

HIERARCHICAL AGGREGATION OF LINEAR
SYSTEMS WITH MULTIPLE TIME SCALES

by

M. Coderch, A.S. Willsky,

S.S. Sastry and D.A. Castanon*

LABORATORY FOR INFORMATION AND DECISIONS SYSTEMS

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

CAMBRIDGE MA 02139 USA

Abstract

In this paper we carry out a detailed analysis of the multiple time scale behavior of singularly perturbed linear systems of the form;

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

where $A(\epsilon)$ is analytic in the small parameter ϵ . Our basic result is a uniform asymptotic approximation to $\exp A(\epsilon)t$ that we obtain under a certain multiple semistability condition. This asymptotic approximation gives a complete multiple time scale decomposition of the above system and specifies a set of reduced order models valid at each time scale.

Our contribution is threefold:

- 1) We do not require that the state variables be chosen so as to display the time scale structure of the system.
- 2) Our formulation can handle systems with multiple (>2) time scales and we obtain uniform asymptotic expansions for their behavior on $[0, \infty[$.
- 3) We give an aggregation method to produce increasingly simplified models valid at progressively slower time scales.

*Research supported in part by the Department of Energy under grant ET-76-C-01-2295 and by the Air Force Office of Scientific Research under Grant AFOSR-82-0258. The first author also acknowledges thankfully the continuing support of the fundacion ITP Madrid-Spain.

Section 1. Introduction

Notions of time-scale separation are commonly used in heuristic model reduction techniques. It is well known that these notions can be formalized using techniques of singular perturbation theory, e.g. [1]. In this paper we carry out a detailed analysis of the multiple time scale behavior of singularly perturbed (defined in section 3) linear systems of the form:

$$\dot{x}^\epsilon(t) = A(\epsilon) x^\epsilon(t), \quad x^\epsilon(0) = x_0. \quad (1.1)$$

where $A(\epsilon)$ is analytic in the small parameter $\epsilon \in]0, \epsilon_0]$. Our analysis gives a complete picture of the relationship between weak couplings, singular perturbations, multiple time scale behavior and reduced order modelling[†] for these systems. Specifically, we give necessary and sufficient conditions under which (1.1) exhibits well defined, non-trivial behavior at several fundamental time scales. We determine these time scales and we associate a reduced order model of (1.1) with each of its fundamental time scales. We then show that these reduced order models can be combined to produce an asymptotic approximation to $x^\epsilon(t)$ uniformly valid on $[0, \infty[$.

In previous work it has generally been assumed that the system under consideration has 'fast' and 'slow' dynamics, and that by a combination of experience and physical insight a choice of state variables is available which displays the two time scale structure of the system. Thus, typically, the starting point for research has been a system of the form:

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

[†] In this paper we use the terms "reduced-order models" and "aggregated models" interchangeably. In many references (such as in the economics literature) in which the latter expression is used what is typically meant by it is a special type of reduced-order model resulting from a procedure which explicitly combines (e.g. adds) groups of variables of the original system. In [4] the results we develop here are taken as a starting point for constructing such an explicit aggregation procedure for singularly perturbed finite state Markov processes.

While this system is not explicitly of the form (1.1), it can be converted into one of the form of (1.1) by rescaling time, $\tau = t/\epsilon$. The rescaling is inconsequential in our development since we recover all the time scales associated with (1.1). We use the form (1.1) throughout our development with the understanding that time has been scaled so that the fastest time-scale associated with the system is the order 1 time scale.

Further, in almost all the available literature the behavior of the system is studied as $\epsilon \downarrow 0$ on intervals of the form $[0, T/\epsilon]$. The existence of non-trivial behavior for times of order $1/\epsilon^k$ or, more generally, on the infinite time interval $[0, \infty[$ is either excluded by assumptions imposed on the matrices A_{ij} , or not considered at all. An example of the former is [2] where in the context of (1.2) it is proved that if A_{22} and $A_{11} - A_{12} A_{22}^{-1} A_{21}$ are stable then (1.2) exhibits only two time scales. Two time scale systems are the only ones studied so far in the context of control and estimation problems (see [3] for a bibliography).

Our main contribution we feel is threefold:

- 1) We relax the requirement that state variables be chosen so as to display the time scale structure of the system.
- 2) Our formulation can handle systems with multiple (> 2) time scales and we obtain uniform asymptotic expansions for their behavior on $[0, \infty[$.
- 3) We give a method of aggregation to produce increasingly simplified models valid at progressively slower time scales. (We have applied this method to hierarchically aggregate finite state Markov processes with rare events. A brief description of this application of our methods is given in section 5, the details appear elsewhere [4].)

Systems with more than two time scales have been studied by other authors in different settings. In [5] the authors considered the asymptotic behavior of the quasi-linear system:

$$\epsilon^2 \dot{x}^\epsilon(t) = A(t) x^\epsilon(t) + \epsilon f(x^\epsilon(t), t, \epsilon) \quad (1.3)$$

on the time interval $[0, T]$ and found that the asymptotic expansion of $x^\epsilon(t)$ requires three series: in t , t/ϵ and t/ϵ^2 respectively. They did not, however,

study the behavior of (1.3) on $[0, \infty]$ and so left open the possibility of additional time scales. More recently, and with a formulation similar to ours, Campbell and Rose ([6] - [8], [23]) have studied the asymptotic behavior of:

$$\dot{x}^\varepsilon(t) = \left(\sum_{n=0}^N \varepsilon^n A_n \right) x^\varepsilon(t) \quad (1.4)$$

For the case $N=1$, they showed that a necessary and sufficient condition for

$$\lim_{\varepsilon \downarrow 0} x^\varepsilon(t/\varepsilon)$$

to exist pointwise (i.e., for fixed t - not uniformly on t) is the semistability of A_0 . For the more general case $N > 1$, they give necessary and sufficient conditions so that

$$\lim_{\varepsilon \downarrow 0} x^\varepsilon(t/\varepsilon^N)$$

exists pointwise and also an expression for the limit. They do not address, however, the question of uniform asymptotic approximations to $x^\varepsilon(t)$ or, equivalently, the question of how to determine the number and the time scales exhibited by (1.4) and how to combine the different pointwise limits to construct a uniform approximation (if possible). Furthermore, it does not seem to be widely appreciated that the system (1.3) may have non-trivial behavior at time scales $t/\varepsilon^2, t/\varepsilon^3, \dots$. In the context of Markov processes with rare events, several authors [9] - [12] have used aggregated models to describe the evolution of these processes. As in the work mentioned before, however, the connection between a hierarchy of increasingly consolidated models and uniform approximation is absent. In this paper we address the foregoing questions within a framework that unifies the partial results cited above. For a more detailed

account the reader is referred to [4] and [20].

Finally in a setting similar to ours, Hoppensteadt [21] studies uniform asymptotic approximations for the dynamics of a system of the form of (1.1). However, he assumes that $A(\epsilon)$ has been decomposed in a form which explicitly displays the time scale structure. Specifically, he assumes that $A(\epsilon)$ is given in the form

$$A(\epsilon) = \bigoplus_{i=1}^M \epsilon^{r_i} A_i(\epsilon) + \tilde{A}(\epsilon) \quad (1.5)$$

and then shows that the dynamics of (1.1) can be uniformly approximated under certain stability conditions by the dynamics given by the $A_i(\epsilon)$ at time scales of order t/ϵ^{r_i} , $i=1, \dots, M$. As we show in this paper the transition from (1.1) to (1.4) is neither obvious nor always possible. In fact from this perspective a major contribution of this paper is in providing an explicit algorithm for determining if a general $A(\epsilon)$ can be put in this form and if a uniform asymptotic approximation exists. This algorithm is constructive and thus if the answers to the questions it answers are in the affirmative, the algorithm will produce the uniform asymptotic approximation and in so doing will in effect produce a transformation which explicitly displays the time scale structure as in (1.4).

The outline of the paper is as follows: In Section 2 we present the basic mathematical machinery for our approach: perturbation theory for linear operators. The fundamental results on perturbation of the resolvent, the eigenvalues and the eigenprojections are stated without proof and are due to Kato [13]. In Section 3 we define regular and singular perturbations, and indicate the difficulties associated with uniform asymptotic approximations. In Section 4 we apply the theory of Section 2 to obtain, under a certain multiple semistability condition, a uniform asymptotic approximation to $\exp\{A(\epsilon)t\}$ that gives a complete multiple time scale decomposition of the system (1.1), and specifies a set of reduced order models valid at each time scale. We then show that our results are tight in that, when the multiple semistability condition is not satisfied the system does not have well defined behavior at

some time scale. A partial time scale decomposition is sometimes possible in this instance and it is carried out in Section 4.5. In Section 5 we summarize our results and explain briefly how they may be applied to the hierarchical aggregation of finite state Markov processes with rare transitions.

Section 2. Mathematical Preliminaries-Perturbation Theory for Linear Operators

We survey here the notation and some results on the perturbation of the eigenvalues, resolvent and eigenprojections of a linear operator $T: \mathbb{C}^n \rightarrow \mathbb{C}^n$ (for details see [13], [22]). These are the major mathematical tools for our development.

2.1. The resolvent

The set of all eigenvalues of T , denoted $\sigma(T)$ is called the spectrum of T . The Function $R(\xi, T): \mathbb{C} - \sigma(T) \rightarrow \mathbb{C}^{n \times n}$ defined by

$$R(\xi, T) := (T - \xi I)^{-1} \tag{2.1}$$

is called the resolvent of T . The resolvent of T is an analytic function with singularities at $\lambda_k \in \sigma(T)$, $k = 0, 1, \dots, s$. The Laurent series of $R(\xi, T)$ at λ_k has the form:

$$R(\xi, T) = - (\xi - \lambda_k)^{-1} P_k - \sum_{i=1}^{m_k-1} (\xi - \lambda_k)^{-i-1} D_k^i + \sum_{i=0}^{\infty} (\xi - \lambda_k)^i S_k^{i+1} \tag{2.2}$$

where

$$P_k := \frac{-1}{2\pi i} \int_{\Gamma_k} R(\xi, T) d\xi \in \mathbb{C}^{n \times n} \tag{2.3}$$

(with Γ_k a positively oriented contour enclosing λ_k but no other eigenvalue of T) is a projection (ie. $P_k^2 = P_k$) called the eigenprojection of the eigenvalue λ_k ;

$$m_k := \dim \mathcal{R}(P_k) \tag{2.4}$$

is the algebraic multiplicity of λ_k .

$$D_k := -\frac{1}{2\pi i} \int_{\Gamma_k} (\xi - \lambda_k) R(\xi, T) d\xi \quad (2.5)$$

is the eigen-nilpotent (ie. $D_k^{m_k} = 0$) for the eigenvalue λ_k ; and

$$S_k = \frac{1}{2\pi i} \int_{\Gamma_k} (\xi - \lambda_k)^{-1} R(\xi, T) d\xi \quad (2.6)$$

The following relations between P_k, S_k, D_k hold

$$P_k S_k = S_k P_k = 0 \quad (2.7)$$

$$P_k D_k = D_k P_k = D_k \quad (2.8)$$

$$P_k T = T P_k \quad (2.9)$$

$$(T - \lambda_k I) S_k = I - P_k \quad (2.10)$$

$$(T - \lambda_k I) P_k = D_k \quad (2.11)$$

$$P_k P_\ell = \delta_{k\ell} P_k \quad (2.12)$$

$$\sum_{k=1}^s P_k = I \quad (2.13)$$

From (2.12) and (2.13) it follows that

$$C^n = \mathcal{R}(P_1) \oplus \dots \oplus \mathcal{R}(P_s)$$

The $\mathcal{R}(P_k)$ is the algebraic eigenspace (or generalized eigenspace) for the eigenvalue λ_k . From (2.8) and (2.11) it follows that

$$T P_k = P_k T = P_k T P_k = \lambda_k P_k + D_k$$

This together with (2.13) yields the spectral representation of T :

$$T = \sum_{k=0}^s (\lambda_k P_k + D_k) = -\frac{1}{2\pi i} \int_{\Gamma} \xi R(\xi, T) d\xi$$

An eigenvalue λ_k is said to be semisimple if the associated eigennilpotent D_k is zero and simple if in addition $m_k = 1$.

