AN ALGEBRAIC APPROACH TO TIME SCALE ANALYSIS AND CONTROL

by

Xi-Cheng Lou

B.S., Qing-hua University, (P.R. China)
(1965)

S.M., Massachusetts Institute of Technology
(1982)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1985

Signature of the Author .....
Department of Electrical Engineering
and Computer Science, May 30 ,1985

Certified by .....
Alan S. Willsky
Thesis Supervisor

Certified by .....
George C. Verghese
Thesis Co-Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Departmental Graduate Committee

AN ALGEBRAIC APPROACH TO

TIME SCALE ANALYSIS AND CONTROL

by

Xi-Cheng Lou

Submitted to the Department of
Electrical Engineering and Computer Science
on October 1, 1985,
in partial fullfillment of the requirement
for the Degree of Doctor of Philosophy

ABSTRACT

An algebraic approach is developed for multiple time scale decomposition of a linear system using the Smith structure of the system matrix viewed as the matrix of functions of a small parameter c. This derivation makes clear that both the necessary and sufficient multiple semi-stability (MSST) condition, which ensures well-defined multiple time scale behavior and the time-scale-decomposed system structure which approximates the original system are closely related to the so-called Schur complements of a certain matrix. Furthermore, this decomposition has been extended to a larger class of systems, satisfying the so-called multiple semi-simple nullstructure (MSSNS) condition.

The algebraic approach is also applied to examine the questions of the feedback control of the linear systems. Specifically we present results on time scale modifications by state feedback.

The characterization of the relationship among the eigenvalues of A(c), its invariant factors and the MSSNS and MSST conditions has been thoroughly studied. It is shown that the MSSNS condition is not only equivalent to the non-singularity of the Schur complements of certain matrix but also equivalent to 1) the eigenvalues and the invariant factors having the same orders and 2) a condition which exposes the relationship among the order of the gcd of all ixi principal minors, the order of their sum and the invariant factors. The MSST condition is equivalent to 1) the Schur complements being Hurwitz and 2) a condition

which reveals the connection between the MSSNS and MSST conditions and the eigenvalues of A(c).

Using the algebraic approach, a scaling procedure is developed to transform a system having no uniform time scale approximation to one which does. This procedure is applied to high-gain feedback problems.

Thesis Supervisor: Alan S. Willsky
Title: Professor of Electrical Engineering and Computer Science

Thesis Supervisor: George C. Verghese
Title: Associate Professor of Electrical Engineering and Computer Science

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to Professor Alan S. Willsky for his guidance and support throughout the course of this research; to Professor George C. Verghese for his valuable insights and constructive criticisms.

Special thanks is due to Prof. M. Vidyasagar whose suuggestions have helped to develop this research.

This work has also benefitted greatly from technical discussions with Prof. Gilbert Strang, Prof. Pamela G. Coxson and Prof. Sanjoy Mitter.

Finally, the support from my wife Kuang-Yi and my mother Pei-De Wang was invaluable for the entire course of the graduate program.

This work was supported in part by the Air Force of Scientific Research under Grant AFOSR-82-0258 and in part by the Army Research Office under Grant DAAG 29-84-k-0005.

TO MY MOTHER

CONTENTS

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1 General Description

The class of systems considered in this thesis are linear, time-invariant systems whose parameters are subject to a small perturbation. Mathematically, this can be stated through the system equation

$$\dot{x}(t) = A(\epsilon)x(t) + B(\epsilon)u(t) \qquad (1.1)$$

Here x and u are n- and m- dimensional state and control vectors respectively; and $A(\epsilon)$, $B(\epsilon)$ are nxn and nxm matrices whose entries are analytic at $\epsilon=0$.

As pointed out by Coderch and et.al. [1], [7], under certain conditions the undriven system

$$\dot{x}(t) = A(\epsilon)x(t) \qquad (1.2)$$

will exhibit multiple time scale behavior. Namely, the state variable x(t) can be "approximated" by a new variable z(t) whose components evolve at several time scales with time constants proportional to

$1, 1/\epsilon, 1/\epsilon^2, \ldots$. In other words, the undriven system which may have very large dimension can be broken down into several subsystems with lower dimensions, each of which focuses on different time scale.

The phenomenon just described is commonly referred to as a consequence of the singular perturbation of the system matrix $A(\epsilon)$. Specifically, as noted by Coderch [7], the system (1.2) may have several time scales only if $A(\epsilon)$ losses rank as ϵ goes to zero[+]. If the rank of $A(0)$ is equal to the rank of $A(\epsilon)$, then $A(\epsilon)$ is regularly perturbed and has only one time scale.

The value of singular perturbation analysis for the system (1.2) rests on the fact that it achieves model order reduction by separating the system's time scales, that is by considering slow and fast phenomena separately. Consequently, problems of analysis and control for systems with very large dimension may boil down to several problems of smaller dimension. In addition, there are situations in which the original system is not singularly perturbed but in which the control which is applied causes the overal closed-loop system to possess several time

---

[+] For most of the discussion in this thesis we will focus on the case in which $A(\epsilon)$ has full rank for $\epsilon \in (0, \epsilon_0]$ but is singular at $\epsilon=0$. In Section 2.8 we show how to extend our results to the case in which $A(\epsilon)$ itself is singular.

scales. For example, singularly perturbed system can result from the optimal control of a system using a quadratic cost functional having a small penalty on the control [2], [13]. The results in these references reveal the existance of multiple time scales in such systems.

There have been numerous papers dedicated to this subject (see survey [10],[22]) among which Kokotovic et.al. have thoroughly studied the two time scale case. Then Coderch, et.al. [1], [7] carried this idea to the multiple time scale case and derived some basic results. In this thesis we develop a new algebraic approach to multiple time scale analysis which allows us to obtain a clearer and deeper understanding of time scale decomposition for the general systems (1.1), (1.2). Not only does this approach allow us to gain more insight into the nature of systems with several time scales but it also provides a framework within which it is possible for us to consider and solve several other important problems. In this chapter we first briefly describe previous work which forms the foundation on which our research is built. We then give the outline of this thesis and summarize its contributions.

1.2 Background

The origin of the multiple time scale problem can be traced to the so-called boundary-layer problem in ordinary differencial equation theory where a small positive number ϵ is incorporated to allow

perturbations. Typically such a problem gives rise to a boundary layer, which is a narrow interval of time close to the origin where the solution of the differential equation changes rapidly. The thickness of the boundary layer approuches zero as $\epsilon \dashrightarrow 0$. Outside the boundary layer, in the outer region, the solution varies slowly. Therefore the system presents two time scales. A simple example will show this phenomenon. Consider the perturbed differential equation

$$\epsilon\ddot{x}(t) + (1+\epsilon)\dot{x}(t) + x(t) = 0, \quad x(0)=0, x(1)=1$$

The exact solution of this problem is

$$x(t) = \frac{e^{-t} - e^{-t/\epsilon}}{e^{-1} - e^{-1/\epsilon}}$$

Therefore this problem exhibits two time scales having time constants of order 1 (slow) and order ϵ (fast). In general, the boundary layer method is based on the fact that if a solution of a differencial equation is slowly varying except in isolated boundary layers, then it may be easy to obtain a leading order approximation without directly solving the equation.

Kokotovic and co-workers studied a special class of systems of singular perturbed linear differential equations [4], [5]

$$\begin{aligned} \dot{x}_1(t) &= A_{11}x_1(t) + A_{12}x_2(t) \\ \epsilon\dot{x}_2(t) &= A_{21}x_1(t) + A_{22}x_2(t) \end{aligned} \qquad (1.3)$$

If we change the argument t to

$$\tau = t/\epsilon$$

then we have

$$\begin{aligned} 1/\epsilon\dot{x}_1(\tau) &= A_{11}x_1(\tau) + A_{12}x_2(\tau) \\ \dot{x}_2(\tau) &= A_{21}x_1(\tau) + A_{22}x_2(\tau) \end{aligned}$$

Or

$$\dot{x} = \begin{bmatrix} \epsilon A_{11} & \epsilon A_{12} \\ A_{21} & A_{22} \end{bmatrix} x = A(\epsilon)x \qquad (1.4)$$

Thus, (1.3), or equivalently (1.4) is a special case of (1.2). If in addition A_{22} and $A_{11}-A_{21}A_{22}^{-1}A_{12}$ have all their eigenvalues strictly in the left-half plane, then system (1.2) has <u>well-defined two time scale behavior</u> (see below) and the eigenvalues of $A(\epsilon)$ will fall in two groups as ϵ approaches zero, one of order 1 the other of order ϵ.

In the previous subsection we gave a verbal, intuitive definition of what we mean by well-defined multiple time scale behavior, namely that the system can be decomposed into several subsystems, one at each time scale. The following is a precise statement of what we mean.

<u>Definition</u>:

System (1.2) <u>has well-defined multiple time scale behavior</u> if there

exist constant matrices A_0, A_1,...., A_{n-1}, T and integers $k_0 \leq k_1 \leq \ldots \leq k_{n-1}$ such that

$$\lim_{\epsilon \to 0} \sup_{t \geq 0} \| \exp\{A(\epsilon)t\} - T \exp\{\mathrm{diag}[\epsilon^{k_0} A_0, \epsilon^{k_1} A_1, \ldots \epsilon^{k_{n-1}} A_{n-1}]\} T^{-1} \| = 0$$

In this case we will frequently say that $[A_0,\ldots,A_{n-1};T]$ defines a time scale decomposition of (1.2).

It is clear from the definition, that if (1.2) has well-defined time scale behavior then the state variable x(t) can be approximated (after a linear transformation) by n components, each evolving at a different time scale. Furthermore, the approximation is uniformely valid on entire half-line $[0,\infty)$[+].

While Kokotovic considered the two time scale case, Coderch, et. al. [1], [7] presented the first complete solution of the general case (1.2). In this work they present necessary and sufficient conditions for the system (1.2) to have well-defined time scale behavior. Their proofs suggest a procedure for extracting and displaying the multiple

---

[+] As point out by Coderch [7], exp{A(0)t} is a good approximation for exp{A(ε)t} on any finite interval [0,T]. Therefore the notion of multiple time scale makes sense only when the concern is the whole interval $[0,\infty)$.

time scale structure of (1.2). However this method is quite involved and does not lend itself to easy interpretation or computation. For example, the relationship between the complex results of [1], [7] and intuitively simple results of [4], [5] is not at all apparent. Making clear this relationship and obtaining a conceptually and computationally simple solution in the general case are two of the objectives achieved in this thesis.

1.3 Outline of This Thesis

The first portion of the research described in this thesis deals with an algebraic approach to the time scale decomposition of (1.2). Specifically, in Chapter 2 we consider the Smith form of $A(\epsilon)$ over the ring T of all functions of ϵ which are analytic at $\epsilon=0$, and based on this form we are able to obtain a more direct and simple description of the multiple time scale decomposition. This derivation makes clear that both the necessary and sufficient multiple semi-stability (MSST) condition, which ensures well-defined multiple time scale behavior and the time-scale-decomposed system structure which approximates the original system are closely related to the so-called Schur complements of a certain matrix. In doing this we also are able to make clear the connection between the general results of Coderch et. al. and those of Kokotovic. Furthermore, having established this framework we are able to use it to solve several additional problems. In particular we are

able to extend our decomposition (with a modified notion of well-defined time scale behavior) to a larger class of systems, satisfying the so-called multiple semisimple nullstructure (MSSNS) condition, and, at the end of Chapter 2, we use our results to examine the questions of the feedback control of the system (1.1). Specifically we present results on time scale modification by state feedback.

Chapter 3 deals with the characterization of the relationship among the eigenvalues of $A(\epsilon)$, its invariant factors and the MSSNS and MSST conditions. It is shown in this chapter that the MSSNS condition is not only equivalent to the non-singularity of the Schur complements of a certain matrix but it is also equivalent to 1) the eigenvalues and the invariant factors having the same orders and 2) a condition which exposes the relationship among the orders of the gcd of all ixi principal minors, the order of their sum, and the invariant factors. These results provide us with a procedure for computing the orders of the eigenvalues and invariant factors and for checking the MSSNS condition. Also, in this chapter we show that the MSST condition is equivalent not only to the Schur complements of a certain matrix being Hurwitz (Chapter 2) but also to a condition on the orders of the real and imaginary parts of the eigenvalues of $A(\epsilon)$, which in turn reveals the connection between the MSSNS and MSST conditions.

In Chapter 4 we use our results on the relationship between the orders of eigenvalues and invariant factors to explore the use of amplitude scaling to transform a system matrix without MSSNS into one

that does have this property. The analysis involved in this investigation is rather delicate as it involves careful examination of the orders of principal minors and the identification of key elements of the matrix that must be scaled. The end result of our efforts is a procedure for determining such a scaling matrix for systems satisfying certain conditions. We then apply our results to interpret and generalize recent results on time scale analysis of high gain or nearly singular optimal feedback systems.

Finally, in Chapter 5 we briefly summarize the main results of the thesis and discuss several directions for further research.

1.4 Main Contributions of This Thesis

We feel the main contributions of this thesis are the following.

1. We present a simpler and clearer picture of the multiple time scale decomposition of a general perturbed linear system based on an algebraic approach which allows further development in several directions.

2. We make clear the connection between Kokotovic's explicit approach for the two time scale decomposition case and Coderch's

elaborate multiple time scale results.

3. We thoroughly study the MSSNS and MSST conditions relating them to the algebraic structure of $A(\epsilon)$. By exposing the important role that the orders of the invariant factors and eigenvalues play, the interpretation of those conditions is clarified.

4. We develop an algorithm to extend time scale decompositions to a larger class of systems which satisfy MSSNS but not MSST.

5. We present results on time scale assignment through state feedback control.

6. We make clear the relationship between the MSSNS condition and the MSST condition. Specifically, $A(\epsilon)$ satisfies the MSST condition if and only if it satisfies the MSSNS condition and the orders of the real parts of its eigenvalues are equal to or less than those of the corresponding imaginary parts.

7. We reveal the role of the gcd's of the <u>principal minors</u> (not <u>all</u> minors as in the general case) in determining the invariant factors of a system with MSSNS and develop an algorithm similar to that of the so-called Newton polygon. We show that if the system has well-defined time scale behavior (or more generally if $A(\epsilon)$ just has MSSNS) then this algorithm determines the orders of the various time scales and the dimensions of the subsystem at each time scale.

8. We develop a procedure for amplitude scaling to transform a

system which does not satisfy the MSSNS condition to a system which does. If the original system satisfies some conditions studied in this thesis, then after scaling the resulting system will have well-defined multiple time scale behavior.

9. We apply the scaling procedure developed in this thesis to high-gain feedback problems, leading to an interpretation and generalization of results in the literature.

CHAPTER 2

ALGEBRAIC ASPECTS OF TIME SCALE BEHAVIOR

2.1 INTRODUCTION

As we pointed out in Chapter 1, the system we shall consider in this thesis is a perturbed, linear, time-invariant system

$$\dot{x}(t) = A(\epsilon)x(t) + B(\epsilon)u(t) \tag{2.1}$$

and its undriven form

$$\dot{x}(t) = A(\epsilon)x(t) \tag{2.2}$$

where $A(\epsilon)$ and $B(\epsilon)$ are nxn and nxm matrices whose entries are functions of a small parameter ϵ analytic at $\epsilon=0$. Kokotovic and co-worker [6, 10] have thoroughly studied a special case of (2.2) in the two time scale case, which, as we discussed in Chapter 1 Eq. (1.4), corresponds to $A(\epsilon)$ having the special form

$$A(\epsilon) = \begin{bmatrix} \epsilon A_{11} & \epsilon A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

with $A_{11}-A_{12}A_{22}^{-1}A_{21}$ Hurwitz.

Coderch et. al. [7,1] studied the general case of (2.2) and derived necessary and sufficient conditions for the system (2.2) to have a well-behaved multiple time scale description. In Section 2.2 we give a more detailed review of their work and point out the limitations of their results which have motivate our work.

In particular, in this chapter we develop an algebraic approach for determining the multiple time scale structure of (2.2). This method involves the examination of the Smith form of $A(\epsilon)$ over a particular local ring T. In Section 2.3 we give the definition and basic feature of this ring. In Section 2.4, we introduce the so-called explicit form of the perturbed system (2.2) which is closely related to the Smith form of $A(\epsilon)$. We also show that after some invertible linear transformations, the solution of this explicit form is a good approximation (in an asymptotic sense) to that of system (2.2) if in fact the original system has well-defined multiple time-scale behavior (a property we define in the sequel). Moreover, using the explicit form we can define a straightforward procedure to check if a system has well-defined time scale behavior. At the end of the section we provide an overview of the major results along these lines that are developed in

the following sections.

In Section 2.5, we give a first derivation of the multiple time scale approximation. This derivation is in fact an extension of the usual two time scale argument. Then in Section 2.6 a proof of the asymptotic properties of this approximation is given based on making explicit the connection with Coderch's results. It turns out that the time-scale-seperated system which approximates the original system is determined by a sequence of Schur complements of the system matrix in explicit form. A major consequence of this is that the computational procedure we derive is far more transparent than that proposed in [7]. The eqivalence between the results obtained in Section 2.5 and 2.6 is established in Section 2.7. This development makes clear the relationship between Kolotovic's two time scale result and Coderch's multiple time scale result. Finally, in Section 2.8 we review our results in order to place our contribution in its proper perspective. Specifically our approach establishes a framework that not only exposes the essential nature of time scale decompositions in a very clear fashion, thereby improving our understanding of such decompositions, but also provides a starting point for posing and solving a variety of problems that are not so easily posed or solved using previously developed approaches. Chapters 3 and 4 contain several important results of this type, as does the end of Section 2.8 in which we solve several problems including one that had been proposed, but not solved, by Coderch [7].

2.2 PREVIOUS WORK

As pointed out in Section 2.1, not only is our work closely related to Kokotovic's two time scale results and Coderch's miltiple time scale approach, but it also establishes the clear and simple relationship between these earlier results. Therefore at the start it is important that we describe their work in some depth. In this section we give an outline of their work and point out several issues.

2.2.1 Two Time Scale Results

The two time scale singular perturbation method developed by Kokotovic, et.al [6.10] is based on a linear time invariant system with a small constant c on the left hand side of its state equation:

$$\begin{bmatrix} \dot{x}_1 \\ \epsilon\dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11}x_1+A_{12}x_2 \\ A_{21}x_1+A_{22}x_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad (2.3)$$

There are two well-known results about this system for small c.

1. If A_{22} and $\tilde{A}_{11}=A_{11}-A_{12}A_{22}^{-1}A_{21}$ are <u>non-singular</u>, then the eigenvalues of (2.3) fall in two groups as $\epsilon \to 0$. One group approaches the eigenvalues of $\tilde{A}_{11}$. The other approaches that of ϵA_{22}. The matrix

$\tilde{A}_{11}$ is commenly referred to as the Schur complement of A_{22} in the matrix A [3]. We will have much more to say about Schur complements as we develop our approach in following sections.

2. If A_{22} and $\tilde{A}_{11}$ are Hurwitz, then the system has "well-defined time-scale structure" as ϵ-->0. The fast time scale is of order 1 and the slow time scale of order $1/\epsilon$.

Mathematically, what "well-defined time-scale-structure" means can be explained as follows.

Define

$$T = \begin{bmatrix} I & 0 \\ -A_{22}^{-1} & I \end{bmatrix} \tag{2.4}$$

and

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = T \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

Then under the condition that A_{22}, A_{11} are Hurwitz, we have

$$\begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{bmatrix} = \begin{bmatrix} y_s(t) \\ y_f(t/\epsilon) \end{bmatrix} + O(\epsilon), \qquad 0 \le t \tag{2.5}$$

where the $O(\epsilon)$ term is bounded uniformly in t on $[0, \infty)$ as ϵ-->0 and

$y_s(t)$ and $y_f(t/\epsilon)$ are called the "slow" and "fast" states respectively and satisfy

$$\dot{y}_s(t) = \tilde{A}_{11}y_s(t), \quad y_s(0) = x_1(0)$$
$$\dot{y}_f(t) = A_{22}y_f(t), \quad y_f(0) = x_2(0)+A_{22}^{-1}A_{21}x_1(0) \qquad (2.6)$$

If we define

$$A(\epsilon) = \begin{bmatrix} A_{11} & A_{12} \\ \epsilon^{-1}A_{21} & \epsilon^{-1}A_{22} \end{bmatrix}$$

in view of (2.4), (2.5) and (2.6), we have

$$\lim_{\epsilon\to 0} \sup_{t\geq 0} ||T\exp\{A(\epsilon)t\}T^{-1} - \exp\{\mathrm{diag}[\tilde{A}_{11}, 1/\epsilon A_{22}]t\}||=0$$

Furthermore, if we change the time scale to $\tau=t/\epsilon$, we have

$$\begin{bmatrix} \dot{x}_1(\tau) \\ \dot{x}_2(\tau) \end{bmatrix} = \begin{bmatrix} \epsilon A_{11} & \epsilon A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1(\tau) \\ x_2(\tau) \end{bmatrix} = A(\epsilon) \begin{bmatrix} x_1(\tau) \\ x_2(\tau) \end{bmatrix}$$

Then similarly,

$$\lim_{\epsilon\to 0} \sup_{\tau\geq 0} ||T\exp\{A(\epsilon)\tau\}T^{-1} - \exp\{\mathrm{diag}[\epsilon\tilde{A}_{11}, A_{22}]\tau\}||=0 \qquad (2.7)$$

In general, as we defined in Chapter 1, if $A(\epsilon)$ is analytic at $\epsilon=0$

and Hurwitz for $\epsilon>0$ and there exist constant matrices A_1, A_2, T such that uniform convergence as in (2.7) holds, then we say that $A(\epsilon)$ has well-defined two time scale behavior. Obviously Kokotovic's results provide a two time scale decomposition of a linear system. Moreover, as mentioned earlier, this approach has been proven to be useful in optimal control [4], stochastic control [8], design of nonlinear regulators [5], analysis of high gain systems [13], cheap control [2, 12, 13] and so on. The limitations of this method consist of two points. First, the assumed form of Eq. (2.3) essentially implies that the seperation of fast and slow state variables has been done beforehand. This is, however, not the general case, especially for complex systems, where the system matrix could be a more general function of ϵ, and state seperation would not be immediately available. Secondly, only two time scales are considered in this model. In the next subsection we introduce the result of Coderch, et.al [7, 1] which successfully solve the general problem of multiple time scale behavior of a perturbed linear system.

2.2.2. Multiple Time Scale Behavior of Singularly Perturbed LTI Systems

The system considered in [7,1] is the singularly perturbed LTI system (2.2) with a slight change in notation whose purpose will become clear shortly.

$$\dot{x}(t) = A_0(\epsilon)x(t) \qquad (2.8)$$

where the nxn matrix $A_0(\epsilon)$ is an analytic function of ϵ. For the present discussion we assume that $A_0(\epsilon)$ is Hurwitz for $\epsilon\in(0,\epsilon_0)$ although Cocerch allows slightly weaker assumptions. In Section 2.8 we will show how our approach can be easily extended to this more general case.

In order to present the results in [7,1], we need several definitions. A matrix H is said to have <u>semisimple nullstructure</u> (SSNS) if its zero eigenvalue has geometric multiplicity equal to its algebraic multiplicity. In other words, every zero eigenvalue of H corresponds to a distinct independent eigenvector. A matrix H is said to be <u>semistable</u> if it has SSNS and all its non-zero eigenvalues have negative real parts. Suppose $A_0(\epsilon)$ has eigenvalues $\lambda_1(\epsilon),.....,\lambda_n(\epsilon)$ where $\lambda_i(\epsilon)->0$, $\epsilon->0$, $i=1,....,m\leq n$. Then the <u>total projection</u> for the zero-group of eigenvalues of $A_0(\epsilon)$, $P_0(\epsilon)$ is the projection onto the subspace spanned by eigenvenctors corresponding to $\lambda_1(\epsilon),....,\lambda_m(\epsilon)$ of $A_0(\epsilon)$ [9].

Since $A_0(\epsilon)$ is analytic at $\epsilon=0$, it has a series expansion of the form

$$A_0(\epsilon) = \sum_{p=0} \epsilon^p F_{0p}$$

It can be proven [7,1] that if F_{00} has SSNS, the matrix

$$A_1(\epsilon) = P_0(\epsilon)A_0(\epsilon)/\epsilon$$

has a series expansion of the form

$$A_1(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p F_{1p}$$

If F_{10} also has SSNS we define $A_2(\epsilon)$ as

$$A_2(\epsilon) = P_1(\epsilon)A_1(\epsilon)/\epsilon = P_1(\epsilon)P_0(\epsilon)A_0(\epsilon)/\epsilon^2$$
$$= \sum_{p=0}^{\infty} \epsilon^p F_{2p}$$

where $P_1(\epsilon)$ is the total projection for the zero-group of eigenvalues of $A_1(\epsilon)$. This process can be continued but it terminates at

$$A_m(\epsilon) = P_{m-1}(\epsilon)A_{m-1}(\epsilon)/\epsilon = P_{m-1}(\epsilon)....P_0(\epsilon)A_0(\epsilon)/\epsilon^m$$
$$= \sum_{p=o}^{\infty} \epsilon^p F_{mp}$$

if the matrix F_{m0} does not have SSNS or if
$\text{rank}F_{00}+\text{rank}F_{10}+...+\text{rank}F_{m0}=n$.

A matrix $A_0(\epsilon)$ is said to satisfy the <u>multiple semisimple null structure</u> (MSSNS) condition if the sequence of matrices $A_k(\epsilon)$ can be contructed up to a stage $k=m$ with all matrices F_{k0}, $k=0,....,m$ having SSNS and $\text{rank}F_{00}+....+\text{rank}F_{m0}=n$. If in addition, all F_{k0} are semistable, then we say that $A_0(\epsilon)$ satisfies <u>multiple semistability</u> (MSST) condition.

The following results determines when an asymptotic approximation to $x(t)$, uniformly valid for $t\geq 0$, can be constructed which clearly displays the multiple time scale behavior of $x(t)$. The main result is the following.

1. System (2.8) has well-defined time-scale behavior if and only if $A_0(\epsilon)$ satisfies MSST condition.

2. If $A_0(\epsilon)$ satisfies the MSST condition. Then

$$\lim_{\epsilon->0} \sup_{t\geq 0} ||\exp\{A(\epsilon)t\} - \exp\{\sum_{k=0}^{m} F_{k0}\epsilon^k t\}|| = 0$$

Furthermore, it can be shown that using a linear transformation T, which is independent of ϵ , we have that

$$T^{-1}F_{k0}T = \text{diag}\{0,0,\ldots,0,\tilde{A}_k,0,\ldots,0\}, \quad k=0,\ldots,m$$

where each of the $\tilde{A}_k$ is Hurwitz. Therefore

$$\lim_{\epsilon->0} \sup_{t\geq 0} ||\exp\{A_0(\epsilon)t\} - T^{-1}\exp\{\sum_{k=0}^{m} TF_{k0}T^{-1}\epsilon^k t\}T|| = 0$$

or

$$\lim_{\epsilon->0} \sup_{t\geq 0} \exp\{TA_0(\epsilon)T^{-1}\} - \text{diag}\{\exp[\tilde{A}_0 t],\ldots,\exp[\tilde{A}_m\epsilon^m t]\} = 0 \qquad (2.10)$$

In other words, to first order approximation, the original system (2.8) can be thought of as being composed of (m+1) uncoupled subsystems (after certain ϵ-independent transformations)

$$d\tilde{y}_k(t)/dt = \tilde{A}_k\tilde{y}_k(t) \quad , \quad k=0,\ldots,m$$

each running at a different time scale. In this case we will say that $(T; \tilde{A}_0, \ldots, \tilde{A}_m)$ defines a time scale decomposition of (2.8).

From this result it can be seen that in order to obtain the uncoupled approximation one must first compute F_{k0}, $k=0,\ldots,m$, then compute T. Although a procedure was proposed in [7], it really should not be viewed as an algorithm, since the computational aspects of the procedure have not been examined, and the procedure is quite involved. For example, F_{10}, F_{20}, F_{30} can be obtained by

$$F_{10} = P_0A_{01}P_0$$

$$F_{20} = P_1P_0(A_{02}-A_{01}A_{00}{}^{\#}A_{01})P_0P_1$$

$$\begin{aligned} F_{30} = P_2P_1P_0(&A_{03}-A_{01}A_{00}{}^{\#}A_{02}-A_{02}A_{00}{}^{\#}A_{01} \\ &+A_{01}A_{00}{}^{\#}A_{01}A_{00}{}^{\#}A_{01}-A_{02}A_{10}{}^{\#}A_{02} \\ &+A_{02}A_{10}{}^{\#}A_{01}A_{00}{}^{\#}A_{01}-A_{01}A_{00}{}^{\#}A_{01}A_{10}{}^{\#}A_{02} \\ &-A_{01}A_{00}{}^{\#}A_{01}A_{10}{}^{\#}A_{01}A_{00}{}^{\#}A_{01})P_0P_1P_2 \end{aligned}$$

where $P_i=\lim_{\epsilon\to 0}P_i(\epsilon)$, $i=0,1,2$, and $A^{\#}$ is the pseudo-inverse of matrix A. The computation of F_{40} is much more complicated. The complexity of these formulas makes it difficult to obtain a deep understanding of multiple time scale behavior and the structure of $A_0(\epsilon)$ or to examine a variety of problems such as the consideration of the effect of control on time scale behavior.

Having these background results we are now in a position to

introduce our main results. As we mentioned in Chapter 1 and earlier in this chapter, an important element in our development is the Smith form of $A(\epsilon)$ over a local ring T of all functions of ϵ analytic at $\epsilon=0$. In next section we give the definition and the major properties of this ring.

2.3 The Local Ring T

In this section we study the ring T of all functions of real variable ϵ which are analytic at $\epsilon=0$. In other words, we study the ring of functions which have Taylor series expansions at $\epsilon=0$. We show that this ring have a Smith form. It is this Smith form that plays an important role in this thesis. We state many of the results concering T and matrices over T without proof because they are immediate extensions or examples of results which can be found in literature [3,11].

Consider a set T which consists of all functions of ϵ which are analytic at $\epsilon=0$. It is easy to show that T is a ring. The units of T are elemlnts of T which do not vavish at $\epsilon=0$,

$$U = \{ x \mid x \in T,\ x(0) \neq 0 \}$$

Therefore any unit of T has the form

$$u = a_0 + a_1\epsilon + a_2\epsilon^2 \ldots .$$

with $a_0 \neq 0$. Define the degree of x, a(x), to be the order of the first nonzere term in its Taylor expansion. (For example, if $x=a_2\epsilon^2+a_3\epsilon^3+....$ and $a_2 \neq 0$, then the degree of x is two). Then T is a Euclidean domain with degree function a(x). If for some x, a(x)=i, then x equals ϵ^i times a unit. In other words, each x is equivalent (modulo units) to one of the element of set $\{1,\epsilon,\epsilon^2,.....\}$. Let f and g be two elements of T. Then it is easy to see that

1. f divides g iff $a(f) \leq a(g)$.

2. f and g are coprime iff at least one of them is not equal to zero at $\epsilon=0$.

Let M(T) denote the set of matrices whose elements belong to T. Then the set of unimodular matrices is defined as

$$U(T) = \{G | G \in M(T), \quad |G(0)| \neq 0\}$$

Because T is a Euclidean domain, it is a principle ideal domain. Therefore any matrix $A \in M(T)$ can be transformed to its Smith form. Namely for any matrix $A \in M(T)$ we have

$$A=PDQ$$

where $P,Q \in U(T)$ and D has the form

$$D = \text{diag } \{\epsilon^{i1},...., \epsilon^{ir}, 0,....,0 \}$$

where the integers i_j are ordered so that $0 \leq i_1 \leq i_2 \leq .... \leq i_r$. The

quantities ϵ^{i1}, ... are called the invariant factors of $A(\epsilon)$. The elementary row and column operations used in bringing A into its Smith form have the form

1. Interchange of any two columns (or rows).

2. Addition to any column (row) of any other column (row) multiplied by an element in T.

3. Scaling any column (row) by any element in U(T).

Let $A, B \in M(T)$ have the same number of rows. Then A and B are left coprime

iff $F=[A\ B]$ has the Smith form $P[I\ 0]Q$.

iff F has full row rank at $\epsilon=0$.

It can be shown that this ring is a local ring because its maximal ideal, namely $\langle\epsilon\rangle$ is unique.

As pointed out in Chapter 1, system (2.2) can be approximated by its explicit form. Further work on our algebraic approach is bassed on this form. Therefore in next section we derive the explicit form of (2.2) using the Smith form of $A(\epsilon)$.

2.4. THE EXPLICIT FORM FOR A SINGULARLY PERTURBED AUTONOMOUS LINEAR SYSTEM

2.4.1. Introduction

In this section we show how a general system (2.2), with $A(\epsilon)$ being a matrix over T, can be put into what we term its explicit form. As mentioned before, it is the explicit form that makes it possible to connect the Smith form of $A(\epsilon)$ with the multiple time scale behavior of the system. This provides considerable insight into the structure of such systems. Also, the explicit form permits us to develop an algorthm for construting the time scale decomposition of (2.2) that makes use of Schur complements and that makes clear the computations required to determine the time-scale decomposition. This form also allows us to pose and answer a variety of questions in subsequent chapters and sections of this chapter.

2.4.2. Explicit Form

As mentioned in Section 2.3, an nxn matrix $A(\epsilon)$ over T has its Smith form

$$A(\epsilon) = P(\epsilon)D(\epsilon)Q(\epsilon)$$

where $P(\epsilon)$ and $Q(\epsilon)$ are unimodular, namely $|P(0)| \neq 0$, $|Q(0)| \neq 0$ and $D(\epsilon)$ is

$$D(\epsilon) = \begin{bmatrix} d_1(\epsilon) & & & 0 \\ & \ddots & & \vdots \\ & & d_r(\epsilon) & \\ 0 & \cdots & & 0 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r \\ \left.\vphantom{\begin{matrix}a\\a\end{matrix}}\right\} n-r \end{matrix}$$

where $d_j(\epsilon)=\epsilon^{i_j}$, $j=1,\ldots,r$, and $0\leq i_j\leq i_k$, $j\leq k$.

consequently we can rewrite the system (2.2) as

$$\dot{x} = P(\epsilon)D(\epsilon)Q(\epsilon)x \qquad (2.11)$$

Because $P(\epsilon)$ is unimodular, $P^{-1}(\epsilon)$ exists in the neighbourhood of $\epsilon=0$. Later on we shall see that we can use $P(\epsilon)$ as a similarity transformation on the state without affecting the time scales of the system. In particular, multiplying by $P^{-1}(\epsilon)$ on the left of both sides of Eq. (2.11) we have

$$P(\epsilon)^{-1}\dot{x} = D(\epsilon)Q(\epsilon)P(\epsilon)P(\epsilon)^{-1}x$$

Let

$$y = P(\epsilon)^{-1}x$$

Then

$$\dot{y} = D(\epsilon)Q(\epsilon)P(\epsilon)y$$

If we define

$$\bar{A}(\epsilon) = Q(\epsilon)P(\epsilon)$$

we have

$$\begin{aligned}\dot{y} &= D(\epsilon)\bar{A}(\epsilon)y \\ &= \text{diag}\{ I_1, \epsilon I_2, \ldots, \epsilon^{n-1} I_{n-1} \}\, \bar{A}(\epsilon)\end{aligned} \qquad (2.12)$$

Eq. (2.12) is called the explicit form because $D(\epsilon)$ explicitly reveals the time scales of thes system. Recall that $Q(\epsilon)$ and $P(\epsilon)$ are unimodular, and therfore so is $A(\epsilon)$.

