022} \equiv \overline{F}^i$ must be stable. Such an implication may be justified as follows. If $F_{i,i+1}$ is neglected and (4.42) is designed, $\overline{S}^i$ is uncontrollable. Thus the stability of $\overline{F}^i$ is not unreasonable as a requirement in order for $(F_c + \Delta F)$ to remain stable for small perturbation $F_{i,i+1}$.

Theorem 4.2 indicates either (ii) or (iii) may be used to prove (i). Without loss of generality, let us guarantee (ii) in order to derive sufficient conditions for the existence of a $M > 0$. The conditions for stability are summarized in Theorem 4.3. The details of the proof may be found in Appendix D.

<u>Theorem 4.3</u>: If

(a) $(\tilde{F}, F_{c21})$ is completely controllable and $(\tilde{F}, E^{i})$ is completely observable

(b) $\mathrm{Re}\lambda(\tilde{F}) < 0$

(c) $I - F_{c21}^{T}(-j\omega I - \tilde{F}^{T})^{-1}\tilde{Q}(j\omega I - \tilde{F})^{-1}F_{c21} > 0 \quad \forall\omega$

where $\tilde{F} \equiv \bar{F}^{i} + (H^{i} + \bar{G}^{i}K)PE^{i}$, $\tilde{Q} \equiv E^{iT}PPE^{i}$, $F_{c21} = H^{i} + \bar{G}^{i}K$. Then the control law (4.43) is robust.

CHAPTER 5

GHR AND CONTROL SYNTHESIS BY FEEDBACK COMPENSATION

Consider a large scale interconnected system

$$\dot{x}_i = A_{ii}x_i + B_{ii}u_i + \mu_i + E_i f \tag{5.1a}$$

$$\mu_i = \sum_{\substack{j=1\\ j\neq i}}^{\nu} A_{ij}y_j \tag{5.1b}$$

$$y_i = C_i x_i \quad , \; i=1,\ldots,\nu \tag{5.1c}$$

where $x_i \in R^{n_i}$, $u_i \in R^{m_i}$, $\mu_i \in R^{n_i}$, $f \in R^q$, and $y_i \in R^{\gamma_i}$ are subsystem state, control, interconnection, disturbance, and output variables respectively. Suppose the system interconnection $y=[y_1^T \ldots y_\nu^T]^T$ are desired to track the output $\hat{y}$ of a reference system

$$\dot{w}' = F_o' w' + G' u_o + W' f \tag{5.2a}$$

$$\hat{y} = D' w' \tag{5.2b}$$

$$u_o = L' w' \tag{5.2c}$$

where $w' \in R^p$, $u_o \in R^m$, $m = \sum_{i=1}^{\nu} m_i$, $f \in R^q$, $\hat{y} \in R^{\gamma}$, $\gamma = \sum_{i=1}^{\nu}\gamma_i$ are the reference system state, control, disturbance and output variables respectively, and u_o is a fixed reference system control. This chapter considers the synthesis of local controls u_i, $i=1,\ldots,\nu$ such that

$$e(t) \equiv \hat{y}(t) - y(t)$$

approaches zero asymptotically. Moreover, u_i assumes the apriori form

$$u_i = \pi_i u_{io} + K_i e + M_{ii} x_i \tag{5.3}$$

where

$$u_o = [u_{1o}^T \ldots u_{\nu o}^T]^T \quad , \; u = [u_1^T \ldots u_\nu^T]^T. \tag{5.4}$$

Also it is desired to obtain (5.3) without excessive information transfer between subsystems and parallel computation by individual subsystem can be achieved.

Note because the local control (5.3) is a function of composite error $e(t)$, some information transfer is required during computation and implementation. We propose a hierarchical structure such that information coordination and decentralization be performed by a higher level coordinator. One motivation stems from the structure of (5.1). If the interconnection input μ_i is viewed as an external disturbance and its dynamic evolution is known to each subsystem, then the computation at local subsystem level can in principle be carried out in parallel and independent of other subsystems. To specify the interconnection y_i, $\forall i$, we propose that a coordinator exists at the secondary level whose main concern is the behavior of y_i according to (5.2). For coordination and control implementation, the coordinator is assumed to be in constant communication with the subsystems so that $e(t)$ is readily computed by the coordinator. Functionally, the local control (5.3) is a composite of the coordinator and decentralized subsystem control. The component $\pi_i u_{io}$ and $K_i e$ are supplied by the coordinator and the local state feedback $M_{ii} x_i$ is generated locally at subsystem level. The basic configuration is shown in Fig. 5.1.

It will be shown subsequently how insights gained from the GHR structure can be utilized to solve the above synthesis problem. The approach is to reformulate the reference system (5.2) into an nth order reference system (having the same input-output behavior between u_o and $\hat{y}$) with a structure similar to the GHR of the composite system formed by subsystems in

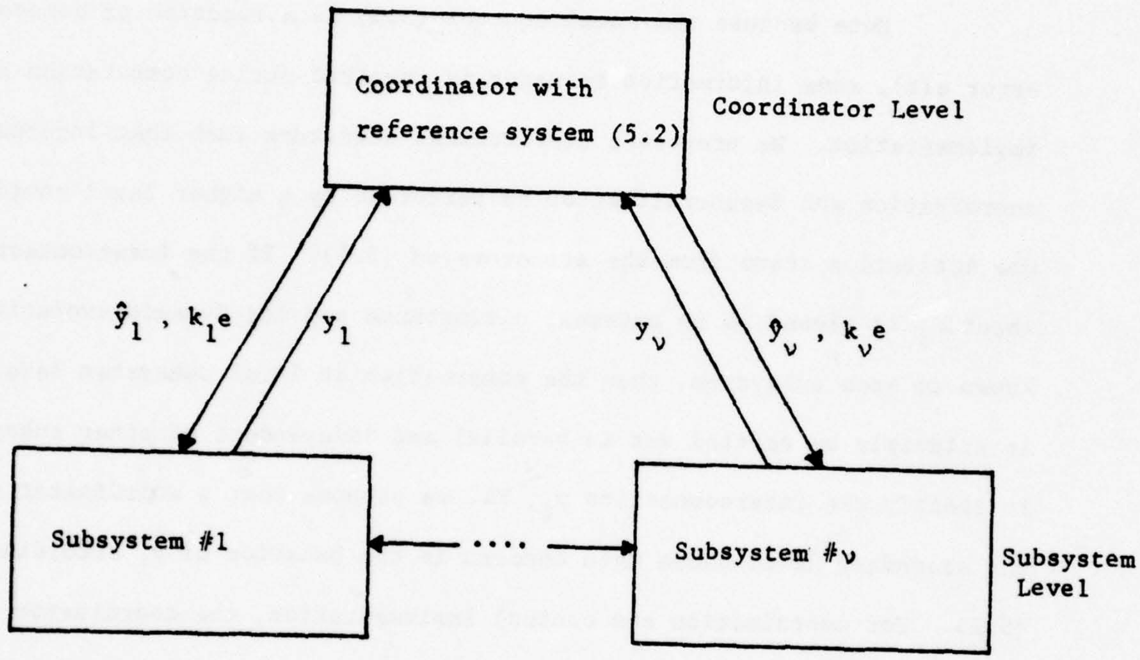


Figure 5.1. Proposed hierarchy of coordinator with reference model (5.2) and local subsystem model (5.1).

(5.1). The role of the coordinator will then be seen to coordinate decoupling of the composite system into an aggregate and a residual subsystems via local state feedback. The tracking of $\hat{y}$ by y is obtained by first compensating the interaction $F_{i,i+1}z_{i+1}$ between the aggregate and the residual and then applying the necessary output following controls. We emphasize the key feature of the proposed method is the separate control of the aggregate and the residual by local state feedback controls once decoupling is introduced. We will proceed to show when this decoupling exists, the local control (5.3) can be decomposed into three basic components. One component provides the necessary feedback compensation to achieve aggregability and decoupling. The second serves to control the aggregate and the third, if exists, represents a freedom in control to shape the dynamics of the residual.

The chapter is organized as follow. In Section 5.1, the synthesis problem is formulated as an output model following problem. In Section 5.2, perfect output following is defined and the existence and properties of the local controller to follow a r_1th order reference system are examined. Section 5.3 extends the results of Section 5.2 to a reference system of order greater than r_1. Finally a decomposed design strategy is investigated in terms of the hierarchy introduced. The result is a precise definition of the coordinator and the subsystem level problems.

5.1 Preliminaries

For the reference system (5.2), we assume $p = p_i = \sum_{j=1}^{i} r_i$ where r_i are the subsystem dimensions of the GHR for the composite system formed by (5.1). Moreover, for the reference system (5.2) to be arbitrary, (F_o', G') is assumed to be controllable so that arbitrary dynamics may be obtained by a suitable reference system control (5.2c). (F_o', D') is assumed to be observable so that arbitrary reference dynamics is seen as arbitrary reference output $\hat{y}$. We shall also assume that the triplet (F_o', G', L') satisfies (for an suitable index i)

$$F^i + G^i L = (F_o' + G'L') \tag{5.5}$$

for some matrix L. Here (F^i, G^i) comes from the aggregate S^i of the GHR of the composite formed by (5.1). The reason for such an assumption is that the reference interconnections $\hat{y}$ should be obtained from a system that exhibits similar behavior between u and y from an input-output point of view, thus giving a degree of mutual harmony between y and $\hat{y}$. Equation (5.5) incorporates this reasoning by relating the reference system to the aggregate S^i which is an approximate reduced order model of the composite system. For simplicity, let $F_o' = F^i$, $G' = G^i$ in (5.5) and $D' = D^i$ in (5.2). Finally, we assume same class of disturbance acts on the reference system as on the aggregate S^i of the GHR (5.8). Thus $W' = W^i$.

The interconnected system (5.1) may be written as a composite system

$$\dot{x} = Ax + \sum_{i=1}^{\nu} B_i u_i + E f \tag{5.6a}$$

$$y = Cx \tag{5.6b}$$

where $x \in R^n$, $n = \sum_{i=1}^{\nu} n_i$, $y \in R^{\gamma}$, $\gamma = \sum_{i=1}^{\nu} \gamma_i$ and

$$A = \begin{bmatrix} A_{11} & A_{12}C_2 & A_{1\nu}C_{\nu\nu} \\ \vdots & \vdots & \vdots \\ A_{\nu 1}C_1 & A_{\nu 2}C_2 \cdots & A_{\nu\nu} \end{bmatrix} \qquad B = \begin{bmatrix} B_{11} & & 0 \\ & \ddots & \\ 0 & & B_{\nu\nu} \end{bmatrix} = [B_1 \; \dots \; B_\nu]$$

$$E = \begin{bmatrix} E_1 \\ \vdots \\ E_\nu \end{bmatrix} \quad C = \begin{bmatrix} C_{11} & & 0 \\ & \ddots & \\ 0 & & C_{\nu\nu} \end{bmatrix} \tag{5.7}$$

Using the chained aggregation procedure of Chapter 2, (5.6) may be equivalently expressed in GHR as

$$\dot{z} = Fz + Gu + Wf \tag{5.8a}$$

$$y = Dz \tag{5.8b}$$

where F, G and D are given by (2.11) and

$$W = HE = [W_1^T \; \dots \; W_k^T]^T \tag{5.9}$$

$$H = [C^T \;\; \bar{H}^T]^T \tag{5.10}$$

H being the generalized Hessenberg transformation matrix for (5.6).

Using (5.5), from an input-output point of view, the reference system (5.2) may be realized equivalently as:

$$\dot{w} = F_o w + Gu_o + Wf \tag{5.11a}$$

$$\hat{y} = D w \tag{5.11b}$$

$$u_o = Lw \tag{5.11c}$$

where $w \in R^n$ and G, W, D are same as in (5.8) except

$$F_o = \begin{bmatrix} F^i & 0 \\ H^i & \bar{F}^i \end{bmatrix} . \tag{5.12}$$

Since the control objective is the regulation of e(t) to zero and y, $\hat{y}$ being outputs of (5.8) and (5.11) respectively, we note y can follow $\hat{y}$ if (5.8) can be made to follow the reference system (5.11). In this sense (5.8) (or equivalently (5.6)) is the system and (5.11) is the reference model. A natural approach to the synthesis of (5.3) is a model following design [50 - 51]. Taking advantage of the GHR structure and y, $\hat{y}$ being the first subvector of z and w respectively, direct results on the existence of (5.3) can be obtained if we define the output error e(t) instead of the state error w(t) - z(t) as done in [50 - 51].

An examination of (5.8) and (5.11) indicates y will follow $\hat{y}$ if the coupling $F_{i,i+1}\, z_{i+1}$ between S^i and $\bar{S}^i$ may be compensated by a suitable local control (5.3) and the substate of z and w corresponding to F^i will track each other. In contrast to Section 4.3 where the coupling is approximated by feedback from the states of the aggregate, i.e. $v = \sum_{j=1}^{i} \delta F_{ij}\, z_j$ such that $\|F_{i,i+1} z_{i+1} - v\|$ is minimized, under appropiate conditions, a complete compensation is possible by feedback from the residual. Thus using (5.5), the tracking of y and $\hat{y}$ is emphasized in terms of the GHR by designing (5.3) to decouple S^i and $\bar{S}^i$ as an initial requirement and then achieving some sort of model following subsequently. From the GHR structure given in (5.8), we note that for S^1 to decouple* from $\bar{S}^1$, there must exists a residual feedback gain K such that

$$F_{12} z_2 = G_1 K z_2 \qquad (5.13)$$

*Note "decoupling" used here means S^i is not driven by $\bar{S}^i$. In general, feedforward exists from S^i to $\bar{S}^i$.

Clearly such a K exists if and only if $\mathcal{R}(F_{12}) \subset \mathcal{R}(G_1)$. On the other hand when (5.13) is not satisfied, it is not sufficient for a gain K to exist such that

$$F_{i,i+1} z_{i+1} = G_i K z_{i+1} , \quad i > 1 \tag{5.14}$$

in order for S^i to decouple from $\bar{S}^i$. Because if $\mathcal{R}\{G_j\} \neq \phi$, $j < i$, then

$$G_j K z_{i+1} \neq 0 , \quad j < i. \tag{5.15}$$

Thus coupling between S^i and $\bar{S}^i$ will be re-introduced in higher level subsystems in the GHR.

In the following section, the model output following problem defined at the beginning of the Chapter will be examined to see (I) Under what condition a local control (5.3) exists such that decoupling (achievement of aggregability) and output y following $\hat{y}$ of the reference system can be achieved? (II) What is the implication of (5.3) in terms of decentralized control and parallel computation such that a separation in design may be carried out? and (III) Is there additional freedom in control synthesis beyond achievement of aggregability and output following, and can it be used to shape the dynamics of the residual subsystem?

5.2 Complete Aggregability and Perfect Output Following by Local State Feedback

This section considers a reference system of dimension $p = r_1$. From Section 5.1, the problem is expressed as to investigate the possibility of compensating the coupling $F_{12} z_2$ at the first GHR layer and then achieving model following between the system and the reference system under a fixed reference control u_o. The extension to a reference system of dimension

$r_1 < p \leq n$ will be considered in Section 5.3. To distinguish between results of Section 5.3 which requires the definition of an extended output, we give the following definition.

Definition 5.1 The local control (5.3) achieves perfect output following if

$$\lim_{t\to\infty} e_i(t) = \lim_{t\to\infty} \hat{y}_i(t) - y_i(t) = 0 \ , \quad i = 1, \ldots \nu \tag{5.16}$$

Recall that the reference interconnection variables $\hat{y}_i$, $i = 1, \ldots \nu$ are governed by (5.2) which under (5.5) is equivalent to (5.11) where $F^i = F_{11}$ in (5.12). The objective now is to investigate the existence of local control (5.3) such that perfect output following results.

5.2.1 Derivation of the Local Control Law

From (5.6) and (5.11), we have

$$\dot{y} = CAx + \sum_{i=1}^{\nu} CB_i u_i + CEf \tag{5.17}$$

$$\dot{\hat{y}} = DF_o w + DGu_o + DWf \tag{5.18}$$

Subtract (5.17) from (5.18)

$$\dot{e} = DF_o w - CAX + DGu_o - \sum_{i=1}^{\nu} CB_i u_i + (DW - CE)f \tag{5.19}$$

since $W = HE$ and H has the general form of $H = [C^T \bar{H}^T]^T$, then $W_1 = CE$. Hence

$$DW - CE \equiv 0 \tag{5.20}$$

Next we can write

$$w = \begin{bmatrix} w_1 \\ \\ \bar{w}_1 \end{bmatrix} = \begin{bmatrix} \hat{y} \\ \\ \bar{w}_1 \end{bmatrix} = \begin{bmatrix} \hat{y} \\ \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \\ \bar{w}_1 \end{bmatrix} = D^T \hat{y} + \xi \tag{5.21}$$

Using (5.21) in (5.19), we finally get

$$\dot{e} = DF_o D^T \hat{y} + DF_o \xi - CAX + DGu_o - \sum_{i=1}^{\nu} CB_i u_i$$

$$= F_{11} e + (F_{11}C - CA)x + (CB\, u_o - \sum_{i=1}^{\nu} CB_i u_i) \tag{5.22}$$

If the imposed control law (5.3) is substituted, $\dot{e}$ becomes

$$\dot{e} = (F_{11} - CBK)e + CBu_o - \sum_{i=1}^{\nu} CB_i \pi_i u_{io} + (F_{11}C - CA)x - \sum_{i=1}^{\nu} CB_i M_{ii} X_i \tag{5.23}$$

where

$$K = \begin{bmatrix} K_1 \\ \vdots \\ K_\nu \end{bmatrix} \tag{5.24}$$

Thus from (5.23), we see perfect output following by control (5.3) requires:

(I) For $K \in R^{m \times r}$, $(F_{11} - CBK)$ be stable.

(II) There exists π_i, $i=1,\ldots\nu$ such that

$$CBu_o - \sum_{i=1}^{\nu} CB_i \pi_i u_{io} \equiv 0 \tag{5.25}$$

(III) There exists M_{ii} $i=1,\ldots\nu$ such that

$$(F_{11}C - CA)x - \sum_{i=1}^{\nu} CB_i M_{ii} x \equiv 0 \tag{5.26}$$

<u>Remark</u> If conditions (II) and (III) holds, then

$$\dot{e} = (F_{11} - CBK)e$$

Thus if $y(0) = \hat{y}(0)$. $e(t) \equiv 0 \ \forall t$. Condition (I) is required if initially $y(0)$ is not equal to $\hat{y}(0)$.

Theorem 5.1. For arbitrary u_o, the matrices π_i, $i=1,\ldots\nu$ always exists such that

$$CBu_o = \sum_{i=1}^{\nu} CB_i \pi_i u_{io} \tag{5.25'}$$

Furthermore

$$\pi_i = (C_i B_{ii})^+ (C_i B_{ii}) + [I - (C_i B_{ii})^+ (C_i B_{ii})] N_i \tag{5.27}$$

where N_i is a $m_i \times m_i$ arbitrary matrix and X^+ denotes the generalized inverse of X [12].

Proof of Theorem 5.1 relies on following Lemmas on generalized inverse.

Lemma 5.1 [12, Theorem 5.19]

If $W = NM$ where W, N and M are respectively $m \times n$, $m \times k$ and $k \times n$ matrices of rank k, then the solution of $Wx = b$, $x,b \in R^n$ which minimizes

(a) the sum of squares of the residual $r^T r$ where $r = b - Wx$

(b) the sum of squares of the unknown $x^T x$ is given by $x = W^+ b$ where

$$W^+ = M^T (MM^T)^{-1} (N^T N)^{-1} N^T \tag{5.28}$$

Remark Lemma 5.1 defines W^+ by (5.28). However, we note $r \equiv 0$ if and only if $b \in \mathcal{R}(W)$. In general, $r \neq 0$. It will be shown that conditions for the existence of freedom component in local control u_i is related to the class of W^+ that makes $r = 0$.

<u>Lemma 5.2.</u> [12, Theorem 5.20]

The generalized inverse W^+ of a $m \times n$ matrix has the following properties

(i) $W^+WW^+ = W^+$

(ii) $WW^+W = W$

(iii) WW^+ and W^+W are symmetric matrices.

The assumption in Lemma 5.1 that W with rank k can be decomposed into a producet of N and M each with maximal rank is assured by Lemma 5.3.

<u>Lemma 5.3</u> Any $m \times n$ matrix W having rank equals to k can be factored as $W = NM$ such that $N \in R^{m \times k}$ and $M \in R^{k \times n}$. Furthermore N, M have maximal rank.

Proof: Since $W \in R^{m \times n}$ has rank k, then there exists a $k \times k$ submatrix of W having rank equals k. Let T_1 and T_2 be Gaussian matrices [12] which transforms W into $\bar{W}$ by

$$T_1WT_2 = \bar{W} = [\bar{W}_{i,j}] \quad , \quad i, j = 1,2 \tag{5.29}$$

where $\bar{W}_{11} \in R^{k \times k}$ is nonsingular. This can always be done. Since T_i, $i = 1,2$ are non-singular, rank $\bar{W} = k$. Then the last $m-k$ rows of $\bar{W}$ are dependent. We have

$$\bar{W}_{21} = P\,\bar{W}_{11} \; , \quad \bar{W}_{22} = P\,\bar{W}_{12} \tag{5.30}$$

for some matrix P. Similarily, the last $n-k$ columns of $\bar{W}$ are linear combinations of the first k columns. Hence there exists a matrix Q such that

$$\bar{W}_{12} = \bar{W}_{11}Q \quad , \quad \bar{W}_{22} = \bar{W}_{21}Q \tag{5.31}$$

From (5.30) - (5.31),

$$P = \bar{W}_{21}\bar{W}_{11}^{-1} \quad , \quad Q = \bar{W}_{11}^{-1}\,\bar{W}_{12} \tag{5.32}$$

giving us

$$\bar{W}_{22} = P\bar{W}_{11}Q \tag{5.33}$$

(5.29) - (5.33) gives three different factorization of $\bar{W}$, namely

$$\bar{W} = \begin{bmatrix} I_k \\ P \end{bmatrix} [\bar{W}_{11}\ \bar{W}_{12}] = \begin{bmatrix} I_k \\ P \end{bmatrix} \bar{W}_{11}[I_k\ Q] = \begin{bmatrix} \bar{W}_{11} \\ \bar{W}_{21} \end{bmatrix} [I_k\ Q] = NM \tag{5.34}$$

where N, M each have rank equal to k. Using (5.29), $W = T_1^{-1}\,\bar{W}\,T_2^{-1} =$ $(T_1^{-1}\ N)(M\,T_2^{-1})$. This completes the proof.

Using Lemma 5.1 to Lemma 5.3, we also have

Lemma 5.4. Given a $m \times n$ matrix W such that rank $W = k$, then a generalized inverse W^+ always exists.

Proof of Theorem 5.1:

Using definition of C, B_i and π_i given by (5.27),

$$\sum_{i=1}^{\nu} CB_i\pi_i u_{io} = \sum_{i=1}^{\nu} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ C_iB_{ii} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \{(C_iB_{ii})^+(C_iB_{ii}) + [I - (C_iB_{ii})^+(C_iB_{ii})]N_i\}u_{io}$$

$$= \sum_{i=1}^{\nu} CB_i\,u_{io} + 0$$

$$= C\,B\,u_o \;.$$

The fact $(C_iB_{ii})^+$ exists is a consequence of Lemma 5.4. Moreover, from Lemma 5.2, $(C_iB_{ii})(C_iB_{ii})^+(C_iB_{ii}) = (C_iB_{ii})$. Hence (5.27) always exists and satisfies (5.25).

<u>Theorem 5.2</u>. If for $i = 1,\dots\nu$

$$[I - (CB_i)(CB_i)^+](F_{11}C - CA)\mathscr{J}_i = 0 \tag{5.35}$$

where $\mathscr{J}_i$ is defined by

$$C\mathscr{J}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ C_i \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5.36}$$

then M_{ii} exists such that

$$(F_{11}C - CA)x = \sum_{i=1}^{\nu} CB_iM_{ii}x_i \tag{5.26'}$$

Furthermore

$$M_{ii} = (CB_i)^+(F_{11}C - CA)\mathscr{J}_i + [I - (C_iB_{ii})^+(C_iB_{ii})]\,P_i \tag{5.37}$$

where P_i is a $m_i \times n_i$ arbitrary matrix.

Proof: (5.26') may be rewritten as

$$\sum_{i=1}^{\nu} (F_{11}C - CA)\mathscr{J}_ix_i = \sum_{i=1}^{\nu} CB_iM_{ii}x_i$$

In order for the equation to hold for arbitrary x_i, we must require

$$(F_{11}C - CA)\mathcal{V}_i = CB_iM_{ii} \tag{5.38}$$

Using (5.37) in (5.38) and applying (5.35), we have

$$\begin{aligned}\sum_{i=1}^{\nu} CB_iM_{ii}x_i &= \sum_{i=1}^{\nu} (CB_i)\{(CB_i)^+(F_{11}C - CA)\mathcal{V}_i + [I - (C_iB_{ii})^+(C_iB_{ii})]P_i\}x_i \\ &= \sum_{i=1}^{\nu} (CB_i)(CB_i)^+(F_{11}C - CA)\mathcal{V}_ix_i \\ &= \sum_{i=1}^{\nu} (F_{11}C - CA)\mathcal{V}_ix_i \\ &= (F_{11}C - CA)x.\end{aligned}$$

Combining Theorem 5.1 and 5.2, we now state the main result of Section 5.2.

Theorem 5.3. Assume the pair (A,C) is not completely aggregable. Consider the LTI system (5.6) expressed in its GHR (5.8) and the reference system (5.11). Then perfect output following is achieved by the local control

$$u_i = \pi_i u_{io} + k_ie + M_{ii}x_i \ , \ i = 1\ldots\nu \tag{5.39}$$

if

(i) (F_{11}, CB) is stabilizable

(ii) $[I - (CB_i)(CB_i)^+](F_{11}C - CA)\mathcal{V}_i = 0, \quad i = 1,\ldots\nu$

Proof: In order to achieve perfect output following, we must have $(F_{11} - CBK)$ be stable for some K and π_i, M_{ii} exists $\forall i$. The stability of $(F_{11} - CBK)$ is guaranteed if (F_{11}, CB) is stabilizable. On the other hand, the existence of π_i and M_{ii} are guaranteed by Theorems 5.1 and 5.2. With π_i and M_{ii} chosen according to Theorems 5.1 and 5.2, (5.23) reduces to

$$\dot{e} = (F_{11} - CBK)e \tag{5.40}$$

Hence $\lim_{t\to\infty} e(t) = 0$ if condition (i) holds. This completes the proof.

Using the expression of π_i and M_{ii} given by (5.27) and (5.37), we note the local control (5.3) is equivalently expressed as

$$u_i = u_{i1} + u_{i2} \tag{5.41}$$

where

$$u_{i1} = (C_iB_{ii})^+(C_iB_{ii})u_{io} + K_ie + (CB_i)^+(F_{11}C - CA)\mathscr{J}_ix_i \tag{5.42}$$

$$u_{i2} = [I - (C_iB_{ii})^+(C_iB_{ii})](N_i u_{io} + P_ix_i). \tag{5.43}$$

u_{i1} is used to achieve complete aggregability and perfect output following while u_{i2} provides freedom in control synthesis that is non-interactive* with u_{i1}.

Corollary 5.3.1. If (A,C) is completely aggregable, then perfect output following is achieved by (5.3) if (F_{11}, CB) is stabilizable.

The proof of Corollary 5.3 follows from that of Theorem 5.3 if we note (A,C) completely aggregable implies $(F_{11}C - CA) \equiv 0$.

Remark. If instead of assuming that e is formed by the coordinator and the component K_ie in u_i is supplied by the coordinator, we can assume e_i is formed locally by individual subsystem. Then a decentralized control

$$u_i = \pi_i u_{io} + K_i e_i + M_{ii}x_i \tag{5.44}$$

can be sought. It can be shown that with π_i, M_{ii} chosen according to Theorems 5.1 and 5.2.

---

*By non-interactive, we mean the control action of u_{i2} will not cancel or affect the control action of u_{i1}. Thus the application of u_{i2} will not destroy aggregability or perfect output following brought about by u_{i1}.

$$\dot{e} = (F_{11} - CBK)e \tag{5.45}$$

where $K \in \mathcal{K} = \{K / K = \text{block diag } (K_1,\ldots K_\nu)\}$. The existence of a stabilizing gain K is studied by Davison [52] who states that if the triple $\{F_{11}, CB, I\}$ has no unstable fixed modes, then a stabilizing $K \in \mathcal{K}$ exists.

5.2.2. Properties of the Local Control Law

Theorem 5.3 and (5.41) indicates that u_i possess a component u_{i2}

$$u_{i2} = [I - (C_i B_{ii})^+ (C_i B_{ii})] (N_i u_{io} + P_i x_i) , \quad i = 1,\ldots \nu \tag{5.46}$$

which, if non-zero, represents the freedom available in shaping the dynamics of the residual subsystem beyond achieving of aggregation and perfect output following by u_{i1}. It is important to examine under what restriction on the system structures that this freedom may not exist. It will be shown that rank of the matrix product CB characterizes the available freedom and the number of inputs in excess the number of outputs is crucial. We mention here that u_{i2} will be denoted as u_{ic} in Section 5.4 where the local control u_i is decomposed into u_{ia} for controlling the aggregate, u_{ib} for achievement of aggregation and u_{ic} for controlling the aggregate.

From (5.46), for a given pair (B,C), $u_{i2} \equiv 0$ for arbitrary matrices N_i and P_i if and only if $[I - (C_i B_{ii})^+ (C_i B_{ii})] = 0$. Three major cases occur: (a) CB has maximal rank which in turn may be subdivided into two subclasses, namely rank $CB = m \leq r$ or rank $CB = r < m$; (b) CB has less than maximal rank, i.e. rank $CB < \min (r,m)$; (c) $CB \equiv 0$. Each class will now be studied separately.

Class I rank (CB) = min(r,m)

Class I-a rank (CB) = $m \leq r$

Since CB has maximal rank, there exists $(CB)^+$ such that

$$(CB)^+ = [(CB)^T(CB)]^{-1}(CB)^T \tag{5.47}$$

and

$$(CB)^+(CB) = I_m \tag{5.48}$$

which implies

$$(C_iB_{ii})^+(C_iB_{ii}) = I_{m_i} \tag{5.49}$$

Using (5.47) - (5.49) in (5.46), clearly

$$u_{i2} \equiv 0 \tag{5.50}$$

Hence the non-interactive freedom component is lost in the sense $u_{i2} = 0$.

Case I-b rank CB = $r < m$

Using the definition of B and C in (5.7),

$$CB = \text{block diag. } \{C_1B_{11}, \ldots C_\nu B_{\nu\nu}\} \tag{5.51}$$

Thus rank $C_iB_{ii} = r_i < m_i$. Denote dim $(\dot{C})$ as the dimension of a subspace C, then

$$\dim \mathcal{R}(C_iB_{ii}) = r_i, \ \dim \ker (C_iB_{ii}) = m_i - r_i. \tag{5.52}$$

In addition, $(C_iB_{ii})^+(C_iB_{ii})$ is a projection matrix [53] into range space of C_iB_{ii} and $[I - (C_iB_{ii})^+(C_iB_{ii})]$ is a projection matrix from a m_i dimensional space into a $m_i - r_i$ dimensional subspace. Thus $[I - (C_iB_{ii})^+(C_iB_{ii})] \neq 0$ in general. Hence $u_{i2} \neq 0$ in general and there are $m_i - r_i$ degrees of freedom in u_i.

We can summarize the results for Class I as follow:

Theorem 5.4. Assume (A,C) is not completely aggregable and CB has maximal rank. Then there exists freedom in local control (5.3) beyond achieving complete aggregability and output following, i.e. $u_{i2} \neq 0$ if and only if $r_i < m_i$.

Class II rank $CB = \ell < \min(r,m)$

(5.51) implies $(C_iB_{ii}) \in R^{r_i \times m_i}$, thus either rank $(C_iB_{ii}) = \min(r_i,m_i)$ or rank $(C_iB_{ii}) = \ell_i < \min(r_i,m_i)$. Let us introduce the index set

$$\mathcal{J}_1 = \{i_t,\ t = 1,2,\ldots s/(C_{i_t}B_{i_ti_t}) \text{ has maximal rank}\}$$

$$\mathcal{J}_2 = \{i_t,\ t = s+1,\ldots k/(C_{i_t}B_{i_ti_t}) \text{ has less than maximal rank}\}. \tag{5.53}$$

Since CB has less than maximal rank, there is at least one index belonging to $\mathcal{J}_2$. In each group, $(C_iB_{ii})^+$ exists by Lemma 5.4. In particular, for $i \in \mathcal{J}_1$, we have either

$$(C_iB_{ii})^+(C_iB_{ii}) = I_{m_i} \qquad \text{if } m_i \leq r_i \tag{5.54}$$

or

$$(C_iB_{ii})(C_iB_{ii})^+ = I_{r_i} \qquad \text{if } r_i < m_i \tag{5.55}$$

In view (5.46), $u_{i2} = 0$ if $i \in \mathcal{I}_1$ and $m_i \leq r_i$ or $u_{i2} \neq 0$ if $i \in \mathcal{I}_1$ and $r_i < m_i$.

On the other hand, $\forall\, i \in \mathcal{I}_2$, $[I_{m_i} - (C_iB_{ii})^+(C_iB_{ii})] \neq 0$. This can be argued as follow. For those $i \in \mathcal{I}_2$ with $r_i \geq m_i$, rank $C_iB_{ii} < \min(m_i, r_i) = m_i$. Thus rank $(C_iB_{ii})^+(C_iB_{ii}) < m_i$ which implies $(C_iB_{ii})^+(C_iB_{ii})$ has less than m_i nonzero eigenvalues. But I_{m_i} has m_i non-zero eigenvalues. Hence $[I_{m_i} - (C_iB_{ii})^+(C_iB_{ii})] \neq 0$. A similar argument holds for $r_i < m_i$.

Theorem 5.5. Assume rank $CB = \ell < \min(r,m)$, then

$$\begin{aligned} u_{i2} &= 0 \ , \quad i \in \mathcal{I}_1 \text{ and } m_i \leq r_i \\ u_{i2} &\neq 0 \ , \quad i \in \mathcal{I}_1 \text{ and } r_i < m_i \text{ or } \forall\, i \in \mathcal{I}_2 \, . \end{aligned} \tag{5.56}$$

Class III $CB \equiv 0$

Since $CB = G_1 = 0$, then from the GHR structure, it is clear that independent of u_o or u, the interconnection $F_{12}z_2$ in GHR cannot be compensated. Hence, any compensation in order to achieve aggregability must be carried out at possibly next or lower layer in the GHR system matrix. In this class, neither perfect output following nor freedom in control synthesis exist.

In summary, the existence of a local control (5.3) is given by Theorem 5.3. In general, (5.3) can be decomposed into $u_i = u_{ia} + u_{ib} + u_{ic}$ where $u_{ia} + u_{ib} = u_{i1}$ and $u_{ic} = u_{i2}$, u_{i1} and u_{i2} being defined in Theorem 5.3. From functional point of view, u_{ia} allows arbitrary shaping of the aggregate dynamics if (F_{11}, CB) is controllable. u_{ib} achieves complete aggregability defined by (A,C) by decoupling S^1 from $\bar{S}^1$. u_{ic}, if exists allows the shaping of the residual dynamics by both local state feedback and changing the effective feedforward term $H^1 z_1 = H^1 y$ in the GHR. Moreover, u_{ic} only exists if $r_i < m_i$ when $C_i B_{ii}$ has maximal rank or if rank $C_i B_{ii} < \min(r_i, m_i)$. The details are given by Theorems 5.4 and 5.5.