Using the resolvent $R(\xi, T)$ and a contour enclosing all the eigenvalues of T in its interior we may define

$$\exp \{Tt\} = -\frac{1}{2\pi i} \int_{\Gamma} \exp(\xi t) R(\xi, T) d\xi$$

2.2 Semisimple and Semistable operators

An operator T is said to have semisimple null structure (SSNS) if zero is a semisimple eigenvalue of T . The following lemma establishes some properties of operators with SSNS.

Lemma 2.1 The following statements are equivalent:

- (i) T has SSNS
- (ii) $C^n = R(T) + N(T)$
- (iii) $R(T) = R(T^2)$
- (iv) $\text{rank } T = \text{rank } T^2$
- (v) $N(T) = N(T^2)$

Proof: See [14]

Comment: When T has SSNS, P_0 the eigenprojection for the zero eigenvalue is the projection onto $N(T)$ along $R(T)$. Further, it follows [13] that if T has SSNS, $T + P_0$ is non-singular. Now, if $T^\#$ is defined to be $(T + P_0)^{-1} - P_0$, then it may be verified that

$$TT^\#T = T, T^\#TT^\# = T^\# \quad (2.14)$$

and

$$T^\#T = TT^\# \quad (2.15)$$

$T^\#$ is thus the group generalized inverse of T (see [14]). Further, if T has SSNS, then P_0 and $T^\#$ determine the Laurent expansion of $R(\lambda, T)$ at zero.

Lemma 2.2 If T has SSNS, then for $\{\lambda : |\lambda| \leq |T^\#|^{-1}\}$

$$R(\lambda, T) = -\frac{P_0}{\lambda} + \sum_{k=0}^{\infty} \lambda^k (T^\#)^{k+1} \quad (2.16)$$

Proof: Using (2.14) and (2.15)

$$\begin{aligned} (T - \lambda I) \left(-\frac{P_0}{\lambda} + \sum_{k=0}^{\infty} \lambda^k (T^\#)^{k+1} \right) \\ = (I - P_0) \sum_{k=0}^{\infty} \lambda^k (T^\#)^k + P_0 - \sum_{k=1}^{\infty} \lambda^k (T^\#)^k \\ = I \end{aligned}$$

Similarly

$$\left(-\frac{P_0}{\lambda} + \sum_{k=0}^{\infty} \lambda^k (T^\#)^{k+1}\right) (T-\lambda I) = I \quad \square$$

Also of interest in the sequel are semistable operators: T is said to be semistable if T has SSNS and all the eigenvalues of T except the zero eigenvalue lie in $\overset{\circ}{\mathbb{C}}_-$ (the open left half plane).

2.2 Perturbation of Eigenvalues

Before we discuss perturbation of the resolvent of an operator T , we discuss perturbation of its eigenvalues, when T is of the form:

$$T(\varepsilon) = T + \sum_{n=1}^{\infty} \varepsilon^n T^{(n)} \quad \varepsilon \in [0, \varepsilon_0] \quad (2.17)$$

Here (2.17) is assumed to be an absolutely convergent power series expansion.

The eigenvalues of $T(\varepsilon)$ satisfy

$$\det(T(\varepsilon) - \xi I) = 0 \quad (2.18)$$

This is an algebraic equation in ξ whose coefficients are ε -analytic.

From elementary analytic function theory e.g. [15]) the roots of (2.18) are branches of analytic functions of ε with only algebraic singularities.

Hence, the number of (distinct) eigenvalues of $T(\varepsilon)$ is a constant s , independent of ε , except at some isolated values of ε . Without loss of generality let $\varepsilon = 0$ be such an exceptional point and further let it be the only such point in $[0, \varepsilon_0]$. In a neighborhood of the exceptional point, the eigenvalues of $T(\varepsilon)$ can be expressed by s distinct, analytic functions.

$\lambda_1(\varepsilon), \dots, \lambda_s(\varepsilon)$. These may be grouped as

$$\{\lambda_1(\varepsilon), \dots, \lambda_p(\varepsilon)\}, \{\lambda_{p+1}(\varepsilon), \dots, \lambda_{p+l}(\varepsilon)\}, \dots \quad (2.19)$$

so that each group has a Puiseux series of the form (written below for the first group)

$$\lambda_h(\varepsilon) = \lambda + \alpha_1 \omega^h \varepsilon^{1/p} + \alpha_2 \omega^{2h} \varepsilon^{2/p} + \dots$$

$h=0,1,\dots,p-1$

where λ is an eigenvalue of the unperturbed operator T and $\omega = \exp\{i2\pi/p\}$. Each group is called a cycle and the number of elements its period. λ is called the center of the cycle and the group of eigenvalues having λ as center is called the λ -group splitting at $\varepsilon = 0$ (the exceptional point).

2.3 Perturbation of the Resolvent

The resolvent of $T(\varepsilon)$ is defined on $\rho(T) = \mathbb{C} - \sigma(T(\varepsilon))$

$$R(\xi, T(\varepsilon)) = (T(\varepsilon) - \xi I)^{-1}$$

Lemma 2.3

If $\xi \in \rho(T)$, then for ε small enough, say $\varepsilon \in [0, \varepsilon_0]$, $\xi \in \rho(T(\varepsilon))$ and

$$R(\xi, T(\varepsilon)) = R(\xi, T) + \sum_{n=1}^{\infty} \varepsilon^n R^{(n)}(\xi) \tag{2.20}$$

where

$$R^{(n)}(\xi) = \sum_{\substack{v_1 + \dots + v_p = n \\ v_i \geq 1}} (-1)^p R(\xi, T) T^{(v_1)} R(\xi, T) T^{(v_2)} \dots T^{(v_p)} R(\xi, T) \tag{2.21}$$

the sum being taken over all integers p and $v_1, \dots, v_p \geq 1$ satisfying $v_1 + \dots + v_p = n$.

The series (2.20) is uniformly convergent on compact subsets of $\rho(T)$.

Proof: See [13]

□

2.4 Perturbation of the eigenprojections

We require first a preliminary lemma:

Lemma 2.4 (Taken verbatim from [13], page 34).

Let $P(t)$ be a projection matrix depending continuously on a parameter t varying in a connected subset of \mathbb{C} . Then the ranges $\mathcal{R}(P(t))$ for different t are isomorphic - i.e. the dimension of $\mathcal{R}(P(t))$ is constant. \square

Let λ be an eigenvalue of $T = T(0)$ with (algebraic) multiplicity m . Let Γ be a closed contour (positively oriented) in $\rho(T)$ enclosing λ but no other eigenvalues of T . From Lemma 2.3 it follows that for ε small enough $\mathcal{R}(\xi, T(\varepsilon))$ exists for $\xi \in \Gamma$ and hence there are no eigenvalues of $T(\varepsilon)$ on Γ . Further the matrix

$$P(\varepsilon) = - \frac{1}{2\pi i} \int_{\Gamma} \mathcal{R}(\xi, T(\varepsilon)) d\xi \quad (2.22)$$

is a projection which is equal to the sum of the eigenprojections for all the eigenvalues of $T(\varepsilon)$ lying inside Γ . Using (2.21) and integrating term by term (recall uniform convergence from Lemma 2.3) we have

$$P(\varepsilon) = P + \sum_{n=1}^{\infty} \varepsilon^n P^{(n)} \quad \varepsilon \in [0, \varepsilon_0] \quad (2.23)$$

where

$$P = - \frac{1}{2\pi i} \int_{\Gamma} \mathcal{R}(\xi, T) d\xi \quad (2.24)$$

and

$$P^{(n)} = - \frac{1}{2\pi i} \int_{\Gamma} \mathcal{R}^{(n)}(\xi) d\xi \quad (2.25)$$

Note that P is the eigenprojection for the eigenvalue λ . Further, note that $P(\varepsilon)$ is continuous in $\varepsilon \in [0, \varepsilon_0]$. By Lemma 2.4,

$$\dim \mathcal{R}(P(\varepsilon)) = \dim \mathcal{R}(P) = m \quad (\text{say}) \quad (2.26)$$

From (2.26) it follows that the eigenvalues of $T(\varepsilon)$ lying inside Γ form

the Γ group. Hence, $P(\epsilon)$ is called the total projection and $R(P(\epsilon))$ the total eigenspace for the λ -group. The following is a central proposition:

Proposition 2.5

Let λ be an eigenvalue of $T = T(\theta)$ of (algebraic) multiplicity m and $P(\epsilon)$ be the total projection for the λ -group of T . Then,

$$\begin{aligned} \frac{(T(\epsilon) - \lambda I) P(\epsilon)}{\epsilon} &= - \frac{1}{\epsilon 2\pi i} \int_{\Gamma} (\xi - \lambda) R(\xi, T(\epsilon)) d\xi \\ &= \frac{D}{\epsilon} + \sum_{n=0}^{\infty} \epsilon^n \tilde{T}^{(n)} \quad \text{for } \epsilon \in [0, \epsilon_0] \end{aligned} \tag{2.26}$$

where Γ is a closed positive contour enclosing λ and no other eigenvalues of T , D is the eigennilpotent for λ and $\tilde{T}^{(n)}$ is given by

$$\begin{aligned} \tilde{T}^{(n)} = - \sum_{p=1}^{n+1} & \sum_{\substack{v_1 + \dots + v_p = n+1 \\ k_1 + \dots + k_{p+1} = p-1 \\ v_i \geq 1, k_j \geq -m+1}} S^{(k_1)} T^{(v_1)} S^{(k_2)} \dots S^{(k_p)} T^{(v_p)} S^{(k_{p+1})} \end{aligned} \tag{2.27}$$

with $S^{(0)} = -P$, $S^{(k)} = -D^k$ for $k > 0$ and

$$S^{(k)} = \left[\frac{1}{2\pi i} \int_{\Gamma} (\xi - \lambda)^{-1} R(\xi, T) d\xi \right]^k \quad \text{for } k > 0. \quad \square$$

Although this result is in [13], the proof given there assumes $\lambda \notin \rho(T(\epsilon))$, a condition violated in some of our applications. A modification of this proof which does not require the condition is given in [20].

Of major interest in later sections is the following special case of Proposition (2.5).

Corrolary 2.6 Let $\lambda = 0$ be a semisimple eigenvalue of T (SSNS):

(2.26) then simplifies to

$$\frac{T(\epsilon) P(\epsilon)}{\epsilon} = \sum_{n=0}^{\infty} \epsilon^n \tilde{T}^{(n)} \tag{2.34}$$

with

$$\tilde{T}^{(n)} = - \sum_{p=1}^{n+1} (-1)^p \sum_{\substack{v_1 + \dots + v_p = n+1 \\ k_1 + \dots + k_{p+1} = p-1 \\ v_i \geq 1, k_j \geq 0}} S^{(k_1)} T^{(v_1)} \dots T^{(v_p)} S^{(k_{p+1})} \quad (2.35)$$

with $S^{(0)} = -P$ and $S^{(k)} = (T^\#)^k$, $k > 0$.