In the next section we shall consider (2.12) and find, under a particular set of conditions, the time scale-decomposed approximation of y (after appropriate similarity transformations). Before we do that, let us overview some of the major results to be devoloped in the following sections.

2.4.3 An Overview

What we are interested in doing is investigating the time scale

decomposition of (2.2), or equivalently of (2.12). Coderch's approach (Section 2.2.2) provides a general method for doing this involving the computation of the total projection for the zero group of eigenvalues of $A(\epsilon)$ and of each of the subsequent system matrices $A_1(\epsilon)$, $A_2(\epsilon)$, ... defined in this procedure. All of these computations involve using the complete ϵ-dependent projection matrices. On the other hand, as we will discuss, if one has a system in a form analogous to that of Kokotovic's treatment (Section 2.2.1), the computation of these projections is straightforward and transparent, as is the check of the MSSNS and MSST conditions. What our approach does is to transform the system so that this straightforward construction can be applied. In this process, we in fact throw away certain parts of the ϵ-dependency of the system matrices that are unimportant in obtaining a time scale decompositions if in fact the original system <u>has</u> a time-scale decomposition. In a sense what this does is to minimize the number of ϵ-dependent computations that must be performed, thereby making far more clear what the critical ϵ-dependencies are in $A(\epsilon)$, its total projection onto the zero group, and those of its successors ($A_1(\epsilon)$, $A_2(\epsilon)$, etc.).

To be more specific, let us assume, both here and in the next three sections that $A(\epsilon)$ in (2.2) is Hurwitz for $\epsilon>0$. Now, if we obtain a time-scale decomposition for y in (2.12), it is straightforward to obtain a time-scale decomposition for x in (2.2) using an <u>ϵ-independent</u> similarity transformation. Specifically,

Lemma 2.1:

Suppose that $(T; \tilde{A}_0, \ldots, \tilde{A}_m)$ defines a time-scale decomposition of the explicit form dynamics (2.12). Then $(TP^{-1}(0); \tilde{A}_0, \ldots, \tilde{A}_m)$ defines a time-scale decomposition of the original system (2.2).

Proof:

It is straightforward to check that this lemma states is equivalent to the following. Note that

$$x(t) = P(\epsilon)y(t)$$

Define

$$x_1(t) = P(0)y(t) \tag{2.13}$$

Then

$$\limsup_{\epsilon \to 0} \sup_{t \geq 0} ||x(t) - x_1(t)|| = 0 \tag{2.14}$$

That is, (2.13) is an asymtotically accurate approximation of the solution of (2.2).

To proceed, note that

$$||x(t) - x_1(t)|| \leq ||P(\epsilon) - P(0)|| \, ||y(t)||$$

So that (2.14) will be proved if we can show that

$$\limsup_{\epsilon \to 0} \sup_{t \geq 0} ||y(t)|| = M < \infty$$

However, by hypothesis

$$y(t) = T^{-1}\text{diag}\{\exp[\tilde{A}_0 t], \ \ldots \ , \ \exp[\tilde{A}_m \epsilon^m t]\}Ty(0)$$

where each of the $\tilde{A}_i$ is Hurwitz[+]. Since

$$\sup_{t\geq 0} \ ||\exp[\tilde{A}_i \epsilon^i t]|| \ = \sup_{t\geq 0} \ ||\exp[\tilde{A}_i t]|| \ < \infty$$

the result is proved.

Note that one consequence of the lemma is that (2.2) has a time-scale decomposition if and only if (2.12) does. However, while this lemma tells us a little bit about those ϵ-dependencies that can be thrown out, there is far more that can be said.

Specifically, recall that $\bar{A}(\epsilon)$ in (2.12) is unimodular. Intuitively, what this means is that $\bar{A}(\epsilon)$ has no "structure at ϵ=0". For notational simplicity, let us denote A(0) as A, and consider the system

$$\dot{z} = D(\epsilon)\bar{A}z \qquad (2.15a)$$

Let us also write $\bar{A}$ In block form compatible with the diagonal block sizes of $D(\epsilon)$

[+] That the $\tilde{A}_i$ are Hurwitz follows from the assumption that $A(\epsilon)$ is Hurwitz for $\epsilon\in(0,\epsilon_0]$.

$$\bar{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots\cdots & A_{1m} \\ A_{21} & A_{22} & \cdots\cdots & A_{2m} \\ & \cdots\cdots\cdots\cdots & & \\ A_{m1} & A_{m2} & \cdots\cdots & A_{mm} \end{bmatrix}$$

So that

$$\dot{z} = \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \cdot \\ \cdot \\ \dot{z}_m \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \cdots\cdots & A_{1m} \\ \epsilon A_{21} & \epsilon A_{22} & \cdots\cdots & \epsilon A_{2m} \\ & \cdots\cdots\cdots\cdots & & \\ \epsilon^m A_{m1} & \epsilon^m A_{m2} & \cdots\cdots & \epsilon^m A_{mm} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ z_m \end{bmatrix} \qquad (2.15b)$$

is in a form very similar to the form (2.7) considered by Kokotovic. We call this the <u>reduced explicit</u> form.

As we saw in Section 2.2.1, the Schur complement play an important role in defining the time-scale decomposition of the system considered by Kokotovic. A similar statement is true here. Specifically, in the next three sections we define a sequence of matrices $\tilde{A}_0, \ldots, \tilde{A}_m$ and a similarity transformation T obtained by successive Schur complementation of $\bar{A}$, and we provide two derivations, one based on the approach of Kokotovic, and a second making explicit contact with the results of Coderch, of the following (again under the assumption that $A(\epsilon)$ is Hurwitz for $\epsilon>0$):

Theorem 2.1:

Consider the original system (2.2) and the explicit form system (2.12). These systems have well-behaved time-scale decompositions if and only if the reduced explicit form system does. Furthermore, these well-behaved time scale decompositions exist *if and only if* $\tilde{A}_0, \ldots, \tilde{A}_m$ mentioned above all exist and are Hurwitz. In this case $(T; \tilde{A}_0, \ldots, \tilde{A}_m)$ defines a time-scale decomposition for *both* (2.12) and (2.15), and (by lemma 2.1), $(TP^{-1}(0); \tilde{A}_0, \ldots, \tilde{A}_m)$ defines one for the original system (2.2).

Let us make several important comments about this result. Specifically, once we have determined $P(0)$ and $\bar{A}=Q(0)P(0)$, *all of the remaining calculations* are *ϵ-independent*. Thus, we have identified the critical ϵ-dependent computations as the determination of $P(0)$ and $Q(0)$ in the Smith decomposition of $A(\epsilon)$. Finally, note that one point of the Theorem is that if (2.15) has a well-behaved time scale decomposition, then

$$\lim_{\epsilon \to 0} \sup_{t \geq 0} \| y(t) - z(t) \| = 0$$

However, as is illustrated in Section 2.7, this *need* not be true if a well-behaved time scale decomposition does *not* exist. In fact, while $A(\epsilon)$ is Hurwitz for $\epsilon>0$ by assumption, and $D(\epsilon)\bar{A}(\epsilon) = P^{-1}(\epsilon)A(\epsilon)P(\epsilon)$ is obviously Hurwitz as well, $D(\epsilon)\bar{A}$ need *not* be Hurwitz for $\epsilon>0$. Thus, if the original system has a well-behaved time scale decomposition, we can

throw away the ϵ-dependent terms in $\bar{A}(\epsilon)$. However, if such a time-scale decomposition does not exist, the ϵ-dependent terms of $\bar{A}(\epsilon)$ represent <u>critical</u> components of the damping in the original system (2.2). In Section 2.8 we use this observation to define time scale decompositions in a slightly weaker sense by pinpointing and keeping these ϵ-dependent terms in $\bar{A}(\epsilon)$ that <u>are</u> critical to system stability. This construction, posed as an open problem by Coderch, is but one example of the problems that our framework allows us to solve.

2.5. DERIVATION FROM TWO TIME SCALE RESULTS

2.5.1. Introduction

As we stated in the last section (Section 2.4.3), finding the time scale decomposition of (2.2) is equivalent to finding that of (2.12). In this section we shall show that if the successive Schur complements $\tilde{A}_{ii}$ are Hurwitz then $(T; \tilde{A}_{11}, \ldots, \tilde{A}_{mm})$ defines a time scale decomposion for both (2.12) and (2.15), and $(TP^{-1}(0); \tilde{A}_{11}, \ldots, \tilde{A}_{mm})$ defines one for the original system (2.2). The proof is in fact an extension of the well-known method used in the two time scale case. Then in the next section we shall relate our approach to the multiple time scale results obtained by Coderch et.al, and this will allow us to prove the full version of Theorem 2.1, that is, that the $\tilde{A}_{ii}$ being Hurwitz is <u>necessary</u> as well as sufficiant for the system (2.2) to have a well-defined time-scale decomposition.

2.5.2. Derivtion From Two Time Scale Results

The basic idea behind this approach is to block-diagonalize $A_1(\epsilon)=D(\epsilon)\bar{A}(\epsilon)$. It uses well known two time-scale results repeatedly to "peel off" each time scale. So first let us review some of those results. Suppose the system we are considering has the following form

$$\dot{x} = A(\epsilon)x$$

where

$$A(\epsilon) = \begin{bmatrix} A_{11} & A_{12} \\ \epsilon A_{21} & \epsilon A_{22} \end{bmatrix}$$

It is proven [17] that if we define

$$T(\epsilon)=\begin{bmatrix} I & -L(\epsilon) \\ \epsilon H(\epsilon) & I-\epsilon H(\epsilon)L(\epsilon) \end{bmatrix} \tag{2.16}$$

where H and L satisfy

$$A_{12}(\epsilon)-A_{11}(\epsilon)L-\epsilon L(A_{22}(\epsilon)+A_{21}(\epsilon)L) = 0 \tag{2.17}$$

$$H(A_{11}(\epsilon)-LA_{21}(\epsilon))-\epsilon(A_{22}(\epsilon)+A_{21}(\epsilon)L)H+\epsilon A_{21}(\epsilon)=0 \tag{2.18}$$

We have

$$T(\epsilon)\begin{bmatrix} A_{11}(\epsilon) & A_{12}(\epsilon) \\ \epsilon A_{21}(\epsilon) & \epsilon A_{22}(\epsilon) \end{bmatrix} T^{-1}(\epsilon)=\begin{bmatrix} G_1(\epsilon) & 0 \\ 0 & G_2(\epsilon) \end{bmatrix} \tag{2.19}$$

where

$$\begin{aligned} G_1(\epsilon)&=A_{11}(\epsilon)-\epsilon L(\epsilon)A_{21}(\epsilon) \\ G_2(\epsilon)&=A_{22}(\epsilon)+A_{21}(\epsilon)L(\epsilon) \end{aligned} \tag{2.20}$$

Solution to (2.17) and (2.18) exist as long as $A_{11}(0)$ is invertible and ϵ is sufficiently small. Note that under these conditions $T(\epsilon)$ is unimodular and, as $\epsilon \to 0$, $L(\epsilon) \to -A_{11}^{-1}(0)A_{12}(0)$, $H(\epsilon) \to 0$.

Now, consider our system in explicit form (Eq. (2.12))

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \vdots \\ \dot{z}_m \end{bmatrix} = \begin{bmatrix} A_{11}(\epsilon) & \cdots\cdots & A_{1m}(\epsilon) \\ \epsilon^{i2}A_{21}(\epsilon) & \cdots\cdots & \epsilon^{i2}A_{2m}(\epsilon) \\ & \cdots\cdots\cdots\cdots & \\ \epsilon^{im}A_{m1}(\epsilon) & \cdots\cdots & \epsilon^{im}A_{mm}(\epsilon) \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{bmatrix} \tag{2.21}$$

Let us apply this procedure m-1 times in succession to $A_1(\epsilon)$, where at each stage we "peel off" one of the time scales of (2.21), starting from the fastest. We begin by assuming that $A_{11}(0)$ in (2.21) is invertible and apply the procedure just described to $A_1(\epsilon)$, with "$\epsilon A_{22}(\epsilon)$" identified as the large lower right-hand block matrix in (3.7) and with "$A_{12}(\epsilon)$" and "$\epsilon A_{21}(\epsilon)$" defined in a corresponding manner. Thus we construct

$$T_1(\epsilon) = \begin{bmatrix} I & -L_1(\epsilon) \\ \epsilon^{i2}H_1(\epsilon) & I-\epsilon^{i2}H_1(\epsilon)L_1(\epsilon) \end{bmatrix} \qquad (2.22)$$

where, as $\epsilon \to 0$, $H(\epsilon) \to 0$, $G_1(\epsilon) \to A_{11}$, and

$$L_1(\epsilon) = [L_{12}(\epsilon) \ldots . L_{1m}(\epsilon)] \to A_{11}^{-1}[A_{12} \ldots . A_{1m}]$$

Also

$$G_2(\epsilon) = \begin{bmatrix} \tilde{A}_{22}(\epsilon) & \ldots\ldots\ldots\ldots & \tilde{A}_{2m}(\epsilon) \\ & \ldots\ldots\ldots\ldots\ldots & \\ \epsilon^{im-i2}\tilde{A}_{m2}(\epsilon) & \ldots . & \epsilon^{im-i2}\tilde{A}_{mm}(\epsilon) \end{bmatrix}$$

where

$$\tilde{A}_{ij}(\epsilon) = A_{ij}(\epsilon) - A_{i1}L_{ij}(\epsilon) \to A_{ij} - A_{i1}A_{11}^{-1}A_{1j} = \tilde{A}_{ij}$$

We now see that if $\tilde{A}_{22}$ is invertible we can repeat this procedure on $G_2(\epsilon)$. Continuing, we finally obtain

$$T_{m-1}(\epsilon)....T_1(\epsilon)A_1(\epsilon)T_1^{-1}(\epsilon)...T_{m-1}^{-1}(\epsilon)$$
$$= \text{diag}\{G_1(\epsilon),\ \epsilon^{i2}G_2(\epsilon),....,\epsilon^{im}G_m(\epsilon)\}$$

provided that $A_{11} = G_1(0)$, $\tilde{A}_{22} = G_2(0),....,\tilde{A}_{m-1,m-1} = G_{m-1}(0)$ are all invertible. Define

$$\hat{A}(\epsilon) = \text{diag}\{\tilde{A}_{11},\ \epsilon^{i2}\tilde{A}_{22},....,\epsilon^{im}\tilde{A}_{mm}\}$$

We shall prove

Theorem 2.2:

If $\tilde{A}_{11},...,\tilde{A}_{mm}$ are all Hurwitz then $(T;\ \tilde{A}_{11},\ ...,\ \tilde{A}_{mm})$ defines a time-scale decomposition for both (2.12) and (2.15) and $(TP^{-1}(0);\ \tilde{A}_{11},...,\ \tilde{A}_{mm})$ defines one for the original system (2.2).

To prove this theorem we need the following lemma.

Lemma 2.2

Let $E(\epsilon)$ be an nxn matrix with entries from T. If $E(0)$ is Hurwitz, then

$$\lim_{\epsilon \to 0} \sup_{t \geq 0} ||\exp\{E(\epsilon)t\} - \exp\{E(0)t\}|| = 0$$

Proof:

It is well known [9] that

$$\exp\{E(\epsilon)t\} = -1/2\pi j \oint_{\Gamma} \exp\{st\}R(s,E(\epsilon))ds$$

where Γ is a positively-oriented contour enclosing all eigenvalues of $E(\epsilon)$ and

$$R(s,E(\epsilon)) = (sI-E(\epsilon))^{-1}$$

Note that for ϵ small enough $E(\epsilon)$ is Hurwitz and Γ can be chosen to lie in the left-half plane and to enclose all eigenvalues of $E(0)$ as well as $E(\epsilon)$. Therefore we have

$$\begin{aligned} &||\exp\{E(\epsilon)t\} - \exp\{E(0)t\}|| = \\ &\quad 1/2\pi|| \oint \exp\{st\}[R(s,E(\epsilon)) - R(s,E(0))]ds|| \\ &\quad \leq 1/2\pi \exp\{-at\}|| \oint_{\Gamma} R(s,E(\epsilon)) - R(s,E(0))||ds \end{aligned}$$

where a is a positive number. The uniform convergence of $R(s,E(\epsilon))$ to $R(s,E(0))$ on Γ [9] then proves the lemma.

Proof of Theorem:

As a first step we note that by lemma and the assumption that $\tilde{A}_{11},\ldots,\tilde{A}_{mm}$ are Hurwitz

$$\lim_{\epsilon\to 0} \sup_{t\geq 0} ||\mathrm{diag}[\exp\{G_1(\epsilon)t,\ldots,\exp\{\epsilon^{im}G_m(\epsilon)t] -$$

$$\text{diag}[\exp\{\tilde{A}_{11}t\},\ldots,\exp\{\epsilon^{im}\tilde{A}_{mm}t\}]\,||$$

$$\leq \lim_{\epsilon\to 0} \sum_{r=1}^{m} \sup_{t\geq 0} ||\exp\{G_r(\epsilon)t\} - \exp\{\tilde{A}_{rr}t\}|| = 0$$

Thus, since $T(\epsilon)$ is unimodular, we have

$$\lim_{\epsilon\to 0} \sup_{t\geq 0} ||\exp\{A_1(\epsilon)t\} - T^{-1}(\epsilon)\exp(\hat{A}(\epsilon)t\}T(\epsilon)|| = 0$$

From the proof we can see that this result holds not only for (2.12), but also for (2.15). Thus we have proved the first part of Theorem 2.2. To prove that $(TP^{-1}(0);\ \tilde{A}_{11},\ldots,\ \tilde{A}_{mm})$ defines a time-scale decomposition for the original system (2.2) we need just invoke Lemma 2.1.

As a final comment we note that the recursive procedure for peeling off successively slower time scales actually yields a sequence of approximations over longer time intervals if $\exp\{A(\epsilon)t\}$ is replaced by $\text{diag}\{\exp[\tilde{A}_{11}t],\ldots,\exp[\tilde{A}_{rr}\epsilon^{ir}t],\ I,\ldots,I\}$, we obtain a uniform approximation over an interval of the form $[0,\ \epsilon^{l-ir+l}]$ (see [7] for a similar comment).

The results in this section show that the Hurwitz condition is a sufficient condition for the system (2.2), (2.12) and (2.15) to have well-behaved time-scale and provide the actural time-scale decomposition. As mentioned previously, in the next section we show, based on Coderch's results, that this condition is also neccessary.

2.6. RELATIONSHIP WITH CODERCH'S RESULTS

2.6.1. Introduction

In Section 2.4 we showed (Lemma 2.1) that a perturbed, linear time-invariant system that exhibits well-defined time scale behavior can be approximated by its explicit form which is closely related to the Smith form of the system matrix over the local ring T provided the system (2.2) has well-behaved time-scales. Then in Section 2.5 we used a method based on the well-known two-time scale results to prove (Theorem 2.2) that the explicit form and the reduced explicit form can be further approximated by a block-diagonalized, time-scale-seperated system if the $\tilde{A}_{ii}$ are Hurwitz. In this section, starting from the multiple-time-scale results in [1,7], we shall prove the necessity of this condition and thus prove Theorem 2.1.

2.6.2 The Schur Complement

Consider the matrix

$$B = \begin{bmatrix} B_{11}\ B_{12}\ \ldots\ B_{1n} \\ \ldots\ldots\ldots \\ B_{n1}\ B_{k2}\ \ldots\ B_{kk} \end{bmatrix}$$

where the B_{ii}, $i=1,\ldots,n$ are square matrices. Define

$$Y_k = \begin{bmatrix} B_{11}\ B_{12}\ \ldots\ B_{1k} \\ \ldots\ldots\ldots\ldots \\ B_{k1}\ B_{k2}\ \ldots\ B_{kk} \end{bmatrix} \qquad k<n$$

Then the Schur complement of Y_k with respect to B_{kk} is

$$\tilde{B}_{kk} = B_{kk} - [B_{k1}\ldots B_{k,k-1}](Y_{k-1})^{-1}[B_{1k}'\ldots B_{k-1,k}']'$$

If we define

$$S_{k1} = (Y_{k-1})^{-1}B_{ik},\ i=1,\ldots,k-1 \tag{2.26}$$

The Schur coplement can be written as

$$\begin{aligned} \tilde{B}_{kk} &= B_{kk} - [B_{k1}\ldots B_{k,k-1}][S_{k1}'\ldots S_{k,k-1}']' \\ &= B_{kk} - \sum_{i=1}^{k-1} B_{ki}S_{ki} \end{aligned} \tag{2.27}$$

A number of properties of Schur complements are described in

Appendix 1, which we shall find useful in proof of our main results. In addition, there is one other relatively simple property of Schur complements that we are in essence using in this section but whose derivation is deferred until Section 2.7 where we derive other related results. Specifically, in Section 2.5 we define the $\tilde{A}_{ii}$ as <u>successive</u> Schur complements. That is $\tilde{A}_{11}$ is simply the upper left-hand block of A. Then we perform a Schur complement of A compatible with the block structure shown below

$$\bar{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots\cdots & A_{1m} \\ A_{21} & A_{22} & \cdots\cdots & A_{2m} \\ \vdots & \vdots & & \\ \vdots & \vdots & & \\ \vdots & \vdots & & \\ A_{m1} & A_{m2} & \cdots\cdots & A_{mm} \end{bmatrix} = \begin{bmatrix} A_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

So that

$$\tilde{C}_{22} = C_{22} - C_{21}A_{11}^{-1}C_{12}$$

and $\tilde{A}_{22}$ is then simply the upper left-hand block of $\tilde{C}_{22}$. The procedure then continues step by step. On the other hand, the Schur complement $\tilde{A}_{kk}$ defined as in (2.27) if the B's are replaced by A's, is obtained in <u>one</u> step. That these two computations yield the same result is shown in Section 2.7.2.

2.6.3. Main results

The main result in this section is the proof of Theorem 2.1, based on Coderch's results. The basic idea behind the proof is to construct the succesive projection matrices and pseudo-inversces used in Coderch's approach. By using the explicit form for these computations we can then directly relate the MSSNS and MSST conditions to the condition that all of the $\tilde{A}_{ii}$ are Hurwitz. Since the proof is tedious we leave it in Appendix A2 at the end of this chapter. One byproduct of this computation is the explicit identification of the similarity transformation T needed in the time scale decomposition. Thus what we actually obtain is the following:

Theorem 2.3:

Consider system (2.2) and its explicit form (2.12) and reduced explicit form (2.15). Let the system matrices of (2.15) be

$$\bar{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}$$

and define

$$\hat{A}(\epsilon) = \begin{bmatrix} \tilde{A}_{11} & & & 0 \\ & \epsilon\tilde{A}_{22} & & \\ & & \ddots & \\ 0 & & & \epsilon^{n-1}\tilde{A}_{nn} \end{bmatrix}$$

where the $\tilde{A}_{ii}$, i=1,....,n are the Schur complements of $\bar{A}$ as defined before. Then we have

1. Systems (2.2), (2.12) and (2.15) have well-behaved time-scale decompositions if and only if the $\tilde{A}_{ii}$, i=1,...,n are all Hurwitz.

2. If the $\tilde{A}_{ii}$ are all Hurwitz, then $(T; \tilde{A}_{11}, \ldots, \tilde{A}_{nn})$ defines a time-scale decomposition for both (2.12) and (2.15), and $(TP^{-1}(0); \tilde{A}_{11}, \ldots, \tilde{A}_{nn})$ defines one for the original system (2.2), where

$$T = \begin{bmatrix} I & S_{21} & S_{31} & \cdots & S_{n1} \\ 0 & I & S_{32} & \cdots & S_{n2} \\ 0 & 0 & \cdots & & \\ 0 & 0 & \cdots & 0 & I \end{bmatrix}$$

and the S_{ij}, i=2, ..., n, j=1, ..., n-1 are defined in Eq.(2.26) if we replace B_{ij} with A_{ij}. A direct consequence of this result is the following.

Corollary:

The following conditions are equivalent:

1. A(ε) and its explicit form satisfy the MSSNS (MSST) condition.

2. $\tilde{A}_{ii}$ are of full rank (Hurwitz).

2.7 Relationship Between The Two Derivations And The Uniqueness Of The Time Scale Decomposition

2.7.1. Introduction

In this section we show first that the $\tilde{A}_{ii}$ in section 2.5 are nothing more than the Schur complements in Section 2.6. We then turn our attention to the question of the uniqueness of the time scale decomposition. Specifically, we show that although the matrices $P(\epsilon)$ and $Q(\epsilon)$ are not unique (because of the non-uniqueness of Smith form decomposition) the Schur complements of $\bar{A}=Q(0)P(0)$ corresponding to different choices of $P(\epsilon)$ and $Q(\epsilon)$ are similar. Therefore different Smith decompositions give the same time scale result up to a similarity transformation at each time scale.

2.7.2. Recursive and nonrecursive computation of Schur complements

To prove that $\tilde{A}_{ii}$ as defined according to Eq.(2.27) equals $\tilde{A}_{ii}$ in Eq.(2.23) we need the following lemma.

Lemma 2.3:

Consider an invertible matrix A

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \tag{2.30}$$

Define

$$F_{ij}=A_{ij}-A_{i1}A_{11}^{-1}A_{1j},\ i,j,=2,3 \tag{2.31}$$

Then the Schur complement of A_{33} is

$$\tilde{A}_{33}=F_{33}-F_{32}F_{22}^{-1}F_{23} \tag{2.32}$$

Proof: Consider the linear equation

$$A[x'\ y'\ z']' = [0\ 0\ b']' \tag{2.33}$$

It is easy to show that the solution is $z=\tilde{A}_{33}^{-1}b$. We can also write the solution as

$$\begin{bmatrix} y \\ z \end{bmatrix} = \left\{ \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix} - \begin{bmatrix} A_{21} \\ A_{31} \end{bmatrix} \cdot A_{11}^{-1}[A_{12}\ A_{13}] \right\}^{-1} \begin{bmatrix} 0 \\ b \end{bmatrix}$$

$$= \begin{bmatrix} F_{22} & F_{23} \\ F_{32} & F_{33} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ b \end{bmatrix}$$

or

$$\begin{bmatrix} F_{22} & F_{23} \\ F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} \tag{2.34}$$

Solving (2.34) for z we have

$$z=[F_{33}-F_{3222}{}^{-1}F_{23}]^{-1}b \tag{2.35}$$

Eq.(2.35) is true for all possible b. Also the solution for Eq.(2.33) is unique. Therefore we have proved the lemma.

Note that $F_{33}-F_{32}F_{22}{}^{-1}F_{23}$ is nothing more than the $\tilde{A}_{33}$ defined in Section 2.5.2. Hence we can conclude that $\tilde{A}_{33}$ defined in Section 2.6 is the same as that defined in (2.26). Repeatedly using Lemma 2.2 it is easy to prove that this is true for all k=1,...,n.

2.7.3. Similarity of Schur complements of different decompositions

In this subsection we show that although for different PDQ decomposition the matrices Q(0)P(0) are different, their Schur complements are similar. We first show how the Schur complement of a matrix is influenced by a left (right) multiplication by a lower (upper)triangular matrix. Then we prove the similarity between Schur complements of different decompositions.

Lemma 2.4: Consider a matrix

$$S=\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where A is invertible. Consider the Schur complement

$$\tilde{D}=D-CA^{-1}B$$

and suppose that D is invertible. Let

$$S_1=\begin{bmatrix} T_1 & 0 \\ U_1 & V_1 \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}=\begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix}$$

where T_1 and V_1 are invertible. Let

$$\tilde{D}_1 = D_1-C_1A_1^{-1}B_1$$

Then

$$\tilde{D}_1 = V_1\tilde{D}$$

Let

$$S_2 = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} T_2 & V_2 \\ 0 & V_2 \end{bmatrix} = \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix}$$

where T_2 and V_2 are invertible. Let

$$\tilde{D}_2 = D_2-C_2A_2^{-1}B_2$$

Then

$$\tilde{D}_2 = \tilde{D}V_2$$

Proof:

Consider the linear equation

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ a \end{bmatrix} \tag{2.36}$$

The solution is

$$y = \tilde{D}^{-1}a \tag{2.37}$$

By left multiplying by $\begin{bmatrix} T_1 & 0 \\ V_1 & V_1 \end{bmatrix}$ on both sides of (2.36) we have

$$\begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ V_1 a \end{bmatrix} \tag{2.38}$$

The solution of (2.38) is

$$y = \tilde{D}_1^{-1} V_1 a \tag{2.39}$$

Combining (2.37) amd (2.39) we have

$$\tilde{D}^{-1} = \tilde{D}_1^{-1} V_1 a \tag{2.40}$$

Because (2.40) holds for all $a \in R^m$, we have

$$\tilde{D}^{-1} = \tilde{D}_1^{-1} V_1$$

or

$$\tilde{D}_1 = V_1 \tilde{D}$$

Now note that (2.36) can be rewritten as

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} T_2 & U_2 \\ 0 & V_2 \end{bmatrix} \begin{bmatrix} T_2^{-1} & -T_2^{-1}U_2V_2^{-1} \\ 0 & V_2^{-1} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ a \end{bmatrix}$$

or

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} T_2 & U_2 \\ 0 & V_2 \end{bmatrix} \begin{bmatrix} T_2^{-1}x - T_2^{-1}U_2V_2^{-1}y \\ V_2^{-1}y \end{bmatrix} = \begin{bmatrix} 0 \\ a \end{bmatrix}$$

Let

$$x_1 = T_2^{-1}x - T_2^{-1}U_2V_2^{-1}y$$

We have

$$\begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} \begin{bmatrix} x_1 \\ V_2^{-1}y \end{bmatrix} = \begin{bmatrix} 0 \\ a \end{bmatrix}$$

Then

$$V_2^{-1}y = \tilde{D}_2^{-1}a = V_2^{-1}\tilde{D}^{-1}a$$

Therefore

$$\tilde{D}_2 = \tilde{D}V_2$$

Suppose $A(\epsilon)$ has two decompositions:

$$A(\epsilon) = P_1(\epsilon)D(\epsilon)Q_1(\epsilon) = P_2(\epsilon)D(\epsilon)Q_2(\epsilon)$$

$$D(\epsilon) = \text{diag}\{I_0, \epsilon I_1, \ldots.\} = \text{diag}\{d_1, \ldots. , d_n\} \quad (2.41)$$

Where $d_i = \epsilon^{a_i}$, $0 \le a_1 \le a_2 \le \ldots. \le a_n$. Then we have

Lemma 2.5:

The Schur complement of $Q_1(0)P_1(0)$ and $Q_2(0)P_2(0)$ are similar.

Proof:

For simplicity, let us write $D(\epsilon)$, $P_i(\epsilon)$ and $Q_i(\epsilon)$ as D, P_i and Q_i repectively for i=1,2. From (2.41) we have

$$D=\underbrace{P_1^{-1}P_2}_{M}D\underbrace{Q_2Q_1^{-1}}_{N}$$

Therefore

$$M^{-1}D=DN$$

Or

$$(M^{-1}(\epsilon))_{ij}=d_i/d_j(N(\epsilon))_{ij} \tag{2.42}$$

Suppose

$$d_1=d_2=\ldots=d_{s_0}=1,\quad d_{s_0+1}=\ldots=d_{s_1}=\epsilon,\ldots,$$
$$d_{s_{m-1}+1}=\ldots=d_n=\epsilon^m,\quad s_m=n$$

Define

$$R_0=\{1,\ldots,s_0\},\ R_1=\{s_0+1,\ \ldots,s_1\},\ldots,\ R_m=\{s_{m-1}+1,\ldots,s_m\}$$

If $i,j\ R_k$, k=0,..., from (2.42) we have

$$(M^{-1}(\epsilon))_{ij}=(N(\epsilon))_{ij}$$

or

$$(M^{-1}(0))_{ij}=(N(0))_{ij}$$

If $i \in R_{k1}$, $j \in R_{k2}$, $k_1 \neq k_2$ then, because $(M^{-1}(0))_{ij}$ and $(N(0))_{ij}$ are finite for all i,j, we have

$$(N(0))_{ij}=N_{ij}=0, \quad k_1<k_2$$
$$(M^{-1}(0))_{ij}=0, \quad k_1>k_2$$

Then N(0) and $M^{-1}(0)$ can be expressed as

$$N(0) = \begin{bmatrix} n_{11} & & 0 \\ n_{21} & n_{22} & \\ \vdots & & \\ n_{m+1,1} & \cdots\cdots & n_{m+1,m+1} \end{bmatrix} \tag{2.43}$$

$$M^{-1}(0) = \begin{bmatrix} n_{11} & & * \\ & n_{22} & \\ 0 & \ddots & \\ & & n_{m+1,m+1} \end{bmatrix}$$

Hence

$$M(0) = \begin{bmatrix} n_{11}^{-1} & & * \\ & n_{22}^{-1} & \\ 0 & \ddots & \\ & & n_{m+1,m+1}^{-1} \end{bmatrix} \tag{2.44}$$

By definition of M and N we have

$$NQ_1P_1M = Q_2P_2$$

Therefore in view of the special forms of (2.43) and (2.44), the k-th

principal sub-block of $Q_1(0)P_1(0)$ and $Q_2(0)P_2(0)$ denoted by $[Q_1P_1]_k$ and $[Q_2P_2]_k$ can be related by

$$\begin{bmatrix} n_{11} & & 0 \\ & \ddots & \\ * & & n_{kk} \end{bmatrix} [Q_1P_1]_k \begin{bmatrix} n_{11}^{-1} & & * \\ & & \\ 0 & & n_{kk}^{-1} \end{bmatrix} = [Q_2P_2]_k \qquad (2.45)$$

Then applying Fact 1 to (2.45) we see that the k-th Schur complement of $[Q_2P_2]_k = H_{k2}$ is related to the k_{th} Schur complement of $[Q_1P_1]_k$ by

$$H_{k2} = n_{kk}H_{k1}n_{kk}^{-1}$$

Thus we have shown that for different decompositions P_1DQ_1 and P_2DQ_2, the Schur complements are related by

$$\tilde{A}_2(\epsilon) = U\tilde{A}_1(\epsilon)U^{-1}$$

where

$$U = \text{diag}\{n_{11},\ldots,n_{mm}\}$$

and n_{ii}, i=1,....,m are the diagonal blocks of the matrix

$$N(0) = Q_2(0)Q_1(0)^{-1} = \text{diag}\{n_{11},\ldots,n_{mm}\}$$

From this lemma we see that although the Q(0)P(0) is not unique, its successive Schur complements are similar. Furthermore, the time-

scale decomposition is determined by the Schur complements and a similarity transformation matrix T. Therefore the similarity of the Schur complements explains the essential uniqueness of the time-scale decomposition.

