Abstract

Title of Dissertation: Optimal Control of Partially Observed Markov Chains Admitting Strong and Weak Interactions

Sheldon Ira Wolk, Doctor of Philosophy, 1988

Dissertation directed by: John S. Baras, Professor of of Electrical Engineering, College of Engineering, University of Maryland

Measurements of the state of a continuous time Markov Chain having an infinitesimal generator that is a function of a small parameter, ϵ, are corrupted by additive white noise. Two cases are considered: (1) Regular perturbations, in which the functional dependence of the infinitesimal generator on ϵ models the effects of a weak signal in noise, and (2) Singular perturbations, in which the small parameter ϵ is introduced to model the behavior of a Markov Chain having two time scales.

The problem of finding approximations to the optimal control of these systems under an integral cost criterion is described. Changing probability measure, the problem is transformed into the equivalent problem of optimal control on a fully observed state space under a linear cost criterion. The dynamics of the transformed problem are described by the Zakai equation for the unnormalized conditional probability which is bilinear in form.

The regular perturbation problem is shown to decompose into a family of optimization problems, the optimizing control for each providing higher order approximations to the optimal control for the original problem. The first problem in this family of optimization problems is deterministic. All other problems have quadratic costs and linear dynamics.

The singular perturbation problem is shown to decompose into two optimization problems. The first of these optimization problems, the "limit problem", captures the effects of slow processes in the original system. The solution to this limit problem approximates the value function for the original problem by $O(\epsilon)$. The rest of the decomposition is the so-called "fast problem", which captures effects of processes that change quickly with respect to those in the limit problem.

Optimal Control of Partially Observed Markov Chains Admitting Strong and Weak Interactions

by

Sheldon Ira Wolk

Dissertation submitted to the Faculty of the Graduate School
of the University of Maryland in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
1988

Advisory Committee:

Professor John Baras, Chairman/Advisor
Professor William Levine
Professor Perinkulam Krishnaprasad
Professor Carlos Berenstein
Associate Professor Evaggelos Geraniotis

ACKNOWLEDGEMENT

I would like to thank my advisor, Dr. John Baras, for his insights and guidance over the past several years. I am also grateful to my professors and fellow graduate students that made my stay at the University of Maryland both enjoyable and rewarding. Finally, special thanks to my family, friends, and coworkers for their continued patience and encouragement.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this thesis we demonstrate how some large scale stochastic control problems involving partial observations of the state space and containing a small parameter, ϵ, can be approximated so as to make them more amenable to solution. In particular, the problem of optimally controlling a partially observed (continuous time) Markov chain admitting two scales of transition probabilities is considered. The small parameter ϵ characterizes the ratio of these time scales.

Problems such as those described above result from discretizing continuous time Markov processes admitting two time scales. Specifically, two sets of dynamical equations are considered. Those associated with regular perturbations:

$$\begin{aligned} dx_t^1 &= \epsilon f(x^1, x^2)dt \\ dx_t^2 &= b(x^1, x^2, u)dt + dw_t \\ x_0^1 &= \xi^1 \\ x_0^2 &= \xi^2 \end{aligned}$$

and those associated with singular perturbations:

$$\begin{aligned} dx_t^1 &= f(x^1, x^2)dt \\ dx_t^2 &= \frac{1}{\epsilon} b(x^1, x^2, u)dt + \frac{1}{\sqrt{\epsilon}} dw_t \end{aligned}$$

$$
\begin{aligned}
x_0^1 &= \xi^1 \\
x_0^2 &= \xi^2
\end{aligned}
$$

The stochastic control problems of minimizing the cost functional

$$J = E\{\int_0^T l(x_t^1, x_t^2)dt\}$$

subject to one of the above sets of dynamics, over a set of controls that have available only the observations

$$y_t = g(x_t^1) + \sqrt{N_0}v_t$$

are approximated by replacing the dynamics with continuous time Markov chains. The resulting Markov chains have two time scales: the "slow" part of the system (i.e. x^1) giving rise to "weak" interactions and the "fast" part of the system (i.e. x^2) giving rise to "strong" interactions. The behavior of the system over a time horizon that is sufficiently long that the weak interactions cannot be neglected is examined.

The motivation for using an approximation for the original diffusion equations is that of reducing the dimension of the state space. It is well known that the problem of optimal control of a partially observed stochastic system can be transformed into a problem in which there is complete information. This equivalence is possible since the probability of the state given the observations is a sufficient statistic [6]. The output of the filter that realizes the conditional probability becomes the state of the transformed problem. For the diffusion equations described above, the Zakai equation describing this filter is infinite dimensional [30], [29], and [4]. The consequence of the infinite dimensionality of the transformed state equations is that in solving the problem of time scale separation and control under an integral cost criterion, one is forced to use equations such as Mortensen's equation, a dynamic programming equation in function space.

In contrast, by first approximating the original dynamics by a continuous time Markov chain, the solution to the Zakai equation that then results lies in a finite dimensional vector space. The approximation of the original dynamics by a Markov chain, therefore, allows the use of the Bellman equation (or of Maximum Principle techniques) in solving the problem of finding time scale decomposition.

Although the conditional probability is a sufficient statistic, and can be used as the state of the transformed problem [6] and [36], in this paper, the unnormalized conditional density is used. This is desirable since the unnormalized density is also a sufficient statistic, and, after a change in probability measures, the resulting Zakai equation is bilinear [16], and the resulting cost function is linear [1] and [17]. These relatively simple forms make the subsequent time scale decompositions more straightforward.

The problem of the optimal control of fully observed Markov chains has been studied by many authors. Applications include management of hydroelectric dams [13], and queueing network models [12]. These and other applications demonstrate the need for reduced order approximations of large scale Markov chains. Among the methods of approximation that are used is a perturbational decomposition-aggregation method [13]. This method recognizes groups of strongly interacting states, and treats the weak interactions between these groups as perturbations. Thus, in the short term, there is a decomposition; and, over longer time periods, as the weak interactions become significant, the coupled states can be replaced by single aggregate states. The resulting optimization algorithms are hierarchical, with fast subsystems coordinated at a slower aggregate level [13] and [31].

In this thesis, the problem of partially observed Markov chains is considered. As in the case of fully observed Markov chains, the goal is to determine groups of strongly interacting states and the weak interactions among these groups. The difference between the fully and the partially observed problem is that for the partially observed problem the dynamics are described by the resultant Zakai equation. It is thus the manifestations of the strong and the weak interactions of the states of the Markov chain on the Zakai equation that determines the form of the time scale decomposition.

The problem of optimally controling a partially observed Markov chain admitting strong and weak interactions is motivated by the problem of an evader being pursued by a much faster, much more maneuverable pursuer. This pursuer has a sensor that provides a continuously updated estimate of the evader's location. The estimate is used to guide the pursuer towards the evader according to a fixed, deterministic strategy. Both the pursuer's sensor law and the pursuer's guidance law are known to the evader. A block diagram of the pursuer's dynamics is given in the following figure.

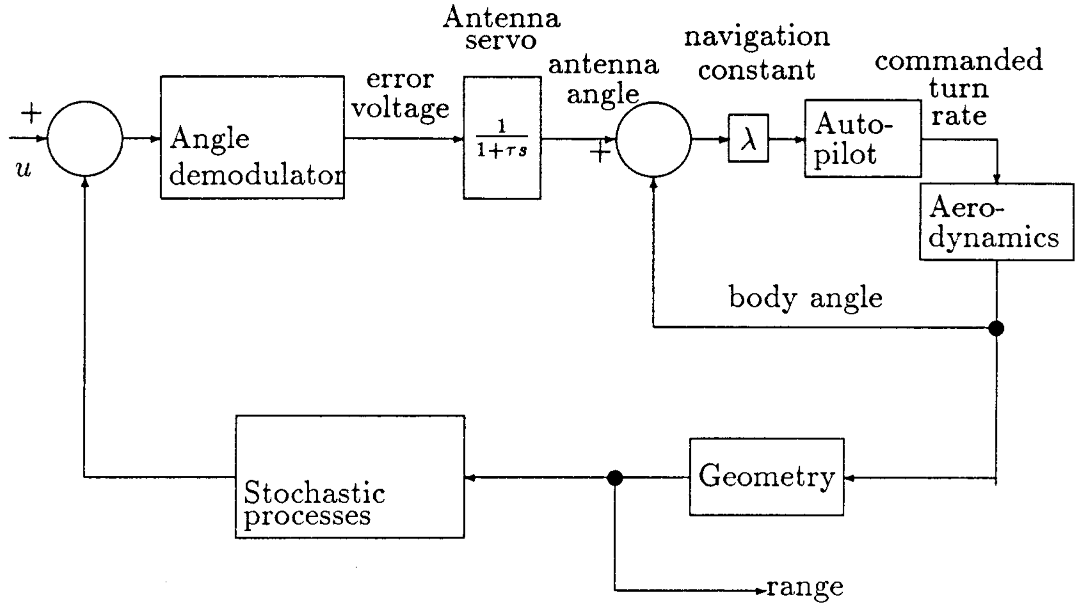


Figure 1.1: Pursuer dynamics

The state of the pursuer's sensor is affected by additive noise. It can also be influenced by the evader by means of a jamming signal. The evader's speed is so small and its agility so poor relative to that of the pursuer that if the evader is to survive, it is largely through the use of jamming signals that it must do so.

The evader also has a sensor. Its sensor is unaffected by the pursuer, and provides a noisy measurement of the pursuer's location in space. It is only through the use of this sensor that the evader can determine the effectiveness of the jamming signal it is applying.

For reasons of stability, the time associated with the determination, by the pursuer, of a position estimate is at least an order of magnitude shorter than that associated with the implementation of a turn rate command (by the pursuer). This is the reason for the time scaling that appears in the dynamics.

In order to simplify the problem further, let the evader be stationary and located at the origin and consider a two state model of the dynamics. One state corresponds to the pursuer's position in space. The other corresponds to the state of its sensor. The single observation equation models the observations available to the evader. That is, using the notation introduced above,

$$x^1 = \text{range} \tag{1.1}$$

$$x^2 = \text{antenna angle} \tag{1.2}$$

$$u = \text{jamming signal} \tag{1.3}$$

$$y = \text{evader's observations} \tag{1.4}$$

Refering then to the block diagram of Figure 1.1, the range, or equivalently, the position of the pursuer is a slowly varying quantity which is essentially a low pass version of the pusuer's estimate of the evader's position. Such guidance laws are also affected by nonlinearities inherent in the system. These nonlinearities typically arise due to limitations in the maneuverability of the pursuer. The time rate of change of x^1 is then a nonlinear function of both x^1 and x^2.

The pursuer's estimate of the evader's position is itself the output of a lowpass filter. This filter is driven by the relative position of the evader to the pursuer and by the evader's jamming signal, after both are processed through memoryless nonlinearities that model the effects of geometry and angle demodulation. As such, the time rate of change of x^2 is a function of x^1, x^2, and u. It is also influenced by an additive noise term.

The filter that produces the pursuer's estimate of the evader's position must be wideband in order to maintain stability. The variable x^2 is thus quickly varying with respect to x^1. (Physically, the time constant of the antenna servo in Figure 1.1 is short compared to that of the autopilot and aerodynamics). It is because the ratio of the time constants of these filters is sufficiently large that a time scale decomposition of this problem is appropriate.

The evader's observations, y are corrupted measurements of x^1. They are, therefore, slowly varying. The jamming signal, the control u here, is a function of $y_s, s \leq t$, at time t. This expresses the fact that the evader uses its sensor to measure the effectiveness of its jamming and modify its jamming accordingly.

It is assumed that the pursuer, although much more maneuverable than the evader, is not capable of turning around once having passed the evader, and is in fact, not capable of any lateral accelerations that can generate lateral velocities that are comparable to its longitudinal velocity within the time horizon of interest. Thus, we may approximate the game as having a fixed time horizon, as opposed to one that ends when the evader has reached some terminal set in the state space.

The problem is to maximize the miss distance of the pursuer. That is, to maximize the value of one of the state variables, x^1, at the terminal time.

We now provide an intuitive (formal) analysis of the two problems described at the beginning of this section, so as to identify the major ingredients of the solution and the associated mathematical difficulties.

First let us consider the regular perturbation problem

$$dx_t^1 \;=\; \epsilon f(x_t^1, x_t^2)dt \tag{1.5}$$

$$dx_t^2 = b(x_t^1, x_t^2, u)dt + dw_t \tag{1.6}$$

$$x_0^1 = \xi^1 \tag{1.7}$$

$$x_0^2 = \xi^2 \tag{1.8}$$

$$y_t = g(x_t^1) + \sqrt{N_0}v_t \tag{1.9}$$

and we want to maximize

$$\max E\{\int_0^T l(x_t^1, x_t^2)dt\} = \max_u J(u) \tag{1.10}$$

Now observe that for $\epsilon = 0$, equation (1.5) results in $x^1 =$ constant. So in this formulation the first order approximation for x_t^1 is that it is constant, a rather uninteresting situation. So the emphasis is more on the analysis of equation (1.6) as higher order approximations of x_t^1 are considered. If we differentiate equation (1.9) we obtain

$$dy_t = \frac{\partial g}{\partial x^1}(x_t^1)\epsilon f(x_t^1, x_t^2)dt + \sqrt{N_0}dv_t \tag{1.11}$$

$$= \epsilon h(x_t^1, x_t^2)dt + \sqrt{N_0}dv_t \tag{1.12}$$

so that the observation depends weakly on both x_t^1 and x_t^2. For $\epsilon = 0$, the observation is just noise and it will be of little value in setting the control u_t. So again the emphasis is on computing higher order control laws which are based on better approximations of the observations. From this point of view one can think of this problem as a "weak signal in noise" (or low signal to noise ratio) problem. This interpretation is apparent from equation (1.12). Thus this model is useful in understanding the practical problem described earlier, when the evader's sensor is of poor quality.

The approach we have taken in this thesis is to first pass to a Markov chain approximation of the problem, so as to overcome the computational difficulties of infinite-dimensional state spaces. Let us briefly see what would be involved otherwise. This stochastic control problem is partially observed, so a sufficient statistic for the controller is given by the unnormalized conditional density of (x_t^1, x_t^2) given $y_s, s \leq t$, [17]. Let us denote this density by $Q(x^1, x^2, t)$. What this means is that instead of considering u_t a function of $y_s, s \leq t$, we can consider it a function of $Q(\cdot, \cdot, t)$. So we need to compute on line the unnormalized density. It is known that $Q(\cdot, \cdot, t)$ is the solution of the stochastic partial differential equation

$$\begin{aligned}\frac{\partial Q(x^1, x^2, t)}{\partial t} &= \frac{1}{2}\frac{\partial^2}{(\partial x^2)^2}Q(x^1, x^2, t) - \frac{\partial}{\partial x^1}(\epsilon f(x^1, x^2)Q(x^1, x^2, t)) \\ &\quad -\frac{\partial}{\partial x^2}(b(x^1, x^2, u)Q(x^1, x^2, t)) \\ &\quad +\frac{1}{\sqrt{N_0}}Q(x^1, x^2, t)\epsilon h(x^1, x^2)\frac{dy}{dt} \qquad (1.13)\end{aligned}$$

$$Q(x^1, x^2, 0) = \Phi(x^1, x^2) \qquad (1.14)$$

where $\Phi(x^1, x^2)$ is a known function.

We note that as a parabolic partial differential equation this is degenerate, a complication we have to deal with following this route. Also note that for $\epsilon = 0$, equation (1.13) reduces to the Fokker-Planck equation for x_t^2, with x^1, u being treated as parameters. This is another way of saying that for $\epsilon = 0$, the sensor data have no value, in reducing uncertainty. One then would proceed to obtain a series (in ϵ) expansion of the solution of equation (1.13), which will provide higher order approximations to the density.

