September 1986 LIDS-P-1604

AN ALGEBRAIC APPROACH TO TIME SCALE ANALYSIS OF SINGULARLY PERTURBED LINEAR SYSTEMS[1]

Xi-Cheng Lou[2]

Alan S. Willsky[2]

George C. Verghese[3]

Abstract

In this paper we develop an algebraic approach to the multiple time scale analysis of perturbed linear systems based on the examination of the Smith form of the system matrix viewed as a matrix over a ring of functions in the perturbation parameter. This perspective allows us to obtain a strengthened version of the results of [3] and to provide a bridge between these complex but general results and previous explicit, conceputally simple, but somewhat restrictive results such as described in [1], [2]. In addition, our algebraic framework allows us to investigate a variety of other problems. In this paper we study the problem of developing valid time scale decompositions in cases in which weak damping terms discarded in the approaches in [1] - [3] must be retained. Also, our approach exposes the role of the invariant factors of the system matrix in determining its time scales. This leads naturally to the problem of time scale modification, i.e. invariant factor placement, via state feedback. We present a result along these lines.

---

[1] Work of the first three authors supported by the Air Force Office of Scientific Research under Grant AFOSR-82-0258, and by the Army Research Office under Grant DAAG-29-84-K-0005.

[2] Department of Electrical Engineering and Computer Science and Laboratory for Information and Decision Systems, MIT, Cambridge, MA. 02139.

[3] Department of Electrical Engineering and Computer Science and Laboratory for Electromagnetic and Electronic Systems, MIT, Cambridge, MA. 02139.

I. Introduction

This paper is concerned with the multiple time scale analysis of the perturbed N-dimensional linear system

$$\dot{x}(t) = A(\epsilon)x(t) \tag{1.1}$$

where $A(\epsilon)$ has a Taylor expansion in the small parameter ϵ. If there is a drop in the rank of $A(\epsilon)$ at $\epsilon = 0$, the system (1.1) is termed singularly perturbed and can exhibit multiple time scale behavior. The analysis of such behavior has been the subject of a number of previous investigations. In particular several researchers [1], [2], [6], [10], [11], [12] have made numerous important contributions by investigating systems in what we will call explicit form[4]:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ \epsilon A_{21} & \epsilon A_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{1.2}$$

Let

$$\bar{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{1.3}$$

and let $\tilde{A}_{22}$ denote the Schur complement of A_{11} in $\bar{A}$:

$$\tilde{A}_{22} = A_{22} - A_{21}A_{11}^{-1}A_{12} \tag{1.4}$$

---

[4] The forms actually considered in [1], [2] have ϵ appearing on the left-hand side rather than the right-hand side. There is no significant difference in considering the form (1.2), since the systems in [1], [2] can be brought to this form by the change of time scale $\tau = t/\epsilon$.

It is known that if A_{11} and $\tilde{A}_{22}$ are nonsingular, the eigenvalues of (1.2) occur in two groups, one being of order 1 and lying "close" to the eigenvalues of A_{11}, and the other being of order ϵ and close to the eigenvalues of $\epsilon\tilde{A}_{22}$. If both the latter matrices are Hurwitz, then the system exhibits well-behaved two-time-scale structure, in the following sense:

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} x_{1f}(t) + x_{1s}(\epsilon t) + O(\epsilon) \\ x_{2s}(\epsilon t) + O(\epsilon) \end{bmatrix}, \quad t \geq 0, \tag{1.5}$$

where

$$\dot{x}_{1f}(t) = A_{11}x_{1f}(t), \qquad x_{1f}(0) = x_1(0) + A_{11}{}^{-1}A_{12}x_2(0)$$

$$x_{1s}(t) = - A_{11}{}^{-1}A_{12}x_{2s}(t), \tag{1.6}$$

and

$$\dot{x}_{2s}(t) = \tilde{A}_{22}\, x_{2s}(t), \qquad x_{2s}(0) = x_2(0).$$

The subcripts s and f denote <u>slow</u> and <u>fast</u> subsystems.

The $O(\epsilon)$ terms in (1.3) are uniform in $t \geq 0$, so that (1.5), (1.6) provide a uniform approximation of the state transition matrix of (1.1). That is,

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} ||e^{A(\epsilon)t} - T^{-1}e^{A_d(\epsilon)t}T|| = 0 \tag{1.7}$$

where

$$A_d(\epsilon) = \text{diag}\ (A_{11}, \epsilon\tilde{A}_{22}) \tag{1.8}$$

and

$$T = \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix} \tag{1.9}$$

The decomposition provided in (1.5) - (1.6) or, equivalently, in (1.7) - (1.9) has found significant applications. One important limitation of these results, however, is the assumption that the system is given in the explicit form (1.2) or its obvious generalizations (e.g. by expanding the A-matrix in (1.2) to include a third row of blocks, each of which is multiplied by ϵ^2). On the other hand, there are several advantages if the system has the form in (1.2). Specifically, there is a simple check to see if the system has a time scale decomposition in the sense of (1.7), (1.8) (namely A_{11} and $\tilde{A}_{22}$ must both be Hurwitz), one immediately knows what the time scales are, and the subsystems describing the behavior at each time scale are easily obtained.

In contrast to the results just described, we have the work of Coderch, et al. [3] which had as its principal objective the development of a general procedure for determining if the system (1.1) has well-defined time scale structure and for constructing a decoupled time-scale decomposition as in (1.7) with

$$A_d(\epsilon) = \text{diag}\,(\epsilon^{k_1}A_1,\ \epsilon^{k_2}A_2,\ldots,\epsilon^{k_m}A_m) \tag{1.10}$$

(and with an appropriate choice for T) without assuming that the system is in the special form of (1.2). This objective is achieved in [3] through a rather elaborate sequence of operations on the Taylor series coefficients of $A(\epsilon)$, involving cascaded projections onto progressively slower subspaces. A major advantage of this result is its generality -- with it we can analyze general systems as in (1.1) without assuming some special form. A price that is paid for this, however, is that the results and procedures developed are rather complicated, involve the computation of numerous pseudo-inverses, and generally do not lend themselves to easy interpretation or computation.

The work presented in this paper bridges the gap between the intuitively and computationally simple but somewhat restrictive results of [1], [2] and the quite general but rather complicated ones in [3]. The key to constructing this bridge is an examination of the algebraic structure of $A(\epsilon)$ considered as a matrix over the ring W of functions of ϵ that are analytic at $\epsilon = 0$. In particular, by considering the Smith form of $A(\epsilon)$ we not only provide the basis for transforming a general system (1.1) to its explicit form, but also make clear the role of the invariant factors of $A(\epsilon)$ in specifying the time scales present in the dynamics (1.1), a role that is suggested but not developed in [3]. This approach provides some valuable additional perspectives on the results in [1] - [3], and it also allows us to consider and solve a number of additional problems. Several of these are presented in the later sections of this paper, while others will be the subject of future papers. We note here that another approach to the main results of [3] is described in [6], which proceeds by transforming $A(\epsilon)$ to a block-diagonal form that is similar to it. There is a clear point of contact between our work and the results in [6], as our proof in Section 4 of the sufficiency of certain conditions for the existence of a time scale approximation is much in the spirit of the methods in [1], [2], and [6]. On the other hand, our results go significantly farther than previous efforts in that we, for the first time, make clear the role of the Smith form and the invariant factors of $A(\epsilon)$ and present a procedure that minimizes the number of ϵ-dependent computations required by identifying and discarding non-critical ϵ-dependent terms in $A(\epsilon)$ and in its explicit form.

In the next section we introduce a new definition of what we call a <u>strong</u> time scale decomposition. Based on this, we present a new result that allows us to state a strengthened version of the main result in [3] and to obtain a criterion for identifying higher-order terms in a system matrix $A(\epsilon)$ that can be discarded without affecting the investigation of the existence of strong time scale behavior. In Section 3 we then introduce the Smith form of $A(\epsilon)$ and use it to transform (1.1) to explicit form. We also perform some initial analysis that allows us to focus subsequent discussions on the case in which $A(\epsilon)$ is Hurwitz for $0 < \epsilon \leq \epsilon_0$ for some $\epsilon_0 > 0$. In Section 4 we develop what can be viewed as a generalization of the procedure in [1], [2] to analyze systems in explicit form. This produces both a set of necessary and sufficient conditions for a system to have a strong time scale decomposition and a procedure for constructing the corresponding strong multiple time scale approximation.

With these results established, we can then consider two important extensions. In Section 5 we consider a generalization of the definition of a time scale approximation that allows us to construct such approximations for a large class of systems violating the conditions of Section 4. In Section 6 we address the problem of modifying and controlling the time scales of the system

$$\dot{x}(t) = A(\epsilon)x(t) + B(\epsilon)u(t) \tag{1.11}$$

through the use of feedback

$$u(t) = K(\epsilon)x(t) \tag{1.12}$$

2. Well-Defined Multiple Time Scale Behavior

To begin this section we give two different definitions of what one might mean by well-defined multiple time scale behavior. The first of these is essentially the standard definition that is stated or implied in previous treatements. The seond, stronger definition is new, as it requires the consideration of an entire family of systems. By introducing this definition we can make several new observations concerning time scale decompositions and can give a stronger interpretation of the results in [3].

Definition 2.1: The system (1.1) has a multiple time scale decomposition if there exist constant matrices $A_1, A_2, \dots, A_n$, T and integers $0 \leq k_1 < k_2 < \dots < k_n$ such that

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} ||e^{A(\epsilon)t} - T^{-1}\exp \{\text{diag} \begin{bmatrix} \epsilon^{k_1}A_1, \epsilon^{k_2}A_2, \dots, \epsilon^{k_n}A_n \end{bmatrix} t\}T|| = 0 \qquad (2.1)$$

In this case we say that $[\{A_i\},\{k_i\},T]$ defines a multiple time scale decomposition of (1.1) or of $A(\epsilon)$.

To introduce the second definition we first need the following:[5]

Definition 2.2: The perturbed family $\mathcal{F}\{A(\epsilon)\}$ associated with the matrix $A(\epsilon)$ is defined as follows:

$$\mathcal{F}\{A(\epsilon)\} = \{U(\epsilon)A(\epsilon)V(\epsilon) | U(0) = V(0) = I\} \qquad (2.2)$$

Definition 2.3: The system (1.1) has a strong multiple time scale decomposition if there exist constant matrices $A_1, A_2, \dots, A_n$, T

---

[5] Throughout this paper we assume that all matrix functions of ϵ are analytic at zero.

and integers $0 \leq k_1 < k_2 < \ldots < k_n$ such that

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \left|\left| e^{F(\epsilon)t} - T^{-1} \exp \left\{ \operatorname{diag} \left[\epsilon^{k_1} A_1, \ldots, \epsilon^{k_n} A_n \right] t \right\} T \right|\right| = 0 \quad (2.3)$$

for <u>all</u> $F(\epsilon) \in \mathcal{F}\{A(\epsilon)\}$. In this case we say that $[\{A_i\},\{k_i\},T]$ <u>defines a strong time scale decomposition of</u> (1.1) <u>or of</u> $A(\epsilon)$.