Proof: The corollary is a straight forward application of Proposition 2.6 with λ set to zero and $D = 0$ (by SSNS). Lemma 2.2 is used to obtain the expression for \tilde{T}^n in terms of $T^\#$. □

Section 3. Regular and Singular Perturbations

We consider linear time invariant systems of the form

$$\dot{x}^\epsilon(t) = A(\epsilon) x^\epsilon(t), \quad x^\epsilon(0) = x_0 \quad (3.1)$$

with $x^\epsilon(t) \in \mathbb{R}^n$ and $\epsilon \in [0, \epsilon_0]$. The matrix $A(\epsilon)$ is assumed to be semistable for each $\epsilon \in [0, \epsilon_0]$ and is assumed to have a convergent power series expansion in ϵ , i.e.,

$$A(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p A_p \quad (3.2)$$

The positive number $\epsilon_0 > 0$ is taken small enough so that $A(\epsilon)$ has constant rank d for $\epsilon \in]0, \epsilon_0[*$. We will refer to d as the normal rank of $A(\epsilon)$ and we will denote it by nrank.

Our objective is to analyze the behavior of $x^\epsilon(t)$ as $\epsilon \downarrow 0$ for $t \in [0, \infty[$. First, it is straightforward to verify that on any time interval of the form $[0, T]$, the system (3.1) can be approximated by

$$\dot{x}^0(t) = A_0 x^0(t), \quad x^0(0) = x_0 \quad (3.3)$$

*The results of our work go through mutatis mutandis when (3.2) is an asymptotic series, provided this rank condition is satisfied.

Precisely,

$$\lim_{\epsilon \downarrow 0} \sup_{t \in [0, T]} \left| \exp\{A(\epsilon)t\} - \exp\{A_0 t\} \right| = o \quad \forall T < \infty$$

However, as simple counterexamples will show, (3.3) is not, in general, a good approximation of (3.1) on the infinite time interval, i.e., (3.4) below is not true in general.

$$\lim_{\epsilon \downarrow 0} \sup_{t > 0} \left| \exp\{A(\epsilon)t\} - \exp\{A_0 t\} \right| = 0 \quad (3.4)$$

If, on the other hand, $A(\epsilon)$ is such that (3.4) is satisfied, we say that (3.1) is a regularly perturbed version of (3.3). Otherwise we call (3.1) a singularly perturbed system. In the literature, systems of the form (3.1) are said to be singularly perturbed if $A(\epsilon)$ has a Laurent series about $\epsilon = 0$,

$$A(\epsilon) = \sum_{p=-r}^{\infty} \epsilon^p A_p \quad (3.5)$$

with $r > 0$, and regularly perturbed if $r = 0$. We find this characterization deficient on two counts:

- i) using this definition, a system is regularly or singularly perturbed depending on the time scale used to write its dynamics; and
- ii) the Laurent series formulation singles out from the very start a certain time scale of interest neglecting the system's evolution at slower and faster time scales.

With a simple normalization of the time variable, a system of the form (3.5) can be rewritten as having a system matrix with a convergent power series as in (3.1). By studying the evolution of the system on the infinite time interval $[0, \infty[$ as in (3.4), we can characterize the perturbation as regular or singular in a more fundamental way which will depend now on the structure of the system matrix. Further, such a study will give equal importance to all time scales present in the system.

In what follows we focus on the singularly perturbed case since failure of (3.4) is symptomatic of distinct behavior at several time scales. Formally,

Definition 3.1 (Time Scale Behavior)

Consider (3.1) and let $\alpha(\epsilon)$ be an order function ($\alpha: [0, \epsilon_0] \rightarrow \mathbb{R}_+$; $\alpha(0) = 0$. and $\alpha(\cdot)$ continuous and monotone increasing), $x^\epsilon(t)$ is said to have well defined behavior at time scale $t/\alpha(\epsilon)$ if there exists a continuous matrix $Y(t)$ such that, for any $\forall \delta > 0, T < \infty$,

$$\lim_{\epsilon \downarrow 0} \sup_{t \in [\delta, T]} \left\| \exp\{A(\epsilon)t/\alpha(\epsilon)\} - Y(t) \right\| = 0$$

The following proposition shows that regularly perturbed (unlike singularly perturbed) systems have extremely simple time scale behavior.

Proposition 3.2 Let (3.1) be a regularly perturbed version of (3.3). Then, for any order function $\alpha(\epsilon)$, $\delta > 0, T < \infty$

$$\lim_{\epsilon \downarrow 0} \sup_{t \in [\delta, T]} \left\| \exp\{A(\epsilon)t/\alpha(\epsilon)\} - P_0 \right\| = 0 \tag{3.6}$$

where P_0 is the eigenprojection for the zero eigenvalue of A_0 .

Proof:

$$\begin{aligned} \left\| \exp\{A(\epsilon)t/\alpha(\epsilon)\} - P_0 \right\| &\leq \left\| \exp\{A(\epsilon)t/\alpha(\epsilon)\} - \exp\{A_0 t/\alpha(\epsilon)\} \right\| + \\ &\left\| \exp\{A_0 t/\alpha(\epsilon)\} - P_0 \right\| \end{aligned} \tag{3.7}$$

By the definition of regular perturbation, the first term of the r.h.s. of (3.7) converges to 0 as $\epsilon \downarrow 0$ uniformly in t . For the second term we write

$$\exp\{A_0 t\} = P_0 - \frac{1}{2\pi i} \int_{\Gamma_0} e^{\lambda t} R(\lambda, A_0) d\lambda \tag{3.8}$$

where Γ_0 is a contour enclosing all nonzero eigenvalues of A_0 . By the assumption of semistability of A_0 , we may choose Γ_0 to be in the left half plane bounded away from the ~~jo~~ $j\omega$ -axis, say by the line $\{\lambda: \text{Re } \lambda = -\beta\}$. Using (3.8) we then

have:

$$\| \exp\{A_0 t/\alpha(\epsilon)\} - P_0 \| \leq K e^{-\beta\delta/\alpha(\epsilon)} \quad \text{for } t \in [\delta, \infty[\quad (3.9)$$

Taking limits on both sides of (3.7) using (3.9) proves (3.6). \square

To complete our discussion of the distinction between regular and singularly perturbed systems, we give a necessary and sufficient condition for (3.1) to be a singularly perturbed version of (3.3)

Proposition 3.3 The system (3.1) is singularly perturbed if and only if $\text{rank } A_0 < \text{nrnk } A(\epsilon)$.

Proof: Necessity is established by contradiction. Let $\text{nrnk } A(\epsilon) = \text{rank } A_0$. Since the set of eigenvalues of $A(\epsilon)$ is a continuous function of ϵ , the zero eigenvalue of $A(\epsilon)$ does not split. Hence, for ϵ small a contour γ_0 enclosing the origin can be found such that it only encloses the zero eigenvalue of $A(\epsilon)$. Since $A(\epsilon)$ is assumed to be semistable, the only singularity of the resolvent $R(\lambda, A(\epsilon))$ within γ_0 is a pole at $\lambda = 0$ with residue $P_0(\epsilon)$ and we obtain

$$-\frac{1}{2\pi i} \int_{\gamma_0} e^{\lambda t} R(\lambda, A(\epsilon)) d\lambda = P_0(\epsilon) \quad (3.10)$$

From section 2, we have that $P_0(\epsilon) \rightarrow P_0$ as $\epsilon \downarrow 0$, where P_0 is the eigenprojection for the zero eigenvalue of A_0 ,

$$P_0 = -\frac{1}{2\pi i} \int_{\gamma_0} e^{\lambda t} R(\lambda, A_0) d\lambda \quad (3.11)$$

Using (3.10), (3.11) we have

$$\| \exp\{A(\epsilon)t\} - \exp\{A_0 t\} \| \leq \frac{1}{2\pi} \int_{\Gamma_0} \| R(\lambda, A(\epsilon)) - R(\lambda, A_0) \| e^{\text{Re}\lambda t} d\lambda + \| P_0(\epsilon) - P_0 \|$$

where Γ_0 is a positive contour enclosing all nonzero eigenvalues of $A_0(\epsilon)$ for ϵ small. Since $R(\lambda, A_0(\epsilon))$ converges uniformly to $R(\lambda, A_0)$ on Γ_0 and Γ_0 can be

chosen to lie in \mathbb{C}_- bounded away from the $j\omega$ -axis (by semistability of A_0), we have

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \|\exp\{A(\epsilon)t\} - \exp\{A_0 t\}\| = 0 \quad (3.12)$$

which establishes the contradiction.

Sufficiency is also established by contradiction. If (3.1) is a regularly perturbed version of (3.3), then

$$\lim_{\epsilon \downarrow 0} P(\epsilon) \triangleq \lim_{\epsilon \downarrow 0} \lim_{t \rightarrow \infty} \exp\{A(\epsilon)t\} = \lim_{t \rightarrow \infty} \exp\{A_0 t\} = P_0$$

But $P(\epsilon)$, P_0 are the eigenprojections for the zero eigenvalue of $A(\epsilon)$, A_0 respectively; and, by Proposition 2.6, $\text{rank } P(\epsilon) = \text{rank } P_0$ thus establishing a contradiction because $\text{rank } P(\epsilon) = \text{null } A(\epsilon)$ and $\text{rank } P_0 = \text{null } A_0$. \square

Remarks:

- 1) If A_0 is asymptotically stable, then any perturbation is regular.
- 2) There is a heuristic connection between the time scale evolution of (3.1) and the eigenvalues of $A(\epsilon)$. In particular, eigenvalues of order ϵ^k are symptomatic of system behavior at time scale t/ϵ^k . However, there are several detailed assumptions and delicate analysis to be performed to validate this heuristic reasoning. This is the focus of our attention in the following sections.

Section 4. Complete Time Scale Decomposition

4.1 Spatial and Temporal Decomposition of $\exp\{A(\epsilon)t\}$. The Multiple Semisimple Null Structure Condition

To facilitate the notation in the development that follows, we choose for the perturbed system (3.1) the notation:

$$\dot{x}^\epsilon(t) = A_0(\epsilon) x^\epsilon(t), \quad x^\epsilon(0) = x_0 \quad (4.1)$$

with

$$A_0(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p A_{0p} \quad (4.2)$$

Of obvious interest here is when (4.1) is singularly perturbed -time scale behavior is trivial when the perturbation is regular as shown by Proposition 3.2. We thus restrict our attention to the case $\text{rank } A_0 < n$ and $\text{rank } A_0(\epsilon)$. For our development we need to construct a sequence of matrices $A_k(\epsilon)$, $k=1, \dots, m$ obtained recursively from $A_0(\epsilon)$ as indicated below.

Recall the notation of Section 2. Let $P_0(\epsilon)$ denote the total projection for the zero group of eigenvalues of $A_0(\epsilon)$. From Corollary 2.8 it follows that if A_{00} has semisimple null structure (SSNS) then the matrix

$$A_1(\epsilon) \triangleq \frac{P_0(\epsilon)A_0(\epsilon)}{\epsilon} = \frac{A_0(\epsilon)P_0(\epsilon)}{\epsilon} = \frac{P_0(\epsilon)A_0(\epsilon)P_0(\epsilon)}{\epsilon}$$

has a series expansion of the form

$$A_1(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p A_{1p} \tag{4.3}$$

If the first term in the series (4.3), namely A_{10} , has SSNS it follows that

$$A_2(\epsilon) \triangleq \frac{P_1(\epsilon)A_1(\epsilon)}{\epsilon} = \frac{P_1(\epsilon)P_0(\epsilon)A_0(\epsilon)}{\epsilon^2}$$

where $P_1(\epsilon)$ is the total projection for the zero group of eigenvalues of $A_1(\epsilon)$, has series expansion

$$A_2(\epsilon) = \sum_{p=0}^{\infty} \epsilon^p A_{2p} \tag{4.4}$$

The recursion ends at step m , i.e., at

$$\begin{aligned} A_m(\epsilon) &\triangleq \frac{P_{m-1}(\epsilon)A_{m-1}(\epsilon)}{\epsilon} = \frac{P_{m-1}(\epsilon)P_{m-2}(\epsilon)\dots P_0(\epsilon)A_0(\epsilon)}{\epsilon^m} \\ &= \sum_{p=0}^{\infty} \epsilon^p A_{mp} \end{aligned} \tag{4.5}$$

if the matrix A_{m0} does not have SSNS. The following proposition establishes several properties of the matrices $A_k(\epsilon)$, $P_k(\epsilon)$. Define $Q_k(\epsilon) = I - P_k(\epsilon)$; note that $Q_k(\epsilon)$ is also a projection (onto the eigenspaces of the nonzero groups of eigenvalues of $A_k(\epsilon)$).