2.8 Summary Plus An Extension

2.8.1 Introduction

In this chapter we have developed a new approach for determining if a system of the form

$$\dot{x} = A(\epsilon)x$$

has well-defined time scale behavior and for constructing the corresponding multiple time scale approximation. As we discussed in Section 2.1, a major motivation for this work was a desire to gain additional insight into the algebraic structure of systems with several time scales and to perhaps develop an approach that is conceptually (and hopefully algorithmically) simpler so that further study of these systems might be facilitated. We feel that the approach described in this chapter accomplishes this, since (1) we have been able to provide a clear bridge between the general work of Coderch, et al. [7] and the far simpler and more transparent results developed by Kokotovic, et. al. [17] for systems in what we have termed "explicit form", and (2) the approach established in this chapter provides the foundation for posing and answering numerous important questions about systems with several time scales. In this section we first briefly review the main results of

this chapter to support the first of these points and we then present two additional results which support the second point. The remaining chapters of this thesis provide further indication of the usefulness of this approach to time scale analysis.

Let us first review the major result of this chapter. Specifically, we have analyzed the multiple time scale structure of a system of the form

$$\dot{x} = A(\epsilon)x \tag{2.45}$$

where $A(\epsilon)$ is assumed to be Hurwitz for $\epsilon \in [0,\epsilon_0]$. Our approach consists first of performing a Smith decomposition of $A(\epsilon)$

$$A(\epsilon) = P(\epsilon)D(\epsilon)Q(\epsilon) \tag{2.46}$$

where $|P(0)| \neq 0$, $|Q(0)| \neq 0$ and

$$D(\epsilon) = \text{diag } (\epsilon^{k0}I_0, \epsilon^{k1}I_1, \ldots, \epsilon^{kn-1}I_{n-1})$$

We then consider a similarity transformation on (2.45) which leads us to the <u>explicit form</u>

$$\dot{y} = D(\epsilon)\bar{A}(\epsilon)y$$

where $\bar{A}(\epsilon)=Q(\epsilon)P(\epsilon)$ and the <u>reduced explicit form</u>

$$\dot{z} = D(\epsilon)\bar{A}z \tag{2.47}$$

where

$$\bar{A} = \bar{A}(0) = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}$$

(here the block dimensions are compatible with those for $D(\epsilon)$ in (2.47)). We have shown that $A(\epsilon)$ satisfies the MSST condition <u>if and only if</u> each of the successive Schur complements $\tilde{A}_{ii}$ (see Section 2.6.2 for definition) of $\bar{A}$ is Hurwitz. In this case we have then shown how to construct a similarity transformation T so that

$$\lim_{\epsilon \to 0} \sup_{t \geq 0} \| \exp\{A(\epsilon)t\} - T\exp\{\hat{A}(\epsilon)t\}T^{-1} \| = 0 \qquad (2.48)$$

where

$$\hat{A}(\epsilon) = \text{diag}[\epsilon^{k0}\tilde{A}_{11}, \epsilon^{k1}\tilde{A}_{22}, \ldots, \epsilon^{kn-1}\tilde{A}_{nn}]$$

That is, $(\tilde{A}_{11}, \ldots, \tilde{A}_{nn}; T)$ define a time-scale decomposition of the original system.

Let us make two comments about this result. The first is that it is straightforward to extend this result to the case in which we assume only that $A(\epsilon)$ is semistable for $\epsilon > 0$, i.e. that

$$\lim_{t \to \infty} \exp\{A(\epsilon)t\}$$

exists (so that $A(\epsilon)$ has no eigenvalues in the right-half plane and the

only eigenvalues on the jw axis are simple eigenvalues at 0). Under this condition, the Smith decomposition of $A(\epsilon)$ is as in (2.46), but $D(\epsilon)$ has the form

$$D(\epsilon) = \text{diag}(\epsilon^{k0}I_0,\ldots,\epsilon^{kn-1}I_{n-1},0) = \text{diag}[D_1(\epsilon),0]$$

In this case the explicit form of $A(\epsilon)$ is

$$D(\epsilon)\bar{A}(\epsilon) = \begin{bmatrix} D_1(\epsilon) & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} A_1(\epsilon) & A_2(\epsilon) \\ 0 & 0 \end{bmatrix}$$

Now $A_1(\epsilon)$ is invertible since the upper block corresponds to the nonzero eigenvalues of $A(\epsilon)$. Furthermore, as direct consequence of the result in this chapter, $A(\epsilon)$ will have MSSNS only if $A_1(0)$ is invertible, i.e. only if $A_1(\epsilon)$ is unimodular[+]. Therefore, let us assume this is the case and define the unimodular matrix

$$R(\epsilon) = \begin{bmatrix} I & -A_1^{-1}(\epsilon)A_2(\epsilon) \\ 0 & I \end{bmatrix}$$

Then

[+] By the assumption that $\lim_{t\to\infty} \exp A(\epsilon)t$ exists, then $A_1(0)$ is in fact Hurwitz.

$$R^{-1}(\epsilon)D(\epsilon)\bar{A}(\epsilon)R(\epsilon) = \begin{bmatrix} D_1(\epsilon)A_1(\epsilon) & 0 \\ 0 & 0 \end{bmatrix}$$

Now apply the time scale decomposition developed in this chapter to the explicit form matrix $D_1(\epsilon)A_1(\epsilon)$. If this does not have well-behaved time scale structure, then neither does the original system. If it does have well-behaved time scale structure, let $(\tilde{A}_{11},\ldots,\tilde{A}_{nn};T)$ denote the time scale decomposition of this system. Then let

$$\hat{A}(\epsilon) = \text{diag}\{\epsilon^{k0}\tilde{A}_{11},\ldots,\ \epsilon^{kn-1}\tilde{A}_{nn},\ 0\}$$

$$T_1 = \text{diag}[T\ \ 0]$$

$$T = P(0)R(0)T_1$$

It is then straightforward to check that $(\hat{A}(\epsilon),\ T)$ define the time scale decomposition of $A(\epsilon)$.

The second point we wish to make is that if $A(\epsilon)$ does not satisfy the MSST condition (i.e. if not all of the $\tilde{A}_{ii}$ are Hurwitz), then (1) the reduced explicit form (2.15) need <u>not</u> be asymptotically equivalent to the explicit form (2.12) or the original system (2.2); and (2) the origenal system <u>does not</u> have well-defined time scale behavior in that sense that we have used so far -- i.e. it is not possible to satisfy (2.10) for <u>any</u> choice of <u>constant</u> matrices $\tilde{A}_{ii}$ (i.e. not restricting attention to simply the definitions of these matrices as successive Schur complements). Coderch [7] conjectured that it might be possible

to find an approximation in the sense of (2.10) even if $A(\epsilon)$ violates the MSST condition if one allows the $\tilde{A}_{ii}$ to be ϵ-dependent. In the next two subsections we show how to construct such an approximation which keeps only those ϵ-dependent terms that are needed to achieve (2.10).

Finally, in Section 2.8.4 we discuss the problem of using feedback to modify the time scales of a perturbed system. Specifically, up to this point we have considered the multiple time scale behavior of the undriven system (2.2). In 2.8.4 we discuss the driven system

$$\dot{x} = A(\epsilon)x + B(\epsilon)u \tag{2.49}$$

with the state feedback

$$u = k(\epsilon)x \tag{2.50}$$

The natural question here is to ask what freedom there is in assigning the time scale structure by application of the state feedback of (2.50). If we insert (2.50) into (2.49), we have

$$\dot{x} = [A(\epsilon) + B(\epsilon)K(\epsilon)]x = F(\epsilon)x \tag{2.51}$$

with

$$F(\epsilon) = A(\epsilon) + B(\epsilon)K(\epsilon)$$

As we have developed in this chapter, the time scales of a system like (2.51) are determined by the invariant factors of $F(\epsilon)$. Thus assigning the time scales of (2.49) by state feedback (2.50) naturally leads to the problem of <u>invariant factor assignment</u> of $F(\epsilon)$ by means of choosing

$K(\epsilon)$. In 2.8.4 we concentrate our attention on the special case where $A(\epsilon)$ and $B(\epsilon)$ are left coprime (over ring T). It is established that the invariant factor assignment is related to the rank of $B(0)$. Specifically, we show that the number of invariant factors which can be arbitrarily assigned is no more than rank of $B(0)$.

2.8.2 Extended Well-Defined Time Scale Decompositions

Let us begin with an example to illustrate some of our ideas.

Example 2.8.1 Consider

$$A(\epsilon) = \begin{bmatrix} -\epsilon-\epsilon^2 & 1 \\ -1 & -\epsilon-\epsilon^2 \end{bmatrix}$$

Here $A(\epsilon)$ is already in explicit form, and

$$A_{11} = A(0) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

is invertible but not stable. In other words, $A(\epsilon)$ satisfies the MSSNS but not the MSST condition. Solving this system we have

$$x_1(t) = x_{10}e^{-\epsilon t-\epsilon^2 t}\sin t$$
$$x_2(t) = x_{20}e^{-\epsilon t-\epsilon^2 t}\cos t$$

Let us show that this system does not have well-defined time scale behavior in the sense we have been using so far. Substituting $\tau=\epsilon t$ we have

$$\begin{aligned} x_1(\tau) &= x_{10}e^{-\tau-\epsilon\tau}\sin\tau/\epsilon \\ x_2(\tau) &= x_{20}e^{-\tau-\epsilon\tau}\cos\tau/\epsilon \end{aligned} \qquad (2.52)$$

Obviously $x(\tau)$ does not have a limit as $\epsilon\to 0$ because of the rapid oscillations. On the other hand, in order to have a well-defined time scale approximation we require that there is some A_0 so that

$$\lim_{\epsilon\to 0}\sup_{t\geq 0} \| \exp\{A(\epsilon)t\} - T\exp\{A_0 t\}T^{-1} \| = 0$$

If this were true, then the following should also be true:

$$\lim_{\epsilon\to 0}\sup_{\tau\geq 0} \| \exp\{A(\epsilon)\tau/\epsilon\} - T\exp\{A_0\tau/\epsilon\}T^{-1} \| = 0 \qquad (2.53)$$

It has been shown in (2.51) that the first term does not have a limit as ϵ goes to zero. So the only way to have (2.53) satisfied is if the left hand side is identically 0, but using a <u>constant</u> A_0 does not allow us to do this.

A further investigation shows that if we slightly change our definition of well-defined time scale behavior, then we may be able to extend our results to a larger class of systems. For example, if we use

$$A_0(\epsilon) = \begin{bmatrix} -\epsilon & 1 \\ -1 & -\epsilon \end{bmatrix}$$

instead of a constant matrix A_0 in (2.48), then a simple computation shows

$$\lim_{\epsilon \to 0} \sup_{t \geq 0} \| \exp\{A(\epsilon)t\} - \exp\{A_0(\epsilon)t\} \|$$

$$= \lim_{\epsilon \to 0} \sup_{t \geq 0} \| \exp(-\epsilon t)[1-\exp(-\epsilon^2 t)] \begin{vmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{vmatrix} \|$$

$$\leq \lim_{\epsilon \to 0} \sup_{t \geq 0} \| \exp(-\epsilon t)[1-\exp(-\epsilon^2 t)] \| = 0$$

This example suggests that if we keep some of the ϵ-dependent terms in our system we may obtain valid time scale decompositions even if the system only satisfies MSSNS but not MSST. In order to study the time scale decomposition process under MSSNS alone, we first propose an extended definition for well-defined time scale behavior.

Definition (Extended well-defined time scale behavior)

Let $A(\epsilon)$ be Hurwitz[+] for $\epsilon_0 > \epsilon > 0$. Then $A(\epsilon)$ has extended well-defined time scale behavior if and only if there exist some matrices $A_0(\epsilon), \ldots, A_{n-1}(\epsilon)$, with $A_i(0)$ invertible, $i=0,\ldots,n-1$, and a

---

[+] As before, it is possible to extend these results to the case when $A(\epsilon)$ is only semistable for $\epsilon > 0$. For simplicity we focus on the case when $A(\epsilon)$ is Hurwitz.

constant invertible matrix T such that

$$\lim_{\epsilon \to 0} \sup_{t \geq 0} \| \exp\{A(\epsilon)t\} - T\exp\{\hat{A}(\epsilon)t\}T^{-1} \| = 0$$

where

$$\hat{A}(\epsilon) = \text{diag}[A_0(\epsilon), \epsilon A_1(\epsilon), \ldots, \epsilon^{n-1} A_{n-1}(\epsilon)]$$

Let us now study the relationship between MSSNS condition and the eigenvalues of $A(\epsilon)$ because a deeper understanding will help us capture the essence of our approach. Recall, $A(\epsilon)$ being Hurwitz for $\epsilon_0 > \epsilon > 0$ is an implicit assumption throughout this Chapter (except at the end of the preceding subsection). We claim that if $A(\epsilon)$ satisfies the MSSNS condition, $\tilde{A}_{ii}$ cannot have eigenvalues with positive real parts. The reason is the following. As we shall see in Chapter 3, Theorem 3.2, the eigenvalues of $A(\epsilon)$ are clustered in n groups, with those in the k-th group lying within $O(\epsilon^{kj+1})$ of the eigenvalues of $\epsilon^{kj}\tilde{A}_{jj}$. Consequently, if $\tilde{A}_{jj}$ for some j has an eigenvalue with positive real part, then one of the eigenvalue of $A(\epsilon)$ must have a positive real part for ϵ sufficiently close to zero. This contridicts the assumption that $A(\epsilon)$ is Hurwitz. What <u>can</u> happen, however, is that for some $\tilde{A}_{ii}$ the real parts of some of their eigenvalues are zero. If $A(\epsilon)$ has MSSNS (so that the $\tilde{A}_{ii}$ are invertible), we must then have that these eigenvalues are purely imaginary (i.e. not equal to zero). The implication of this for $A(\epsilon)$ is that its corresponding eigenvalues have negative real parts of orders that are higher than those of their imaginary parts. By keeping only the dominant terms, then, we in essence throw away these damping

effects. If we can somehow keep the dominant terms of these real parts, we should still have a good approximation as we have seen in Example 2.8.1.

In the next subsection we introduce an iterative approach that guarantees the retention of the dominant ϵ-dependent parts of each mode. This approach is based on the following observations. Consider the two time scale case.

$$A(\epsilon) = \begin{bmatrix} A & B \\ \epsilon C & \epsilon D \end{bmatrix} \tag{2.54}$$

where A is invertible. We have shown that there are a unimodular matrix $T(\epsilon)$ and two matrices $G_1(\epsilon)$, $G_2(\epsilon)$ such that

$$A(\epsilon) = T(\epsilon) \begin{bmatrix} G_1(\epsilon) & 0 \\ 0 & \epsilon G_2(\epsilon) \end{bmatrix} T^{-1}(\epsilon)$$

To really compute $G_1(\epsilon)$ and $G_2(\epsilon)$ is impractical . In the previous sections we used A and $\tilde{D}=D-CA^{-1}B$ to approximate $G_1(\epsilon)$, $G_2(\epsilon)$. But, as we just pointed out, if the MSST condition does not hold then this approximation fails, and we have to retain some ϵ-dependent terms. We also know that $G_1(\epsilon)$ and $G_2(\epsilon)$ can be obtained by an iterative procedure if $A(\epsilon)$ satisfies the MSSNS condition [17]. As we show in the next subsection, if we somehow know the orders of the real parts of the eigenvalues of $A(\epsilon)$, then after a specific finite number of iterations the results capture the dominant parts and we can stop. In the next

subsection, we first describe the iterative procedure proposed in [17]. Then we prove some lemmas which are used to determine the stopping rules. Finally we describe our approach, which includes a method for determining an upper bound on the orders of the real parts of the eigenvalues.

2.8.3 An Iterative Algorithm for Extended Time Scale Decomposition

As mentioned in the previous subsection, we have shown in Chapter 2, Section 2.5 that we can write $A(\epsilon)$ in (2.54) in the form

$$A(\epsilon) = T(\epsilon) \begin{bmatrix} G_1(\epsilon) & 0 \\ 0 & \epsilon G_2(\epsilon) \end{bmatrix} T^{-1}(\epsilon) \qquad (2.55)$$

There is an iterative approach to determine $G_1(\epsilon)$ and $G_2(\epsilon)$ which is presented in the following Lemma.

Let us first define matrices P_1 and P_2 associated with a matrix F where

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$$

(here F may be ϵ-dependent but we suppress this dependency).

Specifically

$$P_1 = \begin{bmatrix} I & F_{11}^{-1}F_{12} \\ 0 & I \end{bmatrix}$$

$$P_2 = \begin{bmatrix} I & 0 \\ -F_{21}F_{11}^{-1} & I \end{bmatrix} \qquad (2.56)$$

if F_{11} is invertible. Then we have

Lemma 2.8.1

Let

$$E_0^{(0)}(\epsilon) = A(\epsilon) = \begin{bmatrix} A(\epsilon) & B(\epsilon) \\ \epsilon C(\epsilon) & \epsilon D(\epsilon) \end{bmatrix} = \begin{bmatrix} A_0^{(0)} & B_0^{(0)} \\ \epsilon C_0^{(0)} & \epsilon D_0^{(0)} \end{bmatrix} \qquad (2.57)$$

where $A_0^{(0)}$, $B_0^{(0)}$, $C_0^{(0)}$ and $D_0^{(0)}$ are generally functions of ϵ and $A_0^{(0)}$ is invertible at $\epsilon=0$. Define

$$\begin{aligned} & E_i^{(j)}(\epsilon) \\ &= P_2^{(j)} \cdots P_2^{(1)} P_1^{(i)} \cdots P_1^{(1)} E_0^{(0)}(\epsilon) (P_1^{(1)})^{-1} \cdots \\ &\quad (P_1^{(i)})^{-1} (P_2^{(1)})^{-1} \cdots (P_2^{(j)})^{-1} \\ &= \begin{bmatrix} A_i^{(j)} & \epsilon^i B_i^{(j)} \\ \epsilon^{j+1} C_i^{(j)} & \epsilon D_i^{(j)} \end{bmatrix} \end{aligned} \qquad (2.58)$$

Then as $i \to \infty$, $j \to \infty$

$$A_i^{(j)} \to G_1, \quad D_i^{(j)} \to G_2$$

where G_1 and G_2 are defined in Section 2.5.2 (see (2.19)) and $P_1^{(i)}$ and

$P_2^{(j)}$ are the matrices associated with $E_{i-1}^{(0)}(\epsilon)$ and $E_i^{(j-1)}(\epsilon)$, respectively, as in (2.56). Again all terms in (2.58) including G_1 and G_2 are functions of ϵ.

Proof : See [17].

Note: From (2.56)-(2.58) we have

$$A_i^{(j+1)} = A_i^{(j)} + \epsilon^{i+j+1} B_i^{(j)} C_i^{(j)} (A_i^{(j)})^{-1} \tag{2.59}$$

$$A_{i+1}^{(j)} = A_i^{(j)} + \epsilon^{i+j+1} (A_i^{(j)})^{-1} B_i^{(j)} C_i^{(j)} \tag{2.60}$$

$$D_i^{(j+1)} = D_i^{(j)} - \epsilon^{i+j} C_i^{(j)} (A_i^{(j)})^{-1} B_i^{(j)} = D_{i+1}^{(j)} \tag{2.61}$$

Thus, after the (i+j+1)-th iteration

$$O(A_i^{(j+1)} - A_i^{(j)}) \geq i+j+1 \tag{2.62}$$

$$O(A_{i+1}^{(j)} - A_i^{(j)}) \geq i+j+1 \tag{2.63}$$

$$O(D_i^{(j+1)} - D_i^{(j)}) = O(D_{i+1}^{(j)} - D_i^{(j)}) \geq i+j \tag{2.64}$$

Therefore we would expect that after a finite number of iterations the $A_i^{(j)}$ and $D_i^{(j)}$ would be "close enough" to G_1 and G_2 so that we can use them to approximate G_1 and G_2. The next two lemmas give the definition of "closeness" and provide conditions based on which we can determine the number of iterations required. The first lemma states the eigenvalue order relations for two matrices $A_1(\epsilon)$ and $A_2(\epsilon)$ under which $\exp\{A_1(\epsilon)t\}$ can be approximated by $\exp\{A_2(\epsilon)t\}$. Let λ_{ij}, i=1,2,

$j=1,\ldots,N$ be eigenvalues of $A_1(\epsilon)$ and $A_2(\epsilon)$, repectively, and let $O(x)$ denote the order of x, then (suppressing the ϵ-dependence of λ_{ij})

Lemma 2.8.2 If $A_1(\epsilon)$ and $A_2(\epsilon)$ satisfy

1. $A_1(0) = A_2(0)$ has SSNS.
2. $O(\mathrm{Re}(\lambda_{1j})) = O(\mathrm{Re}(\lambda_{2j})) = r_j < O(\mathrm{Re}(\lambda_{1j}-\lambda_{2j}))$, $j=1,\ldots,N$.
3. $O(\mathrm{Im}(\lambda_{1j})) = O(\mathrm{Im}(\lambda_{2j})) = s_j$, and if $r_j \geq s_j$ then $s_j < O(\mathrm{Im}(\lambda_{1j}-\lambda_{2j}))$, $i=1,\ldots,N$.

then

$$\lim_{\epsilon \to 0} \sup_{t \geq 0} \| \exp\{A_1(\epsilon)t\} - \exp\{A_2(\epsilon)t\} \| = 0$$

Proof:

Define $R[A(\epsilon),\lambda] = (\lambda I - A(\epsilon))^{-1}$ as before. From Condition 1 we know that the number of the eigenvalues whose real parts have orders higher than 0 is the same for both $A_1(\epsilon)$ and $A_2(\epsilon)$. Denote this number by m and suppose that the first m eigenvalues are these higher-order eigenvalues. From the assumption that $A(\epsilon)$ is Hurwitz for $0<\epsilon<\epsilon_0$, it is also clear that the leading term (i.e. the lowest order term) of the real part of each eigenvalue has a negative coefficient. Now, write $R[A_i(\epsilon),\lambda]$ as

$$R[A_i(\epsilon),\lambda] = \sum_{j=1}^{m} \frac{F_{ij}(\epsilon)}{\lambda - \lambda_{ij}} + \frac{E_i(\epsilon)}{p_i(\lambda,\epsilon)}, \quad i=1,2 \qquad (2.65)$$

where $p_i(\lambda,\epsilon) = (\lambda-\lambda_{i,m+1})(\lambda-\lambda_{i,m+2})\ldots(\lambda-\lambda_{i,N})$ and $F_{ij}(\epsilon)$ and $E_i(\epsilon)$ are functions of ϵ only and are analytic at $\epsilon=0$. We know that

$$\exp\{A_1(\epsilon)t\} - \exp\{A_2(\epsilon)t\}$$
$$= \oint_C \{R[A_1(\epsilon),\lambda] - R[A_2(\epsilon),\lambda]\}e^{\lambda t}d\lambda \qquad (2.66)$$

where C is the contour enclosing every eigenvalues in the complex plane. Using (2.65), (2.66) can be rewritten as

$$\sum_{j=1}^{m} \oint_{C_j} \{\frac{F_{1j}(\epsilon)}{\lambda-\lambda_{1j}} - \frac{F_{2j}(\epsilon)}{\lambda-\lambda_{2j}}\}e^{\lambda t}d\lambda + \oint_{C'} \{\frac{E_1(\epsilon)}{p_1(\lambda,\epsilon)} - \frac{E_2(\epsilon)}{p_2(\lambda,\epsilon)}\}e^{\lambda t}d\lambda \qquad (2.67)$$

where C_j are contours enclosing eigenvalues λ_{1j} and λ_{2j}, $j=1,\ldots,m$ and C' is a contour enclosing all λ_{1j}, λ_{2j}, $j=m+1,\ldots,N$. Since $Re(\lambda_{ij}(\epsilon)) \leftrightarrow Re(\lambda_{ij}(0))<0$, $i=1,2$, $j=m+1,\ldots,N$, C' can be chosen to lie entirely in the left half plane. Therefore, since

$$\left\| \frac{E_1(\epsilon)}{p_1(\lambda,\epsilon)} - \frac{E_2(\epsilon)}{p_2(\lambda,\epsilon)} \right\| \longrightarrow 0, \ \epsilon\longrightarrow 0 \quad \text{(Condition 1)}$$

$$\left\| \oint_{C'} \{\frac{E_1(\epsilon)}{p_1(\lambda,\epsilon)} - \frac{E_2(\epsilon)}{p_2(\lambda,\epsilon)}\} e^{\lambda t}d\lambda \right\| \longrightarrow 0, \ \epsilon\longrightarrow 0$$

$$\leq \oint_{C'} \left\| \frac{E_1(\epsilon)}{p_1(\lambda,\epsilon)} - \frac{E_2(\epsilon)}{p_2(\lambda,\epsilon)} \right\| d\lambda \longrightarrow 0, \ \epsilon\longrightarrow 0 \qquad (2.68)$$

Now consider only one term in the first part of (2.67):

$$\oint_{C_j} \{\frac{F_{1j}(\epsilon)}{\lambda-\lambda_{1j}} - \frac{F_{2j}(\epsilon)}{\lambda-\lambda_{2j}}\} e^{\lambda t} d\lambda \qquad (2.69)$$

Let $\lambda'=\lambda/\epsilon^{rj}$. Then (2.69) can be written as

$$\oint_{C'_j} \{\frac{F_{1j}}{\epsilon^{rj}\lambda'-\lambda_{1j}} - \frac{F_{2j}}{\epsilon^{rj}\lambda'-\lambda_{2j}}\} \exp\{\epsilon^{rj}\lambda' t\} d\lambda' \epsilon^{rj}$$

$$= \oint_{C'_j} \{\frac{F_{1j}}{\lambda'-\lambda'_{1j}} - \frac{F_{2j}}{\lambda'-\lambda'_{2j}}\} \exp\{\lambda' t\} d\lambda' \qquad (2.70)$$

where

$$\lambda'_{ij} = \lambda_{ij}/\epsilon^{rj},\ \tau=\epsilon^{rj}t,\ i=1,2,\ j=1,\ldots,m.$$

From Conditions 2 and 3 and the Hurwitz assumption we have

$$0>\mathrm{Re}(\lambda'_{ij}) = O(1),$$

$$\mathrm{Re}(\lambda'_{1j}-\lambda'_{2j}) = o(1),\quad \mathrm{Im}(\lambda'_{1j}-\lambda'_{2j}) = o(1) \qquad (2.71)$$

Therefore it is obvious that C'_j can be taken as a circle in the left half plane whose diameter is of order 1. Since the order of the imaginary part may be lower or higher than that of the real part, this circle may move up or down vertically as ϵ changes but it remains in the left-hand plane and does not shrink to 0. Also

$$\frac{F_{1j}}{\lambda'-\lambda'_{1j}} - \frac{F_{2j}}{\lambda'-\lambda'_{2j}} = \frac{(\lambda'-\lambda_{2j}')F_{1j}-(\lambda'-\lambda_{1j}')F_{2j}}{(\lambda'-\lambda_{1j}')(\lambda'-\lambda_{2j}')}$$

$$\frac{(\lambda_{1j}'-\lambda_{2j}')F_{2j}-(\lambda'-\lambda_{2j}')(F_{2j}-F_{1j})}{(\lambda'-\lambda_{1j}')(\lambda'-\lambda_{2j}')} \qquad (2.72)$$

On contour C_j', $\lambda_{1j}' \dashrightarrow \lambda_{2j}'$, $F_{2j} \dashrightarrow F_{1j}$ as $\epsilon \dashrightarrow 0$ and the denominator is bounded away from 0. Therefore the integrand of (2.70) goes to zero uniformly on C'_j as $\epsilon \dashrightarrow 0$. Considering (2.68) we can conclude that (2.67), and therefore (2.66), goes to zero uniformly in t as $\epsilon \dashrightarrow 0$.

This Lemma gives us a criterion to judge how "close" the eigenvalues of two matrices should be in order for one to be a good approximation of the other. In the next Lemma we study how a perturbation on a matrix can influence the eigenvalues. To develop a complete picture of how perturbations affect the eigenvalues is beyond the scope of this development. For our purpose of defining a stopping point for our iterative computation of $G_1(\epsilon)$ and $G_2(\epsilon)$, we only need the following result.

Lemma 2.8.3 :

Let $A=B+\epsilon^m C$ where A,B and C are NxN matrix functions of ϵ. Let the eigenvalues of A and B be denoted by $\lambda_i(A)$ and $\lambda_i(B)$, i=1,...,N. Then

$$\min_i O(\lambda_i(A)-\lambda_i(B)) \geq m/p \geq m/N \qquad (2.73)$$

where p is some integer, $1 \leq p \leq N$.

Proof : See [9].

Combining this Lemma with (2.59)-(2.64) it is clear that after the (i+j+1)-th iteration the eigenvalue difference between the (i+j)-th and the (i+j+1)-th iterations must have an order higher than (i+j)/N. Therefore, as the number of iterations becomes larger and larger, the contribution of each additional iteration becomes less and less. Thus, a stopping rule can be established if we in addition take Lemma 2.8.2 into account. This result can be summarized in the following Theorem.

Theorem 2.8.1:

Let

$$A(\epsilon) = \begin{bmatrix} A_0^{(0)}(\epsilon) & B_0^{(0)}(\epsilon) \\ \epsilon C_0^{(0)}(\epsilon) & \epsilon D_0^{(0)}(\epsilon) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \epsilon I \end{bmatrix} \begin{bmatrix} A_0^{(0)}(\epsilon) & B_0^{(0)}(\epsilon) \\ C_0^{(0)}(\epsilon) & D_0^{(0)}(\epsilon) \end{bmatrix}$$

$$= D(\epsilon)\bar{A}(\epsilon)$$

satisfy the MSSNS condition. Define $P_1^{(i)}, P_2^{(j)}, A_i^{(j)}(\epsilon)$ and $D_i^{(j)}(\epsilon)$ as in (2.56),(2.58)-(2.61). Suppose the maximum order of the real parts of the eigenvalues of $A(\epsilon)$ is m. Then

$$\lim_{\epsilon \to 0} \sup_{t \geq 0} \| \exp\{A(\epsilon)t\} - T_1 \exp\{A_1(\epsilon)t\} T_1^{-1} \| = 0 \tag{2.74}$$

where

$$A_1(\epsilon) = \begin{bmatrix} A_q^{(p)}(\epsilon) & 0 \\ 0 & \epsilon D_q^{(p)}(\epsilon) \end{bmatrix} \tag{2.75}$$

and $p+q=mN+1$, where N is the dimension of $A(\epsilon)$, $p\geq 0, q\geq 0$ and T_1 is a constant matrix defined in the proof.

Proof:

For simplicity we will write $A_0^{(0)}(0)$, $B_0^{(0)}(0)$, $C_0^{(0)}(0)$, and $D_0^{(0)}(0)$ as A,B,C, and D, respectively. From Section 2.5 we see that

$$A(\epsilon) = T(\epsilon)\begin{bmatrix} G_1(\epsilon) & 0 \\ 0 & \epsilon G_2(\epsilon) \end{bmatrix} T^{-1}(\epsilon) \qquad (2.76)$$

and $G_1(0)=A$, $G_2(0)=D-CA^{-1}B$. It is also clear from Eq.(2.59)-(2.61) that

$$A_i^{(j)}(0) = A, \quad D_i^{(j)}(0) = D-CA^{-1}B, \quad i+j\geq 1$$

Therefore

$$A_q^{(p)}(0) = G_1(0), \quad D_q^{(p)}(0)=G_2(0)$$

and these matrices have SSNS. Furthermore, from Lemma 2.8.1 we know that

$$A_i^{(j)} \to G_1, \quad D_i^{(j)} \to G_2, \quad i+j \to \infty$$

On the other hand, one more iteration than $p+q=mN+1$ will change the eigenvalues of $A_q^{(p)}$ on the order of $(mN+1)/N=m+1/N>m$. Since the highest order of the real parts of eigenvalues is m, we see that this implies that Condition 2 of Lemma 2.8.2 is satisfied. Furthermore, Condition 3 is automatically satisfied if $r_j\geq s_j$, since as we just

argued, the order change introduced by one more iteration will be higher than r_j and therefore greater than s_j. From Lemma 2.8.2, we therefore have

$$\lim_{\epsilon \to 0} \sup_{t \geq 0} \| \exp\{A_q^{(q)} t\} - \exp\{G_1 t\} \| = 0 \qquad (2.77)$$

A similar equation holds for $D_q^{(p)}$ and G_2.

Now define the product of the iteratively applied $P_1^{(i)}$ and $P_2^{(j)}$ to be $T_1^{-1}(\epsilon)$ which is obviously unimodular and therefore contains no information about the time scales. It is also easy to check that $T_1(0)=T(0)$, where $T(\epsilon)$ is defined in (2.76). Thus, we finally have (2.74) with $T_1=T_1(0)$.

Note:

1. Since $A(\epsilon)$ is Hurwitz for $\epsilon_0>\epsilon>0$, the real parts of its eigenvalues cannot be zero and are in fact negative. Therefore m is finite. In fact, let the order of the determinant of $A(\epsilon)$ be O_d. Then

$$O_d \geq O(|\lambda_i|), \quad i=1,\ldots,N$$

But $|\lambda_i|^2=(\mathrm{Re}\lambda_i)^2+(\mathrm{Im}\lambda_i)^2$. Therefore it is easy to see that $O|\lambda_i| \geq O(\mathrm{Re}\lambda_i)$, $i=1,\ldots,N$. In other words, O_d is an upper bound on m.

2. The integers p and q in (2.75) are not unique but their sum --

the number of total iterations must be sufficiently large. It is quite possible that the iteration can stop much earlier. Nevertheless, $p+q=NO_d$ is a safe bound for stopping.

3. In Theorem 2.8.1 only the first step of the multiple time scale decomposition has been shown. But the general procedure is only a simple extension of this theorem. The detailed procedure is as follows.

(1) Determine the Smith form of $A(\epsilon)=P(\epsilon)D(\epsilon)Q(\epsilon)$.

(2) Let $D(\epsilon)Q(\epsilon)P(\epsilon)=D(\epsilon)\bar{A}(\epsilon)$ where

$$\bar{A}(\epsilon) = \begin{bmatrix} A_{11} & A_{12} & \cdots\cdots & A_{1n} \\ \epsilon A_{21} & \epsilon A_{22} & \cdots\cdots & \epsilon A_{2n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \epsilon^{n-1}A_{n1} & \epsilon^{n-1}A_{n2} & \cdots\cdots & \epsilon^{n-1}A_{nn} \end{bmatrix}$$

(3) Compute the order of the determinant of $A(\epsilon)$ and use it as an upper bound on m.

(4) Similar to the procedure for deriving the multiple time scale decomposition in Chapter 2, we treat A_{11} and the complement lower right principal submatrix as $A_0^{(0)}$ and $D_0^{(0)}$ in Theorem 2.8.1 and apply $p+q=NO_d+1$ iterations to them.

(5) Treat the resulting $D_q^{(p)}$ as the original matrix $A(\epsilon)$ and go back to step (3).

(6) Repeat this procedure until all time scales have been revealed.