Clearly the second of these definitions is significantly stronger that the first. Intuitively the elements of $\mathcal{F}\{A(\epsilon)\}$ should be thought of as mild perturbations of $A(\epsilon)$, and the strong-sense definition requires that any such perturbation must result in a system that has the same time scale decomposition as (1.1). More precisely, an immediate consequence of Definition 2.3 is that if $A(\epsilon)$ has a strong time scale decomposition, then any $G(\epsilon) \in \mathcal{F}\{A(\epsilon)\}$ is <u>asymptotically equivalent</u> to $A(\epsilon)$, i.e.

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \left|\left| e^{A(\epsilon)t} - e^{G(\epsilon)t} \right|\right| = 0 \quad (2.4)$$

To illustrate these ideas let us consider several examples. First, note that the scalar system

$$\dot{x}(t) = x(t) \quad (2.5)$$

trivially has a time scale decomposition according to Definition 2.1. but not according to Definition 2.3 since $(1+\epsilon) \in \mathcal{F}\{1\}$ is not asymptotically equivalent to 1. On the other hand, it is not difficult to check (and is immediate from the results in several papers) that

$$\dot{x}(t) = - x(t) \quad (2.6)$$

<u>does</u> have a strong time scale decomposition.

Consider next the system matrix

$$A(\epsilon) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{2.7}$$

This matrix has a trivial time scale decomposition, but it does not have a strong time scale decomposition, since it is not asymptotically equivalent to the matrix

$$F(\epsilon) = \begin{bmatrix} -\epsilon & 1 \\ -1 & -\epsilon \end{bmatrix} = \begin{bmatrix} 1 & \epsilon \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & \epsilon \\ 0 & 1+\epsilon^2 \end{bmatrix} \tag{2.8}$$

Finally, we note that

$$A(\epsilon) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{2.9}$$

does not have a strong time scale decomposition since it is not asymptotically equivalent to

$$F(\epsilon) = \begin{bmatrix} 0 & 1+\epsilon \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1+\epsilon & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{2.10}$$

These examples indicate that there are problems when there are eigenvalues that are in the right-half plane, are purely imaginary, or are zero with nontrivial Jordan blocks. To see that these examples span all possible cases, we need to briefly re-examine and strengthen the main result in [3]. In particular, although it is not discussed, [3] in fact provides the basis for determining if a system has a <u>strong</u> time scale decomposition and for constructing that decomposition. The system considered in [3] is the singularly perturbed LTI system (1.1) with a slight change in notation whose purpose will become clear shortly.

$$\dot{x}(t) = A_0(\epsilon)x(t) \tag{2.11}$$

where the matrix $A_0(\epsilon)$ is an analytic function of ϵ at $\epsilon = 0$. Suppose $A_0(\epsilon)$ has eigenvalues $\lambda_1(\epsilon), \ldots\ldots, \lambda_n(\epsilon)$ where $\lambda_i(\epsilon) \to 0$, $\epsilon \to 0$, $i=1, \ldots\ldots, m \leq n$. 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 eigenvectors and generalized eigenvectors corresponding to $\lambda_i(\epsilon), \ldots\ldots, \lambda_m(\epsilon)$ of $A_0(\epsilon)$ [7].

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

$$A_0(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p F_{0p} \tag{2.12}$$

It can be proven [3, 7] that if F_{00} has <u>semisimple nullstructure</u> (SSNS) -- i.e. if its zero eigenvalue is semisimple, that is, has geometric multiplicity equal to its algebraic multiplicity -- then the matrix

$$A_1(\epsilon) = P_0(\epsilon)A_0(\epsilon)/\epsilon \tag{2.13}$$

has a series expansion of the form

$$A_1(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p F_{1p} \tag{2.14}$$

(otherwise $A_1(\epsilon)$ will have ϵ^{-1} terms). If F_{10} also has SSNS we define $A_2(\epsilon)$ as

$$\begin{aligned} A_2(\epsilon) &= P_1(\epsilon)A_1(\epsilon)/\epsilon = P_1(\epsilon)P_0(\epsilon)A_0(\epsilon)/\epsilon^2 \\ &= \sum_{p=0}^{\infty} \epsilon^p F_{2p} \end{aligned} \tag{2.15}$$

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

$$A_n(\epsilon) = P_{n-1}(\epsilon)A_{n-1}(\epsilon)/\epsilon = P_{n-1}(\epsilon)\dots P_0(\epsilon)A_0(\epsilon)/\epsilon^n = \sum_{p=0}^{\infty} \epsilon^p F_{np} \qquad (2.16)$$

if the matrix F_{n0} does not have SSNS or if $\text{rank}F_{00}+\text{rank}F_{10}+\dots+\text{rank}F_{n0}$ equals the normal rank of $A_0(\epsilon)$, i.e. the constant rank that $A_0(\epsilon)$ takes on some interval $(0, \epsilon_0]$. A matrix $A_0(\epsilon)$ is said to satisfy the multiple semisimple null structure (MSSNS) condition if the latter of these conditions holds. If in addition, all F_{k0} are semistable -- i.e. if for each k F_{k0} has SSNS and all of its nonzero eigenvalues have strictly negative real parts -- then we say that $A_0(\epsilon)$ satisfies the multiple semistability (MSST) condition.

The main result of [3] is that if $A(\epsilon)$ satisfies MSST, then (i)

$$F_{k0} = \begin{cases} T^{-1}\text{diag}(0,\dots,0,A_i,0,\dots,0)T & , k = k_i \\ 0 & \text{otherwise} \end{cases} \qquad (2.17)$$

for some nonsingular T, semistable A_i, and uniquely determined integers k_i; and (ii) $A(\epsilon)$ has a time scale decomposition in the sense of Definition 2.1. On the other hand, as our examples (2.5), (2.7), (2.9) show, MSST is not necessary for $A(\epsilon)$ to have a time scale decomposition. What we show in Theorem 2.1 is that MSST is necessary and sufficient for $A(\epsilon)$ to have a strong time scale decomposition.

In order to prove our strengthened version of the main result in [3] we need two results.

Proposition 2.1: Let $G(\epsilon) \in \mathcal{F}\{A(\epsilon)\}$. Then

$$F^G_{k0} = F^A_{k0} \qquad \forall k \qquad (2.18)$$

where the superscripts "G" and "A" denote the sequences defined in (2.12) - (2.16) for $G(\epsilon)$ and $A(\epsilon)$, respectively.

Proof: See Appendix A.

Proposition 2.2: Suppose that $[\{A_i\},\{k_i\},T]$ defines a multiple time scale decomposition of $A_0(\epsilon)$ and suppose further that $A_1,\dots,A_n$ are semistable. Then (2.17) holds and hence $A(\epsilon)$ satisfies the MSST condition.

Proof: See Appendix B

We can now state:

Theorem 2.1: The system (2.11) has a strong time scale decomposition if and only if $A_0(\epsilon)$ satisfies the MSST condition.

Proof: As stated previously, it is proved in [3] that the MSST condition is sufficient to satisfy the weaker Definition 2.1. That this condition is sufficient for the stronger definition follows directly from Proposition 2.1. The proof of necessity is also straightforward. Specifically if $A_0(\epsilon)$ has a strong time scale decomposition as in (2.1), then, thanks to Proposition 2.2, all we need to show is that the A_i must be semistable. This can be done by contradiction. Specifically, if A_i is not semistable, then it has a right-half plane eigenvalue, a pair of purely imaginary eigenvalues, or a nontrivial Jordan block corresponding to the 0 eigenvalue. Showing that any of these conditions preclude the existence of a strong time scale decomposition is a minor variation on our previous discussion of the three examples (2.5), (2.7), (2.9).

Note that if $A(\epsilon)$ is invertible for $\epsilon \in (0, \epsilon_0]$, the A_k in the strong time scale decomposition are all Hurwitz.

Finally, it is also shown in [3] that if $A_0(\epsilon)$ does not satisfy MSST, then for some q the limit as $\epsilon \downarrow 0$ of

$$\exp\, [A_0(\epsilon)t/\epsilon^q] \tag{2.19}$$

does not exist. This indicates that a failure of the MSST condition does correspond to some type of noncovergent behavior. However, the precise meaning and interpretation of this could not easily be exposed without the concept of a strong time scale decomposition. Indeed, in addition to providing us with Theorem 2.1, this machinery makes it far simpler to prove the nonexistence of the limit of (2.19). Furthermore, we now see that to verify the MSST condition and to construct a time scale decomposition for $A(\epsilon)$, we can equivalently examine these questions using <u>any</u> element of $\mathcal{F}\{A(\epsilon)\}$ -- i.e. any such element must generate the same sequence F_{k0} if a strong time scale decomposition exists. Of course we can equivalently consider any element of $\mathcal{F}\{SA(\epsilon)S^{-1}\}$ where S is any constant invertible matrix. We make use of these facts in the next section to transform an arbitrary $A(\epsilon)$ to its explicit form.

3. Explicit Form

As mentioned in Section 1, our new approach employs the Smith decomposition of $A(\epsilon)$ over the ring W of functions of ϵ that are analytic at $\epsilon = 0$ (see [4], [5]). The units of W are elements that do not vanish at $\epsilon = 0$. That is, since any element of W can be expanded in a Taylor series,

$$a_0 + a_1\epsilon + a_2\epsilon^2 + \dots \tag{3.1}$$

we see that the set of units are those elements with $a_0 \neq 0$. It is also easily seen that W is a Euclidean domain, with the degree or order, $O(d(\epsilon))$, of any element $d(\epsilon) \in W$ being defined as the order of the first nonzero term in its Taylor expansion. Therefore $A(\epsilon)$ has a Smith decomposition

$$A(\epsilon) = P(\epsilon)D(\epsilon)Q(\epsilon) \tag{3.2}$$

where $P(\epsilon)$ and $Q(\epsilon)$ are unimodular, i.e. $|P(\epsilon)|$ and $|Q(\epsilon)|$ are units (and thus $P^{-1}(\epsilon)$ and $Q^{-1}(\epsilon)$ are matrices over W) or, equivalently

$$|P(0)| \neq 0 \ , \qquad |Q(0)| \neq 0 \tag{3.3}$$

and

$$D(\epsilon) = \text{diag}\ (\epsilon^{k_1}I,\dots,\epsilon^{k_n}I,\ 0) \tag{3.4}$$

where $0 \leq k_1 < k_2 < \dots < k_n$ are integers, the identity matrices I may have different dimensions, and the 0 matrix is only present if $A(\epsilon)$ is singular in a neighborhood of $\epsilon = 0$. The ϵ^{k_i} are called the <u>invariant factors</u> of $A(\epsilon)$ Actual computation of such Smith decompositions is discussed in [4] and [5] (in the terminology of [5], what is required is to transform $A(\epsilon)$ to the matrix $D(\epsilon)Q(\epsilon)$ which is "row-reduced at 0" through row operations embodied in

$P^{-1}(\epsilon)$). Without loss of generality we assume from here on that $k_1 = 0$; this can always be obtained by a change of time scale in (1.1).