Proposition 4.1 For ε small enough, including zero, and $k=1, \dots, m$

- (i) $P_i(\varepsilon) P_j(\varepsilon) = P_j(\varepsilon) P_i(\varepsilon) \quad i, j = 0, 1, \dots, m$
- (ii) $Q_i(\varepsilon) Q_j(\varepsilon) = 0 \quad i \neq j, \quad i, j = 0, 1, \dots, m$
- (iii) $\mathbb{C}^n = R(Q_0(\varepsilon)) \oplus \dots \oplus R(Q_k(\varepsilon)) \oplus R(P_0(\varepsilon) \dots P_k(\varepsilon))$
- (iv) $\text{rank } Q_k(\varepsilon) = \text{rank } A_{k0}$

and for ε small enough but not zero,

$$(v) \quad Q_k(\varepsilon) A_0(\varepsilon) = \varepsilon^k Q_k(\varepsilon) A_k(\varepsilon) = \varepsilon^k A_k(\varepsilon) Q_k(\varepsilon) = A_0(\varepsilon) Q_k(\varepsilon) \quad \square$$

The proof of this result is a modification of results in [1]. See [20] for details.

The following proposition establishes that the sequence $A_k(\varepsilon)$ always terminates at some finite m .

Proposition 4.2 Let $A_k(\varepsilon)$, $k=0, 1, \dots$, be the sequence of matrices defined recursively by (4.5). At least one of the following two conditions (possibly both) are satisfied at some $m < \infty$:

- (i) A_{m0} does not have SSNS
- (ii) $A_{m+1}(\varepsilon) = 0$ or, equivalently,
- (ii') $\sum_{k=0}^m \text{rank } A_{k0} = d$

Proof: It only needs to be shown that (ii) occurs for $m < \infty$ if (i) does not.

From Proposition 4.1, for all $j \geq 0$,

$$\mathbb{C}^n = R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_j(\epsilon)) \oplus R(P_0(\epsilon) \dots P_j(\epsilon)) \quad (4.6)$$

Since $\text{rank } Q_k(\epsilon) = \text{rank } A_{k0}$, only a finite number of A_{k0} 's can be nonzero.

Let m be such that $A_{m0} \neq 0$ and $A_{k0} = 0$ for $k > m$. If $A_{k0} = 0$, $P_k(\epsilon) = I$.

Hence, $A_{k0} = 0$ for $k > m$ implies that $A_{m+1}(\epsilon) = 0$.

To show the equivalence of (ii) and (ii'), note that

$$A_{m+1}(\epsilon) = \frac{A_0(\epsilon) P_0(\epsilon) \dots P_m(\epsilon)}{\epsilon^m}$$

Hence, if $A_{m+1}(\epsilon) = 0$ implies that $R(P_0(\epsilon) \dots P_m(\epsilon)) \subset N(A_0(\epsilon))$. On the

other hand, if $x \in N(A_0(\epsilon))$, then $x \in N(A_k(\epsilon))$ and therefore $P_k(\epsilon)x = x$.

Thus, $N(A_0(\epsilon)) = R(P_0(\epsilon) \dots P_m(\epsilon))$. Using this in (4.6) yields that

(ii) \Rightarrow (ii'). The proof of the converse is similar. \square

Definition 4.3 An analytic matrix function $A_0(\epsilon)$ of ϵ satisfies the multiple semisimple null structure (MSSNS) condition if the sequence of matrices $A_k(\epsilon)$ can be constructed until the stopping condition (ii') of Proposition 4.2 has been met with all the matrices

$$A_{k0} = \lim_{\epsilon \downarrow 0} \frac{P_{k-1}(\epsilon) \dots P_0(\epsilon) A_0(\epsilon)}{\epsilon^k} \quad k = 0, 1, \dots, m$$

having semisimple null structure (SSNS).

Proposition 4.4 If $A_0(\epsilon)$ satisfies the MSSNS condition then, for some

$\epsilon_1 > 0$,

(i) $A_k(\epsilon)$ has SSNS for $\epsilon \in [0, \epsilon_1]$, $k=0, \dots, m$

(ii) For $\epsilon \in]0, \epsilon_1]$

$$R(A_k(\epsilon)) = R(Q_k(\epsilon)) \oplus \dots \oplus R(Q_m(\epsilon)) \quad k=0, \dots, m \quad (4.7)$$

$$N(A_k(\epsilon)) = R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_{k-1}(\epsilon)) \oplus N(A_0(\epsilon)) \quad k=1, \dots, m \quad (4.8)$$

$$N(A_0(\epsilon)) = R(P_0(\epsilon) \dots P_m(\epsilon)) \quad (4.9)$$

(iii) If $\lambda(\epsilon)$ is an eigenvalue of $A_k(\epsilon)$ not belonging to its zero group then $\epsilon^k \lambda(\epsilon)$ is an eigenvalue of $A_0(\epsilon)$ in $R(Q_k(\epsilon))$. Conversely, if $\mu(\epsilon)$ is an eigenvalue of $A_0(\epsilon)$ in $R(Q_k(\epsilon))$ then $\epsilon^{-k} \mu(\epsilon)$ is an eigenvalue of $A_k(\epsilon)$ not belonging to its zero-group.

Proof: Equation (4.9) has been established in the proof of Proposition 4.2. Further, if $y \in R(A_0(\epsilon))$ then $y = A_0(\epsilon)x$ for some x . Now, using (iii) of Proposition 4.1, and (4.9) above

$$y = \sum_{k=0}^m A_0(\epsilon) Q_k(\epsilon) x = \sum_{k=0}^m Q_k(\epsilon) A_0(\epsilon) x$$

so that $R(A_0(\epsilon)) \subset R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_m(\epsilon))$. Equality of the subspaces follows from counting dimensions. Check that this finishes the proof of (i) - (iii) for $k = 0$.

Consider $N(A_k(\epsilon))$. By definition of $A_k(\epsilon)$ we have, for ϵ small enough but nonzero,

$$N(A_k(\epsilon)) \supset N(A_0(\epsilon)) \oplus R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_{k-1}(\epsilon)) \quad (4.10)$$

Establish inclusion in the other direction by contradiction. Let $x \in N(A_k(\epsilon))$ but not the right hand side of (4.10). From (iii) of Proposition 4.1,

$$\mathbb{C}^n = R(Q_0(\epsilon)) \oplus \dots \oplus R(Q_m(\epsilon)) \oplus N(A_0(\epsilon))$$

Hence, write $x = x_1 + x_2$ with $x_1 \in \mathcal{R}(Q_0(\epsilon)) \oplus \dots \oplus \mathcal{R}(Q_{k-1}(\epsilon)) \oplus N(A_0(\epsilon))$ and $0 \neq x_2 \in \mathcal{R}(Q_k(\epsilon)) \oplus \dots \oplus \mathcal{R}(Q_m(\epsilon))$ with $P_l(\epsilon)x = x$ if $l < k$. Now $x \in N(A_k(\epsilon))$ implies that

$$\begin{aligned} 0 = A_k(\epsilon)x &= \frac{A_0(\epsilon) P_0(\epsilon) \dots P_{k-1}(\epsilon)}{\epsilon^k} x \\ &= \frac{A_0(\epsilon) P_0(\epsilon) \dots P_{k-1}(\epsilon)}{\epsilon^k} x_2 \\ &= \frac{A_0(\epsilon)}{\epsilon^k} x_2 \end{aligned}$$

i.e., $x_2 \in N(A_0(\epsilon))$ thereby yielding a contradiction. This establishes (4.8).

To prove (4.7), note that by definition of $A_k(\epsilon)$

$$\mathcal{R}(A_k(\epsilon)) \subset \mathcal{R}(P_0(\epsilon) \dots P_{k-1}(\epsilon)) \cap \mathcal{R}(A_0(\epsilon))$$

and, it follows from Proposition 4.1, and the SSNS of $A_0(\epsilon)$ that

$$\mathcal{R}(P_0(\epsilon) \dots P_{k-1}(\epsilon)) \cap \mathcal{R}(A_0(\epsilon)) = \mathcal{R}(Q_k(\epsilon)) \oplus \dots \oplus \mathcal{R}(Q_m(\epsilon))$$

Equality (4.7) follows now from counting dimensions. To prove (iii) notice that if $A_k(\epsilon)u = \lambda(\epsilon)u$ and $\lambda(\epsilon)$ does not belong to the zero group of eigenvalues of $A_k(\epsilon)$ then $Q_k(\epsilon)u = u$ and therefore it follows from Proposition 4.1- (v) that:

$$A_0(\epsilon)u = A_0(\epsilon) Q_k(\epsilon)u = \epsilon^k A_k(\epsilon)u = \epsilon^k \lambda(\epsilon)u$$

Conversely, let $A_0(\epsilon)u = \mu(\epsilon)u$ with $u \in \mathcal{R}(Q_k(\epsilon))$ then,

$$\epsilon^{-k} \mu(\epsilon)u = \epsilon^{-k} A_0(\epsilon) Q_k(\epsilon)u = A_k(\epsilon)u$$

□

Proposition 4.4 establishes that if $A_0(\epsilon)$ has MSSNS, then it may be decomposed as

$$A_0(\epsilon) = \sum_{k=0}^m \epsilon^k A_k(\epsilon) Q_k(\epsilon) \quad (4.10)$$

and that the eigenvalues of $A_0(\epsilon)$ may be divided into $(m+1)$ groups corresponding to eigenvalues of order ϵ^j , $j=0, \dots, m$, in the invariant subspaces $R(Q_j(\epsilon))$. Further, the eigenvalues of order ϵ^k coincide with ϵ^k times the order one eigenvalues of $A_k(\epsilon)$. The ranges and nullspaces of $A_k(\epsilon)$ are shown in Fig. 1; in addition to $N(A_0(\epsilon))$, $N(A_k(\epsilon))$ includes the eigenspaces of $A_0(\epsilon)$ corresponding to eigenvalues of order $1, \epsilon, \dots, \epsilon^{k-1}$; $R(A_k(\epsilon))$, on the other hand, includes the eigenspaces of $A_0(\epsilon)$ corresponding to all eigenvalues of order ϵ^k .

The following theorem (one of two central results) illustrates the consequences of MSSNS for the time scales behavior of $\exp\{A_0(\epsilon)t\}$.

