LIDS-R-1182

STOCHASTIC STABILITY RESEARCH FOR COMPLEX POWER SYSTEMS

FINAL REPORT FOR DEPT. OF ENERGY CONTRACT ET-76-A-01-2295

TABLE OF CONTENTS

1. Introduction and Overview

1.1 Project Overview

In [WIL 78], we outlined a research plan aimed at the development and study of several mathematical models for the behavior of complex dynamical systems affected by random effects. The models that we proposed to study were motivated by an examination of the key qualitative features of the long-term response of an interconnected power system, in situations initiated by random events.

Among these key features were the following:

1) The system equilibrium point has a domain of attraction that is of finite extent. An excursion of the system variables outside of this domain corresponds to the loss of synchronism in one or more generators.
2) Unpredictable discrete events, e.g. failures of equipment, external events such as lightning stroke, cause some abrupt changes in the structure of the system which may reduce its security, thus making it vulnerable to stress-induced failures (see (3) below) or to other independent events.
3) The system may experience abrupt changes that depend on the system state. A piece of equipment (such as a transmission line) may fail because a particular variable (the load on the line) exceeds a limit (the capacity of the line). Also, the system is equipped with many internal discrete controls such as relays and circuit breakers that respond in an automatic fashion to certain system conditions (note that these relays may fail randomly, and such failures are of the type described in (2) above).
4) A change in system structure by any of the mechanisms described in (2) and (3) leads to a corresponding change in system dynamics and hence to a change in equilibrium conditions.
5) The change in dynamics caused by an initiating event leads to a transient response of the system during which certain variables may exceed their limits. This in turn causes some new changes in the system dynamics which lead to some new transients that can cause further changes. Such a cascade of system failures can lead to a serious loss of integrity in the power network.
6) The notion of stability for power systems is tightly connected with the notion of security. This means that it is not enough to require the system variables to be safely within their limits. The probability of contingencies that would lead to drastic changes in the mode of operation of the system should be kept as low as possible. This implies for example that the system should have some reserve margins that would enable it to withstand a reasonable level of increased demand on it.
7) The system is equipped with a number of human operators whose decisions may affect directly the outcome of emergency situations. The operator is faced with at least two major problems in carrying out his job. The first of these is the one of assimilating all the data that he receives. The information provided to the operator may contain contradictory pieces of information and superfluous details, and it may not contain the most relevant pieces of information for assessing the system status. The second problem faced by the operator is the one of deciding upon the appropriate control action, a choice which is not always clear for complex power systems.

From these features, we isolated a number of characteristics that must appear in mathematical models of power systems and in the formulation of the stability problem, in order to capture the aspects of the behavior of power systems described above.

A) The models must be nonlinear in order to capture the nonlinear dynamics of power systems and in particular, to allow for the consideration of finite extent domains of attraction and multiple equilibria.
B) The models must include a continuous part for those variables that vary continuously (e.g. machine angles and frequencies), and a discrete part to model the system structure (e.g. the topology of the power network) and the abrupt, discrete changes it can undergo. A logical structure for such a model is as a feedback interconnection of a continuous-state system, representing the power system dynamics, and a discrete-state part modeling the system structure. A class of hybrid models having precisely this structure was proposed in [WIL 78], and some of their properties were discussed in [WIL 80]. A large part of this report (see Chapters II and IV) will be devoted to the study of these models.
C) Stochastic effects must be taken into account in both the continuous and discrete models. These effects aim at describing some exogenous, random events such as lightning strikes, generator and transmission line outages, load variations, as well as some overall system variations due to unpredictable human responses throughout the system. When this feature of dynamic models of power systems is combined with the fact that

these models must be nonlinear as mentioned in (A) above, it becomes clear that these models must be subject to stochastic bifurcations. That is to say, under the effect of random events or of a cumulative number of small random changes in the conditions of the system, the model describing the power system can change suddenly from one mode of evolution to another. This phenomenon is frequently observed in the operation of power systems, and a detailed analysis of stochastic bifurcation phenomena will be givne in Chapter 3.

D) To analyze the stability of power systems, we must evaluate the security of operation of the system. To do so, we can compute the difference between some continuous variables and the limits imposed on them. In a probabilistic setting, this implies that we should evaluate containment probabilities, i.e. the probability that the state is within a certain region, or expected first passage times outside a region describing the safe operation of the system. Also, since stochastic bifurcations will occur, we want to compute the probabilities that certain chains of transitions will happen, and among such transitions, which one is the most likely.

E) The dynamics of power systems evolve on several time scales. The continuous states describing the angles and frequencies of generators vary very rapidly (on a time scale of the order of fractions of seconds) whereas the discrete states describing the topology of the network change quite slowly (every hour, or fraction of hour). This motivates the need for the development of a hierarchy of models for power systems, depending on whether we are looking at long-term, mid-term or short-term phenomena.

F) The issues described in D) and E) require the development of methods that could be used to reduce the number of variables needed to describe the system dynamics. This would be particularly useful in helping the system operator obtain a clear picture of the system status. Indeed, if the number of variables needed to describe the system conditions is reduced, the operator will have a better chance of choosing the appropriate course of action in an emergency. We will show in Chapter 2 that this objective can be accomplished by the use of aggregation methods that replace a detailed, complex model of the power system by a more approximate one containing fewer variables.

Our research has been aimed at obtaining a fundamental understanding of the qualitative behavior of mathematical models with these features. We have formulated several different classes of analytical problems that focus on specific aspects of the power system stability problem. In particular, we have devoted a large amount of attention to the impact of random effects on the stability of power systems.

1.2. Report Outline

In order to take some of the previous features of power systems into account, a class of hybrid models was introduced in [WIL 78] and [WIL 80b]. This class of models consists of the feedback interconnection of systems involving continuous state variables with discrete-state systems. However these models are sometimes quite complex, and in Chapter 2 we describe an approximation procedure that can be used to reduce the size of such models. This procedure uses a hierarchical decomposition of these models into simpler models operating at variout time scales. By using the nature of these time scales, some recursive expressions can be obtained to compute approximate, aggregated models valid at each time scale. In Chapter 2, we also analyze the behavior of hybrid models when they are subjected to rare events. To do so, we consider a hybrid model where the continuous state has several equilibrium points and behaves linearly in the domain of attraction of each equilibrium point, and where the discrete process describes the equilibrium location. Under the influence of a small stochastic driving term, the continuous state spends most of its time near equilibrium points with brief transitions to transient conditions and subsequent change to another equilbrium point. This natural separation of time scale between the discrete and continuous parts of the hybrid model is used to approximation the hybrid model by a purely discrete process.

Chapter 3 presents a framework for the study of stochastic bifurcation phenomena in power systems. Bifurcation is the study of branching in the equilibrium behavior of dynamical systems in response to small changes in the parameters of the system. Such a bifurcation study is important in the study of several aspects of the performance and control of power systems, e.g. the load flow equations, transient stability and emergency state control. For a good discussion and relevant research in this direction, see [ARA 81a], [ARA 81b], [SAS 80], [SAS 81]. Nevertheless these methods of deterministic bifurcation are extremely sensitive to the addition of small amounts of noise. Thus, in systems whose macroscopic description arises from an aggregation of microscopically fluctuating dynamics (for example, the load nodes in a power system), the predictions of deterministic bifurcation are incorrect. In the research described in Chapter 3, we seek to remedy this deficiency and develop a theory of stochastic bifurcation.

In Chapter 4, we consider the problem of analyzing and designing control strategies for systems described by hybrid models. This study is motivated by the observation that the local nature of control actions in power systems in a frequent source of instability. Many of these control actions fail to anticipate the potential impact of other disturbances in the network and result in overall instability. The work presented in Chapter 4 is aimed at understanding how contorl actions may affect the stability of the overall system and how control strategies may be designed to ensure stability. A control mechanism that achieves this objective is the one of hedging: the controller identifies regions of hazardous operation and attempts to steer the system away from such regions. The increased security that is achieved by this scheme comes at the expense of performance, but this is a price that one is often willing to pay when the catastrophic consequences of a failure of the system are taken into account.

In Chapters 5 and 6 we discuss the issue of dissipativeness for power system networks and its impact on the stability problem. This is done in

Chapter 5 by studying the thermodynamical properties of finite-state Markov processes and of hybrid models, and in Chapter 6 by characterizing stochastic dynamicla systems which are dissipative.

Finally, Chapter 7 discusses the stability problem for power systems from a deterministic point of view. The presence of relays and logical devices in power system feedback loops gives rise to the possibility of discontinuous switching behavior for the control variables of the system. This switching behavior is examined, and the stability problem is discussed in this context.

Over the three years of the research project discussed here, the following individuals have taken part in research activities:

Prof. Alan S. Willsky (Principal Investigator)
Prof. Sanjoy K. Mitter
Prof. Timothy L. Johnson
Prof. Bernard C. Levy
Prof. Shankar S. Sastry
Prof. John Wyatt
Dr. David A. Castanon
Dr. Lena Valavani
Mr. (now Prof.) Howard Chizeck
Mr. Marcel Coderch
Mr. Peter Doerschuck
Mr. Michael Propp
Mr. Roland Shomber
Dr. Mark H.A. Davis
Prof. Steven I. Marcus
Prof. Jan Willems — consultants
Prof. Wendell Fleming

This project has resulted in the Ph.D. Theses of H. Chizeck and M. Coderch and in the Master's thesis of R. Shomber. In addition, Professors S. Marcus, W. Fleming and Jan Willems, and Dr. Mark Davis have served as consultants on this project and the papers by them listed in the publications, sections of this report cover some of the results of this interaction.

2. Asymptotic Analysis and Approximation of Hybrid Systems

2.1 Introduction

As discussed in the previous chapter, the basic premise of our research effort was to develop analysis methodologies for the study of the properties of stochastic hybrid models. These hybrid models were found to be well-suited for the description of power systems because they included both discrete-state and continuous-state components.

However, one of the undesirable properties associated with such detailed modeling is that the models are so complex that exact analysis of their mathematical properties is seldom feasible. Our objective in this chapter is to develop an approximation methodology which produced simpler models of the hybrid systems that represent accurately the behavior of the exact model.

Based on the properties of these simpler models, we can determine the overall system stability and performance.

As discussed in Willsky et al [WIL 78], the loss of stability in a large power system may be modeled as a rare event, a large deviation of the system state from its nominal behavior. These events may be modeled as occurring in a time scale which is different from the normal time scale of operation of the system. Our approximation approach will try to make use of such time scale separation to develop simpler models which are accurate on specific time scales.

The first paper, contained in section 2.2, studied the asymptotic behavior of finite dimensional linear systems whose coefficients depend smoothly on a small parameter. The properties of linear systems are important for the understanding of stochastic systems, because evolution of the probability density in a stochastic system is described by a linear equation. In this paper, we provide necessary and sufficient conditions which characterize when a linear system can be approximated as a hierarchical composition of simple models operating at various time scales. We describe the nature of these time scales and provide recursive expressions which can be used to compute the approximate, aggregated models valid at each time scale. The aggregate models are shown to be consistent with the approximate models as the small parameter vanishes.

The second paper, contained in section 2.3, uses the results of section 2.2 to analyze the linear models which describe the evolution of the probability distribution of finite state Markov processes. We establish that these linear models satisfy the conditions which establish that a valid decomposition in terms of multiple time scales exists. Furthermore, we describe the resulting simplified models in terms of an aggregation operation, which produces reduced-dimension Markov process, and a slow evolution. This aggregation operation requires de analisis of stochastically discontinuous Markov processes which have not been studied in the literature.

The third paper, included in section 2.4, deals with the analysis of hybrid models subject to rare events. In this paper, the rare events are modeled by a small intensity Brownian motion affecting the behavior of a dynamical system operating in one of several stable modes. The Brownian motion occasionally drives the system to switch between equilibrium points. By properly scaling time, the process is approximated by a finite state model which only follows the changes between equilibrium points. The detailed behavior of the process inside each domain of attraction is aggregated into a single state, resulting in a finite-state description similar to the ones described in section 2.3.

The final paper is this chapter, contained in section 2.5, is an application of the methodology of sections 2.3 and 2.4 to the study of flows in some simple networks. The networks are described by hybrid models, where the continuous variable represent the levels of storage in the system, and the discrete variables describe the connectivity of the network. One of the interesting effects in this model is that of cascading saturation, observed when adjacent storages fill up. Our results provide a simplified, aggregate model which accurately approximates the behavior of the original system.

Although the work we have presented in this section is complete, there are many extensions of these results which should be investigated. The results of section 2.3 apply only to the simplest type of hybrid model, where there are no dynamics involving continuous states. The hierarchical decomposition into aggregate models, together with the algebraic conditions for existence, must be extended to the case where the hybrid models have continuous states. The results of section 2.4 represent an initial effort in this direction, serving to illustrate some of the conceptual differences between the finite-state and continuous state model. The next section contains some applications of these results towards the analysis of bifurcation phenomena in stochastic systems.

The results of section 2.5 should be extended towards the analysis of general interconnected networks. In addition, the special case of power system networks should be considered, where the conservation of flow law is amplified by voltage considerations. This remains the subject of future investigations.

In conclusion, this section describes the foundation of a theory of approximation which can be used to analyze stochastic hybrid models of power systems. These approximate models, based on isolating specific time scales, provide the basis for stability analysis and control system design.

2.2 Multiple Time Scale Behavior of Singularly Perturbed Linear Systems

2.2.1 Introduction

In this section, we discuss the problem of obtaining an accurate approximation of the evolution of the state of a system described by a linear differential equation. Our approximation methods are based on selecting, based on physical properties, such as weak couplings or slow behavior, a small parameter ε which affects the evolution of the system state, as:

$$\frac{d}{dt}\, x^{\varepsilon}(t) = A(\varepsilon)\, x^{\varepsilon}(t) \qquad (2.2.1)$$

It is the purpose of this section to characterize the behavior of equation (2.2.1) at various time scales $\tau = \varepsilon^{s}t$ for positive values of s. In

particular, our goal is to identify conditions on $A(\varepsilon)$ such that the system (2.2.1) has a well-defined approximation which is accurate at given time scales of operation, as ε approaches zero.

The problem of obtaining approximate models of ordinary differential equations based on time scale separation has been studied in the literature under the heading of singular perturbations. A comprehensive review of the literature can be found in the surveys of Nayfeh [NAY 73] and Vasil'eva [VAS 76]. Most of the previous work in this field has concentrated on developing an approximation valid only at a specific time scale, say $\tau = \varepsilon t$. Campbell [CAM 78], and Korolyuk and Turbin [KOR 78] have given algebraic conditions on $A(\varepsilon)$ which characterize the minimal integer $n \geq 1$ such that an accurate approximation exists at the time scales $\tau = \varepsilon^n t$, and provide an expression for these approximation in terms of a reduced-order model. These approximations have been used for the synthesis of two-level controllers and estimators in the literature (see [KOK 80] for a collection of papers).

Our results in this section represent a non-trivial extension of the work in [CAM 78], [KOR 78], in several directions. First, we provide algebraic conditions which characterize <u>all</u> of the possible time scales $\tau = \varepsilon^s t$ for which a reduced-order approximate model can be obtained. This extends the previous results, valid only for the smallest integer time scale. Based on this characterization, we develop a hierarchy of aggregated, approximate models which can be grouped together to obtain a <u>uniformly accurate</u> approximation to the system (2.2.1), valid for all times $t \in [o,\infty)$. Furthermore, we identify conditions where such a uniform decomposition is not possible, and provide restricted approximations in these cases.

For the purposes of conciseness, all of the results are stated without proof. The reader is referred to [COD 81], or [COD 82] for the detailed proofs of these results.

2.2.2 Regular and Singular Perturbations

In this paper, we consider semistable linear, time-invariant (LTI) systems of the form:

$$\frac{dx^\varepsilon(t)}{dt} = A(\varepsilon)\, x^\varepsilon(t) \quad , x^\varepsilon(0) = x_o \qquad (2.2.2)$$

where $\varepsilon \in [0, \varepsilon_o]$, $x^\varepsilon(t) \in \mathbb{R}^n$ and the matrix $A(\varepsilon)$ is assumed to have a power series expansion in ε, i.e.,

$$A(\varepsilon) = \sum_{p=o}^{\infty} \varepsilon^p A_p \qquad (2.2.3)$$

This series can be either a convergent series (and we will refer to this as the analytic case), or an asymptotic series. If (2.2.3) is an asymptotic series, then we will also assume that rank $A(\varepsilon)$ is constant for $\varepsilon \in (0,\varepsilon_o]$. In both cases we will refer to this constant as the <u>normal rank</u> of $A(\varepsilon)$ which we will denote by nrank.

Our objective is to analyze the behavior of $x^\varepsilon(t)$ as $\varepsilon \downarrow 0$ on the time interval $[0,\infty)$. The system (2.2.2) can be viewed as a perturbation of

$$\frac{dx^o(t)}{dt} = A_o x^o(t) \qquad x^o(0) = x_o \qquad (2.2.4)$$

and one of the obvious questions to be addressed is that of the relationship between $x^\varepsilon(t)$ and $x^o(t)$ for small ε. Specifically, under what conditions is $x^o(t)$ a good approximation of $x^\varepsilon(t)$? If it is not, how can we construct such an approximation? The first question is resolved in this subsection and the rest of the section deals with the second question.

The following proposition states that $x^o(t)$ is a uniform asymptotic approximation of $x^\varepsilon(t)$ on any compact time interval $[0,T]$.

<u>Proposition 2.2.1</u>

If $A(\varepsilon)$ is as in (2.2.3) then

$$\lim_{\varepsilon \downarrow 0} \sup_{t \in [0,T]} || \exp\{A(\varepsilon)t\} - \exp\{A_o t\} || = 0$$

for any $T < \infty$.

In general, however, as the next example shows, it is not true that

$$\lim_{\varepsilon \downarrow 0} \sup_{t \geq o} ||\exp\{A(\varepsilon)t\} - \exp\{A_o t\}|| = 0 \qquad (2.2.5)$$

and therefore, in general, $x^o(t)$ is not a good approximation to $x^\varepsilon(t)$ for <u>all times t</u>, no matter how small ε is.

<u>Example 2.2.2</u>

Let $A(\varepsilon)$ be

$$A(\varepsilon) = \begin{bmatrix} -\varepsilon & o \\ o & o \end{bmatrix}$$

then we have:

$$\exp\{A(\varepsilon)t\} = \begin{bmatrix} e^{-\varepsilon t} & o \\ o & 1 \end{bmatrix}, \quad \exp\{A_o t\} = I$$

and therefore

$$\sup_{t \geq o} || \exp\{A(\varepsilon)t\} - \exp\{A_o t\} || = 1 \qquad \square$$

If eq. (2.2.5) is satisfied, we will say that (2.2.2) is a <u>regularly perturbed</u> version of (2.2.4) - otherwise, we will say that is is <u>singularly perturbed</u>. In what follows we will deal primarily with singularly perturbed systems because, as we will now see, it is the only case where we can talk about different behavior at different time scales. Let us formalize the notion of multiple time scale behavior.

<u>Definition 2.2.3</u>

Let $x^\varepsilon(t)$ be the solution of (2.2.2) and let $\alpha(\varepsilon)$ be an order function. We will say that $x^\varepsilon(t)$ has a <u>well defined behavior at time scale $t/\alpha(\varepsilon)$</u> if there exists a bounded continuous function $y(t)$, called the evolution of $x^\varepsilon(t)$ at this time scale, such that:

$$\lim_{\varepsilon \downarrow 0} \sup_{t \in [\delta,T]} ||x^\varepsilon(t/\alpha(\varepsilon)) - y(t)|| = 0 \qquad (2.2.6)$$

for any $\delta > o$, $T < \infty$, x_o.

Equivalently, we will say that the LTI system (2.2.2) has a well defined behavior at time scale $t/\alpha(\varepsilon)$ if there exists a bounded continuous matrix $Y(t)$ such that:

$$\lim_{\varepsilon \downarrow 0} \sup_{t \in [\delta, T]} \| \exp\{A(\varepsilon)\, t/\alpha(\varepsilon)\} - Y(t) \| = 0 \qquad (2.2.7)$$

for any $\delta > 0$, $T < \infty$ □

According to this definition, the system in example 2.2.3 has a well defined behavior at time scale t/ε and its evolution at this time scale is given by:

$$y(t) = \begin{bmatrix} e^{-t} & 0 \\ 0 & 1 \end{bmatrix} x_o$$

Although in this example the convergence condition (2.2.6) is satisfied even for $\delta = 0$ it will become clear later on that, in general, an arbitrarily small interval around zero must be excluded to obtain uniform convergence.

The next proposition shows that semistable regularly perturbed systems have trivial and uninteresting time scale behavior.

Proposition 2.2.4

If $A(\varepsilon)$ is a regularly perturbed matrix which is semistable for $\varepsilon \in [0, \varepsilon_o]$ then, for any order function $\alpha(\varepsilon)$,

$$\lim_{\varepsilon \downarrow 0} \sup_{t \in [\delta, T]} \| \exp\{A(\varepsilon)t/\alpha(\varepsilon)\} - P_o \| = 0 \qquad (2.2.8)$$

for all $\delta > 0$ $t < \infty$

where P_o is the eigenprojection for the zero eigenvalue of A_o, as defined in appendix 2.A. □

It follows from the above proposition that, if properly modeled, a system with a non-trivial multiple time scale behavior will result in a singularly perturbed equation. The next proposition gives necessary and sufficient conditions on $A(\varepsilon)$ for (2.2.2) to be singularly perturbed.

Proposition 2.2.5

The equation (2.2.2), with $A(\varepsilon)$ semistable for $\varepsilon \in [0, \varepsilon_o]$ is a singularly perturbed differential equation if and only if rank $A(\varepsilon)$ is discontinuous at $\varepsilon = 0$. □

As a consequence of the above proposition notice that if A_o is asymptotically stable then any perturbation is regular. In effect, if A_o is asymptotically stable so is $A(\varepsilon)$ for ε small enough and therefore null A_o = null $A(\varepsilon) = 0$ for $\varepsilon \in [0, \varepsilon_o]$.

We have so far established that the analysis of semistable LTI systems with multiple time scale behavior corresponds to the study of singularly perturbed ordinary differential equations (o.d.e.) and that the presence of weak couplings may produce well defined behavior at several time scales only if the perturbation changes the rank of the system matrix $A(\varepsilon)$. To keep this discussion clear we have only considered systems for which $A(\varepsilon)$ is semistable for $\varepsilon \in [0, \varepsilon_o]$. As we will see in the following results, this is a necessary (although not sufficient condition for the system to have well defined behavior at all time scales. Extensions of the results derived for semistable systems to some classes of non-semistable systems are considered in section 2.2.4.

There is an explicit connection between the time scale evolution of the system and the eigenvalues of the perturbed matrix $A(\varepsilon)$. Consider the system (2.2.2), assumed semistable for ε in $[0, \varepsilon_o]$. Assume also, for simplicity, that $A(\varepsilon)$ is diagonalizable and let $\lambda_n(\varepsilon)$, $n = 0,1,\ldots,s$ be its distinct eigenvalues. Using the spectral representation of $\exp\{A(\varepsilon)\}t$ we get:

$$\exp\{A(\varepsilon)t\} = \sum_{n=0}^{s} e^{\lambda_n(\varepsilon)t} P_n(\varepsilon) \qquad (2.2.9)$$

where $P_n(\varepsilon)$ is the eigenprojection for the eigenvalue $\lambda_n(\varepsilon)$ of $A(\varepsilon)$. It is clear from (2.2.9) and the semistability assumption that for $\exp\{A(\varepsilon)t\}$ to have a non-trivial, well defined behavior at time scale $t/\alpha(\varepsilon)$, it is necessary that there exists some eigenvalue $\lambda_n(\varepsilon)$ such that $\lambda_n(\varepsilon)/\alpha(\varepsilon)$ approach a finite value μ_n and $\varepsilon \downarrow 0$. Thus, for a system to have multiple time scale behavior in the sense of definition 2.2.3, it must have eigenvalues of different orders of magnitude in ε. Our formulation is therefore in accordance with the notion of time scales as a manifestation of eigenvalue separation.

This point of view provides some insight into our results. First, the existence of eigenvalues of $A(\varepsilon)$ that converge to zero as $\varepsilon \downarrow 0$ implies that zero itself must be an eigenvalue of A_o and that rank $A(\varepsilon)$ must be discontinuous at zero as stated by propostion 2.2.5. Second, the eigenvalues of $A(\varepsilon)$ have always a power series expansion in fractional power of ε. Therefore, it is logical to conclude that the fundamental time scales of (2.2.9) must be sought as t/ε^q, for some rational q, and that only a finite number of them can exist.

In addition to the eigenvalue structure, the existence of the limit of $\exp\{A(\varepsilon)t/\alpha(\varepsilon)\}$ as $\varepsilon \downarrow 0$ clearly depends on the eigenspaces structure, i.e., on the behavior as $\varepsilon \downarrow 0$ of $P_n(\varepsilon)$ in (2.2.9). For example, the eigenprojections $P_n(\varepsilon)$ (and also the eigennilpotents in the general non-diagonable case) have algebraic singularities at $\varepsilon = 0$ if $\lambda_n(\varepsilon) \to 0$ (see [KAT 66]) and therefore the above limit may not exist even if there are eigenvalues of order $\alpha(\varepsilon)$. It is this aspect of the time scale problem that is overlooked in the heuristic view of time scales as eigenvalue separation and onto which we will focus our attention in the following subsections.

2.2.3 The Multiple Semistability Case: Complete Time Scale Decomposition.

Let $A_o(\varepsilon)$, $\varepsilon \in [0, \varepsilon_o]$ be a semistable matrix with a series expansion (convergent or asymptotic) of the form:

$$A_o(\varepsilon) = \sum_{p=o}^{\infty} \varepsilon^p A_{op} \qquad (2.2.10)$$

and normal rank d. For our development we first need to construct a sequence of matrices $A_k(\varepsilon)$, $R = 1,\ldots,m$, obtained recursively from $A_o(\varepsilon)$ as indicated below.

Let $P_o(\varepsilon)$ denote the total projection for the zero-group of eigenvalues of $A_o(\varepsilon)$. It follows from Appendix 2A that, if A_{oo} has semisimple null structure (SSNS), the matrix

$$A_1(\varepsilon) = \frac{P_o(\varepsilon)A_o(\varepsilon)}{\varepsilon} = \frac{A_o(\varepsilon)P_o(\varepsilon)}{\varepsilon} = \frac{P_o(\varepsilon)A_o(\varepsilon)P_o(\varepsilon)}{\varepsilon} \qquad (2.2.11)$$

has a series expansion of the form:

$$A_1(\varepsilon) = \sum_{p=o}^{\infty} \varepsilon^p A_{1p} \qquad (2.2.12)$$

If the resulting first order term in (2.2.11), A_{10}, also has SSNS we define the next matrix $A_2(\varepsilon)$ as in (2.2.11),i.e.,

$$A_2(\varepsilon) \triangleq \frac{P_1(\varepsilon)A_1(\varepsilon)}{\varepsilon} = \frac{P_1(\varepsilon)P_o(\varepsilon)A_o(\varepsilon)}{\varepsilon^2}$$

$$= \sum_{p=o}^{\infty} \varepsilon^p A_{2p} \qquad (2.2.13)$$

where $P_1(\varepsilon)$ is the total projection for the zero-group of eigenvalues of $A_1(\varepsilon)$. The recursion ends at

$$A_m(\varepsilon) \triangleq \frac{P_{m-1}(\varepsilon)A_{m-1}(\varepsilon)}{\varepsilon} = \frac{P_{m-1}(\varepsilon)P_{m-2}(\varepsilon)\ldots P_o(\varepsilon)A_o(\varepsilon)}{\varepsilon^m}$$

$$= \sum_{p=o}^{\infty} \varepsilon^p A_{mp} \qquad (2.2.14)$$

if the matrix A_{mo} does not have SSNS.

The following proposition establishes several basic properties of the matrices $A_k(\varepsilon)$ and $P_k(\varepsilon)$.

Proposition 2.2.6

Let $A_k(\varepsilon)$, $k=0,1,\ldots,m$, be the sequence of matrices constructed above and let $P_k(\varepsilon)$, $k=0,1,\ldots,m$, be the corresponding total projections for the zero-groups of eigenvalues. Denote $Q_k(\varepsilon) = I-P_k(\varepsilon)$. Then, for ε small enough, we have:

i) $P_i(\varepsilon)P_j(\varepsilon) = P_j(\varepsilon)P_i(\varepsilon) \quad i,j=0,1,\ldots,m$

ii) $Q_i(\varepsilon)Q_j(\varepsilon) = 0 \quad i = j, \quad i,j=0,1,\ldots,m$

iii) $\mathbb{R}^n = \mathcal{R}(Q_o(\varepsilon)) \oplus \ldots \oplus \mathcal{R}(Q_k(\varepsilon)) \oplus \mathcal{R}(P_o(\varepsilon)\ldots P_k(\varepsilon)) \quad k = 0,1,\ldots,m$

iv) $\text{rank } Q_k(\varepsilon) = \text{rank } A_{k,o}$

v) $Q_k(\varepsilon)A_o(\varepsilon) = Q_R(\varepsilon)\varepsilon^k A_k(\varepsilon) = \varepsilon^k A_k(\varepsilon)Q_k(\varepsilon) = A_o(\varepsilon)Q_k(\varepsilon) \quad k = 0,1,\ldots,m$

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

Proposition 2.2.7

Let $A_k(\varepsilon)$, $k = 0,1,\ldots$ be the sequence of matrices defined by (2.2.11-2.2.14). One of the following two conditions (or both) occur at some finite m:

i) $A_{m,o}$ does not have SSNS

ii) $A_{m+1}(\varepsilon) = 0$ (~ 0 if (2.2.10) is asymptotic).

In the analytic case, ii) is equivalent to:

ii) $\sum_{k=o}^{m} \text{rank } A_{k,o} = d$ □

We will say that a matrix $A_o(\varepsilon)$ satisfies the multiple semisimple nullstructure (MSSNS) condition if the sequence of matrices $A_k(\varepsilon)$ can be constructed up to a stage $k = m$ for which condition (ii') of Proposition 2.2.7 is satisfied with all matrices having SSNS. If $A_o(\varepsilon)$ satisfies the MSSNS condition and in addition all matrices $A_{k,o}$, $k,=0,1,..,m$, are semistable we will say that $A_o(\varepsilon)$ satisfied the multiple semistability (MSST) condition. Although we will be interested in matrices which satisfy the MSST condition, all of the results developed in this subsection hold for the less restrictive MSSNS condition.

$$A_{k,o} = \lim_{\varepsilon \downarrow o} \frac{P_{k-1}(\varepsilon)\ldots P_o(\varepsilon)A_o(\varepsilon)}{\varepsilon^k} \quad k=0,1,\ldots,m$$

The following proposition provides some insight into the structure of the matrices $A_k(\varepsilon)$, $k=1,\ldots,m$, and into how they relate to $A_o(\varepsilon)$.

Proposition 2.2.8

If $A_o(\varepsilon)$ satisfies the MSSNS condition then, for some $\varepsilon_1 < 0$,

i) $A_k(\varepsilon)$, $k=0,1,\ldots,m$ have SSNS for $\varepsilon \in [0,\varepsilon_1]$.

ii) For $\varepsilon \in (0,\varepsilon_1]$,

$\mathcal{R}(A_k(\varepsilon)) = \mathcal{R}(Q_k(\varepsilon)) \oplus \ldots \oplus \mathcal{R}(Q_m(\varepsilon)) \quad k=0,\ldots,m$

$\mathcal{N}(A_k(\varepsilon)) = \mathcal{R}(Q_o(\varepsilon)) \oplus \ldots \oplus \mathcal{R}(Q_{R-1}(\varepsilon)) + \mathcal{N}(A_o(\varepsilon))$

$\mathcal{N}(A_o(\varepsilon)) = \mathcal{R}(P_o(\varepsilon)\ldots P_m(\varepsilon)) \quad k=1,\ldots,m,$

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

It follows from Propositions 2.2.6 and 2.2.8 that, if $A_o(\varepsilon)$ has MSSNS, the direct sum

$$\mathbb{R}^n = \mathcal{R}(Q_o(\varepsilon)) \oplus \ldots \oplus \mathcal{R}(Q_m(\varepsilon)) \oplus \mathcal{R}(P_o(\varepsilon)\ldots P_m(\varepsilon)) \qquad (2.2.15)$$

decomposes $A_o(\varepsilon)$ as follows:

$$A_o(\varepsilon) = \hat{A}_o(\varepsilon) \oplus \varepsilon\hat{A}_1(\varepsilon) \oplus \ldots \oplus \varepsilon^m\hat{A}_m(\varepsilon) \oplus 0, \qquad (2.2.16)$$

where $\hat{A}_k(\varepsilon)$ denotes the restriction of $A_k(\varepsilon)$ on $\mathcal{R}(Q_k(\varepsilon))$. Said in another way,

$$A_o(\varepsilon) = \sum_{k=o}^{m} \varepsilon^k A_k(\varepsilon) Q_k(\varepsilon) \qquad (2.2.17)$$

The eigenvalues of $A_o(\varepsilon)$ can thus be divided into (m+1) groups corresponding to the eigenvalues of $A_o(\varepsilon)$ in each of the invariant subspaces $\mathcal{R}(Q_k(\varepsilon))$. Each eigenvalue of $A_o(\varepsilon)$ is of order ε^j for some integer $j \geq o$ and the eigenvalues of order ε^k coincide with ε^k times the eigenvalues of order one of $A_k(\varepsilon)$. Figure 2.2.1 illustrates the structure of the matrix $A_k(\varepsilon)$; its null space includes, in addition to the null space of $A_o(\varepsilon)$, the eigenspaces of $A_o(\varepsilon)$ corresponding to all eigenvalues of order $O(1)$, $O(\varepsilon),\ldots,O(\varepsilon^{k-1})$ while its range includes the eigenspaces of $A_o(\varepsilon)$ for all eigenvalues of order $o(\varepsilon^{k-1})$. The construction of the sequence $A_k(\varepsilon)$ can thus be viewed as a way to separate the eigenvalues of $A_o(\varepsilon)$ in different groups according to their asymptotic order as $\varepsilon \downarrow 0$. The actual calculations required to compute the matrices $A_k(\varepsilon)$ will be discussed later.

The following theorem illustrates the consequences of the MSSNS condition for the multiple time scale behavior of $\exp\{A_o(\varepsilon)\}t$.

Theorem 2.2.9

If $A_o(\varepsilon)$ satisfies the MSSNS condition then:

$$\exp\{A_o(\varepsilon)t\} = \sum_{k=o}^{m} Q_k(\varepsilon)\exp\{A_k(\varepsilon)\varepsilon^k t\} + P_o(\varepsilon)\ldots P_m(\varepsilon)$$

$$= \sum_{k=o}^{m} \exp\{Q_k(\varepsilon)A_k(\varepsilon)\varepsilon^k t\} - mI$$

$$= \prod_{k=o}^{m} \exp\{Q_k(\varepsilon)A_k(\varepsilon)\varepsilon^k t\}$$

$$= \exp\left\{\sum_{k=o}^{m} Q_k(\varepsilon)A_k(\varepsilon)\varepsilon^k t\right\}$$

Theorem 2.2.9 corresponds to the splitting of the evolution $\exp\{A_o(\varepsilon)t\}$ according to the direct sum decomposition of (2.2.15). Under the condition of MSSNS, this splitting corresponds also to a decomposition into parts of $\exp\{A_o(\varepsilon)t\}$ that evolve at different time scales. For example, $Q_k(\varepsilon)\exp\{A_k(\varepsilon)\varepsilon^k t\}$ does not change significantly until t is of order $1/\varepsilon^k$. Theorem 2.2.9 thus gives a consistent spatial and temporal decomposition of eq. (2.2.2), which is very convenient to study the multiple time scale behavior and also to derive uniform asymptotic approximations of $\exp\{A_o(\varepsilon)t\}$.

As we have proved in Prop. 2.2.1, $\exp\{A(o)t\}$ is a uniform asymptotic approximation to $\exp\{A(\varepsilon)t\}$ on any compact time interval $[0,T]$. It is quite clear, however, that this approximation does not capture the multiple time scale behavior of a singularly perturbed system. To construct an approximation which captures this behavior, we have to require it to be uniformly valid over the infinite time interval $[0,\infty)$.

The next theorem gives the desired approximation under the assumption that $A_o(\varepsilon)$ satisifed the MSST condition.

Theorem 2.2.10

Let $A_o(\varepsilon)$ satisfy the MSST condition and let $A_k(\varepsilon)$, $P_k(\varepsilon)$ and $Q_k(\varepsilon)$, $k = 0,1,..,m$, be the sequences of matrices constructed in 2.2.14. Then,

$$\lim_{\varepsilon\downarrow 0} \sup_{t\geq o} \|\exp\{A_o(\varepsilon)t\} - \phi(t,\varepsilon)\| = 0 \qquad (2.2.18)$$

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

$$\phi(t,\varepsilon) = \sum_{k=o}^{m} Q_R \exp\{1=A_{ko}\varepsilon^k t\} + P_o\ldots P_m$$

$$= \sum_{k=o}^{m} \exp\{A_{k\,o}\,\varepsilon^k t\} - mI$$

$$= \prod_{k=o}^{m} \exp\{A_{k\,o}\,\varepsilon^k t\}$$

$$= \exp\left\{\sum_{k=o}^{m} A_{k\,o}\,\varepsilon^k t\right\}$$

with $A_{k\,o} = \lim_{\varepsilon\downarrow o} A_k(\varepsilon)$, $P_k = \lim_{\varepsilon\downarrow 0} P_k(\varepsilon)$ and $Q_k = \lim_{\varepsilon\downarrow 0} Q_k(\varepsilon)$.

Furthermore,

$$\mathbb{R}^n = R(A_{00}) \oplus \ldots \oplus R(A_{m,o}) \oplus \left(\bigcap_{k=o}^{m} N(A_{k,o})\right) \qquad (2.2.19)$$

As the above theorem shows, the sequence of matrices $A_{k,o}$, $k=0,1,\ldots,m$, completely determines an asymptotic approximation to $\exp\{A_o(\varepsilon)t\}$ which captures its multiple time scale behavior. We now use Theorem 2.2.10 to show that systems which satisfy the MSST condition have well defined behavior at all time scales and that the matrices $A_{k,o}$ determine a set of reduced-order models of the system.

The following corollary gives an explicit formula for the evolutions of $\exp\{A_o(\varepsilon)t\}$.

Corollary 2.2.11

Let $A_o(\varepsilon)$ satisfy the MSST condition and let $A_{k,o}$ P_k and Q_k, $k=0,1,\ldots,m$, be the sequence of matrices specified in theorem 2.2.10. Then,

i) $$\lim_{\varepsilon\downarrow 0} \sup_{\delta<t\leq T} \|\exp\{A_o(\varepsilon)t/\varepsilon^k\} - \Phi_k(t)\| = 0$$

$\forall\delta > 0$, $\forall T < \infty$

$k=0,1,\ldots, m-1$

ii) $$\lim_{\varepsilon\downarrow 0} \sup_{\delta<t\leq\infty} \|\exp\{A_o(\varepsilon)t/\varepsilon^m\} - \Phi_m(t)\| = 0$$

$\forall\delta > 0$

where $\Phi_k(t)$ is any of the following expressions:

$$\Phi_k(t) = Q_k \exp\{A_{k,o}t\} + P_o\ldots P_k$$

$$= P_o\ldots P_{k-1} \exp\{A_{k,o}t\} \qquad \square$$

It is now immediate that the evolutions of $\exp\{A_o(\varepsilon)t\}$ at time scales t/ε^k, $\Phi_k(t)$, $k=0,1,..,m$ can be combined to produce a uniform asymptotic approximation to $\exp\{A_o(\varepsilon)t\}$ as follows:

$$\exp\{A_o(\varepsilon)t\} = \sum_{k=o}^{m} \Phi_k(\varepsilon^k t) - \sum_{k=o}^{m-1} P_o P_1\ldots P_k + o(1) \qquad (2.2.20)$$

This equation shows that only the behavior at time scales t/ε^k, $k=0,1,\ldots m$, is needed to capture the main features of the evolution of $\exp\{A_o(\varepsilon)t\}$ over the infinite time interval $[0,\infty)$. It is clear from Theorem 2.2.10, however, that the limit

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

exists for any order function $\alpha(\varepsilon)$. Indeed, if $\alpha_k(\varepsilon) = o(\varepsilon^k)$ and $\varepsilon^{k+1} = o(\alpha_k(\varepsilon))$, $k=0,1,\ldots,m-1$, then

$$\lim_{\varepsilon\downarrow o} \exp\{A_o(\varepsilon)t/\alpha_k(\varepsilon)\} = P_o\ldots P_k$$

and for $\alpha(\varepsilon) = o(\varepsilon^m)$,

$$\lim_{\varepsilon\downarrow o} \exp\{A_o(\varepsilon)t/\alpha(\varepsilon)\} = P_o\ldots P_m$$

Thus the system has well defined behavior at all time scales even though only a finite number of them, that we will call its fundamental(or natural)

time scales, are required to capture the main features of the system's evolution. We now show that a reduced-order model of (2.2.20) can be associated with each fundamental time scale. To interpret the matrices A_{ko} as reduced order models of:

$$\frac{dx^{\varepsilon}(t)}{dt} = A_o(\varepsilon)\, x^{\varepsilon}(t) \qquad x^{\varepsilon}(o) = x_o \qquad (2.2.21)$$

valid at different time scales, notice that the asymptotic approximation

$$\exp\{A_o(\varepsilon)t\} = \sum_{k=o}^{m} Q_k \exp\{A_{k,o}\varepsilon^k t\} + P_o \ldots P_m + o(1) \qquad (2.2.22)$$

and the direct sum decomposition:

$$\mathbb{R}^n = R(Q_o) \oplus \ldots \oplus R(Q_m) \oplus R(P_o \ldots P_m)$$

imply that if $x^{\varepsilon}(0) \in R(Q_k)$, then $x^{\varepsilon}(t)$ remains in $R(Q_k)$ for all $t > o$ except for terms which are uniformly negligible as $\varepsilon \downarrow 0$. Thus $R(Q_k)$, $k=0,1,,,m$ and $R(P_o \ldots P_m)$ can be thought of as almost invariant subspaces of the system (2.2.21) [WILL 80]. Furthermore, the parts of $x^{\varepsilon}(t)$ that evolve in different subspaces do so at different time scales. To describe the part of $x^{\varepsilon}(t)$ that evolves at time scale t/ε^k to first order approximation, the following model can be used:

$$\frac{dy_k(t)}{dt} = A_{k,o}\, y_k(t) \qquad k=0,1,..,m \qquad (2.2.23)$$

If $y_k(o) = Q_k x_o$ then

$$y_k(\varepsilon^k t) = Q_k\, x^{\varepsilon}(t) + o(1) \qquad k=0,1,..,m \qquad (2.2.24)$$

uniformly for $t \geq o$, and once again a uniform approximation of $x^{\varepsilon}(t)$ can be constructed by combining the solutions of the reduced-order models (2.2.23) as follows:

$$x^{\varepsilon}(t) = \sum_{k=o}^{m} y_k(\varepsilon^k t) + P_o \ldots P_m x_o \qquad (2.2.25)$$

Notice also that

$$\sum_{k=o}^{m} \operatorname{rank} A_{k,o} = \operatorname{rank} A_o(\varepsilon) \qquad (2.2.26)$$

and therefore the combined dimensionality of the reduced-order models (2.2.23) equals the dimension of the original model.

This decomposition of (2.2.21) into a set of reduced-order models is more easily visualized using an appropriate change of basis. From Theorem 2.2.10 we have:

$$\mathbb{R}^n = R(A_{o,o}) \oplus \ldots \oplus R(A_{m,o}) \oplus \left(\bigcap_{k=o}^{m} N(A_{k,o})\right) \qquad (2.2.27)$$

and by the SSNS property of the matrices $A_{k,o}$ it follows that:

$$N(A_{k,o}) = R(A_{o,o}) \oplus \ldots \oplus R(A_{k-1,o}) \oplus R(A_{k+1,o}) \oplus \ldots \oplus R(A_{m,o}) \oplus \left(\bigcap_{k=o}^{m} N(A_{k,o})\right)$$

If we now choose a basis adapted to (2.2.27), the matrix $A_{k,o}$ will have, in this new basis, a block diagonal form with only one non-zero block. That is, if T denotes the change of basis matrix, then

$$T A_{k,o} T^{-1} = \operatorname{diag}\{0,0,\ldots,0,\ \tilde{A}_k,\ 0,\ldots,0\}$$

where $\tilde{A}_k$ is a full rank square matirx of dimension equal to rank $A_{k,o}$. Using this change of basis we can write the result of Theorem 2.2.10 as follows

$$\exp\{A_o(\varepsilon)t\} = T^{-1} \exp\Big\{\sum_{k=o}^{m} T A_{k,o} T^{-1} \varepsilon^k t\Big\} T + o(1)$$

$$= T^{-1} \operatorname{diag}\{e^{\tilde{A}_o t},\ e^{\tilde{A}_1 \varepsilon t},\ldots,e^{\tilde{A}_m \varepsilon^m t},\ I\}\ T + o(1) \qquad (2.2.28)$$

showing that, to first order approximation, the system (4.53) can be thought of composed of (m+1) uncoupled subsystems

$$\frac{d\tilde{y}_k(t)}{dt} = \tilde{A}_k \tilde{y}_k(t) \qquad k=0,1,\ldots,m$$

each running at a different time scale.

If the MSST condition is violated, then at least for some time scale. $t/\alpha(\varepsilon)$, the limit

$$\lim_{\varepsilon \downarrow o} \exp\{A_o(\varepsilon)t/\alpha(\varepsilon)\}$$

does not exist. In this case a complete time scale decomposition of the type developed in the previous propositions is not possible. Some partial extensions to systems that violate the MSST condition are, possible and will be considered later.

Two examples will shed some light as to what happens when the MSST condition is violated:

Example 2.2.12

Consider the matrix

$$A_o(\varepsilon) = \begin{bmatrix} \varepsilon & 0 & -2\varepsilon \\ \varepsilon & \varepsilon & -2\varepsilon \\ 1 & 1 & -2 \end{bmatrix}$$

It is semistable for $\varepsilon \in [0,1]$ and it has three real eigenvalues

$$\lambda_o = 0,\ \lambda_1 = -2 + o(1),\ \lambda_2 = -\varepsilon^2 + o(\varepsilon^2)$$

The matrices $A_{o,o}$ and $A_{1,o}$ (see section 3.4.5 for an algorithm to compute them) are given by:

$$A_{o,o} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & -2 \end{bmatrix}, \quad A_{1,o} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & -1/2 & 0 \end{bmatrix}$$

and the MSST condition is violated because $A_{1,0}$ does not have SSNS (it is nilpotent).

A direct computation of $\exp\{A_o(\varepsilon)t\}$ gives

$$\exp\{A_o(\varepsilon)t\} = \frac{1}{\lambda_2-\lambda_1}\left\{\right.$$

$$\begin{bmatrix} \frac{(\lambda_2-\lambda_1)}{2} & -\frac{\varepsilon}{\lambda_1}(e^{\lambda_1 t}-1) & \frac{\varepsilon(\varepsilon-\lambda_1)}{\lambda_1}(e^{\lambda_1 t}-1) \\ 0 & (\varepsilon-\lambda_1-2)e^{\lambda_1 t} & 2\varepsilon e^{\lambda_1 t} \\ 0 & -e^{\lambda_1 t} & (\varepsilon-\lambda_1)e^{\lambda_1 t} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\lambda_2-\lambda_1}{2} & \frac{\varepsilon}{\lambda_2}(e^{\lambda_2 t}-1) & \frac{\varepsilon(\lambda_2-\varepsilon)}{\lambda}(e^{\lambda_2 t}-1) \\ 0 & (\lambda_2-\varepsilon+2)e^{\lambda_2 t} & -2\varepsilon e^{\lambda_2 t} \\ 0 & e^{\lambda_2 t} & (\lambda_2-\varepsilon)e^{\lambda_2 t} \end{bmatrix}\left.\right\}$$

and we have the following time scale behavior:

$$\lim_{\varepsilon\downarrow 0} \exp\{A_o(\varepsilon)t\} =$$

$$\exp\{A_{o,o}t\} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & (1-e^{-2t})/2 & e^{-2t} \end{bmatrix}$$

$$\lim_{\varepsilon\downarrow 0} \exp\{A_o(\varepsilon)t/\varepsilon\} =$$

$$P_o \exp\{A_{o,1}t\} = \begin{bmatrix} 1 & t/2 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 0 \end{bmatrix},$$

To see that the limit

$$\lim_{\varepsilon\downarrow 0} \exp\{A_o(\varepsilon)t/\varepsilon^2\}$$

does not exist, let us look at the entry (1,2) of $\exp\{A_o(\varepsilon)t/\varepsilon^2\}$:

$$(\exp\{A_o(\varepsilon)t/\varepsilon^2\})_{1,2} = [\frac{\varepsilon}{\lambda_2}(e^{\lambda_2 t/\varepsilon^2}-1) - \frac{\varepsilon}{\lambda_1}(e^{\lambda_1 t/\varepsilon^2}-1)]\frac{1}{\lambda_2-\lambda_1}$$

Because $\lambda_2 = -\varepsilon^2 + o(\varepsilon^2)$ the first term is of order $1/\varepsilon^2$ as $\varepsilon\downarrow 0$ and therefore it diverges. Thus, the o.d.e. does not have well defined behavior at time scale t/ε^2 even though it has a real negative eigenvalue of order ε^2. This is so because

$$\| \exp\{A_o(\varepsilon)t/\varepsilon^2\} \| \to \infty \quad \text{as } \varepsilon\downarrow 0$$

This behavior does not contradict the stability properties of $A_o(\varepsilon)$ because even though for every $\varepsilon\in[0,1]$

$$\sup_{t\geq o} \|\exp\{A_o(\varepsilon)t\}\| = K(\varepsilon) < \infty \quad \forall t \geq o$$

the bound $K(\varepsilon) \to \infty$ as $\varepsilon\downarrow 0$.

This example illustrates one reason why even systems which are semistable for $\varepsilon\in[0,\varepsilon]$ may fail to have well defined behavior at some time scales. The next example illustrates another such reason. □

Example 2.2.13

Consider the matrix

$$A_o(\varepsilon) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -\varepsilon^2 & \varepsilon \\ 0 & -\varepsilon & \varepsilon^2 \end{bmatrix}$$

It is semistable for $\varepsilon \geq o$ and it has three eigenvalues

$$\lambda_o = -2, \qquad \lambda_{1,2} = -\varepsilon^2 \pm i\varepsilon$$

The matrices $A_{o,o}$ and $A_{1,0}$ for this example are:

$$A_{o,o} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad A_{1,0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

and the MSST assumption is violated because $A_{1,0}$ has purely imaginary eigenvalues.

A simple calculation gives:

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

Clearly the system has well defined behavior at time scales t and t/ε but $\exp\{A_o(\varepsilon)t/\varepsilon^2\}$ does not have a limit as $\varepsilon\downarrow 0$ because of the presence of terms of the type $e^{-t}\sin t/\varepsilon$.

This example illustrates that the existence of slowly attenuated oscillation (reflected as purely imaginary poles in one of the matrices $A_{k,o}$) impedes the existence of well defined behavior at some time scale.

The next theorem established the fact that for an arbitrary $A_o(\varepsilon)$, MSST is a necessary condition for $\exp\{A_o(\varepsilon)t\}$ to have well defined behavior at all time scales.

Theorem 2.2.14

Let $A_o(\varepsilon)$, $\varepsilon\in[0,\varepsilon_o]$, be a stable matrix with a series expansion in powers of ε and let $A_{k,o}$, $k \geq 0$, be the sequence of matrices contructed in (2.2.10) - (2.2.14). If $A_{o,o}$, $A_{1,o},\ldots,A_{\ell-1,o}$ are semistable but $A_{\ell,o}$ is not, then the limit

$$\lim_{\varepsilon\downarrow 0} \exp\{A_o(\varepsilon)t/\varepsilon^q\} \qquad t > 0$$

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

$$\lim_{\varepsilon \downarrow 0} \sup_{t \geq o} ||\exp\{A_o(\varepsilon)t\}|| = \infty$$

These results indicate that multiple semistability is a necessary and sufficient condition for a system to have well-defined multiple time scale behavior. The behavior of the system at different time scales is then determined by the matrices A_{ko}. We now focus our attention on ways of computing these matrices.

It is convenient to think of the sequence of matrices $A_k(\varepsilon)$, k=0,1,..,m, constructed in (2.2.10)-(2.2.14) as defining a rectangular array of matrices A_{ij}, i=0,1,..,m, j ≥ o, as shown in Fig. 2.2.2. By definition of the matrices $A_k(\varepsilon)$, it follows from Appendix 2.A that the $(i+1)^{\underline{th}}$ row in Fig. 2.2.2 can be computed from the $i^{\underline{th}}$ row using the formula:

$$A_{i+1,j} = -\sum_{p=1}^{j+1}(-1)^p \sum_{\substack{v_1+..+v_p=j \\ k_1+..+k_{p+1}=p-1 \\ v_i \geq 0,\ k_i \geq 0}} S_i^{(R_1)} S_i^{(k_1)} A_{i,v_1} S_i^{(k_2)} \cdots A_{i,v_p} S_i^{(k_{p+1})} \qquad (2.2.29)$$

$$i = 0,1,\ldots,m \quad j \geq 0$$

where

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

$$S_i^{(k)} = (A_{i,o}^{\#})^k \quad k > 0$$

(P_i is the projection on $N(A_{i,o})$ along $R(A_{i,o})$ and $A_{i,o}^{\#}$ devotes the group pseudo-inverse of $A_{i,o}$). Notice that the structure of (2.2.29) permits to grow the array $A_{i,j}$ triangularly: A_{10} is computed from A_{00} and A_{01}; A_{20} requires A_{10} and A_{11} which in turn involve A_{00}, A_{01} and A_{02}, in general, to compute the first column up to A_{ko} involves the matrices A_{ij}, i=0,...,k, j=0,.., k-i. As we have already seen, only a finite number of matrices A_{ko} need to be computed. It follows from (2.2.29) that this requires only a finite amount of computational effort.

Although algorithm (2.2.29) is attrative for its recursive nature, a closer look at its structure reveals that it involves a large number of superflows computations. Without addressing the issue of which is the most efficient way to compute the matrices A_{ko}, we will now give an explicit expression for the matrices A_{00}, A_{10}, A_{20} and A_{30} in terms of the first row in Fig. 2.2.2.

Proposition 2.2.15

Let $A_o(\varepsilon)$ be given by

$$A_o(\varepsilon) = A_{00} + \varepsilon A_{01} + \varepsilon^2 A_{02} + \varepsilon^3 A_{03} + o(\varepsilon^3)$$

then the matrices A_{k0}, k = 0,1,2,3, are given by:

$$A_{00}$$

$$A_{10} = P_o A_{01} P_o$$

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

$$A_{30} = P_2 P_1 P_o (A_{03} - A_{01}A_{00}^{\#}A_{02} - A_{02}A_{00}^{\#}A_{01} + A_{01}A_{00}^{\#}A_{01}A_{00}^{\#}A_{01} - A_{02}A_{10}^{\#}A_{02} + A_{02}A_{10}^{\#}A_{01}A_{00}^{\#}A_{01} - A_{01}A_{00}^{\#}A_{01}A_{10}^{\#}A_{02} - A_{01}A_{00}^{\#}A_{01}A_{10}^{\#}A_{01}A_{00}^{\#}A_{01}) P_o P_1 P_2$$

As shown in Proposition 2.2.15, the computation of A_{ko}, k=0,1,2,3, does not require the construction of the full triangular array shown in Fig. 2.2.2. Instead, the first column is directly derived from the first row. It is reasonable to expect that this pattern of simplifications goes beyond the third step and it thus seems possible to derive a recursive formula for the matrices A_{ko}. Here we do not attempt to derive such a formula nor do we address the question of how (2.2.29) should be computed so that a minimal effort is required. Let us just point out that, the different terms are quite similar and with a good strategy the computational effort required should not be as large as it seems at first sight. For example if,

$$A_o(\varepsilon) = A_o + \varepsilon B$$

we have:

$$A_{00} = A_o$$

$$A_{10} = P_o B P_o$$

$$A_{20} = -P_1 P_o B A_{00}^{\#} B P_o P_1$$

$$A_{30} = P_2 P_1 P_o (B A_{00}^{\#} B A_{00}^{\#} B - B A_{00}^{\#} B A_{10}^{\#} B A_{00}^{\#} B) P_o P_1 P_2$$

Notice that perturbations of order ε can result in well defined behavior at several time scales t/ε^k which is not a commonly recognized fact. In the MSST case, the sequence A_k ends at some m for which

$$\sum_{k=o}^{m} \text{rank } A_o = \text{nrank } A_o(\varepsilon)$$

We now derive an upper bound for m. Consider first the linearly perturbed case $A_o(\varepsilon) = A_o + \varepsilon B$ and let n be the dimension of $A_o(\varepsilon)$. The eigenvalues of $A_o(\varepsilon) = A_o + \varepsilon B$ are the solution of a polynomial of degree n with coefficients that are themselves polynomials in ε of degree ≤ n. A simple argument shows that there can be no eigenvalue, $\lambda(\varepsilon) \neq 0$, of $A_o(\varepsilon)$ such that $\lambda(\varepsilon) = o(\varepsilon^n)$. In effect, let

$$s^n + p_{n-1}(\varepsilon)s^{n-1} + \ldots + p_1(\varepsilon)s + p_o(\varepsilon) = o$$

be the characteristic polynomial of $A_o + \varepsilon B$. The coefficients $P_{n-i}(\varepsilon)$ are polynomials in ε of degree $\leq i$. Then,

$$\lambda(\varepsilon)^{n-1} + p_{n-1}(\varepsilon)\,\lambda(\varepsilon)^{n-2} + \ldots + p_1(\varepsilon) + \frac{p_o(\varepsilon)}{\lambda(\varepsilon)} = 0$$

If $\lambda(\varepsilon)$ were of order $o(\varepsilon^n)$ and $p_o(\varepsilon) \neq 0$ then $p_o(\varepsilon)/\lambda(\varepsilon) \to \infty$ as $\varepsilon \downarrow 0$ and the above equation cannot be satisfied. If $p_o(\varepsilon) = o$ then the same can be set about

$$\lambda(\varepsilon)^{n-2} + p_{n-1}(\varepsilon)\lambda(\varepsilon)^{n-3} + \ldots + p_2(\varepsilon) + \frac{p_1(\varepsilon)}{\lambda(\varepsilon)} = 0$$

which cannot be satisfied unless $p_1(\varepsilon) = o$. Proceeding in this way it is concluded that if $\lambda(\varepsilon) = o(\varepsilon^n)$ then $\lambda(\varepsilon) = o$. It then follows from Proposition 2.2.8 that $A_k(\varepsilon) = o$ for $k > n$ and therefore $m \leq n$. Similarly, in the case of a non-linear perturbation of finite order,

$$A_o(\varepsilon) = \sum_{j=o}^{p} A_{oj}\,\varepsilon^j$$

we have $m \leq n \cdot p$.

In addition to this upper bound on the slowest time scale the number of fundamental time scales (i.e., time scales for which $A_{ko} \neq 0$) can also be easily bounded. From the rank condition:

$$\sum_{k=o}^{m} \text{rank } A_{ko} = \text{rank } A_o(\varepsilon)$$

it follows that the maximum number of non-zero A_{ko}'s among $A_{00}, A_{01}, \ldots, A_{om}$ is equal to n-nrank $A_o(\varepsilon)$.

2.2.4 Partial Time Scale Decomposition

In this subsection we analyze the multiple time scale behavior of singularly perturbed systems that do not satisfy the MSST condition of section 2.2.3. In general, these systems have well defined behavior at some time scales but not at all time scales and their behavior over the infinite time interval $[0, \infty)$ cannot be reconstructed from their evolutions at different time scales. However, it may still be useful to be able to isolate the time scales at which they have well defined behavior, and to compute their evolutions at such time scales. This is the problem we address now.

At a first level we distinguish between systems which, although they do not satisfy the MSST condition, they do satisfy the MSSNS condition, and systems which do not satisfy the MSSNS condition. For system with MSSNS, the sequence of matrices A_{ko}, $k = 0,1,\ldots,m$, can be fully constructed as indicated in section 2.2.3. The MSST condition is violated if one of these matrices has an eigenvalue $\lambda \neq o$ such that $\text{Re}\lambda \geq o$. The strategy to obtain at least a partial time scale decomposition of $\exp\{A_o(\varepsilon)t\}$ is to multiply $\exp\{A_o(\varepsilon)t\}$ by the projections that anhilate the evolution of $\exp\{A_o(\varepsilon)t\}$ at time scales where the reduced-order model is not semistable.

For clarity we treat the case where only one of the matrices A_{Ro} violates the semistability condition. The following proposition gives the partial time scale decomposition for this case.

Proposition 2.2.17

Let $A_o(\varepsilon)$ satisfy the MSSNS condition and let A_{ko} $k = 0,1,\ldots,m$ be the matrices defined in equations (2.2.10) - (2.2.14). Suppose that A_{ko}, $k \neq \ell$, are semistable. Then,

i) $$\lim_{\varepsilon \downarrow 0} \sup_{t>o} \left\| P_\ell(\varepsilon)\exp\{A_o(\varepsilon)t\} - \phi^\ell(\varepsilon,t)\right\| = o$$

where

$$\phi^\ell(\varepsilon,t) = \sum_{\substack{k=o \\ k\neq\ell}}^{m} Q_k \exp\{A_k \varepsilon^k t\} + P_o P_1 \ldots P_m$$

ii) $$\lim_{\varepsilon \downarrow 0} \sup_{\delta \leq t \leq T} \left\| \exp\{A_o(\varepsilon)t/\varepsilon^k\} - \Phi_k(t)\right\| = o$$

for all $\forall \delta > 0$, $T < \infty$

$k = o,\ldots,\ell$

$$\lim_{\varepsilon \downarrow 0} \left\| P_\ell(\varepsilon)\exp\{A_o(\varepsilon)t/\varepsilon^k\} - \Phi_k(t)\right\| = o$$

$\forall \delta > o$, $\forall T < \infty$ $k = \ell+1,\ldots,m$

where

$$\Phi_k(t) = Q_k \exp\{A_{ko}t\} + P_o \ldots P_k$$

$$= P_o \ldots P_{-1} \exp\{A_o t\}$$ □

The above result indicates that, under the conditions of the Proposition, the multiple time scale behavior of $A_o(\varepsilon)$ up to the time scale where the MSST condition is violated is identical to the MSST case. From there on, however, the projection $P_\ell(\varepsilon)$ must be introduced to anihilate the behavior at time scale t/ε^ℓ which involves unstable or oscillatory modes. It is important to note that, in general, the projection $P_\ell(\varepsilon)$ used in (5.1), cannot be substituted by its leading term (nor by any finite number of terms in its power series expansion). Because of this lack of robustness, this result is of minor importance for applications; without some extra conditions, the multiple time scale analysis of MSSNS systems which violate the MSST condition at some time scale is feasible only up to this time scale.

Assume now that the matrix $A_o(\varepsilon)$ is stable for $\varepsilon \in (0,\varepsilon_o]$. Stability implies uniform boundedness of $\exp\{A_o(\varepsilon)t\}$ with respect to t, i.e.,

$$\sup_{t>o} \left\| \exp\{A_o(\varepsilon)t\}\right\| = K(\varepsilon) < \infty \quad \forall t \geq o$$

but, as shown by example 2.2.12 and the example below, for singularly perturbed systems $K(\varepsilon)$ may become unbounded as $\varepsilon \downarrow 0$ (even if $K(0)$ is finite as in example 2.2.12)

Example 2.2.18

Let

$$A_o(\varepsilon) = \begin{bmatrix} -\varepsilon & -\varepsilon \\ 1 & -\varepsilon \end{bmatrix}$$

This is a stable matrix with eigenvalues

$$\lambda_{1,2} = -\varepsilon \pm i\sqrt{\varepsilon}$$

and a simple computation gives:

$$\exp A_o(\varepsilon)t = \begin{bmatrix} \cos\sqrt{\varepsilon}t & -\sqrt{\varepsilon}\sin\sqrt{\varepsilon}t \\ \frac{1}{\sqrt{\varepsilon}}\sin\sqrt{\varepsilon}t & \cos\sqrt{\varepsilon}t \end{bmatrix} e^{-\varepsilon t}$$

Due to the entry $(1/\sqrt{\varepsilon})\sin\sqrt{\varepsilon}t$, the norm of $\exp\{A_o(\varepsilon)t\}$ becomes unbounded as $\varepsilon\downarrow 0$. Notice that, for this example, the unpertubed system

$$A_{oo} = A_o(0) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

is not stable and it is therefore clear that the ε perturbation has stabilizing effects. □

The kind of behavior discussed above indicates that in some systems the ε-dependence, in addition to generating eigenvalues of different orders of magnitude in ε, also models near instabilities. As we have seen the presence of increasingly large amplitudes as $\varepsilon\downarrow 0$ precludes the complete multiple time scale analysis of these systems. We now analyze the multiple time scale behavior of systems in which the ε-dependence does not give rise to increasingly large amplitudes as $\varepsilon\downarrow 0$.

We will say that $A_o(\varepsilon)$ satisfies the uniform stability (US) condition if:

$$\| \exp\{A_o(\varepsilon)t\} \| \leq K \quad \forall t \geq o \qquad \varepsilon\in[0, \varepsilon_o]$$

for some $K > o$ and independent of ε. The following proposition states that US is a sufficient condition for MSSNS.

Proposition 2.2.19

If $A_o(\varepsilon)$ is uniformly stable then it also satisfies the MSSNS condition; if any of the reduced-order models A_{k_o} has eigenvalues on the imaginary axis they must be semisimple. □

Uniform stability guarantees MSSNS but not MSST because some of the matrices A_{k_o} may have purely imaginary eigenvalues. The following proposition shows how to carry out a partial time scale decomposition for US systems.

Proposition 2.2.20

Let $A_o(\varepsilon)$ be uniformly stable and suppose that all the reduced order models A_{k_o}, $k = 0,1,..,m$, are semistable except A_{ℓ_o}. Then:

i) $$\lim_{\varepsilon\downarrow 0} \sup_{t\geq o} \| P_\ell \exp\{A_o(\varepsilon)t\} - \phi^\ell(\varepsilon,t) \| = o$$

where

$$\phi^\ell(\varepsilon,t) = \sum_{\substack{k=o \\ k\neq\ell}}^{m} Q_k \exp\{A_{k_o}\varepsilon^k t\} + P_o P_1 \cdots P_m$$

ii) $$\lim_{\varepsilon\downarrow 0} \sup_{\delta\leq t\leq T} \| \exp\{A_o(\varepsilon)t/\varepsilon^k\} - \Phi_k(t) \| = o$$

$$\forall\delta > o, \forall T < \infty \quad k = 0,1,..,\ell$$

$$\lim_{\varepsilon\downarrow 0} \sup_{\delta\leq t\leq T} \| P_\ell \exp\{A_o(\varepsilon)t/\varepsilon^k\} - \Phi_k(t) \| = o$$

$$\forall\delta > o, \forall T < \infty \quad k = \ell+1,..,m$$

where:

$$\Phi_k(t) = Q_k \exp\{A_{k_o}t\} + P_o \cdots P_k$$

$$= P_o \cdots P_{k-1} \{\exp A_{k_o}t\}$$

It is a simple matter to extend Proposition 2.2.20 to the case where several of the matrices A_{k_o} fail to be semistable.

From the above developments it is clear that uniformly stable systems may fail to have well defined behavior at certain time scales because of the presence of oscillatory modes in some of the reduced-order models A_{k_o}. These oscillations become of infinite frequency when seen at slower time scales. It is important to notice that the appearance of such unattenuated oscillations in some of the reduced order models A_{k_o} does not necessarily imply that the matrix $A_o(\varepsilon)$ has some purely imaginary eigenvalues. Instead, they could as well correspond to eigenvalues with a negative real part that converges to zero faster then its imaginary part. For example take $\lambda(\varepsilon) = -\varepsilon+i$. In this case $\exp\{\lambda(\varepsilon)t\}$ is seen as a pure oscillatory mode at time scale t while at time scale t/ε, when the attenuation effects are beginning to be felt, the oscillations become of infinite frequency. To avoid this lack of well defined limit at slower time scales the oscillatory modes must be excluded from consideration; this can be done using the adequate projections as indicated in Proposition 2.2.20.

The fact that the projections required to anihilate the undesired modes are ε-independent makes Proposition 2.2.30 much more useful the Proposition 2.2.18 for applications.

In all the results derived so far, the decomposition

$$\exp\{A_o(\varepsilon)t\} = \sum_{k=o}^{m} Q_k(\varepsilon)\{\exp A_k(\varepsilon)\varepsilon^k t\} + P_o(\varepsilon)\cdots P_m(\varepsilon) \qquad (2.2.30)$$

together with the fact that

$$\sum_{k=o}^{m} \operatorname{rank} A_{k_o} = \operatorname{nrank} A_o(\varepsilon) \qquad (2.2.31)$$

have played a fundamental role. If the MSSNS condition is violated then this decomposition is not feasible. It is still possible, however, to derive some multiple time scale results from a reduced version of (2.2.30). We briefly discuss this case.

Suppose that $A_o(\varepsilon)$ is such that the sequence of matrices $A_k(\varepsilon)$ constructed in (2.2.10) - (2.2.14) ends for $k = p$ because A_{p_o} does not have SSNS, and that

$$\sum_{k=o}^{p} \text{rank } A_{k_o} < \text{nrank } A_o(\varepsilon)$$

We can decompose $\exp\{A_o(\varepsilon)t\}$ as follows:

$$\exp\{A_o(\varepsilon)t\} = \sum_{k=o}^{p} Q_k(\varepsilon) \exp\{A_k(\varepsilon)\ \varepsilon^k t\}$$

$$+ P_o(\varepsilon)\ldots P_p(\varepsilon) \qquad (2.2.32)$$

and the multiple time scale analysis developed for MSSNS systems can now be performed up to time scale t/ε^p. the fact that $\sum_{k=o}^{p} \text{rank } A_{k_o}$ is not equal to nrank $A_o(\varepsilon)$ indicates, however, that there are eigenvalues of $A_o(\varepsilon)$ that are of order $o(\varepsilon^p)$ whose effect will not be captured by this partial time scale analysis. Furthermore, systems which violate the MSSNS condition at stage p are not uniformly stable and the limit

$$\lim_{\varepsilon \downarrow 0} \exp\{A_o(\varepsilon)t/\varepsilon^{p+1}\}$$

can not exist because $\| \exp A_o(\varepsilon)t/\varepsilon^{p+1} \| \to \infty$ as $\varepsilon \downarrow 0$. If $A_o(\varepsilon)$ is asymptotically stable it may well be possible that the system has well defined behavior at slower time scales t/ε^{ℓ}, $\ell > p+1$, when the effects of high amplitude transients have disappeared. The techniques we have used, however, do not seem adequate to treat this case. More work is needed in this direction.

2.2.5 Summary and Conclusions

In this section we have studied the asymptotic behavior of $\exp\{A_o(\varepsilon)t\}$ over the infinite time interval $[o, \infty)$ for

$$A_o(\varepsilon) = \sum_{k=o}^{\infty} \varepsilon^k A_{ok}$$

We have formalized the notion of multiple time scale behavior and that of reduced order models of the system

$$\frac{dx(t)}{dt} = A_o(\varepsilon)\ x\ (t)$$

valid at different time scales. We have identified several conditions on $A_o(\varepsilon)$ which give rise to qualitatively different asymptotic behaviors. The hierarchical relationship among these conditions is visualized in Fig. 2.2.3.

The most important result is probably the fact that multiple semistability is a necessary and sufficient condition for a system to have well defined behavior at all of its natural time scales. If a system does not satisfy this condition then a time scale analysis of it will not be adequate to capture all of the different system's features. Conversely, if a system has a well defined multiple time scale behavior, then, when it is correctly modelled, this model must result in a system matrix that satisfies the MSST condition.

For MSST systems we have developed a methodology to compute the different reduced-order models of a system which describe its evolution at different time scales and we have shown that these reduced-order models can be combined to approximate the original system. The usefulness of this model simplification technique depends on two counts:

i) It must be shown that problems posed for the original system can be approximately solved by combining the solution of some related problems posed for the reduced order models with some reduction in complexity and

ii) it must also be shown that the calculations required to compute the reduced-order models do not exceed the savings in i).

In the next section, we show that for an important class of systems, the MSST condition is always satisfied indicating that the results developed in this chapter have a wide range of applicability.

2.3 Hierarchical Aggregation of Finite State Markov Process

2.3.1. Introduction

In this section we apply the results presented in section 2.2 to the study of Finite State Markov Processes (FSMP) with rare events.

The presence of rare events in a FSMP is modeled by introducing a small parameter ε in its matrix of transition rates. As an introductory example, consider the process $\eta^{\varepsilon}(t)$ depicted in fig. 2.3.1. Its matrix of transition rates is of the form $A_o(\varepsilon) = A_o + \varepsilon B$ and can be thought of as modeling a system with the following characteristics: it may operate in two different modes corresponding to the two sets of states $E_1 = \{1,2\}$ and $E_2 = \{4,5\}$. A rare event (a failure alarm, for example) puts the system in state 3 where a fast decision is made whether to continue operating in the same mode or whether to change it. While in state 3, another rare event (a wrong decision, for example) may put the system out of operation (state 6). At a first level of simplification we can imagine a reduced-order model which only describes changes in the mode of operation and failures, neglecting the evolution of the system while in a given mode. An even coarser description of the system would be provided by a model which only distinguishes between working and non-working states. Given the structure of the process, transitions between states in E_1 and states in E_2 are likely to occur only for times of order $1/\varepsilon$ while a failure, requiring two consecutive transitions with rates of order ε, will take place at times of order $1/\varepsilon^2$.