Rather than working with the system (1.1), we consider an ϵ-independent change of variables

$$y(t) = P^{-1}(0)x(t) \tag{3.5}$$

so that

$$\dot{y}(t) = P^{-1}(0)P(\epsilon)D(\epsilon)Q(\epsilon)P(0) \tag{3.6}$$

Next we note that if we define the constant matrix

$$\bar{A} = Q(0)P(0) \tag{3.7}$$

then

$$D(\epsilon)\bar{A} \in \mathcal{S}\{P^{-1}(0)P(\epsilon)D(\epsilon)Q(\epsilon)P(0)\} \tag{3.8}$$

(premultiply $P^{-1}(0)P(\epsilon)D(\epsilon)Q(\epsilon)P(0)$ by $P^{-1}(\epsilon)P(0)$ and postmultiply by $P^{-1}(0)Q^{-1}(\epsilon)Q(0)P(0)$). Therefore, we arrive at the <u>explicit form</u> of (1.1):

$$\dot{z} = D(\epsilon)\bar{A}z \tag{3.9}$$

which, if we express $\bar{A}$ in block form with blocks compatible with those in (3.4), can be written as

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \vdots \\ \dot{z}_n \\ \dot{z}_{n+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1,n+1} \\ \epsilon^{k_2} A_{21} & \epsilon^{k_2} A_{22} & \cdots & \epsilon^{k_2} A_{2,n+1} \\ \vdots & \vdots & & \vdots \\ \epsilon^{k_n} A_{n1} & \epsilon^{k_n} A_{n2} & \cdots & \epsilon^{k_n} A_{n,n+1} \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \\ z_{n+1} \end{bmatrix} \tag{3.10}$$

Let us make several comments about the previous transformations. First of all, note that every element of $\mathcal{F}\{A(\epsilon)\}$ has the same explicit form. Secondly, if $A(\epsilon)$ does not have a strong time scale decomposition, then, as in the examples in Section 2, there is no reason to expect that (3.9) is a good approximation of (3.6) (and therefore of (1.1)) in that the two systems need not be asymptotically equivalent. However, if any of the systems (1.1), (3.6), or (3.9) has a strong time scale decomposition, then they all do, and (3.9) is asymptotically equivalent to (3.6). Therefore, we can focus on the explicit form if we are interested in strong time scale decompositions. Finally, note that the system (3.10) is an obvious generalization of (1.2), and this observation provides the basis for our development in the next section. Before doing this, however, we first conclude this section by showing how we can deal with the 0 diagonal block in $D(\epsilon)$ so that hereafter we can focus attention on the case in which there is no such block, i.e. the case in which $A(\epsilon)$ is Hurwitz for $\epsilon \in (0, \epsilon_0]$.

Specifically, let us write $D(\epsilon)$ in (3.4) as

$$D(\epsilon) = \text{diag}\,(D_1(\epsilon),\ 0) \tag{3.11}$$

(so that $D_1(\epsilon)$ consists of all of the nonzero invariant factors), and let us express $\bar{A}$ in (3.7) in blocks compatible with (3.11).

$$\bar{A} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \tag{3.12}$$

We then have that

$$D(\epsilon)\bar{A} = \begin{bmatrix} D_1(\epsilon)G_{11} & D_1(\epsilon)G_{12} \\ 0 & 0 \end{bmatrix} \tag{3.13}$$

Note that $(G_{11}\ G_{12})$ has full row rank since $\bar{A}$ is invertible. In fact, it is immediate from the development in the next section that $D(\epsilon)\bar{A}$ has MSSNS only

<u>if</u> G_{11} is invertible. Therefore, as a first step in our overall procedure, we check the invertibility of G_{11}. If it is not invertible, then we immediately know that (3.9) and hence (1.1) do not have strong time scale decompositions. If G_{11} is invertible, we perform the following ϵ-independent transformation of (3.9)

$$\omega = \begin{bmatrix} I & G_{11}^{-1}G_{12} \\ 0 & I \end{bmatrix} z \tag{3.14}$$

so that

$$\dot{\omega} = \begin{bmatrix} D_1(\epsilon)G_{11} & 0 \\ 0 & 0 \end{bmatrix} \omega \tag{3.15}$$

From this point on we can focus completely on the lower-dimensional, explicit form matrix $D_1(\epsilon)G_{11}$ which is invertible for $\epsilon \in (0, \epsilon_0]$. If this has a strong time scale decomposition, then so do (3.9) and (1.1), and the construction of the time scale approximations for these systems from the one for (3.15) involves the obvious reversal of the steps taken to obtain (3.15) from (1.1).

4. Strong Multiple Time Scale Decompositions of Systems in Explicit Form

Based on the development and discussion in the previous section, we now focus attention on the following system in explicit form

$$\dot{z}(t) = D(\epsilon)\bar{A}z(t) \tag{4.1}$$

where $D(\epsilon) = \text{diag}\ (I,\ \epsilon^{k_2}I,\dots,\epsilon^{k_n}I)$ and

$$\bar{A} = \left[\begin{array}{c|ccc} A_{11} & A_{12} & \cdots & A_{1n} \\ \hline A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{array}\right] = \left[\begin{array}{c|c} \tilde{A}_{11} & R_1 \\ \hline W_1 & Z_1 \end{array}\right] \tag{4.2}$$

is invertible. The reasons for the notation introduced in (4.2) will become clear shortly (here the dashed line in both matrices are in the same locations, so that $\tilde{A}_{11} = A_{11}$, $R_1 = [A_{12}\dots A_{1n}]$. etc).

One direct approach to determining necessary and sufficient conditions under which (4.1) (and thus (1.1)) has a strong time scale decomposition is to identify explicitly the projections and similarity transformations used in [3] to check for MSST and to obtain the multiple time scale decomposition described in Theorem 2.1. This is done in detail in [8]. What we do in this section is to follow an approach that makes use of the results in Section 2 to obtain a set of necessary and sufficient conditions and a procedure for constructing a multiple time scale decomposition that is much more in the spirit of [1] and [2]. Based on our initial review of the analysis of (1.2), it should not come as a surprise that successive Schur complements of $\bar{A}$ play an important role in our development. Also, since we are focusing on strong time scale decompositions, we have the luxury of throwing away many of the ϵ-dependent terms that arise as we proceed. Specifically, whenever we run into a unimodular matrix $U(\epsilon)$ multiplying our system matrix on the left or right, we can replace it by $U(0)$ and continue. Either both of these systems have the same strong time scale decompositions or neither one has such a decomposition.

The basic idea behind the approach we use here is to block-diagonalize $D(\epsilon)\bar{A}$. We do this in stages, "peeling off" one time scale of (4.1) at a time, starting with the fastest. To begin, let us introduce some notation. Specifically, let $D_1(\epsilon) = D(\epsilon)$, $\bar{A}_1 = \bar{A}$, and

$$D(\epsilon)\bar{A} = D_1(\epsilon)\bar{A}_1 = \left[\begin{array}{c:c} \tilde{A}_{11} & R_1 \\ \hline \epsilon^{k_2} S_1(\epsilon) & \epsilon^{k_2} F_1(\epsilon) \end{array}\right] \tag{4.3}$$

$$S_1(\epsilon) = D_2(\epsilon)W_1 \tag{4.4}$$

$$F_1(\epsilon) = D_2(\epsilon)Z_1 \tag{4.5}$$

$$D_2(\epsilon) = \mathrm{diag}(I, \epsilon^{k_3-k_2}I, \ldots, \epsilon^{k_n-k_2}I) \tag{4.6}$$

(here the dimensions of the (n-1) identity matrices in (4.6) are the same as the last (n-1) blocks in $D(\epsilon)$).

As a next step we prove the following:

<u>Lemma 4.1</u>: Consider the constant matrix

$$M = \begin{bmatrix} N & L \\ 0 & 0 \end{bmatrix} \tag{4.7}$$

where (N, L) has full row rank and N is square. Then M has SSNS if and only if N is invertible.

Proof: Suppose N is invertible. Then

$$\begin{bmatrix} I & N^{-1}L \\ 0 & I \end{bmatrix} M \begin{bmatrix} I & -N^{-1}L \\ 0 & I \end{bmatrix} = \begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix} \tag{4.8}$$

which clearly has SSNS. On the other hand, if N is not invertible, then there exists $x \neq 0$ so that $Nx = 0$. Furthermore, since (N, L) has full row rank, we can find x_1 and x_2 so that $Nx_1 + Lx_2 = x$. If we then define

$$z = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4.9 }$$

we have that $Mz \neq 0$ but $M^2z = 0$, showing that M does not have SSNS.

Letting $\epsilon = 0$ in (4.3), we have

$$D_1(0)\bar{A}_1 = \begin{bmatrix} \tilde{A}_{11} & R_1 \\ 0 & 0 \end{bmatrix} \tag{4.10}$$

Since $\bar{A}_1$ is invertible, $[\tilde{A}_{11}, R_1]$ has full row rank. Consequently, from Lemma 4.1 we see that the system matrix (4.10) describing evolution at the fastest time scale has SSNS if and only if $\tilde{A}_{11}$ is invertible. Suppose, then, that $\tilde{A}_{11}$ is invertible. Consider the similarity transformation

$$G(\epsilon) = \begin{bmatrix} I & \tilde{A}_{11}{}^{-1}R_1 \\ 0 & I \end{bmatrix} D_1(\epsilon)\bar{A}_1 \begin{bmatrix} I & -\tilde{A}_{11}R_1 \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} \tilde{A}_{11} + \epsilon^{k_2}\tilde{A}_{11}{}^{-1}R_1D_2(\epsilon)W_1 & \epsilon^{k_2}\tilde{A}_{11}{}^{-1}R_1D_2(\epsilon)\bar{A}_2 \\ \epsilon^{k_2}D_2(\epsilon)W_1 & \epsilon^{k_2}D_2(\epsilon)\bar{A}_2 \end{bmatrix} \tag{4.11}$$

where

$$\bar{A}_2 = Z_1 - W_1\tilde{A}_{11}^{-1}R_1 \tag{4.12}$$

which is invertible (since $\bar{A}_1$ and $\tilde{A}_{11}$ in (4.2) are both invertible). Note further that

$$G(\epsilon) = U(\epsilon) \begin{bmatrix} \tilde{A}_{11} & 0 \\ 0 & \epsilon^{k_2} D_2(\epsilon)\bar{A}_2 \end{bmatrix} V(\epsilon) \tag{4.13}$$

where U(0) = V(0) = I (see Appendix C).

Since we are interested in strong time scale decompositions, we can discard $U(\epsilon)$ and $V(\epsilon)$. From Proposition 2.2 and Theorem 2.1 we can immediately conclude that for diag $(\tilde{A}_{11}, \epsilon^{k_2} D_2(\epsilon)\bar{A}_2)$ to have a strong time scale decomposition, $\tilde{A}_{11}$ must be Hurwitz. Furthermore, we have now reduced the problem to the examination of the explicit form matrix $D_2(\epsilon)\bar{A}_2$ with one fewer time scale.