5.3. Aggregability and Perfect Extended Output Following By Local State Feedback.

The result of Section 5.2 can be extended to include the more general case if compensation of $F_{12} z_2$ in the GHR cannot be achieved at the first layer but may be achieved at next or lower layers. Two possibilities can occur, namely the reference system (5.2) can remain at $p = r_1$th order or increases to higher order p, $p_1 < p \leq n$. The need to increase the order of the reference system (5.2) arises when the system interconnection variables y_i, $i = 1 \ldots \nu$ are best modeled by high order reference system or the GHR structure dictates that compensation of coupling Hessenberg blocks must be carried out at lower layers in the GHR, e.g. when $G_j = 0 \; \forall j < i$, but $G_i \neq 0$. The case that the reference system remains at r_1th order may be reformulated into the case $p > r_1$ by making additional assumption on the range space of F_{12} and G_1 in the GHR. However, we will not treat the former case in this thesis. To simplify the formation of an extented error $\tilde{e} = [e^T \bar{e}^T]^T$ which contains e, we assume $p = p_i = \sum_{j=1}^{i} r_i$ in (5.2). Also in (5.2), define the extended reference output as $\hat{y}_e = w'$. Clearly $\hat{y}_e$ contains $\hat{y}$, i.e. $\hat{y}_e = [\hat{y}^T \hat{y}_o^T]^T$. Assuming (5.5) holds for the triplet (F_o', G', L'), then (5.6) is the system and (5.11) is the equivalent reference model. In terms of the GHR structure, local control (5.3) must compensate (cancel) the coupling $F_{i,i+1} z_{i+1}$ between S^i and $\bar{S}^i$ and achieves model following between z^i and w' subsequently. In the remaining sections, a development similar to Section

5.2 is carried out.

5.3.1. Derivation of the Local Control Law

Definition 5.2. Since

$$z = \begin{bmatrix} z^j \\ \bar{z}^j \end{bmatrix} = H_o x = \begin{bmatrix} C_e \\ \bar{H}_o^j \end{bmatrix} x \qquad (5.57)$$

Then $y_e = C_e x$ will be defined as the extended output of (5.6) induced by y.

Let the reference state w of (5.11) be partitioned similar to (5.57), i.e.

$$w = \begin{bmatrix} w' \\ \bar{w}^j \end{bmatrix} \quad , \quad w' \in R^{p_j}$$

Since y and $\hat{y}$ are respectively the first subvector of y_e and w', then $\lim_{t\to\infty} y(t) = \lim \hat{y}(t)$ if $\lim_{t\to\infty} y_e(t) = \lim_{t\to\infty} w'(t)$. Similar to Section 5.2, we define

$$e_i \equiv \hat{y}_i - y_i \quad , \; i = 1,2,\ldots\nu \qquad (5.58)$$

$$\tilde{e} = y_e - w' \qquad (5.59)$$

Again, we impose the local feedback control structure

$$u_i = \pi_i u_{io} + K_i e + M_{ii} x_i, \; i = 1\ldots\nu \qquad (5.60)$$

Extended output following is defined as follows:

Definition 5.3. The local control (5.60) is said to achieve perfect extended output following if

$$\lim_{t\to\infty} \tilde{e}(t) = \lim_{t\to\infty} w^j(t) - y_e(t) = 0 \tag{5.61}$$

Our objective now is to obtain conditions for the existence of π_i and M_{ii} and to examine the properties of the local control (5.60) so that $\tilde{e}(t)$ approaches zero asymtotically.

Equations (5.6) and (5.55) gives

$$\dot{y}_e = C_eAx + \sum_{i=1}^{\nu} C_eB_iu_i + C_eEf \tag{5.62}$$

$$\dot{w}^j = D^jF_ow + D^jGu_o + D^jWf \tag{5.63}$$

since $W = HE$, then $D^jW = W^j = C_eE$. Hence with u_i defined by (5.60), it can be shown

$$\begin{aligned}\dot{\tilde{e}} &= F^j\tilde{e} + (F^jC_e - C_eA)x + C_eBu_o - \sum_{i=1}^{\nu} C_eB_iu_i \\ &= (F^j - C_eB\tilde{K})\tilde{e} + C_eBu_o - \sum_{i=1}^{\nu} C_eB_i\pi_iu_{io} + (F^jC_e - C_eA)x - \sum_{i=1}^{\nu} C_eB_iM_{ii}x_i\end{aligned} \tag{5.64}$$

where

$$\tilde{K} = [K \quad 0] \tag{5.65}$$

K is a $m \times r$ matrix. Hence perfect extended output following requires:

(I) There exists $K \in R^{m \times r}$ such that $(F^j - C_eB\tilde{K})$ is stable.

(II) There exists matrices π_i, $i = 1, \ldots \nu$ such that

$$C_eBu_o = \sum_{i=1}^{\nu} C_eB_iu_{io} \tag{5.66}$$

(III) There exists matrices M_{ii}, $i = 1, \ldots \nu$ such that

$$(F^j C_e - C_e A)x = \sum_{i=1}^{\nu} C_e B_i M_{ii} x_i \tag{5.67}$$

If π_i and M_{ii} indeed exist and satisfy requirements (II) and (III), then

$$\dot{\tilde{e}} = (F^j - C_e B \tilde{K})\tilde{e} \tag{5.68}$$

which implies $\tilde{e} \to 0$ as $t \to \infty$ if requirement (I) also holds. Following essentially the development of Section 5.2 with appropriate modifications, the following theorems are stated without proofs.

__Theorem 5.6__. The matrices π_i, $i = 1,\ldots\nu$ always exist such that

$$C_e B u_o = \sum_{i=1}^{\nu} C_e B_i \pi_i u_{io}$$

for any u_o. Furthermore,

$$\pi_i = (C_e B_i)^+ (C_e B_i) + [I_{m_i} - (C_e B_i)^+ (C_e B_i)] N_i \tag{5.69}$$

where N_i is any $m_i \times m_i$ arbitrary matrix.

__Theorem 5.7__. If for $i = 1,\ldots\nu$, $\mathscr{I}_i$ defined in (5.36),

$$[I - (C_e B_i)(C_e B_i)^+](F^j C_e - C_e A)\mathscr{I}_i = 0 \tag{5.70}$$

then M_{ii} exists such that

$$(F^j C_e - C_e A)x = \sum_{i=1}^{\nu} C_e B_i M_{ii} x_i$$

Furthermore,

$$M_{ii} = (C_e B_i)^+ (F^j C_e - C_e A)\mathscr{I}_i + [I - (C_e B_i)^+ (C_e B_i)] P_i \tag{5.71}$$

where P_i is a $m_i \times n_i$ arbitrary matrix.

__Theorem 5.8__. Assume (A,C) is non-aggregable. Consider the system (5.6) and the reference model (5.11). Then perfect extended output following is achieved by the local control

$$u_i = \pi_i u_{io} + K_i e + M_{ii} x_i, \quad i = 1,\ldots\nu \tag{5.72}$$

if

(i) $\sigma(F^j/V_1) \cap \sigma(F^j/V_2)$ are stable where V_1 is the smalles F^j - invariant subspace containing $\ker\{D^j\}$ and V_2 is the smallest $(F^j)^T$-invariant subspace containing $\ker\{(C_eB)^T\}$. $\sigma(F^j/V_i)$, $i=1,2$ represents eigenvalues of F^j restricted to V_i.

(ii) $[I - (C_eB_i)(C_eB_i)^+](F^jC_e - C_eA)\mathscr{I}_i = 0 \quad , \quad i=1\ldots\nu$

Furthermore,

$$\pi_i = (C_eB_i)^+(C_eB_i) + [I - (C_eB_i)^+(C_eB_i)]N_i \tag{5.73}$$

$$M_{ii} = (C_eB_i)^+(F^jC_e - C_eA)\mathscr{I}_i + [I - (C_eB_i)^+(C_eB_i)]P_i \tag{5.74}$$

where N_i and P_i are arbitrary matrices. And u_i may be decomposed into

$$u_i = u_{i1} + u_{i2} \tag{5.75}$$

with

$$u_{i1} = (C_eB_i)^+(C_eB_i)u_{io} + K_ie + (C_eB_i)^+(F^jC_e - C_eA)\mathscr{I}_ix_i \tag{5.76}$$

$$u_{i2} = [I - (C_eB_i)^+(C_eB_i)](N_iL_iw^j + P_ix_i) \tag{5.77}$$

Proof: With π_i and M_{ii} chosen according to (5.73) - (5.74), and if condition (ii) holds, then

$$\dot{\tilde{e}} = (F^j - C_e B \tilde{K}) \tilde{e} \tag{5.78}$$

Since $\tilde{K} = [K \quad 0]$, $K \in R^{m \times r}$. Define $D^j = [I_\gamma \quad 0]$ as an output matrix. Then the triple $(F^j, (C_e B), D^j)$ of (5.78) defines an output feedback stabilizability problem. The existence of K is guaranteed by [54] which is interpreted by condition (i). The rest follows.

Theorem 5.8 again indicates a decomposition of u_i into two functional components. The role u_{i1} plays in arriving at (5.78) indicates that it is used to achieve aggregability by compensating the interconnection $F_{j,j+1} z_{j+1}$ between S^j and $\bar{S}^j$. In addition, it also shapes the dynamics of the decoupled aggregate. On the other hand, u_{i2} represents the freedom component in synthesizing the local control u_i. From (5.77), we note u_{i2} has two components. The first component $[I - (C_e B_i)^+(C_e B_i)] N_i L_i w^j$ represents the shaping of effective feedforward term $H^j z^j$ in the residual from the aggregate. This is so because the aggregate state w^j is utilized. The second component in u_{i2}, namely $[I - (C_e B_i)^+(C_e B_i)] P_i x_i$ represents the shaping of the residual dynamics through local feedback. Even though the feedback is introduced through x_i instead of $\bar{w}^j$, we note w^j and $\bar{w}^j$ compliment each other in the state space. Since the aggregate is decoupled from the residual, any state feedback that does not affect the dynamics of the aggregate must not involve the aggregated state w^j.

5.3.2. Properties of the Local Control Law

From (5.77), we observe $u_{i2} = u_{ic} = 0$ if and only if

$$[I - (C_e B_i)^+(C_e B_i)] = 0 . \tag{5.79}$$

Similar to the previous section, we classify $C_e B$ according to (a) $C_e B$ has maximal rank; (b) C_eB has less than maximal rank but nonzero and (c) $C_eB \equiv 0$.

<u>Class I</u> rank $(C_eB) = \min(p,m)$, $p = \sum_{i=1}^{j} r_i = \dim w^j$.

<u>Class I-a</u> rank $(C_eB) = m \leq p$

Since

$$C_eB = [\underbrace{C_eB_1}_{m_1 \text{ columns}} \quad \cdots \quad \underbrace{C_eB_\nu}_{m_\nu \text{ columns}}] \tag{5.80}$$

then

$$\text{rank } C_eB_i = m_i < p \tag{5.81}$$

(5.81) implies $\forall$ i, there exists $(C_eB_i)^+$ such that

$$(C_eB_i)^+(C_eB_i) = I_{m_i} \tag{5.82}$$

Thus $[I - (C_eB_i)^+(C_eB_i)] \equiv 0$ which implies $u_{i2} = 0 \ \forall$ i.

<u>Class I-b</u> rank $C_eB = p < m$

Several possibilities can occur in this class, namely, rank $C_eB_i = m_i \leq p$, rank $C_eB_i = p < m_i$ or rank $C_eB_i < \min(p,m_i)$. Define the index sets

$$\mathcal{L}_1 = \{i/\text{rank } C_eB_i = p < m_i \text{ or rank } C_eB_i < \min(p,m_i)\}$$

$$\mathcal{L}_2 = \{i/\text{rank } C_eB_i = m_i \leq p\}$$

The assumption that rank $C_eB = p < m$ implies $\mathcal{L}_2$ cannot be exhaustive, i.e. not all indices belong to $\mathcal{L}_2$. Due to (5.82), it is easy to see

$$[I-(C_eB_i)^+(C_eB_i)] = \begin{cases} \text{non-zero} & \text{, if } i \in \mathcal{L}_1 \\ 0 & \text{, if } i \in \mathcal{L}_2 \end{cases}$$

Consequently

$$u_{i2} = \begin{cases} \text{non-zero} & \text{, if } i \in \mathcal{L}_1 \\ 0 & \text{, if } i \in \mathcal{L}_2 \end{cases}$$

<u>Class II</u> rank $C_eB < \min(p,m)$

From (5.80), $C_eB_i \in R^{p \times m_i}$. Clearly rank $C_eB_i = p$ is impossible. On the other hand, we may group C_eB_i into two different index sets

$$\mathcal{J}_1 = \{i/\text{rank }(C_eB_i) < \min(m_i,p)\}$$

$$\mathcal{J}_2 = \{i/\text{rank }(C_eB_i) = m_i < p\}$$

Again we note $\mathcal{J}_2$ cannot be exhausive. Hence $\mathcal{J}_1 \neq \phi$. Since

$$[I-(C_eB_i)^+(C_eB_i)] = \begin{matrix} \text{non-zero} & \text{, if } i \in \mathcal{J}_1 \\ 0 & \text{, if } i \in \mathcal{J}_2 \end{matrix}$$

Then

$$u_{i2} = \begin{matrix} \text{non-zero} & \text{, if } i \in \mathcal{J}_1 \\ 0 & \text{, if } i \in \mathcal{J}_2 \end{matrix}$$

<u>Theorem 5.9</u>. Assume (A,C) is non-aggregable. Then there exists freedom in control synthesis in the sense $u_{i2} \neq 0$ if and only if rank $C_eB_i < \min(m_i,p)$ or when the number of extended output is less than number of local controls, i.e.

$$u_{i2} = u_{ic} = \begin{cases} \text{non-zero} & \text{, if } i \in \mathcal{L}_1 \text{ or } i \in \mathcal{J}_1 \\ 0 & \text{, if } i \in \mathcal{L}_2 \text{ or } i \in \mathcal{J}_2 \end{cases}$$

Class III $C_e B \equiv 0$

Independent of u or u_o, the interconnection between S^j and $\bar{S}^j$, $F_{jj+1}z_{j+1}$ cannot be compensated, because the control coefficient matrix for the aggregate subsystem S^j is zero. Thus aggregability cannot be achieved.

Conclusions similar to those at the end of Section 5.2.2 can be summarized. The important point to be emphasized is that u_i again possess three basic components, i.e. $u_i = u_{ia} + u_{ib} + u_{ic}$, $u_{ia} + u_{ib} = u_{i1}$ and $u_{ic} = u_{i2}$. u_{ia}, u_{ib} and u_{ic} are used to control the aggregate, to achieve aggregation and to control the residual respectively.

5.4. A Decomposition in Optimization and Control Synthesis

Consider again the interconnected system (5.1). Suppose that associated with each subsystem, a quadratic cost

$$J_i = \tfrac{1}{2}\int_0^\infty x_i^T Q_i x_i + u_i^T R_i u_i dt \quad , \quad i = 1 \ldots \nu \tag{5.83}$$

is to be minimized. This section investigates the possibility of decomposing J_i into a system level cost to be associated with the coordinator proposed in the beginning of Chapter 5 and a subsystem level cost to be associated with each individual subsystem in (5.1).

From previous sections, we notice u_i may be decomposed into three components

$$u_i = u_{ia} + u_{ib} + u_{ic} \tag{5.84}$$

where the role of u_{ia} is to control the aggregate; the role of u_{ib} is to achieve aggregation and the role of u_{ic} is to provide some freedom, if exists, to control the residual. It will be shown that a suboptimal solution to optimizing (5.83) with respect to (5.1) may be obtained by first solving a coordinator problem, thus defining the component u_{ia}; next specify the resulting optimal dynamic behavior of the system interconnection variable y_i, $i=1...\nu$ under u_{ia} to each local subsystem (5.1); finally with each local subsystem having complete knowledge of this desired interconnection behavior, the remaining component u_{ib} and u_{ic} are obtained at subsystem level simultaneously, thus achieving parallel computations.

We assume here that a first layer aggregate in the GHR will be used to govern the dynamic behavior of the system interaction variables y_i, $i=1 \ldots \nu$. Given the GHR (5.8), if higher order aggregate is selected, the procedure described below can be modified accordingly and thus will not be considered in this thesis. The aggregate in GHR corresponding to the r_1th order reference system is then

$$\dot{y} = F_{11}y + F_{12}z_2 + \sum_{i=1}^{\nu} CB_i u_i \tag{5.85}$$

and the component u_{ib} is used to compensate the coupling term $F_{12}z_2$.

Let $\mathcal{X}_i$ denote the state space corresponding to subsystem state x_i. Since $y_i = C_i x_i$, the subsystem state space $\mathcal{X}_i$ can be decomposed into the range space of C_i and null space of C_i, i.e. $\mathcal{X}_i = \mathcal{X}_{ia} \oplus \mathcal{X}_{ib}$. Thus

$$x_i = x_{ia} + x_{ib} \tag{5.86}$$

where

$$x_{ia} = C_i^+ y_i = C_i^+ C_i x_i \in \mathcal{X}_{ia} \tag{5.87a}$$

$$x_{ib} = x_i - x_{ia} = [I - C_i^+ C_i] x_i \in \mathcal{X}_{ib} \tag{5.88}$$

Clearly $\mathcal{X}_{ia}$ being the range space of C_i, corresponds to the aggregate variable y_i, thus we associate the component u_{ia} with $\mathcal{X}_{ia}$. On the other hand, the complement of $\mathcal{X}_{ia}$ is the residual represented by $\mathcal{X}_{ib}$, since u_{ib} and u_{ic} affects the residual (through effective feedforward from S^i into $\bar{S}^i$ and feedback of residual states), we shall associate u_{ib} and u_{ic} with $\mathcal{X}_{ib}$. Also, in order to make u_{ic} non-interactive, it is necessary to restrict u_{ic} to lie in the null space of CB_i. Thus u_{ic} has the general form

$$u_{ic} = [I - (CB_i)^+(CB_i)](N_i y + P_i x_i) \tag{5.89}$$

where N_i, P_i, $i = 1,\ldots\nu$ are arbitrary matrices. Denote u_{ib} as

$$u_{ib} = M_{ii}x_i \qquad i = 1,\ldots\nu \tag{5.90}$$

where M_{ii} is some fixed gain for the purpose of achieving compensatory action. Since u_{ia} will be designed as a feedback of the aggregate, let

$$u_{ia} = L_i y \tag{5.91}$$

Then using (5.84), (5.86), (5.90) - (5.91) in (5.83), it can be shown

$$J_i = \tfrac{1}{2}\int_o^\infty (x_{ia} + x_{ib})^T Q_i (x_{ia} + x_{ib}) + (u_{ia} + u_{ib} + u_{ic})^T R_i (u_{ia} + u_{ib} + u_{ic})dt$$

$$= J_{ia} + J_{ib} \tag{5.92}$$

where

$$J_{ia} = \tfrac{1}{2}\int_o^\infty y_i^T (C_i^{+T} Q_i C_i^+) y_i + u_{ia}^T R_i u_{ia}\, dt \tag{5.93}$$

$$J_{ib} = \tfrac{1}{2}\int_0^\infty [\eta_i^T \; u_{ic}^T] Q_{io} \begin{bmatrix} \eta_i \\ u_{ic} \end{bmatrix} dt \tag{5.94}$$

$$\eta_i = \begin{bmatrix} y \\ x_i \end{bmatrix} \qquad Q_{io} = \begin{bmatrix} \tilde{Q}_i & \tilde{R}_i \\ \tilde{R}_i^T & R_i \end{bmatrix} \qquad \tilde{Q}_i = \begin{bmatrix} 0 & \tilde{Q}_{i12} \\ \tilde{Q}_{i12}^T & \tilde{Q}_{i22} \end{bmatrix} \tag{5.95}$$

$$\tilde{Q}_{i12} = L_i^T R_i M_{ii} \tag{5.96}$$

$$\tilde{Q}_{i22} = [I - C_i{}^+C_i]Q_i[I - C_i{}^+C_i] + C_i{}^+C_iQ_i[I - C_i{}^+C_i] + [I - C_i{}^+C_i]Q_iC_i{}^+C_i + M_{ii}{}^T R_i M_{ii} \tag{5.97}$$

$$\tilde{R}_i^T = R_i[\; L_i \quad M_{ii}] \tag{5.98}$$

Next we realize the application of u_{ib} causes (5.85) to become

$$\dot{y} = F_{11}y + \sum_{i=1}^{\nu} CB_i(u_{ia} + u_{ic}).$$

Since u_{ic} is to be non-interactive in the sense u_{ic} lies in the ker CB_i, then $\sum_{i=1}^{\nu} CB_i u_{ic} = 0$ which yields

$$\dot{y} = F_{11}y + \sum_{i=1}^{\nu} CB_i u_{ia} \tag{5.99}$$

(5.93) and (5.99) forms the aggregate optimization problem which we define as the coordinator problem.

Coordinator Level Problem

$$\min_{u_{ia}} \; J_a = \sum_{i=1}^{\nu} \alpha_i J_{ia}(u_{ia}) \tag{5.100a}$$

subject to

$$\dot{y} = F_{11}y + \sum_{i=1}^{\nu} CB_i u_{ia} \tag{5.100b}$$

$$J_{ia} = \tfrac{1}{2}\int_0^{\infty} y_i^T (C_i^{+T} Q_i C_i^{+}) y_i + u_{ia}^T R_i u_{ia} dt \tag{5.100c}$$

where $\sum_{i=1}^{\nu} \alpha_i = 1$.

In defining the coordinator problem, co-operation between u_{ia}, $i = 1 \ldots \nu$ is assumed. Thus it can be solved as a Pareto-optimal differential game problem [55].

Let the solution to the coordinator problem be written as in (5.91). The closed-loop aggregate of (5.99) after implementing the solution (5.91) may be expressed as

$$\dot{y} = (F_{11} + CBL)y \tag{5.101}$$

where $L \equiv [L_1^T \ldots L_\nu^T]^T$. On the other hand, with the application of u_{ia} and u_{ib}, the subsystem equation (5.89) may be written as

$$\begin{aligned}\dot{x}_i &= A_{ii}x_i + \sum_{\substack{j=1\\ j\neq i}}^{\nu} A_{ij}y_j + B_{ii}(L_i y + M_{ii}x_i + u_{ic}) \\ &= (A_{ii} + B_{ii}M_{ii})x_i + (\bar{A}_{ii} + B_{ii}L_i)y + B_{ii}u_{ic}, \quad i=1,\ldots\nu\end{aligned} \tag{5.102}$$

where $\sum_{\substack{j=1\\ j\neq i}}^{\nu} A_{ij}y_j \equiv \bar{A}_{ii}y$. Assume the closed-loop aggregate dynamic (5.101) is known to each subsystem (5.102) as a reference dynamic for the system interconnection variable y_i, $i=1,\ldots\nu$, then augmenting (5.101) and (5.102), we get

$$\begin{bmatrix} \dot{y} \\ \dot{x}_i \end{bmatrix} = \begin{bmatrix} F_{11} + CBL & 0 \\ \bar{A}_{ii} + B_{ii}L_i & A_{ii} + B_{ii}M_{ii} \end{bmatrix} \begin{bmatrix} y \\ x_i \end{bmatrix} + \begin{bmatrix} 0 \\ B_{ii} \end{bmatrix} u_{ic} \, , \quad i=1,\dots\nu \tag{5.103}$$

It is interesting to note (5.94) and (5.103) may be solved as a separate optimization problem <u>at the subsystem level without additional information from other subsystems</u>. However, this subsystem level problem is not totally unconstrained, since according to (5.84), u_{ic} must lie in the ker CB_i for each index i in order to be non-interactive. This constraint can be taken into account by modifying the control coefficient matrix B_i. This can be seen if we define

$$u_{ic}' = [N_i \; M_{ii}] \begin{bmatrix} y \\ x_i \end{bmatrix}$$

then

$$B_i u_{ic} = B_i[I - (CB_i)^+(CB_i)]u_{ic}' = \beta_i u_{ic}' \tag{5.104}$$

We note here that there is an implied mutual agreement between the subsystem controller u_i in choosing u_{ib} and the general form of u_{ic}. Although u_{ic} must lie in ker CB_i, it is arbitrary in the sense u_{ic}' is arbitrary as seen in (5.104). The subsystem level optimization problem can then be defined as:

<u>Subsystem Level Problem</u>

$$\min \; J_{ib} = \tfrac{1}{2}\int_0^\infty [\eta_i^T \; u_{ic}'^T] Q_{io} \begin{bmatrix} \eta_i \\ u_{ic}' \end{bmatrix} dt \tag{5.105a}$$

subject to

$$\dot{\eta}_i = \mathcal{A}_i \eta_i + \beta_i u'_{ic} \tag{5.105b}$$

$$y_i = \mathcal{D}_i \eta_i \qquad i=1,\ldots\nu \tag{5.105c}$$

where η_i and Q_{io} are defined in (5.95) - (5.98) and

$$\mathcal{A}_i = \begin{bmatrix} F_{11} + CBL & 0 \\ \bar{A}_{ii} + B_{ii}L_i & A_{ii} + B_{ii}M_{ii} \end{bmatrix} \quad \beta_i = \begin{bmatrix} 0 \\ B_{ii}[I - (CB_i)^+(CB_i)] \end{bmatrix}$$

$$\mathcal{D}_i = [0\ldots0 \ I_{Y_i} \ 0\ldots0] \tag{5.106}$$

Thus by modifying the control coefficient matrix, the subsystem level problem is again an unconstrained optimal regulator problem. Extensive results can be found in [56]. One must emphasize that the subsystem level problem (5.105) may be solved independent of other subsystem's optimization. The only requirement is each subsystem must know the closed-loop dynamics of the aggregate under the influence of u_{ia}. For a truly large scale system of high dimensionality, the order of the aggregate (the reduced order model of the system) can be small. Hence even coupled with each subsystem dynamic, the dimension of η_i, $i=1,\ldots\nu$ is still considerably smaller than the order of the composite system. In addition, with the decomposition suggested, the original problem (5.88) - (5.89) may be solved by a sequential-then-parallel computations. The steps are summarized as follow:

Step I Obain GHR system representation of the composite.

Step II Base on system structural considerations, decide the order of aggregate subsystem desired. Eigen-analysis , GQL algorithm application, capability of compensatory control to achieve aggregation if initially not aggregable may be carried out at this step.

Step III Choose and apply compensatory control u_{ib} for each subsystem to achieve aggregation and decoupling.

Step IV Solve the coordinator level problem (5.100) for u_{ia}.

Step V Apply u_{ia} and obtain closed-loop aggregated system (5.101) and specify closed-loop dynamic (5.101) to each individual subsystem.

Step VI Solve subsystem level optimal control problem for u_{ic} and apply u_{ic}.

Thus at Step VI, we see parallel computations to be carried out locally at each subsystem. Such computation savings can be used to increase efficiency in solving large scale system problem and thus represent a distinct advantage over existing methodologies. The implementation configuration of the local u_i is given in Figure 5.2.

We conclude this section by remarking that the subsystem level problem (5.105) is posed as a full state feedback optimization problem. If $u_{ic}^{\prime}$ is restricted to feedback only locally available informations, e.g. x_i or x_i plus those components of y locally measurable at each subsystem, then the problem becomes a constrained output regulator problem [43,57 - 58]. In Chapter 6, the results of this Chapter will be applied to study two power system examples.

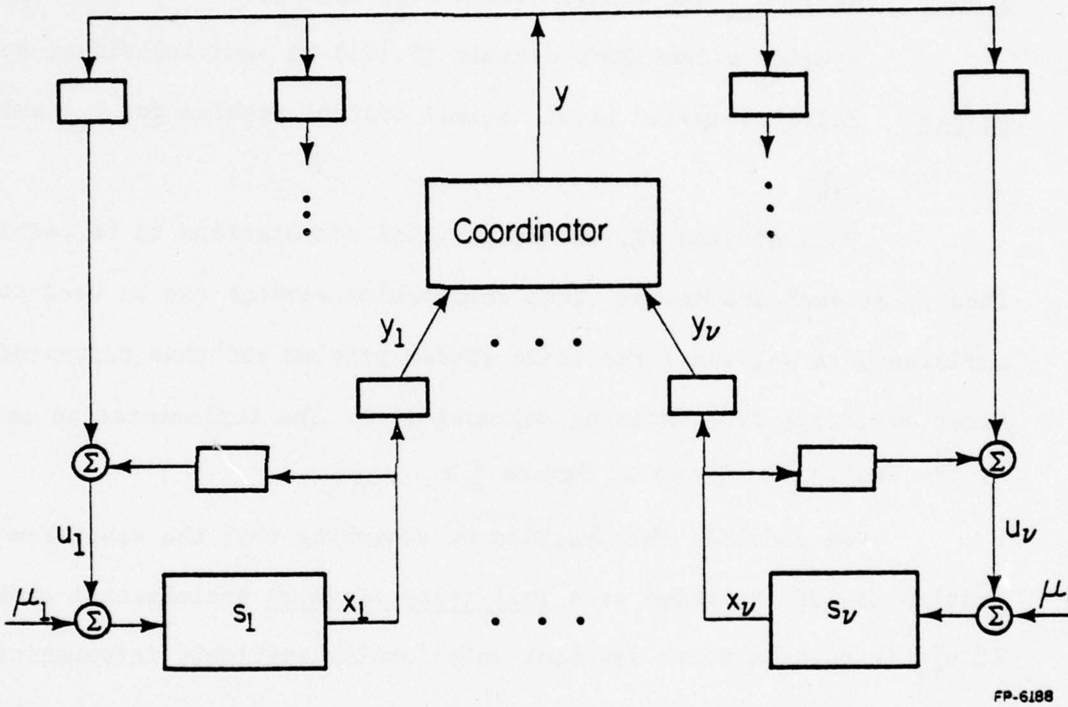


Figure 5.2. Implementation structure of proposed hierarchical control.

CHAPTER 6

APPLICATIONS

Four examples are given in this chapter. Examples 6.1 and 6.2 illustrate the usefulness of GHR and the GQL algorithm in obtaining suitable reduced order models. Examples 6.3 and 6.4 illustrate the role of the GHR in solving the synthesis problem considered in Chapter 5. The existence of local feedback and decomposition of local control to achieve aggregation are shown. The sensitivity of output trajectories with respect to u_{ib} is also studied. To assess the performance of reduced order models, both open-loop and closed-loop trajectories are simulated. For closed loop analysis, we distinguish three different regulator designs:

Optimal Regulator:

Given a LTI system

$$\dot{x} = Ax + Bu \qquad , \; x \in R^n \quad , \quad u \in R^m \tag{6.1a}$$

$$y = Cx \qquad , \; y \in R^r \tag{6.1b}$$

The optimal control $u_o = L_o x$ that minimizes the cost

$$J = \frac{1}{2} \int_o^\infty y^T Q y + u^T R u \; dt \tag{6.2}$$

is obtained. Here Q and R satisfy the usual assumptions of positive semi and positive definiteness. The optimal closed loop system is then

$$\dot{x} = (A + BL_o)x$$

$$y = Cx$$

Reduced Regulator

Consider the GHR of (6.1)

$$\begin{bmatrix} \dot{z}^i \\ \dot{\bar{z}}^i \end{bmatrix} = \begin{bmatrix} F_r & E_r \\ H_r & \bar{F}_r \end{bmatrix} \begin{bmatrix} z \\ \bar{z} \end{bmatrix} + \begin{bmatrix} G_r \\ \bar{G}_r \end{bmatrix} u \equiv Fz + Gu \tag{6.3a}$$

$$y = [D_r \quad 0] \begin{bmatrix} z^i \\ \bar{z}^i \end{bmatrix} \equiv Dz \tag{6.3b}$$

and its approximate aggregated model (reduced order model of (6.1)).

$$\dot{z}_r = F_r z_r + G_r u \quad , \quad z_r \in R^p \tag{6.4a}$$

$$\hat{y} = D_r z_r \tag{6.4b}$$

Let the control $\hat{u}_o = \hat{L}_o z_r$ be the control that minimizes the cost

$$\hat{J} = \frac{1}{2} \int_o^\infty \hat{y}^T Q \hat{y} + u^T R u \, dt \tag{6.5}$$

with respect to (6.4). Again Q and R satisfy the usual assumption and take on same values as in (6.2). The optimal reduced order closed loop system is

$$\dot{z}_r = (F_r + G_r \hat{L}) z_r$$

$$\hat{y} = D_r z_r$$

Suboptimal Regulator

Define the control

$$U_{sub} = [\hat{L} \quad 0] \begin{bmatrix} z^i \\ \bar{z}^i \end{bmatrix} \equiv L_s z \tag{6.6}$$

as the suboptimal control, the application of (6.6) is assumed to give suboptimal behavior to the optimal regulator (6.1) - (6.2). The resulting controller is termed the suboptimal regulator. The suboptimal closed-loop system is

$$\dot{z} = (F + GL_s)z$$

$$y = Dz$$

or equivalently

$$\dot{x} = (A + BL_s H)x$$

$$y = Cx$$

where H is the GHR transformation matrix.

In all output trajectory simulation, O, R, S denote the optimal, reduced and suboptimal regulators respectively.

6.1 Model Reduction of a Single Area Power System

Consider a single area power system consists of two interconnected thermal plants [59]:

$$\dot{x} = Ax + Bu + E\omega \tag{6.7a}$$

$$y = Cx \tag{6.7b}$$

where $x \in R^9$, $u \in R^2$, $y \in R^3$, and $\omega \in R$. The state, control, disturbance and output variables have the following meanings:

$x_1 = \Delta a_1$ - valve position displacement in first unit

$x_2 = \Delta P_{t_1}$ - power output displacement of HP turbine in first unit

$x_3 = \Delta P_{t_2}$ - power output displacement of IP turbine in first unit

$x_4 = \Delta P_{t_3}$ - power output displacement of LP turbine in first unit

$x_5 = \Delta a_2$ - valve position displacement in second unit

$x_6 = \Delta P_{t_1}$ - power output displacement of HP turbine in second unit

$x_7 = \Delta P_{t_2}$ - power output displacement of IP turbine in second unit

$x_8 = \Delta P_{t_3}$ - power output displacement of LP turbine in second unit

$x_9 = \Delta f$ - frequency deviation in the system

$u_1 = \Delta f_{m_1}$ - set point adjustment in first unit

$u_2 = \Delta f_{m_2}$ - set point adjustment in second unit

$\omega = \Delta P_L$ - load disturbance

$y_1 = \Delta P_1$ - total power output of first unit

$y_2 = \Delta P_2$ - total power output of second unit

$y_3 = \Delta f$ - frequency deviation in the system.