Regarding the cost of equation (1.10), one can rewrite it as

$$J(u) = E\{\int_0^T [\int l(x^1, x^2)Q(x^1, x^2, t)dx^1dx^2]dt\} \qquad (1.15)$$

The stochastic control problem described by equations (1.13) and (1.15) is now a fully observed stochastic control problem, but the state equation (1.13) is infinite dimensional. For $\epsilon = 0$, the problem reduces to finding an optimal density for x^2, so as to maximize $J(u)$, via controlling the solution of equation (1.13) from the drift term $b(x^1, x^2, u)$. In this situation x^1 becomes a dummy parameter (recall from a different view it is constant). Then one should proceed to obtain a series expansion (in ϵ) of the optimal performance measure.

Let us next consider the singular perturbation problem

$$
\begin{aligned}
dx_t^1 &= f(x_t^1, x_t^2)dt && (1.16)\\
dx_t^2 &= \frac{1}{\epsilon}b(x_t^1, x_t^2, u_t)dt + \frac{1}{\sqrt{\epsilon}}dw_t && (1.17)\\
x_0^1 &= \xi^1 && (1.18)\\
x_0^2 &= \xi^2 && (1.19)\\
dy_t &= \frac{\partial g}{\partial x^1}(x_t^1)f(x_t^1, x_t^2)dt + \sqrt{N_0}dv_t && (1.20)\\
&= h(x_t^1, x_t^2)dt + \sqrt{N_0}dv_t && (1.21)
\end{aligned}
$$

and again we want to maximize

$$\max E\{\int_0^T l(x_t^1, x_t^2)dt\} = \max_u J(u) \qquad (1.22)$$

The character of this problem is different from the previous one. To see this clearly, let us change time scale from t to $\tau = \frac{t}{\epsilon}$. So as $\epsilon \to 0$, $\tau \to \infty$, i.e. τ is the first time scale. In the first time scale equation (1.17) becomes

$$dx_\tau^2 = b(x_\tau^1, x_\tau^2, u_\tau)d\tau + dw_\tau \qquad (1.23)$$

while equation (1.16) becomes

$$dx_\tau^1 = \epsilon f(x_\tau^1, x_\tau^2)d\tau \qquad (1.24)$$

So as $\epsilon \to 0$, x_τ^1 will approach a constant, that is, x^1 is "slowly" varying with respect to x^2. Therefore, x_τ^1 can be treated as a constant parameter in equation (1.23). Furthermore, in this time scale y behaves as

$$dy_\tau = \epsilon h(x_\tau^1, x_\tau^2)d\tau + \sqrt{N_0}\sqrt{\epsilon}dv_t \qquad (1.25)$$

So, y is also "slowly" varying with respect to x_τ^2, and can be considered constant. But then since u_τ is a function of y, it can also be considered approximately constant in this time scale. So we can rewrite equation (1.23) as

$$dx_\tau^2 = B_{u,x^1}(x_\tau^2)d\tau + dw_\tau \qquad (1.26)$$

where we suppress the dependence of the drift on x^1 and u.

The situation is now clear. As $\epsilon \to 0$, $\tau \to \infty$ and in order for the whole problem to make sense x^2_τ must converge to a random variable, $\overline{x^2}$, with probability density $Q_{u,x^1}(\cdot)$, which depends on the control u. So $x^2_{(\cdot)}$ must be ergodic. Then the admissible controls are all $u(\cdot)$, functions of $y_s, s \leq t$, which make $x^2_{(\cdot)}$ ergodic.

In Q_{u,x^1}, u and x^1 are treated as parameters. If we write the Fokker-Planck equation for x^2_τ

$$\frac{\partial Q}{\partial \tau}(x^2,\tau) = \frac{1}{2}\frac{\partial^2}{(\partial x^2)^2}Q(x^2,\tau) - \frac{\partial}{\partial x^2}(b(x^1,x^2,u)Q(x^2,\tau)) \tag{1.27}$$

for $x^2_{(\cdot)}$ to be ergodic, equation (1.28) must have a unique solution for all x^1 and u (treated as parameters). There are of course conditions on $b(\cdot,\cdot,\cdot)$ which will guarantee this, primarily due to Khasminskii; an expected

$$\frac{1}{2}\frac{\partial^2}{(\partial x^2)^2}Q(x^2) - \frac{\partial}{\partial x^2}(b(x^1,x^2,u)Q(x^2)) = 0 \tag{1.28}$$

relationship since stability is involved. The probability density (or measure) $Q(x^2)$ is known as the "invariant" measure of x^2.

Now let us look at the problem of equations (1.16) through (1.22) in the "slow" time scale t. In this time scale, $x^2_{(\cdot)}$ will have converged to the random variable $\overline{x^2}$, and its effects will appear in equations (1.16) through (1.22). So the "slow" time scale problem is now

$$dx^1_t = f(x^1_t,\overline{x^2})dt \tag{1.29}$$

$$dy_t = h(x^1_t,\overline{x^2})dt + \sqrt{N_0}dv_t \tag{1.30}$$

$$J(u) = E\{\int_0^T l(x^1_t,\overline{x^2})dt\} \tag{1.31}$$

This is a partially observed stochastic control problem, parameterized by the random variable $\overline{x^2}$, with known probability density function $Q_{u,x^1}(\cdot)$, which depends on the control u. So we know that we would have to <u>average</u> the effects of $\overline{x^2}$. The problem, though is not so simple because we have in hand a combined averaging and optimization problem; and the correct averaging is not obvoius. Again one can turn this partially observed problem into a fully observed problem and then perform the required asymptotic analysis on the resulting infinite dimensional fully observed problem. To avoid this, in this thesis we followed a different approach. We first associate with equations (1.16) through (1.21) an approximate Markov chain problem, and then treat this by the method just described. In the end we interpret the results for equations (1.16) through (1.21).

The main contribution of this thesis for the regular perturbation problem is the derivation of a family of optimization problems that give an arbitrary order of ϵ approximation to the actual value function. The first of these optimization problems is deterministic. Each of the others is the problem of minimizing a quadratic cost function subject to linear (and stochastically driven) dynamics.

For the singular perturbation problem, the main contribution of this thesis is the reduction in dimensionality that results from the derived time scale decomposition. It is shown that the stochastic control problem can be reduced to the problem of minimizing an aggregate cost function subject to the dynamics of an aggregate process. The dimension of this aggregate process is that of the number of elements in the grid one uses to approximate the "slow process" (the process x_t^1); a considerable reduction from the dimension of the Markov Chain that approximates both the slow and the fast processes.

This thesis also demonstrates the application of he composite control to obtain an $O(\epsilon)$ approximation to the optimal value function for the singular perturbation problem. The composite control consists of the sum of the minimizing control for the "limit problem" and a corrector term, which is a function of the difference between the value of the "fast" variables and their values in the limit.

We begin with an introductory section that includes some definitions and theorems about stochastic differential equations and the corresponding differential generator. These results are basic to later sections. The diffusion processes that define the plant, the observation equation, and the cost are then described. In the following section the Markov chain approximation to the dynamics is given, a change in probability measure is introduced, and the resulting Zakai equation derived. The cost under this new probability measure is then derived, and, in the following section, the equivalent optimal control problem is described. There are two cases of interest: regular perturbations and singular perturbations.

Sections 5 and 6 outline theorems that are used in subsequent sections. Section 5 contains results dealing with the Maximum Principle and Dynamic Programming, and the relationship between them. Section 6 discusses Volterra expansion for bilinear systems.

The major results of this paper for the regular perturbation problem are contained in Section 7. In that section, using Maximum Principle arguments, an asymptotic expansion reveals the nature of the decomposition for this problem.

The major results for singular perturbations are contained in Section 8. First a transformation is presented which separates the problem into "fast" and "slow" states. The limit problem is then presented. Using an asymptotic analysis, it is then shown that a composite feedback provides a near optimal solution to the problem.

Chapter 2

Stochastic Calculus

The systems under consideration in this paper are described by stochastic differential equations, that is, differential equations driven by Gaussian white noise. The solution of these diffenential equations are defined in terms of a stochastic integral. This section provides the definition of a stochastic (Ito) integral and its relationship to the stochastic differential equation.

The rules of ordinary calculus are not the same as those required for use with the Ito integral. Thus, Ito's lemma, the analog of the fundamental theorem of calculus, is included below.

A stochastic process, X_t is said to be a Gaussian process if every finite linear combination of the form $\sum_{i=1}^{N} \alpha_i X_i$ is a Gaussian random variable. A Brownian motion (or Wiener) process, W_t, is a Gaussian process with zero mean and autocorrelation given by $E\{W_t W_s\} = \min(t, s)$.

Let (Ω, A, P) be a fixed probability space. Let $\{A_t, -\infty < t < \infty\}$ be an increasing family of sub-σ algebras of A, and let $\{W_t, -\infty < t < \infty\}$ be a Brownian motion process such that for each s, the aggregate $\{W_t - W_s, t \geq s\}$ is independent of A_s and W_t is A_t measurable for each t.

Let W_t be a Brownian motion process The stochastic integral, $I(\phi)$, is represented as

$$I(\phi) = \int_a^b \phi(\omega, t) dW(\omega, t) \tag{2.1}$$

Since W_t is not of bounded variation equation (2.1) cannot be interpreted in the usual sense. Assuming that

1. The function ϕ is jointly measurable in (ω, t). (with respect to A in ω and with respect to Lebesgue measure in t). For each t, ϕ_t is measurable with respect to the underlying σ-algebra, A_t.

2. ϕ satisfies $\int_a^b E|\phi_t|^2 dt < \infty$

then the stochastic integral can be defined as follows:

1. If ϕ is an (ω, t)-step function, i.e. if there exist times $t_0, t_1, \ldots, t_n$ independent of ω, such that $a = t_0 < t_1 < \ldots < t_n = b$ and $\phi(\omega, t) = \phi_i(\omega)$ on the interval $[t_i, t_{i+1}), i = 0, \ldots, n-1$, then the stochastic integral is defined by

$$\int_a^b \phi(\omega, t) dW(\omega, t) = \sum_{i=0}^{n-1} \phi_i(\omega)[W(\omega, t_{i+1}) - W(\omega, t_i)] \tag{2.2}$$

2. If ϕ is not an (ω, t)-step function, it can be shown [37] that there exists a sequence of (ω, t)-step functions such that

$$||\phi - \phi_n||^2 = \int_a^b E\{||\phi(\cdot, t) - \phi(\cdot, t)||^2 dt\} \overset{n \to \infty}{\to} 0 \tag{2.3}$$

Since $\int_a^b \phi_n(\omega, t) dW(\omega, t)$ converges in quadratic mean as $n \to \infty$, and the limit does not depend on the chosen sequence of step functions, the stochastic integral can be defined by

$$\int_a^b \phi(\omega, t) dW(\omega, t) = \underset{n \to \infty}{\text{lim in q.m.}} \int_a^b \phi_n(\omega, t) dW(\omega, t) \tag{2.4}$$

where $\{\phi_n\}$ is any sequence satsifying equation (2.3)

Having thus defined the stochastic integral, the definition of a stochastic differential equation will be given. In this paper a stochastic differential equation will be an equation of the form

$$dX(\omega, t) = m(X(\omega, t), t)dt + \sigma(X(\omega, t), t)dW(\omega, t) \tag{2.5}$$

with initial condition $X_a = X$. Equation (2.5) is simply a symbolic way of writing

$$X_t = X_a + \int_a^t m(X_s, s)ds + \int_a^t \sigma(X_s, s)dW_s \tag{2.6}$$

where the stochastic integral $\int_a^t \sigma(X_s, s)dW_s$ was defined above.

An operator, $\mathcal{L}$, is next defined which, when applied to a smooth function (say $V(X_t)$), has, in an average sense a property analogous to a derivative. That is, one recovers $V(X_t)$ by averaging the integral of $\mathcal{L}V(X_t)$. This operator is called the differential generator of the process X_t. Although the lemma below is stated for a scalar function V, it can be extended to vector functions in the obvious way [23].

Lemma 1 (Differential Generator [23]) *Let $V(x)$ be uniformly bounded and have uniformly bounded first, second, and third derivatives. Let X_t satisfy*

$$X_t = X_0 + \int_0^t f(X_s)ds + \int_0^t \sigma(X_s)dZ_s \tag{2.7}$$

where $f(x)$ and $\sigma(x)$ satsify a Lipschitz condition. Define the operator

$$\mathcal{L} = f(x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2} \tag{2.8}$$

Then

$$E_x\{V(X_t) - V(X_0)\} = E_x\{\int_0^t \mathcal{L}V(X_s)ds\} \tag{2.9}$$

$\mathcal{L}$ is known as the differential generator of the process X_t.

If $V(x,t)$ is sufficiently smooth, then

$$\frac{1}{\Delta}[E_{x,t}\{V(X_{t+\Delta}, t+\Delta) - V(x,t)\}] \rightarrow \frac{\partial V(x,t)}{\partial t} + \mathcal{L}V(x,t) \tag{2.10}$$

as $\Delta \rightarrow 0$, where $E_{x,t}$ is the expectation given that $X_t = x$.

Lemma 2 (Ito's Lemma [23]) *Let Z_t be a vector of independent Wiener processes, and let $f(x)$ and $\sigma(x)$ satisfy a vector Lipschitz condition. Let $dx =$*

$f(x)dt + \sigma(x)dZ$, $\sigma\sigma^T = a$, and let $V(x,t)$ have continuous first and second x-derivatives and a continuous t-derivative. Define

$$\mathcal{L} = \sum_i f_i(x) \cdot \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} \tag{2.11}$$

Then

$$dV(X_t,t) = \frac{\partial V(X_t,t)}{\partial t} \cdot dt + \mathcal{L}V(X_t,t)dt + V_x^T(X_t,t)\sigma(X_t)dZ_t \tag{2.12}$$

in the sense that, with probability one,

$$V(X_t,t) = V(X_0,0) + \int_0^t [V_s(X_s,s) + \mathcal{L}V(X_s,s)]ds + \int_0^t [V_x^T(X_s,s)\sigma(X_s)dZ_s] \tag{2.13}$$

If $E\{|V_x^T(X_t,t)\sigma(X_t)|\}^2$ is square integrable on the time interval of interest, then the stochastic integral component of $\int_0^t dV(X_s,s)$ has zero mean value and

$$E_x\{V(X_t,t) - V(X_0,0)\} = E_x\{\int_0^t \mathcal{L}V(X_s,s)ds\} \tag{2.14}$$

Proof: see [23]

Chapter 3

Setting of the problem

This thesis determines decompositions for the optimal control of partially observed Markov chains. Specifically, the chains considered here are approximations to regularly and singularly perturbed diffusions. In the first section below, the diffusions, the observations, the cost, and the admissible control laws are described. In the following sections the approximating Markov chains are found, the Zakai equations are described, a cost function for the transformed problem is derived, and finally, the transformed (equivalent) problem is stated.

Due to the similarity of the derivation of the equivalent problem for the regular and the singular perturbations cases, only for the regular perturbations case will the steps in this derivation be described in detail. A statement of the result for the singular perturbations case will be made following this derivation.