Consider now the following recursion beginning with $\bar{A}_1$ in (4.2) and defined recursively as follows

$$\bar{A}_i = \begin{bmatrix} \tilde{A}_{ii} & R_i \\ W_i & Z_i \end{bmatrix} \tag{4.14}.$$

$$\bar{A}_{i+1} = Z_i - W_i\tilde{A}_{ii}^{-1}R_i \tag{4.15}$$

Here the block size of each $\tilde{A}_{ii}$ is the same as that of the ith block in the original explicit form systems (4.1), (4.2)[6]. Using the results of Section 2

---

[6] Note that at the last step $\bar{A}_n = \tilde{A}_{nn}$.

then yields the following:

Theorem 4.1: The explicit form system (4.2) has MSSNS if and only if each $\tilde{A}_{ii}$ is invertible. Furthermore, the system (4.2) satisfies the MSST condition, and hence has a strong time scale.decomposition if and only if each of the $\tilde{A}_{ii}$ is Hurwitz. In this case

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} ||e^{D(\epsilon)\bar{A}t} - T^{-1} \exp \{\operatorname{diag}[\tilde{A}_{11}, \epsilon^{k_2}\tilde{A}_{22}, \ldots, \epsilon^{k_n}\tilde{A}_{nn}]t\}T|| = 0 \tag{4.16}$$

where

$$T = T_{n-1} \cdots T_1 \tag{4.17}$$

$$T_1 = \begin{bmatrix} I & \tilde{A}_{11}^{-1}R_1 \\ & \\ 0 & I \end{bmatrix} \tag{4.18}$$

$$T_i = \left[\begin{array}{c:cc} I & 0 & \\ \hdashline & I & \tilde{A}_{ii}^{-1}R_i^{-1} \\ 0 & & \\ & 0 & I \end{array}\right], \qquad i > 1 \tag{4.19}$$

(here the upper left-hand identity block is of dimension equal to the first (i-1) blocks of (4.1), (4.2)).

We close this section with several final comments. First, note that the recursive procedure just described for peeling off successively slower time scales actually yields a sequence of approximations over successively longer time intervals, i.e.

$$\lim_{\epsilon \downarrow 0} \sup_{t \in [0, \epsilon^{1-k_{r+1}})} ||e^{D(\epsilon)\bar{A}t} - T^{-1} \exp \{\operatorname{diag}[\tilde{A}_{11}, \ldots, \epsilon^{k_r}\tilde{A}_{rr}, 0 \ldots, 0]t\}T|| = 0 \tag{4.20}$$

(see [3, 6] for similar comments). Secondly, we note that an alternate approach to showing the sufficiency of the conditions in Theorem 4.1 is presented in [9] using an approach much in the spirit of [6]. Specifically, consider the following equations

$$R_1 + \tilde{A}_{11}L_1(\epsilon) - \epsilon^{k_2}L_1(\epsilon)[F_1(\epsilon) + S_1(\epsilon)L_1(\epsilon)] = 0 \tag{4.21}$$

$$H_1(\epsilon)[\tilde{A}_{11} - \epsilon^{k_2}L_1(\epsilon)S_1(\epsilon)] - \epsilon^{k_2}[F_1(\epsilon) + S_1(\epsilon)L_1(\epsilon)]H_1(\epsilon) + S_1(\epsilon) = 0 \tag{4.22}$$

It is straightforward to check that these equations have solutions $L_1(\epsilon)$ and $H_1(\epsilon)$ for ϵ small enough, that

$$L_1(0) = - \tilde{A}_{11}^{-1}R_1, \qquad H_1(0) = - S_1(0)\tilde{A}_{11}^{-1} \tag{4.23}$$

that

$$T_1(\epsilon) = \begin{bmatrix} I & -L_1(\epsilon) \\ \epsilon^{k_2}H_1(\epsilon) & I-\epsilon^{k_2}H_1(\epsilon)L_1(\epsilon) \end{bmatrix} \tag{4.24}$$

is unimodular, and that the similarity transformation specified by $T_1(\epsilon)$ block-diagonalizes $D_1(\epsilon)\bar{A}_1$, i.e.

$$T_1(\epsilon)D_1(\epsilon)\bar{A}_1T^{-1}(\epsilon) = \begin{bmatrix} G_1(\epsilon) & 0 \\ 0 & \epsilon^{k_2}G_2(\epsilon) \end{bmatrix} \tag{4.25}$$

where

$$G_1(\epsilon) = \tilde{A}_{11} - \epsilon^{k_2} L_1(\epsilon) S_1(\epsilon) \qquad G_2(\epsilon) = F_1(\epsilon) + S_1(\epsilon) L_1(\epsilon) \tag{4.26}$$

Noting then that $G_1(0) = \tilde{A}_{11}$ and that $G_2(\epsilon) = D_2(\epsilon)\bar{A}_2 C(\epsilon)$ where $C(0) = I$, we can conclude that $D_1(\epsilon)\bar{A}_1$ has a strong time scale decomposition if and only if

$$T_1^{-1}(0) = \begin{bmatrix} A_{11} & 0 \\ 0 & \epsilon^{k_2} D_2(\epsilon)\bar{A}_2 \end{bmatrix} T_1(0) \tag{4.27}$$

does, where

$$T_1^{-1}(0) = \begin{bmatrix} I & \tilde{A}_{11}^{-1} R_1 \\ 0 & I \end{bmatrix} \tag{4.28}$$

This process can then be iterated to consider the next time scale.

Comparing this procedure to that described previously, and in particular to (4.11) and the subsequent development, we see that, thanks to Theorem 2.1, we do not have to do quite so much work (although, as described in Appendix A, we actually use this full block-diagonalization procedure in the proof of Proposition 2.1). Rather, instead of fully block-diagonalizing $D_1(\epsilon)\bar{A}_1$ using the full $T_1(\epsilon)$, we simply use $T_1(0)$, the key being that we have raised the order of the upper right-hand element of (4.11) *sufficiently* so that (4.13) holds. In a sense what we have done in (4.11) is a first step in an iterative approach to block-diagonalizing $D_1(\epsilon)\bar{A}_1$. Specifically, think of the transformation in (4.11) as an attempt to approximately null out the (1, 2) block of $D_1(\epsilon)\bar{A}_1$ by raising its order. If we then attempt to approximately null out the (2, 1) block of $G(\epsilon)$ (using a lower-block-triangular similarity transformation), we will raise the order of this term. Carrying this process on for a number of steps we obtain better and better approximate block diagonalizations and hence have a series expansion for $T_1(\epsilon)$. What we have shown here is that when looking for strong time scale decompositions, we can stop after the first term in the series. In the next section we describe a procedure for constructing a weaker form of a time scale decomposition for systems not satisfying the MSST condition. This procedure requires keeping additional terms of the series or, equivalently, performing the iterative, approximate block-diagonalization procedure for more than one iteration.

5. Time Scale Decompositions for Systems Without MSST

In this section we describe a procedure for constructing a somewhat weaker time scale decomposition for systems that do not satisfy the MSST condition. To motivate and illustrate the essential ideas behind this procedure, we begin with an example. Specifically, consider the system matrix

$$A(\epsilon) = \begin{bmatrix} -\epsilon & 1 \\ -1 & -\epsilon \end{bmatrix} \tag{5.1}$$

Since A(0) is not semistable we immediately see that this matrix does not have a strong time scale decomposition. In fact, it is not difficult to see that it does not even have a time scale decomposition in the sense of Definition 2.1. The reason for this stems from the requirement that the system matrices A_1, A_2,... in (2.1) be independent of ϵ. Examining $A(\epsilon)$ in (5.1) we see that its eigenvalues $(- \epsilon \pm j)$ have the property that their real parts are of higher order in ϵ than their imaginary parts. Consequently, when we attempt to use a constant system matrix to approximate (5.1) we throw away the crucial damping. From this perspective it seems evident that what one should seek to do in this case is to keep at least some of the ϵ-dependent terms in $A(\epsilon)$ in order to preserve its principal damping characteristics. The procedure we develop in this section does exactly that.

We begin our development with the following

Definition 5.1: Let $A(\epsilon)$ be Hurwitz for $\epsilon \in (0, \epsilon_0]$ and let the Smith form of $A(\epsilon)$ be as in (3.2) with $D(\epsilon) =$ $\text{diag}\,(I, \epsilon I, \ldots, \epsilon^{n-1} I)$. Then $A(\epsilon)$ has a weak multiple time scale decomposition if

$$\lim_{\epsilon\downarrow 0}\ \sup_{t\geq 0} ||e^{A(\epsilon)t} - T^{-1}\exp\{\mathrm{diag}[A_0(\epsilon), \epsilon A_1(\epsilon), \ldots, \epsilon^{n-1}A_n(\epsilon)]t\}T|| = 0 \tag{5.2}$$

where T is a constant matrix and each of the $A_i(\epsilon)$ has the properties that $A_i(0)$ is invertible and each of its purely imaginary eigenvalues is semisimple (i.e. has algebraic multiplicity equal to its geometric multiplicity).

Let us make several comments about this definition. First, using the procedure described at the end of Section 3 we can actually weaken the assumption that $A(\epsilon)$ is Hurwitz by assuming only that $A(\epsilon)$ is semistable for $\epsilon \in (0, \epsilon_0]$ (so that there may be a 0 block in $D(\epsilon)$); however for simplicity here we use the stronger assumption. Also, the assumption that $D(\epsilon)$ has the particular form stated in the definition is no real restriction and again we include it here for convenience only (if some power of ϵ between 0 and n-1 is not an invariant factor, then the corresponding step of our procedure is simply dropped). Finally, let us discuss the assumptions on $A_i(0)$. Note first that requiring $A_i(0)$ to be invertible is equivalent to assuming that $A(\epsilon)$ has MSSNS, while the further semisimplicity assumption eliminates matrices such as

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

which are Hurwitz for $\epsilon > 0$ but for which

$$\sup_{t \geq 0} ||\exp\{F(\epsilon)t\}||$$

grows without bounds as $\epsilon \downarrow 0$. In essence what we are considering in this section is the extension of our theory of time scale decompositions to include $A(\epsilon)$'s with eigenvalues that converge to points on the imaginary axis other than the origin. Consequently, it is not surprising that the multiple semisimplicity condition is extended to include all eigenvalues converging to the imaginary axis.

> Definition 5.2: A matrix $A(\epsilon)$ has multiple semisimple imaginary eigenstructure (MSSIES) if it has MSSNS and if each of the purely imaginary eigenvalues of each of the $\tilde{A}_{ii}$ defined in Theorem 4.1 is semisimple.

Essentially by definition we have that MSSIES is necessary for $A(\epsilon)$ to have a weak time scale decomposition.[7] In fact, the procedure we describe in this section proves the following:

> Theorem 5.1: Let $A(\epsilon)$ be Hurwitz for $\epsilon \in (0, \epsilon_0]$. Then $A(\epsilon)$ has a weak multiple time scale decomposition if and only if it has MSSIES.

---

[7] Indeed if this is not the case, then (5.2) leads to a contradiction, since

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} ||e^{A(\epsilon)t}|| = \infty$$

but $\exp\{\mathrm{diag}[A_0(\epsilon), \ldots, \epsilon^{n-1}A_{n-1}(\epsilon)]t\}$ is uniformly bounded.

For the remainder of this section we assume that $A(\epsilon)$ is Hurwitz and has MSSIES. As a first step in our procedure, we transform the dynamics of (1.1) in a manner similar to that used in Section 2. Specifically, let $A(\epsilon)$ have the Smith form given in (3.2) and define

$$y(t) = P^{-1}(\epsilon)x(t) \tag{5.3}$$

so that

$$\dot{y}(t) = D(\epsilon)\bar{A}(\epsilon)y(t) \tag{5.4}$$

where $\bar{A}(\epsilon) = Q(\epsilon)P(\epsilon)$. In Section 3 and 4 we performed a slightly different similarity transformation and also replaced $\bar{A}(\epsilon)$ by $\bar{A} = \bar{A}(0)$. In the present context we <u>cannot</u> throw away the ϵ-dependent terms in $\bar{A}(\epsilon)$. However, as the following result shows, we can do so in the similarity transformation relating $x(t)$ and $y(t)$.

