This is to certify that the

dissertation entitled

STABILITY AND CONTROL OF NONLINEAR
SINGULARLY PERTURBED STOCHASTIC SYSTEMS

presented by

Mohamed Gamal El-Ansary

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Systems Science

Hassan Khalil
Major professor

Date 1/6/1983

STABILITY AND CONTROL OF NONLINEAR SINGULARLY PERTURBED STOCHASTIC SYSTEMS

By

Mohamed Gamal El-Ansary

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering and
Systems Science

1983

ABSTRACT

STABILITY AND CONTROL OF NONLINEAR SINGULARLY PERTURBED STOCHASTIC SYSTEMS

By

Mohamed Gamal El-Ansary

A class of nonlinear singularly perturbed systems driven by wide-band noise is considered. The probabilistic behavior of the slow variables is studied when the fast variables are sufficiently fast (represented by $\mu \to 0$) and the wide-band noise is sufficiently wide (represented by $\varepsilon \to 0$). The possible interaction between the asymptotic phenomena associated with singular perturbations and the asymptotic phenomena associated with fast stochastic fluctuations, is also considered. The slow state which is, in general, not a Markov process, is shown to converge to a diffusion Markov process in the sense of weak convergence as ε and μ tend to zero and the ratio $\frac{\varepsilon}{\mu}$ tends to a nominal value $\gamma \in [\gamma_1, \infty)$, where $\gamma_1 > 0$ is arbitrary but fixed. This limiting process is the solution of a reduced-order diffusion model which is derived explicitly and the interaction between the two asymptotic phenomena described above, has turned out to be important, as it is revealed from the dependence of the reduced order model, in general, on γ which equals to $\lim_{\varepsilon, \mu \to 0} \frac{\varepsilon}{\mu}$.

The advantages of having a reduced-order Markov model in hand, to approximate the slow states, are displayed by utilizing some of the available work on stability and stabilization of Markov process. Stability properties of the non-Markov slow states are studied through those of the reduced-order Markov states. Design of stabilizing feedback control strategies for the original system is based on well-established stabilization techniques of the reduced-order Markov model.

This Dissertation is Dedicated

to

Hala El-Ansary, my wife for her love
and understanding throughout the long years
of being a student
It is only through her patience
that this work was completed.

to

Tarek, Noha and Sherief, my children
for their love and support.

ACKNOWLEDGMENTS

It is a great pleasure to express my sincere appreciation to my major advisor, Dr. Hassan K. Khalil who introduced me to the fields of singular perturbation and stochastic systems. His patience, valuable guidance, expert advice, stimulating discussions, and useful insights made this work possible.

I would also like to thank the committee members, Dr. Robert Schlueter, Dr. Robert O. Barr, Dr. David H.Y. Yen and Dr. Habib Salehi. These professors offered inspiration as teachers and continued encouragement throughout my graduate study at Michigan State University.

My gratitude extends to Professor John Kreer, the Chairman of the Electrical Engineering and Systems Science Department for his encouragement and support.

Lastly, I would like to thank my typist Tammy Hatfield for her excellent work in typing this manuscript.

TABLE OF CONTENTS

CHAPTER I

LITERATURE SURVEY, BACKGROUND AND INTRODUCTION

1.1. Singular Perturbation Techniques and their Application to Control Systems.

It is a common practice of control engineers to simplify mathematical models which represent physical systems under investigation. The singular perturbation approach outlined in this section provides tools for simplifications in control systems analysis and design. Accordingly a typical simplification is to neglect some small time constants, masses, moments of inertia, some parasitic capacitances and inductances, and a number of unimportant parameters. The presence of such parameters increases the dynamic order of the model and introduces fast modes which make the model stiff, that is, difficult to handle on a digital computer. Consider a dynamic system which is modeled by the following initial value problem:

$$\dot{x}(t) = f(x(t),y(t),u(t)) \qquad x(t_0) = x_0 \qquad (1.1)$$

$$\mu\dot{y}(t) = g(x(t),y(t),u(t)) \qquad y(t_0) = y_0 \qquad (1.2)$$

where μ is a small positive parameter representing parasitic elements, x and y are n- and m-dimensional vectors, respectively, and u is an r-dimensional deterministic input vector. For $\mu = 0$, the order $n+m$ of (1.1) and (1.2) reduces to n, that is (1.2) becomes

$$0 = g(\bar{x}(t),\bar{y}(t),\bar{u}(t)) \tag{1.3}$$

Suppose that (1.3) has an isolated root along which $\frac{\partial g}{\partial y}$ is nonsingular,

$$\bar{y}(t) = h(\bar{x}(t),\bar{u}(t)) \tag{1.4}$$

Substituting (1.4) into (1.1) we obtain the reduced system

$$\dot{\bar{x}} = f(\bar{x}(t),\bar{u}(t)) \qquad \bar{x}(t_0) = x_0 \tag{1.5}$$

Reducing the order $(m+n)$ of (1.1) and (1.2) to n of (1.5) is not the only advantage of (1.5). Another advantage can be realized when we notice that in (1.2) we actually have $\dot{y} = g/\mu$, that is, if μ is very small and $g \neq 0$, then y is increasing very rapidly. This explains, in a sense, what we mean by the stiffness of (1.1) and (1.2) which is eliminated from (1.5).

To see the effect of this simplification procedure on the variable y, which has been excluded from the simplified model (1.5), we notice that $\bar{y}$ which is given by (1.4) starts at t_0 from $\bar{y}(t_0) = \bar{h}(x(t_0), \bar{u}(t_0))$, in contrast to the original variable y which starts at t_0 from a prescribed value y_0, where there may be a large discrepancy between y_0 and $\bar{y}(t_0)$. Thus the best that one can hope for is that $\bar{y}(t)$ is a good approximation to $y(t)$ everywhere except near $t = t_0$ and that $\bar{x}(t)$ is a good approximation to $x(t)$ everywhere. To study the behavior of y near $t = t_0$, the time scale is stretched by introducing the transformations

$$\tau = \frac{t - t_0}{\mu} \tag{1.6}$$

In terms of τ, (1.1) and (1.2) becomes

$$\frac{dx}{d\tau} = \mu f(x,y,u) \qquad x(0) = x_0 \tag{1.7}$$

$$\frac{dy}{d\tau} = g(x,y,u) \qquad y(0) = y_0 \tag{1.8}$$

Setting $\mu = 0$ in (1.7) and (1.8) we get that $x(\tau) \equiv x_0$. Then (1.8) can be written in a more convenient form in terms of $\eta = y - \bar{y}$ as

$$\frac{d\eta}{d\tau} = g(x_0, \bar{y}(t_0) + \eta(\tau), u(t_0)) \tag{1.9}$$

The system (1.9) is called the boundary-layer system and the variable $\eta(\tau)$ is referred to as the boundary-layer correction which is significant only during a short interval $[t_0, t_1]$. A basic result of singular perturbation theory is an initial value theorem due originally to Tihonov (See [1] for references) which spells out conditions under which the solution of the initial value problem for (1.1) and (1.2) as $\mu \to 0$ can be approximated by the solutions of the reduced and boundary-layer systems in the sense that for all $t \in (t_0, t_f]$

$$x(t) \longrightarrow \bar{x}(t), \tag{1.10}$$

$$y(t) \longrightarrow \bar{y}(t) \tag{1.11}$$

We notice that, actually $y(t)$ is approximated by $\bar{y}(t) + \eta(\tau)$ for all $t \in [t_0, t_f]$ but $\eta(\tau) \to 0$ as $\mu \to 0$

(i.e. $\tau \to \infty$). The essential conditions are stability type conditions which are imposed on the boundary-layer system (1.9).

The two time scale phenomena accompanied the solution of the initial value problem is at the heart of the singular perturbation approach to stability and control problems. In a typical control problem one starts by defining separate reduced and boundary-layer problems. Assuming the existence of solutions for these problems, an approximate solution is postulated by combining the separate solutions. The validity of the approximations as $\mu \to 0$ is established via asymptotic analysis (cf. [1-3]).

In general if the singularly perturbed system, which is represented by (1.1) and (1.2) is asymptotically stable, the fast states represented by the vector y are important only during a short initial period. After that period they are negligible and the behavior of the system can be described by its slow states represented by x. In many applications the fast states y are basically parasetics, that is, for example the equation (1.2) can represent the model of an actuator in a control system which can be neglected. Neglecting the fast modes is equivalent to assuming that they are infinitely fast, that is letting $\mu \to 0$ in (1.2).

1.2. Asymptotic Analysis of Systems Driven by Wide-Band Noise:

In this section we study and review some of the work that has been done concerning dynamic systems with external influences which are approximately white noise (wide-band noise). In this thesis, our main concern will be the asymptotic analysis of a class of systems having the above property. Let us first introduce the basic topics and definitions that will be used and then we will review the work done which is related to our work.

Itô's Stochastic Differential Equation:

It is of the form

$$dx = f(t,x)dt + G(t,x)dw(t) \qquad t_0 \leq t \leq T \qquad (1.12)$$

x is a vector (the system state) in R^n, the vector-valued function $f(t,x)$ is usually called the drift coefficient, $G(t,x)$ is an $n \times m$ matrix-valued function and w is a Wiener process, (Brownian motion), usually taken to be Gaussian, in Euclidean m-space. Equation (1.12) was originally studied in [4,5] and later, under less restrictive conditions, in many text books [cf. 6-8]. Equation (1.12) is interpreted as a stochastic integral equation

$$x(t) = x(t_0) + \int_{t_0}^{t} f[s,x(s)]ds + \int_{t_0}^{t} G[s,x(s)]dw(s) \qquad (1.13)$$

It is assumed that f and G are measurable in (t,x)

for $t \in [t_0,T]$, $x \in R^n$; and satisfy (i) a growth condition

$$|f(t,x)| + |G(t,x)| \leq K(1+|x|), \quad t \in [t_0,T], \quad x \in R^n \tag{1.14}$$

and (ii) a uniform Lipschitz condition

$$|f(t,x)-f(t,y)| + |G(t,x)-G(t,y)| \leq K|x-y|, \quad t \in [t_0,T], \quad x,y \in R^n \tag{1.15}$$

In (1.13) $x(t_0)$ is any (finite-valued) random vector independent of the increments dw. Under these conditions (1.13) determines a unique stochastic Markov process x which is also called a diffusion process. For $\varphi \in C^2(R^n)$, the differential operator associated with the process x is defined by:

$$\mathcal{L}\varphi(x) = f'(t,x)\varphi_x(x) + \frac{1}{2}\mathrm{tr}[G'(t,x)\varphi_{xx}(x)G(t,x)] \tag{1.16}$$

Weak Convergence:

The concept of weak convergence can be defined roughly as follows: Suppose that P_n is a sequence of probability measures defined on a metric space S and P is also a probability measure defined on S, then it is said that P_n converges weakly to P, denoted $P_n \Rightarrow P$ if for each continuous function f on S, $\int_S f\,dP_n \to \int_S f\,dP$. Now if $S = C$ = space of continuous functions and if P_n and P are probability measures on S then $P_n \Rightarrow P$ if the finite dimensional distributions corresponding to P_n

converges weakly to those corresponding to P and the sequence of measures $\{P_n\}$ is relatively compact. For more details about this subject see [9], for example. Weak convergence has been used, successfully, as the appropriate type of convergence in asymptotic analysis. In particular, it has been used to prove the convergence of a sequence of non-Markovian or Markovian processes to a Markov process, see for example [10-15].

Stochastic Stability:

As stochastic models have come to be more fully understandable to engineers and scientists, the study of rather important stachastic system properties has become possible. Among these is the property of stability. The literature on the topic is full of many concepts of stability that have been studied, see for example ([16], for a survey). These stability concepts have, in general, been derived for the study of deterministic systems. It follows that there are at least as many stability concepts for the study of stochastic systems as there are for the study of deterministic systems. The reason is that the deterministic concepts of stability have their counterparts in each of the common modes of convergence of probability theory. We may recall that the common modes of convergence, [cf. 17] are convergence in probability, convergence in the mean and almost sure convergence. Thus, it is clear that one has at least three times as many concepts of stability as for the usual deterministic case. Indeed there are

even more. Among those concepts we will state only two definitions, [cf. 16] of the concepts that we are going to adopt in this thesis. We shall refer to the equilibrium or null solution, $x \equiv 0$, as the solution whose stability properties are being tested; x_0 will denote the initial state at the initial time t_0. We will denote the solution with initial state x_0 at time t_0, by $x(t;x_0,t_0)$ or simply $x(t)$, which is assumed to be an n-vector.

Definition of Asymptotic Stability in Probability: The equilibrium solution is said to be asymptotically stable in probability or equivalently, it is uniformly stochastically asymptotically stable if for any $\eta_1 > 0$ and $\eta_2 > 0$ there exists a $\delta > 0$ such that if $|x_0| < \delta$ then:

(i) $P\{|x(t)| \leq \eta_2 e^{-\theta t},\ t \geq 0\} \geq 1 - \eta_1$ for some $\theta > 0$.

(ii) $P\{\lim_{t\to\infty} |x(t)| = 0\} = 1.$

Definition of Asymptotic Stability in the Mean Square: The equilibrium solution is said to be asymptotically stable in mean square if there exists constants $\alpha > 0$, $K_1 \geq 0$ and $K_2 > 0$ such that $E|x(t)|^2 \leq K_1 + K_2 e^{-\alpha t}$ $\forall t \geq 0$ and then the process $x(t)$ is said to be exponentially bounded in mean square with exponent α. This form of the definition is stated in [18] and we are going to use it, as it is, later.

Review of Related Work:

The mathematical theory of stochastic differential equations is concerned almost exclusively with the study of Itô equations and the associated Markov processes. This theory has found many useful applications and has become a powerful tool in the study of diffusion processes (cf. [7],[8]). However, many of its aspects are somewhat drastic idealizations of physical processes in the sense that the noise affecting the physical system is approximated by white noise which is not a physical process but an abstraction. This was the motivation for later work which led to modeling dynamic systems with external noise which are approximately white noise, by systems of ordinary differential equations with wide-band noise as input so that Makov process techniques can be used. Several powerful methods for doing this have been developed. The problem has been initiated by [19] and then developed more (cf. [11-15]). In [19], the Langevin scalar equations:

$$dx_n(t) = m(x_n(t))dt + \sigma(x_n(t))dy_n(t) \tag{1.17}$$

has been considered as a mathematical representation of a physical model, where $y_n(t) \to y(t)$ in the mean square sense as $n \to \infty$, and $y(t)$ is a scalar Brownian motion process. It has been shown that the solutions $x_n(t)$ of (1.17) converge to the diffusion process $x(t)$ as $n \to \infty$, in the mean square sense, where $x(t)$ satisfies the Itô equation:

$$dx = [m(x) + \frac{1}{2}\sigma(x)\frac{d\sigma}{dx}(x)]dt + \sigma(x)dy \tag{1.18}$$

This says that the Langevin equation cannot replaced by an Itô differential equation without realizing the necessity for the correction term, $\frac{1}{2}\sigma(x)\frac{d\sigma}{dx}(x)$ in (1.18). More work has been developed along that line. All the authors in [11-15] have treated the problem of weak convergence of $x^{\epsilon}(\cdot)$ to a diffusion where $x^{\epsilon}(\cdot)$ is defined as the solutions of ordinary differential equations with wide-band random hand sides. More specifically the system that they have all considered is of the form:

$$\begin{aligned} \frac{dx^{\epsilon}(t)}{dt} &= \frac{1}{\epsilon}F(x^{\epsilon}(t),y^{\epsilon}(t)) + G(x^{\epsilon}(t),y^{\epsilon}(t)) \\ x^{\epsilon}(0) &= x_0 \end{aligned} \tag{1.19}$$

$x \in R^n$ and $y \in R^m$.

For each $\epsilon > 0$, $y^{\epsilon}(t) = y(t/\epsilon^2)$ where $y(t)$ has been taken to be, in general, a stationary process (other hypotheses has been introduced in those references). The process $y^{\epsilon}(t)$ is, in a sense (will be made precise in the next chapter), a wide-band noise, and the system (1.20) is a wide-band noise system. The parameter $\epsilon > 0$ measures departure from the white noise. Another interpretation for ϵ is that it differentiates between the time scale of fluctuations of the coefficients and the solution. In [11] and [12] $y(t)$ is considered to be a Markov process, ergodic, bounded and satisfies other assumptions so that under certain smoothness conditions on F and G,

$(x^{\varepsilon}(t), y^{\varepsilon}(t))$ are, an $(n+m)$, jointly Markov process. It has been proved, using partial differential equations and perturbation techniques, that $x^{\varepsilon}(t)$ converges weakly to a diffusion process $x(t)$, as $\varepsilon \to 0$ on $[0,T]$ where $T < \infty$, but arbitrary.

At this point, it seems to be interesting to make analogy between this asymptotic analysis that has been carried out in [11] and [12] and those of the deterministic singular perturbation. We notice that the solution $x^{\varepsilon}(t)$ of (1.19) is not exactly a Markov process, but it can be considered as components of a higher-dimensional Markov process, as was the case when $(x^{\varepsilon}(t), y^{\varepsilon}(t))$ was treated as a jointly Markov process. Then approximating $x^{\varepsilon}(t)$ by the Markov process $x(t)$ explains, in a sense, that an order-reduction procedure has been taken place which is in analogy to the order-reduction that occurs in deterministic singular perturbation.

In [13-15] the same system, which is roughly represented by (1.19) has been studied but with different assumptions on the process $y(t)$. Semigroup techniques due to [10] and Martingale approach have been employed in [13] and [14,15] respectively, to prove that $x^{\varepsilon}(t)$ converges weakly to the diffusion $x(t)$ whose differential operator [cf. 13] takes the form:

$$\begin{aligned} Af(x) &= EG'(x,y(s))f_x(x) \\ &+ \int_0^{\infty} d\tau EF'(x,y(s))(F'(x,y(s+\tau))f_x(x))_x \end{aligned} \tag{1.20}$$

where f is continuous with continuous partial derivatives up to the second order.

All the work that has been done in [11-15] is, of course, closely related to the original problem of [19].

The first attempt, to study stability properties of dynamical systems which are driven by wide band noise, has been made in [12]. Stability results about $x^{\epsilon}(t)$, defined by (1.19), has been established which are based only on conditions upon the approximating diffusion $x(t)$. These are conditions which guarantee that the equilibrium of $x(t)$ is stable in an appropriate stochastic sense. Most of the work that has been done, to study stability properties of stochastic systems, is concerned with the stability of systems represented as an Itô equation. The effective method that has been employed is the stochastic Liapunov method which is analogous to the deterministic Liapunov method. [cf. 7, 16 and 20].

1.3. <u>Stochastic Singularly Perturbed Systems</u>:

Since our work is mainly concerned with stochastic singular perturbations, we will briefly review the prior work that has been done in the linear case while, in the nonlinear case, a considerable detailed review will be established. Singularly perturbed linear differential equations with random forcing functions have been studied as models of control and filtering systems [21-23]. Promising results have been obtained like the two-time

scale linear filter obtained in [21]. However, some difficulties, arising from the idealized behavior of the white noise used in the models, have been encountered especially in the linear quadratic control problem studied in [22] and [23] where the performance index may diverge. Some alternatives have been suggested to overcome these difficulties. Colored noise has been allowed in [24], which in a sense limits the significance of the fast subsystem; a near optimal linear output feedback control is obtained by optimizing a slow subsystem only. In [25], a parameter scaling procedure has been proposed to overcome the difficulties that arise from the unclear behaviour of the fast variables in stochastic singularly perturbed control systems. As a result, the divergence of the performance index has been avoided and a well-posed linear quadratic control problem has been obtained. In a recent study of stochastic linear singularly perturbed systems [cf. 26] a new approach to approximating linear quadratic - Gaussian estimation and control problems has been established.

One of the few attempts which has been made to study nonlinear singularly perturbed systems driven by white noise, is [27], in which a stochastic control problem has been investigated for the system:

$$\begin{aligned} dx &= [a(x) + c(x)z + 2\beta(x)v(t)]dt + \sqrt{2}\,dw_1 \\ x(0) &= x \end{aligned} \tag{1.21}$$

$$\mu dz = [b(x) + d(x)z + 2\alpha(x)v(t)]dt + \sqrt{2}\,\mu dw_2 \qquad (1.22)$$

$$z(0) = z$$

where all variables are scalars, $v(\cdot)$ is a control variable, and w_1 and w_2 are two independent, scalar Wiener processes. The unclear behaviour of the fast variable due to the existence of the white noise, which has been the source of trouble as we pointed out before [cf. 22], has been avoided by multiplying the white noise, in equation (1.22), by μ. However, by modeling the input stochastic process as white noise some model information might have been already lost as a result of the inconsistency encountered in modeling physical systems driven by wide-band noise as systems driven by white noise (see section 1.2).

From the above discussion and from the asymptotic analysis of systems driven by wide-band noise, that has been reviewed in section 1.2, it seems appropriate that in studying singularly perturbed systems the input noise should be modeled as wide-band noise rather than white noise.

In that regard, a study of a nonlinear singularly perturbed system driven by wide-band noise, has been initiated by [28]. The following system has been considered:

$$\dot{x}(t) = J_1^{\varepsilon}(t,x(t),y(t)) + \frac{1}{\sqrt{\varepsilon}} F_1^{\varepsilon}(t,x(t),y(t)) \qquad (1.23)$$

$$\mu\dot{y}(t) = J^{\varepsilon}(t,x(t),y(t)) + \frac{1}{\sqrt{\varepsilon}} F^{\varepsilon}(t,x(t),y(t)) \qquad (1.24)$$

where F_1^ϵ and F^ϵ are fluctuations in the time scale t/ϵ while the natural time scale of the state x is t. Here we write the singular perturbation parameter as μ although in [28] it is written as $\sqrt{\epsilon}$. The asymptotic behavior of the state has been studied as the perturbation parameters tend to zero. Without getting involved in the technical details and assumptions the essential steps of that approach are as follows. First, it is assumed that the equation

$$\sqrt{\epsilon}\, J^\epsilon(t,x,y) + F^\epsilon(t,x,y) = 0 \tag{1.25}$$

has a unique root $y = r^\epsilon(t,x)$ which is used to define an outer solution of x as

$$\begin{aligned} \dot{X}(t) = J_1^\epsilon(t,X(t),r^\epsilon(t,X(t))) + \frac{1}{\sqrt{\epsilon}} F_1^\epsilon(t,X(t),\\ r^\epsilon(t,X(t))) . \end{aligned} \tag{1.26}$$

Second, the asymptotic behavior of the outer solution $X(t)$ as $\epsilon \to 0$ is studied using limit theorems of stochastic processes and conditions are spelled out under which $X(t)$ converges weakly to a diffusion process $\overline{X}(t)$. Third, the diffusion $\overline{X}(t)$ is taken as a candidate for the limit of $x(t)$. To show this, conditions from [29] are imposed to guarantee that $x(t) - X(t) \to 0$ as the singular perturbation parameter (μ in our notation) tends to zero. Implicit in the approach of [28] there is a sequential ordering of the two asymptotic phenomena present in the problem. Since an outer solution is defined first using singular perturbation ideas and then stochastic asymptotic analysis is applied,

it is reasonable to say that this approach assumes that the asymptotic phenomena associated with singular perturbations are faster than the asymptotic phenomena associated with stochastic fluctuations. As it will be seen, our results show that the approach of [28] is valid when $\frac{\mu}{\epsilon} \to 0$ as $\epsilon \to 0$. The approach therefore does not take into consideration the possible interaction between the two asymptotic phenomena when ϵ and μ are of the same order. Such interaction has been brought to attention after a paper by [30]. In that paper the following second-order differential equation, has been considered:

$$\mu\ddot{x}(t) + \dot{x}(t) = a(x(t)) + b(x(t))v^{\epsilon}(t) \tag{1.27}$$

where $v^{\epsilon}(t)$ is exponentially correlated noise with correlation time ϵ. It has been suggested that for sufficiently small ϵ and μ, $x(t)$, the solution of (1.27) can be approximated by a diffusion process, defined by an Itô equation. Moreover, this diffusion process cannot be obtained as the asymptotic limit which results either by letting $\epsilon \to 0$ first then $\mu \to 0$ or by letting $\mu \to 0$ then $\epsilon \to 0$, since two different limits are expected. In deriving the reduced-order model corresponding to (1.27), an intuitive reasoning has been employed. It has been assumed that over a time interval Δt which is very small with respect to the relaxation time of $x(t)$ while very large with respect to μ and ϵ, the process $x(t)$ will behave like a continuous Markov process. With that

assumption the following well known definitions, from the theory of Markov process [cf. 6,7],

$$A(x) = \lim_{\Delta t \to 0} E\{\frac{x(t+\Delta t) - x(t)}{\Delta t} / x(t) = x\}, \tag{1.28}$$

$$B(x) = \lim_{\Delta t \to 0} E\{\frac{[x(t+\Delta t) - x(t)]^2}{\Delta t} / x(t) = x\}, \tag{1.29}$$

has been used to calculate the drift coefficient $A(x)$ of the approximating diffusion and its diffusion coefficient $B(x)$. In the calculations of the conditional moments given by (1.28) and (1.29), there has been no demand for finding an exact solution of (1.27), it has been enough to solve (1.27) on a small interval Δt satisfying

$$\tau_{rel} >> \Delta t >> \max(\mu, \varepsilon) \tag{1.30}$$

where, the relaxation time τ_{rel} of $x(t)$ is defined by:

$$\tau_{rel} \sim \min_x (-\frac{da}{dx}(x))^{-1}$$

Equation (1.27) has been integrated over the interval $[t, t+\Delta t]$ and after applying the basic assumption (1.30), the result of integration has been simplified to:

$$\begin{aligned} x(t+\Delta t) &= x(t) + a(x(t))\Delta t \\ &+ \frac{1}{\mu}\int_t^{t+\Delta t}\int_t^{\lambda} \bar{e}^{(t+\Delta t-\tau)/\mu}\, b(x(\tau))v^{\varepsilon}(\tau)d\tau \end{aligned} \tag{1.31}$$

The integral on the right-hand side of (1.31) considers the correlation of $b(x(\tau))$ with $v^{\varepsilon}(\tau)$. Since this

integral is not a stochastic integral, it has been considered as a Riemann integral. Then successive approximation has been used [cf. 32], with initial solution $x_0 = x(t)$, and Taylor series expansions around x_0, have been employed. With the aid of (1.30), only the terms of order Δt has been retained and a second order approximation has been obtained. It has been claimed that higher order approximations have the same accuracy $O(\Delta t)$ that the second order one has. Finally it has been shown that the results of calculations of (1.28) and (1.29) are:

$$A(x) = a(x) + \frac{.5}{1+\mu/\epsilon} \frac{db}{dx}(x)b(x)S(0), \tag{1.32}$$

$$B(x) = b^2(x)S(0), \tag{1.33}$$

where $S(w/\epsilon)$ is the spectrum of v^ϵ, so that the suggested reduced order-model corresponding to (1.27) has been represented by the following Itô equation:

$$d\bar{x}(t) = A(\bar{x}(t))dt + \sqrt{B(\bar{x}(t))}\ dw(t) \tag{1.34}$$

There has been no rigorous proof, in that paper, to validate that the process $x(t)$, defined by (1.27), converges to the diffusion process $x(t)$, defined by (1.34), as $\epsilon, \mu \to 0$ in any stochastic sense.

The remarkable feature about the suggested reduced-order model (1.35) is its dependence on the ratio $\frac{\epsilon}{\mu}$, as it is apparent from (1.32), hinting to the interaction between the two asymptotic phenomena.

1.4. Objectives of the Thesis:

Our main objective, which has been motivated by [30], is to generalize the reduced-order model, that has been suggested by [30], to a wider class of singularly perturbed systems and to provide a rigorous proof of convergence of the slow states to the diffusion process defined by the reduced-order model and then, to explore the possible application of the reduced-order model in stability and control problems. In this thesis we consider the nonlinear singularly perturbed system:

$$\dot{x}(t) = a_1(x(t)) + A_{12}(x(t))y(t) + B_1(x(t))v^{\varepsilon}(t) \qquad (1.35)$$

$$\mu\dot{y}(t) = a_{21}(x(t)) + A_2 y(t) + B_2(x(t))v^{\varepsilon}(t) \qquad (1.36)$$

where $v^{\varepsilon}(t)$ is a wide-band zero-mean stationary vector process with correlation matrix

$$E\{v^{\varepsilon}(t)v^{\varepsilon}(t+\tau)'\} = \frac{1}{\varepsilon}R\left(\frac{\tau}{\varepsilon}\right)$$

More assumptions will be imposed on the process v^{ε} in the next chapter. This class of singularly perturbed systems is similar to the deterministic one studied in [2] from the viewpoint of allowing nonlinearity in the slow variable x while assuming linear dependence on the fast variable y. We allow the input noise to be state dependent by letting the input matrices B_1 and B_2 be functions in x; we do not, however, allow them to be function in y since that will destroy the linearity in y which is very desirable feature as it is apparent from [2].

The accomplishments reported in this thesis are summarized as follows:

(a) The asymptotic behavior of the slow variables, defined by (1.35) and (1.36), has been studied when the fast variables are sufficiently fast (represented by $\mu \to 0$) and the wide-band noise is sufficiently wide (represented by $\varepsilon \to 0$). A reduced-order model to represent the behavior of the slow variables has been derived. It has been shown that the slow variables converge weakly to the solution of this reduced-order model as $\varepsilon \to 0$ and $\mu \to 0$. However, our proof cover the two cases:

(i) $\frac{\mu}{\varepsilon} \to 0$ as $\varepsilon \to 0$,

(ii) ε and μ of the same order, i.e., there exists positive constants K_1 and K_2 such that $0 < K_1 \leq \frac{\mu}{\varepsilon} \leq K_2 < \infty$.

The third case, namely:

(iii) $\frac{\varepsilon}{\mu} \to 0$ as $\mu \to 0$

Follows essentially as a special case of [33] after applying results of [12] or [13]. This case is briefly discussed in chapter 2. The proof adapts a martingale method developed by [14] for proving weak convergence of a sequence of non-markovian processes to a diffusion process.

(b) The use of the reduced-order model in stability of the full-order system, given by (1.35) and (1.36), has been examined. A result has been obtained which provide stochastic asymptotic stability of the origin of the full system if the origin of the reduced-order model is so, provided that the parameters ε and μ are sufficiently small. The main advantage of using the reduced-order model is that it is a Markov model, and the theory of stochastic stability [cf 16,20] applied to stochastic differential equations of Itô type is rich.

(c) Applying the reduced-order model in control problems has been considered. A stablizing output feedback control has been designed, using a nonlinear observer, for the reduced-order model. We have been motivated by the work of [18], in which a stabilizing feedback control for a system represented by an Itô equation has been designed using an observer. The designed control law has been implemented to the full-order system, with an observer, and conditions, under which the closed loop system is stable, have been spelled out via the use of our stability result.