It is clear from this example (see section 2.3.6 for a full development) that the connection among rare events, multiple time scale behavior and reduced-order (aggregated) models is intuitively appealing in FSMP. In fact, several authors have, in different contexts, used aggregation ideas in the past. Simon and Ando [SIM 61] were probably the first, Curtois [CUR 77] argues in a somewhat heuristical way that aggregated models for Markov chains are useful in the analysis of computer systems while [DEL 81] use aggregated models to simplify control problems for Markov chains. The most complete treatment to date is given in [KOR 78] where the results are not restricted to FSMP but include Markov processes with unbounded infinitesimal generators as well. All these works, however, either introduce

conditions which guarantee that the process under consideration has only a two time scale behavior (i.e. t and t/ε) or restrict their attention to a certain time scale. In any case only one aggregated model is proposed for a given system. In line with the multiple time scale results presented in section 2.2, our results extend previous work on FSMP aggregation by showing that:

(i) FSMP with rare transitions can always be aggregated.

(ii) In general, such processes exhibit multiple time scales and several aggregated models are possible, one for each time scale.

(iii) An approximation to the original process, which is uniformly valid at all times t, can be constructed by combining the set of aggregated models.

The probabilistic evolution of a FSMP $\eta^\varepsilon(t)$ with transition rate matrix $A_o(\varepsilon)$ is completely determined by its matrix of transition probabilities which is given by:

$$P^\varepsilon(t) = \exp\{A_o(\varepsilon)t\}$$

An application of the multiple time scale techniques developed in section 2.2 to the asymptotic analysis of $P^\varepsilon(t)$ is carried out in this section. Specifically, we show that if $A_o(\varepsilon)$ is the matrix of transition rates of a FSMP then it satisfies the MSST condition (see section 2.2.3) and therefore $P^\varepsilon(t)$ has well defined multiple time scale behavior. The reduced-order models that describe the evolution of $P^\varepsilon(t)$ at each of its fundamental time scales are then interpreted as aggregated models of $\eta^\varepsilon(t)$ obtained by collapsing several states of $\eta^\varepsilon(t)$ into a single state of a model that describes events in $\eta^\varepsilon(t)$ of a certain level of rareness. Each of the models is a FSMP and they can be ordered in a hierarchy, where each model describes the evolution of $\eta^\varepsilon(t)$ with a different degree of detail.

The section is ordered as follows: In 2.3.2, 2.3.3 and 2.3.4 we discuss some preliminary material on FSMP. The emphasis is placed on stochastically discontinuous Markov processes, i.e., processes which may undergo an infinite number of instantaneous transitions in a finite time interval. This kind of processes which have received little attention in the past are very useful in interpreting the multiple time scale behavior of $P^\varepsilon(t)$. Basically, if $\eta^\varepsilon(t)$ is analyzed for times of order t/ε^k, all transitions that occur at faster time scales look as instantaneous and a stochastically discontinuous process is adequate to approximate the behavior of $\eta^\varepsilon(t)$ at the time scale of interest. The existence of aggregated models and a method for computing them are discussed in sections 2.3.5 and 2.3.6. One example is given in section 2.3.7.

2.3.2. Ergodic Projections of Markov Chains

A stationary Markov chain $\eta(t)$, t=0,1.,,, taking values in a finite state space $X = \{1,2, \ldots,n\}$ is completely characterized by its transition probability matrix, P, whose elements are the one-step transition probabilities:

$$P_{ij} = P_r\{\eta(t) = j \mid \eta(t-1) = i\} \quad i,j \in X \qquad (2.3.1)$$

A matrix P is the transition probability matrix of a Markov chain if and only if:

$$\text{(i)}\ P \geq o \qquad (2.3.2)$$

$$\text{(ii)}\ P \cdot \mathbb{1} = \mathbb{1} \qquad (2.3.3)$$

where $\mathbb{1}$ is a vector with all components equal to one. A matrix satisfying (2.3.2) and (2.3.3) is called stochastic. For our later work we will be mainly interested in the behavior of P^n as $n \to \infty$. We shall say that a Markov chain $\eta(t)$ with transition matrix P is ergodic if the limit

$$\lim_{n\to\infty} P^n = P$$

exists. Notice that in this case $PP = PP = P$ and that P is a stochastic projection. We shall refer to it as the ergodic projection of $\eta(t)$. The following theorem specifies a canonical form for the transition probability matrix of a ergodic chain and for its ergodic projection.

Theorem 2.3.1 [KEM 60]*

A stochastic matrix P is the transition probability matrix of a ergodic chain iff by an adequate ordering of the states it can be written in the following form:

$$P = \begin{bmatrix} P_{11} & 0 & \cdots & 0 \\ 0 & P_{22} & \cdots & \\ 0 & \cdots & P_{mm} & 0 \\ P_{m+1,1} & \cdots & P_{m+1,m} & P_{m+1,m+1} \end{bmatrix} \qquad (2.3.4)$$

with the submatrices P_{ii} i = 1,2,...m satisfying $P_{ii}^{n_i} > 0$ for some $n_i \geq 1$.

Its ergodic projection then takes the form:

$$P = \begin{bmatrix} P_{11} & 0 & \cdots & 0 \\ 0 & P_{22} & \cdots & \\ 0 & \cdots & P_{mm} & 0 \\ P_{1,m+1} & \cdots & P_{m,m+1} & 0 \end{bmatrix} \qquad (2.3.5)$$

with $P_{kk} = \mathbb{1} \cdot \mu_k^T$, k=1,...,m, for some vector $\mu_k > 0$ such that $\mu_k^T \cdot \mathbb{1} = 1$ and $P_{k,m+1} = \delta_k \cdot \mu_k^T$ for some vector $\delta_k \geq o$ such that $\sum_{k=1}^{m} \delta_k = \mathbb{1}$.

The canonical form (2.3.4) determines a partition of the state space

$$X = \left(\bigcup_{k=1}^{m} X_k\right) \cup X_T \qquad (2.3.6)$$

*Our definition of ergodic chain differs from that of [KEM 60]. Ours is consistent with the notion of ergodicity used for other processes. [KEM 60] allow the existence of cyclic classes and therefore in general no steady state probabilities exist.

The sets X_k are called ergodic classes and X_T is the set of transient states. Once the process enters a ergodic class it never leaves it. If the process starts in a transient state then it leaves X_T in a finite number of steps w.p.1 and X_T is never re-entered. The vectors μ_k $k=1,\ldots,m$ are the ergodic probabilities of the chain P_{kk} with state space X_k and the j^{th} component of the vector δ_k is the probability that the chain will get trapped in ergodic class X_k if it starts at the j^{th} transient state. For latter developments it is convenient to write the ergodic projection P in terms of some of its right and left eigenvectors with eigenvalue 1. Notice that by the structure of P in (2.3.5) there are at least m linearly independent right eigenvectors with eigenvalue 1 given by

$$r_k^T = [0 \ \ldots \ 0 \ \ \mathbb{1}_k^T \ 0 \ \ldots \ 0 \ \delta_k^T]^T \quad k=1,..,m \tag{2.3.7}$$

where $\mathbb{1}_k$ is a vector of ones with dimension equal to that of the block P_{kk}. A complete set of left eigenvectors of P with eigenvalue 1 is readily constructed as follows:

$$\ell_k^T = [0 \ \ 0 \ \ \ldots \ \ 0 \ \ \mu_k^T \ \ 0 \ \ldots \ 0]^T \quad k=1,..,m \tag{2.3.8}$$

Let V deonte the (nxm) matrix whose columns are the vectors r_R and let U be (mxm) matrix whose rows are the ℓ_R's. Then we have:

$$V \cdot U = P \tag{2.3.9}$$

$$U \cdot V = I \tag{2.3.10}$$

We will refer to (2.3.9) as the canonical product decomposition of a ergodic projection.

2.3.3. Finite State Makov processes Basic Definitions

A Markov process $\eta(t)$ taking values in a finite state space $X = \{1,2,\ldots,n\}$ is completely characterized by its transition probability matrix, $P(t)$, whose elements are the transition probabilities:

$$P_{ij}(t) = P_r\{\eta(t) = j \mid \eta(o) = i\}$$

$$i,j \in X, \ t \geq o \tag{2.3.11}$$

An (nxn) matrix-valued function $P(t)$, $t \geq o$ is the transition probability matrix of a Markov process taking values in X iff the following conditions are satisfied:

$$\text{(i)} \quad P(0) = I \tag{2.3.12}$$

$$\text{(ii)} \quad P(t) \geq o, \quad \forall t \geq o \tag{2.3.13}$$

$$\text{(iii)} \quad P(t)\,\mathbb{1} = \mathbb{1} \tag{2.3.14}$$

$$\text{(iv)} \quad P(t)P(\tau) = P(t+\tau), \ \forall t, \tau \geq o \tag{2.3.15}$$

It can be proved ([DOE 38], [D00 42)] that if $P(t)$ is a transition probability matrix then it is continuous for $t > o$ and the limit

$$\lim_{t \downarrow 0} P(t) \triangleq \Pi \tag{2.3.16}$$

always exists. It follows from (2.3.13) - (2.3.15) and the continuity of $P(t)$ that Π satisfies:

$$\Pi \geq 0, \ \Pi \cdot \Pi = \Pi, \ \Pi^2 = \Pi \tag{2.3.17}$$

and

$$\Pi P(t) = P(t)\Pi = P(t) \tag{2.3.18}$$

If $\Pi = I$ then $\eta(t)$ is called stochastically continuous, otherwise it is called stochastically discontinuous. The following theorem gives a unique characterization of $P(t)$ in terms of a set of parameters.

Theorem 2.3.2

If $P(t)$ is the transition probability matrix of a conservative FSMP then

$$P(t) = \Pi \exp\{At\} \qquad t > o \tag{2.3.19}$$

for a pair of matrices Π and A satisfying:

i) $\Pi \geq 0, \ \Pi \cdot \mathbb{1} = \mathbb{1}, \ \Pi^2 = \Pi$

ii) $\Pi A = A\Pi = A$

iii) $A\mathbb{1} = 0$

iv) $A + c\Pi \geq 0$ for some $c > 0$.

Conversely any pair of matrices A, Π satisfying the above properties uniquely determine a FSMP with transition probability matrix given by (2.3.19).

We shall refer to the projection $\Pi = \lim_{t \downarrow 0} P(t)$ as the ergodic projection at zero and to the matrix

$$A = \lim_{t \downarrow 0} \frac{P(t) - \Pi}{t}$$

as the infinitesimal generator of $P(t)$.

Example 2.3.3

Consider a FSMP $\eta(t)$ taking values in $X = \{1,2,3\}$ with transition probability matrix:

$$P(t) = \begin{bmatrix} p_1 e^{-\lambda t} & p_2 e^{-\lambda t} & 1-e^{-\lambda t} \\ p_1 e^{-\lambda t} & p_2 e^{-\lambda t} & 1-e^{-\lambda t} \\ 0 & 0 & 1 \end{bmatrix}$$

with $p_1 + p_2 = 1$. Its initial projection is:

$$\Pi = \lim_{t \downarrow 0} P(t) = \begin{bmatrix} p_1 & p_2 & 0 \\ p_1 & p_2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and its infinitesimal generator:

$$A = \lim_{t\downarrow 0} \frac{P(t)-\Pi}{t} = \begin{bmatrix} -p_1\lambda & -p_2\lambda & \lambda \\ -p_1\lambda & -p_2\lambda & \lambda \\ 0 & 0 & 0 \end{bmatrix}$$

For stochastically continuous processes the elements of the matrix A satisfy $a_{ij} \geq 0$ for $i \neq j$ and they can be interpreted as transition rates in the sense that:

$$P_{ij}(\Delta t) = a_{ij} \cdot \Delta t + o(\Delta t) \qquad i \neq j$$

If we consider a separable version of the process $\eta(t)$ then, as we will see:

$$Pr\{\eta(\tau) = i,\ \tau \in [0,t] \,|\, \eta(o) = i\} = e^{a_{ii}t}$$

which means that the time of first exit out of state i is exponentially distributed with parameter $-a_{ii} \geq 0$. The evolution of $\eta(t)$ can be thought of as a succesion of stays in different states in X, each being or random duration and exponentially distributed with parameter that depends on the state. The sequence of states forms a Markov chain with one-step transition probabilities given by $p_{ij} = -a_{ij}/a_{ii}$ and the sample functions of $\eta(t)$ are easily visualized as piecewise continuous functions taking values in X.

The sample functions of stochastically discontinuous processes are much more irregular. As we will now see, these processes have instantaneous states, i.e., states in which the process spends no time with probability one but in spite of that the process, in general, spends a non-zero amount of time switching among instantaneous states. The sample functions have therefore pieces where they are nowhere continuous. To classify the states of a process $\eta(t)$ with ergodic projection Π at zero and infinitesimal generator A, consider a separable version of the process and let Λ be a separating set. For $t > o$ and $n = 0,1,\ldots$ take $0 = t_{on} < t_{1n} < \ldots < t_{nn} = t$ in such a way that the sets $\Lambda_n = \{t_{on}, t_{1n}, \ldots, t_{nn}\}$ increase monotonically and $\cup \Lambda_n = \Lambda \cap [0,t]$. Then we have:

$$Pr\{\eta(\tau)=i, \tau \in [0,t] \,|\, \eta(0)=i =$$

$$Pr\{\eta(\tau)=i, \tau \in [0,t] \cap \Lambda \,|\, \eta(0)=i\} =$$

$$\lim_{n\uparrow\infty} Pr\{\eta(\tau)=i, \tau \in [0,t] \cap \Lambda_n \,|\, \eta(0)=i\} =$$

$$\lim_{n\uparrow\infty} \prod_{k=o}^{n-1} P_{ii}(t_{k+1,n} - t_{k,n}) =$$

$$\exp\{\lim_{n\uparrow\infty} \sum_{k=o}^{n-1} \log P_{ii}(t_{k+1,n} - t_{k,n})\} \qquad (2.3.20)$$

where $P_{ii}(t)$ are the diagonal elements of $P(t) = \Pi \exp\{At\}$. If $\pi_{ii} = o$ (the diagonal elements of Π) then $\lim_{h\downarrow 0} P_{ii}(h) = o$ and therefore

$$Pr\{\eta(\tau) = i,\ \tau \in [o,t] \,|\, \eta(o)=i\} = o$$

To compute (2.3.20) when $o < \pi_{ii} \leq 1$ write:

$$\frac{P_{ii}(h)}{\pi_{ii}} = 1 + \frac{a_{ii}}{\pi_{ii}} h + o(h)$$

which gives,

$$\log P_{ii}(h) = \log \pi_{ii} + \frac{a_{ii}}{\pi_{ii}} h + o(h)$$

and therefore:

$$Pr\{\eta(\tau) = i,\ \tau \in [o,t] \,|\, \eta(o) = i\} =$$

$$\exp\{\lim_{n\uparrow\infty} [n \log \pi_{ii} + \sum_{k=o}^{n-1} \frac{a_{ii}}{\pi_{ii}} [t_{k+1,n} - t_{k,n}]$$

$$+ o(t_{k+1,n} - t_{k,n})]]\} =$$

$$\begin{cases} 0 & \text{if } \pi_{ii} < 1 \\ \exp\{a_{ii}t\} & \text{if } \pi_{ii} = 1 \end{cases} \qquad (2.3.21)$$

A state i will be called <u>instantaneous</u> if $\pi_{ii} < 1$ and <u>regular</u> if $\pi_{ii} = 1$. An instantaneous state i will be called <u>evanescent</u> if $\pi_{ii} = o$. The sojourn time in instantaneous states is zero with probability one, as indicated by (2.3.21), while for regular states it is exponentially distributed with rate $-a_{ii} \geq o$. In stochastically continuous processes all states are regular. In example 2.3.3 states $\{1,2\}$ are instantaneous while 3 is regular. Notice that even though the duration of stays in instantaneous states is zero w.p.1, there is, in general, a non-zero probability of finding the process in an instantaneous state at any given time as seen in example 2.3.3 for states 1 and 2. The structure of the sample functions for this example is shown in Fig. 2.3.2.

Notice that if the process $\eta(t)$ of example 2.3.3 is aggregated by merging states $\{1,2\}$ into a single state, then the aggregated process $\hat{\eta}(t)$ defined as:

$$\hat{\eta}(t) = \begin{cases} 1 & \text{if } \eta(t) \in \{1,2\} \\ 2 & \text{if } \eta(t) = 3 \end{cases}$$

is a Markov process with transition probability matrix $\hat{P}(t)$ given by:

$$\hat{P}(t) = \begin{bmatrix} e^{-\lambda t} & 1-e^{-\lambda t} \\ 0 & 1 \end{bmatrix}$$

The aggregated process $\hat{\eta}(t)$ is therefore stochastically continuous. In the next section we show that every stochastically discontinuous process uniquely determines a stochastically continuous process obtained by collapsing groups of instantaneous states of the original process into a single state of the aggregated process and conversely that every stochastically discontinuous process is completely determined by its aggregated, stochastically continuous version.

2.3.4. Ergodic projections, state space partitions and aggregation of stochastically discontinuous FSMP.

We prove here that all the probabilistic properties of a stochastically discontinuous process can be derived from an aggregated version of the process that is stochastically continuous.

Let $P(t) = \Pi \exp\{At\}$ be the transition probability matrix of a FSMP $\eta(t)$ taking values in X. It follows from (2.3.17) that the ergodic projection at zero, Π, is also the ergodic projection of some Markov chain and therefore it has the form:

$$\Pi = \begin{bmatrix} I & 0 \cdots\cdots 0 \\ & \Pi_{rr} \cdots\cdots 0 \\ 0 \cdots\cdots & \Pi_{m,m} \quad 0 \\ \Pi_{1,m+1} & \Pi_{m,m+1} \quad 0 \end{bmatrix} \qquad (2.3.22)$$

Let $X = (\bigcup_{k=1}^{m} X^o_k) \cup X^o_T$ denote the partition of X into ergodic classes and the set of transient states determined by (2.3.22). The first group of states in (2.3.22), corresponding to the identity block (i.e., absorbing states of Π), are the regular state of $\eta(t)$. States in X^o_T, i.e., transient states of the chain Π, are the evanescent states of $\eta(t)$ which, as we will now see, can be pruned without affecting the finite dimensional distributions of the process. In effect,

$$P(t) = \Pi \exp\{At\} = \Pi \exp\{At\}\, \Pi \qquad (2.3.23)$$

and given the structure of Π, it follows that:

$$\Pr\{\eta(t) \in X^o_T \,|\eta(o)=i\} = o \quad \forall t > o, \quad \forall i \in X \qquad (2.3.24)$$

Evanescent states can thus be neglected in the sense that there exists a version of the process $\eta(t)$ that has the same finite dimensional distributions but does not take values in X^o_T.

The blocks Π_{ii}, i=r,..,m determine groups of non-evanescent instantaneous states. Such instantaneous states cannot be neglected but each ergodic class can be consolidated into a single state of a stochastically continuous process as stated in the following theorem.

Theorem 2.3.4

Let $P(t) = \Pi \exp\{At\}$ be the transition probability matrix of a FSMP $\eta(t)$ taking values in $X = \{1,2,\ldots,n\}$ and let m be the number of ergodic classes at zero. Let $\Pi = V \cdot U$ be the canonical product decomposition of Π constructed in section 2.3.2. Then:

$$\hat{P}(t) \triangleq U\, P(t) V = \exp\{UAVt\} \qquad (2.3.25)$$

is the transition probability matrix of FSMP $\hat{\eta}(t)$ taking values in $\hat{X} = \{1,2,\ldots,m\}$ and

$$P(t) = V\hat{P}(t)\, U \qquad \forall t > o \qquad (2.3.26)$$

□

The above theorem states that every stochastically discontinuous FSMP uniquely determines a stochastically continuous FSMP in a smaller state space and shows how their transition probability matrices are related. We now show that $\hat{\eta}(t)$ is in fact obtained by neglecting evanescent states and merging states of $\eta(t)$ belonging to a ergodic class at zero into a single state of $\hat{\eta}(t)$. Let $j \in X$ be an arbitrary state of $\eta(t)$ belonging to X^o_ℓ, i.e., to the ℓ^{th} ergodic class at zero, and let X^o_k be any ergodic class at zero. Denote $u_k^T = [0|\,^{k}|0|\Pi_k^T|0|\;\;|0]^T$. Also let 1_j be the vector with the j^{th} component equal to 1 and the rest equal to zero. Then we have:

$$\Pr\{\eta(t) \in X^o_k | \eta(o)=j\} = 1_j^T\, P(t)\, u_k =$$

$$1_j^T\, V \hat{P}(t)\, U\, u_k \qquad (2.3.27)$$

and by construction of V and U (remember $j \in X^o_\ell$),

$$1_j^T V = 1_\ell^T$$

$$U u_k = 1_k \qquad (2.3.28)$$

thus giving:

$$\Pr\{\eta(t) \in X^o_R | \eta(o)=j\} = \hat{p}_{\ell k}(t) = \Pr\{\hat{\eta}(t) =$$

$$k | \hat{\eta}(o) = \ell\} \quad \forall t > o \qquad (2.3.29)$$

The process $\hat{\eta}(t)$ is therefore an aggregated version of $\eta(t)$, i.e.,

$$\hat{\eta}(t) = k \quad \text{if } \eta(t) \in X^o_k \quad k=1,2,\ldots,m \qquad (2.3.30)$$

Notice that, in principle, there is an uncertainty in defining $\hat{\eta}(t)$ as in (2.3.30) because no value is assigned to $\eta(t)$ whenever $\eta(t) \in X^o_T$. As we have said, however, there is a version of $\eta(t)$ which does not take values in X^o_T. Using this version (2.3.30) completely determines $\hat{\eta}(t)$.

In addition to the aggregated probabilities (2.3.29), the aggregated process $\hat{\eta}(t)$ also determines $p_{ij}(t)$, i,j ∈ X as follows:

$$\Pr\{\eta(t) = j | \eta(o)=i\} = 1_j^T \cdot P(t) \cdot 1_i =$$

$$1_j^T\, V \hat{P}(t)\, U 1_i \qquad (2.3.31)$$

By construction of V and u:

$$l_i = \begin{cases} \pi_i^{(k)} & \text{if } i \in X_k^o \quad \text{some } k=1,..,m \\ 0 & \text{if } i \in X_T^o \end{cases}$$

$$l_j^T V = l_j^T \qquad \text{if } i \notin X_T^o$$

where $\Pi_i^{(k)}$ is the $i^{\underline{th}}$ entry of any row of Π_{kk} (see (2.3.22), that is the ergodic probability at zero of state i belonging to class X_k^o. We thus have:

$$\Pr\{\eta(t) = j | \eta(o)=i\} = \Pi_i^{(k)} \quad \hat{p}_{k\ell}(t) =$$

$$\Pi_i^{(k)} \Pr\{\hat{\eta}(t) = \ell | \hat{\eta}(o) = k\} \qquad (2.3.32)$$

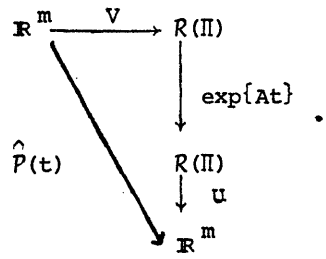
$$\forall t > o \qquad i \in X_k^o$$

$$j \in X_\ell^o$$

We will refer to (2.3.25) as the aggregation operation and to (2.3.26) as the disaggregation operation. These operations can also be interpreted from a geometrical point of view. Notice that the stochastically discontinuous transition probability matrix $P(t) = \Pi \exp\{At\}$ satisfies:

$$P(t)f = \exp\{At\}f \quad \text{for } t \in \mathcal{R}(\Pi) \supset \mathcal{R}(A) \qquad (2.3.33)$$

and therefore it defines a transition matrix on $\mathcal{R}(\Pi)$ that is continuous at zero. Let m be the number of ergodic classes at zero then rank Π = m. The matrix V: $\mathbb{R}^m \to \mathbb{R}^n$ maps $\mathbb{R}^m$ into $\mathcal{R}(\Pi)$ in a one-to-one basis and U: $\mathbb{R}^n \to \mathbb{R}^m$ maps $\mathcal{R}(\Pi)$ back into $\mathbb{R}^m$ also one to one. We thus have the following diagram:



From this point of view the state space aggregation is interpreted as a restriction in the domain of definition of the transition probability matrix.

We next give an example of a stochastically discontinuous process and its aggregated version.

Example 2.3.5

Let $\eta(t)$ be a FSMP taking values in X = {1,2, ...,6} with ergodic projection at zero given by:

$$\Pi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 & 0 \end{bmatrix}$$

and infinitesimal generator:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1/8 & -1/8 & 1/8 & 1/8 & 0 \\ 0 & -1/8 & -1/8 & 1/8 & 1/8 & 0 \\ 0 & 1/8 & 1/8 & -1/8 & -1/8 & 0 \\ 0 & 1/8 & 1/8 & -1/8 & -1/8 & 0 \\ 0 & -1/16 & -1/16 & 1/16 & 1/16 & 0 \end{bmatrix}$$

There is a regular state, {1}, an evanescent state, {6}, and two ergodic classes of instantaneous states, {2,3} and {4,5}.

Notice that A is not a matrix of transition rates of a stochastically continuous FSMP because it has negative elements in off-diagonal positions. However, $A + \Pi \geq 0$ and the pair of matrices A and Π define a FSMP. For this example the matrices V and U are given by:

$$V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1/2 & 1/2 \end{bmatrix} \qquad U = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \end{bmatrix}$$

Notice that the matrix V includes all the information about which states belong to what ergodic class at zero while the matrix U gives the ergodic probabilities for each class. The matrix of transitions rates $\hat{A} = U A V$ is is thus given by:

$$\hat{A} = U A V = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1/4 & 1/4 \\ 0 & 1/4 & -1/4 \end{bmatrix}$$

The aggregation operation corresponds to consolidation of states {2,3} and {4,5} into two states of $\hat{\eta}(t)$. The regular state {1} remain unaffected and the evanescent state {6} is pruned. The consolidated rates are obtained by averaging the coefficients in A corresponding to transitions between states in different ergodic classes against the ergodic probabilities at at zero for each ergodic class.

From the above discussion it is clear that all relevant information about a stochastically discontinuous process is contained in its aggregated version and its initial projection, and therefore the

study of such processes can be reduced, using theorem 2.3.4 to the well known stochastically continuous case.

We analize now the behavior of $P(t) = \Pi \exp\{At\}$ as $t \to \infty$. We will say that a FSMP $\eta(t)$ is ergodic at ∞ if $\lim_{t\to\infty} P(t)$ exists. The following theorem extends to the stochastically discontinuous case a result that for stochastically continuous processes is well known. Namely, that all FSMP are ergodic at ∞.

Theorem 2.3.6

If $P(t)$ is the transition probability matrix of a FSMP with ergodic projection Π_1 at zero and infinitesimal generator A, then:

$$\lim_{t\to\infty} P(t) = P$$

always exists and satisfies:

$$P^2 = P$$

$$P \geq 0, \; P \cdot \Pi = \Pi$$

$$PP(t) = P(t)P = P$$

$$P\Pi = \Pi P = P$$

$$PA = AP = 0$$

The projection $P = \lim_{t\to\infty} P(t)$ will be referred to as the ergodic projection at ∞.

2.3.5 The MSST Property of Singularly Perturbed Finite State Markov Processes.

Consider now a stochastically continuous FSMP $\eta^\varepsilon(t)$ taking values in $X = \{1,2,\ldots,n\}$ with transition probability matrix

$$P^\varepsilon(t) = \exp\{A_o(\varepsilon)t\} \qquad t \geq o \qquad (2.3.34)$$

and infinitesimal generator of the form:

$$A_o(\varepsilon) = \sum_{p=o}^{\infty} \varepsilon^p A_{op} \qquad \varepsilon\in[0, \varepsilon_o] \qquad (2.3.35)$$

It follows from Thm. 2.3.6. that $A_o(\varepsilon)$, $\varepsilon\in[0, \varepsilon_o]$, is a semistable matrix and, as we have seen in section 2.2.2, if rank $A_o(\varepsilon)$ is discontinuous at $\varepsilon = o$ then

$$\lim_{\varepsilon\downarrow 0} \sup_{t\geq o} ||P^\varepsilon(t) - P^o(t)|| \neq 0 \qquad (2.3.36)$$

which means that in this case $\eta^o(t)$ can not be a good approximation of $\eta^\varepsilon(t)$ no matter how small ε is. Accordingly, if

$$n \text{ rank } A_o(\varepsilon) \neq \text{ rank } A_o(o) \qquad (2.3.37)$$

we will say that $\eta^\varepsilon(t)$ is a singularly perturbed FSMP. Because null $A_o(\varepsilon)$ equals the number of ergodic classes of $\eta^\varepsilon(t)$ at ∞, Prop. 2.2.5 has the following probabilistic interpretation:

Proposition 2.3.7

A FSMP $\eta^\varepsilon(t)$, $\varepsilon\in[0,\varepsilon_o]$, is singularly perturbed if and only if the number of ergodic classes at ∞ is discontinuous at $\varepsilon = o$.

The process in Fig. 2.5.1 is singularly perturbed because for $\varepsilon > o$ it has one ergodic class while for $\varepsilon = o$ there are three ergodic classes.

For singularly perturbed FSMP a complete multiple time scale analysis of (2.3.34) is always possible. In effect, it follows from (2.3.13) and (2.3.14) that

$$||\exp\{A_o(\varepsilon)t\}|| = 1$$

and therefore Proposition 2.2.19 guarantees that $A_o(\varepsilon)$ satisfies the MSSNS condition. Let A_{00}, A_{10}, ..., A_{m0} be the sequence of matrices constructed from $A_o(\varepsilon)$ as in section 2.2.1. It is readily seen, by induction, that these matrices are all semistable. First, because $A_{00} = A_o(o)$ is the infinitesimal generator of the stochastically continuous process $\eta^o(t)$, it is semistable. Suppose now that A_{00}, A_{10},..., $A_{\ell-1,0}$ are all semistable then by Corollary 2.2.11 we have

$$\lim_{\varepsilon\downarrow 0} P^\varepsilon(t/\varepsilon^\ell) = \Pi_\ell e^{A_{\ell o} t} \triangleq P_\ell(t) \quad t > o \qquad (2.3.38)$$

where Π_ℓ is a projection such that $\Pi_\ell A_{\ell o} = A_{\ell o} \Pi_\ell = A_{\ell o}$. Notice that $P_\ell(t)$ satisfies the semigroup property $P_\ell(t) P_\ell(\tau) = P_\ell(t+\tau)$, and that $P^\varepsilon(t) \geq o$, $P^\varepsilon(t) \Pi = \Pi$ imply $P_\ell(t) \geq o$ and $P_\ell(t) \Pi = \Pi$. Thus, $P_\ell(t)$ is the transition probability matrix of some FSMP. It then follows from Theorem 2.3.6 that $A_{\ell o}$ is also semistable. This proves the following

Proposition 2.3.8

If $A_o(\varepsilon)$ is the infinitesimal generator of a FSMP then it satisfies the MSST condition.

This result implies that a complete multiple time scale analysis is always possible for signularly perturbed FSMP in the sense that the limit

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

exists for any integer k. In fact it is not difficult to prove that the same result holds for any uniformly stable positive LTI system.

2.3.6. Aggregated Models of Singularly Perturbed Finite State Markov Processes.

In this section we give a probabilistic interpretation of the multiple time scale and reduced-order modelling results presented in section 2.2.2 when the matrix $A_o(\varepsilon)$ is the infinitesimal generator of a singularly perturbed FSMP $\eta^\varepsilon(t)$ taking values in $X = \{1,2,\ldots,n\}$.

From Corollary 2.2.11 and our discussion in section 2.3.5 it follows that

$$\lim_{\varepsilon\downarrow 0} P^\varepsilon(t/\varepsilon^k) = \lim_{\varepsilon\downarrow 0} \exp\{A_o(\varepsilon)t/\varepsilon^k\} = \begin{cases} \Pi_k e^{A_{ko} t} \triangleq P_k(t) & t > o, \; k=0,1,..,m \\ \Pi_{k+1} & t > o, \; k > m \end{cases} \qquad (2.3.39)$$

where $\Pi_o = I$, $\Pi_k = P_o.P_1...P_{k-1}$, $k=1,...,m$, and m is some integer smaller than or equal to the number of ergodic classes of $\eta^o(t)$. The limiting matrices $P_k(t)$, $k=0,1,...,m$, are the transition probability matrices of a collection of FSMP $\eta_k(t)$ which, except for $\eta_o(t) = \eta^o(t)$, are stochastically discontinuous. The stochastically discontinuous nature of the limiting transition probability matrices $P_k(t)$ has a simple interpretation in this context. The time scaling in $P^\varepsilon(t/\varepsilon^k)$ indicates that we focus on rare events, those that are likely to occur only for times of order t/ε^k. In the limit as $\varepsilon\downarrow 0$, all transitions that occur at faster time scales become instantaneous and their net effect on the events that occur at time scale t/ε^k is condensated in the ergodic projection at zero Π_k.

As we have seen in Theorem 2.3.4, each of limiting matrices $P_k(t)$ can be written as

$$P_k(t) = V_k \hat{P}_k(t) U_k \qquad k=1,2,...,m \qquad (2.3.40)$$

where V_k and U_k are matrices obtained from the canonical product decomposition of Π_k (i.e. $\Pi_k = V_k \cdot U_k$) and

$$\hat{P}_k(t) = \exp\{U_k A_{ko} V_k\} \qquad k=1,2,...,m$$

is the transition probability matrix of a stochastically continuous FSMP that we denote by $\hat{\eta}_k(t)$.

According to our discussion in section 2.3.4, the process $\hat{\eta}_k(t)$, with transition probability matrix $\hat{A}_k = U_k A_{ko} V_k$, is an aggregated version of the stochastically discontinuous process $\eta_k(t)$. We will now see that $\hat{\eta}_k(t)$ is also an approximate, aggregated model of $\eta^\varepsilon(t)$ that describes events in $\eta^\varepsilon(t)$ which have non-vanishing probability (as $\varepsilon\downarrow 0$) only in intervals of size $[0, T/\varepsilon^k]$. Each of the projections Π_k, $k=1,...,m$, determines a partition of X as follows:

$$X = (\bigcup_{p=1}^{n_k} E_p^k) \cup E_T^k \qquad k=1,...,m \qquad (2.3.41)$$

where E_p^k, $k=1,...,n_k$, are the ergodic classes at zero of $\eta_k(t)$ and E_T^k is the set of transient states. We have m different partitions of X determined by the process $\eta^\varepsilon(t)$ through the construction indicated above. For each of these partitions define an aggregation mapping as follows:

$$A_k : X \to X_k = \{1,2,...,n_k\} \qquad (2.3.42)$$

$$A_k(i) = j \quad \text{if } i \in E_j^k$$

$$A_k(i) \quad \text{arbitrary if } i \in E_T^k$$

and an aggregated version of $\eta^\varepsilon(t)$:

$$\hat{\eta}_k^\varepsilon(t) \triangleq A_k(\eta^\varepsilon(t)) \qquad (2.3.43)$$

The process $\hat{\eta}_k^\varepsilon(t)$ takes values in $X_k = \{1,2,...,n_k\}$ and it changes only when $\eta^\varepsilon(t)$ jumps from one ergodic class E_i^k to another E_j^k remaining constant while $\eta^\varepsilon(t)$ evolves inside one of the ergodic classes E_i^k. In general, $\hat{\eta}_k^\varepsilon(t)$ will not be a Markov process but, as the following theorem states, it can be approximated by a FSMP in the limit as $\varepsilon\downarrow 0$.

Theorem 2.3.9

Let A_k, $k=1,...,m$ be the aggregation mappings defined in (2.3.42). Then

$$\lim_{\varepsilon\downarrow 0} \eta_k^\varepsilon(t/\varepsilon^k) = A_k(\eta^\varepsilon(t/\varepsilon^k)) = \hat{\eta}_k(t) \qquad (2.3.44)$$

$$k=1,2,...,m$$

the convergence being in the sense of finite dimensional distributions.

This theorem has the following implications:

i)

$$Pr\{\eta^\varepsilon(t) \in E_j^k \,|\, \eta^\varepsilon(o) = \ell \in E_i^k\} = \qquad (2.3.45)$$

$$Pr\{\hat{\eta}_k(\varepsilon^k t) = j \,|\, \hat{\eta}_k(o) = i\} + O(\varepsilon) \quad \forall t \geq o$$

That is, the transition probabilities among classes E_j^k for the process $\eta^\varepsilon(t)$ are asymptotically markovian in the limit as $\varepsilon\downarrow 0$ and they converge to the transition probabilities among states of the process $\hat{\eta}_k(\varepsilon^k t)$.

In terms of the transition probability matrices this can be written as:

$$U_k P^\varepsilon(t) V_k = \exp\{\hat{A}_k \varepsilon^k t\} + O(\varepsilon) \qquad (2.3.46)$$

uniformly on $[0, T/\varepsilon^k]$.

ii) The collection of partitions (2.3.41) determine a classification of events in $\eta^\varepsilon(t)$ into a rarity hierarchy. Transitions among classes E_i^k, $i=1,...,n_k$, may occur on time intervals of the type $[0, T/\varepsilon^k]$ but not on shorter time intervals, i.e.,

$$\lim_{\varepsilon\downarrow 0} Pr\{\eta^\varepsilon(t) \notin E_i^k \text{ some } t \in [0, T/\varepsilon^\ell] \,|\, \eta^\varepsilon(0) \in E_i^k\} = 0 \qquad \forall \ell > k \qquad (2.3.47)$$

In view of Theorem 2.3.9 and the above interpretations we will say that $\hat{\eta}_k(t)$ is an aggregated model of $\eta^\varepsilon(t)$ valid at time scale t/ε^k.

We have thus seen that for a singularly perturbed FSMP it is always possible to construct a sequence of aggregated models each valid at a different time scale. These models form a hierarchy in the sense that the number of states of the aggregated processes $\hat{\eta}_k(t)$ decreases as k goes from 1 to m and consequently the partitions used in the aggregation mappings A_k are increasingly coarse. In effect, $\hat{\eta}_k(t)$ takes values in $X_k = \{1,2,...,n_k\}$ and the number of states n_k is given by:

$$n_k = \dim R(\Pi_k) = \dim R(P_o P_1 \ldots P_{k-1}) \qquad (2.3.48)$$

$k=1,\ldots,m$

it then follows that:

$$n_k = n - \sum_{p=o}^{k-1} \operatorname{rank} A_{ko} \qquad (2.3.49)$$

Also, because

$$\Pi_k \cdot \Pi_\ell = \Pi_\ell \qquad \ell \geq k \qquad k=1,\ldots,m$$

if i and j are two states in X that are aggregated together at stage k then they remain in the same aggregated state at all stages $\ell > k$, i.e.,

$$A_k(i) = A_k(j) \Rightarrow A_\ell(i) = A_\ell(j) \quad \forall \ell > k$$

(strictly speaking this is true only if i and j belong to some ergodic class at both stages k and ℓ. States which belong to some ergodic class at one stage k may become transient states at a posterior stage ℓ in which case the aggregation A_ℓ is not defined for these states. Nevertheless, the hierarchical relation among the mappings A_k, k=1,...,m, remains true). The number of states of $\eta^\varepsilon(t)$ aggregated into a single state of $\hat{\eta}_k(t)$ cannot decrease as k goes from 1 to m. In the next section we present two examples that illustrate the construction of the aggregation mappings and of the aggregated models.

Each of the aggregated models $\hat{\eta}_k(t)$, k=1,...,m, is a simplified model of the process $\eta^\varepsilon(t)$ which accurately describes events that occur at a certain time scale t/ε^k. It is clear, however, from our discussion in section 2.2, that no single aggregated model can accurately describe the evolution of $\eta^\varepsilon(t)$ at all times t. For this it is necessary to combine all them as showed in the following.

Theorem 2.3.10

$$P^\varepsilon(t) = \sum_{k=o}^{m} V_k \exp\{\hat{A}_k \varepsilon^k t\} U_k - \sum_{k=1}^{m} \Pi_k + O(\varepsilon) \qquad (2.3.50)$$

uniformly for $t \in [0,\infty)$.

The transition probability matrix of the singularly perturbed process $\eta^\varepsilon(t)$ is in (2.3.50) uniformly approximated by combining the transition probability matrices of the aggregated models $\hat{\eta}_k(t)$, k=0,1,...,m (here $\hat{\eta}_o(t) \equiv \eta^o(t)$, and in (2.3.50) $\hat{A}_o = A_o(o)$, $V_o = U_o = I$). Theorem 2.3.10 can also be interpreted as an approximation of $\eta^\varepsilon(t)$ by a set of smaller dimensional, independent processes suggesting that events that take place at different time scales can be considered asymptotically independent.

This asymptotic independence of events occuring at different time scales provides the foundation for approximation techniques based on using aggregated models of a complex system. The use of hierarchical aggregation methods in the simplification of filtering and control problems for singularly perturbed FSMP with a large number of states is presently under study and the findings will be reported in [COD 82].

2.3.7 Example

In this section we present an example that illustrates the use of the techniques developed in this chapter for the construction of aggregated models of singularly perturbed FSMP.

Example 2.3.11

Consider the process $\eta^\varepsilon(t)$ depicted in fig. 2.3.1. It has a matrix of transition rates of the form $A_o(\varepsilon) = A_o + \varepsilon B$ with

$$A_o = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The unperturbed process $\eta^o(t)$ with matrix of transition rates A_o is shown in figure 2.3.3. It follows from (2.3.39) that

$$\lim_{\varepsilon \downarrow 0} \exp\{A_o(\varepsilon)\, t/\varepsilon\} = \Pi_1 e^{A_{10} t}$$

where Π_1 is given by:

$$\Pi_1 = \lim_{t \to \infty} e^{A_o t} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$A_{10} = \Pi_1 B \Pi_1 = \begin{bmatrix} -1/8 & -1/8 & 0 & 1/8 & 1/8 & 0 \\ -1/8 & -1/8 & 0 & 1/8 & 1/8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1/8 & 1/8 & 0 & -1/8 & -1/8 & 0 \\ 1/8 & 1/8 & 0 & -1/8 & -1/8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

To determine the aggregated model of $\eta^\varepsilon(t)$ valid at time scale t/ε, $\hat{\eta}_1(t)$, notice that Π_1 determines the following partition of $X = \{1,2,\ldots,6\}$:

$$X = \{1,2\} \cup \{4,5\} \cup \{6\} \cup \{3\}$$

$$\triangleq E_1^1 \cup E_2^1 \cup E_3^1 \cup E_T^1$$

into three ergodic classes and a class with one transient state and that the canonical product decomposition of Π_1 (see section 2.2.3) is:

$$\Pi_1 \triangleq V_1 \cdot U_1 =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad (2.3.51)$$

The aggregated process $\hat{\eta}_1(t)$ thus takes values in $\hat{X}_1 = \{1,2,3\}$ and it has a matrix of transition rates given by:

$$A_1 = U_1 A_{10} V_1 = \begin{bmatrix} -1/4 & 1/4 & 0 \\ 1/4 & -1/4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The state 1 of $\hat{\eta}_1(t)$ corresponds to the set of states $E_1^1 = \{1,2\}$ of $\eta^\varepsilon(t)$ and similarly the state 2 corresponds to the set $E_2^1 = \{4,5\}$ while the state 3 of $\hat{\eta}_1(t)$ corresponds to state 6 of $\eta^\varepsilon(t)$. Only for time intervals of order $1/\varepsilon$, are transition between the classes $\{1,2\}$ and $\{4,5\}$ likely and in the limit as $\varepsilon \downarrow 0$ they follow a markovian law with rate $\varepsilon/2$. Transitions to state 6 are of negligible probability on this time scale.

The next level of aggregation corresponds to the behavior of $\eta^\varepsilon(t)$ at time scale t/ε^2 which is given by:

$$\lim_{\varepsilon \downarrow 0} \exp\{A_o(\varepsilon) t/\varepsilon^2\} = \Pi_2 e^{A_{20} t}$$

where

$$\Pi_2 = \lim_{t \to \infty} \Pi_1 e^{A_{10} t}$$

and $A_{20} = -\Pi_2 B A_o^{\#} B \Pi_2$. The determine the aggregation partition for time scale t/ε^2 and the corresponding aggregated models it is first necessary to compute Π_2. This computation can be simplified by noticing that:

$$\Pi_2 = \lim_{t \to \infty} V_1 \cdot e^{\hat{A}_1 t} \cdot U_1 = V_1 \cdot (\lim_{t \to \infty} e^{\hat{A}_1 t}) \cdot U_1$$

The limit $\hat{\Pi}_2 \triangleq \lim_{t \to \infty} e^{\hat{A}_1 t}$ is the ergodic projection of the aggregated model $\hat{\eta}_1(t)$ which is readily seen to be:

$$\hat{\Pi}_1 = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and combined with (2.3.51) gives

$$\Pi_2 = V_1 \hat{\Pi}_1 U_1 = \begin{bmatrix} 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The aggregation partition determined by Π_2 is

$$X = \{1,2,3,4,5\} \cup \{6\}$$

$$= E_1^2 \cup E_2^2$$

and therefore the aggregated model $\hat{\eta}_2(t)$ takes values in $\hat{X}_2 = \{1,2\}$. The canonical product decomposition of Π_2^2 is given by:

$$\Pi_2 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/4 & 1/4 & 0 & 1/4 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \triangleq V_2 \cdot U_2 \qquad (2.3.52)$$

and to compute the matrix of transition rates of $\hat{\eta}_2(t)$, i.e.,

$$A_2 = U_2 A_{20} V_2 = -u_2 \Pi_2 B A_o^{\#} B \Pi_2 V_2$$

$$= -U_2 B A_o^{\#} B V_2$$

we still need $A_o^{\#}$ which is computed as follows:

$$A_o^{\#} = (A_o + \Pi_1)^{-1} - \Pi_1$$

$$= \begin{bmatrix} -1/4 & 1/4 & 0 & 0 & 0 & 0 \\ 1/4 & -1/4 & 0 & 0 & 0 & 0 \\ 1/4 & 0 & -1/2 & 0 & 1/4 & 0 \\ 0 & 0 & 0 & -1/4 & 1/4 & 0 \\ 0 & 0 & 0 & 1/4 & -1/4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (2.3.53)$$

Finally by combining (2.3.52) and (2.3.53) we get:

$$\hat{A}_2 = \begin{bmatrix} -1/4 & 1/4 \\ 0 & 0 \end{bmatrix}$$

The state 1 of $\hat{\eta}_2(t)$ correspond to the set of states $E_1^2 = \{1,2,3,4,5\}^2$ of $\hat{\eta}_2(t)$ corresponds to the state 6 of $\eta^\epsilon(t)$. For time intervals of order $1/\epsilon^2$ the probability of $\eta^\epsilon(t)$ getting absorbed in state 6 is of order one as $\epsilon \downarrow 0$.

It is clear from the structure of $\eta^\epsilon(t)$ that t, t/ϵ and t/ϵ^2 are the only fundamental time scales. This fact can be readily verified checking that

$$\text{rank } A_o + \text{rank } A_{10} + \text{rank } A_{20} = 5 = \text{rank } A_o(\epsilon)$$

Once the time scale decomposition and aggregation of $\eta^\epsilon(t)$ has been carried out, we can construct a uniform asymptotic approximation to its transition probability matrix $P^\epsilon(t) = \exp\{A_o(\epsilon)t\}$ using Theorem 2.3.10 which only requires the exponentiation of three (2x2) matrices instead of the exponentiation of the (6x6) matrix $A_o(\epsilon)$.

2.4 Aggregation of Diffusion Processes with Multiple Equilibrium Points

2.4.1 Introduction

When a dynamical system with multiple equilibrium points is perturbed by continuous acting wide-band additive noise, it is known that transitions between different equilibrium points occur with probability one. An important problem associated with the analysis of these systems is the statistical characterization of the jump process which represents the transitions between different domains of attraction, in limiting situations. A physical example of a system where multiple equilibrium points are common are interconnected power systems, where the swing equations [EVA 78], Anderson, Fouad [AND 77] represent a dynamical system driven by the power flow equation towards equilibrium. A detailed study of the power flow equation [ARA 80] establishes that there are many possible equilibrium angles in an interconnected network; these angles are defined by a power balance between electrical supply and demand. When the demand fluctuations and unmodeled effects are represented as random, the resulting system has the structure described above.

In this paper, we study the long-term behavior of a subclass of models with multiple equilibrium points and additive white-noise disturbances. These models are characterized by the presence of a parameter ϵ in the description of the process, related to the frequency of transitions between equilibrium points.

The objective of the paper is to obtain a simplified aggregate model of the process. Consistency of the model is established by showing that, in a limiting sense as the parameter ϵ approaches zero, the detailed model converges to the aggregate model. The parameter value ϵ is thus a measure of the accurary of the approximation.

2.4.2 A Diffusion with Small Noise Intensity

Let x(t) be a diffusion process described by the differential equation.

$$dx(t) = \frac{f(x(t))}{a}\, dt + b\,dw \quad (2.4.1)$$

where

$$f(x) = \text{sign } x - x \quad (2.4.2)$$

and a, b > 0, and w is a standard one-dimensional Wiener process.

For any values of a, b, the process given by equation (2.4.1 - 2.4.2) defines a unique probability measure on the space of continuous functions [STR 79]. In addition, the solution to equation (2.4.1) is pathwise unique in the sense of Yamada-Watanabe [YAM 71]. Thus, it can be interpreted as a measurable relation between paths of the Wiener process and the trajectories of x.

For any fixed a and b, there are two domains of attraction, corresponding to the equilibrium points +1, -1. Transitions occur between these domains, as a gets small, the rate of transitions decreases, and the process spends most of its time in small neighborhoods of ±1. As b gets large, the rate of transitions increases. We seek a balance between a and b so that the process, in the limit, has a steady state distribution concentrated on ±1, yet it has a constant rate of transitions between these two points.

The steady state distribution of the process is obtained from the Fokker-Planck equation [WON 71] by solving, for x > 0,

$$(x-1)\frac{dp}{dx} + p + \frac{1}{2}ab^2\frac{d^2p}{dx^2} = 0 \quad (2.4.3)$$

with the boundary conditions

$$p(x) \to 0 \text{ as } x \to \infty \quad (2.4.4)$$

Using the substitution $y = \frac{x-1}{b\sqrt{a}}$, equation (2.4.3) becomes

$$\frac{d\bar{p}}{dy} + \bar{p} + \frac{1}{2}\frac{d^2\bar{p}}{dy^2} = 0$$

$$\bar{p}(y) \to 0 \text{ as } y \to \infty$$

This implies that

$$p(x) = \bar{p}\ \left(\frac{x-1}{b\sqrt{a}}\right) ,$$

so that if $b\sqrt{a} \to 0$, the density $p(x)$ is concentrated on ± 1.

To analyze the transition between regions, let's compute the mean exit time $u(x)$ from $(0,\infty)$, starting at an initial point $x > 0$. From Friedman [FRI 75], this can be obtained from the boundary value problem:

$$\left.\begin{aligned} &\frac{1-x}{a}\ \frac{du}{dx} + \frac{b^2}{2}\ \frac{d^2u}{dx^2} = -1 \\ &u(0) = 0 \\ &\frac{du}{dx} \to 0 \quad \text{as } x \to \infty \end{aligned}\right\} \qquad (2.4.5)$$

It is easy to verify that

$$u(x) = \frac{2\sqrt{a}}{b} \int_0^x e^{\frac{(z-1)^2}{ab^2}} \int_{\frac{z-1}{b\sqrt{a}}}^{\infty} e^{-y^2}\, dy dz \qquad (2.4.6)$$

As $ab^2 \to 0$, the expression for $u(x)$ is nearly constant, except for a small boundary layer near zero. Hence for some $\delta > 0$,

$$u(x) \approx u(1), \quad \delta \leq x < \infty$$

We now proceed to evaluate $u(1)$ for small ab^2.

Lemma 2.4.1: As $ab^2 \to 0$,

$$u(1) \approx \pi\, a^{3/2} b\, e^{1/ab^2}$$

The mean exit time in Lemma 1 suggests the natural scaling choices for a and b. These choices are

$$a^{3/2} b\, e^{1/ab^2} = 1 \qquad (2.4.7)$$

With this scaling between a and b, the frequency of transition between domains of attraction remains constant. Furthermore, by letting ab^2 approach zero, the distribution of $x(t)$ concentrates about the points ± 1. Our objective is to derive an aggregate model for the transitions between domains of attraction as a jump process between two finite states. This result is formalized in Theorem 2.4.2.

Theorem 2.4.2: As $ab^2 \to 0$, with a,b satisfying (2.4.7), the finite dimensional distributions of the process sign $x(t)$ converge to the finite dimensional distributions of a jump process $y(t)$ with two states, ± 1, and transition rates

$$\Pr\{y(t+\Delta) = 1 \mid y(t) = -1\} = \lambda_0 \Delta + o(\Delta)$$

$$\Pr\{y(t+\Delta) = -1 \mid y(t) = 1\} = \lambda_1 \Delta + o(\Delta)$$

where

$$\lambda_0 = \lambda_1 = \frac{1}{2\pi} .$$

Notice that the statement of Theorem 2 is not a statement about the convergence of the $x(t)$ process to a finite jump process. Rather, it looks at the finite state transition process associated with $x(t)$ and establishes convergence to a jump process. It is also possible to study the $x(t)$ process itself. This is the object of Corollary 2.4.3.

Corollary 2.4.3.: Let $y(t)$ be the Markov process in R, defined as follows:

If $y(o) = o$, $y(o+) = 1$ with prob. 1/2, $y(0+) = -1$ with prob. 1/2.

If $y(0) > 0$, then $y(0+) = 1$

If $y(0) < 0$, then $y(0+) = -1$.

For times $t > 0$, the transitions of the $y(t)$ process agree with the jump process defined in Theorem 2.4.2.

As $(ab^2)_n \to 0$ satisfying (2.4.7), the finite dimensional distributions of the $x_n(t)$ process converge to the finite dimensional distributions of $y(t)$.

Corollary 2.4.3 provides an aggregation result. The original multiple equilibrium process can be studied in terms of a finite state jump process.

Note that the results of Theorem 2.4.2 and Corollary 2.4.3 are as strong as can be stated. If one attempts to establish weak convergence of the $x_n(t)$ processes to the $y(t)$ processes, the fact that $x^n(t)$, as a process, has excursions of at least size 1^n from ± 1 preclude weak convergence in the standard spaces one considers.

As a application of this result, consider the process

$$dx(t) = f(x(t))dt + \epsilon\ dw(t) \qquad (2.4.8)$$

with $f(x) = \text{sign } x - x$. The excursions of this process between domains of attration are a rare event, for ϵ small. By compressing the time scale with a transformation

$$\tau = g(\epsilon)t ,$$

we can study the properties of the excursions of the process. Equation (2.4.8) becomes

$$dx(\tau) = \frac{f(x(\tau))}{g(\epsilon)} + \frac{\epsilon}{\sqrt{g(\epsilon)}}\ dw(\tau) \qquad (2.4.9)$$

Comparing (2.4.9) and (2.4.1), we see that (2.4.7) is satisfied when the time scale is given by:

$$g(\epsilon) = \frac{1}{\epsilon}\, e^{-1/\epsilon^2} \qquad (2.4.10)$$

On this time scale, the results of Theorem 2.4.2 and Corollary 2.4.3 apply, so that the finite dimensional distributions of the $x(\tau)$ process converge to those of a jump process as described in Corollary 2.4.3.

2.4.3. The General Scalar Case

Consider the diffusion equation in a one

dimension

$$dx_t = f(x_t)dt + \epsilon\, dw_t \qquad (2.4.11)$$

where $f(x_t)$ is piecewise continuous with a finite number of discontinuities. Define the potential function

$$F(x) = \int_o^x - f(x)\ dx \qquad (2.4.12)$$

Assume that, as $x \to \pm\infty$,

$$F(x) \geq |k\ x^2$$

and that $F(x) \geq o$ for all x. The first assumption guarantees the existence of an ergodic density for (2.4.11), whereas the second assumption represents no loss of generality due to the arbitrariness of the zero reference point.

Equation (2.4.11) can be viewed as the evolution of a Brownian particle in a very steep potential well $F(x)/\epsilon^2$, when a time transformation $\tau = \epsilon^2 t$ is applied. The purpose of this section is to construct a finite state approximation to the evolution of the process at longer time scales, which captures the transitions of the Brownian particle between equilibrium states.

Assume that the function $F(x)$ has a finite number of local maxima and minima. Denote by $x_1, x_3, \ldots, x_{2n-1}$ the local maxima of the function $F(x)$, and $x_o, x_2, \ldots x_{2n}$ the local minima. We will provide an approximation to the evolution of (2.4.11) as a finite-state process whose states are x_i, $i=o,\ldots,2n$. Before we are capable fo doing so, we must perform certain preliminary calculations.

For $\ell = 1,\ldots n-1$, consider the graph of $F(x)$ between $x_{2\ell-1}$ and $x_{2\ell+1}$. Figure 2.4.1 represents a typical graph. We want to compute the transition rates from $x_{2\ell}$ to $x_{2\ell-1}$ and $x_{2\ell+1}$. We proceed as in the previous section, by defining a related boundary value problems.

The mean exit time from $x \in$ in $[x_{2\ell-1}, x_{2\ell+1}]$ is given as the solution of

$$\frac{\epsilon^2}{2}\frac{d^2v}{dx^2} + f(x)\frac{dv}{dx} = -1 \qquad (2.4.13)$$

$$v(x_{2\ell-1}) = v(x_{2\ell+1}) = 0$$

The probability of exiting through $x_{2\ell+1}$ is given as a solution of

$$\frac{\epsilon^2}{2}\frac{d^2u}{dx^2} + f(x)\frac{du}{dx} = o \qquad (2.4.14)$$

$$u(x_{2\ell-1}) = o$$

$$u(x_{2\ell+1}) = 1$$

Solutions to these problems are easy to write in closed form. The solutions are closely related to the scale and speed measures of the diffusion process. In terms of these solutions, we define the transition rates

$$\lambda_{2\ell,2\ell+1} = \frac{u(x_{2\ell})}{v(x_{2\ell})} \qquad (2.4.15)$$

$$\lambda_{2\ell,2\ell-1} = \frac{1-u(x_{2\ell})}{v(x_{2\ell})}$$

Consider now the local graph of $F(x)$ near a local maximum $x_{2\ell-1}$, $\ell=1,..,n$. Figure 2.4.2 illustrates a typical shape of that graph. We are interested in computing the probability of exit from a neighborhood of the point $x_{2\ell-1}$, starting at $x_{2\ell-1}$, as indicated in Figure 2.4.2. The probability of exiting through b satisfies equation (2.4.15) subject to

$$u(a) = o$$

$$u(b) = 1$$

Denote this probability as $p_{2\ell-1,2\ell}$. Define

$$p_{2\ell-1,\ 2\ell-2} = 1 - p_{2\ell-1,2\ell} \qquad (2.4.16)$$

We are now ready to state the main Theorem.

Denote by α the index which minimizes

$$z_\ell = \min\ (F(x_{2\ell-1}) - F(x_{2\ell}),\ F(x_{2\ell+1}) - F(x_{2\ell}))$$

Denote by τ_α the mean exit time from $(x_{2\alpha-1}, x_{2\alpha+1})$ starting at $x_{2\alpha}$. Define a time scale transformation $\tau = \frac{t}{\tau_\alpha}$.

Theorem 2.4.4 In the time scale τ, the finite dimensional distributions of the process x_t converge as $\epsilon \to o$ to the finite dimensional distributions of a stochastically discontinuous finite state Markov process, with states $x_o, x_1 \ldots x_{2n}$. The odd states $x_1, \ldots, x_{2n-1}$ are instantaneous states with transition probabilities given by

$$Pr\{x_{t+} = 2\ell \mid x_t = 2\ell+1\} = p_{2\ell+1,\ 2\ell}$$

$$Pr\{x_{t+} = 2\ell + 2 \mid x_t = 2\ell+1 = 1 - p_{2\ell+1,2\ell}$$

$$Pr\{x_{t+} = k \mid x_t = 2\ell+1\} = o, \quad k \neq 2\ell,\ 2\ell+2$$

The even states are regular states, with transition probabilities given by

$$Pr\{x_{t+\Delta} = 2\ell+1 \mid x_t = 2\ell\} = \lambda_{2\ell,2\ell+1} \cdot \tau\Delta + o(\Delta)$$

$$Pr\{x_{t+\Delta} = 2\ell-1 \mid x_t = 2\ell\} = \lambda_{2\ell,2\ell-1}\,\tau\Delta + o(\Delta)$$

$$\Pr\{x_{t+\Delta} = k \mid x_t = 2\ell,\ x_s = k \text{ or } x_s = k \text{ or } x_x = 2\ell,$$

$$s \in [t,t+\Delta]\} = 0 \quad \text{if } k \neq 2\ell+1 \text{ or } 2\ell-1.$$

The quantities in the description of the finite-state Markov process of Theorem 2.4.4 can be computed exactly. For instance, the exact solution of equation (2.4.13) assuming that $F(x_{2\ell-1}) < F(x_{2\ell+1})$, is given by

$$v(x) = \frac{2}{\epsilon^2} \int_{x_{2\ell-1}}^{x} e^{2F(y)/\epsilon^2} \int_{y}^{x_{2\ell+1}} e^{-2F(z)/\epsilon^2}\, dz\, dy \cdot$$

$$- \frac{2}{\epsilon^2} \int_{x_{2\ell-1}}^{x_{2\ell+1}} e^{2F(y)/\epsilon^2} \int_{y}^{x_{2\ell+1}} e^{-2F(z)/t^2}\, dz\, dy \cdot$$

$$\cdot \frac{\int_{x_{2\ell-1}}^{x} e^{2F(y)/\epsilon^2}\, dy}{\int_{x_{2\ell-1}}^{x_{2\ell+1}} e^{2F(y)/\epsilon^2}\, dy} \qquad (2.4.17)$$

Similarly, equation (2.4.14) is solved by

$$u(x) = \frac{\int_{x_{2\ell-1}}^{x} e^{2F(x)/\epsilon^2}\, dx}{\int_{x_{2\ell-1}}^{x_{2\ell+1}} e^{2F(x)/\epsilon^2}\, dx} \qquad (2.4.18)$$

The expressions (2.4.17) and (2.4.18) can be approximated asymptotically as $\epsilon \to 0$, yielding expressions which depend on the local nature of $F(x)$ around the critical points $x_0, x_1 - x_{2n}$. We do so here for the case where $F(x)$ is twice continuously differentable and $F''(x_\ell) \neq 0$ for $\ell = 0,1,..,2n$.

From the assumptions about F, in a neighborhood of $x_{2\ell-1}$,

$$F(x) \approx F(x_{2\ell-1}) + \frac{1}{2} F''(x_{2\ell-1})\ (x-x_{2\ell-1})^2 + 0(x-x_{2\ell-1})^3 \qquad (2.4.19)$$

From (2.4.16) and (2.4.18), and figure 2.4.2,

$$p_{2\ell-1,2\ell} = \frac{\int_{a}^{x_{2\ell-1}} e^{2F(x)/\epsilon^2}\, dx \approx 1/2}{\int_{a}^{b} e^{2F(x)/\epsilon^2}\, dx}$$

due to the even approximation (2.4.19).

A similar approximation yields

$$u(x_{2\ell}) \approx \frac{A\, e^{2F(x_{2\ell-1})/\epsilon^2}}{B\, e^{2F(x_{2\ell+1})/\epsilon^2} + C\, e^{2F(x_{2\ell-1})/\epsilon^2}}$$

where

$$A = \int_0^\infty e^{F''(x_{2\ell-1})z^2/\epsilon^2}\, dz$$

$$B = \int_0^\infty e^{F''(x_{2\ell+1})z^2/\epsilon^2}\, dz$$

$$C = \int_0^\infty e^{F''(x_{2\ell-1})z^2/\epsilon^2}\, dz$$

The expressions (2.4.17) and (2.4.18) can be approximated asymptotically as $\epsilon \to 0$, yielding expressions which depend on the local nature of $F(x)$ around the critical points $x_0, x_r - x_{2n}$. We do so here for the case where $F(x)$ is twice continuously differentable and $F''(x_\ell) \neq 0$ for $\ell = 0,1,\ldots,2n$.

From the assumptions about F, in a neighborhood of $x_{2\ell-1}$,

$$F(x) \approx F(x_{2\ell-1}) + \frac{1}{2} F''(x_{2\ell-1})\ (x-x_{2\ell-1})^2 + 0(x-x_{2\ell-1})^3 \qquad (2.4.19)$$

and

$$v(x_{2\ell}) \approx \frac{2}{\epsilon^2}\ e^{2F(x_{2\ell-1})/\epsilon^2} \cdot D \cdot E$$

$$D = \int_{-\infty}^{+\infty} e^{-2F(x_{2\ell})/\epsilon^2}\ e^{-F''(x_{2\ell})\ z^2/\epsilon^2}\, dz$$

$$E = \int_0^\infty e F''(x_{2\ell-1}) z^2/\epsilon^2\, dz$$

where $F(x_{2\ell-1}) < F(x_{2\ell+1})$

Using the formula

$$\int_{-\infty}^{\infty} e^{-ax^2} = \sqrt{\pi/a}\ ,$$

we get

$$v(x_{2\ell}) \approx \frac{e^{2/\epsilon^2\ (F(x_{2\ell-1}) - F(x_{2\ell}))}\ \pi}{-F''(x_{2\ell-1})\quad F''(x_{2\ell})}$$

Notice that the expected exit time is a function of the depth of the potential well F(x).

In terms of the scale function τ_α, we have that

$$v(x_{2\ell}) = c\, \tau_\alpha^{\,b}$$

for some constant $b \geq 1$, and $c > 0$. Furthermore

$$\frac{u(x_{2\ell})}{v(x_{2\ell})} \approx \frac{1}{\pi}\; 2(F(x_{2\ell}) - F(x_{2\ell+1}))/\epsilon^2 \cdot \frac{\sqrt{F''(x_{2\ell+1})}}{\sqrt{F''(x_{2\ell})}}$$

If the differences in potential levels were normalized to occur in integer steps (that is, $F(x_{2\ell-1}) - F(s_{2\ell}) = nK$ for some integer n, all ℓ, and some constant k), the resulting transition rates for the approximate finite state Markov process will be of the form

$$\lambda = \lambda(s) = \lambda_o + \lambda_1 s + \ldots + \lambda_n s^n + \ldots$$

where $s = (e^{-2/\epsilon^2})^k$

Finite state Markov processes with this regular dependence on a small parameter s were studied in section 2.2 and 2.3. The resulting Markov process can be further approximated by a hierarchical sequence of simpler models, as described in sections 2.3 and 2.4.

2.4.4 Discussion

The results of section 2.4.3 can be generalized to diffusions with multiple equilibrium points in several dimensions. The major difference in the many dimensional case is that closed form solutions to the partial differential equations of exit times and probabilities are difficult to obtain. However, it is possible to obtain asymptotic estimates for these quantities, as described in Matkowsky and Schuss [MAT 77] and Schuss and Matkowsky [SCH 79]. Using these estimates, a finite state model of the multiple equilibrium process is obtained.

The reason for using instantaneous states in the description of the finite-state Markov process in Theorem 2.4.4 is that setting $\epsilon = o$ in equation (2.4.11) does not accurately capture the evolution of the process (2.4.11) as $\epsilon \to o$ for times of order 1. This is due to the singular nature of the perturbation of the spectrum of the differential operator associated with (2.4.11) [KAT 66]. However, the deterministic flow, together with the instantaneous transition out of the unstable equilibrium states, does capture accurately the limit of the process in (2.4.11) as $\epsilon \to o$, for times of order 1.

The aggregation operation associated with the approximation of Theorem 2.4.4 collapses each domain of attraction onto each equilibrium point x_ℓ; hence, unstable equilibrium points have only their relative domains of attration (in one dimension, only the points) as their aggregate sets.

Another finite state approximation has been proposed by Ventcel and Friedlin [VEN 1970] in order to compute the ergodic distribution of (2.4.11). Their approximation was based on constructing a hierarchical Markov chain based on likelihood of transitions. The approximation developed here contains more information because it includes the effects of the exit times also. In effect, we are developing an approximation which accurately describes the action of the differential operator (2.4.11) when it is restricted to act on the eigenspace corresponding to eigenvalues of magnitude between o and $\frac{1}{\tau_\alpha}$. Due to the discrete nature of the spectrum of (2.4.11), there are at most a finite number of these; the essence of our aggregation result is to identify the eigenprojection which carries a general process into this space, and to establish the right time scale under which the original process, due to its inherent stability approaches this eigenspace.

There are a variety of problems which suggest themselves, based on the results of this paper. In particular, it will be interesting to investigate the consequences of having such approximations available for problems of designing suboptimal controllers and estimations in systems with small dynamical fluctuations.

2.5 Diffusion Approximations of Transfer Lines with Unreliable Links and Finite Storage Elements

2.5.1 Introduction

An important class of systems which arises in manufacturing, chemical processes, computer networks and power systems, is where material moves through a network of unreliable links between storage stations. Transfer lines are networks where all of the storage stations are arranged sequentially; Figure 2.5.1 describes a typical line network. The presence of storage stations serves to compensate for link failures by maintaining the flow upstream and downstream of a failure, thereby descreasing the effect of a failure on the rest of the network. When the operation of a link is modeled as a random process, exact analysis of the flow of material is a difficult task. In this research we develop an aggregate model of the flow through the network based on the physical assumption that the storage capacities are large but finite. This aggregate model is developed as the limit of a sequence of probabilistic models for the flow of material through the line network. Based on this aggregate model, we can approximate properties of the long-term behavior of the line network. Although storage capacities are assumed large, saturation of individual storage stations occurs and is considered in the method here.

Analytical studies of line networks using a probabilistic approach were first studied by Vladzievskii [VLA 52]. A number of authors have studied the flow rates of lines with storages of infinite capacity; some of these are Hunt [HUN 56], Suzuki [SUZ 64], Barlow and Proschan [BAR 75]. Unreliable line networks with one storage station have been studied by a number of authors (Buzacott and Hausifin [BUZ 78], Gershwin and Schick [GER 80a], Gershwin and Berman [GER 81]). These papers have bibliographies of work in this area.

Systems with more storage stations are difficult to analyze because of the complexity of interfaces when storage are either full or empty. For some special systems, Soyster, Schmidt and Rohrer [SOY 79] have obtained exact probabilistic analysis of networks with more than one storage. Gershwin

and Schick's results [GER 80b] are more general, but still limited. Nevertheless, exact analysis of networks with more than one storage is a difficult computational task.

The aggregate model described in this paper is established as a consistent long-term approximation by verifying that an exact model based on the formulation of Gershwin and Schick [GER 80b] converges weakly to the aggregate model in a probabilistic sense. For a discussion of weak convergence of probabilistic measures, the reader should consult Billingsley [BIL 68]. The arguments of convergence depend heavily on the averaging results of Khasminskii [KHA 66a,b].

The aggregate model obtained in this paper is a diffusion process. Diffusion approximations in queueing networks have been studied by a number of authors, notably Borovkov [BOR 65], Iglehart and Whitt [IGL 70], Kobayaski [KOB 74], Reiman [REI 77], Burman [BOR 79] and Harrison [HAR 78]. Although queueing networks feature storages of infinite capacity, many of the techniques used in the analysis of these networks are used here. In particular, the construction of reflected Brownian motion in Harrison and Reimann [HAR 79] provides a valuable introduction to these results.

2.5.2 Mathematical Model of Material Flow

In this paper, we will assume that individual objects are of infinitesimal size, so that when flow of objects through a network is a continuous variable. Using the diagram of figure 2.5.1 as reference, objects flow from an infinite source to an infinite sink across storage stations and unreliable links. The failure and repair processes of the links are assumed to be independent jump processes with constant failure and repair rates. It is also assumed that there is no creation or destruction of objects in the line.

Let x_i, i=1,...k-1 denote the amount of material in storage element i. Let α_j, j=1,...,k denote the state of the link preceding storage element j. The variable α_j can take two values, 1 or 0, indicating respectively that link j is operating or not. By assumption, α_j is a random process, with transition probabilities

$$\Pr\{\alpha_j(t+\Delta) = 1 \mid \alpha_j(t) = 0\} = r_j\Delta + o(\Delta) \quad (2.5.1)$$

$$\Pr\{\alpha_j(t+\Delta) = 0 \mid \alpha_j(t) = 1\} = p_j\Delta + (\Delta)$$

From the theory of representation of jump processes (Davis [DAV 76], we can describe α_j by a stochastic differential equation driven by Poisson processes. Thus, one obtains

$$d\alpha_j(t) = -\alpha_j(t)dF_j(t) + (1 - \alpha_j(t))dR_j(t) \quad (2.5.2)$$

where F_j, R_j, F_i are independent Poisson processes with transition rates p_j, r_j, p_i for any j,i.