Theorem 4.5 If $A_0(\epsilon)$ satisfies the MSSNS condition, then:

$$\exp\{A_0(\epsilon)t\} = \sum_{k=0}^m Q_k(\epsilon) \exp\{A_k(\epsilon)\epsilon^k t\} + P_0(\epsilon) \dots P_m(\epsilon) \quad (4.11)$$

$$= \sum_{k=0}^m \exp\{Q_k(\epsilon) A_k(\epsilon)\epsilon^k t\} - mI \quad (4.12)$$

$$= \prod_{k=0}^m \exp\{Q_k(\epsilon) A_k(\epsilon)\epsilon^k t\} \quad (4.13)$$

Proof: Write

$$\begin{aligned} \exp\{A_0(\epsilon)t\} &= P_0(\epsilon) \exp\{A_0(\epsilon)t\} + Q_0(\epsilon) \exp\{A_0(\epsilon)t\} \\ &= \exp\{A_1(\epsilon)\epsilon t\} - Q_0(\epsilon) + Q_0(\epsilon) \exp\{A_0(\epsilon)t\} \end{aligned}$$

Repeating this manipulation for $\exp\{A_1(\epsilon)\epsilon t\}$ we have

$$\begin{aligned} \exp\{A_0(\epsilon)t\} &= \exp\{A_2(\epsilon)\epsilon^2 t\} + Q_1(\epsilon) \exp\{A_1(\epsilon)\epsilon t\} \\ &\quad + Q_0(\epsilon) \exp\{A_0(\epsilon)t\} - Q_1(\epsilon) - Q_0(\epsilon) \end{aligned}$$

Repeating this procedure m times yields

$$\begin{aligned} \exp\{A_0(\varepsilon)t\} &= \exp\{A_{m+1}(\varepsilon) \varepsilon^{m+1}t\} - \sum_{k=0}^m Q_k(\varepsilon) \\ &+ \sum_{k=0}^m Q_k(\varepsilon) \exp\{A_k(\varepsilon)\varepsilon^k t\} \end{aligned} \quad (4.14)$$

But $A_{m+1}(\varepsilon) = 0$ and $I - \sum_{k=0}^m Q_k(\varepsilon) = P_0(\varepsilon) \dots P_m(\varepsilon)$ so that (4.14) yields (4.11). Use the identity

$$Q_k(\varepsilon) \exp\{A_k(\varepsilon)\varepsilon^k t\} = \exp\{Q_k(\varepsilon) A_k(\varepsilon)\varepsilon^k t\} - I + Q_k(\varepsilon)$$

in (4.16) to obtain (4.17). Equation (4.18) follows directly from (4.10), the property (v) of proposition 4.1 and the fact that:

$$Q_k(\varepsilon) A_k(\varepsilon) \cdot Q_j(\varepsilon) A_j(\varepsilon) = 0 \quad j \neq k \quad \square$$

Remark:

Under the MSSNS condition, equation (4.11) of Theorem 4.5 gives a spatial and temporal decomposition of $\exp\{A_0(\varepsilon)t\}$ -e.g. $Q_k(\varepsilon) \exp\{A_k(\varepsilon)\varepsilon^k t\}$ does not change significantly in time until t is of order $1/\varepsilon^k$. This decomposition is of crucial importance in studying multiple time scale behavior, uniform asymptotic approximations and reduced order models for the system (4.1).

4.2 Uniform Asymptotic Approximation of $\exp\{A_0(\varepsilon)t\}$: The Multiple Semistability Assumption.

As stated in the previous section, $\exp\{A_0(0)t\}$ is a uniform approximation to $\exp\{A_0(\varepsilon)t\}$ on any compact time interval $[0, T]$. To capture all the multiple time scale behavior, however, it is necessary to have a uniform asymptotic approximation on $[0, \infty[$. For this we need the following condition:

Definition 4.6 $A_0(\epsilon)$ satisfies the multiple semistability (MSST) condition if:

- (i) $A_0(\epsilon)$ satisfies the MSSNS condition, and
- (ii) the matrices

$$A_{k_0} = \lim_{\epsilon \downarrow 0} \frac{P_{k-1}(\epsilon) \dots P_0(\epsilon) A_0(\epsilon)}{\epsilon^k}$$

for $k=0,1,\dots,m$ are semistable.

The following is a central result in uniform approximation:

Theorem 4.7 Let $A_0(\epsilon)$ satisfy the MSST condition. Then,

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \|\exp\{A_0(\epsilon)t\} - \phi(t,\epsilon)\| = 0 \quad (4.15)$$

where $\phi(t,\epsilon)$ is any of the following expressions:

$$\phi(t,\epsilon) = \sum_{k=0}^m Q_k \exp A_{k_0} \epsilon^k t + P_0 \dots P_m \quad (4.16)$$

$$= \sum_{k=0}^m \exp\{A_{k_0} \epsilon^k t\} - mI \quad (4.17)$$

$$= \prod_{k=0}^m \exp\{A_{k_0} \epsilon^k t\} \quad (4.18)$$

$$= \exp\left\{\sum_{k=0}^m A_{k_0} \epsilon^k t\right\}$$

where $A_{k_0} = \lim_{\epsilon \downarrow 0} A_k(\epsilon)$, $P_k = \lim_{\epsilon \downarrow 0} P_k(\epsilon)$ and $Q_k = \lim_{\epsilon \downarrow 0} Q_k(\epsilon)$. Furthermore,

$$\mathbb{C}^n = R(A_{0_0}) \oplus \dots \oplus R(A_{m_0}) \oplus \left(\bigcap_{k=0}^m N(A_{k_0})\right) \quad (4.19)$$

Proof: We first establish (4.15) with $\phi(t,\epsilon)$ as in (4.16). Using (4.11)

from Theorem 4.5 for $\exp\{A_0(\epsilon)t\}$

$$\exp\{A_0(\varepsilon)t\} - \phi(t, \varepsilon) = (P_0(\varepsilon) \dots P_m(\varepsilon) - P_0 \dots P_m) + \sum_{k=0}^m (Q_k(\varepsilon) \exp\{A_k(\varepsilon)\varepsilon^k t\} - Q_k \exp\{A_{k0}\varepsilon^k t\})$$

The first term in the above equation tends to zero as $\varepsilon \downarrow 0$ independently of t . For the second term write

$$\begin{aligned} \psi_k(t, \varepsilon) &\triangleq Q_k(\varepsilon) \exp\{A_k(\varepsilon)\varepsilon^k t\} - Q_k \exp\{A_{k0}\varepsilon^k t\} = \\ &= \frac{1}{2\pi i} \int_{\Gamma_k} e^{\lambda \varepsilon^k t} (R(\lambda, A_k(\varepsilon)) - R(\lambda, A_{k0})) d\lambda \end{aligned}$$

where Γ_k is a contour enclosing all nonzero eigenvalues of A_{k0} . By semistability of A_{k0} , Γ_k can be chosen to lie in the left half plane bounded away from the $j\omega$ -axis. Hence, we have for some $\alpha < 0$

$$\begin{aligned} \|\psi_k(t, \varepsilon)\| &\leq \frac{1}{2\pi} e^{\alpha \varepsilon^k t} \int_{\Gamma_k} \|R(\lambda, A_k(\varepsilon)) - R(\lambda, A_{k0})\| d\lambda \leq \\ &\frac{1}{2\pi} \int_{\Gamma_k} \|R(\lambda, A_k(\varepsilon)) - R(\lambda, A_{k0})\| d\lambda \end{aligned}$$

Since $R(\lambda, A_k(\varepsilon))$ converges uniformly to $R(\lambda, A_{k0})$ on compact subsets of C (by Lemma 2.5), we have that $\|\psi_k(t, \varepsilon)\|$ tends to zero as $\varepsilon \downarrow 0$ (uniformly in t). Equality between the different expressions of $\phi(t, \varepsilon)$ is established as in Theorem 4.5.

To establish (4.19) we have from (iii) of Proposition 4.1 that

$$\mathbb{C}^n = R(Q_0(\varepsilon)) \oplus \dots \oplus R(Q_m(\varepsilon)) \oplus R(P_0(\varepsilon) \dots P_m(\varepsilon))$$

and by continuity of the projections $Q_k(\varepsilon)$, $P_k(\varepsilon)$,

$$\mathbb{C}^n = R(Q_0) \oplus \dots \oplus R(Q_m) \oplus R(P_0 \dots P_m)$$

The direct sum decomposition (4.19) follows now from the fact that, by construction, Q_k is the projection on $R(A_{k0})$ along $N(A_{k0})$. □

In the next section we use the result of Theorem 4.7 to determine the complete multiple time scale behavior for (4.1) and obtain a set of reduced order models.

4.3 Multiple Time Scale Behavior and Reduced Order Models.

Multiple time scale behavior is explicated by the following corollary to Theorem 4.7.

Corollary 4.8 Let $A_0(\epsilon)$ satisfy the MSST condition. Then,

(i)

$$\lim_{\epsilon \downarrow 0} \sup_{\delta \leq t \leq T} \left| \exp\{A_0(\epsilon)t/\epsilon^k\} - \Phi_k(t) \right| = 0 \quad (4.20)$$

$$\forall \delta > 0, T < \infty; k=0,1,\dots,m-1$$

(ii)

$$\lim_{\epsilon \downarrow 0} \sup_{\delta \leq t < \infty} \left| \exp\{A_0(\epsilon)t/\epsilon^m\} - \Phi_m(t) \right| = 0 \quad \forall \delta > 0 \quad (4.21)$$

where $\Phi_k(t)$ is given by:

$$\Phi_k(t) = Q_k \exp\{A_{k0} t\} + P_0 \dots P_k \quad (4.22)$$

$$= P_0 \dots P_{k-1} \exp\{A_{k0} t\} \quad k=0,1,\dots,m \quad (4.23)$$

Proof: From Theorem 4.7 we have that

$$\begin{aligned} \exp\{A_0(\epsilon)t/\epsilon^k\} &= \sum_{\ell=0}^{k-1} Q_\ell \exp\{A_{\ell 0} t/\epsilon^{k-\ell}\} + Q_k \exp\{A_{k0} t\} + \\ &\quad \sum_{\ell=k+1}^m Q_\ell \exp\{A_{\ell 0} t/\epsilon^{\ell-k}\} + P_0 \dots P_m + o(1) \end{aligned} \quad (4.24)$$

uniformly for $t \in [0, \infty[$. Now, by the semistability of $A_{\ell 0}$,

$$Q_\ell \exp\{A_{\ell 0} t\} = -\frac{1}{2\pi i} \int_{\Gamma_\ell} e^{\lambda t} R(\lambda, A_{\ell 0}) d\lambda$$

for some Γ_ℓ in \mathbb{C}_- bounded away from the $j\omega$ -axis. By the boundedness of $R(\lambda, A_{k0})$ on Γ_ℓ ,

$$||Q_\ell \exp\{A_{\ell 0} t\}|| \leq M_\ell e^{-\alpha_\ell t} \quad \text{with } \alpha_\ell > 0 .$$

This yields

$$\lim_{\varepsilon \downarrow 0} \sup_{\delta \leq t \leq \infty} \sum_{\ell=0}^{k-1} Q_\ell \exp\{A_{\ell 0} t/\varepsilon^{k-\ell}\} = 0 \quad (4.25)$$

On the other hand, it is clear that

$$\lim_{\varepsilon \downarrow 0} \sup_{0 \leq t \leq T} ||\exp\{A_{\ell 0} \varepsilon^{\ell-k} t\} - I|| = 0 \quad \forall \ell > k, T < \infty \quad (4.26)$$

Using (4.25), (4.26) in (4.24) yields (4.20) and (4.21) with

$$\Phi_k(t) = Q_k \exp\{A_{k0} t\} + \sum_{\ell=k+1}^m Q_\ell + P_0 \dots P_m , \quad (4.27)$$

equality of expressions (4.22) and (4.23) follows from (4.27). \square

Remarks:

1) From (4.22) of Corollary 4.8 and (4.16) of Theorem 4.7 it follows that:

$$\exp\{A_0(\varepsilon)t\} = \sum_{k=0}^m \Phi_k(\varepsilon^k t) - \sum_{k=0}^{m-1} P_0 \dots P_k + o(1) \quad (4.28)$$

uniformly for $t > 0$. Thus, only the behavior at time scales t/ε^k , $k=0, \dots, m$ is needed to capture the evolution of $\exp\{A_0(\varepsilon)t\}$ on $[0, \infty[$. From the proof of Corollary 4.8 it is clear that

$$\lim_{\varepsilon \downarrow 0} \exp\{A_0(\varepsilon)t/\alpha(\varepsilon)\}$$

exists for any order function $\alpha(\varepsilon)$. Indeed if $\alpha_k(\varepsilon) = o(\varepsilon^k)$ and $\varepsilon^{k+1} = o(\alpha_k(\varepsilon))$

then

$$\lim_{\epsilon \downarrow 0} \exp\{A_0(\epsilon)t/\alpha_k(\epsilon)\} = P_0 \dots P_k \quad (4.29)$$

and for $\alpha(\epsilon) = o(\epsilon^m)$

$$\lim_{\epsilon \downarrow 0} \exp\{A_0(\epsilon)t/\alpha(\epsilon)\} = P_0 \dots P_m \quad (4.30)$$

Thus, the system has well defined behavior at all time scales, even though only a finite number of them (t/ϵ^k , $k=0,1,\dots,m$), called the fundamental or natural time scales, are required to capture the system evolution (strictly speaking only those for which $A_{k0} \neq 0$).