CHAPTER II

REDUCED-ORDER MODEL AND CONVERGENCE RESULT

2.1. Introduction.

This chapter is concerned with sutyding the asymptotic behavior of the slow variables of a singularly perturbed system driven by wide-band noise, when the fast dynamics are too fast, represented by $\mu \to 0$, and the wide-band noise is too wide, represented by $\epsilon \to 0$. A reduced-order diffusion model that approximates the behavior of the slow variables is derived together with a rigorous proof of convergence. Our proof covers the two cases $\mu/\epsilon \to 0$ as $\epsilon \to 0$ and ϵ and μ being of the same order of magnitude, i.e, $K_1 \leq \frac{\mu}{\epsilon} \leq K_2$ for some positive constants K_1 and K_2. It is also shown that the case $\frac{\epsilon}{\mu} \to 0$ as $\mu \to 0$, which is not covered by our proof, can be deduced from results already available in the literature.

This chapter is arranged in the following way. In the second section we introduce the singularly perturbed model and list all the assumptions that are needed for the convergence proof. In section 3, the basic theorem is stated and proved. To make the proof more readable some lengthy details which are not very essential to follow the

logic of the proof, have been given in separate appendices at the end of the chapter.

2.2. Problem Formulation and Assumptions:

Consider the singularly perturbed system

$$\dot{x}(t) = a_1(x(t)) + A_{12}(x(t))y(t) + B_1(x(t))v^{\epsilon}(t),\quad x(0) = x_0 \tag{2.1}$$

$$\mu\dot{y}(t) = a_{21}(x(t)) + A_2 y(t) + B_2(x(t))v^{\epsilon}(t),\quad y(0) = y_0 \tag{2.2}$$

where $x \in R^n$, $y \in R^m$ and x_0, y_0 are bounded random vectors. The stochastic process $v^{\epsilon} \in R^r$ is defined as

$$v^{\epsilon}(t) = \frac{1}{\sqrt{\epsilon}}\, v(t/\epsilon) \tag{2.3}$$

where $v(t)$ satisfies

(A1) $v(t)$ is a stationary, zero mean, right continuous, uniformly bounded process on $[0,\infty)$. The σ-algebras induced by $v(t)$ are assumed to have a mixing property with an exponential mixing rate [9], i.e.,

$$\sup_{A_i, t} |P(A_2/A_1) - P(A_2)| \leq e^{-\alpha\tau}$$

for some $\alpha > 0$, where $A_1 \in \sigma\{v(s), s \leq t\}$ and $A_2 \in \sigma\{v(s), s \geq t+\tau\}$. The process $v^{\epsilon}(t)$ is said to

be wide-band noise since its power spectral density matrix $S^{\epsilon}(\omega) = S(\omega/\epsilon)$ will have a frequency band of ω_0/ϵ when $S(\omega)$, the spectral matrix of v, has a frequency band ω_0. Indeed, the process $v^{\epsilon}(t)$ converges to Gaussian white noise by the central limit theorem [11].

The coefficients of (2.1) and (2.2) are assumed to satisfy

(A2) The coefficients a_1, a_{21}, A_{12}, B_1 and B_2 are continuous in x and have continuous partial derivatives up to the second order which are bounded uniformly in x. Moreover, a_{21} and B_2 are bounded uniformly in x.

(A3) The constant matrix A_2 is Hurwitz, i.e. $\mathrm{Re}\lambda(A_2) < 0$.

(A4) The vector $a_1(x)$ and the matrices $A_{12}(x)$ and $B_1(x)$ are required to satisfy

$$|a_1(x)| + |A_{12}(x)| + |B_1(x)| \leq K(1 + |x|) \qquad \forall x \in R^n \tag{2.4}$$

and the vector $a_0(x)$ and the matrix $B_0(x)$ which are defined by:

$$a_0 = a_1 - A_{12}A_2^{-1}a_{21} \tag{2.5}$$

and

$$B_0 = B_1 - A_{12}A_2^{-1}B_2 \tag{2.6}$$

are required to satisfy

$$|a_0(x) - a_0(z)| + |B_0(x) - B_0(z)| \leq K|x - z| \qquad \forall x,z \in R^n$$

for some positive constants K.

We notice that from (A2) and (A4) growth conditions similar to (2.4) will be satisfied for a_0 and B_0.

(A3) is needed to guarantee the asymptotic stability of the boundary layer phenomena associated with y. Under the assumptions (A2)-(A4), the usual existence and uniqueness theory for ordinary differential equations gives us a solution for (2.1) and (2.2) on $[0,T]$ for sufficiently small μ and for each sample path of $v(\cdot)$. This follows by minor modification of the technique used in proving the basic result of [34].

Our objective is to study the asymptotic behavior of $x(\cdot)$ as $\epsilon \to 0$ and $\mu \to 0$. The main result of this chapter shows that $x(\cdot)$ converges weakly to a diffusion process $\bar{x}(\cdot)$. The infinitismal generator associated with $\bar{x}(\cdot)$, whose form will follow from the proof of the result, is given by

$$\begin{aligned} L^{\gamma} f(x) = & \sum_{i=1}^{n} b_i(x) \frac{\partial f}{\partial x_i}(x) \\ & + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \end{aligned} \tag{2.7}$$

where

$$b(x) = a_0(x) + h_1(x) - A_{12}(x)A_2^{-1}h_2(x) + h_3(x), \tag{2.8}$$

$$A(x) = B_0(x)S(0)B_0'(x) \overset{\Delta}{=} [a_{ij}(x)], \tag{2.9}$$

$S(\omega)$ is the spectral matrix of v,

$$h_{1i} = \mathrm{tr}[D_i'B_0W' + D_i'A_{12}A_2^{-1}\Sigma]^1, \tag{2.10}$$

$$h_{2i} = \mathrm{tr}[E_i'B_0W' + E_i'A_{12}A_2^{-1}\Sigma], \tag{2.11}$$

$$\begin{aligned} h_{3i} = \mathrm{tr}[&-F_i'B_0W'B_2'(A_2')^{-1} - F_i'B_0\Sigma'(A_2')^{-1} \\ &+ F_i'A_{12}A_2^{-1}P], \end{aligned} \tag{2.12}$$

$$\begin{aligned} D_i &= [\nabla_x\psi_{i1} \mid \nabla_x\psi_{i2} \mid ---- \mid \nabla_x\psi_{ir}]_{nxr}; \\ B_1 &= [\psi_{ij}]_{nxr}, \end{aligned} \tag{2.13}$$

$$\begin{aligned} E_i &= [\nabla_x\eta_{i1} \mid \nabla_x\eta_{i2} \mid ---- \mid \nabla_x\eta_{ir}]_{nxr}; \\ B_2 &= [\eta_{ij}]_{mxr}, \end{aligned} \tag{2.14}$$

$$\begin{aligned} F_i &= [\nabla_x\xi_{i1} \mid \nabla_x\xi_{i2} \mid ---- \mid \nabla_x\xi_{im}]_{nxm}; \\ A_{12} &= [\xi_{ij}]_{nxm}, \end{aligned} \tag{2.15}$$

---

[1] $(')$ denotes transposition.

$$W = \int_0^\infty R(\tau)\,d\tau$$

$$\Sigma = \int_0^\infty e^{A_2\gamma\tau} B_2 R'(\tau)\,d\tau, \quad \text{for some} \quad \gamma \in [\gamma_1,\infty),$$
$$\gamma_1 > 0 \tag{2.16}$$

and $$P = \int_0^\infty e^{A_2\lambda}(B_2\Sigma' + \Sigma B_2')e^{A_2'\lambda}\,d\lambda. \tag{2.17}$$

We require that the coefficients $A(x)$ and $b(x)$, defined above, satisfy the following conditions

(A5) $|A(x)| \leq C(1 + |x|^2), \qquad x \in R^n$

(A6) $\langle x, b(x)\rangle \leq C(1 + |x|^2), \quad x \in R^n$

where C is some positive constant.

These two conditions in addition to (A2) guarantee that the martingale problem corresponding to (2.7) is well-posed [8].

2.3. The Convergence Theorem:

Theorem: Under the assumptions (A1)-(A6), $x(\cdot)$ converges weakly to $\overline{x}(\cdot)$ as $\epsilon \to 0$, $\mu \to 0$ and $\frac{\epsilon}{\mu} \to \gamma$ where $\gamma \in [\gamma_1,\infty)$, $\gamma_1 > 0$ is arbitrary, but fixed.

Proof: We utilize a technique for proving weak convergence of a sequence of non-Markovian processes to a diffusion process which was introduced in [10] and further developed in [13-15]. The version used here is

due to [14]. The main step in the proof is finding a sequence of test functions $f^{\epsilon,\mu}(t)$ for a given function $f(x)$ such that certain conditions are satisfied. We use the so called perturbed test function which was used for similar purposes in [12-15]. Before we get to the technical details of the proof we need to introduce some definitions and terminology.

Truncated Processes: For every positive integer N, let $S_N = \{x \in R^n, \|x\| \leq N\}$ and define the truncated process $x^{\epsilon,\mu}_N(t)$ to be the solution of

$$\begin{aligned} \dot{x}^{\epsilon,\mu}_N = q_N(x^{\epsilon,\mu}_N)\,[a_1(x^{\epsilon,\mu}_N) + A_{12}(x^{\epsilon,\mu}_N)y^{\epsilon,\mu}_N \\ + B_1(x^{\epsilon,\mu}_N)v^{\epsilon}],\; x^{\epsilon,\mu}_N(0) = x_0 \end{aligned} \tag{2.18}$$

$$\begin{aligned} \mu\dot{y}^{\epsilon,\mu}_N = [a_{21}(x^{\epsilon,\mu}_N) + A_2 y^{\epsilon,\mu} + B_2(x^{\epsilon,\mu}_N)v^{\epsilon}], \\ y^{\epsilon,\mu}_N(0) = y_0 \end{aligned} \tag{2.19}$$

where $q_N(x) = 1$ for $x \in S_N$, $q_N(x) = 0$ for $x \in R^n - S_{N+1}$ and $q_N(x) \in [0,1]$ and has third derivatives that are bounded uniformly in x and N. For each N, $\{x^{\epsilon,\mu}_N(\cdot)\}$ is bounded uniformly in μ and ϵ. As it will be seen the actual technical proof involves only the truncated processes $\{x^{\epsilon,\mu}_N(\cdot)\}$. See [14,15] for similar treatment.

Terminology: Let $(\Omega, P, \mathcal{F})$ be the probability space in which $v(\cdot)$ is defined and let $\mathcal{F}^{\epsilon,\mu}_{t,N}$ be the σ-algebra

induced by $\{x^{\epsilon,\mu}_N(s),\ y^{\epsilon,\mu}_N(s),\ v^{\epsilon}(s),\ 0 \le s \le t\}$ and $E^{\epsilon,\mu}_{t,N}$ the corresponding conditional expectation. Let $\mathcal{L}^0$ be the class of measurable (ω,t) real valued functions such that if $f(\cdot) \in \mathcal{L}^0$ then $E|f(t+s)-f(t)| \to 0$ as $s \to 0^+$ and $\sup_t E|f(t)| < \infty$ and $f(t)$ is adapted to $\mathcal{F}^{\epsilon,\mu}_{t,N}$. We say $\text{p-lim}_{s \to 0} f^s = 0 \Leftrightarrow \sup_{s,t} E|f^s(t)| < \infty$ and $E|f^s(t)| \to 0$ as $s \to 0^+$. Define an operator $A^{\epsilon,\mu}_N$ and its domain $D(A^{\epsilon,\mu}_N)$ as follows: $f \in D(A^{\epsilon,\mu}_N)$ and $A^{\epsilon,\mu}_N f = g \Leftrightarrow f,g \in \mathcal{L}^0$ and

$$\text{p-lim}_{r \to 0^+} \left| \frac{E^{\epsilon,\mu}_{t,N} f(t+r) - f(t)}{r} - g(t) \right| = 0.$$

Let L^{γ}_N be a diffusion operator of the form (2.7) such that the coefficients of L^{γ}_N and L^{γ} are equal for $x \in S_N$. Let $\hat{\mathcal{C}}_0$ be the space of continuous functions $f : R^N \to R$ which have compact support and $\hat{\mathcal{C}}^3_0$ be the space of functions which belongs to $\hat{\mathcal{C}}_0$ together with its partial derivatives up to the third order.

The following Lemma is Theorems (1) and (2) of [14] adapted to our case.

Lemma 1: Assume that the martingale problem associated with (2.7) is well-posed. For each fixed N, let $\{x^{\epsilon,\mu}_N(\cdot)\}$ be the solution of (2.18) and (2.19). Suppose

that for each $f \in \hat{\mathcal{C}}_0^3$, there is a sequence $f_N^{\epsilon,\mu}(\cdot) \in D(A_N^{\epsilon,\mu})$ and a random variable $M_T^{\epsilon,\mu}(f)$, for each $T > 0$, such that

$$\underset{\substack{\epsilon,\mu\to 0\\ \epsilon/\mu\to\gamma}}{\text{p-lim}} \, [f_N^{\epsilon,\mu}(t) - f(x_N^{\epsilon,\mu}(t))] = 0, \tag{2.20}$$

$$\underset{\substack{\epsilon,\mu\to 0\\ \epsilon/\mu\to\gamma}}{\text{p-lim}} \, [A_N^{\epsilon,\mu} f_N^{\epsilon,\mu}(t) - L_N^{\gamma} f(x_N^{\epsilon,\mu}(t))] = 0, \tag{2.21}$$

$$P\{\sup_{t\le T} | f_N^{\epsilon,\mu}(t) - f(x_N^{\epsilon,\mu}(t))| \ge \eta\} \to 0 \quad \text{as} \quad \epsilon,\mu \to 0, \quad \epsilon/\mu \to \gamma \tag{2.22}$$

$$\sup_{t\le T} |A_N^{\epsilon,\mu} f_N^{\epsilon,\mu}(t)| \le M_T^{\epsilon,\mu}(f), \tag{2.23}$$

and

$$\sup_{\epsilon,\mu} P\{M_T^{\epsilon,\mu}(f) \ge K\} \to 0 \quad \text{as} \quad K \to \infty \tag{2.24}$$

then $\{x(\cdot)\}$ converges weakly to $\bar{x}(\cdot)$ as $\epsilon \to 0$ and $\mu \to 0$ and $\epsilon/\mu \to \gamma$.

For notational convenience we write $x(t)$, $y(t)$, $A^{\epsilon,\mu}$, L^{γ}, $f_i(t)$ and E_t instead of $x_N^{\epsilon,\mu}(t)$, $y_N^{\epsilon,\mu}(t)$, $A_N^{\epsilon,\mu}$, L_N^{γ}, $f_{1,N}^{\epsilon,\mu}(t)$ and $E_{t,N}^{\epsilon,\mu}$ respectively but we are always working with the truncated process $\{x_N^{\epsilon,\mu}(\cdot)\}$. Moreover, we omit the q_N terms for further simplification. Now we proceed with the proof of the theorem. Let $f \in \hat{\mathcal{C}}_0^3$ be given, then

$$\begin{aligned} A^{\epsilon,\mu} f(x(t)) = \frac{\partial f}{\partial x} (x(t))[&a_1(x(t)) + A_{12}(x(t))y(t) \\ &+ B_1(x(t))v^{\epsilon}(t)] \end{aligned} \tag{2.25}$$

we observe that $y(t)$, the solution of (2.2) is given by:

$$y(t) = e^{A_2 t/\mu} y_0 + \frac{1}{\mu}\int_0^t e^{A_2(t-\tau)/\mu} a_{21}(x(\tau))d\tau$$

$$+ \frac{1}{\mu\sqrt{\varepsilon}}\int_0^t e^{A_2(t-\tau)/\mu} B_2(x(\tau))v(\tau/\varepsilon)d\tau$$

Now, since a_{21} and B_2 are bounded uniformly in x and $v(t/\varepsilon)$ is uniformly bounded on $[0,\infty)$, we have:

$$|y(t)| \le K_1 e^{-\alpha_2 t/\mu}|y_0| + \frac{K_2}{\mu}\int_0^t e^{-\alpha_2(t-\tau)/\mu} d\tau$$

$$+ \frac{K_3}{\mu\sqrt{\varepsilon}}\int_0^t e^{-\alpha_2(t-\tau)/\mu} d\tau$$

$$\le \bar{K}_1 + \frac{\bar{K}_2}{\sqrt{\varepsilon}}$$

Then, by the compact support of $\frac{\partial f}{\partial x}$, the last two terms on the right hand side of (2.25) are of order $1/\sqrt{\varepsilon}$ and cannot be part of the operator L^{γ}, so they are averaged out by defining $f_1(x,t)$ as:

$$f_1(x,t) = \int_0^{\infty} \frac{\partial f}{\partial x}(x)E_t[A_{12}(x)(\hat{y}(t+s,x) + A_2^{-1}a_{21}(x)) + B_1(x)v^{\varepsilon}(t+s)] \tag{2.26}$$

where

$$\hat{y}(t+s,x) = e^{A_2 s/\mu} y(t) + (e^{A_2 s/\mu} - I)A_2^{-1}a_{21}(x) + \frac{1}{\mu}\int_t^{t+s} e^{A_2(t+s-\tau)/\mu} B_2(x)v^{\varepsilon}(\tau)d\tau \tag{2.27}$$

Subtracting the term $-A_2^{-1}a_{21}(x)$ in (2.26), in a sense, centers $\hat{y}$ at its steady state mean.

Setting $x = x(t)$ in (2.26) and defining $f_1(t) = f_1(x(t),t)$ we claim that

$$|f_1(t)| \leq K_1\sqrt{\varepsilon} + K_2\sqrt{\mu} \tag{2.28}$$

where K_1 and K_2 are positive constants independent of T and ω.

Proof of the claim:

From (2.26) and (2.27) we have:

$$\begin{aligned} f_1(x(t),t) = \int_0^\infty \frac{\partial f}{\partial x}(x(t))E_t[A_{12}(x(t))e^{A_2 s/\mu}(y(t) \\ + A_2^{-1}a_{21}(x(t))) \\ + \frac{1}{\mu}A_{12}(x(t))\int_0^s e^{A_2(s-\lambda)/\mu}B_2(x(t))v^\varepsilon(t+\lambda)d\lambda \\ + B_1(x(t))v^\varepsilon(t+s)]ds \end{aligned}$$

then by using the bound on $|y(t)|$ we have:

$$\begin{aligned} |f_1(x(t),t)| \leq K\int_0^\infty |\tfrac{\partial f}{\partial x}(x(t))|[|A_{12}(x(t))|e^{-\alpha_2 s/\mu}(\overline{K}_1 + \frac{\overline{K}_2}{\sqrt{\varepsilon}}) \\ + \frac{|A_{12}(x(t))|}{\mu}\int_0^s e^{-\alpha_2(s-\lambda)/\mu}|B_2(x(t))||E_t v^\varepsilon(t+\lambda|d\lambda]ds \\ = I_1 + I_2 \end{aligned}$$

Then, from (A2), (A4), the boundedness of the truncated process $x(t)$ and the compact support of $\frac{\partial f}{\partial x}$ we have after integrating with respect to s:

$$|I_1| \le \tilde{K}_1\mu + \tilde{K}_2\frac{\mu}{\sqrt{\varepsilon}} \le \tilde{K}_1\mu + \tilde{K}_2\sqrt{\frac{\mu}{\varepsilon}}\cdot\sqrt{\mu} \le K_2\sqrt{\mu}$$

where we used that $\frac{\mu}{\varepsilon}$ is bounded.

The mixing property implies

$$|E_t v^{\varepsilon}(t+\lambda)| \le \frac{K}{\sqrt{\varepsilon}}\, e^{-\alpha\lambda/\varepsilon}$$

Using that in I_2 we have:

$$|I_2| \le \frac{K}{\mu\sqrt{\varepsilon}} \int_0^{\infty}\int_0^{s} e^{-\alpha_2(s-\lambda)/\mu}\, e^{-\alpha\lambda/\varepsilon}\, d\lambda ds.$$

Changing the order of integration, we get:

$$|I_2| \le \frac{K}{\mu\sqrt{\varepsilon}} \int_0^{\infty}\int_{\lambda}^{\infty} e^{-\alpha_2(s-\lambda)/\mu}\, e^{-\alpha\lambda/\varepsilon}\, ds d\lambda$$

$$= \frac{K}{\sqrt{\varepsilon}\,\alpha_2}\int_0^{\infty} e^{-\alpha\lambda/\varepsilon}\, d\lambda = K_1\sqrt{\varepsilon}$$

$\therefore$ $|f_1(x(t),t)| \le K_1\sqrt{\varepsilon} + K_2\sqrt{\mu}$, which proves (2.28).

We next show that $f_1(t) \in D(A^{\varepsilon,\mu})$. We have [13]

$$A^{\epsilon,\mu}f_1(t) = \underset{\delta \to 0}{\text{p-lim}} \, [E_t f_1(x(t+\delta), t+\delta) - f_1(x(t),t)]/\delta$$

$$= \underset{\delta \to 0}{\text{p-lim}} \, [E_t f_1(x(t+\delta), t+\delta) - f_1(x(t), t+\delta)]/\delta$$

$$+ \underset{\delta \to 0}{\text{p-lim}} \, [E_t f_1(x(t), t+\delta) - f_1(x(t),t)]/\delta \tag{2.29}$$

if the limits exist and are in $\mathcal{L}_0$. We first show that the second limit exists and is in $\mathcal{L}_0$. From (2.26), $f_1(x(t),t)$ can be written in the form:

$$f_1(x(t),t) = \int_0^\infty E_t g_1(x(t), t+s)\,ds$$

where $g_1(x(t), t+s)$ is equal to the integrand in the right-hand side of (2.26). So the second term of (2.29) is

$$I_2 = \underset{\delta \to 0}{\text{p-lim}} \, [E_t f_1(x(t), t+\delta) - f_1(x(t),t)]/\delta$$

$$= \underset{\delta \to 0}{\text{p-lim}} \, [E_t \int_0^\infty E_{t+\delta} g_1(x(t), t+\delta+s)\,ds$$

$$- \int_0^\infty E_t g_1(x(t), t+s)\,ds]/\delta$$

Setting $u = \delta + s$, we get

$$I_2 = \underset{\delta\to 0}{\text{p-lim}} \left[\int_\delta^\infty E_t g_1(x(t),t+u)du - \int_0^\infty E_t g_1(x(t),t+s)ds\right]/\delta$$

$$= -\underset{\delta\to 0}{\text{p-lim}} \frac{1}{\delta}\int_0^\delta E_t g_1(x(t),t+s)ds = -g_1(x(t),t+s)$$

$$= -\frac{\partial f}{\partial x}(x(t))[A_{12}(x(t))y(t) + A_{12}(x(t))A_2^{-1}a_{21}(x(t))$$

$$+ B_1(x(t))v^\epsilon(t)].$$

Therefore, the second limit exists. By the compact support of $\frac{\partial f}{\partial x}$ and the right continuity of $v^\epsilon(t)$ it is obvious that $E|g_1(x(t),t+s) - g_1(x(t),t)| \to 0$ as $s \to 0^+$ and that $\sup_t E|g_1(x(t),t)| < \infty$ and this implies that the second limit in (2.29) belongs to $\mathcal{L}_0$. For the first limit we have:

$$\underset{\delta\to 0}{\text{p-lim}} [E_t f_1(x(t+\delta),t+\delta) - f_1(x(t),t+\delta)]/\delta$$

$$= \underset{\delta\to 0}{\text{p-lim}} \frac{1}{\delta}\int_0^\delta E_t\left[\frac{\partial f_1}{\partial x}(x(t+u),t+\delta)(a_1(x(t+u))\right.$$

$$\left. + A_{12}(x(t+u))y(t+u) + B_1(x(t+u))v^\epsilon(t+u))\right]du$$

$$= \frac{\partial f_1}{\partial x}(x(t),t)(a_1(x(t)) + A_{12}(x(t))y(t) + B_1(x(t))v^\epsilon(t))$$

which shows that the limit exists. See [13] for a similar treatment. Now by an argument similar to the one that has been used to show that the second limit is in $\mathcal{L}_0$, we can show that the first limit also is in $\mathcal{L}_0$. We conclude

that $f_1(t) \in D(A^{\epsilon,\mu})$. Then, from (2.29) and the above limits we have, (with $x(t) = x$)

$$A^{\epsilon,\mu} f_1(t) = -\frac{\partial f}{\partial x}(x)[A_{12}(x)y(t) + A_{12}(x)A_2^{-1}a_{21}(x) + B_1(x)v^{\epsilon}(t)] + \frac{\partial f_1}{\partial x}(x,t)[a_1(x) + A_{12}(x)y(t) + B_1(x)v^{\epsilon}(t)] \tag{2.30}$$

Adding (2.25) to (2.30) we get

$$A^{\epsilon,\mu}(f(x) + f_1(t)) = \frac{\partial f}{\partial x}(x)a_0(x) + \frac{\partial f_1}{\partial x}(x,t)[a_1(x) + A_{12}(x)y(t) + B_1(x)v^{\epsilon}(t)] \tag{2.31}$$

The last two terms of (2.31) cannot be part of the operator L^{γ}, so we average them out by defining f_2, for every $x \in R^n$ and $t \in [0,T]$ as:

$$f_2(x,t) = \int_0^{\infty} [E_t \frac{\partial f_1}{\partial x}(x,t+s)(A_{12}(x)\hat{y}(t+s,x) + A_{12}(x)A_2^{-1}a_{21}(x) + B_1(x)v^{\epsilon}(t+s)) + \frac{\partial f}{\partial x}(x)a_0(x) - L^{(\epsilon/\mu)}f(x)]ds \tag{2.32}$$

The form of $L^{(\epsilon/\mu)}$, as defined by (2.5-2.17) with ϵ/μ replacing γ, results as a by-product of showing that

$|f_2(x(t),t)|$ is $O(\mu+\varepsilon)$, i.e., by identifying the parts of the first three terms on the right-hand side of (2.32) which are not $O(\varepsilon)$ or $O(\mu)$.

Using that f_x and f_{xx} have compact support and the mixing property (2.4), and (A2)-(A4), it is shown in Appendix B that:

$$|f_2(t)| \leq K_3\varepsilon + K_4\mu \tag{2.33}$$

where, $f_2(t) = f_2(x(t),t)$ and K_3 and K_4 are positive constants independent of T and ω. Following the same steps, we used to show that $f_1 \in D(A^{\varepsilon,\mu})$ and $A^{\varepsilon,\mu}f_1 \in \mathcal{L}_0$, it can be shown that $f_2(t) \in D(A^{\varepsilon,\mu})$ and $A^{\varepsilon,\mu}f_2 \in \mathcal{L}_0$, where, (with $x(t) = x$)

$$\begin{aligned} A^{\varepsilon,\mu}f_2(t) &= L^{(\varepsilon/\mu)}f(x) - \frac{\partial f_1}{\partial x}(x,t)(A_{12}(x)y(t) \\ &\quad + A_{12}(x)A_2^{-1}a_{21}(x) + B_1(x)v^{\varepsilon}(t)) - \frac{\partial f}{\partial x}(x)a_0(x) \\ &\quad + \frac{\partial f_2}{\partial x}(x,t)(a_1(x) + A_{12}(x)y(t) + B_1(x)v^{\varepsilon}(t)) \ . \end{aligned} \tag{2.34}$$

Adding (2.31) to (2.34) we get:

$$\begin{aligned} A^{\varepsilon,\mu}(f(x) + f_1(t) + f_2(t)) &= L^{\varepsilon/\mu}f(x) + \frac{\partial f_1}{\partial x}(x,t)a_0(x) \\ &\quad + \frac{\partial f_2}{\partial x}(x,t)[a_1(x) + A_{12}(x)y(t) + B_1(x)v^{\varepsilon}(t)]. \end{aligned} \tag{2.35}$$

We define

$$f^{\epsilon,\mu}(t) = f(x(t)) + f_1(x(t),t) + f_2(x(t),t) \quad \text{for}$$

$$0 \leq t \leq T \tag{2.36}$$

Now we are ready to verify condition (2.20) of lemma (1). From (2.36), (2.28) and (2.33) we have:

$$E|f^{\epsilon,\mu}(t) - f(x(t))| = E|f_1(t) + f_2(t)| \leq K_1\sqrt{\epsilon} + K_2\sqrt{\mu}$$

$$+ K_3\epsilon + K_4\mu \to 0 \quad \text{as} \quad \epsilon,\mu \to 0$$

and it is obvious that, for $t \in [0,T]$ and ϵ,μ small, $\sup\limits_{t,\epsilon,\mu} E|f^{\epsilon,\mu}(t) - f(x(t))| < \infty$.

Then, by the definition of the p-limit, (2.20) of lemma (1) is proved. It is shown in Appendix A that:

$$\left|\frac{\partial f_1}{\partial x}(x(t),t)a_0(x(t))\right| \leq K_5\sqrt{\epsilon} + K_6\sqrt{\mu} \tag{2.37}$$

where K_5 and K_6 are positive constants independent of T and ω. We notice that differentiating $f_1(x,t)$ with respect to x, did not affect the order, i.e. we have the same bounds on f_1 and $\frac{\partial f_1}{\partial x}$. Motivated by this argument and by following the terms that will appear in the calculation of the bound on f_2 which is done in

Appendix B, we can see that differentiating with respect to x will not change the order of the resulting terms and we will get a similar upper bound as in (2.33), taking into account the mixing property of v, the compact support of the partial derivatives of f up to the third order and that $v^{\varepsilon}(t)$ and $y(t)$ are of order $\frac{1}{\sqrt{\varepsilon}}$ we have:

$$\left|\frac{\partial f_2}{\partial x}(x(t),t)a_1(x(t))\right| \le K_7\varepsilon + K_8\mu \tag{2.38}$$

and

$$\begin{aligned}&\left|\frac{\partial f_2}{\partial x}(x(t),t)(A_{12}(x(t))y(t) + B_1(x(t))v^{\varepsilon}(t)\right| \\ &\quad \le K_9\sqrt{\varepsilon} + K_{10}\sqrt{\mu}\end{aligned} \tag{2.39}$$

where the positive constants are independent of ω and T.

By the smooth dependence of L^{γ} on γ (see (2.16)), and by the compact support of f and all of its partial derivatives, there exists a constant $c > 0$ such that:

$$|L^{(\varepsilon/\mu)}f(x) - L^{\gamma}f(x)| \le c\left|\frac{\varepsilon}{\mu} - \gamma\right| \tag{2.40}$$

Now, we verify condition (2.21) of lemma (1) as follows:

$$\begin{aligned}|A^{\varepsilon,\mu}f^{\varepsilon,\mu}(t) - L^{\gamma}f(x(t))| &\le |A^{\varepsilon,\mu}f^{\varepsilon,\mu}(t) - L^{(\varepsilon/\mu)}f(x(t))| \\ &+ |L^{(\varepsilon/\mu)}f(x(t)) - L^{\gamma}f(x(t))|\end{aligned} \tag{2.41}$$

But, from (2.35), (2.36) and by applying (2.37)-(2.39) we get:

$$|A^{\epsilon,\mu}f^{\epsilon,\mu}(t) - L^{(\epsilon/\mu)}f(x(t))| \leq |\frac{\partial f_1}{\partial x}(x(t),t)a_0(x(t))|$$

$$+ |\frac{\partial f_2}{\partial x}(x(t),t)a_1(x(t))|$$

$$+ |\frac{\partial f_2}{\partial x}(x(t),t)(A_{12}(x(t))y(t) + B_1(x(t))v^{\epsilon}(t))|$$

$$\leq \overline{K}_1\sqrt{\epsilon} + \overline{K}_2\sqrt{\mu} + \overline{K}_3\epsilon + \overline{K}_4\mu \qquad (2.42)$$

where all $\overline{K}_i > 0$ and independent of T and w.