> <u>Lemma 5.1</u>: Suppose that $(A_1(\epsilon),\ldots,A_n(\epsilon);\ T)$ defines a weak time scale decomposition of $D(\epsilon)\bar{A}(\epsilon)$. Then $(A_1(\epsilon),\ldots,A_n(\epsilon);\ TP^{-1}(0))$ defines one for $A(\epsilon)$.

<u>Proof</u>: See Appendix D

Let us introduce some notation. Specifically, let

$$A_1^{\,0}(\epsilon) = D(\epsilon)\bar{A}(\epsilon) = \begin{bmatrix} A_{11}^{\,0}(\epsilon) & R_1^{\,0}(\epsilon) \\ \epsilon S_1^{\,0}(\epsilon) & \epsilon F_1^{\,0}(\epsilon) \end{bmatrix} \tag{5.5}$$

where, as in (4.3), the partition indicated is compatible with that of

$$D(\epsilon) = \mathrm{diag}(I,\ \epsilon I,\ldots,\epsilon^{n-1}I) = \mathrm{diag}(I,\ \epsilon D_2(\epsilon)) \tag{5.6}$$

By assumption $A_1^{\,0}(\epsilon)$ has MSSNS, so $A_{11}^{\,0}(\epsilon)$ is unimodular. Consider next an arbitrary (possibly ϵ-dependent) matrix

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \tag{5.7}$$

and define two similarity transformations on F:

$$\Theta F \Theta^{-1} = \begin{bmatrix} I & F_{11}^{-1}F_{12} \\ 0 & I \end{bmatrix} F \begin{bmatrix} I & -F_{11}^{-1}F_{12} \\ 0 & I \end{bmatrix} \tag{5.8}$$

$$\Psi F \Psi^{-1} = \begin{bmatrix} I & 0 \\ -F_{21}F_{11}^{-1} & I \end{bmatrix} F \begin{bmatrix} I & 0 \\ F_{21}F_{11}^{-1} & I \end{bmatrix} \tag{5.9}$$

We also define a thrid similarity transformation, Γ obtained by first applying the Θ-transformation to F and then applying the Ψ transformation to $\Theta F \Theta^{-1}$ (i.e. we construct Ψ using the blocks of $\Theta F \Theta^{-1}$). We can now state the following.

<u>Lemma 5.2:</u> Define the following sequences of matrices:

$$\Gamma_1^{\ i}(\epsilon) = \Gamma\text{-transformation for } A_1^{\ i}(\epsilon) \tag{5.10}$$

$$A_1^{\ i+1}(\epsilon) = \Gamma_1^{\ i}(\epsilon) A_1^{\ i}(\epsilon) \Gamma_1^{\ i}(\epsilon)^{-1} \tag{5.11}$$

where $A_1^{\ 0}(\epsilon)$ is given in (5.5). Then $A_1^{\ i}(\epsilon)$ has the form

$$A_1{}^i(\epsilon) = \begin{bmatrix} A_{11}{}^i(\epsilon) & \epsilon^i R_1{}^i(\epsilon) \\ \epsilon^{i+1} S_1{}^i(\epsilon) & \epsilon F_1{}^i(\epsilon) \end{bmatrix} \tag{5.12}$$

where $A_{11}{}^i(\epsilon)$ is unimodular. Furthermore, $A_{11}{}^i(\epsilon)$ and $\epsilon F_1{}^i(\epsilon)$ converge to the matrices appearing in the block-diagonalization of $A_1(\epsilon)$ obtained as in (4.21) - (4.26).

Proof: Equation (5.12) can be verified by direct calculation. See [6,10,13] for the convergence result (which is not used in what follows).

In Section 4, we contented ourselves both with replacing $\bar{A}(\epsilon)$ in (5.5) with $\bar{A}(0) = \bar{A}$ and with performing only the first step of the iteration. In the present context we can do neither of these. On the other hand, it is still not necessary to go to the limit. To make this precise, we begin with some notation. Specifically, let N denote the dimension of $A(\epsilon)$; $\lambda_i(\epsilon))$ the eigenvalues of $A(\epsilon)$; and M an upper bound on the maximum order of the real parts of the $\lambda_i(A(\epsilon))$, i.e.

$$O(\mathrm{Re}[\lambda_i(A(\epsilon))]) \leq M, \; i=1,\ldots,N \tag{5.13}$$

Since we have assumed that $A(\epsilon)$ is Hurwitz such a bound can be found. For example, if $A(\epsilon)$ is a polynomial matrix, we can take M equal to the highest-order power of ϵ appearing in $|A(\epsilon)|$.

Given N and M, let

$$K = NM + 1 \tag{5.14}$$

and consider carrying out K steps of the iteration described in Lemma 5.2. This produces

$$A_1^K(\epsilon) = \begin{bmatrix} A_{11}(\epsilon) & \epsilon^K R_1^K(\epsilon) \\ \epsilon^{K+1} S_1^{K+1} & \epsilon A_2^0(\epsilon) \end{bmatrix} \tag{5.15}$$

where we have introduced the notation $\tilde{A}_{11}(\epsilon) = A_{11}^K(\epsilon)$ and $A_2^0(\epsilon) = F_1^K(\epsilon)$. Next, we perform the same procedure at the next time scale. That is, write

$$A_2^0(\epsilon) = \begin{bmatrix} A_{22}^0(\epsilon) & R_2^0(\epsilon) \\ \epsilon S_2^0(\epsilon) & \epsilon F_2^0(\epsilon) \end{bmatrix} \tag{5.16}$$

and perform K steps of the interaction in Lemma 5.2 involving the sequence $\Gamma_2^i(\epsilon)$ and producing $\tilde{A}_{22}(\epsilon)$ and $A_3^0(\epsilon)$. Continuing this process we obtain a complete sequence $\tilde{A}_{11}(\epsilon),\ldots,\tilde{A}_{nn}(\epsilon)$ and can state the following

> <u>Theorem 5.2:</u> Suppose that $A(\epsilon)$ is Hurwitz and has MSSIES. Then $D(\epsilon)\bar{A}(\epsilon)$ has a weak time scale decomposition as in (5.2) with $A_i(\epsilon) = \tilde{A}_{ii}(\epsilon)$ and T as in (4.17) - (4.19).

<u>Proof:</u> A straightforward calculation shows that

$$\Sigma(\epsilon)D(\epsilon)\bar{A}(\epsilon)\Sigma^{-1}(\epsilon) = G(\epsilon) + H(\epsilon)$$

where

$$G(\epsilon) = \mathrm{diag}(\tilde{A}_{11}(\epsilon),\ \epsilon\tilde{A}_{22}(\epsilon),\ldots,\epsilon^{n-1}\tilde{A}_{nn}(\epsilon)) \tag{5.17}$$

$O(H(\epsilon)) = K$, and[8]

$$\Sigma(\epsilon) = \Sigma_n^K(\epsilon)\ldots\Sigma_n^0(\epsilon)\Sigma_{n-1}^K(\epsilon)\ldots\Sigma_2^0(\epsilon)\Sigma_1^K(\epsilon)\ldots\Sigma_0^0(\epsilon) \tag{5.18a}$$

where

$$\Sigma_0^k(\epsilon) = \Gamma_0^k(\epsilon) \tag{5.18b}$$

$$\Sigma_i^k(\epsilon) = \begin{bmatrix} I & 0 \\ 0 & \Gamma_i^k(\epsilon) \end{bmatrix} \tag{5.18c}$$

As in Lemma 5.1, we can replace $\Sigma(\epsilon)$ by $\Sigma(0)$. However $\Sigma(0) = T$, since $\Gamma_i^k(\epsilon) = I$ for $k > 0$ and

$$\Gamma_i^0(\epsilon) = \begin{bmatrix} I & A_{ii}^0(0)^{-1}R_i^0(0) \\ 0 & I \end{bmatrix} \tag{5.19}$$

with $A_{ii}^0(0)$ and $R_i^0(0)$ equal to $\tilde{A}_{ii}$ and R_i, respectively, from (4.14), (4.15).

What remains to be shown, then is that $G(\epsilon)$ and $G(\epsilon) + H(\epsilon)$ are asymptotically equivalent. This is done in Appendix E.

The key idea behind this result is that we must approximate the eigenstructure of $A(\epsilon)$ accurately up to at least the order of the damping in each eigenmode. For example, the matrix

$$\begin{bmatrix} -\epsilon-\epsilon^2 & 1 \\ -1 & -\epsilon-\epsilon^2 \end{bmatrix}$$

---

[8] Here $O(H(\epsilon))$ denotes the minimum order of all elements of $H(\epsilon)$. As an aside, note that the diagonal blocks of $H(\epsilon)$ are zero.

is asymptotically equivalent to the matrix in (5.1) -- i.e. it is allowable to neglect the higher order (ϵ^2) damping. On the other hand, the two matrices

$$\begin{bmatrix} -\epsilon^2 & 1 \\ -1 & -\epsilon^2 \end{bmatrix}, \quad \begin{bmatrix} -\epsilon^2 & 1-\epsilon \\ -1+\epsilon & -\epsilon^2 \end{bmatrix}$$

are <u>not</u> asymptotically equivalent since, compared to the order of damping, the difference in frequency (between 1 and $1+\epsilon$) is very significant.

What the procedure we have described does is to perform a sufficient number of iterations to guarantee that the difference between the eigenvalues of $A(\epsilon)$ and its approximant are of higher order than the real (i.e. the damping) part. Admittedly the procedure is conservative -- typically one can get by with fewer iterations and can discard additional higher-order terms retained by the procedure -- but it is guaranteed to work.

6. Assignment of Time Scales by State Feedback

The results of Sections 3 and 4 establish the role of the invariant factors of $A(\epsilon)$ in determining the time scales of the undriven system (1.1). For the driven system (1.11), it is then natural to pose the question of time scale or invariant factor assignment. Specifically it is of interest to determine what freedom there is in assigning the invariant factors of

$$\dot{x}(t) = F(\epsilon)x(t), \quad F(\epsilon) = A(\epsilon) + B(\epsilon)K(\epsilon) \tag{6.1}$$

by application of state feedback as in (1.12). The following is a result in this direction.

Theorem 6.1: Assume that $A(\epsilon)$, $B(\epsilon)$ are left coprime, i.e. that $[A(0)\ B(0)]$ has full row rank. Let b denote the rank of $B(0)$. Then

1. $F(\epsilon)$ can have no more than b non-unit invariant factors.
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 $\epsilon^{\infty} = 0$).

Proof: We first show that we can further assume that

$$A(\epsilon) = \text{diag}\,(1, \ldots, 1, \epsilon^{k_1}, \ldots, \epsilon^{k_L}) \quad , k_i > 0 \tag{6.2}$$

and that $B(\epsilon)$ is upper triangular. Specifically, suppose that $A(\epsilon)$ has the Smith form given in (3.2). We can then write

$$F(\epsilon) = P(\epsilon)[D(\epsilon) + P^{-1}(\epsilon)B(\epsilon)K(\epsilon)Q^{-1}(\epsilon)]Q(\epsilon) \tag{6.3}$$

Thus we can equivalently consider the invariant factors of $D(\epsilon) + \bar{B}(\epsilon)\bar{K}(\epsilon)$, where $\bar{B}(\epsilon) = P^{-1}(\epsilon)B(\epsilon)$, $\bar{K}(\epsilon) = K(\epsilon)Q^{-1}(\epsilon)$. Furthermore, using elementary column operations we can show that $\bar{B}(\epsilon)U(\epsilon) = \hat{B}(\epsilon)$ where $U(\epsilon)$ is unimodular and $\hat{B}(\epsilon)$ is upper triangular. Consequently we have the equivalent problem of invariant factor assignment for $D(\epsilon) + \hat{B}(\epsilon)\hat{K}(\epsilon)$, where $\hat{K}(\epsilon) = U^{-1}(\epsilon)\bar{K}(\epsilon)$.