Let N_j denote the capacity of storage j. Denote by N_j the flow capacity on link j. The flow rate is assumed to be of maximum capacity whenever possible. Since no objects are created or destroyed, we can describe the storage process by differential equation

$$\frac{dx_i}{dt} = \mu_i\alpha_i - \mu_{i+1}\alpha_{i+1}, \quad 0 < x_i < N_i \quad (2.5.3)$$

$$i=1,...k-1$$

Define the vectors $\underline{x} = (x_1,...,x_k)^T$, $\underline{\alpha} = (\alpha_o,...,\alpha_k)^T$ as the state of the system. Let $s = (\underline{x},\underline{\alpha})$. Equations (2.5.2) and (2.5.3) provide a system of stochastic differential equations which describes the evolution of the probabilistic state s(t) whenever all of the storage elements are away from their limits. However, when a storage element is either empty or full, equation (2.5.3) must be modified so that conservation of flow through the line network applies.

Consider the situation when storage i becomes full. Then, equation (2.5.3) must become

$$\frac{dx_i}{di} \leq 0 \quad (2.5.4)$$

Since the storage element filled up, the incoming flow must be reduced to match the outgoing flow. That is, the rate μ_i is modified so that

$$\mu_i'\alpha_i \leq \mu_{i+1}\alpha_{i+1} \quad (2.5.5)$$

This implies

$$\mu_i' \leq \mu_{i+1}\alpha_{i+1} \text{ if } \alpha_{i+1} = 1$$

Consequently

$$\mu_i' = \min(\mu_i,\mu_{i+1}\alpha_{i+1}) \quad (2.5.6)$$

if α_i=1 and x_i=N_i

Notice that $\alpha_i(t)$ cannot equal 0 if storage i just fills up.

When storage i empties, the outgoing flow b_{i+1} must be reduced to match the incoming flow. That is,

$$\mu_{i+1}' = \min(\mu_{i+1}, \mu_i\alpha_i) \quad (2.5.7)$$

when α_{i+1} = 1 and x_i = 0. Note that α_{i+1} is not zero when storage i empties. When more complex combinations of full and empty storages occur, new produciton rates are defined to enforce conservation of flow. The full stochastic differential equations for the $\underline{x}$ process is given by

$$\frac{dx_i}{dt} = \mu_i(s)\alpha_i - \mu_{i+1}(s)\alpha_{i+1} \quad (2.5.8)$$

where $\mu_i(s)$ satisfies the boundary conditions described by equations (2.5.6), (2.5.7) and their extensions to higher order cases.

2.5.3 Scaling

In order to develop an aggregate model of the system, we will assume that all of the storage capacities are large. Mathematically, we assume

$$N_i = \frac{B_i}{\varepsilon}, \quad i=1,...,k-1$$

for small ε, and constants B_j. Without loss of generality, we will assume that all B_i are equal to 1. Otherwise we can introduce constant to keep track of the relative scaling. Define a scaled variable $y_i(t)$ as the fraction of storage used:

$$y_i(t) = \frac{x_i(t)}{N_i}$$

thus, equation (2.5.8) becomes

$$N_i \frac{d}{dt} y_i = \alpha_i \mu_i(s) - \alpha_{i+1}\mu_{i+1}(s) \qquad (2.5.9)$$

Equation (2.5.9) represents a random evolution for the $\underline{y}(t)$ process, with a discontinuity in drift when the process exits the open domain $D = (0,1)^{k-1}$. Aggregation or random evolutions has been studied by a number of authors; Hersh [HER 74] has compliled a comprehensive survey of the work in that area. However, none of that work can incorporate the local discontinuity of the drift as the process reaches the boundary.

The process $\underline{y}(t)$ has coordinates with values between 0 and 1, representing the fraction of capacity used in storage. The boundary effects described in section 2.5.2 will occur whenever one of the coordinates of $\underline{y}(t)$ is either 0 or 1. Let denote he time of first exit of the $\underline{y}(t)$ process from its interior. That is,

$$\gamma(\omega) = \inf\{t > 0 \mid y(t,\omega) \notin D\}$$

We will develop an approximation to the $\underline{y}(t)$ process until its time of first exit from the domain D.

Denote by $\underline{z}(t)$ the process in $\mathbb{R}^{k-1}$ whose evolution described by

$$N_i \frac{dz_i}{dt} = \alpha_i \mu_i - \alpha_{i+1}\mu_{i+1} \qquad (2.5.10)$$

$$z_i(0) = y_i(0)$$

where μ_i are the constant flow rates when y is in D.

Note that the sample paths of the $\underline{z}(t)$ process agree with the sample paths of the $\underline{y}(t)$ process until time $\gamma(w)$. The process $\underline{z}(t)$ represents the evolution of the normalized storage process if no boundary adjustments were made.

Define τ as εt. In this time scale equation (2.5.10) becomes

$$\frac{dz_i}{d\tau}(\tau) = \alpha_i \mu_i - \alpha_{i+1}\mu_{i+1}$$

Notice that the Markov process $(\underline{z}(\tau), \underline{\alpha}(\tau))$ has components varying in two different time scales. The $\underline{z}(\tau)$ process has variations on the slow scale τ, and the $\underline{\alpha}(\tau)$ process has transitions in the t scale. This separation of scales is a consequence of the assumption that N_i is large, and will be exploited to obtain aggregate models. In the next section, we will establish that the process $\underline{z}(\tau)$ can be approximated by a Markov process which does not depend on the jump process $\underline{\alpha}(\tau)$; this approximation can be used in computing expectations of the process $\underline{z}(\tau)$.

2.5.4. AGGREGATION

The $\underline{\alpha}(\tau)$ process described in equation (2.5.2) is a jump process with a finite number of states. Each of the components has independent transitions, and is strongly ergodic. The ergodic measure of the jth component is

$$\bar{P}_j(\alpha_j) = \frac{(1-\alpha_j)p_j + \alpha_j r_j}{p_j + r_j} \qquad (2.5.11)$$

The overall ergodic measure is given by

$$\bar{P}(\underline{\alpha}) = \prod_{j=1}^{k} \bar{P}_j(\alpha_j)$$

As the parameter ε approaches zero, the separation between the time scales τ and t increases. Hence, more transitions of the $\underline{\alpha}$ process occur between significant changes in the $\underline{y}(\tau)$ process. One would expect that a good approximation for the evolution of the $\underline{z}(\tau)$ process would be provided by the expected drift, in terms of the ergodic measure of the $\underline{\alpha}(t)$ process. This result is established in this section.

Define the average drift $\bar{F}_i$ as

$$\bar{F}_i = \sum_{\underline{\alpha}} (\alpha_i \mu_i - \alpha_{i+1}\mu_{i+1}) \bar{P}(\underline{\alpha}) \qquad (2.5.12)$$

Combining equations (2.5.11) and (2.5.12) yields

$$\bar{F}_i = \frac{r_i \mu_i}{r_i + p_i} - \frac{r_{i+1}\mu_{i+1}}{r_{i+1} + p_{i+1}}$$

Define $\underline{z}^0(\tau)$ as

$$z_i^0(\tau) = z_i(0) + \bar{F}_i \tau$$

The processes $\underline{z}^0(\tau)$ represents the average evolution of the $\underline{z}(\tau)$ process. The next results specify the accuracy of this approximation.

Theorem 2.5.1

Let T be an arbitrary finite positive number. Consider the processes $\underline{z}(\tau)$ and $\underline{z}^0(\tau)$, $0 \leq \tau \leq T$. As $\varepsilon \to 0$, the process $\underline{z}(\cdot)$ converges uniformly in the mean to $\underline{z}^0$. That is,

$$\lim_{\varepsilon \to 0} \sup_{0<\tau<T} E\{|\underline{z}(\tau) - \underline{z}^0(\tau)|\} = 0$$

Proof

The proof is a straightforward application of Theorem 1.1 of Khasminskii (KHA 66).

The fact that the rates μ_i are constant enables us to establish a stronger result than uniform convergence in the mean. We can establish that $\underline{z}(\cdot)$ converges to $\underline{z}^0(\cdot)$ almost surely, and examine the distribution of its deviations.

Theorem 2.5.2
Under the conditions of Theorem 4.1, the process $\underline{z}(\cdot)$ converges to the process $\underline{z}^0(\cdot)$ almost surely as $\varepsilon \to 0$. Furthermore, let

$$V_i(\tau) = \frac{1}{\sqrt{\varepsilon}} \left(z_i(\tau) - z_i^0(\tau)\right)$$

The process $\underline{v}(\tau)$ converges weakly to a zero-mean Wiener process $\underline{w}$ with covariance

$$E\underline{w}(\tau)\underline{w}^T(s)\} = \underline{\Sigma} \min(\tau, s)$$

$$\Sigma_{ii} = 2\left\{\frac{\mu_i^2 p_i r_i}{(p_i + r_i)^3} + \frac{\mu_{i+1}^2 p_{i+1} r_{i+1}}{(p_{i+1} + r_{i+1})^3}\right\} \tag{2.5.13}$$

$$\Sigma_{i+1,i} = \Sigma_{i,i+1} = \frac{-2\mu_{i+1}^2 p_{i+1} r_{i+1}}{(p_{i+1} + r_{i+1})^3}$$

$$\Sigma_{ij} = 0, \quad |i-j| \geq 2$$

Theorems 2.51 and 2.52 define aggregate models for the evolution of the $\underline{z}(\tau)$ process. These aggregate models are established as consistent by the convergence of the true process as $\varepsilon \to 0$. The models are developed in the slow time scale $\tau = \varepsilon t$; they are most useful when the line network is unbalanced in the mena. That is, when the average drift in the system, $\underline{\bar{F}}$, is of order 1.

When all of the drifts in the system, $\bar{F}_i$, are of order ε, the approximation given by these theorems is not of much use, because no significant trends occur in times of order $1/\varepsilon$. Such cases are referred to as balanced line networks. However, in a still slower time scale, an aggregate model can be obtained.

Let $\tau_1 = \varepsilon^2 t$ be a slow time scale. In the τ_1 scale, equation (2.5.10) becomes

$$\frac{d}{d\tau_1} z_i(\tau_1) = \frac{-\alpha_{i+1}(\tau_1)\mu_{i+1} + \alpha_i(\tau_1)\mu_i}{\varepsilon} \tag{2.5.14}$$

Assume additionally that

$$\bar{F}_i = \varepsilon f_i, \quad i = 1, \ldots, k-1$$

Then, we can write (2.3.14)

$$\frac{d}{d\tau_1} z_i = \frac{\alpha_i\mu_i - \alpha_{i+1}\mu_{i+1} - \varepsilon f_i}{\varepsilon} + f_i$$

Let Q denote the infinitesimal generator of the Markov process $\underline{\alpha}(\tau)$. The operator Q can be viewed as a singular matrix mapping $R^{2k} \to R^{2k}$. Denote vectors in R^{2k} by the functions $g(\underline{\alpha})$. Suppose that

$$g_i(\underline{\alpha}) = -\alpha_{i+1}\mu_{i+1} + \alpha_i\mu_i - \varepsilon f_i$$

$$h_i(\underline{\alpha}) \frac{\mu_i\alpha_i}{r_i + p_i} - \frac{\mu_{i+1}\alpha_{i+1}}{r_{i+1} + p_{i+1}}$$

By its definition, the matrix Q can be expressed as

$$Qh(\underline{\alpha}) = \sum_i r_i(h(\underline{\alpha}_i^c) - h(\underline{\alpha}))$$

$$\alpha_i = 0$$

$$+ \sum_i p_i(h(\underline{\alpha}_i^c) - h(\underline{\alpha}))$$

$$\alpha_i = 1$$

$$= \sum_j \{r_j(1-\alpha_j) + p_j\alpha_j\}\{h(\alpha_j^c) - h(\alpha)\}$$

where

$$\underline{\alpha}_j^c = (\alpha_1, \ldots, \alpha_{i-1}, 1-\alpha_i, \alpha_{i+1}, \ldots \alpha_k)$$

Then,

$$Qh_i(\underline{\alpha}) = (r_i(1-\alpha_i) + p_i\alpha_i)\frac{\mu_i(1-\alpha_i)}{r_i + p_i} - \frac{\mu_i\alpha_i}{r_i + p_i}$$

$$- \frac{(r_{i+1}(1-\alpha_{i+1}) + p_{i+1}\alpha_{i+1}}{} \left\{\frac{\mu_{i+1}(1-\alpha_{i+1})}{r_{i+1} + p_{i+1}} - \frac{\mu_{i+1}\alpha_{i+1}}{r_{i+1} + p_{i+1}}\right\}$$

$$= \frac{\mu_i r_i}{r_i + p_i} - \mu_i\alpha_i + \mu_{i+1}\alpha_{i+1} - \frac{\mu_{i+1}r_{i+1}}{r_{i+1} + p_{i+1}}$$

$$= -g_i(\underline{\alpha})$$

Consider now an arbitrary bounded function $h(z)$ in $C^2(\mathbb{R}^{k-1})$, the space of real valued, twice continuously differentiable functions of $\mathbb{R}^{k-1}$. Denote by L the infinitesimal generator of the Markov process $(\underline{z}, \underline{\alpha})$ in the τ_1 time scale. Then

$$L = \frac{Q}{\varepsilon^2} + \frac{1}{\varepsilon}\sum_{i=1}^{k-1}(\alpha_i\mu_i - \alpha_{i+1}\mu_{i+1} - \varepsilon f_i)\frac{\partial}{\partial z_i} + \sum_{i=1}^{k-1} f_i \frac{\partial}{\partial z_i}$$

Let $\bar{L}$ denote the diffusion operator

$$\bar{L} = \sum_{i=1}^{k-1} f_i \frac{\partial}{\partial z_i} + \frac{1}{2}\sum_{i=1}^{k-1}\sum_{j=1}^{k-1}\Sigma_{ij}\frac{\partial^2}{\partial z_i \partial z_j} \tag{2.5.15}$$

where $\underline{\Sigma}$ is defined in Theorem 2.5.2.

Notice that $\bar{L}$ is the generator of a pathwise

unique strong Markov process in $\mathbb{R}^{k-1}$ (Stroock-Varadhan [STR 79]).

Theorem 2.5.3

When the link network is nearly balanced, the process $\underline{z}(\tau_1)$, $0 \leq \tau_1 \leq T$, for arbitrary finite T, converges weakly as $\epsilon \to 0$ to the unique diffusion Markov process v whose infinitesimal generator is $\bar{L}$. Moreover, all the moments of $\underline{z}$ converge to the moments of $\underline{v}$ as $\epsilon \to 0$.

The proof of these results is a direct application of Theorem 1 in Papanicolaou-Kohler, [PAP 74] because the $\underline{\alpha}$ process is ergodic, hence it is strongly mixing.

2.5.5. Diffusion Approximations with Boundary Conditions

The results of section 2.5.4 provide an approximation to the normalized storage process $\underline{y}(t)$ until its time of first exit from the interior of the region D. In this section, those approximations will be extended to cover arbitrary intervals of time. In this case, the boundary conditions described in section 2.5.2 have to be explicitly considered.

Consider the process $\underline{z}(\tau)$ defined in section 2.5.4. Define the compensating processes $C_0(t,z)$, $C_1(t,z)$ for any continuous real valued function z as

$$C_0(t,z) = \min_{0 \leq s < t} \{0, z(s)\}$$

$$C_1(t,z) = \max_{0 \leq s \leq} \{0, z(s)-1\}$$

The functions $C_0(t,z)$ and $C_1(t,z)$ represent the excesses of the function z(t) outside the interval [0,1]. Hence, for any function z(t), we can define the compensated function $z^1(t)$ as

$$z^1(t) = z(t) - C_0(t,z) - C_1(t,z) \qquad (2.5.16)$$

The function $z^1(t)$ does not take its values in the unit interval, because the effect of two compensating processes drive the new function outside. However, one can define a sequence of functions $z^j(t)$ inductively as

$$z^{j+1}(t) = z^j(t) - C_0(t,z^j(t)) - C_1(t,z_j(t)) \qquad (2.5.17)$$

For any bounded interval [0,T], and any continuous function z(t) on [0,T], $z_j(t)$ is a continuous function.

Consider the process $\underline{z}(t)$ defined in section 2.5.2. The failure-repair process $\underline{\alpha}(t)$ is a Markov jump process which describes the volution of $\underline{z}(t)$. Since the rates of evolution of $\underline{z}(t)$ are constant except for the effects of $\underline{\alpha}$, the probabilistic distribution of increments of $\underline{z}(t)$ is independent of the value of $\underline{z}(t)$; that is,

$$P_r\{\underline{z}(t+\Delta) - \underline{z}(t) \in B \mid \underline{z}(t), \ \underline{\alpha}(t)\} =$$

$$P_r\{\underline{z}(t+\Delta) - \underline{z}(t) \in B \mid \underline{\alpha}(t)\}$$

The process $\underline{y}(t)$ has a similar property, except for the effects of the boundary conditions. We would like to incorporate the effects of these boundary conditions as compensating processes, in the manner of equations (2.5.16) and (2.5.17). This is the purpose of the next result.

Consider an arbitrary sample path $\underline{z}(t)$, $t \in [0,T]$. Define the sequence of times t_j as

$$t_0 = \inf\{t \mid C_0(t,z_j) \neq 0, \text{ or } C_1(t,z_j) \neq 0 \text{ for some } j\}$$

$$t_i = \inf\{t \mid C_0(t,z_j^i) \neq 0, \text{ or } C_1(t,z_j^i) \neq 0 \text{ for some } j\} \qquad (2.5.18)$$

The times t_i represent times when the compensated processes $_i\underline{z}^i$ would require additional compensation to stay in $\bar{D}$. Now, define an interger valued function on the time sequence t_i as

$$n(t_i) = \max_{1 \leq j \leq k-1} \{j \mid z_j^i(t_i) = 1 \text{ and } C_1(t,z_j^i) > 0, \ t > t_i\} \qquad (2.5.19)$$

If the set of such indices j is empty, let $n(t_j)$ be

$$n(t_i) = \min_{1 \leq j \leq k-1} \{k-1+j \mid z_j^i(t_i) = 0 \text{ and } C_0(t,z_j^i) < 0, \ t > t_i\} \qquad (2.5.20)$$

Notice that, if only one storage level reaches the boundary at time t_i, then $n(t_i)$ identifies that storage, and indicates whether it is empty or full. Whenever two or more storage levels reach the boundary simultaneously at time t_i, the function $n(t_i)$ selects a storage by the following rule:

Select the storage which saturated farthest downstream. If there is no storage which is saturated, the select the storage which emptied farthest upstream.

This selection rule serves to ensure that the compensation process at any one time requires no more than 2k iterations. This is because the effects of saturation propagate upstream, whereas the effects of starvation propogate downstream.

We can now define a sequence of compensated functions $\underline{z}^i$ recursively, as

$$\underline{z}^0(t) = \underline{z}(t)$$

$$\underline{z}_j^{i+1}(t) = z_j^i(t) - C_0(t,z^i)I\{n(t_i) = k-1+j\} - C_i(t,z_j^i)I\{n(t_i) = j\} + C_0(t,z_{j-1}^i)I\{n(t_i) = k+j-z, \ j \neq 1\} + C_1(t,z_{j+1}^i)I\{n(t_i) = j+1, \ j \neq K-1\} \qquad (2.5.21)$$

Equation (2.5.21) expresses the conservation of flow relations. When storage i is full, the flow rate through storage i must be constrained to match its output rate. This effect is modeled by the compensating process $C_1(t,z_i^j)$. However, conservation of flow dictates that the material which does not flow through storage i will accumulate in storage i-1. This is modeled by the coupling term $C_1(t,z_{i+1}^j)$.

The basic claim is that we can express the normalized storage process $\underline{y}(t)$ in terms of the sequence of compensated processes $\underline{z}^i(t)$. Specifically, the result is:

Theorem 2.5.4

For any finite t,

$$\underline{y}(t) = \lim_{i\to\infty} \underline{z}^i(t)$$

where $\underline{z}^i(t)$ is defined by equations (2.5.17-2.5.21).

The proof of this result entails establishing some simple properties of the recursion (2.5.17) - (2.5.21). These properties are summarized in the following lemmas:

Lemma 2.5.5

For any continuous trajectory $\underline{z}(t) \in \mathbb{R}^{k-1}$, $t\in[0,T]$, there is a finite integer $j(t)$ such that

$$\underline{z}^{j(T)}(t) = \underline{z}^{\infty}(t) \text{ for all } t\in[0,T].$$

Lemma 2.5.6

The mapping $G: z \to \lim z^i(\cdot)$ is a continuous map from $C\{[0,T]; \mathbb{R}^N\}$ for any finite T.

Notice that Theorem 2.5.4 and Lemma 2.5.6 establish that the trajectories of the normalized storage process with boundary conditions are a continuous map of the trajectories of the process without boundary. Furthermore, Theorems 2.5.2 and 2.5.3 establish weak convergence, as $\varepsilon \to 0$, of the process without boundary to a diffusion process. Denote this diffusion process as $\underline{v}(t)$, $0 \le t \le T$. Then, Theorem 5.1 of [BIL 68] establishes that, for an arbitrary interval, the process $\underline{y}(t)$ converges weakly as $\varepsilon \to 0$ to the process with support in $C\{[0,T],\bar{D}\}$, whose distributions are given from the map G of Lemma 2.5.6. This discussion can be formalized as

Theorem 2.5.7

Assume that the process $\underline{z}(\cdot;\varepsilon)$ converges weakly in $C\{[0,T]; \mathbb{R}^{k-1}\}$ as $\varepsilon \to 0$ to $\underline{v}(\cdot)$, a diffusion process. Then, the process $\underline{y}(\cdot;\varepsilon)$ converges weakly in $C\{[0,T]; \mathbb{R}^{k-1}\}$ to the process $G(\underline{v})$.

The compensating processes C_0 and C_1 are related to the time that the process $\underline{y}(t;\varepsilon)$ spends on the boundary D. Specifically, for a fixed trajectory of $\underline{z}(t;\varepsilon)$, we can write the process $\underline{y}(t;\varepsilon)$ as

$$\underline{y}(t;\varepsilon) = \underline{z}^i(t;\varepsilon), \quad t < t_i$$

In coordinates, we can write this relationship as

$$y_j(t;\varepsilon) = z_j(t;\varepsilon) - \sum_{\ell=1}^{i} C_0(t,z_j^{\ell-1})\, I\{n(t_{\ell-1}) = k-1+j\}$$

$$- \sum_{\ell=1}^{i} C_1(t,z_j^{\ell-1})\, I\{n(t_{\ell-1}) = j\}$$

$$+ \sum_{\ell=1}^{m} C_0(t,z_{j-1}^{\ell-1})\, I\{n(t_{\ell-1}) = k+j-2,\ j \neq 1\}$$

$$+ \sum_{\ell=1}^{m} C_1(t,z_{j+1}^{\ell-1})\, I\{n(t_{\ell-1}) = j+1,\ j \neq k-1\} \qquad (2.5.22)$$

It is easy to establish inductively that the first sum is constant except when $y_j(t;\varepsilon) = 0$. Similarly, the second, third and fourth terms are constant except when $y_j(t;\varepsilon) = 1$, $y_{j-1}(t;\varepsilon) = 0$, and $y_{j+1}(t;\varepsilon) = 1$ respectively. Hence, we can represent $\underline{y}_j(t;\varepsilon)$ implicitly as

$$y_j(t;\varepsilon) = z_j(t;\varepsilon) + U_j^0(t) - U_j^1(t) - U_{j-1}^0(t) + U_{j+1}^1(t)$$

where $U_j^0(t)$, $U_j^1(t)$ are increasing processes which increase only when $y_j(t) = 0$ or $y_j(t) = 1$.

When the process $\underline{y}(t;\varepsilon)$ is nearly balanced, the limit process becomes a diffusion process with instantaneous oblique reflection at the boundary ∂D. This follows because of the construction of the compensating processes and their relation to the local time of diffusion processes, as defined in Watanabe [WAT 71].

The directions of reflection can be obtained directly from equation (5.9). For instance, on the face

$$y_j = 0$$

the equations for the evolution of $y_j(t,\varepsilon)$ are

$$\left.\begin{aligned} \frac{dy_\ell}{dt}(t,\varepsilon) &= \frac{dz_\ell}{dt}(t,\varepsilon) \quad \ell \neq j,\ j+1 \\ \frac{dy_j}{dt}(t,\varepsilon) &= \frac{dz_j}{dt}(t,\varepsilon) + \frac{dU_j^0}{dt}(t) \\ \frac{dy_{j+1}}{dt}(t;\varepsilon) &= \frac{dz_{j+1}}{dt}(t,\varepsilon) - \frac{d}{dt}\ U_j^0(t) \end{aligned}\right\}$$

Hence, the direction of oblique reflection on the face $y_j = 0$ is given by the effect of the compensating processes U_j^0, corresponding to reflection in the direction

$$\underline{d} = (0,\ldots,0,\underset{j-1}{+1},-1,0,\ldots,0)$$

When the transfer line is nearly balanced, the limiting process spends no scaled time on the boundary, on the time scale $\tau = \varepsilon^2 t$. However, the limiting process has a local time function at the boundary, which can be used to obtain an expression for the real t spent on the boundary. The characterization will be useful in later sections, when we evaluate expressions for the throughput of the

transfer line. From equation (2.5.8), the equation for throughput rate (in normalized units and scaled time) is given by

$$T(\tau) = \frac{1}{\tau}\{\int_0^\tau \mu_k \alpha_k \, dt - U_{k-1}^o(\tau)\}$$

The quantity $\frac{U_{k-1}^o(\tau)}{\tau}$ represents the average lost production rate due to starvation of the last machine.

The result expressed in Theorem 2.5.7 defines a reflected diffusion process as the limit process. This process is defined uniquely in the weak sense, in terms of a continuous mapping on the sample paths of a standard diffusion process. This construction depends strongly on three assumptions: constant flow rates on links, constnat failure and repair rates, and the geometry of line networks. When any of these three conditions are violated, the limit process must be constructed using a different argument. This is a nontrivial problem because of the lack of smoothness of the domain $\bar{D}$, a closed unit cube.

2.5.6 Approximation with Level Dependent Failure Rates

In this formulation of the previous sections, the failure and repair processes of the machines in the transfer line are independent of the levels of storage. However, a common practice in manufacturing networks is to turn off machines which are either starved or blocked, thereby eliminating the possibility of a machine failure during intervals of time when tat machine is not processing any material. A mathematical model with these properties is described in Gershwin and Schick [GER 80b].

The main difference in such a model is to introduce a feedback path from the continuous storage level $\underline{x}$ to the discrete state process $\underline{\alpha}$, occuring when $\underline{x}$ reaches its boundary. In terms of the normalized storage process $\underline{y}$, there are two situations where a machine is on, but not processing any material. The first situation, called blockage, occurs when machine i+1 is off, and storage i full. Then, the adjustment process described by (2.6) yields $\mu_i = 0$. Hence, machine i is assumed not to fail.

The second situation occurs when machine i is off, and storage i is empty. The adjustment process for machine i+1 yields

$$\mu_{i+1} = 0.$$

We call such a machine starved, and assume it cannot fail.

The equations for the $\underline{\alpha}$ process can be modified to describe starvation and blocking as follows.

$$d\alpha_i = (1-\alpha_i)dR_i + \alpha_i(1-I\{\mu_i=0\})dF_i \qquad (2.5.23)$$

where the last term has been modified to prevent failures during non-production intervals. The function $\mu_i(\underline{y},\underline{\alpha})$ depends on the complete state of the system in memoryless, fashion, given by the adjustment rules for conservation of flow.

Essentially, the description of the y process is decomposed into an internal description, describing the evolution of the process away from the boundary, and a boundary description which illustrates what happens to the process near a boundary. Our purpose in this section is to show that the modified $(\underline{y},\underline{\alpha})$ process given by (2.5.22) and (2.5.23) converges weakly to the same diffusion process given in Theorem 2.5.7.

Throughout this section, we assume that the transfer line is nearly balanced, so that the appropriate time scale τ is $\varepsilon^2 t$. Let $\underline{y}^{1\varepsilon}(\tau;\varepsilon)$ denote the scaled process defined in section 5, and $P^{1\varepsilon}$ the induced probability measure on $C\{[0,T]; \mathbb{R}^{k-1}\}$. Similarly, denote by $\underline{y}^{2\varepsilon}(\tau;\varepsilon)$ the resulting scaled process ehn starvation and blockage affect the probability rates, and $P^{2\varepsilon}$ it corresponding measure. The main result of this section is stated in the following theorem.

Theorem 2.5.8

In the topology of weak convergence on $C\{[0,T], \mathbb{R}^{k-1}\}$

$$\lim_{\varepsilon\to 0} P^{1\varepsilon} = \lim_{\varepsilon\to 0} P^{2\varepsilon}$$

Basically, Theorem 2.5.8 is a consequence that, as $\varepsilon \to 0$, the process spends less percent of the time at the boundary. The evolution of $\underline{y}^{1\varepsilon}$ and $\underline{y}^{2\varepsilon}$ are identical outside the boundary, and they leave the boundary in the same direction. Hence, as the time spent on the boundary decays, the two processes approach each other. The differences in the behavior of the α processes associated with $\underline{y}^{1\varepsilon}$ and $\underline{y}^{2\varepsilon}$ do not appear in the slow time scale $\tau = \varepsilon^2 t$. If the transfer line was not nearly balanced, the appropriate time scale would be $\tau = \varepsilon t$, and these differences would be noticeable in the approximate model.

Theorem 2.5.8 has served additionally to establish that the limiting process is instantaneously refelcted at the boundary ∂D, by showing that the Lebesgue measure of the occupation time has expectation zero. This is consistent with the representation of the limiting process as instantaneously reflected Brownian motion.

2.5.7 Ergodic Distribution of Two Machine Transfer Lines using Discussion Approximations

The simplest network one can construct consists of two unreliable links with a storage center in the middle, connecting an infinite source to an infinite sink, as depicted in figure 2.5.2. In the context of manufacturing networks, many authors have studied the long term behavior of this simple network. Gershwin and Schick [GER 80b] provide the basic equations for the description of the Markov processes $(x(t),\alpha_1(t),\alpha_2(t))$.

Assume that the flow rates on each link is equal to 1, that is

$$\mu_1 = \mu_2 = 1.$$

Then, the basic flow equation for the storage process is

$$\frac{dx}{dt} = (\alpha_1 - \alpha_2)$$

when the storage buffer is neither empty nor full. Assuming that the capacity of the storage process N is large, the normalized storage equation is

$$N \frac{dy}{dt} = (\alpha_1 - \alpha_2)$$

$$y = \frac{x}{N}$$

The processes α_i are jump processes with failure and repair rates p_i, r_i respectively, $i = 1,2$.

In Gershwin and Schick [GER 80a], this model is studied in detail, obtaining an exact expression for the ergodic probability distribution of the (x,α_1, α_2) process. We will assume that starvation and blockage prevent machines from failing as in section 2.5.6.

Let $N = 1/\varepsilon$, and $\tau = \varepsilon^2 t$. Then,

$$\frac{dy}{d\tau} = \frac{\alpha_1 - \alpha_2}{\varepsilon}$$

Assume that

$$\frac{r_1}{r_1 + p_1} - \frac{r_2}{r_2 + p_2} = \varepsilon m \qquad (2.5.24)$$

Equation (2.5.24) indicates that the network is nearly balanced, validating the use of the $\varepsilon^2 t$ time scale.

From Gershwin and Schick [GER 80a], the marginal ergodic distribution of the x(t) process is given by:

$$p(x < d) = \int_0^d g(x)dx + P\{x=0\} + P\{x=c\} \cdot I\{c < d\}$$

$$p(x = 0) = C \frac{(r_1+r_2)}{p_2} \left(\frac{1}{r_1} + \frac{1}{p_1+p_2}\right)$$

$$p(x = N)\ C\ e^{\lambda N} \frac{(r_1+r_2)}{p_1} \left(\frac{1}{r_2} + \frac{1}{p_1+p_2}\right)$$

$$g(x) = C\ e^{\lambda x} \left(1 + \frac{r_1+r_2}{p_1+p_2}\right)^2$$

$$\lambda = (p_2 r_1 - p_1 r_2) \left(\frac{1}{p_2+p_1} + \frac{1}{r_2+r_1}\right)$$

Define $\hat{\lambda} = \lambda/\varepsilon$. Then, a simple integral establishes

$$C^{-1} = \frac{r_1 + r_2}{p_2} \left(\frac{1}{r_1} + \frac{1}{p_1+p_2}\right) + \frac{(r_1+r_2)}{p_1} \left(\frac{1}{r_2} + \frac{1}{p_1+p_2}\right) e^{\lambda N}$$

$$+ \frac{1}{\lambda} (e^{\lambda N} - 1) \left(1 + \frac{(r_1+r_2)}{(p_1+p_2)}\right)^2$$

The ergodic distribution of $\underline{y}(t)$ is given in the following equations:

$$P\{y = 0\} = P\{x = 0\}$$

$$P\{y = 1\} = P\{x = N\}$$

$$P\{y \in [y,y+dy)\} = \frac{1}{\varepsilon} g(y/\varepsilon)\ dy$$

Let $N = 1/\varepsilon$. As $\varepsilon \to 0$, λ is of order ε, hence we have

$$\lim_{\varepsilon \to 0} C^{-1} = \lim_{\varepsilon \to 0} \frac{1}{\lambda} (e^{\hat{\lambda}} - 1) \left(1 + \frac{r_1+r_2}{p_1+p_2}\right)^2$$

Thus,

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} g(y/\varepsilon) = \frac{\hat{\lambda}}{e^{\hat{\lambda}} - 1} e^{\hat{\lambda} y}.$$

Furthermore,

$$\lim_{\varepsilon \to 0} P\{y = 0\} = \lim_{\varepsilon \to 0} P\{y = 1\} = 0 \qquad (2.5.25)$$

because λ is of order ε, by assuming that the transfer line is nearly balanced.

The ergodic distribution indicated by equation (2.5.25) reflects the long-term behavior of the $\underline{z}(t)$ process. The diffusion approximation v( ) generated in section 2.5.5 for the balanced line case has as its infinitesimal generator

$$L = m \frac{\partial}{\partial v} + \frac{1}{2} \sigma \frac{\partial^2}{\partial v^2}$$

$$\sigma = \frac{2p_1 r_1}{(p_1 + r_1)^3} + \frac{2p_2 r_2}{(p_2 + r_2)^3}$$

with domain

$$D(L) = \{f \mid f \text{ is bounded, twice differentiable on } (0,1)$$

and

$$\frac{\partial f}{\partial v}(0) = \frac{\partial f}{\partial v}(1) = 0\}.$$

Hence, the ergodic distribution p(v) is given by

$$-m \frac{\partial p}{\partial v} + \frac{1}{2} \sigma \frac{\partial^2 p}{\partial v^2} = 0$$

$$-m\ p(0) + \frac{1}{2} \sigma \frac{\partial p}{\partial v}(0) = 0$$

$$-m\ p(1) + \frac{1}{2} \sigma \frac{\partial p}{\partial v}(1) = 0 \qquad (2.5.26)$$

The solution of (2.5.26) is

$$p(v) = \frac{K}{-1 + e^{+K}} e^{+Ky} \qquad (2.5.27)$$

where $K = \frac{2m}{\sigma}$. To show that the two densities are alike, we have to establish that

$$\lim_{\varepsilon \to 0} | K - \hat{\lambda} | = 0$$

From equation (2.5.24),

$$|K - \hat{\lambda}| = \left| \frac{1}{\varepsilon} (p_2 r_1 - p_1 r_2) \cdot \frac{(p_1+r_1)^2 (p_2+r_2)^2}{p_1 r_1 (p_2+r_2)^3 + p_2 r_2 (p_1+r_1)^3} \right| \cdot$$

$$\left| 1 - \frac{(p_1+p_2+r_1+r_2)}{(p_2+p_1)(r_2+r_1)} \cdot \left(\frac{p_1 r_1 (p_2+r_2)}{(p_1+r_1)^2} + \frac{p_2 r_2}{(p_2+r_2)^2} (p_1+r_1) \right) \right| \qquad (2.5.28)$$

Since the transfer line is nearly balanced, we have

$$\left.\begin{aligned} \frac{r_1}{r_1 + p_1} &= \frac{r_2}{r_2 + p_2} + 0(\varepsilon) \\ r_1 p_2 - r_2 p_1 &= 0(\varepsilon) \\ \frac{r_1}{r_1 + p_1} &= \frac{p_2}{r_2 + p_2} + 0(\varepsilon) \end{aligned}\right\} \qquad (2.5.29)$$

Hence, the first two terms in the right hand side of equation 2.5.28 are bounded as $\varepsilon \to 0$. The last term can be expanded using equation (2.5.29) to give

$$\left| 1 - \frac{(p_1+p_2+r_1+r_2)}{(p_1+p_2)(r_1+r_2)} \cdot \left(\frac{p_1 r_1 (p_2+r_2)}{(p_1+r_1)^2} + \frac{p_2 r_2}{(p_2+r_2)^2} (p_1+r_1) \right) \right| = 0(\varepsilon)$$

which establishes

$$\lim_{\varepsilon \to 0} |K - \hat{\lambda}| = 0.$$

Hence, the ergodic distribution of the diffusion approximation is consistent with the ergodic distribution of the original model. Theorem 2.5.7 indicates that continuous functionals of the process, such as expected exit times, will converge in the same fashion.

2.5.8. Three Machine Transfer Lines

The three machine transfer line is the first nontrivial example of coupling between the storage buffers. Figure 2.5.3 describes a typical three machine transfer line with two storages present. We will assume that blockage and starvation affect machine failure rates, as indicated in section 6.

Assuming that $N_1 = N_2 = 1/\varepsilon$, and that the transfer line is nearly balanced, the normalized equations of flow in the time scale $\tau = \varepsilon^2 t$ are

$$\frac{dy_1}{d\tau} = \frac{1}{\varepsilon} (\mu_1 \alpha_1 - \mu_2 \alpha_2)$$

$$\frac{dy_2}{d\tau} = \frac{1}{\varepsilon} (\mu_2 \alpha_2 - \mu_3 \alpha_3)$$

when $(y_1, y_2) \in (0,1) \times (0,1)$

On the boundary, the adjustment rules for conservation of flow must apply. In terms of the compensating processes, this means

$$\frac{dy_1}{d\tau} = \frac{1}{\varepsilon} (\mu_1 \alpha_1 - \mu_2 \alpha_2) + \frac{d}{d\tau} U_1^0(\tau;\varepsilon) - \frac{d}{d\tau} U_1^1(\tau;\varepsilon) + \frac{d}{d\tau} U_2^1(\tau;\varepsilon)$$

$$\frac{dy_2}{d\tau} = \frac{1}{\varepsilon} (\mu_2 \alpha_2 - \mu_3 \alpha_3) + \frac{d}{d\tau} U_2^0(\tau;\varepsilon) - \frac{d}{d\tau} U_2^1(\tau;\varepsilon) - \frac{d}{d\tau} U_1^0(\tau;\varepsilon)$$

where we have explicitly depicted the dependence of the compensators on . The results of Theorem 2.5.7 and 2.5.8 let us represent the approximating diffusion process as the

$$\begin{aligned} \nu_1(\tau) &= \omega_1(\tau) + U_1^0(\tau) - U_1^1(\tau) + U_2^1(\tau) \\ \nu_2(\tau) &= \omega_2(\tau) + U_2^0(\tau) - U_2^1(\tau) - U_1^0(\tau) \end{aligned} \qquad (2.3.50)$$

where (ω_1, ω_2) is a diffusion process with parameters $(\underline{m}, \underline{\Sigma})$, given by

$$\underline{m} = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} \frac{\mu_1 r_1}{r_1 + p_1} - \frac{\mu_2 r_2}{r_2 + p_2} \\ \frac{\mu_2 r_2}{r_2 + p_2} - \frac{\mu_3 r_3}{r_3 + p_3} \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \qquad (2.5.31)$$

$$\Sigma_{11} = 2 \frac{\mu_1^2 p_1 r_1}{(p_1+r_1)^3} + \frac{2\mu_2^2 p_2 r_2}{(p_2+r_2)^3}$$

$$\Sigma_{12} = \frac{-2\mu_2^2 p_2 r_2}{(p_2+r_2)^3}$$

$$\Sigma_{21} = \frac{-2\mu_2^2 p_2 r_2}{(p_2+r_2)^3}$$

$$\Sigma_{22} = \frac{2\mu_2^2 p_2 r_2}{(p_2+r_2)^3} + \frac{2\mu_3^2 p_3 r_3}{p_3 + r_3}$$

Equation (2.5.30) corresponds to a diffusion

process on the unit square with oblique reflection at the boundaries; the directions of reflection are illustrated in figure 2.5.4.

The processes U_i^0, U_i^1 are continuous, increasing processes, which are bounded almost surely at each time τ. This implies that the processes $v_1(\tau)$, $v_2(\tau)$ are semimartingales, and thus we have a generalization of Ito's formula (Kunita-Watanabe [KUN 67], Harrison-Reiman [HAR 79]). Let f be a twice continuously differentiable function on $\bar{D}$. Denote by $\bar{L}$ the infinitesimal generator of (ω_1,ω_2), that is

$$\bar{L} f(\omega_1,\omega_2) = \sum_{i=1}^{2} m_i \frac{\partial}{\partial \omega_i} f(\omega_1,\omega_2) + \frac{1}{2} \sum_{\substack{i=1 \\ j=1}}^{2} \Sigma_{ij} \frac{\partial^2}{\partial \omega_i \partial \omega_j}$$

Then we have

$$\begin{aligned} & f(v_1(\tau),v_2(\tau)) - f(v_1(0),v_2(0)) = \\ & \int_0^\tau \bar{L} f(v_1(s),v_2(s))ds \\ & + \int_0^\tau \frac{\partial}{\partial v_1} f(v_1(s),v_2(s))d\omega_1(s) \\ & + \int_0^\tau \frac{\partial}{\partial v_2} f(v_1(s),v_2(s))d\omega_2(s) \\ & + \int_0^\tau (\frac{\partial f}{\partial v_1} - \frac{\partial f}{\partial v_2})(v_1(s),v_2(s))dU_1^0(s) \\ & + \int_0^\tau -(\frac{\partial f}{\partial v_1} - \frac{\partial f}{\partial v_2})(v_1(s),v_2(s))dU_2^1(s) \\ & + \int_0^\tau -\frac{\partial f}{\partial v_1}(v_1(s),v_2(s))dU_1^1(s) \\ & + \int_0^\tau \frac{\partial f}{\partial v_2}(v_1(s),v_2(s))dU_2^0(s) \end{aligned} \qquad (2.5.32)$$

where the last four terms represent the contributions of the four parts of ∂D. Notice that, if f were such that $f \in \mathcal{D}^0$, the set of all twice continuously differentiable functions such that

1. $\frac{\partial f}{\partial v_1} - \frac{\partial f}{\partial v_2} = 0$ on $v_1 = 0$

2. $\frac{\partial f}{\partial v_1} - \frac{\partial f}{\partial v_2} = 0$ on $v_2 = 1$

3. $\frac{\partial f}{\partial v_1} = 0$ on $v_1 = 1$

4. $\frac{\partial f}{\partial v_2} = 0$ on $v_2 = 0$,

the equation (2.5.32) implies

$$E\{f(v_1(\tau),v_2(\tau))|v_1(0),v_2(0)\} - f(v_1(0),v_2(0)) = E\{\int_0^\tau \bar{L} f(v_1(s),v_2(s))ds|v_1(0),v_2(0)\} \qquad (2.5.33)$$

The infinitesimal generator of the (v_1,v_2) process is thus seen to be $\bar{L}$, with its domain $\mathcal{D}$ including the class of functions $\mathcal{D}^0$.

The process $(v_1(\tau),v_2(\tau))$ is a diffusion process in a compact domain, with a positive probability of visiting all states except possible the corners of $\bar{D}$. Hence, there exists a unique ergodic probability density function $p^*(v_1,v_2)$ such that

$$L^*p^* = \frac{1}{2} \sum_{\substack{i=1 \\ j=1}}^{2} \Sigma_{ij} \frac{\partial^2}{\partial x_i \partial x_j} p^* - \sum_{i=1}^{2} m_i \frac{\partial p^*}{\partial x_i} = 0 \qquad (2.5.34)$$

$$\begin{aligned} & \frac{1}{2}\Sigma_{11}\frac{\partial p^*}{\partial x} + \frac{1}{2}\Sigma_{11}\frac{\partial p^*}{\partial y} + \Sigma_{12}\frac{\partial p^*}{\partial y} - m_1 p^* = 0 \text{ on } S_1 \\ & \frac{1}{2}\Sigma_{22}\frac{\partial p^*}{\partial y} + \Sigma_{12}\frac{\partial p^*}{\partial x} - m_2 p^* = 0 \text{ on } S_2 \\ & \frac{1}{2}\Sigma_{11}\frac{\partial p^*}{\partial x} + \Sigma_{12}\frac{\partial p^*}{\partial y} - m_1 p^* = 0 \text{ on } S_3 \\ & \frac{1}{2}\Sigma_{22}\frac{\partial p^*}{\partial y} + \frac{1}{2}\Sigma_{22}\frac{\partial p^*}{\partial x} + \Sigma_{12}\frac{\partial p^*}{\partial x} - m_2 p^* = 0 \text{ on } S_4 \end{aligned} \qquad (2.5.35)$$

$$p^*(0,0) = 0 \text{ unless } \Sigma_{12} + \frac{\Sigma_{11}}{2} = 0$$

$$p^*(1,1) = 0 \text{ unless } \frac{\Sigma_{22}}{2} + \Sigma_{12} = 0$$

$$p^*(0,1) = 0 \text{ unless } \Sigma_{11} + 2\Sigma_{12} + \Sigma_{22} = 0$$

$$p^*(1,0) = 0 \text{ unless } \Sigma_{12} = 0$$

From equation (2.5.31), we can verify that $p^*(1,0) = p^*(0,1) = 0$. However, when machines 1 and 2, or machines 2 and 3 have identical failure and repair rates, the values of $p^*(0,0)$ and $p^*(1,1)$ can be nonzero. In these cases, the intensity of the coupling term Σ_{12} matches and cancels the oblique flow along the boundary, resulting in decoupled reflecting conditions. This can be seen from equation (2.5.35), which, when $\Sigma_{11} = \Sigma_{22} = -2\Sigma_{12}$, reduce to

$$\begin{aligned} & \frac{1}{2}\Sigma_{11}\frac{\partial p^*}{\partial x} - m_1 p^* = 0 \text{ on } S_1 \\ & \frac{1}{2}\Sigma_{22}\frac{\partial p^*}{\partial y} - m_2 p^* = 0 \text{ on } S_4 \end{aligned}$$

Exact solutions of equations (2.5.34) with boundary conditions (2.5.35) is a difficult problem, which can seldom be solved in closed form. However, the markov process $(v_1(\tau), v_2(\tau))$ can be approximated in the weak sense by a Markov Chain, as in Kushner [KUS 76], and the ergodic distribution of this chain can be computed as an approximate solution to these equations.

Assume that the stationary probability distribution $p^*(x,y)$ has been determined. Let E^* denote the measure on the path space induced by p^*. Let f be any bounded, twice continuously differentiable function on $\bar{D}$. Then, equation (8.9) implies, from Fubini's theorem,

$$\int_D \bar{L}f\, p^*\, dv_1 dv_2 + \int_0^\tau E^*\{(\frac{\partial f}{\partial v_1} - \frac{\partial f}{\partial v_2}\}dU_1^o(s))$$

$$+ \int_0^\tau E^*\,(\frac{\partial f}{\partial v_1} - \frac{\partial f}{\partial v_2})dU_2^1(s)\} + \int_0^\tau E^*\,\{(-\frac{\partial f}{\partial v_1})\, dU_1^1(s)\}$$

$$+ \int_0^\tau E^*\,\frac{\partial f}{\partial v_2}\, dU_2^o(s)\} = 0 \qquad (2.5.36)$$

Define measures on S_1, S_2, S_3, S_4 as

$$\nu_1(B) = \frac{1}{\tau}\int_0^\tau E^*\{I(v_2(s)\in B)\,|\,dU_1^o(s)\} \text{ for } B \quad S_1$$

$$\nu_2(B) = \frac{1}{\tau}\int_0^\tau E^*\{I(v_1(s)\in B)\,|\,dU_2^o(s)\},\ B \quad S_2$$

$$\nu_3(B) = \frac{1}{\tau}\int_0^\tau E^*\{I(v_2(s)\in B)dU_1^1(s),\ B \quad S_3$$

$$\nu_4(B) = \frac{1}{\tau}\int_0^\tau E^*\{I(v_1(s)\in B)dU_2^1(s)\},\ B \quad S_4$$

The measures ν_i are the occupation time, or local time, measures on the boundary, defined in Donsker-Varadhan [DON 75]. It is easy to show that, for any τ,

$$0 < E^*\{U_j^i(\tau)\} < \infty .$$

Hence, we can use Fubini's theorem to reduce equation (2.5.36) to

$$\int_D \bar{L}\, f \cdot p^*\, dv_1 dv_2 + \int_{S_1} (\frac{\partial f}{\partial v_1} - \frac{\partial f}{\partial v_2})\ \nu_1(dy)$$

$$+ \int_{S_2} \frac{\partial f}{\partial v_2}\ \nu_2(dx) + \int_{S_3} -\frac{\partial f}{\partial v_1}\ \nu_3(dy)$$

$$+ \int_{S_4} (\frac{\partial f}{\partial v_1} - \frac{\partial f}{\partial v_2})\ \nu_4(dx) = 0 \qquad (2.5.37)$$

Equation (2.5.37) can be evaluated for selected functions f, to obtain the properties of the process when it reaches the stationary limit. For instance, recall that the throughput rate in section 5 was given by

$$T(\tau) = \{\frac{1}{\tau}\int_0^\tau \mu_3\alpha_3\, ds - \frac{U_2^o(\tau)}{\tau}\}$$

The expected stationary throughput rate is just

$$T_{SS} = E^*(T(\tau))$$

$$= E^*\left(\frac{\mu_3}{\tau}\int_0^\tau \alpha_3\, ds\right) - E(\{U_2^o(\tau)\}$$

$$= \frac{\mu_3 r_3}{r_3+p_3} - E^*\{\frac{U_2^o(\tau)}{\tau}\}$$

$$= \frac{\mu_3 r_3}{r_3+p_3} - \nu_2(S_2)$$

Let $f = \alpha e^{-v_2/\alpha}$. f is bounded, and smooth, hence (2.5.37) implies

$$\int_D (\frac{1}{2}\ \frac{\Sigma_{22}}{\alpha} - m_2)\ e^{-v_2/\alpha} p^*(v_1,v_2)dv_1$$

$$+ \int_{S_1} e^{-x_2/\alpha}\ \nu_1(dv_2)$$

$$+ \int_{S_2} -\nu_2(dv_1)$$

$$+ \int_{S_2} e^{-1/2}\ \nu_4(dv_1) = 0 \qquad (2.5.38)$$

Letting α approach zero in (8.35) yields

$$\nu_2(S_2) = \int_{S_2} \frac{1}{2}\ \Sigma_{22}\, p^*(v_1,0)\ dv_1 \qquad (2.5.39)$$

2.5.9. Conclusion

In this paper, we have presented a methodology for approximating the flow of material through a transfer line of unreliable machines with finite storage buffers. Under the assumption of large but finite storages, the flow of material is approximated by a diffusion process with reflecting boundary conditions, independent of the process which describes the failures and repairs of the machines. This approximation reduces the number of states which must be considered by a factor of 2^k, where k is the number of machines in the transfer line.

The structure of the approximation was exploited in the case of 2 and 3 machine transfer lines to obtain equations for the stationary distribution of the approximate diffusion process. In the two machine case, these equations were solved explicitly, and found to be consistent with the results of Gershwin and Schick [GER 80a]. The equations for the stationary distribution of the three machine trnasfer line were too complicated to solve in closed form, although numerical algorithms for their solution are currently under study.

The methodology derived in this paper can be applied to transfer lines of arbitrary length without ignoring the coupling effects of starvation and blockage. As such it represents a significant generalization of the previous works mentioned in the introduction. Work is currently in progress to

generalize these results to arbitrary network topologies with nonconstant flow rates and storage dependent failure rates. For these problems, the techniques used in this paper will not apply, because of the dependence of the failure-repair processes on the levels of storage. Key theoretical questions concerning the existence and uniqueness of the limit process must be answered. These problems are currently under investigation, and will be reported in later publications.

Appendix 2.A Linear Operators on Finite-Dimensional Spaces

This appendix contains background material on linear operators on finite-dimensional spaces. Most of the results are available in [KAT 66]. The purpose of this Appendix is to introduce the notation and preliminary results for section 2.2.

Let V and W be two vector spaces and $T: V \to W$ a linear operator. The image of V under T is called the range of T and is denoted by $R(T)$. The dimension of $R(T)$ is called the rank of T; we denote it by rank T. The inverse image of the zero element of W is called the null space of T and is denoted by $N(T)$. The dimension of $N(T)$ is called the nullity of T which we denote by nul T. A basic result in linear algebra is:

$$\text{rank } T + \text{nul } T = \dim V$$

If T maps V on W one to one, the inverse operator $T^{-1}: W \to V$ is well defined and T is said to be non-singlar, otherwise is said to be singular.

Let X and Y be two subspaces of V such that each $u \in V$ can be uniquely decomposed in the form $u = u' + u''$ with $u' \in X$ and $u'' \in Y$, i.e., $V = X \oplus Y$. The linear operator $P: V \to V$, $Pu = u'$ is called the projection on X along Y and we have $R(P) = X$, $N(P) = Y$. P is idempotent, i.e., $P^2 = P$ and conversely any idempotent operator is a projection. More generally

$$V = X_1 \oplus \ldots \oplus X_s \tag{2.A.1}$$

$u = u_1 + \cdot + u_s$, $u_i \in X_i$ and the operator P_j defined by $P_j u = u_j$ is the projection on $X_1 \oplus \ldots \oplus X_{j-1} \oplus X_{j+1} \oplus \ldots \oplus X_s$. Furthermore, we have

$$\sum_{j=1}^{s} P_j = I \tag{2.A.2}$$

$$P_j P_k + \delta_{kj} P_j \tag{2.A.3}$$

Conversely, any set of operators $\{P_j\}$ satisfying (A.2) and (A.3) is a family of projections that determine the direct sum decomposition (2.A.1) with $X_i = R(P_j)$. A basis $\{v_j\}$ of V is said to be adapted to the decomposition (2.A.1) if the first $n_1 = \dim X_1$ elements of $\{v_i\}$ belong to X_1, the following $n_2 = \dim X_2$ ones belong to X_2 and so on.

A subspace X of V is said to be invariant under a linear operator $T: V \to V$ if $TX \subset X$. In this case T induces a linear operator $T_x: X \to X$ defined by $T_x u = Tu$ for $u \in X$ which is called the part of T in X. T is said to be decomposed by a set of subspaces $\{X_i\}$ if 2.A.1 is satisfied and all the X_i are invariant under T.

A linear operator $T: V \to V$ is called niloptent if $T^r = 0$ for some positive integer r. A nilpotent operator is necessarily singular.

The set of all linear operators on V to W is a normed vector space with norm induced by the vector norms in V and W as follows:

$$||T|| = \sup_{\substack{u \in V \\ u \neq 0}} \frac{||Tu||}{||u||} = \sup_{\substack{u \in V \\ ||u||=1}} ||Tu||$$

Operator-valued functions T(t) defined for a real or complex variable t can be defined and treated as vector-valued or scalar functions. The following lemma dealing with projection-valued functions will be useful in latter chapters.

Lemma 2.A.1 (KAT 66] p. 34)

Let P(t) be a projection depending continuously on a parameter t varying in a (connected) region of real or complex numbers. Then the range $R(P(t))$ for different t are isomorphic to one another. In particular, dim $R(P(t))$ is constant.

Let T be a linear operator on V to itself. A complex number λ is called an eigenvalue of T is there exists a non-zero vector u such that

$$Tu = \lambda u$$

u is called an eigenvector of T with eigenvalue λ. The subspace of eigenvectors of T with eigenvalue λ is called the geometric eigenspace for λ and its dimension the geometric multiplicity of λ. The set of all eignevalues of T is called the spectrum of T; we denote it by $\sigma(T)$.

The operator-valued function

$$R(\xi,T) = (T - \xi I)^{-1} \tag{2.A.4}$$

is well defined for any complex number $\xi \in \rho(T) \triangleq C - \sigma(T)$ and it is called the resolvent of T. The set $\rho(T)$ is referred to as the resolvent set of T. $R(\xi,T)$ satisfied the so called resolvent equation:

$$R(\xi_1,T) - R(\xi_2,T) = (\xi_1 - \xi_2)\, R(\xi_1,T)\, R(\xi_2,T) \tag{2.A.5}$$

which, in particular, implies that $R(\xi_1,T)$ and $R(\xi_2,T)$ conmute. The resolvent is an analytic function with isolated singularities at precisely the eigenvalues λ_k, $k = 0,1,\ldots,s$, of T.

The Laurent series of $R(\xi,T)$ at λ_k has the form:

$$R(\xi,T) = -(\xi - \lambda_k)^{-1} P_k - \sum_{n-1}^{m_k-1} (\xi - \lambda_k)^{-n-1} D_k^n + \sum_{n=0}^{\infty} (\xi - \lambda_k)^n S_k^{n+1} \tag{2.A.6}$$

where

$$P_k = -\frac{1}{2\pi i} \int_{\Gamma_k} R(\xi,T)\, d\xi \tag{2.A.7}$$

(with Γ a positively oriented contour enclosed λ_k but no other eigenvalue of T) is a projection called the eigenprojection for the eigenvalue λ_k of T; $m_k = \dim R(P_k)$ is the algebraic multiplicity of λ_k;

$$D_k = -\frac{1}{2\pi i}\int_{\Gamma_k}(\xi-\lambda_k)k(\xi,T)d\xi \qquad (2.A.8)$$

is the eigennilpotent ($D_k^{m_k} = 0$) for the eigenvalue λ_k of T; and

$$S_k = \frac{1}{2\pi i}\int_{\Gamma_k}(\xi-\lambda_k)^{-1}\,R(\xi,T)d\xi \qquad (2.A.9)$$

It is not difficult to see that the following relations hold:

$$P_k S_k = S_k P_k = 0$$

$$P_k D_k = D_k P_k = D_k$$

$$P_k T = TP_k$$

$$(T-\lambda_k I)\,S_k = I - P_k$$

$$(T-\lambda_k I)\,P_k = D_k$$

$$P_k P_\ell = \delta_{k\ell} P_k$$

$$\sum_{k=1}^{s} P_k = I$$

$$V = M_1 \oplus \ldots \oplus M_s$$

with $M_k = R(P_k)$. M_k is called the algebraic eigenspace for the eigenvalue λ_k of T. It follows that

$$TP_k = P_k T = P_k T P_k = \lambda_k P_k + D_k \qquad (2.A.10)$$

which gives the canonical form or spectral representation of T:

$$T = \sum_{k=0}^{s}(\lambda_k P_k + D_k) \qquad (2.A.11)$$

An eigenvalue λ_k is said to be semisimple if the associated eigenniloptent D_k is zero and simple if in additoin $m_k = 1$. T is said to be diagonabiable if all its eigenvalues are semisimple.

Definition 2.A.2

A linear operator T_o on V to itself is said to have semisimple null structure (SSNS) if zero is a semisimple eigenvalue of T_o.

Lemma 2.A.3

The following are equivalent statements:

i) To has SSNS
ii) $V = R(T_o) \oplus N(T_o)$
iii) $R(T_o) = R(T_o^2)$
iv) rank T_o = rank T_o^2
v) $N(T_o) = N(T_o^2)$

It follows that if T_o has SSNS then P_o, the eigenprojection for the zero eigenvalue of T_o, is also the projection on $N(T_o)$ along $R(T_o)$. Let $Q_o = I - P_o$.

Therem 2.A.4.

If T_o has SSNS then $T_o + P_o$ is non-singular. Define the operator $T_o^\#$ by $T_o^\# = (T_o + P_o)^{-1} - P_o$. The following lemma gives several properties of this operator.

Lemma 2.A.5.

i) $P_o T_o^\# = T_o^\# P_o = 0$
ii) $Q_o T_o^\# = T_o^\# Q_o = T_o^\#$
iii) $T_o^\# T_o = T_o T_o^\# = Q_o$
iv) $\| T_o^\# \| = \| \hat{T}_o^{-1} \|$

It follows from lemma 2.7 that $T_o T_o^\# T_o = T_o$, $T_o^\# T_o T_o^\# = T_o^\#$ and $T_o T_o^\# = T_o^\# T_o$. $T_o^\#$ is thus the generalized Group inverse of T_o (see [Cam 79a]; we will refer to it simply as the generalized inverse of T_o. The following lemma shows that if T_o has SSNS then P_o and $T_o^\#$ fully determine the Laurent expansion of the resolvent $R(\lambda,T_o)$ at zero.

Lemma 2.A.6.

If T_o has SSNS, then for $0 < |\lambda| < \| T_o^\# \|^{-1}$ we have

$$R(\lambda,T_o) = -\frac{P_o}{\lambda} + \sum_{k=0}^{\infty}\lambda^k\,(T_o\#)^{k+1}$$

$$= -\frac{P_o}{\lambda} + T_o\#\,(I-\lambda T_o\#)^{-1}$$

Assume that an operator-valued function $T(\varepsilon)$ is given which is continuous in a neighborhood of $\varepsilon = 0$, say for $\varepsilon \in [0,\varepsilon^1]$. We will distinguish two cases:

a) $T(\varepsilon)$ has an absolutely convergent power series expansion for $\varepsilon \in (0,\varepsilon^1]$, i.e.,

$$T(\varepsilon) = T + \sum_{n=1}^{\infty}\varepsilon^n T^{(n)} \quad \text{for } \varepsilon \in [0,\varepsilon^1] \qquad (2.A.11)$$

$$\| T(\varepsilon) \| \le \| T \| + \sum_{n=1}^{\infty}\varepsilon^n \| T^{(n)} \| \triangleq \gamma(\varepsilon) < \infty$$

$$\text{for } \varepsilon \in [0,\varepsilon^1] \qquad (2.A.12)$$

b) $T(\varepsilon)$ has an asymptotic expansion in powers of ε for $\varepsilon \downarrow 0$, i.e.,

$$\lim_{\varepsilon\downarrow 0} \| T(\varepsilon) - T - \sum_{n=1}^{N}\varepsilon^n T^{(n)} \| \cdot \varepsilon^{-N} = 0 \quad \forall N \ge 0$$

$$(2.A.13)$$

The eigenvlaues of $T(\varepsilon)$ satisfy the characteristic equation:

$$\det\,(T(\varepsilon) - \xi I) = 0 \qquad (2.A.14)$$

If $T(\varepsilon)$ is analytic this is an algebraic equation in ξ of degree $n = \dim V$, with coefficients which are analytic in ε. It follows that the roots of (2.A.14) are branches of analytic functions of ε with only algebraic singularities and therefore the number of (distinct) eigenvalues of $T(\varepsilon)$ is a constant s

independent of ε, except at some special values of ε. There are only a finite number of such exceptional points in a compact interval $\varepsilon\in[0,\varepsilon']$. We will assume that ε' is small enough so that $[0,\varepsilon']$ contains only one exceptional point which, without loss of generality, we take as $\varepsilon=0$.

In a neighborhood of the exceptional point, the eigenvalues of $T(\varepsilon)$ can be expressed by s analytic functions $\lambda_1(\varepsilon),\ldots,\lambda_s(\varepsilon)$ with $\lambda_n(\varepsilon) \neq \lambda_k(\varepsilon)$ for $n \neq k$ which can be grouped in the manner:

$$\{\lambda_1(\varepsilon),\ldots,\lambda_p(\varepsilon)\},\ \{\lambda_{p+1}(\varepsilon),\ldots,\ _{p+q}(\varepsilon)\},\ldots$$

in such a way that we have the Puiseux series:

$$\lambda_n(\varepsilon) = \lambda + \alpha_1\,\omega^n\,\varepsilon^{1/p} + \alpha_2\,\omega^{2n}\,\varepsilon^{2/p} \qquad (2.A.15)$$

$$n = 0,1,\ldots,p-1$$

where λ is an eigenvalue of the unperturbed operator $T(o)$ and $\omega = \exp\{2\pi i/p\}$. Each group is called a cycle and the number of elements its period. It should be noticed that the $\lambda_n(\varepsilon)$ are continuous at $\varepsilon=0$; $\lambda = \lambda_n(0)$ will be called the center of the cycle under consideration.

In general there are several cycles with the same center λ. All eigenvalues belonging to cycles with center λ are said to depart from the unperturbed eigenvalue λ by splitting at $\varepsilon=0$. The set of these eigenvalues will be called the λ-group since they cluster around λ for ε small. Equation (2.A.15) shows that $\lambda_n(\varepsilon) - \lambda = 0_s(\varepsilon^{1/p})$.

In the asymptotic case (b), the number of eigenvalues may change with ε quite irregularly; the splitting and coalescence of eigenvalues taking place in a very complicated nammer. It may even happen that in no interval of the form $(0,\varepsilon']$ is the number of eigenvalues constant. In dealing with non-analytic perturbations we will restrict ourselves to the case of constant numbers of eigenvalues for $\varepsilon \in (0,\varepsilon']$.

The resolvent of $T(\varepsilon)$ is

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

Lemma 2.A.7
Let

$$T(\varepsilon) = T + \sum_{n=1}^{\infty} \varepsilon^n\ T^{(n)}$$

If $\xi \in \rho(T)$ then for, ε small enough, $\xi \in \rho(T(\varepsilon))$ and

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

where

$$R^{(n)}(\xi) = \sum_{\substack{\nu_1+\ldots+\nu_p=n \\ \nu_i \geq 1}} (-1)^p\ R(\xi,T)T^{(\nu_1)}$$

$$R(\xi,T)T^{(1/2)}\ldots T^{(\xi_p)}R(\xi,T)$$

the sum being taken for all combinations of positive integers p and $\nu_1,\ldots,\nu_p$ such that $1 \leq p \leq n$, $\nu_1+\ldots+\nu_p = n$. The series is uniformly convergent on compact subsets of $\rho(T)$ in case (a) and it is a uniform asymptotic series for $R(\xi,T(\varepsilon))$ in compact subsets of $\rho(T(\varepsilon))$ in case (b).

Let λ be an eigenvalue of $T = T(0)$, with algebraic multiplicity m. Let Γ be a closed positive contour in $\rho(T)$ enclosing λ but no other eigenvalues of T. It follows from lemma 4.1 that for ε small enough $R(\xi,T(\varepsilon))$ exists for $\xi \in \Gamma$ and therefore there are no eigenvalues of $T(\varepsilon)$ on Γ.

The operator

$$P(\varepsilon) = -\frac{1}{2\pi i}\int_\Gamma R(\xi,T(\varepsilon)d\xi \qquad (2.A16)$$

is a projection that commutes with $T(\varepsilon)$ and is equal to the sum of the eigenprojections for all the eigenvalues of $T(\varepsilon)$ lying inside Γ. From Lemma A.2.7, integrating term by term yields

$$P(\varepsilon) = P + \sum_{n=1}^{\infty} \varepsilon^n\ P^{(n)} \qquad \varepsilon\in[0,\varepsilon_o] \qquad (2.4.17)$$

where

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

is the eigenprojection for the eigenvalue λ of T, and

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

$P(\varepsilon)$ is continuous in a neighborhood of zero and it follows from lemma 2.A.1 that the range of $P(\varepsilon)$ is isonorphic to the range of P; in particular,

$$\dim R(P(\varepsilon)) = \dim R(P) = m\ .$$

that all eigenvalues of $T(\varepsilon)$ lying inside Γ form exactly the λ-group. For this reason $P(\varepsilon)$ will be called the total projection, and $R(P(\varepsilon))$ the total eigenspace for the λ-group.

Lemma 2.A.8
Let

$$T(\varepsilon) = T + \sum_{n=1}^{\infty} \varepsilon^n T^{(n)}$$

Let λ be an eigenvalue of T with multiplicity m and let $P(\varepsilon)$ denote the total projection for the λ-group. Then

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

$$= \frac{D}{\varepsilon} + \sum_{n=D}^{\infty} \varepsilon^n\ \tilde{T}^{(n)} \qquad \varepsilon\in(0,\varepsilon_o]$$

where Γ is a closed positive contour enclosing

λ but no other eigenvalues of T, D is the eigennilpotent for λ and $\tilde{T}^{(n)}$ is given by:

$$T^{(n)} = - \sum_{p=1} (-1)^p \sum_{\substack{\nu_1+\ldots+\nu_p=n \\ k_1+\ldots+k_{p+1}=p-1 \\ \nu_i \geq 1,\ k_j \geq -m+1}} S^{(k_1)} T^{(\nu_1)}$$

$$S^{(k_2)} \ldots S^{(k_p)} T^{(\nu_p)} S^{(k_{p+1})}$$

with $S^{(o)} = - P(0) = - P$, $S^{(-k)} = - D^k$, $k \geq 1$ and $S^{(k)} = S^k$, $k \geq 1$ for

$$S = \frac{1}{2\pi i} \int_\Gamma (\xi-\lambda)^{-1} R(\xi,T)\, d\xi$$

$$\mathbb{R}^n = \mathcal{R}(Q_m(\varepsilon)) \oplus \ldots\ldots \oplus \mathcal{R}(Q_1(\varepsilon)) \oplus \mathcal{R}(Q_0(\varepsilon)) \oplus \mathcal{N}(A_0(\varepsilon))$$

$$\mathcal{R}(A_0(\varepsilon)) \quad \mathcal{N}(A_0(\varepsilon))$$

$$\mathcal{R}(A_1(\varepsilon)) \quad \mathcal{N}(A_1(\varepsilon))$$

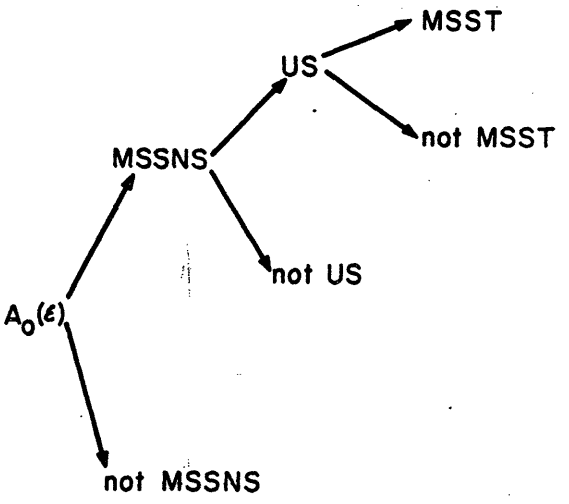
$$\mathcal{R}(A_2(\varepsilon)) \quad \mathcal{N}(A_2(\varepsilon))$$

$$\mathcal{R}(A_m(\varepsilon)) \quad \mathcal{N}(A_m(\varepsilon))$$

Figure 2.2.1. Structure of the matrices $A_k(\varepsilon)$.

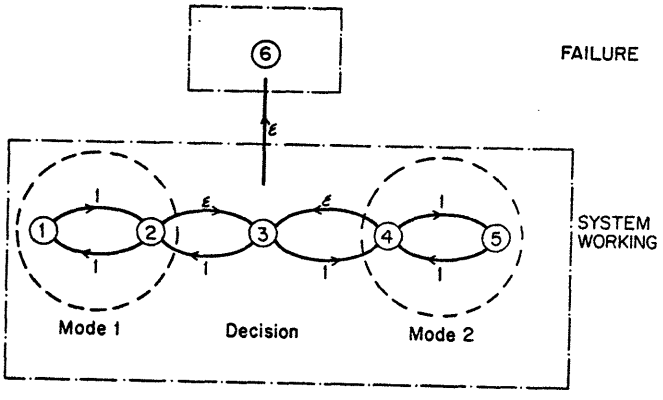
$$\begin{array}{lll}
A_0(\varepsilon) & A_{00}\ A_{01}\ A_{02} \cdots\cdots & A_{0m} \cdots\cdots \\
A_1(\varepsilon) & A_{10}\ A_{11} \cdots\cdots\cdots & A_{0,m-1} \cdots\cdots \\
A_2(\varepsilon) & A_{20} \cdots\cdots\cdots & A_{0,m-2} \cdots\cdots \\
\vdots & \vdots & \\
A_m(\varepsilon) & A_{m0}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

Figure 2.2.2. Table of matrices A_{ij}.



MSSNS = Multiple Semisimple Nullstructure
US = Uniform Stability
MSST = Multiple Semistability

Figure 2.2.3 Different cases studied in section 2.



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

Figure 2.3.1 A FSMP with rare events.

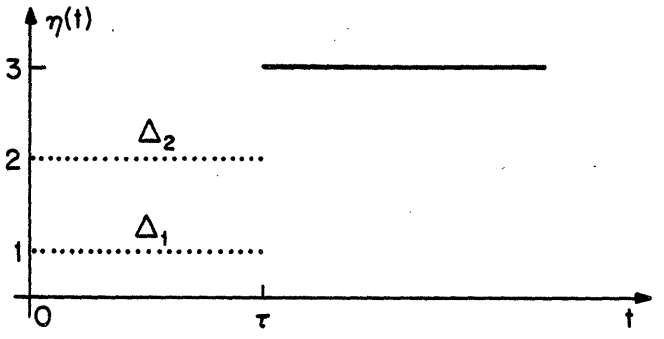


Figure 2.3.2. Sample functions of the process $\eta(t)$ of example 2.3.3.