3.1 System description

Consider the functions

$$f(x^1, x^2) : \mathcal{R}^1 \times \mathcal{R}^1 \to \mathcal{R}^1 \tag{3.1}$$

$$b(x^1, x^2, u) : \mathcal{R}^1 \times \mathcal{R}^1 \times U \to \mathcal{R}^1 \tag{3.2}$$

$$l(x^1, x^2) : \mathcal{R}^1 \times \mathcal{R}^1 \to \mathcal{R}^1 \tag{3.3}$$

$$g(x^1) : \mathcal{R}^1 \to \mathcal{R}^1 \tag{3.4}$$

$$U \text{ is a metric space} \tag{3.5}$$

$$f, b, l, g \text{ are continuous and bounded} \tag{3.6}$$

Let

$$U_{ad} \text{ be a compact nonempty subset of } U \tag{3.7}$$

Let $(\Omega, \mathcal{A}, P)$ be a probability space on which there is a filtration $\mathcal{F}_t$ and on which two independent scalar standard Wiener processes, w_t and y_t are constructed. Also let ξ^1 and ξ^2 be scalar random variables, $\mathcal{F}_0$-measurable and independent of y_s and w_s, $\forall s \geq 0$.

An admissible control is a process u_t with values in U_{ad} and adapted to the observation process $A_t^y = \sigma(y_s, s \leq t)$. For such a control, the following state space equations can be solved

$$dx_t^1 = \epsilon f(x^1, x^2)dt \tag{3.8}$$

$$dx_t^2 = b(x^1, x^2, u)dt + dw_t \tag{3.9}$$

$$x_0^1 = \xi^1 \tag{3.10}$$

$$x_0^2 = \xi^2 \tag{3.11}$$

As indicated above, scalar observations of this system, corrupted by additive noise, are available. They are given by

$$y_t = g(x_t^1) + \sqrt{N_0} v_t \tag{3.12}$$

where v is a scalar white noise and N_0 is a constant.

The objective is to minimize the cost

$$J = E\{\int_0^T l(x_t^1, x_t^2)dt\} \tag{3.13}$$

3.2 Markov chain approximation

In general we can consider a diffusion process in $\mathcal{R}^n$, x_t, with infinitesimal generator A and with domain $\mathcal{D}(A) \subset C_b(\mathcal{R}^n)$, where the latter space is the space of continuous bounded functions in $\mathcal{R}^n$. For each space discretization mesh ρ, let $x^\rho \in \mathcal{R}^n$, denote the generic mesh points. We typically restrict x^ρ to a "cube" so that we can get a finite state space for the resulting Markov chain. So to each ρ there corresponds a finite Markov chain X_t^ρ, with state space $\mathcal{S}_\rho = \{1, 2, \ldots, N_\rho\}$.

To each state i in $\mathcal{S}_\rho$, we associate a point x_i^ρ in $\mathcal{R}^n$. To each function $f : D \to \mathcal{R}$ (where D is the "cube" considered above) there corresponds an approximation f_ρ, where $f(x_i^\rho) = f_\rho(i)$, for $i \in \mathcal{S}_\rho$. This operation is denoted by the projection P_ρ. That is, $P_\rho f = f_\rho$.

Let now $\mathcal{S}$ be a dense subspace of $C_b(\mathcal{R}^n)$ invariant under e^{At}. Let A_ρ be the transition probability matrix of the resulting Markov chain in $\mathcal{S}_\rho$; e.g. the matrices we constructed above. Then the way we constructed A_ρ above generalizes to the condition of requiring A_ρ to be such that

$$\lim_{\rho \to 0} \| A_\rho P_\rho f - P_\rho A f \| = 0 \tag{3.14}$$

for all f in any $\mathcal{S}$ with the properties as above. There are many ways to construct such A_ρ from A and are well known (see [24] for details).

We want to apply an implicit discretization scheme to the Zakai equation to obtain the discrete contraction semigroup

$$J_{\rho,\Delta}^k = (I - \Delta A_\rho)^{-k} \tag{3.15}$$

where k is an integer and Δ is the time step. Then by [32, Theorem 2.4] we have

Theorem 1 *If A_ρ is constructed on a finite dimensional space, $\mathcal{R}_\rho^N$, so that* *for any subspace $\mathcal{S} \subset C_b(\mathcal{R}^n)$ as above, and any $f \in \mathcal{S}$, we have*

$$\lim_{\rho \to 0} \parallel A_\rho P_\rho f - P_\rho A f \parallel = 0 \tag{3.16}$$

Then for any $f \in C_b(\mathcal{R}^n)$ and any $T > 0$,

$$\lim_{(\rho,\Delta) \to 0} \sup_{t \in [0,T]} \parallel J_{\rho,\Delta}^{[\frac{t}{\Delta}]} P_\rho f - P_\rho e^{At} f \parallel = 0 \tag{3.17}$$

The interpretation of Theorem 1 is that the solution of the Fokker-Planck equation for the Markov chain converges weakly to the solution of the Fokker-Planck equation for the diffusion.

Continuous time Markov chain approximations to the dynamics of equations (3.8) and (3.9) will be used in this thesis so as to take advantage of the simplicity of the form of the resulting Zakai equation. The infinitesimal generator for these approximations will be determined by finding a discrete time Markov Chain approximation (c.f. [24] and [28]) to the dynamics, dividing the transition probabilities for this chain by its time increment, and then taking the limit as that time increment approaches zero [20].

The discrete time Markov chain approximation that will be used is due to Kushner [24]. Given a diffusion of the form $X(t) = x + \int_0^t f(X(s))ds + \int_0^t \sigma(X(s))dw(s)$, Kushner seeks an approximating chain $\{\xi_n^\rho\}$ which, when suitably interpolated, converges weakly to the original process, $X(\cdot)$ as the state space step size, $\rho \to 0$. In particular, Kushner requires that at least the first and second order moments of the increment $\xi_{n+1}^\rho - \xi_n^\rho$, conditioned on ξ_n^ρ, be consistent with those of the diffusion and that the approximation not have large jumps. That is

$$\begin{aligned} E_x[\xi_{n+1}^\rho - \xi_n^\rho | \xi_n^\rho = y] &= f(y)\Delta t^\rho(y) + o(\Delta t^\rho(y)) & (3.18)\\ cov_x[\xi_{n+1}^\rho - \xi_n^\rho | \xi_n^\rho = y] &= \sigma\sigma^T \Delta t^\rho(y) + o(\Delta t^\rho(y)) & (3.19)\\ P_x[|\xi_{n+1}^\rho - \xi_n^\rho| \geq \varepsilon | \xi_n^\rho = y] &= o(\Delta t^\rho(y)) & (3.20) \end{aligned}$$

for some function (depending on ρ) which goes to zero as $\rho \to 0$.

The discrete time Markov chain will now be described. Following [24], the state space is divided into a uniform grid with increments of size ρ. As indicated above, Kushner allows the time increment to vary with both position in the state space and with gridsize. Defining the quantity Q_ρ by $Q_\rho = 1 + \rho b + \epsilon \rho f$, the time increment $\Delta t^\rho(x)$ is defined by $\Delta t^\rho(x) = \frac{\rho^2}{Q_\rho}$.

We begin by quantizing the state space of the diffusions x^1 and x^2 into uniform increments, ρ, and ordering the resulting grid points in accordance with Figure 3.1 (or equivalently, Table 3.1) below. In both the table and the figure, z denotes the state of the Markov Chain, subscripted values of z correspond to values of the ordered pair (x^1, x^2) in the obvious way.

Table 3.1: Ordering of states in the Markov chain approximation

value of x^1	value of x^2	value of z
β_0	ζ_0	z_0
	ζ_1	z_1
	$\vdots$	$\vdots$
	ζ_k	z_k
$\vdots$	$\vdots$	$\vdots$
β_m	ζ_0	$z_{m(k+1)}$
	$\vdots$	$\vdots$
	ζ_k	$z_{m(k+1)+k}$

Let $\nu^+(x) = \max\{0, \nu(x)\}$ and $\nu^-(x) = \max\{0, -\nu(x)\}$ for any function $\nu(x)$. Then the transitions rates for the regular and the singular perturbation problems are illustrated in Figures 3.2 and 3.3 respectively.

Then the infinitesimal generator of the Markov chain approximation to the regular perturbation problem is given by

$$Q_r^\epsilon = A + \epsilon B \tag{3.21}$$

and the infinitesimal generator of the Markov chain approximation to the singular perturbation problem is given by

$$Q_s^\epsilon = \frac{A}{\epsilon} + B \tag{3.22}$$

In both of these equations A is a block diagonal matrix which is the infinitesimal generator of a Markov chain characterizing the fast transitions. That is, it contains the transition rates associated with x^2 (with x^1 fixed).

$$A = \begin{pmatrix} A_0 & & & 0 \\ & A_1 & & \\ & & \ddots & \\ 0 & & & A_m \end{pmatrix} \tag{3.23}$$

where the matrices A_i are tri-diagonal, and are of the form

$$A_i = \frac{1}{\rho^2} \begin{pmatrix} \alpha_{11}^i & \alpha_{12}^i & & 0 \\ \alpha_{21}^i & \alpha_{22}^i & \ddots & \\ & \ddots & \ddots & \alpha_{k,k+1}^i \\ 0 & & \alpha_{k+1,k}^i & \alpha_{k+1,k+1}^i \end{pmatrix} \tag{3.24}$$

The first two elements in the first row of A_i are defined by

$$\alpha_{11}^i = -\frac{1}{2} - \rho b^+(z_{im}, u) \tag{3.25}$$

$$\alpha_{12}^i = \frac{1}{2} + \rho b^+(z_{im}, u) \tag{3.26}$$

For $0 < j < k + 1$, the non-zero elements of A_i are given by

$$\alpha_{j,j-1}^i = \frac{1}{2} + \rho b^-(z_{im+j-1}, u) \tag{3.27}$$

$$\alpha_{j,j}^i = -1 - \rho|b(z_{im+j-1}, u)| \tag{3.28}$$

$$\alpha_{j,j+1}^i = \frac{1}{2} + \rho b^+(z_{im+j-1}, u) \tag{3.29}$$

The last two elements of the last row of A_i are given by

$$\alpha_{k+1,k}^i = \frac{1}{2} + \rho b^-(z_{im+k}, u) \tag{3.30}$$

$$\alpha_{k+1,k+1}^i = -\frac{1}{2} - \rho b^-(z_{im+k}, u) \tag{3.31}$$

The slow transitions of the original diffusion are captured in the matrix B, which may also be considered to be an infinitesimal generator.

$$B = \begin{pmatrix} -B_0^+ & B_0^+ & & 0 \\ B_1^- & B_1 & \ddots & \\ & \ddots & \ddots & B_{m-1}^+ \\ 0 & & B_m^- & -B_m^- \end{pmatrix} \tag{3.32}$$

where the diagonal matrices B_i^+, B_i^-, and B_i are given by

$$B_i^+ = \frac{1}{\rho} \begin{pmatrix} f^+(z_{im}) & & 0 \\ & \ddots & \\ 0 & & f^+(z_{im+k}) \end{pmatrix} \tag{3.33}$$

$$B_i^- = \frac{1}{\rho} \begin{pmatrix} f^-(z_{im}) & & 0 \\ & \ddots & \\ 0 & & f^-(z_{im+k}) \end{pmatrix} \tag{3.34}$$

$$B_i = \frac{1}{\rho} \begin{pmatrix} -|f(z_{im})| & & 0 \\ & \ddots & \\ 0 & & -|f(z_{im+k})| \end{pmatrix} \tag{3.35}$$

This completes the approximation of the dynamics by a continous time Markov chain.

A derivation of the Zakai equation follows.

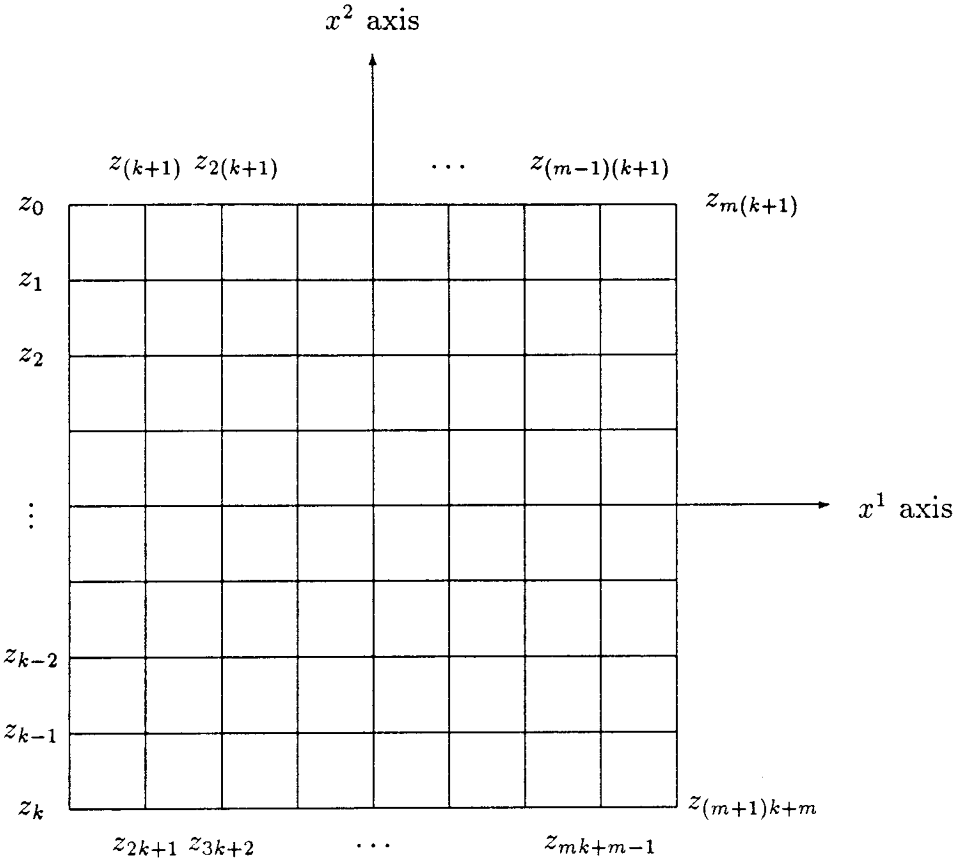


Figure 3.1: Discretized state space

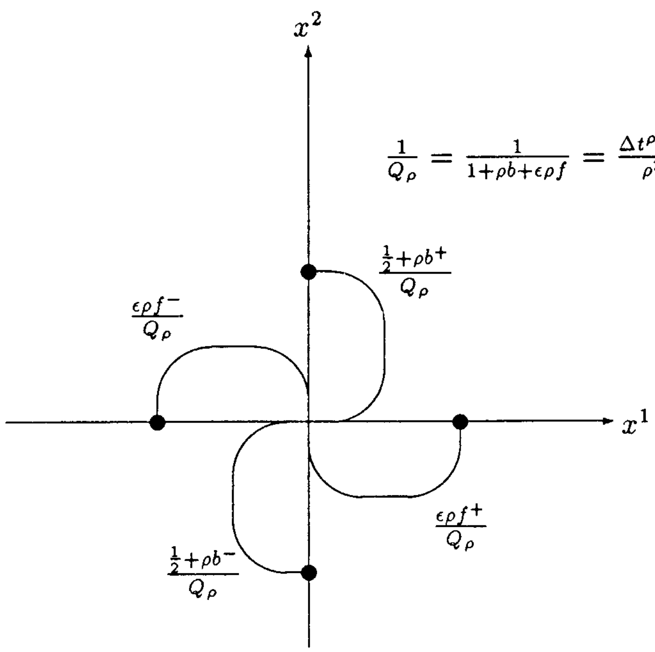


Figure 3.2: State transitions in the Markov chain approximation to the regular perturbation problem

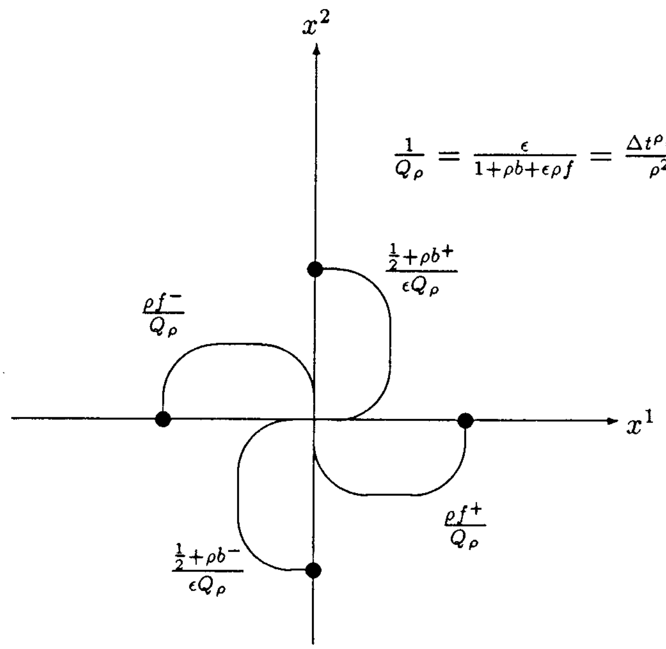


Figure 3.3: State transitions in the Markov chain approximation to the singular perturbation problem

3.3 Zakai equation

The purpose of this section is to outline the derivation of the Zakai equation for the unnormalized conditional density of the dynamics described by equation (3.21) given the observations of equation (3.12). The presentation in [37] is essentially followed.