Suppose then that $A(\epsilon)$ is given by (6.2) and $B(\epsilon)$ is upper triangular. Furthermore, for notational simplicity and without loss of generality we assume that both $A(\epsilon)$ and $B(\epsilon)$ are NxN. Let us first prove the second part of the theorem statement. Note first that for $[A(0), B(0)]$ to have full row rank it must be ture that $L \leq b$, and $\bar{B}$ must have the form[9]

$$B(\epsilon) = \left[\begin{array}{cccccc} * & * & \cdots & & & * \\ & * & & & & \vdots \\ & & * & & & \vdots \\ & & & 1 & & * \\ & & & & \ddots & * \\ 0 & & & & \ddots & * \\ & & & & & 1 \end{array}\right] \begin{array}{l} \\ \\ \\ \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} L \end{array} \tag{6.4}$$

where $*$ represents an aribtrary element in W.

Assume first that $L = b$. Than we can construct a unimodular matrix $V(\epsilon)$ so that

$$B(\epsilon)V(\epsilon) = \left[\begin{array}{ccc} * & * & * \\ & \ddots & \\ 0 & \ddots\, * & \\ \hline 0 & & I \end{array}\right] \begin{array}{l} \\ \\ \\ \}b \end{array} \tag{6.5}$$

and let

$$K(\epsilon) = V(\epsilon)\mathrm{diag}\ (0,\ldots,0,\epsilon^{j_1} - \epsilon^{k_1},\ldots,\epsilon^{j_b} - \epsilon^{k_b}) \tag{6.6}$$

---

[9] Actually what we can conclude is that the last L diagonal elements of B(0) are nonzero. By right-multiplications we can make these values unity.

It is straightforward then to show that $A(\epsilon) + B(\epsilon)K(\epsilon)$ has the desired invariant factors. If $L < b$, we are in essence replacing some of the unit invariant factors of $A(\epsilon)$ with nonunit invariant factors. Since rank $B(0) = b$, $b - L$ of the first $N - L$ columns of (6.4) are linearly independent at $\epsilon = 0$. Then, just as in constructing (6.5), we can construct a unimodular matrix $V(\epsilon)$ so that

$$B(\epsilon)V(\epsilon) = \left[\begin{array}{l} * \qquad * \ldots\ldots\ldots\ldots * \\ 0 \ldots 0 \quad 1 \quad 0 \ldots\ldots 0 \\ 0 \ldots 0 \; 1 \quad 0 \ldots\ldots\ldots 0 \\ * \quad * \ldots\ldots\ldots\ldots\ldots * \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \hdashline \qquad\quad 0 \qquad\qquad I \end{array}\right] \tag{6.7}$$

i.e. so that b-L of the first N-L rows are zero except for a single entry of unity, and so that these rows and the last L rows are linearly independent. In this case, it is then simply a matter of performing a permutation similarity transformation so that the transformed versions of $A(\epsilon)$ is as in (6.2) with some of the $k_i = 0$, while the transformed version of $B(\epsilon)V(\epsilon)$ is given by (6.5). From this point on the construction is the same as before.

To prove the first statement in the theorem, let $M = \text{rank}(A(0) + B(0)K(0))$ = number of unit invariant factors of $F(\epsilon)$. Also, assume that $V(\epsilon)$ has been constructed so that (6.5) holds (perhaps after the permutation similarity transformation described previously if $L < b$). Letting $\hat{K}(\epsilon) = V^{-1}(\epsilon)K(\epsilon)$, we see that

$A(0) + B(0)K(0) =$

$$N-L\left\{\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right.\!\!\left[\begin{array}{ccccc} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & 0 & \\ & & & & \ddots \\ 0 & & & & 0 \end{array}\right] + \left[\begin{array}{c:c} \begin{matrix} * & \ddots & * \\ 0 & & * \end{matrix} & * \\ \hdashline 0 & I \end{array}\right]\hat{K}(0) \qquad (6.8)$$

$$\hphantom{N-L\left[\begin{array}{ccccc}1&&&&0\end{array}\right]+\left[\begin{array}{c}* \ddots *\end{array}\right.}\underbrace{\hphantom{I}}_{L}$$

where the $*$ terms may be zero or nonzero; however, since $b \geq L$, there are b-L independnet column vectors in the first n-L columns of the matrix multiplying $\hat{K}(0)$ in (6.8). Consequently, adding $B(0)K(0)$ to $A(0)$ can reduce the rank of $A(0)$ by at most b-L. Thus

$$M \geq N - L - (b-L) = N-b \qquad (6.9)$$

Some results are also available for the case of non-coprime $A(\epsilon),B(\epsilon)$. In this case $F(\epsilon)$ is of the form

$$F(\epsilon) = W(\epsilon)\bar{F}(\epsilon) \qquad (6.10)$$

where

$$\bar{F}(\epsilon) = \bar{A}(\epsilon) + \bar{B}(\epsilon)K(\epsilon) \qquad (6.11)$$

Here $W(\epsilon)$ is a greatest common left division of $A(\epsilon),B(\epsilon)$, and $\bar{A}(\epsilon),\bar{B}(\epsilon)$ are left coprime. If the invariant factors of $F(\epsilon)$, $W(\epsilon)$, and $\bar{F}(\epsilon)$ are denoted by $f_i(\epsilon)$, $w_i(\epsilon)$, and $\bar{f}_i(\epsilon)$ and ordered such that the i^{th} one divides the $(i+1)^{th}$, we have (thanks to the Binet-Cauchy formula [14])

$$w_i(\epsilon)\,|\,f_i(\epsilon) \qquad \text{and} \qquad \bar{f}_i(\epsilon)\,|\,f_i(\epsilon) \qquad (6.12)$$

The first divisibility condition in (6.12) shows that every invariant factor of $F(\epsilon)$ must contain the corresponding invariant factor of $W(\epsilon)$. The $\bar{f}_i(\epsilon)$ are governed by Theorem 6.1, and conclusions about the $f_i(\epsilon)$ can then be drawn from the second divisibility condition in (6.12).

7. Conclusions

In this paper we have developed an algebraic approach to time scale analysis of singularly perturbed linear systems that exposes the role played by the Smith form of $A(\epsilon)$ viewed as a matrix over the ring of functions analytic at $\epsilon = 0$. This approach bridges the gap between previous easily interpreted but restricted results [1], [2] and more recent results [3] that are completely general but quite intricate. Our work not only provides a simple interpretation of the MSSNS condition introduced in [3] in terms of the invertibility of successive Schur complements of a particular matrix but also allows us to state and prove a strengthened and more precise version of the main result of [3] using the new concept of a strong multiple time scale decomposition.

The framework and concepts introduced in this paper also open the way for the investigation of additional questions. Several of these we have considered here as well. In particular, we have investigated the relaxing of the so-called MSST condition by developing a procedure involving iterated Schur complementation in order to guarantee that weak but essential damping terms are retained. In addition, we have investigated the problem of time-scale modification via state feedback, which in our context corresponds to changing the invariant factors of the system matrix. Another question that can be asked concerns the fact that the Smith decomposition is not unique. As shown in [8], while the use of different Smith decompositions leads to different time scale approximations, the successive Schur complements in these approximations are similar. Also, there is the problem of computing the Smith decomposition of $A(\epsilon)$. Some ideas related to this are given in [8], but these remain to be developed. In a sense we have traded the difficult tasks of computing ϵ-dependent projections and pseudo-inverses that are needed in the approach in [3] for the Smith form computation in our approach. However, in our work this computation is identified as a separate task which need not be carried through the remaining analysis and therefore does not obscure the intuition behind our results.

Finally, note that in [3] the orders of the various time scales of (1.1) are shown to correspond to the orders of the eigenvalues of $A(\epsilon)$. On the other hand, in this paper we have shown that the orders of the invariant factors determine the time scales. It should not come as too much of a surprise that there is a relationship between the orders of eigenvalues and invariant factors and that the MSSNS condition plays a central role in this relationship. This is the subject of a forthcoming paper.

Acknowledgements

We would like to thank P.G. Coxson, J.R. Rohlicek, and M. Vidyasagar for numerous valuable discussions on the subject of this paper and related topics.

References

1. P.V. Kokotovic, "Subsystems, time scales, and multimodeling," Automatica, Vol. 17, 1981, pp. 789-795.

2. J.H. Chow (Ed.), Time-Scale Modeling of Dynamic Networks, with Applications to Power Systems, Berlin: Springer-Verlag, 1982.

3. M. Coderch, A.S. Willsky, S.S. Sastry and D.A. Castanon, Hierarchical Aggregation of Linear Systems with Multiple Time Scales," IEEE Trans. Automatic Control, Vol. AC-28, No. 11, Nov. 1983, pp. 1017-1030.

4. P. Van Dooren, P. Dewilde and J. Vandewalle, "On the determination of the Smith-McMillan form of a rational matrix from its Laurent expansion," IEEE Trans. Circuits and Systems, March 1979, pp. 180-189.

5. G. Verghese and T. Kailath, "Rational matrix structure," IEEE Trans. Automatic Control, Vol. AC-26, No. 2, April 1981, pp. 434-439.

6. H.K. Khalil, "Time Scale Decomposition of Linear Implicit Singularly Perturbed Systems," IEEE Trans. Automatic Control, Vol. AC-29, No. 11, Nov. 1984, pp. 1054-1056.

7. T. Kato, Perturbation Theory for Linear Operators, Berlin: Springer Verlag, 1966.

8. X.-C. Lou, An Algebraic Approach to the Analysis and Control of Time Scales," Ph.D. Thesis, M.I.T. Dept. of Elec. Eng. and Comp. Sci., Sept. 1985.

9. X.-C. Lou, G.C. Verghese, A.S. Willsky, and M. Vidyasagar, "An Algebraic Approach to Analysis and Control of Time Scales," 1984 American Control Conference, San Deigo, CA, June 1984.

10. P.V. Kokotovic, "A Riccati Equation for Block Diagonalization of Ill-Conditioned Systems," IEEE Trans. Aut. Control. Vol. AC-20, 1975, pp. 812-814.

11. S.L. Campbell, "Singular Perturbations of Autonomous Linear Systems II," J. Diff. Equations. Vol. 29, 1978, pp. 362-373.

12. S.L. Campbell and N.J. Rose, "Singular Perturbations of Autonomous Linear Systems," SIAM J. Math. Anal. Vol. 10, No. 3, 1979, pp. 542-551.

13. P.V. Kokotovic, J.J. Allemong, J.R. Winkelman, and J.H. Chow, "Singular Perturbation and Iterative Separation of Time Scales," Automatica, Vol. 16, 1980, pp. 23-33.