Then, from (2.40) and (2.42) and by taking expectation, we get:

$$E|A^{\epsilon,\mu}f^{\epsilon,\mu}(t) - L^{\gamma}f(x(t))| \leq \overline{K}_1\sqrt{\epsilon} + \overline{K}_2\sqrt{\mu} + \overline{K}_3\epsilon + \overline{K}_4\mu$$

$$+ c|\frac{\epsilon}{\mu} - \gamma| \to 0 \quad \text{as} \quad \epsilon,\mu \to 0 \quad \text{and} \quad \frac{\epsilon}{\mu} \to \gamma,$$

and since this expected value is finite for all t, condition (2.21) of lemma (1) is verified. From (2.36), (2.28) and (2.33) it is obvious that:

$$\{\sup_{t\leq T} |f^{\epsilon,\mu}(t) - f(x(t))| \geq \eta\} \subset \{(K_1\sqrt{\epsilon} + K_2\sqrt{\mu} + K_3\epsilon + K_4\mu) \geq \eta\}$$

Therefore, it follows immediately that

$$P\{\sup_{t\leq T} |f^{\epsilon,\mu}(t) - f(x(t))| \geq \eta\} \to 0 \quad \text{as} \quad \epsilon,\mu \to 0 \quad \text{and} \quad \frac{\epsilon}{\mu} \to \gamma$$

and this proves (2.22) of lemma (1). From (2.42), definition of $L^{\varepsilon/\mu}$ and the compact support of f and its partial derivatives, condition (2.23) and (2.24) of lemma (1) follow directly.

Then applying lemma (1) the proof of the theorem is completed.

Remark 2.1: The above theorem does not cover the case $\frac{\varepsilon}{\mu} \to 0$ as $\mu \to 0$. This case however can be treated in the following manner.

Asymptotic analysis can be applied in two steps by letting $\varepsilon \to 0$ followed by $\mu \to 0$. For each $\mu > 0$ it follows from [12] or [13] that if the coefficients of (2.1) and (2.2) and the process $v(t)$ satisfy the appropriate assumptions then the corresponding solutions of those equations converge weakly as $\varepsilon \to 0$ to the solutions of the singularly pertubed Ito model: (See Appendix C for derivations).

$$dx = [\bar{a}_1(x) + A_{12}(x)y]dt + B_1(x)\sqrt{S(0)}\,dw \tag{2.43}$$

$$\mu dy = [\bar{a}_{21}(x) + A_2 y]dt + B_2(x)\sqrt{S(0)}\,dw \tag{2.44}$$

where

$$\bar{a}_1(x) = a_1(x) + \bar{h}_1(x), \tag{2.45}$$

$$\bar{a}_{21}(x) = a_{21}(x) + \bar{h}_2(x), \tag{2.46}$$

$$\bar{h}_{1i} = tr[D_i' B_1 W'], \tag{2.47}$$

and

$$\bar{h}_{2i} = \mathrm{tr}[E_i' B_1 W'], \tag{2.48}$$

and D_i, E_i and W are as defined in Section 2.2. The system (2.43) and (2.44) is a special case of the system studied in [33] (See (2.1.1) of [33]). Under the assumptions of [33] the process $X(\cdot)$ converges weakly to the diffusion process $X^0(\cdot)$ with differential operator L given by (See Appendix C for details)

$$Lf(x) = \sum_{i=1}^{n} \bar{b}_i(x) \frac{\partial f}{\partial x_i}(x) + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \tag{2.49}$$

where

$$\bar{b}(x) = \bar{a}_1(x) - A_{12}(x) A_2^{-1} \bar{a}_{21}(x) + \bar{h}_3(x), \tag{2.50}$$

$$A(x) = [a_{ij}(x)] = B_0(x) S(0) B_0'(x), \tag{2.51}$$

$$\bar{h}_3(x) = \mathrm{tr}[-F_i' B_1 S(0) B_2' (A_2')^{-1} - F_i' A_{12} \bar{P} (A_2')^{-1}] \tag{2.52}$$

and $\bar{P}$ satisfies

$$\bar{P} A_2' + A_2 \bar{P} = -B_2 S(0) B_2' \tag{2.53}$$

It is interesting to notice that the reduced order model corresponding to the operator L in (2.49) can be obtained from L^Y in our reduced-order model (2.7) by letting

$\gamma \to 0$ (or $\frac{\varepsilon}{\mu} \to 0$). So, generally speaking, we can say that the operator L^{γ} of (2.7) gives the right form of the reduced-order model for all values of γ, i.e. $\gamma \in [0,\infty)$.

Remark 2.2: There are special cases where the infinitismal generator associated with $\bar{x}$, given by (2.7), is independent of the parameter γ. Such special cases and their significance will be discussed in Chapter 5. Here we would like to point out that in such cases the convergence theorem reduces to the statement: "under the assumptions (A1)-(A6), $x(\cdot)$ converges weakly to $\bar{x}(\cdot)$ as $\varepsilon \to 0$ and $\mu \to 0$, provided that $\frac{\varepsilon}{\mu} \geq \gamma_1 > 0$.

APPENDIX A

To verify inequality (2.37), let us consider $f_1(x,t)$ as defined by (2.26) and (2.27). So, we have:

$$
\begin{aligned}
f_1(x,t) &= \int_0^\infty \frac{\partial f}{\partial x}(x)\, A_{12}(x) e^{A_2 s/\mu} ds\,(y(t) + A_2^{-1} a_{21}(x)) \\
&+ \frac{\partial f}{\partial x}(x) A_{12}(x) \frac{1}{\mu} \int_0^\infty \int_t^{t+s} e^{A_2(t+s-\lambda)/\mu} B_2(x) E_t v^\epsilon(\lambda)\, d\lambda\, ds \\
&+ \frac{\partial f}{\partial x}(x) B_1(x) \int_0^\infty E_t v^\epsilon(t+s)\, ds \\
&= -\mu \frac{\partial f}{\partial x}(x) A_{12}(x) A_2^{-1}(y(t) + A_2^{-1} a_{21}(x)) \\
&+ \frac{\partial f}{\partial x}(x) A_{12}(x) \frac{1}{\mu} \int_t^\infty \int_{\lambda-t}^\infty e^{A_2(t+s-\lambda)/\mu} ds\, B_2(x) E_t v^\epsilon(\lambda)\, d\lambda \\
&+ \frac{\partial f}{\partial x}(x) B_1(x) \int_0^\infty E_t v^\epsilon(t+s)\, ds
\end{aligned}
$$

The second term, after integrating with respect to s will be

$$
-\frac{\partial f}{\partial x}(x) A_{12}(x) A_2^{-1} B_2(x) \int_t^\infty E_t v^\epsilon(\lambda)\, d\lambda \ .
$$

Thus we have:

$$f_1(x,t) = -\mu \frac{\partial f}{\partial x}(x)A_{12}(x)A_2^{-1}(y(t) + A_2^{-1}a_{21}(x))$$

$$+ \frac{\partial f}{\partial x}(x)B_0(x) \int_t^\infty E_t v^\epsilon(\lambda)d\lambda \qquad \text{(A-1)}$$

where $B_0(x)$ is defined by (2.6).

Let

$$a_{21}(x) = [\zeta_i(x)]_{m\times 1},$$

$$A_2^{-1} = [\alpha_{ij}]_{m\times m}$$

and

$$B_0(x) = [\theta_{ij}(x)]_{n\times m}$$

then $f_1(x,t)$ in (A-1) can be expressed as:

$$f_1(x,t) = \sum_{j=1}^{n} \sum_{k,\ell=1}^{m} [-\mu \frac{\partial f}{\partial x_j}(x)\xi_{j\ell}(x)\alpha_{\ell k}(y_k(t)$$

$$+ \sum_{\nu=1}^{m} \alpha_{k\nu}\zeta_\nu(x))] + \sum_{j=1}^{n} \sum_{k=1}^{r} (\frac{\partial f}{\partial x_j}(x)\theta_{jk}(x) \int_t^\infty E_t v_k^\epsilon(\lambda)d\lambda)$$

$$\text{(A-2)}$$

Then, differentiating $f_1(x,t)$ with respect to x, we get, for the ith component of the gradient $\frac{\partial f_1}{\partial x}(x,t)$, $i = 1,2,\cdots,n$

$$\frac{\partial f_1}{\partial x_i}(x,t) = \sum_{j=1}^{n} \sum_{k,\ell=1}^{m} [-\mu \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \xi_{j\ell}(x) \alpha_{\ell k}(y_k(t)$$

$$+ \sum_{\nu=1}^{m} \alpha_{k\nu} \zeta_\nu(x)) - \mu \frac{\partial f}{\partial x_j}(x) \frac{\partial \xi_{j\ell}(x)}{\partial x_i} \alpha_{\ell k}(y_k(t)$$

$$+ \sum_{\nu=1}^{m} \alpha_{k\nu} \zeta_\nu(x)) - \mu \frac{\partial f}{\partial x_j}(x) \xi_{j\ell}(x) \alpha_{\ell k} \sum_{\nu=1}^{m} \alpha_{k\nu} \frac{\partial \zeta_\nu}{\partial x_i}(x)]$$

$$+ \sum_{j=1}^{n} \sum_{k=1}^{r} [\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \theta_{jk}(x) \int_t^\infty E_t v_k^\epsilon(\lambda) d\lambda$$

$$+ \frac{\partial f}{\partial x_j}(x) \frac{\partial \theta_{jk}}{\partial x_i}(x) \int_t^\infty E_t v_k^\epsilon(\lambda) d\lambda] \qquad (A-3)$$

where $[\xi_{ij}(x)]_{n \times m} = A_{12}(x)$ as defined by (2.15).

For simplicity we are going to use the same symbol K to denote different constants. Now setting $x = x(t)$ in (A-3), and by the use of the following facts:

1. $|y_k(t)| \le \|y(t)\| \le K_1 + \frac{K_2}{\sqrt{\varepsilon}} \qquad \forall k = 1,2,\cdots,m.$

2. $|\xi_{j\ell}(x)\xi \le \|A_{12}(x)\| \le K(1 + |x|) \le K$

   $\forall j = 1,2,\cdots,n$ and $\ell = 1,2,\cdots,m$

   where $|x|$ is bounded as a result of the truncation. Same is true for θ_{jk}.

3. $a_{21}(x)$ is uniformly bounded in x.

4. First and second partial derivatives of $f(x)$ have compact support.

5. From the mixing property (2.4), we have:

$$|E_t v_k^\varepsilon(\lambda)| \leq \|E_t v_k^\varepsilon(\lambda)\| \leq \frac{K}{\sqrt{\varepsilon}}\, e^{-\alpha(\lambda - t)/\varepsilon}$$

we conclude that

$$|f_1(x(t),t)| \leq K_1\sqrt{\varepsilon} + K_2\sqrt{\mu}\ ,$$ and since

$$|a_0(x)| \leq K(1 + |x|)$$

and $|x|$ is bounded, (2.37) follows.

APPENDIX B

We need to verify the inequality given by (2.33). The important fact behind doing that is showing how the form of $L^{\epsilon/\mu}$, as defined by (2.5-2.17), results as a by product of showing that (2.33) is valid. Our goal is to show that

$$\int_0^\infty [E_t \frac{\partial f_1}{\partial x} (x,t+s)(A_{12}(x)\hat{y}(t+s,x) + A_{12}(x)A^{-1}_2 a_{21}(x) + B_1(x)v^\epsilon(t+s)]ds$$

$$= \int_0^\infty [L^{\epsilon/\mu} f(x) - \frac{\partial f}{\partial x}(x) a_0(x)]ds + O(\epsilon) + O(\mu) \qquad \text{(B-1)}$$

so that, if we define $f_2(x,t)$ by (2.32), (2.33) will be satisfied. Let I = the left hand-side of (B-1), and let

$$g(x,t+s) = A_{12}(x)\hat{y}(t+s,x) + A_{12}(x)A^{-1}_2 a_{21}(x) + B_1(x)v^\epsilon(t+s) \qquad \text{(B-2)}$$

$$I = \int_0^\infty E_t[\frac{\partial f_1}{\partial x}(x,t+s)\cdot g(x,t+s)]ds$$
$$= \int_0^\infty E_t[g'(x,t+s)\cdot\nabla_x f_1(x,t+s)]ds \qquad \text{(B-3)}$$

where $\nabla_x' = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \cdots, \frac{\partial}{\partial x_n})$. From (A-3), the inner product $g' \cdot \nabla_x f_1$ can be expressed as:

$$
\begin{aligned}
& g'(x,t+s) \cdot \nabla_x f_1(x,t+s) \\
& \quad = \sum_{i,j=1}^{n} \sum_{k,\ell=1}^{m} [-\mu g_i(x,t+s) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \xi_{j\ell}(x) \alpha_{\ell k} (\hat{y}_k(t+s,x) \\
& \quad + \sum_{\nu=1}^{m} \alpha_{k\nu} \zeta_\nu(x)) - \mu g_i(x,t+s) \frac{\partial f}{\partial x_j}(x) \frac{\partial \xi_{j\ell}}{\partial x_i}(x) \alpha_{\ell k} (\hat{y}_k(t+s,x) \\
& \quad + \sum_{\nu=1}^{m} \alpha_{k\nu} \zeta_\nu(x)) - \mu g_i(x,t+s) \frac{\partial f}{\partial x_j}(x) \xi_{j\ell}(x) \alpha_{\ell k} \\
& \qquad \cdot \sum_{\nu=1}^{m} \alpha_{k\nu} \frac{\partial \zeta_\nu}{\partial x_i}(x)] \\
& \quad + \sum_{i,j=1}^{n} \sum_{k=1}^{r} [g_i(x,t+s) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \theta_{jk}(x) \int_{t+s}^{\infty} E_{t+s} v_k^{\epsilon}(\lambda) d\lambda \\
& \quad + g_i(x,t+s) \frac{\partial f}{\partial x_j}(x) \frac{\partial \theta_{jk}}{\partial x_i}(x) \int_{t+s}^{\infty} E_{t+s} v_k^{\epsilon}(\lambda) d\lambda]
\end{aligned}
\tag{B-4}
$$

Let us define

$$\bar{y}(t+s,x) = \hat{y}(t+s,x) + A_2^{-1} a_{21}(x) \tag{B-5}$$

Then, it can be seen easily that the first term of (B-4) takes the form:

$$- \mu g'(x,t+s) f_{xx}(x) A_{12}(x) A_2^{-1} \bar{y}(t+s,x) \tag{B-6}$$

where $f_{xx}(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)_{n \times n}$. The second term of

$$(B\text{-}4) = -\mu \sum_{j=1}^{n} \sum_{\ell=1}^{m} \frac{\partial f}{\partial x_j}(x) g'(x,t+s) \nabla_x \xi_{j\ell}(x) (A_2^{-1} \bar{y}(t+s,x))_\ell$$

where $(A_2^{-1}\bar{y}(t+s,x))_\ell$ is the ℓ^{th} component of the m-vector $A_2^{-1}\bar{y}(t+s,x)$. Then, summing over ℓ and using (2.15) we get: The second term of (B-4)

$$= \sum_{j=1}^{n} -\mu \frac{\partial f}{\partial x_j} g'(x,t+s) F_j(x) A_2^{-1} \bar{y}(t+s,x) \tag{B-7}$$

where $F_j(x)$ is defined by (2.15). The third term of (B-4)

$$= -\mu g'(x,t+s)(a_{21}(x))'_x (A_2')^{-2} A_{12}'(x) \nabla_x f(x) \tag{B-8}$$

The fourth term of (B-4) can be reduced to

$$g'(x,t+s) f_{xx}(x) B_0(x) \int_{t+s}^{\infty} E_{t+s} v^{\epsilon}(\lambda) d\lambda \tag{B-9}$$

For the last term of (B-4), we need to look at $\frac{\partial \theta_{jk}}{\partial x_i}(x)$. Since $B_0(x) = B_1(x) - A_{12}(x) A_2^{-1} B_2(x) = (\theta_{jk}(x))_{n \times m}$ it is seen from (2.13)-(2.15) that

$$\theta_{jk}(x) = \psi_{jk}(x) - \sum_{\ell=1}^{m} \sum_{q=1}^{m} \xi_{j\ell}(x) \alpha_{\ell q} \eta_{qk}(x),$$ and therefore

$$\frac{\partial\theta_{jk}}{\partial x_i}(x) = \frac{\partial\psi_{jk}}{\partial x_i}(x) - \sum_{\ell,q=1}^{m} \frac{\partial\xi_{j\ell}}{\partial x_i} \alpha_{\ell q}\eta_{qk}(x)$$

$$- \sum_{\ell,q=1}^{m} \zeta_{j\ell}(x)\alpha_{\ell} \frac{\partial\eta_{qk}}{\partial x_i}(x)$$

Denoting the last term of (B-4) by I_4 we have:

$$I_4 = \sum_{i,j=1}^{n} \sum_{k=1}^{r} [g_i(x,t+s) \frac{\partial f}{\partial x_j}(x) \frac{\partial\psi_{jk}}{\partial x_i}(x) \int_{t+s}^{\infty} v_k^{\epsilon}(\lambda)d\lambda$$

$$- \sum_{\rho,q=1}^{m} g_i(x,t+s) \frac{\partial f}{\partial x_j}(x) \frac{\partial\xi_{j\rho}}{\partial x_i}(x) \alpha_{\rho_q}\eta_{qk}(x)$$

$$\int_{t+s}^{\infty} E_{t+s}v_k^{\epsilon}(\lambda)d\lambda$$

$$- \sum_{\rho,q=1}^{m} g_i(x,t+s) \frac{\partial f}{\partial x_j}(x) \alpha_{\rho_q} \frac{\partial\eta_{qk}}{\partial x_i}(x) \int_{t+s}^{\infty} E_{t+s}v_k^{\epsilon}(\lambda)d\lambda]$$

$$= I_{41} + I_{42} + I_{43} \tag{B-10}$$

$$I_{41} = \sum_{j=1}^{n} \sum_{k=1}^{r} \frac{\partial f}{\partial x_j}(x) g'(x,t+s)\nabla_x\psi_{jk}(x) \int_{t+s}^{\infty} E_{t+s}v_k^{\epsilon}(\lambda)d\lambda$$

but from (2.13) and by summing over k we get

$$I_{41} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) g'(x,t+s)D_j(x) \int_{t+s}^{\infty} E_{t+s}v^{\epsilon}(\lambda)d\lambda \tag{B-11}$$

where $D_j(x)$ is defined by (2.13)

$$I_{42} = - \sum_{i,j=1}^{n} \sum_{\rho=1}^{m} \frac{\partial f}{\partial x_j}(x)g_i(x,t+s)\frac{\partial\xi_{j\rho}}{\partial x_i}(x) (A_2^{-1}B_2(x)$$

$$\int_{t+s}^{\infty} E_{t+s}v^{\epsilon}(\lambda)d\lambda)_{\rho}$$

where $(A_2^{-1}B_2(x)\int_{t+s}^{\infty} v^{\epsilon}(\lambda)d\lambda)_{\ell}$ is the ℓ^{th} component of the m-vector $A_2^{-1}B_2(x)\int_{t+s}^{\infty} E_{t+s}v^{\epsilon}(\lambda)d\lambda$. Then by summing over i and then over ρ, and using (2.15) we get:

$$I_{42} = -\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)g'(x,t+s)F_j(x)A_2^{-1}B_2(x)\int_{t+s}^{\infty} E_{t+s}v^{\epsilon}(\lambda)d\lambda \tag{B-12}$$

$$I_{43} = -\sum_{j=1}^{n}\sum_{\rho,q=1}^{m} \frac{\partial f}{\partial x_j}(x)\xi_{j\rho}(x)\alpha_{\rho_q}(g'(x,t+s)E_q \int_{t+s}^{\infty} E_{t+s}v^{\epsilon}(\lambda)d\lambda)$$

where E_q is given by (2.14).

$$I_{43} = -\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)(A_{12}(x)A_2^{-1}w(x,t+s))_j \tag{B-13}$$

where $w(x,t+s)$ is a vector whose ith component is given by

$$g'(x,t+s)E_i\int_{t+s}^{\infty} v^{\epsilon}(\lambda)d\lambda \tag{B-14}$$

Then, from (B-3), (B-4) and (B-6)-(B-13), we have:

$$I = \int_0^{\infty} E_t[g'(x,t+s)\cdot\nabla_x f_1(x,t+s)]ds$$

$$= \int_0^{\infty} E_t[-\mu g'(x,t+s)f_{xx}(x)A_{12}(x)A_2^{-1}\bar{y}(t+s,x)$$

$$-\sum_{j=1}^{n}\mu\frac{\partial f}{\partial x_j}(x)g'(x,t+s)F_j(x)A_2^{-1}\bar{y}(t+s,x)$$

$$- \mu g'(x,t+s)(a_{21}(x))'_x(A_2')^{-2}A_{12}'(x)\nabla_x f(x)$$

$$+ g'(x,t+s)f_{xx}(x)B_0(x) \int_{t+s}^{\infty} E_{t+s}v^{\epsilon}(\lambda)d\lambda$$

$$+ \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)g'(x,t+s)D_j(x) \int_{t+s}^{\infty} E_{t+s}v^{\epsilon}(\lambda)d\lambda$$

$$- \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)g'(x,t+s)F_j(x)A_2^{-1}B_2(x) \int_{t+s}^{\infty} E_{t+s}v^{\epsilon}(\lambda)d\lambda$$

$$- \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)(A_{12}(x)A_2^{-1}w(x,t+s))_j ds \qquad \text{(B-15)}$$

From (B-2) and (B-5) we have:

$$g(x,t+s) = A_{12}(x)\overline{y}(t+s,x) + B_1(x)v^{\epsilon}(t+s)$$

Substituting g into (B-15) and then after simple manipulation we have:

$$I = \int_0^{\infty} [-\mu tr(f_{xx}(x)A_{12}(x)A_2^{-1}E_t(\overline{y}(t+s,x)\overline{y}'(t+s,x))A_{12}'(x))$$

$$- \mu tr(f_{xx}(x)A_{12}(x)A_2^{-1}E_t(\overline{y}(t+s,x)v'^{\epsilon}(t+s))B_1'(x))$$

$$- \mu tr(\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)F_j'(x)A_{12}E_t(\overline{y}(t+s,x)\overline{y}'(t+s,x))(A_2^{-1})')$$

$$- \mu tr(\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)F_j'(x)B_1(x)E_t(v^{\epsilon}(t+s)\overline{y}'(t+s,x))(A_2^{-1})')$$

$$+ \operatorname{tr}(f_{xx}(x)B_O(x) \int_{t+s}^{\infty} E_t v^{\varepsilon}(\lambda)\, d\lambda \bar{y}'(t+s,x)A'_{12}(x))$$

$$+ \operatorname{tr}(f_{xx}(x)B_O(x) \int_{t+s}^{\infty} E_t v^{\varepsilon}(\lambda) v'^{\varepsilon}(t+s)\, d\lambda B'_1(x))$$

$$+ \operatorname{tr}(\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) D'_j(x) A_{12}(x) \int_{t+s}^{\infty} E_t \bar{y}(t+s,x) v'^{\varepsilon}(\lambda)\, d\lambda$$

$$+ \operatorname{tr}(\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) D'_j(x) B_1(x) \int_{t+s}^{\infty} E_t v^{\varepsilon}(t+s) v'^{\varepsilon}(\lambda)\, d\lambda)$$

$$- \operatorname{tr}(\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) F'_j(x) A_{12}(x) \int_{t+s}^{\infty} E_t \bar{y}(t+s,x) v'^{\varepsilon}(\lambda)\, d\lambda$$

$$B'_2(x)(A^{-\frac{1}{2}})')$$

$$- \operatorname{tr}(\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) F'_j(x) B_1(x) \int_{t+s}^{\infty} E_t v^{\varepsilon}(t+s) v'^{\varepsilon}(\lambda)\, d\lambda$$

$$B'_2(x)(A-\tfrac{1}{2})')$$

$$- \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} E_t (A_{12}(x) A^{-\frac{1}{2}} w(x,t+s))_j$$

$$- \mu[A_{12}(x) E_t \bar{y}(t+s,x) + B_1(x) E_t v^{\varepsilon}(t+s)]'$$

$$(a_{21})'_x (A'_2)^{-2} f_x(x)]ds$$

$$= \sum_{i=1}^{12} I_i \qquad \text{(B-16)}$$

Now let us consider $\int_0^\infty E_t\bar{y}(t+s,x)\bar{y}'(t+s,x)ds$, where we set $\bar{y}(t) = y(t) + A_2^{-1}a_{21}(x)$

$$\int_0^\infty E_t\bar{y}(t+s,x)\bar{y}'(t+s,x)ds$$

$$= \int_0^\infty E_t[e^{A_2s/\mu}\bar{y}(t)\bar{y}'(t)e^{A_2's/\mu}$$

$$+ \frac{1}{\mu}\int_0^s e^{A_2(s-\tau)/\mu}B_2(x)v^\epsilon(t+\tau)\bar{y}'(t)e^{A_2's/\mu}d\tau$$

$$+ \frac{1}{\mu}e^{A_2s/\mu}y(t)\int_0^s v'^\epsilon(t+\lambda)B_2'(x)e^{A_2'(s-\lambda)/\mu}d\lambda$$

$$+ \frac{1}{\mu^2}\int_0^s\int_0^s e^{A_2(s-\tau)/\mu}B_2(x)v^\epsilon(t+\tau)v'^\epsilon(t+\lambda)$$

$$B_2'(x)e^{A_2'(s-\lambda)/\mu}d\tau d\lambda]ds \tag{B-17}$$

From the estimate on $y(t)$ and that $a_{21}(x)$ is bounded, we get

$$\left|\int_0^\infty e^{A_2s/\mu}\bar{y}(t)\bar{y}'(t)e^{A_2's/\mu}ds\right| \leq \left(K_1 + \frac{K_2}{\sqrt{\epsilon}}\right)^2 \int_0^\infty e^{-2\alpha_2s/\mu}ds \leq \frac{K\mu}{\epsilon} \leq K' \tag{B-18}$$

and for the second term in (B-17) we have, (notice that B_2 is bounded)

$$\left|\frac{1}{\mu}\int_0^\infty\int_0^s e^{A_2(s-\tau)/\mu} B_2(x)E_t v^\epsilon(t+\tau)d\tau\bar{y}'(t)e^{A_2's/\mu}ds\right|$$

$$\leq \frac{K}{\mu\epsilon}\int_0^\infty\int_0^s e^{-\alpha_2(2s-\tau)/\mu}d\tau ds = \frac{K}{\epsilon}\int_0^\infty [e^{-\alpha_2 s/\mu} - e^{-2\alpha_2 s/\mu}]ds$$

$$\leq \frac{K\mu}{\epsilon} \leq K \qquad \text{(B-19)}$$

We used that $|v^\epsilon(t+\tau)| \leq \frac{K}{\sqrt{\epsilon}}$ and that $\frac{\mu}{\epsilon} \leq K$.

From (B-17) and (B-16), the significant term of I_1 is:

$$L_1 = -\frac{\mu}{\mu^2}\int_0^\infty \operatorname{tr} f_{xx}(x)A_{12}A_2^{-1}\int_0^s\int_0^s e^{A_2(s-\tau)/\mu}$$

$$B_2(x)E_t(v^\epsilon(t+\tau)v'^\epsilon(t+\lambda))B_2'(x)e^{A_2'(s-\lambda)/\mu}A_{12}'(x)d\tau d\lambda ds$$

$$\text{(B-20)}$$

Subtracting and adding to (B-20), a term equal to the one that appears on its right hand side but with $E(v^\epsilon(t+\tau)v'^\epsilon(t+\lambda))$ replacing $E_t(v^\epsilon(t+\tau)v'^\epsilon(t+\lambda))$, L_1 can be written as:

$$L_1 = -\frac{1}{\mu}\int_0^\infty\int_0^s\int_0^s \operatorname{tr}(E_t v^\epsilon(t+\tau)v'^\epsilon(t+\lambda) - Ev^\epsilon(t+\tau)v'^\epsilon(t+\lambda))$$

$$B_2'(x)e^{A_2'(s-\lambda)/\mu}A_{12}'(x)f_{xx}(x)A_{12}(x)e^{A_2(s-\tau)/\mu}B_2(x)d\tau d\lambda ds$$

$$-\frac{1}{\mu}\int_0^\infty\int_0^s\int_0^s f_{xx}(x)A_{12}(x)A_2^{-1}e^{A_2(s-\tau)/\mu}$$

$$B_2(x)E(v^\epsilon(t+\tau)v'^\epsilon(t+\lambda))B_2'(x)e^{A_2'(s-\lambda)/\mu}A_{12}'(x)d\tau d\lambda ds$$

$$\text{(B-21)}$$

$$= L_{11} + L_{12}$$

From the mixing property (2.4), it follows [c.f. 36] that

$$|E_t v^\varepsilon(\tau)v^\varepsilon(\lambda) - Ev^\varepsilon(\tau)v^\varepsilon(\lambda)| \leq \frac{K}{\varepsilon} e^{-\alpha(\lambda-t)/\varepsilon} \tag{B-22}$$

for every $0 \leq t \leq \tau \leq \lambda$.