2) The behavior of the system as given by (4.23) is canonic in the following sense: at a given time scale, say t/ϵ^k , all faster time scales t/ϵ^ℓ for $\ell < k$ have come to their equilibria (respectively P_ℓ); and all slower time scales t/ϵ^ℓ for $\ell > k$ have yet to evolve. \square

To interpret the matrices A_{k0} as reduced order models of (4.1), notice that the uniform asymptotic approximation

$$\exp\{A_0(\epsilon)t\} = \sum_{k=0}^m Q_k \exp\{A_{k0} \epsilon^k t\} + P_0 \dots P_m + o(1)$$

together with (4.19) imply that the subspaces $R(Q_k)$, $k=0,\dots,m$, are almost invariant subspaces (or ϵ -invariant subspaces, as defined in [16]) of (4.1). The parts of $x^\epsilon(t)$ that evolve in different subspaces do so at different time scales.

Corollary 4.9 Consider the linear systems

$$\dot{y}_k = A_{k0} y_k, \quad y_k(0) = Q_k x_0, \quad k=0,1,\dots,m \quad (4.31)$$

Then,

$$Q_k x^\epsilon(t) = y_k(\epsilon^k t) + o(1), \quad k=0,1,\dots,m$$

and

$$x^\varepsilon(t) = \sum_{k=0}^m y_k(\varepsilon^k t) + P_0 \dots P_m x_0 + o(1) \quad (4.32)$$

Proof: A straightforward modification of Corollary 4.8. □

Remarks:

1) The linear systems (4.31) though written as equations on \mathbb{R}^n are really reduced order models since each evolves in $\mathcal{R}(Q_k)$, $k=0,1,\dots,m$

and

$$\sum_{k=0}^m \dim \mathcal{R}(Q_k) = n - \text{rank } A_0(\varepsilon)$$

2) In particular, choosing a basis adapted to the direct sum decomposition (4.19) it is possible to asymptotically decouple (4.1) into a set of lower dimensional systems each evolving at a different time scale as follows:

Let V be the (ε -independent) change of basis mentioned above; then Theorem 4.7 can be written as:

$$\begin{aligned} \exp\{A_0(\varepsilon)t\} &= V^{-1} \left\{ \exp \sum_{k=0}^m V A_{k0} V^{-1} \varepsilon^k t \right\} V + o(1) \\ &= V^{-1} \text{diag} \{ e^{\hat{A}_0 t}, e^{\hat{A}_1 \varepsilon t}, \dots, e^{\hat{A}_m \varepsilon^m t}, I \} V + o(1) \end{aligned}$$

where the matrices \hat{A}_k are full rank square matrices with dimension equal to $\text{rank } A_{k0}$ corresponding to the nonzero part of A_{k0} in the new basis. □

In the next section we show that such a complete decomposition and simplification as has been elaborated here is possible only if $A_0(\varepsilon)$ satisfies the MSST condition.

4.4 Necessity of the Multiple Semistability Condition

We have shown in Sections 4.2 and 4.3 the existence of well defined behavior at several time scales under the MSST condition. If MSST is not satisfied then, at least for some order function $\alpha(\epsilon)$, the limit

$$\lim_{\epsilon \downarrow 0} \exp\{A_0(\epsilon)t/\alpha(\epsilon)\}$$

does not exist. To illustrate this, consider the following examples:

Example 4.9 (A_{k_0} not SSNS) Consider the matrix

$$A_0(\epsilon) = \begin{bmatrix} \epsilon & 0 & -2\epsilon \\ \epsilon & \epsilon & -2\epsilon \\ 1 & 1 & -2 \end{bmatrix}$$

semistable for $\epsilon \in [0,1]$ with eigenvalues $\lambda_0 = 0$, $\lambda_1 = -2 + o(1)$ and $\lambda_2 = -\epsilon^2 + o(\epsilon^2)$. This matrix does not satisfy the MSSNS condition; as may be verified (for a systematic procedure to do this calculation see section 4.5) that:

$$A_{10} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & -1/2 & 0 \end{bmatrix}$$

is nilpotent. Also, by direct computation, it is found that

$$\exp\{A_0(\epsilon)t\} = \frac{1}{\lambda_2 - \lambda_1} \sum_{i=1}^2 \begin{bmatrix} (\lambda_2 - \lambda_1) \frac{(-1)^i}{\lambda_i} (e^{\lambda_i t} - 1) & \frac{(-1)^i}{\lambda_i} \epsilon (\lambda_i - \epsilon) (e^{\lambda_i t} - 1) \\ 0 & (-1)^k (\lambda_i - \epsilon + 2) e^{\lambda_i t} & (-1)^{i-1} 2\epsilon e^{\lambda_i t} \\ 0 & (-1)^i e^{\lambda_i t} & (\lambda_i - \epsilon) e^{\lambda_i t} \end{bmatrix}$$

with the following time scale behavior

$$\lim_{\epsilon \downarrow 0} \exp\{A_0(\epsilon)t\} = \exp\{A_{00}t\} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & (1-e^{-2t})/2 & e^{-2t} \end{bmatrix}$$

and

$$\lim_{\epsilon \downarrow 0} \exp\{A_0(\epsilon)t/\epsilon\} = P_0 \exp\{A_{10}t\} = \begin{bmatrix} 1 & t/2 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 0 \end{bmatrix}$$

To see that the limit

$$\lim_{\epsilon \downarrow 0} \exp\{A_0(\epsilon)t/\epsilon^2\}$$

does not exist, consider the (1,2) entry of $\exp\{A_0(\epsilon)t/\epsilon^2\}$:

$$\frac{1}{\lambda_2 - \lambda_1} \left[\frac{\epsilon}{\lambda_2} (e^{\lambda_2 t/\epsilon^2} - 1) - \frac{\epsilon}{\lambda_1} (e^{\lambda_1 t/\epsilon^2} - 1) \right] \quad (4.33)$$

Since $\lambda_2 = -\epsilon^2 + o(\epsilon^2)$, the first term in equation (4.33) is of order $1/\epsilon$ as $\epsilon \downarrow 0$. Thus, the system does not have well defined behavior at time scale t/ϵ^2 even though it has a negative real eigenvalue of order ϵ^2 .

Example 4.10 (A_{KO} not MSST). Consider the matrix

$$A_0(\epsilon) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -\epsilon^2 & \epsilon \\ 0 & -\epsilon & -\epsilon^2 \end{bmatrix}$$

This matrix is semistable for $\epsilon \geq 0$ and it has the three eigenvalues $\lambda_0 = -2$, $\lambda_1 = -\epsilon^2 + i\epsilon$ and $\lambda_2 = -\epsilon^2 - i\epsilon$. Also,

$$A_{00} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad A_{10} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

MSST is violated since A_{10} has purely imaginary eigenvalues. Calculation yields

$$\exp\{A_0(\epsilon)t\} = \begin{bmatrix} e^{-2t} & 0 & 0 \\ 0 & e^{-\epsilon^2 t} \cos \epsilon t & -e^{-\epsilon^2 t} \sin \epsilon t \\ 0 & e^{-\epsilon^2 t} \sin \epsilon t & e^{-\epsilon^2 t} \cos \epsilon t \end{bmatrix}$$

The system has well defined behavior at time scales t and t/ϵ but $\exp\{A_0(\epsilon)t/\epsilon^2\}$ does not have a limit as $\epsilon \downarrow 0$ because of the presence of terms of the form $e^{-t} \sin t/\epsilon$. (The attenuation is slower than the frequency of oscillation.)

In fact, the MSST condition is a necessary and sufficient condition for the existence of multiple time scale behavior.

Theorem 4.11 Let $A_0(\epsilon)$ be semistable for $\epsilon \in [0, \epsilon_0]$ and let A_{k0} , $k \geq 0$ be the sequence of matrices constructed in Section 4.1. If $A_{00}, A_{10}, \dots, A_{\ell-1,0}$ are semistable but $A_{\ell 0}$ is not, then the limit as $\epsilon \downarrow 0$ of

$$\exp\{A_0(\epsilon)t/\epsilon^q\} \tag{4.34}$$

does not exist for any $\ell < q \leq \ell + 1$. Further, if $A_{\ell 0}$ has a pole on the imaginary axis (including zero) which is not semisimple, then

$$\limsup_{\epsilon \downarrow 0} \sup_{t \geq 0} \left\| \exp\left\{ \frac{A_0(\epsilon)t}{\epsilon^q} \right\} \right\| = \infty$$

Proof: We construct the proof for $\ell = 0$ by contradiction. Assume that the limit

$$\lim_{\varepsilon \downarrow 0} \exp\{A_0(\varepsilon)t/\varepsilon^q\}$$

exists for $t > 0$ and some $q \in]0,1]$. If this limit exists, then so does the limit of $\exp\{P_0(\varepsilon)A_0(\varepsilon)t/\varepsilon^q\}$ as $\varepsilon \downarrow 0$ because:

$$\lim_{\varepsilon \downarrow 0} P_0(\varepsilon) \exp\{A_0(\varepsilon)t/\varepsilon^q\} = \lim_{\varepsilon \downarrow 0} \exp\{P_0(\varepsilon)A_0(\varepsilon)t/\varepsilon^q\} = Q_0.$$

Define,

$$F_0(\varepsilon) = \frac{P_0(\varepsilon)A_0(\varepsilon)}{\varepsilon^q}$$

The next step is to prove that $\sigma(F_0(\varepsilon))$ remains bounded as $\varepsilon \downarrow 0$. Take $0 \neq \lambda(\varepsilon) \in \sigma(P_0(\varepsilon)A_0(\varepsilon))$ and let $\phi(\varepsilon)$ be a corresponding eigenvector with $\|\phi(\varepsilon)\| = 1$. Then,

$$\exp\{P_0(\varepsilon)A_0(\varepsilon)t/\varepsilon^q\} \phi(\varepsilon) = \exp\{\operatorname{Re}\lambda(\varepsilon)t/\varepsilon^q\} \cdot \exp\{i \operatorname{Im}\lambda(\varepsilon)t/\varepsilon^q\} \phi(\varepsilon)$$

and if $\varepsilon_m \downarrow 0$ is a sequence for which $\phi(\varepsilon_m)$ converges,

$$\exp\{\operatorname{Re}\lambda(\varepsilon_m)t/\varepsilon_m^q\} \cdot \exp\{i \operatorname{Im}\lambda(\varepsilon_m)t/\varepsilon_m^q\} \tag{4.35}$$

must also converge as $m \rightarrow \infty$. Now, since the trace of $A_0(\varepsilon)$ has a series expansion in integer powers of ε and the eigenvalues in the zero group of $A_0(\varepsilon)$, $\mu(\varepsilon)$, have non-positive real parts,

$$\operatorname{Re}\mu(\varepsilon)/\varepsilon^k \rightarrow \mu \quad \text{as } \varepsilon \downarrow 0$$

for some integer $k \geq 1$ and some constant μ . We thus conclude that $\operatorname{Re}\lambda(\varepsilon_m)/\varepsilon_m^q$ must converge as $m \rightarrow \infty$. Further, by the convergence of equation (4.35), $\operatorname{Im}\lambda(\varepsilon_m)/\varepsilon_m^q$ must also converge as $\varepsilon \downarrow 0$. Because $\sigma(F_0(\varepsilon))$ remains bounded

as $\varepsilon \downarrow 0$ we can choose t_1 such that

$$|\operatorname{Im} \sigma(F_0(\varepsilon)t_1)| < \pi$$

for ε small enough. Hence, if Ln denotes the principal branch of the logarithmic function, we obtain:

$$\operatorname{Ln} \exp\{F_0(\varepsilon)t_1\} = F_0(\varepsilon)t_1 = \frac{D_0}{\varepsilon^q} + G_\theta(\varepsilon)t_1 \quad (4.36)$$

where the last equality follows from Proposition 2.7 with D_0 being the eigennilpotent for the zero eigenvalue of A_{00} and $G_\theta(\varepsilon)$ a continuous function of ε . The limit

$$\lim_{\varepsilon \downarrow 0} \operatorname{Ln}\{\exp F_0(\varepsilon)t_1\} \triangleq B(t_1)$$

is well defined by the boundedness of $\sigma(F_0(\varepsilon))$ and therefore, by (4.36), $D_0 = 0$, i.e., if the limit of (4.34) exists then A_{00} must be SSNS.