14. M. Vidyasagar, Control System Systhesis: A Factorization Approach, The M.I.T. Press, Cambridge, MA, 1985.

Appendix A: Proof of Proposition 2.1

This proof of (2.17) uses several of the ideas introduced and developed in Sections 3 and 4. We first need the following

Lemma A.1: Let $H(\epsilon)$ be obtained form $A(\epsilon)$ by a similarity transformation

$$H(\epsilon) = S(\epsilon)A(\epsilon)S^{-1}(\epsilon) \qquad (A.1)$$

where $S(\epsilon)$ is unimodular. Then

$$F_{k0}^{H} = S(0)F_{k0}^{A}S^{-1}(0) \qquad (A.2)$$

This result follows easily from the fact that the sequence of eigenprojections and successive system matrices defined as in (2.12) - (2.16) for $A(\epsilon)$ and $H(\epsilon)$ are all related by the same similarity transformation. Equation (A.2) then follows on examination of the leading-order terms of the successive system matrices.

Consider next any $G(\epsilon) \in \mathcal{F}\{A(\epsilon)\}$, i.e.

$$G(\epsilon) = U(\epsilon)A(\epsilon)V(\epsilon) \qquad (A.3)$$

with

$$U(0) = V(0) = I \qquad (A.4)$$

Then by performing similarity transformations, it is straightforward to check that Proposition 2.1 will be proved if we can verify

Lemma A.2: Proposition 2.1 holds if

$$A(\epsilon) = D(\epsilon)\bar{A} \qquad (A.5)$$

$$G(\epsilon) = D(\epsilon)\bar{A}(\epsilon) \qquad (A.6)$$

with $\bar{A}(0) = \bar{A}$, which is invertible.

The proof of this lemma is a straightforward variation on the development in Section 4. As in Section 3, let us assume, without loss of generality that $k_1 = 0$ (since otherwise we can divide (A.5), (A.6) by ϵ^{k_1}). The result is then proved by induction on n, the number of time scales. For n=1 the result is immediate, since

$$A(\epsilon) = \mathrm{diag}(I,\ 0)\bar{A} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix} \tag{A.7}$$

$$G(\epsilon) = \mathrm{diag}(I,\ 0)\bar{A}(\epsilon) = \begin{bmatrix} A_{11}(\epsilon) & A_{12}(\epsilon) \\ 0 & 0 \end{bmatrix} \tag{A.8}$$

Clearly

$$F_{00}{}^{A} = F_{00}{}^{G} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix} \tag{A.9}$$

Furthermore, thanks to Lemma 4.1, $F_{00}{}^{A} = F_{00}{}^{G}$ has SSNS if and only if A_{11} is invertible. If A_{11} is singular, the procedure stops. If A_{11} is invertible, we have already achieved the normal rank of $A(\epsilon)$ (and $G(\epsilon)$) so that all subsequent F_{k0}'s are equal to 0. In either case the lemma is verified.

If $n > 1$, then

$$A(\epsilon) = \begin{bmatrix} A_{11} & A_{12} \\ \epsilon^{k_2} D_*(\epsilon) A_{21} & \epsilon^{k_2} D_*(\epsilon) A_{22} \end{bmatrix} \tag{A.10}$$

with

$$D_*(\epsilon) = \mathrm{diag}\ (I,\ \epsilon^{k_3-k_2} I, \ldots, \epsilon^{k_n-k_2} I,\ 0) \tag{A.11}$$

and $G(\epsilon)$ has a form analogous to (A.10) wiht $A_{ij}(\epsilon)$ replacing A_{ij}. Again (A.9) holds, and, as before, the procedure stops if A_{11} is singular. If A_{11} is invertible, we can use the same procedure as sketched at the end of Section 4 to block diagonalize $A(\epsilon)$ and $G(\epsilon)$. Specifically, consider equations (4.21) and (4.22), where, we replace $\tilde{A}_{11}$ by A_{11}, R_1 by A_{12}, $S_1(\epsilon)$ by $D_*(\epsilon)A_{21}$, and $F_1(\epsilon)$ by $D_*(\epsilon)A_{22}$. Again because of the invertibility of A_{11}, solutions $L_1^A(\epsilon)$ and $H_1^A(\epsilon)$ exist to these equations, with

$$L_1^A(0) = -A_{11}^{-1}A_{12} \qquad H_1^A(0) = D_*(0)A_{21}A_{11}^{-1} \tag{A.12}$$

Similarly, we can solve (4.21), (4.22) with analogous replacements but with $A_{ij}(\epsilon)$ substituted for A_{ij}. This yields solutions $L_1^G(\epsilon)$, $H_1^G(\epsilon)$. Applying the corresponding diagonalizing similarity transformations (4.24), (4.25) to $A(\epsilon)$ and $G(\epsilon)$, and noting that $T_1^A(0) = T_1^G(0)$, we see that, thanks to Lemma A.1, we have reduced the problem to one with one fewer time scale -- i.e. we are left to examine

$$G_2^A(\epsilon) = D_*(\epsilon)[A_{22} + A_{21}L_1^A(\epsilon)] \tag{A.13}$$

$$G_2^G(\epsilon) = D_*(\epsilon)[A_{22}(\epsilon) + A_{21}(\epsilon)L_1^G(\epsilon)] \tag{A.14}$$

From the invertibility of $\bar{A}$ and A_{11} we can immediately deduce the invertibility of $[A_{22} + A_{21}L_1^A(\epsilon)]$ and $[A_{22}(\epsilon) + A_{21}(\epsilon)L_1^G(\epsilon)]$ in a neighborhood of $\epsilon = 0$. Since these matrices are equal at $\epsilon = 0$, the result is proved by induction.

Appendix B: Proof of Proposition 2.2

Without loss of generality we assume that the similarity transformation T in (2.1) is the identity -- if this is not the case we can simply perform an initial ϵ-independent similarity transformation on $A_0(\epsilon)$. Furthermore, since $A_1,\ldots,A_n$ are assumed to be semistable, we can perform another ϵ-independent similarity transformation so that what we are given are Hurwitz matrices $G_1,\ldots,G_n$ so that

$$\limsup_{\substack{\epsilon\downarrow 0 \\ t\geq 0}} \| e^{A_0(\epsilon)t} - \exp\{\text{diag }(0, \epsilon^{k_1}G_1,\ldots,\epsilon^{k_n}G_n)t\}\| = 0 \tag{B.1}$$

and what we would like to show is that

$$F_{k0} = \begin{cases} \text{diag }(0,\ldots,0, G_i,0,\ldots,0) & k=k_i \\ 0 & \text{otherwise} \end{cases} \tag{B.2}$$

As a first step, note that (B.1) implies that for any integer r

$$\limsup_{\substack{\epsilon\downarrow 0 \\ t\in[0,\epsilon^{-r})}} \| e^{A_0(\epsilon)t} - \exp\{\text{diag }(0, \epsilon^{k_1}G_1,\ldots,\epsilon^{k_j}G_j,0,\ldots,0\}t\| = 0 \tag{B.3}$$

where

$$k_j \leq r < k_{j+1} \ , \quad j=0,\ldots,n \tag{B.4}$$

(here, for completeness $k_0 = 0$, $k_{n+1} = \infty$). Note also that, since $F_{00} = A_0(0)$

$$\limsup_{\substack{\epsilon\downarrow 0 \\ t\in[0,1)}} \| e^{A_0(\epsilon)t} - e^{F_{00}t}\| = 0 \tag{B.5}$$

From (B.3) - (B.5) we can conclude that if $k_1 > 0$, $F_{00} = 0$, $P_0(\epsilon) = I$, and $A_1(\epsilon) = A_0(\epsilon)/\epsilon$. Consequently, we can simply replace $A_0(\epsilon)$ in (A.1) with $A_1(\epsilon)$ and reduce each of the k_i by 1. Continuing in this fashion we find that $F_{k0} = 0$, $k < k_1$. From [3] we then have

$$\lim_{\epsilon \downarrow 0} \sup_{t \in [0, \epsilon^{-k_1})} || e^{A_0(\epsilon)t} - \exp\{F_{k_1 0} \epsilon^{k_1} t\} || = 0 \tag{B.6}$$

and from (B.3), (B.6) we conclude that

$$F_{k_1 0} = \mathrm{diag}(0, G_1, 0, \ldots, 0) \tag{B.7}$$

The remainder of the proof proceeds by induction on n. The case of n=1 is essentially complete, since in this case the sup on the left-hand side of (A.6) can be taken over $[0, 1/\epsilon^r)$ for any $r \geq k_1$. Consequently an argument identical to the one used in the preceding paragraph shows that $F_{k0} = 0$ for all $k > k_1$. To consider the case of $n > 1$, we assume, without loss of generality, that $k_1 = 0$ (since as we have seen, if $k_1 > 0$ then $A_0(\epsilon)$ is divisible by ϵ^{k_1} so we can rescale time to eliminate this factor). Next, write $A_0(\epsilon)$ as the sum of two commuting matrices

$$\begin{aligned} A_0(\epsilon) &= P_0(\epsilon)A_0(\epsilon) + [I-P_0(\epsilon)]A_0(\epsilon) \\ &= \epsilon A_1(\epsilon) + [I - P_0(\epsilon)]A_0(\epsilon) \end{aligned} \tag{B.8}$$

Note that, from [3] and (B.7)

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} || e^{[I-P_0(\epsilon)]A_0(\epsilon)t} - \exp\{\mathrm{diag}(0, G_1, 0, \ldots, 0)t\} || = 0 \tag{B.9}$$

Then, using (B.8) and performing several standard manipulations we obtain the following

$$||e^{A_0(\epsilon)t} - \exp\{\mathrm{diag}\ (0,G_1,\epsilon^{k_2}G_2,\dots,\epsilon^{k_n}G_n)t\}||$$

$$\leq ||e^{A_1(\epsilon)\epsilon t} - \exp\{\mathrm{diag}(0,0,\epsilon^{k_2}G_2,\dots,\epsilon^{k_n}G_n)t\}||.||e^{[I-P_0(\epsilon)]A_0(\epsilon)t}||$$

$$+ ||e^{[I-P_0(\epsilon)]A_0(\epsilon)t} - \exp\{\mathrm{diag}(0,G_1,0,\dots,0)t\}||.$$

$$||\exp\{\mathrm{diag}(0,0,\epsilon^{k_2}G_2,\dots,\epsilon^{k_n}G_n)t\}|| \qquad \text{(B.10)}$$

Note that since n > 1, (B.9) implies that

$$||e^{[I-P_0(\epsilon)]A_0(\epsilon)t}||$$

is bounded away from zero uniformly in t. Consequently (B.1), (B.9), (B.10), and the semistability of $G_2,\dots,G_n$ imply that

$$\lim_{\epsilon\downarrow 0}\ \sup_{t\geq 0}||e^{A_1(\epsilon)t} - \exp\{\mathrm{diag}\ (0,0,\epsilon^{k_2-1}G_2,\dots,\epsilon^{k_n-1}G_n)t\}|| = 0 \qquad \text{(B.11)}$$

and consequently (B.2) follows by induction.