Then from (B-22),(A4) the bounded truncation state $x(t) = x$ and the compact support of $f_{xx}(x)$, we have:

$$|L_{11}| \leq \frac{K}{\mu\varepsilon} \int_0^\infty \int_0^s \int_0^\lambda e^{-\alpha(\lambda/\varepsilon)} e^{-\alpha_2(s-\lambda)/\mu} e^{-\alpha_2(s-\tau)/\mu} d\tau d\lambda ds$$

$$+ \frac{K}{\mu\varepsilon} \int_0^\infty \int_0^s \int_0^\tau e^{-\alpha(\tau/\varepsilon)} e^{-\alpha_2(s-\lambda)/\mu} e^{-\alpha_2(s-\tau)/\mu} d\lambda d\tau ds \tag{B-23}$$

changing the order of integration in (B-23). Then, for example, the first term in the right hand side becomes:

$$\frac{K}{\mu\varepsilon} \int_0^\infty \int_0^\lambda \int_\lambda^\infty e^{-\alpha\lambda/\varepsilon} e^{-\alpha_2(s-\lambda)/\mu} e^{-\alpha_2(s-\tau)/\mu} ds d\tau d\lambda$$

$$= \frac{K\mu}{2\alpha_2\mu\varepsilon} \int_0^\infty \int_0^\lambda e^{-(\frac{\alpha_2}{\mu} + \frac{\alpha}{\varepsilon})\lambda} e^{\alpha_2\tau/\mu} d\tau d\lambda$$

$$= \frac{K'\mu}{\varepsilon} \int_0^\infty \left(e^{-\frac{\alpha\lambda}{\varepsilon}} - e^{-(\frac{\alpha_2}{\mu} + \frac{\alpha}{\varepsilon})\lambda}\right) d\lambda \leq K\varepsilon$$

Similar estimate can be obtained for the second term of (B-23). Then, we have: (we used that $\mu/\varepsilon \leq K_1$)

$$|L_{11}| \leq K\varepsilon \tag{B-24}$$

Now we consider

$$L_{12} = -\frac{1}{\mu\epsilon}\int_0^\infty \int_0^s \int_0^\lambda \operatorname{tr} f_{xx}(x) A_{12}(x) A_2^{-1} e^{A_2(s-\tau)/\mu} B_2(x) R'\left(\frac{\lambda-\tau}{\epsilon}\right)$$

$$B_2'(x) e^{A_2'(s-\lambda)/\mu} A_{12}'(x)\, d\tau\, d\lambda\, ds$$

$$-\frac{1}{\mu\epsilon}\int_0^\infty \int_0^s \int_0^\tau \operatorname{tr} f_{xx}(x) A_{12}(x) A_2^{-1} e^{A_2(s-\tau)/\mu} B_2(x) R\left(\frac{\lambda-\tau}{\epsilon}\right)$$

$$B_2'(x) e^{A_2'(s-\lambda)/\mu} A_{12}'(x)\, d\lambda\, d\tau\, ds$$

$$= T_1 + T_2$$

where, the correlation matrix $R(\tau) = E(v(t+\tau)v'(t))$ satisfies, (also follows from (2.4) and [36]):

$$|R(\tau)| \le Ke^{-\alpha\tau} \tag{B-26}$$

setting $\lambda - \tau = w$ in T_1 and changing the order of integrabon, we get:

$$T_1 = -\frac{1}{\mu\epsilon}\int_0^\infty \int_0^s \int_w^s \operatorname{tr} f_{xx}(x) A_{12}(x) A_2^{-1} e^{A_2(s+w-\lambda)/\mu} B_2(x) R'\left(\frac{w}{\epsilon}\right)$$

$$B_2'(x) e^{A_2'(s-\lambda)/\mu} A_{12}'(x)\, d\lambda\, dw\, ds$$

Integrating by parts just once, gives

$$T_1 = \frac{1}{\varepsilon}\int_0^\infty \int_0^s \mathrm{tr} f_{xx}(x) A_{12}(x) A_2^{-1} e^{A_2 w/\mu} B_2(x) R'(\tfrac{w}{\varepsilon})$$

$$B_2'(x)(A_2')^{-1} A_{12}'(x)\, dw ds$$

$$- \frac{1}{\varepsilon}\int_0^\infty \int_0^s \mathrm{tr} f_{xx}(x) A_{12}(x) A_2^{-1} e^{A_2 s/\mu} B_2(x) R'(\tfrac{w}{\varepsilon})$$

$$B_2(x) e^{A_2'(s-w)/\mu} (A_2')^{-1} A_{12}'(x)\, dw ds$$

$$+ \frac{1}{\mu\varepsilon}\int_0^\infty \int_0^s \int_w^s \mathrm{tr} f_{xx}(x) A_{12}(x) e^{A_2(s+w-\lambda)/\mu} B_2(x) R'(\tfrac{w}{\varepsilon})$$

$$B_2'(x) e^{A_2'(s-\lambda)/\mu} (A_2')^{-1} A_{12}'(x)\, d\lambda dw ds \qquad \text{(B-27)}$$

Similarly, setting $\tau - \lambda = w$ in T_2, changing the order of integration and then replacing the dummy variable τ by λ we get:

$$T_2 = -\frac{1}{\mu\varepsilon}\int_0^\infty \int_0^s \int_w^s \mathrm{tr} f_{xx}(x) A_{12}(x) A_2^{-1} e^{A_2(s-\lambda)/\mu} B_2(x) R(\tfrac{w}{\varepsilon})$$

$$B_2'(x) e^{A_2'(s+w-\lambda)/\mu} A_{12}'(x)\, d\lambda dw ds$$

and integration by parts once implies:

$$T_2 = \frac{1}{\varepsilon}\int_0^\infty \int_0^s \mathrm{tr} f_{xx}(x) A_{12}(x) A_2^{-1} B_2(x) R(\tfrac{w}{\varepsilon}) B_2'(x) e^{A_2' w/\mu} (A_2')^{-1}$$

$$A_{12}'(x)\, dw ds$$

$$-\frac{1}{\varepsilon}\int_0^\infty \int_0^s \mathrm{tr} f_{xx}(x) A_{12}(x) A_2^{-1} e^{A_2(s-w)/\mu} B_2(x) R(\tfrac{w}{\varepsilon}) B_2'(x) e^{A_2 s/\mu}$$

$$(A_2')^{-1}A_{12}'(x)\,dwds$$

$$+ \frac{1}{\mu\epsilon}\int_0^\infty \int_0^s \int_w^s \mathrm{tr} f_{xx}(x)A_{12}(x)e^{A_2(s-\lambda)/\mu}B_2(x)R(\tfrac{w}{\epsilon})$$

$$B_2'(x)e^{A_2(s+w-\lambda)/\mu}(A_2')^{-1}A_{12}'(x)\,d\lambda dwds \qquad \text{(B-28)}$$

Employing the facts that for any matrix A, $\mathrm{tr}A = \mathrm{tr}A'$ and for any matrices $X_{n\times m}$ and $Y_{m\times n}$, $\mathrm{tr}(XY) = \mathrm{tr}(X'Y')$ $= \mathrm{tr}(YX)$. We observe the following:

(i) The first two terms in (B-27) are equal to the first two terms in (B-28) respectively.

(ii) The third term of (B-27) = $-T_2$ and the third term of (B-28) = $-T_1$.

From these observations and (B-25), we conclude that:

$$L_{12} = \frac{1}{\epsilon}\int_0^\infty \int_0^s \mathrm{tr} f_{xx}(x)A_{12}(x)A_2^{-1}e^{A_2 w/\mu}B_2(x)R'(\tfrac{w}{\epsilon})B_2'(x)$$

$$(A_2')^{-1}A_{12}(x)\,dwds$$

$$- \frac{1}{\epsilon}\int_0^\infty \int_0^s \mathrm{tr} f_{xx}(x)A_{12}(x)A_2^{-1}e^{A_2 s/\mu}B_2(x)R'(\tfrac{w}{\epsilon})B_2'(x)$$

$$e^{A_2(s-w)/\mu}(A_2')^{-1}A_{12}'(x)\,dwds \qquad \text{(B-29)}$$

The first term in B-29

$$= \int_0^\infty \mathrm{tr} f_{xx}(x) A_{12}(x) A_2^{-1} \Sigma B_2'(x) (A_2')^{-1} A_{12}(x)\, ds$$

$$+ \int_0^\infty \int_{s/\epsilon}^0 \mathrm{tr} f_{xx}(x) A_{12}(x) A_2^{-1} e^{A_2 w(\frac{\epsilon}{\mu})} B_2(x) R'(w) B_2'(x)$$

$$(A_2')^{-1} A_{12}(x)\, dwds \tag{B-31}$$

It can be seen that the second term in (B-30) $\leq K\mu$. And the second term in (B-29) is bounded by $K_1\epsilon + K_2\mu$ where we used (B-26), (A4), boundendess of $x(t)$, compact support of $f_{xx}(x)$ and that $\frac{\mu}{\epsilon} \leq K$. So, from (B-16), (B-18), (B-19), (B-21), (B-24), (B-29), (B-30) and (B-31) we conclude that:

$$I_1 = \int_0^\infty [-\mu \mathrm{tr}(f_{xx}(x) A_{12}(x) A_2^{-1} E_t(\bar{y}(t+s,x)\bar{y}'(t+s,x)) A_{12}'(x))\, ds$$

$$= \int_0^\infty \mathrm{tr}(f_{xx}(x) A_{12}(x) A_2^{-1} \Sigma B_2'(x) (A_2')^{-1} A_{12}(x))\, ds + e_1 \tag{B-32}$$

where, Σ is defined by (2.16) and

$$|e_1| \leq K_1\epsilon + K_2\mu, \tag{B-33}$$

where

K_1, K_2 are some positive constants independent of

T and α

Now, let us consider I_3 of (B-16). From (B-17) we see that the important term in I_3 is similar to (B-20), namely,

$$L_2 = -\frac{1}{\mu}\int_0^\infty tr\Big(\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)F_j'(x)A_{12}(x)\int_0^s\int_0^s e^{A_2(s-\tau)/\mu}$$

$$B_2(x)E_t(v^\epsilon(t+\tau)v'^\epsilon(t+\lambda))B_2'(x)e^{A_2'(s-\lambda)/\mu}(A_2')^{-1}\Big)d\tau d\lambda ds$$

(B-34)

Repeating similar steps to the ones that has been made to get from (B-20) to (B-21), we can express (B-34) as the sum of two terms, i.e.,

$$L_2 = L_{21} + L_{22} \tag{B-35}$$

and from (B-22), the boundedness of $F_j(x)$ and the same reasoning as before, we get (note: L_{21} involves the expression appearing in (B-22))

$$|L_{21}| \leq K\epsilon$$

and similar to (B-25), we get:

$$L_{22} = -\frac{1}{\mu\epsilon}\int_0^\infty\int_0^s\int_0^\lambda \sum_{j=1}^{n} \cdot\frac{\partial f}{\partial x_j}(x)\,trF_j'(x)A_{12}(x)e^{A_2(s-\tau)/\mu}$$

$$B_2(x)R'\Big(\frac{\lambda-\tau}{\epsilon}\Big)B_2'(x)e^{A_2'(s-\lambda)/\mu}(A_2')^{-1}d\tau d\lambda ds$$

$$- \frac{1}{\mu\epsilon} \int_0^\infty \int_0^s \int_0^\tau \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) \operatorname{tr} F_j'(x) A_{12}(x) e^{A_2(s-\tau)/\mu}$$

$$B_2(x) R'\left(\frac{\lambda-\tau}{\epsilon}\right) B_2'(x) e^{A_2'(s-\lambda)/\mu} (A_2')^{-1} d\tau d\lambda ds$$

$$- \frac{1}{\mu\epsilon} \int_0^\infty \int_0^s \int_0^\tau \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) \operatorname{tr} F_j'(x) A_{12}(x) e^{A_2(s-\tau)/\mu}$$

$$B_2(x) R\left(\frac{\tau-\lambda}{\epsilon}\right) B_2'(x) e^{A_2'(s-\lambda)/\mu} (A_2')^{-1} d\tau d\lambda ds \qquad \text{(B-37)}$$

After some manipulations it can be shown that

$$L_{22} = - \int_0^\infty \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) \operatorname{tr} F_j'(x) A_{12}(x) \int_0^\infty e^{A_2\lambda} \Big(\int_0^\infty e^{A_2(\frac{\epsilon}{\mu})w}$$

$$B_2(x) R'(w) dw B_2'(x) + B_2(x) \int_0^\infty R(w) B_2'(x) e^{A_2'(\frac{\epsilon}{\mu})w} dw \Big)$$

$$e^{A_2'\lambda} d\lambda (A_2')^{-1} ds + e_2$$

$$= - \int_0^\infty \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) \operatorname{tr} F_j'(x) A_{12}(x) \int_0^\infty e^{A_2\lambda} (\Sigma B_2'(x) + B(x) \Sigma')$$

$$e^{A_2'\lambda} d\lambda (A_2')^{-1} ds + e_2$$

$$= - \int_0^\infty \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) \operatorname{tr} F_j'(x) A_{12}(x) P (A_2^{-1})' ds + e_2 \qquad \text{(B-38)}$$

where P is defined by (2.17) and

$$|e_2| \leq K\mu \qquad \text{(B-39)}$$

Similar to (B-18) and (B-19), it follows that the contribution of the first three terms of (B-17) to I_3 of (B-16) is $O(\mu)$. Combining this, with the results of (B-35)-(B-39) we get the following estimate:

$$I_3 = \int_0^\infty -\mu \operatorname{tr}\Big(\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)F_j'(x)A_{12}(x)$$

$$E_t(\bar{y}(t+s,x)\bar{y}'(t+s,x)(A_2^{-1})')ds$$

$$= -\operatorname{tr}\int_0^\infty \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)F_j'(x)A_{12}'(x)P(A_2^{-1})'ds + e_3 \qquad \text{(B-40)}$$

where

$$|e_3| \leq K_1\varepsilon + K_2\mu \qquad \text{(B-41)}$$

for some positive constants K_1 and K_2 independent of T and ω. Now let us consider

$$\int_0^\infty E_t\bar{y}(t+s,x)v'^{\varepsilon}(t+s)ds = \int_0^\infty e^{A_2 s/\mu}\bar{y}(t)E_t v'^{\varepsilon}(t+s)ds$$

$$+ \frac{1}{\mu}\int_0^\infty \int_0^s e^{A_2(s-\tau)/\mu} B_2(x)E_t v^{\varepsilon}(t+\tau)v'^{\varepsilon}(t+s)d\tau ds$$

$$\text{(B-42)}$$

Using the mixing property and that $\bar{y}(t)$ is $O(\frac{1}{\sqrt{\varepsilon}})$ we get:

$$\Big|\int_0^\infty e^{A_2 s/\mu}\bar{y}(t)E_t v^{\varepsilon}(t+s)ds\Big| \leq \frac{K}{\varepsilon}\int_0^\infty e^{-(\alpha_2/\mu+\frac{\alpha}{\varepsilon})s}ds \leq \frac{K\mu}{\alpha_2\varepsilon+\alpha\mu} \leq K$$

$$\text{(B-43)}$$

Let T^* equal to the second term of (B-42) then we have:

$$T^* = \frac{1}{\mu}\int_0^\infty \int_0^s e^{A_2(s-\tau)/\mu} B_2(x)(E_t v^\varepsilon(t+\tau)v'^\varepsilon(t+s)$$

$$- Ev^\varepsilon(t+\tau)v^\varepsilon(t+s))d\tau ds + \frac{1}{\mu}\int_0^\infty \int_0^s e^{A_2(s-\tau)/\mu}$$

$$B_2(x)E(v^\varepsilon(t+\tau)v'^\varepsilon(t+s))d\tau ds$$

$$= T_1 + T_2 \tag{B-44}$$

Using (B-22) and that $B_2(x)$ is bounded we have:

$$|T_1| \leq \frac{K}{\mu\varepsilon}\int_0^\infty \int_0^s e^{-\alpha_2(s-\tau)/\mu} e^{-\alpha s/\varepsilon}\, d\tau ds \leq K \tag{B-45}$$

If we consider I_2 in (B-16) with (B-42)-(B-45) we get:

$$I_2 = -\mu \int_0^\infty tr(f_{xx}(x)A_{12}(x)A_2^{-1}E_t(\bar{y}(t+s,x)v'^\varepsilon(t+s))B_1'(x))ds$$

$$= -\mu \int_0^\infty tr(f_{xx}(x)A_{12}(x)A_2^{-1} \cdot \frac{1}{\mu}\int_0^s e^{A_2(s-\tau)/\mu} B_2(x)$$

$$E(v^\varepsilon(t+\tau)v'^\varepsilon(t+s))B_1'(x)d\tau ds) + e_3 \tag{B-46}$$

where $|e_3| \leq K_1\varepsilon + K_2\mu$.

First term of (B-46)

$$= -tr\int_0^\infty \int_0^\infty f_{xx}(x)A_{12}(x)A_2^{-1}e^{A_2(\frac{\varepsilon}{\mu})w} B_2(x)R'(w)B_1'(x)dwds$$

$$+ \operatorname{tr} \int_0^\infty \int_{s/\epsilon}^\infty f_{xx}(x) A_{12}(x) A_2^{-1} e^{A_2(\frac{\epsilon}{\mu})w} B_2(x) R'(w) B_1'(x)\,dw\,ds$$

$$= - \int_0^\infty \operatorname{tr} f_{xx}(x) A_{12}(x) A_2^{-1} \Sigma B_1'(x)\,ds + e_4 \qquad \text{(B-47)}$$

where

$$|e_4| \leq K_1 \epsilon + K_2 \mu \qquad \text{(B-48)}$$

Hence, it follows from (B-46)-(B-48) that:

$$I_2 = - \int_0^\infty \operatorname{tr} f_{xx}(x) A_{12}(x) A_2^{-1} \Sigma B_1'(x)\,ds + e_5, \qquad \text{(B-49)}$$

$$|e_5| \leq K_1 \epsilon + K_2 \mu \qquad \text{(B-50)}$$

Following steps similar to what has been done in (B-42)-(B-50) we conclude that:

$$I_4 = - \mu \int_0^\infty \operatorname{tr}\Big(\sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) F_j'(x) B_1(x) E_t(v^\epsilon(t+s)\bar{y}'(t+s,x))$$

$$(A_2^{-1})')\,ds$$

$$= - \int_0^\infty \operatorname{tr}\Big[\sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) F_j'(x) B_1'(x)\, \Sigma' \,(A_2^{-1})'\Big]\,ds + e_6 \qquad \text{(B-51)}$$

where

$$|e_6| \leq K_1 \epsilon + K_2 \mu . \qquad \text{(B-52)}$$

Then, we consider the integral:

$$\int_0^\infty \int_{t+s}^\infty E_t\bar{y}(t+s,x)v'^{\varepsilon}(\lambda)d\lambda ds = \int_0^\infty e^{A_2 s/\mu}\bar{y}(t)\int_{t+s}^\infty E_t v'^{\varepsilon}(\lambda)d\lambda ds$$

$$+\frac{1}{\mu}\int_0^\infty \int_s^\infty \int_0^s e^{A_2(s-\tau)/\mu} B_2(x)E_t v^{\varepsilon}(t+\tau)v'^{\varepsilon}(t+\lambda)d\tau d\lambda ds$$

(B-53)

$$= \varphi_1 + \varphi_2$$

$$|\varphi_1| \le \frac{K}{\varepsilon}\int_0^\infty \int_s^\infty e^{-\alpha_2 s/\mu}\, e^{-\alpha\lambda/\varepsilon}\, d\lambda ds \le K_1\varepsilon + K_2\mu \qquad \text{(B-54}$$

where we have used the mixing property and that $\bar{y}(t)$ is $O(\frac{1}{\sqrt{\varepsilon}})$.

$$\varphi_2 = \frac{1}{\mu}\int_0^\infty \int_s^\infty \int_0^s e^{A_2(s-\tau)/\mu} B_2(x)\,(E_t v^{\varepsilon}(t+\tau)v'^{\varepsilon}(t+\lambda)$$

$$- Ev^{\varepsilon}(t+\tau)v'^{\varepsilon}(t+\lambda))d\tau d\lambda ds$$

$$+\frac{1}{\mu}\int_0^\infty \int_s^\infty \int_0^s e^{A_2(s-\tau)/\mu} B_2(x)E(v^{\varepsilon}(t+\tau)v'^{\varepsilon}(t+\lambda))d\tau d\lambda ds$$

$$= \varphi_{21} + \varphi_{22} \qquad \text{(B-55)}$$

From (B-22) we get:

$$|\varphi_{21}| \le \frac{1}{\mu}\int_0^\infty \int_s^\infty \int_0^s e^{-\alpha_2(s-\tau)/\mu}\, e^{-\alpha\lambda/\varepsilon}\, d\tau d\lambda ds \le K_1\varepsilon + K_2\mu$$

(B-56)

$$\varphi_{22} = \frac{1}{\mu\varepsilon}\int_0^\infty \int_s^\infty \int_0^s e^{A_2(s-\tau)/\mu} B_2(x)R'(\frac{\lambda-\tau}{\varepsilon})d\tau d\lambda ds$$

Changing order of integration, we get

$$\varphi_{22} = \frac{1}{\mu\varepsilon}\int_0^\infty \int_0^\lambda \int_\tau^\lambda e^{A_2(s-\tau)/\mu} B_2(x)R'(\frac{\lambda-\tau}{\varepsilon})\,ds\,d\tau\,d\lambda$$

$$= \int_0^\infty (A_2^{-1}\Sigma - A_2^{-1}W')\,ds + e_7 \tag{B-57}$$

where

$$|e_7| \leq K_1\varepsilon + K_2\mu \tag{B-58}$$

and W is defined by $\int_0^\infty R(\tau)d\tau$. Then I_7 in (B-16), in view of (B-53)-(B-58) can be written as:

$$I_7 = \int_0^\infty tr(\sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)D_j'(x)A_{12}(x) \int_{t+s}^\infty E_t\bar{y}(t+s,x)v'^{\varepsilon}(\lambda)\,d\lambda$$

$$= \int_0^\infty [tr(\sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)D_j'(x)A_{12}(x)A_2^{-1}\Sigma - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)D_j'(x)A_{12}(x)$$

$$A_2^{-1}B_2(x)W')]ds + e_8 \tag{B-59}$$

where

$$|e_8| \leq K_1\varepsilon + K_2\mu \tag{B-60}$$

where K_1, K_2 are some positive constants independent of T and α. We have used that $f_x(x)$ has compact support, (A2) and (A4). Similarly I_5 can be handeled as I_7 and we get:

$$I_5 = \int_0^\infty tr(f_{xx}(x)B_0(x) \int_{t+s}^\infty E_t v^\varepsilon(\lambda)\bar{y}'(t+s,x)d\lambda A_{12}'(x))ds$$

$$= \int_0^\infty tr[f_{xx}(x)A_{12}(x)A_2^{-1}\Sigma B_0'(x) - f_{xx}(x)A_{12}(x)A_2^{-1}W'B_0'(x)]ds$$

$$+ e_9 \tag{B-61}$$

where

$$|e_9| \leq K_1\varepsilon + K_2\mu \tag{B-62}$$

and K_1, K_2 are some positive constants independent of T and ω. Also, in the same way:

$$I_9 = \int_0^\infty tr[-\sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)F_j'(x)A_{12}(x) \int_{t+s}^\infty E_t\bar{y}(t+s,x)v'^\varepsilon(\lambda)d\lambda$$

$$B_2'(x)(A_2')^{-1}]ds$$

$$= -\int_0^\infty tr \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)F_j'(x)A_{12}(x)A_2^{-1}\Sigma B_2'(x)(A'_2)^{-1}ds$$

$$+ \int_0^\infty tr \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)F_j'(x)A_{12}(x)A_2^{-1}W'B_2'(x)(A_2')^{-1}ds + e_{10}$$

$$\tag{B-63}$$

where

$$|e_{10}| \leq K_1\varepsilon + K_2\mu \tag{B-64}$$

Let us consider the integral

$$J = \int_0^\infty \int_{t+s}^\infty E_t v^\epsilon(t+s) v'^\epsilon(\lambda)\, d\lambda ds$$

$$= \int_0^\infty \int_s^\infty E_t v^\epsilon(t+s) v'^\epsilon(t+\lambda)\, d\lambda ds$$

$$= \int_0^\infty \int_s^\infty [E_t v^\epsilon(t+s) v'^\epsilon(t+\lambda) - Ev^\epsilon(t+s) v'^\epsilon(t+\lambda)]\, d\lambda ds$$

$$+ \int_0^\infty \int_s^\infty Ev^\epsilon(t+s) v'^\epsilon(t+\lambda)\, d\lambda ds = J_1 + J_2 \qquad \text{(B-65)}$$

From (B-22), we have

$$|J_1| \leq K\epsilon \qquad \text{(B-66)}$$

$$J_2 = \frac{1}{\epsilon} \int_0^\infty \int_0^\infty R'\left(\frac{w}{\epsilon}\right) dw ds = \int_0^\infty W' ds \qquad \text{(B-67)}$$

Then, from (B-65)-(B-67), (B-16), compact support of f_{xx}, (A2) and (A4), we get:

$$I_6 = \int_0^\infty tr(f_{xx}(x) B_0(x) \int_{t+s}^\infty E_t v^\epsilon(\lambda) v'^\epsilon(t+s)\, d\lambda B_1'(x))\, ds$$

$$= \int_0^\infty tr(f_{xx}(x) B_1(x) W' B_0'(x))\, ds + e_{11} \qquad \text{(B-68)}$$

where

$$|e_{11}| \leq K_1\epsilon + K_2\mu \qquad \text{(B-69)}$$

and K_1, K_2 are some positive constants independent of T and μ. Also, we have from (B-65)-(B-67), (B-16),

compact support of f_x, (A2) and (A4) we get:

$$I_8 = \int_0^\infty \operatorname{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) \; D_j'(x) B_1(x) \int_{t+s}^\infty E_t v^\epsilon(t+s) v'^\epsilon(\lambda) d\lambda$$

$$= \int_0^\infty \operatorname{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) \; D_j'(x) B_1(x) W' ds + e_{12} \qquad \text{(B-70)}$$

where

$$|e_{12}| \leq K_1 \epsilon + K_2 \mu \qquad \text{(B-71)}$$

and K_1, K_2 are some positive constants independent of T, ω. From (B-65)-(B-67), (B-16) and the same assumptions as before, we get:

$$I_{10} = -\int_0^\infty \operatorname{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) \; F_j'(x) B_1(x) \int_{t+s}^\infty E_t v^\epsilon(t+s) v'^\epsilon(\lambda) d\lambda$$

$$B_2'(x) (A_2^{-1})' ds$$

$$= -\int_0^\infty \operatorname{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) \; F_j'(x) B_1(x) W' B_2'(x) (A_2')^{-1} ds + e_{13}$$

(B-72)

where

$$|e_{13}| \leq K_1 \epsilon + K_2 \mu \qquad \text{(B-73)}$$

and K_1, K_2 are some positive constants independent of T and ω. From (B-14) and (B-16)

$$I_{11} = -\int_0^\infty \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) E_t (A_{12}(x) A_2^{-1} w(x, t+s))_j$$

$$= -\int_0^\infty \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)\,(A_{12}(x) A_2^{-1} h_2(x))_j ds + e_{14} \qquad \text{(B-74)}$$

where

$$|e_{14}| \leq K_1 \epsilon + K_2 \mu \qquad \text{(B-75)}$$

and $w(x,t+s)$, $h_{2i}(x)$ are defined by (B-14) and (2.11) respectively. Finally, it can be shown that

$$|I_{12}| \leq K_1 \epsilon + K_2 \mu$$

for some positive constants K_1, K_2 independent of T and ω. Now we add (B-32), (B-40), (B-49), (B-51), (B-59), (B-61), (B-63), (B-68), (B-70), (B-72) and (B-74), and let e denote the sum of all e_i that appear in the above equations and then from (B-3) and (B-16) we have:

$$I = \sum_{i=1}^{12} I_i = \int_0^\infty E_t [\frac{\partial f_1}{\partial x}(x, t+s) \cdot g(x, t+s)] ds$$

$$= \int_0^\infty [\mathrm{tr} f_{xx}(x) A_{12}(x) A_2^{-1} \Sigma B_2'(x) (A_2')^{-1} A_{12}'(x)$$

$$- \mathrm{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) F_j'(x) A_{12}'(x) P (A_2')^{-1}$$

$$- \mathrm{tr} f_{xx}(x) A_{12}(x) A_2^{-1} \Sigma B_1'(x) - \mathrm{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) F_j'(x)$$

$$B_1(x)\, \Sigma' (A_2^{-1})'$$

$$+ \operatorname{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) D_j'(x) A_{12}(x) A_2^{-1} \Sigma$$

$$- \operatorname{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) D_j'(x) A_{12}(x) A_2^{-1} B_2(x) W'$$

$$+ \operatorname{tr} f_{xx}(x) A_{12}(x) A_2^{-1} \Sigma B_0'(x) - \operatorname{tr} f_{xx}(x) A_{12}(x) A_2^{-1} B_2(x) W' B_0'(x)$$

$$- \operatorname{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) F_j'(x) A_{12}(x) A_2^{-1} \Sigma B_2'(x) (A_2')^{-1}$$

$$+ \operatorname{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) F_j'(x) A_{12}(x) A_2^{-1} W' B_2'(x) (A_2')^{-1}$$

$$+ \operatorname{tr} f_{xx}(x) B_1(x) W \; B_0'(x) + \operatorname{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) D_j'(x) B_1(x) W'$$

$$- \operatorname{tr} \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) F_j'(x) B_1(x) W' B'_2(x) (A')^{-1}$$

$$- \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) (A_{12}(x) A_2^{-1} h_2(x))_j + e \qquad \text{(B-77)}$$

where e is the sum of all e_k which appears in the right hand side of each I_i for $i = 1, 2, \cdots, 12$, and it satisfies

$$|e| \leq K_1 \epsilon + K_2 \mu \qquad \text{(B-78)}$$

where K_1, K_2 are some positive constants independent of T and ω.

From the definition of $B_0(x)$, the first term, the third term and the seventh term of (B-77) will cancel.

Moreover, we notice that

$$\mathrm{tr} f_{xx}(x)[B_1(x)W'B_0'(x) - A_{12}(x)A_2^{-1}B_2(x)W'B_0'(x)]$$

$$= \mathrm{tr} f_{xx}(x)B_0(x)W'B_0'(x) = \frac{1}{2}\,\mathrm{tr} f_{xx}(x)B_0(x)S(0)B_0'(x)$$

$$= \frac{1}{2}\,\mathrm{tr} f_{xx}(x)A(x) \qquad \text{(B-79)}$$

where $A(x)$ is defined by (2.9)

From the definition of P in (2.17), P satisfies the Lyapunov equation

$$PA_2' + A_2P = -(\Sigma B_2' + B_2\Sigma') \qquad \text{(B-80)}$$

From (B-80) and (B-77) we have:

$$\mathrm{tr}[-F_j'(x)A_{12}(x)A_2^{-1}\Sigma B_2'(x)(A_2^{-1})' - F_j'(x)A_{12}(x)P(A_2^{-1})'$$

$$- F_j'(x)B_1(x)\,\Sigma'\,(A_2^{-1})']$$

$$= \mathrm{tr}[-F_j'(x)B_0(x)\,\Sigma'\,(A_2^{-1})' + F_j'(x)A_{12}(x)A_2^{-1}P] \qquad \text{(B-81)}$$

Then, from (B-77) and (B-79)-(B-81) we have:

$$I = \int_0^\infty [\sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)(h_{1j}(x) - (A_{12}(x)A_2^{-1}h_2(x))_j + h_{3j}(x))$$

$$+ \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x)\,\frac{\partial^2 f}{\partial x_i \partial x_j}(x)]ds + e \qquad \text{(B-82)}$$

where h_{ij}, h_{2j} and h_{3j} are given by (2.10), (2.11) and (2.12) respectively. Then from (2.7)-(2.9) and (B-82) we have:

$$I = \int_0^\infty (L^{\epsilon/\mu} f(x) - \frac{\partial f}{\partial x}(x) a_0(x))ds + e \tag{B-83}$$

Therefore, by defining $(x = x(t))$

$$f_2(x,t) = \int_0^\infty [E_t[\frac{\partial f_1}{\partial x}(x,t+s)g(x,t+s)] + \frac{\partial f}{\partial x}(x)a_0(x)$$
$$- L^{\epsilon/\mu} f(x)]ds$$

It follows directly from (B-83) and (B-78) that

$$|f_2(x,t)| \leq K_1 \epsilon + K_2 \mu$$

where K_1 and K_2 are some positive constants independent of T and ω as required.