Using Ito's rule, equation (3.12) can be rewritten as

$$dy_t = \epsilon \frac{\partial g}{\partial x^1} f(x^1, x^2) dt + \sqrt{N_0} dv_t \tag{3.36}$$

So as to put this equation in a more convenient form (i.e. one in which the coefficient of the noise is unity), the following functions will be defined

$$h(z) = \frac{\partial g}{\partial x^1} f(x^1, x^2) \tag{3.37}$$

and

$$\tau = N_0 t \tag{3.38}$$

In the τ time scale the observation equation is

$$dy_\tau = \frac{\epsilon}{N_0} h(z_\tau) d\tau + dv_\tau \tag{3.39}$$

In this new time scale, the transition rates of the Markov chain become

$$\frac{dp}{d\tau} = \frac{1}{N_0} p Q^\epsilon \tag{3.40}$$

where p is a vector, each element of which is the probability of the Markov chain being in the state corresponding to that element.

A new measure, P_0, is introduced on $(\Omega, \mathcal{A})$ by the formula

$$\frac{dP_0}{dP} = \exp(-\int_0^{N_0 T} h(z_\tau) dv_\tau - \frac{1}{2} \int_0^{N_0 T} |h(z_\tau)|^2 d\tau) \tag{3.41}$$

Under the measure P_0, y has components that are independent Brownian Motions. Let

A^y_τ denote the minimal σ algebra with respect to which $\{y_s, 0 \leq s \leq \tau\}$ is measurable

A_τ^z denote the minimal σ algebra with respect to which $\{z_s, 0 \le s \le \tau\}$ is measurable

A_τ be defined by $A_\tau^z \vee A_\tau^y$

And also defining the relative conditional expection [26, pp 10–11] of the Radon-Nikodym derivative,

$$\Lambda_\tau = E_0^{A_\tau}\{\frac{dP}{dP_0}\} \tag{3.42}$$

the sigma algebra of observations,

$$Y_0^\tau = \{Y_s, 0 \le s \le \tau\} \tag{3.43}$$

and the unnormalized conditional density of z_τ given Y_0^τ,

$$R_\tau(z_\tau, Y_0^\tau) = E_0(\Lambda_\tau | A(z_\tau) \vee A_\tau^y) \tag{3.44}$$

Then the (normalized) conditional expectation of a function of the state given the observations is given by (c.f. [37])

$$E^{A_\tau^y}\{l(z_\tau)\} = \frac{E_0^{A_\tau^y}\{\Lambda_\tau l(z_\tau)\}}{E_0^{A_\tau^y}\{\Lambda_\tau\}} \tag{3.45}$$

$$= \frac{\int_{\mathcal{R}^n} l(z) R_\tau(z, Y_0^\tau) P(dz, \tau)}{\int_{\mathcal{R}^n} R_\tau(z, Y_0^\tau) P(dz, \tau)} \tag{3.46}$$

where $n = (m+1)(k+1)$ and $P(dz, \tau) = P(z_\tau \in dz) = P_0(z_\tau \in dz)$. The above relationship is the justification for the interpretation of R_τ as the unnormalized conditional expectation.

In [37] it is shown that the unnormalized conditional density is given by

$$R_\tau(z_\tau, Y_0^\tau) = 1 + \frac{\epsilon}{N_0} \int_0^\tau \int_{\mathcal{R}^n} h(v) R_s(v, Y_0^s) P(dv, s | z, \tau) dY_s \tag{3.47}$$

The transition density function satisfies a Kolmogorov equation. Define

$$p(z, \tau | v, s) dz \triangleq P(z_\tau \in dz | z_s = v) \tag{3.48}$$

$$p(z, \tau) dz \triangleq P(z_\tau \in dz) \tag{3.49}$$

$$K_\tau(z, Y_0^\tau) \triangleq p(z, \tau) R_\tau(z, Y_0^\tau) \tag{3.50}$$

Then from equation (3.47)

$$K_\tau(z, Y_0^\tau) = p(z,\tau) + \frac{\epsilon}{N_0}\int_0^\tau \int_{\mathcal{R}^n} h(v)p(z,\tau|v,s)K_s(v, Y_0^s)dv dY_s \qquad (3.51)$$

But z_τ can assume only finitely many values, $Z_0, Z_1, \ldots, Z_{m(k+1)+k}$. So equation (3.51) can be rewritten as

$$K_\tau(Z_j, Y_0^\tau) = p(z_\tau = Z_j) + \frac{\epsilon}{N_0}\int_0^\tau \sum_{i=0}^{m(k+1)+k} h(Z_i)p(z_\tau = Z_j|z_s = Z_i)K_s(Z_i, Y_0^s)dY_s \qquad (3.52)$$

In order to write this equation in vector form, define

$$\Pi(s,\tau) \triangleq \text{matrix with the i,j element } p(z_\tau = Z_j|z_s = Z_i) \qquad (3.53)$$

$$q_\tau(Y_0^\tau) \triangleq \text{vector with the j th element } K_\tau(Z_j, Y_0^\tau) \qquad (3.54)$$

$$H \triangleq \begin{pmatrix} h(z_1) & & 0 \\ & \ddots & \\ 0 & & h(z_{m(k+1)+k}) \end{pmatrix} \qquad (3.55)$$

$$\pi_\tau \triangleq \text{vector with j th element } p(z_\tau = Z_j) \qquad (3.56)$$

By direct substitution, then

$$q_\tau(Y_0^\tau) = \pi_\tau + \frac{\epsilon}{N_0}\int_0^\tau q_s(Y_0^\tau)H\Pi(s,\tau)dY_s \qquad (3.57)$$

Using equation (3.57) to find an increment in q with respect to time, substituting the relationship implied by the Kolmogorov equation

$$\frac{\partial}{\partial\tau}\Pi(s,\tau) = \frac{1}{N_0}\Pi(s,\tau)Q^\epsilon \qquad (3.58)$$

and taking the limit as $d\tau \to 0$, the resulting Zakai equation is found to be given by [37]

$$dq_\tau(Y_0^\tau) = \frac{1}{N_0}(Q^\epsilon)^T q_\tau d\tau + \frac{\epsilon}{N_0}Hq_\tau dy_\tau \qquad (3.59)$$

In the original time scale, then,

$$dq_t = (Q^\epsilon)^T q_t dt + \frac{\epsilon}{\sqrt{N_0}}Hq_t dy_t \qquad (3.60)$$

with initial condition

$$q_0 = \overline{q} \tag{3.61}$$

where $\overline{q} = (0\ldots010\ldots0)^T$. That is, $\overline{q}$ is a vector of zeros with a one in the column which, using the reordering indicated in Table 1, corresponds to $(x_0^1, x_0^2) = (\xi^1, \xi^2)$.

Equations (3.60) and (3.61) are the nonlinear filtering equations that define the evolution of the unnormalized conditional density. It will be shown in the next section that the cost given by Equation (3.13) can be written as a function of this density, q_t. Thus, q_t is a sufficient statistic for the problem defined in the previous two sections; the problem of optimally controlling a partially observed Markov chain. Equations (3.60) and (3.61) are the dynamics of the resultant transformed problem.

The (normalized) conditional density is also a sufficient statistic and has oftentimes been used as such to solve partially observed optimal control problems (c.f. [36]). However, equations (3.60) and (3.61) are considerably simpler than the differential equation that describes the normalized conditional density [38]. Thus, throughout this paper, the unnormalized conditional density is used to describe the dynamics of the equivalent control problem.

The simplification achieved by using Markov chains (as opposed to diffusions) for the original dynamics is apparent in equations (3.60) and (3.61). In contrast to the above results, the Zakai equation for partially observed diffusions is a partial differential equation with a stochastic driving term. The solution to this p.d.e., i.e. the conditional probability density for the current state, is a functional of the entire past history of observations. That is, the state is infinite dimensional. One is forced to deal with state functions rather than state vectors; and the optimization problem must be handled using function space techniques [29]. In order to avoid these mathematical difficulties, the Markov chain approximation of Section 2.2 was utilized in this thesis, resulting in the dynamics of equations (3.60) and (3.61).

3.4 Derivation of equivalent cost

In this section an expression for the cost given by equation (3.13) is derived. This equivalent cost is expressed as a function of q_t, the unnormalized conditional density, under measure P_0. Following [1], [17], and [18], define $\psi_\tau(\cdot)$ by

$$E^{A_\tau^y}\{l(z_\tau)\} = \frac{E_0^{A_\tau^y}\{\Lambda_\tau l(z_\tau)\}}{E_0^{A_\tau^y}\{\Lambda_\tau\}} \tag{3.62}$$

$$\triangleq \frac{\psi(l)}{\psi(1)} \tag{3.63}$$

It then follows that

$E\{l(z_\tau)\}$	$= E\{E^{A_\tau^y}\{l(z_\tau)\}\}$	Chain rule
	$= E_0\{\Lambda_{N_0 T} E^{A_\tau^y}\{l(z_\tau)\}\}$	Radon-Nikodym Theorem
	$= E_0\{\Lambda_{N_0 T}\frac{\psi(l)}{\psi(1)}\}$	Definition of ψ
	$= E_0\{E_0^{A_\tau^y}\{\Lambda_{N_0 T}\frac{\psi(l)}{\psi(1)}\}\}$	Chain rule
	$= E_0\{[E_0^{A_\tau^y}\Lambda_{N_0 T}]\frac{\psi(l)}{\psi(1)}\}$	$\psi(l)$ and $\psi(1)$ are measurable with respect to A_τ^y
	$= E_0\{[E_0^{A_\tau^y}E_0^{A_\tau}\Lambda_{N_0 T}]\frac{\psi(l)}{\psi(1)}\}$	$A_\tau \supset A_\tau^y$
	$= E_0\{[E_0^{A_\tau^y}\Lambda_\tau]\frac{\psi(l)}{\psi(1)}\}$	Λ_τ is a $P_0 - A_\tau$ martingale
	$= E_0\{\psi(1)\frac{\psi(l)}{\psi(1)}\}$	Definition of ψ
	$= E_0\{\psi(l)\}$	

It has thus been established that

$$E\{l(z_\tau)\} = E_0\{\psi(l)\} \tag{3.64}$$

But it has previously been shown that

$\psi(l)$	$= E_0^{A_\tau^y}\{\Lambda_\tau l(z_\tau)\}$	Definition of ψ
	$= \int_{\mathcal{R}^n} l(z)R_\tau(z, Y_0^\tau)P(dz,\tau)$	from equation (3.46)
	$= \int_{\mathcal{R}^n} l(z)R_\tau(z, Y_0^\tau)p(z,\tau)dz$	
	$= \int_{\mathcal{R}^n} l(z)K_\tau(z, Y_0^\tau)dz$	Definition of K
	$= C^T q_\tau(Y_0^\tau)$	Definition of q

where

$$C \triangleq \begin{pmatrix} l(Z_1) \\ \vdots \\ l(Z_{m(k+1)+k}) \end{pmatrix} \tag{3.65}$$

Thus, for the integral cost of equation (3.13),

$$\begin{aligned} J &= E\{\int_0^T l(z_t)dt\} && (3.66) \\ &= \frac{1}{N_0} E\{\int_0^{N_0 T} l(z_\tau)d\tau\} && (3.67) \\ &= \frac{1}{N_0} E_0\{\int_0^{N_0 T} C^T q_\tau d\tau\} && (3.68) \\ &= E_0\{\int_0^T C^T q_t dt\} && (3.69) \end{aligned}$$

Equation (3.69) is the desired expression for the equivalent cost.

Chapter 4

The equivalent stochastic control problem

In this section the equivalent stochastic control problem is stated for the case of regular perturbations. For the case of singular perturbations, a different infinitesimal generator is used in the approximation to the diffusion. However, in all other respects the derivation of an equivalent problem, and hence, the resultant problem is identical. Thus, for the singular perturbations case, a simple statement of the equivalent problem is made.

4.1 Regular perturbations

In summary, then, the problem of regular perturbations that is considered in this thesis is that of minimizing the cost

$$J = E_0\{\int_0^T C^T q_t dt\} \tag{4.1}$$

subject to the constraint

$$dq_t = (A + \epsilon B)^T q_t dt + \frac{\epsilon}{\sqrt{N_0}} H q_t dy_t \tag{4.2}$$

with initial condition

$$q_0 = \overline{q} \tag{4.3}$$

where under P_0 the observations y_t are an independent scalar Brownian process. Here A and B are given by equations (3.23) and (3.32) and C by equation (3.65).

4.2 Singular perturbations

In the case of singular perturbations, the original dynamics are given by

$$dx_t^1 = f(x^1, x^2)dt \tag{4.4}$$

$$dx_t^2 = \frac{1}{\epsilon}b(x^1, x^2, u)dt + \frac{1}{\sqrt{\epsilon}}dw_t \tag{4.5}$$

$$x_0^1 = \xi^1 \tag{4.6}$$

$$x_0^2 = \xi^2 \tag{4.7}$$

Consequently, the approximating Markov chain has an infintesimal generator of the form

$$\frac{A}{\epsilon} + B \tag{4.8}$$

where A, B and C are given by equations (3.23), (3.32) and (3.65) respectively.

Thus the Zakai equation for the singular perturbations case is given by

$$dq_t = (\frac{A^T}{\epsilon} + B^T)q_t dt + \frac{1}{\sqrt{N_0}}Hq_t dy_t \tag{4.9}$$

with initial condition

$$q_0 = \overline{q} \tag{4.10}$$

where under P_0 the observations y_t are an independent scalar Brownian process.

In summary, then, the problem of singular perturbations that is considered in this paper is that of minimizing the cost

$$J = E_0\{\int_0^T C^T q_t dt\} \tag{4.11}$$

subject to the constraint of equation (4.9)

Chapter 5

Maximum principle and Dynamic programming

The purpose of this section is to describe the stochastic maximum principle and the dynamic programming equations as they relate to the problem defined in this paper. To this end, after stating the usual assumptions regarding the dynamics and the cost function, a derivation of the necessary conditions for optimality is presented. The proofs for these results may be found in [2].