And the system matrices have the structure:

$$A = \begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{19} \\ a_{21} & a_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{43} & a_{44} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{55} & 0 & 0 & 0 & a_{59} \\ 0 & 0 & 0 & 0 & a_{65} & a_{66} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{76} & a_{77} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{87} & a_{88} & 0 \\ 0 & a_{92} & a_{93} & a_{94} & 0 & a_{96} & a_{97} & a_{98} & a_{99} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & b_{52} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad E = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ f_{19} \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & c_{12} & c_{13} & c_{14} & 0 & 0 & 0 & 0 & c_{19} \\ 0 & 0 & 0 & 0 & 0 & c_{26} & c_{27} & c_{28} & c_{29} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} . \qquad (6.8)$$

The entries of (A,B,E,C) in terms of physical parameters and their values are given in Appendix E.

Let us initially carry out the chained aggregation procedure described in Chapter 2 in order to obtain the GHR. Note $y \in R^3$, and $C = [C_{11} \ \ \bar{C}_{11}]$, from (6.8), C_{11} is singular. A initial state permutation is necessary. From C, we pick columns four, eight and nine as the first three columns of the permuted C. In doing so, we have placed the outputs of the last stage of turbine in each plant and the system frequency as the leading state components. Formally this is achieved by using the permutation matrix P_1 of the chained aggregation transformation. Then the permuted system matrices are:

$$A_1 = \begin{bmatrix} a_{44} & 0 & 0 & 0 & 0 & a_{43} & 0 & 0 & 0 \\ 0 & a_{88} & 0 & 0 & 0 & 0 & 0 & 0 & a_{87} \\ a_{94} & a_{98} & a_{99} & 0 & a_{92} & a_{93} & 0 & a_{96} & a_{97} \\ 0 & 0 & a_{19} & a_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{21} & a_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{32} & a_{33} & 0 & 0 & 0 \\ 0 & 0 & a_{59} & 0 & 0 & 0 & a_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{65} & a_{66} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{76} & a_{77} \end{bmatrix} , \quad B_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ b_{11} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & b_{52} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$E_1 = \begin{bmatrix} 0 \\ 0 \\ f_{19} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{6.9}$$

$$C_1 = \left[\begin{array}{ccc:cccccc} c_{14} & 0 & c_{19} & 0 & c_{12} & c_{13} & 0 & 0 & 0 \\ 0 & c_{28} & c_{29} & 0 & 0 & 0 & 0 & c_{26} & c_{27} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] .$$

$$\longleftarrow C_{11} \longrightarrow \qquad \longleftarrow \bar{C}_{11} \longrightarrow$$

The second step in the chained aggregation procedure consists in applying the transformation

$$T_1 = \begin{bmatrix} c_{14} & 0 & c_{19} & 0 & c_{12} & c_{13} & 0 & 0 & 0 \\ 0 & c_{28} & c_{29} & 0 & 0 & 0 & 0 & c_{26} & c_{27} \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

to the system (6.9) which results in the representation

$$
\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \\ \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \end{bmatrix}
=
\begin{bmatrix}
f_{11} & f_{12} & f_{13} & p_{14} & p_{15} & p_{16} & 0 & 0 & 0 \\
f_{21} & f_{22} & f_{23} & 0 & 0 & 0 & p_{27} & p_{28} & p_{29} \\
f_{31} & f_{32} & f_{33} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & a_{19} & a_{11} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & a_{21} & a_{22} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & a_{32} & a_{33} & 0 & 0 & 0 \\
0 & 0 & a_{59} & 0 & 0 & 0 & a_{55} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & a_{65} & a_{66} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{76} & a_{77}
\end{bmatrix}
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ x_1 \\ x_2 \\ x_3 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}
+
\begin{bmatrix}
0 & 0 \\ 0 & 0 \\ 0 & 0 \\ b_{11} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & b_{52} \\ 0 & 0 \\ 0 & 0
\end{bmatrix}
\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}
+
\begin{bmatrix} h_{11} \\ h_{21} \\ h_{31} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\omega
\tag{6.10}
$$

while the output equation becomes

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ x_1 \\ x_2 \\ x_3 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} . \qquad (6.11)$$

Note that in (6.10) the submatrix F'_{12},

$$F'_{12} = \begin{bmatrix} p_{14} & p_{15} & p_{16} & 0 & 0 & 0 \\ 0 & 0 & 0 & p_{27} & p_{28} & p_{29} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

is not a zero matrix, so a complete aggregation with respect to C is not possible. We proceed to the next step in the chained aggregation procedure. Note that rank $F'_{12} = 2$ and the interaction term entering the first aggregate has two independent components. These become the outputs from the residue and are denoted as y_4 and y_5:

$$\begin{bmatrix} y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} p_{14} & p_{15} & p_{16} & 0 & 0 & 0 \\ 0 & 0 & 0 & p_{27} & p_{28} & p_{29} \end{bmatrix} x^r . \qquad (6.12)$$

The residual system becomes

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{55} & 0 & 0 \\ 0 & 0 & 0 & a_{65} & a_{66} & 0 \\ 0 & 0 & 0 & 0 & a_{76} & a_{77} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} + \begin{bmatrix} 0 & 0 & a_{19} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a_{59} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} + \begin{bmatrix} b_{11} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & b_{51} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \omega. \qquad (6.13)$$

Observe that (6.13) represents two completely uncoupled third order systems. Moreover, (6.12) shows that each of the two outputs comes from a separate subsystem. Such uncoupling is consistent with the fact the physical system is an interconnection of two identical thermal plants. It is these system structural features that are retained and exhibited by the GHR. The use of given output information structure to induce reduced order modeling is thus appealing. Because of this total decoupling in the residual, it is possible to proceed in either of two possible directions. The first is to introduce the second permutation on (6.13) which will arrange the components in the measurement matrix (6.12) so that C_{22} is nonsingular. The procedure is then repeated as described above by defining the new transformation matrix T_2, forming the transformed state equations of this residual and proceeding further as described. The other possibility is to consider separately the two decoupled systems and obtain their GHR with respect to their measurement vectors, and then compose these into the GHR of the coupled system by permuting the order of the state variables. In both cases, (6.7) is brought into the following GHR:

$$
\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \\ \hline \dot{y}_4 \\ \dot{y}_5 \\ \hline \dot{y}_6 \\ \dot{y}_7 \\ \hline \dot{y}_8 \\ \dot{y}_9 \end{bmatrix}
=
\left[\begin{array}{ccc|cc|cc|cc}
f_{11} & f_{12} & f_{13} & 1 & 0 & 0 & 0 & 0 & 0 \\
f_{21} & f_{22} & f_{23} & 0 & 1 & 0 & 0 & 0 & 0 \\
f_{31} & f_{32} & f_{33} & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & f_{43} & f_{44} & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & f_{53} & 0 & f_{55} & 0 & 1 & 0 & 0 \\
\hline
0 & 0 & 0 & f_{64} & 0 & f_{66} & 0 & f_{68} & 0 \\
0 & 0 & 0 & 0 & f_{75} & 0 & f_{77} & 0 & f_{79} \\
\hline
0 & 0 & 0 & 0 & 0 & f_{86} & 0 & f_{88} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & f_{97} & 0 & f_{99}
\end{array}\right]
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \hline y_4 \\ y_5 \\ \hline y_6 \\ y_7 \\ \hline y_8 \\ y_9 \end{bmatrix}
+
\left[\begin{array}{cc}
0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \hline g_{41} & 0 \\ 0 & g_{52} \\ \hline 0 & 0 \\ 0 & 0 \\ \hline 0 & 0 \\ 0 & 0
\end{array}\right]
\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}
+
\begin{bmatrix} h_{11} \\ h_{21} \\ h_{31} \\ \hline 0 \\ 0 \\ \hline 0 \\ 0 \\ \hline 0 \\ 0 \end{bmatrix} \omega. \tag{6.14}
$$

The block structure induced by the chained aggregation with respect to the available measurements is shown by the dashed lines. The output equation associated with (6.14) is given by (6.11).

Having obtained the GHR we now exploit its structural properties to select a reduced order model. It is immediately seen that aggregation to a third order model is unsatisfactory. The result would be a system unaffected by the control, and acted upon by the external disturbance. Clearly the interconnection into the third order aggregate from the residual is significant. On the other hand, a fifth order reduced order model is justifiable. Separating the composite system into the aggregate subsystem and the residual subsystem, and denoting the states of the aggregate model by z, we have

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ \dot{z}_4 \\ \dot{z}_5 \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & f_{13} & 1 & 0 \\ f_{21} & f_{22} & f_{33} & 0 & 1 \\ f_{31} & f_{32} & f_{33} & 0 & 0 \\ 0 & 0 & f_{43} & f_{44} & 0 \\ 0 & 0 & f_{53} & 0 & f_{55} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_6 \\ y_7 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ g_{41} & 0 \\ 0 & g_{52} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} h_{11} \\ h_{21} \\ h_{31} \\ 0 \\ 0 \end{bmatrix} \omega .$$

and for the residual system

$$\begin{bmatrix} \dot{y}_6 \\ \dot{y}_7 \\ \dot{y}_8 \\ \dot{y}_9 \end{bmatrix} = \begin{bmatrix} f_{66} & 0 & f_{68} & 0 \\ 0 & f_{77} & 0 & f_{79} \\ f_{86} & 0 & f_{88} & 0 \\ 0 & f_{97} & 0 & f_{99} \end{bmatrix} \begin{bmatrix} y_6 \\ y_7 \\ y_8 \\ y_9 \end{bmatrix} + \begin{bmatrix} f_{64} & 0 \\ 0 & f_{75} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_4 \\ y_5 \end{bmatrix} .$$

Note first that all controls and external disturbances enter only into the aggregate model and not into the residual directly. Although the indirect effect of the disturbance (through feedforward from the aggregate and then as feedback from the residual to the aggregate) from the residual is coupled back to the aggregate by the Hessenberg block, such effect is secondary and if the coupling between the aggregate and the residual is weak, we may assume most effects of external disturbance inputs is accounted for in the aggregate model. Second, the aggregate model is completely controllable and, third, the interconnection variables which come from the residual may be cancelled by residual state feedback because $R(F_{23}) \subset R(G_2)$. Thus it is conjectured based on structural considerations that the reduced order model

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ \dot{z}_4 \\ \dot{z}_5 \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & f_{13} & 1 & 0 \\ f_{21} & f_{22} & f_{23} & 0 & 1 \\ f_{31} & f_{32} & f_{33} & 0 & 0 \\ 0 & 0 & f_{43} & f_{44} & 0 \\ 0 & 0 & f_{53} & 0 & f_{55} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix} + \begin{bmatrix} h_{11} \\ h_{21} \\ h_{31} \\ 0 \\ 0 \end{bmatrix} \omega + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ g_{41} & 0 \\ 0 & g_{52} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

will produce the basic features of the higher order model (6.7). Using the parameter values given in Appendix E, the GHR system matrices (F G,W,D) are given in Table 6.1a and the reduced order model matrices (F^2, G^2, W^2, D^2) are given in Table 6.1b.

With the assumption of a fifth order reduced order model, the Hessenberg block F_{23} is initially neglected to obtain an approximate GHR without compensation or parameter adjustment. The open-loop eigenvalues are compared in Table 6.2. The lack of resemblance in open-loop eigenvalues can be attributed to the fact that we are asking five modes to model the input-output behavior of a 9th order system. Thus an algebraic combination of the original system modes results in general. To assess the closed loop behavior, the three regulator designs given at the beginning of Chapter 6 are implemented. Two cases are considered. Case A assumes $Q = I$, $R = I$ in (6.2) and (6.5) while case B assumes $Q = 10\,I$, $R = I$ for the costs. The corresponding closed loop eigenvalue for the optimal and suboptimal regulators are presented in Table 6.3 and 6.4 respectively. In case A, it is seen that both the optimal and the suboptimal design show the same general pattern of pole locations, exhibiting a cluster of five dominant modes and four faster modes clustered in two pairs. We

Table 6.1a. GHR system representation of Example 6.1

$$F = \left[\begin{array}{ccc|cc|cc|cc}
-1.99375 & 0.00625 & 0.28125 & 1.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00625 & -1.99375 & 0.28125 & 0.00000 & 1.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.04167 & 0.04167 & -0.12500 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
\hdashline
0.00000 & 0.00000 & -5.71200 & -3.04440 & 0.00000 & 1.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & -5.71200 & 0.00000 & -3.04440 & 0.00000 & 1.00000 & 0.00000 & 0.00000 \\
\hdashline
0.00000 & 0.00000 & 0.00000 & 6.31430 & 0.00000 & -1.81294 & 0.00000 & -2.13398 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 6.31430 & 0.00000 & -1.81294 & 0.00000 & -2.13398 \\
\hdashline
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & +0.087985 & 0.00000 & -0.50933 & 0.00000 \\
0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & +0.087985 & 0.00000 & -0.50933
\end{array}\right]$$

$$G = \begin{bmatrix}
0.00000 & 0.00000 \\
0.00000 & 0.00000 \\
0.00000 & 0.00000 \\
5.71200 & 0.00000 \\
0.00000 & 5.71200 \\
0.00000 & 0.00000 \\
0.00000 & 0.00000 \\
0.00000 & 0.00000 \\
0.00000 & 0.00000
\end{bmatrix} \qquad W = \begin{bmatrix}
-0.01250 \\
-0.01250 \\
-0.08333 \\
0.00000 \\
0.00000 \\
0.00000 \\
0.00000 \\
0.00000 \\
0.00000
\end{bmatrix}$$

$$D = \begin{bmatrix}
1.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 1.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 \\
0.00000 & 0.00000 & 1.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000 & 0.00000
\end{bmatrix}$$

Table 6.1b. System matrices (F^2, G^2, W^2, D^2) of the aggregated model

$$F^2 = \begin{bmatrix} -1.99375 & 0.00625 & 0.28125 & 1.0000 & 0.0000 \\ 0.00625 & -1.99375 & 0.28125 & 0.0000 & 1.0000 \\ 0.04167 & 0.04167 & -0.12500 & 0.0000 & 0.0000 \\ 0.000 & 0.000 & -5.712 & -3.0444 & 0.000 \\ 0.000 & 0.000 & -5.712 & 0.000 & -3.0444 \end{bmatrix}$$

$$G^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 5.712 & 0 \\ 0 & 5.712 \end{bmatrix} \qquad W^2 = \begin{bmatrix} -0.0125 \\ -0.0125 \\ -0.0833 \\ 0.0 \\ 0.0 \end{bmatrix}$$

$$D^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Table 6.2. Open-loop eigenvalues of the original system and its approximate GHR

Open-loop system	Aggregate+residual system	
-0.01947+j0.36203	-0.205975	(A)
-0.01947-j0.36203	-0.67422	(R)
-0.16678	-0.67422	(R)
-0.19990	-1.64800	(R)
-0.43169	-1.64800	(R)
-1.98840	-1.77447	(A)
-2.0000	-2.00000	(A)
-5.00000	-3.04440	(A)
-5.02010	-3.17645	(A)

Table 6.3. Closed-loop eigenvalues of the optimal and suboptimal regulators for Case A

Closed-loop optimal system	Closed-loop suboptimal system
-0.17287	-0.27271+j0.43602
-0.69534+j0.31615	-0.27271-j0.43602
-0.69534-j0.31615	-0.30669
-0.73189+j0.27399	-0.37160+j0.37009
-0.73189-j0.27399	-0.37160-j0.37009
-2.1228	-2.0638
-2.1445	-2.0788
-4.8526	-4.9743
-4.8761	-4.9958

Table 6.4. Closed-loop eigenvalues of optimal and suboptimal regulators for Case B

Closed-loop optimal system	Closed-loop Suboptimal system
-0.13855	-0.14182
-0.67578	-0.70893
-0.67608	-0.71098
-1.63910	-1.57030
-1.63940	-1.57030
-5.92460+j4.70560	-5.67730+j4.38030
-5.92460-j4.70560	-5.67730-j4.38030
-5.92660+j4.70780	-5.68090+j4.38470
-5.92660-j4.70780	-5.68090-j4.38470

note even the closed-loop eigenvalues are not very close in values, it is the input-output behavior that is stressed by the GHR approach to model reduction. This is exhibited in Fig. 6.1a for the power outputs y_1 and y_2 under an initial condition disturbance of $y_o = (0.1, 0.1, 0)^T$ and in Fig. 6.1b for the system frequency y_3 under a 10% frequency deviation $y_o = (0,0,0.1)^T$. In case B, because of the larger-valued weighting matrix Q,the output deviation from zero is penalized more, thus more control effect is allowed, resulting a good agreement in closed loop eigenvalue approximation. The analogous plots for case B is given in Fig. 6.2a and Fig. 6.2b with same initial condition as in Fig. 6.1.

Thus far, we have seen (Table 6.2) that by simply neglecting the coupling $F_{23}z_3$ between S^2 and $\bar{S}^2$, while the reduced aggregated model is able to give good approximation to input-output behavior in closed loop, it fails to predict the open-loop behavior. Next consider the compensation of $F_{23}z_2$ by parameter adjustment as indicated by the RQL algorithm. The result obtained after 5 RQL iterations is a model characterized by the same matrices as in Table 6.1b except that parameters f_{44} and f_{55} in F^i have been changed from -3.0444 to -0.1520. Table 6.5 contains the open-loop eigenvalues of the original and the reduced order model as well as the closed-loop eigenvalues of the optimal and suboptimal regulators. Here we have assume $Q = I$ and $R = I$. The open loop output trajectories are plotted in Fig. 6.3 for an initial condition of $y_o = (0.1, 0.1, 0)^T$. Notice the oscillatory behavior of the full order open loop system is now retained in the 5th order reduced model after the parameter adjustment by the RQL algorithm. The closed loop behavior of the optimal,reduced,and suboptimal regulators using the same initial

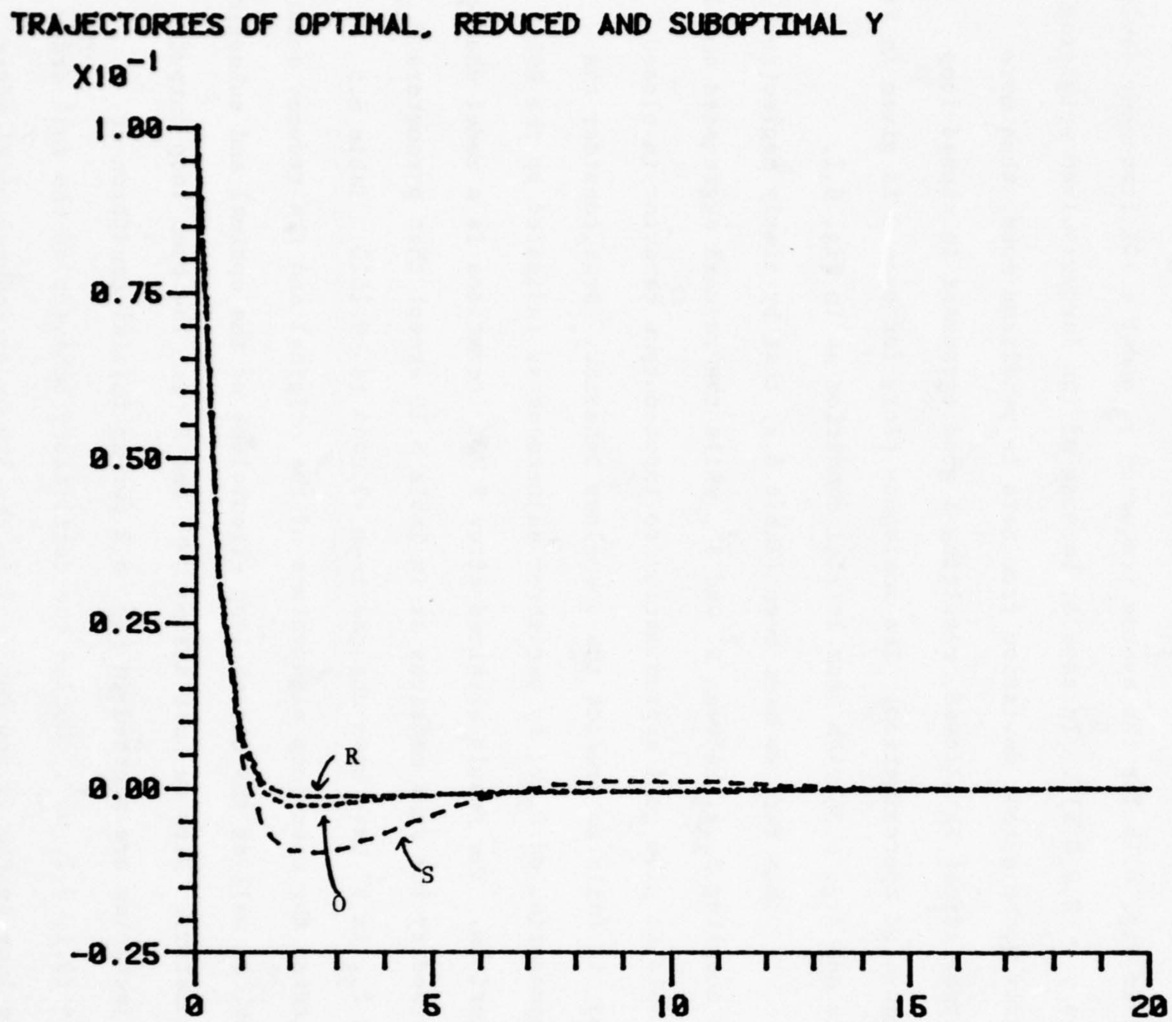


Figure 6.1a. Power outputs (y_1, y_2) of thermal units #1 and #2 for case A.

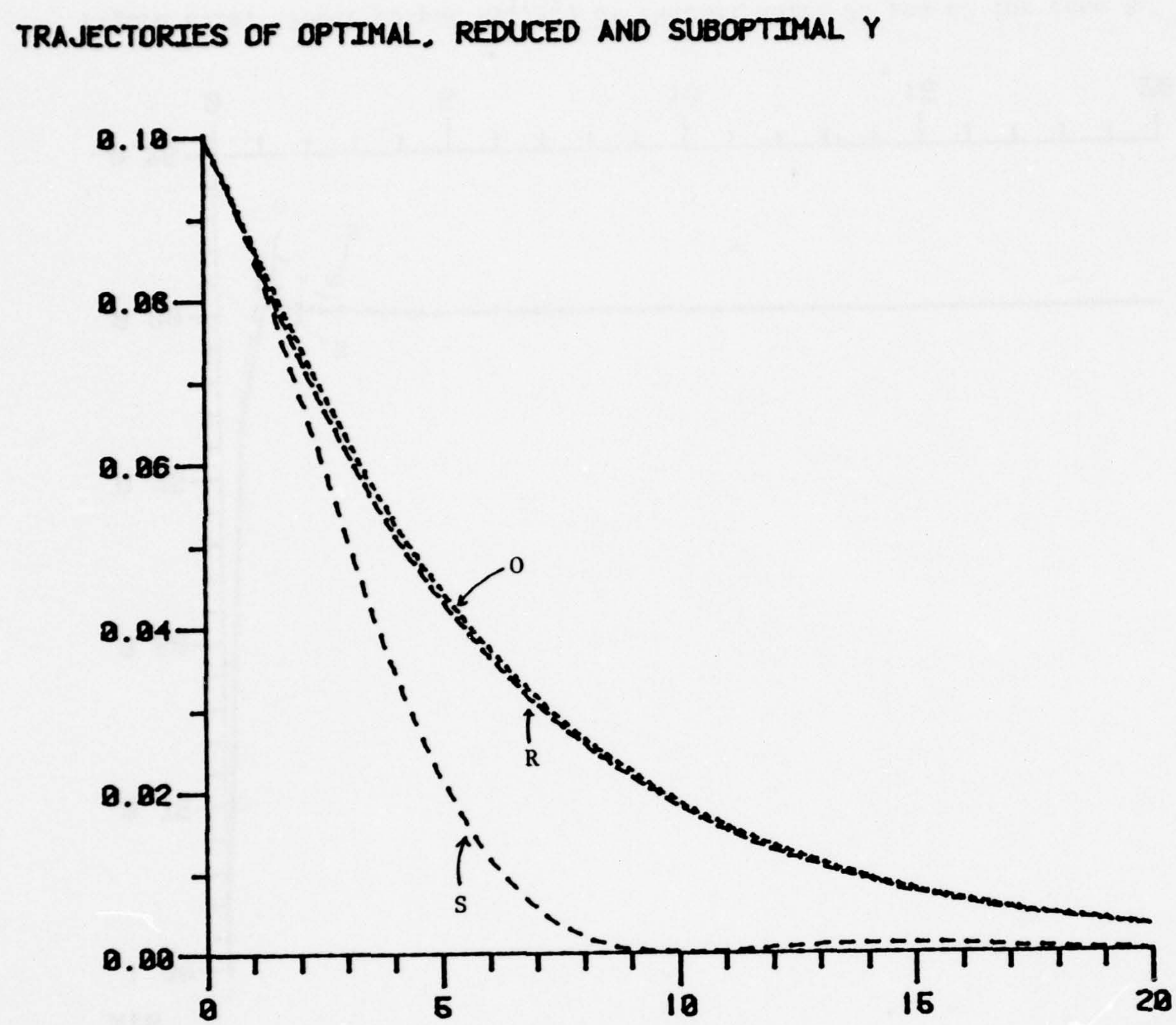


Figure 6.1b. System frequency (y_3) for case A.

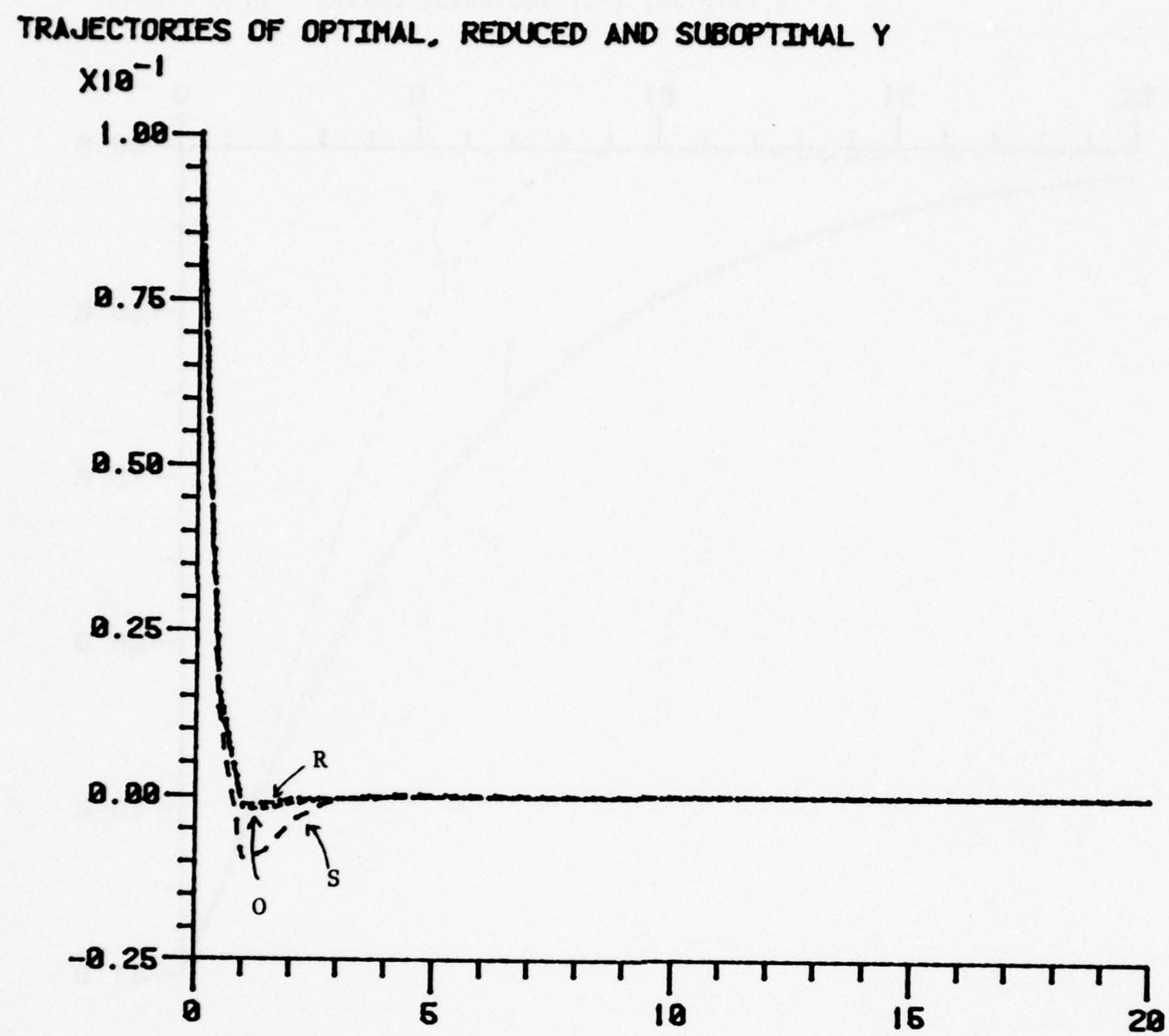


Figure 6.2a. Power output (y_1, y_2) of thermal units #1 and #2 for case B.

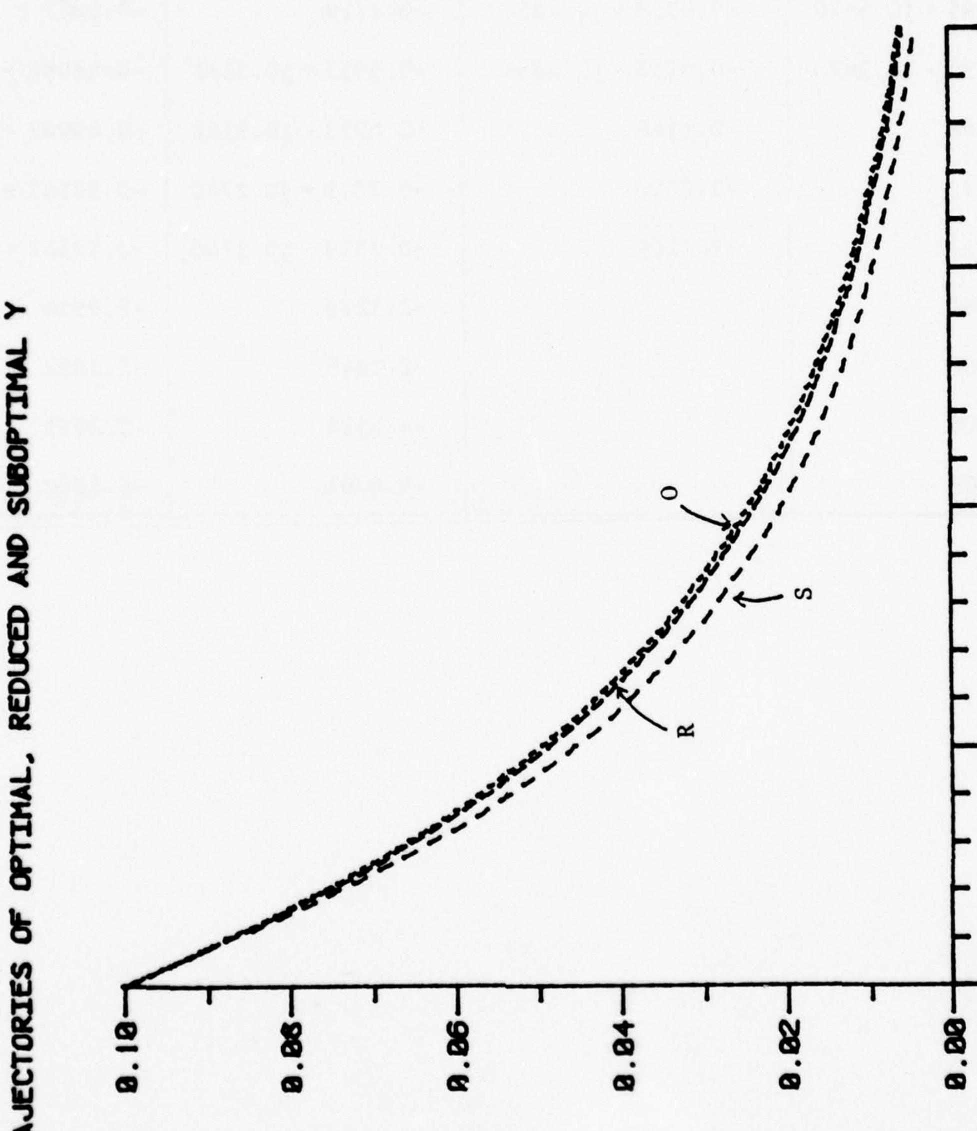


Figure 6.2b. System frequency (y_3) for case B.

Table 6.5. Eigenvalues of the RQL based reduced order model and full order system.

Open-loop system	Open-loop Reduced order model	Closed-loop optimal system	Closed-loop sub-optimal system
-0.0195 + j0.3620	-0.0758 + j0.4854	-0.1729	-0.1827
-0.0195 - j0.3620	-0.0758 - j0.4854	-0.6953 + j0.3162	-0.49082 + j0.4426
-0.1668	-0.1668	-0.6953 - j0.3162	-0.49082 - j0.4426
-0.1999	-2.0000	-0.7319 + j0.2740	-0.52547 + j0.42175
-0.4317	-2.1209	-0.7319 - j0.2740	-0.52547 - j0.42175
-1.9884		-2.1228	-2.0954
-2.0000		-2.1445	-2.1052
-5.0000		-4.8526	-5.4971
-5.0201		-4.8761	-5.5098

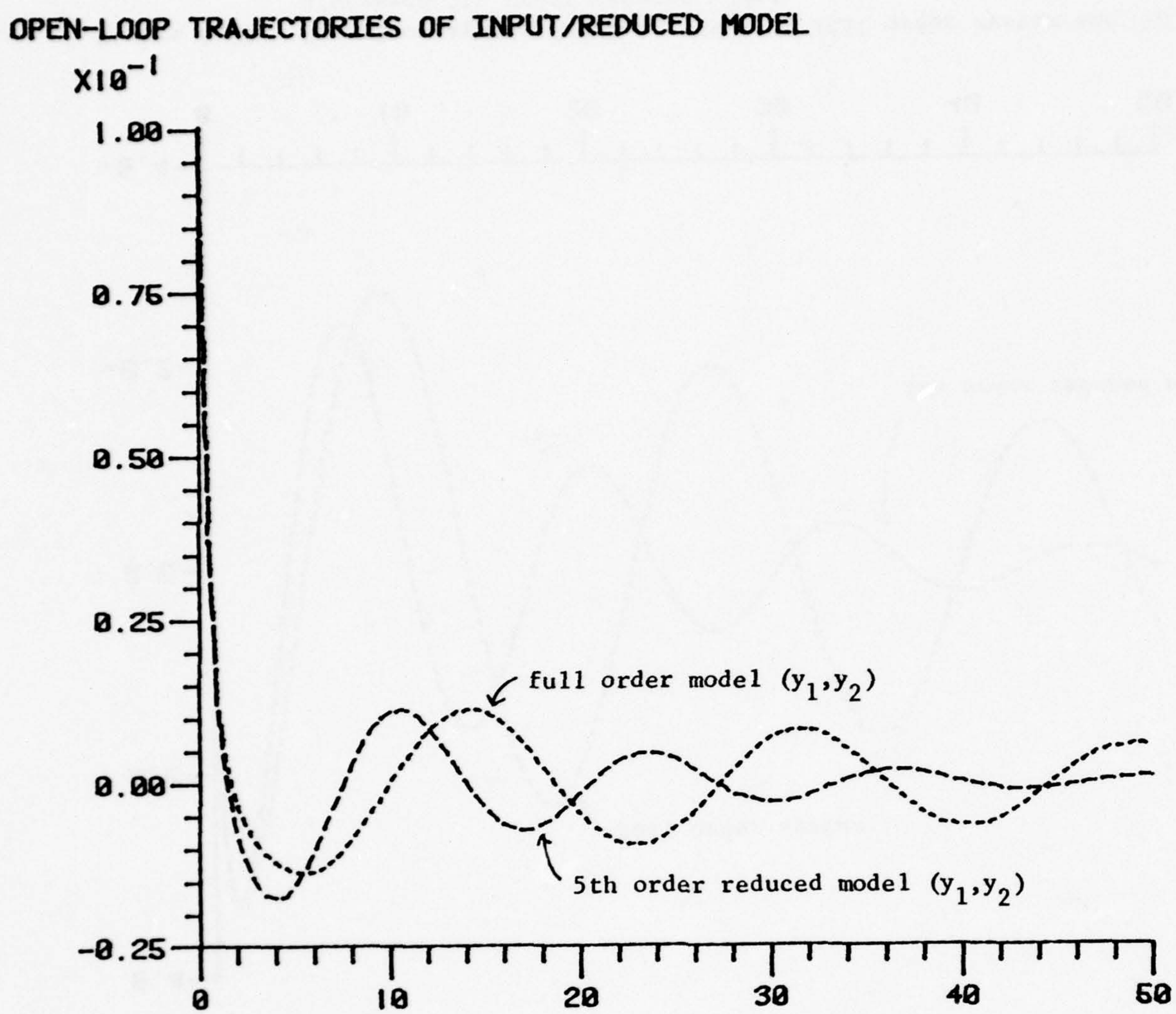


Figure 6.3a. Open-loop power output (y_1,y_2) of full order system and RQL processed 5th order reduced model.