APPENDIX C

We first derive equation (2.43) and (2.44):

Let $X = \begin{pmatrix} x \\ y \end{pmatrix}$, then (2.1) and (2.2) can be written in the form:

$$\dot{X} = \begin{pmatrix} a_1(x) + A_{12}(x)y \\ \\ \frac{1}{\mu}(a_{21}(x) + A_2 y) \end{pmatrix} + \frac{1}{\sqrt{\varepsilon}} \begin{pmatrix} B_1(x)v^{\varepsilon} \\ \\ \tilde{B}_2(x)v^{\varepsilon} \end{pmatrix} \tag{C-1}$$

where $\mu > 0$ is small, arbitrary but fixed and $\tilde{B}_2(x) = \frac{1}{\mu}B_2(x)$. Equation (C-1) is of the form:

$$\dot{X} = \frac{1}{\sqrt{\varepsilon}} F(X,v) + G(X,v) \tag{C-2}$$

which has been considered in [11-13], but let us apply, for example, the result of [13] on the system (C-2), where

$$F(X,v) = \begin{pmatrix} B_1(x)v^{\varepsilon} \\ \tilde{B}_2(x)v^{\varepsilon} \end{pmatrix} \quad , \quad v^{\varepsilon}(t) = v(t/\varepsilon) \tag{C-3}$$

and

$$G(X,v) = \begin{pmatrix} a_1(x) + A_{12}(x)y \\ \frac{1}{\mu}(a_{21}(x) + A_2 y) \end{pmatrix} \tag{C-4}$$

The convergence result of [13], says that under certain conditions $X(\cdot)$ converges to a diffusion process $X^o(\cdot)$ whose differential operator is given by:

$$\begin{aligned} Af(X) &= EG'(X,v(s))f_s(X) \\ &+ \int_0^\infty EF'(X,v(s))(F'(X,v(s+\tau))f_X(X))_X)\,d\tau \qquad \text{(C-5)} \\ &= I_1 + I_2 \end{aligned}$$

where

$$f_X = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)'$$

$$I_2 = E\int_0^\infty (v'(s)B_1'(x)v'(s)\tilde{B}_2'(x))$$

$$\begin{pmatrix} \nabla_x(v'(\tau+s)B_1'(x)f_x(X) + v'(\tau+s)\tilde{B}_2'(x)f_y(X)) \\ \nabla_y(v'(\tau+s)B_1'(x)f_x(X) + v'(\tau+s)\tilde{B}_2'(x)f_y(X)) \end{pmatrix} d\tau$$

$$\begin{aligned} &= E\int_0^\infty (v'(s)B_1'(x)\nabla_x(v'(\tau+s)B_1'(x)f_x(X)) \\ &+ v'(s)B_1'(x)\nabla_x(v'(\tau+s)\tilde{B}_2'(x)f_y(X)) \\ &+ v'(s)\tilde{B}_2'(x)\nabla_y(v'(\tau+s)B_1'(x)f_x(X)) \\ &+ v'(s)\tilde{B}_2'(x)\nabla_y(v'(\tau+s)\tilde{B}_2'(x)f_y(X)))\,d\tau \qquad \text{(C-6)} \end{aligned}$$

Consider the first term of, I_2

$$\int_0^\infty Ev'(s)B_1'(x)\nabla_x(v'(\tau+s)B_1'(x)f_x(X))d\tau$$

$$= \mathrm{tr}\int_0^\infty EB_1(x)v(s)(\nabla_x(f_x'(X)B_1(x)v(\tau+s))'d\tau$$

$$= \mathrm{tr}\int_0^\infty EB_1(x)v(s)(f_{xx}(x)B_1(x)v(\tau+s)$$

$$+ \sum_{i,j=1}^{n} \frac{\partial f}{\partial x_i}(X)\nabla_x\psi_{ij}(x)v_j(\tau+s))'d\tau$$

$$= \mathrm{tr}\int_0^\infty EB_1(x)v(s)(f_{xx}(X)B_1(x)v(\tau+s)$$

$$+ \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X)D_i(x)v(\tau+s))'d\tau$$

$$= \mathrm{tr}\int_0^\infty EB_1(x)v(s)(v'(\tau+s)B_1'(x)f_{xx}(X)$$

$$+ \sum_{i=1}^{n} v'(\tau+s)D_i'(x)\frac{\partial f}{\partial x_i}(X))d\tau$$

$$= \mathrm{tr}(B_1(x)W'B_1'(x)f_{xx}(X)) + \mathrm{tr}\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(X)B_1(x)W'D_i'(x) \tag{C-7}$$

where ψ_{ij}, D_i and W are defined as before.

The second term of I_2 is given by:

$$\int_0^\infty Ev'(s)B_1'(x)\nabla_x(v'(\tau+s)\tilde{B}_2'(x)\frac{\partial f}{\partial y}(X))d\tau$$

$$= \mathrm{tr}\int_0^\infty [EB_1(x)v(s)(f_{yx}(X)\tilde{B}_2(x)v(\tau+s))'$$

$$+ \sum_{i=1}^{m}\frac{\partial f}{\partial y_i}(X)E_iv(\tau+s))']d\tau$$

$$= \mathrm{tr}(B_1(x)W'\tilde{B}_2'(x)f_{yx}'(X)) + \mathrm{tr}\sum_{i=1}^{m}\frac{\partial f}{\partial y_i}(X)B_1(x)W'E_i'$$

(C-8)

The third term of I_2 is given by:

$$\int_0^\infty Ev'(s)\tilde{B}_2'(x)\nabla_y(v'(\tau+s)B_1'(x)\frac{\partial f}{\partial x}(X))d\tau$$

$$= \mathrm{tr}\int_0^\infty E\tilde{B}_2(x)v(s)(f_{xy}(X)B_1(x)v(\tau+s))'d\tau$$

$$= \mathrm{tr}\tilde{B}_2(x)W'B_1'(x)f_{xy}'(X) \qquad \text{(C-9)}$$

The fourth term is

$$\int_0^\infty E(v'(s)\tilde{B}_2'(x)\nabla_y(v'(\tau+s)\tilde{B}_2'(x)\frac{\partial f}{\partial y}(X)))d\tau$$

$$= \mathrm{tr}(\tilde{B}_2(x)W'\tilde{B}_2'(x)f_{yy}(X)) \qquad \text{(C-10)}$$

From (C-5), we have

$$I_1 = (a_1(x)+A_{12}(x)y)'f_x(X) + \frac{1}{\mu}(a_{21}(x)+A_2y)'f_y(X) \qquad \text{(C-11)}$$

and therefore, by adding the expressions in (C-7)-(C-11), we get: that the operator A, corresponding to X^0, which is defined by (C-5) is

$$Af(X) = (\bar{a}_1(x) + A_{12}(x)y)f_x(X) + (\bar{a}_{21}(x) + A_2y)f_y(X)$$

$$+ \frac{1}{2}\, tr[B_1(x)S(0)B_1'(x)f_{xx}(x) + 2B_1(x)S(0)\tilde{B}_2'(x)f_{xy}(x)$$

$$+ \tilde{B}_2(x)S(0)\tilde{B}_2'(x)f_{yy}(x)] \qquad \text{(C-12)}$$

where $\bar{a}_1(x)$ and $\bar{a}_{21}(x)$ are defined through (2.45)-(2.48). We used the following identity:

$$tr[B_1(x)W'\tilde{B}_2'(x)f_{yx}'(X) + \tilde{B}_2(x)W'B_1'(x)f_{xy}'(X)]$$

$$= tr[B_1(x)S(0)\tilde{B}_2'(x)f_{xy}(X)$$

Notice that:

$$W + W' = S(0) \qquad \text{(C-13)}$$

It follows from (C-12) that X^0 satisfies the Itô - differential equations

$$\begin{aligned} dx &= (\bar{a}_1(x) + A_{12}(x)y)dt + B_1(x)\sqrt{S(0)}\; dw \\ \mu dy &= (\bar{a}_{21}(x) + A_{2y})dt + B_2(x)\sqrt{S(0)}\; dw \end{aligned} \qquad \text{(C-14)}$$

(C-14) has been introduced in (2.43) and (2.44).

Now we would like to apply the result of [33] to the singularly perturbed system (C-14). To do that, we need to modify (C-14) to be expressed in the same form of the system considered by [33], see equation (2.1.1) in [33].

So, let us set $\mu = \nu^2$ and $\tilde{y} = \nu y$ then (C-14) becomes (for simplifying the notation, we use y again instead of $\tilde{y}$ but it will be understood that we are working with $\tilde{y}$)

$$dx = \bar{a}_1(x)dt + \frac{1}{\nu} A_{12}(x)ydt + B_1(x)\sqrt{S(0)}\, dw \tag{C-15}$$

$$dy = \frac{1}{\nu}\bar{a}_{21}(x)dt + \frac{1}{\nu^2} A_2 ydt + \frac{1}{\nu} B_2(x)\sqrt{S(0)}\, dw \tag{C-16}$$

In [33], Theorem 2.1 says roughly that under certain assumptions the slow states represented by $x^{\nu}(t)$, and defined by (2.1.1) there, converges weakly as $\nu \to 0$ to the diffusion Markov process $x(t)$ generated by $\bar{\mathcal{L}}$ of (2.2.5).

Our goal here is to use the above result to derive the operator L corresponding to the limiting diffusion x^0 of the reduced-order model corresponding to (C-15) and (C-16), when $\mu(=\nu^2) \to 0$. Then we compare the form of L as given by (2.49)-(2.53) with the formula of L^{γ}, given by (2.7), and conclude that L can be obtained from L^{γ} by letting $\gamma \to 0$ (or $\frac{\varepsilon}{\mu} \to 0$). So we proceed by writting down the expressions for the operators $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L}_3$ defined by (2.1.3)-(2.1.6) in [33].

$$\mathcal{L}_1 = (A_2 y)' \cdot \nabla_y + \frac{1}{2}\, \mathrm{tr}(B_2(x)S(0)B_2'(x) \cdot \nabla_{yy}) \tag{C-17}$$

$$\mathcal{L}_2 = (A_{12}(x)y)' \cdot \nabla_x + \bar{a}_{21}'(x) \cdot \nabla_y + \mathrm{tr}(B_1(x)S(0)B_2'(x)\nabla_{xy}) \tag{C-18}$$

$$\mathcal{L}_3 = (\bar{a}_1(x))' \cdot \nabla_x + \frac{1}{2}\, \mathrm{tr}(B_1(x)S(0)B_1'(x) \cdot \nabla_{xx}) \tag{C-19}$$

where $\nabla_x = (\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n})'$, $\nabla_y = (\frac{\partial}{\partial y_1}, \cdots, \frac{\partial}{\partial y_m})'$,

$\nabla_{xx} = \nabla_x(\nabla_x) = (\frac{\partial^2}{\partial x_i \partial x_j})_{n\times m}$, etc.

For any smooth function $f(x)$, $x \in R^n$, which has compact support, the function $\psi_f^{(i)}(x,y)$ has been introduced in [33] and it has been required that $\psi_f^{(1)}$ must satisfy:

$$\mathcal{L}_1 \psi_f^{(1)}(x,y) + \mathcal{L}_2 f(x) = 0 \tag{C-20}$$

which can be written, via (C-17) and (C-18), as

$$(A_2 y)' \cdot \nabla_y \psi_f^{(1)}(x,y) + \frac{1}{2}\, \mathrm{tr}(B_2(x)S(0)B_2'(x)\nabla_{yy}\psi_f^{(1)}(x,y))$$
$$+ (A_{12}(x)y)' f_x(x) = 0 \tag{C-21}$$

By means of the linearity in y and the fact that the constant matrix A_2 is nonsingular, it can be seen that $\psi_f^1(x,y)$ has to be linear in y. So we suggest the following form for ψ_f^1

$$\psi_f^{(1)}(x,y) = g'(x)y \tag{C-22}$$

where $g(x)$ is a vector function of x, to be determined later. Substituting from (C-22) into (C-21) we get:

$$(A_2 y)'g(x) + (A_{12}(x)y)'f_x(x) = 0$$

$$y'(A_2' g(x) + A_{12}'(x)f_x(x)) = 0 \tag{C-23}$$

(C-23) has to be true $\forall y \in R^m$, this implies that

$$g(x) = -(A_2')^{-1}A_{12}'(x)f_x(x)$$

Then, from (C-22) we have:

$$\psi_f^{(1)}(x,y) = -f_x'(x)A_{12}(x)A_2^{-1}y \tag{C-24}$$

which is defined up to an additive constant. We proceed as in [33], by defining $Y(t;x)$ to be the diffusion process in R^m generated by $\mathcal{L}_1$ given by (C-17). This process is actually a Gaussian Brownian motion process which possess an invariant measure given by:

$$\bar{P}(A;x) = \int_A \frac{1}{((2\pi)^m \det Q)^{\frac{1}{2}}} e^{-\frac{1}{2}(y'Q^{-1}y)} dy \quad \forall A \subset R^m \tag{C-25}$$

Q is the variance matrix of $Y(t;x)$ and it is dependent on the parameter $x \in R^n$. Q is in fact the solution $\overline{P}$ of the Lyapunov equation (2.53). We notice that

$$\int_{R^m} \overline{P}(dy;x) = 1 \tag{C-26}$$

As in [33] the diffusion operator L (see (2.2.5), (2.2.6) in [33]), is defined by:

$$\begin{aligned} Lf(x) = \int_{R^m} & \overline{P}(dy;x)[\overline{a}_{21}'(x) \cdot \nabla_y \psi_f^{(1)}(x,y) \\ & + tr(B_1(x)S(0)B_2'(x)\nabla_{xy}\psi_f^{(1)}(x,y)) \\ & + (A_{12}(x)y)' \cdot \nabla_x \psi_f^{(1)}(x,y) + \overline{a}_1(x) \cdot f_x(x) \\ & + \frac{1}{2} tr(B_1(x)S(0)B_1'(x) \cdot f_{xx}(x))] \end{aligned} \tag{C-27}$$

Substituting from (C-24) into (C-27) and making use of (C-26), then, the first term of (C-27) is

$$\begin{aligned} & = - \int_{R^m} \overline{P}(dy;x)\overline{a}_{21}'(x)\nabla_y(f_x'(x)A_{12}(x)A_2^{-1}y) \\ & = - \int_{R^m} \overline{P}(dy;x)\overline{a}_{21}'(x)(f_x'(x)A_{12}(x)A_2^{-1})' \\ & = - f_x'(x)A_{12}(x)A_2^{-1}\overline{a}_{21}(x) \end{aligned} \tag{C-28}$$

It can be shown after simple manipulations that the second term of (C-27) is:

$$= -\int_{R^m} \overline{P}(dy;x)\,tr(B_1(x)S(O)B_2'(x)$$

$$\cdot\ \nabla_{xy}(f_x'(x)A_{12}(x)A_2^{-1}y))$$

$$= -\ trf_{xx}(x)B_1(x)S(O)B_2'(x)\,(A_2^{-1})'A_{12}(x)$$

$$-\sum_{i=1}^{n} tr\ \frac{\partial f}{\partial x_i}(x)B(x)S(O)B_2'(x)\,(A_2^{-1})'F_i'(x) \qquad \text{(C-29)}$$

The third term:

$$\int_{R^m} \overline{P}(dy;x)\,(-y'A_{12}'(x)f_{xx}(x)A_{12}(x)A_2^{-1}y$$

$$-\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)y'A_{12}'(x)F_i(x)A_2^{-1}y$$

$$= tr(-Q(A_2^{-1})'A_{12}'(x)f_{xx}(x)A_{12}(x)$$

$$-\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)Q(A_2^{-1})Q(A_2^{-1})'F_i'(x)A_{12}(x))$$

$$= -\frac{1}{2}\ tr((Q(A_2^{-1})' + A_2^{-1}Q)A_{12}'(x)f_{xx}(x)A_{12}(x))$$

$$-\ tr\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)Q(A_2^{-1})'F_i'(x)A_{12}(x)$$

$$= -\frac{1}{2}\ tr\ A_2^{-1}B_2(x)S(O)B_2'(x)\,(A_2^{-1})'A_{12}'(x)f_{xx}(x)A_{12}(x)$$

$$-\ tr\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)Q(A_2^{-1})'F_i'(x)A_{12}(x) \qquad \text{(C-30)}$$

We used the fact that Q satisfies the Lyapunov equation (2.53).

The last two terms on the right-hand sides of (C-27) are given by

$$\bar{a}_1'(x)f_x(x) + \frac{1}{2}\,\mathrm{tr}(B_1(x)S(O)B_1'(x)f_{xx}(x)) \tag{C-31}$$

So, from (C-27)-(C-31) and by replacing Q by $\bar{P}$, we get

$$\begin{aligned} Lf(x) = {} & (\bar{a}_1(x) - A_{12}(x)A_2^{-1}\bar{a}_{21}(x) + \bar{h}_3(x))'f_x(x) \\ & + \frac{1}{2}\,\mathrm{tr}(B_O(x)S(O)B_O'(x)f_{xx}(x)) \end{aligned} \tag{C-32}$$

where $\bar{a}_1$, $\bar{a}_{21}$ and $\bar{h}_3(x)$ are given by (2.45), (2.46) and (2.52) respectively. The form of L in (C-32) is exactly the same as if we let ε/μ (or $\gamma \to 0$) in (2.7).

CHAPTER III

STABILITY

3.1. Introduction:

Let us consider again the singularly perturbed system that has been studied in Chapter two, namely:

$$\dot{x} = a_1(x) + A_{12}(x)y + B_1(x)v^{\varepsilon}, \quad x(0) = x_0 \tag{3.1}$$

$$\mu\dot{y} = a_{21}(x) + A_2 y + B_2(x)v^{\varepsilon}, \quad y(0) = y_0 \tag{3.2}$$

which defines $(x(t),y(t))$ under the hypotheses of Chapter II. As we have shown, $x(t)$ converges weakly to the diffusion process $\bar{x}(t)$ generated by L^{γ} of (2.1)-(2.17). We shall study properties of $x(t)$ as $t \to \infty$ with $0 < \varepsilon \leq \varepsilon_0$, $0 < \mu \leq \mu_0$ and $|\frac{\varepsilon}{\mu} - \gamma| \leq \gamma_0$; ε,μ are fixed and ε_0,μ_0 and γ_0 are sufficiently small. Our objective is to establish stability results about $x(t)$ which are based only on conditions upon the approximating diffusion $\bar{x}(t)$. So stability analysis can be performed as if $x(t)$ was given by the reduced order model. This will lead to a considerable simplification due to the smaller size of the model as well as the fact that the reduced-order model is a diffusion one. To justify this approximation, analytical study has to be performed to guarantee that the behaviour of the actual $x(t)$ as defined

by (3.1) and (3.2) can be asymptotically predicted by the reduced order model. Similar studies have been carried out in [12] for systems driven by wide-band noise, as we pointed out in Chapter I. Our job is essentially to extend the methods of [12] to the singularly perturbed case. In section 3.2 we state and prove theorem (1) which gives us sufficient conditions that guarantee stochastic asymptotic stability of the origin $x = 0$ of (3.1) if the origin $\bar{x} = 0$ of the reduced-order model is so. To show that, we will proceed in a way similar to [12] except that the averaging of the Lyapunov function is done in a way that is similar to the averaging of $f(x)$ in the proof of the convergence theorem of Chapter II. Since the operator $A^{\epsilon,\mu}$ cannot be applied to unbounded functions, the Lyapunov function $V(x)$, whose existence is required for the stability of the reduced-order model, has to be artifically bounded using truncations. Recently, in [31], truncations have been employed for a similar purpose. In Section 3-3, we allow the coefficients of a_1, A_{12} and B_1 not to vanish at $x = 0$, which is an essential requirement in Theorem 1, and prove theorem 2 which shows that the mean square of $x(t)$ is bounded on the entire time interval. The basic steps of the proof of Theorem 1 are used again in the proof of Theorem 2, and only differences between the two proofs will be emphasized. In Section 3.4 illustrative examples are explored.

3.2. Stochastic Asymptotic Stability:

Before we state the basic theorem of this section we state all the assumptions needed:

(1) The process $v^{\epsilon}(t)$ is defined and is assumed to satisfy the same conditions exactly as in Chapter II.

(2) The coefficients a_1, A_{12}, B_1, a_{21} and B_2 are continuous in x and have continuous partial derivatives up to the second order which are uniformly bounded in $x \in R^n$. Moreover a_{21} and B_2 are required to be bounded uniformly in x.

(3) A_2 satisfies condition (A3).

(4) The coefficients $a_1(x), A_{12}(x)$ and $B_1(x)$ vanish at $x = 0$ and for every $x \in R^n$ and for some $M > 0$,

$$|a_1(x)| + |A_{12}(x)| + |B_1(x)| \leq M|x| \quad .$$

(5) The coefficients $a_0(x)$ and $B_0(x)$ which are defined by (2.5) and (2.6) are required to satisfy:

$$|a_0(x) - a_0(z)| + |B_0(x) - B_0(z)| \leq K|x - z| \quad \forall x, z \in R^n$$

and for some $K > 0$.

Now we may consider the diffusion process $\bar{x}(t)$, generated by L^{γ} and given by (2.7)-(2.17), to be the solution of the Ito equation:

$$d\bar{x}(t) = b(\bar{x}(t)) + \sigma(\bar{x}(t))dw(t), \ \bar{x}(0) = x_0 \tag{3.3}$$

where $b(x)$ is defined by (2.8) and

$$\sigma(x) = B_0(x)\sqrt{S(0)} \tag{3.4}$$

and we require:

(6) The coefficients $b(x)$ and $\sigma(x)$ satisfy

$$|b(x)| + |\sigma(x)| \leq M|x| \quad \forall x \in R^n, \text{ for some } M > 0$$

and

$$|b(x) - b(z)| + |\sigma(x) - \sigma(z)| \leq K|x - z| \quad \forall x,z \in R^n$$

for some $K > 0$.

We will consider functions $V(x)$, $x \in R^n$ with the following properties:

(a) $V(x)$ is real-valued, positive definite, $V(x) = 0 \Rightarrow x = 0$, $V(x) \to \infty$ as $|x| \to \infty$ and has continuous partial derivatives up to the third order.

(b) For any vector or matrix valued function $g(x,t) = O(x)$ for $t \in [0,T]$ we have:

$$|V_x'(x)g(x,t)| \leq KV(x), \qquad \forall x \in R^n \tag{3.5}$$

$$|(V_x'(x)g(x,t))_x g(x,t)| \leq KV(x), \qquad \forall x \in R^n \tag{3.6}$$

$$\left|\frac{\partial^3 V(x)}{\partial x_i \partial x_j \partial x_k} g_i^1(x,t) g_j^2(x,t) g_k^3(x,t)\right| \leq KV(x) \qquad \forall x \in R^n \tag{3.7}$$

$\forall i,j,k = 1,2,\cdots,n$ where g_i^1, g_j^2 and g_k^3 are components of vectors or matrices which are $O(x)$. The constant K in (3.5-3.7) may not be the same

and it is independent of T. For example if $g(x) = x$ and $V(x) = x^2$ or $V(x) = x^4$ then it is obvious that (3.5)-(3.7) are satisfied.

Truncated Lyapunov function: Let $V(x)$ be a Lyapunov function which satisfies (a) and (b), stated above. For each positive integer N, define $S_N = \{x : |x| \leq N\}$.

Define $V_N(x) = V(x)q_N(x)$ $\forall x \in R^n$, where $q_N(x) = 1$ in S_N, $q_N(x) = 0$ in $R^n - S_{N+1}$ and $q_N(x) \in [0,1]$ is smooth in x and have partial derivatives up to the third order which are bounded uniformly in x,N.

In the following theorem we write L instead of L^γ, where it will be understood that L depends on the prespecified number $\gamma \in [\gamma_1,\infty)$, $\gamma_1 > 0$ is arbitrary but fixed. Moreover L is defined by:

$$L(\cdot) = \sum_{i=1}^{n} b_i(x) \frac{\partial}{\partial x_i}(\cdot) + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j}(\cdot) \tag{3.8}$$

where

$$A(x) = B_0(x)S(0)B_0'(x) \stackrel{\Delta}{=} [a_{ij}(x)], \tag{3.9}$$

The vector $b(x)$ is defined by (2.8)-(2.17).

Now we state the theorem:

Theorem 1: Suppose that there exists a Lyapunov function $V(x)$ on R^n satisfying (a),(b) and, for some $\lambda > 0$,

$$LV(x) \leq -\lambda V(x) \qquad \forall x \in R^n \tag{3.10}$$

Suppose that all the assumptions (1)-(6) are satisfied. Then, there exist ε_0, μ_0 and γ_0 such that for all ε, μ satisfying $0 < \varepsilon \leq \varepsilon_0$, $0 < \mu \leq \mu_0$ and $|\frac{\varepsilon}{\mu} - \gamma| \leq \gamma_0$, $x(t)$, the solution of (3.1) and (3.2), is uniformly stochastically asymptotically stable as $t \to \infty$, i.e. for any $\eta_1 > 0$ and $\eta_2 > 0$ there is a $\delta > 0$ such that if $|x_0| < \delta$ then:

(I) $P\{|x(t)| \leq \eta_2 e^{-\theta t},\ t \geq 0\} \geq 1 - \eta_1$ for some $\theta > 0$.

(II) $P\{\lim\limits_{t \to \infty} |x(t)| = 0\} = 1$.

Moreover, if $V(x)$ satisfies, in addition,

$$c_2 |x|^{n_2} \leq V(x) \leq c_1 |x|^{n_1} \qquad \forall x \in R^n$$

for some positive constants c_1 and c_2 and some positive integers n_1 and n_2 then (I) and (II) will be satisfied in the large, i.e. independent of the initial condition x_0.

Remark: The condition (3.10), under the assumptions stated on the coefficients of L, guarantees that the limiting diffusion process $\bar{x}(t)$ is uniformly stochastically asymptotically stable as $t \to \infty$, see [7] or [20].

Proof of Theorem 1: In the proof we adopt the same terminology and definitions, concerning the operator $A^{\varepsilon,\mu}$, which has been used in Chapter II. In fact we are going to

repeat the averaging method but with the truncated function $V_N(x)$ instead of $f(x)$. We follow the basic idea of the proof of [12 Ch. 5].