4. The result in the previous sections of this Chapter can be viewed as a special case of the theorem in this section. That is, if $A(\epsilon)$ satisfies the MSST condition, then for each time scale, only one iteration is required. But if we do not know if the system satisfies MSST, we have to do the following. First we apply the procedure described in the preceding section, which corresponds to performing only one step of the iteration described in this section at each time scale. If for all time scales the resulting deagonal blocks (i.e. $\tilde{A}_{ii}$ as we denoted before) are Hurwitz at $\epsilon=0$, then we are done. Otherwise, if at some step i, the resulting $\tilde{A}_{ii}$ is not Hurwitz at $\epsilon=0$, we have to go back to the very beginning and perform the full set of iterations at each time scale.

2.8.4 Invariant Factor Assignment When $A(\epsilon)$ And $B(\epsilon)$ Are Left Coprime

In this subsection we solve the problem of invariant factor assignment by state feedback when $A(\epsilon)$ and $B(\epsilon)$ are left coprime over the ring T. To prove the main results we need the following Lemma.

Lemma 2.8.4

Let $A(\epsilon)$, $B(\epsilon)$ and $K(\epsilon)$ be nxn, nx1 and 1xn matrices over T and

$$A(\epsilon) \sim \text{diag}[a_1,\ldots,a_n] = S_A$$

where $a_1,\ldots,a_n$ are invariant factors of $A(\epsilon)$ and "$\sim$" denotes equivalence via multiplication (on the right and left) by unimodular transformations. Let

$$F(\epsilon) = A(\epsilon) + B(\epsilon)K(\epsilon) \tag{2.78}$$

Then for any $K(\epsilon) \in T^{mxn}$ there is a $\bar{K}(\epsilon) \in T^{mxn}$ such that

$$F(\epsilon) \sim S_A + \bar{B}(\epsilon)\bar{K}(\epsilon)$$

where $\bar{B}(\epsilon)$ is an upper triangular matrix and $B(\epsilon) \sim \bar{B}(\epsilon)$. Furthermore there exist unimodular matrices R and S such that

$$K = R\bar{K}S$$

Proof:

We know that there exist unimodular matrices P and Q such that

$$A = PS_AQ$$

Therefore

$$A + BK = PS_AQ + BK = P(S_A + P^{-1}BKQ^{-1})Q$$

and

$$A + BK \sim S_A + P^{-1}BKQ^{-1} \triangleq W \tag{2.79}$$

Furthermore, by elementary column operations we can show that

$$P^{-1}BU = \bar{B}$$

where U is a unimodular matrix and $\bar{B}$ is an upper triangular matrix.

Therefore (2.79) can be rewritten as

$$W = S_A + \bar{B}U^{-1}KQ^{-1}$$

Let

$$\bar{K} = U^{-1}KQ^{-1}$$

We have

$$W = S_A + \bar{B}\bar{K}$$

and the lemma is proved.

What Lemma 2.8.4 tells us is that instead of considering the invariant factor assignment of $F(\epsilon)$ we may work on $S_A+\bar{B}(\epsilon)\bar{K}(\epsilon)$. Because of the special form of S_A and $\bar{B}(\epsilon)$, the original problem becomes much easier. We shall see this in the proof of our main result Theorem 2.8.2.

Theorem 2.8.2

Let $A(\epsilon) \in T^{nxn}$, $B(\epsilon) \in T^{nxr}$, and let b denote the rank of $B(0)$. Assume $A(\epsilon)$ and $B(\epsilon)$ are left coprime. Then

1. $F(\epsilon)$ defined in (2.78) can have no more than b non-unit invariant factors for any choice of $K(\epsilon)$.

2. There exists a $K(\epsilon)$ such that $F(\epsilon)$ has $\epsilon^{j_1},\ldots,\epsilon^{j_b}$ as its invariant factors, for arbitrary non-negative integers $j_1,\ldots,j_b$ (with the convention that $\epsilon^{\infty}=0$).

Note

1. The left coprimeness of two matrices $A(\epsilon)$ and $B(\epsilon)$ over the local ring T means that if there are matrices $D(\epsilon)$ T^{nxn}, $\bar{A}(\epsilon)$ T^{nxn} and $\bar{B}(\epsilon) \in T^{nx}$ such that

$$A(\epsilon) = D(\epsilon)\bar{A}(\epsilon), \quad B(\epsilon) = D(\epsilon)\bar{B}(\epsilon) \tag{2.80}$$

then $D(\epsilon)$ is unimodular. That is $D(0)$ is nonsingular. A simple criterian for the left coprimeness of $A(\epsilon)$ and $B(\epsilon)$ is to check if $[A(0)\ B(0)]$ is of full row rank. For example,

$$A(\epsilon) = \begin{bmatrix} 1 & 0 \\ 1 & \epsilon \end{bmatrix}, \qquad B(\epsilon) = \begin{bmatrix} \epsilon & 0 \\ 0 & 1 \end{bmatrix}$$

are coprime because

$$[A(0)\ \ B(0)] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

has full rank. But

$$A(\epsilon) = \begin{bmatrix} 1 & 0 \\ 1 & \epsilon \end{bmatrix}, \qquad B(\epsilon) = \begin{bmatrix} 1 & 0 \\ 1+\epsilon^2 & \epsilon \end{bmatrix}$$

are not because there exists a nonunimodular matrix

$$D(\epsilon) = \begin{bmatrix} 1 & 0 \\ 1 & \epsilon \end{bmatrix}$$

and matrices

$$\bar{A}(\epsilon) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \bar{B}(\epsilon) = \begin{bmatrix} 1 & 0 \\ \epsilon & 1 \end{bmatrix}$$

satisfying (2.80). Alternatively, we note that

$$[A(0) \quad B(0)] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

is not full rank.

2. It is also easy to see that in order to ensure the full rank of $[A(0) \quad B(0)]$, the rank of $A(0)$ must be greater than $n-b$. Or in other words, $A(\epsilon)$ itself has at most b invariant factors other than 1.

Proof:

As in Lemma 2.8.4, let

$$A = PS_AQ, \quad \bar{B} = P^{-1}BU$$

where $\bar{B}$ is an upper triangular matrix. Consider

$$W = S_A + \bar{B}K$$

Suppose A has $n-m$ unit invariant factors.

$$S_A = \text{diag}[\underbrace{1,..,1}_{n-m}, \underbrace{e^{il},.....,e^{im}}_{m}]$$

Then in view of left coprimeness of A and B, $\bar{B}$ must have the form of

$$\bar{B}(\epsilon) = \begin{bmatrix} x & \cdots\cdots\cdots & x \\ \vdots & & \vdots \\ x & \cdots\cdots\cdots & x \\ & \vdots & \\ & x \quad x & \vdots \\ & 1 & \vdots \\ 0 & \quad \ddots & \vdots \\ & & 1 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n-m \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \end{matrix}$$

where x represents and arbitrary element in T.

First, let us assume that m=b. Then we can construct a unimodular matrix V such that

$$\bar{B}V = \begin{bmatrix} x & \cdots\cdots\cdots & x \\ \vdots & & \vdots \\ x & \cdots\cdots\cdots & x \\ & \vdots & \\ & x \quad \cdots\cdots & x \\ & 1 & \\ 0 & \quad \ddots & \\ & & 1 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n-m \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \end{matrix}$$

Suppose $g_1,...g_m$ are the desired closed-loop invariant factors. Let

$$\bar{K} = V\begin{bmatrix} 0 & 0 \\ 0 & G_1 \end{bmatrix}$$

where $G_1 = \text{diag}[g_1-\epsilon^{il},...,g_m-\epsilon^{im}]$. Then

$$S_A + \bar{B}\bar{K} = \begin{bmatrix} I & W \\ 0 & G_2 \end{bmatrix}$$

where $G_2 = \text{diag}[g_1,\ldots,g_m]$ and W is some $(r-m)\times m$ matrix. After some elementary operation, we can show that

$$S_A + \bar{B}\bar{K} \sim \text{diag}[1,..,1,g_1,\ldots,g_m]$$

By Lemma 2.8.4 we know that A+BK has invariant factors $g_1,\ldots,g_m$ where

$$K = U\bar{K}Q$$

Now, assume $m \leq b$. Then after some elementary column operations we can always achieve

$$\bar{B}V \triangleq \begin{array}{c} \\ k_1 \\ \\ k_2 \\ \\ \\ \\ \\ \end{array} \begin{array}{c} \begin{matrix} s_1 & & & s_2 \end{matrix} \\ \left[\begin{array}{cccc} x & \cdots & \cdots & x \\ 0\ldots0 & 1 & 0 & \cdots\cdots 0 \\ & \cdots & \cdots & \\ 0..0 & 1 & 0 & \cdots\cdots 0 \\ x & \cdots & \cdots & x \\ \cdot & & & \cdot \\ & x & \cdots\cdots & x \\ & & 1 & \\ 0 & & \cdot & \\ & & & 1 \end{array}\right] \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{c}a\\a\\a\\a\\a\\a\\a\\a\end{array}}\right\} n-m \\ \left.\vphantom{\begin{array}{c}a\\a\\a\\a\end{array}}\right\} m \end{array}$$

Namely there will be b-m rows (say, rows $k_1,\ldots,k_{b-m}$) among the first n-m rows, each of which has one unit and r-1 zeros. Furthermore if the unit in k_j-th row is located in the s_j-th column, then $s_j \neq s_p$, if $j \neq p$. That is, the units for different rows are located in different columns.

If $g_1,\ldots,g_b$ are the desired invariant factors, we can construct a

$\bar{K}$ in a similar manner to that used previously:

$$\bar{K} = V\begin{bmatrix} K_1 & 0 \\ 0 & G_3 \end{bmatrix}$$

where

$$K_1 = \begin{matrix} \\ j_2 \\ j_1 \end{matrix}\begin{matrix} \begin{matrix} k_1 & k_2 \end{matrix} \\ \begin{bmatrix} 0 & 0.....0....0 \\ 0 & 0.....0....0 \\ 0 & 0....g_2\text{-}1..0 \\ 0 & g_1\text{-}1\ ..0....0 \\ 0 & 0\ ..... \ 0 \end{bmatrix} \end{matrix}$$

and

$$G_3 = \text{diag}[g_{b-m+1}-\epsilon^{il},....,g_b-\epsilon^{im}]$$

similarly we have

$$S_A + \bar{B}\bar{K} = \begin{bmatrix} K_2 & k_3 \\ 0 & G_4 \end{bmatrix}$$

where

$$K_2 = \begin{bmatrix} 1 & x & x & ..... & x \\ & & ..... & & \\ 0....0 & g_1 & 0.....0 & & \\ x....x & x & 1 & x...x & \\ & & ..... & & \\ 0 & ..... & 0 & g_2 & \end{bmatrix}$$

$$G_4 = \text{diag}[g_{b-m+1},....,g_b]$$

and K_3 is some arbitriry matrix. After some column operations we have

$$S_A + \bar{B}\bar{K} \sim \text{diag}[1,\ldots,1,g_1,1,\ldots,1,g_2,g_{b-m+1},\ldots,g_b]$$

By Lemma 2.8.4 we know A+BK has invariant factors $g_1,\ldots,g_b$. This proves the second part of Theorem 2.8.2.

Let N be the number of unit invariant factors of A. It is easy to see, that

$$N = \text{rank}[A(0)+B(0)K(0)]$$

Using Lemma 2.8.4 this is equivalent to

$$N = \text{rank}[S_A(0)+\bar{B}\bar{K}(0)]$$

And as pointed out before

$$S_A(0) + \bar{B}(0)\bar{K}(0) = \text{diag}[1,\ldots,1,0,\ldots,0]$$

$$+ \begin{bmatrix} x & \cdots & x \\ \vdots & & \vdots \\ x & \cdots & x \\ & x \;\; \cdots \;\; x & \\ & 1 & \\ 0 & \ddots & \\ & & 1 \end{bmatrix} \begin{matrix} \bar{K}(0) \\ \Big\} m \end{matrix}$$

where x represents either 1 or 0. Since the rank of B(0) is $b \geq m$, there must exist b-m independent column vectors in the first r-m columns in $\bar{B}(0)$. Since left multiplication by $\bar{K}(0)$ is nothing more than column operations and the rank of $S_A(0)$ is n-m, adding $\bar{B}(0)\bar{K}(0)$ to $S_A(0)$ can reduce the rank of $S_A(0)$ by at most b-m. In other words, the rank of $S_A(0)+\bar{B}(0)\bar{K}(0)$ is greater than n-m-(b-m)=n-b. Or

$$N \geq n-b$$

which proves part 1 of the theorem.

(Note: There is no page 98)

CHAPTER 3

EIGENVALUE AND INVARIANT FACTOR STRUCTURES

3.1 INTRODUCTION

So far we have shown that the MSSNS of a matrix $A(\epsilon)$ implies that the Schur complements $\tilde{A}_{ii}$ are non-singular[+] and MSST implies that $\tilde{A}_{11}$ are Hurwitz. In this chapter we shall introduce new but equivalent criteria for the MSSNS and MSST conditions. These criteria are important not only because they provide new insight into these conditions and their relationship to the eigenvalues of $A(\epsilon)$, but also because they make it possible to introduce a new approach, developed in Chapter 4, for defining time and amplitude scaling when the MSSNS condition is not satisfied so that the traditional time-scale

---

[+] As in most of our development in Chapter 2, we assume here that $A(\epsilon)$ is invertible and in fact Hurwitz for $\epsilon \in (0,\epsilon_0]$. The extension ofthe results of this chapter to the case when this is not true can be accomplished in a straightforward manner, such as in Section 2.8.

decomposition method cannot be applied.

First, in Section 3.2, we show several equivalent conditions for MSSNS. In particular, it is proven in Theorem 3.1 that the MSSNS condition is not only equivalent to $\tilde{A}_{ii}$ being non-singular as seen in last chapter, but also equivalent to 1) the eigenvalues and invariant factors of $A(\epsilon)$ having the same orders and to 2) a "no-cancellation" condition, namely that for a certain set of values of i the order of the gcd of all ixi principal minors must be the same as that of their sum, together with the condition that this set of the gcd's uniquely determines the invariant factors in a particular fashion. The plausibility of the first result can be seen from the following observation. Suppose $A(\epsilon)$ not only has MSSNS but also has MSST. Then it has been shown in last chapter that the time scales are determined by its invariant factors. On the other hand, the eigenvalues of $A(\epsilon)$ determine its modes. Therefore we would expect that if the system has well behaved time-scales, i.e. $A(\epsilon)$ has MSST, the eigenvalues and invariant factors of $A(\epsilon)$ should have the same orders. Since the MSST condition differs from MSSNS condition only by adding an extra stability conditions (which we shall touch on in a moment) MSSNS should also ensure this purely algebraic equality. It is shown in Theorem 3.1 that the equality of the orders of the eigenvalues and invariant factors is not only a necessary condition for $A(\epsilon)$ to have MSSNS, as we just argued, it is also a sufficient condition. To illustrate these ideas, let us consider two examples. The first one does not have MSSNS but the

second one does.

Example 3.1.1

In this example we consider a system that does not have MSSNS. As we will see, its eigenvalues and invariant factors do not have the same orders either. The system considered is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -\epsilon & 1 \\ 0 & -\epsilon \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A_1(\epsilon)x$$

This system does not have well-defined time scale behavior which can be seen directly from the solution of this system.

$$x_1(t) = x_{10}e^{-\epsilon t} + x_{20}te^{-\epsilon t}$$
$$x_2(t) = x_{20}e^{-\epsilon t}$$

If we substitute $\tau=\epsilon t$, we have

$$x_1(\tau/\epsilon) = x_{10}e^{-\tau} + x_{20}\tau/\epsilon e^{-\tau}$$
$$x_2(\tau/\epsilon) = x_{20}e^{-\tau}$$

It is clear that at the time scale $\tau=\epsilon t$, the second term in the first equation becomes unbounded as ϵ goes to zero. As pointed out in [7], this is an evidence that the system does not have well-defined time scale behavior. It is easy to see that this system does not have MSSNS either. Furthermore, the eigenvalues of $A_1(\epsilon)$ are ϵ and ϵ. On the other hand, the Smith form of $A_1(\epsilon)$ is

$$A_1(\epsilon) = \begin{bmatrix} 1 & 0 \\ \epsilon & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \epsilon^2 \end{bmatrix} \begin{bmatrix} -\epsilon & 1 \\ -1 & 0 \end{bmatrix}$$

Therefore the orders of eigenvalues are 1 and 1 and the orders of the invariant factors are 0 and 2.

Example 3.1.2

Now consider the system matrix

$$A_2(\epsilon) = \begin{bmatrix} -1 & 1 \\ 0 & -\epsilon \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$$

We already know that this system has well-behaved time-scales. It can be seen that the orders of the eigenvalues and the invariant factors are the same, i.e. 0 and 1.

Based on the proof of this result we shall observe that under MSSNS the eigenvalues of $A(\epsilon)$ are clustered in n groups with those in the k-th group close to the eigenvalues of $\epsilon^{jk}\tilde{A}_{kk}$. This block diagonal dominant phenomenon is expected in view of the appoximation of $\exp\{A(\epsilon)t\}$ by $T\exp\{\hat{A}(\epsilon)t\}T^{-1}$ if the system satisfies MSST. It is also true when MSSNS condition is satisfied. This observation is summarized in Theorem 3.2.

The second result in Theorem 3.1 first says A(ε) satisfies MSSNS if and only if the set of gcd's (greatest common divisors) of all ixi principal minors, i=1,...,n uniquely determine the invariant factors and in addition we have the no-cancellation condition mentioned earlier. This is different from the general case where the orders of all ixi minors, instead of principal minors, determine the invariant factors. An algorithm is also given to actually compute the invariant factors. This gives us another potentially very useful criterion for checking the MSSNS condition.

In Section 3.3, we present Theorem 3.3, which defines conditions equivalent to MSST. They include two known conditions, i.e. 1) the $\tilde{A}_{ii}$ being Hurwitz and 2) the system having well-behaved time-scales, and 3) a new condition which reveals the connection between the MSSNS and MSST conditions and the eigenvalues of A(ε). Specifically, if A(ε) satisfies MSSNS and the orders of the real parts of its eigenvalues are equal to or less than those of the corresponding imaginary parts then A(ε) has MSST and vice versa.[+] This condition implies that in addition to MSSNS, if the damping rates of the system modes are equal to or faster than the oscillation rates, then well-behaved time-scales are ensured and vice versa. For example, consider the following system,

---

[+] Because of our assumption that A(ε) is Hurwitz for $\varepsilon\in(0,\varepsilon_0]$, we know that the real parts of the eigenvalues are negative.

$$\dot{x} = \begin{bmatrix} -\epsilon & 1 \\ -1 & -\epsilon \end{bmatrix} x \qquad (3.1)$$

where $A(\epsilon)$ is stable and satisfies MSSNS since

$$\tilde{A}_{11} = A(0) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

which obviously is invertible. Note, however that $\lambda_i = -\epsilon \pm 1$. Therefore the damping rate here is of order 1 but the oscillation rate is of order 0. Thus the condition we have just stated is violated so that the system does not have well-behaved time-scales. Indeed, solving (3.1) we have

$$x_1(t) = x_{10}e^{-\epsilon t}\sin t$$

$$x_2(t) = x_{20}e^{-\epsilon t}\cos t$$

Substituting $\tau=\epsilon t$ we have

$$x_1(\tau) = x_{10}e^{-\tau}\sin t/\epsilon$$

$$x_2(\tau) = x_{20}e^{-\tau}\cos t/\epsilon$$

Thus the time behavior of $x(\tau)$ as $\epsilon \to 0$ shows oscillation but no damping, while the process $x(\tau)$ does not have a limit as $\epsilon \to 0$

because of rapid oscillations. Note, from the fact that $\tilde{A}_{11}$ is not Hurwitz we can come to the same conclusion.

In Section 3.4, we present an algorithm for checking MSSNS based on the results of Section 3.2. Finally, in Section 3.5 we summarize our results and present an additional result dealing with MSSNS and the detailed eigenstructure of $A(\epsilon)$.

3.2 Equivalent Statements of the MSSNS and MSST conditions

In this Section we prove several statements which are equivalent to MSSNS condition. We have already shown in Chapter 2. that the MSSNS condition is equivalent to $\tilde{A}_{ii}$ being non-singular. In this section we shall prove two new equivalent conditions -- 1) the orders of the eigenvalues of $A(\epsilon)$ being equal to the orders of the corresponding invariant factors and 2) the orders of the gcd's of the ixi principal minors of $A(\epsilon)$, i=1,...,n, uniquely determining the invariant factors and certain of these orders being equal to the orders of the sums of the corresponding principal minors. As mentioned in Section 3.1, the intuition behind the first result is based on the following observations. We have shown that the invariant factors of $A(\epsilon)$ determine the time-scale structure of the system if the system has well-behaved time-scale structure. On the other hand, the eigenvalues determine the modes of a system. Therefore if the system has well-

behaved time-scale behavior, the invariant factors and eigenvalues should have the same orders. In this section we prove this is true if and only if the system has MSSNS.

The second result is closely related to the first one. We shall see that the orders of the sums of the principal minors uniquely determine the orders of the eigenvalues. Thus if these orders are equal to the gcd's of the principal minors, in other words if there is no complete cancellation of the lowest order terms in the summation, then the latter will determine the orders of the eigenvalues too. Therefore from the first result we see that if in addition the gcd's of the principal minors also uniquely determine the orders of the invariant factors then $A(\epsilon)$ must have MSSNS because its eigenvalues and invariant factors must have the same orders. We show the first result in section 3.2.1. Then in section 3.2.2 we prove the second result and state the entire theorem.

3.2.1 Orders of eigenvalues and invariant factors

To show the first result we need the following lemma.

Lemma 3.1

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

Then, if A_{11} is invertible

$$\det A = \det A_{11} \det \tilde{A}_{22}$$

where

$$\tilde{A}_{22} = A_{22} - A_{21} A_{11}^{-1} A_{12}$$

Proof: see [ ].

Consider a matrix $A(\epsilon)$ over the ring T. Denote the orders of the eigenvalues and invariant factors by $b_1 \leq b_2 \leq \ldots \leq b_n$ and $a_1 \leq a_2 \leq \ldots \leq a_n$ respectively and the corresponding eigenvalues are $\lambda_1, \ldots, \lambda_n$. The Smith form of $A(\epsilon)$ is

$$A(\epsilon) = P(\epsilon) D(\epsilon) Q(\epsilon)$$

Let

$$A_1(\epsilon) = P^{-1}(\epsilon) A(\epsilon) P(\epsilon) = D(\epsilon) Q(\epsilon) P(\epsilon) = D(\epsilon) \bar{A}(\epsilon)$$

$$= \begin{bmatrix} I_0 & & & \\ & \epsilon^{k_1} I_1 & & \\ & & \ddots & \\ & & & \epsilon^{k_{n-1}} I_{n-1} \end{bmatrix} \begin{bmatrix} A_{11}(\epsilon) \ldots A_{1n}(\epsilon) \\ A_{21}(\epsilon) \ldots A_{2n}(\epsilon) \\ \ldots\ldots\ldots\ldots \\ A_{n1}(\epsilon) \ldots A_{nn}(\epsilon) \end{bmatrix}$$

where $I_0,\ldots,I_{n-1}$ are $n_0 x n_0$, $n_1 x n_1,\ldots,n_{n-1} x n_{n-1}$ identity matrices. Let $\tilde{A}_{ii}$, $i=1,\ldots,n$ denote the successive Schur complements of A(0) as described in Section 2. Then we state the first result as the following lemma.

Lemma 3.2

A(ε) has MSSNS if and only if The eigenvalues and invariant factors have the same orders; i.e. $b_i = a_i$, $i=1,\ldots,N$.

Proof

Since $A_1(\epsilon)$ is similar to A(ε), they have the same eigenvalues. Furthermore, because P(ε) is unimodular, it will not effect the invariant factors. Therefore we shall consider $A_1(\epsilon)$ instead of A(ε) in our proof. We shall use the fact that A(ε) having MSSNS is equivalent to the $\tilde{A}_{ii}$, $i=1,\ldots,n$ having full rank, a fact proven in last section.

We first prove the "only if" part.

Define

$$d(\epsilon,\lambda) = \det(\epsilon^{-kj} A_1(\epsilon) - \lambda I)$$

$$= \det \begin{bmatrix} \epsilon^{-kj}A_{11}(\epsilon) - I & \cdots & \epsilon^{-kj}A_{1n}(\epsilon) \\ \cdots & \cdots & \cdots \\ A_{j1}(\epsilon) & \cdots & A_{jj}(\epsilon) - I \quad \cdots \quad A_{jn}(\epsilon) \\ \cdots & \cdots & \cdots \\ \epsilon^{n-kj-1}A_{n1}(\epsilon) & \cdots & \epsilon^{n-kj-1}A_{nn}(\epsilon) - I \end{bmatrix} \qquad (3.2)$$

and

$$d_1(\epsilon,\lambda) = \det \begin{bmatrix} A_{11}(\epsilon)-\epsilon^{kj} I & \cdots & A_{1n}(\epsilon) \\ A_{21}(\epsilon) & \cdots A_{22}(\epsilon)-\epsilon^{kj-kl} I \cdots & A_{2n}(\epsilon) \\ \cdots & \cdots & \cdots \\ A_{j1}(\epsilon) & \cdots & A_{jj}(\epsilon) - I \quad \cdots \quad A_{jn}(\epsilon) \\ \cdots & \cdots & \cdots \\ \epsilon^{n-kj-1}A_{n1}(\epsilon) & \cdots & \epsilon^{n-kj-1}A_{nn}(\epsilon) - I \end{bmatrix}$$

Denote $d(\epsilon,\lambda)$ and $d_1(\epsilon,\lambda)$ by d and d_1 respectively. Then the difference between them is that in d_1 the first j-1 blocks of rows of the matrix have been mutiplied by ϵ^{kj}, ϵ^{kj-kl}, ...,$\epsilon^{kj-kj-1}$ respectively. Therefore

$$d_1 = \epsilon^{v} d$$

where

$$v = n_0k_j+n_1(k_j-k_1)+\ldots+n_{j-1}(k_j-k_{j-1})$$

Since ϵ is a positive number, d_1 and d will have the same zeros.

It is easy to see that d_1 is a continuous function of ϵ at

$\epsilon=0$ and

$$d_1(0,\lambda) = \det \begin{bmatrix} A_{11} & A_{12} & \cdots\cdots\cdots & A_{1n} \\ A_{21} & A_{22} & \cdots\cdots\cdots & A_{2n} \\ & \cdots\cdots\cdots\cdots\cdots & & \\ A_{j1} & A_{j2} & \cdots A_{jj}-\lambda I \cdots & A_{jn} \\ & \cdots\cdots\cdots\cdots\cdots & & \\ 0 & 0 & \cdots\cdots\cdots\ 0 & \lambda I \end{bmatrix}$$

since the $\tilde{A}_{ii}$ are of full rank for i=1,...,n. Repeatedly using Lemma 3.1, one can easily see that

$$d_1(0,\lambda) = \lambda^t |A_{jj} - I| \prod_{i=1}^{j-1} |\tilde{A}_{ii}| \qquad (3.3)$$

where

$$t = \sum_{i=1}^{n-1-n_j} n_j+i$$

Therefore $d_1(0,\lambda)$ has n_j zeros which are identical to the eigenvalues of $\tilde{A}_{jj}$.

Furthermore, since the zeros of d_1 are continuous functions of ϵ, they can be approximated by the eigenvalues of $\tilde{A}_{jj}$ for ϵ is small enough. Because the zeros of d_1 are the same as the zeros of d, in view of (3.2) the eigenvalues of $\epsilon^{-kj}A_1(\epsilon)$ can be approximated by the eigenvalues of $\tilde{A}_{jj}$. In other words, denote the eigenvalues of $A_1(\epsilon)$ by λ_i, i=1,...,N then there are n_j eigenvalues of $A_1(\epsilon)$ for which the

following inequalities hold. Given any positive number δ there is an c_j such that

$$|\lambda_{Mj+i}/\epsilon^{kj} - \lambda_i(\tilde{A}_{jj})| \leq \delta \quad , \quad i=1,\ldots,n_j, \text{ if } 0<\epsilon<\epsilon_j \tag{3.4}$$

where

$$M_j = \sum_{i=1}^{j-1} n_i \quad \text{and } \lambda_{Mj+i} \text{ is the } (M_j+i)\text{-th eigenvalue of } A_1(\epsilon).$$

Repeating this procedure for $j=0,\ldots,n-1$ we can finally prove that for any $\delta>0$, there is an ϵ_n

$$\epsilon_n = \min\{\epsilon_0,\ldots, \epsilon_{n-1}\}$$

such that if $0<\epsilon<\epsilon_0$ then

$$|\lambda_{Mj+i}/\epsilon^{kj} - \lambda_i(\tilde{A}_{jj})| < \delta \quad , \quad i=1,\ldots,n_j, \quad j=0,\ldots,n-1 \tag{3.5}$$

Since $\lambda_i(\tilde{A}_{jj}) \neq 0$, for $j=0,\ldots,n-1$, $i=1,\ldots,n_j$, these inequalities tell us that there are n_0 eigenvalues of order k_{j-1}. These are exactly the orders of the invariant factors of $A(\epsilon)$.

Now let us prove the "if" part.

Suppose that $\tilde{A}_{j,j}$ is not of full rank but $\tilde{A}_{ii}$, $i=0,\ldots,j-1$ are. Then from (3.4) we see that there is at least one eigenvalue of $\tilde{A}_{jj}$, say $\lambda_{i0}(\tilde{A}_{jj})=0$ and the corresponding eigenvalue of $A_1(\epsilon)$, λ_{i0}, has an order

higher than k_j. This implies that there are fewer eigenvalues with the order k_j than there are invariant factors.

Now, let us show the second result in Theorem 3.1 which relates the invariant factors with MSSNS condition.

3.2.2 The Invariant Factors and Principal Minors of A(ϵ)

As mentioned in 3.2.1, in this subsection we investigate some properties relating the MSSNS condition with the invariant factors and principal minors of A(ϵ). We prove in Lemma 3.4, that if A(ϵ) satisfies the MSSNS condition then the invariant factors will be uniquely determined by the orders of the gcd's of the principal minors and these orders must be equal to the orders of the sums of the principal minors. In Theorem 3.1 we show that the reverse statement is also true. Also a corollary of Theorem 3.1 shows that if the gcd of the principal minors are computed from the explicit form of A(ϵ) then the following statement is true: if the orders of the gcd of the principal minors uniquely determine the invariant factors then A(ϵ) has MSSNS. That is, in explicit form the no-cancellation condition is automatically satisfied if the invariant factors are uniquely determined by the orders of the gcd's of the principal minors.

It is well known that the eigenvalues of a matrix are determined by its principal minors of all sizes. Then it is obvious from the previous subsection that under MSSNS the principal minors also determine the invariant factors of A(ϵ). For a general matrix A(ϵ) (possibly not possessing MSSNS) the invariant factors can be determined from the gcd's of <u>all</u> ixi minors, i=1,...,n. We shall prove that under MSSNS, the <u>gcd</u> of all ixi <u>principal</u> minors, i=1,....,n, determine the invariant factors.

To show this, we first prove Lemma 3.3. Then we show that the orders of the gcd's of ixi principal minors and the orders of i-th invariant factors satisfy the conditions of Lemma 3.3. Based on this fact we prove the result.

Now let us first state and prove Lemma 3.3.

LEMMA 3.3.

Given a set of real numbers p_i, i=1,...,n, there is a unique set of real nubers s_i, i=1,...,n such that

F1 $s_1 \leq s_2 \leq \ldots \leq s_n$

F2 $\sum_{i=1}^{j} s_i \leq p_j$, j=1,...,n

F3 $\sum_{i=1}^{j} s_i = p_j$ if $s_j \neq s_{j+1}$ or j=n.

Proof:

This result is most easily seen pictorially. In Fig. 3.2.1, the x's denote the p_i's. First, we draw a line starting from origin and passing through at least one p_i and leaving all p_i's either on or above this line. Obviously there is one and only one line

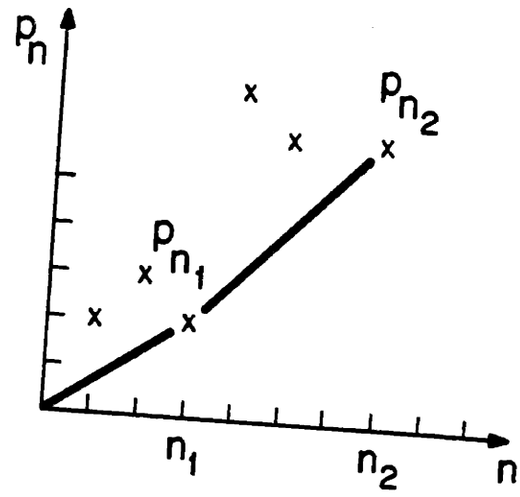


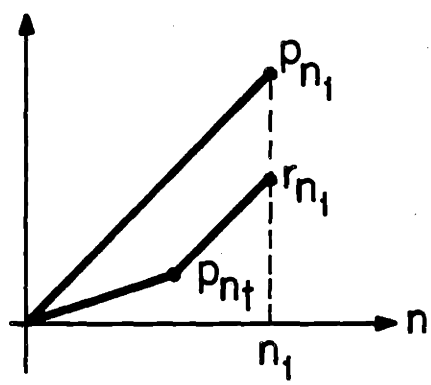
Figure 3.1

which satisfies this requirement. Let p_{n1} be the last of the p_j that lie on this line. If $p_{n1}=p_n$ the procedure is terminated. If not, then starting from p_{n1} we draw the second line which passes through at least one of the remaining p_i's ($n \geq i \geq n_1$) and leaves the other p_i's either on or above it. Let p_{n2} denote the last point through which this second line passes. The procedure is continuned until some line created passes p_n. We claim that the slope of these segments are the s_i's. Namely, $s_1=s_2=\ldots=s_{n1}=p_{n1}/n_1$, $s_{n1+1}=\ldots.=s_{n2}=(p_{n2}-p_{n1})/(n_2-n_1)$, etc. The points p_{n1}, p_{n2}, etc. will be called <u>turning</u> points.

We first prove that this is one solution. First, the slope of each succesive line segment is always larger than the previous one because otherwise the previous one must leave some points below it. This proves condition F1. Condition F2 follows from the fact that all p_i's are either on or above these segments. Conditon F3 holds because, for example, the segments op_{n1} and $p_{n1}p_{n2}$ must have different slopes by

definition of these segments.

To show the uniqueness, suppose we have another solution s'_i, $i=1,\ldots,n$. Let $\sum_{j=1}^{i} s'_j = r_i$. Obviously, $r_{n1} \leq p_{n1}$. Otherwise F2 will be violated. Suppose $r_{n1} < p_{n1}$. Since $p_1,\ldots,p_{n1}$ are either on or above op_{n1} by definition, the line or_{n1} must be a straight line too. If this were not the case, say, at some point n_t the slope changed, then p_{nt} would be on the line op_{nt} because of F3 and p_{nt} would be below op_{n1}, contradicting the construction of this line (see Fig. 3.2.2). In a similar fashion we can deduce that the curve $O-r_1-r_2-\ldots-r_m$ cannot change slope at <u>any</u> point. But we know that our curve $O-p_{m1}-p_{n2}\ldots$ meet at the end since $r_n=p_n$ by condition F3. This contradicts the hypotheses that $r_{n1} < p_{n1}$. The same argument works if $r_{ni}=p_{ni}$ $i=1,\ldots,k-1$ but $r_{nk} < p_{nk}$.