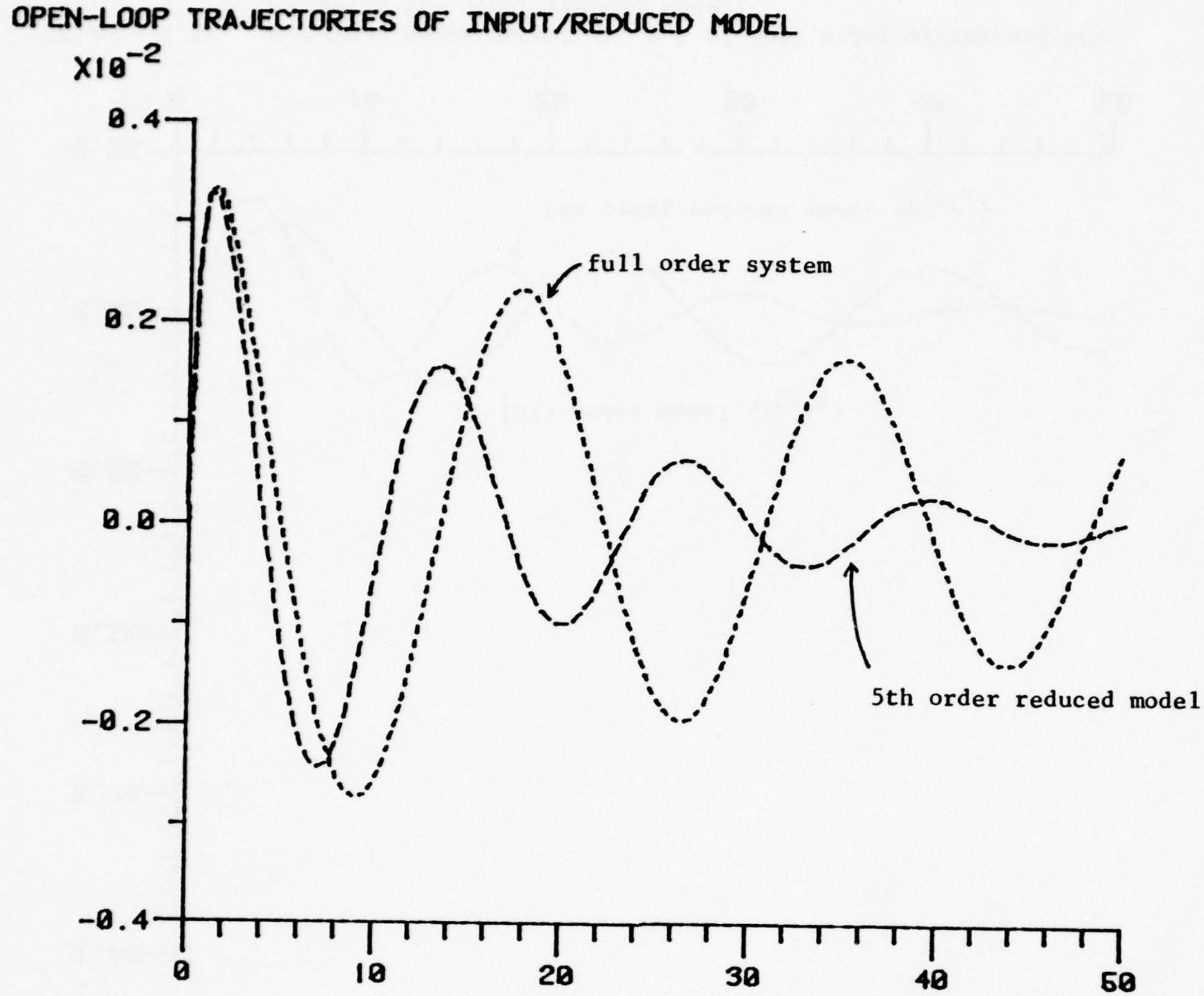


Figure 6.3b. Open-loop system frequency (y_3) of full order system and RQL processed 5th order reduced model.

condition are plotted in Fig. 6.4. Good agreement is evident. Hence the RQL processed GHR leads to a good reduced order model that is able to approximate both open loop and closed loop input-out behavior.

6.2 Model Reduction of Pitch Dynamic of Flexible Rocket

This example shows the simplicity of model reduction using the GHR. A 6th order system representing the pitch dynamic of a flexible rocket [60] is considered. In [60], Rogers and Sworder obtained a 2nd order reduced model using projection theory in Hilbert space. Their procedure involves evaluation of inner products, solution of non-linear matrix equations and use of gradient search in order to obtain reduced order model parameters, thus is numerically involved. We shall attempt a comparison by showing a second order model may be obtain via GHR using only linear transformations. The full order system of 6th order is given by:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} -.21053 & -.10526 & -.0007378 & 0 & .0706 & 0 \\ 1.0 & -.03537 & -.000118 & 0 & .0004 & 0 \\ 0 & 0 & 0 & 1.0 & 0 & 0 \\ 0 & 0 & -605.16 & -4.92 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.0 \\ 0 & 0 & 0 & 0 & -3906.25 & -12.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} + \begin{bmatrix} -7.211 \\ -.05232 \\ 0 \\ 794.7 \\ 0 \\ -448.5 \end{bmatrix} u \tag{6.15a}$$

and

$$\begin{bmatrix} y_{m1} \\ y_{m2} \end{bmatrix} = \begin{bmatrix} 1.0 & 0 & .000334 & 0 & -.007728 \\ 0 & 1.0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} \tag{6.15b}$$

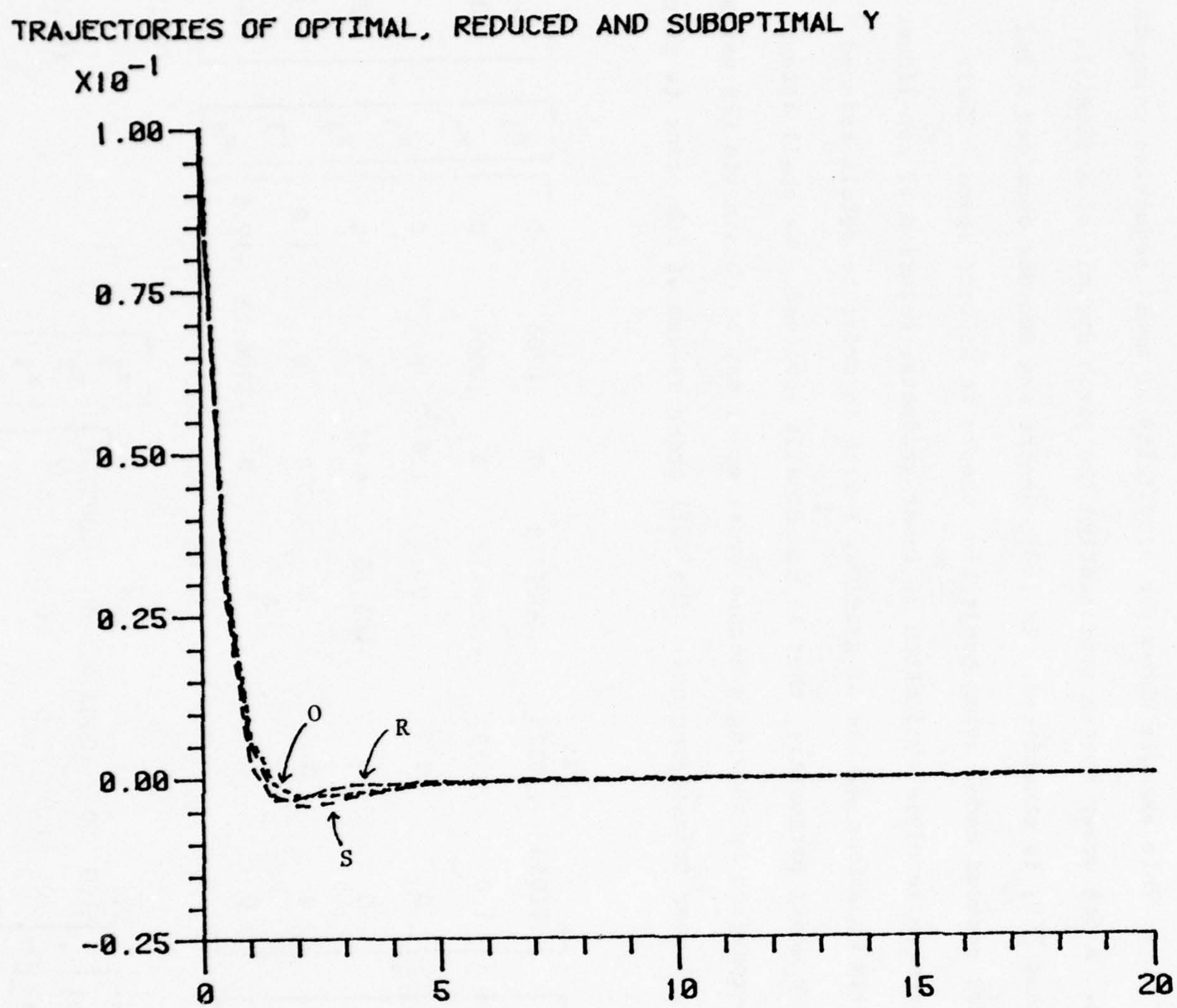


Figure 6.4a. Power output (y_1, y_2) of optimal, reduced and suboptimal regulators using 5th order RQL processed reduced model.

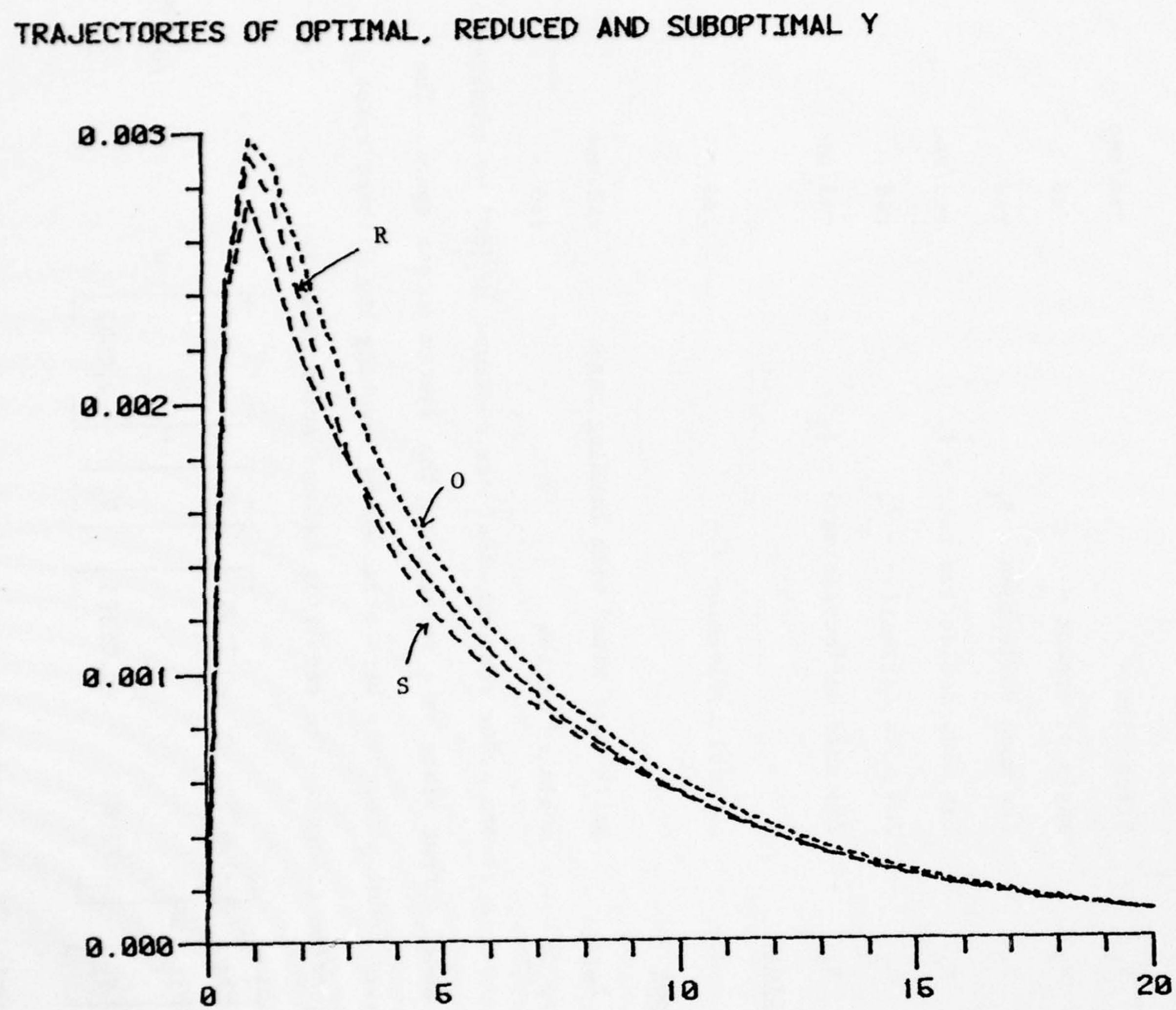


Figure 6.4b. System frequency (y_3) of the optimal, reduced and suboptimal regulators using 5th order RQL processed reduced model.

$$J = \frac{1}{2}\int_o^\infty (y^TQy + u^TRu)dt \tag{6.16}$$

The state and control vectors have the following physical meaning:

<u>state vector</u>

x_1:	-	pitch rate $= \dot{\theta}$	rad/sec
x_2:	-	angle of attack $= \alpha$	rad
x_3:	-	1st mode deflection $= \xi_1$	rad
x_4:	-	1st mode deflection rate $= \dot{\xi}_1$	rad/sec
x_5:	-	2nd mode deflection $= \xi_2$	rad
x_6:	-	2nd mode deflection rate $= \dot{\xi}_2$	rad/sec

<u>control</u>

u:	-	control deflection $= \delta$	rad

<u>output</u>

y_{m1}:	-	pitch rate mixed with bending rate	rad/sec
y_{m2}:	-	angle of attack	rad

A second order reduced model was obtained in [60] by considering the model output space as a subspace of the system output space. The reduced model parameters were obtained by invoking the Hilbert space projection theorem. The resulting reduced order model is:

$$\begin{bmatrix} \dot{\tilde{y}}_1 \\ \dot{\tilde{y}}_2 \end{bmatrix} = \begin{bmatrix} -1.415 & -0.75758 \\ 0.75081 & -0.02573 \end{bmatrix} \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} + \begin{bmatrix} -6.7317 \\ -0.0002 \end{bmatrix} u \tag{6.17}$$

The model (6.17) approximates the actual system output (6.15b). Using chained aggregation of Chapter 2, the 6th order system (6.15) is given equivalently in GHR as:

$$\begin{bmatrix} \dot{y}_1 \\ \cdots \\ \dot{y}_2 \\ \cdots \\ \dot{y}_3 \end{bmatrix} = \left[\begin{array}{cc:cc:cc} -.21053 & -.10526 & 1.0 & 0 & 0 & 0 \\ 1.0 & -.03537 & 0 & 1.0 & 0 & 0 \\ \hdashline 0 & 0 & -5.0461 & 607.96 & -218.55 & 24.852 \\ 0 & 0 & -.99644 & .12612 & -.037194 & -.0025395 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 1.0 \\ 0 & 0 & 0 & 0 & -3906.3 & -12.5 \end{array}\right] \begin{bmatrix} y_1 \\ \cdots \\ y_2 \\ \cdots \\ y_3 \end{bmatrix} + \left[\begin{array}{c} -3.4796 \\ -.05232 \\ \hdashline -43.845 \\ -3.7314 \\ \hdashline 0 \\ -448.5 \end{array}\right] u \tag{6.18a}$$

$$y_m = \left[\begin{array}{cc:cc:cc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} y_1 \\ \cdots \\ y_2 \\ \cdots \\ y_3 \end{bmatrix} \tag{6.18b}$$

We point that the GHR (6.18) is particularly simple to obtain. Only one step of chained aggregation is required. In this representation, the substate vector y_1 represents the original system measurements. As discussed in Chapter 4, different order reduced models may be obtained by neglecting either one of the Hessenberg blocks F_{12} or F_{23}. To compare with reduced model of Rogers and Sworder, the crudest aggregated model is obtained by simply neglecting F_{12} in the Hessenberg system matrix without any compensation by feedback or parameter adjustment. This is represented as a second order aggregated model:

$$\dot{y}_1 = \begin{bmatrix} -.21053 & -.10526 \\ 1.0 & -.03537 \end{bmatrix} y_1 + \begin{bmatrix} -3.4796 \\ -.05232 \end{bmatrix} u \tag{6.19}$$

Note the reduced order model is completely controllable in this case. When F_{23} is neglected, a 4th order reduced order model is obtained:

$$\begin{bmatrix} \dot{y}_1 \\ \cdots \\ \dot{y}_2 \end{bmatrix} = \left[\begin{array}{cc:cc} -.21053 & -.10526 & 1.0 & 0 \\ 1.0 & -.03537 & 0 & 1.0 \\ \hdashline 0 & 0 & -5.0461 & 607.96 \\ 0 & 0 & -.99644 & .12612 \end{array}\right] \begin{bmatrix} y_1 \\ \cdots \\ y_2 \end{bmatrix} + \begin{bmatrix} -3.4796 \\ -.05232 \\ \cdots \\ -43.845 \\ -3.7314 \end{bmatrix} u \tag{6.20a}$$

$$y_m = \left[\begin{array}{cc:cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] \begin{bmatrix} y_1 \\ \cdots \\ y_2 \end{bmatrix} . \tag{6.20b}$$

Again, this reduced model is controllable as well as observable. Equation (6.20b) now represents an approximation of the actual system output.

To illustrate the ability of reduced models to approximate open loop behavior of the full order system, the open loop eigenvalues are given in Table 6.6. Notice the two dominant open loop eigenvalues $-0.12 \pm j0.31$ of the full order system are retained in both the 2nd order and the 4th order aggregated model whereas the open loop poles of model of [60] do not approximate any actual system poles. The open loop output trajectories of the system (6.18) and its two reduced order models, i.e. (6.17) and (6.19) are plotted in Fig. 6.5 for an initial condition of $y_o = (1\ ,\ 1)^T$. Notice the 2nd order aggregate and the full order system have essentially some open-loop output trajectories while those of Rogers and Sworder show no resemblance at all.

Next consider the closed loop behavior. Same as in [60], the objective is to minimize (6.16) for $Q = I$, $R = 1$. The closed loop eigenvalues

Table 6.6. Comparison of open-loop eigenvalues of flexible rocket

Full Order System	Rogers and Sworder	2nd Order Aggregate	4th Order Aggregate
$-0.12 \pm j\ 0.31$	$-0.84 \pm j\ 0.48$	$-0.12 \pm j\ 0.31$	$-0.12 \pm j\ 0.31$
$-2.46 \pm j\ 24.48$			$-2.46 \pm j\ 24.48$
$-6.25 \pm j\ 62.19$			

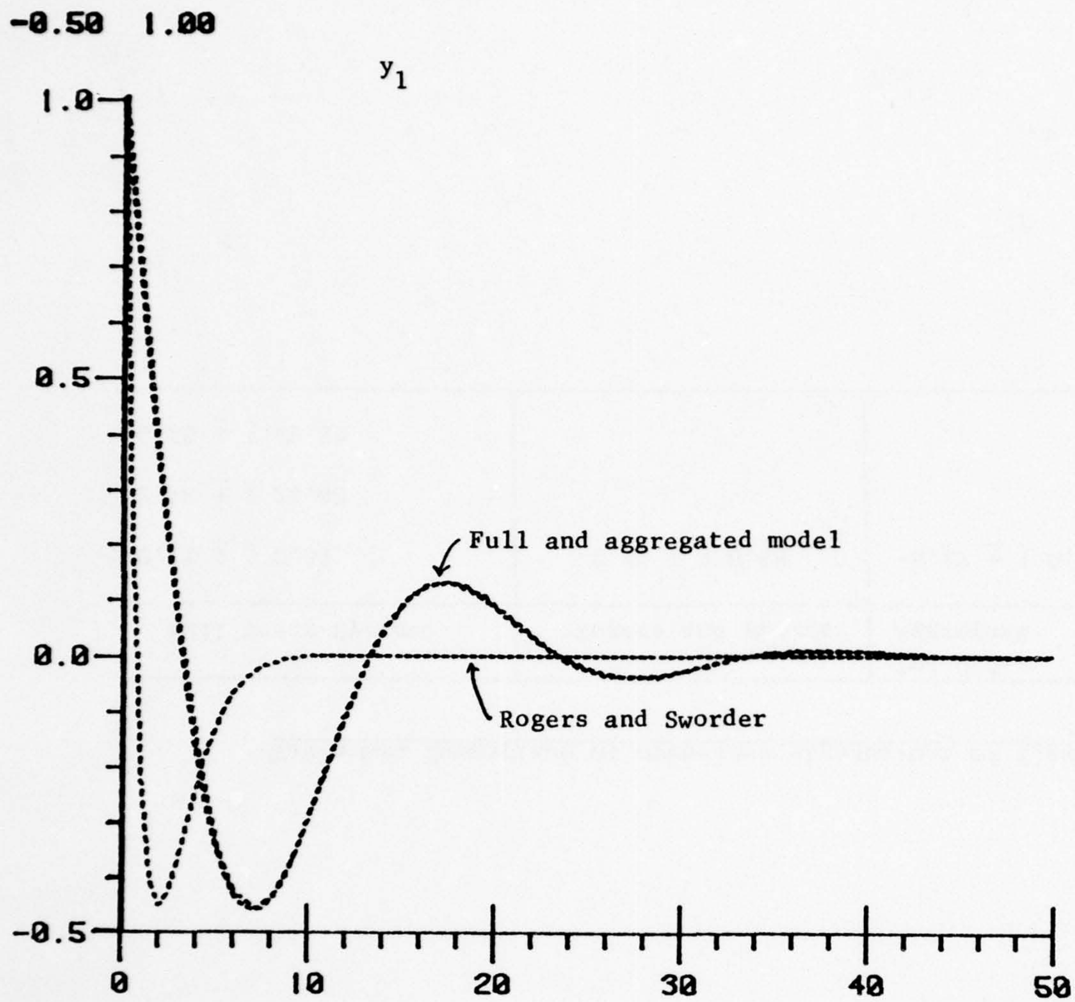


Figure 6.5. Open loop output trajectory comparison of full order model and 2nd order reduced order models from [60] and by GHR.

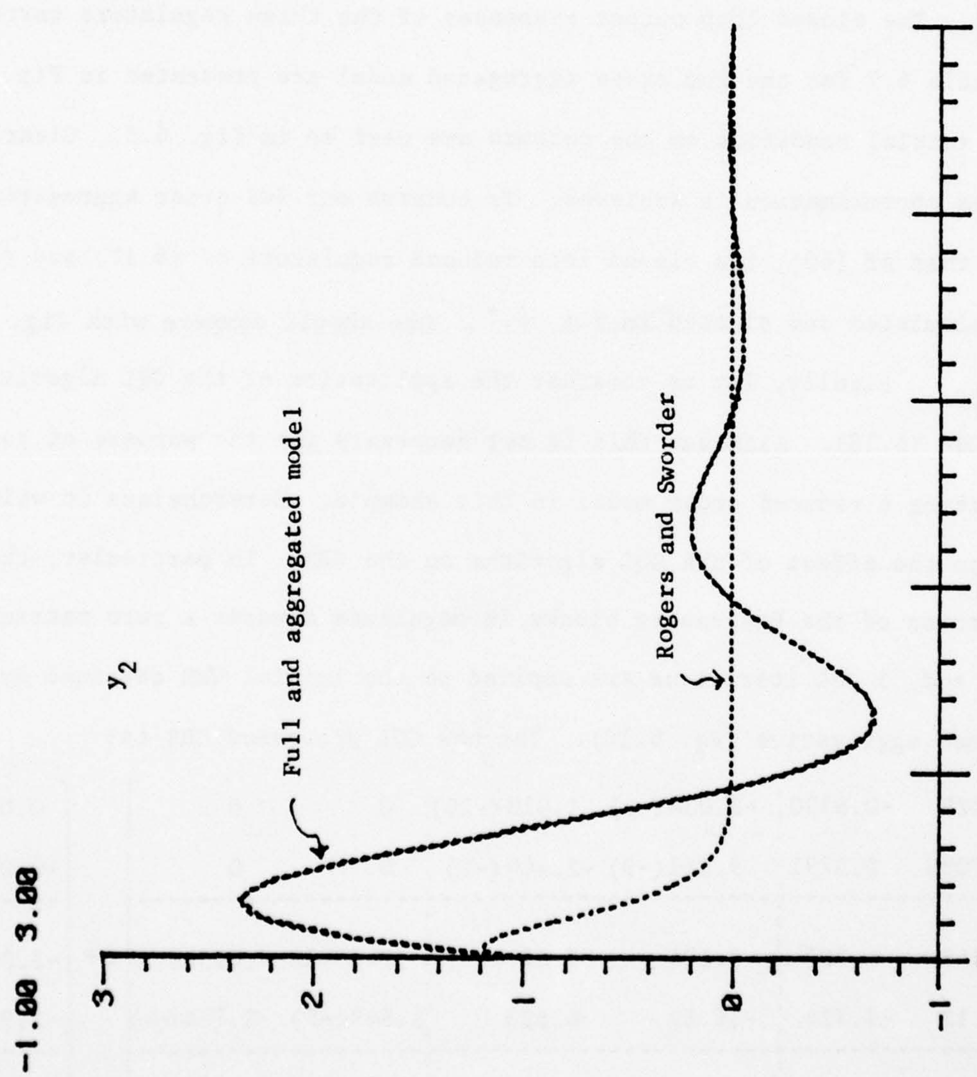


Figure 6.5. continued

are compared in Table 6.7 and 6.8 respectively for the 2nd and the 4th order reduced models. It is interesting to note that the closed loop poles of both aggregated models (6.19)-(6.20) obtained from the GHR give a much closer approximation to the optimal closed loop poles than the model by Rogers and Sworder, thus indicating the advantage of using the given information structure to induce a structurally meaningful reduced order model. The closed loop output responses of the three regulators correspond to Table 6.7 for the 2nd order aggregated model are presented in Fig. 6.6. Same initial condition on the outputs are used as in Fig. 6.5. Clearly, a good approximation is achieved. To compare our 2nd order aggregated model with that of [60], the closed loop reduced regulators of (6.17) and (6.19) are simulated and plotted in Fig. 6.7. One should compare with Fig. 6.6.

Finally, let us consider the application of the GQL algorithm to the GHR (6.18). Although this is not necessary for the purpose of constructing a reduced order model in this example. Nevertheless it will demonstrate the effect of the GQL algorithm on the GHR. In particular, the reduction of the Hessenberg blocks in magnitude towards a zero matrix. To this end, 5 GQL iterations are applied to the initial GHR obtained by chained aggregation (Eq. 6.18). The new GQL processed GHR is:

$$\dot{z} = \left[\begin{array}{cc:cc:cc} -0.825 & -0.8330 & -2.034(-9) & 1.818(-10) & 0 & 0 \\ 0.7038 & 0.5791 & 9.561(-9) & -2.404(-9) & 0 & 0 \\ \hdashline 4.348 & 4.307 & -9.174 & 3.90 & -0.00412 & .02563 \\ -10.13 & -9.724 & -16.32 & 4.523 & 3.649(-5) & -2.784(-4) \\ \hdashline 84.83 & 80.29 & 1421 & -3322 & -36.59 & 119.90 \\ -310.3 & -299.8 & 57.32 & -1281 & -39.91 & 23.82 \end{array}\right] z + \left[\begin{array}{c} 0.003622 \\ -0.0164 \\ \hdashline -3.005(7) \\ -2.973(7) \\ \hdashline -1.659(11) \\ -2.389(11) \end{array}\right] u \qquad (6.21a)$$

Table 6.7. Comparison of closed-loop eigenvalues for 2nd order reduced order model

Optimal Reduced Regulator		Optimal Full Regulator	Suboptimal Regulator	
2nd Order Aggregate	Rogers and Sworder		2nd Order Aggregate	Rogers and Sworder
		-1.01	-0.91	-0.79
-1.06	-0.8	$-2.46 \pm j\,24.51$	$-2.32 \pm j24.54$	$-1.38 \pm j24.88$
-3.29	-6.75	$-6.48 \pm j62.35$	$-4.25 \pm j62.62$	$-4.73 \pm j62.41$
		-7.07	-7.69	-5.64

Table 6.8 Comparison of losed-loop igenvalues for 4th rder educed rder odel

Optimal Reduced Regulator	Optimal Full Regulator	Suboptimal Regulator
-1.01	-1.01	-1.02
$-3.16 \pm j\ 24.92$	$-2.46 \pm j\ 24.51$	$-2.49 \pm j\ 24.52$
-6.84	$-6.48 \pm j\ 62.35$	-6.71
	-7.07	$-6.97 + j\ 63.79$

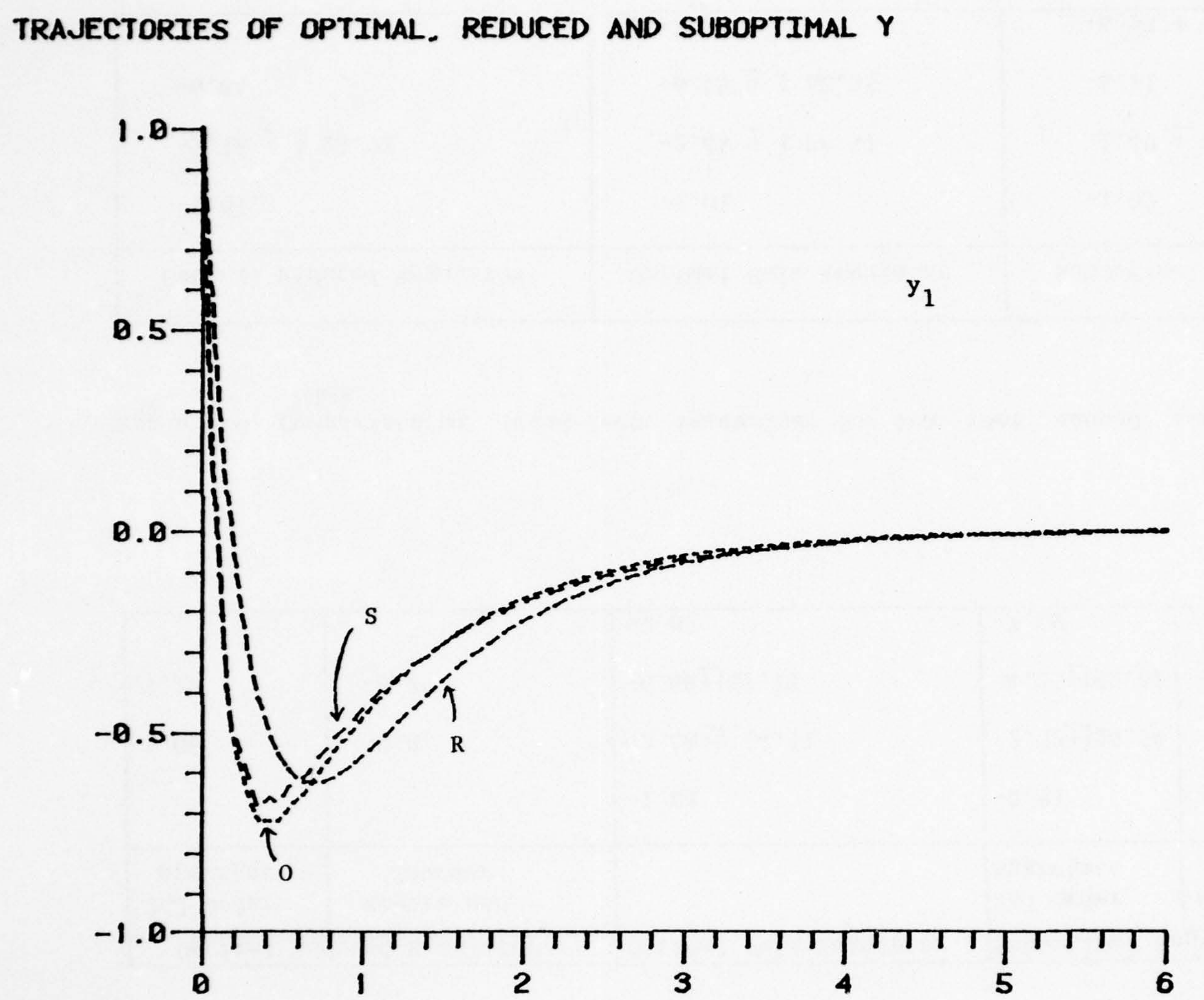


Figure 6.6. Closed-loop output trajectories of optimal, reduced and suboptimal regulators using 2nd order aggregated model.