Suppose now that A_{00} has some purely imaginary eigenvalue μ . Then there exists at least one eigenvalue $\mu(\varepsilon)$ of $A_0(\varepsilon)$ such that $\mu(\varepsilon) \rightarrow \mu$ as $\varepsilon \downarrow 0$. Let $\phi(\varepsilon)$ be a corresponding eigenvector with $\|\phi(\varepsilon)\| = 1$ and $\varepsilon_m \downarrow 0$ a sequence for which $\phi(\varepsilon_m)$ converges. Then, if (4.34) has a limit so does

$$\phi(\varepsilon_m)^T \exp\{A_0(\varepsilon_m) t/\varepsilon_m^q\} \phi(\varepsilon_m) = e^{\mu(\varepsilon_m)t/\varepsilon_m^q}$$

which is a contradiction. We have thus shown that if A_{00} is not semistable, (4.34) cannot have a limit as $\varepsilon \downarrow 0$.

To prove the theorem for an arbitrary ℓ , notice that using the same algebraic manipulation as in the proof of Theorem 4.5, we can write

$$\begin{aligned} \exp\{A_0(\epsilon)t/\epsilon^q\} &= \exp\{A_\ell(\epsilon)t/\epsilon^{q-\ell}\} = \sum_{k=0}^{\ell-1} Q_k(\epsilon) + \\ &\sum_{k=0}^{\ell-1} Q_k(\epsilon) \exp\{A_k(\epsilon)t/\epsilon^{q-k}\} \end{aligned} \quad (4.37)$$

By semistability of $A_{00}, \dots, A_{\ell-1,0}$ the second and third sums in the right hand side of (4.37) have well defined limits as $\epsilon \downarrow 0$. Thus, assuming that

$$\exp\{A_0(\epsilon)t/\epsilon^q\}$$

has a limit as $\epsilon \downarrow 0$ so does

$$\exp\{A_\ell(\epsilon)t/\epsilon^{q-\ell}\} \quad \ell < q \leq \ell + 1$$

implying, as proved previously for $\ell = 0$, that $A_{\ell 0}$ is semistable, a contradiction.

To prove the second part of the theorem, suppose that $A_{\ell 0}$ has as eigenvalue on the imaginary axis which is not semisimple. The $\forall M < \infty$ there exists a $T < \infty$ such that

$$\|\exp\{A_{\ell 0} T\}\| > M$$

and because $\exp\{A_\ell(\epsilon)T\}$ converges to $\exp\{A_{\ell 0}T\}$ as $\epsilon \downarrow 0$, we conclude that

$$\lim_{\epsilon \downarrow 0} \sup_{t \geq 0} \|\exp\{A_\ell(\epsilon)t\}\| = \infty$$

The desired result follows now from (4.37). □

The above theorem can be interpreted as saying that if a system has well defined behavior at all time scales then its system matrix must be MSST. As we will discuss in Section 5, there are systems for which this condition is always satisfied. In general, however, the sequence of matrices A_{k0} will have to be computed to check for semistability.

Section 4.5 Computation of the Multiple Time Scale Behavior

Theorem 4.7 reveals that the matrices A_{k0} play a fundamental role in the asymptotic analysis of singularly perturbed systems. These are the leading terms in the series expansions of the matrices

$$A_k(\epsilon) = \frac{P_{k-1}(\epsilon) \dots P_0(\epsilon) A_0(\epsilon)}{\epsilon^k}$$

for $k = 0, 1, \dots, m$. The matrices in the series expansion of $A_k(\epsilon)$ are shown in Figure 3. The $(i+1)$ th row of Figure 2 is computed from the i th row using Corollary 2.8, i.e.,

$$A_{i+1,j} = \sum_{p=1}^{j+1} (-1)^p \sum_{\substack{v_1 + \dots + v_p = j+1 \\ k_1 + \dots + k_{p+1} = p-1 \\ v_i \geq 1, k_1 \geq 0}} S_i^{(k_1)} A_{i v_1} S_i^{(k_2)} \dots A_{i v_p} S_i^{(k_{p+1})} \quad (4.38)$$

$i = 0, 1, \dots, m$
 $j \geq 0$

where

$$S_i^{(0)} = -P_i$$

$$S_i^{(p)} = (A_{i0}^\#)^p, \quad p > 0$$

The formula (4.38) enables the array of matrices A_{ij} to be computed triangularly, so that computation of A_{k0} requires only the computation of A_{ij} for $i = 0, \dots, k-1$ and $j = 0, \dots, k-i$. Thus, the algorithm contained in equation (4.38) may be implemented recursively. In the following proposition we illustrate the complexity of the expressions for the A_{k0} in terms of the given data A_{00}, A_{01}, \dots (i.e. $A_0(\epsilon)$). We note also from (4.34) that the computation of the A_{k0} 's

and hence the asymptotic limit of $\exp\{A_0(\epsilon)t\}$ involves only A_{00}, \dots, A_{0m} (only finitely many matrices in the asymptotic expansion of $A_0(\epsilon)$).

Proposition 4.12 The matrices A_{k0} for $k = 0, 1, 2$ and 3 are given by:

$$A_{00}$$

$$A_{10} = P_0 A_{01} P_0$$

$$A_{20} = P_1 P_0 (A_{02} - A_{01} A_{00}^\# A_{01}) P_0 P_1$$

$$A_{30} = P_2 P_1 P_0 (A_{03} - A_{01} A_{00}^\# A_{02} - A_{02} A_{00}^\# A_{01} + A_{01} A_{00}^\# A_{01} A_{00}^\# A_{01}$$

$$- A_{02} (P_0 A_{01} P_0)^\# A_{02} + A_{02} (P_0 A_{01} P_0)^\# A_{01} A_{00}^\# A_{01}$$

$$- A_{01} A_{00}^\# A_{01} (P_0 A_{01} P_0)^\# A_{02} - A_{01} A_{00}^\# A_{01} (P_0 A_{01} P_0)^\#$$

$$A_{01} A_{00}^\# A_{01}) P_0 P_1 P_2$$

Proof: By somewhat laborious calculation. □

Remarks:

1) If $A(\epsilon)$ is of the form $A + \epsilon B$, then we have

$$A_{00} = A$$

$$A_{10} = P_0 B P_0$$

$$A_{20} = - P_1 P_0 B A^\# B P_0 P_1$$

$$A_{30} = P_2 P_1 P_0 (BA^\# BA^\# B - BA^\# B (P_0 B P_0)^\# B A^\# B) P_0 P_1 P_2$$

Thus, a system of the form

$$\dot{x}^\epsilon(t) = (A + \epsilon B) x^\epsilon(t)$$

may exhibit time scale behavior at time scales of order $1/\epsilon, 1/\epsilon^2, \dots, 1/\epsilon^m$,

a fact that is not widely appreciated in the literature. As examination of the characteristic polynomial of $A + \epsilon B$ will show, no eigenvalue of $A + \epsilon B$ can be of $o(\epsilon^n)$ so that at most $m = n$. Similar reasoning leads us to the conclusion that for

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

m can at most be np .

2) For the classical two time scale formulation of (1.2), normalized to the form (1.1), we have

$$A = A_{00} = \begin{bmatrix} 0 & 0 \\ A_{21} & A_{22} \end{bmatrix} \quad \text{and} \quad B = A_{01} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix}$$

If A_{22} is stable then A is semistable with

$$P_0 = \begin{bmatrix} I & 0 \\ -A_{22}^{-1} A_{21} & 0 \end{bmatrix}$$

so that

$$A_{10} = \begin{bmatrix} A_{11} - A_{12} A_{22}^{-1} A_{21} & 0 \\ -A_{22}^{-1} A_{21} (A_{11} - A_{12} A_{22}^{-1} A_{21}) & 0 \end{bmatrix}$$

From this, we see that the model at the fast time scale is

$$\dot{x}_2 = A_{22} x_2 + A_{21} x_1$$

$$\dot{x}_1 = 0$$

and the reduced order model for the slower dynamics is

$$\dot{x}_1 = (A_{11} - A_{12} A_{22}^{-1} A_{21}) x_1$$

$$x_2 = -A_{22}^{-1} A_{21} x_1$$

If, as is usually assumed, $A_{11} - A_{12} A_{22}^{-1} A_{21}$ is also stable then $\text{rank } A_{22} + \text{rank } (A_{11} - A_{12} A_{22}^{-1} A_{21}) = \text{rank } (A + \epsilon B)$ so that the system (1.2) has only two time scales.

3) If $A_0(\epsilon)$ is a rational function of ϵ , with Taylor series about $\epsilon = 0$ given by

$$A_0(\epsilon) = \sum_{k=0}^{\infty} \epsilon^k A_{0k}$$

Then it is well known (e.g. [17], [18], [19],) that m is the order of the Smith McMillan zero of $A_0(\epsilon)$ at $m = 0$. It may also be established that the matrices A_{k0} defined above are related to block Toeplitz matrices of the form

$$\begin{bmatrix} A_{00} & 0 \\ A_{01} & A_{00} \end{bmatrix} \quad \begin{bmatrix} A_{00} & 0 & 0 \\ A_{01} & A_{00} & 0 \\ A_{02} & A_{01} & A_{00} \end{bmatrix}, \dots \quad \begin{bmatrix} A_{00} & 0 & & 0 \\ & \ddots & & \\ A_{01} & \ddots & \ddots & \\ \vdots & \ddots & \ddots & \ddots \\ A_{0m} & \dots & \dots & A_{01} & \dots & A_{00} \end{bmatrix}$$

The details of this connection will be presented elsewhere, since it is not in the mainstream of our development here.