3.1. Stochastic Bifurcation and Singular Perturbation

3.1. Introduction

Bifurcation is the study of branching in the equilibrium behavior of dynamical systems in response to small changes in the parameters of the system. In our past work on stability of large scale interconnected power systems at the University of California, Berkeley we have shown the importance of the study of static bifurcations of the load-flow equations and dynamic bifurcations of the swing equations (see [SAS-80], [SAS-81], [ARA-81a], [ARA-81b]).

Nevertheless the methods of deterministic bifurcation are extremely sensitive to the addition of small amounts of noise, as has been noticed before by some researchers. Thus, in systems whose macroscopic description arises from an aggregation of microscopically fluctuating dynamics; for example, the power demanded at a load node (PV or PQ) of a power systems, the predictions of deterministic bifurcation are incorrect. In Section 3.2, our aim is to remedy this deficiency and set down a theory of bifurcation in the presence of noise - referred to as stochastic bifurcation. Section 3.2 is organized as follows:

Section 3.2.1 is the introduction to this section and states a problem from thermodynamics that led to this development-phase transitions in Van der Waals gases by the Maxwell equal area rule. Section 3.2.2 discusses the Maxwell's equal area rule. In Section 3.2.3 we compare and highlight the differences in the predictions of deterministic and stochastic bifurcation for a class of gradient systems. Finally, we apply this theory to the study of noisy constrained systems: The dynamics of several engineering systems are not described explicitly by differential equations; but rather implicitly with a combination of algebraic and differential equations. An example is the dynamics of a power system: the swing equations coupled with algebraic equations at each load node. These equations (deterministically) admit of jump at points of bifurcation of the algebraic equation. The nature of the jump can be altered quite dramatically by the presence of noise - as we discuss in Section 3.2.4.

Section 3.3 is titled Singular Perturbation, State Aggregation and Non-linear Filtering. In situations where the structure of a dynamical system varies with time it is often the case that the (random) structural changes occur on a time scale that is much slower than the dynamics in any given mode of operation. For example in the study of power system dyanamics, the swing equations are sometimes thought of as occuring on a fast time scale compared to the relatively slow time scale of random faults or breakdowns. The purpose of this section is to study the asymptotic behavior of the resulting hybrid system in the limit that the two time scales mentioned above are singularly perturbed to the slow time scale. Thus, returning to the example of the power system dynamics, we are interested in the asymptotic behavior of the transitions between different faulted states in the limit that the swing dynamics become infinitely fast compared to the incidence of (say) lightning strikes.

In addition to studying the asymptotic behavior of the hybrid process in the slow time scale we show asymptotically that the problem of estimating the projection onto the slow time-scale of the state of the process given noisy, non-linear observations reduces to a finite-state Wonham filter. The layout of Section 3.3 is as follows: In 3.3.1 we set out the preliminary definitions and discuss variable structure systems and give the flavor of the results to be expected. Sections 3.3.2, 3.3.3 develop the probabilistic and deterministic mathematical machinery needed for the study. Section 3.3.4 contains the main results on the asymptotic behavior of the hybrid process. Asymptotic filtering is discussed in Section 3.3.5.

The link between the two sections is in that they both study the asymptotic qualitative behavior of two-time scale systems and hybrid systems in different contexts that arise in the study of the dynamics of power systems.

3.2. Bifurcation in the Presence of Small Noise

3.2.1 Introduction

Bifurcation is the study of branching in the equilibrium behavior of a dynamical system in response to small changes in the parameters of the system. Deterministic bifurcation, using as foundations singularity theory, has been fairly successful at explaining a wide variety of such phenomena in fluid mechanics, optics, elastic structures, laser physics and ecology [11]. Nevertheless, the methods of deterministic bifurcation are extremely sensitive to the addition of small amounts of noise. Thus, in systems whose macroscopic description arises from an aggregation of microscopically fluctuating dynamics, thermodynamic systems for example, the predictions of deterministic bifurcation may be incorrect. We seek in this paper to set down, a mathematically rigorous theory of stochastic bifurcation - i.e. bifurcation in the presence of small additive noise. We show repeatedly that, in the limit as the intensity of the additive noise tends to zero, the conclusions of this theory are rather different from those of deterministic (or no-noise) bifurcation.

In section 3.2.2 we discuss as motivation an example of a thermodynamical phenomenon, isothermic phase transition in Van der Waals gases in which the predictions of deterministic bifurcation theory are incorrect. We then indicate how the addition of small noise predicts the experimentally observed phase transition first studied by Maxwell [May 1875]. In passing we should mention that since the advent of quantum mechanics physicists have been concerned with the derivation of the Van der Waals equation from first principles (i.e. quantum statistical mechanics). The derivation was first done in 1963 [KAC 63], showing that the Van der Waals equation together with Maxwell's rule are consequences of the quantum theory. The notion that "fluctuations play an especially important role near bifurcation points" has been noticed and elaborated by the Brussels school, Prigogine, Nicolis, and others (see [NIC 77] and references contained therein).

In section 3.2.3 we compare deterministic and stochastic bifurcation. We use Laplace's method of steepest descent to compare the two theories in the limit that the noise intensity goes to zero.

In section 3.2.4 we apply the theory of noisy bifurcation to the study of noisy constrained or implicitly defined dynamical systems resulting from the singular perturbation of fast (or 'parasitic') dynamics on some coordinates of the system. The deterministic solution of these systems admit jump discontinuities, including possibly relaxation oscillations, as studied in [SAS 81]. The addition

of noise however changes the nature of the jump and can in some instances result in the destruction of relaxation oscillations. This is shown explicitly in the case of the degenerate Van der Pol oscillator equation (see for example [2EE 72]).

3.2.2. Phase Transitions for Van der Waals Gases

One of the models used in the study of phase transition from liquid to gas in thermodynamics is the Van der Waals equation [CAL 60] relating the pressure P, the volume V, and the absolute temperature T:

$$rT - (P + \frac{a}{V^2})(V-b) = 0 \qquad (3.2.1)$$

where a,b,r are positive constants depending on the gas. Loosely speaking, the surface satisfying (3.2.1) in (P,V,T) space (see figure 1) is a smooth two-dimensional manifold with two fold lines meeting in a cusp at (P_c,V_c,T_c). Isotherms in the (P,V) plane are drawn in Figure 2. For temperatures less than T_c the upper left hand corner of figure 2 represents the liquid phase. We study here phase transitions from liquid to gas phase at constant temperature:

3.2.2.1 Isothermic Phase Transitions

For temperatures above T_c (supercritical) in figure 2 there is no phase transition (only gas phase). At the critical temperature T_c, the portion of the isotherm to the right (left) of (P_c, V_c) represents the gas (liquid) phase. At subcritical temperatures phase transition is more subtle. If the liquid were allowed to expand 'quasistatically' and isothermally at T_2 by decreasing the pressure, the variation of pressure and volume is described by

$$\dot{P} = f(t) \qquad P(0) = P_o \qquad (3.2.2)$$

$$g(P,V,T_2) \triangleq rT_2-(P+\frac{a}{V^2})(V-b) = 0 \qquad V(0) = V_o > b \qquad (3.2.3)$$

where f(t) is a negative function of time. Note that equations (3.2.2) and (3.2.3) describe the variation in time of P and V so long as

$$\frac{\partial}{\partial V} g(P,V,T_2) \neq 0 \qquad (3.2.4)$$

since we may then obtain V(t) as a function of P(t) from the implicit function theorem applied to (3.2.3). At points (P_1,V_1), (P_2,V_2) shown in figure 2, (3.2.4) is not satisfied and equation (3.2.3) is singular. A regularization we suggest accounts for the fact that 'quasistatic' expansion of the liquid neglects some 'fast' dynamics (see [SAS 81] for details):

$$\dot{P} = f(t) \qquad P(0) = P_o$$

$$\varepsilon\dot{V} = g(P,V,T_2) \qquad V(0) = V_o \qquad (3.2.5)$$

The limit of the trajectories of (2.2) (2.5) as $\varepsilon \downarrow 0$ yields a discontinuous change (jump) in volume from (P_1,V_1) to (P_3,V_3) as shown by the dotted line in figure 2. This is the predicted liquid to gas transition. For the converse phase transition choose f(t) to be positive and the jump transition predicted is from (P_2,V_2) to (P_4,V_4).

Thus the predicted deterministic phase transitions are hysteretic (figure 3). Unfortunately, the observed phase transitions are non-hysteretic as shown by the solid line. This line is drawn according to Maxwell's equal area rule [MAS 1875] - equality of the shaded areas in figure 3.

3.2.2.2. Noisy Isothermic Phase Transitions

To explain the observed phase transition we propose to account for noise in equations (3.2.2), (3.2.5) stemming from the fact that P, V are aggregations of microscopically stochastic behavior. In section 3.2.4.2 we present rigorous results justifying the manipulations outlined here. We replace (3.2.2), (3.2.5) by

$$\dot{P} = f(t) + \sqrt{\mu}\,\xi(t) \qquad P(0) = P_o \qquad (3.2,6)$$

$$\varepsilon\dot{V} = g(P,V,T_2) + \sqrt{\varepsilon\lambda}\,\eta(t) \quad V(0) = V_o \qquad (3.2.7)$$

where $\xi(\cdot)$ and $\eta(\cdot)$ are standard independent white noises and $\lambda > 0$, $\mu > 0$ scale their variance. For each ε, μ, λ the above equations generate a diffusion $t \to (P(t),V(t))$ in the plane. The evolution of the corresponding probability density $p^{\lambda}_{\mu,\varepsilon}(P,V,t)$ is then given by the Fokker-Planck equation

$$\frac{\partial}{\partial t} p^{\lambda}_{\mu,\varepsilon} = \frac{\mu}{2}\frac{\partial^2}{\partial P^2} p^{\lambda}_{\mu,\varepsilon} - \frac{\partial}{\partial P} p^{\lambda}_{\mu,\varepsilon} f(t) + \frac{\lambda}{2\varepsilon}\frac{\partial^2}{\partial V^2} p^{\lambda}_{\mu,\varepsilon} - \frac{1}{\varepsilon}\frac{\partial}{\partial V} p^{\lambda}_{\mu,\varepsilon} g(P,V,T_2) \qquad (3.2.8)$$

To study (3.2.8) in the limit $\varepsilon \downarrow 0$, multiply (2.8) by ε and let $\varepsilon \downarrow 0$. Then the limit p^{λ}_{μ} of $p^{\lambda}_{\mu,\varepsilon}$, provided it exists, satisfied

$$\frac{\lambda}{2}\frac{\partial^2}{\partial V^2} p^{\lambda}_{\mu} - \frac{\partial}{\partial V} p^{\lambda}_{\mu} g(P,V,T_2) = 0 \,. \qquad (3.2.9)$$

Solving (3.2.9) yields

$$p^{\lambda}_{\mu} = k^{\lambda}_{\mu} \exp[-S(P,V,T_2)/\lambda] \qquad (3.2.10)$$

where $k^{\lambda}_{\mu} = k^{\lambda}_{\mu}(P,t)$ and

$$S(P,V,T) = -2\int g(P,V,T)\,dV$$

$$= PV^2+2a \log V-2PbV+ \frac{2ab}{V} -2rVT \,.$$

Substituting back in (2.8) and then integrating over V yields

$$\frac{\partial}{\partial t}(K^{\lambda}_{\mu}c^{\lambda}) = \frac{\mu}{2}\frac{\partial^2}{\partial P^2}(k^{\lambda}_{\mu}c^{\lambda}) - \frac{\partial}{\partial P}(k^{\lambda}_{\mu}c^{\lambda}) f(t) \qquad (3.2.11)$$

where

$$c^{\lambda}(P) = \int_b^{\infty} \exp(-S(P,V,T_2)/\lambda)\,dV < \infty \,.$$

We see therefore that in the limit $\varepsilon \downarrow 0$ the P process converges to a diffusion whose probability density $k^{\lambda}_{\mu}c^{\lambda}$ satisfies (3.2.11). Further in the limit $\varepsilon \downarrow 0$, the conditional density $p^{\lambda}(V|P)$ is given by

$$\frac{1}{c^{\lambda}(P)} \exp(-S(P,V,T_2)/\lambda) \,.$$

This is plotted in figure 4 for different values of P. Note that the critical points of $p^{\lambda}(V|P)$ are exactly the solutions of (3.2.1) with $T = T_2$. For

$P > P_2$ there is only one critical point. For $P \leq P_2$ two additional critical points - a local minimum and a local maximum - appear from a fold bifurcation. The new maximum grows in height so that for $P \leq P_4$ it is the global maximum. The old local maximum shrinks and annihilates the minimum again in a fold at P_6. Now Laplace's method of steepest descent (next section) shows that as $\lambda \downarrow 0$ the conditional density $p^\lambda(V|P)$ coverges to delta functions supported at the global maxima of $p^\lambda(V|P)$; these 'densities' are plotted in figure 5. The pressure at which the limiting conditional densities jump is the pressure P_4 at which the two local maxima of $p^\lambda(V|P)$ are equal i.e. Maxwell's equal-area rule. Thus the limiting behavior of (3.2.6), (3.2.7) as $\varepsilon \downarrow 0$, $\lambda \downarrow 0$, and $\mu \downarrow 0$, in that order, is the deterministic system

$$\begin{aligned} \dot{P} &= f(t) \\ V &= \psi(P) \end{aligned} \qquad (3.2.12)$$

where ψ is given by figure 5 for $V < \bar{V}$ and $V > \bar{V}_4$ and $\psi(P_4) = V_4$ or $\bar{V}$ with probability $\frac{1}{2}$ each.

3.2.3.1 Deterministic Bifurcation

Roughly speaking, bifurcation is the study of branching in the equilibrium behavior of a dynamical system in response to small changes in the parameters of the system. Consider, for example, the class of gradient systems*

$$\dot{x} = -\frac{1}{2} \operatorname{grad} S(x,u) \qquad (3.2.13)$$

with $S(x,u)$ a smooth (C^∞) proper function growing sufficiently rapidly as $||x|| \to \infty$ in $\mathbb{R}^n$, for each fixed $u_o \in \mathbb{R}^m$. It is well-known [HIR 75] that every trajectory of (3.2.13) converges to an equilibrium point of (3.2.13) and that every critical point of $S(x,u_o)$ is an equilibrium point of (3.2.13). Further, if for some u_o, $S(x,u_o)$ is a Morse function [ABR 78] then every stable equilibrium of (3.1) is a strict local minimum of $S(x,u_o)$ and conversely. In several practical problems [POS 78] it is of interest to study the variation of the critical points of $S(x,u)$ with the parameter u or, in other words, to study solutions of**

$$D_1 S(x,u) = 0 \qquad (3.2.14)$$

as u varies. If x^*_o is a nondegenerate critical point of $S(x,u_o)$ then there exists a smooth function $x^*(u)$ defined in a neighborhood of u_o such that $x^*(u_o) = x^*_o$ and $x^*(u)$ is the only critical point of $S(x,u)$ in some fixed neighborhood of x^*_o. This is the implicit function theorem. Thus smooth continuation of critical points is locally possible from a nondegenerate critical point.

Now suppose x^*_o is a degenerate critical point i.e.

$$\operatorname{rank} D_1^2 S(x^*_o,u_o) = r < n \quad . \qquad (3.2.15)$$

By using the implicit function theorem on the r-dimensional range space of $D_1^2 S(x^*_o,u_o)$ one can show (the method of Lyapunov-Schmidt [HAL 77]) that the study of the n equations in n unknows (3.2.14) reduces to the study of n-r equations in n-r unknowns

$$N(q,u) = 0 \quad ; \qquad (3.2.16)$$

here $q = Px$ where P is the projection onto the kernel of $D_1^2 S(x^*_o,u_o)$ and the first derivative of the bifurcation function N vanishes at (q^*_o,u_o) where $q^*_o = Px^*_o$. The nature of the solution set of (3.2.14) is thus dependent on the function N. For example suppose $r = n-1$ (the codimension one case); the function N is then a scalar function of a scalar variable that is at least quadratic near q^*_o. If

$$\frac{\partial^2}{\partial q^2} N(q^*_o,u_o) \neq 0$$

then in a sufficiently small neighborhood of (q^*_o,u_o) there is a unique $q^*(u)$ such that

$$\frac{\partial}{\partial q} N(q^*(u),u) = 0 \quad ,$$

i.e. N has a local maximum or minimum at $(q^*(u),u)$; for definiteness assume that it is a maximum. It can then be shown [HAL 77] that if

$$\xi(u) = N(q^*(u),u)$$

then the equations (3.2.14)

(i) have no solution if $\xi(u) > 0$

(ii) have one solution if $\xi(u) = 0$

(iii) have two solutions if $\xi(u) < 0$,

in a neighborhood of u_o. This is the fold bifurcation and is visualized in figure 6. If

$$\frac{\partial^2}{\partial q^2} N(q^*_o,u_o) = 0 \text{ but } \frac{\partial^3}{\partial q^3} N(q^*_o,u_o) \neq 0$$

then $N(q,u_o)$ is at least cubic in $q-q^*_o$ and the bifurcation is a cusp bifurcation (figure 7). Equations (3.2.14) then have one, two, or three solutions in a neighborhood of u_o.

We do not discuss other bifurcations here. Suffice it to say that the normal form theorems of singularity theory (see Thom [THO 75] Hale [HAL 77]) yield universal unfoldings of singularities; the bifurcations of $N(q,u)$ are in general sections of one of these unfoldings.

3.2.3.2 Bifurcation in the Presence of Small Noise

Consider, for example, equation (3.2.13) with added white noise

$$\dot{x} = -\frac{1}{2} \operatorname{grad} S(x,u) + \sqrt{\lambda}\, \xi(t) \qquad (3.2.17)$$

where $\xi(\cdot)$ is an u-dimensional standard white noise process and λ is a positive constant scaling its variance. For each u in $\mathbb{R}^m$, equation (3.2.17) generates a diffusion whose probability density $p^\lambda(t,x,u)$ satisfies the Fokker-Planck equation

$$\frac{\partial}{\partial t} p^\lambda = \frac{1}{2} \sum_{i=1}^{n} \left[\lambda \frac{\partial^2}{\partial x_i^2} + \frac{\partial}{\partial x_i} (\operatorname{grad} S)_i\right] p^\lambda \qquad (3.2.18)$$

where $(\operatorname{grad} S)_i$ is the ith component of the vector grad S. We now assume that the derivatives of S grow rapidly enough at ∞ such that as $t \uparrow \infty$, the

---

* Here grad stands for gradient with respect to x, using the standard inner product on $\mathbb{R}^n$.

** $D_1 S(x,u)$ ($D_1^2(x,u)$) stands for the first (second) derivative of S with respect to x, while $D_2 g(x,y)$ stands for the first derivative of g with respect to y.

density $\bar{p}^{\lambda}(t,x,u)$ converges exponentially [PAP 77] to a unique invariant density $\bar{p}^{\lambda}(x,u)$; the density $\bar{p}^{\lambda}$ is then given by

$$\bar{p}^{\lambda}(x,u) = \exp[-S(x,u)/\lambda]/c^{\lambda}(u) \qquad (3.2.19)$$

where $c^{\lambda}(u)$ is chosen such that $\bar{p}^{\lambda}$ integrates to 1.

Note that for all $\lambda > 0$ and u in $\mathbb{R}^m$ the critical points of $\bar{p}^{\lambda}(x,u)$ are precisely the equilibrium points of the deterministic system (3.2.13). Further, if for some u_o, $S(x,u_o)$ is a Morse function then for all $\lambda > 0$ every local maximum of $\bar{p}^{\lambda}(x,u)$ is a stable equilibrium of (3.2.13). Thus the study of the bifurcations of the critical points of $S(x,u)$ also yields the bifurcations of the critical points of $\bar{p}^{\lambda}(x.u)$. Here we study the bifurcations of $\bar{p}^{\lambda}(x,u)$ in the limit as $\lambda \downarrow 0$ using Laplace's method of steepest descent (for more information see chapter 4 of [HIJ 80]. We will need the following version of the method.

Theorem. Let, for each $u \in \mathbb{R}^m$, $S(x,u)$ have global minima at $x_1^*(u),\ldots, x_N^*(u)$, where N may depend on u. Let them all be nondegenerate. Let $S(x,u)$ have at least quadratic growth as $x \to \infty$. Then in the limit as $\lambda \downarrow 0$, $\bar{p}^{\lambda}(x,u)$ converges to

$$\sum_{i=1}^{N} a_i \delta(x-x_i^*) \Big/ \sum_{i=1}^{N} a_i$$

where $a_i(u) = \det(D_1^2 S(x_i^*,u))^{-\frac{1}{2}}$. More precisely if $\phi(x,u)$ is a smooth function having polynomial growth as $x \to \infty$, then

$$\lim_{\lambda \downarrow 0} \bar{\phi}_{\lambda}(u) \equiv \lim_{\lambda \downarrow 0} \int_{\mathbb{R}^n} \phi(x,u)\bar{p}^{\lambda}(x,u)dx =$$

$$\sum_{i=1}^{N} a_i \phi(x_i^*,u) \Big/ \sum_{i=1}^{N} a_i \equiv \bar{\phi}_o(u) \ .$$

Moreover if the above growth conditions on S and ϕ are uniform in u, for $|u| \leq R$, then $\bar{\phi}_{\lambda}$ is bounded on $|u| \leq R$ uniformly in $\lambda > 0$ and

$$\int_{|u| \leq R} |\bar{\phi}_{\lambda}(u) - \bar{\phi}_o(u)|^p du \to 0 \quad ,$$

for all $R > 0$ and $p \geq 1$.

Proof. The proof that $\bar{\phi}_{\lambda}(u) \to \bar{\phi}_o(u)$ pointwise in u is a slight modification of that appearing in [HIJ 80]. For the L^p convergence, use the dominated convergence and Egoroff's theorems.

We therefore see that bifurcations in the presence of small noise are obtained from the study of changes in (non-degenerate) global minima of $S(x,u)$ as a function of u. In contrast to section 3.3.1 appearance and disappearance of global minima will be points of bifurcation in this context. We close this section with the remark that if the drift in equation (3.2.17) were not a gradient, then although the invariant density $\bar{p}^{\lambda}(x,u)$ will not necessarily be of the form (3.2.19), it will be so asymptotically as $\lambda \downarrow 0$, which is enough for the above theorem. This will be developed elsewhere.

3.2.4. Noisy Constrained Dynamical Systems

3.2.4.1 Deterministic Constrained Dynamical Systems

Consider a constained or implicitly denfined dynamical system of the form:

$$\dot{x} = f(x,y) \qquad (3.2.20)$$

$$0 = g(x,y) \qquad (3.2.21)$$

Here $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, f and g are smooth maps $\mathbb{R}^n \times \mathbb{R}^m$ into $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Assume that 0 is a regular value of g. We interpret (3.2.20), (3.2.21) as describing implicitly a dynamical system on the n-dimensional manifold

$$M = \{(x,y) \,|\, g(x,y) = 0\} \subset \mathbb{R}^{n+m} \qquad (3.2.22)$$

Alternatively, one can think of (3.2.20) as describing a control system with control variable y and (3.2.21) describing implicitly a feedback control law, a situation that arises frequently in optimal control for example; (3.2.21) would then correspond to the Euler-Lagrange equations. The venctor field X(x,y) on M is defined by specifying its projection along the x-axis, namely

$$\pi X(x,y) = f(x,y) \quad , \qquad (3.2.23)$$

where π is the projection map $(x,y) \to x$. At points where the mxm matrix $D_2 g(x,y)$ is full rank, it is clear that (3.2.23) uniquely specifies X(x,y). Difficulties occur at points (x,y) where $D_2 g(x,y)$ is not of full rank and f(x.y) is transverse to $\pi TM(x,y)$. It would seem then that the trajectory should instantaneously jump off the manifold from (x,y) to some other point (x,y') on M (x is constrained to vary continuously by (3.2.20). (see figure 8). This intuition is made precise in [SAS 80] by interpreting solution trajectories of (3.2.20), (3.2.21) to be the 'degenerate' limit as $\varepsilon \downarrow 0$ of the trajectories of

$$\dot{x} = f(x,y)$$

$$\varepsilon \dot{y} = g(x,y) \quad , \qquad (3.2.24)$$

provided the limits exist. Under certain conditions this regularization has been shown to make physical sense in the context of electrical circuits in [SAS 80]. We illustrate the theory with an example.

Consider the system of equations in the plane given by

$$\dot{x} = y \qquad (3.2.25)$$

$$0 = -x - y^3 + y \quad , \qquad (3.2.26)$$

the degenerate Van der Pol oscillator [ZEE 72]. The phase portrait for the degenerate system including jumps from two fold singularities of the projection map $\pi : (x,y) \to x$ is shown in figure 9. Note the relaxation oscillation formed by including the two jumps.

3.2.4.2 Noisy Constrained Systems

We add noise to the system (3.2.20), (3.2.24) in such a way as to obtain

$$\dot{x} = f(x,y) + \sqrt{\mu}\, \xi(t) \quad x(0) = x_o \qquad (3.2.27)$$

$$\varepsilon \dot{y} = g(x,y) + \sqrt{\lambda\varepsilon}\, \eta(t) \qquad (3.2.28)$$

Here $\xi(\cdot)$ and $\eta(\cdot)$ are, as before, independent $\mathbb{R}^n$-valued and $\mathbb{R}^m$-valued white noise processes

and λ, μ scale their variance. For each μ, $\lambda > 0$ systems of this kind have been studied in the limit $\varepsilon \downarrow 0$ extensively by Papanicolaou, Strook and Varadhan [PAP 77]. Our contribution here is to study the behavior of (3.2.27), (3.2.28) in the further limit $\lambda \downarrow 0$ followed by $\mu \downarrow 0$, and to present it in contrast to the results of section 3.4.1. We note in passing that the results remain unchanged if the order in which $\lambda, \mu \downarrow 0$ is interchanged. However, if $\lambda \downarrow 0$ before $\varepsilon \downarrow 0$ then we are reduced to the case in section 3.2.4.1. In order to apply Laplace's method, restrict attention to the case

$$g(x,y) = -\frac{1}{2} \operatorname{grad} S(x,y) \quad ,$$

the gradient taken with respect to y, for some S smooth proper and growing sufficiently rapidly as $y \to \infty$, uniformly for $|x| \leq R$, for all $R > 0$. Note that here x and y play the roles of u and x in section 3.2.3.2. For each ε, λ, $\mu > 0$, the evolution of the corresponding probability density $p^{\lambda}_{\mu,\varepsilon}$ is governed by

$$\frac{\partial}{\partial t} p^{\lambda}_{\mu,\varepsilon} = (L_o^* + \frac{1}{\varepsilon} L_1^*) p_{\mu,\varepsilon}$$

where L_o^*, L_1^* are formal adjoints of the operators L_o, L_1 given by

$$L_o p = \sum_{i=1}^{n} [\frac{\mu}{2} \frac{\partial^2}{\partial x_i^2} + f_i \frac{\partial}{\partial x_i}] p$$

and

$$L_1 p = \sum_{i=1}^{m} [\frac{\lambda}{2} \frac{\partial^2}{\partial y_i^2} + g_i \frac{\partial}{\partial y_i}] p \quad .$$

The conditional density of y given x is, in the limit $\varepsilon \downarrow 0$, given by

$$\bar{p}^{\lambda} = \exp[-S(x,y)/\lambda]/c^{\lambda}(x)$$

where $c^{\lambda}(x)$ is chosen so that $\bar{p}^{\lambda}$ integrates to one. Set

$$\bar{f}_{\lambda}(x) = \int_{\mathbb{R}^m} f(x,y) \bar{p}^{\lambda}(x,y) dy \quad ,$$

for $\lambda > 0$. Assume that $f(x,y)$ has at most polynomial growth as $y \to \infty$, uniformly in $|x| \leq R$, for all $R > 0$. It then follows that $\bar{f}_{\lambda}(x)$ is bounded on $|x| \leq R$ uniformly in $0 < \lambda < \lambda_o$, for some λ_o sufficiently small. We assume in what follows that $\bar{f}_{\lambda}(x)$ has linear growth uniformly in $0 < \lambda \leq \lambda_o$. Set

$$\bar{L}_o p = \sum_{i=1}^{n} [\frac{\mu}{2} \frac{\partial^2}{\partial x_i^2} + (\bar{f}_{\lambda})_i \frac{\partial}{\partial x_i}] p \quad .$$

The operator $\bar{L}_o$ is the operator L_o averaged over the invariant density $\bar{p}^{\lambda}$ of y given x.

Theorem [10] As $\varepsilon \downarrow 0$ the first component $t \to x(t)$ of the solutions of equations (3.2.27), (3.2.28) converges weakly in $C([0,T]; \mathbb{R}^N)$ to the unique diffusion, denoted by $t \to x_{\lambda,\mu}(t)$, governed by $\bar{L}_o$.

We now have the following

Theorem. As $\lambda \downarrow 0$ the diffusion $t \to x_{\lambda,\mu}(t)$ converges weakly in $C([0,T]; \mathbb{R}^n)$ to the unique diffusion $t \to x_{\mu}(t)$ satisfying in law

$$\dot{x} = \bar{f}_o(x) + \sqrt{\mu}\, \xi(t) \quad x(0) = x_o \qquad (3.2.29)$$

where

$$f_o(x) = \sum_{i=1}^{N} a_i(x) f(x, y_i^*(x)) / \sum_{i=1}^{N} a_i(x)$$

as in section 3.2.3.2, and $y_1^*(x), \ldots, y_N^*(x)$ are the nondegenerate global minima of $S(x,\cdot)$.

Proof. The proof of this and the next theorem mimic that of the previous theorem and so we only outline the proof. The reader may perfer to master the proof of the preceding theorem appearing in [PAP77] first. We first show that the measures on $C([0,T]; \mathbb{R}^n)$ corresponding to $t \to x_{\lambda,\mu}(t)$ are relatively weakly compact in $C([0,T]; \mathbb{R}^n)$. The second step is to identify any limiting measure of the family $x_{\lambda,\mu}$, $\lambda > 0$, as the unique solution of the martingale problem associated to (4.10). It is essential in both steps 1 and 2 that μ remain fixed and positive.

To prove compactness in $C([0,T]; \mathbb{R}^n)$ it suffices to show that

$$\limsup_{\delta \downarrow 0} \sup_{\lambda > 0} P\Big(\sup_{\substack{|t-s| \leq \delta \\ 0 \leq t,s \leq T}} |x_{\lambda,\mu}(t) - x_{\lambda,\mu}(s)| \geq \varepsilon \Big) = 0$$

for all $\varepsilon > 0$ ([13] page 351). This follows from the fact that

$$\dot{x}_{\lambda,\mu} = \bar{f}_{\lambda}(x_{\lambda,\mu}) + \sqrt{\mu}\, \xi(t) \quad x(0) = x_o$$

in law (which is all we need here), and the assumption on $\bar{f}_{\lambda}(x)$ [KRY 80] page 120).

For step 2, take a C^{∞} function ϕ whose support is in $|x| < R$ and consider

$$\phi(x_{\lambda,\mu}(t)) - \phi(x_o) - \int_0^t f_{\lambda}(\phi)\ (x_{\lambda,\mu}(s)) ds$$

$$- \frac{\mu}{2} \int_0^t \Delta(\phi)(x_{\lambda,\mu}(s)) ds \ ,$$

where $f(\phi)$ is the directional derivative of ϕ in the direction of the vector field f and Δ is the Laplacian. For each $\lambda > 0$ this expression is a martingale and the idea is that the limit is also a martingale. Note that to get the correct expression in the limit, one has to show that

$$E\Big(\int_0^T |\bar{f}_{\lambda}(\phi)(x_{\lambda,\mu}(s)) - \bar{f}_o(\phi)(x_{\lambda,\mu}(s))| ds \Big) \qquad (3.2.30)$$

goes to zero as $\lambda \downarrow 0$. This follows from the fact that for all p sufficiently large there is a $K > 0$ depending only on p, a uniform bound for $\bar{f}_{\lambda}(x)$, $|x| \leq R$, $0 < \lambda \leq \lambda_o$, and T, such that (3.2.30) is bounded by

$$K \| \bar{f}_{\lambda}(\phi) - \bar{f}_o(\phi) \|_p \qquad (3.2.31)$$

where the L^p norm is over $|x| \leq R$ ([KRY 80] page 52). But from section 3.2.3.2, we know that (3.2.31) goes to zero as $\lambda \downarrow 0$. This completes the proof.

Finally we can let $\mu \downarrow 0$ to obtain

Theorem. The family $t \to x_{\mu}(t)$, $\mu > 0$ is relatively

weakly compact in $C([0,T];\mathbb{R}^n)$. Any limiting diffusion of the diffusions $t \to x_\mu(t)$, $\mu > 0$, as $\mu \downarrow 0$ then satisfies the ordinary differential equation

$$\dot{x} = \bar{f}_o(x) \quad , \quad x(0) = x_o \quad . \tag{3.2.32}$$

Proof. As before one checks that $t \to x_\mu(t)$, $\mu > 0$, is relatively weakly compact, as before any limiting diffusion is then governed by $\bar{f}_o$, i.e.

$$\phi(x(t)) = \phi(x) + \int_0^t \bar{f}_o(\phi)(x(s))ds + \text{martingale} .$$

Since the variance of the martingale is

$$E\int_0^t [\bar{f}_o(\phi^2) - 2\phi\bar{f}_o(\phi)](x(s))ds$$

and $f(\phi^2) - 2\phi f(\phi)$ is zero for any vector field f, it follows that the martingale part is identically zero and (3.2.32) holds. □

Let us consider, as an example, the noisy version of the degenerate Van der Pol oscillator. Consider

$$\dot{x} = y + \sqrt{\mu}\,\xi(t)$$

$$\varepsilon\dot{y} = -x - y^3 + y + \sqrt{\lambda\varepsilon}\,\eta(t)$$

For $\lambda, \mu > 0$ as $\varepsilon \downarrow 0$ the x-process converges to one satisfying

$$\dot{x} = \bar{y}^\lambda(x) + \sqrt{\mu}\,\xi(t)$$

where

$$\bar{y}^\lambda(x) = \int_{-\infty}^{+\infty} y \exp\frac{2}{\lambda}\left(-xy - \frac{y^4}{4} + \frac{y^2}{2}\right)dy \cdot \left[\int_{-\infty}^{\infty} \exp\frac{2}{\lambda}\left(-xy - \frac{y^4}{4} + \frac{y^2}{2}\right)dy\right]^{-1} .$$

In figure 10, $\bar{y}^\lambda(x)$ is plotted for $\lambda_1 > \lambda_2 > 0$.

In the further limit that $\lambda \downarrow 0$ followed by $\mu \downarrow 0$, x satisfies

$$\dot{x} = \psi(x) \qquad x \neq 0$$

$$= 0 \qquad x = 0$$

where $\psi(x)$ is shown heavy in figure 10. Note that ψ is discontinuous at $x = 0$, since the support of the conditional density jumps from one leg of the curve $x = y - y^3$ to the other leg as shown in figure 10. Consequently the relaxation oscillation is broken up by the presence of small noise.

3.3 Singular Perturbation, State Aggregation and Nonlinear Filtering

3.3.1. Variable Structure Systems

In situations where the structure of a dynamical system varies in time, it is often the case that the structural changes occur on a time scale that is much slower than the dynamics in any given mode of operation. For example, in the study of power systems, the swing equations are sometimes thought of as occuring on a fast time scale when compared to the relatively slow time scale of random faults or breakdowns.

Suppose that $g_1, \ldots, g_N$ are vector fields on $\mathbb{R}^n$, suppose that $A(x) = (a_{ij}(x))$, $1 \le i,j \le N$, is an intensity matrix for each x in $\mathbb{R}^n$. If there are N possible modes of operation of the system and if $a_{jk}(x)$ represents the infinitesimal transition probability that a structural change from mode j to mode k happens when the system is in state x, then a natural formulation of the above situation is to consider the trajectories of

$$\dot{x} = (1/\varepsilon)\, g_i(x) \tag{3.3.1}$$

where $x = x(t)$ is the state at time t and $i = i(t)$, a jump process on $\{1,\ldots,N\}$ governed by $(a_{jk}(x(t))$, represents the mode in operation at time t. More accurately and more concisely, a natural formulation of the above situation is to say that $t \to (x(t), i(t))$ is a Markov process on $X = \mathbb{R}^n \times \{1,\ldots,N\}$ governed by $A + (1/\varepsilon)g$.

One can generalize (3.3.1) in various directions. For the purposes of system identification one may replace $\{1,\ldots,N\}$ by an arbitrary parameter space Λ (for related work see [HIJ 81b]). Alternatively, the state space need not be $\mathbb{R}^n$ and may be replaced by any smooth manifold. In fact all of these situations are subsumed by the following set-up:

Let X be a smooth manifold and let g be a smooth vector field on X. Let A be an integral operator on X given by

$$A\Phi(x) = \int_X (\Phi(x') - \Phi(x))\,\mu(x,dx'), \tag{3.3.2}$$

for some measures $B \to \mu(x,B)$ depending on x in X. For each $\varepsilon > 0$ let $t \to x^\varepsilon(t)$ be a Markov process on X governed by $A + (1/\varepsilon)g$.

The purpose of this paper is to study the limiting behavior of these processes as $\varepsilon \downarrow 0$. Our main result is that while the original motion $t \to x^\varepsilon(t)$ clearly blows up as $\varepsilon \downarrow 0$, in certain cases there is a reduced-order state space $\bar{X}$ and a projection $\pi: X \to \bar{X}$ such that $t \to \bar{x}^\varepsilon(t) \equiv \pi(x^\varepsilon(t))$ converges to a well-defined limit as $\varepsilon \downarrow 0$. Thus X may be regarded as the full-order state space while A and g are the generators of the slow and fast dynamics respectively.

In general $\bar{X}$ should be chosen to be the limit set of the vector field g. In this paper we deal with the simplest kind of limiting behavior, when the limit set of g is given by a finite number of states $\bar{X} = \{\bar{x}_1, \ldots, \bar{x}_N\}$ in X. Even in this case there are a number of novel features. Viewed as a singular perturbation problem, here there is no "fast variable" and thus the state space X is not a product of a fast variable and the slow variable $\bar{x} = \pi(x)$. Viewed as a state aggregation problem, here we have aggregation from a continuum of states X to a finite state space $\bar{X}$, a fact that radically changes the level of computational difficulty of nonlinear filtering and (partially or fully observable) stochastic control problems associated to the processes $t \to x^\varepsilon(t)$.

As an application of our main result we shall see that while the nonlinear filters corresponding to the problem of estimating $t \to x^\varepsilon(t)$ in the presence of additive white noise do not converge as $\varepsilon \downarrow 0$, it turns out that the projected filters do in fact converge to well-defined object, the (finite-dimensionally computable) finite-state Wonham filter.

Our treatment here is based on the martingale formulation an analogous theorem due to Papanicolaou, Stroock and Varadhan [PAP 77]. We therefore begin, for the sake of completeness, with a review of the martingale problem for $A + (1/\varepsilon)g$. For a general treatment of the martingale problem for Levy

processes, see [STR 75].

3.3.2. The Martingale Problem for A + (1/ε)g

Let X denote a smooth manifold and leg g be a smooth complete vector field on X. Let B(X) be the space of all bounded Borel functions on X, and let α_t, $-\infty < t < \infty$, denote the flow of g. The domain of g is the set $\mathcal{D}$ of all functions Φ in B(X) such that there is a ψ in B(X) satisfying

$$\Phi(\alpha_t(x)) - \Phi(\alpha_s(x)) = \int_s^t \psi(\alpha_r(x))dr$$

for all x in X and $0 \leq s \leq t \leq T$. Any such ψ is then denoted by $g(\Phi)$ and we emphasize that there may be more than one $g(\Phi)$ associatied to a given Φ. If Φ is sufficiently smooth, however, then there is a natural choice of $g(\Phi)$ given by

$$g(\Phi)(x) = \frac{d}{dt}\Big|_{t=0} \Phi(\alpha_t(x)).$$

We note for future reference that $\mathcal{D}$ is a vector space.

Let Ω denote the space of all right-continuous paths $\omega : [0,T] \to X$ having only a finite number of discontinuities of the first kind in any compact time interval. For each $0 \leq t \leq T$ let $x(t) : \Omega \to X$ be the evaluation map at time t: $x(t,\omega) = \omega(t)$. The Borel σ-algebra of Ω is then given by $\mathcal{F}_T$, where $\mathcal{F}_t$ is the σ-algebra generated by the maps x(s), $0 \leq s \leq t$.

If $B \to \mu(x,B)$ is a finite positive Borel measure on X for each x in X such that $x \to \mu(x,B)$ is in B(X) for each Borel set $B \subset X$, let $A\Phi$ be given by (2), for any Φ in B(X). A is then a bounded linear operator on B(X) whose norm is less than or equal to twice the sup norm of λ where $\lambda(x) \equiv \mu x,X)$. Let $C_o^\infty(X)$ denote the space of all smooth functions of compact support on X. Let $L_t : C_o^\infty(X) \to B(X)$ be a linear operator depending on t. We use the standard martingale definition of a Markov process [5]:

Definition. A Markov process on X governed by L_t is a probability measure P on Ω satisfying

$$E(\Phi(x(t)) - \Phi(x(s)) - \int_s^t L_r(\Phi)(x(r))dr \,|\, \mathcal{F}_s) = 0 \qquad (3.3.3)$$

for all Φ in $C_o^\infty(X)$ and all $0 \leq s \leq t \leq T$. Recall that this equivalent to the statement that for any bounded $\mathcal{F}_s$ - measurable $\Phi : \Omega \to R$,

$$E(\{\Phi(x(t)) - \Phi(x(s)) - \int_s^t L_r(\Phi)(x(r))dr\}\ \Phi\ ) = 0$$

for $0 \leq s \leq t \leq T$. By abuse of notation, the measure P is referred to as the distribution of the Markov process $t \to x(t)$.

Let $G_t\Phi(x) \equiv \Phi(\alpha_t(x))$ and set $A_t = G_t A\, G_{-t}$. Consider the map $\alpha_\varepsilon : \Omega \to \Omega$ given by $\alpha_\varepsilon(\omega)(t) \equiv \alpha_{-t/\varepsilon}(\omega(t))$ and let $\tilde{P}_\varepsilon$ be the image of a given measure P_ε under the map α_ε.

Lemma. P_ε is governed by A + (1/ε)g if and only if $\tilde{P}$ is governed by $A_{t/\varepsilon}$.

This lemma is proven using integration by parts in (3) exactly as in the proof of theorem (2.1) of [STR 71]. Since $A_{t/\varepsilon}$ is an integral operator the methods of chapter 3, [2] yield the fact that there is one and only one measure $\tilde{P}$ for any given initial distribution governed by ε the operator $A_{t/\varepsilon}$. Thus

Proposition. There is one and only one Markov process on X governed by A + (1/ε)g, for any given initial distribution. Moreover (3) above holds with L_t = A + (1/ε)g for any Φ in the domain $\mathcal{D}$ of g, and

$$P_\varepsilon(t \to x(t) \text{ is a finite disjoint union of compact trajectories of g}) = 1.$$

Proof. The sample paths of P_ε are as stated because P_ε is the image of the measure $\tilde{P}_\varepsilon$ under the map α_ε^{-1} and the sample paths of $\tilde{P}_\varepsilon$ are piecewise constant. To see that (3.3.3) holds with L_t = A + (1/ε)g for all Φ in $\mathcal{D}$ first note that P_ε can be constructed so that (3.3.3) holds for all Φ in B(X), when $L_t = A_{t/\varepsilon}$, and then note that the integration by parts trick referred to above still holds when Φ is in $\mathcal{D}$.

Thus the martingale problem for A + (1/ε)g is well-posed. In particular if $\bar{X} = \{\bar{x}_1,\ldots,\bar{x}_N\}$ is a finite set then $\bar{X}$ can be considered to be a zero-dimensional manifold. Thus suppose (μ_{ij}) $1 \leq i, j \leq N$ is given and set

$$\bar{A}\bar{\Phi}(\bar{x}_i) = \Sigma\, (\bar{\Phi}(\bar{x}_j) - \bar{\Phi}(\bar{x}_i))\, \mu_{ij} \qquad (3.3.4)$$

where the sum is over j, for all $\bar{\Phi}$ in $B(\bar{X}) = C_o^\infty(\bar{X})$. If in the above proposition we make the replacements $X \leftarrow \bar{X}$, $A \leftarrow \bar{A}$, $g \leftarrow 0$ then we conclude that the martingale problem for $\bar{A}$ is also well-posed. In closing this section, we note that the only property of g that we have used is the existence and uniqueness of an associated flow satisfying $C_o^\infty(X) \subset \mathcal{D}$.

3.3.3. Gradient-like Vector Fields

Recall that g is a complete smooth vector field on X with flow α_t. We assume that there are a finite number of points $\bar{x}_1,\ldots,\bar{x}_N$ in X such that for all x in X, $\alpha_t(x)$ converges to one of $\bar{x}_1,\ldots,\bar{x}_N$ as $t \uparrow \infty$. The set $\bar{X} = \{\bar{x}_1,\ldots,\bar{x}_N\}$ represents the reduced order state space. Let $B_i \subset X$ be the ith basin of attraction: B_i is the Borel set of all x in X such that $\alpha_t(x)$ converges to $\bar{x}_i$ as $t \uparrow \infty$. For x in B_i set $\pi(x) = \bar{x}_i$. The map $\pi : X \to \bar{X}$ is then in $\mathcal{D}$ and one choise of $g(\pi)$ is given by the zero function. For Φ in B(X) let $\bar{\Phi}$ denote the restriction of Φ to $\bar{X}$.

Definition. The Fredholm alternative holds for Φ in B(X) iff there is a ψ in $\mathcal{D}$ satisfying

$$g(\psi) = \bar{\Phi}o\pi - \Phi\ . \qquad (3.3.5)$$

Consider the following assumption.

(A) There is an integrable function R(t), $0 \leq t < \infty$ such that

$$|\mu(\alpha_t(x),B_j) - \mu(\pi(x),B_j)| < R(t)$$

for $1 \leq j \leq N$, x in X, and $t \geq 0$, and $\int_0^\infty R(t)dt$ is infinite.

Proposition. Under assumption (A), the Fredholm alternative holds for all functions of the form $(A(\bar{\Phi}o\pi)$, for any given $\bar{\Phi}$ in $B(\bar{X})$.

Proof. Set $\Phi \equiv \bar{\Phi}o\pi$ and

$$\psi(x) \equiv \int_o^\infty A\Phi(\alpha_s(x)) - A\Phi(\pi(x))\, ds.$$

Since $A\Phi$ is a finite linear combination of the functions $x \to \mu(x,B_j)$, assumption (A) guarantees that ψ is in B(X). The rest of the proof follows

m the above formula for $\bar{\psi}$, p as here

If the conclusion of this last proposition is ue, then we shall say simply that the Fredholm ternative for g holds.

In what follows $\bar{\Omega}$ denotes the right-continuous th space of $\bar{X}$. Since the trajectories of P_ε are finite disjoint union of compact trajectories of we see that $\pi : X \to \bar{X}$ includes a well-defined p $\Omega \to \bar{\Omega}$ given by $\omega \to \omega'$ where $\omega'(t) = \pi(\omega(t))$.

The swing equations arising in the study of wer systems can be thought of as a vector field on $X = T^n \times R^n$. For a study of equilibrium ints of this vector field g, see [ARA 81b].

3.4. Singular Perturbation t first. Please leave ½" ace between and of Abstract and first line of text.

Let A and g be as before, and define $\bar{A}$ by

$$\bar{A}\bar{\phi} = A(\bar{\phi}\circ\pi)\Big|_{\bar{X}} .$$

r linear map $\bar{A}: B(\bar{X}) \to B(\bar{X})$ is then given by uation (4) where

$$\mu_{ij} = \mu(\bar{x}_i, B_j), \quad 1 \le i,j \le N.$$

t follows is the main result of the paper.

orem. Assume that the Fredholm alternative for olds. Let $t \to x^\varepsilon(t)$ be Markov processes on X erned by $A + (1/\varepsilon)g$, all having a common initial tribution on X. Then the Markov processes $t \to$ t) coverage in distribution to the unique Markov cess on $\bar{X}$ governed by $\bar{A}$ and having the projected tial distribution, as $\varepsilon \downarrow 0$. This means that for bounded continuous functional $\bar{\Phi} : \bar{\Omega} \to R$

$$\bar{E}_\varepsilon(\bar{\Phi}) \to E(\Phi)$$

$\varepsilon \downarrow 0$.

The proof of this theorem is analogous to that a theorem due to Papanicolaou, Stroock and adhan [4], and breaks naturally into two steps. first step consists of showing that the distri- ions $\{\bar{P}_\varepsilon\}$ of $t \to \bar{x}^\varepsilon(t)$ are a relatively weakly pact family of measures on $\bar{\Omega}$, while the second p is the identification of the limiting dis- bution $\bar{P}$ via the Fredholm alternative and the l-posedness of the martingale problem for $\bar{A}$.

The topology on $\bar{\Omega}$ is the Skorokhod topology. s turns $\bar{\Omega}$ into a complete metric space and thus Prohorov theory applies: A family of measures on $\bar{\Omega}$ is relatively weakly compact if $\{\bar{P}_\varepsilon\}$ is formly tight: For each $\alpha > 0$ there is a compact $K \subset \bar{\Omega}$ such that $\bar{P}_\varepsilon(K) > 1 - \alpha$ for all $\varepsilon > 0$. e $\bar{X}$ may be considered as embedded in a real the standard theory applies and so we conclude $\{\bar{P}_\varepsilon\}$ is relatively weakly compact, using a ial case of proposition (A.1) or [STR 75].

Now suppose that $\varepsilon_k \downarrow 0$ and $\bar{P}_{\varepsilon_k} \to$ some P' on $\bar{\Omega}$. $\bar{\phi}$ be in $B(\bar{X})$ and choose ψ in D such that

$$g(\psi) = \bar{A}(\phi)\circ\pi - A(\phi\circ\pi) . \qquad (3.3.6)$$

e P_{ε_k} is governed by $A + (1/\varepsilon)g$ and $\phi + \varepsilon_k\psi$ is we have ($\phi \equiv \bar{\phi}\circ\pi$)

$$= E_{\varepsilon_k}\left([(\phi+\varepsilon_k\psi)(x(t)) - (\phi+\varepsilon_k\psi)(x(s))\right.$$

Start second integral on next line if necessary

$$\left. - \int_s^t (A+(1/\varepsilon_k)g)(\phi+\varepsilon_k\psi)(x(r))dr\right]\bar{\Phi}\circ\pi)$$

for all bounded $\bar{\mathcal{F}}_s$ - measurable $\bar{\Phi} : \bar{\Omega} \to R$. Thus by expanding and using (3.3.6) we see that

$$E_{\varepsilon_k}([\bar{\Phi}(\bar{x}(t)) - \bar{\Phi}(\bar{x}(s)) - \int_s^t \bar{A}\bar{\Phi}(\bar{x}(r))dr]\,\bar{\Phi})$$

is $0(\varepsilon_k)$ as $\varepsilon \downarrow 0$, since ψ, $A(\psi)$, and $\bar{\Phi}$ are all bounded. Thus letting $k \uparrow \infty$, we see that any limiting probability measure of the set $\{\bar{P}_\varepsilon\}$ is a Markov process governed by $\bar{A}$, and since there is a unique such Markov process, this shows that $\bar{P}_\varepsilon \to \bar{P}$, the Markov process governed by $\bar{A}$.

Begin With a little extra work, the above result still holds for all bounded $\bar{\Phi}$ that are continuous off a set of $\bar{P}$ measure zero. See [PAR 67].

3.3.5. Filtering

Suppose we are given noisy observations

$$y(t) = h(x^\varepsilon(t)) + \text{white noise} \qquad (3.2.7)$$

of the Markov process $t \to x^\varepsilon(t)$. The nonlinear filter corresponding to (3.2.7) is a well-defined map given by the Kallianpur-Striebel formula for example. Rather than use this formula, we shall use the robust form of the filter and simply define it to be the expectation of a certain bounded functional on $\bar{\Omega}$. For each y in $C([0,T])$ and Φ, h in $B(X)$ let

$$\Phi_y(\Phi) = (x(t))\exp(-\int_0^t V(s,x(s))ds)$$

where

$$V(t,x) = \tfrac{1}{2}(y(t) - h(x))^2 .$$

It can be shown that Φ_y is a continuous map $\Omega \to R$ off a set of P_ε-measure zero.

Definition. The filter corresponding to a Markov process with distribution P is given by the map

$$C([0,T]) \to C([0,T])$$

$$y \to E(\Phi_y(\Phi))/E(\Phi_y(1)) .$$

Since the distribution P_ε of the Markov processes $t \to x^\varepsilon(t)$ do not converge, we do not expect the corresponding filters to converge. However the projected filters, obtained by replacing Φ by $\bar{\Phi}\circ\pi$, h by $\bar{h}\circ\pi$, do in fact converge as $\varepsilon \downarrow 0$:

Theorem. The projected filters converge to the finite-state Wonham filter corresponding to the problem of estimating the finite state process governed by $\bar{A}$, in the presence of additive white noise.

This follows immediately from our main result.

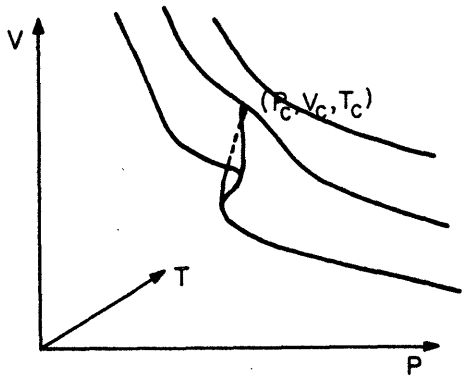


Figure 3.1. The Surface Satisfying (2.1) Showing a Cusp point at (P_c, V_c, T_c)

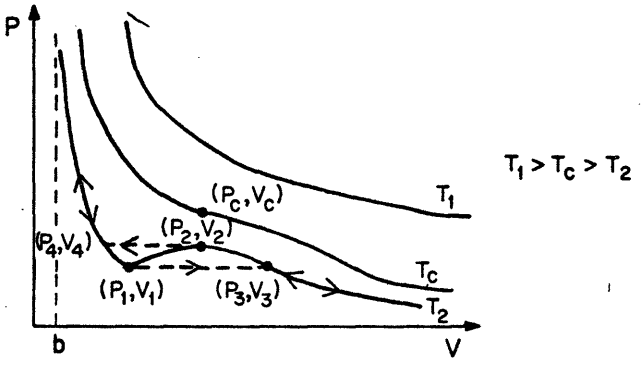


Figure 3.2. Isotherms of the Van der Vaals Gas Equation

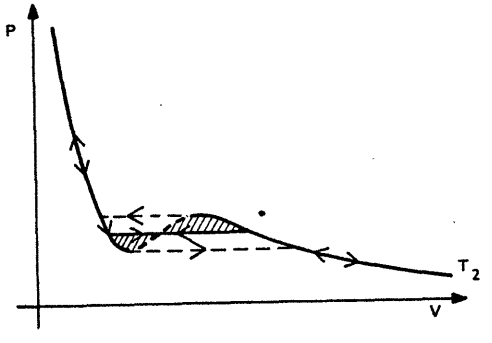


Figure 3.3. Showing the Discrepancy Between the Deterministic Prediction (Dotted) and Experimentally Observed (Solid) Phase Transtion.

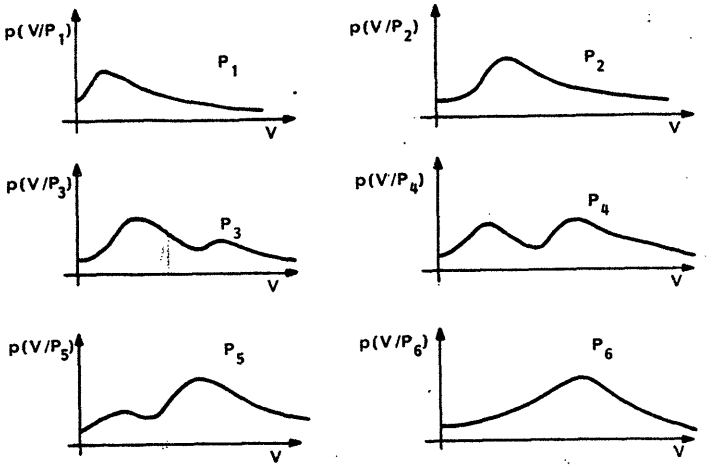


Figure 3.4. The Conditional Density p(V/P) Plotted for Decreasing Values of Pressure $P_1 > P_2 > P_3 > P_4 > P_5 > P_6$.

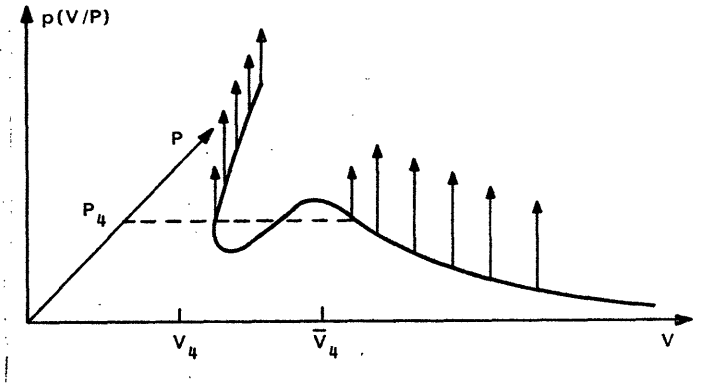


Figure 3.5. The Limit of the Conditional Density p(V/P) when $\lambda \downarrow 0$.

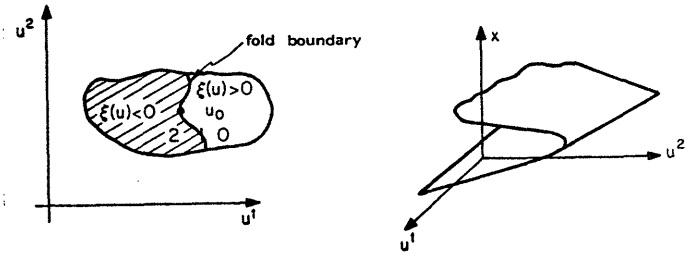


Figure 3.6. Visualization of the Fold Bifurcation

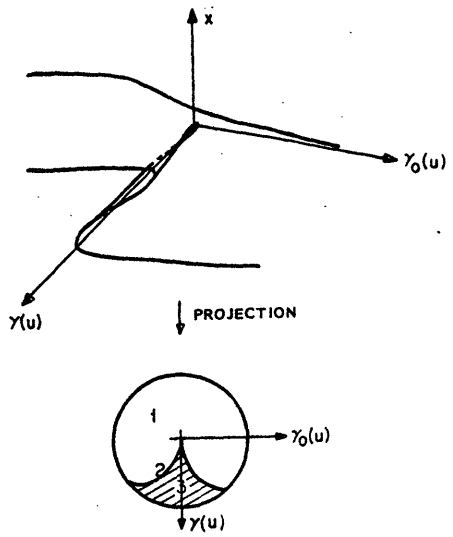


Figure 3.7. The Cusp Birfuraction

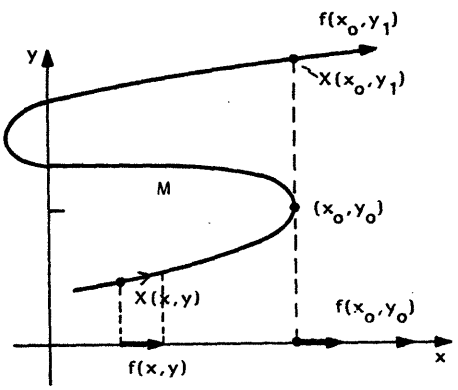


Figure 3.8. Illustrating Jump Behavior in Constrained Systems

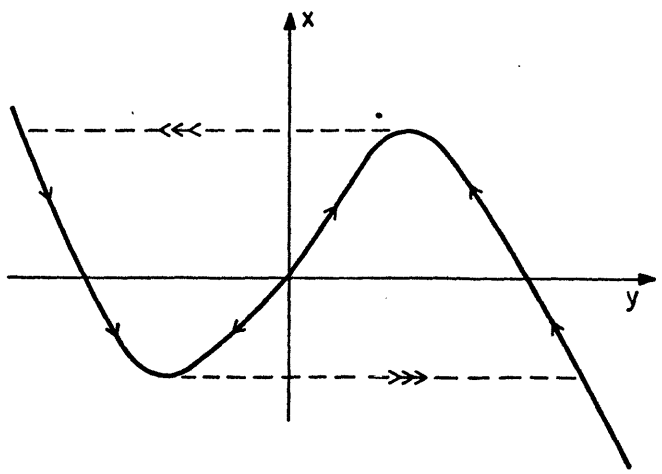


Figure 3.9. Degenerate Form of the Van Der Pol Oscillator Showing Jump Behavior

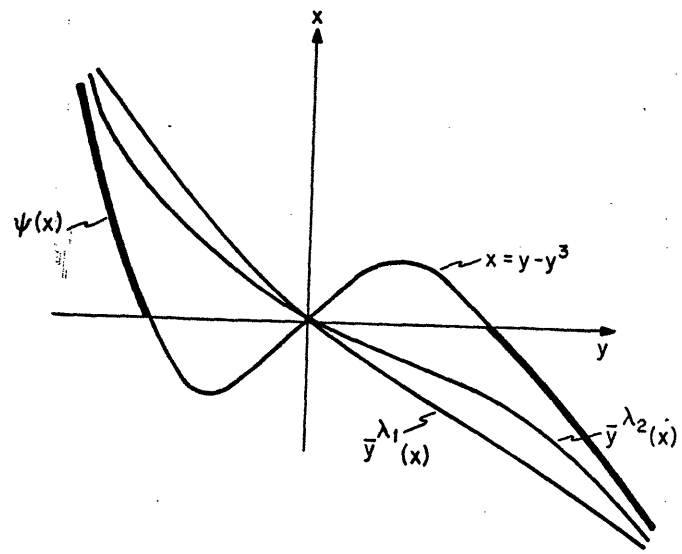


Figure 3.10. The Drift $\bar{y}^{\lambda}(x)$ for the Limit Diffusion of the Degenerate Van Der Pol Oscillator.

4. STABILITY AND CONTROL OF HYBRID SYSTEMS

4.1 Introduction

The material in this chapter deals with the analysis and design of decision-making policies in systems described by stochastic hybrid models. One of our conjectures in the proposal was that a cause of instability in large-scale power systems consists of the local nature of some of the discrete feedback control systems. Many of these control actions fail to anticipate the potential effects of disturbances elsewhere in the region, resulting in overall instability. Most of the work in this chapter is aimed at understanding how discrete control actions may result in instabilities, and how to properly modify them to ensure stability of the overall stochastic system.

In section 4.2, we introduce a control-theoretic formulation suitable for hybrid models: The JLQ optimal control problem. This is the simplest nontrivial control problem involving stochastic hybrid systems. We derive the detailed form of the optimal control law, thereby obtaining valuable insight into the qualitative properties which controller designs for hybrid systems must have. In particular, we establish that an optimal controller identifies regions of hazardous operations, and will attempt to steer the system away from such regions. We describe two separate mechanisms for such steering: passive hedging and active hedging. The results presented in this section are a condensed version of the doctoral thesis of Dr. H. Chizeck, which is under preparation.

In section 4.3, we study in detail the stochastic stability of a specific control scheme for a continuous system. This control scheme known as the Multiple Model Adaptive Control concept, results in a control law which switches between a finite number of possible candidates. The mechanism for the switch is a nonlinear dynamical relation, driven by the output of the system itself. Hence, the resulting controlled system can be represented as a stochastic hybrid system. We provide a detailed analysis of the stability of such a system, isolating classes of behavior which may be typical of more general stochastic hybrid systems.

The work presented in section 4.4 deals with the stability analysis of controlled hybrid systems after an abrupt change in the parameters of the system (such as a generator outage) has occurred. In this setting, a critical factor in maintaining overall stability is to identify the change in the system, as soon as possible, while trying to preserve the stability and integrity of the system.

Basically, new models of the system must be developed quickly; this may involve applying specific controls to determine the effects of the abrupt changes. Furthermore, one must also account for the possibility of further changes due to the initiating event.

In section 4.4, we approach this problem using the principles of adaptive control. We study the problem of how to design control inputs, based on noisy observations, which will provide us with an accurate description of the abrupt change. These two objectives are often contraditory in nature. With this study, we hope to develop analysis methodologies which can be used to guide the actions of controllers when failures occur, so that control actions do not add to the severity of the disturbance, but rather serve to identify quickly a new operating condition.

Our results in this chapter provide an understanding of the principles whereby good control systems for stochastic hybrid models may be designed. Although the complexity of the optimization formulation in section 4.2 appears discouraging, we must point out that there are simpler structures which can be determined using suboptimal approximation procedures. These are the first results which have been obtained with regards to the control of nontrivial stochastic hybrid systems. There are many open problems which remain in this area, which should be the subject of further research.

4.2 Control of Hybrid Systems

One of our conjectures concerning the sources of instability in large-scale power systems was that the actions taken to control the instability would instead promote further spreading of unstable behavior throughout the interconnected system. For example, prolongued operation in a vulnerable emergency state is likely to result in additional breakdown of the system. In order to understand these mechanisms for instability, we studied the abstract problem of how to control a stochastic hybrid system to achieve optimal performance. Due to the interacting nature of the discrete transitions and the continuous evolution of hybrid models, the control strategies would have to compensate for the existence of vulnerable regions of operation, anticipating potential structural changes. By studying the qualitative features of some "optimal" control policies, we were able to identify several important characteristics which well-designed control strategies should possess. This work will be reported in detail in the upcoming thesis of Mr. Howard Chizen [Chi 82]. Preliminary versions of these results have appeared in [Chi 79], [Chi 80a], [Chi 80 b]. In this section, we provide an overview of our results in this area.

We can divide the research into two parts, referring to the class of hybrid systems which each part studied. The first part of the research studied the control of hybrid systems with a hierarchical structure, depicted in figure 4.2.1. The discrete part of the system state evolves independently of the continuous part, and cannot be affected by control actions. Its state, however, affects the evolution of the continuous part. The basic mathematical model for this class of hybrid systems, assuming that the evolution occurs in discrete time, is given by:

$$x(t+1) = A(\rho(t))\, x(t) + B(\rho(t))\, u(t) + G(\rho(t))\, w(t) \qquad (4.2.1)$$

where $\rho(t)$ is the state of a finite-state Markov chain whose transition matrix at time t is denoted as $P(t)$. If $p(t)$ denotes the distribution of $\rho(t)$, then

$$p(t+1) = P(t)\, p(t) \qquad (4.2.2)$$

The matrices A, B, G belong to a finite set, indexed by the discrete state $\rho(t)$. The sequence $w_i(t)$,

$w_j(t_1)$ is a sequence of independent identically distributed, normal random variables with zero mean and unit variance.

The objective function for control design was chosen to be quadratic, of the form

$$J = E\left\{\sum_{t=o}^{T-1} x(t)'\, Q(t,\rho(t))\, x(t) + u'(t)R(t,\rho(t))\, u(t) + x(T)'\, Q(T,p(t))\, x(T)\right\} \quad (4.2.3)$$

Notice that the penalty matrices depend on the discrete state $\rho(t)$; hence, operation under certain undesirable structures can be penalized more. Control problems of this type are referred to as Jump Linear Quadratic (JLQ) control problems.

The study of finite-horizon control problems of the form (4.2.1) - (4.2.3) was initiated by Krasovskii and Lidskii [Kra 61]. Similar formulations in continuous time were studied by Wonham [WON 70], Sworder [SWO 69], Birdwell et. al,. [BIR 78], and Chizeck and Willsky [Chi 70, Chi 80a]. In all of these works, the assumption is made that the full state is observed exactly; that is, $x(t)$ and $\rho(t)$ are known. Under this assumption, a straightforward application of dynamic programming yields the following result:

Proposition 4.2.1 [Chi 80a] The optimal control law $u^*(t)$ which minimizes (4.2.3) subject to (4.2.1) - (4.2.2) is

$$u^*(t,x(t),\rho(t)) = -\, S^{-1}(t,\rho(t))\, K(t,\rho(t))\, x(t) \quad (4.2.4)$$

where

P_{ij} = i - j entry of P. M = dimension P .

$$T(t,\rho(t)) = \sum_{j=1}^{M} P_{p(t),j}\,(t)\; k(t+1,j) \quad (4.2.5)$$

$$k(t,j) = Q(t,j) + A'(j)\, T(t,j)\, A(j) - K(t,j)\, S^{-1}(t,j)\, K(t,j) \quad (4.2.6)$$

$$S(t,j) = R(t,j) + B'(j)\, T(t,j)\, B(j) \quad (4.2.7)$$

$$K(t,j) = B'(j)\, T(t,j)\, A(j) \quad (4.2.8)$$

$$K(T,p) = Q(T,\rho) \quad (4.2.9)$$

Proposition 4.2.1 displays the explicit dependence of the control law on the costs associated with operating in a given form (discrete state). The optimal control gains will hedge against the possibility of a transition to a form of expensive operation by over controlling when the system structure is in favorable configurations. Several examples which illustrate this conclusion are included in [Bir 78] and [Chi 82]. These examples illustrate how instabilities can occur even though for each individual, discrete state ρ, the resulting system is completely controllable.

Our major contribution to the study of these problems consists of examining the infinite horizon control problem, thus searching for time invariant control strategies which stabilize the system, and result in a finite cost. The close relationship between stability and the existence of steady state control laws is studied in detail. Necessary and sufficient conditions for the existence of steady-state control laws which result in finite cost are given in [Chi 82], where their relationships to mean-square stability is explicitly displayed. We summarize these results in the next proposition.

Before stating this result, we recall the following terminology pertaining to finite-state Markov chains:

- A state is transient if a return to it is not guaranteed.
- A state i is recurrent if an eventual return to i is guaranteed. If the state set is finite, the mean time until return is finite.
- state i is accessible from state j if it is possible to begin in j and arrive in i in some finite number of steps.
- states i and j are said to communicate if each is accessible from the other.
- A communicating class is closed if there are no possible transitions from inside the class to any state outside of it.
- A closed communicating class containing only one member, j, is an absorbing state. That is, $P_{jj}=1$.
- A Markov chain state set can be divided into disjoint sets $\underline{T}$, $C_1,\ldots,C_s$, where all of the states in $\underline{T}$ are transient, and each C_j is a closed communicating class (of recurrent states).[1]

Define the cover c_j of a form $j \in \underline{M}$ to be the set of all forms accessible from j in one time step. That is,

$$c_j = \{1 \le i \le M:\; P_{ji} \neq 0\}$$

For the purposes of notational simplicity, denote form auguments as subscripts, such as $A(i) = A_i$.

Proposition 4.2.2 Consider the time-invariant Markovian JLQ problem (4.2.1) - (4.2.3). Suppose that there exist feedback control laws

$$u(t) = -\, F_i\, x(t) \quad \text{for each} \quad 1 \le i \le M.$$

such that the following conditions hold:

(1) For each absorbing form i (ie: $P_{ii}=1$) the (deterministic) cost-to-go from (x, i) at time t remains finite (for any finite x) as $(N-t)\to\infty$. This is true if and only if

$$\sum_{t=0}^{\infty} (A_i - B_i F_i)'^{t} (Q_i + F_i' R_i F_i)(A_i - B_i F_i)^{t} < \infty \quad (4.2.10)$$

(each element finite).

(2) For each closed communicating class C_j (having two or more members) the expected cost-to-go from (x, $j \in C_j$) at time t remains finite (for any finite x and each $i \in C_j$) as $(N-t) \to \infty$. This will be true if and only if for each such class C_j there exists a set of finite positive-definite nxn matrices $\{Z_1,\ldots,Z_{|C_j|}\}$ satisyfing

$$Z_i = (1-p_{ii}) \sum_{t=1}^{\infty} p_{ii}^{t-1} (A_i-B_iF_i)'^t \left[\begin{array}{l} Q_i+F_i'R_iF_i \\ \quad + \\ \sum_{\substack{\ell\in C_j \\ \ell\neq i}} \frac{P_{i\ell}}{1-P_{ii}} Z_\ell \end{array} \right] (A_i-B_iF_i)^t$$

for all $i \in C_j$. (4.2.11)

(3) For each transient form $i \in T \quad M$, the expected cost-to-go unitl the form process leaves T (that is, until a closed communicating class is entered) is finite. This is true if and only if there exist finite positive-definite nxn matrices $\{G_1,\ldots,G_T\}$ satisfying:

$$G_i = (1-P_{ii}) \sum_{t=1}^{\infty} p_{ii}^{t-1} (A_i-B_iF_i)'^t \left[\begin{array}{l} Q_i+F_i'R_iF_i \\ \quad + \\ \sum_{\substack{\ell\in T \\ \ell\neq i}} \frac{P_{i\ell}}{1-P_{ii}} G_\ell \end{array} \right] (A_i-B_iF_i)^t$$

for all $i \in T$. (4.2.12)

Conditions (1) to (3) are necessary and sufficient for the existence of steady-state solutions to the equations of Proposition 4.2.1, yielding an optimal steady-state gain which results in a finite cost. Furthermore, suppose that,

(4) For at least one form i in each closed communicating subset of $1 \leq j \leq M$, if L_i is is the steady-state gain,

$$\text{null}\{Q_i\} \cap \text{null}\{L_i\} = \{0\} .$$

Then,

$$\lim_{t\to\infty} E\{x'(t)\, x(t)\} = 0 .$$

The conditions (2) - (3) take into account

- the probability of being in forms that have unstable closed loop dynamics
- the relative expansion and contraction effects of unstable and stable form dynamics, and how the eigenvectors of accessible forms are "aligned." That is, it is not necessary or sufficient for all forms to be stable, since the interaction of different expected form dynamics determines the behavior of $E\{x'(t)\, x(t)\}$.