5.1 Assumptions

Since the problem of interest for this paper has no terminal cost, and involves only a scalar Wiener process and a scalar control, in order to simplify the presentation, the assumptions below will conform to these conditions. Thus, for the purposes of this section the following relationships will be defined

$$m(q,v) \; : \; \mathcal{R}^n \times \mathcal{R} \rightarrow \mathcal{R}^n \tag{5.1}$$

$$\sigma(q,v) \; : \; \mathcal{R}^n \times \mathcal{R} \rightarrow \mathcal{R}^n \tag{5.2}$$

$$m, \sigma \text{ are continuously differentiable} \tag{5.3}$$

$$m, \sigma \text{ have bounded derivatives} \tag{5.4}$$

$$|m(q,v)| \leq \overline{m}(1+|q|+|v|) \quad (5.5)$$

$$|\sigma(q,v)| \leq \overline{\sigma}(1+|q|+|v|) \quad (5.6)$$

$$G(q,v) : \mathcal{R}^n \times \mathcal{R} \to \mathcal{R} \quad (5.7)$$

$$G \text{ is continuously differentiable} \quad (5.8)$$

$$|G(q,v)| \leq \overline{G}(1+|q|^2+|v|^2) \quad (5.9)$$

$$|G_q(q,v)|, |G_v(q,v)| \leq \overline{G}(1+|q|+|v|) \quad (5.10)$$

Given a filtration F_t on (Ω, A, P) and an F_t measurable standard Wiener process $y(t)$, an admissible control is a process satisfying

$$v(\cdot) \in L^2_F(0,T;\mathcal{R}) \quad (5.11)$$

$$v(t) \in U_{ad} \quad (5.12)$$

$$U_{ad} \text{ convex, closed, and non-empty subset of } \mathcal{R} \quad (5.13)$$

For all admissible controls, the state of the system is defined by

$$dq(t) = m(q(t),v(t))dt + \sigma(q(t),v(t))dy(t) \quad (5.14)$$

$$q(0) = q_0. \quad (5.15)$$

The following cost criterion will be considered

$$J(v(\cdot)) = E\{\int_0^T G(q(t),v(t))dt\} \quad (5.16)$$

5.2 Extremality condition

Let $u(\cdot)$ denote an optimal control, and let $x(\cdot)$ denote the corresponding state defined by

$$dx = m(x,u)dt + \sigma(x,u)dy \quad (5.17)$$

$$x(0) = q_0 \quad (5.18)$$

Let $v(\cdot)$ be an admissible control and

$$u_\theta(t) = u(t) + \theta v(t) \tag{5.19}$$

We will calculate $\frac{d}{d\theta}J(u_\theta)|_{\theta=0}$.

Lemma 3

$$\frac{d}{d\theta}J(u_\theta)|_{\theta=0} = E\int_0^T [G_q(x,u)\cdot z + G_v(x,u)v]dt \tag{5.20}$$

Proof: see [2]

Lemma 4 *Equation (5.20) can be written as*

$$\frac{d}{d\theta}J(u_\theta)|_{\theta=0} = E\{\int_0^T H_v(x(t),u(t),\lambda(t),r(t))\cdot v(t)dt\} \tag{5.21}$$

where the hamiltonian, H, is defined by

$$H(q,v,\lambda,r) = G(q,v) + \lambda\cdot m(q,v) + r\cdot\sigma(q,v) \tag{5.22}$$

and r is a column vector.

Proof: Given arbitrary functions $\phi \in L^2_F(0,T;R^n)$ and $\psi \in L^2_F(0,T;R^n)$, consider the functions ζ which is the solution to

$$d\zeta = (m_q(x,u)\zeta + \phi)dt + (\sigma_q(x,u)\zeta + \psi)dy \tag{5.23}$$

$$\zeta(0) = 0 \tag{5.24}$$

Then the mapping

$$\phi,\psi \rightarrow E\{\int_0^T G_q(x,u)\cdot\zeta dt\} \tag{5.25}$$

is a linear and continuous mapping of $L^2_F(0,T;R^n)$ into R. Thus, there exists a unique $\lambda \in L^2_F(0,T;R^n)$ and $r \in L^2_F(0,T;R^n)$ such that

$$E\{\int_0^T \phi(t)\cdot\lambda(t)dt + \int_0^T \psi(t)\cdot r(t)dt\} = E\{\int_0^T G_q(x,u)\zeta dt\} \tag{5.26}$$

Combining this result with Lemma 3 gives the desired result.

Theorem 2 *Under the hypotheses of equations (5.1) through (5.13), if $u(\cdot)$ is an optimal control for equations (5.14) through (5.16) then the following necessary condition is true*

$$H_v(q(t), u(t), \lambda(t), r(t)) \cdot (v - u(t)) \geq 0 \tag{5.27}$$

a.s. $\forall v \in U_{ad}$

Proof: For any admissible control $v(\cdot)$, the control $u(\cdot) + \theta(v(\cdot) - u(\cdot))$ is also admissible. Thus,

$$J(u(\cdot) + \theta(v(\cdot) - u(\cdot)) \geq J(u(\cdot)) \tag{5.28}$$

Then, applying lemma 4,

$$E\{\int_0^T H_v(x(t), u(t), \lambda(t), r(t)) \cdot (v(t) - u(t))dt\} \geq 0 \tag{5.29}$$

In order to localize this condition, let

$$\gamma(t, \omega) = H_v(x(t), u(t), \lambda(t), r(t)) \cdot (v(t) - u(t)) \tag{5.30}$$

where $v \in U_{ad}$ is fixed. Let

$$S = \{(t, \omega) | \gamma(t, \omega) < 0\} \tag{5.31}$$

Also, let $v(t, \omega)$ be v in S and be $u(t, \omega)$ outside of S. Let

$$S_t = \{(\omega) | \gamma(t, \omega) < 0\} \tag{5.32}$$

and $v(t)$ be v in S_t and be $u(t)$ outside of S_t. Then, by equation (5.29),

$$\int_S \gamma(t, \omega) dt dP \geq 0 \tag{5.33}$$

which contradicts the definition of S, unless the measure of S is zero. This proves the theorem.

Theorem 3 *Under the hypotheses of Theorem 2 as well as the condition that*

$$F_t = \sigma(y(s), s \leq t) \tag{5.34}$$

the vectors λ and r are uniquely characterized by the equation

$$- d\lambda = [m_q^*(x,u)\lambda + G_q(x,u) + \sigma_q^*(x,u)r]dt - r(t)dy \tag{5.35}$$

$$\lambda(T) = 0 \tag{5.36}$$

Proof: It is readily shown that if λ and r satisfy equations (5.35) and (5.36), then they also satisfy equation (5.26). In fact, applying these equations and using Ito's formula,

$$\begin{aligned} d[\lambda(t)\cdot\zeta(t)] &= \lambda\cdot(m_q\zeta+\phi)dt + \lambda\cdot(\sigma_q\zeta+\psi)dy + \\ &\quad -\zeta\cdot[(m_q^*\lambda + G_q + \sigma_q^* r)dt - rdy] + r\cdot(\sigma_q\zeta+\psi)dt \end{aligned} \tag{5.37}$$

Therefore, combining terms,

$$E\{\lambda(T)\cdot\zeta(T)\} = E\{\int_0^T (\lambda\cdot\phi - \zeta\cdot G_q + r\cdot\psi)dt\} \tag{5.38}$$

which, using the terminal conditions, gives equation (5.26).

It now only remains to show that equation (5.35) and (5.36) has a solution. To do this, the matrices Φ and Ψ are defined as follows

$$d\Phi = m_q(x,u)\Phi dt + \sigma_q(x,u)\Phi dy \tag{5.39}$$

$$\Phi(0) = I \tag{5.40}$$

$$- d\Psi = (\Psi m_q(x,u) - \Psi\sigma_q\sigma_q(x,u))dt + \Psi\sigma_q(x,u)dy \tag{5.41}$$

$$\Psi(0) = I \tag{5.42}$$

By differentiating $\Phi\Psi$ a differential equation in $\Phi\Psi$ is obtained with initial condition $\Phi(0)\Psi(0) = I$. It is readily verified that the solution to this equation is $\Phi(t)\Psi(t) = I$. Similarly, differentiating $\Psi\Phi$ it is determined that

$$\Phi(t)\Psi(t) = \Psi(t)\Phi(t) = I \tag{5.43}$$

Now, letting $\psi = 0$ in equation (5.23), and computing the differential $d\Psi(t)\zeta(t)$, it is readily shown that

$$\zeta(t) = \Phi(t)\int_0^t \Psi(s)\phi(s)ds \tag{5.44}$$

By substitution into equation (5.26), then it quickly follows that

$$E\{\int_0^T \phi(t)\lambda(t)dt\} = E\{\int_0^T \phi(t)\cdot\Psi^*(t)[\int_0^T \Phi^*(s)G_q(x(s),u(s))ds - \int_0^t \Phi^*(s)G_q(x(s),u)ds]dt\} \quad (5.45)$$

And so,

$$\lambda(t) = \Psi^*(t)[-\int_0^t \Phi^*(s)G_q(x(s),u(s))ds + E^{F_t}\{X\}] \quad (5.46)$$

where

$$X = \int_0^T \Phi^*(s)G_q(x(s),u(s))ds \quad (5.47)$$

The process $E^{F_t}\{X\}$ is an F_t martingale. It therefore has the representation

$$E^{F_t}\{X\} = \int_0^T Mdy + E\{X\} \quad (5.48)$$

where $M \in L^2_F(0,T;R^n)$ is uniquely defined.

Note that

$$\lambda(T) = 0 \quad (5.49)$$

Transposing the equation defining Ψ, defining the process ξ by

$$d\xi = -\Phi^*(t)G_q(x(t),u(t))dt + Mdy \quad (5.50)$$

and assuming the relationship

$$\lambda(t) = \Psi^*(t)\xi(t) \quad (5.51)$$

it is readily verified by a direct computation of $d\lambda$ that equations (5.35) and (5.36) have a solution if

$$r(t) = -\sigma_q^*(x,u)\lambda(t) + \Psi^*(t)M(t) \quad (5.52)$$

Equations (5.46) and (5.52) thus define a solution to equations (5.35) and (5.36).

5.3 Dynamic Programming

The Stochastic Maximum Principle, derived in the previous section, consists of a set of stochastic differential equations which are necessary conditions for optimality. In this section attention is turned to dynamic programming as an approach to determining the optimal control. Dynamic programming, for the problem of the previous section, will provide a partial differential equation which is a sufficient condition for optimality. A formal derivation of this partial differential equation, the Hamilton-Jacobi-Bellman equation, follows below [23]. More rigorous derivations may be found in [3] and [25].

Let the symbol $E^v_{q,s}$ denote the expectation given that $q_s = q$ and the control $v(q,t)$ is used. Similarly, let the differential operator for the process of equation (5.14) with control $v(q,t)$ be denoted

$$\mathcal{L}^v = \sum_i m_i(q,v) \cdot \frac{\partial}{\partial q_i} + \frac{1}{2}\sum_{i,j}(\sigma\sigma^*)_{ij}(q,v)\frac{\partial^2}{\partial q_i \partial q_j} \tag{5.53}$$

The cost associated with the application of some control, v, is defined as

$$V^v(q, T-t) \triangleq E^v_{q,t} \int_t^T G(q_s, v_s)ds \tag{5.54}$$

The cost associated with the application of the optimal control is defined as

$$V(q, T-t) \triangleq \inf_{\substack{v(q,s) \in U_{ad} \\ t \le s \le T}} V^v(q, T-t) \tag{5.55}$$

Applying the principle of optimality, that is, the property that subpaths of optimal paths are themselves optimal, yields the relationship

$$V(q, T-t) = \inf_{\substack{v(q,s) \in U_{ad} \\ t \le s \le t+\Delta}} E^v_{q,t}\{V(q_{t+\Delta}, T-t-\Delta) + \int_t^{t+\Delta} G(q_s, v_s)ds\} \tag{5.56}$$

This equation indicates that the optimal cost-to-go from time t is the minimum of the sum of the incremental cost (i.e. the cost associated with the time interval $(t, t+\Delta)$) and the optimal cost-to-go from time $t+\Delta$. This result must be so since, by the principle of optimality, whatever state is reached at time $t+\Delta$, if the path is to be optimal, it must also be an optimal path from that state at time $t+\Delta$ onwards.

Rearranging terms and dividing by Δ, equation (5.56) can be written equivalently as

$$0 = \inf_{v \in U_{ad}} \{\frac{1}{\Delta}[E^v_{q,t} V(q_{t+\Delta}, T-t-\Delta) - V(q, T-t)] + \frac{E^v_{q,t}}{\Delta} \int_t^{t+\Delta} G(q_s, v_s) ds\} \tag{5.57}$$

By expanding this equation and taking the limit as $\Delta \to 0$, one obtains

$$0 = \inf_{v(q,s) \in U_{ad}} [V_t(q, T-t) + \mathcal{L}^v V(q, T-t) + G(q, v)] \tag{5.58}$$

It is useful to define another cost-to-go function, ϕ. This new cost-to-go will be defined as a function of the current time, rather than a function of the time remaining. That is,

$$\phi(q, t) \triangleq V(q, T-t) \tag{5.59}$$

Thus, equation (5.58) may be written as

$$\frac{\partial \phi}{\partial t} + \inf_{v \in U_{ad}} \{G(q, v) + D\phi \cdot m(q, v) + \frac{1}{2} tr D^2 \phi \sigma \sigma^*(q, v)\} = 0 \tag{5.60}$$

$$\phi(q, T) = 0 \tag{5.61}$$

In order to show sufficiency, suppose that the infimum is achieved at some value $u = v(q, t)$. Then the realtionship between u and any other control function v is given by

$$V_t(q, T-t) + \mathcal{L}^u V(q, T-t) + G(q, u) = 0 \tag{5.62}$$

$$V_t(q, T-t) + \mathcal{L}^v V(q, T-t) + G(q, v) \geq 0 \tag{5.63}$$

Integrating these equations and combining the results yields

$$E^u_{q,t}\{\int_t^T [V_s(q_s,T-s)+\mathcal{L}^u V(q_s,T-s)+G(q_s,u_s)]ds\} \leq$$
$$E^v_{q,t}\{\int_t^T [V_s(q_s,T-s)+\mathcal{L}^v V(q_s,T-s)+G(q_s,u_s)]ds\} \qquad (5.64)$$

Then, applying Ito's lemma and the initial condition $V(q,0)=0$, one obtains

$$V(q,t)=E^u_{q,t}\{\int_t^T G(q_s,u_s)ds\} \leq E^v_{q,t}\{\int_t^T G(q_s,v_s)ds\} \qquad (5.65)$$

Thus, equation (5.60) provides a sufficient condition for optimality, in the sense that if it has a solution, it is the optimal cost. That is why such theorems are called "verification" theorems.

5.4 Relationship between Dynamic Programming and the Maximum Principle

Let $u(\cdot)$ denote an optimal control and let $x(\cdot)$ be the corresponding trajectory. Also, let

$$\lambda(t)=D\phi(x(t),t) \qquad (5.66)$$

By Ito's formula

$$d\lambda = D^2\phi(x(t),t)\cdot(m(x,u)dt+\sigma(x)dy)+\frac{1}{2}trD^3\phi\sigma\sigma^*(x,u)dt+D\frac{\partial\phi}{\partial t}dt \quad (5.67)$$

Differentiating equation (5.60),

$$D\frac{\partial\phi}{\partial t}+G_q(x,u)+D^2\phi\cdot m(x,u)+m^*_q(x,u)D\phi+$$
$$\frac{1}{2}trD^3\phi\sigma\sigma^*(x,u)+\sigma^*_q D^2\phi\sigma(x,u)=0 \qquad (5.68)$$

The following relationship follows from equations (5.67) and (5.68)

$$-d\lambda=(m^*_q(x,u)\lambda(t)+G_q(x,u)+\sigma^*_q D^2\phi\sigma(x,u)dt-D^2\phi(x(t),t)\sigma(x,u)dy$$
$$(5.69)$$

$$\lambda(T) = 0 \tag{5.70}$$

But this is simply a statement of Theorem 3 if

$$r(t) = D^2\phi(x(t), t)\sigma(x(t), u(t)) \tag{5.71}$$

Chapter 6

Stochastic differential equations and Bilinear systems

The systems of interest in this paper are bilinear. In this section, a theorem is stated which describes the Volterra expansion of such systems, as well as theorems which describe an equivalence between stochastic differential equations (i.e. differential equations driven by a Wiener process), and the corresponding differential equations driven by white noise. These theorems will be used to demonstrate convergence of the decompositions derived below.