Operating, now on $V_N(x)$ by $A^{\varepsilon,\mu}$, (using from now on x for $x(t)$, for $0 \leq t \leq T$ where T is arbitrary), we get

$$A^{\varepsilon,\mu}V_N(x) = \frac{\partial V_N}{\partial x}(x)\,[a_1(x) + A_{12}(x)y(t) + B_1(x)v^{\varepsilon}(t)] \tag{3.11}$$

Averaging out the last two terms by defining

$$\begin{aligned} V_{N,1}(x,t) = \frac{\partial V_N}{\partial x}(x) \int_0^{\infty} [&A_{12}(x)(E_t\hat{y}(t+s,x) + A_2^{-1}a_{21}(x)) \\ &+ B_1(x)E_t v^{\varepsilon}(t+s)]ds \end{aligned} \tag{3.12}$$

where

$$\begin{aligned} \hat{y}(t+s,x) = {}& e^{A_2 s/\mu}(y(t) + A_2^{-1}a_{21}(x)) - A_2^{-1}a_{21}(x) \\ &+ \frac{1}{\mu}\int_t^{t+s} e^{A_2(t+s-\tau)/\mu} B_2(x)v^{\varepsilon}(\tau)d\tau \end{aligned} \tag{3.13}$$

For $x \in S_N$ $\quad (\Rightarrow \frac{\partial V_N}{\partial x}(x) = \frac{\partial V}{\partial x}(x))$, we have

$$\begin{aligned} |V_{N,1}(x,t)| \leq \Big|\frac{\partial V}{\partial x}(x) \int_0^{\infty} [&A_{12}(x)(E_t\hat{y}(t+s,x) + A_2^{-1}a_{21}(x)) \\ &+ B_1(x)E_t v^{\varepsilon}(t+s)]ds\Big|, \end{aligned}$$

but from (1)-(4) and (3.13) the integral is bounded by $(K_1\sqrt{\varepsilon} + K_2\sqrt{\mu})|x|$. So it follows from condition (b) that:

$$|V_{N,1}(x,t)| \leq (K_1\sqrt{\epsilon} + K_2\sqrt{\mu})V(x) \quad \text{for} \quad x \in S_N \tag{3.14}$$

where K_1 and K_2 are independent of N and T. Operating on $V_{N,1}(x,t)$ by $A^{\epsilon,\mu}$ we get

$$A^{\epsilon,\mu}V_{N,1}(x,t) = -\frac{\partial V_N}{\partial x}(x)[A_{12}(x)(y(t) + A_2^{-1}a_{21}(x)) + B_1(x)v^{\epsilon}(t)] + \frac{\partial V_{N,1}}{\partial x}(x,t)[a_1(x) + A_{12}(x)y(t) + B_1(x)v^{\epsilon}(t)] \tag{3.15}$$

Adding (3.11) and (3.15) yields

$$A^{\epsilon,\mu}(V_N(x) + V_{N,1}(x,t)) = \frac{\partial V_N}{\partial x}(x)a_0(x) + \frac{\partial V_{N,1}}{\partial x}(x,t)[a_1(x) + A_{12}(x)y(t) + B_1(x)v^{\epsilon}(t)] \tag{3.16}$$

We average out the last two terms of (3.16) by defining:

$$V_{N,2}(x,t) = \int_0^{\infty} [E_t \frac{\partial V_{N,1}}{\partial x}(x,t+s)(A_{12}(x)(\hat{y}(t+s,x) + B_1(x)v^{\epsilon}(t+s)) + \frac{\partial V_N}{\partial x}(x)a_0(x) - L^{\epsilon/\mu}V_N(x)]ds \tag{3.17}$$

Following the steps of the convergence proof of Chapter 2, in which we have shown that $|f_2(x,t)| = O(\epsilon+\mu)$, and by replacing $f_1(t,x)$ by $V_{N,1}$ and $f(x)$ by $V_N(x)$ we see that each term appearing in $V_{N,2}(x,t)$ is identified, using the assumptions (1)-(4) and (3.13), with the left

hand sides of either (3.5) or (3.6). Thus, those inequalities imply that

$$|V_{N,2}(x,t)| \leq (K_3\varepsilon + K_4\mu)V(x) \quad \text{for} \quad x \in S_N \tag{3.18}$$

where K_3 and K_4 are independent of N and T.

$$\begin{aligned} A^{\varepsilon,\mu}V_{N,2}(x,t) &= L^{\varepsilon/\mu}V_N(x) - \frac{\partial V_{N,1}}{\partial x}(x,t)(A_{12}(x)(y(t) + A_2^{-1}a_{21}(x)) \\ &\quad + B_1(x)v^{\varepsilon}(t)) - \frac{\partial V_N}{\partial x}(x)a_0(x) + \frac{\partial V_{N,2}}{\partial x}(a_1(x) + A_{12}(x)y(t) \\ &\quad + B_1(x)v^{\varepsilon}(t)) \end{aligned} \tag{3.19}$$

Adding (3.16) to (3.19) we get

$$\begin{aligned} A^{\varepsilon,\mu}V_N(x,t) &= L^{\varepsilon/\mu}V_N(x) + \frac{\partial V_{N,1}}{\partial x}(x,t)a_0(x) \\ &\quad + \frac{\partial V_{N,2}}{\partial x}(x,t)(a_1(x) + A_{12}(x)y(t) + B_1(x)v^{\varepsilon}(t)) \end{aligned} \tag{3.20}$$

where

$$V_N(x,t) = V_N(x) + V_{N,1}(x,t) + V_{N,2}(x,t) \tag{3.21}$$

Now from (3.20) and (3.21) and for any $\hat{\lambda} > 0$, to be determined later, we have:

$$\begin{aligned} (A^{\varepsilon,\mu} + \hat{\lambda})V_N(x,t) &= L^{\varepsilon/\mu}V_N(x) + \hat{\lambda}V_N(x,t) + \frac{\partial V_{N,1}}{\partial x}(x,t)a_0(x) \\ &\quad + \frac{\partial V_{N,2}}{\partial x}(x,t)(a_1(x) + A_{12}(x)y(t) + B_1(x)v^{\varepsilon}(t)) \end{aligned} \tag{3.22}$$

Our goal at this point is to choose $\hat{\lambda}$ appropriately such that $(A^{\varepsilon,\mu} + \hat{\lambda})V_N(x,t) \leq 0$. From the definition of $L^{\varepsilon/\mu}$, (1)-(4), (3.5), (3.6) and the smooth dependence of $L^{\varepsilon/\mu}$ on $\varepsilon/\mu \in [\gamma_1,\infty)$, there exists a constant $c > 0$ such that for $x \in S_N$

$$|L^{\varepsilon/\mu}V_N(x) - LV_N(x)| \leq c|\frac{\varepsilon}{\mu} - \gamma|V(x) \tag{3.23}$$

From (3.14) and (3.18) we have for $x \in S_N$:

$$V_N(x,t) \geq (1 - K_1\sqrt{\varepsilon} - K_2\sqrt{\mu} - K_3\varepsilon - K_4\mu)V(x) \tag{3.24}$$

Similarly we have for $x \in S_N$:

$$V_N(x,t) \leq (1 + K_1\sqrt{\varepsilon} + K_2\sqrt{\mu} + K_3\varepsilon + K_4\mu)V(x) \tag{3.25}$$

Now we want to find upper bounds for the last two terms in (3.22) similar to the upper bounds in (3.14) and (3.18), and this follows from the definitions of $V_{N,1}(x,t)$, $V_{N,2}(x,t)$, from the assumptions (1)-(4) and from (3.5)-(3.7). So we have:

$$|\frac{\partial V_{N,1}}{\partial x}(x,t)a_0(x)| \leq (K_5\sqrt{\varepsilon} + K_6\sqrt{\mu})V(x) \qquad \forall x \in S_N \tag{3.26}$$

$$|\frac{\partial V_{N,2}}{\partial x}(x,t)a_1(x)| + |\frac{\partial V_{N,2}}{\partial x}(x,t)(A_{12}(x)y(t) + B_1(x)v^{\varepsilon}(t))|$$

$$\leq (K_7\varepsilon + K_8\mu + K_9\sqrt{\varepsilon} + K_{10}\sqrt{\mu})V(x) \qquad \forall x \in S_N \tag{3.27}$$

All the K_i in (3.26) and (3.27) are independent of N and T. Now it follows from (3.22),(3.23),(3.26),(3.27), (3.25) and (3.10) that for $x \in S_N$ we have:

$$(A^{\varepsilon,\mu} + \hat{\lambda})V_N(x,t) \leq [-\lambda + c|\tfrac{\varepsilon}{\mu} - \gamma| + \bar{K}_1\sqrt{\varepsilon} + \bar{K}_2\sqrt{\mu} + \bar{K}_3\varepsilon + \bar{K}_4\mu$$
$$+ \hat{\lambda}(1 + K_1\sqrt{\varepsilon} + K_2\sqrt{\mu} + K_3\varepsilon + K_4\mu)]V(x) \qquad (3.28)$$

There exist $\varepsilon_0 > 0$, $\mu_0 > 0$ and $\gamma_0 > 0$ sufficiently small such that the following conditions are satisfied:

(i) $1 - K_1\sqrt{\varepsilon_0} - K_2\sqrt{\mu_0} - K_3\varepsilon_0 - K_4\mu_0 = c_1 > 0$

(ii) For all ε and μ satisfying $0 < \varepsilon \leq \varepsilon_0$, $0 < \mu \leq \mu_0$ and $|\frac{\varepsilon}{\mu} - \gamma| \leq \gamma_0$, $\hat{\lambda}$ can be taken small enough such that:

$$-\lambda + c|\tfrac{\varepsilon}{\mu} - \gamma| + \bar{K}_1\sqrt{\varepsilon} + \bar{K}_2\sqrt{\mu} + \bar{K}_3\varepsilon + \bar{K}_4\mu$$
$$+ \hat{\lambda}(1 + K_1\sqrt{\varepsilon} + K_2\sqrt{\mu} + K_3\varepsilon + K_4\mu) \leq 0. \qquad (3.29)$$

Then by this choice of ε_0, μ_0 and γ_0, (3.28) reduces to:

$$(A^{\varepsilon,\mu} + \hat{\lambda})V_N(x,t) \leq 0 \quad \text{for} \quad x \in S_N \qquad (3.30)$$

For $0 \leq t \leq T$ and for each N, the function

$$M_N(t) = V_N(x(t),t) - V_N(x(0),0) - \int_0^t A^{\varepsilon,\mu}V_N(x(s),s)\,ds \qquad (3.31)$$

is a zero mean martingale even if t is replaced by a stopping time τ with $E(\tau) < \infty$, [c.f. 10].

Let us now redefine M_N as in (3.31), but with V_N replaced by $e^{\hat{\lambda} t} V_N$. We have:

$$M_N(t) = V_N(x(t),t)e^{\hat{\lambda} t} - V_N(x(o),0) - \int_0^t A^{\epsilon,\mu} V_N(x(s),s)e^{\hat{\lambda} s} ds \tag{3.32}$$

But from the definition of A, we have:

$$A^{\epsilon,\mu} V_N(x(s),s)e^{\hat{\lambda} s}$$

$$= \underset{\delta \to 0}{\text{p-lim}}\ E_t \frac{(e^{\hat{\lambda}(s+\delta)} V_N(x(s+\delta),s+\delta) - e^{\hat{\lambda} s} V_N(x(s),s))}{\delta}$$

$$= \underset{\delta \to 0}{\text{p-lim}}\ \frac{1}{\delta}[e^{\hat{\lambda}(s+\delta)} - e^{\hat{\lambda} s}]E_t V_N(x(s+\delta),s+\delta)$$

$$+ e^{\hat{\lambda} s}\ \underset{\delta \to 0}{\text{p-lim}}\ \frac{1}{\delta}[E_t V_N(x(s+\delta),s+\delta) - V_N(x(s),s)]$$

$$= \frac{d}{ds} e^{\hat{\lambda} s} V_N(x(s),s) + e^{\hat{\lambda} s} A^{\epsilon,\mu} V_N(x(s),s)$$

$$= e^{\hat{\lambda} s}(A^{\epsilon,\mu} + \hat{\lambda})V_N(x(s),s)$$

Then, we replace t by $t \cap \tau_N = \min(t,\tau_N)$ in (3.32) where $\tau_N = \inf\{t : x(t) \notin S_N\}$. It is obvious that $E(t \cap \tau_N) < \infty$. Then (3.32) becomes:

$$M_N(t \cap \tau_N) = V_N(x(t \cap \tau_N), t \cap \tau_N) e^{\hat{\lambda}(t \cap \tau_N)} - V_N(x(0), 0)$$
$$- \int_0^{t \cap \tau_N} e^{\hat{\lambda} s}(A^{\varepsilon,\mu} + \hat{\lambda}) V_N(x(s), s)\, ds$$

from which we have:

$$e^{\hat{\lambda}(\tau_N \cap t)} V_N(x(t \cap \tau_N), t \cap \tau_N) = V_N(x(0), 0)$$
$$+ \int_0^{t \cap \tau_N} e^{s\hat{\lambda}}(A^{\varepsilon,\mu} + \hat{\lambda}) V_N(x(s), s)\, ds + M_N(t \cap \tau_N) \tag{3.33}$$

Since $x(s) \in S_N$ for $0 \leq s \leq t \cap \tau_N$, (3.30) and (3.32) give

$$e^{\hat{\lambda}(\tau_N \cap t)} V_N(x(t \cap \tau_N), t \cap \tau_N) \leq V_N(x(0), 0) + M_N(t \cap \tau_N) \tag{3.34}$$

Let $c_2 = 1 + K_1\sqrt{\varepsilon_0} + K_2\sqrt{\mu_0} + K_3\varepsilon_0 + K_4\mu_0$ then from (3.24), condition (i), and (3.25) we have

$$c_1 V(x) \leq V_N(x,t) \leq c_2 V(x) \quad \text{for} \quad x \in S_N \tag{3.35}$$

So (3.34) and (3.35) imply:

$$0 \leq c_1 e^{\hat{\lambda}(\tau_N \cap t)} V(x(t \cap \tau_N)) \leq c_2 V(x(0)) + M_N(t \cap \tau_N) \tag{3.36}$$

The right hand side of (3.36) is a nonnegative integrable martingale. Using Kolomogrov's inequality for nonnegative martingales and that $E(M_N(t \cap \tau_N)) = 0$, (3.36) gives

$$P\{c_1 e^{\hat{\lambda}(\tau_N \cap t)} V(x(t \cap \tau_N)) > \eta, \quad 0 \le t \le T\}$$

$$\le P\{c_2 V(x(0)) + M_N(t \cap \tau_N) > \eta, \quad 0 \le t \le T\}$$

$$\le c_2 E(V(x(0)))/\eta \tag{3.37}$$

Letting $N \to \infty$ implies that $\tau_N \to \infty$, since by the linear growth assumptions on the coefficients there is no finite escape time. Then from (3.37) we have:

$$P\{c_1 e^{\hat{\lambda} t} V(x(t)) > \eta, \quad 0 \le t \le T\} \le c_2 E(V(x_0))/\eta \tag{3.38}$$

Letting $T \to \infty$ we get:

$$P\{c_1 e^{\hat{\lambda} t} V(x(t)) > \eta, \; t \ge 0\} \le c_2 E(V(x_0))/\eta \tag{3.39}$$

By the smoothness assumptions on $V(x)$ we have:

$$\bar{c}_1 |x|^{n_1} \le V(x) \le \bar{c}_2 |x|^{n_2} \quad \text{for} \quad |x| \le r_0 \tag{3.40}$$

for some $r_0 > 0$, $\bar{c}_1 > 0$, $\bar{c}_2 > 0$ and some positive integers n_1 and n_2. Then

$$\{\bar{c}_1 |x(t)|^{n_1} \le \frac{-e^{-\hat{\lambda} t}}{c_1} \eta, \quad t \ge 0\} \supset \{V(x(t)) \le \frac{e^{-\hat{\lambda} t}}{c_1} \eta, \quad t \ge 0\}$$

Hence by (3.39) and (3.40) we have:

$$P\{|x(t)| \le e^{-\hat{\lambda}t/n_1}\left(\frac{\eta}{c_1\bar{c}_1}\right)^{1/n_1}, \quad t \ge 0\} \ge 1 - c_2 E(V(x_0))/\eta$$

$$\ge 1 - \frac{c_2\bar{c}_2}{\eta} E(|x_0|^{n_2}) \tag{3.41}$$

For any η_1 and η_2, choose η so small that (3.41) gives

$$P\{|x(t)| \le e^{-\theta t}\eta_2, \quad t \ge 0\} \ge 1 - \frac{c_2\bar{c}_2}{\eta} E(|x_0|^{n_2}) \tag{3.42}$$

where $\theta = \frac{\hat{\lambda}}{n_1}$. Choosing δ so small that $c_2\bar{c}_2\delta^{n_2}/\eta < \eta_1$, we get that for all x_0 with $|x_0| < \delta$

$$P\{|x(t)| \le e^{-\theta t}\eta_2, \quad t \ge 0\} \ge 1 - \eta_1 \tag{3.43}$$

and this proves (I).

Now since

$$\{\lim_{t\to\infty} |x(t)| = 0\} = \{\lim_{t\to\infty} V(x(t)) = 0\}$$

$$\supset \{\sup_{t\ge 0} e^{\hat{\lambda}t} c_1 V(x(t)) \le C\}$$

($V(x)$ being radially unbounded is necessary for the validity of the above statement) where C is any positive constant, (3.39), (3.40) imply that for $|x_0| < \delta$ we have

$$P\{\lim_{t\nearrow\infty} |x(t)| = 0\} \ge 1 - \frac{c_2\bar{c}_2\delta^{n_2}}{C}$$

Then as $C \to \infty$ we get

$$P\{\lim_{t \nearrow \infty} |x(t)| = 0\} = 1 \tag{3.44}$$

which proves II. Now if $V(x)$ satisfies (3.40) $\forall x \in R^n$ and if $|x_0| < \infty$, (3.43) and (3.44) follow in the large.

3.3. Mean Square Boundedness:

The stability result presented in section 3.2 was concerned with establishing the asymptotic stability of the origin in a stochastic sense where the origin $x = 0$ is an equilibrium point of the system for any driving input noise. That is, if the initial condition $x_0 = 0$ then $x(t) = 0$ for all $t \geq 0$. A key assumption there was the requirement that $a_1(x)$, $A_{12}(x)$ and $B_1(x)$ vanish at $x = 0$. While requiring $a_1(x)$ to vanish at $x = 0$ is a typical and acceptable assumption because it can be always achieved by shifting the origin to the equilibrium point of the unforced system, requiring $A_{12}(x)$ and $B_1(x)$ to vanish at $x = 0$ is not always valid. In many cases driving inputs do not vanish at $x = 0$ and one cannot discuss asymptotic stability of $x = 0$ because $x(t)$ does not necessarily tend to $x = 0$ as $t \to \infty$. For deterministic systems the appropriate concept of stability is bounded-input bounded-output stability, i.e., to establish that for any bounded input the trajectories of the system remain in a bounded set.

A stochastic version of this concept is the mean-square boundedness where one shows that $x(t)$ has a bounded mean square. The objective of this section is to study the stability of $x(t)$ when $A_{12}(x)$ and $B_1(x)$ do not necessarily vanish at $x = 0$. We now state the following alternative assumption.

(4$'$) The coefficients $a_1(x)$, $A_{12}(x)$ and $B_1(x)$ are required to satisfy, for every $x \in R^n$ and for some positive constants M_1 and M_2,

$$|a_1(x)| + |A_{12}(x)| + |B_1(x)| \leq M_1|x| + M_2.$$

The Lyapunov function $V(x)$ will be taken to be a quadratic form, namely $V(x) = x'Qx$, $Q > 0$. So it is obvious that $V(x)$ in this form satisfies conditions (a) and (b) including (3.5)-(3.7), which are stated in section 1. Now we are ready to state theorem 2.

<u>Theorem 2</u>: Suppose that there exists a positive definite $n \times n$ matrix Q such that $V(x) = x'Qx$ satisfies

$$LV(x) \leq K - \lambda V(x) \qquad \forall x \in R^n \tag{3.45}$$

for some $K \geq 0$ and $\lambda > 0$. Moreover assume that all the assumptions (1)-(6), with (4$'$) replacing (4), are satisfied. Then there exist ε_0, μ_0 and γ_0 such that for all ε, μ satisfying $0 < \varepsilon \leq \varepsilon_0$, $0 < \mu \leq \mu_0$ and $|\frac{\varepsilon}{\mu} - \gamma| \leq \gamma_0$, the process $x(t)$, defined by (3.1) and (3.2), is bounded in the mean square, i.e., there exists a positive constant $K > 0$ such that $E\{|x(t)|^2\} \leq K$, provided that $E\{|x_0|^2\} < \infty$.

Remark 2: The quadratic Lyapunov function $V(x)$ satisfies the following conditions:

$$E\{V(x(t))\}, \quad E\{|LV(x(t))|\} \quad \text{and} \quad E\{|\frac{\partial V}{\partial x_i}(x(t))\sigma_{ij}(x(t))|^2\}$$

are bounded in t in any bounded time interval, and that $V(x) \geq c|x|^2 \quad \forall x \in R^n$ and for some $c > 0$. Similar conditions to these and to (3.45) have been required [c.f. 18,31] to guarantee that, the solution to an Itô equation is exponentially bounded in mean square with some positive exponent, i.e., $E\{|x(t)|^2 \leq K_1 + K_2 e^{-\bar{\alpha}t}$ for some $K_1 \geq 0$, $K_2 > 0$ and $\bar{\alpha} > 0$, where $x(t)$ is the solution of an Itô equation. This is actually the case for $\bar{x}(t)$, the solution of our reduced-order model. The above conditions are valid because $V, |LV|$ and $|(\frac{\partial V}{\partial x_i})\sigma_{ij}|$ are dominated by polynomials and c is in fact equal to $\lambda_{min}(Q)$ which is positive since Q is positive definite.

Proof of theorem 2: Using the fact that $V(x)$ satisfies

$$c_1|x|^2 \leq V(x) \leq c_2|x|^2 \quad \forall x \in R^n \tag{3.46}$$

for some positive constants c_1 and c_2, and then preceeding in the same way as in the proof of theorem 1 of section 3.2, we get the following inequalities which are similar to the ones given by (3.14), (3.18), (3.23), (3.26) and (3.27) respectively. The new inequalities take the forms:

$$|V_{N,1}(x,t)| \leq (K_1\sqrt{\varepsilon} + K_2\sqrt{\mu})(|x|^2 + |x|), \tag{3.47}$$

$$|V_{N,2}(x,t)| \leq (K_3\varepsilon + K_4\mu)(|x|^2 + |x| + 1), \tag{3.48}$$

For some $c > 0$

$$|L^{\varepsilon/\mu}V_N(x) - LV_N(x)| \leq c\left|\frac{\varepsilon}{\mu} - \gamma\right|(|x|^2 + |x|), \tag{3.49}$$

$$\left|\frac{\partial V_{N,1}}{\partial x}(x,t)a_0(x)\right| \leq (K_5\sqrt{\varepsilon} + K_6\sqrt{\mu})(|x|^2 + |x|), \tag{3.50}$$

and finally

$$\left|\frac{\partial V_{N,2}}{\partial x}(x,t)a_1(x)\right| + \left|\frac{\partial V_{N,2}}{\partial x}(x,t)(A_{12}(x)y(t) + B_1(x)v^{\varepsilon}(t))\right|$$

$$\leq (K_7\varepsilon + K_8\mu + K_9\sqrt{\varepsilon} + K_{10}\sqrt{\mu})(|x|^2 + |x| + 1) \tag{3.51}$$

All the positive constants K_i in (3.47)-(3.51) are independent of N and T. Defining $V_N(x,t)$ as in (3.21), we have

$$A^{\varepsilon,\mu}V_N(x,t) = L^{\varepsilon/\mu}V_N(x) + \frac{\partial V_{N,1}}{\partial x}(x,t)a_0(x) + \frac{\partial V_{N,2}}{\partial x}(x,t)(a_1(x)$$

$$+ A_{12}(x)y(t) + B_1(x)v^{\varepsilon}(t)) \tag{3.52}$$

Then from (3.49), (3.45), (3.50) and (3.51), (3.52) implies that

$$A^{\varepsilon,\mu}V_N(x,t) \le K - \lambda c_1|x|^2 + c\left|\frac{\varepsilon}{\mu} - \gamma\right|(|x|^2 + |x|)$$

$$+ \ (K_7\varepsilon + K_8\mu + \bar{K}_9\sqrt{\varepsilon} + \bar{K}_{10}\sqrt{\mu})(|x|^2 + |x| + 1)$$

$$\le K + (\bar{c}_1\left|\frac{\varepsilon}{\mu} - \gamma\right| + c_1(\varepsilon,\mu) - \lambda c_1)|x|^2 + \bar{c}_2\left|\frac{\varepsilon}{\mu} - \gamma\right| + c_2(\varepsilon,\mu) \tag{3.54}$$

There exist $\varepsilon_0 > 0$, $\mu_0 > 0$ and $\gamma_0 > 0$, sufficiently small, such that for $0 < \varepsilon \le \varepsilon_0$, $0 < \mu \le \mu_0$ and $\left|\frac{\varepsilon}{\mu} - \gamma\right| < \gamma_0$ we have

$$\bar{c}_1\left|\frac{\varepsilon}{\mu} - \gamma\right| + c_1(\varepsilon,\mu) - \lambda c_1 = -\bar{K}_2 \quad \text{for some } \bar{K}_2 > 0.$$

Then it follows from (3.54) that

$$A^{\varepsilon,\mu}V_N(x,t) \le \bar{K}_1 - \bar{K}_2|x|^2 \tag{3.55}$$

for some positive constants $\bar{K}_1$ and $\bar{K}_2$. Also the above choice of ε_0, μ_0 and γ_0 can be made small enough that by the aid of (3.47), (3.48) and (3.46) we have

$$\alpha_1|x|^2 - \alpha_2 \le V_N(x,t) \le \beta_1|x|^2 + \beta_2 \qquad \forall x \in S_N \tag{3.56}$$

where α_i and β_i are some positive constants.

Now let us introduce a set Q_0 as follows:

$$Q_0 = \{x \in R^n : |x| < K_0\} \quad \text{where} \quad K_0 = \left(\frac{\bar{K}_1}{\bar{K}_2}\right)^{\frac{1}{2}}$$

We define the starting time τ_0 as follows:

$$\tau_0 = 0 \quad \text{if} \quad x(0) \notin Q_0$$

$$= \inf_{t \geq 0} \{t : |x(t)| = K\} \quad \text{if} \quad x(0) \in Q_0$$

and let τ_1 be defined as the first time that $x(t)$ enter Q_0 for $t \geq \tau_0$, i..e., $\tau_1 = \inf\{t : t \geq \tau_0, x(t) \in Q_0\}$. Finally we define τ_N as before (see section 3.2), that is, $\tau_N = \inf\{t : x(t) \notin S_N\}$. Without loss of generality we assume that $S_N \supset Q_0$ so that $\tau_N \geq \tau_0$. Then similar to (3.31), and for $0 \leq t \leq T$, $t \geq \tau_0$

$$V_N(x(t \cap \tau_N \cap \tau_1), t \cap \tau_N \cap \tau_1) = V_N(x(\tau_0), \tau_0) + \int_{\tau_0}^{t \cap \tau_N \cap \tau_1} A^{\epsilon,\mu} V_N(x(s), s)\, ds + M_N(t \cap \tau_N \cap \tau_1) \tag{3.57}$$

where M_N is the zero mean martingale defined similar to (3.31) except that the lower limit of integration is τ_0. It is obvious by the definitions of τ_0, τ_1 and τ_N that for $\tau_0 \leq s \leq t \cap \tau_N \cap \tau_1$, $x(s) \in S_N - Q_0$ and then (3.55) gives

$$A^{\epsilon,\mu} V_N(x(s), s) \leq 0 \tag{3.58}$$

Therefore, (3.57) implies

$$V_N(x(t \cap \tau_N \cap \tau_1), t \cap \tau_N \cap \tau_1) \leq V_N(x(\tau_0), \tau_0) + M_N(t \cap \tau_N \cap \tau_1) \tag{3.59}$$

But from (3.56) we have:

$$\alpha_1 |x(t \cap \tau_N \cap \tau_1)|^2 \leq \alpha_2 + \beta_1 |x(\tau_0)|^2 + \beta_2 + M_N(t \cap \tau_N \cap \tau_1)$$

Then by taking conditional expectation E_{τ_0}, we have:

$$E_{\tau_0}|x(t \cap \tau_N \cap \tau_1)|^2 \leq \frac{\alpha_2}{\alpha_1} + \frac{\beta_2}{\alpha_1} + \frac{\beta_1}{\alpha_1} E_{\tau_0}|x(\tau_0)|^2 \leq K$$

for some positive constant K independent of N. And then by taking unconditional expectation, we get

$$E|x(t \cap \tau_N \cap \tau_1)|^2 \leq K$$

Since K is independent of N, letting $N \to \infty$ as we get

$$E|x(t \cap \tau_1)|^2 \leq K \quad \forall\, 0 \leq t \leq T \tag{3.60}$$

Now, we consider the following cases:

(i) If $\tau_1 \geq T$ then $E|x(t)|^2 \leq K \quad \forall\, 0 \leq t \leq T$ (3.61)

(ii) If $\tau_1 < T$ we redefine $\tau_0' = \inf_{t \geq \tau_1} \{t : |x(t)| = K\}$

If $\tau_0' \geq T$ then $E|x(t)|^2 \leq C \quad \forall\, 0 \leq t \leq T$ where $C = \max(K, K_0)$.

If $\tau_0' < T$ then τ_0' is taken as the starting time and we repeat the whole process again starting from (3.57) and with $\tau_1' = \inf\{t : t \geq \tau_0', x(t) \in Q_0\}$, and then similar to (3.60) we get:

$$E|x(t \cap \tau_1')|^2 \leq K' \quad \forall\, t \in [\tau_0', T \cap \tau_1']$$

and so on. Then we conclude that

$$E|x(t)|^2 \leq \overline{K}, \quad 0 \leq t \leq T \quad \text{for some } \overline{K} > 0,$$

independent of T. Therefore letting $T \to \infty$ we get

$$E|x(t)|^2 \leq \bar{K} \qquad \forall\, t \geq 0$$

and this completes our proof.

Remark 3: Suppose that $K = 0$ in (3.45), i.e., $LV(x) \leq -\lambda V(x)$, and that the constant M_2 in (4′) vanishes, i.e. the coefficients $a_1(x)$, $A_{12}(x)$ and $B_1(x)$ vanish at the origin and satisfy

$$|a_1(x)| + |A_{12}(x)| + |B_1(x)| \leq M_1 |x| \qquad \forall\, x \in R^n.$$

Then, if we proceed in a way similar to the steps of the proof of theorem (1) we conclude the following inequalities which are similar to (3.30) and (3.35) respectively

$$(A^{\epsilon,\mu} + \hat{\lambda}) V_N(x,t) \leq 0 \qquad \text{for} \quad x \in S_N \tag{3.62}$$

$$c_1 V(x) \leq V_N(x,t) \leq c_2 V(x) \qquad \text{for} \quad x \in S_N \tag{3.63}$$

c_1 and c_2 are positive constants independent of T and N. But since $V(x) = x'Qx$, we have;

$$\bar{c}_1 |x|^2 \leq V_N(x,t) \leq \bar{c}_2 |x|^2 \qquad \text{for} \quad x \in S_N \tag{3.64}$$

Then similar to (3.34) and by the aid of (3.62) we have:

$$e^{\hat{\lambda}(\tau_N \cap t)} V_N(x(t \cap \tau_N), t \cap \tau_N) \leq V_N(x(0),0) + M_N(t \cap \tau_N) \tag{3.65}$$

then from (3.64) we have:

$$0 \le \bar{c}_1 e^{\hat{\lambda}(\tau_N \cap t)} |x(t \cap \tau_N)|^2 \le \bar{c}_2 |x_0|^2 + M_N(t \cap \tau_N) \tag{3.66}$$

Taking unconditional expectation we get:

$$\bar{c}_1 E(e^{\hat{\lambda}(\tau_N \cap t)} |x(t \cap \tau_N)|^2) \le \bar{c}_2 E|x_0|^2 \le \bar{C}$$

By the monotonic convergence theorem we have:

$$\lim_{N\to\infty} E(e^{\hat{\lambda}(\tau_N \cap t)} |x(t \cap \tau_N)|^2) = E(e^{\hat{\lambda}t}|x(t)|^2) \le \frac{\bar{C}}{\bar{c}_1}$$

then $\qquad E(|x(t)|^2) \le Ke^{-\hat{\lambda}t} \qquad \forall\, 0 \le t \le T$

since K is independent of T, it follows that

$$E(|x(t)|^2) \le Ke^{-\hat{\lambda}t} \qquad \forall\, t \ge 0.$$

3.4. Examples:

Example 1. Consider the system:

$$\dot{x} = -5x + xy + xv^{\varepsilon} \tag{3.67}$$

$$\mu\dot{y} = -y + e^{-x^2} v^{\varepsilon} \tag{3.68}$$

We would like to study the stability of $x(t)$ when ε and μ are sufficiently small, and let us take $\gamma = .1$, as a nominal value of the ratio ε/μ. $v^{\varepsilon}(t) = \frac{1}{\sqrt{\varepsilon}} v(t/\varepsilon)$ and $v(t)$ is a zero mean, stationary, uniformly bounded process for $t \in [0,\infty)$ and satisfies a mixing condition with decaying exponential, so that, if $R(\tau)$ is the correlation function, then we have

$$|R(\tau)| \leq \bar{e}^{\tau} \quad \text{(taking the exponent to be 1 here).}$$

According to the results of Chapter 2 the state $x(t)$ of (3.67) and (3.68) can be approximated by a diffusion process $\bar{x}(t)$ whose diffusion operator is given by: (Let us assume that $S(0) = 1$, where $S(w)$ is the spectrum of v)

$$L(\cdot) = \{-5x + \frac{1}{2}[x + x\bar{e}^{x^2} - 2x^3\bar{e}^{x^2} + 2x^3\bar{e}^{2x^2} + x\bar{e}^{2x^2} + x\bar{e}^{x^2}]$$
$$+ 2x^3\bar{e}^{2x^2} \int_0^\infty \bar{e}^{.1(\tau)} R(\tau)\, d\tau\} \frac{d}{dx}(\cdot) + \frac{1}{2}x^2(1 + \bar{e}x^2)^2 \frac{d2}{dx^2}(\cdot) \tag{3.70}$$

To establish the stability properties of $x(t)$ we need to study the stability of the diffusion $\bar{x}(t)$ and this can be done, if we can find a Lyapunov function $V(x)$ which satisfies $LV(x) \leq -\lambda V(x)$, $\forall\, x \in \mathbb{R}$ and for some $\lambda > 0$. Let us choose $V(x) = x^2$, then from (3.70) we have:

$$LV(x) = -10x^2 + [x^2 + x^2\bar{e}^{x^2} - 2x^4\bar{e}^{x^2} + 2x^4\bar{e}^{2x^2} + x^2\bar{e}^{2x^2} + x^2\bar{e}^{x^2}$$
$$+ 4\int_0^\infty \bar{e}^{.1(\tau)} R(\tau)\, d\tau \cdot x^4\bar{e}^{2x^2} + x^2(1 + \bar{e}^{x^2})^2 \tag{3.71}$$
$$\leq -10x^2 + 4.36x^2 + .75x^2 \int_0^\infty \bar{e}^{(1.1)\tau} d\tau + 4x^2$$
$$\leq -x^2 = -V(x) \qquad \forall\, x \in \mathbb{R}$$

We have used that $\max_x (x^2\bar{e}^{2x^2}) \simeq .18$.