Proposition 4.2.3 A sufficient condition for the existence of a steady-state optimal control law yielding a finite cost is that, for each form i, there exists a feedback gain F_i such that

1. For each absorbing form i

$$\sum_{t=o}^{\infty} ||(A_i - B_iF_i)^t||^2 < \infty$$

2. For each recurrent, nonabsorbing form i

$$(1-P_{ii}) \sum_{t=1}^{\infty} P_{ii}^{t-1} ||(A_i-B_iF_i)^t||^2 < c < 1$$

3. For each transient form i that is accessible from a form j in its cover c_i,

$$(1-P_{ii}) \sum_{t=1}^{\infty} P_{ii}^{t-1} ||(A_i-B_iF_i)^t||^2 < c < 1$$

4. For each transient form $i \in T$ that is not accessible from any form j in its cover c_i,

$$(1-P_{ii}) \sum_{t=1}^{\infty} P_{ii}^{t-1} ||(A_i-B_iF_i)^t||^2 < \infty$$

The proofs of these propositions are contained in [Chi 82], as well as some examples which illustrate the possible cases when some assumptions are relaxed. The work in [Chi 80a] and [Chi 82] also contains several extensions of the problem described in this section, incorporating additional terms into the cost function, or considering noisy observations of the state. The qualitative results remain the same: The optimal control gains hedge against the possiblity of being transferred into a discrete state (form) where the cost of control is very expensive. We have labelled this effect as passive hedging, to indicate that there is no direct influence by the control action on the transitions of the form process.

The second part of the research studied the control of fully interconnected hybrid systems, in contrast with the hierarchical hybrid structure of figure 4.2.1. Figure 4.2.2 describes the structure of the hybrid system under consideration. The key difference between the structures of figures 4.2.1 and 4.2.2 is that the current continuous state affects the evolution of the discrete state of the system in figure 4.2.2. This permits the possibility of steering the continuous state of the system into regions where expensive form transitions are unlikely to occur.

In order to understand the issues associated with obtaining an optimal control strategy for hybrid systems with structures akin to figure 4.2.2, we formulated and solved a simple example. The continuous state was described by a scalar equation, as

$$x(t+1) = a(\rho(t))\, x(t) + b(\rho(t))\, u(t) \qquad (4.2.13)$$

The discrete state p(t) could take values on {1,2}, with transition matrix P(x), given by

$$P(x) = \begin{pmatrix} 1-\lambda(x) & \lambda(x) \\ 0 & 1 \end{pmatrix} \qquad (4.2.14)$$

The rate $\lambda(x)$ was assumed to be piecewise constant, for the sake of simplicity, we will assume that there are only two pieces,

$$\begin{aligned} \lambda(x) &= \lambda_1 \qquad x \geq v \\ &= \lambda_2 \qquad x < v \end{aligned} \qquad (4.2.15)$$

The objective of the control action was to minimize

the expected cost:

$$J = E \left\{ \sum_{t=1}^{T=1} (x^2(t)\, q(t,\rho(t)) + u^2(t)\, R(t,\rho(T)) + x^2(T)\, Q(t,\rho(T)) \right\} \qquad (4.2.16)$$

Notice that, although equation 4.2.13 is linear in x, the true dynamics in x are non-linear due to the x-dependence of the ρ evolution (4.2.15). Similarly, the cost (4.2.16) is not quadratic in x.

In [Chi 80b], [Chi 82], we have solved this problem using a dynamic programming approach [Bel 57]. The fact that the form state ρ = 2 is an absorbing state permits us to decompose the problem. When ρ(t) = 2, the remaining control problem is a standard optimal control problem, with a solution available from classical theory. When ρ(t) = 1, the solution is more complex, because the control strategy can control the probability that a form transition occurs. In order to solve this problem, we developed an extension of dynamic programming to hybrid systems, denoted as hybrid dynamic programming, which exploits the special structure of the hybrid system. Essentially, this algorithm divides the standard search procedure associated with dynamic programming into a hierarchical operation: at the lowest level, optimal strategies are computed which result in specific descrete form transitions. These are a finite number of these strategies, due to the finite number of discrete states, and finite pieces in (4.2.15). At the highest level, the costs of these strategies are compared, and the optimal one is selected.

The basic advantage of this algorithms is that it divides the difficult task of the optimal control of a hybrid system into two easier tasks: conventional minimization of the cost for a several continuous-variable systems, and a comparison over a finite set of options. Upon applying the algorithm to the problem (4.2.13) - (4.2.16), we developed the following results:

1. The expected cost-to-go from state x and ρ=1 at time t is piecewise quadratic. The number of pieces increases linearly with T-t. The optimal control law is piecewise linear, over the same regions.
2. We have characterized conditions which guarantee that the expected cost-to-go is monotone for x > 0. These conditions involve the relative magnitudes of the costs associated with each discrete form, and the switching rate λ. Most important, there are conditions which establish that the optimal cost-to-go is <u>not</u> monotone for x > 0. Typical cost-to-go functions are depicted in figure 4.2.3.
3. Recursive Riccati-type equations which describe the pieces of the control law and the cost-to-go function are given in [Chi 80b], [Chi 82]. Upper and lower bounds for the cost-to-go functions are derived there.
4. The optimal control law exhibits <u>active hedging</u>: There is a preferred side of v to be in, and the optimal control strategy has regions where the state is "over-controlled" to ensure that the next state is on the preferred side. This results in discontinuous control laws, although the number of discontinuities is equal to the number of pieces in λ(x).

In the remainder of [Chi 82], we studied several ways in which to extend these results to more general hybrid systems. In particular, we studied possible extensions to hybrid systems where the continuous state equations have driving noises, and obtained some analytical characterizations of the structure of the optimal strategies. Another extension was to consider hybrid systems with a multidimensional continuous state. The complexity of the solution algorithm in the scalar case increased manifold when applied to the vector case. The basic reason is that $\mathbb{R}^m$ is not totally ordered; hence, it is very difficult to evaluate the number of different pieces which arise from the discrete comparison part of the hybrid dynamic programming algorithm.

Some approximation methods were also explored, which yield suboptimal strategies. In particular, it was of interest to characterize the existence of a steady-state optimal strategy, and approximate it by an easily-implemented strategy. The reader is referred to [Chi 82] for the details of these results.

Overall, our conclusions were that good control strategies for stochastic hybrid systems arise from a blend of these major factors: The desire to regulate the continuous state process, the desire to influence the transitions of the discrete state process, and the intent of compensating for the probability of undesirable discrete-state transitions in the future. A good control policy is a compromise between these three factors, partitioning the space into regions where one of these factors is the dominant consideration. The precise nature of such a partition is of mathematical interest, but, in terms of real applications, it remains a question of sound engineering judgement.

These results have a lot of implications for the design of operator policies in the control of power systems in an emergency or an in-extremis state. It emphasizes that regulating the continuous variables (power surges, swing angles) alone is too much of a myopic policy. One must also recognize the possible effects of these actions on creating new structural transitions due to overloads, and to hedge against the possibility of additional failures. These results provide the foundation of an analytic methodology which can characterize regions where one of these factors should be the primary concern, hence providing guidelines as to when a system in an emergency state should be islanded. However, it must be pointed out that a problem of such scope is far beyond the reach of existing methodology; we have developed a conceptual methodology, and successfully applied it to the study of some simple examples. There is a lot of additional research which must be completed to establish these theoretical developments as useful analytical design tools.

4.3. <u>Qualitative Analysis of a Switch-Like Adaptive System</u>

As discussed in the Introduction, one of the principal directions for research in this project has been the qualitative analysis of stochastic hybrid systems, where a significant portion of the discrete dynamics represents discrete feedback mechanisms (e.g. protective devices). It is an important fact that many adaptive control systems

which do not explicitly include such discrete feedback mechanisms behave essentially as if they do. For example, the work of Greene in [GRE 78], some initial deterministic analysis of the multiple model adaptive control (MMAC) explored and exposed this possibility for MMAC algorithms. Motivated by this work and by the general hybrid system charter of this DOE project, we undertook an extension of the work of Greene. A thorough treatment of the results of this extension can be found in the S.M. thesis of Mr. H.R. Shomber [SHO 80]. In this section we will provide an overview of the results of this work.

In the basic MMAC formulation (written in discrete-time) the open-loop system is assumed to be linear

$$x(k+1) = Ax(k) + Bu(k) + w(k) \tag{4.3.1}$$

$$y(k) = Cx(k) + v(k) \tag{4.3.2}$$

where w and v are independent white noise processes, with

$$E[w(k)w'(j)] = Q\delta_{kj},\ E[v(k)v'(j)] = R\delta_{kj} \tag{4.3.3}$$

While the open-loop system is assumed to be linear, it is not assumed to be known. Instead, a set of possible models are postulated

$$x_i(k+1) = A_i x_i(k) + B_i u(k) + w_i(k) \tag{4.3.4}$$

$$y(k) = C_i x_i(k) + v_i(k) \tag{4.3.5}$$

$$w_i \sim Q_i \quad , \quad v_i \sim R_i \tag{4.3.6}$$

$i = 1,\ldots,N$.

If one designs Kalman filters for each of these models, one can use the filter innovations processes to compute the conditional probability $p_i(k)$ for the validity of the ith model given all of the data up to time k and assuming that one of the N models is correct. If we use steady-state filters, we obtain

$$\hat{x}_i(k+1) = A_i\hat{x}_i(k) + B_i u(k) + H_i r_i(k+1) \tag{4.3.7}$$

$$r_i(k+1) = y(k+1) - C_i[A_i\hat{x}_i(k) + B_i u(k)] \tag{4.3.8}$$

where H_i is the Kalman gain for the ith model

$$H_i = P_i C_i' R_i^{-1} \tag{4.3.9}$$

and P_i is the solution of the Riccati equation

$$P_i = [C_i' R_i^{-1} C_i + (A_i P_i A_i' + Q_i)^{-1}]^{-1} \tag{4.3.10}$$

Assuming that the actual system matches the ith model, then $r_i(k)$ is a zero mean, white process with covariance

$$V_i = C_i P_i C_i' + R_i \tag{4.3.11}$$

and the probabilites are obtained from the recursive equation

$$p_i(k+1) = \frac{p_i(k)\ f_i[r_i(k+1)]}{\sum_{j=1}^{N} p_j(k)\ f_j[r_j(k+1)]} \tag{4.3.12}$$

where $f_i(\cdot)$ is the probability density function for r_i assuming that the ith model is correct:

$$f_i(r) = (\sqrt{(2\pi)^m \det\ (V_i)})^{-1} \exp -\{\tfrac{1}{2} r'\ V_i^{-1}\ r\} \tag{4.3.13}$$

Suppose we now assume that with each model we have associated a feedback control law

$$u_i(k) = G_i\hat{x}_i(k) \tag{4.3.14}$$

Then, the MMAC algorithm specified that the actual control as a probabilitically weighted sum of these

$$u_i(k) = \sum_{i=1}^{N} p_i(k)\ G_i\hat{x}_i(k) \tag{4.3.15}$$

The impetus for the work in [GRE 78] and in [SHO 80] was the experience obtained from several applications of MMAC. What was observed was that the probabilities $p_i(k)$ behaved in an essentially switch-like manner -- i.e. they were approximately piecewise constant. Consequently the system behaves very much like a stochastic hybrid system. To begin to gain an understanding of this qualitative property of some MMAC Systems and of its implications for overall system behavior, design choices such as model selection, etc., Greene focused attention on the case when there are two models (N = 2), neither of which may concide with the true system. In addition, Greene focused attention entirely on the deterministic analysis of the MMAC algorithm.

In his work Greene was able to isolate several modes of behavior for the MMAC system. To describe these (and also for later use) it is convenient to rewrite the overall closed-loop dynamics for the special case considered by Greene. In this case $N = 2$, $B = B_1 = B_2 = C = C_1 = C_2 = I$, $R_1 = R_2 = R$, $Q_1 = Q_2 = Q$. Thus, the only differences between the true system and the models are in the matrices A, A_1, A_2. Also when $N = 2$, $p_2(k) = 1 - p_1(k)$. Thus, if we define

$$w(k) = \begin{bmatrix} x(k) \\ r_1(k) \\ r_2(k) \end{bmatrix} \tag{4.3.16}$$

the MMAC system can be written as

$$w(k+1) = A(p_1(k))\ w(k) \tag{4.3.17}$$

$$p_o(k+1) = \frac{p_1(k)\ f_1[r_1(k+1)]}{p_1(k)\ f_1[r_1(k+1)] + [1-p_1(k)]\ f_2[r_2(k+1)]} \tag{4.3.18}$$

where

$$\tilde{A}(p_1) = \begin{bmatrix} A-p_1G_1-(1-P_1)G_2 & p_1G_1(I-H_1) & (1-p_1)G_2(I-H_2) \\ A-A_1 & A_1(I-H_1) & 0 \\ A-A_2 & 0 & A_2(I-H_2) \end{bmatrix} \quad (4.3.19)$$

A sepcific choice was made for the A's: for many of the numerical exaamples:

$$A = \begin{bmatrix} a & o \\ o & a \end{bmatrix}, \quad A_1 = \begin{bmatrix} a & o \\ o & a_1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} a_1 & o \\ o & a \end{bmatrix} \quad (4.3.20)$$

which, by symmetry, leads to

$$H_1 = \begin{bmatrix} h & o \\ o & h_1 \end{bmatrix}, \quad H_2 = \begin{bmatrix} h_1 & o \\ o & h \end{bmatrix}, \quad (4.3.21)$$

$$G_1 = \begin{bmatrix} g & o \\ o & g_1 \end{bmatrix}, \quad G_2 = \begin{bmatrix} g_1 & o \\ o & g \end{bmatrix} \quad (4.3.22)$$

This case allows one to expose a wide variety of behaviors by simply adjusting one or two parameters.

Greene's analysis led to modes of behavior specified in terms of the stability properties of the matrix $\tilde{A}(p_1)$. Specifically:

(1) If $\tilde{A}(P_1)$ is a stable matrix for all values of p_1, then all components of the system state decay exponentially. In this case the probability $p_1(k)$ may behave in a switch-like manner in the initial transient period if the initial states are large enough (and r_1 and r_2 are of different sizes), but the probability will eventually settle to some value depending completely on the initial condition.

(2) If $\tilde{A}(p_1)$ is an unstable matrix for all values p_1, then the probability will behave in a switch-like fashion for all time. Even in this case, it is possible to have an overall response that is stable. For the example specified by (4.3.20) - (4.3.22). In this case the model 1 controlles stabilizes x_2 but not x_1. Thus one can imagine a stable overall response in which p_1 switches between 0 and 1, alternately stabilizing x_1 and x_2. Greene defined the notion of hyperbolic stability, which is motivated by the fact that at any time one state is decaying exponentially and the other is exponentially diverging. Consequently $\ln(x_1(t)x_2(t))$ is essentially a linear function of t: if it is increasing, then the system is stable; if it is zero, then the system essentially limit cycles; if it is decreasing, then states decay more during periods in which they are stabilized than they diverge during destabilizing intervals. In these cases the trajectories of components of the state are rather peculiar, as states alternately diverge and decay. For this case Greene performed some detailed approximate analysis of stability and also determined approximate analysis of stability and also determined approximate expressions for the switch times for p_1. As one would expect for a hybrid system, the intervals between switchings are x-dependent, as predicted by Greene's analysis and simulations.

(3) If $\tilde{A}(p_1)$ is stable for some range of p_1 but not for all p_1, then one can obtain trajectories that display either of the types of behavior -- hyperbolic oscillations or exponential decay -- depending upon the size of the initial condition. Eventually, however, the system would settle down and decay exponentially with p_1 in the range for which $\tilde{A}(p_1)$ is stable.

(4) Greene also analyzed the behavior of a limited memory version of the MMAC algorithm, where the $p_i(k)$ are based only on a window of the most recent measurements. In this case one observes an increase in the frequency of oscillations in p_1 and x and a decrease in the peaks of x. In addition, if $\tilde{A}(1/2)$ is unstable, limit cycles are guaranteed to exist.

In our more recent work, as reported in [SHO 80], we have expanded upon Greene's deterministic analysis. Specifically:

(1) Greene's approximations for the switch times and for state magnitudes at these times were modified to obtain significant improvements in accuracy when compared to simulations. While the approximation is not exact, it does provide a reasonably accurate measure of the rate of decay or divergence of $x_1(k)$. $x_2(k)$.

(2) Analogous approximations for switch times and corresponding state sizes were derived for limited memory MMAC and also for MMAC with nonzero set points. Again these approximations provide reasonably accurate predictions of behavior with one exception which is easily understood. Specifically, the switch-time approximations are based on the assumption that $p_1(k)$ is either o or 1 at any time. However, in the case when the state is stabilized and limited memory MMAC is used, $p_1(k)$ tends toward 1/2. Thus, intermediate values of $p_1(k)$ will appear.

The major component of the work developed in [SHO 80] is aimed at analyzing the effects of noise on MMAC behavior. The basic approach used is to perform covariance analysis of (4.3.17) - (4.3.19) (with noise included in (4.3.17)) using random input describing functions(RIDF's). As the resulting quasi-linear equations are quite involved, we will not repeat them here and refer the reader to [SHO 80] for the revelant equations. Using these equations, we have investigated the stochastic (specifically mean-square) stability of MMAC systems and have compared our results and predictions to Monte Carlo simulations. Again the nature of the results is best explained in terms of $\tilde{A}(p_1)$:

(1) Consider the case in which $\tilde{A}(p_1)$ is stable for all values of p_1. In this case the RIDF analysis predicts mean square stability in that the second moments of the states remain bounded. This was confirmed by Monte Carlo simulations which also indicated

that the RIDF approximation was exceptionally accurate in predicting the transient behavior of the state variances as well as their steady-state values.

(2) Consider the case in which $\tilde{A}(p_1)$ is unstable for all values of p_1. As indicated earlier, Greene's deterministic analysis indicated three modes of behavior: hyberbolic asymptotic stability, neutral stability, and instability. The RIDF analysis and Monte Carlo simulations indicate that the state variances grow exponentially in all of these cases, suggesting the singular nature of deterministic hyperbolic asymptotic stability. A proof of mean square stability has not been obtained, however.

(3) If $\tilde{A}(p_1)$ is stable for p_1 in an interval of the form $[\varepsilon_1, \varepsilon_2]$ with $\varepsilon_1 > o$, $\varepsilon_2 < 1$, the state variances remain stable for a while and then grow significantly, indicating instability. This is again verified by simulations, although in this case the RIDF is in error by a substantially greater margin.

(4) If $\tilde{A}(p_1)$ is stable for p_1 in an interval of the form $[o, \varepsilon_1]$ or $[\varepsilon_2, 1]$ or both, the RIDF analysis indicates instability but Monte Carlo simulations did not support this prediction.

Based on these results some subsequent analysis, conjectures, and directions for further work were developed. In particular:

(a) Write

$$p_1(k) = \frac{1}{1 - \frac{1-p_1(o)}{p_1(o)} \frac{\det V_1}{\det V_2} \exp -\frac{1}{2}\alpha(k)} \qquad (4.3.23)$$

where

$$\alpha(k+1) = \alpha(h) + r_1(k+1)'V_1^{-1}r_1(k+1) - r_2(k+1)'V_2^{-1}r_2(k+1) \qquad (4.3.24)$$

In all of the RIDF analysis performed, the variance of $\alpha(k)$ remained bounded. On the other hand, for the case specified by (4.3.20) - (4.3.22) we have performed auxiliary analysis which indicates that the variance of $\alpha(k)$ diverges. Given this, it is not difficult to check that the distribution for $p_1(k)$ approaches one concentrated completely at o and 1. This provides an understanding of why stochastic instability results in Case (3) described earlier and why there is no instability in Case (4) although the RIDF predicts instability based on an erroneous calculation of the spread of $\alpha(k)$ and thus of $p_1(k)$. A modification to the RIDF which corrects the erroneous prediction for the variance of α is needed.

(b) This analysis is incomplete in another sense in that what is needed is a probabilistic description of the temporal behavior of $\alpha(k)$ and thus of $p_1(k)$. Specifically, we would like to prove in Case (2) that systems which are deterministically hyperbolically stable are stochastically unstable. To do this one must determine that noise (even small amounts) disrupts the delicate switching mechanism by which states are alternately stabilized and destabilized.

(c) It is conjectured that it may be possible for, both the RIDF indication of instability and the apparent stability from simulations of Case (4) to be correct. That is, it may be that the system is stable with probability one but not in mean square. Again the nature of switches in the probabilities in this stochastic setting must be investigated. Intuitively, if $[o, \varepsilon_1]$ is the range of p_1 for which $\tilde{A}(p_1)$is stable, then p_1 will spend most of its time in this range with rare occasions on which p_1 leaves this region due to stochastic effects. For any sample path p_1 will return to the stable range, but there may be the rare possibility of arbitrarily long excursions from this range, leading to moment instability.

In conclusion, the analysis we have performed on MMAC systems has provided us with signficant amount of insight into and some useful tools for analyzing stochastic hybrid systems. We have had significant success in performing approximation analysis for the prediction of deterministic and stochastic behavior and have determined or conjectured the essential causes for the few signficant discrepancies between our predictions and the results of simulations.

These determiniations and conjectures suggest several promising directions for further work.

4.4 Convergence Issues in Stochastic Adaptive Control

4.4.1 Introduction

The development of a systematic design methodology for the synthesis of practical self-adjusting control systems which can maintain first stability and second performance improvement, in the presence of rapid and large variations in the open-loop dynamics, represents a very important generic goal in control system engineering, in view of its wide applicability to industrial and defense applications. The so-called "adaptive control problem" has received attention by theoreticians and practitioners alike for the past twenty five years. About a dozen books and hundreds of articles have been devoted to the subject, different philosophies have been developed (model reference adaptive control, self-tuning regulators, dual-control methods, mulitple model adaptive control etc.) and a variety of (mostly academic) examples have been simulated.

In spite of the intense research activity there seems to be a significant gap between the availabe methodologies and the potential applications. To put it bluntly, none of the available adaptive control algorithms can be routinely implemented on a real system and guarantee even the stability of the closed loop process, especially if the physical process is characterized by oscillatory or unstable dynamics and/or unmodeled high frequency dynamics, and/or significant stochastic disturbances and noisy sensor measurements.

One should not blame the theory for this state of affairs. Elegant and useful theoretical advances have been made in the last decade, and especially in the past three years, that have unified diverse approaches. The difficulty appears to be that some of the hypotheses needed to rigorously prove the theoretical results are too restrictive from a practical point of view. Hence, new advances in the theory are necessary, by making different assumptions which better reflect the desired properties of physical control systems.

Results merging deterministic stability approaches (e.g. model reference adaptive control), and stochastic optimization approaches (self-tuning regulators, dual control), together with structural assumptions upon the nature of the adaptive compensator which will hopefully eliminate some of the undesirable (from the applications point of view) characteristics of currently available adaptive control algorithms are presently needed.

Moreover, despite recent rigorous theoretical developments in adaptive control, [NAR 80a], [CRU 79], [GOO 80c], the status of stochastic adaptive systems has been little advanced. Both in identification and primarily in control, the presence of even only observation noise complicates immensely the stability and convergence analysis of such schemes. There is still no general global convergence theory available to date for stochastic adaptive schemes and the results obtained so far are only valid locally, and/or asymptotically-without concern for the transient behavior (stability) characteristics of the adaptive process. For the most part, the presence of noise has been dealt with in a rather ad hoc manner with "common sense" modifications to already existing deterministic adaptive control algorithms. Prefiltering (of the output error) and use of stochastic approximation techniques have been the dominant approaches in this direction. The former, introduces extra delays in the adaptive process with obviously adverse effects on the speed of response and the convergence characteristics of the overall system. The rationale behind the latter was that, when adaptation is completed, the effect of noise should be removed from the adaptation mechanism. However, it is not clear that the parameters will actually converge to the "desired" values before adaptation is stopped.

The most important issue of our current work that must be understood and appreciated relates to the possibility of obtaining global stability results. Every adaptive control algorithm involves a dynamic compensator whose parameters are adjusted in real time based upon output measurements. If the measurements are stochastic processes (due to stochastic disturbance and/or noisy sensor measurements), it follows that the parameters of the adaptive dynamic compensator will be stochastic processes. Thus, to answer stability results in a global sense, one should ideally be able to analyze the global stability properties of differential equations involving multiplicative noise. No general mathematical theory is currently available for such stochastic differential equations. A more promising alternative is to exploit frequency separation properties and band-limited signals as well as generalizing the "passivity inequality" to encompass stochastic quantities also. Although at present no final (or rigorous) results have been obtained along the above lines, cons derable progress has been made in understanding the fundamental concepts underlying adaptive control theory, that unify both deterministic as well as stochastic algorithms, and in formulating clearly the problems that are to be addressed. The long range objective of this research is to develop a methodology of design for adaptive control systems, by attempting to unify promising concepts based upon hyperstability (passivity) theory and stochastic optimal control, respectively, with some common sense engineering techniques related to multivariable loop shaping ideas in the frequency domain which include issues of good command following, integral control, disturbance rejection, bandwidth control, and high frequency roll-off characteristics. A more basic understanding of robustness properties of multivariable systems that has recently been developed through the use of singular value diagrams, serves as a useful tool here.

The emphasis in this
senting a finished body of rigorous theory, is in developing the basic motivation for the avenues of research pursued currently. We attempt this by first presenting succinctly, and without proofs existing theoretical results concerning the (local) stochastic stability of adaptive systems in such a way as to render transparent the points of tangency and intersections among them, from which our current research direction arises as a natural extension. For more details and proofs the reader is referred to the literature cited. Our current results and a novel method of analysis of adaptive algorithms follow next along with our conclusions.

4.4.2 A Brief Review of Existing Convergence Results for Recursive Stochastic Adaptive Algorithms

Two are the main methods used to prove (local) convergence in recursive stochastic algorithms: the associated Ordinary Differential Equation (ODE) approach developed by L. Ljung and the Martingale approach, first developed by [GOO 79a] and [GOO 79b] and later followed by [CAI 80].

The self-tuning regulator (STURE) of [AST 73] was the first recursive stochastic system whose convergence had to be analyzed. Self-tuning regulators were in general designed from an optimization point of view, the objective being to minimize output variance, without any explicit stability considerations at the outset. By using the (ODE method, [JU 77] showed that the stochastic convergence analysis of the STURE could be reduced, under certain associated differential equation, in a deterministic framework thus bringing the stability issue also into this class of adaptive controllers.

Later, the close relationship between the self-tuning regulator and the deterministic model reference schemes was understood, [NAR 79b], [EGA 79]. Partly as a consequence of this, another approach to the convergence analysis emerged, for the MRAS with observation noise, which has become known as the martingale approach. According to this, near-supermartingales are constructed corresponding to stochastic Lyapunov functions, in terms of the state variables of the overall closed loop system (parameter and state errors), which are now random variables. Then again, under certain rather restrictive conditions, the martingale convergence theorem is employed

to prove convergence of the recursive scheme.

Although at first glance the above two approaches appear to be widely different, they both make use of the positive realness of a transfer function that describes the (output) error equation corresponding to a particular recursive scheme. This realization is remarkable, since it has its exact counterpart in the deterministic case, where it was recently shown [VAL 80], that positive realness is a unifying underlying factor in all deterministic schemes to date, either implicitly or explicitly.

Moreover, even in the deterministic case, aymptotic stability of the nonlinear time-varying differential or difference equations, that describe the overall adaptive system, has been proven only for those schemes where the parameter adjustment law is a vector of square integrable functions. This condition is also present in some form, either explicitly or implicitly both in the ODE as well as in the martingale approaches for the convergence analysis of stochastic adaptive schemes.

(i) The ODE Approach

According to this approach, an ordinary differential euqation is associated to the two sets of equations that represent the adaptive algorithm. These are typically of the following form:

$$x(t) = x(t-1) + \gamma(t)Q(t;x(t-1),y(t)) \quad (4.4.1a)$$

$$\phi(t) = A(x(t-1)\phi(t-1) + B(x(t-1))e(t) \quad (4.4.1b)$$

where the set of equations (4.4.1a) represents the parameter adjustment laws and (4.4.1b) the observations (auxiliary state variable generation). $\gamma(t)$ is a decreasing adaptation gain, $Q(.,.)$ represents the correction to the parameter estimate of the previous time instant and $e(t)$ is the stochastic disturbance. The results obtained are "probability one" results and link the convergence properties of the associated ODE with those of the recursive algorithm.

a. Associating an ODE with (4.4.1a) & (4.4.1b)

The differential equation corresponding to the recursive algorithm in (4.4.1a) and (4.4.1b) is obtained as an asymptotic approximation after the system has more or less reached steady state and under the assumption that (4.4.1b) describes an aymptotically stable equation. Using standard arguments to approximate the solutions to (4.4.1a) and (4.4.1b) under such assumptions, one can evaluate

$$x(t+s) = x(t) + \sum_{k=t+1}^{t+s} \gamma(k)Q(x(k-1),\phi(k))$$

$$\approx x(t) + f(x(t)) \sum_{k=t+1}^{t+s} \gamma(k) + \sum_{k=t+1}^{t+s} \gamma(k)w(k)$$

$$\tilde{=} x(t) = f(x(t)) \sum_{t+1}^{t+s} \gamma(k) \quad (4.4.2)$$

where: $Q(x(k-1),\phi(k) \approx Q(x(t), \bar{\phi}(k;x(t))) = f(x(t)) + w(k)$ for $t \geq k$,

$\bar{\phi}(k;x)$ is a steady state approximation of $\bar{\phi}(k)$,

$f(x) = EQ(x,\bar{\phi}(k;x))$

and hence $w(k)$ is a random variable with zero mean.

Equation (4.4.2) then suggests that the sequence of estimates more or less follows the difference equation

$$x^{\Delta}(\tau+\Delta\tau) = x^{\Delta}(\tau) + \Delta_{\tau} f(x^{\Delta}(\tau))$$

$$\text{where } \Delta\tau \to \sum_{t+1}^{t+s} \gamma(k).$$

and hence, for $\Delta\tau$ small enough,

$$\frac{d}{d\tau} x^{\Delta}(\tau) = f(x^{\Delta}(\tau)) \quad (4.4.3)$$

where the fictitions time τ relates to the real time t by

$$\tau_t = \sum_{k=1}^{t} \gamma(k).$$

For more details of the above derivation we refer the reader to [LJU 77].

b. Assumptions on the Algorithm and Related Theorems [LJU 77]

The approximation arrived at in α. and the subsequent analysis of the convergence properties of the recursive algorithm given by the generic form in equations (4.4.1a) and (4.4.1b) depends on certain regularity conditions on the functions Q, A and B and on the driving "noise" term e. Several sets of assumptions are possible, derived from one another by differing restrictions and tradeoffs on Q, A, B and e. We list here only one such set, since it will be enough to see from it the limitations under which the ODE methods is valid.

Let us first define

$$D_s = \{x | A(x) \text{ has all eigenvalues inside the unit circle}\}.$$

Then (4.4.1b) defines an asymptotically stable system for all $x \in D_s$ and hence we can write

$$|A(x)^k| < \lambda(x)^k, \qquad \lambda(x) < 1,$$

Then take $\bar{x} \in D_s$ and define the random variable $\bar{\phi}(t,\bar{x})$ s.t. $\bar{\phi}(t,\bar{x}) = A(\bar{x})\bar{\phi}(t-1,\bar{x}) + B(\bar{x})e(t)$ $\bar{\phi}(0,\bar{x}) = 0$. Let D_R be an open, connected subset of D_s. The regularity conditions will be assumed to be valid in D_R. We tehn impose the following set of assumptions:

1. $e(\cdot)$ is a sequence of independent random variables (not necessarily stationary or with zero mean).
2. $|e(t)| < C$ with probability one for all t.
3. The function $Q(t,x,\phi)$ is continuously differentiable w.r.t. x and ϕ for $x \in D_R$ and the derivatives are bounded in t, for fixed x and ϕ.
4. The matrix functions $A(\cdot)$ and $B(\cdot)$ are Lipschitz continuous in D_R.
5. $\lim_{t\to\infty} E\, Q(t,\bar{x},\bar{\phi}(t,\bar{x}))$ exists for $\bar{x} \in D_R$ and is denoted by $f(\bar{x})$. The expectation is over $e(\cdot)$.
6. $\sum_1^{\infty} \gamma(t) = \infty$
7. $\sum^{\infty} \gamma(t)^p < \infty$ for some p.
8. $\gamma(\cdot)$ is a decreasing sequence

9. $\limsup_{t\to\infty} \frac{1}{\gamma(t)} - \frac{1}{\gamma(t-1)} < \infty$.

Assumptions 6,7,8,9 come directly from the analysis of stochastic approximation convergence techniques and are rather needed for technical points in the proof; thye are also the easiest to be satisfied. Although assumption 2 includes the common Gaussian models of noise, it may not be unreasonable for practical purposes. The regularity conditions represented in 3 and 4 are reasonable, once one is willing to accept that 1-9 are valid only in $D_R \subset D_s$, i.e. in a stability domain in parameter space. This is clearly very restrictive and unrealistic, since it is precisely the boundedness of the resulting closed loop system that is under question, even in the relatively simpler deterministic case. While it may be possible to assure that $x \in D_s$ in an open-loop (identification) scheme, by using appropriate projections on D_s, this never turns out to be possible in a closed-loop (control) situation. Besides D_s is never known in an adaptive control problem. Lastly, Assumption 5 is clearly the one that allows the association of an ODE with the recursive stochastic algorithm, once the rest of the assumptions hold.

The whole ODE approach then heavily hinges on three theorems. We state them below for the sake of completeness. For proofs and more details we again refer the reader to (LJU 77).

Theorem 1: Consider the algorithm (4.4.1a) and (4.4.1b) subject to the assumptions above. Let $\bar{D}$ be a compact subset of D_R such that the trajectories of (4.4.3) that start in $\bar{D}$ remain in a closed subset $\bar{D}_R$ of D_R for $\tau > 0$. Assume that

1. there is a random variable such that (4) $x(t) \in \bar{D}$ and $|\phi(t)| < C$ infinitely often w.p.1.
2. the differential equation (4.4.3) has an invariant set D_c with domain of attraction $D_A \supset \bar{D}$.

Then $x(t) \to D_c$ w.p.1. as $t \to \infty$.

Theorem 2: For the same algorithm and given assumptions, suppose $x^* \in D_R$ has the property

$$P(x(t) \to \bar{B}(x^*,\rho)) > 0 \qquad \forall\ \rho > 0$$

where $\bar{B}(x(,\rho)$ denotes a ρ-neighborhood of x^*.

Also, suppose that

$Q(t,x^*, \bar{\phi}(t,x^*))$ has a covariance matrix bounded from below by a strictly positive definite matrix and that $E\, Q(t,x,\bar{\phi}(t,x))$ is continuously differentiable w.r.t. x in a neighborhood of x^* and the derivatives converge uniformly in this neighborhood as $t > \infty$.
Then

$$f(x^*) = 0 \quad \text{and}$$

$$H(x^*) = \frac{d}{dx} f(x)\Big|_{x=x^*} \quad \text{has all eigenvalues in the LHP.}$$

Theorem 3: For eqns. (4.4.1a) and (4.4.1b) assume 1-9, and also that $f(x)$ is continuously differentiable and that (4) holds. Assume that the solutions to (4.4.3) with initial conditions in $\bar{D}$ are exponentially stable and let I be a set of integers such that

$$\inf |\tau_i - \tau_j| = \delta_o > 0.$$

where $i \neq j$ and $i, j \in I$.

The for any $p \geq 1$ there exist constants K, ε_o and T_o depend on p, $\bar{D}$, δ_o such that for

$$\varepsilon \leq \varepsilon_o \quad \text{and} \quad t_o > T_o$$

$$P\left\{\sup_{\substack{t\in I \\ t\geq t_o}} |x(t) - x^D(\tau_t,\tau_{t_o},x(t_o))| > \varepsilon\right\} \leq \frac{K}{\varepsilon 4p} \sum_{j=t_o}^{N} \gamma(j)^p.$$

where N = sup i; $i \in I$ which may be ∞.
Theorems 1-3 above can be expressed in a somewhat more intuitive language as follows:

1. x(t) can converge only to stable stationary points of the ODE.
2. If x(•) belonds to the domain of attraction of a stable stationary point x^* of the ODE infinitely often, then x(t) converges w.p.1. to x^* as $t \to \infty$.
3. The trajectories of the ODE are "the asymptotic paths" of the estimates x(•).

C. Examples

1. Åström and Wittenmark's Self-Tuning Regulator [AST 73]. The data generating process is described by an ARMA model as follows:

$$A(q^{-1})\, y(t) = B(q^{-1})\, u(t) + C(q^{-1})e(t) \qquad (4.4.4)$$

where q^{-1} is the delay operator and

$$A(q^{-1}) = 1 + a_1 q^{-1} + \ldots + a_n q^{-n}$$

$$B(q^{-1}) = b_o + b_1 q^{-1} + \ldots + b_m q^{-m},\ b_o = 0,\ b_1 \text{ known.}$$

$$C(q^{-1}) = 1 + c_1 q^{-1} + \ldots + c_n q^{-n}$$

$\{e(t)\}$ is a stationary random sequence of independent random variables, such that all its moments exist. Also (4.4.4) is minimum phase.

Then, if the parameters were known, a minimum variance controller would be given by

$$u(t-1) = -\frac{1}{b_1}\, \theta_{MV}^T\, \phi(t)$$

where $\theta_{MV}^T = (a_1 - c_1, \ldots, a_n - c_n, b_2, \ldots, b_m)$

$$\phi^T(t) = (-y(t-1), \ldots, -y(t-n), u(t-2), \ldots, u(t-m)).$$

Since the vector of parameters is not known, it is estimated on-line according to the recursion

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{1}{t}\frac{1}{r(t-1)}\, \phi(t)\, \varepsilon(t) \qquad (4.4.5)$$

$$r(t) = r(t-1) + \frac{1}{t}\,(\phi(t)^T \phi(t) - r(t-1)) \qquad (4.4.6)$$

where $\varepsilon(t) = y(t) - b_1 u(t-1) - \hat{\theta}(t-1)^T \phi(t)$.

Both $y(t)$ and $u(t)$ are influenced by the sequence of estimates $\{\hat{\theta}(t)\}$ through the control law $u(t)$. Define $y(t,\theta)$, $u(t,\theta)$, $\phi(t,\theta)$, $\varepsilon(t,\theta)$ to be the stationary processes which would be obtained with a control law corresponding to some fixed θ.

Then $\bar{u}(t-1,\theta) = \frac{1}{b_1} \theta^T \bar{\phi}(t,\theta)$

$$\bar{\varepsilon}(t,\theta) = \bar{y}(t,\theta) - b_1 \bar{u}(t-1,\theta) -$$

$$\theta^T \bar{\phi}(t,\theta) = \bar{y}(t,\theta).$$

We note here that the process is defined only for those θ in the stability region D_s, where also the stationarity assumption is valid.
Furthermore, we can write

$$\bar{\varepsilon}(t,\theta) = \bar{y}(t,\theta) = \theta_o^T \bar{\phi}(t,\theta) + b_1 \bar{u}(t-1,\theta) + C(q^{-1})$$

$$e(t) =$$

$$= (\theta_o - \theta)^T \bar{\phi}(t,\theta) + C(q^{-1}) e(t) =$$

$$= (\theta_o - \theta_{MV})^T \bar{\phi}(t,\theta) + \bar{\phi}(t,\theta)^T (\theta_{MV} - \theta) +$$

$$C(q^{-1}) e(t)$$

where θ_o is the vector of (original) process parameters.
Hence, since

$$(\theta_o - \theta_{MV})^T \bar{\phi}(t,\theta) = [1 - C(q^{-1})]\, \bar{y}(t,\theta) =$$

$$(1 - C(q^{-1}))\, \bar{\varepsilon}(t,\theta)$$

it follows that

$$C(q^{-1})\, \bar{\varepsilon}(t,\theta) = \bar{\phi}(t,\theta)^T (\theta_{MV} - \theta) + C(q^{-1}) e(t) \qquad (4.4.7)$$

and

$$\bar{\varepsilon}(t,\theta) = \tilde{\phi}(t,\theta)^T (\theta_{MV} - \theta) + e(t) \qquad (4.4.8)$$

where $\tilde{\phi}(t,\theta) = H(q^{-1}) \bar{\phi}(t,\theta)$

$$H(q^{-1}) = \frac{1}{C(q^{-1})}$$

Also, define

$$\tilde{G}(\theta) = E\bar{\phi}(t,\theta)\tilde{\phi}^T(t,\theta) \qquad (4.4.9)$$

$$\theta^* = \theta_{MV}$$

Then, since $e(t)$ is independent of $y(s)$, $u(s)$ for $s < t$, we have from eqns. (4.4.8) and (4.4.9)

$$f(\theta) = E\bar{\phi}(t,\theta)\, \bar{\varepsilon}(t,\theta) = \tilde{G}(\theta)\, (\theta^* - \theta)$$

and we can now associate with the algorithm (4.4.5), (4.4.6) the following differential equation:

$$\frac{d}{d\tau} \theta(\tau) = \frac{1}{r(\tau)} \tilde{G}(\theta(\tau))\, (\theta^* - \theta(\tau))$$

$$\frac{d}{d\tau} r(\tau) = g(\theta(\tau)) - r(\tau) \qquad (4.4.11)$$

which is defined for $\theta(\tau) \in D(s)$, $r(\tau) > 0$ and where

$$g(\theta) = E\bar{\phi}^T(t,\theta)\, \bar{\phi}(t,\theta).$$

Global stability of (4.4.10) follows if

$$\tilde{G}(\theta) + \tilde{G}^T(\theta) \text{ is } > 0.$$

Taking into account eqns. (4.4.8) and (4.4.9) it can be shown quite easily that the above holds <u>iff</u> $H(q^{-1})$ is strictly positive real.
<u>Remark 1</u>. In this approach, strict positive realness of the transfer function associated with the error, however it may be defined in any particular algorithm, is a necessary condition for global stability of the ODE, which then implies local convergence of the recursive algorithm. In deterministic stability analysis, strict positive realness was a sufficient condition for global stability of the adaptive system.
<u>Remark 2</u>. We note here from eqns. (4.4.5) and (4.4.6) that the parameter adjustment low is a square integrable function.

2. General Adaptive Algorithms

In general, the process is given by

$$A(q^{-1})y(t) = q^{-k} B(q^{-1})\, u(t) + w(t)$$

where q^{-k} represents a pure time-delay (relative degree) $m \geq k \geq 1$, A and B as defined before and $w(t)$ represents the disturbance.

A reference model may be specified by

$$A^M(q^{-1})y^M(t) = q^{-k} \frac{B^M(q^{-1})}{A^M(q^{-1})}\, u^M(t) .$$

It can be shown {[EGA 79]} that the output error defined as

$$e(t) = y(t) - y^M(t)$$

satisfies

$$\Lambda A^M e(t) = q^{-K}[\theta^T \phi(t)] + P\, w(t)$$

where Λ, A^M, P, B^M, B are polynomials of compatible degrees and such that the polynomial equation

$$\Lambda A^M = A\, P + q^{-K} Q$$

where the parameter in the polynomials P and Q can be chosen, Λ is any arbitrary Hurwitz polynomial and B^M is absorbed in Λ. θ represents the unknown parameter vector that has to be adjusted and ϕ is the state vector, both defined in an analogous manner as in example a.

When the relative degree (pure time delay) is greater than one, extra filtering has to be introduced through a strictly stable and inversely stable

rational transfer function $L(q^{-1})$. Then the augmented error can be written as

$$e_a(t) - \frac{L}{\Lambda A^M}\,[\theta^T L^{-1}\phi(t-k)] + \frac{P}{\Lambda A^M}\,w(t)$$

A typical recursive adaptive algorithm is then given by

$$\hat{\theta}(t) = \hat{\theta}(t-1) + [L^{-1}\phi(t-k)]\,\frac{\varepsilon(t)}{r(t)} \qquad (4.4.12)$$

$$r(t) = r(t-1) + |L^{-1}\phi(t-k)|^2 \qquad (4.4.13)$$

where $\varepsilon(t) = e_a(t) - e(t)$

$$e(t) = \frac{L}{\Lambda A^M}\,[\hat{\theta}^T(t-1)\,L^{-1}\phi(t-k)]$$

and the control law is chosen to satisfy

$$\hat{\theta}^T[L^{-1}\phi(t)] = 0.$$

For more details of the above derivations, the reader is referred to [EGA 79]. Following the same method of analysis as in example 1, it can be shown again that the algorithm (4.4.12), (4.4.13) converges locally, provided $\frac{L}{\Lambda A^M}$ is SPR.

Remark: For (4.4.12), (4.4.13) we see that the parameter adjustment law (12) is an L^2 function here also.

(ii) The Martingale Approach

The basic proof technique in [GOO 79a] was the use of the Lyapunov $V_t = ||\theta - \hat{\theta}_t||^2$, t=1,2,..., where θ was the nominal (actual) parameter value and $\hat{\theta}_t$ its estimated value. V_t can be called a stochastic Lyapunov function since one attempts to show V_t is a super-martingale. In fact, in [GOO 79a], $V_t + \frac{1}{r_{t-1}} \cdot S_t$ was shown to be a "near super-martingale", i.e., a positive super-martingale less a negative quantity plus a positive quantity, the latter being a.s. summable. The term S_t was introduced in an apparently arbitrary manner in order to deal with a cross-term arising in the expansion of V_t via the equation defining the parameter adjustment law. S_t is positive by virtue of a strict positive real condition required of the transfer function corresponding to the error generating dynamics. [SOL 79] also showed how to exploit this property in parameter estimation. The whole technique is based on the Martingale Convergence Theorem, which we state below for the sake of completeness. Its proof can be found in [NEV 75], [SOL 79].

Martingale Convergence Theorem: Let $\{T_n\}$, $\{\alpha_n\}$, $\{\beta_n\}$ be sequences of non-negative random variables adapted to an increasing sequence of σ-algebras F_n such that

$$E[T_n|F_{n-1}] \leq R_{n-1} - \alpha_{n-1} + \beta_{n-1}\,.$$

If $\Sigma_1^\infty \beta_n < \infty$ a.s., then T_n converge almost surely to a finite random variable T and $\Sigma_1^\infty \alpha < \infty$ a.s. Then, for the general problem stated in example b, with k=1, and for recursive adaptive algorithms in [GOO 79a] which are very similar in nature to the ones already discussed, the following result was obtained using the Martingale Convergence Theorem.

Result: If the noise generating dynamics satisifies the following conditions

$$\frac{1}{C(z)} - \frac{1}{2} \text{ is P.R.} \qquad (4.4.14)$$

$$C(z) - \frac{1}{2} \text{ is SPR}$$

then the algorithms in [GOO 79a] ensure that with probability one:

$$(1)\quad \sup_N \frac{1}{N} \sum_{t=1}^{N} y(t)^2 < \infty \qquad (4.4.15)$$

$$(2)\quad \sup_N \frac{1}{N} \sum_{t=1}^{N} u(t)^2 < \infty \qquad (4.4.16)$$

$$(3)\quad \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} E\{[y(t) - y_M(t)]^2|F_{t-1}\} = \gamma^2 \qquad (4.4.17)$$

where γ^2 is the minimum possible mean square control error achievable with any causal linear feedback. (This includes feedback designed using the true parameters).

Remark 1. Strict positive realness - of the "noise transfer function" - is also a requirement using the martingale approach. However, this requirement on the noise dynamics is very unrealistic, since they are even unknown.

Remark 2. No *a priori* boundedness of the adaptive signals is assumed.

Remark 3. The martingale approach is much less systematic than the ODE method. One of the reason is that convergence proofs require first to find an appropriate super martingale which to some extent can be considered as a stochastic Lyapunov function. The search for the appropriate near supermartingale can be lengthy, the rationale behind its derivation not clear and the conditions on the relevant adaptive signals unrealistic or not verifiable. Clearly, a more systematic approach has to be developed.

1. An alternative Formulation of the Structure of Stochastic Adaptive Algorithms

It is a well established fact to date that most -- if not all -- of the currently existing recursive identification and adaptive estimation and control schemes can be equivalently represented by a feedback system with a linear time-invariant feedforward path and a feedback time-varying path. The forward path represents the generalized (state) error equations of the adaptive system, while the feedback path corresponds to the parameter adjustment mechanism. Hyperstability theory has then been employed to prove global asymptotic stability of the overall feedback system, which can be guaranteed -- under additional conditions ([VAL 80]) -- if the forward block consists of a strictly positive real transfer function and the feedback part represents a passive system. In fact, it can be readily seen that this equivalent feedback configuration appears implicitly in the design approach using Lyapunov functions, since the Lyapunov function candidate always consists of two terms; one quadratic

form of the state errors and another one of the parameter errors.

When sensor disturbances are present, the output of the forward path is contaminated with noise, w(t), while the rest of the structure of the (now) stochastic adaptive system remains the same. However, deterministic stability theory no longer holds, since now the parameters in the feedback path contain random quantities and the requirements for passivity can no longer be assumed to be satisfied. Global stability results for such a stochastic adaptive system have been obtained using martingales by [GOO 79a] and [GOO 79b] under the very restrictive and rather unrealistic assumption of positive realness of the noise generating dynamics. A much more realistic assumption is to assume knowledge of the power spectrum of the noise and attempt to make use of such data to derive the adaptive laws within the existing structure (feedforward and feedback). This will guarantee convergence of the stochastic adaptive system under more realistic and less restrictive assumptions than have been made to date.

We show below a general method of proof, using the Martingale Convergence Theorem, for this structure. Unfortunately, so far, we have had to use a very restrictive assumption, also, but one whose form allows for more general extensions and motivates our search for some sort of a "stochastic storage function" or "energy indexing function" which lies at the center of our research efforts at present. Besides, this helps clarify and state more precisely the problems we are currently addressing.

2. Use of the MCT in the Passivity Framework for Convergence of Adaptive Schemes

[LAN 79a] generalized the classes of systems for which the structure described in 1. above results in global asymptotic stability for deterministic adaptive schemes. We present below their definitions which we shall employ later in the proofs.

Consider the following discrete time, completely controllable (and/or observable) linear time-invariant system

$$\begin{aligned} x(t+1) &= A\,x(t) + B\,u(t) \\ y(t) &= C\,x(t) + D\,u(t) \end{aligned} \qquad S_1$$

where x is the state vector of dimension n, u is the input m-vector, y the output m-vector and A, B, C, D are matrices of compatible dimensions.

Definition 1. [The Class $L(\Lambda)$]: Let Λ be an arbitrary symmetric matrix. The above system is said to belong to the class $L(\Lambda)$ if the resulting system obtianed by its parallel connection with a gain matrix $-\frac{1}{2}\Lambda$ is characterized by a strictly positive real discrete transfer matrix. The resulting system, consequently, is described by

$$\begin{aligned} x(t+1) &= A\,x(t) + B\,u(t) \\ y(t) &= C\,x(t) + (D - \tfrac{1}{2}\Lambda)\,u(t) \end{aligned}$$

and its transfer matrix is given by

$$H(z) = D - \tfrac{1}{2}\Lambda + C(zI-A)^{-1}B.$$

Consider now the discrete linear time-varying system described by

$$\begin{aligned} \tilde{x}(t) &= A(t)\,\tilde{x}(t) + B(t)\,\tilde{u}(t) \\ y(t) &= C(t)\,\tilde{x}(t) + D(t)\,\tilde{u}(t) \end{aligned} \qquad S_2$$

where the vectors $\tilde{x}$, $\tilde{y}$, $\tilde{u}$ and the matrices A,B,C,D are defined as the analogous quantities as in the system of definition 1.

Definition 2. [The Class $N(\Gamma)$]. Let $\Gamma(t)$ be an arbitrary sequency of symmetric matrices. The system S_2 is said to belong to the class $N(\Gamma)$ if the resulting system obtained by its feedback connection with a gain matrix $\frac{1}{2}\Gamma(t)$ satisfies the inequality

$$\sum_{t=t_o}^{t_1} \bar{y}(t)^T \bar{u}(t) \geq -\mu_o^2 \qquad \forall\, t_1 \geq t_o$$

where $\Gamma(t)$ is chosen to satisfy $\Lambda - \Gamma(t) \geq 0$.

Remark: The above inequality is often called the Popov inequality and its interpretation is taken as a passivity condition.

The resulting system is then described as

$$\begin{aligned} \bar{x}(t+1) &= A(t)\,\bar{x}(t) + B(t)\,\bar{u}(t) \\ \bar{y}(t) &= C\,\bar{x}(t) + D(t)\,\bar{u}(t) \\ \bar{u}(t) &= \tilde{u}(t) - \tfrac{1}{2}\Gamma(t)\,\tilde{y}(t) \end{aligned} \qquad S_3$$

and, consequently, the passivity inequality can be expressed as

$$\sum_{t=t_o}^{t_1} \bar{y}(t)^T \bar{u}(t) = \sum_{t=t_o}^{t_1} \tilde{y}(t)\,\tilde{u}(t) + \frac{1}{2}\sum_{t=t_o}^{t_1} \bar{y}^T(t)\,\Gamma(t)\,\tilde{y}(t)$$

Note: For the case of single-input single-output systems, the matrix Λ becomes a scalar denoted by λ and the matrix sequence $\Gamma(t)$ becomes a scalar sequence $\gamma(t)$.

System S_1 in the forward path with S_3 in a feedback connection with it is the most general form of recursive adaptive algorithms of the type described by equations (4.4.12) and (4.4.13). In fact, all the existing recursive adaptive algorithms, with appropriate interpretations of their state, input and output variables, can be cast in such an equivalent feedback connection, that furthermore satisfy conditions for classes $L(\Lambda)$ and $N(\Gamma)$. Our purpose is not to show this here, since it can be found in the existing literature. We just want to mention again at this point that the state variables of S_1 are the generalized error states, while those in S_3 are the corresponding parameter errors. In the present discussion, the input to system S_3 is

$$\bar{u}(t) = y(t) + w(t) \qquad (4.4.18)$$

where y is the output of S_1 and w the measurement noise. The input to system S_1 is

$$u(t) = -\bar{y}(t) \qquad (4.4.19)$$

where $\bar{y}(t)$ is the output of S_3.

We will now prove that the quantity Z(t) to be defined in the sequel in terms of the states of systems S_1 and S_3 is a near-supermartingale, which in turn is crucial in the proofs of convergence, i.e. obtaining results of the type of equns. (4.4.15), (4.4.16), (4.4.17) following from this point on very similar arguments as in [GOO 79b], without, however, needing to involve positive realness of the noise generating dynamics.

$$\text{Define } Z(t) = \frac{V(t)}{t} + \sum_{k=1}^{t-1} \frac{1}{k+1} \frac{V(k)}{k} \qquad (4.4.20a)$$

$$\text{where } V(t) = x^T(t)\, P\, x(t) + \bar{x}^T(t)\, \bar{P}(t)\, \bar{x}(t) + \sum_{k=0}^{t-1} [x^T(k),\, u^T(k)]\, M_o \begin{matrix} x(k) \\ \hline u(k) \end{matrix}, \qquad (4.4.20b)$$

where $\bar{P}(t) \geq 0$, $M_o > 0$.

Proof: If can be shown {[LAN 79]} from the definitions for the classes $L(\Lambda)$ and $N(\Gamma)$, that the following two relations hold true for S_1 and S_3, respectively:

$$\sum_{k=0}^{t} y^T(k)\, u(k) = \frac{1}{2} x^T(t+1)\, P\, x(t+1) - \frac{1}{2} x^T(o)\, P\, x(o) + \frac{1}{2} \sum_{k=0}^{t} [x^T(k),\, u^T(k)] M_o \begin{matrix} x(k) \\ \hline u(k) \end{matrix} + \frac{1}{2} \sum_{k=0}^{t} u^T(k)\, \Lambda\, u(k) \qquad (4.4.21)$$

where $P > 0$; $M_o > 0$

$$\sum_{k=0}^{t} \bar{y}^T(k)\bar{u}(k) = \frac{1}{2} \bar{x}^T(t+1)\bar{P}(t+1)\bar{x}(t+1) - \frac{1}{2} \bar{x}^T(o)\bar{P}(o)\bar{x}(o) + \frac{1}{2} \sum_{k=0}^{t} [\bar{x}^T(k), \bar{u}^T(k)]\bar{M}(k) \begin{matrix} \bar{x}(k) \\ \hline \bar{u}(k) \end{matrix} - \frac{1}{2} \sum_{k=0}^{t} \bar{y}^T(k)\Gamma(k)\bar{y}(k);$$

where $\bar{P}(t) \geq 0$, $\bar{M}(t) \geq 0$.

Also, from eqn. (4.4.18) and (4.4.19) we obtain

$$y^T(t)u(t) = -\bar{y}(t)\bar{u}(t) + \bar{y}^T(t)\, w(t) \quad \rightsquigarrow$$

$$\rightsquigarrow \sum_{k=0}^{t} y^T(k)u(k) = -\sum_{k=0}^{t} \bar{y}^T(k)\bar{u}(k) + \sum_{k=0}^{t} \bar{y}^T(k)w(k) \rightsquigarrow$$

$$\rightsquigarrow \sum_{k=0}^{t} y^T(k)u(k) + \sum_{k=0}^{t} \bar{y}^T(k)\bar{u}(k) = \sum_{k=0}^{t} \bar{y}^T(k)w(k) \qquad (4.4.23)$$

Adding the LHS and RHS respectively of equs [4.4.21] and [4.4.22] and using (4.4.23) we obtain

$$\frac{1}{2} x^T(t+1)P\, x(t+1) - \frac{1}{2}\alpha + \frac{1}{2} \sum_{k=0}^{t} \beta^T(k) M_o \beta(k) + \frac{1}{2} \sum_{k=0}^{t} u^T(k)\, \Lambda u(k) + \frac{1}{2} \bar{x}^T(t+1)\bar{P}(t+1)\bar{x}(t+1) - \frac{1}{2}\gamma + \frac{1}{2} \sum_{k=0}^{t} \delta^T(k)\, \bar{M}(k)\delta(k) - \frac{1}{2} \sum_{k=0}^{t} \bar{y}^T(k)\Gamma(k)\bar{y}(k) = \sum_{k=0}^{t} \bar{y}^T(k)w(k) . \qquad (4.4.24)$$

where

$$\alpha = x^T(o)Px(o),\ \beta^T(k) = [x^T(k),\, u^T(k)]$$

$$\gamma = \bar{x}^T(o)\bar{P}\bar{x}(o),\ \delta^T(k) = [\bar{x}^T(k),\, \bar{u}^T(k)]$$

Using now the definition of V(t) from (4.4.20b) we can rewrite the above as

$$V(t+1) - \alpha + \sum_{k=0}^{t} \beta^T(k)\, M_o\beta(k) + \sum_{k=0}^{t} u^T(k)\, \Lambda u(k) - \gamma + \sum_{k=0}^{t} \delta^T(k)\bar{M}(k)\delta(k) - \sum_{k=0}^{t} \bar{y}^T(k)\, \Gamma(k) = 2 \sum_{k=0}^{t} \bar{y}^T(k)w(k) \qquad (4.4.25)$$

In a completely analogous manner as done for eqn. (4.4.25) we also find

$$V(t) - \alpha + \sum_{k=0}^{t-1} \beta^T(k)\, M_o\beta(k) + \sum_{k=0}^{t-1} u^T(k)\, \Lambda u(k) - \gamma + \sum_{k=0}^{t-1} \delta^T(k)\, \bar{M}(k)\, \delta(k) - \sum_{k=0}^{t-1} \bar{y}^T(k)\, \Gamma(k)\, \bar{y}(k) = 2 \sum_{k=0}^{t-1} \bar{y}^T(k)w(k) \qquad (4.4.26)$$

Comparing (4.4.29) and (4.4.26) we get

$$V(t+1) = V(t) + 2\bar{y}^T(t)w(t) - \beta^T(t)\, M_o\beta(t) - u^T(t)\, \Lambda u(t) - \delta^T(t)\, \bar{M}(t)\, \delta(t) + \bar{y}^T(t)\, \Gamma(t)\bar{y}(t) \qquad (4.4.27)$$

and since $M_o > 0$, $\bar{M}(t) \geq 0$

$$V(t+1) \leq V(t) + 2\bar{y}^T(t)w(t) - u^T(t)\, \Lambda u(t) + \bar{y}(t)\, \Gamma(t)\, \bar{y}(t). \qquad (4.4.28)$$

But from (4.4.19), $u(t) = -\bar{y}(t)$ and since $\Lambda - \Gamma(t) \geq 0$, the inequality above becomes

$$V(t+1) \leq V(t) + 2\bar{y}^T(t)w(t) \quad \rightsquigarrow \qquad (4.4.29)$$

$$\rightsquigarrow \frac{V(t+1)}{t+1} \leq \frac{V(t)}{t+1} + \frac{2}{t+1}\bar{y}(t)w(t) = \frac{V(t)}{t} -$$

$$- \frac{t}{t+1}\frac{V(t)}{t+1} + \frac{2}{t+1}\bar{y}^T(t)w(t)$$

$$\rightsquigarrow \frac{V(t+1)}{t+1} + \sum_{k=1}^{t}\frac{1}{k+1}\frac{V(k)}{k} \leq \frac{V(t)}{t+1} + \frac{2}{t+1}\bar{y}(t)w(t) +$$

$$+ \sum_{k=1}^{t}\frac{1}{k+1}\frac{V(k)}{k} \quad \rightsquigarrow$$

$$\rightsquigarrow Z(t+1) \leq Z(t) + \frac{2}{t+1}\bar{y}^T(t)w(t) \qquad (4.4.30)$$

and

$$E\{Z(t+1)\,|\,F_t\} \leq Z(t) + 2\frac{1}{t+1}E\{\bar{y}^T(t)w(t)\,|\,F_t\} \qquad (4.4.31)$$

In order for Z(t) to satisfy the conditions of the Martingale theorem in [NEV 75], we need to have

$$\frac{1}{t+1}E\{\bar{y}^T(t)w(t)\,|F_t\} = \varepsilon(t+1) \geq 0,$$

where $\{E\,\bar{y}^T(t)w(t)\,|F_t\} = \varepsilon(t+1)$ (4.4.32)

and

$$\sum_{t=1}^{\infty}\varepsilon(t) < \infty.$$

Then,

$$\lim_{t\to\infty} Z(t) = \lim_{t\to\infty}\frac{V(t)}{t} + \sum_{k=1}^{t-1}\frac{1}{k+1}\frac{V(k)}{k} < \infty \quad a.s. \qquad (4.4.33)$$

From this point, the rest of the convergence results are obtained by using the MCT repeatedly to obtain inequalities of the form (4.4.15) - (4.4.17).

The passivity inequality as implied by eq. 4.4.32] clearly is unrealistic to assume and very hard to guarantee in the time-domain. It does, however, suggest that a frequency domain equivalent may be possible, given that, usually, the power spectrum of the noise can be assumed known. In the frequency domain, then, it can be interpreted as the "tolerable energy" that can be injected into the system due to the presence of noise without risk of instability for the overall feedback adaptive loop. One can then define "admissible frequency sectors" within which (4.4.32) is satisfied, given the noise characteristics. This, in turn, points to a way of choosing $\bar{y}$ - and hence the adaptive laws - such that (4.4.32) can be satisfied taking into account the noise characteristics.

An alternative way of looking at the problem is to see what modifications and/or tradeoffs are possible in the conditions required from the forward and feedback blocks, i.e., possible redefinitions of the classes $L(\Lambda)$ and $N(\Gamma)$ in the original feedback adaptive structure, in order to take into account the random effects introduced by the noise. For example, the class $N(\Gamma)$ implies some weak positivity condition for the system S_3. This can be expressed as the following set of conditions, in terms of the system matrices A(t), B(t), C(t), D(t):

There exists a sequence of matrices $\bar{P}(t) \geq I\delta > 0$, $\Lambda(t), \bar{M}(t)$ such that

$$A^T(t-1)\,\bar{P}(t)\,A(t-1) - \bar{P}(t-1) = -\Lambda(t)\,\Lambda(t)^T$$

$$A^T(t-1)\,\bar{P}(t)\,B(t-1) = C^T(t-1) - \Lambda(t)\bar{M}(t)$$

$$\bar{M}^T(t)\,\bar{M}(t) = D(t-1) + D^T(t-1) - B^T(t-1)\bar{P}(t)B(t-1)$$

The above set of conditions can be thought of as the time-varying version of the positive real lemma. Then, by the implicit function theorem applied to the eqns. (4.4.34), the exists a non-empty open set D_r around any parameter corresponding to a positive real system such that a (possibly) random evolution of parameters, due to the presence of noise, is possible within D_r given some reasonable assumptions about noise characteristics. Hence non-constant random uniformly positive real systems exist.

The method of proof in (i) and (ii) above explicitly makes use of positive reality or a passivity condition that has to be satisfied within the overall adaptive loop. Furthermore, in (i) the additional assumption of stationarity of the adaptive signals is necessary for the local convergence analysis; unfortunately, stationarity does not hold in a closed loop framework, where the adaptation mechanism is still ongoing. In (iii) following, we prove that the output and parameter errors of a representative discrete-time algorithm remain bounded in a mean-square sense, when the plant output is corrupted by a measurement noise sequence, assumed to be white, Gaussian, of zero mean and arbitrary constant variance. The proof does not rely on any more restrictive assumptions other than the requirement that the noise samples be independent and uncorrelated at time t with the output(s) of the preceding stages, up to time t-1. This latter part can be ensured by construction of the algorithm per se, as will be seen next. The proof in (iii) is the first available in the existing literature that an adaptive control algorithm with observation noise is mean-square stable not only in the output but in the parameter errors as well, independent of the choice of a reference input. This confirms the often expressed belief that the output noise will in fact provide the "sufficient excitation" necessary for parameter error boundedness at least in the case considered next.

(iii) Mean Square Boundedness of a Discrete-Time Stochastic Adaptive Algorithm [ROH 81]

The proof in this subsection makes use of the ideas of Bitmead and Anderson in [BIT 80 a & b] and

Anderson and Johnson [AND 81] and although it is given here for a first order plant it is extendable to the multivariable case as well. The algorithm analyzed is the discrete-time version of that in [NAE 80a] and is contained in [NAR 80b]. We describe it briefly below.

The plant-model representation is given by two first order difference equations (4.4.34) and (4.4.35).

Actual Plant: $y(t+1) = \alpha y(t) + \beta u(t+1)$ (4.4.34)

Reference Model: $y^*(t+1) = ay(t) + br(t+1)$ (4.4.35)

where $|a| < 1$, $b > 0$, α and β are unknown, y and u are the plant state and control input and y^* and r are the model state and reference input respectively. The plant output $y(t)$ is contaminated by observation noise $n(t)$ and is described by

$$z(t) = y(t) + n(t) \tag{4.4.36}$$

The plant input $u(t)$ is synthesized recursively according to equation (4.4.37)

$$u(t+1) = \theta_1(t+1)\, z(t) + \theta_2(t+1)\, r(t+1)$$

$$- \rho\{\gamma_{11}\, z^2(t) + (\gamma_{12}+\gamma_{21})\, z(t)\, r(t+1)$$

$$+ \gamma_{22}\, r^2(t+1)\}\, e(t+1)$$

$$= \frac{1}{\beta}\,[(a-\alpha+\phi_1(t+1)]\, z(t) + [b+\phi_2(t+1)]r(t+1)$$

$$-\rho\beta[\gamma_{11}\, z^2(t) + (\gamma_{12}+\gamma_{21})\, z(t)\, r(t+1) +$$

$$+ \gamma_{22}\, r^2(t+1)]\, e(t+1)]$$

where θ_1, θ_2 are the adjustable parameters

ϕ_1, ϕ_2 are the associated parameter errors

$\{\gamma_{ij}\}$ constitute a gain matrix $\Gamma = \Gamma^T > 0$

$\frac{1}{2} < \rho \leq 1$, and we choose $\rho = 1$ here for simplicity ($\gamma_{12} = \gamma_{21} = 0$)

The error equation for this system is

$$e(t) = z(t) - y^*(t)$$

$$= \frac{ae(t-1) + \phi_1(t)\, z(t-1) + \phi_2(t)\, r(t)}{1 + \beta d(t)} \tag{4.4.38}$$

where $d(t) = \gamma_1\, z^2(t-1) + \gamma_2\, r^2(t)$

and $\gamma_{11} \triangleq \gamma_1$, $\gamma_{22} \triangleq \gamma_2$.

The overall error system is now described by the following set of equations:

$$\begin{bmatrix} e(t) \\ \frac{\phi_1(t+1)}{\beta} \\ \frac{\phi_2(t+1)}{\beta} \end{bmatrix} = \left(\frac{1}{1+\beta d(t)}\right) \begin{bmatrix} & A & \end{bmatrix} \begin{bmatrix} e(t-1) \\ \frac{\phi_1(t)}{\beta} \\ \frac{\phi_1(t)}{\beta} \end{bmatrix}$$

$$+ \frac{1}{1+\beta d(t)} \begin{bmatrix} \beta d(t) \\ \gamma_1 z(t-1) \\ \gamma_2 r(t) \end{bmatrix} n(t) \tag{4.4.38}$$

where $[A] =$

$$\begin{bmatrix} a & \beta z(t-1) & r(t) \\ -\gamma_1 a z(t-1) & 1+\beta d(t)-\beta\gamma_1 z^2(t-1) & -\beta\gamma_1 z(t-1) r(t) \\ -\gamma_2 a r(t) & -\beta\gamma_2 z(t-1) r(t) & 1+\beta d(t)-\gamma_2\beta r^2(t) \end{bmatrix}$$

and the last two equations correspond to the parameter adjustment and $\{n(t),\ t=0,\ldots,\infty\}$ is a zero-mean white noise sequence with each sample having variance σ^2.

Note that, in this algorithm, the error at time t is multiplied by the noise corrupted plant output at time (t-1) for the parameter adjustment laws. Since the additive noise samples at those two times are assumed to be uncorrelated, the expected value of the noise driving term in equation (4.4.38) is zero. Equations (4.4.38) can alternately be written as follows:

$$x(t+1) = A(t)x(t) = B(t)n(t) \tag{4.4.39}$$

where the correspondence of A(t), B(t), and x(t) with the elements of eqn. (4.4.38) is self-evident.

The weighted mean square error for a particular time t can now be written as $E[x'(t)Px(t)]$ where $P = P^T > 0$

Similarly, at time 2(t+1) the corresponding error - before taking expected values - is expressed as

$$x'[2(t+1)Px[2(t+1)] \quad . \tag{4.4.40}$$

Substitution of eqn. (4.4.39) in expression (4.4.40), in turn, yields:

$$x'(2(t+1))Px(2(t+1)) =$$

$$= x'(2)A'(2t)A'(2t+1)PA(2t+1)A(2t)x(t)t$$

$$+ 2x'(2t)A'(2t)A'(2t+1)PA(2t+1)B(2t)n(2t) +$$

$$+ 2x(2t)A'(2t)A'(2t+1)PB(2t+1)n(2t+1) +$$

$$+ n(2t)B'(2t)A'(2t+1)PA(2t+1)B(2t)n(2t) +$$

$$+ 2n(2t)B'(2t)A'(2t+1)PB(2t+1)n(2t+1) +$$

$$+ n(2t+1)B'(2t+1)PB(2t+1)n(2t+1) \quad . \tag{4.4.41}$$

Subtracting the term $x'(2t)A'(2t)PA(2t)x(2t)$ from both sides of the above equation and taking expectations we get

$$E[x'2(t+1))Px(t+1))] - E[x'(2t)Px'(2t)] =$$

$$= E[x'(2t)\{A'(2t)(A'(2t+1)PA(2t+1)-P)A(2t) +$$

$$+ A'(2t)PA(2t)-P\}x(2t)] +$$

$$+ 2E[x'(2t)A'(2t)A'(2t+1)PA(2t+1)B(2t)n(2t)] +$$

$$+ E[n(2t)B'(2t)A'(2t+1)PA(2t+1)B(2t)n(2t)]$$

$$+ E[n(2t+1)B'(2t+1)PB(2t+1)n(2t+1)] . \quad (4.4.42)$$

Straightforward algebraic manipulations for the algorithm of eqn. (4.4.38) shows in turn, that the following equality holds for all t.

$$A'(t)PA(t)-P = -H(t)H'(t) \quad (4.4.43)$$

where

$$P = \begin{bmatrix} 1/\beta & 0 & 0 \\ 0 & 1/\gamma_1 & 0 \\ 0 & 0 & 1/\gamma_2 \end{bmatrix} \quad (4.4.44)$$

and $H(t) =$

$$= \frac{1}{1+\beta d(t)} \begin{bmatrix} \frac{\sqrt{1-a^2}}{\beta}(1+\beta d(t)) & a\sqrt{\beta}\, d(t) & \sqrt{d(t)}\,a \\ 0 & -\sqrt{\beta}\, z(t-1) & \sqrt{d(t)}\beta z(t-1) \\ 0 & -\sqrt{\beta}\, r(t) & \sqrt{d(t)}\beta r(t) \end{bmatrix} \quad (4.4.45)$$

Finally, substitution of eqn. (4.4.43) in the first term on the RHS of eqn. (4.4.41) allows it to be rewritten as

$$-E[x'(2t)\{A'(t)H(2t+1)H'(2t+1)A(2t) +$$

$$+ H(2t)H'(2t)\}x(2t)] = -E[x'(2t)W(2t)x(2t)] \quad (4.4.46)$$

where

$$W(2t) \triangleq \{A'(2t)H(2t+1)H'(2t+1)A(2t) + H(2t)H'(2t)\}$$

in what follows we prove next that $W(2t) \geq \mu(2t)P$ where $\mu(2t)>0$ for all t within two consecutive time steps.