Appendix C: Verification of Equation (4.13)

Let us rewrite (4.11) as

$$G(\epsilon) = \begin{bmatrix} G_{11}(\epsilon) & G_{12}(\epsilon) \\ G_{21}(\epsilon) & G_{22}(\epsilon) \end{bmatrix} \tag{C.1}$$

where

$$G_{11}(\epsilon) = \tilde{A}_{11} + \epsilon^{k_2}\tilde{A}_{11}^{-1}R_1D_2(\epsilon)W_1 \tag{C.2a}$$

$$G_{12}(\epsilon) = \epsilon^{k_2}\tilde{A}_{11}^{-1}R_1D_2(\epsilon)\bar{A}_2 \tag{C.2b}$$

$$G_{21}(\epsilon) = \epsilon^{k_2}D_2(\epsilon)W_1 \tag{C.2c}$$

$$G_{22}(\epsilon) = \epsilon^{k_2}D_2(\epsilon)\bar{A}_2 \tag{C.2d}$$

Note that $G_{11}(\epsilon)$ is invertible in a neighborhood of $\epsilon = 0$. Let

$$C_1(\epsilon) = \begin{bmatrix} I & 0 \\ -G_{21}(\epsilon)G_{11}^{-1}(\epsilon) & I \end{bmatrix} \tag{C.3}$$

$$E(\epsilon) = \begin{bmatrix} I & -G_{11}^{-1}(\epsilon)G_{12}(\epsilon) \\ 0 & I \end{bmatrix} \tag{C.4}$$

From (C.2) we see that $C_1(0) = E(0) = I$, and a straightforward calculation yields

$$H(\epsilon) = C_1(\epsilon)G(\epsilon)E(\epsilon) = \begin{bmatrix} G_{11}(\epsilon) & 0 \\ 0 & G_{22}(\epsilon)-G_{21}(\epsilon)G_{11}{}^{-1}(\epsilon)G_{12}(\epsilon) \end{bmatrix} \qquad (C.5)$$

Note that

$$G_{22}(\epsilon)-G_{21}(\epsilon)G_{11}{}^{-1}(\epsilon)G_{12}(\epsilon) = [I-\epsilon^{k_2}D_2(\epsilon)W_1G_{11}{}^{-1}(\epsilon)\tilde{A}_{11}{}^{-1}R_1]\ G_{22}(\epsilon) \qquad (C.6)$$

and the quantity in brackets on the right-hand side of (C.6) is obviously invertible in a neighborhood of $\epsilon = 0$. Let

$$C_2(\epsilon) = \begin{bmatrix} \tilde{A}_{11}G_{11}{}^{-1} & 0 \\ 0 & [I-\epsilon^{k_2}D_2(\epsilon)W_1G_{11}{}^{-1}(\epsilon)\tilde{A}_{11}{}^{-1}R_1]^{-1} \end{bmatrix} \qquad (C.7)$$

Again we can check that $U_2(0) = I$ and

$$C_2(\epsilon)H(\epsilon) = \begin{bmatrix} \tilde{A}_{11} & 0 \\ 0 & \epsilon^{k_2}D_2(\epsilon)\bar{A}_2 \end{bmatrix} \qquad (C.7)$$

so that (4.13) is verified with $U(\epsilon) = C_1(\epsilon)^{-1}C_2{}^{-1}(\epsilon)$ and $V(\epsilon) = E^{-1}(\epsilon)$.

Appendix D: Proof of Lemma 5.1

We have that

$$\lim_{\epsilon\downarrow 0}\sup_{t\geq 0} ||e^{D(\epsilon)\bar{A}(\epsilon)t} - T^{-1}\exp\{\mathrm{diag}[A_0(\epsilon),\epsilon A_1(\epsilon),\ldots,\epsilon^{n-1}A_{n-1}(\epsilon)]t\}T|| = 0 \tag{D.1}$$

Therefore

$$\lim_{\epsilon\downarrow 0}\sup_{t\geq 0}||e^{A(\epsilon)t}-P(\epsilon)T^{-1}\exp\{\mathrm{diag}[A_0(\epsilon),\epsilon A_1(\epsilon),\ldots,\epsilon^{n-1}A_{n-1}(\epsilon)]t\}TP^{-1}(\epsilon)|| = 0 \tag{D.2}$$

What we must show is

$$\lim_{\epsilon\downarrow 0}\sup_{t\geq 0} ||P(\epsilon)T^{-1}\exp\{\mathrm{diag}[A_0(\epsilon),\ldots,\epsilon^{n-1}A_{n-1}(\epsilon)]t\}TP^{-1}(\epsilon)$$
$$- P(0)T^{-1}\exp\{\mathrm{diag}[A_0(\epsilon),\ldots,\epsilon^{n-1}A_{n-1}(\epsilon)]t\}T^{-1}P(0))|| = 0 \tag{D.3}$$

A simple triangle inequality argument shows that the left-hand side of (D.3) is bounded above by

$$\lim_{\epsilon\downarrow 0}\sup_{t\geq 0}||(P(\epsilon)-P(0))T^{-1}\exp\{\mathrm{diag}[A_0(\epsilon),\ldots,\epsilon^{n-1}A_{n-1}(\epsilon)]t\}TP^{-1}(\epsilon)|| +$$
$$\lim_{\epsilon\downarrow 0}\sup_{t\geq 0} ||P(0)T^{-1}\exp\{\mathrm{diag}[A_0(\epsilon),\ldots,\epsilon^{n-1}A_{n-1}(\epsilon)]t\}T(P^{-1}(\epsilon)-P^{-1}(0))|| \tag{D.4}$$

The first term in (D.4) is in turn bounded above by

$$\lim_{\epsilon\downarrow 0}\{||P(\epsilon)-P(0)||\,||P^{-1}(\epsilon)||\bullet||T||\bullet||T^{-1}||$$
$$\bullet \sup_{t\geq 0}||\exp\{\mathrm{diag}[A_0(\epsilon),\ldots,\epsilon^{n-1}A_{n-1}(\epsilon)]t\}|| \tag{D.5}$$

From the construction in Section 5, we know that each $A_i(\epsilon)$ is Hurwitz for $\epsilon > 0$ and, since $A_i(0)$ has MSSIES, we know that $||\exp\{\mathrm{diag}[A_0(0),\ldots,\epsilon^{n-1}A_{n-1}(0)]t\}||$ is bounded. From this we conclude that the limit in (D.5) is zero. Obviously a similar argument works for the other term in (D.4), and the lemma is proved.

Appendix E: Completion of the Proof of Theorem 5.2

The result we need to prove is the following

Proposition E.1: Suppose that the NxN matrix $G(\epsilon)$ is Hurwitz. Suppose further that

$$O[\mathrm{Re}(\lambda_i(G(\epsilon))] \leq M \quad i=1,\ldots,N \tag{E.1}$$

and let $K = MN + 1$. Then $G(\epsilon)$ is asymptotically equivalent to $G(\epsilon) + H(\epsilon)$, where $H(\epsilon)$ is any matrix with $O(H(\epsilon)) = K$.

Proof: The proof is a variation on the methods in [3, 7, 8]. First from [7] we have the following

Lemma E.1: Let $A(\epsilon) = B(\epsilon) + \epsilon^p C(\epsilon)$ be an NxN matrix. Then

$$\min_i O[\lambda_i(A(\epsilon)) - \lambda_i(B(\epsilon))] \geq p/N \tag{E.2}$$

Consequently in our case

$$O[\lambda_i(G(\epsilon)) - \lambda_i(G(\epsilon) + H(\epsilon))] > O[\mathrm{Re}(\lambda_i(G(\epsilon)))] \tag{E.3}$$

Next, recall the definition of the resolvent of a matrix $A(\epsilon)$

$$R(\lambda,A) = [A(\epsilon) - \lambda I]^{-1} \tag{E.4}$$

so that

$$e^{A(\epsilon)t} = -\frac{1}{2\pi i}\sum_k \oint_{\Gamma_k} e^{\lambda t}R(\lambda,A)d\lambda \tag{E.5}$$

where the Γ_k are positively-oriented contours enclosing disjoint portions of the complex plane and all of the eigenvalues of $A(\epsilon)$. Consider, then

$$2\pi i[e^{[G(\epsilon)+H(\epsilon)]t} - e^{G(\epsilon)t}] = \sum_k \oint_{\Gamma_k} e^{\lambda t}[R(\lambda,G) - R(\lambda,G+H)]d\lambda \tag{E.6}$$

where we choose the Γ_k carefully. Specifically Γ_1 is a circle centered at

$\lambda_1(G(\epsilon))$, of radius of order $O[Re(\lambda_1(G(\epsilon)))]$, and completely contained in the left-half plane $\{Re(\lambda) < 0\}$. More precisely, we require the maximum value of $Re(\lambda)$ on Γ_k to also be of order $O[Re\ (\lambda_1(G(\epsilon)))]$.[10] Also, for ϵ small enough (E.3) guarantees that this circle includes $\lambda_1(G(\epsilon) + H(\epsilon))$. The circle may also include other pairs of eigenvalues, but for ϵ sufficiently small this happens only if

$$O[\lambda_1(G(\epsilon))-\lambda_j(G(\epsilon))] < \min \{O[Re(\lambda_1(G(\epsilon)))], O[Re(\lambda_j(G(\epsilon)))]\} \qquad (E.7)$$

Consider next a single term in (E.6) and suppose that the radius of Γ_k is of order m. If we let $\lambda' = \lambda/\epsilon^m$ we can rewrite this term as

$$\oint_{\Gamma_k'} e^{\epsilon^m\lambda' t}[R(\epsilon^m\lambda',G) - R(\epsilon^m\lambda', G+H)]\epsilon^m d\lambda' \qquad (E.8)$$

where Γ_k', the image of Γ_k under this mapping, has radius of order 1, is completely contained in the left-half plane, and in fact consists of points with negative real parts of order 1. Consequently, the norm of (E.8) is bounded above by

$$\oint_{\Gamma_k'} ||R(\epsilon^m\lambda',G) - R(\epsilon^m\lambda',G+H)||\epsilon^m d\lambda' \qquad (E.9)$$

Also, we can write

$$R(\lambda,G) - R(\lambda,G+H) = R(\lambda,G)\{I - [I+HR(\lambda,G)]^{-1}\} \qquad (E.10)$$

Note that, thanks to (E.3) and (E.7), $R(\lambda,G)$ is of order $1/\epsilon^m$ on Γ^k. Consequently (since $m \leq M$) $HR(\lambda,G)$ is of order at least $m(N-1) + 1$, and we can write the series

---

[10] For example, the circle $\{\lambda: |\lambda-\lambda_1(G(\epsilon))| = \frac{1}{2} Re(\lambda_1(G(\epsilon)))\}$ will due unless another singularity lies on it.

$$R(\lambda,G) - R(\lambda,G+H) = R(\lambda,G) \sum_{n=1}^{\infty} (-1)^n [HR(\lambda,G)]^n \tag{E.11}$$

which converges uniformly for $\lambda \epsilon \Gamma_k$. Obviously the same statements can be made for $R(\epsilon^m\lambda', G)$ and $HR(\epsilon^m\lambda', G)$ on Γ_k'. and therefore we conclude that

$$O(||R(\epsilon^m\lambda',G) - R(\epsilon^m\lambda',G+H)||\epsilon^m) \geq m(N-1) + 1 \tag{E.12}$$

uniformly on Γ_k'. Since Γ_k' has perimeter of order 1 in length, (E.9) converges to 0 as $\epsilon\downarrow 0$, and the result follows.