6.1 Stochastic differential equations

The following theorem gives the stochastic differential equation that corresponds to a differential equation driven by white noise. It may be found in [37].

Theorem 4 (Wong Zakai Correction) *Let*

$$\frac{d}{dt}x(t) = m(x(t),t) + \sigma(x(t),t)\xi_t \tag{6.1}$$

be a stochastic differential equation such that $m(x,t)$*,* $\sigma(x,t)$*,* $\frac{\partial}{\partial x_i}\sigma(x,t), i = 1,\ldots,n$*, and* $\frac{\partial}{\partial t}\sigma(x,t)$ *are continuous on* $-\infty < x_i < \infty, a \leq t \leq b$*. Also let*

m, σ, and $\sigma\frac{\partial\sigma}{\partial x_i}$ satisfy a uniform Lipschitz condition:

$$\|m(x,t)-m(y,t)\| \leq K\|x-y\| \tag{6.2}$$

$$\|\sigma(x,t)-\sigma(y,t)\| \leq K\|x-y\| \tag{6.3}$$

$$\|\sigma\frac{\partial\sigma}{\partial x_i}(x,t)-\sigma\frac{\partial\sigma}{\partial x_i}(y,t)\| \leq K\|x-y\| \tag{6.4}$$

Let $\|\sigma\| \geq K_1 > 0$ and $\|\sigma\| < K_2\|\sigma\|^2$. If $\xi \to n(t)$, a white noise with unity spectral density, and $dw(t)$ is the corresponding Wiener increments process, then the solution x converges almost surely on $a \leq t \leq b$ to the solution of

$$dx_t = [m(x_t,t) + \frac{1}{2}\sum_{l,m}\frac{\partial\sigma_{km}}{\partial x_l}(x_t,t)\sigma_{lm}(x_t,t)]dt + \sigma(x_t,t)dw_t \tag{6.5}$$

The following theorem provides the formula for converting differential equations driven by white noise into stochastic differential equations.

Theorem 5 (Application to bilinear systems) *Let*

$$\frac{d}{dt}x(t) = Ax(t) + Nx(t)\xi(t) + B\xi(t) + D \tag{6.6}$$

If $\xi \to n(t)$, a white noise with unity spetral density, then x converges to the solution of

$$dx_t = (Ax_t + \frac{1}{2}NNx_t + \frac{1}{2}NB + D)dt + Nx_t dw_t + Bdw_t \tag{6.7}$$

Proof: Immediate from Theorem 4

Next, a formula for the conversion of a Wiener driven stochastic differential equation into white noise differential equations will be provided.

Theorem 6 *Let*

$$dx_t = Ax_t dt + Nx_t dw_t + Bdw_t \tag{6.8}$$

If $\xi \to n(t)$, with unity spetral density, then x converges to the solution of

$$\frac{d}{dt}x_t = \overline{A}x_t + Nx_t n(t) + Bn(t) + D \tag{6.9}$$

where

$$\overline{A} = A - \frac{1}{2}NN \qquad (6.10)$$
$$D = -\frac{1}{2}NB \qquad (6.11)$$

Proof: Immediate from Theorem 4

6.2 Bilinear systems

Theorem 7 *The bilinear system with state equation given by*

$$\dot{x}(t) = A(t)x(t) + D(t)x(t)u(t) + b(t)u(t) \qquad (6.12)$$

$$x(0) = x_0 \qquad (6.13)$$

has the Volterra system representation

$$\begin{aligned} x(t) &= \Phi(t,0)x_0 + \sum_{k=1}^{\infty} \int_0^t \int_0^{\sigma_1} \cdots \int_0^{\sigma_{k-1}} \Phi(t,\sigma_1)D(\sigma_1)\Phi(\sigma_1,\sigma_2)D(\sigma_2)\cdots \\ &\quad \cdots D(\sigma_k)\Phi(\sigma_k,0)x_0 u(\sigma_1)\cdots u(\sigma_k)d\sigma_k \cdots d\sigma_1 + \\ &\quad \sum_{k=1}^{\infty} \int_0^t \int_0^{\sigma_1} \cdots \int_0^{\sigma k-1} \Phi(t,\sigma_1)D(\sigma_1)\Phi(\sigma_1,\sigma_2)D(\sigma_2)\cdots \\ &\quad \cdots D(\sigma_{k-1})\Phi(\sigma_{k-1},\sigma_k)b(\sigma_k)u(\sigma_1)\cdots u(\sigma_k)d\sigma_k \cdots d\sigma_1 \end{aligned} \qquad (6.14)$$

which also converges uniformly on any interval.

Proof: see [34].

Chapter 7

Regular perturbations

In this section an approximate solution to the problem of equations (4.1) and (4.2) is found by means of the stochastic maximum principle. The resultant control, $u^M = u_0 + \epsilon u_1 + \cdots + \epsilon^m u_m$, consists of controls which are optimal for a set of simpler control problems. In particular, u_0 is optimal for a deterministic optimal control problem, while $u_i, i \geq 1$ are solutions to problems having linear dynamics and quadratic cost criteria.

Each of these latter problems result in an affine stochastic feedback. In fact

$$u^M(t) = u_0(t) + \sum_{m=1}^{M} \epsilon^m (K q_m + \beta_m) \tag{7.1}$$

where q_m is the state that corresponds to a linear quadratic problem with control u_m applied. In addition the related Ricatti equations have stochastic input for $m \geq 2$ which implies that the corresponding β must be computed online. However, in order to have a control law that is more closely related to the state of the actual system, the following feedback control may be used

$$\overline{u}^M(q;t) = u_0(t) + K(t)(q - q_0(t)) + \sum_{m=1}^{M} \epsilon^m \beta_m \tag{7.2}$$

This feedback law gives nearly the performance of the control of equation (7.1), the difference being only in high order terms.

The feedback control law is perhaps most useful for the case in which $M = 1$ since for this case the functions $K(t)$ and $\beta(t)$ can be computed offline.

In the sections below, first an asymptotic analysis is performed in order to describe each of the successive problems. This is followed by a proof of convergence.

7.1 Asymptotic analysis

The Maximum Principle provides a set of necessary conditions for the problem described by equations (4.1) and (4.2). Thus, the following Hamiltonians are defined

$$H_0(q, u, \lambda) \triangleq C^T q + \lambda^T [A^T(u) q] \tag{7.3}$$

$$H_1(q, \lambda, r) \triangleq \lambda^T (B^T q) + \frac{1}{\sqrt{N_0}} r^T (Hq) \tag{7.4}$$

$$H^\epsilon \triangleq H_0 + \epsilon H_1 \tag{7.5}$$

We assume that

$$A \in C^\infty \tag{7.6}$$

The necessary conditions for optimality are given by

$$dq^\epsilon = [A^T(u^\epsilon) + \epsilon B^T] q^\epsilon dt + \frac{\epsilon}{\sqrt{N_0}} H q^\epsilon dy_t \tag{7.7}$$

$$-d\lambda^\epsilon = [C + (A(u^\epsilon) + \epsilon B)\lambda^\epsilon + \frac{\epsilon}{\sqrt{N_0}} H r^\epsilon] dt - r^\epsilon(t) dy \tag{7.8}$$

$$q^\epsilon(0) = \overline{q} \tag{7.9}$$

$$\lambda^\epsilon(T) = 0 \tag{7.10}$$

$$H_u^\epsilon(q^\epsilon, u^\epsilon, \lambda^\epsilon, r^\epsilon) = 0 \tag{7.11}$$

That is,

$$(\lambda^\epsilon)^T A_u^T(u^\epsilon) q^\epsilon = 0 \tag{7.12}$$

Let

$$q^\epsilon = q_0 + \sum_{j=1}^{\infty} \epsilon^j q_j \tag{7.13}$$

$$u^\epsilon = u_0 + \sum_{j=1}^{\infty} \epsilon^j u_j \tag{7.14}$$

$$r^\epsilon = r_0 + \sum_{j=1}^{\infty} \epsilon^j r_j \tag{7.15}$$

$$\lambda^\epsilon = \lambda_0 + \sum_{j=1}^{\infty} \epsilon^j \lambda_j \tag{7.16}$$

Using a Taylor series, the matrix $A(u)$ can then be expanded as

$$A(u^\epsilon) = A(u_0 + \sum_{j=1}^{\infty} \epsilon^j u_j) \tag{7.17}$$

$$= A(u_0) + \sum_{j=1}^{\infty} \frac{D^{(j)}A(u_0)(\sum_{i=1}^{\infty} \epsilon^i u_i)^j}{j!} \tag{7.18}$$

$$= A(u_0) + A_u(u_0)(\sum_{i=1}^{\infty} \epsilon^i u_i) + \frac{1}{2}A_{uu}(u_0)(\sum_{i=1}^{\infty} \epsilon^i u_i)^2 + \cdots \tag{7.19}$$

Or, equivalently, regrouping in like powers of ϵ,

$$A(u^\epsilon) = A(u_0) + \sum_{j=1}^{\infty} \sum_{l_1,\ldots,l_j=1}^{\infty} \frac{D^{(j)}A(u_0)u_{l_1}\ldots u_{l_j}}{j!}\epsilon^{l_1+\cdots+l_j} \tag{7.20}$$

$$= A(u_0) + \sum_{k=1}^{\infty} \epsilon^k \sum_{j=1}^{k} \sum_{\substack{l_1+\cdots+l_j=k \\ l_1,\ldots,l_j \geq 1}} \frac{D^{(j)}A(u_0)u_{l_1}\ldots u_{l_j}}{j!} \tag{7.21}$$

From equation (5.71) and from the form of the diffusion term in equation (7.7),

$$r_0 = 0 \tag{7.22}$$

Substitution of these expansions into the equations expressing the necessary conditions yields

$$d(\sum_{i=0}^{\infty} \epsilon^i q_i) =$$

$$= \{A^T(u_0) + \sum_{k=1}^{\infty} \epsilon^k \sum_{j=1}^{k} \sum_{\substack{l_1+\cdots+l_j=k \\ l_1,\ldots,l_j \geq 1}} \frac{D^{(j)}A^T(u_0)u_{l_1}\ldots u_{l_j}}{j!} + \epsilon B^T\}$$

$$\cdot[\sum_{j=0}^{\infty} \epsilon^j q_j]dt + \frac{\epsilon}{\sqrt{N_0}} H \sum_{j=0}^{\infty} \epsilon^j q_j \tag{7.23}$$

$$-d(\sum_{i=0}^{\infty} \epsilon^i \lambda_i) =$$

$$= \{C + [A(u_0) + \sum_{k=1}^{\infty} \epsilon^k \sum_{j=1}^{k} \sum_{\substack{l_1 + \cdots + l_j = k \\ l_1, \ldots, l_j \geq 1}} \frac{D^{(j)} A(u_0) u_{l_1} \ldots u_{l_j}}{j!} + \epsilon B]$$

$$\cdot[\sum_{j=0}^{\infty} \epsilon^j \lambda_j] + \frac{\epsilon}{\sqrt{N_0}} H \sum_{j=1}^{\infty} \epsilon^j r_j\}dt - \sum_{j=1}^{\infty} \epsilon^j r_j dy \tag{7.24}$$

from equations (7.7) and (7.8) respectively. From equation (7.12) we have, for $m \geq 1$,

$$\epsilon^m \{\sum_{k=1}^{m} \sum_{i=1}^{m-k} (\lambda_i^T)[\sum_{j=1}^{k} \sum_{\substack{l_1 + \cdots + l_j = k \\ l_1, \ldots, l_j \geq 1}} \frac{D^{(j)} A_u^T(u_0) u_{l_1} \ldots u_{l_j}}{j!}] q_{m-i-k}$$

$$+ \sum_{i=0}^{m} (\lambda_i^T) A_u^T(u_0) q_{m-i}\} = 0 \tag{7.25}$$

and, from equation (7.12),

$$\lambda_0^T A_u^T(u_0) q_0 = 0 \tag{7.26}$$

We now collect terms involving like powers of ϵ. Equating the coefficients of ϵ^0 yields

$$\dot{q}_0 = A^T(u_0) q_0 \tag{7.27}$$

$$-\dot{\lambda}_0 = C + A(u_0)\lambda_0 \tag{7.28}$$

$$q_0(0) = \bar{q} \tag{7.29}$$

$$\lambda_0(T) = 0 \tag{7.30}$$

$$\lambda_0^T A_u^T(u_0) q_0 = 0 \tag{7.31}$$

Equating the coefficients of ϵ^1 yields

$$dq_1 = [A^T(u_0)q_1 + A_u^T(u_0)q_0u_1 + B^Tq_0]dt + \frac{1}{\sqrt{N_0}}Hq_0dy_t \quad (7.32)$$

$$-d\lambda_1 = [A(u_0)\lambda_1 + A_u(u_0)\lambda_0u_1 + B\lambda_0]dt - r_1dy_t \quad (7.33)$$

$$q_1(0) = 0 \quad (7.34)$$

$$\lambda_1(T) = 0 \quad (7.35)$$

$$\lambda_0^TA_u^T(u_0)q_1 + \lambda_0^TA_{uu}^T(u_0)q_0u_1 + \lambda_1^TA_u^T(u_0)q_0 = 0 \quad (7.36)$$

Equating the coefficients of $\epsilon^m, (m \geq 2)$ yields

$$dq_m = [A^T(u_0)q_m + A_u^T(u_0)q_0u_m + X_m]dt + \frac{1}{\sqrt{N_0}}Hq_{m-1}dy \quad (7.37)$$

$$-d\lambda_m = [A(u_0)\lambda_m + A_u(u_0)\lambda_0u_m + \Lambda_m]dt - r_m(t)dy_t \quad (7.38)$$

$$q_m(0) = 0 \quad (7.39)$$

$$\lambda_m(T) = 0 \quad (7.40)$$

$$\lambda_m^TA_u^T(u_0)q_0 + \lambda_0^TA_u^T(u_0)q_m + \lambda_0^TA_{uu}^T(u_0)q_0u_m + N_m = 0 \quad (7.41)$$

where X_m, Λ_m, and N_m are terms not involving q_m, λ_m, or u_m and are defined as follows

$$X_m(q_0, \ldots, q_{m-1}, u_0, \ldots, u_{m-1}) = [\sum_{k=1}^{m-1}\sum_{j=1}^{k} \sum_{\substack{l_1+\cdots+l_j=k \\ l_1,\ldots,l_j\geq 1}} \frac{D^{(j)}A^T(u_0)u_{l_1}\ldots u_{l_j}}{j!}q_{m-k}] + $$
$$+[\sum_{j=2}^{m} \sum_{\substack{l_1+\cdots+l_j=m \\ l_1,\ldots,l_j\geq 1}} \frac{D^{(j)}A^T(u_0)u_{l_1}\ldots u_{l_j}}{j!}q_0] + B^Tq_{m-1} \quad (7.42)$$

$$\Lambda_m(\lambda_0, \ldots, \lambda_{m-1}, r_{m-1}, u_0, \ldots, u_{m-1}) = [\sum_{k=1}^{m-1}\sum_{j=1}^{k} \sum_{\substack{l_1+\cdots+l_j=k \\ l_1,\ldots,l_j\geq 1}} \frac{D^{(j)}A(u_0)u_{l_1}\ldots u_{l_j}}{j!}\lambda_{m-k}] +$$

$$+[\sum_{j=2}^{m} \sum_{\substack{l_1+\cdots+l_j=m \\ l_1,\ldots,l_j\geq 1}} \frac{D^{(j)}A(u_0)u_{l_1}\ldots u_{l_j}}{j!}\lambda_0]+$$

$$+\frac{1}{\sqrt{N_0}}(Hr_{m-1})+B\lambda_{m-1} \tag{7.43}$$

$$N_m(q_0,\ldots,q_{m-1},\lambda_0,\ldots,\lambda_{m-1},u_0,\ldots,u_{m-1}) =$$

$$[\sum_{i=1}^{m-1}\lambda_i^T A_u^T(u_0)q_{m-i}]+$$

$$+[\sum_{k=1}^{m-1}\sum_{i=0}^{m-k}\lambda_i^T\sum_{j=1}^{k} \sum_{\substack{l_1+\cdots+l_j=k \\ l_1,\ldots,l_j\geq 1}} \frac{D^{(j)}A_u^T(u_0)u_{l_1}\ldots u_{l_j}}{j!}q_{m-i-k}]+$$

$$+\lambda_0^T\sum_{j=2}^{m} \sum_{\substack{l_1+\cdots+l_j=m \\ l_1,\ldots,l_j\geq 1}} \frac{D^{(j)}A_u^T(u_0)u_{l_1}\ldots u_{l_j}}{j!}q_0 \tag{7.44}$$