Let $W(2t) \triangleq L(2t)L'(2t)$ (4.4.47)

where $L(2t) = [H(2t) \quad A'(2t)H(2t+1)]$ (4.4.48)

Also, define $K(2t) =$

$$\begin{bmatrix} 0 & \sqrt{\beta} & \\ 0 & 0 & \frac{\gamma_1 z(2t-1)}{\sqrt{d(2t)}} \\ 0 & 0 & \frac{\gamma_2 r(2t)}{\sqrt{d(2t)}} \end{bmatrix}$$

$$T(2t) = \begin{bmatrix} I & K'(2t)H(2t+1) \\ 0 & I \end{bmatrix} \quad (4.4.49)$$

and

$$\bar{W}(2t) = L(2t)T(2t)T'(2t)L(2t) \quad (4.4.50)$$

Then

$$W(2t) \geq \frac{1}{\lambda_{max}(T(2t)T'(2t)}\, \bar{W}(2t) \quad (4.4.51)$$

where $\lambda_{max}(TT')$ is the maximum eigenvalue of TT'. Direct calculation shows that

$$\lambda_{max}(T(2t)T'(2t)) \leq 2\max\, 3, \frac{I+\gamma_1+\gamma_2}{\gamma_1\gamma_2} \text{ for all t.} \quad (4.4.52)$$

Further, it can be straightforwardly shown that

$$\bar{W}(2t) \geq \frac{\gamma_1\gamma_2}{2} \frac{(\beta z(2t)r(2t)-\beta z(2t-1)r(2t+1))^2}{(1+\beta d(2t))(1+\beta d(2t+1))}\, P \quad (4.4.53)$$

with P given by (4.4.44) since

$$\bar{W}(2t) = \begin{bmatrix} a^\dagger & 0 & 0 \\ 0 & b^\dagger & c^\dagger \\ 0 & b^{\dagger\dagger} & c^{\dagger\dagger} \end{bmatrix}$$

where $a^\dagger = \frac{1+\beta d(2t)-a^2}{\beta(1+\beta d(2t))} + \frac{a^2(1+\beta d(2t+1)-a^2)}{\beta(1+\beta d(2t+1))}$

$$b^\dagger = \frac{\beta z^2(2t-1)}{1+\beta d(2t)} + \frac{\beta z^2(2t)}{1+\beta d(2t+1)}$$

$$b^{\dagger\dagger} = \frac{\beta z(2t-1)r(2t)}{1+\beta d(2t)} + \frac{\beta z(2t)r(2t+1)}{1+\beta d(2t+1)}$$

$$c^\dagger = \frac{\beta z(2t-1)r(2t)}{1+\beta d(2t)} + \frac{\beta z(2t)r(2t+1)}{1+\beta d(2t+1)}$$

$$c^{\dagger\dagger} = \frac{\beta r^2(2t)}{1+\beta d(2t)} + \frac{\beta r^2(2t+1)}{1+\beta d(2t+1)}$$

Combining eqns.(4.4.51),(4.4.52) and (4.4.53) yields that

$$W(2t) \geq \mu(2t)P \tag{4.4.54}$$

where $\mu(2t) =$

$$= \frac{\gamma_1\gamma_2}{2 \max 3, \frac{1+\gamma_1+\gamma_2}{\gamma\ \gamma}} \frac{(\beta z(2t)r(2t)-\beta z(2t-1)r(2t+1))^2}{1+\beta d(2t))(1+\beta d(2t+1))} \tag{4.4.55}$$

We note that $\mu > 0$ unless

$$\{y(2t)+n(2t)\}r(2t) = \{y(2t-1)+n(2t-1)\}r(2t+1) \tag{4.4.56}$$

an event which occurs with zero probability. Also

$$E[x'2t)W(2t)x(2t)] \geq E[\mu(2t)x'(2t)Px(2t)] \tag{4.4.57}$$

$$\geq E(\mu(2t)]E[x'(2t)Px(2t)]$$

and clearly $E[\mu(2t)] > 0$ (4.4.58)

Substitution of eqns. (4.4.46), (4.4.57) in eqn. (4.4.41) results in the following inequality:

$$E[x'(2(t+1)Px(2(t+1))] \leq (1-E[\mu(2t)])E[x'2t(Px(2t)] +$$

$$+2Ex'(2t)A'(2t)A'(2t+1)PA(2t+1)B(2t)n(2t)] +$$

$$+ E[x'(2t)A'(2t)A'(2t+1)PA(2t+1)B(2t)n(2t)] +$$

$$+ E[n'2t+1)B'(2t+1)PB(2t+1)n(2t+1)] \tag{4.4.59}$$

Next, using eqn. (4.4.43) we prove below that the second term on the RHS of ineq. (4.4.59) is less than or equal to zero, independently of the fact that A(2t+1) depends on n(2t).

Let $D = E[x'(2t)A'(2t)A'(2t+1)PA(2t+1)B(2t)n(2t)]$

$$\leq E[x'(2t)A(2t)PB(2t)n(2t)] =$$

$$= E[x'(2t)A(2t)PB(2t)]E[n(2t)] = 0 \tag{4.4.60}$$

Similarly, for the third term on the RHS of the same inequality, (4.4.60), it can be shown that

$$E[n(2t)B'(2t)A'(2t+1)PB(2t+1)B(2t)n(2t)]$$

$$\leq E[n'(2t)B'(2t)PN(2t)n(2t)] \tag{4.4.61}$$

and substituting the values for B and P from eqns. (4.4.38) and (4.4.44) we further get

$$E[n'(2t)B'(2t)PB(2t)n(2t)] =$$

$$= E\left[\frac{\beta d^2(2t)+\gamma_1 z^2(2t-1)+\gamma_2 r^2(2t)}{\{1+\beta d(2t)\}^2}\right]E[n^2(2t)] =$$

$$= E\left[\frac{d(2t)(1+\beta d(2t))}{\{1+\beta d(2t)\}^2}\right]E[n^2(2t)] \leq \frac{\sigma^2}{\beta} \tag{4.4.61}$$

In an exactly analogous manner it can be shown that

$$E[n'(2t+1)B'(2t+1)PB(2t+1)n(2t+1)] \leq \frac{\sigma^2}{\beta} \tag{4.4.62}$$

Combining equations (4.4.60), (4.4.61),(4.4.62) eqn. (4.4.59) becomes:

$$E[x'(2(t+1))Px(2(t+1))] \leq (1-E[\mu(2t)])E[x'(2t)Px(2t)]+$$

$$+ \frac{2\sigma^2}{\beta} + D. \tag{4.4.63}$$

where $D \leq 0$ from eqn. (4.4.60).

Equation (4.4.63) states that indeed the system (4.4.38) is mean square stable in both output and parameter errors. Moreover, the steady state mean square error is bounded above by

$$\lim_{t\to\infty} E[x'(2t)Px(2t)] \leq \frac{\frac{2\sigma^2}{\beta}}{\lim_{t\to\infty} \inf E[\mu(2t)]} \tag{4.4.64}$$

where $\mu(2t)$ is as defined in eqn. (4.4.55).

Although the last proof seems to be more generally valid than the two preceding it, it does not solve the problem for all existing stochastic adaptive algorithms than may have a different structure in their information pattern. We simply refer the reader to [ROH 81], section 6, for a case in point. The continuous-time stochastic adaptive control convergence problem is considerably more complicated and the most rigorous results were obtained by Ljung in [LJU 77 a b] and surveyed in (i) of this section.

Our research in the last year, however, has shown that performance rather than stability has emerged as a more important issue, particularly during the transient adaptation phase, at least under the assumption of exact modeling of the process. Digital simulation studies of various adaptive control algorithms have uncovered some very undesirable characteristics exhibited by these algorithms, with instability occurring only in the presence of unmodelled dynamics. It has been our experience so far, that observation noise has not caused the adaptive loop to become unstable, but has contributed to degradation in the performance of the overall system. We discuss our findings in section 4.4.3. In the second part of the same section, we also provide

an analytical verification of the observed undesirable properties of adaptive algorithms.

4.4.3 Performance of Adaptive Algorithms

(i) Experimental Results

An intesive study of characteristics of existing direct adaptive control algorithms was conducted. The initial emphasis was to understand the transient behavior of such algorithms as well as their robustness to unmodeled dynamics and observation noise. Although the simulation results showed that no consistent pattern could be predicted, they nonetheless confirmed our suspicion that the majority of adaptive algorithms are characterized by

(1) high-frequency control signals characteristic of a high-bandwidth system

(2) the extreme sensitivity of the algorithms to unmodeled high-frequency dynamics which can result in unstable closed-loop behavior

(3) lack of robustness to observation noise; i.e.the presence of even a small amount of observation noise caused the closed loop system not to converge to the model but to slowly drift away to an increasingly higher bandwidth system.

We present below some specific results obtained for a second order system that address the above points. The results are typical of the behavior of most (if not all) adaptive algorithms whether discrete or continuous time and have been observed even for first order systems quite dramatically. The analysis in subsection (ii) will in fact concentrate on first order systems alone.

Example.

In what follows, convergence patterns for a plant whose transfer function contained an unstable pole are described. The plant transfer function was taken to be $W_p(s) = \frac{s + .5}{(s+1.5)(s-1)}$ and the reference model was chosen as $W_m(s) = \frac{5(s+1.5)}{(s+1)(s+2)}$. The adaptive algorithm in [NAR 78] was employed to control the plant. The reference input was square waves of amplitude 5 and 8 units at frequencies $\omega = .5$ and 1.5 rad/sec. "Snapshot" plots were taken of the closed loop poles of the controlled process with the following assumptions made; (1) if "snapshot" poles move slowly between consecutive snapshots, these poles should indicate approximate response to inputs at that time; (2) if poles move considerably between snapshots, they are then meaningless. The frequency content of the control input was also analyzed by the use of Fast Fourier Transforms (FFT).

a. Control with Process as Described Above

Two phases were exhibited during adaptation. In phase I the high frequency shape was maintained and the root locus at different "snapshots" follows successively the patterns depicted in Fig.4.4.1

It is interesting to note here that the distance between poles A and B remained constant at the different time instants when "snapshots" were taken. In phase II the poles move as if one were adjusting the loop gain and the pattern is shown in Fig.4.4.2 The control input in the first 2.6 secs. contained a frequency of 12 rad/sec at the 10 db down point (of the nominal system) and 50 rad/sec at the 20 db point. Its log magnitude plot exhibited a large hump between 65-100 rad/sec. At $\omega = 120$ rad/sec the magnitude went down to the 10 db point, while at 270 rad/sec it was reduced to the 20 db point.

Unmodeled Pole

In this experiment, an unmodeled pole was "added" to the plant so that the actual transfer function was $Wp(s) = \frac{s + .5}{(s+1.5)(s-1)} \frac{50}{s+50}$. The adaptive system was found to swing between two configurations in the initial stages, and up to t=7 secs, as shown in Fig. 4.4.3. Up to 7 secs the system parameters change smoothly and $e_1 < 1$ (unit). After t = 7 secs the controller fails completely in its objective and the overall system becomes unstable with large unbounded amplitudes in the adaptive signals up to 100 rad/sec frequency range. The same results were obtained when the unmodeled pole was placed at 150 rad/sec below the 10 db down point of the nominal system. In the log magnitude plot of the control input the peak was up by 6 db at 62 rad/sec and was 10 db down at 252 rad/sec.

Observation Noise

The same system, without the unmodeled dynamics, was now controlled with observation noise n(t) present at the output. The noise used was white with variance equal to 1 unit, and with $|n(t)|/r(t) \sim 1/10$. The evolution of the controlled system is shown successively in the diagrams shown in Fig. 4.4.4.

From the above, it is evident that as adaptation progresses, the system bandwidth is increased. It is interesting to observe in this experiment that the adaptive controller in the presence of observation noise performs much better- at least maintains stability of the process - than in the case where unmodeled dynamics are present.

(ii) Analytical Verification of the Undesirable Properties

In this subsection the undesirable characteristics discussed in the foregoing are demonstrated analytically and some insight as to what causes them is attained. The basic problem is that large reference inputs force the adaptive system to try to react too quickly. This results in a large bandwidth system and consequently in the excitation of unmodeled dynamics, which brings about instability. The additional effect of observation noise on the increase of bandwidth intensifies the problem further.

The analysis technique employs linearization of the nonlinear time-varying equations that describe the closed-loop system; this analysis technique is refered to as "final approach analysis" because the linearization is valid when the system and

reference model outputs are close to each other, a fact that occurs during the final phases of adaptation. In the final approach analysis it is assumed that the parameters of the controlled plant are very close to those parameters which would make the closed loop characteristics of the process the same as those of the reference model. Such a situation could develop when the asymptotically stable adaptive controller has already been operating for a long period of time with sufficiently rich inputs and therefore is close to final convergence. It could also arise when the plant parameters are fairly well known under reasonable a priori knowledge of the plant parameter values and the adaptation is just employed as a fine-tuning mechanism. Surely, if an algorithm behaves poorly under these mild conditions, it certainly cannot be expected to be useful as a practical control design.

Existing adaptive algorithms, under this analysis, are found to suffer (more or less) from the same basic problem: they lead to high-gain designs with large bandwidth. In what follows, we examine the characteristics of the discrete-time algorithm, discussed in 4.4.2(iii), under the light of the final approach analysis; no observation noise is assumed present. The method of analysis applies equivalently to continuous-time systems also; for a more detailed study of various adaptive algorithms using this approach, the reader is referred to [ROH 81].

Proper Modeling

The discrete-time system and model are exactly as described by eqns (4.4.34) and (4.4.35). The control input u(t) is generated in an exactly analogous manner as in eqn. (4.4.37), the only difference being that the output y(t) is not noise-corrupted, i.e.

$$u(t+1) = \theta_1(t+1)y(t) + \theta_2(t+1)r(t+1) - \rho\{\gamma_{11}y^2(t) + (\gamma_{12}+\gamma_{21})y(t)r(t+1) + \gamma_{22}r^2(t+1)\}e(t+1) \quad (4.4.65)$$

with all quantities defined as before.

The adaptive laws, similarly, are given by:

$$\frac{\phi_1(t+1)}{\beta} = \frac{\phi_1(t)}{\beta} - \gamma_{11}y(t-1)e(t) - \gamma_{12}r(t)\,e(t) \quad (4.4.66a)$$

$$\frac{\phi_2(t+1)}{\beta} = \frac{\phi_2(t)}{\beta} - \gamma_{21}y(t-1)e(t) - \gamma_{22}\,r(t)\,e(t) \quad (4.4.66b)$$

The resulting error equation for this system is then:

$$e(t) = y(t) - y^*(t) = \frac{ae(t-1) + \phi(t)y(t-1) + \phi_2(t)r(t)}{1 + \rho\beta d(t)} \quad (4.4.67)$$

where $d(t) = \gamma_{11}y^2(t-1) + (\gamma_{12}+\gamma_{21})y(t-1)r(t) + \gamma_{22}r^2(t)$ (4.4.68)

We further assume that r(t) = r = a constant, so that at the final stage, $y^* = y^*(t-1)$ and $d(t) = d^*$, as defined by eqn. (4.4.68) with y^* replacing y.

Then, the overall system error equations can be represented in state-space form as follows:

$$\begin{bmatrix} e(t) \\ \frac{\phi_1(t+1)}{\beta} \\ \frac{\phi_2(t+1)}{\beta} \end{bmatrix} = \left[\begin{array}{c:c:c} a^1 & a^2 & a^3 \end{array}\right] \begin{bmatrix} e(t-1) \\ \frac{\phi_1(t)}{\beta} \\ \frac{\phi_2(t)}{\beta} \end{bmatrix} \quad (4.4.69)$$

where the column vectors a^1, a^2, a^3 are given by:

$$a^1 = \frac{1}{1+\rho\beta d^*}\begin{bmatrix} a \\ \gamma_{11}ay^* - \gamma_{12}ar \\ -\gamma_{21}ay^* - \gamma_{22}ar \end{bmatrix} \quad (4.4.70a)$$

$$a^2 = \frac{1}{1+\rho\beta d^*}\begin{bmatrix} \beta y^* \\ 1+\rho\beta d^* + (-\gamma_{11}\beta y^{*2} - \gamma_{12}\beta y^* r) \\ -\gamma_{21}\beta y^{*2} - \gamma_{22}\beta y^* r \end{bmatrix} \quad (4.4.70b)$$

$$a^3 = \frac{1}{1+\rho\beta d^*}\begin{bmatrix} \beta r \\ -\gamma_{12}\beta y^* r - \gamma_{22}\beta r^2 \\ 1+\rho\beta d^* + (-\gamma_{21}\beta y^* r - \gamma_{22}\beta r^2) \end{bmatrix} \quad (4.4.70c)$$

The above system has the characteristic equation:

$$(z-1)[(z-1)(z-a) + \beta d^*\rho(z(z+\tfrac{\rho-1}{\rho}))] = 0 \quad (4.4.71)$$

There is a marginally stable pole frozen at z = 1 associated with the eigenvector

$$e(t-1) = 0;\quad \phi_2(t) = \frac{-y^*}{r}\,\phi_1(t) \quad (4.4.72)$$

Two other poles appear in a d* - root locus as shown in figure 4.4.5. One pole starts at z=a

and the other at z=1 and, with increasing d*, move towards the zeros at z=0 and z= $\frac{1-\rho}{\rho}$.

The latter zero is, however, under the designer's control, by use of the parameter ρ.

If the additional error feedback terms were not present in eqn. (4.4.65), i.e. if ρ = 0, the zero at z = $\frac{1-\rho}{\rho}$ would be missing. As a result, one of the poles would move along the negative real axis towards infinity causing a chatter type instability, characteristic of discrete-time systems. Incidentally, this additional term is customarily not present in many adaptive algorithms [NAR 78], [FEU 78], etc.

Analysis with an Unmodeled Pole

In order to investigate the effects of unmodeled dynamics, the actual plant is augmented to have two poles, located at $-\alpha_1$ and $-\alpha_2$ respectively. It is also assumed that there exists a second order model with poles at $-a_1$, $-a_2$ as eqns. (4.4.73) show:

Actual Plant: $y(t+2) - (\alpha_1+\alpha_2)y(t+1) +$

$$+ \alpha_1\alpha_2 y(t = \beta u(t+2)$$

Reference Model: $y^*(t+2) - (a_1+a_2)y^*(t+1) +$

$$+ a_1a_2y^*(t) = b_2r(t+2) \qquad (4.4.73)$$

In addition, the following conditions are required for the analysis:

(i) The reference model is stable.

(ii) $a_1+a_2 = \alpha_1+\alpha_2$

(iii) $\frac{b_2}{1-(a_1+a_2)+a_1a_2} = \frac{b}{1-a}$ of eqn.(4.4.35)

Conditions(i) and (iii) allow the substitution of eqn. (4.4.35) with eqn. (4.4.73b) with no change in (steady-state) response for constant reference inputs. Condition (ii) is necessary for the analysis and is somewhat restrictive in the plants and models that can be studied comparatively particularly more so in the discrete-time case. We also note that although the ensuing analysis is carried out for constant reference input, it is actually valid over the range of reference input frequencies where (4.4.73b) matches (4.4.35).

The control input is chosen exactly as in (4.4.65), as if the plant were first order, and the adaptive laws are described by eqns. (4.4.66). The output error equations then become:

$$e(t) = \frac{(a_1+a_2)e(t-1) - a_1a_2e(t-2) + \phi_1(t)y(t-1)}{1 + \rho\,\beta\,d(t)} + \frac{\phi_2(t)r(t)}{1+\rho\beta d(t)} \qquad (4.4.74)$$

The final approach analysis yields the following characteristic equation:

$$(z-1)[(z-1)(z^2-(a_1+a_2)z + a_1a_2) + \beta d^*(z(z+\rho-1))]=0 \qquad (4.4.75)$$

The d* - root locus in the figure 4.4.6 shows that the error system will become unstable for large reference inputs.

In some cases it is possible to choose the adaptive gains in such a way, so as to artificially slow down the adaptive process when the reference inputs are large, resulting in a smaller bandwidth system and improved final approach behavior.

In order to achieve this the adaptive gains must be (nonlinear) functions of the reference input (and the adaptive signals). However, global asymptotic stability has only been proven for the case of constant adaptive gains or certain restricted types of time-varying gains. Thus, the approach that follows is only theoretically valid for the case of constant reference inputs. However, its validity is not limited to the final approach domain but extends to the entire duration of the adaptation process.

To improve the final approach characteristics, the constant gain matrix $\Gamma_{old} = \{\gamma_{ij}\}$ is replaced by the following matrix

$$\Gamma_{new} = \frac{1}{\gamma+r^2+y^{*2}} \cdot \Gamma_{old}$$

Then, d* becomes

$$d^* = \begin{bmatrix} y^* \\ r \end{bmatrix}^T \Gamma_{new} \begin{bmatrix} y^* \\ r \end{bmatrix}$$

with the condition

$$d^* \leq \frac{y^{*2} + r^2}{\gamma + y^{*2} + r^2}\ \sigma_{max}(\Gamma_{old}) \leq \sigma_{max}(\Gamma_{old}).$$

where $\sigma_{max}(\Gamma_{old})$ is the maximum eigenvalue of Γ_{old} and $\gamma > 0$. Both $\sigma_{max}(\Gamma_{old})$ and γ are under designer's control.

Thus, given an upper bound on β, $\sigma_{max}(\Gamma_{old})$ can be chosen to limit how far along the d*-root locus the corresponding system roots can travel. Consequently, the maximum frequency of parameter error variation in the final approach is under the direct control of the designer. Also, with an upper bound on d*, the adaptive system is able to handle any number of high frequency unmodeled poles while retaining final approach stability.

Observation Noise

We have discussed extensively in (iii) of the preceding subsection as well as in (i) of this subsection the case of observation noise in an otherwise properly modeled system. In (iii) of 4.4.2 we in fact gave a proof for the mean-square boundedness of output and parameter errors of the

particular discrete-time algorithm studied there. We will not, therefore, go into any further detailed discussion here, since the final approach analysis corroborates what has already been stated in the above mentioned subsections. The reader is referred to [ROH 81] for more details, if desired. We would only like to direct the reader's attention once again to the importance of the linear noise forcing term in eqn. 4.4.38 of subsection 4.4.2, which effectively accounts for the increase in bandwidth of the overall adaptive system.

4.4.4 Conclusions and Directions for Further Research

We have examined in the preceding convergence issues in stochastic adaptive control algorithm. Special emphasis was placed on the performance of such algorithms which seems to be of crucial importance, particularly in the transient phases of adaptation and in cases where the proper modeling assumption in not valid, as is usually the case in any engineering application. Instability has quite often resulted under such circumstances, within the framework of an algorithm which was designed at the outset to be (globally) stable.

A new method, called final approach analysis was introduced to analyze the dynamic properties of adaptive control algorithms with special emphasis on their robustness to

(a) generation of high frequencies in the plant control signal,

(b) excessive bandwidth of the adaptive control loop resulting in excitation of unmodeled dynamics and, consequently, leading to dynamic instability of the closed-loop adaptive system,

(c) noise corrupted measurements.

An adaptive control algorithm must have reasonable tolerance to such modeling error and stochastic uncertainties before it can be used routinely in practical applications. However, the final approach analysis has shown that, currently, such algorithms have unacceptable dynamic characteristics.

The final approach analysis is useful because it can be used in a constructive way to adjust the adaptivegains so as to limit the closed-loop system bandwidth and to ameliorate some of the undesirable characteristics of existing adaptive algorithms. However, though necessary, the final approach analysis is by no means sufficient in the analysis and design of adaptive systems. The technique is limited to the cases in which the output error is small and does not change rapidly, so that dynamic linearization of the complex non-linear (differential or) difference equations that describe the adaptation process makes sense. By itself, it cannot predict what happens in the truly transient phase; the simulation results presented in 4.4.3(i) suggest that even more complex and undesirable effects are present.

It is our opinion that a great deal of additional basic research is needed in the area of adaptive control. Future theoretical investigation must, however, take drastically new directions than those reported in the recent literature. The existence of unmodeled dynamics and stochastic effects must be an integral part of the theoretical problem formulation. In addition, future adaptive algorithms must be able to deal with problems in which partial knowledge of the system dynamics is available so that at the very least the intentional augmentation of the controlled plant dynamics with roll-off and noise rejection transfer functions can be handled without confusing the adaptation process. Such augmentation of the plant dynamics (loop-shaping in the frequency domain) is necessary even in non-adaptive modern control systems [DOY 81] for good performance and stability; clearly, the same techniques must be used in adaptive systems. The adaptive algorithms currently available cannot handle the additional dynamics because the existence of the latter violates the theoretical assumptions necessary to assure (global) stability and stochastic convergence.

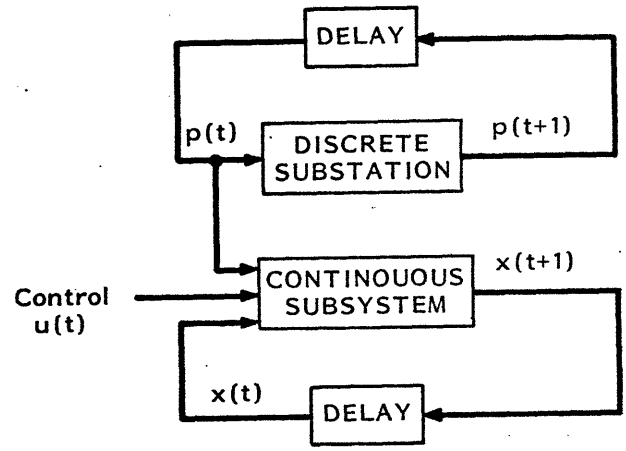


Figure 4.2.1 Hierarchical Hybrid Systems

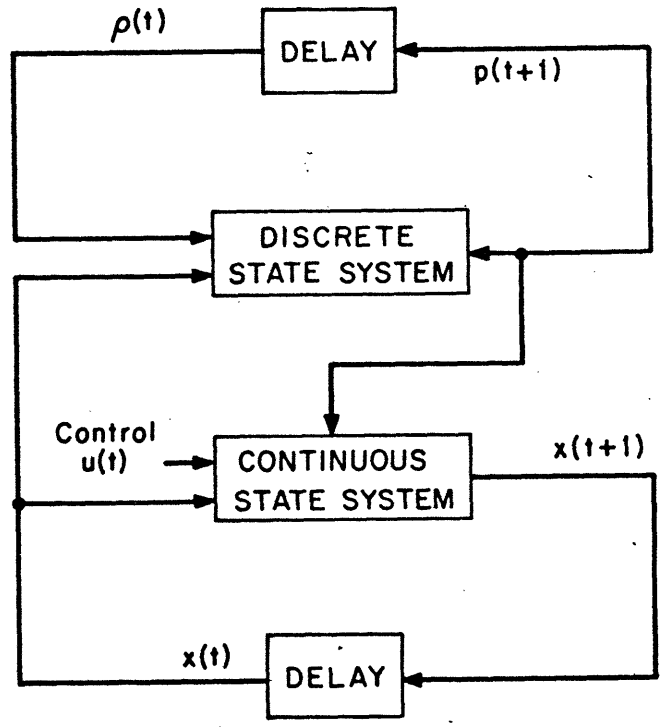


Figure 4.2.2 Fully interconnected hybrid systems

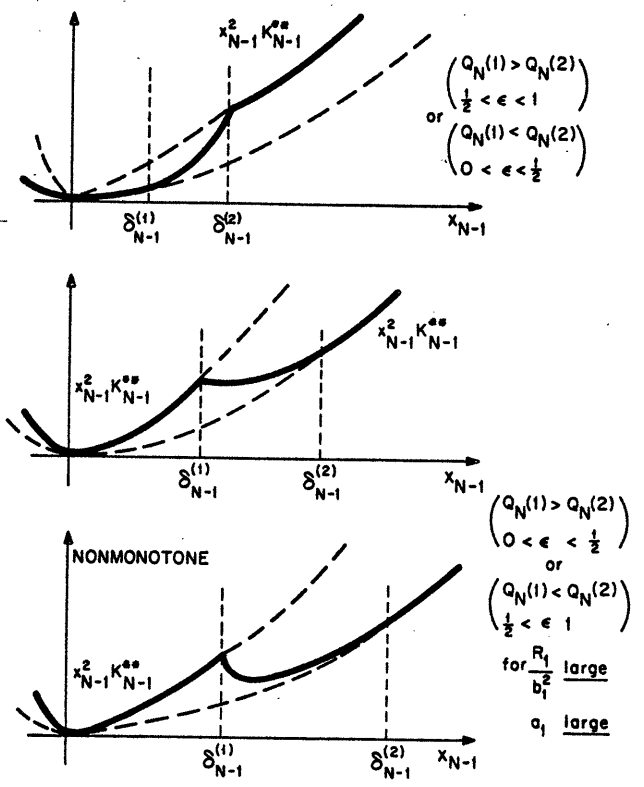


Figure 4.2.3

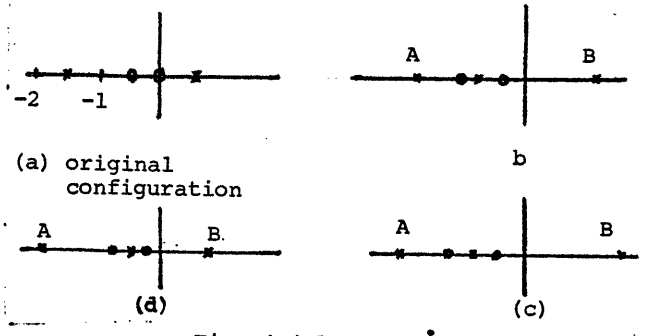


Fig. 4.4.1 - Phase I

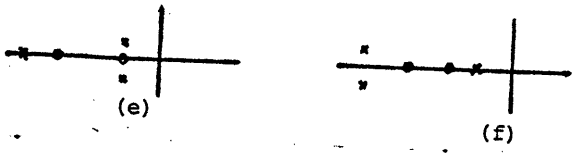


Fig. 4.4.2 - Phase II

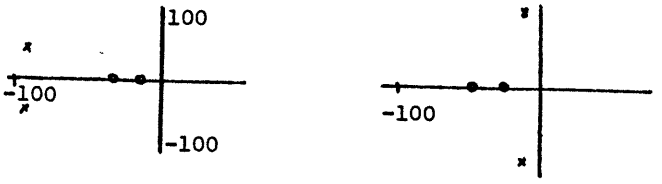


Fig. 4.4.3 - Unmodeled Pole

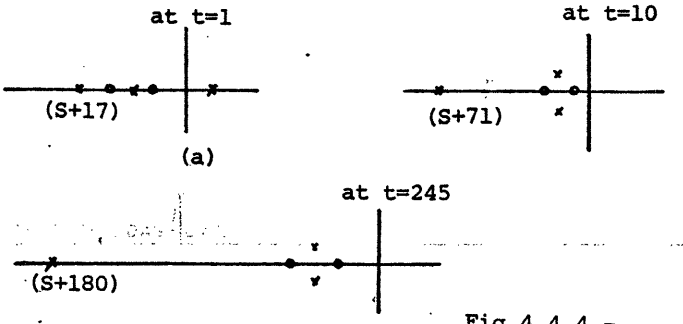


Fig.4.4.4 - Observation Noise

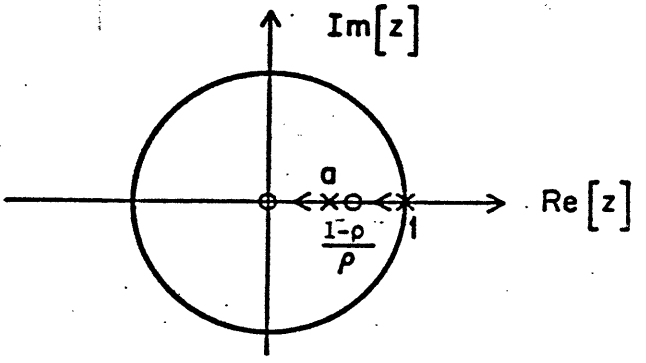


Fig.4.4.5 -d* root locus of eqn (4.4.71)

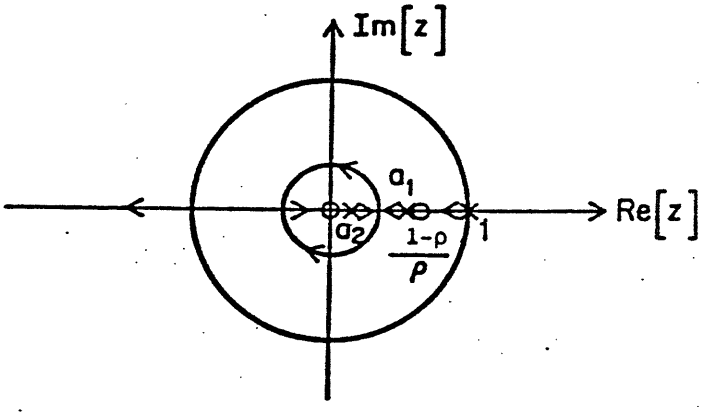


Fig.4.4.6 - d*-root locus of eqn. (4.4.75)

5. The Thermodynamic Properties of Markov Processes

5.1 Introduction

Classical thermodynamics developed as a description of systems in equilibrium. Through Onsager's reciprocity relations, the treatment was extended to nonequilibrium conditions where perturbations about equilibrium preserved a form of linearity. The theory had a strong physical and a weak mathematical basis.

Attempts have been made to further extend the theory to describe systems far from equilibrium and reconstruct it on a firm mathematical foundation. The need exists for a general axiomatic development incorporating only those features of macroscopic processes which yield both their equilibrium and nonequilibrium thermodynamic properties. In this chapter we show that, using stochastic system theory, such an axiomatic development is possible. Our approach rests upon the physically reasonable assumption that for the purposes of a thermodynamic description of a system, the system can be viewed as an ensemble of particles, each representable by a stochastic differential equation describing the evolution of a Markov process.

In Section 5.2, we present an overview of the principles of thermodynamics, in the context of nonequilibrium thermodynamics, or the thermodynamics of irreversible processes. These ideas -- a characterization of a thermodynamic system, the first and second laws, a definition of the equilibrium state and equivalent conditions, a decomposition of the flows of work and the entropy production into inner products of forces and fluxes, and Onsager's reciprocity relations -- are captured by the theory outlined in Sections 5.3 through 5.5.

Two examples developed in the literature serve to illustrate the generality of our approach. As done by Brockett and Willems [BRO 79], in Section 5.3.1 we analyze the equations of motion for a system constructed from a Nyquist-Johnson resistor and a linear capacitor of time-varying capacitance.

Particularly in developing a model of muscle contraction assuming the cross-bridge theory, our formalism is invaluable. Hill [HIL 74, HIL 77], using statistical mechanical theory, has outlined a formalism which relates the rate constants and free energy changes determined biochemically to the mechanical and thermal properties of contracting muscle. As we show in a Section 5.3.2 our approach yields the same results. Furthermore, without additional work, we have proven reciprocity of the model near equilibrium.

In Section 5.4 we present an outline of the results which applies to both continuous-state Markov processes (Markov diffusion processes) and discrete-state Markov processes (Markov chains).

In Section 5.5 we develop an axiomatic framework for thermodynamics assuming a Markov chain description of a thermodynamic system. As shown in the previous section, the results can be generalized to broader classes of Markov processes, but the proofs are most easily followed in the context of Markov chain theory.

Directions for further research are briefly discussed in Section 5.6.

5.2 Thermodynamics

In this section, we present an overview of the principles of thermodynamics, in the context of nonequilibrium thermodynamics, or the thermodynamics of irreversible processes. The concepts presented in this section can be found in Callen [CAL 60], Katchalsky and Curran [KAT 65], Lavenda [LAV 78], or Nicolis and Prigogine [NIC 77].

5.2.1 A Characterization of a Thermodynamic System

A system is that subset of the universe we isolate from its surroundings for thermodynamic study. An isolated system can exchange neither energy nor matter with its surroundings; a closed system can exchange energy but not matter with its surroundings; an open system can exchange both energy and matter with its surroundings. A system is described by a set of internal and external parameters. The state of the system can be specified by a set of values for these parameters.

5.2.2 The First and Second Laws

With a change in state of a system, the change in a state function depends only on the initial and final states, while a path function depends on the path through which the change in state is effected. Thus, the differential of a state function is exact and the differential of a path function is inexact. The differential of a state function X can be decomposed (formally) into the sum of the differentials of two path functions, as

$$dX = d_e X + d_i X \ ,$$

where $d_e X$ represents changes in X due to the flow of X between the system and its surroundings, and $d_i X$ represents the production of X in the system.

Using the terminology of Willems' work on dissipative systems [WIL 72], we say a thermodynamic system with storage function X is dissipative with respect to the supply rate $\frac{d_e X}{dt}$ and lossless with respect to the sum of the supply rate and the dissipation rate $\frac{-d_i X}{dt}$. The corresponding mathematical statements are:

$$\frac{dX}{dt} = \frac{d_e X}{dt} + \frac{d_i X}{dt} \leq \frac{d_e X}{dt} \ ; \ \frac{d_i X}{dt} \geq 0 \ . \qquad (5.2.1)$$

The first law of thermodynamics postulates the existence of a state function E called the internal energy such that

$$\frac{dE}{dt} = \frac{d_e E}{dt}$$

$$\frac{d_i E}{dt} = 0 \ , \qquad (5.2.2)$$

i.e., the internal energy of a system is conserved. In a closed system, the flow of energy between the system and its surroundings is equal to the difference of the heat flow $\frac{đq}{dt}$ and the work flow $\frac{đw}{dt}$, as

$$\frac{dE}{dt} = \frac{đq}{dt} - \frac{đw}{dt} \ . \qquad (5.2.3)$$

By convention, a flow from the system to the surroundings is positive for work and negative for heat. In an open system, the heat and work terms include contributions due to the flow of matter between the system and its surroundings.

The second law of thermodynamics postulates the existence of a state function S called the entropy, such that

$$\frac{dS}{dt} = \frac{d_e S}{dt} + \frac{d_i S}{dt} \geq \frac{d_e S}{dt}$$

$$\frac{d_i S}{dt} \geq 0 \quad , \tag{5.2.4}$$

i.e., the entropy produced inside a system is nondecreasing as a function of time. A temperature T can be assigned to a system. Then,

$$\frac{d_e S}{dt} = \frac{1}{T} \frac{đq}{dt} \tag{5.2.5}$$

For an isothermal change of state, at a constant temperature T,

$$\frac{d}{dt} E - T \frac{d}{dt} S = \frac{d}{dt}(E - TS) \quad .$$

The (Helmholtz) free energy F is defined as

$$F = E - TS \quad . \tag{5.2.6}$$

We define heat flow, work flow, dissipation (entropy production) rate, internal energy, and free energy as

$$Q = \frac{1}{T} \frac{đq}{dt}$$

$$W = \frac{1}{T} \frac{đw}{dt}$$

$$D = \frac{d_i S}{dt} \tag{5.2.7}$$

$$U = \frac{1}{T} E$$

$$G = \frac{1}{T} F$$

and henceforth assume that the temperature T is independent of time. We summarize our discussion of the first and second laws of thermodynamics by:

$$\frac{d}{dt} U = Q - W$$

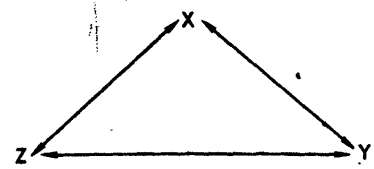
$$\frac{d}{dt} S = Q + D \geq Q \tag{5.2.8}$$

$$\frac{d}{dt} G = -W - D \leq -W \quad .$$

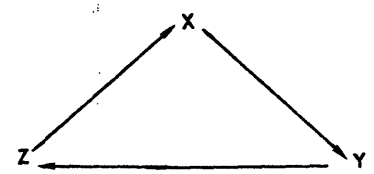
The heat flow Q and work flow W are observable only through effects produced in the surroundings; the internal entropy production D is not observable through effects produced in the surroundings. In an isolated system $W = Q = \frac{dU}{dt} = 0$ and $\frac{dS}{dt} \geq 0$. In a system with $W = 0$, $\frac{dG}{dt} \leq 0$. In steady state, all state functions are independent of time, so in a system where either $W = 0$ or $Q = 0$, $D = 0$ in steady state.

5.2.3 The Equilibrium State

A steady state can be maintained by a cyclic mechanism. For example, in a cyclic reaction



a steady state could be maintained by



where a steady flow exists around a cycle. When detailed balance holds, there is "on the average, ... the same frequency of transition from the condition [X] to [Y] as from the condition [Y] to [X]. ... the transitions from [X] to [Y] do not have to be thought as balanced with the help of some indirect route such as [Y] to [Z] to [X]," as expressed by Tolman [TOL 38] When microscopic reversibility holds, as formulated by Tolman [TOL 38], "any molecular process and the reverse of that process will be taking place on the average at the same rate," i.e., on the average, a microscopic process is invariant under time-reversal. Equilibrium is defined as that steady state for which detailed balance holds, and is equivalent to that steady state for which microscopic reversibility holds. In a reaction system, there are no oscillations in the approach to equilibrium.

The equivalence of detailed balance, microscopic reversibility, and equilibrium are quantum principles. Thermodynamics assumes $D = 0$ in a steady state if and only if the steady state is an equilibrium state. An equilibrium state is completely determined by the external parameters of the system. There is an equilibrium state in which a system is at a uniform composition.

5.2.4 Forces and Fluxes

Nonequilibrium thermodynamics gives a structure for the work flow W and the entropy production D, which is nonzero for nonequilibrium states. The work flow W can be written as the inner product of a vector of fluxes $\underline{I}$ and a vector of forces $\underline{F}$, usually as

$$W = \sum_i I_i F_i = \langle \underline{I}, \underline{F} \rangle$$

For example, the flow of electrical work is the product of current and voltage. The entropy production D can be written as the inner product of a vector of fluxes $\underline{J}$ and a vector of forces $\underline{X}$, $D = \sum_i J_i X_i = \langle \underline{J}, \underline{X} \rangle$. A steady state is an equilibrium state if and only if all forces and fluxes vanish identically. Close to equilibrium, a linear relation is assumed between the forces

and fluxes, as

$$\underline{J} = \underline{L}\,\underline{F}$$

Onsager showed that the matrix $\underline{L}$ is symmetric for processes sufficiently closed to equilibrium (in the linear range) as a consequence of the principles of detailed balance and microscopic reversibility. The relations $L_{ij} = L_{ji}$ are known as Onsager's reciprocity relations.

There have been attempts to describe the evolution of a system toward steady state, far from the equilibrium state. For a system in which the entropy flow is only at the physical boundary of the system, Nicolis and Prigogine [NIC 77] define a quantity P as the rate of entropy production. For zero fluxes or time-independent concentrations at the boundary, $\frac{d}{dt}P$ is decomposed as

$$\frac{d}{dt}P = \sum_i J_i \frac{d}{dt}X_i + \sum_i X_i \frac{d}{dt}J_i$$

$$= \frac{d_x P}{dt} + \frac{d_J P}{dt} \quad . \tag{5.2.9}$$

For a system far from equilibrium (in the nonlinear range) $\frac{d_x P}{dt}$ is nonpositive, with equality if and only if the system is in steady state, in what has been called the universal evolution criterion. Note that $d_x P$ is an inexact differential.

The thermodynamic concepts presented here are based upon a weak mathematical foundation. We shall show that, by assuming a thermodynamic system evolves as a Markov process, these ideas can be interpreted in terms of mathematical properties of Markov processes.

5.3 Examples

5.3.1 The Example of Brockett and Willems

A Nyquist-Johnson resistor connected to a capacitor is a simple example of a thermodynamic system. Brockett and Willems [BRO 79] have obtained some intuitively-appealing results by analyzing the model. Our analysis shall serve as an introduction to a general treatment of the thermodynamic properties of Markov processes.

Consider a Nyquist-Johnson resistor at temperature T and of conductance g connected to a linear capacitor of capacitance c. Allow the temperature T and capacitance c to be functions of the time $t \in [t_0, \infty)$. The equation of motion for the capacitor charge q is the Itô equation

$$dq(t) = \frac{g}{c(t)}\,q(t) + \sqrt{2kT(t)g}\;\; dw(t)$$

where k is Boltzmann's constant and w(t) is Brownian motion. Assume the capacitance c is determined by a control u

$$\frac{dc(t)}{dt} = u(t).$$

Then the state equations for the network are

$$\begin{aligned} dq &= -\frac{g}{c}\,q\,dt + \sqrt{2kTg}\;\; dw \\ dc &= u\,dt \quad . \end{aligned} \tag{5.3.1}$$

Our thermodynamic system is a collection of independent copies of the process $\{q(t),\ t \in [t_0,\infty)\}$. Let p(t,q) be the density function of q. Assume $q(t_0) = 0$ w.p.1. As the density p(t,q) can then be written as the Gaussian density

$$p(t,q) = \frac{1}{\sqrt{2\pi\,\Sigma(t)}}\; e^{-\frac{q^2}{2\Sigma(t)}} \quad ,$$

either the variance $\Sigma(t)$ or the density function p(t,q) can be taken as the state of the system at time t.

Definition 5.3.1. The fluxes of the system are

$$\begin{aligned} j(t,q) &= -\frac{g}{c(t)}\,q\,p(t,q) - kT(t)g\,\frac{\partial p(t,q)}{\partial q} \\ u(t) &= \frac{dc(t)}{dt} \quad . \end{aligned} \tag{5.3.2}$$

■

Note that Kolmogorov's forward equation can then be written as

$$\frac{\partial p}{\partial t} = -\frac{\partial j}{\partial q} \quad .$$

Definition 5.3.2. A periodic trajectory of period τ is a solution $\Sigma(\cdot)$ of

$$\frac{d\Sigma(t)}{dt} = -2\,\frac{g}{c(t)}\,\Sigma(t) + 2k\,T(t)g \;, \tag{5.3.3}$$

where $c(t) = c(t+\tau)$ and $T(t) = T(t+\tau)$, with the property that $\Sigma(t+\tau) = \Sigma(t)$ for some $\tau < \infty$. ■

Definition 5.3.3. An equilibrium state is a state of zero flux. Let $p^e(t,q)$ denote the equilibrium state of the system with time-independent parameters $c = c(t)$ and $T = T(t)$. ■

If the fluxes $j = 0$ and $u = 0$ then

$$\frac{\partial}{\partial q}\ln p = -\frac{q}{kTc}$$

by manipulation of (5.3.2) and the capacitance c is constant. Thus

$$\Sigma = kTc$$

and by (5.3.3) $\frac{d\Sigma}{dt} = 0$. So an equilibrium state $p^e(t,q)$ is the unique invariant solution of Kolmogorov's forward equation with the parameters fixed for all times $s \geq t$.

Definition 5.3.4. Let the external parameter F be defined as

$$F(t,q) = \frac{q}{kT(t)c(t)} \quad . \qquad ■$$

For an equilibrium state p^e

$$-\frac{\partial}{\partial q}\ln p^e = F \quad .$$

Thus in equilibrium, the state of the system can be written solely in terms of the external parameter F.

Definition 5.3.5.

Let the internal force X be defined as

$$X(t,q) = \frac{1}{kT(t)q} \frac{j(t,q)}{p(t,q)} . \quad \blacksquare$$

The internal force X = 0 iff the flux j = 0.

Lemma 5.3.6.

$$-\frac{\partial}{\partial q} \ln p = X + F$$

Proof: Follows by calculation. ■

Definition 5.3.7. Let the entropy S be defined as

$$S(t) = \frac{k}{2}(\ln 2\pi + \ln \Sigma(t) + 1) . \quad \blacksquare$$

Definition 5.3.8. Let the heat flow Q be defined as

$$Q(t) = kT(t) \langle j,F \rangle = kT(t) \int j(t,q)F(t,q)dq$$

and the entropy production D as

$$D(t) = \frac{dS(t)}{dt} - \frac{Q(t)}{T(t)} . \quad \blacksquare$$

By calculation, $Q = \frac{kq}{c}(T - \frac{\Sigma}{kc})$.

Proposition 5.3.9. The entropy production

$$D = k \langle j,X \rangle \geq 0,$$

with equality iff the system is at equilibrium.

Proof:

$$\frac{dS}{dt} - \frac{Q}{T}$$

$$= -\frac{d}{dt} k \int p \ln p \, dq - k \langle j,F \rangle$$

$$= -k \langle \frac{\partial p}{\partial t} , \ln p \rangle - k \langle j,F \rangle$$

$$= -k \langle -\frac{\partial j}{\partial q} , \ln p \rangle - k \langle j,F \rangle$$

$$= k \langle j, -\frac{\partial}{\partial q} \ln p - F \rangle$$

$$= k \langle j, X \rangle$$

$$= k \langle \frac{1}{kTg} \frac{j^2}{p} \rangle \geq 0,$$

with equality iff equilibrium. ■

We thus have the "second law of thermodynamics,"

$$\frac{dS}{dt} \geq \frac{Q}{T} .$$

Definition 5.3.10. The energy E is defined as

$$E(t) = \frac{1}{2} \frac{\Sigma(t)}{c(t)} . \quad \blacksquare$$

Definition 5.3.11. The work flow W is defined as

$$W(t) = \frac{1}{2} u(t) \frac{\Sigma(t)}{c^2(t)} . \quad \blacksquare$$

Proposition 5.3.12. The thermodynamic functions E, Q, and W satisfy the "first law of thermodynamics,"

$$\frac{dE}{dt} = Q - W .$$

Proof:

$$E = \langle p, \frac{1}{2} \frac{q^2}{c} \rangle .$$

So

$$\frac{dE}{dt} = \langle \frac{\partial p}{\partial t}, \frac{1}{2} \frac{q^2}{c} \rangle + \langle p, -\frac{1}{2} \frac{q^2}{c^2} u \rangle$$

$$= \langle j, \frac{q}{c} \rangle - \langle u, \frac{1}{2} \frac{q^2}{c^2} p \rangle$$

$$= Q - W . \quad \blacksquare$$

Definition 5.3.13. For a system at constant temperature T, the free energy G is defined as

$$G = E - TS . \quad \blacksquare$$

Proposition 5.3.14. For a system at a constant temperature $\frac{dG}{dt} \leq -W$, with equality iff the system is at equilibrium.

Proof:

$$\frac{dG}{dt} = \frac{dE}{dt} - T \frac{dS}{dt}$$

$$= Q - W - TD - Q$$

$$= -W - TD \leq -W \quad \blacksquare$$

Proposition 5.3.15. For any periodic trajectory of period τ

$$\int_0^\tau \frac{Q(t)}{T} dt \leq 0 .$$

Proof:

$$\int_0^\tau \frac{Q}{T} dt \leq \int_0^\tau \frac{dS}{dt} dt = 0 . \quad \blacksquare$$

Proposition 5.3.16. For any periodic trajectory for which the temperature is constant

$$\int_0^\tau W dt \leq 0 .$$

Proof:

$$\int_0^\tau W dt \leq -\int_0^\tau \frac{dG}{dt}\, dt = 0 \quad . \quad \blacksquare$$

The above propositions capture the classical thermodynamic ideas governing the behaviour of heat engines. If we define the notation f^+ to mean the positive part of a function f, i.e., $f^+(t) = \max\{f(t), 0\}$, then Proposition 5.3.15 implies

$$0 \geq \int_0^\tau \frac{Q^+}{T}\, dt + \int_0^\tau \frac{Q - Q^+}{T}\, dt \quad .$$

Let $T_{max} = \sup_{0\leq t\leq\tau} \{T(t)\}$ and $T_{min} = \inf_{0\leq t\leq\tau} \{T(t)\}$. Then

$$0 \geq \frac{1}{T_{max}} \int_0^\tau Q^+ dt + \frac{1}{T_{min}} \int_0^\tau (Q - Q^+) dt$$

and

$$\frac{\int_0^\tau (Q-Q^+)\, dt}{\int_0^\tau Q^+ dt} \leq -\frac{T_{min}}{T_{max}} \quad .$$

But, using Proposition 5.3.12,

$$\int_0^\tau Q dt = \int_0^\tau W dt \quad .$$

So

$$\frac{\int_0^\tau W dt}{\int_0^\tau Q^+ dt} \leq \frac{T_{max} - T_{min}}{T_{max}} \quad ,$$

which is the classical bound on the efficiency of a heat engine.

Define an equilibrium trajectory $\Sigma^e(\cdot)$ as $\Sigma^e(t) = kT(t)\, c(t)$. Under suitable smoothness assumptions, the function

$$\Sigma^*(t) \triangleq \lim_{\varepsilon\to 0} kT\left(\frac{t}{\varepsilon}\right) c\left(\frac{t}{\varepsilon}\right)$$

$$= kT^*(t) c^*(t)$$

is a solution of

$$-\frac{d\Sigma(t)}{dt} = -2\frac{g}{c^*(t)}\, \Sigma(t) + 2kT^*(t) g$$

where $c^*(t) \triangleq \lim_{\varepsilon\to 0} c\left(\frac{t}{\varepsilon}\right)$ and $T^*(t) \triangleq \lim_{\varepsilon\to 0} T\left(\frac{t}{\varepsilon}\right)$ [WIL 79].

For a system at a constant temperature T the work extracted from an equilibrium trajectory

$$W^e \triangleq \frac{1}{2} u \frac{\Sigma^e}{c^2} = \frac{1}{2} kT \frac{d}{dt} \ln c$$

$$= \frac{1}{2} kT \frac{d}{dt} \ln \Sigma^e$$

$$= -\frac{dG^e}{dt} \quad ,$$

where

$$G^e \triangleq E^e - TS^e$$

$$= \frac{1}{2} \frac{\Sigma^e}{c} - \frac{1}{2} kT(\ln 2\pi + \ln \Sigma^e + 1)$$

$$= -\frac{1}{2} kT(\ln 2\pi + \ln \Sigma^e) \quad .$$

But

$$kT \langle p, -\ln p^e \rangle$$

$$= \frac{1}{2} kT \left(\ln 2\pi + \ln kTc + \frac{\Sigma}{kTc}\right)$$

$$= E + \frac{1}{2} kT(\ln 2\pi + \ln \Sigma^e) \quad .$$

$$= E - G^e \quad .$$

So

$$\frac{d}{dt}\, kT \langle p, -\ln p^e \rangle$$

$$= \frac{dE}{dt} - \frac{dG^e}{dt}$$

$$= Q - (W - W^e) \quad .$$

By calculation

$$S = -k \langle p, \ln p \rangle \quad .$$

Thus for a system at a constant temperature T

$$\frac{d}{dt}(G - G^e) = \frac{d}{dt} kT \langle p, \ln \frac{p}{p^e} \rangle$$

$$= -(W - W^e) - D \leq -(W - W^e) \quad .$$

and for any periodic trajectory

$$\int_0^\tau W dt \leq \int_0^\tau W^e dt \quad .$$

Through the appropriate definitions, the diffusion process q has acquired a thermodynamic flavour. In the next section, we show such is also the case for a Markov chain model of muscle contraction.

5.3.2 A Model of Muscle

The mechanism of striated muscle contraction is explained by 2 hypotheses [HUX 74]:

(1) The length of striated muscle changes through a sliding movement of thin (actin) filaments relative to thick (myosin) filaments.

(2) Active movement results from a cyclical interaction of projections of the myosin filaments -- cross-bridges--with actin sites. The cross-bridges act independently of each other and can be modeled as passive instantaneous elastic elements in series with active force generators.

The first is the sliding-filament theory, now universally accepted; the second is the cross-bridge theory, now widely accepted.

There are constraints upon any model of muscle contraction. Hill [HIL 74, HIL 75, HIL 77] has presented a theoretical formalism combining statistical mechanics, thermodynamics, and chemical kinetics to delineate the class of models which can describe energy transduction in muscle. We have been able to develop a theory which yields the same results, but rests on a solid mathematical foundation. In [PRO 81], we have presented a detailed treatment of a biochemically- and structurally-based model of muscle contraction (which contains proofs of many of the theorems quoted here).

We consider a muscle fibre as a collection of independent units. A unit is one cross-bridge associated with a one-dimensional periodic array of actin sites. A cross-bridge can bind to only a group of 2 sites per period. The site separation s = 55 Å, the period λ = 385 Å, and the separation between the cross-bridge joint and the actin site ℓ = 150 Å. As in Figure 5.3.1, we let x be the position of a point on the actin filament equidistant from the nearest 2 sites relative to the ψ = 90° position of a cross-bridge (which we shall henceforth call the actin position), $x \in [-\lambda/2, \lambda/2)$. Due to the lack of register between the myosin and actin periods, the actin positions x should be distributed uniformly over a 385 Å interval.

We shall characterize our structural unit of a cross-bridge and its associated array of actin sites by the cross-bridge conformation (the biochemical state of the cross-bridge) and the actin position (the position of the nearest set of actin sites relative to the cross-bridge). In a contraction, as in Figure 5.3.2, we can imagine a set of actin sites appearing to the right of an unattached cross-bridge. The cross-bridge will execute a random walk through a sequence of biochemical states. In any attached state, the cross-bridge will assume an angle determined by the actin position and the site to which it is attached. However, the angle at which a cross-bridge will be found attached with greatest probability depends upon the biochemical state.

Consider a single structural unit of a cross-bridge and its associated array of actin sites. Let

$$T = [t_0, \infty)$$

be the time set of interest. Let the conformation of the cross-bridge at time t be given by

$$\xi(t) \in E \triangleq I \cup I' \cup I'' ,$$

where $I \triangleq \{1,2,\ldots,9\}$ is the set of unattached states, $I' \triangleq \{1',2',\ldots,9'\}$ is the set of actin-site-1-attached states, and $I'' \triangleq \{1'',2'',\ldots,9''\}$ is the set of actin-site-2-attached states. Let N be the cardinality of the state space E (which, for the *in vivo* reaction mechanism of [PRO 81], is 27). Let the position (relative to the cross-bridge) at time t of the set of actin sites which were nearest the cross-bridge at time t_0 be given by

$$\chi(t) \in \mathbb{R} .$$

Define an equivalence relation $\equiv$ on $\mathbb{R}$ by

$$x \equiv y \text{ if and only if } |x-y| = n\lambda, \; n=0,1,2,\ldots ,$$

where λ = 385 Å is the actin period. The set of equivalence classes of $\mathbb{R}$ modulo $\equiv$ is

$$\mathbb{R}/\equiv \; = \Lambda \triangleq [-\lambda/2, \lambda/2) .$$

For any $x \in \mathbb{R}$, we shall denote the equivalence class containing x by [x] and consider [x] as an element of Λ. Then the equivalence class $[\chi(t)] \in \Lambda$ is the position (relative to the cross-bridge) of the set of actin sites nearest the cross-bridge at time t--the actin position at time t.

Given the initial actin position x_0 at time t_0, the positional process $\{\chi(t), t \in T\}$ is deterministic, i.e., it is described by the ordinary differential equation

$$\frac{dx(t)}{dt} = -v(t); \qquad x(t_0) = x_0,$$

where v(t) is the shortening velocity. Define the density function p(t,x) by

$$\int_{x_1}^{x_2} p(t,x)dx = \Pr\{x_1 \leq [\chi(t)] \leq x_2\} ,$$

for $x_1, x_2 \in \Lambda$, $x_1 < x_2$. Then

Lemma 5.3.17.

The density function p(t,x) satisfies

$$\frac{\partial p(t,x)}{\partial t} = v(t) \frac{\partial p(t,x)}{\partial x} ; \quad p(t_0,x) = p_0(x) . \blacksquare$$

There is thus no diffusion term in the positional process. Let

$$P_x = \{p: p(x) \geq 0 \text{ for all } x \in \Lambda, \int_{-\lambda/2}^{\lambda/2} p(x)dx = 1\} .$$

Then

Theorem 5.3.18.

Given (5.1) with $p_0 \in P_x$, the solution $p(t,x) \in P_x$, $t \geq t_0$. $\blacksquare$

Given the initial actin position x_0, or equivalently, given the positional process $\{\chi(t), t \in T\}$, we shall assume the conformational process $\{\xi(t), t \in T\}$ is a Markov step process characterized by a set of N(N-1) transition rate functions (a_{ij} the rate for the transition $j \to i$)

$$\{a_{ij}(t) = a_{ij}(\chi(t)) : i,j \in E,\ i \neq j,\ t \in T\}$$

associated with the directed graph for the state space E (Figure 5.3.3). Assume intercommunicating states with $a_{ij} \neq 0$ if and only if $a_{ji} \neq 0$. Let

$$\underline{A} \triangleq (a_{ij}) \quad ,$$

where $a_{jj} = -\sum_{i \neq j} a_{ij}$. The transition rates will have different values at different temperatures and at different concentrations of ATP, ADP, and P_i. In active muscle, on the time-scale of cross-bridge cycling, nucleotide concentrations are maintained constant by external reactions. Thus to describe the thermodynamic properties of muscle we can consider the transition rates at a fixed temperature and fixed concentrations of ADP and P_i, but parameterized by the time-independent ATP concentration $\alpha \in \mathbb{R}_+$. So

$$a_{ij} = a_{ij}(\alpha,t) = a_{ij}(\alpha, \chi(t))$$

$$\underline{A} = \underline{A}(\alpha,t) = \underline{A}(\alpha, \chi(t)) \quad .$$

Specifically, referring to the directed graph for E,

$$a_{ij}(\alpha,\chi(t)) \begin{cases} \kappa_{ji}(\chi(t))\alpha, & (i,j) \in \{(2,1),(2',1'), \\ & \quad (2'',1'')\} \\ \kappa_{ji}(\chi(t))[ADP], & (i,j) \in \{(9,1),(9',1'), \\ & \quad (9'',1'')\} \\ \kappa_{ji}(\chi(t))[P_i], & (i,j) \in \{(6,7),(6',7'), \\ & \quad (6'',7'')\} \\ \kappa_{ji}(\chi(t)), & \text{otherwise} \ , \end{cases} \tag{5.3.4}$$

where κ_{ji} is the rate constant for the transition $j \to i$.

Muscle structure dictates that a cross-bridge can bind to only the nearest set of actin sites. With finite detachment rate constants, there will be a nonzero probability that a cross-bridge will not detach from an actin site by the time it has passed into the range of the next set of actin sites. To ensure detachment with probability one and avoid introducing singularities in the parameters, we let

$$\tau_{x_0} = \{t : [\chi(t)] = -\lambda/2\} = \{t : [x_0 - \int_{t_0}^{t} v(\tau)d\tau] = -\lambda/2\}$$

and require boundary conditions at any $t \in \tau_{x_0}$. The set τ_{x_0} is the set of times at which the cross-bridge passes from the range of one set of actin sites to the next.

By definition, detachment with probability one is the occurrence with probability one of the transitions $i' \to i$ and $i'' \to i$. Thus we require

$$p_{ij}^{+} \triangleq \Pr\{\xi(t^{+}) = i \,|\, \xi(t^{-}) = j\},\ t \in \tau_{x_0},$$

to be given by

$$p_{ij}^{+} \triangleq \begin{cases} 1 \ , & i \in I,\ j \in \{i,i',i''\} \\ 0 \ , & \text{otherwise} \ . \end{cases}$$

Let

$$p_i(t|x_0) = p(t,i|x_0) = \Pr\{\xi(t) = i \,|\, \chi(t_0) = x_0\}.$$

Then $\underline{p}(t|x_0) \triangleq (p_i(t|x_0))$ satisfies Kolmogorov's forward equation

$$\frac{d\underline{p}(t|x_0)}{dt} = \underline{A}(\alpha, [x_0 - \int_{t_0}^{t} v(\tau)d\tau])\underline{p}(t|x_0) \ ,$$

$$t \in T \setminus \tau_{x_0} \ ; \ \underline{p}(t_0|x_0) = \underline{p}_0 \tag{5.3.5a}$$

and the boundary conditions

$$\underline{p}(t^{+}|x_0) = \underline{P}^{+}\underline{p}(t^{-}|x_0), \quad t \in \tau_{x_0}, \tag{5.3.5b}$$

where

$$\underline{P}^{+} = (p_{ij}^{+}) \quad .$$

Let

$$x(t) = [x_0 - \int_{t_0}^{t} v(\tau)d\tau]$$

and

$$p_i(t|x(t)) = p(t,i|x(t)) = \Pr\{\xi(t) = i \,|\, [\chi(t)] = x(t)\}$$

$$= \Pr\{\xi(t) = i \,|\, \chi(t_0) = x_0\}$$

$$= \Pr\{\xi(t) = i \,|\, \chi(t_0) = [x(t) + \int_{t_0}^{t} v(\tau)d\tau]\}$$

$$= p_i(t|x_0)\Big|_{x_0 = [x(t) + \int_{t_0}^{t} v(\tau)d\tau]} \quad .$$

Then $\underline{p}(t|x(t)) \triangleq (p_i(t|x(t)))$ also satisfies

$$\frac{d\underline{p}(t|x(t))}{dt} = \underline{A}(\alpha,x(t))\underline{p}(t|x(t)),\ t \in \tau_{x_0} \ ;$$

$$\underline{p}(t_0|x(t_0)) = \underline{p}_0$$

and the boundary conditions

$$\underline{p}(t^{+}|x(t)) = \underline{P}^{+}\underline{p}(t^{-}|x(t)), \quad t \in \tau_{x_0} \ .$$

Let

$$P_E = \{\underline{p} : p_i \geq 0 \text{ for all } i \in E,\ \sum_i p_i = 1\}.$$

Then

Theorem 5.3.19.

Given (5.3.5) with $\underline{p}_0 \in P_E$, the solution $\underline{p}(t|x_0) = \underline{p}(t|x(t)) \in P_E$, $t \geq t_0$. ■

Note that

$$\frac{d\underline{p}(t|x)}{dt} = \frac{\partial \underline{p}(t|x)}{\partial t} + \frac{\partial \underline{p}(t|x)}{\partial x}\frac{dx}{dt}$$

$$= \frac{\partial \underline{p}(t|x)}{\partial t} - v(t)\frac{\partial \underline{p}(t|x)}{\partial x} .$$

So for $x \in \Lambda$ $\underline{p}(t|x)$ satisfies

$$\frac{\partial \underline{p}(t|x)}{\partial t} = v(t)\frac{\partial \underline{p}(t|x)}{\partial x} + \underline{A}(\alpha,x)\underline{p}(t|x);$$

$$\underline{p}(t_0|x) = \underline{p}_0.$$

The boundary conditions (5.3.5b) become

$$p_i(t,-\lambda/2) = p_i(t,\lambda/2) = 0, \quad i \in I'UI'' .$$

Note that we can distinguish between $p_i(t|x) = \Pr\{\xi(t)=i|\chi(t)=x\}$ and $p_i(t|x_0) = \Pr\{\xi(t)=i|\chi(t_0)=x_0\}$ by writing

$$\underline{p}(t|x) = \underline{p}(t|x_0)\Big|_{x_0=[x+\int_{t_0}^{t} v(\tau)d\tau]} .$$

$$\underline{p}(t|x_0) = \underline{p}(t|x)\Big|_{x=[x_0-\int_{t_0}^{t} v(\tau)d\tau]} .$$

We now consider the joint process $\{\xi(t), \chi(t), t \in T\}$.

Proposition 5.3.20.

The joint process $\{\xi(t),\chi(t),t \in T\}$ is Markov. ■

Define the density function $p_i(t,x) = p(t,i,x)$ for the joint process $\{\xi(t), \chi(t), t \in T\}$ by

$$\int_{x_1}^{x_2} p_i(t,x)dx = \Pr\{\xi(t) = i,\ x_1 \leq [\chi(t)] \leq x_2\} .$$

Then

Lemma 5.3.21.

The density function $\underline{p}(t,x) \triangleq (p_i(t,x))$ satisfies

$$\frac{\partial \underline{p}(t,x)}{\partial t} = v(t)\frac{\partial \underline{p}(t,x)}{\partial x} + \underline{A}(\alpha,x)\underline{p}(t,x) \quad ;$$

$$\underline{p}(t_0,x) = \underline{p}_0(x) \triangleq \underline{p}_0 p_0(x) , \qquad (5.3.6)$$

$$p_i(t,-\lambda/2) = p_i(t,\lambda/2) = 0, \quad i \in I'UI'' . \quad ■$$

Let

$$P = \{\underline{p} : p_i(x) \geq 0 \text{ for all } i \in E,\ x \in \Lambda,$$

$$\sum_i \int_{-\lambda/2}^{\lambda/2} p_i(x)dx = 1\} .$$

Then

Theorem 5.3.22.

Given (5.3.6) with $\underline{p}_0 p_0(x) \in P$, $\underline{p}_0 \in P_E$, and $p_0(x) \in P_x$, the solution $\underline{p}(t,x) \in P$, $t \geq t_0$. ■

A single structural unit is characterized at time t by the cross-bridge conformation $\xi(t)$ and the actin position $\chi(t)$. The joint process $\{\xi(t), \chi(t), t \in T\}$ is Markov. Assuming a lack of coupling between structural units, a muscle fibre becomes an ensemble of independent copies of the process $\{\xi(t), \chi(t), t \in T\}$ and the properties of a structural unit are the properties of the ensemble.

A muscle fibre is a thermodynamic system. The system Σ is thus defined as (5.3.6). The state of the system Σ at time t is the vector probability density function $\underline{p}(t,x)$. The state space is the space P of all density functions of the process. We now show the system Σ has the expected thermodynamic properties.

Definition 5.3.23.

We shall call the unique solution of

$$v^s\frac{\partial \underline{p}(x)}{\partial x} + \underline{A}(\alpha,x)\underline{p}(x) = \underline{0}; \quad p_i(-\frac{\lambda}{2}) = p_i(\frac{\lambda}{2}) = 0,$$

$$i \in I'UI'' ,$$

where v^s and α are independent of time, a steady state. We shall say the system Σ evolves to the steady state $\underline{p}^s(\cdot)$ if, given the state trajectory $\{\underline{p}(t,\cdot), t \in T\}$,

$$\lim_{t\to\infty} \underline{p}(t,\cdot) = \underline{p}^s(\cdot) . \quad ■$$

Proposition 5.3.24.

If the system Σ evolves to a steady state, then $p_0(\cdot)$ is the uniform PDF. ■

If $p_0(\cdot)$ were not a uniform PDF, then $p(t,x) = p_0([x+v^s(t-t_0)])$ would be periodic with period $\frac{\lambda}{v^s} \sim 10^{-1}$ sec, for v^s on the order of a velocity of contraction. The force (5.3.10) exerted by a muscle fibre in steady state during a constant velocity contraction shows no such oscillations [JUL 75]. Furthermore, the lack of register between the myosin and actin periods suggests the function $p(t,x)$ describing the distribution at time t of actin positions x is the uniform PDF. We shall therefore

take $p_0(x)$ as the uniform PDF and show that the system Σ evolves to a steady state when $v=v^s$ and α are time-independent.

With $p(t,x)$ the uniform PDF,

$$\underline{p}(t,x) = \underline{p}(t|x)p(t,x)$$
$$= \frac{1}{\lambda}\,\underline{p}(t|x)$$
$$= \frac{1}{\lambda}\,\underline{p}(t|x_0)\Big|_{x_0=[x(t)+\int_{t_0}^{t} v(\tau)d\tau]} .$$

With $v=v^s$ time-independent,

$$\underline{A}(\alpha,[x_0 - \int_{t_0}^{t} v(\tau)d\tau]) = \underline{A}(\alpha,[x_0-v^s(t-t_0)])$$
$$= \underline{A}(\alpha,[x_o-v_s((t+\frac{\lambda}{v^s})-t_0)]).$$

Theorem 5.3.25.

For $v = v^s \neq 0$ time-independent,

$$\lim_{t\to\infty} \underline{p}(t|x_0) = \underline{p}^s(t|x_0) = \underline{p}^s(t+\frac{\lambda}{v^s}\,|x_0)$$

and $\underline{p}^s(t|x_0)$ is independent of the initial distribution $\underline{p}(t_0|x_0)$. ■

Theorem 5.3.26.

For $v=v^s$ time-independent and $p(t,x)$ the uniform PDF,

$$\lim_{t\to\infty} \underline{p}(t,x) = \underline{p}^s(x) = \frac{1}{\lambda}\,\underline{p}^s(t_0|x_0)\Big|_{x_0=x} .$$

and $\underline{p}^s(x)$ is independent of the initial distribution $\underline{p}_0(x)$ and the time t. ■

In a steady state, the average behaviour of a single cross-bridge over the time period $\frac{\lambda}{v^s}$ yields the time-independent state of the system.

Corollary 5.3.27.

The limiting distribution

$$\underline{p}^s(t|x_0) = \lambda\underline{p}^s([x_0-v^s(t-t_0)]) \quad . \blacksquare$$

We can distinguish the equilibrium state as a particular steady state, as follows.

Definition 5.3.28.