TRAJECTORIES OF OPTIMAL, REDUCED AND SUBOPTIMAL Y

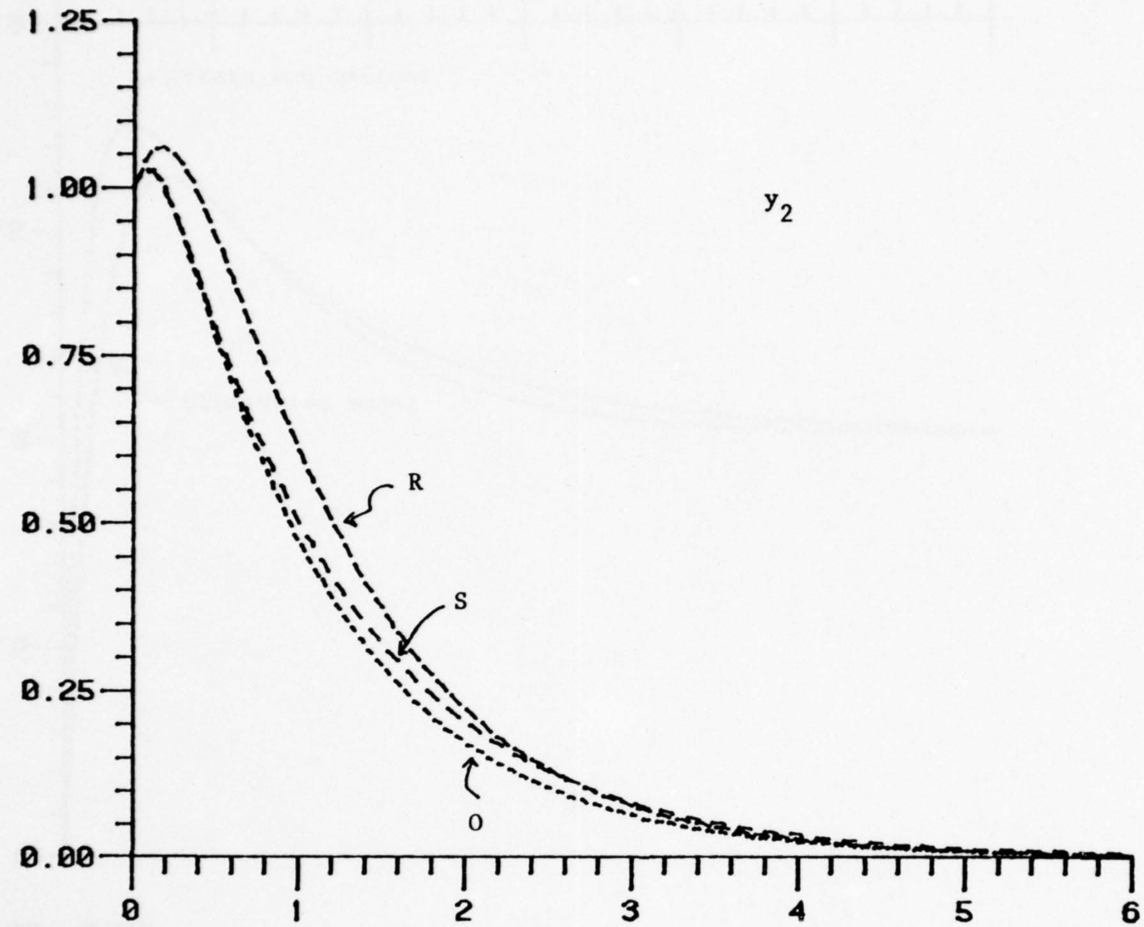


Figure 6.6. continued

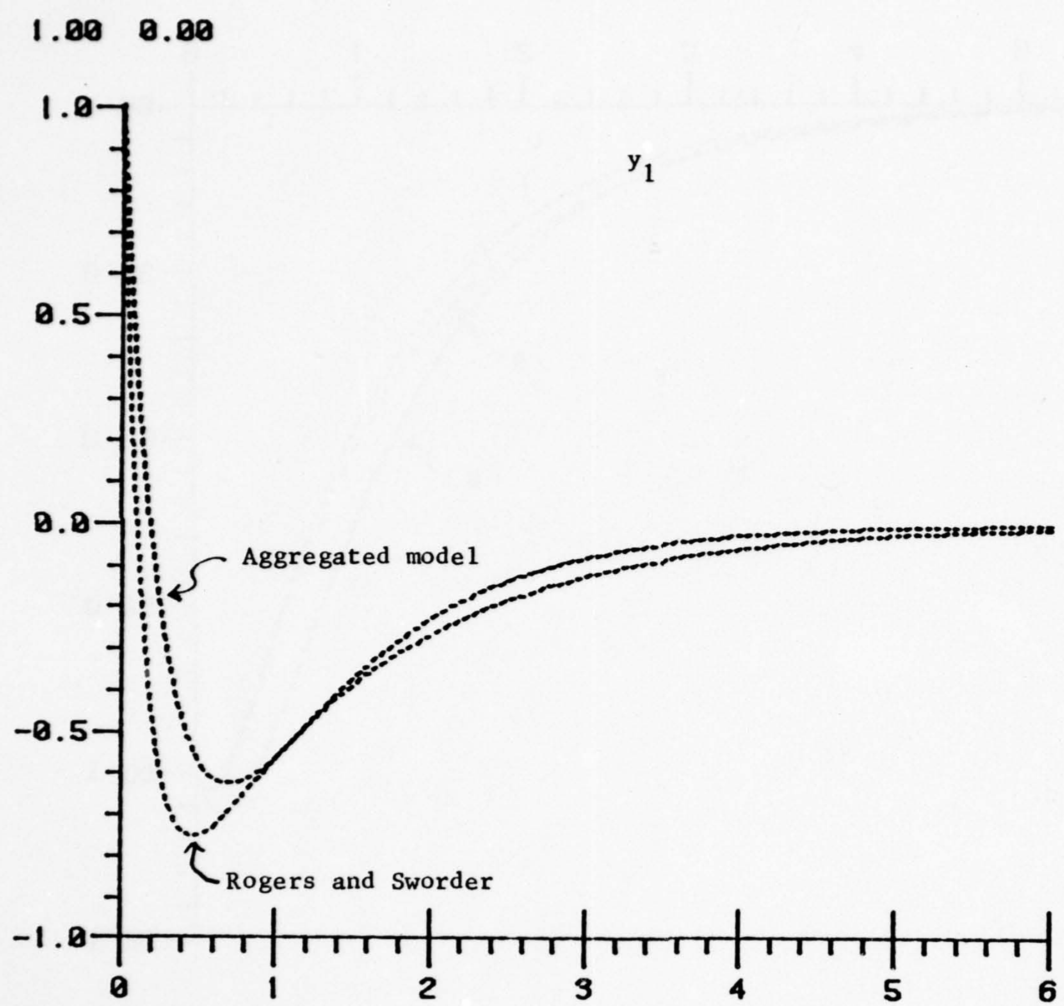


Figure 6.7. Closed-loop output trajectory comparison of 2nd order model by Rogers and Sworder and by GHR.

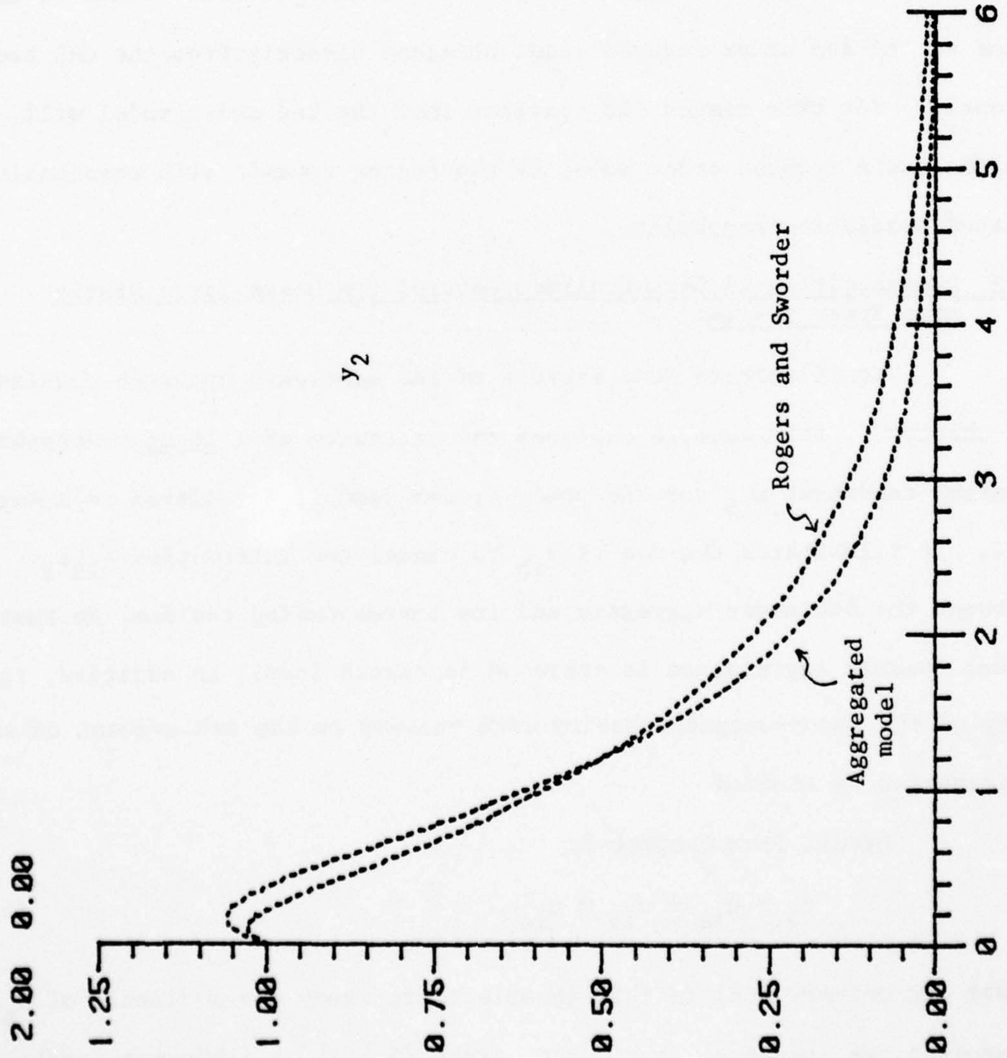


Figure 6.7. continued

$$y_m = \begin{bmatrix} 241.6 & 265.6 & 0 & 0 & 0 & 0 \\ -378.0 & -80.31 & 0 & 0 & 0 & 0 \end{bmatrix} z \qquad (6.21b)$$

From (6.21a), it is noted that the Hessenberg block F_{12} is converging to zero much faster than the other Hessenberg block F_{23}. In addition, the eigenvalues of the leading 2x2 block are $-0.1229 \pm j0.3124$. This is the same as the 2nd order reduced model obtained directly from the GHR previously. For this reason, it confirms that the 2nd order model will suffice as a reduced order model of the rocket dynamic with reasonable output prediction capability.

6.3 Decomposition in Decentralized Control Synthesis for a Single Area Power System.

To illustrate some aspects of the synthesis approach considered in Chapter 5, this example explores the existence of a local compensatory control component u_{ib} for the power system example considered in Example 6.1. It illustrates the use of u_{ib} to cancel the interaction $F_{23}z_3$ between the 5th order aggregate and its corresponding residual so that exact dynamic aggregation is achieved in closed loop. In addition, the sensitivity of the input-output behavior with respect to the achievement of exact aggregation is studied.

Recall from Chapter 5,

$$u_i = u_{ia} + u_{ib} + u_{ic}. \qquad (6.22)$$

Since our primary goal in this example is to study the influence of u_{ib} component, we assume $u_{ia} = u_{ic} = 0$. From (6.14), to achieve aggregability so that $z^2 = [z_1\ z_2\ \dots\ z_5]^T$ is decoupled from its residual $\bar{z}^2 = [z_6\ \dots\ z_9]^T$

it is sufficient to apply a feedback compensation in GHR of the form:

$$\begin{bmatrix} u_{1b} \\ u_{2b} \end{bmatrix} = \left[\begin{array}{ccccc:cccc} 0 & 0 & 0 & 0 & 0 & -0.17507 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.17507 & 0 & 0 \end{array}\right] \begin{bmatrix} z^2 \\ \hdashline \bar{z}^2 \end{bmatrix} \tag{6.23}$$

Using the GHR transformation matrix H given in Table 6.9 which transforms (6.7)-(6.8) to (6.14), the compensatory control component u_b that cancels the interaction $F_{23}z_3$ is expressed in original system representation as

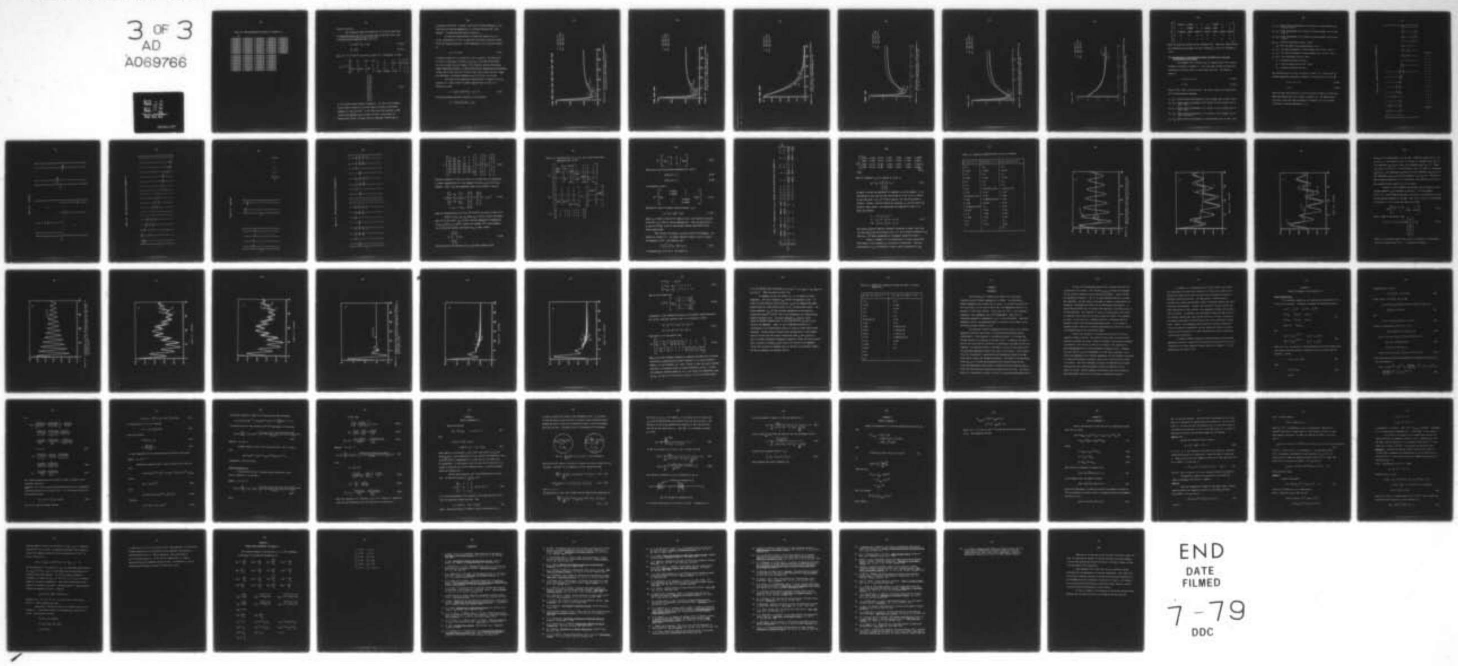
$$\begin{bmatrix} u_{1b} \\ u_{2b} \end{bmatrix} = \begin{bmatrix} 0 & -0.33164 & -0.68179 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.33164 & -0.68179 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \hdashline x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{bmatrix} \tag{6.24}$$

where the state x_i has the physical meaning given in Example 6.1. Note we have partitioned the state into two sub-states. The first substate is composed of state variables of the first thermal plant while the second sub-state is composted of states x_5 to x_9 which are the state variables of the second thermal plant plus the system frequency deviation. Equation (6.24) clearly indicates that feedback compensation to achieve aggregability can be obtained by local subsystem feedback. This is much

AD-A069 766 ILLINOIS UNIV AT URBANA-CHAMPAIGN COORDINATED SCIENCE LAB F/G 12/1
MODEL REDUCTION AND DECENTRALIZED CONTROL OF LARGE SCALE SYSTEM--ETC(U)
AUG 78 E C TSE DAAB07-72-C-0259

UNCLASSIFIED R-820 NL

3 OF 3
AD
A069766



END
DATE
FILMED
7-79
DDC

Table 6.9. GHR transformation matrix H of example 6.3

H

0.00000E+00	0.30000E+00	0.28000E+00	0.42000E+00	0.00000E+00
0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00
0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00
0.14280E+01	-0.85332E+00	0.13533E+01	0.00000E+00	0.00000E+00
0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00	0.14280E+01
0.00000E+00	0.18943E+01	0.38944E+01	0.00000E+00	0.00000E+00
0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00
0.00000E+00	0.00000E+00	0.10000E+01	0.00000E+00	0.00000E+00
0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00

0.00000E+00	0.00000E+00	0.00000E+00	0.15000E+00
0.30000E+00	0.28000E+00	0.42000E+00	0.15000E+00
0.00000E+00	0.00000E+00	0.00000E+00	0.10000E+01
0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00
-0.85332E+00	0.13533E+01	0.00000E+00	0.00000E+00
0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00
0.18943E+01	0.38944E+01	0.00000E+00	0.00000E+00
0.00000E+00	0.00000E+00	0.00000E+00	0.00000E+00
0.00000E+00	0.10000E+01	0.00000E+00	0.00000E+00

desired in practice.

Let us emphasize that the application of (6.24) has the effect of replacing elements f_{46} and f_{57} by hard zeros in the GHR (6.14), thus resulting an exact aggregated model for (6.7)

$$\dot{z}^2 = F^2 z^2 + G^2 u + W^2 \omega \tag{6.25a}$$

$$\hat{y} = D^2 z^2 \tag{6.25b}$$

where (F^2, G^2, W^2 and D^2) are given in Table 6.1b. Furthermore, we have

$$y_e = z^2 = \begin{bmatrix} 0 & 0.3 & 0.28 & 0.42 & 0 & 0 & 0 & 0 & 0.15 \\ 0 & 0 & 0 & 0 & 0 & 0.3 & 0.28 & 0.42 & 0.15 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1.428 & -0.85332 & 1.3533 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.428 & -0.85332 & 1.3533 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{bmatrix} \tag{6.26}$$

as the extended output defined in Chapter 5. We note in this example, because numer of inputs is less than number of outputs, the freedom component u_c does not exist. On the other hand, the component u_a that controls the aggregate may be sought as either a centralized or a decentralized control by proper choice of aggregate feedback gain in conjunction with (6.26). Because of the lack of freedom component u_c, the price paid in achieving aggregation is a residual subsystem with fixed dynamics. Its eigenvalues are given in Table 6.2.

To study the sensitivities of (6.25a) with respect to u_c of (6.24), (equivalently (6.23)), we argue that the output trajectory (6.25b) should not change excessively if 100% compensation is not achieved, namely, if

$$u_c = (L + \Delta L)\,z \tag{6.27}$$

is applied instead of $u_c = Lz$ given by (6.23) is applied. The application of (6.27) is equivalent to replace f_{46} and f_{57} of the GHR system matrix in (6.14) by small non-zero numbers. The simulation results of the three output trajectories subject to complete and partial compensations are presented in Fig. 6.8a to Fig. 6.8c for $\omega = 0$ and in Fig. 6.9a to Fig. 6.9c for ω equal to a unit step. An initial condition of $z_o = [1,1,0.5,0...0]$ is used.

As a check to the sufficient condition given in Chapter 5 for the existence of a local control to achieve aggregability, we note it is sufficient to have

$$[I - (C_eB)\,(C_eB)^+](F^2C_e - C_eA) \equiv 0 \tag{6.28}$$

Using the appropiate matrices involved, it is found that

$$[I - (C_eB)(C_eB)^+](F^5C_e - C_eA)$$

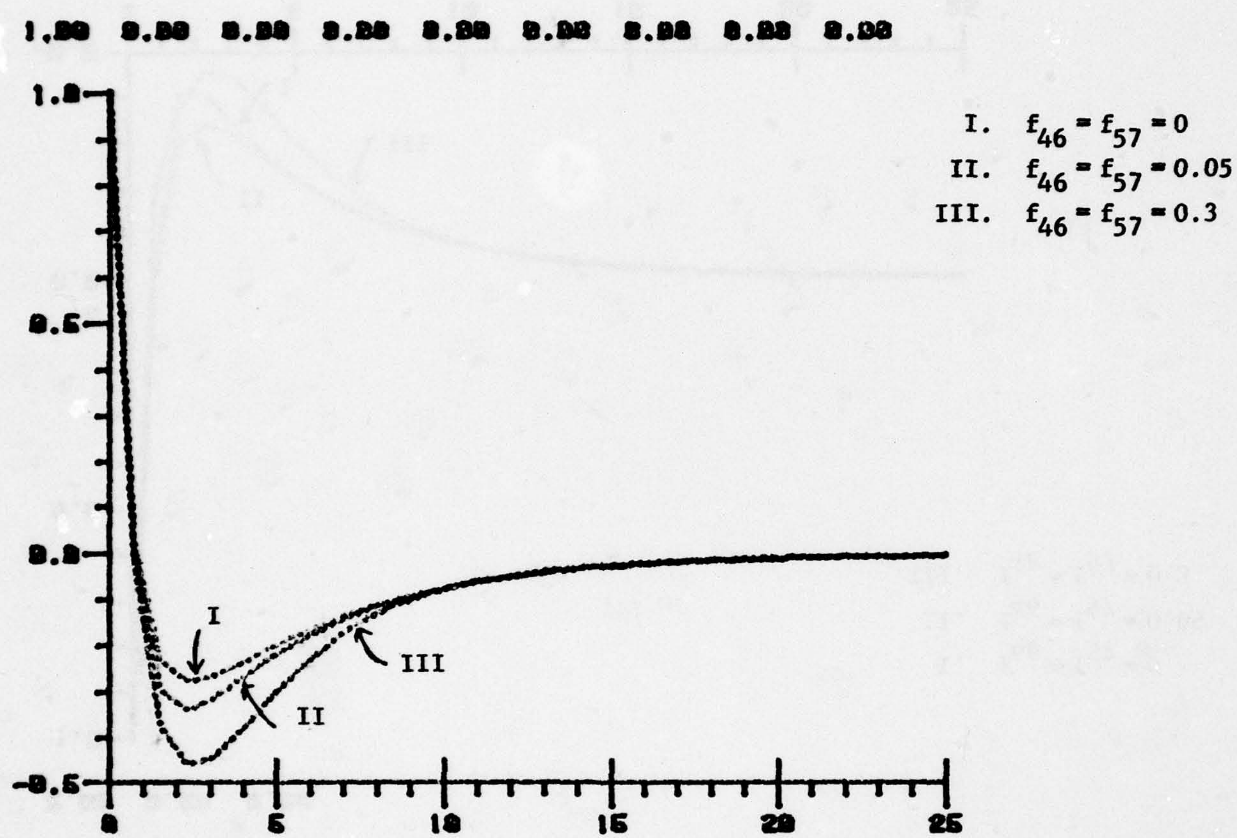


Figure 6.8a. Sensitivities of power output y_1 under perturbation of f_{46} and f_{57} in GHR system matrix, $\omega = 0$.

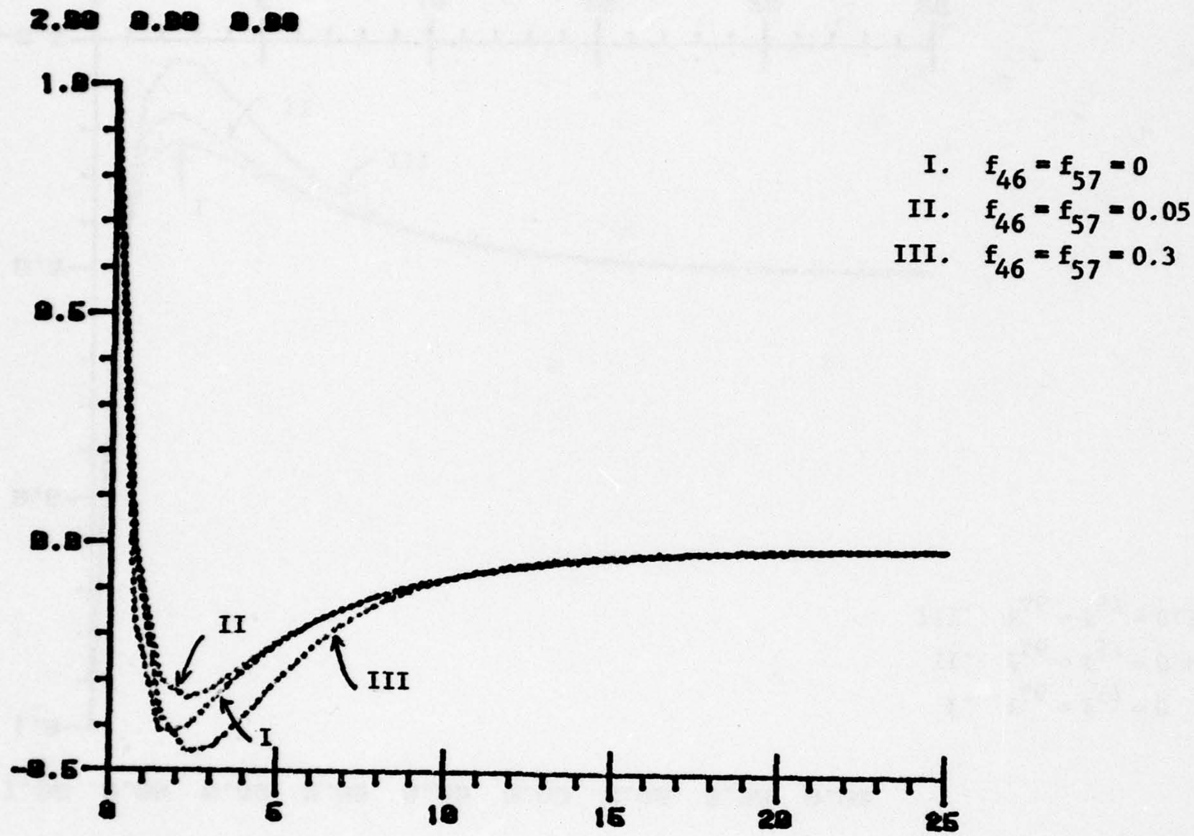


Figure 6.8b. Sensitivities of power output y_2 under perturbation of f_{46} and f_{57} in GHR system matrix, $\omega = 0$.

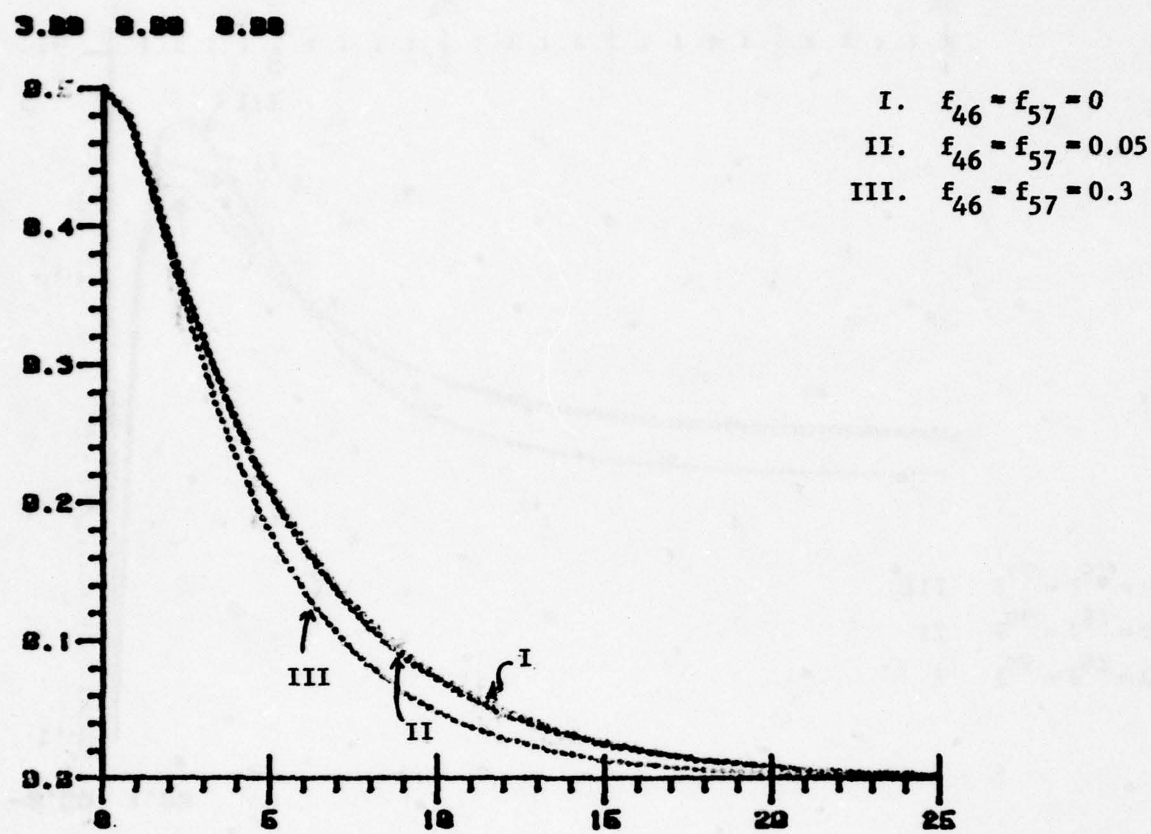


Figure 6.8c. Sensitivities of system frequency y_3 under perturbation of f_{46} and f_{57} in GHR system matrix, $\omega = 0$.

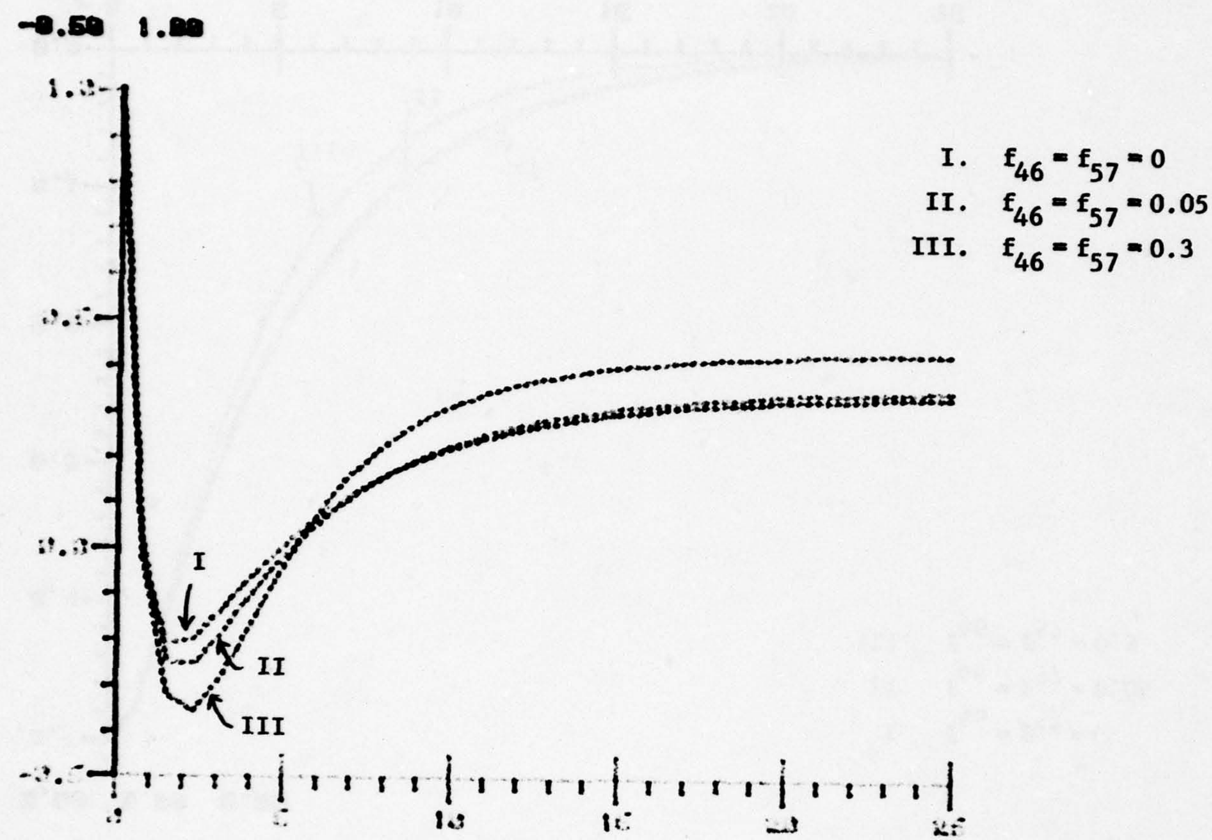


Figure 6.9a. Sensitivities of power output y_1 under perturbation of f_{46} and f_{57} in GHR system matrix, ω = unit step.

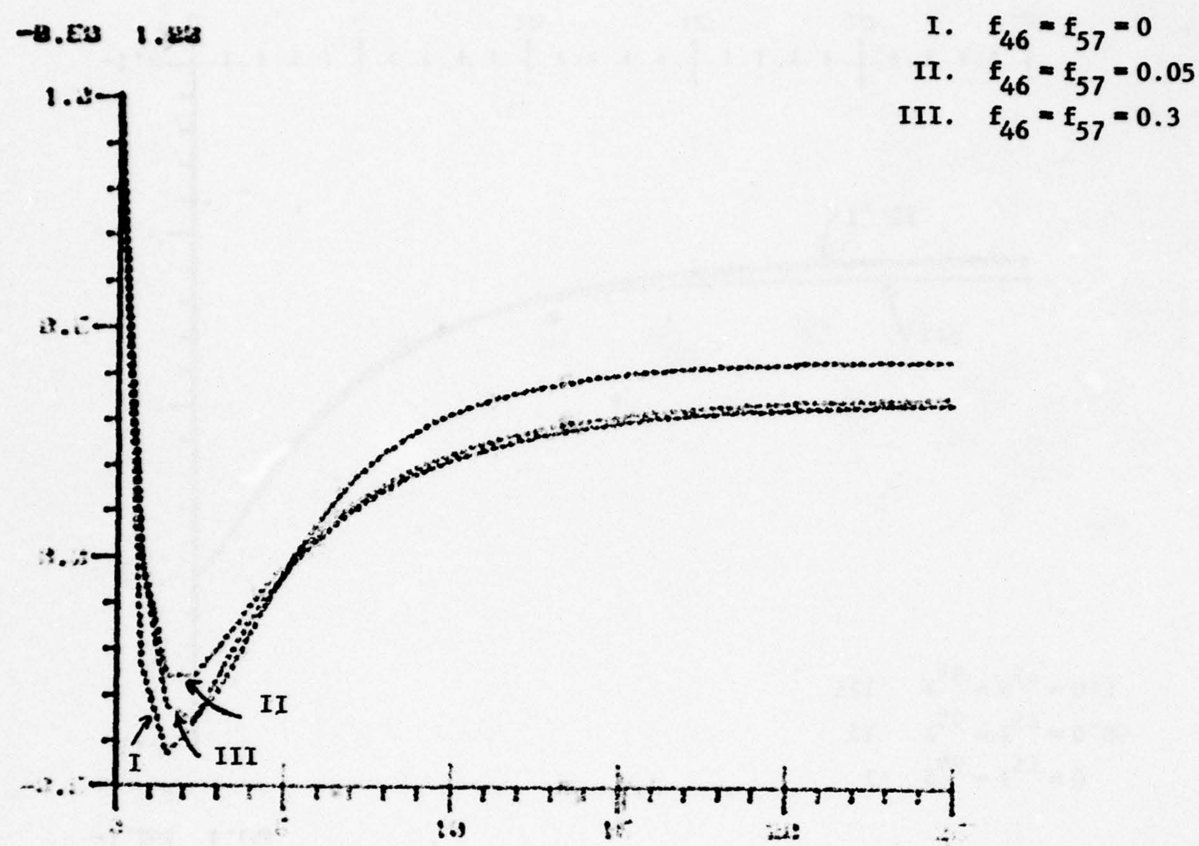


Figure 6.9b. Sensitivities of power output y_2 under perturbation of f_{46} and f_{57} in GHR system matrix, ω = unit step.