Then it follows that, [c.f. 7,20], the solution $\bar{x}$ of the reduced-order model given by L of (3.70), is stochastically asymptotically stable. Then, it follows by theorem 1 of this chapter, since all the assumptions of the theorem are satisfied, that the process $x(t)$ is stochastically asymptotically stable for ε and μ sufficiently small and for $\frac{\varepsilon}{\mu}$ sufficiently close to $\gamma = .1$.

Remark: It is interesting to notice that, if we allow $\frac{\varepsilon}{\mu}$, for example, to take values in $[.05,\infty)$ (say), then we see from (3.71), that the term:

$$4x^4 e^{-2x^2} \int_0^\infty e^{-(\frac{\varepsilon}{\mu})\tau} R(\tau)d\tau \leq 4x^4 e^{-2x^2} \int_0^\infty e^{-(\frac{\varepsilon}{\mu}+1)} = \frac{4x^4 e^{-2x^2}}{(\frac{\varepsilon}{\mu}+1)}$$

$$\leq \frac{4x^4 e^{-2x^2}}{1.05}$$

Since $\frac{\varepsilon}{\mu} \geq .05$ and the above conclusion is valid for sufficiently small ε and μ and for any $\frac{\varepsilon}{\mu}$ in $[.05,\infty)$. Although theorem 1 is valid only for the case when $\frac{\varepsilon}{\mu}$ is close to a norminal value γ in $[\gamma_1,\infty)$ for some $\gamma_1 > 0$, the proof can be modified to show that if $LV \leq -\lambda V$ is satisfied uniformly in γ then the statement of the theorem holds for all $\frac{\varepsilon}{\mu} \geq \gamma_1 > 0$.

Example 2: Consider the system:

$$\dot{x} = -2x + xy + v^{\varepsilon} \qquad (3.71)$$

$$\mu\dot{y} = -y + v^{\varepsilon} \qquad (3.72)$$

This system is different from the one which has been considered in example 1 in that, the right-hand side of (3.71) does not vanish at $x = 0$ which means that $x = 0$

is not an equilibrium point and the best that we can hope to establish is to show that $x(t)$ is bounded in the mean square, for all $t \geq 0$ and for ε and μ sufficiently small and for $\frac{\varepsilon}{\mu}$ sufficiently close to a nominal value $\gamma = .5$ (say). Let us assume that v^{ε} satisfies the same assumptions as in example 1. The process $x(t)$ of (3.71) and (3.72) can be approximated by a diffusion process $\bar{x}$ whose differential operator is given by: (We take $S(0) = 1$).

$$L(\cdot) = \{-2x + \frac{1}{2}[x+1] + \Sigma\} \frac{d}{dx}(\cdot) + \frac{1}{2}(1+x)^2 \frac{d^2}{dx^2}(\cdot) \tag{3.73}$$

where

$$\Sigma = \int_0^{\infty} R(\tau)\bar{e}^{.5\tau} d\tau \tag{3.74}$$

Then if we choose the Lyapunov function $V(x) = x^2$, (3.73) implies:

$$\begin{aligned} LV(x) &= -4x^2 + x^2 + x + 2x\Sigma + (1+x)^2 \\ &= -3x^2 + (x^2 + (3+2\Sigma)x + 1) \\ &= -3x^2 + \left(x + \frac{3+2\Sigma}{2}\right)^2 - \frac{(3+2\Sigma)^2}{4} + 1 \end{aligned}$$

Using the fact that $(a+b)^2 \leq 2(a^2+b^2)$ for any real numbers a and b and that

$$|\Sigma| \leq \int_0^{\infty} \bar{e}^{(1.5)\tau} d\tau = \frac{2}{3}$$

we get

$$LV(x) \leq 12 - x^2 = 12 - V(x) \qquad \forall x \in \mathbb{R}$$

This implies [c.f. 18] that $\bar{x}(t)$ is exponentially bounded in mean square with exponent 1, i.e. $E(|\bar{x}(t)|^2) \leq K_1 + K_2 e^{-t}$ $\forall\ t \geq 0$, for some $K_1 \geq 0$ and $K_2 > 0$. Then, since all the assumptions of theorem 2, of this chapter, are satisfied, it follows that the process $x(t)$ is bounded in the mean square sense, $\forall\ t \geq 0$, for sufficiently small ϵ and μ and for $\frac{\epsilon}{\mu}$ sufficiently close to 0.5.

CHAPTER IV

STABILIZING CONTROL

4.1. Introduction:

It is a well known fact that an important aspect of feedback design, is the stability of the control system. Whatever has to be achieved with the control system, its stability must be assured. Actually, sometimes, the main goal of a feedback design is to stabilize a system if it is initially unstable. Let us recall that the two types of feedback designs are the state feedback, in which it is assumed that the complete state of the system can be accurately measured at all times and is available for feedback, and the output feedback, which is the much more realistic case where there is an observed variable whose dimension is, in general, less than that of the states and it serves as input to the controller. The observed variable is usually corrputed by an observation noise. The states of the system, which cannot be measured accurately in this case, can be reconstructed from the observed variables and the feedback control, in this case, is a function of the reconstructed states. For example, in the case of linear systems, where both the state equation and all the output variables are corrupted by

additive white noise (the state equation is an Itô equation) one can use a Kalman filter [c.f. 37] for a state reconstruction and then a state feedback control can be designed to achieve certain prespecified objectives. Stabilizing nonlinear stochastic systems via the use of an asymptotically stable stochastic observer have been considered recently [18]. The work which is done in that paper is a generalization of the Kalman filter structure.

Until recently, singular perturbation techniques have primarily focused on state feedback design of linear systems. Advantages of these techniques, such as order reduction and separation of time scales, are expected to have a more dramatic effect on feedback design of nonlinear systems. Stabilizing deterministic nonlinear singularly perturbed systems have been considered, for example, in [2] and [38]. In this chapter we consider the stochastic stabilization problem for nonlinear singularly perturbed systems driven by wide-band noise. We consider the following system:

$$\dot{x} = a_1(x) + A_{12}(x)y + B_1(x)v_1^{\varepsilon_1} + G_1(x)u \tag{4.1}$$

$$\mu\dot{y} = a_{21}(x) + A_2 y + B_2(x)v_1^{\varepsilon_1} + G_2(x)u \tag{4.2}$$

$$z = c_1(x) + c_2 y + B_3(x)v_2^{\varepsilon_2} \tag{4.3}$$

where u is a control vector in R^p, z is the output vector in R^q $(q \leq n)$, $v_1^{\varepsilon_1} \in R^r$ and $v_2^{\varepsilon_2} \in R^s$ are independent and have the same properties, as v^{ε} defined in Chapter 2,

where ε_1 and ε_2 are different in general, so that, if the observation noise $v_2^{\varepsilon_2}$ has spectrum which is wider than that of the system noise $v_1^{\varepsilon_1}$ then we expect ε_2 to be much smaller than ε_1. The matrices G_1, G_2, C_1 and B_3 are, in general, functions in x and are required to satisfy certain smoothness conditions which are specified later. The outline of this chapter is, roughly, as follows:

1. We begin with the open-loop full-order system (4.1)-(4.3) and we aticipate an open-loop reduced order model (OLROM) in the form of an Itô equation.

2. We design a stabilizing feedback control for the above (OLROM) model which will result in a stochastically asymptotically stable closed-loop reduced order model (CLROM). Work similar to that of [18] has been done, in that regard.

3. We apply the feedback control which we obtained in step 2 to the full-order open-loop system (4.1),(4.2), and we obtain a full-order closed loop system (FOCLS) which will be of the form (2.1) and (2.2).

4. We apply results of Chapter 2 to identify the reduced-order closed-loop model (ROCLM) corresponding to (FOCLS) which has been obtained in step 3.

5. We require that the (CLROM) be the same as the (ROCLM) and this results in some conditions which will be referred to as the consistency conditions under which the OLROM will be identified completely.

6. We apply the results of Chapter 3 to obtain conditions, under which the (FOCLS) is stoachastically asymptotically stable.

Remark: In the second section we will study the case when all the states of (4.1) and (4.2) are available for perfect measurement and a stabilizing feedback controller has been designed according to the above outline. In the third section we repeat the same procedure but in this case we assume that the states are not available for perfect measurement and an output feedback controller via an observer is employed. In section 4 we illustrate the procedure by an example.

4.2. State Feedback Stabilizing Control: Let us write again the full-order system (4.1) and (4.2) with $v_1^{\varepsilon_1}$ written simply as v^{ε}

$$\dot{x} = a_1(x) + A_{12}(x)y + B_1(x)v^{\varepsilon} + G_1(x)u \tag{4.4}$$

$$\mu\dot{y} = a_{21}(x) + A_2 y + B_2(x)v^{\varepsilon} + G_2(x)u$$

We assume that the slow state variables $x(t)$ are available for measurement. Since the results of Chapter 2 indicate that $x(t)$ tends in the limit to a diffusion process, it is reasonable to anticipate that the open-loop reduced-order model corresponding to (4.4) and (4.5) takes the Itô form:

$$d\bar{x} = \tilde{b}(\bar{x})dt + \tilde{G}(\bar{x})udt + \tilde{\sigma}(\bar{x})dw \tag{4.6}$$

where the exact forms of the vector $\tilde{b}$ and the matrices $\tilde{G}$ and $\tilde{\sigma}$ will be determined later. Let us assume that there exists a sufficiently smooth function $g(x)$ such that, when the feedback control $u = g(\bar{x})$ is applied to the system (4.6), the resulting closed-loop reduced-order model

$$d\bar{x} = (\tilde{b}(\bar{x}) + \tilde{G}(\bar{x})g(\bar{x}))dt + \tilde{\sigma}(\bar{x})dw \tag{4.7}$$

is asymptotically stable in some stochastic sense (see Chapter 3). Then, we apply the control law $u = g(x)$ to the open-loop full order system (4.4) and (4.5) to obtain the closed-loop full order system

$$\dot{x} = \tilde{a}_1(x) + A_{12}(x)y + B_1(x)v^{\varepsilon} \tag{4.8}$$

$$\mu\dot{y} = \tilde{a}_{21}(x) + A_2 y + B_2(x)v^{\varepsilon} \tag{4.9}$$

where

$$\tilde{a}_1 = a_1 + G_1 g \tag{4.10}$$

and

$$\tilde{a}_{21} = a_{21} + G_2 g \tag{4.11}$$

Equations (4.8) and (4.9) are in the form of (2.1) and (2.2) respectively. Then, the reduced-order closed-loop model corresponding to (4.8) and (4.9) can be obtained by applying the results of Chapter 2, assuming that the coefficients and the process v^{ε} satisfy the required

assumptions which will be spelled out later. So, the reduced-order closed-loop model corresponding to (4.8) and (4.9) is:

$$d\bar{x} = b(\bar{x})dt + (G_1(\bar{x}) - A_{12}(\bar{x})A_2^{-1}G_2(\bar{x}))g(\bar{x})dt + \sqrt{A(\bar{x})}\, dw \tag{4.12}$$

where $b(x)$ and $A(x)$ are defined similar to (2.8) and (2.9). Now we impose a consistency condition which is stated as follows: The closed-loop reduced-order model, which is obtained by applying the control $u = g(\bar{x})$ to the open-loop reduced-order model, is the same as the reduced-order closed-loop model corresponding to the closed-loop full-order system (obtained by applying the same control $u = g(x)$ to the open-loop full-order system).

This, condition says that the coefficients of (4.6) must be the same as those of (4.12) for any $g(x)$ and this implies:

$$\tilde{b}(x) = b(x) \qquad \forall\, x \in R^n \tag{4.13}$$

$$\begin{aligned} \tilde{G}(x) &= G_1(x) - A_{12}(x)A_2^{-1}G_2(x) \qquad \forall\, x \in R^n \\ &\triangleq G_0(x) \end{aligned} \tag{4.14}$$

$$\tilde{\sigma}(x) = \sqrt{A(x)} \tag{4.15}$$

Hence the open-loop reduced-order Itô model that approximates the slow states of the non-Markov open-loop full-order system (4.4) and (4.5) is given by:

$$d\bar{x} = b(\bar{x})dt + G_0(\bar{x})udt + \sqrt{A(\bar{x})}\,dw \tag{4.16}$$

With the (OLROM) (4.16) in hand we can proceed now to design the feedback control $u = g(x)$ to stabilize x of (4.16).

This control task is much simpler than the original task of stabilizing x of (4.4), (4.5) since now we are dealing with the Itô equation (4.16) for which stability and stabilization techniques exist in the literature [c.f. 7, 16, 18, 20]. Suppose now that we succeeded in finding a sufficiently smooth function $g(x)$ with $g(0) = 0$ such that the application of the feedback control $u = g(\bar{x})$ to (4.16) results in a stochastically asymptotically stable (CLROM) with a diffusion operator $\tilde{L}$ given by:

$$\tilde{L}f(x) = (b(x) + G_0(x)g(x))' \cdot f_x(x) + \frac{1}{2}\,\mathrm{tr}\,(A(x)f_{xx}(x)) \qquad (4.17)$$

The use of the feedback control $u = g(x)$ with the full system (4.4) and (4.5) is justified by the following theorem whose proof is a straight forward application of Theorem 1 of Chapter 3.

<u>Theorem 1</u>: Suppose that there exists a Lyapunov function $V(x)$ for $x \in R^n$ which satisfies all the assumptions of Theorem 1 Chapter 3 and that $\tilde{L}V(x) \leq -\lambda V(x)$ for some $\lambda > 0$. Moreover, suppose that assumptions similar to (1)-(6) stated in Chapter 3 are satisfied where, a_1, a_{21} and a_0 are replaced by $\tilde{a}_1, \tilde{a}_{21}$ and $\tilde{a}_0 = \tilde{a}_1 - A_{12}A_2^{-1}\tilde{a}_{21}$, respectively. Then the solution, $x(t)$, of (4.4) and (4.5), with the control $u = g(x)$, is uniformly stochastically

asymptotically stable as $t \to \infty$, for sufficiently small ϵ and μ and for $\frac{\epsilon}{\mu}$ sufficiently close to a normal value $\gamma \in [\gamma_1, \infty)$ for some $\gamma_1 > 0$.

4.3. Output Feedback Stabilizing Control:

Let us consider the full-order open loop system (4.1) and (4.2) with z, given by (4.3), representing the observed variables, where $v_1^{\epsilon_1}(t) = \frac{1}{\sqrt{\epsilon_1}} v_1(t/\epsilon_1)$, $v_2^{\epsilon_2}(t) = \frac{1}{\sqrt{\epsilon_2}} v_2(t/\epsilon_2)$, $\frac{\epsilon_1}{\mu} \in [\gamma_1, \infty)$ for $\gamma_1 > 0$ arbitrary but fixed and $v_1(t)$ and $v_2(t)$ satisfy all the assumptions given in Chapter 2. Moreover, let $R_1(\tau) = E(v_1(t)v_1'(t+\tau))$, $R_2(\tau) = E(v_2(t)v_2'(t+\tau))$, $W_1 = \int_0^\infty R_1(\tau)d\tau$, $W_2 = \int_0^\infty R_2(\tau)d\tau$ and $S_1(w)$ and $S_2(w)$ denote spectrum matrices of v_1 and v_2, respectively. The main assumption in this section is that the states of the system are not available for perfect measurement. Then, we follow essentially the basic steps of section 4.2 to stabilize the initially unstable system (4.1) and (4.2). So we proceed in doing that as follows: We anticipate that the open-loop reduced-order model of (4.1)-(4.3) takes the form:

$$d\bar{x} = f_1(\bar{x})dt + F_1(\bar{x})udt + \sigma_1(\bar{x})dw_1 \qquad (4.18)$$

$$d\bar{Z} = f_2(\bar{x})dt + F_2(\bar{x})udt + \sigma_2(\bar{x})dw_1 + \sigma_3(\bar{x})dw_2 \qquad (4.19)$$

We consider a controller of the form $u = g(\hat{x})$ where $\hat{x}$ is the output of the observer:

$$d\hat{x} = f_1(\hat{x})dt + F_1(\hat{x})g(\hat{x})dt + K[dz - f_2(\hat{x})dt - F_2(\hat{x})g(\hat{x})dt] \tag{4.20}$$

for some constant gain matrix K. The vectors $f_1(x)$, $f_2(x)$ and the matrices $F_1, F_2, \sigma_1, \sigma_2$ and σ_3 are to be determined by applying a similar consistency condition to the one stated in section 4.2. The closed-loop reduced-order augumented model which follows from (4.18)-(4.19) with $u = g(\hat{x})$ is:

$$d\bar{x} = f_1(\bar{x})dt + F_1(\bar{x})g(\hat{x})dt + \sigma_1(\bar{x})dw_1 \tag{4.21}$$

$$\begin{aligned} d\hat{x} &= f_1(\hat{x})dt + F_1(\hat{x})g(\hat{x})dt + K(f_2(\bar{x}) - f_2(\hat{x}))dt \\ &\quad K(F_2(\bar{x}) - F_2(\hat{x}))g(\hat{x})dt + K\sigma_2(\bar{x})dw_1 + K\sigma_3(\bar{x})dw_2 \end{aligned} \tag{4.22}$$

To determine the exact form of $f_1, f_2, F_1, F_2, \sigma_1, \sigma_2$ and σ_3 we propose an observer for the full system (4.1)-(4.3), to reconstruct the states x, in the form:

$$\dot{\tilde{x}} = f_1(\tilde{x}) + F_1(\tilde{x})u + K(z - f_2(\tilde{x}) - F_2(\tilde{x})u) \tag{4.23}$$

The gain matrix K is the same as the one appearing in (4.22). Now applying the same control law $u = g(\tilde{x})$ to (4.1), (4.2) and (4.23), as a function of the reconstructed states $\tilde{x}$, then the closed-loop full-order augumented system takes the form:

$$\begin{aligned} \dot{x} &= a_1(x) + G_1(x)g(\tilde{x}) + A_{12}(x)y + B_1(x)v_1^{\epsilon_1} \\ \dot{\tilde{x}} &= f_1(\tilde{x}) + F_1(\tilde{x})g(x) + K(c_1(x) - f_2(\tilde{x}) - F_2(\tilde{x})g(\tilde{x}) \\ &\quad + Kc_2y + KB_3(x)v_2^{\epsilon_2} \end{aligned}$$

$$\mu\dot{y} = a_{21}(x) + G_2(x)g(\tilde{x}) + A_2 y + B_2(x)v_1^{\epsilon_1}$$

which can be simplified to

$$\dot{X} = \tilde{a}_1(X) + \tilde{A}_{12}(X)y + \tilde{B}_1(X)v^{\epsilon} \tag{4.24}$$

$$\mu\dot{y} = \tilde{a}_{21}(X) + A_2 y + \tilde{B}_2(X)v^{\epsilon} \tag{4.25}$$

where $X = \begin{pmatrix} x \\ \tilde{x} \end{pmatrix}$, $v^{\epsilon} = \begin{pmatrix} v_1^{\epsilon_1} \\ v_2^{\epsilon_2} \end{pmatrix}$

$$\tilde{a}_1(X) = \begin{pmatrix} a_1(x) + G_1(x)g(\tilde{x}) \\ f_1(\tilde{x}) + F_1(\tilde{x})g(\tilde{x}) + K(c_1(x) - f_2(\tilde{x}) - F_2(\tilde{x})g(\tilde{x})) \end{pmatrix} \tag{4.26}$$

$$\tilde{a}_{21}(X) = (a_{21}(x) + G_2(x)g(\tilde{x})) \tag{4.27}$$

$$\tilde{A}_{12}(X) = \begin{pmatrix} A_{12}(x) \\ KC_2 \end{pmatrix} \tag{4.28}$$

$$B_1(X) = \begin{pmatrix} B_1(x) & O \\ O & KB_3(x) \end{pmatrix} \tag{4.29}$$

and

$$B_2(X) = (B_2(x) \quad O) \tag{4.30}$$

System (4.24) and (4.25) are basically in the form of (2.1) and (2.2) and under additional assumptions, which will be stated later, it can be shown by the convergence result of Chapter 2 that $X(t)$ of (4.24) and (4.25) converges weakly to a diffusion $X = \begin{pmatrix} \bar{x} \\ \hat{x} \end{pmatrix}$ as $\epsilon_1 \to 0$, $\epsilon_2 \to 0$, $\mu \to 0$ and $\frac{\epsilon_1}{\mu} \to \gamma$. If we trace the steps of the convergence proof in the case of only two parameters ϵ and μ we will find that, in the case of three parameters, the ratio is given by $\frac{\epsilon_1}{\mu}$ and it does not depend on ϵ_2, and so we will require $\frac{\epsilon_1}{\mu} \to \gamma \in [\gamma_1, \infty)$. Moreover all the upper bounds that we established in the steps of the proof which were $O(\mu + \epsilon)$ or $O(\sqrt{\mu} + \sqrt{\epsilon})$ will depend here on $\epsilon_1, \epsilon_2, \mu$ and on the fact that $\frac{\epsilon_1}{\mu} \geq \gamma_1$. Let us derive the differential operator corresponding to $\bar{X}$ with the aid of (2.7)-(2.17), where the assumptions that will be listed later, will validate this derivation. As in (2.8), the drift coefficient is

$$b(X) = \tilde{a}_0(X) + \tilde{h}_1(X) - \tilde{A}_{12}(X) A_2^{-1} \tilde{h}_2(X) + \tilde{h}_3(X) \tag{4.31}$$

The diffusion coefficient

$$\tilde{A}(X) = \tilde{\sigma}(X) \tilde{\sigma}'(X) \tag{4.32}$$

where

$$\tilde{\sigma} = \begin{pmatrix} \tilde{\sigma}_{11} & \tilde{\sigma}_{12} \\ \tilde{\sigma}_{21} & \tilde{\sigma}_{22} \end{pmatrix} \tag{4.33}$$

$$\tilde{a}_0(X) = \tilde{a}_1(X) - \tilde{A}_{12}(X)A_2^{-1}\tilde{a}_{21}(X)$$

$$= \begin{pmatrix} a_0(\bar{x}) + G_0(\bar{x})g(\hat{x}) \\ \\ f_1(\hat{x}) + (F_1(\hat{x}) - KF_2(\hat{x}) - Kc_2A_2^{-1}G_2(\bar{x}))g(\hat{x}) + K(c_1(\bar{x}) - f_2(\hat{x}) - c_2A_2^{-1}a_{21}(\bar{x})) \end{pmatrix}$$

(4.34)

where

$$G_0(x) = G_1(x) - A_{12}(x)A_2^{-1}G_2(x) \qquad (4.35)$$

and $a_0(x)$ is defined as in (2.5):

$$B_0(X) = \begin{pmatrix} B_0(\bar{x}) & 0 \\ \\ -Kc_2A_2^{-1}B_2(\bar{x}) & KB_3(\bar{x}) \end{pmatrix} \qquad (4.36)$$

where $B_0(x)$ is defined as in (2.6)

$$\tilde{W} = \begin{pmatrix} W_1 & 0 \\ \\ 0 & W_2 \end{pmatrix} \qquad (4.37)$$

$$\tilde{\Sigma} = (\Sigma \qquad 0) \qquad (4.38)$$

where

$$\Sigma = \int_0^\infty e^{A_2(\frac{\varepsilon_1}{\mu})\tau} B_2(x)R_1'(\tau)d\tau \qquad (4.39)$$

From (4.29) and similar to (2.13) we have: (for $X = \begin{pmatrix} x \\ \hat{x} \end{pmatrix}$)

$$\tilde{D}_i(X) = \begin{pmatrix} D_i(x) & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{for} \quad i = 1,2,\cdots,n \tag{4.40}$$

where

$D_i(x)$ is defined in (2.13)

and

$$\tilde{D}_i(X) = \begin{pmatrix} 0 & \bar{D}_i(x) \\ 0 & 0 \end{pmatrix} \qquad \text{for} \quad i = n+1,\cdots,2n \tag{4.41}$$

where

$$\bar{D}_i(x) = [\nabla_x \alpha_{i1}(x) \quad \nabla_x \alpha_{is}(x)]_{n\times s}; \; KB_3(x) = [\alpha_{ij}(x)]_{n\times s}$$

$$\text{for} \quad i = n+1,\cdots,2n \tag{4.42}$$

From (4.28) and similar to (2.15) we have

$$\tilde{F}_i(X) = \begin{pmatrix} F_i(x) \\ 0 \end{pmatrix} \qquad i = 1,2,\cdots,n \tag{4.43}$$

where F_i is defined as in (2.15) and

$$\tilde{F}_i(X) = 0_{2n\times m} \qquad i = n+1,\cdots,2n \tag{4.44}$$

and from (4.30) and similar to (2.14) we have:

$$\tilde{E}_i(X) = \begin{pmatrix} (E_i)_{n\times r} & 0 \\ 0 & 0 \end{pmatrix} \quad i = 1,2,\cdots,m \tag{4.45}$$

where E_i is defined as in (2.14).

Then similar to (2.10) and from (4.36), (4.37), (4.38), (4.40) and (4.41) we have:

$$\begin{aligned} \tilde{h}_{1i} &= \mathrm{tr}[\tilde{D}_i\tilde{B}_O\tilde{W}' + \tilde{D}_i'\tilde{A}_{12}A_2^{-1}\tilde{\Sigma}] \\ &= \mathrm{tr}[D_i'B_O(x)W + D_i'A_{12}A_2^{-1}\Sigma] = h_{1i}(x) \\ &\qquad \text{for} \quad i = 1,2,\cdots,n \end{aligned} \tag{4.46}$$

where h_{1i} is defined as in (2.10) and from (4.41) we get

$$\tilde{h}_{1i} = 0 \quad \text{for} \quad i = n+1,\cdots,2n \tag{4.47}$$

Hence

$$\tilde{h}_1(X) = \begin{pmatrix} h_1(x) \\ 0 \end{pmatrix} \tag{4.48}$$

Similar to (2.11), we have:

$$\begin{aligned} \tilde{h}_{2i}(X) &= \mathrm{tr}[\tilde{E}_i'\tilde{B}_O\tilde{W}' + \tilde{E}_i'\tilde{A}_{12}A_2^{-1}\tilde{\Sigma}] = h_2(x) \\ &\qquad \text{for} \quad i = 1,2,\cdots,m \end{aligned} \tag{4.49}$$

where $h_2(x)$ is defined as in (2.11), and similar to (2.12) we have

$$\tilde{h}_{3i}(X) = \mathrm{tr}[-\tilde{F}_i'\tilde{B}_O\tilde{W}'\tilde{B}_2'A^{-\frac{1}{2}} - \tilde{F}_i'\tilde{B}_O\tilde{\Sigma}'(A^{-\frac{1}{2}})' + \tilde{F}_i'\tilde{A}_{12}A^{-\frac{1}{2}}\tilde{P}]$$

$$= h_{3i}(x) \quad \text{for} \quad i = 1,2,\cdots,n \tag{4.50}$$

and

$$h_{3i}(X) = 0 \quad \text{for} \quad i = n+1,\cdots,2n \tag{4.51}$$

$$\text{Hence} \quad \tilde{h}_3(X) = \begin{pmatrix} h_3(X) \\ \\ 0 \end{pmatrix} \tag{4.52}$$

where $h_{3i}(x)$ is defined as in (2.12). Notice that, it can be verified that $\tilde{P} = P$, where P is defined as in (2.17).

$$\text{Now, since} \quad \tilde{A}(X) = \tilde{B}_O(X)\tilde{S}(O)\tilde{B}_O'(X) \tag{4.53}$$

$$\text{where} \quad \tilde{S}(O) = \begin{pmatrix} S_1(O) & O \\ \\ O & S_2(O) \end{pmatrix}$$

Then from (4.32),(4.33),(4.36) and (4.53) we get:

$$\tilde{\sigma}(X) = \begin{pmatrix} B_O(x)\sqrt{S_1(O)} & O \\ \\ -Kc_2A^{-\frac{1}{2}}B_2(x)\sqrt{S_1(O)} & KB_3(x)\sqrt{S_2(O)} \end{pmatrix} \tag{4.54}$$

Then the reduced-order closed-loop model corresponding to (4.24) and (4.25) takes the form:

$$d\bar{x} = (b(\bar{x}) + G_0(\bar{x})g(\hat{x}))dt + B_0(x)\sqrt{S(0)}\ dw_1 \tag{4.55}$$

$$\begin{aligned} d\hat{x} = {} & [f_1(\hat{x}) + (F_1(\hat{x}) - KF_2(\hat{x}) - Kc_2 - Kc_2A_2^{-1}G_2(\bar{x}))g(\hat{x}) \\ & + K(c_1(\bar{x}) - f_2(\hat{x}) - c_2A_2^{-1}a_{21}(\bar{x}) - c_2A_2^{-1}h_2(\bar{x})]dt \\ & - Kc_2A_2^{-1}B_2(\bar{x})\sqrt{S_1(0)}\ dw_1 + KB_3(\bar{x})\sqrt{S_2(0)}\ dw_2 \end{aligned} \tag{4.56}$$

Applying the consistency requirement, as stated in section 4.2, the reduced-order closed-loop model (4.55) and (4.56) has to be the same as the closed loop reduced-order model given by (4.21) and (4.22). Hence by comparing the coefficients in the two systems and insisting that they must be equal for all K and for all functions $g(x)$, we have:

$$f_1(x) = b(x) \tag{4.57}$$

$$F_1(x) = G_0(x) \tag{4.58}$$

$$\sigma_1(x) = B_0(x)\sqrt{S_1(0)} \tag{4.59}$$

$$f_2(x) = c_1(x) - c_2A_2^{-1}a_{21}(x) - c_2A_2^{-1}h_2(x) \tag{4.60}$$

$$F_2(x) = -c_2A_2^{-1}G_2(x) \tag{4.61}$$

Hence, the open-loop reduced-order model (4.18) and (4.19) can be written in the form.

$$d\bar{x} = b(\bar{x})dt + G_0(\bar{x})udt + \sigma_1(\bar{x})dw_1 \tag{4.62}$$

$$d\bar{z} = c_0(x)dt + F_0(x)udt + \sigma_2(\bar{x})dw_1 + \sigma_3(\bar{x})dw_2 \tag{4.63}$$

where $b(x)$ and $G_0(x)$ are defined by (2.8) and (4.35) respectively,

$$c_0(x) = c_1(x) - c_2A_2^{-1}a_{21}(x) - c_2A_2^{-1}h_2(x) \tag{4.64}$$

$$F_0(x) = -c_2A_2^{-1}G_2(x) \tag{4.65}$$

$$\sigma_1(x) = B_0(x)\sqrt{S_1(0)} \tag{4.66}$$

$$\sigma_2(x) = -c_2A_2^{-1}B_2(x)\sqrt{S_1(0)} \tag{4.67}$$

$$\sigma_3(x) = B_3(x)\sqrt{S_2(0)} \tag{4.68}$$

and the proposed observer (4.20) takes the form:

$$\begin{aligned} d\hat{x} &= (b(\hat{x}) + G_0(\hat{x})g(\hat{x}))dt + K(c_0(\bar{x}) - c_0(\hat{x}))dt \\ &\quad + K(F_0(\bar{x}) - F_0(\hat{x}))g(\hat{x})dt + K\sigma_2(\bar{x})dw_1 + K\sigma_3(\bar{x})dw_2 \end{aligned} \tag{4.69}$$

The design problem, is to choose a function $g(x)$ which is smooth enough and a constant matrix K such that both the state x and the error $e = \bar{x} - \hat{x}$ will be stochastically asymptotically stable. Let us write the Itô equations (4.62) and (4.69) in the form (using $X = \begin{pmatrix} \bar{x} \\ \hat{x} \end{pmatrix}$)

$$d\bar{X} = \tilde{b}(\bar{X})dt + \tilde{\sigma}(\bar{X})dw \tag{4.70}$$

$$\tilde{b}(X) = \begin{pmatrix} b(\bar{x}) + G_0(\bar{x})g(\hat{x}) \\ b(\hat{x}) + G_0(\hat{x})g(\hat{x}) + K(c_0(\bar{x}) - c_0(\hat{x})) + K(F_0(\bar{x}) - F_0(\hat{x}))g(\hat{x}) \end{pmatrix} \tag{4.71}$$

and

$$\bar{\sigma}(X) = \begin{pmatrix} \sigma_1(\bar{x}) & 0 \\ K\sigma_2(\bar{x}) & K\sigma_3(\bar{x}) \end{pmatrix} \tag{4.72}$$

Suppose we succeed in finding $u = g(x)$ and the gain matrix K to stabilize (4.70), then the next step is to apply the same control law to the open-loop full order system (4.1)-(4.3) where $u = g(\tilde{x})$ and $\tilde{x}$ is the reconstructed states and satisfies the equation of the following observer:

$$\dot{\tilde{x}} = b(\tilde{x}) + G_0(\tilde{x})g(\tilde{x}) + K(c_1(x) - c_0(\tilde{x})) - KF_0(\tilde{x})g(\tilde{x}) + Kc_2y + KB_3(\tilde{x})v_2^{\epsilon_2} \; . \tag{4.73}$$

where K in (4.73) is the same gain matrix obtained above. Then we would like to spell out all the conditions, under which the stability of (4.72) would imply that of (4.1), (4.73) and (4.2), when $u = g(\tilde{x})$ is applied to (4.1) and (4.2). This will be done with the aid of the results of Chapter 3. We state here the assumptions that will imply asymptotic stability in probability according to Theorem 1 of Chapter 3. This will require that we consider the case when $c_2 \equiv 0$ (in the case when $c_2 \not\equiv 0$ assumptions can be made to show boundedness in the mean square sense according to Theorem 2 of Chapter 3.) So considering $c_2 \equiv 0$ we require the following assumptions:

(A) $B_0(0) = 0$, $B_3(0) = 0$, $A_{12}(0) = 0$, $C_1(0) = 0$, $a_1(0) = 0$, and $b(0) = 0$.