<u>Figure 3.2</u> : Illustrating a contradiction:
If $r_{n1} < p_{n1}$ and if the postulated curve from O to r_{n1} changes slope, then necessarily $p_{nt} < p_{n1}$, contradicting the construction of this line

Note: It should be pointed out that the algorithm proposed in this section is actually a "Newton polygon" concept. This is to be expected in the context of the present problem, c.f.[19]. However, we have not encountered in the literature a statement as simple as that given in the Lemma or to be given in statement 4 of that of Theorem 3.1.

Now we turn to the relation between gcd's of principal minors and invariant factors. We shall show in the following Lemma that their orders satisfy F1-F3 in Lemma 3.3.

Let us first define:

p_i = the order of the gcd of all ixi principal minors.

a_i = the order of the i-th invariant factor.

M_i^m, $i=1,\ldots,\binom{m}{n}$, $m=1,\ldots,n$ = mxm principal minors of $A(\epsilon)$.

$r_m = \text{Order}(\sum_i M_i^m)$.

Then we have

Lemma 3.4

If $A(\epsilon)$ T^{nxn} has MSSNS, then p_i and a_i satisfy F1-F3 (with the a_i playing the role of the s_i). Furthermore $r_m=p_m$ if $a_m \neq a_{m+1}$ (the "no-cancellation" condition).

Proof:

By definition,

$$a_1 \leq a_2 \leq \ldots \leq a_n \tag{3.6}$$

Also $\sum_{i=1}^{j} a_i$ is the order of the gcd's of all jxj minors. Therefore p_i and a_i satisfy F1 and F2. To show F3, let us recall that A(ϵ) having MSSNS also means that the invariant factors of A(ϵ) have the same orders as the corresponding eigenvalues. In other words,

$$b_i = a_i, \quad i=1,\ldots,n \tag{3.7}$$

where b_i are orders of the eigenvalues, λ_i.

It is known that

$$\det(\lambda I - A(\epsilon)) = \lambda^n - \lambda^{n-1}(\Sigma \lambda_i) + \lambda^{n-2}(\Sigma \lambda_i \lambda_j) + \ldots + \lambda_1 \cdots \lambda_n$$

$$= \lambda^n - \lambda^{n-1}(\Sigma M_i^1) + \lambda^{n-2}(\Sigma M_i^2) + \ldots + M_1^n \tag{3.8}$$

Where M_i^m, $i=1,\ldots,\binom{m}{n}$, $m=1,\ldots,n$ are principal minors of A(ϵ).

Compare the coefficients on both sides of (3.8) we see that

$$\sum_i M_i^m = \Sigma \lambda_{i1} \cdots \lambda_{im} \tag{3.9}$$

Now, suppose $a_m \neq a_{m+1}$. Under the MSSNS condition, in view of (3.6), (3.7), we see that among all the terms on the right hand side of (3.9),

$\lambda_1 \cdots \lambda_m$ has the smallast degree, say d_m, and any term in (3.9) other than $\lambda_1 \cdots \lambda_m$ must have an order greater than d_m. Therefore

$$r_m = \text{Order}\left(\sum_i M_i^m\right) = d_m = \sum_{i=1}^{m} a_i$$

$$= \text{order of gcd of all mxm minors} = p_m \qquad (3.10)$$

Since the order of the gcd of M_i^m (i.e. p_m) must be greater than or equal to d_m (since p_m is the gcd of a smaller set of minors), this equality in turn means that p_m is equal to d_m. That is

$$p_m = \sum_{i=1}^{m} a_i, \quad \text{if } a_m \neq a_{m+1}$$

This, together with the fact that $\sum_{i=1}^{n} a_i = \text{order}(M_1^n)$ yields F3, and (3.10) gives us the no-cancellation condition.

Before we state and prove Theorem 3.1, we still need the following lemma.

Lemma 3.5 The sets of numbers r_i and b_i satisfy F1 to F3, with the r_i playing the role of the p_i and the b_i the role of the s_i.

Proof: The conditions F1 and F2 are obtained simply by cheking the two sides of (3.8). The derivation of F3 is similar to that of (3.10).

Lemma 3.6 Define q_i i=1,...,n to be the orders of the gcd of all ixi minors. Then $r_i \geq p_i \geq q_i$.

Note: $q_i=a_1+....+a_n$ by definition of the invariant factors.

The proof is also trivial. Since the q_i are computed from a larger set of minors, the second inequality is then proved. The first inequality is due to the fact that the order of the sum of some terms is always greater than or equal to the order of the gcd of those terms.

Based on lemma 3.1 to 3.6 we can prove the following theorem.

Theorem 3.1

Suppose $A(\epsilon)\epsilon T^{nxn}$. Then the following statements are equivalent.

1. $A(\epsilon)$ satisfies the MSSNS condition.
2. The $\tilde{A}_{ii}$, i=1,...,n are invertible.
3. The orders of the eigenvalues are equal to that of the invariant factors, i. e. $a_i=b_i$.
4. (a) p_i and a_i i=1,...,n satisfy F1, F2 and F3.
   (b) $r_m=p_m$ if $a_m \neq a_{m+1}$ (the no-cancellation condition).

Proof:

1 <===> 2: This has been proven in the last section.

2 <===> 3: This has been proven in Lemma 3.2.

3 ===> 4: This is Lemma 3.4.

4 ===> 3: Define p_i^c the lower bound of the convex hull of p_i. Refering to Figure 3.2.1, p_i^c, i=1,...,n are the corresponding points lying on the solid line in the figure. We will on occasion refer to p_i^c

as the lower hull of p_i. Then from statement 4a we can conclude that $p_i^c=a_1+\ldots+a_i$. But this is in fact the definition of q_i. Therefore we have proved $p_i^c=q_i$. Define r_i^c the lower bound of the convex hull of r_i. Then Lemma 3.6 and statement 4b imply that $r_i^c=p_i^c$. Using Lemma 3.5 and statement 4a we then have that $a_i=b_i$, for $i=1,\ldots,n$.

Corollary If $A(\epsilon)$ is in explicit form, then 4a implies 4b in Theorem 3.5, so that 4a by itself is equivalent to MSSNS.

Proof: Since $A(\epsilon)$ satisfying the MSSNS condition implies that its explicit form $A_1(\epsilon)$ satisfies the same condition and vice versa, we shall consider $A_1(\epsilon)$ in the proof. Let

$$A_1(\epsilon) = \begin{bmatrix} A_{11} & A_{12} \cdots\cdots & A_{1n} \\ \epsilon^{k1}A_{21} & \epsilon^{k1}A_{22} \cdots\cdots & \epsilon^{k1}A_{2n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \epsilon^{kn-1}A_{n1} & \epsilon^{kn-1}A_{n2} \cdots\cdots & \epsilon^{kn-1}A_{nn} \end{bmatrix} \begin{matrix} \} n_0 \\ \} n_1 \\ \\ \} n_{n-1} \end{matrix}$$

Then from 4a we know

$$\sum_{j=1}^{i} a_j = p_i, \quad i=n_0, n_0+n_1,\ldots,n_0+\ldots+n_{n-1}, N$$

because $a_{n0}=a_{n0+1}$, $a_{n0+n1}=a_{n0+n1+1}$,.... That is, the order of the gcd

of all $n_0 x n_0$ principal minors is 0. The order of the gcd of all $(n_0+n_1)x(n_0+n_1)$ principal minors is $n_1 i_1$ and so on. It is easy to see that $r_{n0}=p_{no}=0$ since A_{11} is the only principal minor of order n_0 whose order is 0. Similarly, the submatrix containing A_{11}, A_{12}, $\epsilon^{il}A_{21}$, $\epsilon^{il}A_{22}$ is the only $(n_0+n_1)x(n_0+n_1)$ principal submatrix whose determinant can be equal to $n_1 k_1$. Thus $r_{n0+nl}=p_{n0+nl}$. Continune this argument we can finally prove this corollary.

Note:

Here is an example to show that in general condition 4a alone is not sufficient to ensure the MSSNS condition. Let

$$A(\epsilon) = \begin{bmatrix} 1 & -1 \\ 1 & -1+\epsilon^2 \end{bmatrix}$$

Here

$$p_i = 0,2$$

$$a_i = 0,2$$

So that p_i and a_i satisfy F1-F3. But A(0) does not have SSNS as is easily seen by checking that both its eigenvalues are 0 but A(0)=0. This is due to the fact that 2b is violated: $r_1=2\neq p_1=0$ while $a_1\neq a_2$.

From the proof of Theorem 3.1, especially (3.3), we see that if $A(\epsilon)$ satisfies MSSNS, then for any $\delta>0$, there is an ϵ_0, such that if $0<\epsilon<\epsilon_0$ then

$$|\lambda_{M_j+i}/\epsilon^k - \lambda_i(\tilde{A}_{jj})| \leq \delta, \quad \begin{matrix} i=1,\ldots,n_j \\ j=0,\ldots,n-1 \end{matrix}$$

or

$$|\lambda_{M_j+i}/\epsilon^k - \lambda_i(\tilde{A}_{jj})| = O(\epsilon), \quad \begin{matrix} i=1,\ldots,n_j \\ j=0,\ldots,n-1 \end{matrix}$$

or

$$|\lambda_{M_j+i} - \lambda_i(\epsilon^k \tilde{A}_{jj})| = O(\epsilon^{kj+1}), \quad \begin{matrix} i=1,\ldots,n_j \\ j=0,\ldots,n-1 \end{matrix}$$

Therefore we have the following theorem.

Theorem 3.2:

If $A(\epsilon)$ has MSSNS, the eigenvalues of $A(\epsilon)$ are clustered in n groups, with those in the k-th group lying within $O(\epsilon^{kj+1})$ of the eigenvalues of $\epsilon^{kj}\tilde{A}_{jj}$.

In Section 3.4, we use the results of this section to define an algorithm to determine the invariant factors of $A(\epsilon)$ from the gcd's of its principal minors under the condition that $A(\epsilon)$ satisfies the MSSNS condition. Before doing that, let us first consider some equivalent conditions for MSST.

3.3 Equivalent statements for MSST condition

Having seen the equivalent conditions for MSSNS, one would expect to see the similar conditions for MSST. As we will see however, there is one additional property of the eigenvalues that is required for MSST,

and this allows us to gain a deeper understanding of the MSST condition.

Theorem 3.3:

The following statements are equivalent.

1. $A(\epsilon)$ has MSST.

2. The $\tilde{A}_{ii}$, $i=1,\ldots,n$ are Hurwitz.

3. The system $\dot{x}=A(\epsilon)x$ has well-behaved time-scale structure.

4. $A(\epsilon)$ has MSSNS and $O[Re(\lambda_i)] < O[Im(\lambda_i)]$ for $i=1,\ldots,n$.

where the λ_i are the eigenvalues of $A(\epsilon)$ and the $O[Re(\lambda_i)]$ and $O[Im(\lambda_i)]$ are the orders of the real and imaginary parts of the eigenvalues of $A(\epsilon)$ respectively.

Proof:

The first three statements were proven in the last chapter. Here we only need to prove 2 <==> 4. First let us prove 4 ===> 2.

Since $A(\epsilon)$ has MSSNS, from theorem 3.1 we know that the eigenvalues of $A(\epsilon)$ are clustered in n groups. Denote the eigenvalues of $A(\epsilon)$ by $\lambda_{11},\ldots,\lambda_{1n_1},\lambda_{21},\ldots,\lambda_{k1},\ldots,\lambda_{kn_k},\ldots,\lambda_{nn}$, where, from

Theorem 3.2 and the MSSNS assumption

$$\lambda_{ji} = \lambda_i(\epsilon^{kj}\tilde{A}_{jj}) + O(\epsilon^{kj+1}), \quad \begin{matrix} i=1,\ldots,n_j \\ j=1,\ldots,n \end{matrix}$$

Since MSSNS implies that $\tilde{A}_{jj}$ is invertible, we see that

$$\lim_{\epsilon->0}(\lambda_{ji}/\epsilon^{kj}) = \lambda_i(\tilde{A}_{jj}) \neq 0 \qquad (3.11)$$

Furthermore, since we have assumed the $O[Re(\lambda_{ki})] \leq O[Im(\lambda_{ki})]$, we can immediately conclude that

$$c_{ji} = \lim(Re(\lambda_{ji})/\epsilon^{kj}) = Re(\lambda_i(\tilde{A}_{jj})) \neq 0$$

Since $A(\epsilon)$ is Hurwitz for $\epsilon\epsilon(0,\epsilon_0)$ we have that $c_{ji}<0$, which in turn implies that $\tilde{A}_{jj}$ is Hurwitz.

2 ==> 4

We have already seen that 2 ==> that $A(\epsilon)$ has MSSNS. Since (3.11) still holds and $Re[\lambda_i(\tilde{A}_{jj})] < 0$ by assumption, we immediately have that $O[Re(\lambda_{ji})] \leq O[Im(\lambda_{ji}]$.

As commented in Section 3.1, the fourth statement in this theorem states that if the system has MSSNS and if in addition the damping rate is at least as fast as the oscillation rate (i.e. the orders of the real parts of the eigenvalues are equal to or less than the orders of the corresponding imaginary parts) then the system has well-behaved

time-scales. The reverse is also true.

Thus, in this section we have proven a result which relates the MSST condition with several other conditions. In the following section we use the results of Section 3.2 to define an algorithm to determine the invariant factors of $A(\epsilon)$ from the gcd's of its principal minors under the condition that $A(\epsilon)$ satisfies the MSSNS condition.

3.4 An Algorithm For Determing The Invariant Factors Of A(ε)

Given certain real numbers p_i, i=1,...,n, there is a simple algorithm to determine a_i which satisfy F1-F3 in Lemma 3.3. The algorithm proceeds as follows. The first objective is to find the slope of the initial line segment and the first "turning point". We do this by computing for each p_i the slope $u_i=b_i/i$ of the line from the origin to the point p_i. Then the initial slope is the smallest of the u_i and the first turning point corresponds to the largest value of i that has this minimum value of u_i. From this point we essentially repeat the process by looking at the slopes of lines from this turning point to each of the p_i corresponding to subsequent values of i. A flow graph of an algorithm that accomplishes this is depicted in Figure 3.3. As indicated in this flow graph, each subsequent slope and turning point computation can be reduced to the original one by shifting the previous turning point to the origin.

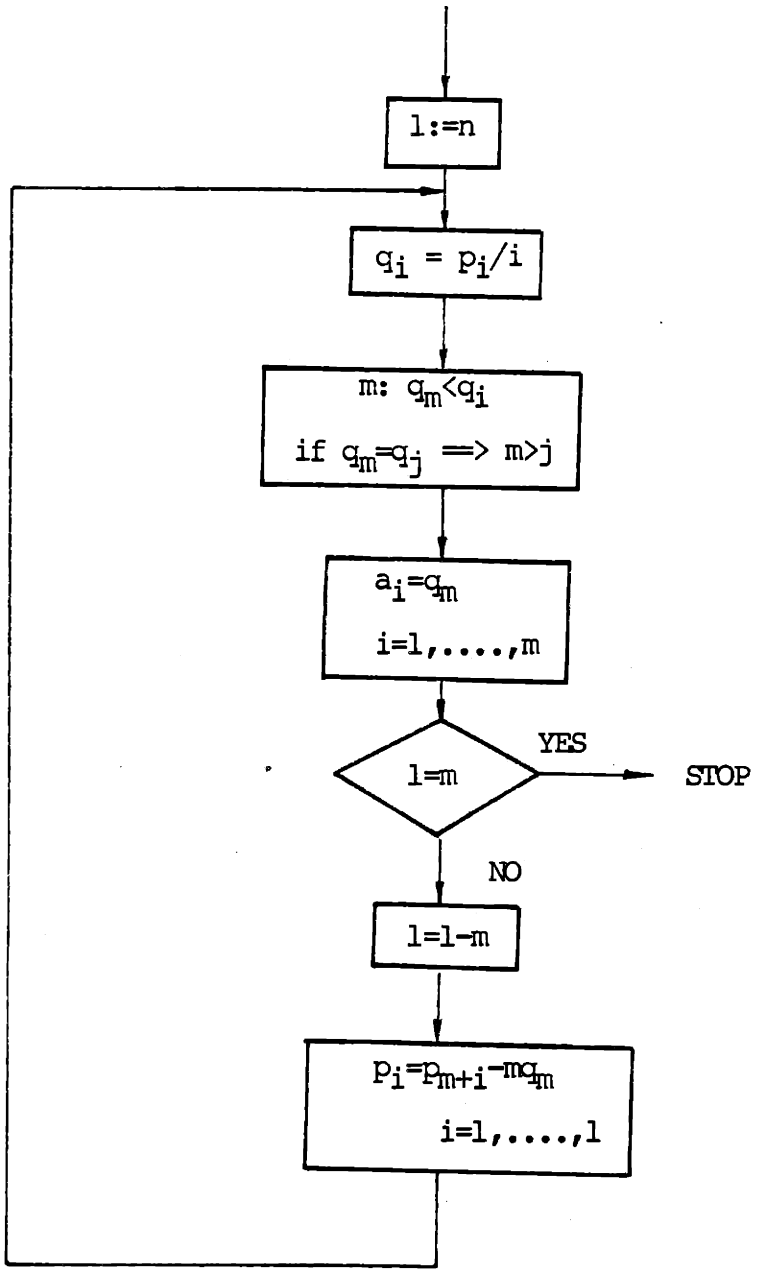


Figure 3.4.1 Flow graph of an algorithm

The following are some examples showing how this algorithm works.

Example 3.4.1

In this example, p_i=3,2,2,3,4,6 are as shown in Fig. 3.4.

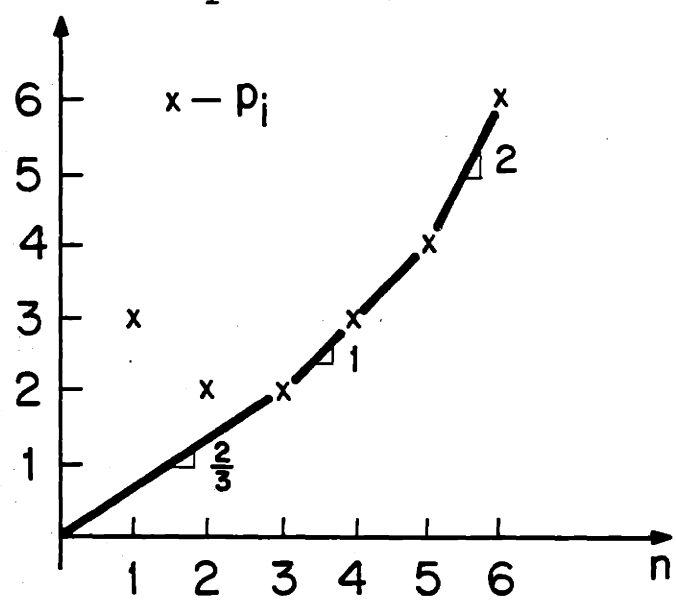


Fig. 3.4

The algorithm, illustrated in Figure 3.4.1, proceeds as follows

1. l=6
2. q_i= 3, 1, 2/3, 3/4, 4/5, 1
3. It is easy to check that m=3, q_m=2/3.
4. a_1=a_2=a_3=2/3.
5. m=n=6. Therefore continue.

6. $l=6-3=3$, $p_i=1,2,4$.

7. Repeat 1-4. We have $a_4=a_5=1$.

8. This time, $l=1$, $p_i=2=a_6$. Stop.

So we have $a_i = 2/3, 2/3, 2/3, 1, 1, 2$.

The next two examples show how to actually compute the invariant factors from a matrix satisfying the MSSNS condition.

Example 3.4.2

$$A(\epsilon) = \begin{bmatrix} \epsilon^2 & 1 & c & \epsilon^3 \\ \epsilon^2 & c & 1 & \epsilon^2 \\ \epsilon & 1 & 1 & \epsilon \\ \epsilon^5 & \epsilon & \epsilon^3 & \epsilon^5 \end{bmatrix}$$

The order of the gcd of M_i^1 is obviously 0, because there are diagonal elements of $A(\epsilon)$ that have order 0. Since one of the 2x2 principal minor is 1, we also have that $p_2=0$. We then find that p_3 is determined by the upper left 3x3 minor which is ϵ. Also $\det(A(\epsilon))=\epsilon^6$. So that

$$p_1=p_2=0,\ p_3=1,\ p_4=5.$$

If we also know that $A(\epsilon)$ satisfies MSSNS condition (this is true for this example), then using our algorithm we have

$$a_1=a_2=b_1=b_2=0$$

$$a_3=b_3=1,\ a_4=b_4=5.$$

This algorithm in fact also provides a way to check the MSSNS condition. We first observe that $r_i=p_i$. Then, since there is a 3x3 principal minor whose order is 1 and there is no other 3x3 minor of lower order, we know that the invariant factors are 0,0,1 and 5. The numbers determined by p_i according to the algorithm are 0,0,1 and 5 which are equal to the orders of the invariant factors. Therefore statement 4 of Theorem 3.1 tells us that $A(\epsilon)$ has MSSNS.

Example 3.4.3

$$A(\epsilon) = \begin{bmatrix} \epsilon & 1 & 1 & 1 \\ 1 & \epsilon^7 & \epsilon & \epsilon \\ \epsilon^2 & \epsilon^3 & 0 & \epsilon^3 \\ \epsilon^6 & \epsilon^{11} & \epsilon^{11} & \epsilon^{11} \end{bmatrix}$$

It can be seen that

$$p_1=1,\ p_2=0,\ p_3=3,\ p_4=14.$$

As in the preceding example, we can check that this $A(\epsilon)$ satisfies MSSNS, and our algorithm then yields the invariant factors:

$$a_1=a_2=0,\ a_3=3,\ a_4=11.$$

3.5 Conclusion

The new results in this Chapter are the following.

1. New equivalent conditions for MSSNS — 1) the orders of the eigenvalues being equal to those of the invariant factors. 2) $A(\epsilon)$ satisfies the no-cancellation condition and the p_i uniquely determine a_i (Theorem 3.1).

2. A new equivalent condition for MSST — $A(\epsilon)$ satisfies MSSNS and in addition the orders of the real parts of the eigenvalues are equal to or less than those of the corresponding imaginary parts (Theorem 3.3).

3. If $A(\epsilon)$ has MSSNS, then the eigenvalues of $A(\epsilon)$ are clustered in n groups (Theorem 3.2).

Based on these new results one can derive several conclusions.

1. The new condition for MSSNS does not require us to compute the successive Schur complements to determine if a system satisfies this condition. Instead, if the orders of the eigenvalues and invariant factors have been computed, then we can easily see if $A(\epsilon)$ has MSSNS as pointed out in Theorem 3.1, statement 3. This result is in fact the basis for our approach to scaling developed in the next chapter. This scaling involves the construction of a non-unimodular similarity

transformation so that the orders of the invariant factors can be changed to be eaual to the orders of the eigenvalues (which of course are left unchanged). If in addition the real and imaginary parts of the eigenvalues of $A(\epsilon)$ satisfy the order condition as stated in Theorem 3.3, then after scaling the new system matrix will not only have MSSNS but also MSST. So that, as stated in Section 2.5 and 2.6,the system will have well-behaved time-scales and the results in those sections can apply. If this condition is not satisfied, then using the results in Chapter 2, Section 2.8, after scaling we can still have the extended well-defined time scale behavior.

The second approach for checking the MSSNS condition is to use statement 4 in Theorem 3.1. Namely, we first check if $r_m=p_m$ when $a_m \neq a_{m+1}$,(the no-cancellation condition). If the answer is yes then we conclude that the system does not have MSSNS. If, on the other hand, they are equal, we proceed to check if the gcd's of the principal minors and the invariant factors satisfy F1-F3. The third method is to make use of the corollary of Theorem 3.1. That is, we first obtain the explicit form of $A(\epsilon)$. Then compute the p_i from the explicit form and check if they and a_i satisfy F1-F3.

2. The new condition for MSST is not only a new criterian for MSST, it is also a bridge between MSSNS and MSST. In other words, if $A(\epsilon)$ has MSSNS then we should only check the orders of the real and imaginary parts of its eigenvalues to make sure if it has MSST.

3. Theorem 3.2 provides a way to check the approximate locations of the eigenvalues if $A(\epsilon)$ satisfies MSSNS. In principal, if $A(\epsilon)$ does not satisfy MSSNS, but we can find a scaling to bring it to MSSNS, then this theorem can still apply to the transformed system. It holds regardless of the stability of the system. It also suggests a connection to frequency domain consideration. That is, even if $A(\epsilon)$ does not satisfies the stability condition of Theorem 3.3, as long as it satisfies the MSSNS condition, we should expect, in view of its clustered eigenvalues, well-defined multiple frequency-scales (cf. [20],[21]). In other words, the frequency response of the original system should be able to be approximated by the frequency response of a decoulpled frequency-scale separated system as the one suggested in Chapter 2, Theorem 2.1.

This Chapter has dealt to a great extent with the relationship between time scales, invariant factors, and eigenvalues. A natural question to ask is the relationship between time scales and the complete eigenstructure of $A(\epsilon)$. In particular, let us point out one result on the relationship between the Jordan form of $A(\epsilon)$ and the MSSNS condition. Let

$$A(\epsilon) = M(\epsilon)\Lambda(\epsilon)M^{-1}(\epsilon)$$

where $\Lambda(\epsilon)$ is the Jordan form of $A(\epsilon)$ and $M(\epsilon)$ is normalized (i.e. the eigenvectors are chosen to have the unit length). We then have the following reault. Suppose that $M(\epsilon)$ is unimodular. Then $A(\epsilon)$ has MSSNS

if and only if $\Lambda(0)$ has SSNS (i.e. if it contains no diagonal blocks of dimension greater than 1 with 0's on the diagonal). This follows from the observations that (1) $\Lambda(\epsilon)$ has MSSNS if and only if $\Lambda(0)$ has SSNS and (2) $M(\epsilon)$ is unimodular. Let us give some examples and make several comments about this result. Consider the matrix

$$A(\epsilon) = \begin{bmatrix} \epsilon & 0 \\ 0 & \epsilon \end{bmatrix}$$

In this case $M(\epsilon)=I$ and $\Lambda(0)=0$, from which we can deduce that $A(\epsilon)$ has MSSNS. On the other hand, consider

$$A(\epsilon) = \begin{bmatrix} \epsilon & 1 \\ 0 & \epsilon \end{bmatrix}$$

In this case $M(\epsilon)=I$ but $\Lambda(0)$ does not have MSSNS.

It is important to point out that there are many cases in which $M(\epsilon)$ is <u>not</u> unimodular. In this case it is perfectly possible for $A(\epsilon)$ to have MSSNS even if $\Lambda(0)$ does <u>not</u> have SSNS or for $A(\epsilon)$ <u>not</u> to have MSSNS even if $\Lambda(0)$ <u>does</u> have SSNS. Let us illustrate this with two examples.

Ex 3.5.1 Consider

$$A(\epsilon) = \begin{bmatrix} \epsilon & 1 \\ 0 & \epsilon^2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1/(\epsilon^2-\epsilon) \\ 0 & 1 \end{bmatrix}}_{M(\epsilon)} \underbrace{\begin{bmatrix} \epsilon & 0 \\ 0 & \epsilon^2 \end{bmatrix}}_{\Lambda(\epsilon)} \begin{bmatrix} 1 & 1/(\epsilon-\epsilon^2) \\ 0 & 1 \end{bmatrix}$$

Here $A(\epsilon)$ does not have MSSNS and $M(\epsilon)$ is not unimodular, while $\Lambda(0)$ has SSNS.

Ex 3.5.2 The matrix

$$A(\epsilon) = \begin{bmatrix} -\epsilon & \epsilon \\ 0 & -\epsilon \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & \epsilon^{-1} \end{bmatrix}}_{M(\epsilon)} \underbrace{\begin{bmatrix} -\epsilon & 1 \\ 0 & -\epsilon \end{bmatrix}}_{\Lambda(\epsilon)} \begin{bmatrix} 1 & 0 \\ 0 & \epsilon \end{bmatrix}$$

satisfies MSSNS condition and has one time scale ϵt. (This can be seen by checking the orders of the eigenvalues and invariant factors. They are 1 and 1 in this example). But obviously $\Lambda(0)$ does not have SSNS and $M(\epsilon)$ is not unimodular.

As seen from these examples, the MSSNS condition does not require $\Lambda(0)$ to have SSNS or $M(\epsilon)$ to be unimodular. It is an open questionto investigate the full relationship between MSSNS and the eigenstructure of $A(\epsilon)$.

CHAPTER 4

SCALING

4.1 INTRODUCTION

4.1.1 Motivation

Having seen the decisive role that MSSNS and MSST play, we now consider systems that do not have MSSNS. In this chapter we show that under specified conditions a scaling transformation on the states produces a system that has MSSNS, so that the formulation in previous chapters can apply.

We have already seen in Chapter 3 that, when the system does not have MSSNS, the orders of eigenvalues and invariant factors are not equal, and conversely. The only way to make the system have MSSNS is therefore to change the orders of its eigenvalues and/or those of its invariant factors. We choose the latter, since this will not change the system's dynamics and is more direct to implement. In fact, a similarity transformation through a <u>non-unimodular</u> matrix has the potential to meet this requirement, since it will not change the

eigenvalues but can modify the invariant factors.

This can be clarified by a simple example that we have inspected before, namely

$$\dot{x} = \begin{bmatrix} -\epsilon & 1 \\ 0 & -\epsilon \end{bmatrix} x$$

As mentioned in Chapter 3, Section 3.1, this system does not have MSSNS since the orders $\{b_i\}$ of its eigenvalues are 1,1 but the orders $\{a_i\}$ of its invariant factors are 0,2. This failure to satisfy the MSSNS condition could also be deduced from Fig. 4.1. From this figure, we see that the slopes of the lower bound of the $\{p_i\}$ are 1 and 1. Recall that points on the lower bound of the $\{p_i\}$ were defined as p_i^c in the last chapter. Recall also that we defined r_i as the order of the sum of all ixi principal minors, and r_i^c as points on their lower bound. Since in this example $r_i^c = p_i^c$, these slopes are also the $\{b_i\}$. Also, the invariant factor orders are 0 and 2, corresponding to the slopes of the dotted line segments making up the plot of q_i versus i, where we recall that q_i was defined as the gcd of all ixi minors.

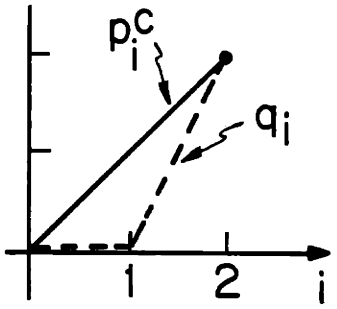


Figure 4.1 A simple example

Absence of MSSNS implies that the system does not have well-behaved time scales. For this example, as we have seen in Section 3.1, $x_1(t)$ becomes unbounded at the time scale ϵt as ϵ goes to zero.

However, if we apply the following linear transformation

$$S(\epsilon) = \begin{bmatrix} -\epsilon & 0 \\ 0 & 1 \end{bmatrix}$$

which is simply a (diagonal) scaling of state variables, we obtain the system

$$\dot{y} = S\dot{x} = S \begin{bmatrix} -\epsilon & 1 \\ 0 & -\epsilon \end{bmatrix} S^{-1}y$$

or $$\dot{y} = \begin{bmatrix} -\epsilon & \epsilon \\ 0 & -\epsilon \end{bmatrix} y = \epsilon \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} y$$

For this new system, it can be seen that the orders of eigenvalues and invariant factors are the same, so the system satisfies MSSNS. It is furthermore evident that the system satisfies MSST as well, so results we obtained in previous chapters apply to it. Returning to Fig. 4.1.1, we note that the orders $\{p_i\}$ of the gcd's of the principal minors are unchanged by diagonal scaling, but the orders $\{q_i\}$ of the gcd's of all minors have been raised by scaling, so that the two curves now coincide. The solution of the scaled system is given by:

$$y_1(t) = \epsilon x_1(t) = \epsilon x_{10} e^{-\epsilon t} + x_{20} \epsilon t e^{-\epsilon t}$$

$$y_2(t) = x_2(t) = x_{20} e^{-\epsilon t} \qquad (4.1)$$

Substitute τ for ϵt to get

$$y_1(\tau) = \epsilon x_{10} e^{-\tau} + x_{20} \tau e^{-\tau}$$

$$y_2(\tau) = x_{20} e^{-\tau}$$

Obviously this scaling has served to avoid the problem of the unscaled system, in that the result remains bounded for time scales τ and ϵt as ϵ goes to 0.

The significance of scaling is two-fold. First, as we have seen in Chapter 3, Theorem 3.2, if the system matrix satisfies the MSSNS condition, then its eigenvalues can be approximated by the eigenvalues of its successive Schur complements. Hence, if $A(\epsilon)$ does not have MSSNS

but we can find a scaling so that the resulting matrix does, then we can still obtain a good approximation to its eigenvalues via reduced-order computations, because scaling preserves eigenvalues. Second, since the scaled system has MSSNS, the procedure described in the previous chapter will provide a time scale decomposed system to approximate it, provided MSST is also satisfied. We conjecture that the solutions of this system can provide a good approximation to those of the original system, in a sense that will be discussed at the conclusion of this chapter.

4.1.2 Summary of the Scaling Procedure

This chapter concentrates on the development of a scaling procedure that can be applied to a large class of systems. The procedure involves the following steps. The first step is to transform $A(\epsilon)$ to its explicit form, $A_e(\epsilon) = P^{-1}(\epsilon)A(\epsilon)P(\epsilon) = D(\epsilon)\bar{A}(\epsilon)$. This step does not change either the invariant factors or the eigenvalues of $A(\epsilon)$.

The second step involves identifying what we term a skeleton in the explicit form. A skeleton consists of n entries of $A_e(\epsilon)$, precisely one from each row and column, with orders equal to the orders of the rows in which they are located. By the order of the i-th row, we mean the order a_i of the invariant factor of this row of the explicit form matrix $A_e(\epsilon)$. Each skeleton element therefore has minimal order for its row. Since $A(\epsilon)$ has full rank, it is easy to see that there is at least one skeleton in $A_e(\epsilon)$.

Now identify the skeleton above with the nxn permutation matrix that has 1's at the locations of the skeleton elements and 0's elsewhere. It is really only this pattern of the skeleton's placement within $A_e(\epsilon)$ that we shall be using in what follows, so we shall usually talk of the skeleton as if it is the permutation. Any permutation can be uniquely expressed as a product of disjoint cycles (see subsection 4.2.1). It follows from this that, perhaps after some re-ordering of the variables associated with our system (which corresponds to similarity transformation by a symmetric permutation), the elements of

the skeleton can be brought to the positions occupied by the 1's in a block diagonal canonical circulant matrix (BDCM), whose diagonal blocks take the form:

$$C = \begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & 0 \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ 1 & & & & 0 \end{bmatrix},$$

or simply [1] for a scalar block. This re-ordering of variables is the third step of our procedure. Let $\bar{a}_i$ denote the order of the skeleton element in the i-th row, following the re-ordering of variables. The set $\{\bar{a}_i\}$ is the same, therefore, as the set $\{a_1 \leq a_2 \leq \ldots \leq a_N\}$ of invariant factor orders.