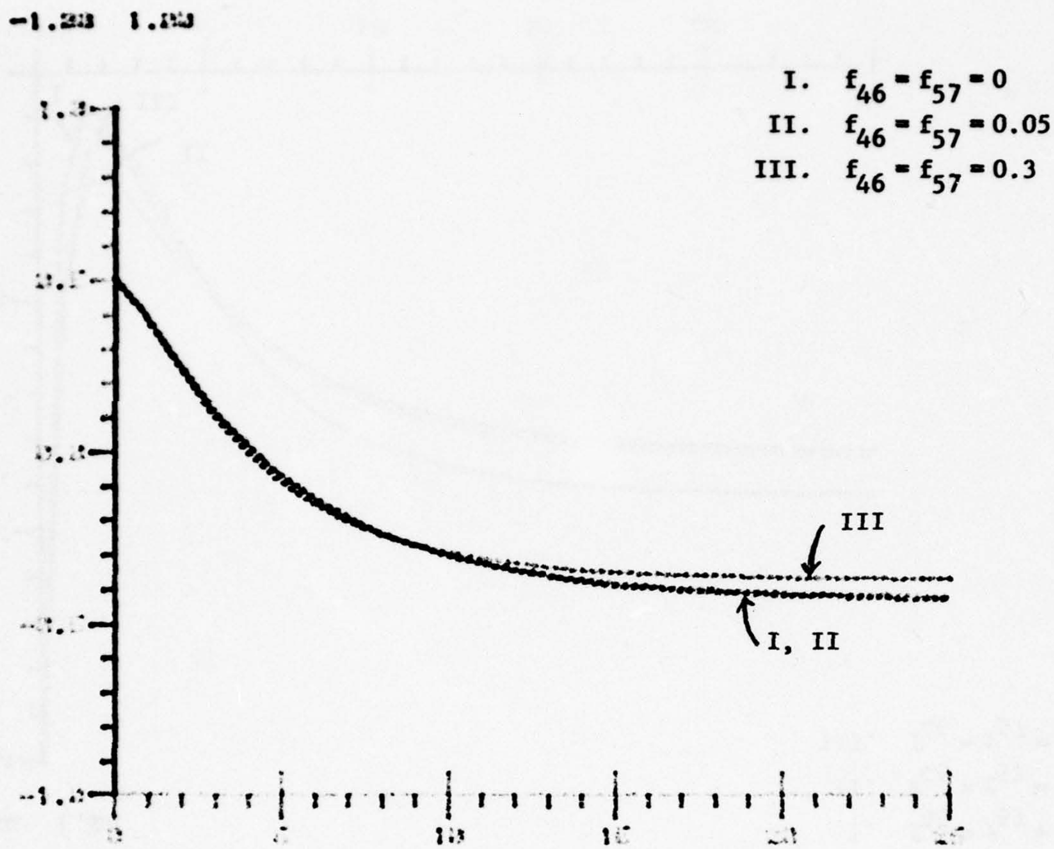


Figure 6.9c. Sensitivities of system frequency y_3 under perturbation of f_{46} and f_{57} in GHR system matrix, ω = unit step.

$$= \begin{bmatrix} 0 & 0.00004 & -0.00002 & 0 & 0 & 0 & -.000005 & 0 & 0 \\ 0 & 0 & -.000005 & 0 & 0 & .000004 & -.00002 & 0 & 0 \\ 0 & 0.000001 & -.00003 & .000001 & 0 & .000001 & -.00003 & .000001 & .000001 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

which for practical purpose can be considered zero. Hence the output following conditions are satisfied, and local feedback u_{ic} exists as confirmed by (6.24).

6.4 Decomposition in Decentralized Control Synthesis for a Two Area Interconnected Power System.

This example will illustrate the full capabilities of the control synthesis considered in Chapter 5. A two area power system with each area containing two thermal plants is constructed from [59]. The system is model by

$$\dot{x} = Ax + Bu + E\omega \tag{6.29a}$$

$$y = Cx \tag{6.29b}$$

where $x \in R^{19}$, $u \in R^4$, $\omega \in R^2$ and $y \in R^3$. The state, control and outputs have the following physical meanings:

x_1, x_{12} = valve position displacement in first thermal unit of area 1 and 2.

x_2, x_{13} = power output displacement of HP turbine in first thermal unit of area 1 and 2.

x_3, x_{14} = power output displacement of IP turbine in first thermal unit of area 1 and 2.

x_4, x_{15} = power output displacement of LP turbine in first thermal unit of area 1 and area 2.

x_5, x_{16} = valve position displacement in second thermal unit of area 1 and 2.

x_6, x_{17} = power output displacement of HP turbine in second thermal unit of area 1 and 2.

x_7, x_{18} = power displacement of IP turbine in second thermal unit of area 1 and 2.

x_8, x_{19} = power displacement of LP turbine in second thermal unit of area 1 and 2.

x_9, x_{11} = frequency deviation of area 1 and 2.

x_{10} = tie-line power flow connecting area 1 and 2.

u_1, u_3 = set point adjustment of first thermal unit in area 1 and 2.

u_2, u_4 = set point adjustement of second thermal unit in area 1 and 2.

ω_1, ω_2 = load disturbance of area 1 and 2.

y_1 = frequency deviation of area 1.

y_2 = tie-line power flow of area 1 and 2.

y_3 = frequency deviation of area 2.

The system matrices (A,B,E,C) are given in Table 6.10. Apply first the chained aggregation procedure of Chapter 2 to bring (6.29) into the GHR

$$\dot{z} = Fz + Gu + W\omega \qquad (6.30a)$$

$$y = Dz \qquad (6.30b)$$

where the GHR system matrices (F,G,W,D) are given in Table 6.11 while the GHR transformation matrix H is given in Table 6.12. The GHR structure indicates 6 subsystem (GHR subsystems) of dimension 3,2,2,2,2,8 after five steps of chained aggregations, i.e.

Table 6.10. System matrices (A,B,E,C) of Example 6.4

$$
A = \left[\begin{array}{ccccccccc:c:ccccccccc}
-0.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 & & & & & & & & & \\
4.75 & -5.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & & & & & & & & \\
0 & .1667 & -.1667 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & & & & & & & & \\
0 & 0 & 2.0 & -2.0 & 0 & 0 & 0 & 0 & 0 & 0 & & & & 0 & & & & & \\
0 & 0 & 0 & 0 & -0.2 & 0 & 0 & 0 & -4 & 0 & & & & & & & & & \\
0 & 0 & 0 & 0 & 4.75 & -5 & 0 & 0 & 0 & 0 & & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & .1667 & -.1667 & 0 & 0 & 0 & & & & & & & & & \\
0 & 0 & 0 & 0 & 0 & 0 & 2.0 & -2.0 & 0 & 0 & & & & & & & & & \\
0 & .0125 & .0117 & .0175 & 0 & .0125 & .0117 & .0175 & -.1125 & -0.0833 & & & & & & & & & \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 22.2144 & 0 & -22.2144 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
 & & & & & & & & & 0.0833 & -.1125 & 0 & .0125 & .0117 & .0175 & 0 & .0125 & .0117 & .0175 \\
 & & & & & & & & & 0 & -4 & -.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 0 & 0 & 4.75 & -5.0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 0 & 0 & 0 & .1667 & -.1667 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 0 & 0 & 0 & 0 & 2 & -2.0 & 0 & 0 & 0 & 0 \\
 & & & & 0 & & & & & 0 & -4 & 0 & 0 & 0 & 0 & -0.2 & 0 & 0 & 0 \\
 & & & & & & & & & 0 & 0 & 0 & 0 & 0 & 0 & 4.75 & -5.0 & 0 & 0 \\
 & & & & & & & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .1667 & -.1667 & 0 \\
 & & & & & & & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2.0 & -2.0
\end{array}\right]
$$

$$
C = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

Table 6.10. continued

$$
B = \left[\begin{array}{cc:cc}
4.0 & 0 & & \\
0 & 0 & & \\
0 & 0 & & \\
0 & 0 & & \\
0 & 4.0 & 0 & \\
0 & 0 & & \\
0 & 0 & & \\
0 & 0 & & \\
0 & 0 & & \\
0 & 0 & 0 & 0 \\
\hdashline
 & & 0 & 0 \\
\hdashline
 & & 4.0 & 0 \\
 & & 0 & 0 \\
 & & 0 & 0 \\
 & 0 & 0 & 0 \\
 & & 0 & 4.0 \\
 & & 0 & 0 \\
 & & 0 & 0 \\
 & & 0 & 0
\end{array}\right]
\qquad
E = \begin{bmatrix}
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
-0.0833 & 0 \\
0 & 0 \\
0 & -0.0833 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0
\end{bmatrix}
$$

Table 6.11. GHR representation (F,G,W,D) of Example 6.4

$$F = \begin{bmatrix}
0 & 22.21 & -22.21 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.0833 & -.1125 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.0833 & 0 & -.1125 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -4.844 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -4.844 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -.475 & 0 & 1.197 & 0 & -.2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -.475 & 0 & 1.197 & 0 & -.2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1.181 & 0 & 0 & 0 & -2.82 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1.181 & 0 & 0 & 0 & -2.82 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -2.514 & 0 & .497 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2.514 & 0 & .497 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2.0 & 2.0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -.1667 & .1667 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.75 & 0 & 0 & -5.0 & 0 & 0 & 0 & 0 \\
0 & 0 & -4.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.75 & -5.0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .1667 & -.1667 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2.0 & -2.0
\end{bmatrix}$$

Table 6.11. continued

$$
G = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
.2375 & .2375 & 0 & 0 \\
0 & 0 & .2375 & .2375 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 4 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 4 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\qquad
W = \begin{bmatrix}
0 & 0 \\
-0.0833 & 0 \\
0 & -0.0833 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0
\end{bmatrix}
$$

$$
D = \left[\begin{array}{ccc:c}
0 & 1 & C & \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 &
\end{array}\right]
$$

Table 6.12. GHR transformation matrix H of Example 6.4

$$
H = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & .0125 & .0117 & .0175 & 0 & .0125 & .0117 & .0175 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .0125 & .0117 & .0175 & 0 & .0125 & .0117 & .0175 \\
.05938 & 0 & .08972 & .04977 & .05938 & 0 & .08972 & .04977 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .05938 & 0 & .08972 & .04977 & .05938 & 0 & .08972 & .04977 \\
0 & 0 & .08853 & -.1105 & 0 & 0 & .08853 & -.1105 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .08853 & -.1105 & 0 & 0 & .08853 & -.1105 \\
0 & 0 & 0 & -.1113 & 0 & 0 & 0 & -.1113 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -.1113 & 0 & 0 & 0 & -.1113 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

$$F = \begin{bmatrix} F_{11} & F_{12} & 0 & 0 & 0 & 0 \\ F_{21} & F_{22} & F_{23} & 0 & 0 & 0 \\ F_{31} & F_{32} & F_{33} & F_{34} & 0 & 0 \\ F_{41} & F_{42} & F_{43} & F_{44} & F_{45} & 0 \\ F_{51} & F_{52} & F_{53} & F_{54} & F_{55} & 0 \\ F_{61} & F_{62} & F_{63} & F_{64} & F_{65} & F_{66} \end{bmatrix} \quad G = \begin{bmatrix} 0 \\ 0 \\ G_3 \\ 0 \\ 0 \\ G_6 \end{bmatrix} \quad W = \begin{bmatrix} W_1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{6.31}$$

From the structure of F and G in (6.31), it is clear that one possibility to achieve aggreability is to use feedback so that $F_{3,4}z_4$ is fully compensated. Thus a 7th order aggregated model of the system is desired.

$$\dot{z}_r = \begin{bmatrix} F_{11} & F_{12} & 0 \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} z_r + \begin{bmatrix} 0 \\ 0 \\ G_3 \end{bmatrix} u + \begin{bmatrix} W_1 \\ 0 \\ 0 \end{bmatrix} \omega, \quad z_r \in R^7 \tag{6.32}$$

$$\hat{y} = [I_1 \quad 0 \quad 0] z_r, \quad \hat{y} \in R^3$$

where the system matrices of (6.32) (F^r, G^r, W^r, D^r) are given by Table 6.13.

Recall $u_i = u_{ia} + u_{ib} + u_{ic}$ where u_{ia} is used to control the aggregate (z^3 in this case), u_{ib} is used to achieve aggregation (compensation of $F_{3,4}z_4$) and u_{ic} is used to control the residual ($\bar{z}^3$ in this example). Let us initially consider the design of u_{ib} in GHR. Denote

$$u_b = \begin{bmatrix} u_{1b} \\ \\ u_{2b} \end{bmatrix} = L_b z \tag{6.33}$$

Note that due to the structure of G, L_b can have a general form:

Table 6.13 System matrices (F_r, G_r, W_r, D_r) of the reduced order aggregated model (6.32).

$$F_r = \left[\begin{array}{ccc|cc|cc}
0 & 22.21 & -22.21 & 0 & 0 & 0 & 0 \\
-.0833 & -.1125 & 0 & 1 & 0 & 0 & 0 \\
.0833 & 0 & -.1125 & 0 & 1 & 0 & 0 \\
\hline
0 & 0 & 0 & -4.844 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & -4.844 & 0 & 1 \\
\hline
0 & -.475 & 0 & 1.197 & 0 & -0.2 & 0 \\
0 & 0 & -.475 & 0 & 1.197 & 0 & -0.2
\end{array}\right]$$

$$G_r = \left[\begin{array}{cccc}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\hline
.2375 & .2375 & 0 & 0 \\
0 & 0 & .2375 & .2375
\end{array}\right]
\qquad
W_r = \left[\begin{array}{cc}
0 & 0 \\
-0.0833 & 0 \\
0 & -0.0833 \\
\hline
0 & 0 \\
0 & 0 \\
\hline
0 & 0 \\
0 & 0
\end{array}\right]$$

$$D_r = \left[\begin{array}{ccc|cc|cc}
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0
\end{array}\right]$$

$$L_b = \left[\begin{array}{c:cc:c} & a & \gamma & \\ & b & -\gamma & \\ 0_{4x7} & & & 0_{4x10} \\ & c & \alpha & \\ & -c & \beta & \end{array}\right] \tag{6.34}$$

where $a,b,c,\alpha,\beta,\gamma$ are arbitrary parameters as long as:

$$.2375(a+b) = -1 \tag{6.35}$$

$$.2375(\alpha+\beta) = -1 \tag{6.36}$$

One possible L_b can be

$$L_b = \begin{bmatrix} & -2.105263 & 0 & \\ & -2.105263 & 0 & \\ 0_{4x7} & & & 0_{4x10} \\ & 0 & -2.105263 & \\ & 0 & -2.105263 & \end{bmatrix} \tag{6.37}$$

Expressed in terms of original system variables, u_b is

$$u_b = L_b z = L_b Ex = L_{bo} x \tag{6.38}$$

where L_{bo} is given in Table 6.14. Again we see a local feedback structure exists for u_b in order to achieve aggregability. With the application of u_b given by (6.38), (6.32) is now an exact reduced order model of the power system (6.29).

Next consider the design of u_a which controls the aggregate. For simplicity, assume $\omega = o$. An optimal regulator design is used to control the aggregate (6.32). The quadratic cost

$$J = \frac{1}{2}\int_o^\infty z_r^T Q\, z_r + u_a^T R\, u_a \, dt \tag{6.39}$$

is optimized with $Q = 9I$, $R = I$. The result is:

Table 6.14. Feedback gain of u_b in original system representation for Example 6.4

$$L_{bo} = \left[\begin{array}{ccccccccc|c|cccccccccc} 0 & 0 & -.186378 & .23263 & 0 & 0 & -.186378 & .23263 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -.186378 & .23263 & 0 & 0 & -.186378 & .23263 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -.186378 & .23263 & 0 & 0 & -.186378 & & .23263 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -.186378 & .23263 & 0 & 0 & -.186378 & & .23263 \end{array}\right]$$

$$u_a = \begin{bmatrix} .87991 & -11.066 & 10.714 & -3.5577 & 2.9453 & -3.7568 & 1.4842 \\ .87991 & -11.066 & 10.714 & -3.5577 & 2.9453 & -3.7568 & 1.4842 \\ -.87991 & 10.714 & -11.066 & 2.9453 & -3.5577 & 1.4842 & -3.7568 \\ -.87991 & 10.714 & -11.066 & 2.9453 & -3.5577 & 1.4842 & -3.7568 \end{bmatrix} z_r$$

$$= L_a' z_r \qquad (6.40)$$

Hence the component u_a to be applied to (6.30) is

$$u_a = [L_a' \quad 0] \begin{bmatrix} z_r \\ \bar{z}_r \end{bmatrix} = L_a z \qquad (6.41)$$

We pause to stress the importance of component u_a in this example. It is interesting to note that the open loop system of (6.30) ($u \equiv o$) is stable. On the other hand, if u_b of (6.38) is applied, the resulting system is unstable. However, with the application component u_a, the new closed loop system is again stable. The eigenvalues are summarized in Table 6.15. Adopt the notation:

$$\begin{aligned} \text{I} &: u_a = u_b = u_c = 0 \\ \text{II} &: u_a = 0, \quad u_b \neq 0, \quad u_c = 0 \\ \text{III} &: u_a \neq 0, \quad u_b \neq 0, \quad u_c = 0 \end{aligned} \qquad (6.42)$$

The output trajectory behavior (frequency deviations of areas 1 and 2 and tie-line power flow) are plotted in Fig. 6.10 for a initial condition of $z_o = (0,0.5,0,...0)^T$ which corresponds to a frequency variation in area 1.

Similar to Example 6.3, the sensentivity of output trajectories with respect to the component u_b is studied via simulations. Again the perturbation of u_b are expressed in terms of small perturbation of f_{68}

Table 6.15. Eigenvalue comparison when u_a and u_b are applied

$u_a = u_b = u_c = 0$	$u_a = 0, u_b \neq 0, u_c = 0$	$u_a \neq 0, u_b \neq 0, u_c = 0$
-2.0	-2.0	-2.0
-0.1667	-0.1667	-0.1667
-2.0	-2.0	-2.0
-0.1667	-0.1667	-0.1667
-5.0	-5.0	-5.0
-0.2	-0.2	-0.2
-5.02	$-0.04764 \pm j1.944$	$-0.541 \pm j2.152$
-5.017	-5.107	-1.457
$-0.04723 \pm j1.943$	-5.105	-1.648
-1.994	$-0.02464 \pm j0.2956$	-0.831
-1.988	+0.04367	-0.675
$-0.01946 \pm j0.3618$	-5.0	-5.107
-0.4318	-1.648	-5.114
-0.14	-1.648	-0.2912
-0.2332	-0.675	-1.648
-5.0	-0.2	-5.0
-0.2	-0.675	-0.675
		-0.2

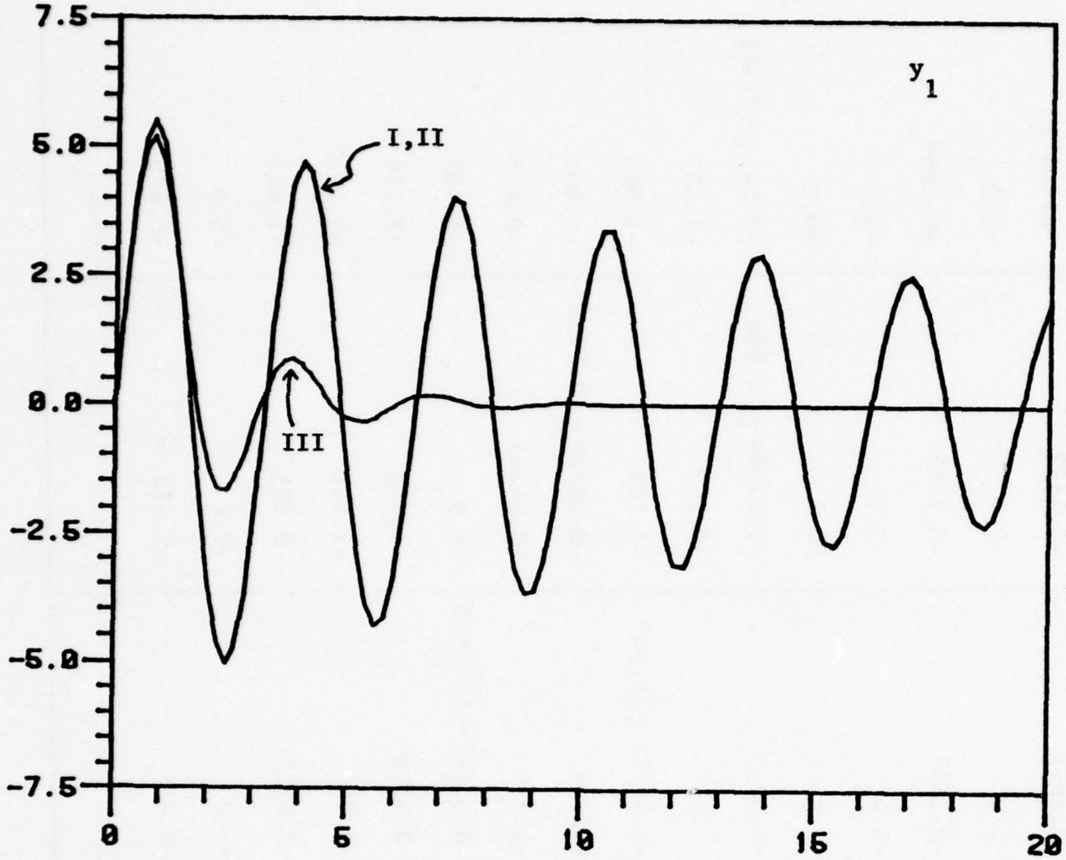


Figure 6.10. Output trajectories (y_1, y_2, y_3) under (I) $u_a = u_b = u_c = 0$ (II) $u_a = 0$, $u_b \neq 0$, $u_c = 0$, (III) $u_a \neq 0$, $u_b \neq 0$, $u_c = 0$.

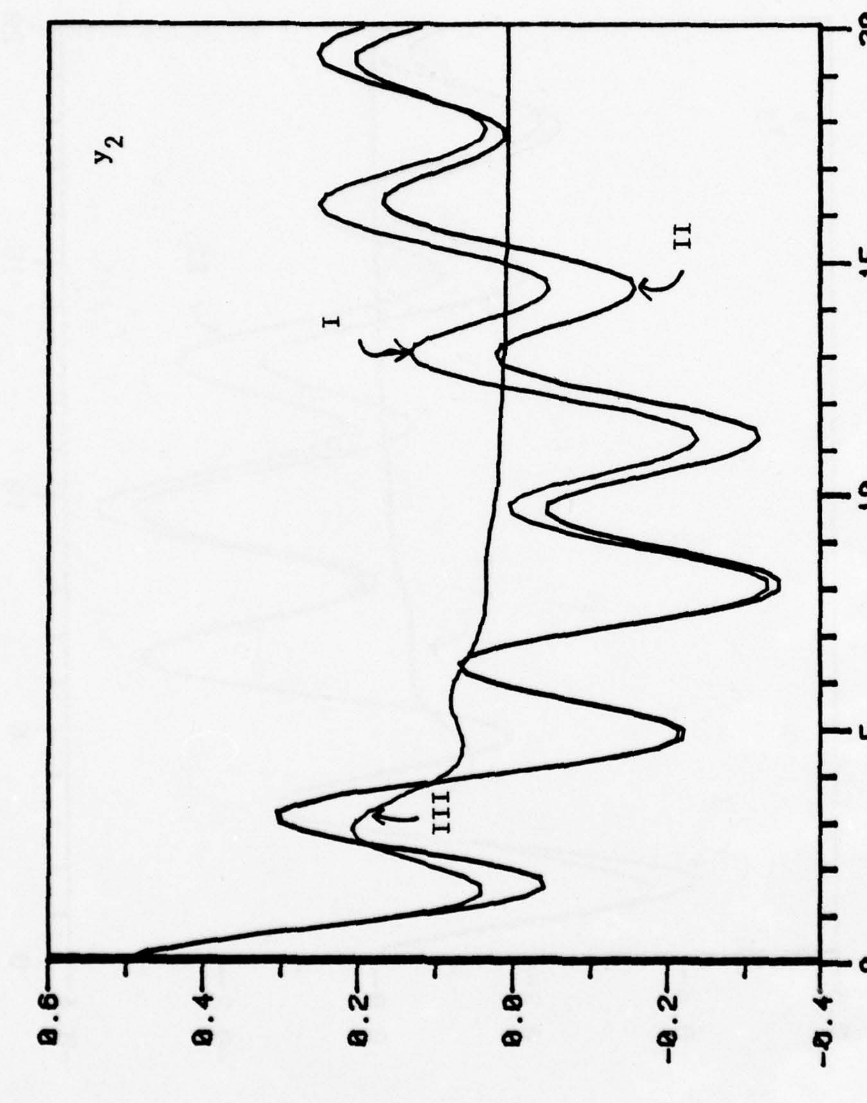


Figure 6.10. continued

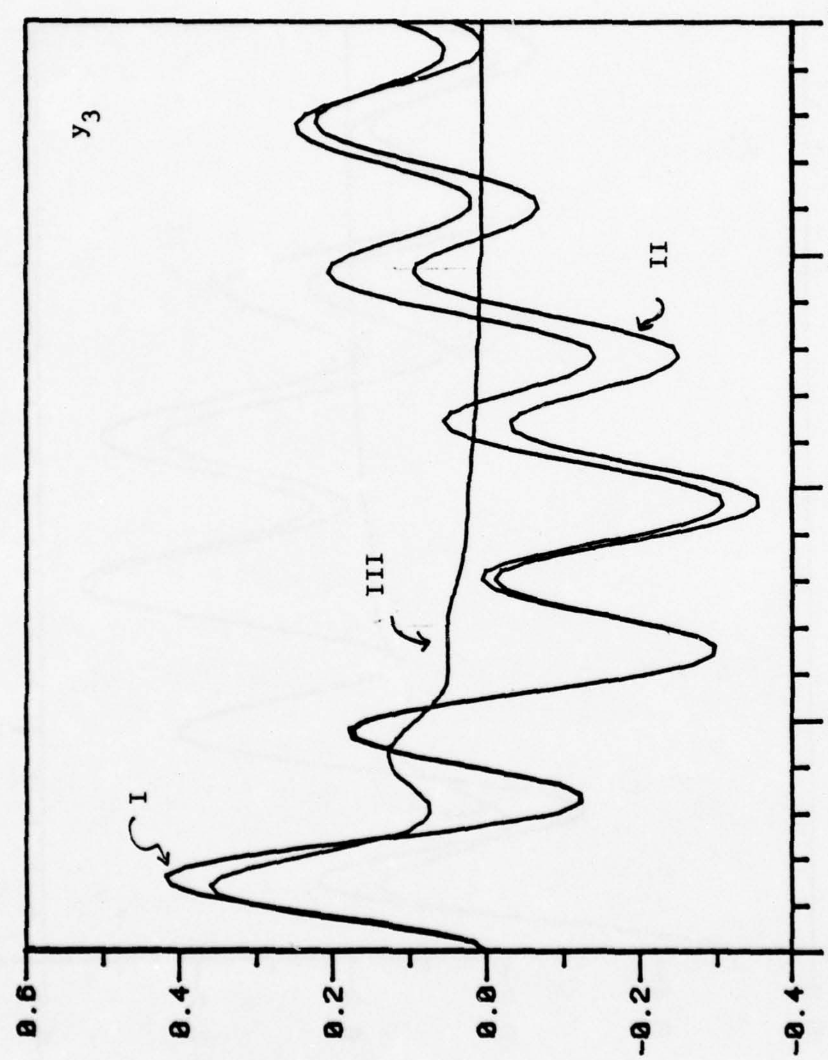


Figure 6.10. continued

and f_{79} of the system matrix F in the GHR. Simulation results for $u_a = 0$, $u_b \neq 0$, $u_c = 0$ are plotted in Fig. 6.11 where (I) represents $f_{68} = f_{79} = 1.0$, (II) represents $f_{68} = f_{79} = 1.5$ and (III) represents $f_{68} = f_{79} = 1.2$. Those for $u_a \neq 0$, $u_b \neq 0$ and $u_c = 0$ are plotted in Fig. 6.12 where (I) represents $f_{68} = f_{79} = 1$, (II) represents $f_{68} = f_{79} = 0.5$ (III) represents $f_{68} = f_{79} = -1.0$, (IV) represents $f_{68} = f_{79} = 0$ and (V) represents $f_{68} = f_{79} = 2.5$. Even when a fairly large perturbation in L_b is introduced, the sensitivity of the output trajectory with respect to u_b is not severe.

Finally, let us illustrate the existence and utilization of the freedom component u_c as discussed in Chapter 5. For this example, $m = 4 > r = 3$. Thus freedom should exist, resulting a non-interactive control of the residual subsystem in the GHR. From (6.31), we see it is sufficient to restrict u_c to lie in the null space of G_3 where

$$G_3 = \begin{bmatrix} 0.2375 & 0.2375 & 0 & 0 \\ 0 & 0 & 0.2375 & 0.2375 \end{bmatrix} \tag{6.43}$$

Thus u_c takes on the general form:

$$u_c = [0 \quad 0 \quad 0 \ \vdots \ (I_4 - G_3^{+}G_3)K] \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \cdots \\ z_4 \\ z_5 \\ z_6 \end{bmatrix} \equiv L_c z \tag{6.44}$$

Where K is an arbitrary matrix and z_i are the substates of the GHR partitioned in accordance with (6.31). Introduce the notations

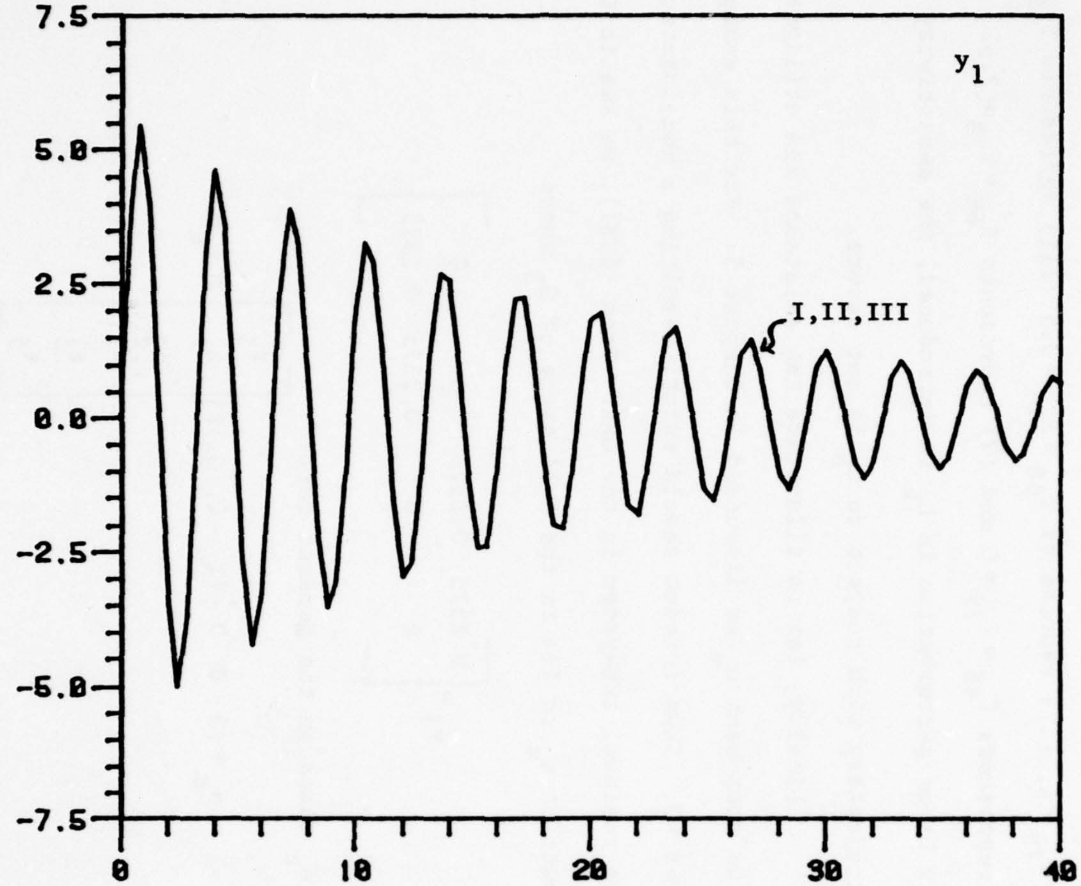


Figure 6.11. Sensitivities of output trajectory (y_1, y_2, y_3) with respect to perturbation of f_{68}, f_{79} in system matrix F of GHR, $\omega = 0$, $u_a = u_c = 0$, $u_b \neq 0$.