7.2 Resolution of the successive problems

Equations (7.27) through (7.31) are precisely the necessary conditions for the deterministic optimal control problem of minimizing the cost functional

$$J_0 = \int_0^T C^T q dt \tag{7.45}$$

subject to the dynamics

$$\dot{q} = A^T(u)q \tag{7.46}$$

$$q(0) = \overline{q} \tag{7.47}$$

This is the discrete analog of the intuitive picture we gave in the introduction (c.f. the discussion of equation (1.13)). Indeed equation (7.46) is the discretization of equation (1.13), when $\epsilon = 0$. In this case since $Q(x^1, x^2, t)$ is not random, equation (1.15) is discretized to equation (7.45) and the resulting problem is deterministic. As in equations (7.27) through (7.31), we denote by q_0, λ_0, and u_0 the optimal state, costate, and control for this problem.

Equations (7.32) through (7.36) are the necessary conditions for the stochastic optimal control problem of minimizing the quadratic cost functional

$$J_1(u) = \frac{1}{2}E\int_0^T 2\lambda_0^T A_u^T(u_0)qu + \lambda_0^T A_{uu}^T(u_0)q_0u^2 + 2\lambda_0^T B^T q dt \qquad (7.48)$$

subject to the linear dynamics

$$dq = [A_u^T(u_0)q_0u + A^T(u_0)q + B^T q_0]dt + \frac{1}{\sqrt{N_0}}Hq_0dy_t \qquad (7.49)$$

$$q(0) = 0 \qquad (7.50)$$

where u_0, λ_0, and q_0 are determined by the solution to equations (7.45), (7.46), and (7.47).

Similarly, equations (7.37) through (7.41) are the necessary conditions for the stochastic optimal control problem of minimizing the quadratic cost functional

$$J_m(u) = \frac{1}{2}E\int_0^T 2\lambda_0^T A_u^T(u_0)qu + \lambda_0^T A_{uu}^T(u_0)q_0u^2 + 2\Lambda_m q + 2N_m u dt \qquad (7.51)$$

subject to the linear dynamics

$$dq = [A_u^T(u_0)q_0u + A^T(u_0)q + X_m]dt + \frac{1}{\sqrt{N_0}}Hq_{m-1}dy_t \qquad (7.52)$$

$$q(0) = 0 \qquad (7.53)$$

where terms with indices less than m are determined from earlier problems.

We can verify that the necessary conditions for optimality for the problem described by equations (7.51) through (7.53) are given by equations (7.37) through (7.41). The Hamiltonian for the problem of equations (7.51) through (7.53) is

$$\begin{aligned}&\lambda_0^T A_u^T(u_0)qu + \frac{1}{2}\lambda_0^T A_{uu}^T(u_0)q_0u^2 + \Lambda_m q + N_m u + \\ &\lambda^T[A^T(u_0)q + A_u^T(u_0)q_0u + X_m] + \frac{1}{\sqrt{N_0}}r^T Hq_{m-1}\end{aligned} \qquad (7.54)$$

Thus by application of the Stochastic Maximum Principle we have

$$-d\lambda = [A(u_0)\lambda + A_u(u_0)\lambda_0 u + \Lambda_m]dt - r dy \tag{7.55}$$

$$\lambda_0^T A_u^T(u_0)q + \lambda_0^T A_{uu}^T(u_0)q_0 u + N_m + \lambda^T A_u^T(u_0)q_0 = 0 \tag{7.56}$$

$$q_m(0) = 0 \tag{7.57}$$

$$\lambda_m(T) = 0 \tag{7.58}$$

which is exactly equations (7.37) through (7.41).

It will henceforth be assumed that the problem described by equations (7.48) through (7.50) is not singular. That is,

$$\lambda_0^T A_{uu}^T(u_0)q_0 \neq 0 \text{ on any interval} \tag{7.59}$$

The optimal control for equations (7.48) through (7.50) is given by an affine feedback law. Equation (7.36) gives

$$u_1 = \frac{-1}{\lambda_0^T A_{uu}^T(u_0)q_0}(\lambda_0^T A_u^T(u_0)q_1 + q_0^T A_u(u_0)\lambda_1) \tag{7.60}$$

By substitution into equations (7.32) and (7.33), the following relations may be deduced

$$\begin{pmatrix} dq_1 \\ d\lambda_1 \end{pmatrix} = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \begin{pmatrix} q_1 \\ \lambda_1 \end{pmatrix} dt + \begin{pmatrix} B^T q_0 \\ -B\lambda_0 \end{pmatrix} dt + \begin{pmatrix} \frac{1}{\sqrt{N_0}} H q_0 \\ r_1 \end{pmatrix} dy_t \tag{7.61}$$

$$q_1(0) = 0 \tag{7.62}$$

$$\lambda_1(T) = 0 \tag{7.63}$$

where

$$\Gamma_{11} = A^T(u_0) - \frac{A_u^T(u_0)q_0\lambda_0^T A_u^T(u_0)}{\lambda_0^T A_{uu}^T(u_0)q_0} \tag{7.64}$$

$$\Gamma_{12} = -\frac{A_u^T(u_0)q_0 q_0^T A_u(u_0)}{\lambda_0^T A_{uu}^T(u_0)q_0} \tag{7.65}$$

$$\Gamma_{21} = \frac{A_u(u_0)\lambda_0\lambda_0^T A_u^T(u_0)}{\lambda_0^T A_{uu}^T(u_0)q_0} \tag{7.66}$$

$$\Gamma_{22} = -A(u_0) + \frac{A_u(u_0)\lambda_0 q_0^T A_u(u_0)}{\lambda_0^T A_{uu}^T(u_0)q_0} \tag{7.67}$$

Differentiating both sides of the relationship

$$\lambda_1 = Rq_1 + \rho_1 \tag{7.68}$$

yields

$$d\lambda_1 = \dot{R}q_1 dt + Rdq_1 + d\rho_1 \tag{7.69}$$

Substitution of equation (7.61) then produces the relationship

$$\begin{aligned}[\Gamma_{21}q_1 + \Gamma_{22}\lambda_1 - B\lambda_0]dt + r_1 dy = \\ \dot{R}q_1 dt + R[\Gamma_{11}q_1 + \Gamma_{12}\lambda_1 + B^T q_0]dt + \\ \frac{1}{\sqrt{N_0}}Hq_0 dy + d\rho_1\end{aligned} \tag{7.70}$$

By substitution of equation (7.68) and collection of terms we have

$$[\Gamma_{21} + \Gamma_{22}R - R\Gamma_{11} - R\Gamma_{12}R]q_1 = \dot{R}q_1 \tag{7.71}$$

$$\Gamma_{22}\rho_1 - B\lambda_0 = R\Gamma_{12}\rho_1 + B^T q_0 + \dot{\rho}_1 \tag{7.72}$$

$$r_1 = \frac{1}{\sqrt{N_0}}Hq_0 \tag{7.73}$$

Thus R is the solution to the Ricatti equation

$$\dot{R} = \Gamma_{21} + \Gamma_{22}R - R\Gamma_{11} - R\Gamma_{12}R \tag{7.74}$$

$$R(T) = 0, \tag{7.75}$$

ρ_1 is the solution to the deterministic differential equation

$$\dot{\rho}_1 = -B\lambda_0 - RBq_0 + (\Gamma_{22} - R\Gamma_{12})\rho_1 \tag{7.76}$$

$$\rho_1(T) = 0, \tag{7.77}$$

and $r_1(t)$ is deterministic and is given by

$$r_1(t) = \frac{1}{\sqrt{N_0}}R(t)Hq_0(t) \tag{7.78}$$

Substituting (7.68) into (7.60) gives

$$\begin{aligned} u_1 &= \frac{-1}{\lambda_0^T A_{uu}^T(u_0)q_0}[(\lambda_0^T A_u^T(u_0) + q_0^T A_u(u_0)R)q_1 + q_0^T A_u(u_0)\rho_1] & (7.79) \\ &= K(t)q_1 + \beta(t) & (7.80) \end{aligned}$$

Similarly, the optimal control for equations (7.51) through (7.53) is given by an affine feedback law. Equation (7.41) gives

$$u_m = \frac{-1}{\lambda_0^T A_{uu}^T(u_0)q_0}(\lambda_0^T A_u^T(u_0)q_m + q_0^T A_u(u_0)\lambda_m + N_m) \qquad (7.81)$$

By substitution into equations (7.37) and (7.38), the following relations may be deduced

$$\begin{pmatrix} dq_m \\ d\lambda_m \end{pmatrix} = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}\begin{pmatrix} q_m \\ \lambda_m \end{pmatrix} dt + \begin{pmatrix} -\frac{A_u^T(u_0)q_0N_m}{\lambda_0^T A_{uu}^T(u_0)q_0} + X_m \\ \frac{A_u(u_0)\lambda_0 N_m}{\lambda_0^T A_{uu}^T(u_0)q_0} - \Lambda_m \end{pmatrix} dt + \begin{pmatrix} \frac{1}{\sqrt{N_0}}Hq_{m-1} \\ r_m \end{pmatrix} dy_t \qquad (7.82)$$

$$\begin{aligned} q_m(0) &= 0 & (7.83) \\ \lambda_m(T) &= 0 & (7.84) \end{aligned}$$

Using the relationship

$$\lambda_m(t) = R(t)q_m(t) + \rho_m(t) \qquad (7.85)$$

we have

$$d\lambda_m = \dot{R}q_m dt + R dq_m + d\rho_m \qquad (7.86)$$

By substitution of equation (7.82), $R(t)$ may be shown to be the solution of the Ricatti equation described by equations (7.74) and (7.75). Indeed, substitution of equation (7.82) into equation (7.86) yields

$$\begin{aligned} &[\Gamma_{21}q_m + \Gamma_{22}\lambda_m + \frac{A_u(u_0)\lambda_0 N_m}{\lambda_0^T A_{uu}^T(u_0)q_0} - \Lambda_m]dt + r_m dy = \\ &\dot{R}q_m dt + R\{(\Gamma_{11}q_m + \Gamma_{12}\lambda_m - \frac{A_u^T(u_0)q_0N_m}{\lambda_0^T A_{uu}^T(u_0)q_0} + X_m)dt \\ &+\frac{1}{\sqrt{N_0}}Hq_{m-1}dy\} + d\rho_m \qquad (7.87) \end{aligned}$$

Substitution equation (7.85),

$$[\Gamma_{21}q_m + \Gamma_{22}(Rq_m + \rho_m) + \frac{A_u(u_0)\lambda_0 N_m}{\lambda_0^T A_{uu}^T(u_0)q_0} - \Lambda_m]dt + r_m dy =$$

$$\dot{R}q_m dt + R\{(\Gamma_{11}q_m + \Gamma_{12}(Rq_m + \rho_m) - \frac{A_u^T(u_0)q_0 N_m}{\lambda_0^T A_{uu}^T(u_0)q_0} + X_m)dt$$

$$+\frac{1}{\sqrt{N_0}}Hq_{m-1}dy\} + d\rho_m \tag{7.88}$$

Collecting terms,

$$[\dot{R} - \Gamma_{22}R + R\Gamma_{11} + R\Gamma_{12}R - \Gamma_{21}]q_m = 0 \tag{7.89}$$

Thus $R(t)$ satisfies the Ricatti equation described by equations (7.74) and (7.75).

The variable ρ_m is somewhat more difficult to determine for $m \geq 2$ than for the cases described above. This is due to the fact that there is no guarantee of separation, and so, ρ must be considered to be stochastically driven [2]. From equation (7.88) collecting terms we have

$$\begin{aligned} d\rho_m &= [\frac{RA_u^T(u_0)q_0 N_m + A_u(u_0)\lambda_0 N_m}{\lambda_0^T A_{uu}^T(u_0)q_0} - RX_m - \Lambda_m + (\Gamma_{22} - R\Gamma_{12})]dt + \\ &+(r_m - \frac{1}{\sqrt{N_0}}RHq_{m-1})dy(t) \end{aligned} \tag{7.90}$$

$$\rho_m(T) = 0 \tag{7.91}$$

And, again as in the case of first order terms, the control u_m may be determined by the substitution of equation (7.85) into equation (7.81). The result is

$$\begin{aligned} u_m &= \frac{-1}{\lambda_0^T A_{uu}^T(u_0)q_0}[(\lambda_0^T A_u^T(u_0) + q_0^T A_u(u_0)R)q_m + \\ &\quad q_0^T A_u(u_0)\rho_m + N_m] \end{aligned} \tag{7.92}$$

$$= K(t)q_m + \beta_m(t) \tag{7.93}$$

It can therefore be seen that by defining

$$u^M(t) = u_0(t) + \sum_{m=1}^{M} \epsilon^m u_m(t) \tag{7.94}$$

$$= u_0(t) + \sum_{m=1}^{M} \epsilon^m (Kq_m + \beta_m) \tag{7.95}$$

an approximation to the optimal control law, u^ϵ is achieved.

The following figure contains a system block diagram illustrating the derivation of the controller u^M for the case in which $M = 1$. In that figure, the plant provides the observations that are used as inputs to both the filter that realizes the Zakai equation and a pair of filters that approximate the output of the Zakai equation (i.e. $z^1 \cong q_0 + \epsilon q_1$). These latter filters, labeled the controller dynamics, are also driven by the controls u_0 and u_1. The control u_0 is the solution to the optimization problem of equations (7.45) through (7.47). Although the control u_1 is the solution to the optimization problem of equations (7.48) through (7.50), that solution is realized as a feedback law. Thus, the near-optimal controller u^1 can be obtained though use of this block diagram by solving the optimization problem of equations (7.45) through (7.47) and the requisite differential equations for R and ρ_1 that determine K and β_1 respectively. Note that since the limit problem is deterministic, the ouput of the differential equations that determine R and ρ_1, and all of the quantities associated with the limit problem (i.e. u_0, λ_0, and q_0) can be computed offline.

The case in which $M > 1$ is somewhat more complicated. In that situation filters that approximate higher order terms in the expansion of the ouput of the Zakai equation are functionally dependent on all of the lower order filters. Moreover, with the exception of the zeroth order term (i.e. q_0), all of these filter outputs are stochastic. For these controllers, only the function K and the parameters associated with the zeroth order filter can be computed offline.