For the moment, let us assume that there is only one block in the BDCM, so as to simplify our introduction. Then, under some assumptions on these integers $\bar{a}_i$ and on the principal minors of $A_\epsilon(\epsilon)$, the following scaling can be shown to transform the matrix to one that satisfies MSSNS:

$$S(\epsilon) = \text{diag}\{\epsilon^{s_1}, \epsilon^{s_2}, \ldots, \epsilon^{s_{N-1}}, 1\}, \qquad (4.2)$$

where

$$s_i = s_{i+1} + b_i - \bar{a}_i, \quad s_N = 0$$

and the b_i are, as before, the orders of the eigenvalues. The fourth (and final) step of our scaling procedure is thus the application of

this diagonal scaling transformation to the system.

In summary, the procedure for finding a diagonal scaling that induces MSSNS is

1. Transform $A(\epsilon)$ to its explicit form $A_e(\epsilon)$.
2. Identify a skeleton in $A_e(\epsilon)$.
3. Apply a symmetric permutation T to $A_e(\epsilon)$ so that the skeleton of $A_1(\epsilon) = TA_e(\epsilon)T^{-1}$ corresponds to a BDCM.
4. Calculate $S(\epsilon) = \text{diag}\{\epsilon^{s1},\ldots\epsilon^{sN}\}$ for $A_1(\epsilon)$ so that $A_2(\epsilon) = S(\epsilon)A_1(\epsilon)S^{-1}(\epsilon) = S(\epsilon)TA_e(\epsilon)T^{-1}S^{-1}(\epsilon)$ has MSSNS.

A precise statement of sufficient conditions for this procedure to succeed, as well as the details of Step 4 for a system with a multiple-block skeleton, are given in Section 4.2, along with proofs of the main assertions.

Note 1 We shall suppose in the remainder of this chapter (unless otherwise specified) that $A(\epsilon)$ is already in its explicit form, i.e. that step 1 has already been carried out.

Note 2 Rather than using the symbols $\{\bar{a}_i\}$ for the reordered set $\{a_i\}$, we shall from now on in this chapter use $\{a_i\}$ as generic symbols for invariant factors, with ordering determined by the context.

Note 3 In the next section we shall state several conditions that are together sufficient for our procedure to work. Before doing this,

we first note an important necessary condition for our scaling procedure to succeed. From the description of the procedure above, we see that only symmetric permutations and diagonal scaling will be applied to the explicit form $A_e(\epsilon)$. Obviously these operations have no effect on principal minors. However, from Theorem 3.1 we also know the condition that $p_i^c = r_i^c$, i=1,...N, i.e. that the $\{p_i\}$ uniquely determine the $\{b_i\}$, is a necessary condition for $A_e(\epsilon)$ to have MSSNS. (This condition was called the "no cancellation" condition in the previous chapter.) Therefore, if after scaling we have the MSSNS condition, so that the no cancellation condition holds, it must hold for the original matrix. We shall take this necessary no-cancellation condition as a standing assumption throughout this chapter.

4.1.3 Outline of this chapter

Section 4.2 is devoted to filling out the outline above. In Section 4.2.1 we briefly discuss some properties of permutations; further details can be found in Appendix 4A. Then in Section 4.2.2 we introduce the essential idea of our approach through a key example. Section 4.2.3 deals with diagonal scaling for the simplest case where there is only one cycle in the skeleton. Section 4.2.4 extends this result to the more general case where the skeleton has several cycles.

Since our scaling procedure is derived under some assumptions, it

is important to know that there are interesting problems where these assumptions are indeed satisfied. We describe, in Section 4.3, the application of our scaling procedure to high-gain feedback problems. In particular, we show that our procedure leads to scalings used by Sannuti in [14]. (This paper, along with [15], motivated our study of scaling in the first place.)

4.2 Diagonal Scaling

4.2.1 Introduction

In this section, we examine matrices $A(\epsilon)$ with structural features that cause them to not have MSSNS. We show that it is possible (under certain conditions) to choose a scaling of such a matrix that induces MSSNS.

Without loss of generality, it is assumed that $A(\epsilon)$ is in its explicit form (see Note 1 of Section 4.1). As mentioned in that section, we shall be considering transformations of the form $S(\epsilon)T$, where $S(\epsilon)$ is a diagonal scaling matrix, T is a permutation matrix, and

$$\bar{A}(\epsilon) = S(\epsilon)TA(\epsilon)T^{-1}S(\epsilon)^{-1} = \bar{S}(\epsilon)A(\epsilon)\bar{S}(\epsilon)^{-1} \qquad (4.3)$$

Note that $A(\epsilon)$ and $\bar{A}(\epsilon)$ have the same eigenvalues but may have different invariant factors if $\bar{S}(\epsilon)$ is not unimodular. We shall show how $\bar{S}(\epsilon)$ can be chosen so that the invariant factors and eigenvalues of $\bar{A}(\epsilon)$ have the same orders, even when $A(\epsilon)$ does not possess this property.

The development of our scaling procedure involves identifying and manipulating permutation matrices that reflect important structural features of $A(\epsilon)$. Before we begin this development, therefore, let us first review some properties of permutation matrices, i.e. of matrices

obtained by permuting the rows and/or columns of the identity matrix I.

A permutation matrix, say

$$P = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

can also be represented in terms of the cycles that comprise it. For the example above, the decomposition of P into cycles is given by

$$\theta_p = (1,4)(2,3,6)(5)$$

This notation serves to indicate that if P is applied to a vector $x=(x_1,x_2,\ldots,x_6)$, then the following cycling of elements occurs: $[x_1 \text{--}> x_4, x_4 \text{--}> x_1], [x_2 \text{--}> x_3, x_3 \text{--}> x_6, x_6 \text{--}> x_2], [x_5 \text{--}> x_5]$. Here θ_P consists of three cycles, which are uniquely defined.

The canonical circulant matrix is a permutation of the form

$$\theta_c = (n, n-1, n-2, \ldots, 2, 1)$$

This corresponds to the permutation or circulant matrix

$$\begin{bmatrix} 0 & 1 & 0 & \ldots\ldots & 0 \\ 0 & 0 & 1 & \ldots\ldots & 0 \\ 0 & 0 & 0 & \ldots\ldots & 0 \\ & \ldots\ldots\ldots\ldots & & & \\ 0 & 0 & 0 & \ldots\ldots & 1 \\ 1 & 0 & 0 & \ldots\ldots & 0 \end{bmatrix}$$

A block diagonal canonical circulant matrix (BDCM) is defined as

$$C = \text{block diagonal } \{C_1, C_2, \ldots, C_m\}$$

where C_i, $i=1,\ldots, m$ are canonical circulant matrices.

It can be shown (see Appendix 4A) that any permutation matrix can be brought to a BDCM by a symmetric similarity transformation T that is itself a permutation matrix:

$$\bar{P} = TPT^{-1}, \quad \bar{P} \text{ is BDCM form.}$$

We shall call this transformation T a symmetric permutation. Each block of $\bar{P}$ corresponds to one cycle of θ_p.

In the next subsection we introduce our approach to scaling through an example that also serves to illustrate the role of circulant matrices in our development.

4.2.2 A Key Example

The essential ideas underlying the scaling approach proposed in this section are exposed by the following key example. To proceed, let us first recall some definitions from the previous chapter:

a_i = orders of invariant factors, $i=1,\ldots,N$

b_i = orders of eigenvalues, $\quad i=1,\ldots,N$

$$q_i = \sum_{j=1}^{i} a_j$$

and, based on the no cancellation assumption, the points on the lower bound of the $\{p_i\}$ are given by

$$p_i{}^c = r_i{}^c = \sum_{j=1}^{i} b_j,$$

Now, suppose $A(\epsilon)$ has the following special form:

$$A(\epsilon) = \begin{bmatrix} 0 & \epsilon^{a_1} & 0 & 0 & \ldots\ldots & 0 \\ 0 & 0 & \epsilon^{a_2} & 0 & \ldots\ldots & 0 \\ & & \ldots\ldots\ldots\ldots\ldots & & & \\ 0 & 0 & 0 & 0 & \ldots\ldots & \epsilon^{a_{n-1}} \\ \epsilon^{a_n} & 0 & 0 & 0 & \ldots\ldots & 0 \end{bmatrix} \qquad (4.4)$$

Note that A becomes the canonical circulant matrix if nonzero entries are replaced by 1; we shall therefore say A has the structure of a circulant matrix. It is easy to see that A is in explicit form, and that the orders of its invariant factors are a_i, $i=1,\ldots,n$. (Recall from Note 2 of the previous section that we are still using $\{a_i\}$ to denote the orders of the invariant factors, but they may <u>not</u> be ordered as in earlier chapters.) Since the ixi pricipal minors with $i<n$ are all zero (this implies $r_i=p_i=$, $i=1,\ldots,n-1$) and $r_n=p_n=q_n=a_1+\ldots+a_n$, we can

draw the following picture:

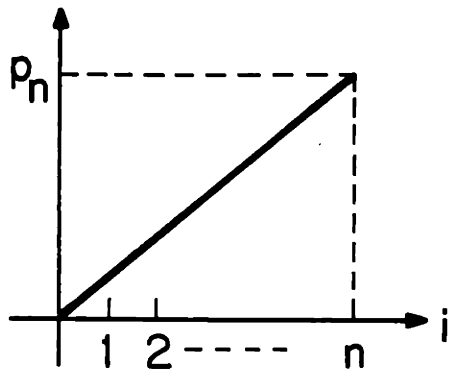


Figure 4.2 A key example

It is easy to see from this picture that the orders of the eigenvalues b_i, which are determined by $r_i{}^c=p_i{}^c$, are equal to $b=q_n/n$ for all i. This can also be obtained by checking the characteristic polynomial of A, which is

$$\lambda^n - \epsilon^v$$

where $v=q_n$. This again shows that all of the b_i are equal, namely $b_i=b=q_n/n$ for $i=1,\ldots,n$. Thus A will not have MSSNS unless all of the a_i are the same.

Now let

$$S = \text{diag}\{\epsilon^{s_1}, \epsilon^{s_2}, \ldots., \epsilon^{s_{n-1}}, 1\}$$

with

$$s_i - s_{i+1} = b - a_i, \quad i=1,\ldots.,n-1,$$

$$s_n = 0$$

Then after scaling

$$\bar{A} = SAS^{-1} = \begin{bmatrix} 0 & \epsilon^b & 0 & 0 & \cdots\cdots & 0 \\ 0 & 0 & \epsilon^b & 0 & \cdots\cdots & 0 \\ & & \cdots\cdots\cdots\cdots & & & \\ 0 & 0 & 0 & 0 & \cdots\cdots & \epsilon^b \\ \epsilon^b & 0 & 0 & 0 & \cdots\cdots & 0 \end{bmatrix}$$

and the invariant factors of $\bar{A}$ have the same orders as its eigenvalues.

If A does not have the structure of a circulant matrix, but of a more general permutation matrix, then we can apply a symmetric permutation to bring it to block circulant structure. For example, consider

$$A = \begin{bmatrix} 0 & 0 & \epsilon^{a1} \\ \epsilon^{a2} & 0 & 0 \\ 0 & \epsilon^{a3} & 0 \end{bmatrix}$$

As with the last example, A is in explicit form and the orders of the invariant factors are a_1, a_2 and a_3. The orders of the eigenvalues are all equal to $b=(a_1+a_2+a_3)/3$. Choosing the symmetric permutation matrix

$$T = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

we obtain

$$A_1 = TAT^{-1} = \begin{bmatrix} 0 & \epsilon^{a2} & 0 \\ 0 & 0 & \epsilon^{a1} \\ \epsilon^{a3} & 0 & 0 \end{bmatrix}$$

To bring A_1 to MSSNS, we can choose

$$S = \text{diag}\{\epsilon^{s1}, \epsilon^{s2}, 1\}$$

where $s_1=2b-a_1-a_2$, $s_2=b-a_1$. Applying S to A_1, we obtain

$$STAT^{-1}S^{-1} = \underbrace{\begin{bmatrix} 0 & \epsilon^{s2} & 0 \\ \epsilon^{s1} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{S} A \underbrace{\begin{bmatrix} 0 & \epsilon^{-s1} & 0 \\ \epsilon^{-s2} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{S^{-1}}$$

$$= SAS^{-1} = \begin{bmatrix} 0 & \epsilon^{b} & 0 \\ 0 & 0 & \epsilon^{b} \\ \epsilon^{b} & 0 & 0 \end{bmatrix}$$

In general, the entries of A that were zero in this example will not be zero. Nevertheless, the essential idea can still be carried over to the more general situation, under some conditions. In the next subsection we start with the special case where A has the form of

(4.4), except that the zero entries of (4.4) might no longer be zero (though they are still, in some sense, subsidiary). In other words, we shall assume that there is a skeleton in A, as defined in Section 4.1, whose elements correspond to the 1's in a circulant matix with one cycle. We move to the more general case, where the skeleton elements form a permutation matrix with several cycles, in the subsection after the next one.

4.2.3 Scaling for matrices with one-cycle skeleton

We shall now prove that the scaling procedure described in Section 4.1.2 works for the simplest case, where a skeleton of $A(\epsilon)$ (in its explicit form, as assumed in Note 1 of Section 4.2.1) has only one cycle. We shall then move to the more general case where all skeletons in $A(\epsilon)$ has multiple cycles. Recall that the skeleton of $A(\epsilon)$ having one cycle means that, after a symmetric permutation, the skeleton will be in BDCM form with only one block, while multiple cycles implies that the BDCM form will have multiple blocks.

Before stating the three assumptions which are retained throughout this and the next subsections, we state some definitions that we need.

The <u>dimension</u> of a principal submatrix P of matrix A is the number of rows in P. The <u>order</u> of a principal submatrix is the order of its determinant, i.e. the order of the associated principal minor. A <u>diagonal</u> of a principal submatrix of dimension n is defined as a set of n elements containing precisely one entry from each row and each column of P. The diagonal containing the entries $a_{m1,m2}$, $a_{m2,m3}$,....,$a_{mn,m1}$ will be represented as the ordered set

$$S = (m_1, m_2, \ldots, m_n)$$

There are n! diagonals in P. If P is the matrix A itself (in explicit form), then some diagonals will be the skeletons defined before. The

order of a diagonal is defined as the order of the product of all its entries. There is one special diagonal of any leading principal submatrix, namely

$$S_c = (1, 2, \ldots, n)$$

that will be called the canonical diagonal. A contiguous principal submatrix is defined as a principal submatrix with contiguous rows. It will be represented by two numbers, the idex of its last row i and the index of its first column j, and denoted as M[i,j], with a subscript on M to indicated dimension when needed. The overlap of two contiguous principal submatrices is defined as the number of common rows. It is an integer ranging from 0 (no overlap) to m_1, the smaller of the dimensions of the two submatrices. Finally, the closure of a principal submatrix M is defined as the smallest dimension contiguous principal submatrix containing M.

We can now state three assumptions under which our scaling procedure is derived, and the two lemmas required to derive the scaling.

Assumption 4.1. $b_i \geq a_j$, for i,j=1,...,N-1.

If we draw the curves for p_i^c and q_i as in Fig. 4.2.2, this assumption means that the slope of p_i^c is larger than that of q_i except for the last segment. Note that this is not a necessary condition for A(ε) to have a diagonal scaling. For example, consider Eq. (4.2.2). We

have seen that the diagonal scaling defined in that section always work no matter b (the order of the eigenvalues in this example) is smaller than a_i or not. But this is the condition under which our algorithm has been shown to work.

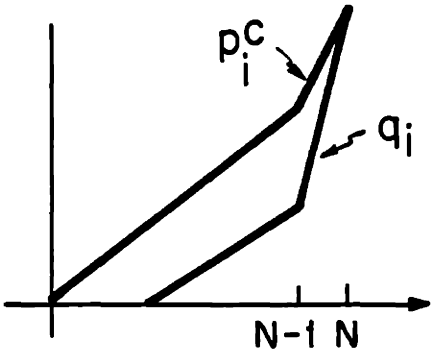


Fig. 4.3 p_i^c and q_i for Assumption 1

Assumption 4.2:

The order of any principal submatrix of $A(\epsilon)$ is equal to the smallest order of its diagonals.

This assumpotion serves to rule out certain cancellation when summing procucts of elements in different diagonals, ensuring that the order of any principal minor is not greater than the smallest order of the diagonals that determine the principal minor. Again, this is not a necessary condition for a diagonal scaling to exest.

The third assumption is motivated by the following example. Consider

$$A(\epsilon) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \epsilon^3 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \epsilon^3 & 0 & 0 & 1 \\ \epsilon^9 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

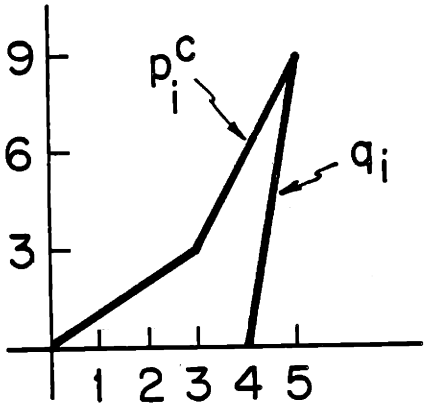


Figure 4.4

There are two principal submatrices M_1 and M_2 whose orders are 3. In order to bring q_i to p_i^c, we ought to scale three of the 1's, e.g. a_{12}, a_{23} and a_{34}, to ϵ and scale the remaining two 1's, namely a_{45}, a_{56}, to ϵ^2. Since the diagonal scaling will not change the principal minors, we shall end up with some entries of order less than 1, no matter how we arrange the scaling. (After scaling the order of the gcd of all the 1x1 minors should be equal to 1 since the order of b_1 is 1 in this example.) The source of the problem here is the overlap of two principal submatrices whose orders are equal to p_3^c, where 3 is the first turning point in the graph of Fig.4.4. To avoid this situation, we impose two more assumptions.

Assumption 4.3:

Suppose the order of an mxm contiguous principal submatrix M_1 is $p_m{}^c$. If there is an sxs contiguous principal submatrix M_2 that overlaps with M_1 and the overlap is t (an integer from 0 to min{m,s}), then the order of M_2 satisfies

$$O(M_2) \geq b_{m-t+1}+b_{m-t+2}+\ldots+b_{m-t+s+1}$$

Consider the previous example again. If A(ε) satisfied this assumption, i.e. if one of the M_i, i=1,2 had order 1+2+2=5 instead of 3, the difficulty encountered in that example would have been eliminated. From this assumption we also can see that any contiguous M_i whose order is $p_{mi}{}^c$ must be the unique contiguous principal submatrix that has this order.

Assumption 4.4:

For each i=1,2,.... the principal submatrix M_i of A(ε) whose determinant is $p_{mi}{}^c$ is a contiguous principal submatrix. Furthermore, M_i M_{i+1}, i=1,...., i.e. the rows of the M_i are contained in those of the M_{i+1}.

The following theorem now states that the scaling procedure described in Section 4.1.2 does indeed induce MSSNS under Assumptions 4.1-4.4. Since the steps 1-3 can be applied to any A(ε), we shall suppose that the A(ε) discussed in Theorems 4.1 and 4.2 has already gone through these steps. In other words, we suppose that some skeleton of

$A(\epsilon)$ already forms a BDCM of one cycle.

Define

m_i : turning points of q_j^c.

M_i : principal submatrix whose determinant is equal to p_{mi}^c.

f_i : slope of segement of q_j^c between m_{i-1} and m_i.

I_j : set of row indices of M_j.

We now have the following theorem.

Theorem 4.1: Suppose $A(\epsilon)$ satisfies Assumptions 4.1-4.4, and has a skeleton that forms a BDCM of one cycle. Then the scaling matrix

$$S = \text{diag}\{\epsilon^{s1}, \ldots, \epsilon^{sN}\} \tag{4.5}$$

causes SAS^{-1} to have MSSNS, where $s_1, \ldots, s_N$ are defined by

$s_i - s_{i+1} = f_j - a_i$, if $i \in I_j$, $i=1,\ldots,N-1$.

$s_N = 0$.

Proof:

According to Assumption 4.4, the M_i are contiguous. Now, consider M_1 and M_2. As a consequence of Assumption 4.4, M_1 is contained in M_2 and each entry a_{ij} of M_1, $i>j$, must belong to a contiguous principal submatrix $M[i,j]$ containing $a_{j,j+1}, \ldots, a_{i-1,j}$, which are the elements

of the canonical diagonal of the original matrix A and have orders $a_j, a_{j+1}, \ldots, a_{i-1}$ respectively. We then have

$$O(a_{ij}) + a_j + a_{j+1} + \ldots + a_{i-1} \geq (i-j+1)f_1, \quad i>j$$

The scaling defined in this theorem satisfies

$$s_j - s_{j+1} = f_1 - a_j, \ldots\ldots, \; s_{i-1} - s_i = f_1 - a_{i-1}$$

Denote the order of the (i,j)-th entry after scaling as $O(\bar{a}_{ij})$. Considering the fact that the diagonal scaling does not change the principal minors, we have

$$\begin{aligned} & O(\bar{a}_{ij}) + (a_j + s_j - s_{j+1}) + \ldots + (a_{i-1} + s_{i-1} - s_i) \\ & = O(\bar{a}_{ij}) + (a_j + f_1 - a_j) + \ldots + (a_{i-1} + f_1 - a_{i-1}) \geq (i-j+1)f_1 \end{aligned}$$

or $$O(\bar{a}_{ij}) \geq f_1, \quad i>j \qquad (4.6)$$

The entry a_{ij}, $i<j$, in the upper right triangle of M_1 has an order equal to or higher than f_1 before scaling. After scaling, for $i<j$, and considering $f_1 \geq a_i$, $i=1,\ldots,N-1$, we have

$$\begin{aligned} O(\bar{a}_{ij}) &= O(a_{ij}) + s_i - s_j = O(a_{ij}) + (s_i - s_{i+1}) + \ldots + (s_{j-1} - s_j) \\ &= O(a_{ij}) + (j-i)f_1 - a_i - \ldots - a_{j-1} \\ &\geq f_1, \quad i<j \end{aligned} \qquad (4.7)$$

Also $$O(a_{ii}) = O(\bar{a}_{ii}) \geq f_1 \qquad (4.8)$$

Combining (4.6)-(4.8) we can conclude that, after scaling, every entry in M_1 has an order equal to or higher than f_1.

Now consider the entries contained in M_2 but not in M_1 (see Fig. 4.5). Using the same argument as before and the Assumption 4.3, we have for a_{ij}, $i>j$, $i \notin I_1$,

$$O(a_{ij})+a_j+\ldots.+a_{i-1} \geq tf_1+(j-i+1-t)f_2 \qquad (4.9)$$

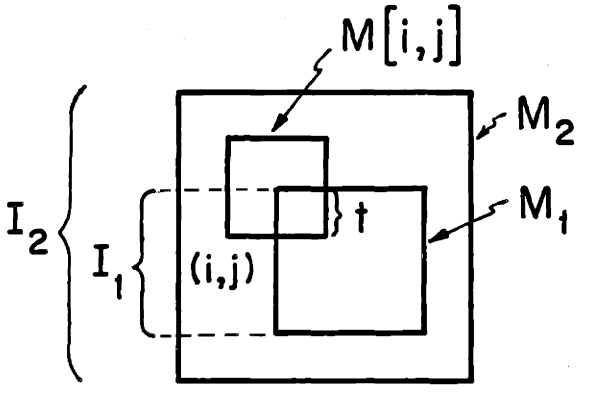


Figure 4.5

where t is the overlap of M_1 and M[i,j]. Since, after the scaling defined in this theorem, there will be exactly (j-i+1-t) canonical diagonal elements of M[i,j] having order f_2, and t-1 having order f_1, from (4.9) we have

$$O(\bar{a}_{ij}) \geq f_1, \quad i>j, \; i \in I_1 \qquad (4.9)$$

Arguments similar to those used in deriving (4.7) will prove $O(\bar{a}_{ij}) \geq f_1$, $i<j$, $i \in I_1$. Repeating this procedure for $\bar{a}_{ij}$, $i \in I_2 \backslash I_1$, we can prove

$$O(\bar{a}_{ij}) \geq f_2, \quad i \in I_2 \backslash I_1$$

This procedure can easily be extended to M_3, M_4, ... and finally we can show

$$O(\bar{a}_{ij}) \geq f_k, \quad i \in I_k \backslash I_{k-1}$$

Since there is one set of ixi minors, i=1,...,N, whose orders are exactly q_i (the $a_{i,i+1}$, $i \in I_k \backslash I_{k-1}$), the theorem is proved.

The following example illustrates our approach.

Example 4.1 Consider

$$A(\epsilon) = \begin{bmatrix} \epsilon^3 & \epsilon^4 & \epsilon^5 & \textcircled{1} \\ \textcircled{\epsilon} & \epsilon^3 & \epsilon^3 & \epsilon \\ \epsilon^3 & \textcircled{\epsilon} & \epsilon^2 & \epsilon^7 \\ \epsilon^6 & \epsilon^8 & \textcircled{\epsilon^6} & \epsilon^7 \end{bmatrix}$$

Before applying Theorem 4.1, we have to apply steps 1-3 to $A(\epsilon)$. First we identify the only skeleton, which comprises the circled elements in the matrix above, and corresponds to the permutation (1,2,3,4). To transform this permutation to canonical circulant form, we use the symmetric permutation θ_t=(1,3). Then

$$A_1(\epsilon) = TA(\epsilon)T^{-1} = \begin{bmatrix} \epsilon^2 & \epsilon & \epsilon^3 & \epsilon^7 \\ \epsilon^3 & \epsilon^3 & \epsilon & \epsilon \\ \epsilon^5 & \epsilon^4 & \epsilon^3 & 1 \\ \epsilon^6 & \epsilon^8 & \epsilon^6 & \epsilon^7 \end{bmatrix}$$

This matrix has a_i=1,1,0,6, and b_i=2,2,2,2. In other words, m_1=4, f_1=2. Hence, using Theorem 4.1, we have $s_1-s_2=f_1-a_1=1$, $s_2-s_3=f_1-a_2=1$, $s_3-s_4=f_1-a_3=2$, or s_1=4, s_2=3, s_3=2, s_4=0. This leads to

$$A_2(\epsilon) = S(\epsilon)A_1(\epsilon)S(\epsilon)^{-1} = STA(\epsilon)T^{-1}S^{-1} = \begin{bmatrix} \epsilon^2 & \epsilon^2 & \epsilon^5 & \epsilon^{11} \\ \epsilon^2 & \epsilon^3 & \epsilon^2 & \epsilon^4 \\ \epsilon^3 & \epsilon^3 & \epsilon^2 & \epsilon^2 \\ \epsilon^2 & \epsilon^5 & \epsilon^4 & \epsilon^7 \end{bmatrix}$$

It is easy to check that $A_2(\epsilon)$ has MSSNS, since its eigenvalues (which are the same as those of $A(\epsilon)$) and invariant factors have the same orders.

4.2.4 Scaling For Matrices With Two-Cycle Skeleton

We have derived the scaling of Theorem 4.1 for the case where the skeleton only has one cycle. Now we shall extend this result to a more general case where the skeleton has two cycles. From the results in Section 4.1, we know that after some symmetric permutation $A(\epsilon)$ can always be brought to the following form

$$A(\epsilon) = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{4.10}$$

where the skeletons of A_{11} and A_{22} are BDCM with one cycle and A_{11} is an mxm matrix. We then have the following lemma.

Lemma 4.1:

If $A(\epsilon)$ has the form of (4.10) and satisfies Assumptions 4.1 and 4.2, then

1. $a_1=b_1=....=a_m=b_m = b$ (4.11)

2. $a_j \leq a_i$, $j=m+1,....,N-1$, $i=1,....,m$. (4.12)

Proof:

From the assumptions of this lemma, we know that A_{11} is a BDCM with one cysle and that its determinant is a principal minor of order $a_1+....+a_m$. But any mxm principal minor must be equal to or larger than $p_m{}^c=b_1+b_2+....+b_m$. Therefore

$$a_1+....+a_m \geq b_1+....+b_m \tag{4.13}$$

On the other hand, from Assumption 4.1 we know that

$$a_i \leq b_j, \quad i,j=1,....,m-1 \tag{4.14}$$

Combining (4.13) and (4.14) we have (4.11). Using Assumption 4.1 and (4.11) we have (4.12).

From now on we shall suppose that A(c) has the form of (4.10) and satisfies (Assumptions 4.1 and 4.2, and therefore) (4.11) and (4.12). The required scaling for this case will be stated in the following theorem. In order to prove this theorem, we need to augment Assumption 4.3 by the following Assumption 4.3a. We also introduce two new assumptions, Assumption 4.5 and 4.6.

Assumption 4.3a

Suppose the order of an mxm contiguous principal submatrix M_1 is $p_m{}^c$. If there is an sxs principal submatrix N_1 whose closure M_2 of dimension r overlaps M_1, and if the overlap is t, then the order of N_1 satisfies

$$O(N_1) \geq b_{m+r-t-s}+b_{m+r-t-s+1}+\ldots+b_{m+r-t-1} = h$$

where the sum above that we have denoted by h is as marked on the figure.

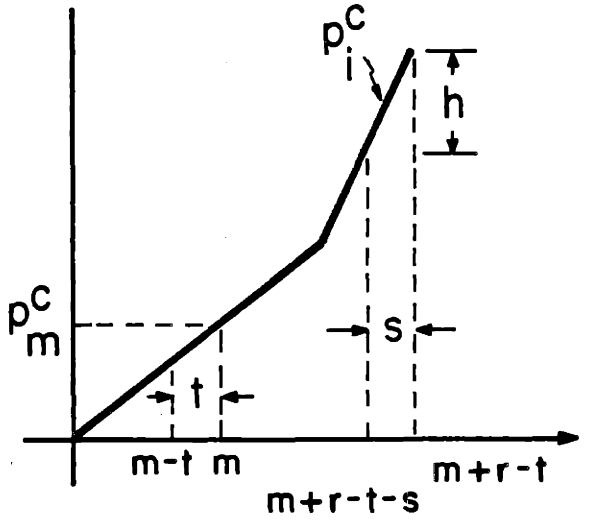


Figure 4.6

Assumption 4.5:

A(ϵ) has the form of (4.10) and the diagonals contained in A_{11} and A_{22} determine the $q_{mi}{}^{c}$.

This assumption implies that there is at least one diagonal of order m_i that is contained in A_{11} and A_{22} and has the order of $q_{mi}{}^{c}$ for each i.

To simplify our proof, we impose the following assumption.

Assumption 4.6:

The M_i contain $a_{m+1,m+1}$ for i=2,3,....

Note: We believe that without this assumption the proof still can be extended but the notation will get very complicated.

Theorem 4.2:

Let M_i, m_i, I_j and f_i be defined as before. Suppose A(ϵ) satisfies Assumptions 4.1-4.6 and 4.3a. Then the scaling matrix is

$$S = \text{diag}\{\epsilon^{s1}, \ldots, \epsilon^{sN}\} \tag{4.15}$$

$$s_i - s_{i+1} = \begin{cases} x, & i=m \\ f_j - a_i, & i \neq m \text{ and } i \in I_j \end{cases} \tag{4.16}$$

where

$$x = \max_{\substack{i \in J_1 \\ j \in J_2}} \{b_i - \sum_{k=m+1}^{i-1} (a_k - b_k) + O(a_{ij})\} \tag{4.17}$$

$$J_1 = \{m+1,\ldots,N\}, \quad J_2 = \{1,\ldots,m\}$$

Let us see one example of a multiple-cycle skeleton before the proof. The following three-cycle example not only shows the procedure of this theorem but also suggests how to extend this theorem to the case whire there are more than two cycles.

Example 4.2 Consider

$$A(\epsilon) = \begin{bmatrix} \epsilon & 1 & 1 & \textcircled{1} \\ \epsilon^2 & \textcircled{2\epsilon} & \epsilon & \epsilon^2 \\ \epsilon^2 & \epsilon^2 & \textcircled{\epsilon} & \epsilon^2 \\ \textcircled{\epsilon^2} & \epsilon^3 & \epsilon^4 & \epsilon^5 \end{bmatrix}$$

It is easy to identify the skeleton in this example. The permutation formed by this skeleton is (2)(3)(1,4). The symmetric permutation should therefore be (1,3), i.e. switching the 1-st and 3-rd rows and columns. We have

$$T = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$A_1(\epsilon) = TA(\epsilon)T^{-1} = \begin{bmatrix} \textcircled{\epsilon} & \epsilon^2 & \epsilon^2 & \epsilon^2 \\ \epsilon & \textcircled{2\epsilon} & \epsilon^2 & \epsilon^2 \\ 1 & 1 & \epsilon & \textcircled{1} \\ \epsilon^4 & \epsilon^3 & \textcircled{\epsilon^2} & \epsilon^5 \end{bmatrix}$$

Here a_i=1,1,0,2, b_i=1,1,1,1. In the following, we shall repeatedly aplly Theorem 4.2 to $A_1(\epsilon)$. Since the skeleton forms a BDCM with three blocks, we start from the lower-right 3x3 principal submatrix, which consists of two blocks. For this submatrix, m=1, b_i=1, 1, 1, a_i=1, 0, 2, $b=b_1=1$, $J_1=\{2, 3\}$, $J_2=\{1\}$.

$$x = \max_{\substack{i\in\{2,3\} \\ j=1}} \{b_i - \sum_{k=2}^{i-1} (a_k - b_k) - O(a_{ij})\} = 1, \quad g=2, \ h=1.$$

Therefore a scaling for this 3x3 principal submatrix of $A_1(\epsilon)$ follows from the calculation

$$i_1 - i_2 = 0, \ i_2 - i_3 = -x = -1, \ i_3 - i_4 = b_3 - a_3 = 1.$$

or $$i_1 = i_2 = i_4 = 0, \ i_3 = 1.$$

Applying this scaling to $A_1(\epsilon)$, we obtain

$$A_2(\epsilon) = \text{diag}\{1,1,\epsilon,1\} A_1(\epsilon) \text{diag}\{1,1,\epsilon^{-1},1\}$$

$$= \begin{bmatrix} \epsilon & \epsilon^2 & \epsilon & \epsilon^2 \\ \epsilon & 2\epsilon & \epsilon & \epsilon^2 \\ \epsilon & \epsilon & \epsilon & \epsilon \\ \epsilon^4 & \epsilon^3 & \epsilon & \epsilon^5 \end{bmatrix}$$

Now, apply Theorem 4.2 to $A_2(\epsilon)$ again. Taking the (1,1) entry as A_{11} and the complementary 3x3 principal submatrix as A_{22} of Lemma 4.1, we see that $A_2(\epsilon)$ satisfies the requirements for Theorem 4.2. Therefore we have $m=1$, $J_1=\{2,3,4\}$, $J_2=\{1\}$, and

$$x = \max_{i,j} \{b_i - O(a_{ij})\}=0,$$

So the scaling remains the same, i.e.

$$S = \text{diag}\{1,1,\epsilon,1\}$$

and

$$A_2(\epsilon) = SA_1(\epsilon)S^{-1} = STA(\epsilon)T^{-1}S^{-1}$$

$$= \begin{bmatrix} \epsilon & \epsilon^2 & \epsilon & \epsilon^2 \\ \epsilon & 2\epsilon & \epsilon & \epsilon^2 \\ \epsilon & \epsilon & \epsilon & \epsilon \\ \epsilon^4 & \epsilon^3 & \epsilon & \epsilon^5 \end{bmatrix}$$

which has $a_i=b_i=1,1,1,1$. Therefore $A_2(\epsilon)$ has MSSNS.