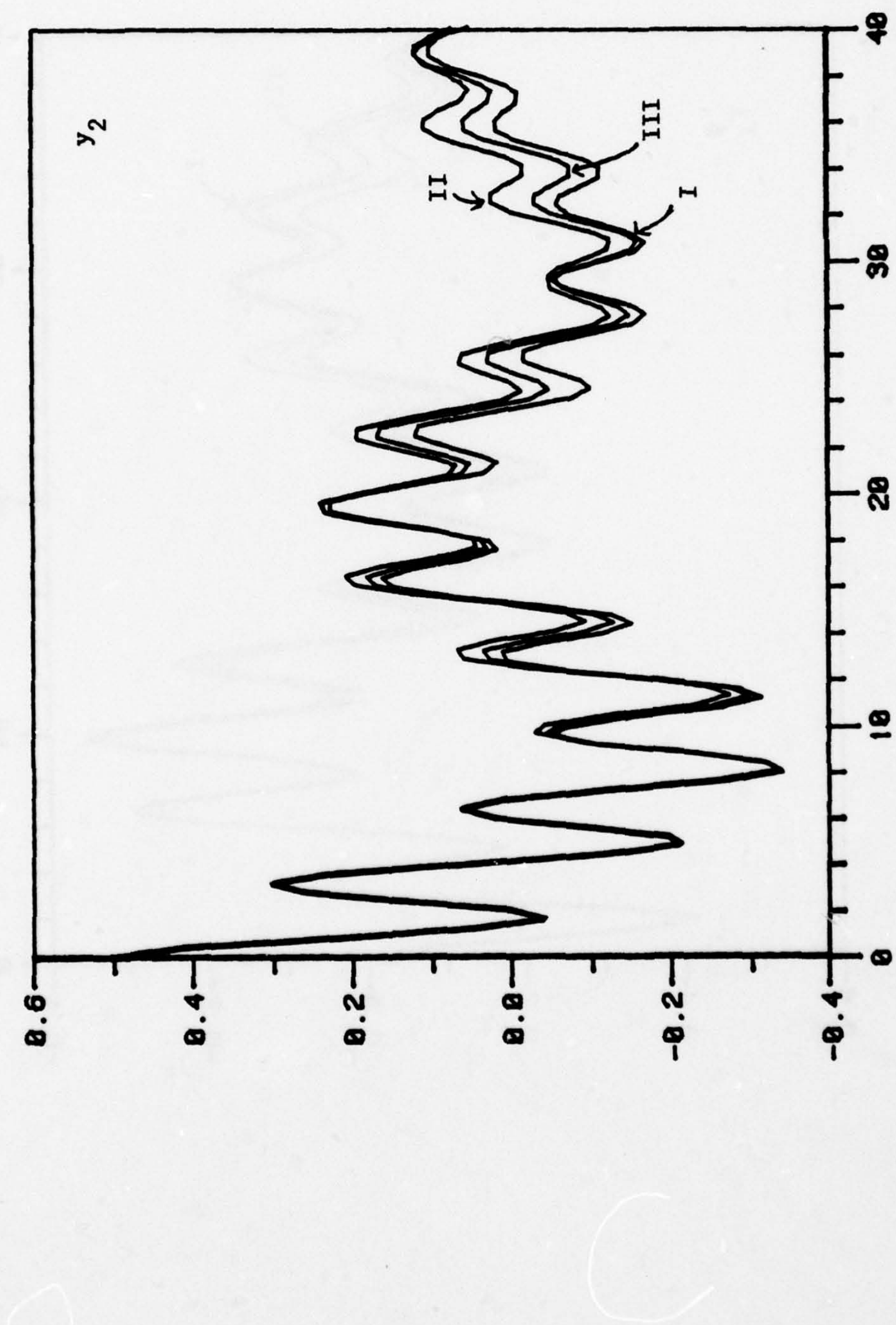


Figure 6.11. continued

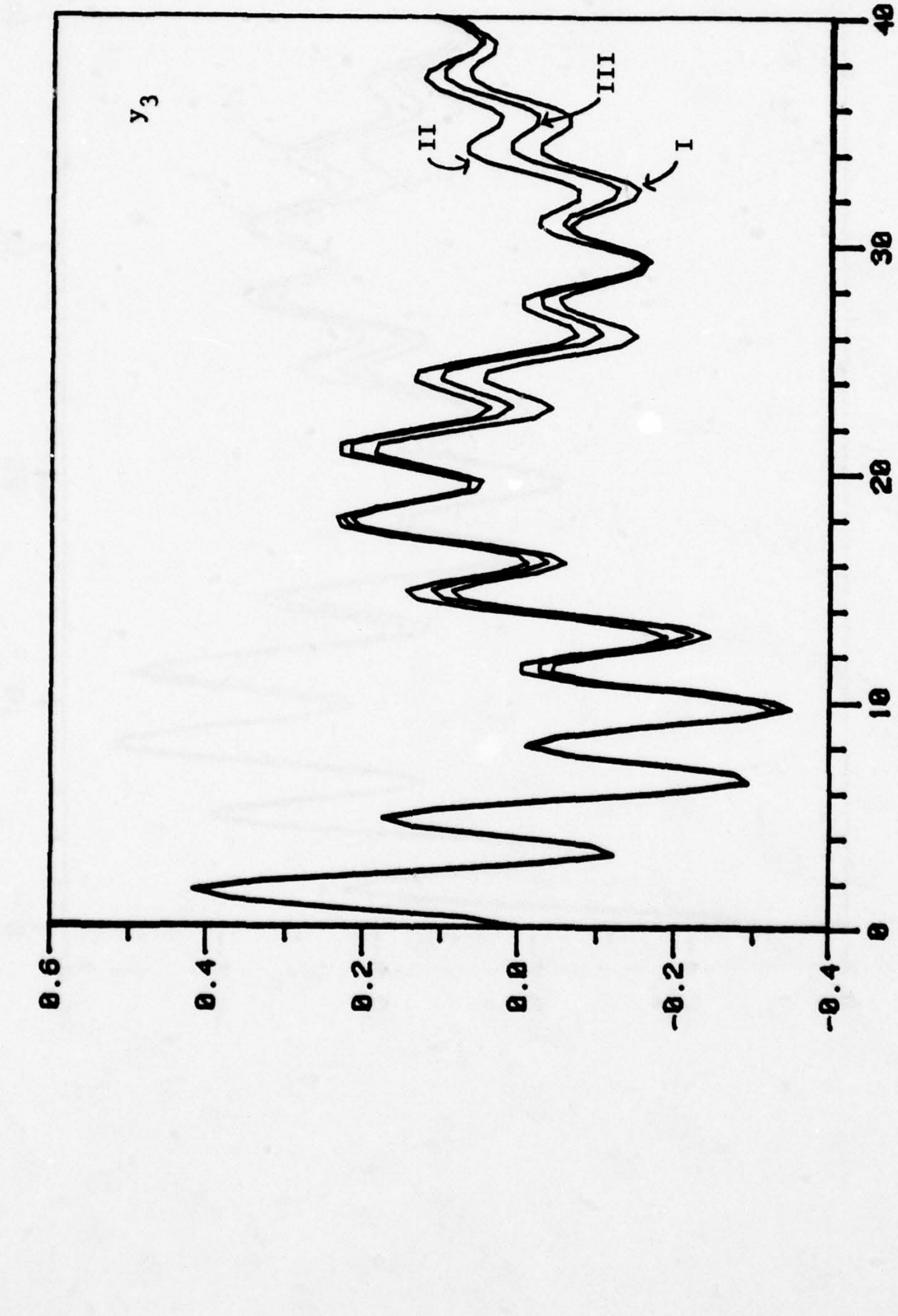


Figure 6.11. continued

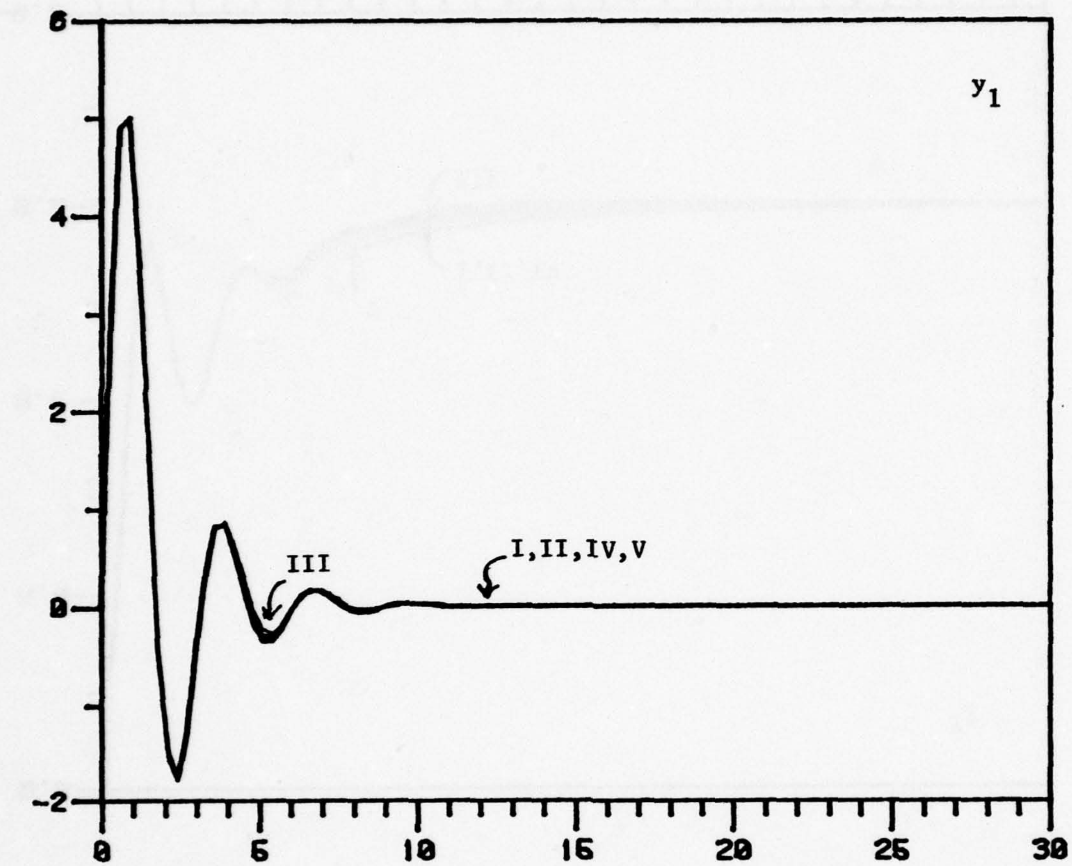


Figure 6.12. Sensitivity output trajectory (y_1,y_2,y_3) with respect to perturbation of f_{68} and f_{79} in system matrix F of GHR, $\omega = 0$, $u_a \neq 0$, $u_b \neq 0$, $u_c = 0$.

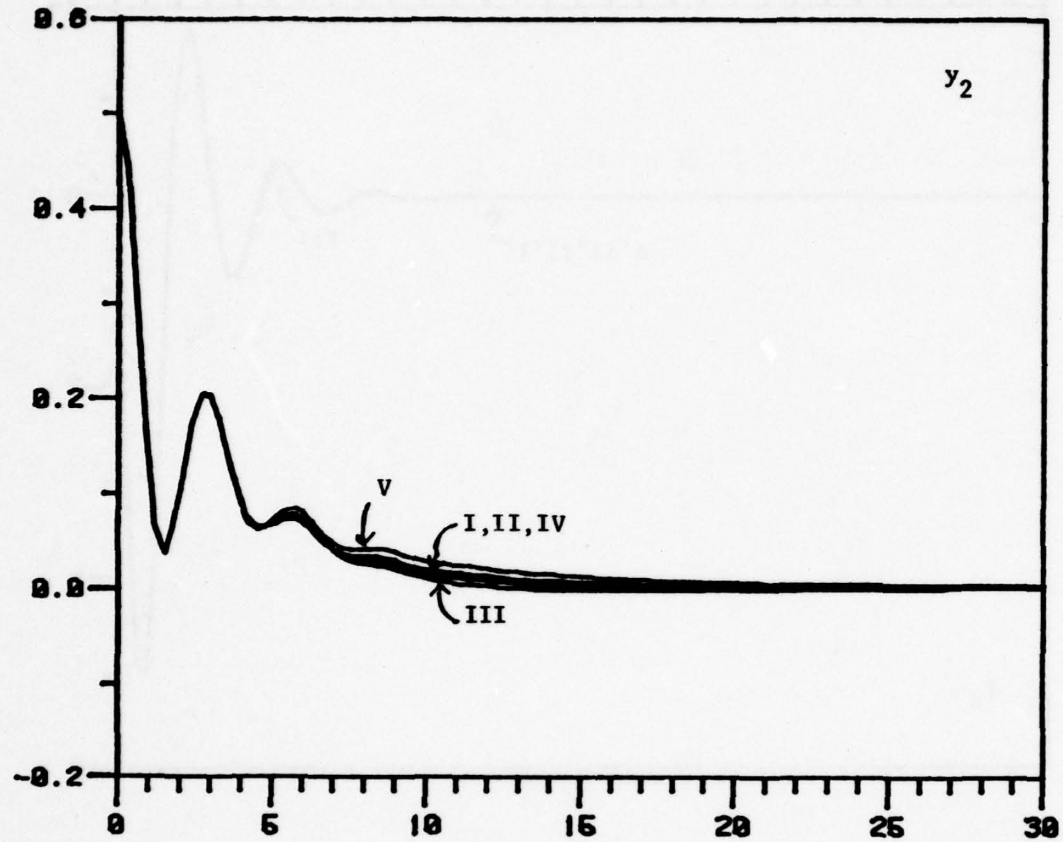


Figure 6.12. continued

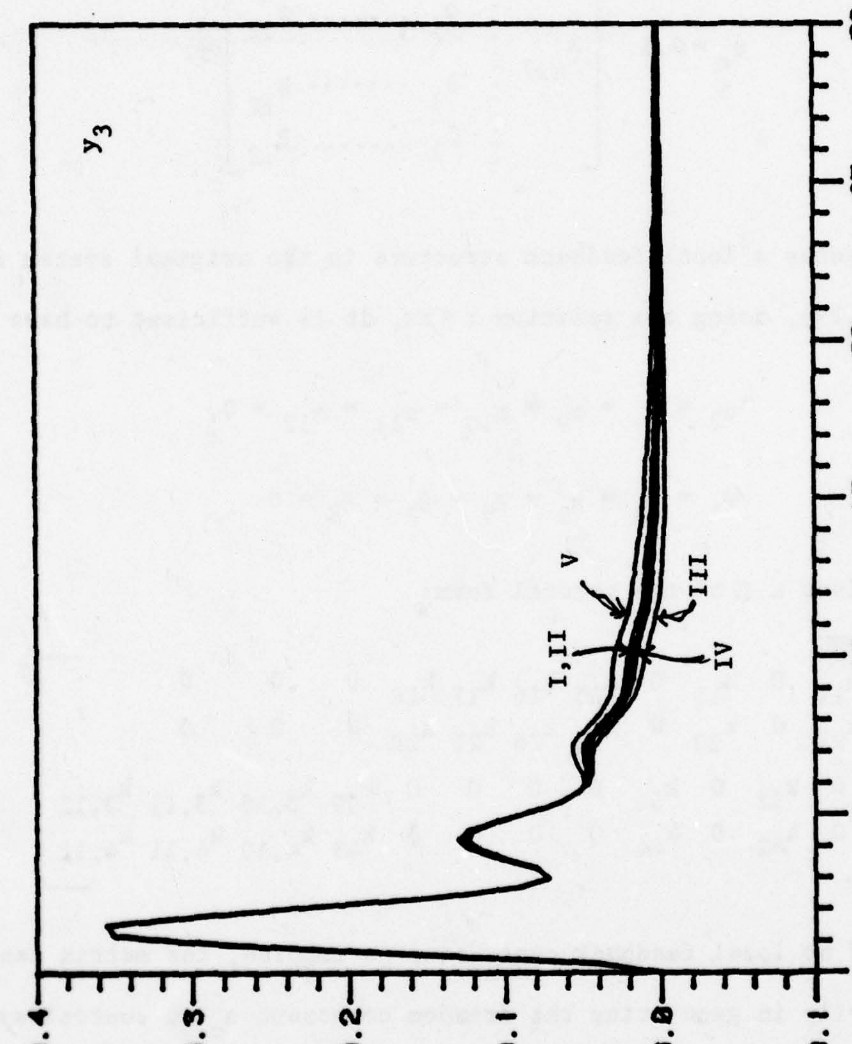


Figure 6.12. continued

$$K = [k_{ij}] \quad , \quad k_{ij} \in R$$

$$\alpha_j = k_{1j} - k_{2j} \,, \quad i = 1, 2, 3, 4 \tag{6.45}$$

$$\beta_j = k_{3j} - k_{4j} \,, \quad j = 1, 2, \ldots 12$$

Then it can be shown that

$$u_c = 0.5 \begin{bmatrix} & \vdots & \alpha_1 \cdots\cdots \alpha_{12} \\ & \vdots & \alpha_1 \cdots\cdots \alpha_{12} \\ 0_{4x7} & \vdots & \\ & \vdots & \beta_1 \cdots\cdots \beta_{12} \\ & \vdots & \beta_1 \cdots\cdots \beta_{12} \end{bmatrix} z. \tag{6.46}$$

To guarantee a local feedback structure in the original system representation (6.29), using the relation $z = Hx$, it is sufficient to have

$$\alpha_2 = \alpha_4 = \alpha_9 = \alpha_{10} = \alpha_{11} = \alpha_{12} = 0$$
$$\beta_1 = \beta_3 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = 0 \tag{6.47}$$

which gives a K of the general form:

$$K = \begin{bmatrix} k_{11} & 0 & k_{13} & 0 & k_{15} & k_{16} & k_{17} & k_{18} & 0 & 0 & 0 & \\ k_{21} & 0 & k_{23} & 0 & k_{25} & k_{26} & k_{27} & k_{28} & 0 & 0 & 0 & \\ 0 & k_{32} & 0 & k_{34} & 0 & 0 & 0 & 0 & k_{39} & k_{3,10} & k_{3,11} & k_{3,12} \\ 0 & k_{42} & 0 & k_{44} & 0 & 0 & 0 & 0 & k_{49} & k_{4,10} & k_{4,11} & k_{4,12} \end{bmatrix} \tag{6.48}$$

Hence if no local feedback constraint is imposed, the matrix can be choosen arbitrarily in generating the freedom component u_c in control synthesis. However, if local feedback (i.e. area i control u_i uses only state variables from area i) is imposed, K must be chosen according to (6.48). To show the eigenvalue shifting capability of u_c, the closed loop eigenvalues under $u_a = L_a z$, $u_b = L_b z$, $u_c = 0$ and $u_a = L_a z$, $u_b = L_b z$, $u_c = L_c z$ are obtained where

L_c is the feedback gain correspond to $k_{11} = k_{32} = -2$, $k_{21} = k_{42} = 2$, $k_{13} = k_{34} = -5$, $k_{23} = k_{44} = 5$. These are given in Table 6.16.

In summary, we note the control u_i, $i = 1,2$ consists of three components. The first component u_{ia} controls the aggregate (6.32). In terms of actual physical system variables, u_{ia} may be designed and implemented either as a centralized control or as a decentralized control. The second component, u_{ib} ((6.38)) achieves aggregation by canceling the interaction between S^3 and $\bar{S}^3$, and is seen implemented as a decentralized control (see Table 6.14). The third component u_{ic} given by (6.44) offers an freedom component that is non-iteractive with u_{ia} and u_{ib} and controls the aggregate. Again, it can be implemented either as a centralized or as a decentralized control in terms of actual power system variables. We have chosen a decentralized implementation in the example. Finally, we note that the output trajectories under u_a and u_b given in Fig. 6.10 shows considerable damping as compared to those in actual practice. This is because in forming u_a, five states (all states of the aggregate (6.32)) are utilized for feedback while in reality, only available outputs are used to generate the feedback control.

Table 6.16. Closed loop eigenvalues showing the effect of control component u_c

$u_a = L_a z$ $u_b = L_b z$, $u_c = 0$	$u_a = L_a z$, $u_b = L_b z$, $u_c = L_c z$
-2.0	-2.0
-0.1667	-0.1667
-2.0	-0.1667
-5.0	-5.0
-0.2	-0.2
$-0.541 \pm j2.152$	-5.339
-1.457	-5.30
-1.648	$-0.5355 \pm j2.08$
-0.831	$-1.726 \pm j1.362$
-0.675	$-1.594 \pm j1.23$
-5.107	$-0.03584 \pm j0.1515$
-5.114	-0.106
-0.2912	-5.0
-1.648	-0.2
-5.0	
-0.675	
-0.2	

CHAPTER 7

CONCLUSION

The construction of reduced order models of a linear time invariant system by dynamic aggregation is studied. The relationship between aggregation and observability is given. It is shown that given the pair (A,C) where A is the system matrix and C the aggregation matrix (C is assumed to be the output matrix). Three cases can occur: (A,C) completely aggregable, (A,C) aggregable and (A,C) nonaggregable. When (A,C) is completely aggregable or aggregable, (A,C) is also unobservable. When exact aggregation occurs, the aggregated model is the model of the system in the observable subspace defined by (A,C).

The extension of Aoki's aggregation method results in the chained aggregation procedure which leads to the generalized Hessenberg system representation. The properties of the GHR are studied. It is shown that the GHR structure is a property of the pair (A,C). In addition, the observability and controllability properties of subsystems in the GHR structure are examined. One important feature of the GHR is its use for the construction of approximate aggregated reduced order model of the system. In the crudest form, this corresponds to neglecting certain Hessenberg blocks in the GHR system matrix without any parameter adjustment. The effect of the Hessenberg blocks $F_{i,i+1}$ in introducing approximation error between actual system output and approximate system output is studied both from trajectory perturbation and from open-loop eigenvalue perturbation point of view. The perturbations are approximately bounded in terms of the perturbing Hessenberg block.

If none of the Hessenberg blocks can be considered weak and thus be neglected at the outset. The coupling $F_{i,i+1}z_{i+1}$ between the aggregate and the residual subsystem must be adjusted accordingly. Several methods are presented in Chapter 4. One of the major methods proposed is the GQL-RQL algorithms. The GQL theory is developed in Chapter 3 giving mainly its properties, convergence and effects on the GHR system representation. A specialization adopted for our reduced order modeling via the GHR results in the RQL algorithm. The robustness of using a selected reduced order model in designing control laws for the actual system is studied. Two approaches are used. In the first one, a Lyapunov function is introduced and then bounded in terms of the perturbing Hessenberg blocks. In the second, a frequency domain condition is obtained guaranteeing the stability of control law designed based on the reduced order model.

Finally the synthesis of a decentralized control for an interconnected system is studied. The conditions under which the system interconnection variables may track a system of reference interconnection dynamics are obtained. The solution is shown to be related to the capability of achieving aggregability by using compensatory feedback control. Moreover, the local control u_i can be decomposed into three basic components, one to control the aggregate, one to achieve aggregation of some variables containing the system interconnection variables, and one for controlling the residual if it exists. This third component is shown to be dependent on the system structure and only exists when number of inputs are greater than the number of outputs. Several examples using typical large scale systems as bench mark models are given to illustrate the methodology proposed.

In summary, it is demonstrated that "good" reduced order models that approximate the input-output behavior of a large scale system can be obtained by taking into account the given output information structure in defining a reduction procedure. The GHR approach to model reduction provides both a simple and systematic means of obtaining reduced order model structures and parameters. However, the implication of weak coupling and weak observability of the residual by the aggregate subsystem must be given future attention. In addition, the relationship between the GHR structure, the obtained reduced order model and the stability of control law based on the reduced order model in stabilizing the full order system must be understood. Finally, the decomposed control problem in terms of hierarchical control and parallel computations requires additional studies in order to realize its full potential as a viable methodology to solve a large scale system problem efficiently.

A complete computer software has been developed for the chained aggregation procedure, generalized and restricted QL algorithms and for the simulation of the three regulator designs described in Chapter 5. A documentation may be found in [13].

APPENDIX A

PROOFS OF THEOREM 2.9 AND COROLLARY 2.9

Proof of Theorem 2.9

For notational simplicity, all subscript and superscript "i" in (2.21) will be omitted. In addition, introduce the following notations:

$$
\begin{aligned}
&x_1 = z^i, \quad x_2 = \bar{z}^i, \quad \hat{x}_1 = \hat{z}^i, \quad \hat{x}_2 = \hat{\bar{z}}^i \\
&\|F\| = m_1, \quad \|E\| = m_2, \quad \|G\| = m_3 \\
&\|H\| = m_5, \quad \|\bar{F}\| = m_4, \quad \|\bar{G}\| = m_6 \\
&\|X_{10}\| = \eta_1, \quad \|X_{20}\| \eta_2
\end{aligned}
\tag{A1}
$$

and

$$
\begin{aligned}
&\|\phi(t,\tau)\| \le \rho_1 e^{\alpha_1(t-\tau)} \qquad \rho_1 \ge 0, \quad \rho_1, \alpha_1 \in R \\
&\|\bar{\phi}(t,\tau)\| \le \rho_2 e^{\alpha_2(t-\tau)} \qquad \rho_2 \ge 0, \quad \rho_2, \alpha_2 \in R
\end{aligned}
\tag{A2}
$$

where $\phi(\cdot)$, $\bar{\phi}(\cdot)$ are the state transition matrices for F and $\bar{F}$ respectively. Note also no sign definiteness is assumed on α_1 and α_2 except they are constants. Define

$$e(t) = x_1(t) - \hat{x}_1(t) \tag{A3}$$

then

$$
\begin{aligned}
&\dot{e} = Fe + E x_2 \\
&e(0) = 0
\end{aligned}
\tag{A4}
$$

Integrating (A4), we have

$$e(t) = \int_o^t \phi(t-\tau)\ E\, x_2(\tau)dt \tag{A5}$$

Using (2.56a), (2.57b) and (A3), we get

$$x_2(t) = \bar{\phi}(t,o)x_{20} + \int_o^t \bar{\phi}(t,\tau)\ H\, e(\tau)d\tau + \int_o^t \bar{\phi}(t,\tau)\ H\,\hat{x}_1(\tau)d\tau$$
$$+ \int_o^t \bar{\phi}(t,\tau)\ \bar{G}\ u(\tau)d\tau \tag{A6}$$

and

$$\hat{x}_1(t) = \phi(t,o)x_{10} + \int_o^t \phi(t,\tau)\ G\ u(\tau)d\tau \tag{A7}$$

Taking the instantaneous norm of (A5) - (A7),

$$\|e(t)\| \le \int_o^t \|\phi(t,o)\|\ \|E\|\ \|x_2(\tau)\|d\tau \tag{A7a}$$

$$\|x_2(\tau)\| \le \|\bar{\phi}(\tau,o)\|\ \|x_{20}\| + \int_o^t \|\bar{\phi}(\tau,\sigma)\|\ \|H\|\ \|e(\sigma)\|\, d\sigma$$
$$+ \int_o^\tau \|\bar{\phi}(\tau,\sigma)\|\ \|H\|\ \|\hat{x}_1(\sigma)\|d\sigma \tag{A7b}$$
$$+ \int_o^\tau \|\phi(\tau,\sigma)\|\ \|\bar{G}\|\ \|u(\sigma)\|\, d\sigma$$

$$\|\hat{x}_1(\sigma)\| \le \|\phi(\sigma,o)\|\ \|x_{10}\| + \int_o^\sigma \|\phi(\sigma,s)\|\ \|G\|\ \|u(s)\|ds \tag{A7c}$$

After some tedious algebraic manipulation and integration and using (A1) - (A2), we arrive at:

$$\|e(t)\| \le K_1(t)e^{\alpha_1 t} + K_2 e^{\alpha_2 t} + K_3 - \frac{m_2\rho_1\rho_2 m_5}{(\alpha_1-\alpha_2)}\ e^{\alpha_2 t}\cdot \int_o^t \|e(\tau)\|e^{-\alpha_2\tau}\, d\tau$$
$$+ \frac{m_2\rho_1\rho_2 m_5}{(\alpha_1-\alpha_2)}\ e^{\alpha_1 t}\cdot \int_o^t \|e(\tau)\|e^{-\alpha_1\tau}\, d\tau \tag{A8}$$

where

$$K_1(t) = \left[\frac{m_2\rho_1^2 m_5 \eta_1 \rho_2}{(\alpha_1-\alpha_2)} + \frac{m_2\rho_1^2\rho_2 m_3 m_5 K}{\alpha_1(\alpha_1-\alpha_2)}\right] t + \frac{m_2\eta_2\rho_1\rho_2}{(\alpha_1-\alpha_2)}$$

$$- \frac{m_2\rho_1^2 m_5\eta_1\rho_2}{(\alpha_1-\alpha_2)^2} - \frac{m_2\rho_1^2\rho_2 m_3 m_5 K}{\alpha_1(\alpha_1-\alpha_2)^2} + \frac{m_2\rho_1 m_6 K\rho_2}{\alpha_2(\alpha_1-\alpha_2)}$$

$$- \frac{m_2\rho_1 m_6 K\rho_2}{\alpha_1\alpha_2} + \frac{m_2\rho_1^2\rho_2 m_3 m_5 K}{\alpha_1^2\alpha_2} - \frac{m_2\rho_1^2 m_3 m_5\rho_2 K}{\alpha_1\alpha_2(\alpha_1-\alpha_2)}$$

$$= K_1' + K_1'' t \tag{A9a}$$

$$K_2 = \frac{m_2\rho_1^2 m_5\eta_1\rho_2}{(\alpha_1-\alpha_2)^2} - \frac{m_2\eta_2\rho_1\rho_2}{(\alpha_1-\alpha_2)} + \frac{m_2\rho_1^2\rho_2 m_3 m_5 K}{\alpha_1(\alpha_1-\alpha_2)^2}$$

$$- \frac{m_2\rho_1 m_6 K\rho_2}{\alpha_2(\alpha_1-\alpha_2)} + \frac{m_2\rho_1^2 m_3 m_5\rho_2 K}{\alpha_1\alpha_2(\alpha_1-\alpha_2)} \tag{A9b}$$

$$K_3 = \frac{m_2\rho_1\rho_2 m_6 K}{\alpha_1\alpha_2} - \frac{m_2\rho_1^2\rho_2 m_3 m_5 K}{\alpha_1^2\alpha_2} \tag{A9c}$$

The following Lemma known as the Gronwall's lemma is required for the subsequent derivation.

Lemma [16] Let $\lambda(t)$ be a real continuous function and $\mu(t)$ a nonnegative continuous function on the interval $[a,b]$. If a continuous function $y(t)$ has the property that

$$y(t) \leq \lambda(t) + \int_a^t \mu(s)\, y(s)ds \tag{A10}$$

for $a \leq t \leq b$, then on the same interval

$$y(t) \leq \lambda(t) + \int_a^t \lambda(s)\ \mu(s)\ \exp\ (\int_s^t \mu(\tau)d\tau)ds \tag{A11}$$

In particular if $\lambda(t) \equiv \lambda$ is a constant

$$y(t) \leq \lambda\ \exp\ (\int_a^t \mu(s)ds). \tag{A12}$$

Next let us define

$$\alpha = \max\ (\alpha_1,\ \alpha_2) \tag{A13}$$

$$\mu = \frac{m_2 \rho_1 \rho_2 m_5}{|\alpha_1 - \alpha_2|} \geq 0 \tag{A14}$$

A closer examination of (A8) indicates that two distinct cases exists

<u>Case I</u>: $(\alpha_1 - \alpha_2) > 0$

Dropping the negative terms in (A8) and using (A13) - (A14), we have

$$\|e(t)\| \leq K_1(t)e^{\alpha t} + K_2 e^{\alpha t} + K_3 + \mu e^{\alpha t} \cdot \int_o^t \|e(\tau)\| e^{-\alpha \tau} d\tau \tag{A15}$$

Define

$$w(t) = \|e(t)\| e^{-\alpha t} \tag{A16}$$

then

$$w(t) \leq K_1(t) + K_2 + K_3 e^{-\alpha t} + \mu \int_o^t w(\tau)d\tau \tag{A17}$$

Defining

$$\lambda(t) = K_1(t) + K_2 + K_3 e^{-\alpha t} \tag{A18}$$

We may apply Gronwall's Lemma to (A17) which yields upon integration

$$w(t) \leq K_1' + K_2 + K_3 e^{-\alpha t} + (K_1' + K_2)e^{\mu t} + \frac{1}{\mu} K_1'' e^{\mu t} + \frac{\mu K_3}{\alpha + \mu}(e^{\mu t} - e^{-\alpha t})$$

or equivalently after post multiply by $e^{\alpha t}$ and regrouping, we finally get

$$\|e(t)\| \leq \frac{K_3 \alpha}{\alpha + \mu} + (K_1' + K_2)e^{\alpha t} + \left[\frac{(K_1' + K_2 + K_3)\mu^2 + (K_1'\alpha + K_2\alpha + K_1'')\mu + K_1''\alpha}{\mu(\alpha + \mu)}\right] e^{(\mu + \alpha)t} \tag{A19}$$

<u>Case II</u>: $(\alpha_1 - \alpha_2) < 0$

Dropping negative terms in (A8) and using (A13) - (A14), we again have

$$\|e(t)\| \leq K_1 e^{\alpha t} + K_2 e^{\alpha t} + K_3 + \mu e^{\alpha t} \int_0^t \|e(\tau)\| e^{-\alpha \tau} d\tau$$

Consequently, (A19) also holds.

<u>Proof of Corollary 2.9</u>

Starting from (A8) and following similar procedure as in the proof of Theorem 2.9, we can show

<u>Case I'</u>: $(\alpha_1 - \alpha_2) > 0$

$$\|e(t)\| \leq \frac{K_3 \alpha}{\alpha + \mu} + (\tilde{K}_1' + \tilde{K}_2)e^{\alpha t} + \left[\frac{(\tilde{K}_1' + \tilde{K}_2 + \tilde{K}_3)\mu^2 + (\tilde{K}_1'\alpha + \tilde{K}_2\alpha + \tilde{K}_1'')\mu + \tilde{K}_1''\alpha}{\mu(\alpha + \mu)}\right] e^{(\alpha + \mu)t} \tag{A20}$$

where

$$\tilde{K}_1 = \tilde{K}_1' + \tilde{K}_2'' t$$

$$= \left[\frac{\mu \eta_2}{m_5} - \frac{\mu \rho_1 m_3 K}{\alpha_1 (\alpha_1 - \alpha_2)} \right] + \mu \rho_1 \eta_1 t \quad \text{(A21a)}$$

$$\tilde{K}_2 = \frac{\mu \rho_1 \eta_1}{\alpha_1 - \alpha_2} - \frac{\mu m_6 K}{\alpha_2 m_5} + \frac{\mu \rho_1 m_3 K}{\alpha_1 \alpha_2} \quad \text{(A21b)}$$

$$\tilde{K}_3 = K_3 = \frac{(\alpha_1 - \alpha_2) \mu m_6 K}{\alpha_1 \alpha_2 m_5} - \frac{\mu \rho_1 m_3 K (\alpha_1 - \alpha_2)}{\alpha_1^2 \alpha_2} \quad \text{(A21c)}$$

<u>Case II'</u>: $(\alpha_1 - \alpha_2) < 0$

$$\|e(t)\| \le \frac{\hat{K}_3 \alpha}{\alpha+\mu} + (\hat{K}_1' + \hat{K}_2) e^{\alpha t} + \left[\frac{(\hat{K}_1' + \hat{K}_2 + \hat{K}_3)\mu^2 + (\hat{K}_1'\alpha + \hat{K}_2 \alpha + \hat{K}_1'')\mu + \hat{K}_1''\alpha}{\mu(\alpha+\mu)} \right] e^{(\alpha+\mu)t} \quad \text{(A22)}$$

where

$$\hat{K}_1 = \hat{K}_1' + \hat{K}_1'' t$$

$$= \left[\frac{\mu \rho_1 m_3 K}{\alpha_2 (\alpha_1 - \alpha_2)} - \frac{\mu m_6 K}{m_5 \alpha_2} \right] - \frac{\mu \rho_1 m_3 K}{\alpha_1} t \quad \text{(A23a)}$$

$$\hat{K}_2 = \frac{\mu \eta_2}{m_5} - \frac{\mu \rho_1 \eta_1}{(\alpha_1 - \alpha_2)} \quad \text{(A23b)}$$

$$\hat{K}_3 = K_3 = \frac{m_2 \rho_1 m_6 K \rho_2}{\alpha_1 \alpha_2} - \frac{m_2 \rho_1^2 \rho_2 m_3 m_5 K}{\alpha_1^2 \alpha_2} \quad \text{(A23c)}$$

Under the assumption of F, $\bar{F}$ stable, $\alpha_i < 0$, $i = 1,2$, thus $\alpha < 0$. Equations (A20) and (A22) immediately give the conclusion of Corollary 2.9.