The fluxes of the system Σ are the probabilistic fluxes

$$J_{ij}(t,x) = a_{ij}(\alpha,x)p_j(t,x) - a_{ji}(\alpha,x)p_i(t,x),$$

$$i,\ j\varepsilon E,\ x\varepsilon\Lambda$$

and the deterministic flux $v(t)$. ■

The fluxes J_{ij} are elements of the matrix $\underline{J} = (J_{ij})$. Note that $J_{ji} = -J_{ij}$. We can rewrite (5.3.6) as

$$\frac{\partial p_i(t,x)}{\partial t} = v(t)\,\frac{\partial p_i(t,x)}{\partial x} + \sum_j J_{ij}(t,x),\ i\varepsilon E.$$

The steady state condition $v^s\,\frac{\partial p_i}{\partial x} + \sum_j J_{ij} = 0$ or $\sum_j J_{ij} = 0$ for $v^s = 0$ does not exclude coupling between fluxes and rotational fluxes. When the fluxes are identically zero, a steady state cannot be maintained by any cyclic mechanism. We distinguish such a steady state as an equilibrium state.

Definition 5.3.29.

The equilibrium state is that steady state for which $v = v^s = 0$ and $\underline{J} = \underline{0}$. We shall use the notation $\underline{p}^e$ for the probability density function of the equilibrium state. ■

The thermodynamic condition of detailed balance is equivalent to equilibrium by definition. The thermodynamic condition of microscopic reversibility is that of time-reversal invariance, as follows.

Definition 5.3.30.

We shall say the joint process $\{\xi(t),\ \chi(t),\ t\varepsilon T\}$ is invariant under time reversal if, for all $t,\ s\ \varepsilon\ T$, $i,\ j\varepsilon E$, $x,\ y\varepsilon\Lambda$,

$$p(t,i,x|s,j,y) = p(s,i,x|t,j,y),$$

where $p(t,i,x|s,j,y)$ is defined by

$$\int_{x_1}^{x_2} p(t,i,x|s,j,y)dx = \Pr\{\xi(t)=i,\ x_1\leq\chi(t)\leq x_2\,|\,\xi(s)=j,\chi(s)=y\}$$

for $x_1,x_2\varepsilon\Lambda,\ x_1<x_2$. ■

Theorem 5.3.31.

The system Σ is in the equilibrium state if and only if the process $\{\xi(t),\ \chi(t),\ t\varepsilon T\}$ is invariant under time reversal. ■

As is to be expected from kinetic theory, we can show

Theorem 5.3.32.

A steady state is the equilibrium state if and only if $v=0$ and $\sum_{\text{closed path}} \ln\frac{a_{ij}(\alpha,x)}{a_{ji}(\alpha,x)} = 0$ for all $x\varepsilon\Lambda$ and for all closed paths of the directed graph associated with the state space E. ■

The equilibrium constant for the reaction $i\leftrightarrow j$

$$K_{ij} = \frac{\kappa_{ij}}{\kappa_{ji}} .$$

Corollary 5.3.33.

A steady state of (5.3.6), with $\underline{A}$ given by (5.3.4), is an equilibrium state if and only if $v=0$ and for

all $x \in \Lambda$,

$$\frac{K_{ij}K_{jj'}}{K_{ii'}K_{i'j'}} = 1 \quad \text{for the closed paths} \qquad (5.3.7a)$$

$$\begin{array}{ccc} i & \leftrightarrow & j \\ \updownarrow & & \updownarrow \\ i' & \leftrightarrow & j' \end{array} \qquad (5.3.7b)$$

$$\frac{K_{ij}K_{jj''}}{K_{ii''}K_{i''j''}} = 1 \quad \text{for the closed paths} \qquad (5.3.8a)$$

$$\begin{array}{ccc} i & \leftrightarrow & j \\ \updownarrow & & \updownarrow \\ i'' & \leftrightarrow & j'' \end{array} \quad , \qquad (5.3.8b)$$

and

$$\alpha = \alpha_e \triangleq \frac{[ADP][P_i]}{K_{HYD}}$$

where $K_{HYD} = K_{1'1''}$. ■

In equilibrium, we have the relation

$$J_{ij}(x) = 0 = a_{ij}(x)p_j^e(x) - a_{ji}(x)p_i^e(x) \quad .$$

So the equilibrium state $\underline{p}^e(x)$ is determined by

$$\frac{p_i^e(x)}{p_j^e(x)} = \frac{a_{ij}(\alpha^e,x)}{a_{ji}(\alpha^e,x)} \quad .$$

$$\sum_i p_i^e(x) = \frac{1}{\lambda} \quad .$$

Note that Theorem 5.3.32 guarantees

$$\prod_{\text{closed path}} \frac{a_{ij}(\alpha^e,x)}{a_{ji}(\alpha^e,x)} = 1 = \prod_{\text{closed path}} \frac{p_i^e(x)}{p_j^e(x)} \quad .$$

In order for the system Σ to evolve to the equilibrium state at some ATP concentration and zero shortening velocity, we must choose the x-dependence of the parameters κ_{ij} such that (5.3.7) and (5.3.8) are satisfied for all $x \in \Lambda$. As physical observation confirms the existence of the equilibrium state, i.e., a steady state of zero force and zero flux (Proposition 5.3.36), we shall so choose the parameters κ_{ij}. Then by defining $u \triangleq \ln \frac{\alpha}{\alpha_e}$, the system Σ will evolve to the equilibrium state $\underline{p}^e$ if and only if u=0 and v=0. We shall consider the system as controlled by inputs u and v.

Given that the uncontrolled system evolves to the equilibrium state, one might wish to know how far the system Σ is from equilibrium. The state functions of thermodynamics can be thought of as distance measures on $\mathcal{P}$, the space of all density functions of the process $\{\xi(t), \chi(t), t \in T\}$. (We choose a temperature scale such that Boltzmann's constant k is unity and by considering the system Σ at a constant temperature take all quantities in units of entropy.) We define the free energy G of the system Σ as

$$G(t) = \sum_i \int_{-\lambda/2}^{\lambda/2} p_i(t,x) \ln \frac{p_i(t,x)}{p_i^e(x)} dx \quad , \qquad (5.3.9)$$

the entropy S of the system Σ as

$$S(t) = -\sum_i \int_{-\lambda/2}^{\lambda/2} p_i(t,x) \ln \frac{p_i(t,x)}{p_i^e(x)} dx \quad ,$$

and the internal energy U of the system Σ as

$$U(t) = -\sum_i \int_{-\lambda/2}^{\lambda/2} p_i(t,x) \ln p_i^e(x) dx \quad .$$

The input u is a force -- the chemical potential of ATP--and the input v is a flux--the shortening velocity.

Definition 5.3.34.

The flux $\bar{J}$ conjugate to the force u is defined as

$$\bar{J} \triangleq \frac{\partial}{\partial u} \left(\frac{d}{dt} G(t)\right)\Big|_{\underline{J}=\underline{0},\ v=0}$$

and the force $\bar{F}$ conjugate to the flux v is defined as

$$\bar{F} \triangleq -\frac{\partial}{\partial v} \left(\frac{d}{dt} G(t)\right)\Big|_{\underline{J}=\underline{0},\ v=0} \quad . \quad \blacksquare \qquad (5.3.10)$$

From (5.3.9) we compute the rate of change of the free energy as

$$\frac{d}{dt}G(t) = \sum_i \int_{-\lambda/2}^{\lambda/2} \frac{\partial}{\partial t} p_i(t,x) \ln \frac{p_i(t,x)}{p_i^e(x)} dx$$

$$+ \sum_i \int_{-\lambda/2}^{\lambda/2} \frac{\partial p_i(t,x)}{\partial t} dx$$

$$= \sum_i \int_{-\lambda/2}^{\lambda/2} \Big[v(t) \frac{\partial}{\partial x} p_i(t,x)$$

$$+ \sum_j J_{ij}(t,x) \Big] \ln \frac{p_i(t,x)}{p_i^e(x)} dx \quad .$$

$$= \sum_i \sum_{j<i} \int_{-\lambda/2}^{\lambda/2} J_{ij}(t,x)$$

$$\cdot \ln \frac{a_{ij}(\alpha,x)/a_{ji}(\alpha,x)}{a_{ij}(\alpha^e,x)/a_{ji}(\alpha^e,x)} dx$$

$$-v(t)\sum_i \int_{-\lambda/2}^{\lambda/2} p_i(t,x)\,\frac{\partial}{\partial x}\ln\frac{p_i(t,x)}{p_i^e(x)}dx$$

$$+\sum_i\sum_{j<i}\int_{-\lambda/2}^{\lambda/2} J_{ij}(t,x)\ln\left[1+\frac{J_{ij}(t,x)}{a_{ji}(\alpha,x)p_i(t,x)}\right]dx$$

$$= u\int_{-\lambda/2}^{\lambda/2}(J_{21}(t,x)+J_{2'1'}(t,x)+J_{2''1''}(t,x))dx$$

$$-v(t)\sum_i\int_{-\lambda/2}^{\lambda/2}p_i(t,x)\cdot\left[-\frac{\partial}{\partial x}\ln p_i^e(x)\right]dx$$

$$-\sum_i\sum_{j<i}\int_{-\lambda/2}^{\lambda/2}J_{ij}(t,x)\ln\left[1+\frac{J_{ij}(t,x)}{a_{ji}(\alpha,x)p_i(t,x)}\right]dx.$$

From Definition 5.3.34 we obtain

$$\bar{J}(t) = \int_{-\lambda/2}^{\lambda/2}(J_{21}(t,x)+J_{2'1'}(t,x)+J_{2''1''}(t,x))dx$$

$$\bar{F}(t) = \sum_i\int_{-\lambda/2}^{\lambda/2}p_i(t,x)\cdot\left[-\frac{\partial}{\partial x}\ln p_i^e(x)\right]dx\ .$$

As

$\bar{J} = \frac{d}{dt}E\{$(Total transitions from 1→2, 1'→2', 1''→2'')

$-$(Total transitions from 2→1, 2'→1', 2''→1'')$\}$,

$\bar{J}$ is the mean ATP flux, which can be observed as an output of the system Σ [PRO 81]. As

$$\bar{F}(t) = \sum_i\int_{-\lambda/2}^{\lambda/2}p_i(t,x)\cdot\left[-\frac{\partial}{\partial x}\ln p_i^e(x)\,dx\right]$$

$$= \sum_i\int_{-\lambda/2}^{\lambda/2}p_i(t,x)\left[\frac{\partial}{\partial x}\ln\frac{p_i(t,x)}{p_i^e(x)}\right]dx$$

$$= E\left\{\frac{\partial}{\partial x}\ln\frac{p_{\xi(t)}(t,x)}{p^e_{\xi(t)}(x)}\bigg|_{x=\chi(t)}\right\}$$

and

$$G(t) = \sum_i\int_{-\lambda/2}^{\lambda/2}p_i(t,x)\ln\frac{p_i(t,x)}{p_i^e(x)}\,dx$$

$$= E\left\{\ln\left\{\frac{p_{\xi(t)}(t,\chi(t))}{p^e_{\xi(t)}(\chi(t))}\right\}\right\},$$

we can consider $\ln\frac{p_{i(t,x)}}{p_i^e(t,x)}$ as the free energy of a cross-bridge in state i and at position x and

$$\frac{\partial}{\partial y}\ln\frac{p_i(t,y)}{p_i^e(y)}\bigg|_{y=x}$$ as

the force exerted on the actin filament by a cross-bridge in state i and at position x. Then

$$\bar{F} = E\left\{\frac{\partial}{\partial x}\ln\frac{p_{\xi(t)}(t,x)}{p^e_{\xi(t)}(x)}\bigg|_{x=\chi(t)}\right\}$$

becomes the mean force exerted on the actin filament, which can be observed as an output of the system Σ [PRO 81]. Thus the mean ATP flux $\bar{J}$ and the mean force $\bar{F}$ are outputs of the system Σ, observable as expected values.

Definition 5.3.35.

The work flow W is given by

$$W = -\bar{J}u + \bar{F}v\ . \quad \blacksquare$$

Define the external forces F_{ij} as

$$F_{ij}(x) = -\ln\frac{a_{ij}(\alpha,x)}{a_{ji}(\alpha,x)}\ ,\ i,\ j\in E,\ j<i,$$

the internal forces X_{ij} as

$$X_{ij}(t,x) = \ln\left[1+\frac{J_{ij}(t,x)}{a_{ji}(\alpha,x)p_i(t,x)}\right],\ i,\ j\in E,\ j<i,$$

the heat flow Q as

$$Q = \sum_i\sum_{j<i}\int_{-\lambda/2}^{\lambda/2}J_{ij}(t,x)F_{ij}(x)dx\ ,$$

and the entropy production D as

$$D = \sum_i\sum_{j<i}\int_{-\lambda/2}^{\lambda/2}J_{ij}(t,x)X_{ij}(t,x)dx\ .$$

The entropy production $D\geq 0$, with equality if and only if the system Σ is at equilibrium. The first and second laws of theromodynamics follow:

$$\frac{d}{dt}U = Q - W$$

$$\frac{d}{dt}S = Q + D \geq Q$$

$$\frac{d}{dt}G = -W - D \leq -W\ .$$

In a steady state, the state functions U, S, and G are time-independent. So in steady state $W = Q = -D \leq 0$, with equality if and only if the steady state is the equilibrium state. As the heat flow

Q and the entropy production D are independent of the equilibrium state, the work flow W is independent of the equilibrium state in steady state. In steady state we can write

$$\bar{J}u = \bar{F}v - Q$$

$$= \bar{F}v + D \geq \bar{F}v ,$$

i.e., the free energy of ATP (u) is both dissipated as heat ($-Q = D$) and converted into external work ($\bar{F}v$). The efficiency of muscle is the efficiency of the transduction of chemical work flow from the reservoir controlling the ATP concentration to the system ($\bar{J}u$) into mechanical work flow from the system to a weight in the surroundings ($\bar{F}v$). We thus define the efficiency η as the ratio of the mechanical work flow to the sum of the mechanical work flow and the heat flow from the system to the surroundings ($-Q$). So

$$\eta = \frac{\bar{F}v}{\bar{F}v - Q} = \frac{\bar{F}v}{\bar{J}u} \leq 1 \text{ in steady state} .$$

Although $\bar{J}$ and $\bar{F}$ are averaged quantities, we can still show that they both vanish only in a steady state that is the equilibrium state.

Proposition 5.3.36.

In a steady state, $\begin{bmatrix} \bar{J} \\ \bar{F} \end{bmatrix} = \underline{0}$ if and only if the steady state is the equilibrium state. ■

Using Corollary 5.3.27, we can transform the expressions for the mean ATP flux and mean force to integrals over the time period λ/v^s.

Proposition 5.3.37.

In a steady state, for any $\tau \in T$

$$\bar{J} = \frac{v^s}{\lambda} \sum_{(i,j)\in\{(2,1),\,(2'1'),(2'',1'')\}} \int_{\tau}^{\tau+(\lambda/v^s)} (a_{ij}(\alpha,[x_0 - v^s(t-t_0)]) \cdot p_j^s(t|x_0) - a_{ji}(\alpha,[x_0 - v^s(t-t_0)]) p_i^s(t|x_0)) dt$$

$$\bar{F} = \frac{v^s}{\lambda} \sum_i \int_{\tau}^{\tau+(\lambda/v^s)} p_i^s(t|x_0) \cdot \left[-\frac{\partial}{\partial x} \ln p_i^e(x) \Big|_{x=[x_0 - v^s(t-t_0)]} \right] dt . \quad ■$$

Close to equilibrium, we can show that the linearized input-output relationship of the system Σ is reciprocal -- thus the model statisfies the reciprocal relations of thermodynamics.

Definition 5.3.38.

We shall say the system linearized about a steady state $\underline{p}^s(\cdot)$ is reciprocal if the operator

$$H^s : (\delta v, \delta u) \to (\delta\bar{J}, \delta\bar{F})$$

is self-adjoint, where $(\delta v, \delta u)$ and $(\delta\bar{J}, \delta\bar{F})$ are related by the dynamical system

$$\frac{\partial \delta\underline{p}}{\partial t} = A\delta\underline{p} + \underline{B}_1 \delta u + \underline{B}_2 \delta v ; \; \delta\underline{p}(t_0) = \underline{0} \qquad (5.3.11a)$$

$$\delta\bar{J} = C_1 \delta\underline{p} + D\delta u \qquad (5.3.11b)$$

$$\delta\bar{F} = C_2 \delta\underline{p} , \qquad (5.3.11c)$$

where $\delta\underline{p}$ is defined as the solution of (5.3.11a) and the parameters of the system are defined by

$$A\delta\underline{p} = v^s \frac{\partial \delta\underline{p}(t,x)}{\partial x} + \underline{A}(\alpha,x)\delta\underline{p}(t,x)$$

$$\underline{B}_1 = (b_i^1)$$

$$b_i^1(x) = \begin{cases} -\kappa_{ij}(x)\alpha p_i^s(x), & (i,j)\in\{(1,2),(1'2'),(1'',2'')\} \\ \kappa_{ji}(x)\alpha p_j^s(x), & (i,j)\in\{(2,1),(2'1'),(2'',1'')\} \\ 0, & \text{otherwise} \end{cases}$$

$$\underline{B}_2 = \frac{\partial p^s(x)}{\partial x}$$

$$C_1\delta\underline{p} = \sum_{(i,j)\in\{(2,1),\,(2',1')(2'',1'')\}} \int_{-\lambda/2}^{\lambda/2} (\kappa_{ji}(x)\alpha\delta p_j(t,x) - \kappa_{ij}(x)\delta p_i(t,x)dx$$

$$C_2\delta\underline{p} = \sum_i \int_{-\lambda/2}^{\lambda/2} \delta p_i(t,x) \cdot \left[-\frac{\partial}{\partial x} \ln p_i^e(x) \; dx \right]$$

$$D = \alpha \int_{-\lambda/2}^{\lambda/2} (\kappa_{12}(x)p_1^s(x) + \kappa_{1'2'}(x)p_{1'}^s(x) + \kappa_{1''2''}(x) \cdot p_{1''}^s(x))dx . \quad ■$$

Let the operators

$$B = (\underline{B}_1 \; \underline{B}_2) : \mathbb{R}^2 \to L^2(\Lambda, \mathbb{R}^N)$$

$$C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} : L^2(\Lambda, \mathbb{R}^N) \to \mathbb{R}^2 .$$

Then

$$B^* : L^2(\Lambda, \mathbb{R}^N) \to \mathbb{R}^2$$

$$B^*\underline{\phi} = \begin{bmatrix} \sum_i \int_{-\lambda/2}^{\lambda/2} b_i^1(x)\phi_i(x)dx \\ \sum_i \int_{-\lambda/2}^{\lambda/2} \frac{\partial p_i^s(x)}{\partial s}\phi_i(x)dx \end{bmatrix}, \quad \underline{\phi} \in L^2(\Lambda, \mathbb{R}^N)$$

$$C^* = (C_1^* \; C_2^*) : \mathbb{R}^2 \to L^2(\Lambda, \mathbb{R}^N)$$

$$C_1^* = \underline{C}_1 = (c_i^1)$$

$$c_i^1(x) = \begin{cases} \kappa_{ij}(x)\alpha, & (i,j) \in \{(1,2),(1',2'),(1'',2'')\} \\ -\kappa_{ij}(x), & (i,j) \in \{(2,1),(2',1'),(2'',1'')\} \\ 0, & \text{otherwise} \end{cases}$$

$$C_2^* = \underline{C}_2 = \left[-\frac{\partial}{\partial x} \ell n p_i^e(x)\right] \quad .$$

Lemma 5.3.39.

If the linearized system (5.3.11) is a continuously controllable and observable realization on $L^2(\Lambda, \mathbb{R}^{N-1})$ of the input-output relation between $(\delta u, \delta v)$ and $\delta \bar{J}, \delta \bar{F})$, then reciprocity implies the existence of an operation T such that

$$AT = TA^*$$

$$CT = B^* \quad . \quad \blacksquare$$

Theorem 5.3.40.

The system Σ linearized about the equilibrium state is reciprocal. Furthermore, if the linearized system (5.3.11) is a continuously controllable and observable realization on $L^2(\Lambda, \mathbb{R}^{N-1})$ of H^s, then the system Σ linearized about a steady state is reciprocal if and only if the steady state is the equilibrium state.

Proof:

($\Longleftarrow$)

For the system Σ linearized about an equilibrium state, we find

$$b_i'(x) = -p_i^e(x)\, c_i'(x)$$

$$\frac{\partial p_i^e(x)}{\partial x} = -p_i^e(x)\left(-\frac{\partial}{\partial x} \ln p_i^e(x)\right) \quad .$$

So

$$B = -\mathrm{diag}(p_i^e)C^*$$

$$C = B^*\left[-\mathrm{diag}\left[\frac{1}{p_i^e}\right]\right] .$$

Also by calculation

$$A = \underline{A}(\alpha^e, x) = \mathrm{diag}(p_i^e(x))\underline{A}'(\alpha^e, x)\mathrm{diag}\left[\frac{1}{p_i^e(x)}\right]$$

$$= \mathrm{diag}(p_i^e)A^*\mathrm{diag}\left[\frac{1}{p_i^e}\right] .$$

Hence, for the system Σ linearized about an equilibrium state

$$H^e(t) = Ce^{At}B + \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}\delta(t)$$

$$= B^*\left[-\mathrm{diag}\left[\frac{1}{p_i^e}\right]\right]\mathrm{diag}(p_i^e)e^{A^*t}\mathrm{diag}\left[\frac{1}{p_i^e}\right]$$

$$\cdot\left[-\mathrm{diag}(p_i^e)\right]C^* + \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}\delta(t)$$

$$= B^*e^{A^*t}C^* + \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}\delta(t)$$

$$= H^{e^*}(t) \; ,$$

where $H^e(t)$ is defined for $t \in \mathbb{R}_+$.

($\Longrightarrow$)

We prove sufficiency, given the linearized system (5.3.11) is a continuously controllable and observable realization on $L^2(\Lambda, \mathbb{R}^{N-1})$ of H^s, by proving its contrapositive, i.e., we prove that if the steady state is not an equilibrium state, then the system Σ linearized about the steady state is not reciprocal. Using Lemma 5.3.39, we need only show that no T exists such that

$$AT = TA^*$$

$$CT = B^*$$

for a nonequilibrium steady state.

If such a T exists, then

$$A(T\underline{1}) = TA^*\underline{1} = \underline{0}$$

By Theorem 5.3.26, the steady state solution $\underline{p}^s$ is unique. So

$$T\underline{1} = \underline{p}^s \quad .$$

But then,

$$B^*\underline{1} = \underline{0}$$

while

$$C(T\underline{1}) = C\underline{p}^s$$

$$= \begin{bmatrix} \underline{J}^s \\ \overline{F}^s \end{bmatrix} \neq \underline{0}$$

for all nonequilibrium steady states, by Proposition 5.3.36. So no such T exists for a nonequilibrium steady state. ■

Equilibrium is equivalent to time reversibility defined in terms of the conditional probabilities, as in Theorem 5.3.31. We now relate reciprocity of the system Σ linearized about equilibrium to a form of time reversibility defined in terms of the work flow, using a result of Day [DAY 71].

As the linearized system is time-invariant, we let $t_0 = -\infty$ without loss of generality. Then for first-order deviations of the inputs u and v about their equilibrium values (of zero) the work flow

$$W(t) = -\overline{J}(t)u(t) + \overline{F}(t)v(t)$$

$$= \int_{-\infty}^{t} \begin{bmatrix} -u(t) \\ v(t) \end{bmatrix}' H^e(t-s) \begin{bmatrix} u(s) \\ v(s) \end{bmatrix} ds \quad .$$

The total work done by the system Σ is thus

$$W(u(\cdot),v(\cdot)) = \int_{-\infty}^{\infty} W(t)dt \quad .$$

We shall say the inputs $\tilde{u}(\cdot)$ and $\tilde{v}(\cdot)$ are the dynamical reversals of the inputs $u(\cdot)$ and $v(\cdot)$ if for all $t \in \mathbb{R}$,

$$\tilde{u}(t) = u(-t)$$

$$\tilde{\chi}(t) = \chi(-t) \quad .$$

So the inputs $\tilde{u}(\cdot)$ and $\tilde{v}(\cdot)$ are the dynamical reversals of the inputs $u(\cdot)$ and $v(\cdot)$ if for all $t \in \mathbb{R}$,

$$\tilde{u}(t) = u(-t)$$

$$\tilde{v}(t) = -\frac{d\tilde{\chi}(t)}{dt} = -\left[-\frac{d\chi(\tau)}{d\tau}\bigg|_{\tau=-t}\right] = -v(-t).$$

Then

Theorem 5.3.41.

If $H^e(t) = H^{e*}(t)$ for all $t \in \mathbb{R}_+$, then for the first-order deviations of the inputs u and v about their equilibrium values and for all piecewise continuous u and v the total work done by the system Σ is invariant under the dynamical reversal of the inputs. ■

In this example, we have shown that our class of models which describe energy transduction in muscle exhibits the properties required of any model of a thermodynamic system--the evolution of the system to a steady state, the equivalence of time-reversal invariance and the equilibrium state, reciprocity of the linearized system if and only if the linearization is about equilibrium, and dissipation inequalities related to the second law. We have derived expressions for the ATPase rate and contractile force and have shown a simple method for calculating the ATPase rate and contractile force in steady state. In [PRO 81] we show how the values found for the structural parameters and rate constants can be incorporated into a model to predict the physiological properties of muscle. Thus here we have shown that our class of models is plausible as an explanation of energy transduction in muscle. In [PRO 81] we prove that the model in our class of models which has its parameters determined by biochemical and structural data mirrors reality by predicting the physiological data.

5.4 An Overview of the Subject

5.4.1 Preliminaries

In this section, we present the outline of an axiomatic framework for thermodynamics assuming a Markov process description of a thermodynamic system Σ. Our purpose here is to show how the results apply to both continuous-state Markov processes (Markov diffusion processes) and discrete-state Markov processes (Markov step processes) and demonstrate how seemingly-unrelated examples fit rather nicely into our framework.

Consider a Markov process $\{\chi(t), t \in T\}$, where $T = [t_0,\infty)$ is the time set of interest. Let the process χ describe the motion of a particle in a state space X. As the thermodynamic system Σ, take an ensemble of particles -- a collection of independent copies of the process χ. The state at time t_Δ of the system Σ is then the density function $p_t(\cdot) \triangleq p(t,\cdot)$. The state space is the space P of all density functions of the process χ.

Statistical mechanics considers a theromdynamic system as an ensemble of particles. The theory of the Brownian motion of a particle in an external field of force is based on the Langevin equation, which describes the evolution of a Markov process (e.g., [KAC 69] and [NEL 67]). Thus the ideas that a macroscopic thermodynamic system is an ensemble of microscopic particles and that the motion of a microscopic particle can be considered a Markov process have a solid basis in physics. Starting with this assumption -- that a thermodynamic system is a collection of independent copies of a Markov process -- we proceed to develop the theory of thermodynamics rigorously, which is an original approach.

The transition density function $p(t,x|s,y)$ satisfies Kolmogorov's forward and backward equations. The distribution function $P(t,\Gamma) = Pr\{\chi(t) \in \Gamma\} = \int_\Gamma p(t,x)dV$. As the density function $p(t,x) = E\{p(t,x|s,y)\}$, $s<t$, Kolmogorov's forward equation

$$\frac{\partial p_t}{\partial t} = Ap_t$$

defines the evolution of the state p_t on P. The operator of the backward equation is the adjoint of A -- A*.

The state space X can be continuous or discrete. For a discrete state space, we give a topology to the set X by associating a directed graph with X. The expressions of vector calculus can then be defined for a discrete state space.

5.4.2 Characterizations of Equilibrium

We can write Kolmogorov's forward equation as

$$\frac{d}{dt}\int_{\Gamma} p(t,x)\,dV = -\int_{\partial\Gamma} n \cdot J\,dS$$

for all regions $\Gamma \subset X$, where we define probabilistic fluxes J in terms of the state p_t and the parameters of the operator A of the forward equation. If a unique invariant solution of Kolmogorov's forward equation is the initial state, we say the system Σ is in a steady state. The steady-state condition $\int_{\partial\Gamma} n \cdot J\,dS = 0$ for any region $\Gamma \subset X$ does not exclude circulation around $\partial\Gamma$, the boundary of Γ. When the fluxes J are identically zero, a steady state cannot be maintained by any cyclic mechanism. We thus call that steady state for which all fluxes vanish identically an equilibrium state. The thermodynamic condition of detailed balance is equivalent to equilibrium by definition. We prove a state is an equilibrium state if and only if the intertwining operator equation

$$AM_{p_t} = M_{p_t} A^*$$

holds, where M_{p_t} is the operator multiplication by p_t. We then use this result to show the system Σ is in equilibrium if and only if the condition of time reversal invariance holds, defined as $p(t,x|s,y) = p(s,x|t,y)$ for all $t, s \in T$, $x, y \in X$.

When $\mu(X) \triangleq \int_X dV < \infty$, there is an equilibrium state which is independent of x, which we call the homogeneous state. The homogeneous state corresponds to the thermodynamic condition of uniform composition. Using the result that the system Σ will evolve to the homogeneous state if and only if the operator A is self-adjoint, we show the system will evolve to the homogeneous state if and only if the condition of positional reversal invariance holds, defined as $p(t,x|s,y) = p(t,y|s,x)$ for all $t, s \in T$, $x, y \in X$.

The condition of time reversal invariance can be stated in terms of the joint probabilities as $p(t,x;\, s,y) = p(s,x;t,y)$. This gives a physical meaning to the condition of zero flux -- namely, the system is in the equilibrium state if and only if for any regions $\Gamma_1, \Gamma_2 \subset X$,

$$E\{(\text{Total transitions from } \Gamma_1 \text{ to } \Gamma_2) - (\text{Total transitions from } \Gamma_2 \text{ to } \Gamma_1)\} = 0.$$

We define external parameters F and internal parameters A of the thermodynamic system Σ. Knowledge of the internal and external parameters is equivalent to knowledge of the operator A. For a system which evolves to an equilibrium state p^e, $-\nabla \ln p^e = F$. Thus an equilibrium state can be written solely in terms of the external parameters. As $\oint_C \ln p^e(x)\cdot dx = 0$, for all closed paths $C \subset X$, if a steady state is an equilibrium state, then $\oint_C F(x)\cdot dx = 0$, for all closed paths $C \subset X$. We prove a steady state is an equilibrium state if and only if $\oint_C F(x)\cdot dx = 0$, for all closed paths $C \subset X$ and an equilibrium state is a homogeneous state if and only if $F=0$.

The connection between the intertwining equation time-reversal invariance, and the condition on the external parameters was first studied by Kolmogorov in the context of Markov process theory (and using a different terminology). Recent references are [KEI 79] for Markov chains and [KEN 78] for Markov diffusion processes. However, these ideas are discussed in the framework of time-reversible Markov processes and not connected to thermodynamics.

For continuous-state Markov processes, we are able to characterize processes for which $p(t,x|s,y) = p(s, Rx|t, Ry)$ in steady-state, where R is a linear map from X into itself. If R is a diagonal matrix with entries either 1 or -1 (a signature matrix), then $p(t,x|s,y) = p(s, Rx|t, Ry)$ corresponds to the condition of dynamic reversibility, a special case of which is discussed by Anderson [AND 80].

5.4.3 Interactions of a Thermodynamic System with Its Surroundings

To model interactions of a thermodynamic system with its environment, we describe systems with time-dependent parameters. The evolution of a Markov process can be described by a stochastic differential equation. We shall say two thermodynamic systems are interconnected if the parameters of the stochastic differential equation describing the evolution of one process (or both) can be written as functions of the other process (or both). The state of an interconnected thermodynamic system is the probability density function for the joint process.

The external and internal parameters of the joint process are functions of (x_1, x_2), where $x_1 \in X_1$, $x_2 \in X_2$, and $X_1 \times X_2$ is the state space of the joint process. We in addition take the parameters as functions of $u(t, x_1, x_2)$, where u is an external control applied to one or both systems -- we thus allow both open-loop and feedback controls. In the absence of either work or heat flow, thermodynamics suggests a system will evolve to equilibrium. We formalize this thermodynamic idea with the assumption that a thermodynamic system will evolve to equilibrium if and only if the system is uncontrolled.

Our formalism allows us to control a thermodynamic system of interest either directly or by controlling a thermodynamic system to which it is connected. For that part of the process which evolves deterministically through the action of the control, we define deterministic fluxes V such that for the system to evolve to equilibrium, V must be equal to zero.

5.4.4 The Dissipation Inequalities of Thermodynamics

Let us now consider a thermodynamic system Σ, which may be an interconnection of thermodynamic systems and acted upon by external controls. Given the uncontrolled system evolves to the equilibrium state, one might wish to know how far the system Σ is from equilibrium. The state functions of thermodynamics can be thought of as distance measures on $\mathcal{P}$, the space of all density functions of the process.

Define the distance between states p_t^1 and p_t^2,

given that $\chi(t) = x$, as $\Lambda(p_t^1, p_t^2) = \ln\left(\frac{p_t^1(x)}{p_t^2(x)}\right)$.

The function Λ has two valuable properties: If $p_t^1 = p_t$, the state of the system Σ, and $p_t^2 = p^s$, a steady state, then $E\{\frac{d}{dt}\Lambda(p_t^1(x), p_t^2(x))\} = 0$. Furthermore, the function $|\Lambda|$ is a metric on $\mathbb{R}_+$. It can be verified that only multiples of Λ possess these two properties.

We define the free energy of the system Σ at time t, G(t), as

$$G(t) = E\{\Lambda(p_t(x(t)), p_t^e(x(t)))\},$$

the entropy of the system Σ at time t, S(t), as

$$S(t) = E\{\Lambda(1, p_t(x(t)))\},$$

and the internal energy of the system Σ at time t, U(t), as

$$U(t) = E\{\Lambda(1, p_t^e(x(t)))\},$$

where p_t^e is the equilibrium state to which the system will evolve when the control is set to zero at time t. We find

$$G(t) = \int_X p_t(x) \ln \frac{p_t(x)}{p_t^e(x)} dV \geq 0,$$

with equality iff $p_t = p_t^e$,

$$S(t) = -\int_X p_t(x) \ln p_t(x) dV \leq \ln \mu(x),$$

with equality iff p_t is the homogeneous state,

$$U(t) = -\int_X p_t(x) \ln p_t^e(x) dV \geq 0 .$$

We interpret the free energy as the distance of the system Σ from the equilibrium state, the entropy as the distance of the system Σ from the homogeneous state, and the internal energy as the distance of the equilibrium state from the homogeneous state.

The function $\int_X p^{(H_1)}(x) \ln \frac{p^{(H_1)}(x)}{p^{(H_2)}(x)} dV$ appears frequently in the statistics and information theory literature (e.g., see [WIL 76]) and is known by many names, including the "average weight of evidence in favor of H_1 against H_2, given H_1" [WIL 76]. Aside from incorporating this function, our approach is original.

External work forces are defined through the Fréchet derivatives of the free energy flow $\frac{dG}{dt}$ with respect to the fluxes J and V, evaluated at equilibrium conditions. Similarly external heat forces are defined through the Fréchet derivatives of the entropy flow $\frac{dS}{dt}$ with respect to the fluxes J and V, evaluated at equilibrium conditions. The external forces conjugate to the probabilistic fluxes J are equivalent to the external parameters F in the case of heat forces and are the difference of the external parameters F and their equilibrium values F^e in the case of work flow. The external forces conjugate to the deterministic fluxes V are $p_t \nabla_y \Lambda(p_t, p_t^e)$ in the case of work forces and $p_t \nabla_y \Lambda(1, p_t)$ in the case of heat forces, where $\frac{dy}{dt} = -V$ describes that part of the system which evolves deterministically. If V is independent of x, these forces can be defined as expected values. If $\nabla_y \cdot V = 0$, the heat forces vanish.

We define the work flow W and heat flow Q as inner products of the fluxes and their conugate forces. The dissipation rate D, defined as $\frac{dS}{dt} - Q$, can be expressed as the inner product of the probabilistic fluxes J and appropriately-defined conjugate internal forces. The free energy G, entropy S, internal energy U, work flow W, heat flow Q, and dissipation rate D satisfy the dissipation inequalities of thermodynamics and have the required values in steady state, equilibrium, and the homogeneous state, which we interpret as the equilibrium state of an isolated system.

5.4.5 Reciprocity

The deterministic fluxes V and external work forces determined by the external parameters $F - F^e$ can be considered as inputs to the thermodynamic system Σ. The outputs of the system Σ are then the probabilistic fluxes J and the external work forces $p_t \nabla_y \Lambda(p_t, p_t^e)$ computed from the system state and in many cases expressible as expected values.

We say the thermodynamic system Σ linearized about a steady state is externally reciprocal if the operator mapping the inputs to the outputs obtained by linearizing the system is self-adjoint. We prove the system Σ linearized about a steady state is externally reciprocal if and only if that steady state is the equilibrium state. A similar result is obtained for internal reciprocity when we consider the map from the probabilistic fluxes, internal parameters, and state to the internal forces. Thus under our assumption that a thermodynamic system evolves as a Markov process, Onsager's reciprocity theorem can be stated precisely and proven.

Using a result of Day [DAY 71], we show that, just as equilibrium is equivalent to time reversal invariance expressed in terms of the conditional transition probabilities, external reciprocity is equivalent to the invariance of the work done by the system Σ under time reversal of the inputs.

5.5 The Thermodynamic Properties of Markov Step Processes

In this section, we develop an axiomatic framework for thermodynamics assuming a Markov step process description of the system. The results presented in this section can be generalized to broader classes of Markov processes, as indicated in the previous section.

Consider a Markov step process $\xi(t)$, $t \in T = [t_o, \infty)$, which describes the motion of a particle. A thermodynamic system Σ is a collection of independent copies of this process, represented mathematically as an ensemble of particles. The state of the system Σ at time t is its distribution function $\underline{p}(t)$. The state space will be the space of all distribution functions of the

$\xi(t)$ process.

Assume that $\{\xi(t), t\in T\}$ is a conservative Markov step process with state space $E = \{1,\ldots N\}$, with infinitesimal generator the bounded operator $\underline{A}^*$, whose entries are $a^*_{ij}(t)$. The entries a_{ij} can be interpreted as infinitesimal transition rates, as

$$\Pr\{\xi(t+s) = j \mid \xi(t) = i\} = a^*_{ij}(t)s + o(s) \quad (5.5.1)$$

The probability distribution of $\xi(t)$ is described by the Fokker-Planck equation, as

$$\frac{d}{dt}\underline{p}(t) = \underline{A}(t)\underline{p}(t);\ \underline{p}(t_0) = \underline{p}_o\ , \quad (5.5.2)$$

where the entries of the $\underline{A}(t)$ matrix are $a_{ij}(t) = a^*_{ji}(t)$. Let $\underline{P}_E$ denote the space of all probability measures on E. Then, equation (5.5.2) defines an evolution on $\underline{P}_E$ because the Markov process is conservative. We will assume the following conditions on the process ξ.

Assumption 5.5.1. The Markov chain of transitions associated with the process ξ is irreducible. Every state is reachable with finite probability from every other state. ■

Assumption 5.5.1 guarantees the following conditions [KEI 72]:

(1) A unique solution of (5.5.2) exists which is a steady state solution, denoted $\underline{p}^s$, with all entries strictly positive.

(2) The nullspace of $\underline{A}$ is of dimension 1.

We define a thermodynamic system Σ as (5.5.2). The state of the system is $\underline{p}$ and the state space is $\underline{P}_E$.

Definition 5.5.2. The fluxes of the system Σ given by (3.2) are the elements of the matrix $\underline{J}$, where

$$J_{ij} = a_{ij}p_j - a_{ji}p_i\ . \ \blacksquare$$

We can interpret the fluxes as outputs of the thermodynamical system, from the equation

$$J_{ij} = \frac{d}{dt}\ E\{\text{Total transitions from } j \text{ to } i\} - E\{\text{Total transitions from } i \text{ to } j\}$$

Notice that Equation (5.5.2) can be rewritten as:

$$\frac{d}{dt}p_i(t) = \sum_j J_{ij}(t)\ . \quad (5.5.3)$$

Since the fluxes can be written as expected values, we consider them externally observable; that is, the fluxes are outputs of the system Σ. Consider now a thermodynamic system which does not interact with the environment. We model these interactions as affecting the parameters of $\underline{A}$. When the parameters of $\underline{A}$ are time invariant, we define a steady state as follows.

Definition 5.5.3 A steady state $\underline{p}^s$ is the unique solution of the equation

$$\underline{A}\underline{p}^s = \underline{0}\ . \ \blacksquare$$

The steady state conditions $\sum_j J_{ij} = 0$ do not exclude rotational fluxes. When the matrix of fluxes $\underline{J}(t)$ is identically 0 for $t \geq t_o$, a steady state cannot be maintained by any cyclic mechanism. We distinguish these steady states as equilibrium states.

Definition 5.5.4. An equilibrium state is a steady state for which $\underline{J} = \underline{0}$. We use the notation $\underline{p}^e$ for the probability distribution of an equilibrium state. ■

Definition 5.5.5. The process ξ is invariant under time reversal if, for all t, s ε T,

$$\underline{P}(t|s) = \underline{P}(s|t),$$

where $(\underline{P}(t|s))_{ij} = \Pr\{\xi(t) = j \mid \xi(s) = i\}$. ■

With this background, we can establish a number of results, which can be easily verified:

Lemma 5.5.6. The state $\underline{p}(t)$ is an equilibrium state if and only if

$$\underline{A} = \{\text{diag } p_i(t)\}\underline{A}^*\ \{\text{diag } 1/p_i(t)\}\ . \ \blacksquare$$

Corollary 5.5.7. [KEI 79] If the steady state $\underline{s}$ is an equilibrium state, the N eigenvalues of the matrix $\underline{A}$ are real and

$$o = \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_N\ . \ \blacksquare$$

Corollary 5.5.7 shows that the approach to equilibrium is without oscillations as all the eigenvalues are real.

Theorem 5.5.8. The system Σ is in an equilibrium state if and only if the system has time reversal invariance. ■

Note that Theorem 5.5.8 is the mathematical version of the statement in Section 5.2 that microscopic reversibility holds at thermodynamic equilibrium. In addition to the concept of equilibrium, we want to identify the equilibrium state which corresponds to a uniform distribution. This is done in the following results.

Definition 5.5.9. The homogeneous state is that equilibrium state where

$$p^e_i = \frac{1}{N}\ . \ \blacksquare$$

Definition 5.5.10. The process ξ is invariant under positional reversal if, for all t, s ε T,

$$\underline{P}(t|s) = \underline{P}'(t|s),$$

where ' denotes transpose. ■

Lemma 5.5.11. The system Σ will evolve to the homogeneous state if and only if $\underline{A} = \underline{A}'$. ■

Notice that Lemma 5.5.11 and Lemma 5.5.6 provide a characterization of systems which evolve to homogeneous states. Namely, the system will evolve to the homogeneous state if and only if the conditions of position reversal invariance hold. To have time reversal invariance, the system must be in an

equilibrium state, from Theorem 5.5.8. To have position reversal invariance, the system must evolve to a homogeneous state, but it does not have to be there yet. Thus, we can have position reversal invariance at any time if $\underline{A}$ is symmetric, whereas time reversal invariance depends on having achieved a steady state.

Adopt the convention that if $a_{ij} = a_{ji} = 0$, the ratio $\frac{a_{ij}}{a_{ji}} = 1$. With this convention, we define the forces F_{ij} as:

Definition 5.5.12. The force F_{ij} is defined as

$$F_{ij} = -\ln \frac{a_{ij}}{a_{ji}} .$$

The elements of the set $\{F_{ij};\ i, j \in E,\ i > j\}$ are called the external parameters of the system Σ. The elements of the set $\{a_{ji};\ i,j \in E,\ i>j\}$ are called the internal parameters of the system Σ. ■

Knowledge of the internal and external parameters of the system is equivalent to knowledge of $\underline{A}$, provided that $a_{ji} \neq 0$ for $j < i$.

In equilibrium, we have the relation

$$J^e_{ij} = 0 = a_{ij}p^e_j - a_{ji}p^e_i .$$

Hence,

$$F_{ij} = -\ln \left(\frac{p^e_i}{p^e_j}\right).$$

That is, for systems which achieve thermodynamic equilibrium, the forces are always finite and the parameter set $\{F_{ij},\ a_{ji};\ i>j\}$ completely specifies the process. In addition, the equilibrium state of the system can be written solely in terms of the external parameters F_{ij}.

Definition 5.5.13. Let X_{ij} be defined as

$$X_{ij} = \ln\left(1 + \frac{J_{ij}}{a_{ji}p_i}\right) .$$

The elements of the set $\{X_{ij};\ i>j\}$ are called internal forces. ■

It is easy to verify that the internal force X_{ij} is zero if and only if the corresponding flux J_{ij} is zero. Hence, at steady state, the internal forces are all zero if and only if the steady state is an equilibrium state.

Lemma 5.5.14. Consider the transitions of the ξ process as a directed graph between the states in E. Consider any cycle in the graph, also referred to as a closed path. Then

$$\sum_{\text{closed path}} X_{ij} + \sum_{\text{closed path}} F_{ij} = 0.$$

Proof: The proof of this result is a typical proof of results in this section. By definition,

$$J_{ij} = a_{ij}p_j - a_{ji}p_i .$$

Hence,

$$\ln \frac{p_i}{p_j} = -\ln\left(1 + \frac{J_{ij}}{a_{ji}p_i}\right) + \ln \frac{a_{ij}}{a_{ji}}$$

$$= - X_{ij} - F_{ij} .$$

In addition,

$$\sum_{\text{closed path}} \ln \frac{p_i}{p_j} = \ln 1 = 0 .$$

Hence,

$$\sum_{\text{closed path}} (X_{ij} + F_{ij}) = 0 . \quad ■$$

With this result, we can characterize the forces F_{ij} associated with a system approaching equilibrium as conservative. This result is expressed as:

Theorem 5.5.15. A steady state is an equilibrium state if and only if

(a) $\sum_{\text{closed path}} F_{ij} = 0$ for any closed path of the directed graph of the process.

(b) $a_{ij} = 0$ implies $a_{ji} = 0$ for all $i, j \in E$.

If, in addition, $F_{ij} = 0$ for all $i, j \in E$, the equilibrium state is the homogeneous state. ■

The concepts and results described so far apply primarily to systems whose parameters $\underline{A}$ are constant, so that no interactions with the environment are obtained. For nonequilibrium thermodynamics, these interactions must be modeled explicitly. This is done in the rest of this section.

The interactions are modeled through the time dependence of the matrix $\underline{A}$. We assume that this time dependence is caused by an external deterministic process controlled in the surroundings of the system. Let $\underline{\alpha}(t)$ be the state vector of this dynamical system. We assume that the coupled evolution of the thermodynamic system Σ and the state $\underline{\alpha}$ is given by

$$\frac{d\underline{p}}{dt} = \underline{A}(\underline{F}(t), \underline{a}(t))\underline{p}(t); \quad \underline{p}(t_o) = \underline{p}_o$$

$$\underline{J} = \underline{J}(\underline{F}(t), \underline{a}(t), \underline{p}(t))$$

$$\frac{d}{dt}\underline{\alpha} = -\underline{V}(\underline{\alpha}(t), \underline{J}(t), \underline{u}(t)); \quad \underline{\alpha}(t_o) = 0 \qquad (5.5.4)$$

$$\underline{F}(t) = \underline{F}(\underline{\alpha}(t), \underline{u}(t))$$

$$\underline{a}(t) = \underline{a}(\underline{\alpha}(t), \underline{u}(t)),$$

where $\underline{F}$, $\underline{a}$, $\underline{J}$ are vectors composed of F_{ij}, a_{ji}, J_{ij},

$i > j$, and $\underline{u}(t)$ is an external control. We have assumed that $a_{ij} = 0$ implies $a_{ji} = 0$, so that F_{ij} is finite and $\underline{F}$, $\underline{a}$ describe the system.

Assumption 5.5.16. For any pair of initial conditions $(\underline{\alpha}_o, \underline{p}_o)$, and any bounded, integrable control $\underline{u}(t)$, there is a unique solution to Equations (5.5.4). When $\underline{u}(t)$ is zero everywhere, let

$$\underline{\alpha}^s(\underline{\alpha}_o) = \lim_{t\to\infty} \underline{\alpha}(t\,;\,\underline{\alpha}_o, \underline{p}_o) \ .$$

Assume that $\underline{\alpha}^s$ is independent of the initial distribution $\underline{p}_o$. Assume that the unique solution of

$$\underline{A}(\underline{F}(\underline{\alpha}^s(\underline{\alpha}_o),\underline{0}), \underline{a}(\underline{\alpha}^s(\underline{\alpha}_o),\underline{0}))\underline{p}^s(\underline{\alpha}_o) = \underline{0}$$

is an equilibrium state, and that, for $\underline{u}(t)$ not identically zero, there exists no steady state which is an equilibrium state. ■

By assumption, the system Σ will evolve to equilibrium if and only if the process which determines the parameter values is uncontrolled. The equilibrium state is unique and depends only on the values to which the parameters are driven before the controls are released.

Definition 5.5.17. Let $\underline{p}^e(t)$ denote the equilibrium state solution of the uncontrolled system (5.5.4) with the initial condition $\underline{\alpha}_o = \underline{\alpha}(t)$. We call $\underline{p}^e(t)$ the refenence equilibrium state at time t for the thermodynamic system. ■

Since the fluxes $\underline{J}(t)$ are externally observable, the process $\underline{\alpha}(t)$ can be controlled externally. Thus, the parameter set

$$\{\underline{F}(t), \underline{a}(t), \underline{p}^e(t), \frac{\partial \underline{p}^e}{\partial \underline{\alpha}}, \underline{V}(t), t \in T\}$$

can be viewed as the external interactions of the surroundings with the system Σ. In what follows, we show how these interactions can be interpreted in terms of work flow, heat flow and other thermodynamic concepts.

The state functions of thermodynamics can be thought of as distance measures on P_E. Define the distance between states $\underline{p}^1(t)$ and $\underline{p}^2(t)$, given that $\xi(t) = i$, as

$$\Lambda(p_i^1(t), p_i^2(t)) = \ln\left(\frac{p_i^1(t)}{p_i^2(t)}\right) \ .$$

The function Λ has two valuable properties: If $\underline{p}^1(t) = \underline{p}(t)$, the state of the thermodynamic system Σ, and $\underline{p}^2(t) = \underline{p}^s$, its steady state, then

$$E\{\frac{d}{dt}\Lambda(p_i^1(t), p_i^2(t))\} = 0.$$

Furthermore, the function $|\Lambda|$ is a metric on $[0,1]$. It can be verified that only multiples of Λ satisfy these two properties. Using this conditional distance, we define the thermodynamical properties of Σ.

Definition 5.5.18. The free energy of the system Σ at time t, G(t), is given by

$$G(t) = E\{\Lambda(p_i(t), p_i^e(t))\} \ .$$

The entropy of the system Σ at time t, S(t), is given by

$$S(t) = E\{\Lambda(1, p_i(t))\} \ .$$

The internal energy of the system Σ at time t, U(t) is given by

$$U(t) = E\{\Lambda(1, p_i^e(t))\} \ . \quad ■$$

Notice that $p_i^e(t)$ depends functionally on $\underline{\alpha}(t)$, from Definition 5.5.17. Hence, all of the state functions are dependent both on the state of the system $\underline{p}(t)$, and the initial value of the parameters $\underline{\alpha}(t)$, when the system is left uncontrolled. In other words, the state functions depend on the complete state, $\underline{p}(t)$ and $\underline{\alpha}(t)$. Using the functional expression for Λ, we can evaluate the concepts in Definition 5.5.18, as follows, in terms of the state $\underline{p}(t)$ and the reference equilibrium state $\underline{p}^e(t)$.

$$G(t) = \sum_i p_i(t) \ln\left(\frac{p_i(t)}{p_i^e(t)}\right) \qquad (5.5.5)$$

$$S(t) = \sum_i -p_i(t) \ln p_i(t) \qquad (5.5.6)$$

$$U(t) = \sum_i -p_i(t) \ln p_i^e(t) \ . \qquad (5.5.7)$$

These expressions lead to some simple relations between these functions, due to the properties of the logarithm function. These relations are summarized in the following theorem.

Theorem 5.5.19. The thermodynamic functions G, S, and U satisfy:

(a) $G = U - S$.

(b) $G(t) \geq 0$, with equality if and only if $\underline{p}(t) = \underline{p}^e(t)$.

(c) $S(t) \leq \ln N$, with equality if and only if $\underline{p}(t)$ is a homogeneous state.

(d) $U(t) \geq 0$. ■

It is easy to see that the relations between the state functions described in Theorem 5.5.19 correspond to the classical relations of thermodynamics, summarized in Section 5.2, Equations (5.2.1) - (5.2.6). In the next definition we identify the thermodynamic concepts of interactions with the environment.

Definition 5.5.20. The deterministic fluxes at time t are the elements of the vector $\underline{V}(t)$ which describe the evolution of $\underline{\alpha}(t)$. The external forces $\underline{F}^{w,J}$, $\underline{F}^{w,V}$, and $\underline{F}^{q,J}$ are defined as:

$$\underline{F}^{w,J} = -\frac{\partial}{\partial \underline{J}}\left(\frac{d}{dt}G(t)\right)\Big|_{\underline{J}=0,\underline{V}=0}$$

$$\underline{F}^{w,V} = -\frac{\partial}{\partial \underline{V}}\left(\frac{d}{dt}G(t)\right)\Big|_{\underline{J}=0,\underline{V}=0}$$

$$\underline{F}^{q,J} = -\frac{\partial}{\partial \underline{J}}\left(\frac{d}{dt}S(t)\right)\Big|_{\underline{J}=0,\underline{V}=0} \quad . \blacksquare$$

These forces are defined at equilibrium conditions, $\underline{J} = 0$, $\underline{V} = 0$. The forces represent the effect of variations in the external input set when the system is at equilibrium. A stronger characterization of these forces can be derived using the expressions in Equations (5.5.5) - (5.5.7). Define $\underline{F}^e(t)$ as the vector whose elements $F^e_{ij}(t)$ are arranged similarly to $\underline{F}$, and are defined by

$$F^e_{ij}(t) = -\ln\left(\frac{p^e_i(t)}{p^e_j(t)}\right) \quad . \tag{5.5.8}$$

Then, from (5.5.5), we can compute the rate of change of free energy as

$$\frac{d}{dt}G = \sum_i \left(\left(\frac{d}{dt}p_i\right)\ln\left(\frac{p_i}{p^e_i}\right) + \frac{d}{dt}p_i - p_i\frac{d}{dt}\ln p^e_i\right). \tag{5.5.9}$$

The second term in (5.5.9) drops out because total probability is conserved. The remainder can be written as

$$\frac{d}{dt}G = \sum_i \sum_{j<i} J_{ij}\ln\left(\frac{p_i}{p_j}\right) - \sum_i \sum_{j<i} J_{ij}\ln\left(\frac{p^e_i}{p^e_j}\right) + \sum_i p_i\left(\frac{\partial}{\partial \underline{\alpha}}\ln p^e_i\Big|_{\underline{\alpha}(t)} \cdot \underline{V}(t)\right) \quad . \tag{5.5.10}$$

From Equation (5.5.10) and Definition 5.5.20 we obtain

$$\underline{F}^{w,J}(t) = \underline{F}(t) - \underline{F}^e(t) \tag{5.5.11}$$

$$\underline{F}^{w,V}(t) = -\sum_i p_i(t)\frac{\partial}{\partial \underline{\alpha}}\ln p^e_i(t)\Big|_{\underline{\alpha}(t)} \tag{5.5.12}$$

$$\underline{F}^{q,J}(t) = \underline{F}(t) \quad . \tag{5.5.13}$$

The next result characterizes the effects of these forces.

Theorem 5.5.21. Assume $\underline{\alpha}$ is constant for $t \geq s$. Then,

(a) $\sum_{\text{closed path}} F^{q,J}_{ij}(s) = 0$ if and only if the system will evolve to the equilibrium state.

(b) $\underline{F}^{w,J}(s) = \underline{0}$ if and only if the system Σ will evolve to the equilibrium state.

(c) $\underline{F}^{q,J}(s) = \underline{0}$ if and only if the system Σ will evolve to the homogeneous state.

(d) if $\left\{\frac{\partial}{\partial \alpha_1}\underline{A}, \ldots \frac{\partial}{\partial \alpha_N}\underline{A}\right\}$ is of rank N-1, then $F^{w,V}(t) = \underline{0}$ if and only if the system Σ is in equilibrium at time t. $\blacksquare$

Notice that the assumptions on $\underline{\alpha}$ imply that the inputs $\underline{F}(t)$, $\underline{F}^e(t)$, $\underline{p}^e(t)$ and $\frac{\partial}{\partial \underline{\alpha}}\ln p^e_i(t)$ are constant. The proof follows from Lemma 5.5.14 and Theorem 5.5.15.

Recall that $\underline{J}$, $\underline{F}$ are conformally arranged vectors with N(N-1) entries. We now define work flow as the effect of forces and fluxes.

Definition 5.5.22. The work flow W is given by

$$W = \langle \underline{J}, \underline{F}^{w,J}\rangle + \langle \underline{V}, \underline{F}^{w,V}\rangle \quad .$$

The heat flow Q is defined as

$$Q = \langle \underline{J}, \underline{F}^{q,J}\rangle \quad .$$

The entropy production D is defined as

$$D = \frac{d}{dt}S - Q \quad . \blacksquare$$

Lemma 5.5.23. $D = \langle \underline{J}, \underline{X}\rangle \geq 0$, with equality if and only if the system is at equilibrium. $\blacksquare$

The proof follows from the identities

$$\frac{d}{dt}G = -\langle \underline{J}, \underline{F}^{w,J}\rangle - \langle \underline{V}, F^{w,V}\rangle - \langle \underline{J}, \underline{X}\rangle \tag{5.5.14}$$

$$\frac{d}{dt}U = -\langle \underline{J}, \underline{F}^e\rangle - \langle \underline{V}, \underline{F}^{w,V}\rangle \tag{5.5.15}$$

and from the fact that J_{ij} and X_{ij} have the same sign. Lemma 5.5.23 verifies the thermodynamic statement that entropy production D is positive, and zero only at equilibrium, as stated in Equation (5.2.4).

Using the results of Lemma 5.5.23 and Equations (5.5.11) and (5.5.12), we obtain the first and second law of thermodynamics, Equation (5.2.8):

$$\frac{d}{dt}U = Q - W$$

$$\frac{d}{dt}S = Q + D \geq Q$$

$$\frac{d}{dt}G = -W - D \leq -W \quad .$$

The thermodynamic state functions U, G, S, and D can be interpreted in terms of the theory of dissipative systems as described in [WIL 72]. The free energy of the system is a storage function, G. The entropy production D is the dissipation rate of the thermodynamic system, which is nonnegative. Hence,

the thermodynamic system is dissipative.

In steady state, all of the state functions are constant. Hence, the work flow W, heat flow Q, and the negative of the entropy production D are all equal and negative. If the steady state is an equilibrium state, all of these quantities are zero. This implies that W must be independent of the reference equilibrium state, in any steady state. We will write W^S explicitly as a product of forces and fluxes, with the forces independent of the steady state, using concepts from graph theory and circuit theory.

Note that, in steady state, $\sum_j J_{ij} = 0$, so Kirchoff's current law is satisfied for the graph of the Markov process, with J_{ij} considered as a current. These currents can be written as a set of independent loop currents flowing around closed paths of the graph [BOS 65]. Thus,

$$J_{ij} = \sum_{\substack{\text{all loop} \\ \text{currents } k}} \sigma_k^{ij} I_k ,$$

where $\sigma_k^{ij} = \{0, -1, 1\}$ depending on whether link ij is in loop k or not and its orientation. Note that $J_{ij} = -J_{ji}$.

With this notation, we can write W^S as

$$W^S = Q = \langle \underline{J}, \underline{F} \rangle \quad \text{by (5.5.13)}$$

$$= \sum_i \sum_{j<i} J_{ij} F_{ij}$$

$$= \sum_i \sum_{j<i} \sum_{\substack{\text{loop} \\ \text{currents} \\ k}} \sigma_k^{ij} I_k F_{ij}$$

$$= \sum_{\substack{\text{loop} \\ \text{currents} \\ k}} I_k \sum_{\substack{\text{closed path} \\ \text{of } k}} F_{ij} .$$

Notice that W^S is independent of the reference equilibrium state, because the sum around any closed path of F^e_{ij} is zero.

Consider now an uncontrolled system with time-independent parameters. By Assumption 5.5.16, the system approaches an equilibrium state. By Theorem 5.5.21 $\underline{F}^{W,J}$ is zero; hence, the work flow W is zero. However, the heat flow is zero only in steady state. This leads to the definition of an isolated system.

Definition 5.5.24. An isolated system is an uncontrolled system with time independent external parameters and constant internal energy U. ■

Lemma 5.5.25. The internal energy of the system Σ is constant if and only if the reference equilibrium state is homogeneous. ■

Theorem 5.5.26. If a system is an isolated system, then

(a) $W = Q = 0$

(b) $\underline{F}^{q,J} = \underline{F}^{W,J} = \underline{0}$

(c) $\frac{d}{dt} S \geq 0$. ■

We can also obtain the equivalent statement of the universal evolution criterion of Prigogine [NIC 77]. Define a variable Z as the rate of entropy change. That is,

$$Z = \frac{d}{dt} S .$$

Then, Equations (5.5.14) and (5.5.15) imply

$$\frac{d}{dt} Z = - \sum_i \sum_{j<i} J_{ij} \frac{d}{dt} \left(\ln \frac{p_i}{p_j}\right) + \sum_i \sum_{j<i} \ln\left(\frac{p_i}{p_j}\right) \frac{d}{dt} J_{ij}$$

$$= \frac{d_F Z}{dt} + \frac{d_J Z}{dt} ,$$

where

$$\frac{d_F Z}{dt} = - \sum_i \sum_{j<i} J_{ij} \frac{d}{dt} \left(\ln \left(\frac{p_i}{p_j}\right)\right)$$

$$= - \sum_i \frac{1}{p_i} \left(\frac{d}{dt} p_i\right)^2 > 0,$$

with equality if and only if the system has reached steady state. This is the equivalent statement of Equation (5.2.9).

In a macroscopic sense, we can consider the fluxes $\underline{V}$ and the forces $\underline{F}^{W,J}$ as inputs to the thermodynamic system Σ. The outputs of the system can be viewed as the fluxes $\underline{J}$ and the forces $\underline{F}^{W,V}$. In other words, the inputs to the system are the external parameters and the deterministic fluxes, while the outputs are expected values computed from the system state. In addition, the fluxes $\underline{J}$, the internal parameters $\underline{a}$, and the state $\underline{p}$ define the internal forces $\underline{X}$. Concepts of reciprocity involve these input-output relationships, as well as the internal relationship between state and internal forces. Close to equilibrium these relationships can be expressed as linear, leading to the next result.

Definition 5.5.27. The thermodynamic system Σ linearized about a steady state $\underline{p}^S$ is externally reciprocal if the operator $H^S : (\underline{v},\underline{w}) \dashrightarrow (\underline{x},\underline{y})$ is self-adjoint, where $(\underline{v},\underline{w})$ and $(\underline{x},\underline{y})$ are related by the dynamical system

$$\frac{d}{dt} \underline{u} = \underline{A}^S \underline{u} + \underline{B}_1 \underline{v} + \underline{B}_2 \underline{w} .$$

$$\underline{x} = \underline{C}_1 \underline{u} + \underline{D}_1 \underline{v}$$

$$\underline{y} = \underline{C}_2 \underline{u} + \underline{D}_2 \underline{w} ,$$

and the parameters of the system are defined by

$$\underline{A}^S = \underline{A}\Big|_{ss}$$

$$\underline{B}_1 = \frac{\partial}{\partial \underline{F}^{W,J}} (\underline{A}\underline{p})\Big|_{ss}$$

$$\underline{B}_2 = \frac{\partial}{\partial \underline{V}} (\underline{A}\underline{p})\Big|_{ss}$$

$$\underline{C}_1 = \frac{\partial}{\partial \underline{p}} \underline{J}\Big|_{ss}$$

$$\underline{C}_2 = \frac{\partial}{\partial \underline{p}} \underline{F}^{w,v}\Big|_{ss}$$

$$\underline{D}_1 = \frac{\partial}{\partial \underline{F}^{w,J}} \underline{J}\Big|_{ss}$$

$$\underline{D}_2 = \frac{\partial}{\partial \underline{v}} \underline{F}^{w,v}\Big|_{ss} .$$

The system Σ is internally reciprocal if the operator L: $(\underline{J}, \underline{a}, \underline{p}) \dashrightarrow \underline{X}$ has

$$\frac{\partial}{\partial \underline{a}} \underline{X}\Big|_{ss} = 0 = \frac{\partial}{\partial \underline{p}} \underline{X}\Big|_{ss}, \text{ and}$$

$$\frac{\partial}{\partial \underline{J}} \underline{X}\Big|_{ss} \text{ is symmetric} . \quad \blacksquare$$

External reciprocity involves the input-output relations of the thermodynamical system, linearized about a steady state. Internal reciprocity looks at the relation between the internal forces $\underline{X}$ and the internal parameters $\underline{J}$, $\underline{a}$, and $\underline{p}$. We can now characterize reciprocity in terms of equilibrium concepts.

By definition, we know

$$X_{ij} = \ln\left(1 + \frac{J_{ij}}{a_{ji}p_i}\right) .$$

Therefore,

$$\frac{\partial X_{ij}}{\partial J_{kl}}\Big|_{ss} = \frac{\delta(i-k)\,\delta(j-l)}{a^s_{ji}p^s_i + J^s_{ij}}$$

$$\frac{\partial X_{ij}}{\partial a_{kl}}\Big|_{ss} = -\frac{J^s_{ij}\,\delta(k-j)\,\delta(l-i)}{p^s_i a^s_{ji}\left(1 + \dfrac{J^s_{ij}}{a^s_{ji}p^s_i}\right)}$$

$$\frac{\partial X_{ij}}{\partial p_k}\Big|_{ss} = -\frac{J^s_{ij}\delta(i-k)}{p^s_i(a^s_{ji}p^s_i + J^s_{ij})} .$$

Hence, for $\dfrac{\partial \underline{X}}{\partial \underline{p}}\Big|_{ss}$ and $\dfrac{\partial \underline{X}}{\partial \underline{a}}\Big|_{ss}$ to be 0, it is necessary and sufficient that the fluxes $\underline{J}$ be zero, at steady state. This condition also guarantees symmetry because $\dfrac{\partial \underline{X}}{\partial \underline{J}}$ is diagonal. Since $\underline{J}$ is zero at equilibrium states, we have proven:

Theorem 5.5.28. The system Σ linearized about a steady state is internally reciprocal if and only if the steady state is an equilibrium state. ■

If the linearized system H^s in Definition 5.5.27 is a minimal realization on $\mathbb{R}^{N-1}$ of the input output relation between $(\underline{v},\underline{w})$ and $(\underline{x},\underline{y})$, then external reciprocity implies the existence of a matrix $\underline{T}$ such that

$$\underline{A}^s\underline{T} = \underline{T}(\underline{A}^s)'$$

$$\begin{bmatrix}\underline{C}_1\\ \underline{C}_2\end{bmatrix}\underline{T} = \begin{bmatrix}\underline{B}'_1\\ \underline{B}'_2\end{bmatrix} .$$

This follows from the assumed symmetry of H^s and a representation result of Willems [WIL 72]. An additional characterization of external reciprocity is provided in the next theorem.

Theorem 5.5.29. Assume that the linearized system H^s is minimal on $\mathbb{R}^{N-1}$. Then, the system Σ is externally reciprocal when linearized about a steady state if and only if that steady state is an equilibrium state. ■

Earlier in this section, we showed that equilibrium was equivalent to time reversibility in terms of the conditional transition probabilities. It turns out that external reciprocity is equivalent to a form of time reversibility in terms of the work flow. This is based on a result of Day [DAY 71].

Let $t_o = -\infty$. For first-order deviations of the forces $\underline{F}^{w,J}$, denoted by $\underline{f}^{w,j}$, and fluxes $\underline{V}$, denoted by $\underline{v}$, about their equilibrium values, the instantaneous work flow at time t is

$$W(t) = \int_{-\infty}^{t} \left\langle \begin{bmatrix}\underline{f}^{w,j}\\ \underline{v}\end{bmatrix}, H^s(t-s)\begin{bmatrix}\underline{f}^{w,j}\\ \underline{v}\end{bmatrix} \right\rangle ds .$$

The total work done by the system is thus

$$TW(\underline{f}^{w,j}, \underline{v}) = \int_{-\infty}^{\infty} W(t)dt .$$

Define the time reversal of a function f(t) as f*(t). Then, Day's result implies that, for all piecewise continuous $\underline{f}^{w,j}$, $\underline{v}$, if the linearized system is externally reciprocal, then

$$TW(\underline{f}^{w,j}, \underline{v}) = TW((\underline{f}^{w,j})*, \underline{v}*)$$

which corresponds to time reversal of the inputs.

For a minimal system, Willems [WIL 72] established that the matrix $\underline{T}$ is symmetric. In our case, whenever the system is linearized about an equilibrium state, the matrix $\underline{T}$ will be given by

$$\underline{T} = \text{diag}(p^e_i) .$$

We can consider linearization of G for small deviations of $\underline{p}$ about $\underline{p}^e$. For Equation (5.5.5), let g be the perturbation of G; then,

$$g = \sum_i \frac{1}{2}\frac{(p_i - p^e_i)^2}{p^e_i} = \frac{1}{2}((\underline{p}-\underline{p}^e)'\,\underline{T}^{-1}(\underline{p} - \underline{p}^e)).$$

This function is called the coenergy of the system in [WIL 72], and serves the role of a storage function in establishing stability results.

5.6 Directions of Future Research

We have made the assumption that a system consisting of an ensemble of identical particles, corresponding to a thermodynamic system, can be studied in terms of the properties of a Markov process. This leads to the development of a theory which corresponds rather nicely to thermodynamics, describing both equilibrium and nonequilibrium, linear and nonlinear processes. This correspondence provides a basis for the analysis of complex physical systems operating under thermodynamic constraints and hope of applying a theory of stochastic control in a physically-meaningful manner to such systems.

The above results have been developed rigorously for Markov step processes with a finite state space and formally for Markov diffusion processes.

A program for the future should include:

(1) A rigorous development of the results for Markov diffusion processes.

(2) Better characterizations of interconnections of thermodynamic systems and the reference equilibrium state.

(3) A generalizaton of the performance bounds for periodic trajectories of Example 5.3.1 to the general case.

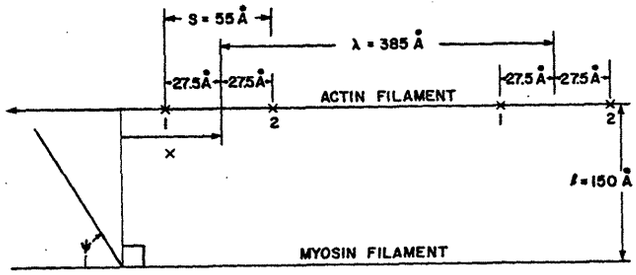


Fig. 5.3.1 The actin position x and the cross-bridge tilt angle ψ. For unattached states ψ is independent of x; for attached states

$$\begin{cases} \pi/2 + \tan^{-1}(x - s/2)/\ell, \\ \quad \text{CB attachment to site 1} \\ \pi/2 + \tan^{-1}(x + s/2)/\ell, \\ \quad \text{CB attachment to site 2} \end{cases}$$

By convention, x decreases in contraction, when the actin filament moves to the left relative to the myosin filament (indicated by the arrow).

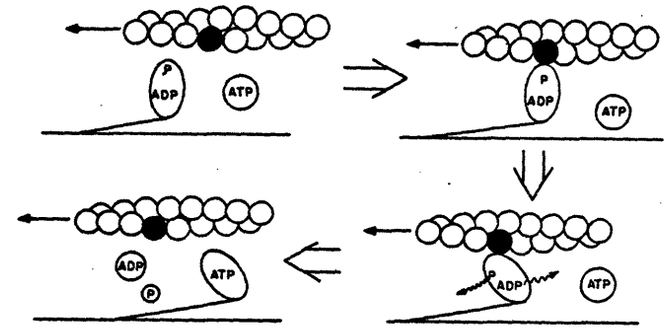


Fig. 5.3.2. A set of actin sites appears to the right of an unattached cross-bridge. The cross-bridges attaches and cycles through 2 attached states, each with different preferred angles. As the angle at which the cross-bridge is attached decreases, the probability of attachment becomes sufficiently small that the cross-bridge detaches.

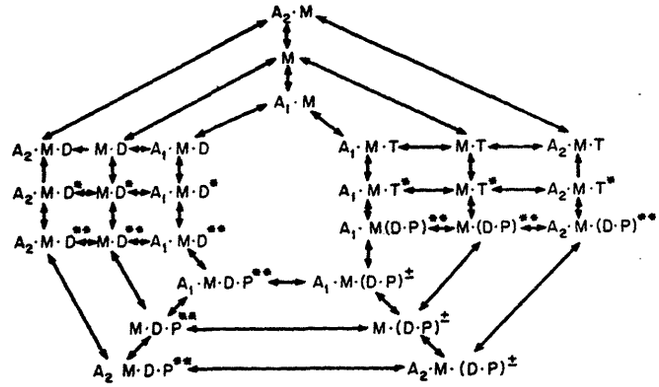


Fig. 5.3.2 A directed graph for the state space E. The states are numbered clockwise, from M for state 1, A_1·M for state 1', and A_2·M for state 1''. We have let M=myosin, A_1 = actin-site-1, A_2 = actin-site-2, T = ATP, D = ADP, and P = P_i. The super=scripts *, **, and ± distinguish biochemical states.