Now, let us prove Theorem 4.2.

Proof:

According to Theorem 4.1, the scaling defined in (4.15)-(4.17) will ensure that the orders of the entries in the i-th row of A_{11}'(where

A_{ij}' denotes the submatrices after scaling) have order equal to or larger than f_1 and the orders of $a_{12}', \ldots, a_{m-1,m}'$ and a_{m1}' which, in this case, are equal to f_1. The entries in i-th row of A_{22}' have order equal to or larger than b_i and the orders of $a_{m+1,m+2}', \ldots, a_{N-1,N}'$ and $a_{N,m+1}'$ are equal to $b_{m+1}, \ldots, b_N$. Now let us check the orders of the entries in A_{21}' and A_{12}'.

It is easy to see that after scaling the order of the (i,j)-th entry in A_{21}' is

$$O(\bar{a}_{ij}) = O(a_{ij}) + \sum_{k=m+1}^{i-1} (a_k - b_k) + x$$

$$= O(a_{ij}) + \sum_{k=m+1}^{i-1} (a_k - b_k) + \max_{\substack{i \in J_1 \\ j \in J_2}} \{b_i - O(a_{ij}) - \sum_{k=m+1}^{i-1} (a_k - b_k)\}$$

$$\geq b_i, \quad i \in J_1$$

Also there exist integers g and h such that

$$O(\bar{a}_{gh}) = b_g$$

Now define $d_{ij} = O(\bar{a}_{ij}) - O(a_{ij})$ and consider the entries in A_{12}'. For columns $g+1, \ldots, N$ the order increment d_{ij} are

$$d_{ij} = - \sum_{k=m+1}^{j-1} (a_k - b_k) - x, \quad j = g+1, \ldots, N$$

$$= \sum_{k=m+1}^{j-1} (b_k - a_k) - b_g + \sum_{k=m+1}^{g-1} (a_k - b_k) + O(a_{gh})$$

$$= \sum_{k=g}^{j-1} (b_k - a_k) - b_g + O(a_{gh})$$

$$= \sum_{k=g+1}^{j-1} (b_k - a_k) + b_g - a_g + O(a_{gh}) - b_g$$

But

$$O(a_{gh}) \geq a_g, \; b_k \geq a_k, \text{ for any } k$$

Therefore

$$d_{ij} \geq 0$$

and

$$O(a_{ij}') = O(a_{ij}) + d_{ij} \geq O(a_{ij}) \geq f_1, \; i=1,\ldots,N, \; j=g+1,\ldots,N$$

For the remaining part of A_{12}', i.e. columns $m+1,\ldots,g-1$, we can use the following arguments.

In this area a_{ij} and a_{gh} must be contained in a diagonal of some principal minor containing only canonical diagonal elements of A_{11} and A_{22} apart from a_{ij} and a_{gh}. In other words, if $i \leq h$, the diagonal can be expressed as $(g,h,g+1,\ldots,i-1,i,j,j+1,\ldots,g-1)$, and contains $h+g-i-j+2$ elements. Since the principal minor is enclosed by $M[g,h]$, its order must be (according to Assumption 4.4) equal to or higher than

$$\sum_{k=i+j-h-2}^{g} b_k$$

Therefore, it is easy to check that $O(a_{ij}')>f_1$ after scaling. For the case where i>h we can apply similar arguments.

The above scaling theorems have been obtained under rather strong assumptions. It is therefore important to know that they are applicable in certain situations of interest. In the next section we show how to apply our results to the scaling problem in a high gain feedback system.

4.3 Scaling in High Gain Feedback systems

The problem discussed in this section is the time scale analysis of feedback systems. In particular, we consider a high gain feedback control problem, of the type studied by Young, Kokotovic and in more detail by Sanutti. [14], [15], [13]. We shall see that the methodology developed in the previous subsections serves to extend the scaling results of these references to a more general case.

To simplify our discussion, let us restrict ourselves to the simple situation proposed in [14], where an N-state, m-input, m-output system with output feedback is considered:

$$\begin{aligned} \hat{\dot{x}} &= \hat{A}\hat{x} + \hat{B}\hat{u} \\ \hat{y} &= \hat{C}\hat{x} \\ \hat{u} &= K(\epsilon)\hat{y} \end{aligned} \qquad (4.18)$$

with $\hat{A}$, $\hat{B}$, $\hat{C}$, $K(\epsilon)$ being matrices of appropriate dimensions.

It is assumed that system (4.18) satisfies

$\hat{C}\hat{A}^{i-1}\hat{B}=0$, $i=1,\ldots,q$

$\hat{C}\hat{A}^{q}\hat{B}$ is nonsingular for some $q>1$.

Under this assumption, there is a linear transformation to transfer (4.18) to the following form [14].

$$\begin{bmatrix} \dot{x}_0 \\ \dot{x}_1 \\ \vdots \\ \dot{x}_q \end{bmatrix} = \left[\begin{array}{ccc:c} A_0 & A_{01} & & 0 \\ \cdots & \cdots & & \cdots \\ 0 & 0 & & \\ & & & I \\ \cdots & \cdots & & \cdots \\ A_{q0} & A_{q1} & A_{q2} \cdots\cdots & A_{qq} \end{array}\right] \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_q \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ \\ B_q \end{bmatrix} u$$

$$y = [0\ I\ 0\ \ldots\ 0]x$$

$$u = K(\epsilon)y$$

or

$$\begin{bmatrix} \dot{x}_0 \\ \dot{x}_1 \\ \vdots \\ \dot{x}_q \end{bmatrix} = \left[\begin{array}{cc:c} A_0 & A_{01} & 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & \\ & & I \\ \cdots & \cdots & \cdots \\ A_{q0} & A_{q1}+KB_q & A_{q2} \cdots\cdots A_{qq} \end{array}\right] \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_q \end{bmatrix}$$

Suppose that we have a high gain feedback problem where

$$K(\epsilon) = \epsilon^{-q}I = p^{q}I, \quad p=\epsilon^{-1}$$

Then

$$A(\epsilon) = \begin{bmatrix} A_0 & A_{01} & & 0 \\ 0 & 0 & & \\ & & & I \\ A_{q0} & A_{q1} & A_{q2} \cdots\cdots & A_{qq} \\ & +p^q B_q & & \end{bmatrix}$$

If we scale $A(\epsilon)$ by ϵ^q (which corresponds to time scaling) and denote the resultant matrix as $A_1(\epsilon)$, then we see that $A_1(0)$ does not have SSNS. Therefore a scaling is necessary in order to have a standard two-time-scale form.

We can also see that $A_1(\epsilon)$ is in the <u>block</u> explicit form with one cycle (i.e. the same form as that in the previous sections but now each entry is a constant matrix multiplied by a power of ϵ). Applying the scaling in Theorem 4.1 but treating each block as one entry of matrix $A(\epsilon)$ in the theorem, we can see that a scaling

$$S = \text{diag}\{p^q I, p^q I, p^{q-1} I, p^{q-2} I, \ldots, I\}$$

will bring $A(\epsilon)$ to

$$A_1(\epsilon) = SA(\epsilon)S^{-1} = \begin{bmatrix} A_0 & A_{01} & & 0 \\ 0 & 0 & & \\ & & & pI \\ p^{-q}A_{q0} & p^{-q}A_{q1} & p^{-q+1}A_{q2} \cdots\cdots & A_{qq} \\ & +pB_q & & \end{bmatrix}$$

$$= \begin{bmatrix} I & & & \\ & \epsilon^{-1}I & & \\ & & \ddots & \\ & & & \epsilon^{-1}I \end{bmatrix} \begin{bmatrix} A_0 & A_{01} & & 0 \\ 0 & 0 & & \\ & & & I \\ \epsilon^{q+1}A_{q0} & \epsilon^{q+1}A_{q1}+B_q & \epsilon^{q}A_{q2} \cdots\cdots & A_{qq} \end{bmatrix}$$

$$= D(\epsilon)\bar{A}(\epsilon)$$

As in Chapter 3, we check the Schur complements of $\bar{A}(0)$ to see if A_1 satisfies the MSSNS condition or not. Note that there are two time scales in this example so we need to check A_0 (i.e. the $\tilde{A}_{11}$ in Chapter 3) and the Schur conplement of A_0 (i.e. the $\tilde{A}_{22}$ in Chapter 3). Here we only consider the case where A_0 is of full rank. Otherwise the scaling will be too complex. Since B_q is full rank, we can see that the two Schur complements are full rank. Therefore $A_1(\epsilon)$ has MSSNS.

Now consider a more general cases where K(p) is an analytic function of p at $p=\infty$. This is a case that has not been considered before in the literature.

Let us first consider the $(n_2+n_3)x(n_2+n_3)$ submatrix

$$\begin{bmatrix} 0 & I & & \\ \vdots & & \ddots & 0 \\ & & \ddots & \\ 0 & & & I \\ A_q+K(p)B_q & x & \cdots\cdots & x \end{bmatrix}$$

of $A(\epsilon)$. Now suppose the Smith form of $A_q+K(p)B_q$ is

$$A_q+K(p)B_q = P(p)D(p)Q(p) = P\,\mathrm{diag}\{I_1, pI_2, \ldots\}Q$$

where P(p) and Q(p) are unimodular. Then by a linear transformation, we get

$$A_1 = \text{diag}\{Q,I,I,P^{-1}\}\ A\ \text{diag}\{Q^{-1},I,I,P\}$$

$$= \begin{bmatrix} 0 & Q & & \\ \cdot & & I & 0 \\ \cdot & & & \cdot \\ \cdot & 0 & & I \\ 0 & & & P \\ D(p) & x & \cdots\cdots & x \end{bmatrix}$$

To see the necessary scaling for A_1, let us consider an 8x8 example.

Example:

$$\text{Let } Q=\begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix},\quad P=\begin{bmatrix} 0 & e \\ f & 0 \end{bmatrix},\quad D(p)=\begin{bmatrix} p^4 & 0 \\ 0 & p^8 \end{bmatrix}.$$

Then

$$A_1 = \begin{bmatrix} 0 & 0 & 0 & a & 0 & 0 & 0 & 0 \\ 0 & 0 & b & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & e \\ 0 & 0 & 0 & 0 & 0 & 0 & f & 0 \\ p^4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & p^8 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

After some symmetric transformation we have

$$A_1' = \begin{bmatrix} 0 & a & 0 & 0 & & & & \\ 0 & 0 & 1 & 0 & & 0 & & \\ 0 & 0 & 0 & f & & & & \\ p^4 & 0 & 0 & 0 & & & & \\ & & & & 0 & b & 0 & 0 \\ & & & & 0 & 0 & 1 & 0 \\ & & 0 & & 0 & 0 & 0 & e \\ & & & & p^8 & 0 & 0 & 0 \end{bmatrix}$$

This is a two-cycle case. Applying Theorem 4.2, we have the scaling for A_1'

$$S' = \mathrm{diag}\{p^9, p^8, p^7, p^6, p^6, p^4, p^2, 1\}$$

Then

$$S'A_1'S'^{-1} = \begin{bmatrix} pI & 0 \\ 0 & p^2I \end{bmatrix} \begin{bmatrix} 0 & a & 0 & 0 & & & & \\ 0 & 0 & 1 & 0 & & 0 & & \\ 0 & 0 & 0 & f & & & & \\ 1 & 0 & 0 & 0 & & & & \\ & & & & 0 & b & 0 & 0 \\ & & & & 0 & 0 & 1 & 0 \\ & & 0 & & 0 & 0 & 0 & e \\ & & & & 1 & 0 & 0 & 0 \end{bmatrix}$$

which has MSSNS. It is easy then to derive the scaling for the original system (simply keep the (1,1) block unchanged) and find that the system has three time scales. From this example we notice that, unlike in the simple case considered just prior to it, the scaling is determined by the structure of Q(p) and P(p). Therefore, in general, if K(p) is a function of p, one should first apply the Smith decomposition on K(p) to find Q(p) and P(p) and then use the scaling approach in the previous sections on A_1 to obtain the necessary scaling.

From the results in this section we see that the algebraic approach developed in this thesis not only gives a clearer picture of the scaling proposed in [14], e.g. the justification of the scaling used, but also provides a tool to treat the more general case where K(p) is an analytic

function of p.

4.4 Conclusion

In this chapter we have discussed the situation where the MSSNS condition is violated. We first observe that the lack of MSSNS is equivalent to inequality of the orders of eigenvalues and orders of invariant factors. This suggests the use of a non-unimodular similarity transformation that keeps the eigenvalues the same while changing the invariant factors. We focus attention on a special transformation, namely scaling, i.e. a diagonal similarity transformation matrix with the diagonal entries being powers of ϵ.

In general, diagonal scaling may not be successful in inducing MSSNS if applied to the system matrix $A(\epsilon)$ directly. Our procedure requires us to first transform $A(\epsilon)$ to its Smith form $A_1(\epsilon)$. We then identify a skeleton in $A_1(\epsilon)$, as defined in Section 4.1. The next step is to apply a symmetric permutation to $A_1(\epsilon)$ so that its skeleton forms a block diagonal circulant matrix. The diagonal scaling can then be derived fairly easily from this form, under certain assumptions.

Several points should be made here.

1. The scaling is <u>not</u> unique. This is because, in the first place, the skeleton is not unique. For example

$$A(\epsilon) = \begin{bmatrix} 5\epsilon & \epsilon & \epsilon & \epsilon^5 \\ \epsilon & \epsilon & \epsilon^2 & \epsilon \\ \epsilon^4 & \epsilon & \epsilon & 1 \\ \epsilon^2 & \epsilon^6 & \epsilon^4 & \epsilon^4 \end{bmatrix}$$

has two skeletons. One is a_{22}, a_{13}, a_{41}, a_{34}, which has the structure (2)(3,1,4). Another is a_{12}, a_{23}, a_{34}, a_{41}, with structure (3,2,1,4). For the first one, we need a symmetric permutation (2,1). After the symmetric permutation we have

$$A_1(\epsilon) = \begin{bmatrix} \epsilon & \epsilon & \epsilon^2 & \epsilon \\ \epsilon & 5\epsilon & \epsilon & \epsilon^5 \\ \epsilon & \epsilon^4 & \epsilon & 1 \\ \epsilon^6 & \epsilon^2 & \epsilon^4 & \epsilon^4 \end{bmatrix}$$

The scaling for $A_1(\epsilon)$ is $S=\text{diag}\{\epsilon,\epsilon,\epsilon,1\}$ according to Theorem 4.2. Therefore the overall transformation is

$$S_1 = \begin{bmatrix} 0 & \epsilon & 0 & 0 \\ \epsilon & 0 & 0 & 0 \\ 0 & 0 & \epsilon & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

On the other hand, for the second one the scaling is simply

$$S_2 = \text{diag}\{\epsilon,\epsilon,\epsilon,1\} = S$$

2. The approach stated in this chapter is of course not the only possible one. For example, suppose

$$A(\epsilon) = \begin{bmatrix} \epsilon & 1 \\ 0 & \epsilon \end{bmatrix}$$

Let

$$P = \begin{bmatrix} 1 & 0 \\ \epsilon & 1 \end{bmatrix}, \qquad P^{-1} = \begin{bmatrix} 1 & 0 \\ -\epsilon & 1 \end{bmatrix}, \quad \text{unimodular.}$$

Then

$$P^{-1}AP = \begin{bmatrix} 2\epsilon & 1 \\ -\epsilon^2 & 0 \end{bmatrix}$$

is in explicit form and the scaling is S=diag{ϵ,1}. The overall transformation is

$$\begin{bmatrix} \epsilon & 0 \\ -\epsilon & 1 \end{bmatrix} A \begin{bmatrix} \epsilon^{-1} & 0 \\ 1 & 1 \end{bmatrix} = \epsilon \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$$

By inspection, however, we see that $S_2 = \text{diag}\{\epsilon^n, 1\}$ is also a candidate

scaling for the original system, because

$$A_1 = S_2 A S_2^{-1} = \begin{bmatrix} \epsilon & \epsilon^n \\ 0 & \epsilon \end{bmatrix} = \epsilon \begin{bmatrix} 1 & \epsilon^{n-1} \\ 0 & 1 \end{bmatrix}$$

where $n>1$ and A_1 has MSSNS.

3. In general, a diagonal scaling may not be sufficient because diagonal scaling will not change the principal minors of $A(\epsilon)$. As pointed out earlier in Chapter 3, the no cancellation condition is a necessary condition for $A(\epsilon)$ to have MSSNS. Therefore, if this condition is not satisfied, i.e. if there are cancellations among the lowest order principal minors such that the order of the gcd of the principal minors is not equal to the order of the sum of the principal minors, then diagonal scaling will not work. The reason is that the diagonal scaling proposed in this chapter does not change the principal minors, so that after scaling $A(\epsilon)$ still does not have MSSNS. Note that this cannot be remedied even if we first transform $A(\epsilon)$ into its explicit form $A_e(\epsilon)$, as we have done in this chapter, since the explicit form does not ensure this condition.

4. We have mentioned at the begining of this chapter that scaling firstly provided a way to find an approximation of the eigenvalues without computing them. The second role of the scaling is to transform a system that does not have MSSNS to a system that does. We can then (if the scaled system satisfies MSST as well) find a time scale

decomposed system—denote its system matrix by $\hat{A}(\epsilon)$—to approximate the scaled new system using the procedure described in Chapter 4.2. Intuitively, if we apply a reverse scaling (i.e. S^{-1}) to this time-scale decomposed system we should obtain an approximation of the original system. Our conjecture is that the difference between $\exp\{\hat{A}(\epsilon)t\}$ and $S^{-1}\exp\{A(\epsilon)t\}S$ must have an order higher than the order of $\exp\{A(\epsilon)t\}$, i.e. the norm

$$||\exp\{A(\epsilon)t\}-S^{-1}\exp\{A(\epsilon)t\}S||$$

must have an order higher than that of $||\exp A(\epsilon)t\}||$. Mathematically, it may be true that the norm

$$||I-\exp\{-A(\epsilon)t\}S^{-1}\exp\{A(\epsilon)t\}S||$$

approaches zero as ϵ goes to zero, but further study of this possibility is left for further work.

5. Assumption 4.4 can actually be derived from a revised version of Assumption 4.3, which constrains the orders of the overlapped principal submatrices (possibly several in contrast to two in Assumption 4.3). Unfortunately, the revision of Assumption 4.3 is almost impossible to check in a practical situation and the subsequent derivation of Assumption 4.4 is also very complicated. This is why we have stated Assumption 4.3 and 4.4 as two independent assumptions in this chapter.

Several examples are given in this chapter, including the

application to the high gain feedback control problem. The scaling approach proposed in this chapter provides a better understanding and approach to treat a more general case of the type considered in Section 4.3.

CHAPTER 5 CONCLUSION

5.1 Contributions of This Thesis

The principal contribution of this thesis is the development of an algebraic approach to multiple time scale decomposition of perturbed linear systems. Based on this approach, we have been able not only to obtain a more direct and simple description of the multiple time scale decomposition of such systems but also to handle number of problems that are difficult to solve using previously developed results.

In Chapter 2, using this approach we make clear the connection between Kokotovic's explicit two time scale approach and Coderch's multiple time scale results and in the process we provide a far more straightforward procedure than that of Coderch that makes clear the role of Smith decompositions and the Schur complements of a certain matrix. Using this machinery we are then able to extend the multiple time scale result to a larger class of systems that satisfy the MSSNS but not the MSST condition and to solve a feedback time scale assignment problem for the system $\dot{x}=A(\epsilon)x+B(\epsilon)u$ when $A(\epsilon)$ and $B(\epsilon)$ are coprime.

In Chapter 3, the further study of the algebraic structure of $A(\epsilon)$ proved useful in deriving equivalent conditions for MSSNS and MSST. In

Chapter 2 we had shown that the MSSNS (MSST) condition is equivalent to the successive Schur complements of a particular matrix being full rank (Hurwitz). In Chapter 3 we proved that the MSSNS condition is also equivalent to the eigenvalues and invariant factors of A(ε) having the same orders. This result not only simplifies the interpretation of the MSSNS condition but also provides a new and perhaps the only practical criterion to actually verify this condition. Related to this result is the relationship between the MSSNS and MSST conditions. We showed that A(ε) satisfies the MSST condition if and only if it satisfies the MSSNS condition and the orders of the real parts of its eigenvalues are equal to or less than those of the corresponding imaginary parts. Therefore, to check the MSST condition we may first check the relationship between the orders of the real and imaginary parts of the eigenvalues of the matrix. If they satisfy the condition we have just mentioned, then we can proceed to check the MSSNS condition. The final and perhaps most important contribution in Chapter 3 is the clarification of the relationship among the principal minors of A(ε), the orders of the invariant factors, and the MSSNS condition. It is shown in this chapter, that unlike the usual cases, if the system satisfies the MSSNS condition, then the orders of the gcd's of the principal minors of each size instead of all minors of each size determine the invariant factors. Based on this observation, 1) an algorithm, related to the Newton polygon, was developed to determine the invariant factors from the gcd's of the principal minors when the system has MSSNS; and 2) an equivalent condition for MSSNS was obtained which emphasizes the role of the

principal minors and a so-called "no cancellation" condition. Therefore, if a system has well defined time scale behavior (or more generally, $A(\epsilon)$ satisfies the MSSNS condition) then we can use this algorithm to determine the orders of the various time scales. Finally, if the system has MSSNS, we provide an approximation for the eigenvalues of $A(\epsilon)$ in terms of the eigenvalues of the successive Schur complements. As we point out in Chapter 4, this is a rather general result because if the system does not have MSSNS, we may be able to apply some scaling so that the scaled system does.

The results in Chapter 3 are used in Chapter 4 to develop an amplitude scaling procedure to transform a system matrix without MSSNS but satisfying certain conditions, into one that does have the MSSNS property. The scaling procedure requires first identifying a skeleton of the matrix to be scaled, that is identifying a set of critical elements to be scaled. Then, using a symmetric permutation to transform the matrix to a particular canonic form, we can determine the appropriate scaling matrix. Finally, we use our result on amplitude scaling to interpret and generalize recent results on time scale analysis of high gain or nearly singular optimal feedback systems.

5.2 Suggestions For Further Research

1. The procedure developed in Section 2.8 for time scale decomposition of a system with MSSNS but not MSST requires an iterative

algorithm. A bound for stopping this iteration is given in Section 2.8 as the order of the determinant of $A(\epsilon)$ multiplied by the dimension of $A(\epsilon)$. However it is possible--and is typically the case--that the iteration can be stopped much earlier. Therefore, one interesting question is to find a tighter stopping point. Related to this is the question of possibly developing a more efficient procedure for time scale decomposition than that given in Section 2.8. In particular, if the system satisfies the MSSNS condition, then only one step of the iteration is needed at each step. Thus if <u>all</u> of the successive Schur complements in Chapter 2 are Hurwitz, we are finished. However, if at <u>any</u> stage one of these matrices is not Hurwitz, then we must go back to the <u>beginning</u> and do the full set of iterations at each stage. It would be desirable, if possible, to obtain a procedure with a more recursive flavor than this.

2. Feedback time scale assingment is discussed only for the simple case where $A(\epsilon)$ and $B(\epsilon)$ are coprime. At least two extensions of our result are of interest.

1) The consideration of the time scale assignment problem that we study for the general case where $A(\epsilon)$ and $B(\epsilon)$ are not coprime.

2) A generalized cheap control problem, namely the problem of minimizing

$$J = \int x'Q(\epsilon)x + u'R(\epsilon)u$$

for the system

$$\dot{x} = A(\epsilon)x + B(\epsilon)u$$

It is possible that the Hamiltonian matrix for this problem contains the neccessary information for the time scale decomposition. Also, it would be desirable to make contact explicitly with the amplitude scaling results of Chapter 4 and in particular with their uses in interpreting and extending recent results on the time scale structure of more standard cheap control and high-gain feedback problems, i.e. the case in which only $R(\epsilon)$ depends on ϵ and in fact is of the form ϵR.

3. In Chapter 4 we derive a diagonal scaling procedure under somewhat restrictive conditions. The reason for these conditions was primarily due to the absence of simple methods for calculating the invariant factors for an arbitrary matrix. That is we have had to impose conditions so that the scaled system matrix has a special form and the invariat factors can be readily determined. Further research could be directed in at least three directions:

1) Finding simpler procedures for calculating the invariant factors.

2) Relaxing the conditions that were imposed on the matrix $A(\epsilon)$ in order to verify the validity of the scaling approach proposed in this thesis work.

3) Considering the possibility of the non-diagonal scaling. In this thesis, only the diagonal scaling of a system in so-called explicit

form is considered. More generally, a similarity transformation which changes the invariant factors but not eigenvalues can be represented as a non unimodular matrix. By means of the Smith form of this matrix, the general form of the similarity transformation can be treated as a unimodular transformation plus a diagonal scaling. Therefore, an interesting question is what role can be played by the unimodular transformation before diagonal scaling.

4. In Chapter 2 we have given a definition of the well-defined time scale behavior and an extended definition in order to include systems that are Hurwitz for some $\epsilon \in (0, \epsilon_0]$ but do not satisfy the MSST condition. The Hurwitz condition (or, as pointed out in Section 2.8, a slightly weeker conditin that allows $A(\epsilon)$ to have some zero eigenvalues) is a requirement throughout this thesis. However, intuitively it should also be possible to define meaningful approximations for some unstable systems as well. For example, the systems $\dot{x}=x+u$ and $\dot{x}=(1+\epsilon)x+u$ should be thought of as being close. Although the difference between $\exp\{t\}$ and $\exp\{(1+\epsilon)t\}$ grows without bound, the <u>leading order</u> terms in the two systems' dynamics are the same so that, for example, any feedback $u=-kx$ which stabilizes the leading-order term of one system does the same for the other. For this reason it would be desirable to extend further our definition of well-defined time scale behavior so that a much larger class of systems can be considered. Also another possible extension of this definition is proposed at the end of Chapter 4 to define the way in which an amplitude-scaled system approximates the original system.

5. It has been pointed out in Chapter 3 that the eigenvalues of $A(\epsilon)$ are clustered into several groups according to the eigenvalues of the successive Schur complements. It is possible that a frequency scale decomposition, i.e. the decomposition of the transfer function of the system, could be achieved based on the machienary provided in this thesis.

6. As we have seen in Chapter 3, Section 3.5, if the Jordan form of $A(\epsilon)$ does not have nontrivial Jordan blocks having eigenvalues of order higher that 0 and if the similarity transformation matrix $M(\epsilon)$ is also unimodular, then $A(\epsilon)$ has MSSNS. However as the examples in that section also have shown, the reverse is not true. In other words, the matrix $A(\epsilon)$ can have MSSNS but at the same time $M(\epsilon)$ can be non-unimodular and there can be a nontrivial Jordan block whose eigenvalue is of order higher than zero. Further research is required to investigate the relationship between the Jordan form of $A(\epsilon)$ and the MSSNS condition so that a neccessary and sufficient condition can be derived.

7. It would be useful to develop numerical algorithms for the time scale decomposition results in this thesis. The crucial problem in this regard is the development of a numerically sound procedure for computing the Smith decomposition of $A(\epsilon)$. Although an algorithm based on Verhgese's result [18] can be constructed, further investigation is still required to find a well posed algorithm. Note that thanks to the results in Chapter 3, it is possible to derive an algorithm for directly

determining the invariant factors of $A(\epsilon)$. An interesting question is the development of a Smith decomposition algorithm that makes use of prior knowledge of the invariant factors.

Appendix 1

In order to obtain a recursive method for the computation of the S_{ki} as well as the machinary required to prove Theorem 2.3, we introduce some definitions and properties related to the Schur complement.

Consider a matrix A

$$A = \begin{bmatrix} A_{11} \cdots\cdots A_{1n} \\ \cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots \\ A_{n1} \cdots\cdots A_{nn} \end{bmatrix}$$

where the A_{ii}, $i=1,\ldots,n$ are square matrices. Define matrices $C_{i,j}^{k,l}$ and $B_{i,j}^{k,l}$, as follows

$$C_{i,j}^{k,0} = B_{i,j}^{k,0} = A_{ij} \tag{A1.1}$$

$$C_{i,j}^{k,1} = C_{i,j}^{k-1,1} = C_{i,j}^{1,1} \tag{A1.2}$$

$$B_{i,j}^{k,l} = B_{i,j}^{k,l-1} - A_{i,j-l-k+1}(C_{j-l-k+1}^{1,l})^{-1} C_{j-l-k+1,j}^{k,l} \tag{A.3}$$

$$C_{i,j}^{k,l} = B_{i,j}^{k,l-1} - [A_{i1}\ldots A_{i,i-1}](Y_{i-1})^{-1} \begin{bmatrix} B_{1,j}^{k,l-1} \\ \vdots \\ B_{i-1,j}^{k,l-1} \end{bmatrix} \tag{A1.4}$$

$$Y_0 = 0$$

The following property states the relationship between $C_{i,j}^{k,l}$ for different l.

Property 1

$$C_{i,j}^{k,l} = C_{i,j}^{k,l-1} - C_{i,j-l-k+2}^{i,l}(C_{j-l-k+2}^{l,l})^{-1}C_{j-l-k+2,j}^{k,l-1}$$

Here we denote $C_{j-l+1,j-l+1}^{k,l}$ by $C_{j-l+1}^{k,l}$ for convenience.

Proof: By definition (A1.3),(A1.4)

$$C_{ij}^{k,l} = B_{ij}^{k,l-1} - (A_{il}\cdots A_{i,j-1})(Y_{i-1})^{-1}\begin{bmatrix} B_{ij}^{k,l-1} \\ \vdots \\ B_{i-1,j}^{k,l-1} \end{bmatrix}$$

$$= B_{ij}^{k,l-2} = A_{i,j-l-k+2}(C_{j-l-k+2})C_{j-l-k+2}^{k,l-1} -$$

$$(A_{i1}\cdots A_{i,i-1})(Y_{i-1})^{-1}\begin{bmatrix} B_{ij}^{k,l-2} \\ \vdots \\ B_{i-1,j}^{k,l-2} \end{bmatrix}$$

$$+(A_{il}\cdots A_{i,i-1})(Y_{i-1})^{-1}\begin{bmatrix} A_{1,j-l-k+2} \\ \vdots \\ A_{i-1,j-l-k+2} \end{bmatrix}(C_{j-l-k+2}^{i,l})$$

$$C_{j-l-k+2}^{k,l-1} = C_{ij}^{k,l-1}(C_{j-l-k+2})^{-1}C_{j-l-k+2}^{k,l-1}$$

Based on this fact we can prove the following property that relates the $S_{i,j}$ with .

Define

$$S^{l}_{k,1} \triangleq A_{11}^{-1} B^{l}_{1,k+l-1} \qquad \text{(A1.6)}$$

$$S^{l}_{k,k-i} \triangleq (C^{1,1}_{k-i})^{-1} C^{l,i-l+1}_{k-i,k+l-1}$$

We have

Property 2:

$$C^{1,1}_{k,k+1} = A_{k,k+1} - \sum_{i=1}^{k-1} A_{ki} S^{l+1}_{k,k-i}$$

In particular, when l=0, we have

$$C^{1,1}_{k,k} = \tilde{A}_{kk} = A_{kk} - \sum_{i=1}^{k-1} A_{ki} S^{1}_{k,k-i}$$

which is the same as (2.27) in Chapter 2.

Proof: We proof this property by induction. Suppose for some $1 \le i \le k-2$ we have

$$C^{i,i}_{k,k+l} = A_{k,k+l} - \left[A_{k1} \cdots A_{k,k-i-1}\right] \left[Y_{k-i-1}\right]^{-1} \begin{bmatrix} B^{l+1,i}_{1,k+l} \\ \vdots \\ B^{l+1,i}_{k-i-1,k+l} \end{bmatrix}$$

$$\underbrace{- \sum_{j=1}^{i} A_{k,k-j} (C^{1,1}_{k-j})^{-1} C^{l+1,j}_{k-j,k}}_{I_i}$$

Using the matrix inversion lemma, (A1.7) can be rewritten as

$$C^{1,1}_{k,k+1} = A_{k,k+1} - \left[A_{k1} \cdots A_{k,k-i-1} \right]$$

$$\begin{bmatrix} (Y_{k-i-2})^{-1} + (Y_{k-i-2})^{-1} \begin{bmatrix} A_{1,k-i-1} \\ \vdots \\ A_{k-i-2,k-i-1} \end{bmatrix} (C^{1,1}_{k-i-1})^{-1} \left[A_{k-i-1,1} \cdots A_{k-i-1,k-i-2} \right] (Y_{k-i-2})^{-1} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ -(C^{1,1}_{k-i-1})^{-1} \left[A_{k-i-1,1} \cdots A_{k-i-1,k-i-2} \right] (Y_{k-i-2})^{-1} \end{bmatrix}$$

$$\begin{bmatrix} -(Y_{k-i-2})^{-1} \begin{bmatrix} A_{1,k-1} \\ \vdots \\ A_{k-i-2,k-1} \end{bmatrix} (C^{1,1}_{k-i-1})^{-1} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ (C^{1,1}_{k-i-1})^{-1} \end{bmatrix} \begin{bmatrix} B^{l+1,i}_{1,k+1} \\ \vdots \\ \vdots \\ \vdots \\ B^{l+1,i}_{k-i-1,k+1} \end{bmatrix} - I_j$$

$$= A_{k,k+1} - \left[A_{k1} \cdots A_{k,k-i-1} \right] \cdot$$

$$\begin{bmatrix} (Y_{k-i-2})^{-1} \left\{ \begin{bmatrix} B^{l+1,i}_{1,k+1} \\ \vdots \\ B^{l+1,i}_{k-i-2,k+1} \end{bmatrix} + \begin{bmatrix} A_{1,k-i-1} \\ \vdots \\ A_{k-i-2,k-i-1} \end{bmatrix} C^{1,1\,-1}_{k-i-1} \overbrace{\left(\left[A_{k-i-1,1} \cdots A_{k-i-1,k-i-2} \right] Y_{k-i-2}^{-1} \begin{bmatrix} B^{l+1,i}_{1,k+1} \\ \vdots \\ B^{l+1,i}_{k-i-2,k-1} \end{bmatrix} - B^{l+1,i}_{k-i-1,k+1} \right)}^{-C^{l+1,i+1}_{k-i-1,k+1}} \right\} \\ \cdots\cdots \\ (C^{1,1}_{k-i-1})^{-1} \underbrace{\left(B^{l+1,i}_{k-i-1,k+1} - \left[A_{k-i-1,1} \cdots A_{k-i-1,k-i-2} \right] (Y_{k-i-2})^{-1} \begin{bmatrix} B^{l+1,i}_{1,k+1} \\ \vdots \\ B^{l+1,i}_{k-i-2,k+1} \end{bmatrix} \right)}_{C^{l+1,i+1}_{k-i-1,k+1}} \end{bmatrix}$$

$$-I_j$$

Using (Al.4) we have

$$C^{1,1}_{k,k+1} = A_{k,k+1} - \left[A_{k1} \cdots A_{k,k-i-1} \right] \begin{bmatrix} (Y_{k-i-2})^{-1} \left(\begin{bmatrix} B^{l+1,i}_{1,k+1} \\ \vdots \\ B^{l+1,i}_{k-i-2,k} \end{bmatrix} + \begin{bmatrix} A_{1,k-i-1} \\ \vdots \\ A_{k-1-2,k-i-1} \end{bmatrix} C^{1,1\,-1}_{k-i-1} C^{l+1,i+1}_{k-i-1,k+1} \right) \\ \cdots\cdots \\ (C^{1,1}_{k-i-1})^{-1} C^{l+1,i+1}_{k-i-1,k+1} \end{bmatrix} - I_j$$

$$=A_{k,k+1}-\left[A_{k1}\cdots A_{k,k-i-2}\right](Y_{k-i-2})^{-1}\begin{bmatrix}B^{l+1,i+1}_{1,k+2}\\ \vdots\\ B^{l+1,i+1}_{k-i-2,k+1}\end{bmatrix}-A_{k,k-i-1}(C^{1,1}_{k-i-1})^{-1}C^{l+1,i+1}_{k-i-1,k}-I_j$$

$$=A_{k,k+1}-\left[A_{k1}\cdots A_{k,k-i-2}\right](Y_{k-i-2})^{-1}\begin{bmatrix}B^{l+1,i+1}_{1,k+1}\\ \vdots\\ B^{l+1,i+1}_{k-i-2,k+1}\end{bmatrix}-\sum_{j=1}^{i+1}A_{k,k-j}(C^{1,1}_{k-j})^{-1}C^{l+1,j}_{k-j,k+1} \qquad \textbf{(A1.8)}$$

Comparing (A1.8) and (A1.7) we see that (A1.6) is true for i+1.