APPENDIX B

PROOF OF THEOREM 2.10

Define the quantity

$$\beta_{ij} = y_i^T \Delta A_c x_i \qquad i,j = 1,2,\ldots n \tag{B1}$$

then

$$X^{-1} A_c X = X^{-1}(\hat{A}_c + \Delta A_c) X$$

$$= \mathrm{diag}\,(\hat{\lambda}_1, \hat{\lambda}_2 \ldots \hat{\lambda}_n) + [\beta_{ij}] \tag{B2}$$

Apply Lemma 2.3, we conclude λ_i lies in disc with center $\hat{\lambda}_i + \beta_{ii}$ and radius $\sum_{j \neq i} |\beta_{ij}|$. Since eigenvalues of A_c are distinct by assumption, there exists distinct ϵ - neighborhoods of λ_i such that λ_i lies within this ϵ - neighborhood. If the radius of the i-th disc can be reduced so as to isolate it from the rest of the Gorschgorin discs, λ_i may be determined within the isolated disc.

Consider the isolation of the i-th Gorschgorin disc from the rest. To reduce the radius $r_i = \sum_{j \neq i} |\beta_{ij}|$, let

$$T_i = \begin{vmatrix} I_{i-1} & 0 & 0 \\ 0 & \epsilon_i & 0 \\ 0 & 0 & I_{n-i} \end{vmatrix}, \quad \epsilon_i \in R, \quad \epsilon_i << 1 \tag{B3}$$

be a linear transformation to be applied to the right hand side of (B2). Thus the eigenvalues remain invariant. Then

$$A_c \sim \mathrm{diag}\,(\hat{\lambda}_1, \ldots \hat{\lambda}_n) + T_i\, \beta\, T_i^{-1} \tag{B4}$$

where $\sim$ denotes equivalency in terms of linear transformation on A_c.

T_i serves to modify the radius of the Gorschgorin discs. ϵ_i is chosen so that the radius of the i-th disc is reduced as much as possible while keeping the radii of other discs sufficiently small to avoid overlapping with the i-th disc. Pictorally, this is illustrated by the following

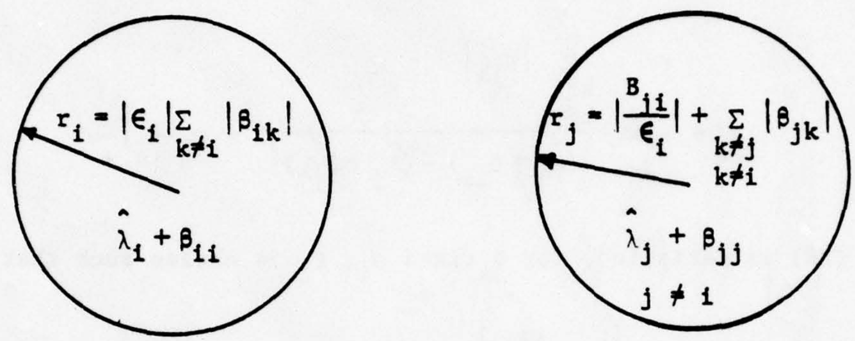


Fig. B1. The effect of T_i on the i, j-th Gorschgorin disc.

Notice while the radius of i-th disc is reduced, the radii of other disc inc increases. From Fig. B1, we observe ϵ_i is to be chosen such that

$$\left|\frac{\beta_{ji}}{\epsilon_i}\right| + \sum_{\substack{k \neq j \\ k \neq i}} |\beta_{jk}| \leq \rho_i \left|(\hat{\lambda}_i + \beta_{ii}) - (\hat{\lambda}_j + \beta_{jj})\right| \quad \forall j \neq i \quad \text{(B5)}$$

$$|\epsilon_i| \sum_{k \neq i} |\beta_{ik}| \leq (1 - \rho_i) \left|(\hat{\lambda}_i + \beta_{ii}) - (\hat{\lambda}_j + \beta_{jj})\right| \quad \forall j \neq i \quad \text{(B6)}$$

where $0 < \rho_i < 1$.

The implication of (B5) - (B6) is made clear by combining the equations into

$$\left|\frac{\beta_{ji}}{\epsilon_i}\right| + \sum_{\substack{k \neq j \\ k \neq i}} |\beta_{jk}| + |\epsilon_i| \sum_{k \neq i} |\beta_{ik}| \leq \left|(\hat{\lambda}_i + \beta_{ii}) - (\hat{\lambda}_j + \beta_{jj})\right| \quad \forall j \neq i \quad \text{(B7)}$$

(B7) says $r_i + r_j \le d_{ji}$ $\forall j \neq i$ where r_i is the radius of the i-th disc and d_{ji} is the distance between the centers of the i-th and j-th discs. The existence of such an ϵ_i guarantees the isolation of the i-th disc from the rest for the chosen value ρ_i. From (B5), it is necessary to choose ρ_i so that

$$\underline{\rho}_i = \max_j \frac{\sum\limits_{\substack{k\neq j\\ k\neq i}} |\beta_{jk}|}{|(\hat{\lambda}_i + \beta_{ii}) - (\hat{\lambda}_j + \beta_{jj})|} \le \rho_i < 1 \tag{B8}$$

If (B8) is satisfied, for a fixed ρ_i, ϵ_i is chosen such that

$$|\epsilon_i| \ge \max_{\substack{j\\ j\neq i}} \frac{|\beta_{ji}|}{\rho_i|(\hat{\lambda}_i + \beta_{ii}) - (\hat{\lambda}_j + \beta_{jj})| - \sum\limits_{\substack{k\neq j\\ k\neq i}} |\beta_{jk}|} \equiv \underline{\gamma}_i \tag{B9}$$

$$|\epsilon_i| \le \min_{\substack{j\\ j\neq i}} \frac{(1-\rho_i)|(\hat{\lambda}_i + \beta_{ii}) - (\hat{\lambda}_j + \beta_{jj})|}{\sum\limits_{k\neq i} |\beta_{ik}|} \equiv \bar{\gamma}_i \tag{B10}$$

The interval of existence of ϵ_i is illustrated by Fig. B2.

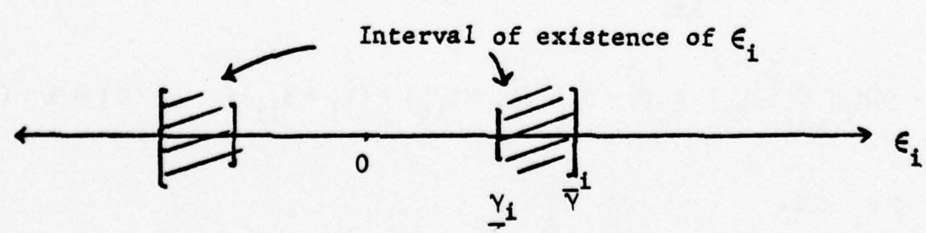


Fig. B2 Interval of existence of ϵ_i.

It is obvious that ϵ_i exists if and only if $\underline{\gamma}_i \le \bar{\gamma}_i$. Assuming this is true (see example in Chapter 2), then the smallest ϵ_i is

$$\epsilon_i = \max_{\substack{j \\ j \neq i}} \frac{|\beta_{ji}|}{\rho_i |(\hat{\lambda}_i + \beta_{ii}) - (\hat{\lambda}_j + \beta_{jj})| - \sum\limits_{\substack{k \neq j \\ k \neq i}} |\beta_{jk}|} > 0 \qquad \text{(B11)}$$

Since $r_i = |\epsilon_i| \sum\limits_{k \neq i} |\beta_{ik}|$, then the radius of the i-th Gorschgorin disc is

$$r_i = \max_{\substack{j \\ j \neq i}} \frac{|\beta_{ji}| \sum\limits_{k \neq i} |\beta_{ik}|}{\rho_i |(\hat{\lambda}_i + \beta_{ii}) - (\hat{\lambda}_j + \beta_{jj})| - \sum\limits_{\substack{k \neq j \\ k \neq i}} |\beta_{jk}|} \qquad \text{(B12)}$$

Giving us the necessary bound on λ_i as

$$\hat{\lambda}_i + \beta_{ii} - r_i \leq \lambda_i \leq \hat{\lambda}_i + \beta_{ii} + r_i \qquad \text{(B13)}$$

This completes the proof of Theorem 2.10.

APPENDIX C

PROOF OF THEOREM 4.1

Taking time derivative of (4.48) and evaluating along (4.45), we get

$$\dot{v}\big|_{(4.45)} = \dot{z}^T Pz + z^T P\dot{z}$$
$$= z^T(F_c^T P + PF_c)z + 2z^T \Delta F^T Pz$$
$$= -z^T Qz + 2\, z^T \Delta F^T Pz$$

But

$$z^T \Delta F^T Pz \le \|z\|\, \|\Delta F^T\|\, \|Pz\| \le \|\Delta F\|\, \|P\|_s \|z\|^2 \qquad \text{(C1)}$$

If

$$0 \le \|\Delta F\| \le \frac{1}{2}\, \frac{\lambda_{min}(Q)}{\lambda_{max}(P)}$$

Using the fact

$$\|P\|_s \equiv \{\lambda_{max}(PP^T)\}^{1/2}$$
$$= \{\lambda^2_{max}(P)\}^{1/2}$$
$$= \lambda_{max}(P)$$

then (C1) becomes

$$z^T \Delta FPz \le \frac{1}{2}\lambda_{min}(Q)\, \|z\|^2$$

which implies

$$\dot{v}\Big|_{(4.45)} \le -z^T Q\, z + \lambda_{min}(Q)\| z \|^2$$

$$\le -z^T\ [Q - \lambda_{min}(Q)\cdot I]\ z$$

Hence $\dot{v} \le 0$. It is easy to see $\dot{v} \not\equiv 0$ along any non-zero trajectory of (4.45). Thus stability follows

APPENDIX D

PROOF OF THEOREM 4.3

Recall the definition of R(M) from (4.57), expanding the right hand side, we obtain

$$R(M) = M[F_{022} + F_{c21}Q^{-1}PF_{012}] + [F_{022} + F_{c21}Q^{-1}PF_{012}]^T M$$
$$+ MF_{c21}Q^{-1}F_{c21}^T M + F_{012}^T PQ^{-1}PF_{012} \tag{D1}$$

Define

$$\tilde{F} = F_{022} + F_{c21}Q^{-1}PF_{012} \tag{D2a}$$

$$\tilde{S} = F_{c21}Q^{-1}F_{c21}^T \tag{D2b}$$

$$\tilde{Q} = F_{012}^T PQ^{-1}PF_{012} \tag{D2c}$$

Then (D1) may be rewritten in simplier form:

$$R(M) = M\tilde{F} + \tilde{F}^T M + M\tilde{S}M + \tilde{Q} < 0 \tag{D3}$$

Let us examine first the equality instead:

$$R(M) = M\tilde{F} + \tilde{F}^T M + M\tilde{S}M + \tilde{Q} = 0 \tag{D4}$$

Notice the requirement of a positive definite and symmetric solution to (D4) is equivalent to asking if there is a negative definite and symmetric solution $N = -M$ to

$$R_1(N) = N\tilde{F} + \tilde{F}^T N - N\tilde{S}N - \tilde{Q} = 0 \tag{D5}$$

(D5) is a Riccati equation. Thus the initial requirement of a $M>0$ with $R(M)<0$ is equivalent to guaranteeing the existence of a $N<0$ such that $R_1(N)>0$. It turns out that the existence of a $N<0$ such that $R_1(N)>0$ is related to the existence of a $N_1<0$ such that $R_1(N_1)=0$. In order to facilitate the proofs, the following lemmas will be required:

Lemma D1 [48]

Consider the algebraic Riccati equation

$$KA + A^TK - KBB^TK - C^TC = 0 \qquad \text{(ARE)}$$

If (A, B), (A, C) are completely controllable and completely observable respectively and $\text{Re}\,\lambda(A)<0$ (where $\text{Re}\,\lambda(\cdot)$ means real part of eigenvalue λ). Then there exists a real symmetric solution K to the (ARE) such that $\text{Re}\,\lambda(A-BB^TK)\leq 0$ if and only if

$$I - B^T(-j\omega I - A^T)^{-1}\, C^TC(j\omega I - A)^{-1}B \geq 0 \quad \forall \text{ real } \omega \qquad \text{(D6)}$$

Moreover, such a solution K is unique, negative definite and symmetric.

Unfortunately, Lemma D1 only guarantees $\text{Re}\,\lambda(A-BB^TK)\leq 0$. Lemma D2 strengthens the results of Lemma D1.

Lemma D2

Under the assumptions of Lemma D1, then there exists a unique negative definite and symmetric solution K to the (ARE) such that $\text{Re}\,\lambda(A-BB^TK)<0$ if and only if

$$I - B^T(-j\omega I - A^T)^{-1}\, C^TC(j\omega I - A)^{-1}B > 0 \qquad \text{(D7)}$$

$\forall$ real ω.

Proof: Rewrite (ARE) as

$$A^TK + KA - KBB^TK + Q_o = 0$$

where $Q_o = -C^TC$. By hypothesis, (A,B) is controllable. Theorem 6 of Monilari [49] states if there exists a F satisfying Re (A - BF) < 0, then there exists a solution K^+ of (ARE) such that $\mathrm{Re}\,\lambda(A - BB^TK^+) < 0$ if and only if

$$\phi_F(j\omega) = [I - F(-j\omega I - A)^{-1}B]^T[B^T(-j\omega I - A^T)^{-1}Q_o(j\omega I - A)^{-1}B + I][I - F(j\omega I - A)^{-1}B] > 0 \tag{D8}$$

for all ω. Since $\mathrm{Re}\,\lambda(A) < 0$ by assumption, it is sufficient to take $F = 0$. In addition, from Remark 13 of Willems [48], if $Q_o = -C^TC$, all solutions to the (ARE) are negative definite, thus $K = K^+ < 0$. Apply (D8) with $F \equiv 0$, we conclude $\mathrm{Re}\,\lambda(A - BB^TK) < 0$ if and only if

$$\phi_o(j\omega) = I - B^T(-j\omega I - A^T)^{-1}C^TC(j\omega I - A)^{-1}B > 0 \qquad \forall\omega$$

This proves the lemma.

<u>Lemma D3</u>

Consider the quantity

$$H(\omega) = B^T(-j\omega I - A^T)^{-1}\,(j\omega I - A)^{-1}\,B \qquad \forall\omega \tag{D9}$$

Then $H(\omega)$ is bounded if A is a stable matrix.

Proof: Taking the norm of (D9), we have

$$\|H(\omega)\| = \|B^T(-j\omega I - A^T)^{-1}\,(j\omega I - A)^{-1}\,B\|$$

$$\leq \|B^T(-j\omega I - A^T)^{-1}\|\,\|(j\omega I - A)^{-1}\,B\|$$

$$\leq \|(j\omega I - A)^{-1} B\|^2$$

$$\leq \|(j\omega I - A)^{-1}\|^2 \|B\|^2$$

By assumption, A is stable. Hence $\|(SI - A)\|_{s = j\omega}$ is bounded. Futhermore, B is a constant matrix independent of ω. Thus $\|H(\omega)\|$ is also bounded.

Let us recall the objective is to examine the existence of a negative definite and symmetric solution to (D5). Comparing (D5) with the (ARE), they are identical if we associate $\tilde{S} = BB^T$ and $\tilde{Q} = C^T C$. Hence Lemma D1 to D3 may be utilized. We now prove the following result.

<u>Theorem D1</u>: Consider the algebraic Riccati equation (D5), $R_1(N) = 0$. If there exist a symmetric equilibrium point $N_1 < 0$ such that $R_1(N_1) = 0$ and Re $\lambda(\tilde{F} - \tilde{S} N_1) < 0$, then there also exist a symmetric matrix $N_2 < N_1 < 0$ such that $R_1(N_2) > 0$.

Proof: By hypothesis, $R_1(N_1) = 0$. Denote

$$N_2 \equiv N_1 - N^* < 0, \; N^* > 0$$

Consider $R_1(N_2)$, we have

$$R_1(N_2) = (N_1 - N^*)\tilde{F} + \tilde{F}^T(N_1 - N^*) + (N_1 - N^*)\tilde{S}(N_1 - N^*) - \tilde{Q}$$

$$= (-N^*)(\tilde{F} - \tilde{S} N_1) + (\tilde{F} - \tilde{S} N_1)^T(-N^*) - (-N^*)\tilde{S}(-N^*)$$

$$= M^* F_1 + F_1^T M^* - M^* \tilde{S} M^* \tag{D10}$$

where $F_1 \equiv F - \tilde{S} N_1$ is a stable matrix and $M^* \equiv -N^* < 0$. Next consider the perturbed Riccati equation for some constant $k > 0$:

$$M^* F_1 + F_1^T M^* - M^* \tilde{S} M^* - kI = 0 \tag{D11}$$

Applying Lemma D2 to (D11), we see that if (i) (F_1, F_{c21}) is completely controllable; (ii) $(F_1, \sqrt{k}I)$ is completely observable, then a negative definite and symmetric matrix M^* with the property $\mathrm{Re}\,\lambda(F_1 - \tilde{S}M^*) < 0$ exists if and only if

$$\text{(iii)}\quad I - kF_{c21}^T(-j\omega I - F_1^T)^{-1}(j\omega I - F_1)^{-1}F_{c21} > 0 \qquad \forall\omega$$

But the existence of a N_1 implies $(\tilde{F}, F_{c21})$ is controllable which implies $(F_1, F_{c21}) = (\tilde{F} - \tilde{S}N_1, F_{c21})$ is controllable because controllability is preserved under state feedback. $(F_1, \sqrt{k}I)$ is clearly observable. Furthermore, by Lemma D3, $F_{c21}^T(-j\omega I - F_1^T)^{-1}(j\omega I - F_1)^{-1}F_{c21} \geq 0$ is bounded because F_1 is stable by hypothesis. Hence there always exists a real number $k > 0$ such that (iii) holds. Applying Lemma D3, a unique negative definite and symmetric M^* exists. Moreover,

$$R_1(M^*) = M^*F_1 + F_1^TM^* - M^*\tilde{S}M^* = kI > 0$$

giving $N_2 = N_1 - N^* = N_1 + M^* < 0$. In view of (D10), then $R_1(N_2) = R_1(M^*) > 0$. This completes the proof.

Application of Theorem D1 allows us to complete the proof of Theorem 4.3. To see this, without loss of generality, choose $Q = I$ in (4.50). Then from (D2), we have

$$\tilde{F} = \bar{F}^i + (H^i + \bar{G}^iK)PE^i$$

$$\tilde{S} = (H^i + \bar{G}^iK)\,(H^i + \bar{G}^iK)^T$$

$$\tilde{Q} = E^{i\,T}PPE^i$$

In order for $N_1 < 0$ to be a solution of (D5), from Lemma D2, it is sufficient to have conditions (a)-(c)of Theorem 4.3 hold. Moreover, the solution N_1 satisfies $\mathrm{Re}(\tilde{F} - \tilde{S} N_1) < 0$. Then by Theorem D1, there also exists a $N_2 < 0$ such that $R_1(N_2) > 0$. Let $M = -N_2 > 0$, clearly $R(M) < 0$. Thus a positive definite and symmetric matrix M exists. By Theorem 4.2, $\Lambda(M) < 0$. Stability and robustness conclusion follows from (4.54).

APPENDIX E

POWER SYSTEM PARAMETERS OF EXAMPLE 6.1

The nonzero elements of the matrices A, B, C, and F expressed as functions of the physical parameters are:

$$a_{11} = -\frac{r_1}{T_{s_1}} \qquad a_{19} = -\frac{1}{T_{s_1}} \qquad a_{21} = \frac{k_{t_1}}{T_{u_1}} \qquad a_{22} = -\frac{1}{T_{u_1}}$$

$$a_{32} = \frac{1}{T_{r_1}} \qquad a_{33} = -\frac{1}{T_{r_1}} \qquad a_{43} = \frac{1}{T_{n_1}} \qquad a_{44} = -\frac{1}{T_{n_1}}$$

$$a_{55} = -\frac{r_2}{T_{s_2}} \qquad a_{59} = -\frac{1}{T_{s_2}} \qquad a_{65} = \frac{k_{t_2}}{T_{u_2}} \qquad a_{66} = -\frac{1}{T_{u_2}}$$

$$a_{76} = \frac{1}{T_{r_2}} \qquad a_{77} = -\frac{1}{T_{r_2}} \qquad a_{87} = \frac{1}{T_{n_2}} \qquad a_{88} = -\frac{1}{T_{n_2}}$$

$$a_{92} = \frac{e_1 c_{v_1}}{T} \qquad a_{93} = \frac{e_1 c_{s_1}(1-c_{v_1})}{T} \qquad a_{94} = \frac{e_1(1-c_{s_1})(1-c_{v_1})}{T}$$

$$a_{96} = \frac{e_2 c_{v_2}}{T} \qquad a_{97} = \frac{e_2 c_{s_2}(1-c_{v_2})}{T} \qquad a_{98} = \frac{e_2(1-c_{s_2})(1-c_{v_2})}{T}$$

$$a_{99} = -e/T$$

$$b_{11} = \frac{1}{T_{s_1}} \qquad b_{52} = \frac{1}{T_{s_2}}$$

$$c_{12} = c_{v_1} \qquad c_{13} = (1-c_{v_1})c_{s_1} \qquad c_{14} = (1-c_{v_1})(1-c_{s_1})$$

$$c_{26} = c_{v_2} \qquad c_{27} = (1-c_{v_2})c_{s_2} \qquad c_{28} = (1-c_{v_2})(1-c_{s_2})$$

$$c_{19} = e_{T_1} \qquad c_{29} = e_{T_2} \qquad c_{39} = 1$$

$$f_{19} = -\frac{1}{T}.$$

$$
\begin{array}{ll}
r = 0.05 & c_v = 0.3 \\
T_s = 0.25 & c_s = 0.4 \\
T_u = 0.20 & e_T = 0.15 \\
T_r = 6.00 & k_t = 0.95 \\
T_n = 0.50 & e_p = 1.50 \\
T = 12.0 & e_1 = e_2 = 0.5.
\end{array}
$$

REFERENCES

1. M. Aoki, "Control of Large-Scale Dynamic Systems by Aggregation," IEEE Trans. on Automatic Control, Vol. AC-13, No. 3, pp. 246-253, June 1968.

2 M. Aoki, Optimization Methods for Large-Scale Systems, David A. Wismer, Ed., Chapter 5, McGraw-Hill Book Co., New York, 1971.

3. J. Hickin and N. K. Sinha, "Aggregation Matrices for a Class of Low-Order Models for Large-Scale Systems," Electronics Letters, Vol. 11, No. 9, p. 186, May 1975.

4. S. S. Lamba and S. Vittal Rao, "On Suboptimal Control Via the Simplified Model of Davison," IEEE Trans. on Automatic Control, Vol. 19, No. 4, pp. 448-450, August 1974.

5. N. K. Sinha and Z. Elrazaz, "Eigenvalue Sensitivity of Aggregated Models," Proc. Fifteenth Annual Allerton Conference on Communications, Control, and Computing, pp. 305-313, September 1977.

6. J. M. Siret, G. Michailesco,and P. Bertrand, "Optimal Approximation of High-Order Systems Subject to Polynomial Inputs," Int. J. Control, Vol. 26, No. 6, pp. 963-971, December 1977.

7. J. Hickin and N. K. Sinha, "Eigenvalue Assignment by Reduced Order Models," Electronic Letters, Vol. 11, No. 15, pp. 318-319, July 1975.

8. M. Aoki, "Aggregation and Informationally Decentralized Linear Dynamic Systems," Large-Scale Systems Theory and Applications, G. Guardabassi and A. Locatelli, Eds., pp. 449-455, 1976.

9. C. T. Chen, Introduction to Linear System Theory, pp. 200-201, Holt, Rinehart and Winston, Inc., New York, 1970.

10. E. C. Y. Tse, J. V. Medanic, and W. R. Perkins, "Generalized Hessenberg Transformation for Reduced Order Modeling of Large Scale Systems," Int. J. Control, Vol. 27, No. 4, pp. 493-512, April 1978.

11. E. C. Y. Tse, J. V. Medanic, and W. R. Perkins, "Chained Aggregation of Linear Time Invariant Systems," Proc. JACC, pp. 550-555, 1977.

12. B. Noble, Applied Linear Algebra, Prentice Hall, Inc., Englewood Cliffs, NJ, 1969.

13. J. V. Medanic, W. R. Perkins, and E. Tse, Reduced Order Modeling by the Restricted QL Algorithm, Report DC-10, Decision and Control Laboratory, University of Illinois, Urbana, IL, March 1978.

14. M. Aoki, "On Subspaces Associated with Partial Reconstruction of State Vectors, the Structure Algorithm and the Predictable Directions of Riccati Equations," IEEE Trans. on Automatic Control, Vol. AC-18, No. 4, pp. 399-400, August 1973.

15. L. M. Silverman and H. J. Payne, "Input Output Structure of Linear Systems with Application to the Decoupling Problem," SIAM J. Control, Vol. 9, pp. 199-233, May 1971.

16. W. A. Coppel, Stability and Asymptotic Behavior of Differential Equations, Boston, MA: D. C. Heath, 1965.

17. E. J. Davison, "A Method for Simplifying Linear Dynamic Systems," IEEE Trans. on Automatic Control, Vol. AC-11, pp. 93-101, January 1966.

18. M. R. Chidambara, "On a Method for Simplifying Linear Dynamic Systems," IEEE Trans. on Automatic Control, Vol. AC-12, pp. 213-214, April 1967.

19. G. Hirzinger and G. Kreisselmeier, "On Optimal Approximation of High-Order Linear Systems by Low-Order Models," Int. J. Control, Vol. 22, No. 3, pp. 399-408, 1975.

20. P. V. Kokotovic and P. Sannuti, "Singular Perturbation Method for Reducing the Model Order in Optimal Control Design," IEEE Trans. on Automatic Control, Vol. AC-13, No. 4, pp. 377-384, August 1968.

21. C. F. Chen and L. S. Shieh, "A Novel Approach to Linear Model Simplifications, Int. J. Control, Vol. 8, No. 6, pp. 561-570, 1968.

22. R. D. Milne, "The Analysis of Weakly Coupled Dynamical Systems," Int. J. Control, Vol. 2, pp. 171-199, August 1965.

23. J. H. Wilkinson, The Algebraic Eigenvalue Problem, Oxford University Press, 1965.

24. J. E. Van Ness, "Sensitivities of Large, Multiple-Loop Control Systems," IEEE Trans. on Automatic Control, Vol. AC-10, No. 3, pp. 308-315, July 1965.

25. A. S. Householder, The Theory of Matrices in Numerical Analysis, Blaisdell Publishing Co., 1965.

26. A. R. Gourlay and G. A. Watson, Computational Methods for Matrix Eigenproblem, John Wiley and Sons, New York, 1973.

27. G. W. Stewart, Introduction to Matrix Computations, Academic Press, New York, 1973.

28. J. G. F. Francis, "The QR Transformation, Parts I and II," The Computer Journal, Vol. 4, pp. 265-271, and pp. 332-341, 1961.

29. B. N. Parlett and W. G. Poole, Jr., "A Geometric Theory for the QR, LU and Power Iteration," SIAM J. of Numerical Analysis, Vol. 10, No. 2, pp. 389-412, April 1973.

30. D. M. Wiberg, Theory and Problems of State Space Linear Systems, Schaum's Outline Series, McGraw-Hill Book Co., New York, 1971.

31. J. H. Anderson, "Geometrical Approach to Reduction of Dynamical Systems," Proc. IEE, Vol. 114, No. 7, pp. 1014-1018, July 1967.

32. L. S. Shieh and M. J. Goldman, "A Mixed Cauer Form for Linear System Reduction," IEEE Trans. Systems, Man, and Cybernetics, Vol. SMC-4, pp. 584-588, Nov. 1974.

33. C. F. Chen, "Model Reduction of Multivariable Control Systems By Means of Matrix Continued Fractions," Int. J. Control, Vol. 20, No. 2, pp. 225-238, 1974.

34. P. V. Kokotovic, W. R. Perkins, J. B. Cruz, Jr. and G. D'Ans, "ϵ - Coupling Method for Near-Optimum Design of Large Scale Linear Systems," Proc. IEE, Vol. 116, No. 5, pp. 889-892, May 1969.

35. D. A. Wilson, "Optimum Solution of Model Reduction Problem," Proc. IEE, Vol. 117, pp. 1161-1165, 1970.

36. R. Genesio and M. Milanese, "A Note on the Derivation and Use of Reduced-Order Models," IEEE Trans. on Automatic Control, Vol. AC-21, No. 1, pp. 118-122, Feb. 1976.

37. M. J. Bosley and F. P. Lees, "A Survey of Simple Tranfer-Function Derivations from High-Order State Variable Models," Automatica, Vol. 8, pp. 765-775, 1972.

38. N. R. Sandell, Jr., P. Varaiya, and M. Athans, "A Survey of Dencentralized Control Methods for Large Scale Systems," System Engineering for Power: Status and Prospects, U.S. Energy Research and Development Administration, 1975.

39. N. R. Sandell, Jr., P. Varaiya, M. Athans, and M. G. Safonov, "Survey of Dencentralized Control Methods for Large Scale Systems," IEEE Transaction on Automatic Control, Vol. AC-23, No. 2, pp. 108-128, April 1978.

40. U. Shaked and N. Karcamias, "The Use of Zero and Zero-Direction in Model Reduction," Int. J. Control, Vol. 23, No. 1, pp. 113-135, Jan. 1976.

41. J. D. Simon, "Theory and Application on Modal Control," Case Western Reserve University, System Report SRC 104-A-67-46, 1967.

42. Leonard M. Silverman, "Realization of Linear Dynamical Systems," IEEE Transaction on Automatic Control, Vol. AC-16, No. 6, pp. 554-567, December 1971.

43. W. S. Levine and M. Athans, "On the Determination of the Optimal Constant Output Feedback Gains for Linear Multivariable Systems," IEEE Trans. Automatic Control, Vol. AC-15, No. 1, pp. 44-48, Feb. 1970.

44. S. P. Singulac, N. M. Cuk, and M. S. Calovic, "A Sequential Algorithm for Calculation of Optimal Output-Constrained Feedback Regulators," Proc. Eleventh Annual Allerton Conf. on Circuit and System Theory, pp. 373-382, Oct. 1973.

45. S. Fujishige, H. Nagai, and Y. Sawaragi, "System-Theoretical Approach to Model Reduction and System-Order Determination," Int. J. Control, Vol. 22, No. 6, pp. 807-819, 1975.

46. R. Genesio, and R. Pome, "Indentification of Reduced Models from Noisy Data," Int. J. Control, Vol. 21, No. 2, pp. 203-211, 1975.

47. W. W. Price, E. M. Gulachenski, and R. F. Silva," Dynamic Equivalents From On-Line Measurements," IEEE Trans. on Power Apparatus and Systems, Vol. PAS-94, No. 4, pp. 1349-1357, July/August 1975.

48. J. C. Willems, "Least Squares Stationary Optimal Control and the Algebraic Riccati Equation," IEEE Trans. on Automatic Control, Vol. AC-16, No. 6, pp. 621-634, December 1971.

49. B. P. Molinari, "The Time-Invariant Linear-Quadratic Optimal Control Problem," Automatica, Vol. 13, pp. 347-357, 1977.

50. H. Erzberger, "Analysis and Design of Model Following Control System by State Space Techniques," Proc. JACC, pp. 572-581, June, 1968.

51. Y. T. Chan, "Perfect Model Following with A Real Model," Proc. JACC, paper 10-5, pp. 287-293, 1973.

52. S. H. Wang and E. J. Davison, "On the Stabilization of Decentralized Control Systems," IEEE Trans. on Automatic Control, Vol. AC-18, pp. 473-478, Oct. 1973.

53. A. Ben-Israel, "On the Geometry of Subspaces in Euclidean N-Spaces," SIAM J. Applied Math, Vol. 15, No. 5, pp. 1184-1198, September 1967.

54. M. T. Li, "On Output Feedback Stabilizability of Linear Systems," IEEE Trans. on Automatic Control, Vol. AC-17, pp. 408-410, June 1972.

55. J. Medanic and M. Andjelic, "On a Class of Differential Game Without Saddle Point Solutions," Journal of Optimization Theory and Applications, Vol. 8, pp. 413-430, 1971.

56. B. D. O. Anderson and J. B. Moore, Linear Optimal Control, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.

57. Robert L. Kosut, "Suboptimal Control of Linear Time-Invariant Systems Subject to Control Structure Constraints," IEEE Trans. on Automatic Control, Vol. 15, No. 5, pp. 557-563, October 1970.

58. Timothy L. Johnson and Michael Athans, "On the Design of Optimal Constrained Dynamic Compensators for Linear Constant Systems," IEEE Trans. on Automatic Control, Vol. 15, No. 6, pp. 658-660, December, 1970.

59. M. Calovic, "Dynamic State-Space Models of Electric Power Systems," Report, Depts. of Electrical and Mechanical Engineering, Univ. of Illinois, Urbana, Illinois, 1971.

60. James E. Potter, "Matrix Quadratic Solutions," SIAM J. of Applied Math., Vol. 14, No. 3, pp. 496-501, May 1966.

61. E. J. Davison, "The Robust Control of a Servo Mechanism Problem for Linear Time Invariant Multivariable Systems," IEEE Trans. on Automatic Control, Vol. AC-21, No. 1, pp. 25-34, February 1976.

62. R. W. Shields and J. B. Pearson, "Structural Controllability of Multi-input Linear Systems," IEEE Trans. on Automatic Control, Vol. AC-21, No. 2, pp. 203-212, April 1976.

63. J. P. Corfmat and A. S. Morse, "Decentralized Control of Linear Multi-variable Systems," Automatica, Vol. 12, pp. 479-495, 1976.

64. M. Aoki, "On Feedback Stabilizability of Decentralized Dynamic Systems," Automatica, Vol. 8, pp. 163-173, 1972.

65. S. H. Wang and E. J. Davison, "On the Stabilization of Decentralized Control Systems," IEEE Trans. Aut. Control, Vol. AC-18, pp. 473-478, 1973.

66. E. J. Davison and N. K Tripathi, "The Optimal Dencentralized Control of a Large Power System: Load and Frequency Control," IEEE Trans. Aut. Control, Vol. AC-23, No. 2, pp. 312-325, 1976.

67. N. R. Sandell, Jr., "Decomposition vs Decentralization in Large Scale System Theory," Proc. JACC, pp. 1043-1046, 1976.

68. B. C. Moore, "Singular Value Analysis of Linear Systems, Part I: External Variables," Systems Control Report No. 7801, (April 1978), Department of Electrical Engineering, University of Toronto, Toronto, Ontario, M5S 1A4.

69. B. C. Moore, "Singular Value Analysis in Linear Systems, Part II: Controllability,Observability, and Model Reduction," Systems Control Report No. 7802, (April 1978), Department of Electrical Engineering, University of Toronto, Toronto, Ontario, M5S 1A4.

VITA

Edwin Chi Yu Tse was born in New York City, New York, on April 20, 1951. He received his Bachelor of Science and Master of Science degrees in Electrical Engineering from the University of Illinois, Urbana, in June 1973 and January 1975 respectively.

From September 1973 to May 1975, he was a graduate teaching assistant with the Department of Electrical Engineering. Since June 1975, he has been a research assistant in the Decision and Control Laboratory at the Coordinated Science Laboratory, University of Illinois. His interests are in large scale system theories, particularly model reduction, decentralized control and hierarchical control methods.

Mr. Tse is a member of the Institute of Electrical and Electronics Engineers and the honor societies of Eta Kappa Nu and Phi Kappa Phi.