4.6 Partial Time Scale Decomposition

We discuss here the multiple time scale behavior of systems that do not satisfy the MSST condition of Section 4.2. Such systems have well defined behavior at some, but not all time-scales, and it may be useful to be able to isolate the time scales at which they have well defined behavior. Consider, for example, the case when time $\{A_{k0}\}_{k=0}^m$ have SST, but A_{l0} for some $l \leq m$

violates the semistability condition - i.e., it has at least one non-zero eigenvalue λ , with $\text{Re } \lambda \geq 0$. Then, we have

Proposition 4.13

Let the matrix $A_0(\epsilon)$ satisfy the MSSNS condition and let the matrices A_{k0} , $k = 0, 1, \dots, m$ for $k \neq \ell$ be semistable. Then, $\forall \delta > 0$, $T < \infty$

$$\lim_{\epsilon \downarrow 0} \sup_{\delta \leq t \leq T} \left\| \exp\{A_0(\epsilon)t/\epsilon^k\} - \Phi_k(t) \right\| = 0 \quad \text{for } k=1, \dots, \ell \quad (4.39)$$

where

$$\begin{aligned} \Phi_k(t) &= P_0 \dots P_{k-1} \exp\{A_{k0}t\} \\ &= Q_k \exp\{A_{k0}t\} + P_0 \dots P_k \end{aligned} \quad (4.40)$$

Proof: Follows readily from the proof of Theorem 4.7. □

Remarks: (i) Equation (4.39) does not hold for $k > \ell$. For these values of k we however have

$$\lim_{\epsilon \downarrow 0} \sup_{\delta \leq t \leq T} \left\| P_\ell(\epsilon) \exp\{A_0(\epsilon)t/\epsilon^k\} - \Phi_k(t) \right\| = 0 \quad (4.41)$$

Note that in equation (4.41) the projection matrix $P_\ell(\epsilon)$ annihilates behavior at time scale t/ϵ^ℓ involving unstable or oscillatory modes. In general, however, $P_\ell(\epsilon)$ in (4.41) cannot be replaced by $P_\ell(0)$ so that (4.41) is of limited use in obtaining a uniform asymptotic series for $\exp\{A_0(\epsilon)t\}$.

ii) Sometimes in applications, $A_0(\epsilon)$ satisfies a uniform stability condition, viz

$$\left\| \exp\{A_0(\epsilon)t\} \right\| \leq \kappa \quad \forall t \geq 0, \quad \epsilon \in [0, \epsilon_0] \quad (4.42)$$

Although (4.42) guarantees that $A_0(\varepsilon)$ satisfies the MSSNS condition and that any purely imaginary eigenvalue of the matrices A_{k0} is semisimple, it is not enough to guarantee MSST. A uniformly stable system may not have well defined behavior at some time scales because of the presence of slightly attenuated oscillations that when seen at slower time scales present infinite frequency. For uniformly stable systems, however, Proposition 4.13 can be strengthened.

Proposition 4.14

Let the matrix $A_0(\varepsilon)$ satisfy (4.42) and let the matrices A_{k0} , $k = 0, 1, \dots, m$, be semistable for $k \neq \ell$. Then $\forall \delta > 0, T < \infty$

$$\lim_{\varepsilon \downarrow 0} \sup_{\delta \leq t \leq T} \left\| \exp\{A_0(\varepsilon)t/\varepsilon^k\} - \Phi_k(t) \right\| = 0 \quad k = 1, \dots, \ell$$

$$\lim_{\varepsilon \downarrow 0} \sup_{\delta \leq t \leq T} \left\| P_\ell \exp\{A_0(\varepsilon)t/\varepsilon^k\} - \Phi_k(t) \right\| = 0 \quad k = \ell+1, \dots, m$$

where $\Phi_k(t)$ is as in (4.40) and T can be taken equal to ∞ for $k=m$.

Proof Follows readily from Theorem 4.7 and the properties of uniformly stable systems mentioned above. □

Section 5. Conclusions and Application of our results to the Hierarchical Aggregation of Finite State Markov Processes

We have studied the asymptotic behavior of $\exp\{A_0(\varepsilon)t\}$ over the time interval $[0, \infty)$. We have formalized the notions of multiple time scales and reduced order models valid at different time scales. The most important conclusion is that a certain multiple stability condition referred to as the MSST condition is necessary and sufficient for $\exp\{A_0(\varepsilon)t\}$ to have well defined multiple time scales behavior. We feel that our results will have important

computational consequences for the simulation of large-scale linear systems with weak couplings, but this has yet to be explored.

An application of particular interest to us is the hierarchical aggregation of finite state Markov processes (FSMP) with some rare events. The presence of rare events in a FSMP is modelled by a small parameter ϵ in its matrix of transition rates, e.g. $A_0(\epsilon) = A_0 + \epsilon B$. The matrix of transition probabilities for the FSMP, $\eta_\epsilon(t)$, is then given by

$$P^\epsilon(t) = \exp\{A_0(\epsilon)t\} \tag{5.1}$$

It is shown by us in [4] that when $A_0(\epsilon)$ is a matrix of transition rates then $A_0(\epsilon)$ satisfies the MSST condition so that $P^\epsilon(t)$ always has well defined multiple time scale behavior. The reduced order models that describe the evolution of $P^\epsilon(t)$ at each of its fundamental time scales are then interpreted as increasingly simplified aggregated models of $\eta^\epsilon(t)$ obtained by collapsing several states of $\eta^\epsilon(t)$ into single states of the reduced order model. The aggregation is hierarchical so that the model at a time scale, say, t/ϵ^ℓ can be obtained by coalescing some states of the (already simplified) model valid at time scale $t/\epsilon^{\ell-1}$.

This problem has also been studied in detail in [12] where a sequence of aggregation models is also obtained. However, the question of a uniform asymptotic approximation to (5.1) was not studied in [12]. In [4], we develop the hierarchy of approximations as a uniform asymptotic expansion to (5.1) and relate our results to those of [12].

Acknowledgements

We acknowledge gratefully many helpful discussions with Profs. P.V. Kokotovic, B.C. Levy, S.I. Marcus, G. Verghese and J.C. Willems.

References

- [1] P.V. Kokotovic, R.E. O'Malley, Jr, and P. Sannuti, "Singular Perturbations and Order Reduction in Control Theory - An Overview", Automatica, Vol. 12, pp. 123-132, 1976.
- [2] F.C. Hoppensteadt, "Singular Perturbations on the Infinite Interval", Trans. Amer. Math. Soc., Vol. 123, No. 2, pp. 521-535, 1966.
- [3] P.V. Kokotovic Ed., "Singular Perturbations and Time Scales in Modelling and Control of Dynamic Systems", Univ. of Illinois, Coordinated Science Lab., Rep. #R.901, 1980.
- [4] M. Coderch, A.S. Willsky and S.S. Sastry, and D. Castanon, "Hierarchical Aggregation of Singularly Perturbed Finite State Markov Processes", to appear in Stochastics.
- [5] V.F. Butuzov and N.N. Nefedov, "A Problem in Singular Perturbation Theory", Diff. Eq., Vol. 10 pp. 1219-1227, 1976.
- [6] S.L. Campbell and N.J. Rose, "Singular Perturbation of Autonomous Linear Systems", SIAM J. Math. Anal., Vol. 10, No. 3, pp. 542-551, 1979.
- [7] S.L. Campbell, "Singular Perturbation of Autonomous Linear Systems, II", J. Diff. Eq., Vol. 29, pp. 362-373, 1978.
- [8] S.L. Campbell and N.J. Rose, "Singular Perturbation of Autonomous Linear Systems, III", Houston J. Math., Vol. 4, No. 4, 1978.
- [9] V.S. Korolyuk and A.F. Turbin, "Mathematical Foundations of Phase Consolidation for Complex Systems", Naukova Dumka, Kiev, 1978 (In Russian).
- [10] P.J. Curtois, "Decomposability: Queueing and Computer System Applications", Academic Press, New York, 1977.
- [11] F. Delebecque and J.P. Quadrat, "Optimal Control of Markov Chains Admitting Strong and Weak Interactions", Automatica, Vol. 17, No. 2, pp. 281-296, 1981.
- [12] F. Delebeque, "A Reduction Process for Perturbed Markov Chains", submitted to SIAM J. Appl. Math.
- [13] T. Kato, Perturbation Theory for Linear Operators, Springer Verlag, Berlin, 1966.
- [14] S.L. Campbell and C.D. Meyer, Jr., "Generalized Inverses of Linear Transformations", Pitman, London, 1979.

- [15] K. Knopp, "Theory of Functions", Dover, New York, 1945.
- [16] J.C. Willems, "Almost Invariant Subspaces; An Approach to High-Gain Feedback Design, Part I: Almost Controlled Invariant Subspaces", IEEE Transactions on Automatic Control, Vol. AC-26, pp. 235-253, 1981.
- [17] G. Vergheese and T. Kailath, "Comments and Structural Invariants and Root Loci of Linear Multivariable Systems"; International Journal of Control, Vol. 29, pp. 1077-1080, 1979.
- [18] P. M. van Dooren, P. Dewilde and J. Vandewalle, "On the Determination of the Smith MacMillan Form of a Rational Matrix from its Laurent Expansion"; IEEE Transactions on Circuits and Systems, Vol. CAS-26, pp. 180-189, 1979.
- [19] S.S. Sastry and C.A. Desoer, "Asymptotic Unbounded Root Loci-Formula and Computation", L.I.D.S. P-1065, M.I.T., Jan. 1981, to appear IEEE Trans. Aut. Control, April 1983.
- [20] M. Coderch, "Multiple Time Scale Approach to Hierarchical Aggregation of Linear Systems and Finite State Markov Process", Ph.D. Thesis, MIT Dep. of Elec. Eng. and Comp. Sci., August 1982.
- [21] F.C. Hoppensteadt, "Analysis of Some Problems having Matched Asymptotic Expansion Solutions", SIAM Review, Vol. 17 (1974) pp. 123-135.
- [22] E.B. Davies, "One-Parameter Semi-Groups", Academic Press, 1980, esp. Chapters 2,3,5.
- [23] S.L. Campbell, "Singular Systems of Differential Equations", Pitman, 1980.

$$\begin{array}{c}
 \mathbb{R}^n = \mathcal{R}(Q_m(\varepsilon)) \oplus \dots \oplus \mathcal{R}(Q_1(\varepsilon)) \oplus \mathcal{R}(Q_0(\varepsilon)) \oplus \mathcal{N}(A_0(\varepsilon)) \\
 \underbrace{\hspace{10em}}_{\mathcal{R}(A_0(\varepsilon))} \quad \underbrace{\hspace{10em}}_{\mathcal{N}(A_0(\varepsilon))} \\
 \underbrace{\hspace{10em}}_{\mathcal{R}(A_1(\varepsilon))} \quad \underbrace{\hspace{10em}}_{\mathcal{N}(A_1(\varepsilon))} \\
 \underbrace{\hspace{10em}}_{\mathcal{R}(A_2(\varepsilon))} \quad \underbrace{\hspace{10em}}_{\mathcal{N}(A_2(\varepsilon))} \\
 \vdots \qquad \qquad \qquad \vdots \\
 \vdots \qquad \qquad \qquad \vdots \\
 \vdots \qquad \qquad \qquad \vdots \\
 \underbrace{\hspace{10em}}_{\mathcal{R}(A_m(\varepsilon))} \quad \underbrace{\hspace{10em}}_{\mathcal{N}(A_m(\varepsilon))}
 \end{array}$$

Fig. 1 Geometric Content of Proposition 4.4

$$\begin{array}{ccccccc}
A_0(\varepsilon) & & A_{00} & A_{01} & A_{02} & \cdots & A_{0m} & \cdots \\
A_1(\varepsilon) & & A_{10} & A_{11} & \cdots & \cdots & A_{1,m-1} & \cdots \\
A_2(\varepsilon) & & A_{20} & \cdots & \cdots & \cdots & A_{2,m-2} & \cdots \\
\vdots & & \vdots & & & & & \\
\vdots & & \vdots & & & & & \\
A_m(\varepsilon) & & A_{m0} & \cdots & \cdots & \cdots & &
\end{array}$$

Fig. 2 Triangular array of matrices A_{ij}