6. Conservative and Dissipative Systems

6.1 Introduction

One of the important problems in the study of interconnected power systems is the transient stability problem. This problem has attracted considerable attention recently and methods have been proposed using Lyapunov function ideas. In constructing the Lyapunov function the idea of energy functions serves as an important guide. One of the objectives of the present work is to develop a theory of stability of stochastic differential systems based on ideas of conservation and dissipation of energy which would be applicable to stability (stochastic) problems for power systems. At the same time, this theory must contain as a special case the deterministic theory of stability based on ideas of dissipation as the variance of the noise processes involved tends to zero. In the deterministic case such a theory was developed by J.C. Willems [WIL 79].

As a result of these considerations it was decided to take a closer look at conservative and dissipative linear systems, even in the deterministic situation. It is shown later in this chapter that the theory of conservative linear systems is intimitely connected to Scattering Theory as developed by Lax and Phillips [LAX 67] and many of these ideas extend to linear gaussian systems. We also show in this chapter how recent results in stochastic differential equations provide the right framework for extending the ideas of Willems for dissipative systems to a stochastic setting.

6.2 Recent Developments in Stochastic Differential Equations (Kunita) [KUN 81]).

Suppose we are given

(i) $B_t = (B_t^1, \ldots, B_t^m)$, $t \in [o,T]$ which is an m-dimensional Brownian motion.

(ii) $X_o(t,x), \ldots, X_m(t,x)$, $t \in [o,T]$, $x \in \mathbb{R}^d$ are d-vector-valued Lipschitz continuous functions i.e. $X_j(t,x) = (X_j^1(t,x), \ldots, X_j^d(t,x))$.

Consider an Ito equation

$$d\xi_t = X_o(x,\xi_t)dt + \sum_{j=1}^{m} X_j(t,\xi_t)dB_t^j. \qquad (6.2.1)$$

A continuous d-vector process ξ_t is called a solution of (6.2.1) with the given initial condition $\xi_s = x$, if it is F_s^t measureable for each $t > s$ and satisfies

$$\xi_t = x + \int_s^t X_o(r, \xi_r)dr + \sum_{j=1}^{m} \int_s^t X_j(r,\xi_r)dB_r^j \qquad (6.2.2)$$

The solution of (6.2.1) is denoted by $\xi_{s,t}(x,\omega)$. We then have:

Theorem: (Kunita)

We can choose a modification of the solution in a way such that for almost all ω

(i) $\xi_{s,t}(x,\omega)$ is continuous in (s,t,x).

(ii) $\xi_{t_1,t_3}(x,\omega) = \xi_{t_2,t_3}(\xi_{t_1,t_2}(x,\omega))\omega)$. holds for any $t_1 \le t_2 \le t_3$ and x.

(iii) The map $x \to \xi_{s,t}(x,\omega)$: $\mathbb{R}^d \to \mathbb{R}^d$ is a homeomorphism for any $s \le t$.

The significance of this theorem is that stochastic differential equation behave exactly like ordinary differential equations except on a null set.

6.3. Dissipative Systems

In order to extend the definition of dissipativeness to a stochastic setting we adapt the framework of Willems [WIL 79]. Using the notation of Willems, a dissipative system (deterministic) in state space form is defined by a triple $\{\Sigma,\nu,S\}$, with Σ a system in state space form on $X \times W$ (X = state space, W = space of external variables), ν: $W \times \mathbb{R} \to \mathbb{R}$ such that $\nu(w(\cdot);) \in L^1_{loc}$ for all $w \in \Sigma_e$ (Σ_e = external behavior) called the supply rate and S: $X \times \mathbb{R} \to \mathbb{R}$ a function called the storage function, such that the dissipation inequality

$$S(x(t_o),t_o) + \int_{t_o}^{t_1} \nu(w(t),t)dt \ge S(x(t_1),t_1) \qquad (6.3.1)$$

in satisfied for all $t_1 \quad t_o$ and $(x,w) \in \Sigma$.

If the above inequality holds with equality the system is said to be conservative.

For the generalization to stochastic systems, consider equation (6.2.1) together with the observation equation

$$dy_t = h(\xi_t)\,dt + d\eta_t\,, \qquad (6.3.2)$$

where y_t is a p-vector-valued random variable and h a map from $\mathbb{R}^d$ to $\mathbb{R}^p$ which is assumed to be square integrable and η_t a p-vector-valued Brownian motion independent of ξ_t.

We then extend the definition of dissipativeness to a stochastic setting by refering the dissipation inequality to hold in one of two following ways:

$$S(\xi_{t_o},t_o) + \int_{t_o}^{t_1} \nu(dB_t,y_t)dt \ge S(\xi_{t_1},t_1) \qquad (6.3.3)$$

for almost all ω

or

$$E[S(\xi_{t_o},t_o) - S(\xi_{t_1},t_1)] + \int_{t_o}^{t_1} E\nu(dB_t,y_t)dt \ge 0 \qquad (6.3.4)$$

(E denotes expectation).

The two definitions of dissipativeness a priori lead to different results. Nevertheless they need not be contradictory. These ideas will be developed further in a forthcoming paper [MIT 82] using the theory developed by Bismut [BIS 81]. It leads to the study of stochastic variational problems involving stochastic integrals in the cost functions and makes essential use of the ideas of Kunita cited earlier.

6.4. Stochastic Calculus of Variations

The idea developed in the previous section is related to the study of stochastic calculus of variations and stochastic mechanics as developed by

Fleming [FLE 81a]. In this work a certain logarithmic transformation relating linear parabolic equations and Bellman-Hamilton-Jacobi equations plays an important role. These ideas have also been extended to the context of non-linear filtering by Fleming and Mitter [FLE 81b].

It is felt that these ideas will play an important role in the work of Propp and Wyatt, (see Chapter 5) since it gives stochastic variational interpretations of Fokker-Planck equations.

6.5. Linear Conservative Systems and Scattering Theory

In this section we discuss the relationship between linear conservative dynamical systems and Scattering Theory as developed by Lax and Phillips. For simplicity we consider the discrete-time situation. The continuous-time case can be handled by an appropriate conformal transformation. For all definitions and notation see either [NAG 70] or [LAX 67].

Consider the system

$$(\Sigma) \qquad \begin{aligned} x(t+1) &= Fx(t) + Gu(t) \\ y(t) &= Hx(t) + Ju(t) \end{aligned} \qquad (6.5.1)$$

where $u(t) \in U$, $x(t) \in X$ and $y(t) \in Y$ where U, X and Y are finite dimensional spaces and we have arranged matters such that U and Y are isomosphic to subspaces of X, and $F: X \to X$, $G: U \to X$, $H: X \to Y$ and $J: U \to Y$ are linear transformations.

We assume that (Σ) is controllable and observable and it is also conservative, that is,

$$||x(t)||^2 + ||u(t)||^2 = ||x(t+1)||^2 + ||y(t)||^2, \qquad (6.5.2)$$

$\forall t \geq 0$ and $\forall x(o)$. This is equivalent to the matrix

$$\begin{bmatrix} F & G \\ H & D \end{bmatrix}$$

being unitary. It can also be verified that F is a contraction. It can then be proved

Theorem: (Σ) is unitarily equivalent to

$$(N\text{-}F) \qquad \begin{bmatrix} T & D_T^* \\ D_T & -T^*|_{\mathcal{D}_{T^*}} \end{bmatrix} \qquad (5.3)$$

where $D_{T^*}: \mathcal{D}_{T^*} \to X$, $\mathcal{D}_{T^*}$ being an invariant subspace of X and $D_T: X \to \mathcal{D}_T$, $\mathcal{D}_T$ subspace of X. Here T is a contraction.

For (5.3), form the Nagy-Foias characteristic function

$$\phi_T(z) = -T^*|_{\mathcal{D}_{T^*}} + z\, D_{T^*} (T - zI)^{-1} D_T$$

It can be shown that $\phi_T(z)$ is a contraction in the interior of the unit disc where it is analytic. System (N-F) is controllable and observable if and only if T and T* are asymptotically stable.

Define

$$\Theta_T(e^{i\phi}) = \lim_{z \to e^{i\phi}} \phi_T(z)$$

For a (N-F) system which further satisfies T and T* are asymptotically stable, Θ_T is unitary. Now consider the unitary map

$$U \begin{bmatrix} \ddots & & & \\ & I_{D_T} & -T^* & \\ & & & \\ & T & D_{T^*} & \\ & & & I \ddots \end{bmatrix} \quad \text{on}$$

$$K = \mathcal{D}_{T^*} \oplus \mathcal{D}_{T^*} \oplus X \oplus \mathcal{D}_T \oplus \mathcal{D}_T \oplus$$

Let

$$\mathcal{D}_- = \ldots \oplus \mathcal{D}_{T^*} \oplus \ldots \oplus \mathcal{D}_{T^*}$$

$$\mathcal{D}_+ = \mathcal{D}_T \oplus \mathcal{D}_T \oplus \ldots$$

Firstly, U is the minimal unitary dilation of T. Now the quadruple $(U, \mathcal{D}_-, \mathcal{D}_+, X)$ is a Lax-Phillips system.

Define

$$S(z) = P_{\mathcal{D}_{T^*}} (U - zI)^{-1}|_{\mathcal{D}_T},$$

where $\mathcal{D}_{T^*}$ and $\mathcal{D}_T$ considered as embedded in K, and P denotes projection. This is termed the Heisenberg Scattering matrix. This coincides with the Nagy-Foias characteristic function considered earlier. The significance of this result is that the spectral structure of T gives all the information about conservative systems.

These ideas can be generalized to a stochastic setting, where u(t) is now thought of as a white Gaussian sequence, by working with spectral densities. Conversely starting from a Gaussian process y(t) it is possible to define the scattering matrix by means of a canonical procedure. The definition of the scattering matrix may be illustrated by means of the following diagram. The details of this work will be avialable in the forthcoming report be Avniel and Mitter [AVN 82].

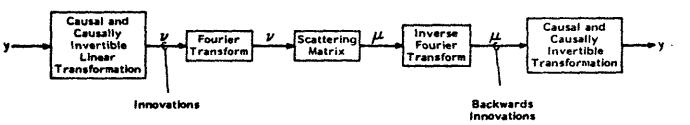


Figure 6.1.

7. Deterministic Analysis of Stability

7.1 Introduction

The presence of relays and logical devices in power system feedback loops at all levels gives rise to the possibility of discontinuous (switching) behavior of the control variables of the system as functions of time. While the methods of stochastic stability provide a more general framework for the evalution of stability of discontinuous systems, they do not in many cases lead to a useful characterization of sample path properties. Yet these are often essential in assessing the validity of a complex model, since they may be readily compared with actual operating records. In most cases of practical interest a deterministic model or models can be derived from a stochastic model by setting noise terms to their mean values or considering each possible value taken by a discrete random variable, etc. A differential-equation or difference-equation model is obtained. Classical stability theory deals with the asympotitc properties of trajectories generated by a system in a given initial state or set of initial states. Classical stability theory for discontinuous systems is not well-developed, and even basic questions of existence and uniqueness need to be resolved; the means for resolving such questions are often different for deterministic and stochastic systems.

In order to highlight the major effects of discontinuities on stability a class of piecewise-constant systems, termed diced systems, have been examined in detail. This analysis, described below, leads to the conclusion that, in addition to the usual issues of stability theory, discontinuities may lead to a fracturing of trajectories and a growth in the number of possible solutions to the dynamic equation. This "chain-reaction" effect has not been traditionally studied by stability theorists: it may be viewed as a new type of instability. In other words, not only the magnitude but also the cardinality of trajectories may be dynamically unstable. This calls into serious question many of the proposed methods for predicting the future behavior of a power system; in the presence of discontinuities, trajectories may be unpredictable.

Other qualitative aspects of deterministic power system models have been studied in related works ([JOH 81a,b], [WIM 81]). These include representation of asynchrony, dynamic modelling of multitasking, and realization theory.

7.2. Stability of Diced Systems - Preliminaries

Diced systems, as defined here, are finite-dimensional, autonomous, continuous-time dynamic systems governed by equations of the form $\frac{dx}{dt}(t) = f(x(t))$; $x_o(t_o) = x_o \epsilon R^n$, $t \geq t_o$, where $f : R^n \to R^n$ is piecewise-constant with discontinuities only on the surfaces where one or more coordinates of R^n take integer values. A diced system in R^2 is very easy to illustrate: the plane can be divided into a uniform gird, and within each square a vector representing the magnitude and direction of f is shown (Figure 1).

Existence and uniqueness of a solution for any fixed initial state, x_o can be studied using a generalization of the method introduced by [FIL 64]; traectories may exhibit sliding mode segments and higher-order non-differentiable behavior as illustrated in Figure 1. In order to obtain existence of solutions, multivalued extensions of f onto its discontinuity surfaces are required. Every trajectory can be represented by a sequence of transition points and times, $\{x(t_i), t_i\}$.

Definitions of various types of stability and instability can be constructed from an examination of the invariant limit sets of the trajectories. For diced systems, the range of asymptotic behavior of trajectories starting from different initial conditions can be exceedingly rich. The possibility of approximate global stability analysis using non-deterministic automata is examined and its limitations are indicated.

In practice, diced systems might be viewed as approximations of continuous or discontinuous systems. In the former case, for instance, we might seek the best piecewise-constant (finite-element) approximation to a continuous system. Wang [3] has presented an application of this type for solving partial differential equations. In the latter case, a state space diffeomorphism might be used first to transform the discontinuities of a system to lie along coordiante axes, and then a diced approximation could be developed which would preserve the discontinuous behavior of such systems. The potential practical advantages of diced approximations lie in a reduction of information storage required to characterize a system and the possibility of assessing its approximate asymptotic behavior without a detailed simulation.

For example, at the time of a known failure of a power system, it is often desirable to predict the long-term consequences of various control strategies so that an operator can decide among them. Yet the system is too big to store all possible consequences in advance. A practice which has thus been followed in some cases [EWA 68] is to run a simulation "faster than real-time" for each control strategy. While the issue of approximation accuracy is not treated here, the results suggest that significant economy of real-time computation might be achieved by approximating the dynamics of a diced system. However, they also suggest that the patterns of stability and instability exhibited by such discontinuous systems may be highly complex and that analytical methods are not likely to yield clear-cut predictions about global stability.

Let $i = [i_1,\ldots,i_n] \epsilon Z^n$ be a multi-index on the n-tuples of integers (Z). Let $b = [b_1,\ldots,b_n] \epsilon B^n$ represent an n-tuple of binary numbers ($B = \{0,1\}$). Let $\chi_i(x)$: $R^n \to R$ be the characteristic function of the open set $\{x = [x_1,\ldots,x_n] \epsilon R^n \mid i_k < x_k < i_k + 1,\ k = 1,2,\ldots,n\}$

Definition: A diced initial value problem (IVP) is specified by a system of ordinary differential equations

$$\dot{x}(t) = f(x(t)) \ ; \ x(t_o) = x_o \epsilon r^n; \ t \geq t_o \qquad (7.2.1)$$

where $f : R^n \to R^n$ has the particular form

$$f(x) = \sum_{i \epsilon z^n} f_{bi} \chi_i(x) \ ; \ b = 0 \epsilon B^n \qquad (7.2.2)$$

and $f_{oi} \epsilon R^n$ for each multi-index i.

The surfaces of discontinuity of f may be classified by their dimension. Let $1(b) : B^n \to \{1,\ldots,n\}$ be a function denoting the number of "1"s in the binary n-tuple b. For fixed $i \epsilon Z^n$, consider the sets

$$S_{bi} = \{x \epsilon R^n \mid i_k < x_k < i_{k+1} \text{ if } b_i = 0 \quad i_k = x_k \text{ if } b_k = 1,\ k = 1,2,\ldots,n\} \qquad (7.2.3)$$

These may be viewed as the set of submanifolds "attached to" the point x = i.* For example S_{oi} is the interior of the n-dimensional cube indexed by its vertex at x=i; S_{1i} (the shorthand 1 denoting b = [1,1,...,1]) is the single point x=i. The submanifolds of dimension p associated with x=i are

$$S_i^p = \{S_{bi} \mid 1(b) = n-p\} \quad p = 0,1,\ldots,n. \quad (7.2.4)$$

This notation provides a compact classification of all of the subsets of R^n which are of interest.

In the next section, conditions for well-posedness of a DIVP are examined. This is done by extending f to its discontinuity surfaces (from $\{f_{oi}\}$, we generate $\{f_{bi}\}$, $b \neq 0 \varepsilon B^n$). Then a constructive procedure can be used to generate solutions $\dot{x}(t) = \phi(t,t_o,x_o)$ for each $x_o \varepsilon R^n$, $t_o \varepsilon R$ and hence to define the transition map ϕ: R x R x R^n → R^n. Let χ denote the function space in which trajectories are defined. This leads to the following

Definition: A diced system is an autonomous dynamical system (χ, R^n, ϕ) (See [WIL 71]).

Stability has been viewed as a qualitative property of a dynamical system, and concerns the asymptotic behaviors of trajectories $x(\cdot) = \phi(\cdot, t_o, x_o)$ as $x_o \varepsilon X$ is varied. Stability of diced systems is discussed in Section IV. Two useful notions will be those of the positive limit set and the invariant set [WIL 70].

Definition: The set $\Omega \subset R^n$ is invariant with respect to the system $\dot{x}(t) = \tilde{f}(x(t),t)$ if for any $x_o \varepsilon \Omega$ there is a t_o such that the motion $\tilde{\phi}(t,t_o,x_o)$ belongs to Ω for all $t \geq t_o$.

Definition: The set $\Pi \subset R^n$ is called the positive limit set of a bounded motion $\phi(t;t_o,x_o)$ if, for any point $p \in \Pi$, there exits a sequence of times $\{t_n\}$ tending to infinity as $n \to \infty$, so that

$$\lim_{n\to\infty} ||\phi(t_n,t_o,x_o) - p|| = 0 \quad (7.2.5)$$

In applying these definitions it will be useful to recall that a function x(t) is periodic of period T > 0 if x(t) = x(t+T) for all t; "the" period of a periodic function is defined as the least T for which this equality holds.

7.3. Existence and Uniqueness

consider the DIVP (7.2.1), (7.2.2). Defining solutions within the cubes S_{oi} by integration is entirely straightforward; all difficulties arise in attempting to extend solutions across the discontinuity surfaces of f; in general, there is no unique continuation. Various possibilities are

(a) To restrict the class of f so that continuations are always unique (this is very restrictive indeed, and essentially eliminates many interesting phenomena from consideration).

(b) To eliminate the non-continuable surfaces from the domain of f; however, then all points on all trajectories leading to such surfaces must also be eliminated, and a large part of the original domain of definition may ultimately be excluded.

(c) To choose an ad hoc rule for continuation of solutions; however, it proves difficult or impossible to do this in a self-consistent and unbiased manner.

A fourth alternative has been selected here:

(d) To sacrifice uniqueness and continue all solutions through a discontinuity.

In this way a viable deterministic existence theory can be developed, at the cost of considering a countable number of alternative solutions. A "physical" justification for adopting this approach is that in the presence of small perturbations of the initial conditions, a solution near to at least one alternative solution will occur.

A constructive procedure is given for defining solutions. To simplify its presentation, a multivalued continuation of f to the surfaces S_{bi}, $b \neq 0$, is first defined. Initially, f is specified on the submanifolds $S_i^n = \{S_{oi}\}$ of dimension n. The continuation proceeds recursively to submanifolds, S_i^p of dimension n-1, n-2,...,0. Recall that S_i^o is the point set $\{x \varepsilon R^n \mid x = i_k, i_k \text{ an integer}\}$. Notationally, a single valued f_{bi} will not be distinguished from a multivalued f_{bi}, the implication being that the prescribed rule is applied to each possible value of f_{bi} in turn, and the set of all results is retained. Let p=n. Suppose f_{bi} are known on S_i^q, $p \leq q \leq n$. Then f_{bi} can be extended to S_i^{p-1} as follows, for each $i \varepsilon Z^n$.

Suppose $S_{bi} \varepsilon S_i^{p-1}$. Let indices $j_j \ldots j_{n-(p-1)}$ denote the ordered nonzero positions of b, i.e., $b_{j_k} = 1$, k=1,..., n-(p-1) and b_j=0 otherwise. The j_k neighborhoods of S_{bi} of dimension q, $p \leq q \leq n$, can be defined as follows. For q=n, consider all indices $\tilde{i}$ formed by decrementing i_{jk} by one for any subset of the subindices k=1,..., n-(p-1), including the null-set; then $S_{\tilde{b}\tilde{i}} \varepsilon S_{\tilde{i}}^{n-1}$ is a neighborhood of S_{bi} where $\tilde{b} \neq 0$. For q = n-1, consider all values b having a single "one" in one of the positions $j_1 \ldots j_{n-(p-1)}$ and for each $\tilde{b}$, form $\tilde{i}$ from the remaining n-(p-1)-1 indices as above; then $S_{\tilde{b}\tilde{i}} \varepsilon S_{\tilde{i}}^{n-1}$ is a neighborhood of S_{bi}. For q = n-2, consider all values b having "ones" in any two of the positions $j_1, \ldots, j_{n-(p-1)}$ and from each $\tilde{b}$ and $\tilde{i}$ from the remaining n-(p-1)-2 indices as above; then $S_{\tilde{b}\tilde{i}} \varepsilon S_{\tilde{i}}^{n-2}$ is a neighborhood of S_{bi}. This procedure is continued until q=p.

The values of f_{bi} on $S_{bi} \varepsilon S_i^{p-1}$ are determined from the values of $f_{\tilde{b}\tilde{i}}$ on each of its neighborhoods $S_{\tilde{b}\tilde{i}} \varepsilon S_{\tilde{i}}^q$, $p \leq q \leq n$. It is thus sufficient to give the procedure for determining f_{bi}, assuming that these values on higher-dimensional submanifolds are known (i.e., the values can be determined recursively). Define $S_{\tilde{b}\tilde{i}}$ to be an input submanifold to S_{bi} if $(f_{\tilde{b}\tilde{i}})_\ell = 0$ for all ℓ such that $\tilde{b}_\ell = 1$, and for all remaining ℓ in the set $j_1,\ldots,j_{n-(p-1)}$, $(f_{\tilde{b}\tilde{i}}) < 0$ for those ℓ such that $\tilde{i}_\ell = i_\ell$, while $(f_{\tilde{b}\tilde{i}}) > 0$ for those ℓ such that $\tilde{i}_\ell = i_\ell - 1$. Define $S_{\tilde{b}\tilde{i}}$ to be an output submanifold if $(f_{\tilde{b}\tilde{i}})_\ell = 0$ for all ℓ such that $b_\ell = 1$, and for all remaining ℓ in the set $j_1,\ldots,j_{n-(p-1)}$, $(f_{\tilde{b}\tilde{i}})_\ell \geq 0$ for those ℓ such that $\tilde{i}_\ell = i_\ell$ while $(f_{\tilde{b}\tilde{i}})_\ell >$ for those such that $\tilde{i}_\ell = i_\ell - 1$. Note that those sets for which $(f_{\tilde{b}\tilde{i}})_\ell \neq 0$ when $\tilde{b}_\ell = 1$ need not be considered. So long as the set of output submanifolds of S_{bi} is non-empty, f_{bi} is assigned the set of all values f_{bi} on the output submanifolds. If the set of output submanifolds is empty, f_{bi} is assigned the set of all values f_{bi} on the output submanifolds. If the set of output submanifolds is empty, S_{bi} is a generalized sliding surface. Consider $f_{\tilde{b}\tilde{i}}$ on $S_{\tilde{b}\tilde{i}} \varepsilon S_{\tilde{i}}^p$ in the input set. If this set is empty, set $f_{\tilde{b}\tilde{i}} = 0$. Recall that $S_{\tilde{b}\tilde{i}} \varepsilon S_{\tilde{i}}^p$ is formed by keeping i unchanged in all but one position, say j_k, of b, so $\tilde{b} = [b_1,\ldots,b_{j_k}-1, 0, b_{j_k+1},\ldots,b_n]$ and either $\tilde{i} = i$ or $\tilde{i} = [i_1,\ldots,i_{j_k}-1,$

*The obvious injection of the integers into the reals is implied.

$i_{j_{k+1}}, i_n]$. Thus there are a maximum of $2(n-(p-1))$ surfaces in this subset of the input set. These surfaces are considered in pairs to determine the admissable values of f_{bi}; using the example above, if $S_{\tilde{b}i}$ is in the input set then $(f_{\tilde{b}i})_{j_k} < 0$ and if $S_{\tilde{\tilde{b}}i}$ is the input set $(f_{\tilde{\tilde{b}}i})_{j_k} > 0$. If both elements are members of the input set then

$$f_{bi} = [(f_{\tilde{\tilde{b}}i})_{j_k} f_{\tilde{b}i} - (f_{\tilde{b}i})_{j_k} f_{\tilde{\tilde{b}}i}]/[(f_{\tilde{\tilde{b}}i})_{j_k} - (f_{\tilde{b}i})_{j_k}] \tag{7.3.1}$$

while if only one is in the input set, let

$$f_{bi} = 0$$

The set of possible values of f_{bi} on a generalized sliding mode is completed by considering each $S_{b\tilde{\tilde{\tilde{b}}}i} \in S^p_{\tilde{\tilde{\tilde{b}}}i}$ in this manner. In all such cases, $(f_{bi})_{j_k}$; $k = 1,\ldots,n-(p-1)$ are zero, so that further motion occurs on S_{bi} itself.

Thus, the procedure for extending the function f to all of R^n is completed. The complexity of the procedure arises from the large number of possibilities which can arise. A number of such special cases are illustrated on Figure 2. Evidently, the procedure for extending f is not the only one which could be devised. In the next step, construction of solutions, however, it will become apparent that the underlying principle has been to define f in a manner which preserves <u>all</u> trajectories that might arise from each initial condition.

Let $x_o \in R^n$ be given as the initial condition of (7.2.1) at $t = t_o$; let $S_{bi} \in S^p_i$ be the smallest submanifold containing x_o. Let f_{bi} denote one of the extended values of f on S_{bi}. Define

$$\phi(\tau,t_o,x_o) = x_o + f_{bi}(\tau - t_o) \; ; \; t_o < \tau \leq t_1 \tag{7.3.2}$$

The time t_1 is defined as follows: for each ℓ such that $(f_{bi})_\ell$ is nonzero, let $(t_1)_\ell$ denote the first $\tau > t_o$ such that $[\phi(\tau,t_o,x_o)]_\ell$ is an integer; then $t_1 = \min[(t_1)_\ell]$ and $t_1 = \phi(t_1,t_o,x_o)$. If $f_{bi} = 0$, then $t_1 = x_o$, and this solution terminates. Otherwise, x_1 defines new values of b, i, and p, and the solution process continues:

$$\phi(\tau,t_k,x_k) = x_k + f_{bi}(\tau - t_k) \; ; \; t_k < \tau < t_{k+1} \tag{7.3.3}$$

On those surfaces where f_{bi} is multivalued, each possibility must be examined in turn; in this sense, ϕ is also multivalued. Each trajectory pieced-together in this fashion can be summarized by a sequence $\{x_k,t_k\}$, $k = 0,1,\ldots$ in some cases, these sequences are finite and in other cases infinite. By inspection of $\{x_k\}$ alone, a corresponding sequence of regions $\{\sigma_k\}$, where $\sigma_k \in \{S_{bi}\}$ is the minimal submanifold containing x_k, can be constructed.

A <u>solution</u> of (7.2.1), (7.2.2) is then defined in the obvious manner, as any $\phi(t,t_o,x_o)$ constructed by the continuation procedure (7.3.3). It has the property that for any finite admissable k, $\phi(t,t_o,x_o)$ is piecewise continuous on $[t_o,t_k]$. This solution by continuation is said to be <u>asymptotic</u> if $\lim_k t_k = \infty$. An asymptotic solution is piecewise continuous. For purpose of the present work, a solution will be said to exist if the state-space continuation is asymptotic.*

*Moreover if $\lim t_k \neq \infty$, solutions by time-continuation could be defined; however, their properties will not be explored here.

Asymptotic solutions need not be unique, but the rate of growth in the number of solutions can be bounded as a function of k, since the maximum number of output submanifolds can be bounded above for any S_{bi}. If there is only one asymptotic solution through (x_o,t_o), it is said to be unique. Continuous dependence of $\phi(t,t_o,x_o)$ with respect to x_o, of course, is not to be expected for $t > t_1$.

7.4. Stability

The usual definitions of stability presuppose a solution which is well-posed in the sense of existence, uniqueness, and continuous dependence on the initial data. Diced systems, in general, do not possess the last two properties. One alternative is to nevertheless use the standard notions of stability, restricting their domain of application to those initial states for which the usual notions of well-posedness are (locally) satisfied. Unfortunately, the set of such initial states appears quite difficult to characterize and thus imposes an awkward restriction on the applicability of this alternative.

Another alternative, introduced here, does not impose such restrictions, but weakens the notion of stability that is employed. Stability is viewed as a qualitative property of a trajectory, and a system is then said to be stable when all of its trajectories share this property.

<u>Definition</u>: The motion of diced system (7.2.1), (7.2.2) initiated at (t_o,x_o) is

$$M(t_o,x_o) = \{\phi(t,t_o,x_o),\ t > t_o \mid \phi \text{ is a transition function initiated at } (t_o,x_o)\}$$

which is the set of all trajectories originating at (t_o,x_o).

<u>Definition</u>: The motion $M(t_o,x_o)$ of a diced system is said to be

(a) <u>Bounded in magnitude</u> if there is a constant $\Phi > 0$ such that

$$\max_{\phi \in M(t_o,x_o)} \{\sup_{t \geq t_o} ||\phi(t,t_o,x_o)||\} < \Phi$$

(b) <u>Bounded in cardinality</u> if there exists a constant N such that

$$\sup_{t>t_o} \{\text{cardinality of } M(t_o,x_o)\} < N$$

The concepts of boundedness in magnitude and cardinality are independent. In both cases, the only difficulties occur at $t \to \infty$, since (a) any $\phi(t,t_o,x_o)$ is by construction bounded for all finite t, and (b) the cardinality of $\phi(t,t_o,x_o)$ is finite, by construction, for all finite t. The following propositions are almost immediate.

<u>Proposition 7.4.1</u>: In (7.2.1), suppose $||f_{io}|| < F$ for all i, then $||\phi(t,t_o,x_o) - x_o|| < F(t-t_o)$ for all $\phi \in M(t_o,x_o)$.

<u>Proof</u>: The extension of f_{oi} to f_{bi} always guaranteed that $||f_{bi}|| < F$, and the construction procedure (3.3) guaranteed that the estimate of the proposition held for each t. q.e.d.

<u>Proposition 7.4.2</u>: Let $|i| = |i_1| + \ldots + |i_n|$. Suppose for system (2.1) there exists $B > 0$ such that for all $|i| > B$, and $k = 1,\ldots,n$, $(f_{oi})_{i_k} < 0$. Then $M(t_o,x_o)$ is bounded in magnitude.

<u>Proof</u>. For any i such that $|i| > B$, every

set S_{bi} contains output submanifolds with the same $|i|$ or smaller $|i|$, and input submanifolds with the same $|i|$ or larger $|i|$; further more, S_{ij} always outputs to $S_{o\tilde{i}}$ with $|\tilde{i}| < |i|$. Thus the construction process cannot terminate for $|i| > B$, and for such i, $|i|$ is reduced at least once every n intervals; hence every solution satisfies $|\phi(t,t_o,x_o)| < B$ for t sufficiently large. Thus $M(t_o,x_o)$ is magnitude-bounded.

Proposition 7.4.3: Suppose that for every $i\epsilon Z^n$, $b\epsilon B^n$, S_{bi} has at most one output submanifold. Then the motion $M(t_o,x_o)$ of (7.2.1), (7.2.2) is bounded in cardinality.

Proof: The extension procedure of Section 7.3 shows that in this case f_{bi} takes the value on its output submanifold or the value zero. If a trajectory enters S_{bi}, it either continues uniquely to the output submanifold, or terminates at S_{bi}. In either case, the cardinality of the solution cannot increase during its construction.

Thus there are two notions of instability for diced systems: solutions may become unbounded in magnitude, and/or they may become unbounded in cardinality. This second form of instability is new: a trajectory can fracture and a chain reaction of subsequent fractures may ensue--the complexity of the process grows without bound.

Next, a notion of stability is put forth. Suppose that the motion $M(t_o,x_o)$ of a diced system is bounded in magnitude and cardinality (or simply "bounded"). Then a set $S \subset R^n$ consisting of a finite union of the submanifolds S_{bi} is termed a positive limit set of a (bounded) trajectory $\phi(t, t_o,x_o)$ if for any point $x\epsilon S$, there exists a sequence of times $\{\tau_k\}$, tending to infinity as $k\to\infty$, so that

$$\lim_{k\to\infty})\phi(\tau_k,t_o,x_o) - x(= 0 \qquad (7.4.1)$$

where $)(\cdot)($ denotes the set-membership metric, i.e., if $x\epsilon S_{bi}$,

$$)y-x(= \begin{cases} 0 & y\epsilon S_{bi} \\ 1 & y\notin S_{bi} \end{cases}$$

In applying this definition, it is important to recall the standing assumption from Section 7.3, that all trajectories are asymptotic, so that such sequences $\{\tau_k\}$ exist.

Definition: A bounded motion $M(t_o,x_o)$ of a diced system is termed pointwise stable if all trajectories $\phi(t,t_o,x_o)\epsilon M(t_o,x_o)$ have the same positive limit set. The motion is locally stable for $x_o\epsilon S_{bi}$ if all trajectories $\phi(t,t_o,x)\epsilon M(t_o,x)$, $x\epsilon S_{bi}$, have the same positive limit set. The motion is globally stable if all trajectories $\phi(t, t_o,x)$ have the same positive limit set.

Concepts of uniform stability will not be discussed since only time-invariant diced systems are considered in the present account.*

*The results could be extended in this direction for systems with continuous time-variation; however discontinuously time-varying systems may not be continuable, as Filippov pointed out.

In fact, the evaluation of stability, according to the definitions given, can be based merely on knowledge of the sequence $\{\sigma_k\}$ of submanifolds containing $\{x_k\}$, since it is known from the construction procedure that $t_{k+1} > t_k$ and from the asymptotic assumption that $\lim_{k\to\infty} t_k = \infty$. This suggests that a way to generate the sequence $\{\sigma_k\}$ autonomously without explicit integration and generation of $\{x_k, t_k\}$ would be particularly valuable in the assessment of stability. This has not been achieved yet.

Knowledge of the time-structure $\{t_k\}$ of individual solutions can be of further value in refining stability notions. To simplify the remaining concepts it is now assumed that the trajectories are uniquely-defined (e.g., as occurs in Proposition 7.4.3) and bounded. Suppose Π is a positive limit set of such a solution in the conventional sense of Section 7.2 (eq. 7.2.5)). Then in the usual manner it can be shown that Π is bounded, closed, non-empty and invariant, the last property being a consequence of time-invariance. In fact, as a consequence of finite-dimensionality of R^n, all such solutions are asymptotically almost-periodic [DAF 74]. Two cases of special interest are the asymptotically constant (equilibrium) solution and the asymptotically periodic solution. These can be identified directly from the sequence $\{x_k, t_k\}$ characterizing $\phi(t,t_o,x_o)$.

Proposition 7.4.4: If the sequence $\{x_k,t_k\}$ is finite of length N, the positive-invariant limit set consists of one point, the last value x_N (for which $t_N = \infty$). If the sequence $\{x_k,t_k\}$ is jointly periodic of period m for $k > N$, then the positive-invariant limit set is a cycle (closed curve) in R^n.

Proof: For the first case, note that the construction procedure automatically defines $t_N = \infty$ when the sequence is finite, and this implies a constant solution for $t \geq t_N$. In the second case, note that since $\{x_k,t_k\}$ completely specify $\phi(t,t_o, x_o)$, ϕ must be periodic $t_{k+m}-t_k$, $k > N$, whenever $\{x_k,t_k\}$ is periodic (in fact, the solution is a linear interpolation between these points).

It is interesting to note that for diced systems, the establishment of an equilibrium or periodic solution after a finite time (t_N) is often to be expected (whereas this would be considered exceptional in the case of continuous differential equations); however, in some cases almost periodic solutions may also exist.

7.5. Discussion and Conclusions

The present account of the stability of diced systems leaves a number of questions unanswered and raises some new ones. A study of methods for temporal continuation of non-asymptotic solutions is needed; such solutions may represent a new sort of sliding mode which can arise in higher dimensional spaces, as suggested by an example of [UTK 78]. The possibility of extending the techniques developed here to time-varying systems has been mentioned; Filippov's general existence results apply to this problem. A study of the partitioning of initial states which is implied by the proposed stability definition would also be fruitful; what properties are shared by initial state sets giving rise to the same asymptotic solution? In general, it would appear that the initial states

within a given region S_{bi} can ultimately end up widely dispersed. The possibility of using an automaton to simplify the propagation of solutions has also been raised. The approximation of continuous systems by diced systems has not been explored, but under appropriate conditions, a bound on the approximation error should be achievable.

In spite of the questions that are unanswered, some modest progress has been made toward defining the stability properties of diced systems. First, a constructive continuation procedure for higher dimensions has been found; the problem readily evades one's intuition above n = 1,2 and even 3 as endless combinations of difficult situations may occur. Second, a compromise on the issue of uniqueness has been put forth: the number of admissable solutions at any finite time is bounded. Third, the concepts of stability have been generalized to provide meaningful criteria for discontinuous systems of diced type.

Returning to the electric power system example cited in the opening section, it would appear that the implications of the research might be very disturbing, for two primary reasons. First, a new type of instability--an unbounded growth in the number of possible solutions with time-has been identified. Second, and independently, the partitioning of the initial state--at least in worst-case situations--based on asymptotic properties, appears to be very fine and irregular; thus a small perturbation in the initial state may give rise to completely different asymptotic behavior than is found for the unperturbed initial state. Both of these phenomena imply that the future behavior of a diced system with a (approximately) specified initial state may be fundamentally unpredictable; if the long-term future consequences of a present control policy are unpredictable, the problem of choosing the best policy becomes more difficult and planning must be done with a shorter horizon.

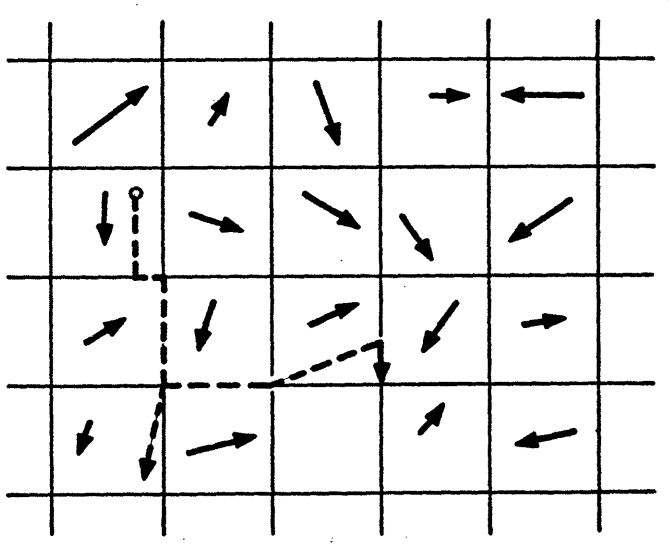


Figure 7.1: A diced System in R^2.

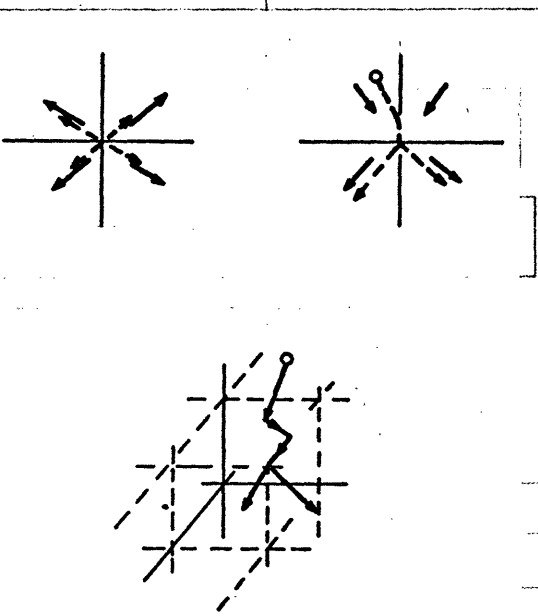


Figure 7.2: Illustration of Nonuiqueness of Solutions

REFERENCES

[ABR 78] Abraham, R., and Marsden, J. Foundation of Mechanics, 2nd ed., Addison-Wesley, Reading, MA 1978.

[AND 77] Anderson, P.M. and Fouad, A.A., Power System Control and Stability, Iowa State Univ. Press, Ames, Iowa.

[AND 80] Anderson, B.D.O., "Nonlinear Networks and Onsager-Casimer Reversibility," IEEE Trans. Circuits and Systems, 27, 1051-1058, 1980.

[AND 81] Anderson, B.D.O. and Johnson, C.R., (1981), "Exponential Convergence of Adaptive Identification and Control Algorithms," Automatica, September.

[ARA 80] Arapostathis, A., Sastry, S. and Varaiya, P., "Analysis of the Power Flow Equation," ERL Memorandum M80/35, U. of Calif., Berkeley, August.

[ARA 81a] Arapostathis, A., Sastry, S., and Varaiya, P., "Analysis of the Power Flow Equation", Int. Journal of Electrical Power and Energy Systems, Vol. 3 (1981) pp. 115-126.

[ARA 81b] Arapostathis, A., Sastry, S., and Varaiya, P., "Global Analysis of Swing Dynamics", ERL Memo UCB/ERL 81/3, University of California, Berkeley.

[AST 73] Astrom, K.J., and Wittenmark, B., (1973), "On Self-Tuning-Regulators," Automatica, No. 8, pp. 185-199.

[AVN 82] Avniel, Y. and Mitter, S.K., "Scattering Theory for Deterministic and Stochastic Linear Systems," to appear.

[BIR 78] Birdwell, J.D., Castanon, D. and Athans, M., "On Reliable Control System Designs With and Without Feedback Reconfiguration," Proc. 1978 IEEE Conf. on Decision and Control, San Diego, CA.

[BIS 81] Bismut, J.M., Stochastic Mechanics, Springer Lecture Notes in Mathematics, Vol. 866, 1981. Springer-Verlag,Berlin, New York.

[BIT 80a] Bitmead, R.R. and Anderson, B.D.O., (1980a), "Lyapunov Techniques for the Exponential Stability of Linear Difference Equations with Random Coefficients," IEEE Trans. Automat. Contr., Vol. AC-25, pp. 782-787, August.

[BIT 80b] Bitmead, R.R. and Anderson, B.D.O, (1980b), "Performance of Adaptive Estimation Algorithms in Dependent Random Environments," IEEE Trans. Automat. Contr., Vol. AC-25, pp. 788-793, August.

[BOR 65] Borovkov, A.A., "Some Limit Theorems in the Theory of Mass Services, II," Theory Probability Appl., Vol. 10, pp. 375-400.

[BOS 65] Bose, A.G. and Stevens, K.N., Introductory Network Theory, Harper and Row, New York, 1965.

[BRO 79] Brockett, R.W. and Willems, J.C., "Stochastic Control and the Second Law of Thermodynamics," Proc. of the 1978 Conf. on Decision and Control, San Diego, Calif., 1007-1011, Jan. 1979.

[BUR 79] Burman, D.Y., "An Analytic Approach to Diffusion Approximations in Queueing," Ph.D. Thesis, Courant Inst. Math. Sci., New York.

[BUZ 78] Buzacott, J.A. and Hanifin, L.E., "Models of Automatic Transfer Lines with Inventory Banks - A Review and Comparison," AIIE Trans., 10, 2, 197-207.

[CAI 80] Caines, P.E. and Dorer, D. (1980), "Adaptive Control of Systems Subject to a Class of Random Parameter Variations and Disturbances," Proc. of 1981 IFAC Congress, Japan, August 1981.

[CAL 60] Callen, H.B., Thermodynamics, Wiley, New York, 1960.

[CAM 79] Campbell, S.L. "Singular Perturbation of Autonomous Linear Systems", SIAM J. Math. Anal., Vol. 10, No. 3, pp. 542-551, 1979.

[CHI 79] Chizeck, H.J. and Willsky, A.S., "Towards Fault-Tolerant Optimal Control, Proc. 1978 IEEE Conf. on Decision and Control, San Diego, CA.

[CHI 80a] Chizeck, H.J. and Willsky, A.S., "Jump Linear Quadratic Problems with State-Independent Rates," Rpt. LIDS-R-1053, Lab. for Information and Decision Systems, MIT, Camb., MA., Oct.

[CHI 80b] Chizeck, H.J., Castanon, D.A. and Willsky, A.S., "Detailed Analysis of a Scalar Control Problem Containing Abrupt Changes with State Dependent Rates," M.I.T Lab. for Information and Decision Systems, Rpt. LIDS-R-1057, Camb., MA. Dec.

[CHI 82] Chizeck, H.J., Fault-Tolerant Optimal Control, Ph.D thesis, M.I.T. Dept. of Elec. Eng. and Comp. Sci., Camb.,Mass., to be completed in 1982.

[COD 81] Coderch, M., Sastry, S.S, Willsky, A.S., and Castanon, D.A., "Hierarchical Aggregation of Linear Systems with Multiple Time Scales," Proc. of the 20th IEEE Conference on Decision and Control, San Diego, California, Dec. 1981.

[COD 82] Coderch, M., "Multiple Time Scale and Aggregation Methods with Filtering Applications", Ph.D. Thesis MIT-LIDS (to be submitted June 1982).

[CRU 79] Cruz, J.B. (1979), Workshop on Adaptive Control, R-79-1, Champaign, Illinois, May, sponsored by AFOSR.

[CUR 77] Curtois, J.P., "Decomposability: Queueing and Computer Systems Applications," Academic Press, New York, 1977.

[DAF 74] Dafermos, C.M., "Semiflows Associated with Compact and Uniform Processes," Math. Sys. Theory, Vol. 8, No. 2.

[DAY 71] Day, W.A., "Time-Reversal and the Symmetry of the Relaxation Function of a Linear Viscoelastic Material," Arch. Rational Mech. Anal., 40, 155-159, 1971.

[DEL 81] Delebecque, F., and Quadrat, J.P., "Optimal Control of Markov CHains Admitting Strong and Weak Interactions," Automatica, Vol. 17, No. 2, pp. 281-296, 1981.

[DOE 38] Doeblin, W., "Sur l'Equation Matricielle $A(t+s) = A(t)\cdot A(s)$ et ses Applications aux Probabilites en Chaine", Bulletin des Sciences Mathematiques, (2), Vol. 62, pp. 21-32, 1938.

[DOO 42] Doob, J.L., "Topics in the Theory of Markoff Chains", Trans. Amer. Math. Soc., Vol. 42, pp. 37-64, 1942.

[DOY 81] Doyle, J.C. and Stein, G. (1981) "Multivariable Feedback Design: Concepts for a Classical Modern Synthesis," IEEE Trans. Automat. Contr., Vol. AC-26, No. 1, pp. 4-17, February.

[EGA 79] Egardt, B. (1979), "Unification of Some Continuous-Time Adaptive Control Schemes," IEEE Trans. on Auto. Control, Vol. AC-24, No. 4, pp. 588-592, August.

[EVA 78] Evans, F.J. (1978), "Prospects for Dynamic Security Monitoring in Large Scale Electric Power Systems," Proc. 1978 IFAC, Helsinki, Finland, June.

[EWA 68] Ewart, D.N., "Whys and Wherefores of Power System Blackouts," IEEE Spectrum, April 1968, pp. 36-41.

[FEU 78] Feuer, A. and Morse, S. (1978), "Adaptive Control of Single-Input Single-Output Linear Systems," IEEE Trans. Automat. Contr., Vol. AC-23, pp. 557-570, August.

[FIL 64] Filippov, A.F., "Differential Equations with Discontinuous Right-Hand Sides," Amer. Math. Soc. Trans., Ser. 2, Vol. 42, 1964, pp. 199-231.

[FLE 75] Fleming, W., and Rishel, R., Deterministic and Stochastic Optimal Control, Springer-Verlag, N.Y. 1975.

[FLE 81a] Fleming, W.H., "Stochastic Calculus of Variations and Mechanics," LIDS-P-1070, M.I.T. Laboratory for Information and Decision Systems, Submitted to J. of Optimization Theory, 1981

[FLE 81b] Fleming, W.H., and Mitter, S.K., "Optimal Control and Nonlinear Filtering for Nondegenerate Diffusion Processes" Submitted to Stochastics.

[FRI 75] Friedman, A., Stochastic Differential Equations and Applications, Vol. I (1975), Vol. II (1976), Academic Press, New York.

[GER 80] Gershwin S.B., and Schick, I., "Modeling and Analysis of Two- and Three- Stage Transfer Lines with Unreliable Machines and Finite Buffers," LIDS-R-979, Lab. for Information and Decision Systems, MIT, Camb., MA.

[GER 81] Gershwin, S.B., and Berman, O., "Analysis of Transfer Lines Consisting of Two Unreliable Machines with Random Processing Times and Finite Storage Buffers," AIIE Transactions, Vol. 13, No. 1, 1981, pp. 2-11.

[GIH 75] Gihman, I.I., and Skorokhod, A.V., The Theory of Stochastic Processes II, Springer-Verlag, Berlin 1975.

[GOO 79a] Goodwin, G.C., Ramadge, P.J., Caines, P.E. (1979), "Discrete Time Stochastic Adaptive Control," Proc. Conf. on Decision and Control, Ft. Lauderdale, FA, December.

[GOO 79b] Goodwin, G.C. and Sin, K.S. (1979), "Stochastic Adaptive Control Using a Modified Least Squares Algorithm," Tech. Rept. EE 7907, U. of Newcastle, Australia, June.

[GOO 80] Goodwin, G.C., Ramadge, P.J. and Caines, P. E.(1980), "Discrete Time Multivariable Adaptive Control," IEEE Trans. Auto. Control, Vol. AC-25, No. 3, pp. 449-456, June.

[GRE 78] Greene, C.S., "An Analysis of the Multiple Model Adaptive Control Algorithm, "Ph.D. Thesis, M.I.T. Dept. of Elec. Eng. and Comp. Sci., Aug. 1978.

[HAL 77] Hale, J.K., "Generic Bifurcation with Applications", in Nonlinear Analysis and Mechanics, Heriot-Watt Symposium, v.1, Pitman 1977.

[HAR 78] Harrison, J.M., "The Diffusion Approximation for Tandem Queues in Heavy Traffic," Advances in Applied Probability, Vol. 10, No. 4, pp. 886-905, Dec.

[HAR 79] Harrison, J.M. and Reiman, M., "On Reflected Brownian Motion," personal communication.

[HER 74] Hersh, R. "Random Evolutions: A Survey of Results and Problems," Rocky Mountain Math. J., Vol. 14, pp. 443-496.

[HIJ 80] Hijab, O., "Minimum Energy Estimation" Ph.D. Dissertation, University of California, Berkeley, December 1980.

[HIJ 81] Hijab, O., "The Adaptive LQG Problem", submitted to the IEEE Trans. Aut. Control.

[HIL 74] Hill, T.L., "Theoretical Formalism for the Sliding Filament Model of Contraction of Striated Muscle: Part I," Prog. Boiphys. Mol. Biol., 28, 267-340, 1974.

[HIL 75] Hill, T.L., "Theoretical Formalism for the Sliding Filament Model of Contraction of Striated Muscle: Part II," Prog. Biophys. Mol. Biol., 29, 105-159, 1975.

[HIL 77] Hill, T.L., Free Energy Transduction in Biology Academic Press, New York, 1977.

[HIR 75] Hirsch, M.W. and Smale, S., Differential Equations, Dynamical Systems, and Linear Algebra, Academic Press, 1975.

[HUN 56] Hunt, G.C., "Sequential Arrays of Waiting Lines," Op. Res., 4, 6, pp. 674-684.

[HUX 74] Huxley, A.F., "Review Lecture: Muscular Contraction," J. Physiol. 243, 1-43, 1974.

[IGL 70] Iglehart, D.L. and Whitt, W., "Multiple Channel Queues in Heavy Traffic, I," Advances in Applied Prob., Vol. 2, pp. 150-177.

[JOH 80] Johnson, T.L., "Stability of Diced Systems," Proc. 19th IEEE Conf. on Decision and Control, pp. 1110-1115, Dec. 1980, (Also M.I.T. LIDS-P-980, March, 1980).

[JOH 81a] Johnson, T.L. "Analytic Models of Multitask Processes," Proc. 20th IEEE Conf. on Decision and Control, (to appear), Dec. 1981.

[JOH 81b] Johnson, T.L., "Multitask Control of Distributed Processes" Proc. 3rd IFAC Symposuim on Distributed Parameter Systems, Toulouse, France, June 1981 (to appear).

[JUL 75] Julian, F.J. and Salleni, M.R., "Variation of Muscle Stiffness with Force at Increasing Speeds of Shortening," J. Gen. Physiol., 66, 287-302, 1975.

[KAC 63] Kac, M., Uhlenbeck, G.E., Hemmer, P.C., "On the Van der Waals Theory of the Vapor-Liquid Equilibrium," Journal of Mathematical Physics 4, #2, February 1963, pp. 216-247.

[KAC 69] Kac, M., "The Physical Background of Langevin's Equation," in: Aziz, A.K. (ed.), Lecture Series in Differential Equations: Vol. II, Van Nostrand, New York, 1969.

[KAT 65] Katchalsky, A. and Curran, P.F., Nonequilibrium Thermodynamics in Biophysics, Harvard Univ. Press, Cambridge, 1965.

[KAT 66] Kato, T., "Perturbation Theory for Linear Operators," Springer Verlag, New York, 1966 (2nd edition 1976).

[KEI 72] Keizer, J., "On the Solutions and the Steady States of a Master Equation," J. Stat. Phys., 6, 67-72, 1972.

[KEI 79] Keilson, J., "Markov Chain Models -- Rarity and Exponentiality, Springer-Verlag, New York, 1979.

[KEM 60] Kemeny, J.G., Snell, J.L., "Finite Markov Chains," Springer Verlag, 2nd ed., New York, 1976.

[KEN 78] Kent, J., "Time-Reversible Diffusions," Adv. Appl. Prob., 10, 819-835, 1978.

[KHA 63] Khasminskii, R.Z., "The Averaging Principle for Parabolic and Elliptic Differential Equations and Markov Processes with Small Diffusion," Theory Probability Appl., Vol. 8, pp. 1-21.

[KHA 66a] Khasminskii, R.Z., "On Stochastic Processes Defined by Differential Equations with a Small Parameter," Theory Probability Appl., Vol. 11, No. 2, pp. 211-228.

[KHA 66b] Khasminskii, R.Z., "A Limit Theorem for the Solutions of Differential Equations with Random Right-Hand Sides," Theory Probability Appl., Vol. 11, No. 3, pp. 390-406.

[KOB 74] Kobayashi, H., "Application of the Diffusion Approximation to Queueing Networks, I: Equilibrium Queue Distributions," J. Assoc. Computing Machinery, Vol. 21, No. 2, April, pp. 316-328.

[KOK 80] Kokotovic, P.V., ed., "Singular Perturbations and Time Scales in Modelling and Control of Dynamic Systems", Univ. of Illinois, Coordinated Science Lab., Rep. #R.901.

[KOR 78] Korolyuk, V.S., and Turbin, A.F., "Mathematical Foundations of Complex Systems Aggregation", Nauka, Moscow, 1978 (In Russian).

[KRA 61] Krasovskii, N.N. and Lidskii, E.A., "Analytical Design of Controllers in Systems with Random Attributes, I-III," Automation and Remote Control, Vol. 22, pp. 1021-1025, 1141-1146, 1289-1294.

[KRY 80] Krylov, N.V., Controlled Diffusion Processes, Springer-Verlag, N.Y., 1980.

[KUN 67] Kunita, H., and Watanabe, S., "On Square Integrable Martingales" Nagoya Math. Journal, Vol. 30, 1967.

[KUN 81] Kunita, H. Lectures Given at Interational School on Non-Linear Filtering and Stochastic Control, Cartona, Italy, July 1981.

[KUS 76] Kushner, H.J., "Probabilistic Methods for Finite Difference Approximations to Degenerate Elliptic and Parabolic Equations with Neumann and Dirichlet Boundary Conditions", Journal of Math. Analysis and Appl., Vol. 53, 1976.

[LAN 79a] Landau, I.D., and Silveira, H.M. (1979), "A Stability Theorem with Applications to Adaptive Control," IEEE Trans. Auto. Control, Vol AC-24, No. 2, pp. 305-312, April.

[LAN 80] Landau, I.D., (1980), "An Extension of a Stability Theorem Applicable to Adaptive Control," IEEE Trans. Autom. Contr., Vol. AC-25, pp. 814-817, August.

[LAV 78] Lavenda, B.H., Thermodynamics of Irreversible Processes, Wiley, New York, 1978.

[LAX 67] Lax, P.D., and Phillips, R.S., Scattering Theory, Academic Press, New York, N.Y. 167.

[LJU 77] Ljung, L. (1977) "Analysis of Recursive Algorithms," IEEE Trans. Auto. Control, Vol. AC-22, No. 4, pp. 551-575, August.

[MAT 77] Matkowsky, B.J., and Schuss, Z., "The Exit Problem for Randomly Perturbed Dynamical Systems" SIAM J. Appl. Math., Vol. 33, No. 2, 1977.

[MAX 1875] Maxwell, J.C., "On the Dynamical Evidence of the Molecular Constitution of Bodies", Nature 11, (1875), pp. 357-374.

[MIT 82] Mitter, S.K., "Conservative and Dissipative Stochastic Systems", to appear.

[MOR 80] Morse, A.S., (1980), "Global Stability of Parameter Adaptive Control Systems," IEEE Trans. Automat. Contr., Vol. AC-25, pp. 433-440, June.

[NAG 70] Sz-Nagy, B., and Foias, C., Harmonic Analysis of Operators on Hilbert Space, North Holland, Amsterdam-London, 1970.

[NAR 78] Narendra, K.S., and Valavani, L.S., (1978) "Stable Adaptive Controller Design-Direct Control", IEEE Trans. Automat. Contr., Vol. AC-23, pp. 570-583, August.

[NAR 79] Narendra, K.S., and Valavani, L.S. (1979), "Direct and Indirect Model Reference Adaptive Control," Automatica, Nov. 1979.

[NAR 80a] Narendra, K.S., Lin, Y.H., and Valavani, L.S., (1980), "Stable Adaptive Controller Design, Part II", IEEE Trans. on Automatic Control, Vol. AC-25, No. 3, pp. 440-449, June.

[NAR 80b] Narendra, K.S., and Lin, Y.H., (1980), "Stable Discrete Adapative Control," IEEE Trans. Automat. Contr., Vol. AC-25, pp. 456-461, June.

[NAY 73] Nayfeh, A.H., "Perturbation Methods", Wiley, New York, 1973.

[NEL 67] Nelson, E., Dynamical Theories of Brownian Motion, Princeton Univ. Press, Princeton, 1967.

[NEV 75] Neveu, J. (1975), Discrete Parameter Super-Martingale, North Holland, N.Y.

[NIC 77] Nicolis, G. and Prigogine, I. Self Organization in Non-Equilibrium Systems, Wiley, New York, 1977.

[PAP 77] Papanicolaou, G., Stoock, D., Varadhan, S.R.S., " Martingale Approach to Some Limit Theorems", in Statistical Mechanics, Dynamical Systems, and the Duke Turbulence Conference, ed.D. Ruelle, Duke University Math. Series, Durham, NC 1977.

[PAR 67] Parthsarathy, K.R., Probability Measures on Metric Spaces, Academic Press, NY 1967.

[POS 78] Poston, T., and Stewart, I.W., Catastrophe Theory and Its Applications, Pitman, 1978.

[PRO 81] Propp, M.B., A Biochemically- and Structurally- Based Model of Striated Muscle Contraction, S.M. dissertation, Department of Electrical Engineering and Computer Science, M.I.T., Cambridge, MA., May 1981.

[REI 77] Reiman, M.I., "Queueing Networks in Heavy Traffic", Tech. Report 46, Dept. of Oper. Res., Stanford University, September.

[ROH 80a] Rohrs, C., Valavani, L., Athans, M. (1980a), "Convergence Studies of Adaptive Control Algorithms, Part I: Analysis,".Proc. 1980 Cont. on Decision and Control, Albuquerque, NM.

[SAF 76] Safonov, M.G. and Athans, M. (1976), "Gain and Phase Margins for Multiloop LQG Regulators," Proc. 1976 IEEE Conf. on Dec. & Control, Clearwater, Florida.

[SAF 78] Safonov, M.G. and Athans, M. (1978), "Robustness and Computational Aspects of Nonlinear Stochastic Estimators and REgulators," IEEE Trans. Auto. Control, Vol. AC-23, pp. 717-725, August.

[SAF 80] Safonov, M.G. (1980), Robustness and Stability of Stochastic Multivariable Feedback System Design, M.I.T Press, Cambridge.

[SAS 80] Sastry, S., and Varaiya, P., "Hierarchical Stability and Alert State Steering Control of Interconnected Power Systems", IEEE Trans. CAS Vol. 27 (1980) pp. 1102-1112.

[SAS 81] Sastry, S.S., Desoer, C.A., and Varaiya, P.P., "Jump Behavior of Circuits and Systems," ERL Memo No. M80/44, University of California Berkeley (To appear IEEE Trans. Circuits and Systems, Dec. 1981).

[SCH 79] Schuss, Z., and Matkowsky, B.J., "The Exit Problem: A New Approach to Diffusion Across Potential Barriers" SIAM J. Applied Math., Vol. 36, June 1979.

[SHO 80] Shomber, H.R., "An Extended Analysis of the Multiple Model Adaptive Control Algorithm," S.M. Thesis, M.I.T. Dept. of Elec. Eng. and Comp. Sci.; also Rept. LIDS-TH-973 of the M.I.T. Lab. for Inf. and Dec. Sys; Feb. 1980.

[SIM 61] Simon, H.A., and Ando, A., "Aggregation of Variables in Dynamic Systems," Econometrica, Vol. 29, No. 2, pp. 111-138, 1961.

[SOL 79] Solo, V. (1979) "The Convergence of AML", IEEE Trans. Auto. Control, Vol. AC-24, No. 6, pp. 958-963.

[SOY 79] Soyster, A.L., Schmidt, J.W., and Rohrer, M.W., "Allocation of Buffer Capacities for a Class of Fixed Cycle Production Lines," AIEE Trans. Vol. 11, No. 2, June.

[STR 69] Strook, D. and Varadhan, S.R.S., "Diffusion Processes with Continuous Coefficients, I", Comm. Pure and Applied Math. XXII, pp. 345-400, (1969).

[STR 71] Stoock, D. and Varadhan, S.R.S., "Diffusion Processes with Boundary Conditions", Comm. Pure Applied Math., XXIV, 147-225 (1971).

[STR 75] Stroock, D., "Diffusion Processes Associated with Levy Generators", Z. Wahrscheinlichkeitstheorie verw. Gebiete, 32, 209-244 (1975).

[STR 79] Stroock, D.W., and Varadhan, S.R.S, Multidimensional Diffusion Processes, Springer Verlag, New York.

[SUZ 64] Suzuki, T., "On a Tandem Queue with Blocking," J. Op. Res. Japan, Vol. 6, pp. 137-257.

[SWO 69] Sworder, D.D., "Feedback Control of a Class of Linear Systems with Jump Parameters," IEEE Trans. Auto. Control, Vol. AC-14, No. 1., pp. 9-14.

[THO 75] Thom, R., Structural Stability and Morphogenesis, Benjamin 1975.

[TOL 38] Tolman, R.C., The Principles of Statistical Mechanics, Oxford Univ. Press, New York, 1938.

[UTK 78] Utkin, V.I., Sliding Modes and Their Application in Variable Structure Systems, Mir Publishers, Moscow.

[VAL 80] Valavani, L.S. (1980), "Stability and Convergence of Adaptive Control Algorithms: A Survey and Some New Results," Proc. 1980 JACC, San Francisco, CA.

[VAS 76] Vasileva, A.B., "The Development of the Theory of Ordinary Differential Equations with a Small Parameter Multiplying the Highest Derivatives During the Period 1966-1976", Russian Math. Surveys, Vol. 31, No. 6, pp. 109-131, 1976.

[VEN 70] Ventsel, A.D. and Freidlin M.I., "On Small Random Perturbations of Dynamical Systems," Russian Math. Surveys, Vol. 25, pp. 1-55.

[WAN 68] Wang, P.K.C., "A Method for Approximating Dynamical Processes by Finite-State Systems," Int. J. Control, Vol. 8, No. 3, 1968, pp. 285-296.

[WAT 71] Watanabe, S., "On Stochastic Differential Equations for Multi-dimensional Diffusion Processes with Boundary Conditions, I & II," J. Math. Kyoto University, Vol. II, pp. 169-180, pp. 545-551.

[WIL 70] Willems, J.L., Stability Theory of Dynamical Systems, Nelson and Sons, Ltd., London, 1970.

[WIL 71] Willems, J.C. and Mitter S.K., "Controllability, Pole Allocation and State Reconstruction," IEEE Trans. Auto. Control, Vol. AC-16, No. 6, pp. 582-595.

[WIL 72] Willems, J.C., "Dissipative Dynamical Systems, Part I: General Theory, Part II: Linear Systems with Quadratic Supply Rates," Arch. Rational Mech. Anal., 45, 321-393, 1972.

[WIL 76] Willems, J.C., "Lyapunov Functions for Diagonally Dominant Systems," Automatica, 12, 519-523, 1976.

[WIL 79] Willems, J.C., "System Theoretic Models For the Analysis of Physical Systems" Ricerche di Automatica, Vol. 10, No. 2, 1979 pp. 71-106.

[WIL 80a] Willems, J.C., "Feedforward Control, PID Control Laws, and Almost Invariant Subspaces", MIT-LIDS Report # P-1120, Aug. 1981.

[WIl 78] Willsky, A.S., et. al., "Stochastic Stability Research for Complex Power Systems," Research Proposal for DOE Contract No. ET-76-C-01-2295.

[WIL 80b] Willsky, A.S., et. al., "Midterm Report for DOE Contract ET-76-A-01-2295", M.I.T. Laboratory for Information and Decision Systems, Rept. No. LIDS-R-1057, Nov. 1980.

[WIM 81] Wimpey, D.N., Johnson, T.L., and Kaliski, M.E., "Synthesis of A/D and D/A Coders", Proc. 1981 Joint Auto. Control Conf., June 1981 (Session WP-5).

[WON 71] Wong, E., Stochastic Processes in Information and Dynamical Systems, McGraw-Hill, New York.

[WON 70] Wonham, W.M., "Random Differential Equations in Control Theory," in Probabilistic Methods in Applied Mathematics, Vol. 2, A.T. Bharucha-Reid, ed., Academic Press, N.Y. pp. 131-212.

[YAM 71] Yamada, R. and Watanabe, S., "On the Uniqueness of Solutions of Stochastic Differential Equations," J. Math. Kyoto Univ., Vol. 11, pp. 155-167.

[ZEE 72] Zeeman, E.C., "Differential Equations for the Heart Beat and Nerve Impulse", in Towards a Theoretical Biology 4, ed. C.H. Waddington, Edinburgh University Press 1972.

PUBLICATIONS

The following papers were supported in part by DOE Contract ET-76-C-01-2295.

1. M.H.A. Davis, "Martingale Methods in Stochastic Control," Proc. Workshop on Stochastic Control Theory and Stochastic Differential Systems, Univ. of Bonn, Jan. 1979, to be published in the Springer-Verlag Lecture Notes in Control and Systems Sciences Series, M. Kohlmann, editor.

2. M.H.A. Davis, "A Pathwise Solution of the Equations of Nonlinear Filtering," Teoria Veroyatnostei i ee Primeneniya, to appear.

3. S.I. Marcus, "Modeling and Approximation of Stochastic Differential Equations Driven by Semimartingales," Stochastics, Vol. 4, No. 3, 1981, pp. 233-245.

4. T.L. Johnson, "Stability of Diced Systems," Proc. of 19th IEEE Conference on Decision and Control and the IEEE Transactions on Automatic Control, Dec. 1980; also M.I.T. Rept. LIDS-P-980.

5. H.R. Shomber, "An Extended Analysis of the Multiple Model Adaptive Control Algorithm," Report No. LIDS-TH-973, M.I.T. Laboratory for Information and Decision Systems, Cambridge, Mass., Feb. 1980.

6. D.A. Castanon, M. Coderch, B.C. Levy, and A.S. Willsky, "Asymptotic Analysis, Approximation, and Aggregation Methods for Stochastic Hybrid Systems," Proc. 1980 Joint Automatic Control Conf., San Francisco, August 1980.

7. D.A. Castanon, H.J. Chizeck, and A.S. Willsky, "Discrete-Time Stochastic Control of Hybrid Systems," Proc. 1980 Joint Automatic Control Conf., San Francisco, August 1980.

8. B.C. Levy, D.A. Castanon, and S.B. Gershwin, "The Asymptotic Analysis of Unreliable Distribution Networks," 19th IEEE Conference on Decision and Control, Albuquerque, N.M., December 1980.

9. D.A. Castanon, "Reverse Time Diffusion Processes," IEEE Trans. on Information Theory, accepted for publication.

10. H.J. Chizeck and A.S. Willsky, "Jump Linear Quadratic Problems with State-Independent Rates," Rept. LIDS-R-1053, M.I.T. Laboratory for Information and Decision Systems, Cambridge, Mass., Oct. 1980.

11. J.C. Willems, "Feedforward Control, PID Control Laws, and Almost Invariant Subspaces," Systems and Control Letters, to appear.

12. M. Coderch, A.S. Willsky, S.S. Sastry, and D.A. Castanon, "Hierarchical Aggregation of Linear Systems with Multiple Time Scales," to be submitted to IEEE Trans. Automatic Control.

13. M. Coderch, A.S. Willsky, S.S. Sastry, and D.A. Castanon, "Multiple Time Scale Behavior and Aggregation of Finite-State Markov Processes Subject to Rare Transitions," to be submitted for publication.

14. M.Coderch, A.S. Willsky, S.S. Sastry and D.A. Castanon, "Hierarchical Aggregation of Linear Systems with Multiple Time Scales", Proc. 1981 CDC Conference, San Diego, California.

15. M. Coderch, "Multiple Time Scale Approach to Aggregation with Filtering and Control Applications", Phd Thesis, MIT-EECS, to be submitted May 1982.

16. D.A. Castanon, B.C. Levy, S.B. Gershwin, "Diffusion Approximations of Transfer Lines with Unreliable Links and Finite Storage Elements", LIDS R-1041, November 1980.

17. S.K. Mitter, "Nonlinear Filtering and Stochastic Mechanics", in M. Hazewinkel and J.C. Willems (eds.) Stochastic Systems: The Mathematics of Filtering and Identification and Applications, pp. 479-503. D. Reidel Publishing Company, 1981.

18. Y. Avniel and S.K. Mitter, "Scattering Theory for Deterministic and Stochastic Linear Systems," to appear.

19. W.H. Fleming, "Stochastic Calculus of Variation and Mechanics," Rept. LIDS-P-1070, M.I.T. Laboratory for Information and Decision Systems; also submitted to J. of Optimization Theory, 1981.

20. W.H. Fleming and S.K. Mitter, "Optimal Control and Nonlinear Filtering for Nondegenerate Diffusion Processes, "submitted to Stochastics.

21. S.K. Mitter, "Conservative and Dissipative Stochastic Systems," to appear.

22. S.S. Sastry and O. Hijab, "Bifurcation in the Presence of Small Noise," Systems and Control Letters, to appear.

23. O. Hijab and S.S. Sastry, "Singular Perturbation, State Aggregation and Nonlinear Filtering, "Proc. 20th IEEE Conf. on Decision and Control, Dec. 1981.

24. M.B. Propp and T.L. Johnson, "A Biochemically and Structurally-Based Model of Striated Muscle Contraction," Biophysical Journal, Vol. 33, 1981, page 25a.

25. M.B. Propp, A Biochemically - and Structurally-Based Model of Striated Muscle Contraction, S.M. Dissertation, Department of Electrical Engineering and Computer Science, M.I.T., Cambridge, MA., May 1981.

26. M.B. Propp, The Thermodynamic Properties of Markov Processes, Ph.D. dissertation, Department of Electrical Engineering and Computer Science, M.I.T., Cambridge, MA, in preparation.

27. M.B. Propp, "The Thermodynamic Properties of Markov Processes," in preparation.

28. S.I. Marcus, "Low Dimensional Filters for a Class of Finite State Estimation Problems with Poisson Observations," submitted to Systems and Control Letters.

29. J.L. Willems and J.C. Willems, "Robust Stabilization of Uncertain Systems," submitted to SIAM Journal on Control and Optimization.

30. H.J. Chizeck, Fault-Tolerant Optimal Control, Ph.D. dissertation, Department of Electrical Engineering and Computer Science, M.I.T., Cambridge, MA., May 1982.

31. M. Coderch, D.A. Castanon, and A.S. Willsky, "A Note on Stochastically Discontinous Processes," to be submitted for publication.

32. D.A. Castanon, M. Coderch, and A.S. Willsky, "Aggregation of Diffusion Processes with Multiple Equilibrium Points," to be submitted for publication