(B) $\tilde{b}(X)$ and $\tilde{\sigma}(X)$ are required to satisfy:

$$|\tilde{b}(X) - \tilde{b}(Y)| + |\tilde{\sigma}(X) - \tilde{\sigma}(Y)| \leq C|X - Y| \quad \forall X,Y \in R^{2n}$$

and

$$|\tilde{b}(X)|^2 + |\tilde{\sigma}(X)|^2 \leq C(1 + |X|^2) \quad \forall X \in R^{2n}$$

(C) The coefficients $\tilde{a}_1(X)$, $\tilde{A}_{12}(X)$, $\tilde{B}_1(X)$, $\tilde{a}_{21}(X)$ and $\tilde{B}_2(X)$ are continuous and have continuous partial derivatives up to the second order which are uniformly bounded in $X \in R^{2n}$ in addition to $\tilde{a}_{21}(X)$ and $\tilde{B}_2(X)$.

(D) $|\tilde{a}_1(X)| + |\tilde{A}_{12}(X)| + |\tilde{B}_1(X)| \leq K|X| \quad \forall X \in R^{2n}$

and for some $K > 0$.

(E) $|\tilde{a}_0(X) - \tilde{a}_0(Y)| + |\tilde{B}_0(X) - \tilde{B}_0(Y)| \leq K|X - Y| \quad \forall X,Y \in R^{2n}$

for some $K > 0$.

(F) $v_1(t)$ and $v_2(t)$ satisfy the same type of conditions as stated in (A1) of Chapter 2.

(G) The constant matrix A_2 is Hurwitz, i.e., $\text{Re}\,\lambda(A_2) < 0$.

Now we state the following theorem:

<u>Theorem 2</u>: Suppose that there exists a Lyapunov function $V(X)$ on R^{2n} which satisfies the same type of assumptions as in (a) and (b) of section (3.2), but for $X \in R^{2n}$. Moreover suppose that the assumptions (A)-(G) are satisfied and let $\tilde{L}$ be the diffusion operator corresponding to the diffusion process $\bar{X}$ defined by (4.70), and for some $\lambda > 0$

$$\tilde{L}V(X) \leq -\lambda V(X) \qquad \forall X \in R^{2n} \tag{4.74}$$

Then there exist ϵ_o^i, μ_o, and γ_o such that for all $0 < \epsilon_i \leq \epsilon_o^i$, $0 < \mu \leq \mu_o$ and $|\frac{\epsilon_1}{\mu} - \gamma| \leq \gamma_o$, $X(t) = \begin{pmatrix} x(t) \\ \tilde{x}(t) \end{pmatrix}$, the solution of the closed-loop system resulting from (4.1), (4.2) and (4.73) after applying the feedback control $u = g(\tilde{x})$ is stochastically asymptotically stable as $t \to \infty$.

Proof: From the assumptions and (4.76), the solution of the closed-loop reduced-order model represented by (4.72) is stochastically asymptotically stable then following exactly the steps of the proof of Theorem 1 of Chapter 3 after the necessary modification concerning the following estimates, which are similar to the estimates given by (3.14), (3.18), (3.23), (3.26) and (3.27) respectively. (S_N here is subset of R^{2n})

$$|V_{1,N}(X,t)| \leq (K_1\sqrt{\epsilon} + K_2\sqrt{\epsilon_2} + K_3\sqrt{\mu})V(X),$$

$$|V_{2,N}(X,t)| \leq (K_4\epsilon_1 + K_5\epsilon_2 + K_6\mu)V(X),$$

$$|L^{\epsilon_1/\mu}V_N(X) - LV_N(X)| \leq c|\frac{\epsilon_1}{\mu} - \gamma|V(X),$$

$$|\frac{\partial V_{N,1}}{\partial X}(X,t)\tilde{a}_o(X)| \leq (K_7\sqrt{\epsilon_1} + K_8\sqrt{\epsilon_2} + K_9\sqrt{\mu})\ V(X),$$

and

$$|\frac{\partial V_{N,2}}{\partial X}(X,t)\tilde{a}_1(X)| + |\frac{\partial V_{N,2}}{\partial X}(X,t)\ (\tilde{A}_{12}(X)y + \tilde{B}_1(x)\ (v_1^{\epsilon_1}, v_2^{\epsilon_2})')|$$

$$\leq (K_{10}\sqrt{\epsilon_1} + K_{11}\sqrt{\epsilon_2} + K_{12}\sqrt{\mu} + K_{13}\epsilon_1 + K_{14}\epsilon_2 + K_{15}\mu)V(X)$$

all the above inequalities are true for $X \in S_N$. Also the above estimates depend on the fact that $\frac{\epsilon_1}{\mu} \geq \gamma_1$. Then the result of the theorem follows.

4.4. Example:

Consider the system

$$\dot{x} = x + xy + xv_1^{\epsilon_1} + u \tag{4.75}$$

$$\mu\dot{y} = -\frac{9}{8} - y - \frac{1}{2}v_1^{\epsilon_1} \tag{4.76}$$

with observed variable

$$Z = x + xv_2^{\epsilon_2} \tag{4.77}$$

Our objective is to design an output feedback control u, as a function of the recontructed state $\tilde{x}$, to stabilize the above system. We assume that

$$v_1^{\epsilon_1}(t) = \frac{1}{\sqrt{\epsilon_1}} v_1(t/\epsilon_1) \quad \text{and} \quad v_2^{\epsilon_2}(t) = \frac{1}{\sqrt{\epsilon_2}} v_2(t/\epsilon_2)$$

where $v_1(t)$ and $v_2(t)$ are taken to be independent, zero mean, stationary, uniformly bounded processes for $t \in [0,\infty)$ and satisfy mixing conditions with decaying exponentials. Let $R_1(\tau)$ and $R_2(\tau)$ denote the correlation functions of v_1 and v_2 respectively, then we have:

$$|R_1(\tau)| \leq e^{-\alpha_1\tau} \quad \text{and} \quad |R_2(\tau)| \leq e^{-\alpha_2\tau}$$

for some $\alpha_1 > 0$ and $\alpha_2 > 0$. Let $S_1(w)$ and $S_2(w)$ denote

the corresponding power spectrum functions respectively. The open-loop reduced-order model is given by (we take $S_1(0) = S_2(0) = 1$).

$$d\bar{x} = udt + \tfrac{1}{2}\bar{x}dw_1 \tag{4.78}$$

$$dz = xdt + xdw_2 \tag{4.79}$$

We want to find a control $u = g(\hat{x})$ where $\hat{x}$ satisfies the equation of the observer

$$d\hat{x} = udt + K[dz - \hat{x}dt] \tag{4.80}$$

The design problem is to choose an appropriate function $g(\hat{x})$ and a constant K such that $\bar{x}$ and $\hat{x}$ are stochastically asymptotically stable. So if we choose $g(\hat{x}) = -F\hat{x}$ and we define $e = \bar{x} - \hat{x}$ then from (4.78)-(4.80) we have:

$$\begin{pmatrix} d\bar{x} \\ de \end{pmatrix} = \begin{pmatrix} -F & F \\ 0 & -K \end{pmatrix}\begin{pmatrix} \bar{x} \\ e \end{pmatrix} dt + \begin{pmatrix} \tfrac{1}{2}\bar{x} & 0 \\ \tfrac{1}{2}\bar{x} & -K\bar{x} \end{pmatrix} dw \tag{4.81}$$

where $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$

By taking $V(x,e) = x^2 + e^2$, we have:

$$\mathcal{L}V(\bar{x},e) = \begin{pmatrix} \bar{x} \\ e \end{pmatrix}' \begin{pmatrix} -F & F \\ 0 & -K \end{pmatrix}\begin{pmatrix} \bar{x} \\ e \end{pmatrix} + \tfrac{1}{4}\bar{x}^2 + \tfrac{1}{4}\bar{x}^2 + K^2\bar{x}^2$$

where $\mathcal{L}$ is the diffusion operator corresponding to the Itô equation (4.81). Choosing $K = 1$, $F = 2$ we get

$$\mathcal{L}V(\bar{x},e) = -(\tfrac{5}{2}\bar{x}^2 - 4\bar{x}e + 2e^2) \leq -.23(\bar{x}^2 + e^2) = -.23V(\bar{x},e)$$

$$\forall (\bar{x},e) \in \mathbb{R}^2$$

Then it follows that, the solution $\begin{pmatrix} \bar{x} \\ e \end{pmatrix}$ of (4.81) is stochastically asymptotically stable. Then applying the control law $u = -2\tilde{x}$ to the open-loop full-order system (4.75)-(4.77), where $\tilde{x}$ satisfies the equation of the observer

$$\dot{\tilde{x}} = -3\tilde{x} + u + (z - \tilde{x}) \tag{4.85}$$

will stabilize the system according to theorem 2 of this chapter, since all the assumptions of the theorem are obviously satisfied.

This conclusion holds for sufficiently small $\varepsilon_1, \varepsilon_2$ and μ. We notice that in this particular example we do not have to require that the ratio ε_1/μ be sufficiently close to a nominal value γ since the reduced-order model is independent of γ.

CHAPTER 5

DISCUSSION AND CONCLUSION

5.1. Discussion:

In this section we discuss the reduced-order model defined by the operator L^{γ} of (2.7) and explore various special cases of practical significance. Inspecting the drift coefficient $b(x)$ and the diffusion coefficient $A(x)$, given in chapter 2, shows that the wide-band nature of $v^{\varepsilon}(t)$ affects only the drift coefficient. In other words, if one had tried to obtain a reduced-order model by following the intuitively appealing, but wrong, procedure of simply replacing the wide-band noise by its limit white noise and then applying the order reduction procedure of singularly perturbed deterministic systems [1], he would have obtained a reduced-order model with drift coefficient $a_0(x)$ and diffusion coefficient $A(x)$. The differences between the two drift coefficients are the terms h_1, $-A_{12}A_{22}^{-1}h_2$ and h_3. These terms depend, respectively, on the partial derivatives of B_1, B_2 and A_{12} with respect to x. The appearance of the partial derivatives of B_1 and B_2 should be expected in view of the asymptotic analysis of nonlinear systems driven by wide-band noise [11-13]. The appearance of the partial derivatives of

A_{12} is less obvious. However, if we take into consideration that as ϵ and μ tend to zero the process $y(t)$ itself tends to white noise, we can see that A_{12} plays the role of an input matrix multiplying wide-band process, similar to the roles played by B_1 and B_2. It is interesting to notice that if the matrices A_{12}, B_1 and B_2 are constant (independent of x), the terms h_1, h_2 and h_3 will vanish. In this special case applying the intuitive procedure of formally setting $\mu = 0$ and formally replacing the wide-band noise by its limit white noise, would lead to the correct reduced-order model.

One disturbing fact about the reduced-order model (2.7) is that the drift coefficient $b(x)$ depends on $\gamma = \lim_{\epsilon,\mu \to 0} \frac{\epsilon}{\mu}$ through the matrices Σ and P. This is the consequence of the interaction between the asymptotic phenomena associated with singular perturbations on one hand, and the asymptotic phenomena associated with rapid stochastic fluctuations on the other hand. The dependence of L^{γ} on γ has important impact on the engineering practice of neglecting parasitic elements when writing down differential equations representing electrical networks, mechanical systems, etc. It is apparent now that if one would be interested in solving those equations when driven by wide-band noise and using the usual white noise approximations, the parasitic elements should not be neglected from the outset. Rather, they should be included in the system description and their relationship with the wide-band noise be studied in order to obtain the right reduced-order

diffusion model.

Fortunately, there are interesting classes of systems for which the engineering practice will work out without causing trouble. These are systems for which the operator L^{γ} will be independent of γ. Using the explicit form of the operator L^{γ} given by (2.7)-(2.17), we can easily identify classes of systems for which this is true. Essentially, we need to look for special cases when $\Sigma = 0$ or when the partial derivatives multiplying Σ and P vanish. For example, when $B_2 = 0$, the matrix $\Sigma = 0$. That is intuitively clear since $B_2 = 0$ means that $y(t)$ would be a smooth process whose elimination from (2.1) can be done using the usual singular perturbation routine. Indeed, we do not need $B_2 = 0$ for $y(t)$ to be a smooth process in the limit. We only need that B_2 takes the special form $B_2(x) = \mu^{\alpha}\tilde{B}_2(x)$ or $B_2(x) = \varepsilon^{\alpha}\tilde{B}_2(x)$ for some constant $\alpha > \frac{1}{2}$. Checking the proof of the theorem, it can be seen that the terms containing B_2, Σ or P drop out. In addition, we have already seen that for the class of systems in which A_{12}, B_1 and B_2 are constant matrices; the terms h_1, h_2 and h_3 vanish and the drift coefficient $b(x)$ reduces to $a_0(x)$ which is independent of γ.

In Chapter 1 we have outlined Blankenship's ppproach [28] and implied that it is valid when $\frac{\mu}{\varepsilon} \to 0$ as $\varepsilon \to 0$. This can be verified for our problem by applying Blankenship's procedure to our system. The algebraic equation

$$a_{21}(x) + A_2 y + B_2(x) v^{\varepsilon} = 0 \tag{5.1}$$

has the unique solution

$$y = -A_2^{-1}[a_{21}(x) + B_2(x) v^{\varepsilon}]. \tag{5.2}$$

Using (5.2), an outer solution for x is defined by

$$\dot{X}(t) = a_0(X(t)) + B_0(X(t)) v^{\varepsilon}(t). \tag{5.3}$$

As $\varepsilon \to 0$, $X(t)$ tends to a diffusion process $\bar{X}(t)$ whose infinitismal generator has drift and diffusion coefficients defined by

$$\tilde{b} = a_0 + \tilde{h}_1 - A_{12} A_2^{-1} \tilde{h}_2 + \tilde{h}_3, \tag{5.4}$$

and

$$\tilde{A} = B_0 S(0) B_0' \tag{5.5}$$

where

$$\tilde{h}_1 = \mathrm{tr}[D_i' B_0 W], \tag{5.6}$$

$$\tilde{h}_2 = \mathrm{tr}[E_i' B_0 W], \tag{5.7}$$

$$\tilde{h}_3 = -\mathrm{tr}[F_i' B_0 W B_2' A_2^{-1}] \tag{5.8}$$

It can be easily verified that this is exactly our reduced-order model when $\gamma \to \infty$ $(\frac{\mu}{\varepsilon} \to 0)$.

5.2. Conclusions and Future Research:

In this thesis, a class of nonlinear singularly perturbed systems driven by wide-band noise has been considered. It has been shown that the probabilistic behavior of the slow variables can be predicted from a reduced-order diffusion model which has been derived explicitly. The use of the reduced-order model in studying stability of the full-order system, has been examined. Then the possible application of the reduced-order model in control problems has been considered. Stabilizing state feedback and output feedback controls have been designed, where for the latter a nonlinear stochastic observer for the reduced-order model has been used.

The importance of these results is that of getting an explicit form of a reduced-order model, where the use of this reduced-order model may lead to considerable simplification in solving problems. It is obvious, for example, that any simulation involving the full-order systems will, computationally, be much more difficult than working with the reduced-order model, because of the higher dimension and the ill-conditioning caused by the small parameters ε and μ. The reduced-order model, being a Markov model, has an important significance in its own, and this stems from the fact that the mathematical theory of stochastic differential equations is concerned mainly with the study of Itô equations and the associated Markov processes. This Markov model and its dependence, in general, on the ratio

$\frac{\epsilon}{\mu}$ contradicts the engineering practice of neglecting parasitic elements when writing differential equations representing physical systems as we pointed out in the above discussion. What this reduced-order model tells us is that before neglecting any parasitic elements one has to study their relationship with the wide-band noise.

An important step towards the effective use of the reduced-order model has been explored in chapters 3 and 4, where stability properties of the non-Markov full-order system has been established from stability properties of the reduced-order model, and stabilization of the full-system via the use of the Markov reduced-order model has been, also, estabilished.

Several additional work and topics are worth of future study, among those are:

1 . Studying nonlinear systems which are more general than the one that has been considered here and represented by (2.1) and (2.2), in the sense that the system may be non-linear in y and also nonlinear in the driving noise. In this case different stability conditions have to be imposed to guarantee the stability of the boundary layer system. The time varying case may also be considered. We studied here the case when the input noise is bounded and satisfies a certain mixing condition, so one may consider the case when the noise is unbounded in addition to some different conditions other than the mixing one, for example, the case

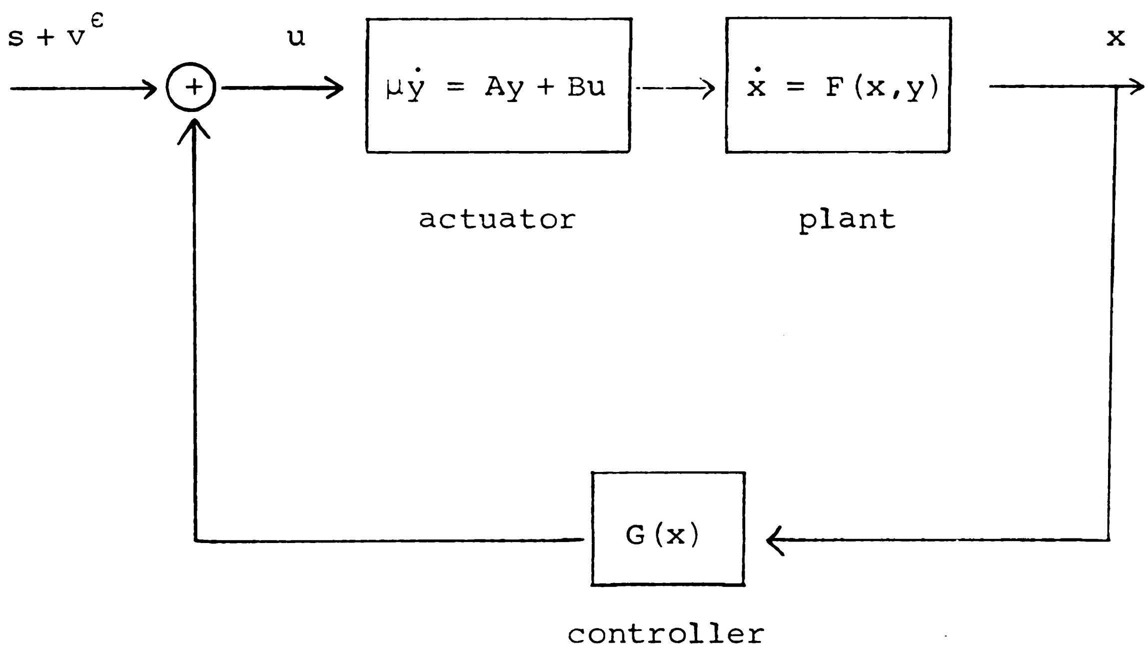
when $v(t)$ is the output of a linear system driven by white noise.

2. Studying systems of the form (2.1) and (2.2) but with $\sqrt{\mu}$ multiplying the coefficient $B_2(x)$ in (2.2). This case does not have the trouble caused by the fast variable $y(t)$ which, in our case, tends to white noise as μ and $\epsilon \to 0$, [c.f. 27]. One may also be able to say more about the asymptotic behavior of $y(t)$ as μ and $\epsilon \to 0$.

3. Studying the possibility of obtaining a near optimal control by optimizing an appropriate cost function for the corresponding reduced-order model. One may also consider an approach to the output feedback problem different from the one that has been studied in chapter 4 of this thesis. The suggested approach is as follows: 1) Design a stabilizing control law for the open-loop reduced-order model based on state feedback, assuming the states $\bar{x}$, of that system, can be measured. 2) Construct an observer which generates a vector $\hat{x}$ such that for any u the error $\bar{x}(t)-\hat{x}(t) \to 0$ as $t \to \infty$ in some stochastic sense. 3) Apply the previously determined control law to $\hat{x}(t)$ then a stability result may be established for the augmented system including the states $\begin{pmatrix} \bar{x} \\ \hat{x} \end{pmatrix}$. If this scheme works for the open-loop reduced-order model then this control law may be applied to the full order model in a way similar to what we have established in chapter 4 or in a way similar to the above procedure. This suggested approach is well established for linear time

invariant systems [c.f. 37], and it is done in the spirit of the separation principle. A recent work following the above procedure has been done by [39] for deterministic nonlinear systems.

4. An important task for future research is identifying physical systems that can be treated using the results that has been established in this thesis. A conceivable class of a systems, may be represented by the following blockdiagram



where s is some reference input and v^{ε} is a wide-band stochastic process.

BIBLIOGRAPHY

1. P.V. Kokovtovic, R.E. O'Malley, Jr. and P. Sannuti, "Singular Perturbation and Order Reduction in Control Theory--An Overview," Automatica, Vol. 12, pp. 123-132, 1976.

2. J.H. Chow and P.V. Kokotovic, "A Two-Stage Lyapunov-Bellman Feedback Design of a Class of Nonlinear Systems," IEEE Trans. Automatic Control, Vol. AC-26, pp. 656-663, June 1981.

3. A. Saberi and H. Khalil, "Quadratic-Type Lyapunov Functions for Singularly Perturbed Systems," IEEE DCD Conference, San Diego, California, Dec. 1981.

4. K. Ito, "Stochastic Differential Equations in a Differentiable Manifolds", Nagoya Math. J. 1(1950), 35-47.

5. K. Ito, "On Stochastic Differential Equations", Mem. Amer. Math. Sco. 4, 1951.

6. E.B. Dynkin, "Markov Processes", Academic Press, New York, 1965.

7. I.I. Grihman and A.V. Skorohod, "Stochastic Differential Equations" Springer-Verlag Berlin Heidelberg New York, 1972.

8. D.W. Stroock and S.R.S. Varadhan, "Multidimensional Diffusion Processes," Springer-Verlag New York Inc., 1979.

9. P. Billingsley, "Convergence of Probability Measures," John Wiley, New York, 1968.

10. T.G. Kurtz, "Semigroups of Conditioned Shifts and Approximation of Markov Processes, Ann. Probability, 1975, Vol. 3, pp. 618-642.

11. G.C. Papanicolaou and W. Kohler, "Asymptotic Theory of Mixing Stochastic Ordinary Differential Equations," Comm. Pure Appl. Math., Vol. 27, pp. 641-668, 1974.

12. G. Blankenship and G.C. Papanicolaou, "Stability and Control of Stochastic Systems with Wide-Band Noise Disturbance," SIAM J. Appl. Math., Vol. 34, pp. 437-476, May 1978.

13. H.J. Kushner, "Jump-Diffusion Approximations for Ordinary Differential Equations with Wide-Band Random Right Hand Side," SIAM J. Control and Optimization, Vol. 17, pp. 729-744, November 1979.

14. H.J. Kushner, "A Martingale Method for the Convergence of a Sequence of Processes to a Jump-Diffusion Process," Z. Wahrscheinlichkeitstheorie Verw. Gebiete, 53, 207-219, 1980.

15. H.J. Kushner and Y. Bar-Ness, "Analysis of Nonlinear Stochastic Systems with Wide-Band Inputs," IEEE Trans. Automatic Control, Vol. AC-25, pp. 1072-1078, 1980.

16. F. Kozin "A survey of Stability of Stochastic Systems," Automatica, J. IFAC, 5, 1969, pp.95-112.

17. E. Lukacs "Stochastic Convergence," Academic press, 1975.

18. T. Tarn and Y. Rasis, "Observers for Nonlinear Stochastic Systems," IEEE Trans. Automatic Control, AC-21, pp. 441-448 (1976).

19. E. Wong and M. Zakai, "On the Relation between Ordinary and Stochastic Differential Equations," Internat. J. Engin. Science, 3, pp. 213-229, 1965.

20. H.J. Kushner, "Stochastic Stability and Control," Academic press, New York, 1967.

21. A. Haddad, "Linear Filtering of Singularly Perturbed Systems," IEEE Trans. Automatic Control, AC-21, pp. 515-519, 1976.

22. A Haddad, P. Kokotovic, "Stochastic Control of Linear Singularly Perturbed Systems," IEEE Trans. Automatic Control, AC-22, pp. 815-821, 1977.

23. D. Taneketzis and N. Sandell, "Linear Regulator Design for Stochastic Systems by Multiple Time-Scale Method," IEEE Trans. Automatic Control, AC-22, pp. 615-621, 1977.

24. H. Khalil, "Control of Linear Singularly Perturbed Systems with Colored Noise Disturbance," Automatica, Vol. 14, pp. 153-156, 1978.

25. H. Khalil, A. Haddad, and G. Blankenship," Parameter Scaling and Well-Posedness of Stochastic Singularly Perturbed Control Systems," Proc. Twelfth Asilomar Conference, Pacific Grove, California, Nov. 6-8, 1978.

26. H. Khalil and Z. Grajic, "Near-Optimum Regulators for Stochastic Linear Singularly Perturbed Systems," The 21st IEEE Conference on Decision and Control, Orlando, Florida, pp. 1317-1321, December 1982.

27. A. Bensoussan, "Singular Perturbation Results for a Class of Stochastic Control Problems," IEEE Trans. Automatic Control, Vol. AC-26, pp. 1071-1080, October 1981.

28. G. Blankenship, "On the Separation of Time Scales in Stochastic Differential Equations," Proc the 7th IFAC Congress, pp. 937-944, Helsinky, 1978.

29. Hoppensteadt, "Properties of Solutions of Ordinary Differential Equations with Small Parameters," Comm. Pure Appl. Math., 24, pp. 807-840, 1971.

30. V.D. Razvig, "Reduction of Stochastic Differential Equations with Small Parameters and Stochastic Integrals," Int. J. Control, Vol. 28, pp. 707-720, 1978.

31. H.J. Kushner, "Asymptotic Distribution of Solutions of Ordinary Differential Equations with Wide-Band Noise Inputs: Approximate Invariant Measures," Stochastics, 6, pp. 259-277, 1982.

32. A.G. Korn and T.M. Korn, "Mathematical Handbook for Scientists and Engineers," McGraw Hill Co., New York, 1968.

33. G.C. Papanicolaou, D. Stroock and S.R.S. Varadhan, "Martingale Approach to Some Limit Theorems," Statistical Mechanics and Dynamical Systems, Duke Turbulence Conference, M. Reed, ed., Duke University Mathematics Series, Vol. 3, Durham, NC, 1977.

34. J.J. Levin and N. Levinson, "Singular Perturbations of Nonlinear Systems of Differential Equations and an Associated Boundary layer Equation," J. Ration. Mech. Anal., Vol. 3, pp. 247-270, 1954.

35. A.P. Sage and J.L. Melsa, "Estimation Theory with Applications to Communications and Control," McGraw-Hill Book Co. 1979.

36. P. Hall and C.C. Heyde, "Martingale Limit Theory and its Application," Academic Press, 1980.

37. H. Kwakernaak and R. Sivan, "Linear Optimal Control Systems," John Wiley and Sons, Inc., 1972.

38. J. Chow and P. Kokotovic, "Two-Time Scale Feedback Design of a Class of Nonlinear Systems," IEEE Trans. Automatic Control, AC-23, pp. 438-443, 1978.

39. M. Vidyasagar, "On the Stabilization of Nonlinear Systems Using State Detection," IEEE Trans. on Automatic Control, Vol. AC-25, No 3, June 1980.