For i=1, by definitions (A1.1) and (A1.4)

$$C^{1,1}_{k,k+1} = A_{kk+1}-\left[A_{k1}\cdots A_{k,k-1}\right](Y_{k-1})^{-1}\begin{bmatrix}A_{1,k+1}\\ \vdots\\ A_{k-1,k+1}\end{bmatrix}$$

$$= A_{k,k+1}-\left[A_{k1}\cdots A_{k,k-1}\right]\begin{cases}(Y_{k-2})^{-1}+(Y_{k-2})^{-1}\begin{bmatrix}A_{1,k-1}\\ \vdots\\ A_{k-2,k-1}\end{bmatrix}(C^{1,1}_{k-1})^{-1}\left[A_{k-1,1}\cdots A_{k-1,k-2}\right](Y_{k-2})^{-1} & \vdots\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \vdots\\ -(C^{1,1}_{k-1})^{-1}\left[A_{k-1,1}\cdots A_{k-1,k-2}\right](Y_{k-2})^{-1} & \vdots\end{cases}$$

$$-(Y_{k-2})^{-1}\begin{bmatrix}A_{1,k-1}\\ \vdots\\ A_{k-2,k-1}\end{bmatrix}(C^{1,1}_{k-1})^{-1}\begin{bmatrix}A_{1,k+1}\\ \vdots\\ A_{k-1,k+1}\end{bmatrix}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$(C^{1,1}_{k-1})^{-1}$$

$$=A_{k,k+1}-\{A_{k1},\cdots A_{k,k-2}\}(Y_{k-2})^{-1}\left\{\begin{bmatrix}A_{1,k+1}\\ \vdots\\ A_{k-2,k+1}\end{bmatrix}+\begin{bmatrix}A_{1,k-1}\\ \vdots\\ A_{k-2,k-1}\end{bmatrix}\right.,$$

$$\overbrace{-C^{1,1}_{k-1,k+1}}$$

$$\left.\cdot(C^{1,1}_{k-1})^{-1}\left[(A_{k-1,1}\cdots A_{k-1,k-2})(Y_{k-2})^{-1}\begin{matrix}A_{1,k+1}\\ \vdots\\ A_{k-2,k+1}\end{matrix}-A_{k-1,k+1}\right]\right\}.$$

$$\overbrace{-C^{1,1}_{k-1,k+1}=-C^{l+1,1}_{k-1,k+1}}$$

$$+A_{k,k-1}(C^{1,1}_{k-1})^{-1}\left\{\left[A_{k-1,1}\cdots A_{k-1,k-2}\right](Y_{k-2})^{-1}\begin{bmatrix}A_{1,k+1}\\ \vdots\\ A_{k+2,k+1}\end{bmatrix}-A_{k-1,k+1}\right\} \quad (A1\cdot9)$$

Considering (A1.1) and (A1.2) we have

$$A_{i,k+1} = B^{l+1,0}_{i,k+1} \qquad i=1,\ldots,k+2$$

and

$$C^{1,1}_{k-1,k+1} = C^{l+1,1}_{k-1,k+1}$$

Therefore by (A1.3)

$$\begin{aligned} A_{i,k+1} &= A_{1,k-1}(C^{1,1}_{k-1})^{-1}C^{1,1}_{k-1,k+1} \\ &= B^{l+1,0}_{i,k+1} + A_{i,k-1}(C^{1,1}_{k-1})^{-1}C^{l+1,1}_{k-1,k+1} \qquad \text{(A1.10)} \\ &= B^{l+1,1}_{i,k+1} \end{aligned}$$

Inserting (A1.10) into (A1.9) we have

$$C^{1,1}_{k,k+1} = A_{k,k+1} - [A_{k1}\cdots A_{k,k-2}](Y_{k-2})^{-1}\begin{bmatrix} B^{l+1,1}_{1,k+1} \\ \vdots \\ B^{l+1,1}_{k+2,k+1} \end{bmatrix}$$

$$-A_{k,k-1}(C^{1,1}_{k-1})^{-1}C^{l+1,1}_{k-1,k+1} \qquad \text{(A1.11)}$$

which is the same as (A1.7) when i=1.

For i=k-2 we have from (A1.7)

$$C^{1,1}_{k,k+1} = A_{k,k+1} - A_{k1}A_{11}^{-1}B^{l+1,k-2}_{1,k+1} - \sum_{j=1}^{k-2} A_{k,k-j} \cdot (C^{1,1}_{k-j})^{-1}C^{l+1,j}_{k-j,k+1}$$

This completes the proof of (A1.7).

Property 2 will be used to prove our main result, Theorem 2.3. For the same reason we need the following property.

Property 3:

$$C^{1,i-k+1}_{k-i,k} = C^{1+1,i-1}_{k-i,k} - C^{1,i-1}_{k-i,k-1}(C^{1,1}_{k-1})^{-1}C^{1,1}_{k-1,k} \qquad \text{(A1.12)}$$

Proof:

Suppose (A1.12) is true for i=1,...,m-1, we prove it is true for i=m.

Let us consider $C^{1,2}_{k-m,k}$, $1 \le m-1$. According to (A1.2) and (A1.5)

$$C^{1,2}_{k-m,k} = C^{1,1}_{k-m,k} - C^{1,1}_{k-m,k-1}(C^{1,1}_{k-1})^{-1}C^{1,1}_{k-1,k}$$

$$= C^{1+1,1}_{k-m,k} - C^{1,1}_{k-m,k-1}(C^{1,1}_{k-1})^{-1}C^{1,1}_{k-1,k}$$

Similarly by (A1.5)

$$C^{1,3}_{k-m,k} = C^{1,2}_{k-m,k} - C^{1,1}_{k-m,k-1-1}(C^{1,1}_{k-1-1})^{-1}C^{1,2}_{k-1-1,k}$$

Using (A1.2) to find an expression for $C^{1,2}_{k-1-1,k}$ we obtain

$$C^{1,3}_{k-m,k} = C^{l+1,1}_{k-m,k} - C^{1,1}_{k-m,k-1}(C^{1,1}_{k-1})^{-1}C^{l,1}_{k-1,k} - C^{1,1}_{k-m,k-1-1}(C^{1,1}_{k-1-1})^{-1}C^{l+1,1}_{k-1-1,k}$$

$$- C^{1,1}_{k-m,k-1-1}(C^{1,1}_{k-1-1})^{-1}C^{1,1}_{k-1-1,k-1}(C^{1,1}_{k-1})^{-1}C^{1,1}_{k-1,k}$$

$$= C^{l+1,2}_{k-m,k} - C^{1,2}_{k-m,k-1}(C^{1,1}_{k-1})^{-1}C^{l,1}_{k-1,k}$$

Then

$$C^{1,4}_{k-m,k} = C^{l,3}_{k-m,k} - C^{1,1}_{k-m,k-1-2}(C^{1,1}_{k-1-2})^{-1}C^{l,3}_{k-1-2,k}$$

$$= C^{l+1,2}_{k-m,k} - C^{1,2}_{k-m,k-1}(C^{1,1}_{k-1})^{-1}C^{l,1}_{k-1,k} - C^{1,1}_{k-m,k-1-2}(C^{1,1}_{k-1-2})^{-1}C^{l+1,2}_{k-1-2,k}$$

$$+ C^{1,1}_{k-m,k-1-2}(C^{1,1}_{k-1-2})^{-1}C^{1,2}_{k-1-2,k-1}(C^{1,1}_{k-1})^{-1}C^{l,1}_{k-1,k}$$

$$= C^{l+1,3}_{k-m,k} - C^{1,3}_{k-m,k-1}(C^{1,1}_{k-1})^{-1}C^{l,1}_{k-1,k}$$

Similarly we have for $C^{l,m-l+1}_{k-m,k}$

$$C^{l,m-l+1}_{k-m,k} = C^{l,1}_{k-m,k} - \sum_{j=2}^{m-1} C^{1,1}_{k-m,k-j}(C^{1,1}_{k-j})^{-1}C^{l+1,j-1}_{k-j,k}$$

$$- C^{1,m-l}_{k-m,k-1}(C^{1,1}_{k-1})^{-1}C^{l,1}_{k-1,k}$$

$$= C^{l+1,m-l}_{k-m,k} - C^{1,m-l}_{k-m,k-1}(C^{1,1}_{k-1})^{-1}C^{l,1}_{k-1,k}$$

Thus we have that (A1·12) is true for i=m.

For i=l+1 we have according to(A1.5)

$$C^{l,2}_{k-l-1,k} = C^{l,1}_{k-l-1,k} - C^{l,1}_{k-l-1,k-l}(C^{l,1}_{k-l})^{-1}C^{l,1}_{k-l,k}$$

Because

$$C^{l,1}_{k-l-1,k} = C^{l+1,1}_{k-l-1,k}$$

we have

$$C^{l,2}_{k-l-1,k} = C^{l+1,1}_{k-l-1,k} - C^{l,1}_{k-l-1,k-l}(C^{l,1}_{k-l})^{-1}C^{l,1}_{k-l,k} \qquad \square$$

APPENDIX 2

To prove Theorem 2.3 we first need two lemmas. The first lemma provides an explicit form of the projection along the range space onto the null space of a specific matrix A. It also displays the three equivalent conditions for a matrix to have semi-simple null structure. This lemma is important because the results in [1] are based on the eigenprojection of zero eigenvalues of the matrices $F_{i,0}$ (see Section 2.2 for details).

Lemma 1

Let

$$A = \begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix} \qquad \text{(A2.1)}$$

where $A_1 \in R^{n \times n}$, $A_2 \in R^{n \times m}$. Suppose that $[A_1 \; A_2]$ has full rank. Then the three statements

1. A_1 is nonsingular.
2. $R(A) \oplus N(A) = R^{n+m}$.
3. A has SSNS.

are equivalent. Furthermore, if A_1 is nonsingular, the projection along R(A) onto N(A) is given by

$$P_0 = \begin{bmatrix} 0 & -A_1^{-1}A_2 \\ 0 & I \end{bmatrix} \qquad \text{(A2.2)}$$

Proof:

1==>2:

It is easy to see that

$$R(A) = \{x|\ x=[z'\ 0]',\ z \in R^n\}$$

If $x \in R(A)$, $x \neq 0$, then

$$Ax = \begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z \\ 0 \end{bmatrix} = A_1 z \neq 0$$

Thus

$$x \notin N(A) \Longrightarrow N(A) \cap R(A) = \emptyset$$

and

$N(A) \oplus R(A) = R^{n+m}$ follows by counting dimensions.

2==>1

If $R(A) \oplus N(A) = R^{n+m}$ but A_1 is singular, then there is some $z \neq 0$ such that $A_1 z = 0$. Let $x=[z'\ 0]'$. Then

$$Ax = 0 \Longrightarrow x \in N(A)$$

But because $[A_1\ A_2]$ has full row rank, we can find a y such that $[A_1\ A_2]y = z$. Therefore

$$Ay = [z'\ \ 0]' = x$$

or $0 \neq x \in R(A) \cap N(A)$. This is a contradiction.

2==>3 See reference [1], [16].

Now consider the similarity transformation

$$Q = \begin{bmatrix} I & A_1^{-1}A_2 \\ 0 & I \end{bmatrix}$$

$$Q^{-1} = \begin{bmatrix} I & -A_1^{-1}A_2 \\ 0 & I \end{bmatrix}$$

We have that

$$QAQ^{-1} = \begin{bmatrix} I & A_1^{-1}A_2 \\ 0 & I \end{bmatrix}\begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} I & -A_1^{-1}A2 \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} I & -A_1^{-1}A_2 \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$$

Clearly the projection onto $N(QAQ^{-1})$ along $R(QAQ^{-1})$ is

$$\begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}$$

Therefore

$$P_0 = Q^{-1}\begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}Q$$

$$= \begin{bmatrix} 0 & -A_1^{-1}A_2 \\ 0 & I \end{bmatrix}\begin{bmatrix} I & A_1^{-1}A_2 \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -A_1^{-1}A_2 \\ 0 & I \end{bmatrix}$$

This lemma shows that if a matrix A has the form of (A2.1), A

having SSNS is equivalent to A_1 being nonsingular. Furthermore, the projection along R(A) onto N(A) is given by (A2.2). In other words, the eigenprojection of A for the zero eigenvalue can be computed directly and easily.

Lemma 2 computes matrices $F_{i,0}$, i=0,...,m.

Lemma 2

Define

$$P_{i,0} = P_i P_{i-1} \cdots P_0 \tag{A2.3}$$

where P_i is the projection along $R(F_{i,0})$ onto $N(F_{i,0})$. Then

$$P_{i-2,0} = \begin{bmatrix} \overbrace{0 \cdots 0}^{i-1} & -S_{i1}^1 & -S_{i1}^2 & \cdots & -S_{i1}^{n-i+1} \\ \vdots & -S_{i2}^1 & -S_{i2}^1 & \cdots & -S_{i2}^{n-i+1} \\ \vdots & \vdots & & & \\ \vdots & -S_{i,i-1}^1 & -S_{i,i-1}^1 & \cdots & -S_{i,i-1}^{n-i+1} \\ \vdots & 1 & & 0 & \\ \vdots & & 1 & \ddots & \\ \vdots & & & \ddots & \\ 0 \cdots 0 & 0 & & & 1 \end{bmatrix} \tag{A2.4}$$

where the $S_{i,j}^1$ are defined as before. Also

$$F_{i-1,0} = \begin{bmatrix} \overbrace{0 \cdots 0}^{i-1} & s^1_{i1}c^{1,1}_{i,i} & s^1_{i1}c^{1,1}_{i,i+1} \cdots s^1_{i1}c^{1,1}_{i,n} \\ \vdots & \vdots & \vdots \\ & s^1_{i2}c^{1,1}_{i,i} & s^1_{i2}c^{1,1}_{i,i+1} \cdots s^1_{i2}c^{1,1}_{in} \\ & \vdots & \vdots \\ & s^1_{i,j-1}c^{1,1}_{i,i} & s^1_{i,j}c^{1,1}_{i,j+1} \cdots s^1_{i,j-1}c^{1,1}_{in} \\ & c^{1,1}_{i,i} & c^{1,1}_{i,j+1} \cdots c^{1,1}_{in} \\ & & 0 \\ & & \vdots \\ 0 \cdots 0 & & 0 \end{bmatrix} \quad \text{(A2.5)}$$

Proof:

Let the Taylor expansion of A(c) be

$$A(c) = F_{00} + cF_{01} + c^2F_{02} + \cdots$$

From the special form of A(c) we can observe that

$$F_{00} = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{bmatrix}$$

Similarly,

$$F_{i,i-1} = \begin{bmatrix} x & x & \cdots & x \\ x & x & \cdots & x \\ A_{i1} & A_{i2} & \cdots & A_{in} \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

where x's represent some terms which may be non-zero.

Now suppose that (A2.4) is valid for $i \leq k$. We would like to prove

that if is valid for i=k+1. From the special form of (A2.4), namely the first i-1 columns are zero vectors, and the special form of $F_{k,0}$ shown in Coderch's thesis [1] we can observe that

$$\begin{aligned} F_{k-1,0} &= P_{k-2}P_{k-3}\cdots P_0 F_{0,k-1} P_0 \cdots P_{k-3}P_{k-2} \\ &= P_{k-2}P_{k-3}\cdots P_0 F_{0,k-1} P_{k-2}P_{k-3}\cdots P_0 \\ &= P_{k-2,0} F_{0,k-1} P_{k-2,0} \end{aligned} \qquad \text{(A2.6)}$$

The other terms in $F_{k-1,0}$ which are related to $F_{0,k-2}$, $F_{0,k-3}$,.... vanish because of the special forms of $P_{k-2,0}$ and $F_{0,i}$, i=k-2,k-3,.... Also $F_{k-1,0}$ can be rewritten as

$$F_{k-1,0} = P_{k-2,0} E_{0,k-1} P_{k-2,0} \qquad \text{(A2.6)}$$

where

$$E_{0,k-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ A_{k1} & A_{k2} & \cdots & A_{kn} \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Then using Eq.(A2.2) of Property 2 we have

$$F_{k-1,0} = P_{k-2,0} \begin{bmatrix} \overbrace{0 \cdots 0}^{k-1} & & 0 & \\ \vdots & & \vdots & \\ \vdots & C^{1,1}_{kk} & 0 & C^{1,1}_{k,k+1} \cdots C^{1,1}_{k,n} \\ \vdots & & \vdots & \\ 0 & & 0 & \end{bmatrix}$$

$$= \begin{bmatrix} \overbrace{0 \cdots 0}^{k-1} & -S^1_{k1} C^{1,1}_{kk} & -S^1_{k1} C^{1,1}_{k,k+1} \cdots & -S^1_{k1} C^{1,1}_{kn} \\ \vdots & \vdots & & \\ & -S^1_{k,k-1} C^{1,1}_{k,k} & -S^1_{k,k-1} C^{1,1}_{k,k+1} \cdots & -S^1_{k,k-1} C^{1,1}_{k,n} \\ & C^{1,1}_{kk} & C^{1,1}_{k,k+1} \cdots & C^{1,1}_{k,n} \\ & & 0 & \\ & & \vdots & \\ 0 \cdots 0 & & 0 & \end{bmatrix} \quad \text{(A2.7)}$$

which proves (A2.5).

Next we prove that the projection P_{k-1} which is the projection along $R(F_{k-1,0})$ onto $N(F_{k-1,0})$ is of the form

$$P_{k-1} = \begin{bmatrix} \overbrace{1 \quad 0}^{k-1} & S^1_{k1} & S^1_{k1}(C^{1,1}_{kk})^{-1} C^{1,1}_{k,k+1} \cdots & S^1_{k1}(C^{1,1}_{kk})^{-1} C^{1,1}_{k,n} \\ \ddots & \vdots & & \vdots \\ 0 \quad 1 & S^1_{k,k-1} & S^1_{k,k-1}(C^{1,1}_{kk})^{-1} C^{1,1}_{k,k+1} \cdots & S^1_{k,k-1}(C^{1,1}_{kk}) C^{1,1}_{k,n} \\ & 0 & -(C^{1,1}_{kk})^{-1} C^{1,1}_{k,k+1} \cdots & -(C^{1,1}_{kk})^{-1} \\ & & 1 \ddots & \\ & & & \ddots 1 \end{bmatrix}$$

(A2.8)

It is easy to see that

$$P_{k-1}^2 = P_{k-1}$$

Therefore P_{k-1} is a projection. Also it is easy to check that

$$P_{k-1}F_{k-1,0} = 0$$

Then if we can prove that

$$x \in N(F_{k-1,0}) \Longrightarrow P_{k-1}x = x$$

we are done. Suppose

$$F_{k-1,0}x = 0, \quad x' = (x_1 \ldots\ldots x_n),$$

We have

$$\sum_{i=0}^{n-k} C_{k,k+i}^{1,1} x_{k+i} = 0 \qquad \text{(A2.9)}$$

Then let

$$y = P_{k-1}x = [y_1' \ldots\ldots y_n']'$$

We have for $i \leq k-1$

$$Y_i = X_i + S_{k,i}\left[\sum_{j=1}^{n-k} (C_{k,k}^{1,1})^{-1} C_{k,k+j}^{1,1} X_{k+j} + X_k\right]$$

$$= X_i + S_{k,i}(C_{k,k}^{1,1})^{-1} \underbrace{\sum_{j=0}^{n-k} C_{k,k+j}^{1,1} X_{k+j}}_{0}$$

Because of (A2.9),

$$y_i = x_i, \quad i \leq k-1$$

For i=k

$$y_k = -(C^{1,1}_{kk})^{-1} \sum_{j=1}^{n-k} C^{1,1}_{k,k+j} x_{k+j} \qquad \text{(A2.10)}$$

Left multiplying $(C^{1,1}_{k,k})^{-1}$ on both sides of (A2.9) and adding the resulting equation to (A2.10) we have

$$y_k = x_k$$

It is straightforward then to see that

$$y_i = x_i, \quad n > i \geq k+1$$

Therefore we have proved that if x $N(F_{k-1,0})$ then $P_{k-1}x=x$. This proves that the P_{k-1} geven in (A2.8) is the projection along $R(F_{k-1,0})$ onto $N(F_{k-1,0})$.

Then using (A2.4),(A2.8) and property 3, Eq.(A2.7)

$$P_{k-1,0} = P_{k-1}P_{k-2,0}$$

$$= \begin{bmatrix} 0 \cdots 0 & -S^2_{k1}+S^1_{k1}(C^{1,1}_{kk})^{-1}C^{1,1}_{k,k+1} & \cdots & -S^{n-k+1}_{k1}+S^1_{k1}(C^{1,1}_{k,k})^{-1}C^{1,1}_{kn} \\ \vdots & \vdots & & \vdots \\ & -S^2_{k,i}+S^1_{ki}(C^{1,1}_{k,k})^{-1}C^{1,1}_{k,k+1} & & -S^{n-k+1}_{k,i}+S^1_{ki}(C^{1,1}_{k,k})^{-1}C^{1,1}_{kn} \\ & \vdots & & \vdots \\ & -(C^{1,1}_{k,k})^{-1}C^{1,1}_{k,k+1} & \cdots & -(C^{1,1}_{kk})C^{1,1}_{k,n} \\ & 1 \ddots & & \\ & & \ddots & \\ 0 \quad 0 & & & 1 \end{bmatrix}$$

By definition (A1.6) and Property 3, (A2.7)

$$-S^2_{ki}+S^1_{ki}(C^{1,1}_{k,k})^{-1}C^{1,1}_{k,k+1}$$

$$= -(C^{1,1}_{ii})^{-1}C^{2,k-i-1}_{i,k}+(C^{1,1}_{kk})^{-1}C^{1,k-i}_{i,k}(C^{1,1}_{kk})^{-1}C^{1,1}_{k,k+1}$$

$$= -(C^{1,1}_{ii})^{-1}C^{2,k-i-1}_{i,k+1} + C^{1,k-i}_{ik}(C^{1,1}_{kk})^{-1}C^{1,1}_{k,k+1}$$

$$= -(C^{1,1}_{ii})^{-1}C^{1,k-i}_{i,k+1} = -S^1_{k+1,i}$$

Then we have

$$P_{k-1,0} = \begin{bmatrix} \overset{k}{0 \cdots 0} & -S^{1}_{k+1,1} & & -S^{n-k}_{k+1,1} \\ \vdots & \vdots & & \vdots \\ & -S^{1}_{k+1,k} & \cdots & -S^{n-k}_{k+1,k} \\ & 1 & \ddots & \\ & & \ddots & \\ 0 \cdots 0 & & & 1 \end{bmatrix} \qquad \text{(A2.11)}$$

Comparing (A2.11) and (A2.4) we see that (A2.4) is true for i=k+1. For i=2,

$$F_{99} = \begin{bmatrix} A_{11} A_{12} \cdots A_{1n} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Using Lemma 1, we see that A_{11} must be nonsingular. Under this assumption and (A2.1), (A1.1)

$$P_{0,0} = P_0 = \begin{bmatrix} 0 & -A_n^{-1}A_n & \cdots & -A_{11}^{-1}A_{1n} \\ \vdots & 1 & \ddots & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix} = \begin{bmatrix} 0 & -S_{21}^1 & -S_{21}^2 & \cdots & -S_{21}^{n-1} \\ & 1 & \ddots & & 0 \\ & & \ddots & \ddots & \\ 0 & & 0 & & 1 \end{bmatrix}$$

which is of the form of (A2.4).

So far we have obtained the explicit expression for $F_{i-1,0}$, $i=1,\ldots,n$. Based on this result we are able to derive the time scale decomposition. That is, we can compute

$$\exp\{\sum_{k=0}^{n-1} F_{k0}\epsilon^k t\} \qquad \text{(A2.12)}$$

as the approximation to $\exp\{A(\epsilon)t\}$. In order to prove Theorem 2.3 we need only diagonalize (A2.12).

Proof of Theorem 2.3:

From the special form of $F_{i-1,0}$ of (A2.5) and T^{-1} as defined in Theorem 2.3, we see that the j-th column of $F_{i-1,0}$ ($j \geq i$) is nothing more than the i-th column of T^{-1} multiplied on the right by $C_{k,j}$. Because any row of T except the i-th row times i-th column of T^{-1} must be zero and the i-th row of T times i-th column of T^{-1} is 1, we have

$$G_i = TF_{i-1,0}T^{-1} = \begin{bmatrix} \overset{i-1}{\begin{matrix}0\\ \vdots \\ 0\end{matrix}} & \begin{matrix} c^{1,1}_{ii} & \cdots & c^{1,1}_{i1} \end{matrix} \\ 0 & \begin{matrix} 0 & \cdots & 0 \end{matrix} \end{bmatrix} T^{-1}$$

The only row of G_i which can be non-zero is the i-th row. It is easy to see that

$$(G_i)_{ii} = c^{1,1}_{ii}$$

and $$(G_i)_{ij} = 0, \quad j<i.$$

Now consider $(G)_{ij}$, $j>i$. It can be seen that

$$(G_i)_{ij} = c^{1,1}_{ii}s^1_{ij} + c^{1,1}_{i,j+1}s_{j,i+1} + \cdots + c^{1,1}_{i,j-1}s^1_{j,j-1} + c^{1,1}_{ij}$$

$$= -c^{1,j-i}_{ij} + (\underbrace{c^{1,1}_{i,j} - c^{1,1}_{i,j-1}(c^{1,1}_{j-1})^{-1}c^{1,1}_{j-1,j}}_{c^{1,2}_{ij}} - c^{1,1}_{i,j-2}(c^{1,1}_{j-2})^{-1}c^{1,2}_{j-2,j}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}_{c^{1,3}_{ij}}$$

$$- \cdots - c^{1,1}_{i,i+1}(c^{1,1}_{i+1})^{-1}c^{1,j-i-1}_{i+1,j})$$

$$= -c^{1,j-i}_{ij} + c^{1,j-i}_{ij} = 0$$

Then

$$G_i = \text{diag}\{0,\ldots,0,\tilde{A}_{ii},0,\ldots,0\}$$

Therefore we have

$$T\exp\{\sum_{k=0}^{n-1} F_{k0}\epsilon^k t\}T^{-1}$$

$$= \exp\{\sum_{k=0}^{n-1} TF_{k0}T^{-1}\epsilon^k t\}$$

$$= \exp\{\sum_{k=0}^{n-1} G_i\epsilon^k t\}$$

$$= \exp\{\text{diag}[A_{11},\ldots,\epsilon^n\tilde{A}_{nn}]t\}$$

But we know that

$$\lim_{\epsilon\to 0}\sup_{t\geq 0} ||\exp\{A(\epsilon)t\} - \exp\{\sum_{k=0}^{n-1} F_{k0}\epsilon^k t\}|| = 0 \qquad \text{(A2.14)}$$

Combining (A2.13) and (A2.14) we have

$$\lim_{\epsilon\to 0}\sup_{t\geq 0} ||\exp\{A(\epsilon)t\} - T^{-1}\exp\{\text{diag}[A_{11},\ldots,\epsilon^n\tilde{A}_{nn}]t\}T|| = 0$$

which is Theorem 2.3.

APPENDIX 3

In this appendix we describe some basic properties of permutation matrix.

Lemma 1 If T is a permutation matrix and A is a matrix, then

1. TAT^{-1} is a symmetric permutation on A.
2. If TAQ is a symmetric permutation on A then $T=Q^{-1}$.

Proof:

1. T can be expressed as product of finite number of 2 cycles. Suppose $T=t_1*t_2*....*t_m$. Then it is easy to see that $T^{-1}=t_m*t_{m-1}*...*t_2*t_1$ and TAT^{-1} is a symmetric permutation on A.

2. Since $TQ=TIQ=I$ by definition of symmetric permutation, $T=Q^{-1}$.

Lemma 2:

Any permutation matrix can be brought to BDC form by similarity transformation

$$TAT^{-1} = \text{diag}\{C_1, C_2, \ldots, C_n\}$$

where T is a permutation matrix and TAT^{-1} is a symmetric permutation on A.

Proof:

We already showed that a permutation can be represented by its

cycles and this representation is unique.

The idea behind this proof is similar to that of bringing a matrix into its Jordan form where a new basis which span invariant subspaces is chosen. A linear transformation T serves as the basis transform. In the new basis, the matrix will have block diagonal form.

$$\bar{A} = TAT^{-1} = \text{diag}\{J_1,\ldots,J_n\}$$

where J_i are Jordan blocks.

Similarly, here a simple reordering (a permutation) will do. For example, suppose A has only one cycle.

$$0_A = (x_3,x_1,x_2) \quad (\text{or } 0_A=(3,1,2) \text{ for simplicity})$$

Here 0_A is not a circulant permutation. But it is easy to see that if we switch x_3 and x_2, then in the new coordinate system 0_A will be a circulant permutation. Namely, if

$$\bar{A}=TAT^{-1}$$

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

then $0_{\bar{A}}=(2,1,3)$ is a circulant permutation.

In general if T is a permutation matrix which permutes x_i and x_j and $\bar{A}=TAT^{-1}$ then it is easy to check that $0_{\bar{A}}$ will be the same as 0_A except x_i and x_j have switched positions. We also know that by finite

number, say m, of switches of two entries, we can "sort" an arbitrary n entry array in the canonical form. Suppose product of permutation matrices is still a permutation matrix, we proved the lemma if there is only one cycle.

If there are several cycles, one can first regroup the entries and then use the same schema as for single cycle.

REFERENCES

[1] M.Coderch, A. Willsky, S.Sastry and D.Castanon.
Hierarchical Agregation of Linear Systems with Multiple Time Scales.
Conf. on Decosopm amd Dontr., San Diego : 930-935, Dec.,1981

[2] R.E.O'Maley, and A.Jameson.
Singular Perturbations and Singular Arcs-part 1.
IEEE Tran. AC 20(2):218-226, Apr, 1975.

[3] G.Birkhoff and S.McLane.
A Survey of Modern Algebra (third edition).
The Macmillan Company, 1965.

[4] J.H.Chow, and P.V.Kokotovic.
A Decomposition of Near-Optimum Regulators for Systems with Slow and Fast Modes.
IEEE Tran. AC 23:701-705, Oct., 1976.

[5] J.H.Chow and P.V.Kokotovic.
Two-Time-Scale Feedback Design of a Class of Nonlinear Systems.
IEEE Tran. AC 23(3):438-443, June, 1978.

[6] J.H.Chow.
Time-Scale Modeling of Dynamic Networks With Applications to Power Systems.
Spring-Verlag, 1982.

[7] M.Coderch,
Multiple Time Scale Approach to Hierarchical Aggregation of Linear Systems and Finite State Markov Processes.
Ph.D. thesis, M.I.T., 1982.

[8] A.H.Haddad.
On Singular Perturbation in Stochastic Dynamic Systems.
10th Asilomar Conf. on Circuits, systems, and Computers, Nov., 1976.

[9] T.Kato.
Perturbation Theorey for Linear Operators.
Springer Verlag, Berlin, 1966.

[10] P.V.Kokotovic, R.E.O'Malley,Jr., and P.Sannuti,
Singular Perturbation and Order Reduction in Control Theory-
An Overview.
Automatica 12:123-132, 1976.

[11] S.Lang.
Algebra.
Addison-Wesley, 1965.

[12] R.e.O'Malley.
A More Direct Solution of the Nearly Singular Linear
Regulator Problem.
SIAM J.Control and Optimization 14(6):1063-1077, Nov., 1976.

[13] K.D.Young, P.V.Kokotovic and V.I.Utkin.
A Singular Perturbation Analysis of High-gain Feedback
Systems.
IEEE Tran. on Automatic Control, 22(6):931-937,Dec., 1977.

[14] P.Sannuti.
Direct Singular Perturbation Analysis of High-gain and Cheep
Control Problems.
Automatica, Vol. 19, No 1, Jan.,1983. pp41-51.

[15] P.Sannuti and H.Wason.
A songular Perturbation Canonical Form of Invartible
Systems: Determination of Multiple Root-Loci.
Int.J.Contr., 37(6):1259-1286, 1983.

[16] R.J.Walker.
Algebraic Curves.
Spriger-Verlag, New York, 1978.

[17] P.V.Kokotovic.
An Introduction to singular Perturbations.
Systems Engi. for Power, Organizational Forms for Large
Scale Systems, Vol.2, Oct. 1979.

[18] G.Verghese and T. Kailath, "Rational Matrix Structure", IEEE, AC-26, No. 2, Apr. 81.

[19] C. I. Byrnes and P.K. Stevens, "The McMillan and Newton Polygons of a Feedback System and the Construction of Rootloci", Int. J. Control, 1982, Vol. 35, No. 1, 29-53.

[20] P.W.Luse and H.K.Khalil, "A Frequency Domain Approach For Systems With Slow and Fast Modes", Proceedings of Automatic Control Conference, San Francisco, June 1983, 443-444

[21] R. Silva-Madriz and S.S.Sastry, "Input-output Description of Linear Systems With Multiple Time Scales" Memorandum N. UCB/ERL, M83/77, December, 1983

[22] P.V.Kokotovic, "Applications of Singular Perturbation Techniques to Control Problems" SIAM Review, Vol. 26, No.4, Oct. 1984