An approximation to the control law of equation (7.95) which is perhaps more closely related to the state of the original problem is the following feedback law

$$\overline{u}^M(q;t) = u_0(t) + K(t)(q - q_0(t)) + \sum_{m=1}^{M} \epsilon^m \beta_m \tag{7.96}$$

The feedback control $\overline{u}^M$ is of course only an approximation to the law u_M. However, the resultant trajectories will differ only in high order terms. A realization of the feedback control $\overline{u}^M$ is illustrated in the following figure for the case in which $M = 1$.

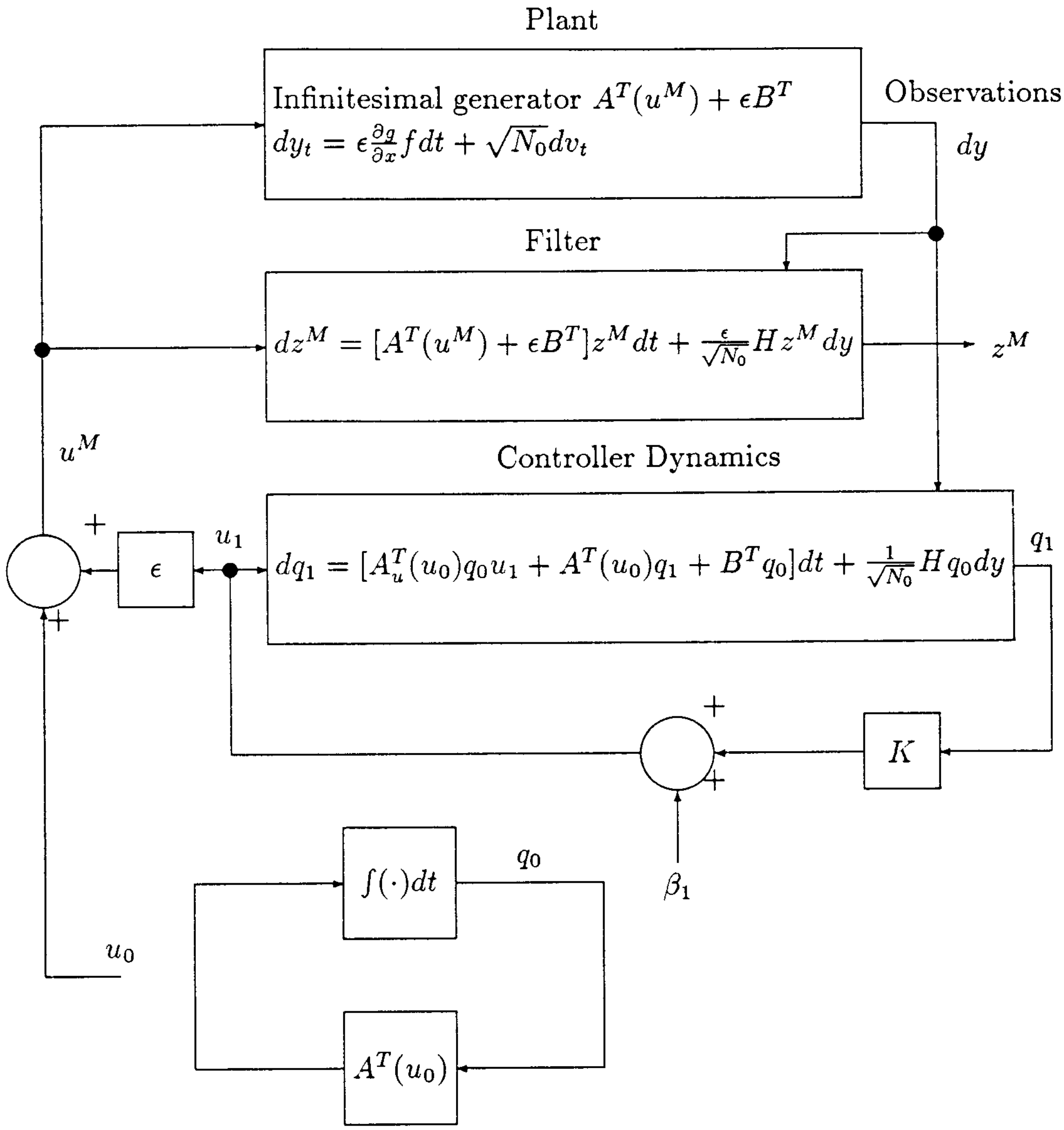


Figure 7.1: System block diagram for the controller u^1

In contrast to the control law u^1, the feedback control $\overline{u}^1$ uses the output of the Zakai equation (z) to determine the near-optimal control. Although this appears to be a great simplification, in reality, if one wishes to obtain higher order approximations to the optimal control u^ϵ, one must build higher order approximations to the Zakai equation. This is the case because the functions β_i are functions of these filter outputs (equation 7.90).

The remarks about offline computations that were made with respect to the control law u^1 are also relevant to the feedback law $\overline{u}^1$. Note also, that both control laws are linear and deterministic if $M = 1$.

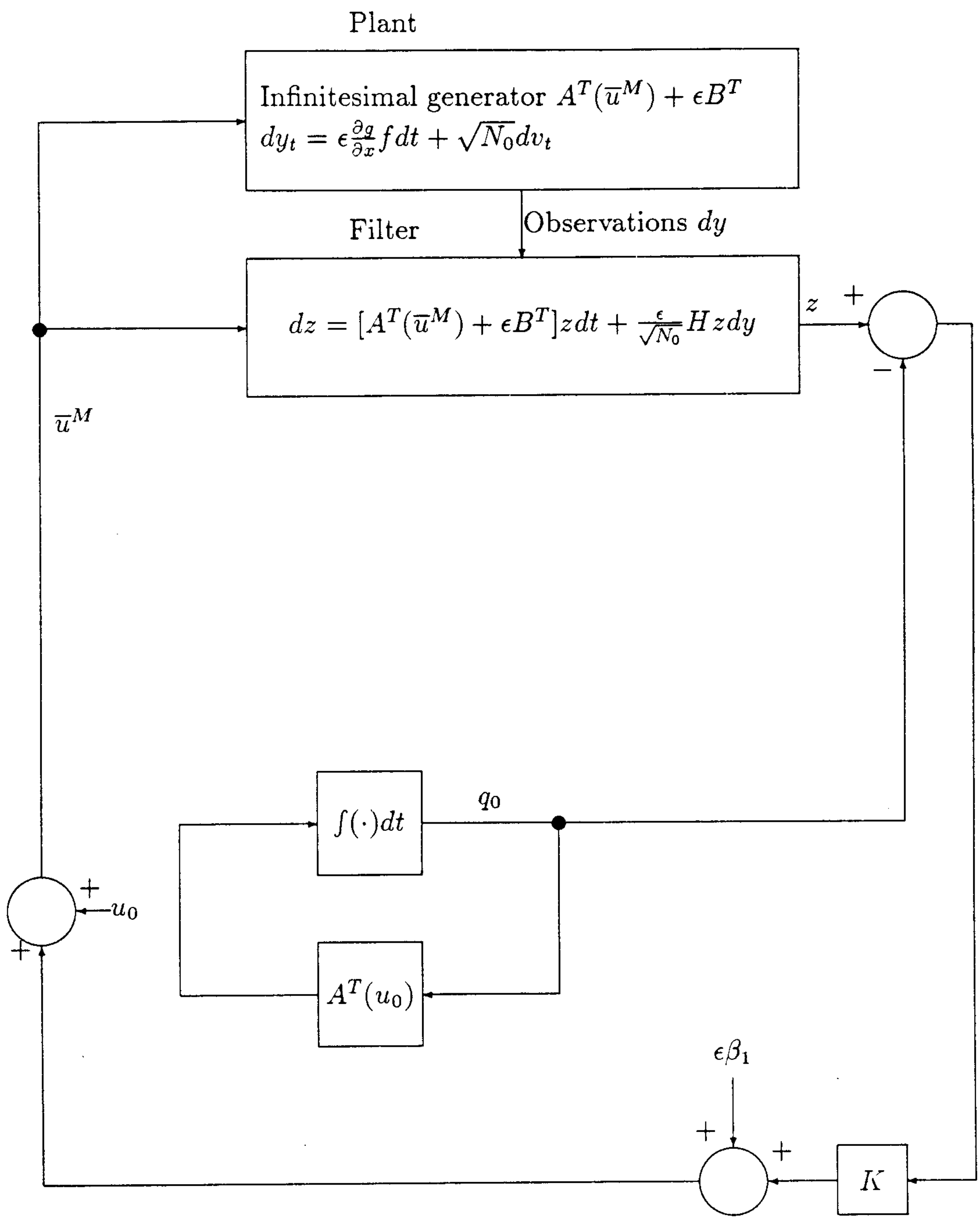


Figure 7.2: System block diagram for the controller $\overline{u}^1$

7.3 Proof of convergence

Theorem 8 *If there exists a solution to the problem of equations (4.1) and (4.2), then*

$$|J(u^\epsilon) - J(u_0)| = O(\epsilon) \tag{7.97}$$

Proof. By Theorem 4 (the Wong-Zakai Correction), equations (7.7) and (7.9) may be rewritten as

$$\dot{q}^\epsilon(t) = [A^T(u^\epsilon(t)) + \epsilon B^T - \frac{\epsilon^2}{2N_0}H^2]q^\epsilon(t) + \frac{\epsilon}{\sqrt{N_0}}Hq^\epsilon(t)\dot{y}(t) \tag{7.98}$$

$$q^\epsilon(0) = \overline{q} \tag{7.99}$$

Let $N^\epsilon(t) = A^T(u^\epsilon(t)) + \epsilon B^T - \frac{\epsilon^2}{2N_0}H^2$. Then the corresponding fundamental transition matrix, Φ^ϵ is described by a Peano-Baker series [7].

$$\Phi^\epsilon(t, t_0) = I + \int_{t_0}^{t} N^\epsilon(\sigma_1)d\sigma_1 + \int_{t_0}^{t} N^\epsilon(\sigma_1)\int_{t_0}^{\sigma_1} N^\epsilon(\sigma_2)d\sigma_2 d\sigma_1 + \cdots \tag{7.100}$$

A new coordinate system will now be introduced in order to simplify the algebra required in applying Theorem 7. Let

$$\tilde{q}^\epsilon(t) = q^\epsilon(t) - \Phi^\epsilon(t, 0)\overline{q} \tag{7.101}$$

Then, by differentiation of both sides of this equation,

$$\dot{\tilde{q}}^\epsilon(t) = \dot{q}^\epsilon(t) - N^\epsilon(t)\Phi^\epsilon(t, 0)\overline{q} \tag{7.102}$$

Substituting equation (7.98) into this equation, we have

$$\dot{\tilde{q}}^\epsilon(t) = N^\epsilon(t)q^\epsilon(t) + \frac{\epsilon}{\sqrt{N_0}}Hq^\epsilon(t)\dot{y}(t) - N^\epsilon(t)\Phi^\epsilon(t, 0)\overline{q} \tag{7.103}$$

Then applying the definition of $\tilde{q}^\epsilon(t)$ of equation (7.101) we obtain

$$\dot{\tilde{q}}^\epsilon(t) = N^\epsilon(t)\tilde{q}^\epsilon(t) + \frac{\epsilon}{\sqrt{N_0}}H\tilde{q}^\epsilon(t)\dot{y}(t) + \frac{\epsilon}{\sqrt{N_0}}H\Phi^\epsilon(t, 0)\overline{q}\dot{y}(t) \tag{7.104}$$

with initial condition

$$\tilde{q}^\epsilon(0) = 0 \tag{7.105}$$

Applying Theorem 7 to obtain an expression for $\tilde{q}^\epsilon$ and then returning to the original coordinates yields

$$\begin{aligned} q^\epsilon(t) &= \{\Phi^\epsilon(t,0) + \frac{\epsilon}{\sqrt{N_0}}\int_0^t \Phi^\epsilon(t,\sigma_1)H\Phi^\epsilon(\sigma_1,0)\dot{y}_{\sigma_1}d\sigma_1 + \\ &+\frac{\epsilon^2}{N_0}\int_0^t \Phi^\epsilon(t,\sigma_1)H\int_0^{\sigma_1}\Phi^\epsilon(\sigma_1,\sigma_2)H\Phi^\epsilon(\sigma_2,0)\dot{y}_{\sigma_2}\dot{y}_{\sigma_1}d\sigma_2 d\sigma_1 + \cdots\}\overline{q} \end{aligned} \tag{7.106}$$

But $A(\cdot)$ is bounded. Thus, we have

$$J(u^\epsilon) = E\{\int_0^T C^T\Phi^\epsilon(t,0)\overline{q}dt + O(\epsilon)\} \tag{7.107}$$

And for an arbitrary control $\tilde{u}$,

$$J(\tilde{u}) = E\{\int_0^T C^T\Phi^{\tilde{u}}(t,0)\overline{q}dt + O(\epsilon)\} \tag{7.108}$$

where

$$\Phi^{\tilde{u}}(t,t_0) = I + \int_{t_0}^t N^{\tilde{u}}(\sigma_1)d\sigma_1 + \int_{t_0}^t N^{\tilde{u}}(\sigma_1)\int_{t_0}^{\sigma_1} N^{\tilde{u}}(\sigma_2)d\sigma_2 d\sigma_1 + \cdots \tag{7.109}$$

and

$$N^{\tilde{u}}(t) = A^T(\tilde{u}(t)) + \epsilon B^T - \frac{\epsilon^2}{2N_0}H^2 \tag{7.110}$$

Thus

$$J(\tilde{u}) - J(u^\epsilon) = E\{\int_0^T C^T[\Phi^{\tilde{u}}(t,0) - \Phi^\epsilon(t,0)]\overline{q}dt + O(\epsilon)\} \tag{7.111}$$

And since $J(u^\epsilon)$ is optimal,

$$J(\tilde{u}) - J(u^\epsilon) \geq 0 \tag{7.112}$$

Combining equations (7.111) and (7.112) and taking the limit as $\epsilon \to 0$, we have

$$E\{\int_0^T C^T \lim_{\epsilon \to 0} [\Phi^{\tilde{u}}(t,0) - \Phi^\epsilon(t,0)]\overline{q}dt\} \geq 0 \tag{7.113}$$

If the problem of optimizing the cost of equation (4.1) subject to the dynamics of equation (4.2) has a solution, the control

$$\tilde{u}(t) = \lim_{\epsilon \to 0} u^{\epsilon}(t) \tag{7.114}$$

satisfies equation (7.113) with equality. That is,

$$\lim_{\epsilon \to 0} J(\tilde{u}) - J(u^{\epsilon})|_{\tilde{u} = \lim_{\epsilon \to 0} u^{\epsilon}} = 0 \tag{7.115}$$

But note that by equation (7.113), the control $\tilde{u}$ also has the property of minimizing

$$E\{\int_0^T \lim_{\epsilon \to 0} C^T \Phi^{\tilde{u}}(t,0)\overline{q}dt\}. \tag{7.116}$$

Thus, using equations (7.109) and (7.110), $\tilde{u}$ is the solution to the problem of minimizing

$$\int_0^T C^T q_0(t)dt \tag{7.117}$$

subject to the dynamics

$$\frac{dq_0(t)}{dt} = A^T(\tilde{u})q_0(t) \tag{7.118}$$

$$q_0(0) = \overline{q} \tag{7.119}$$

That is, $\tilde{u}$ is given by some u_0 which satisfies the necessary conditions of equations (7.27) through (7.31). Applying this observation to equation (7.115) proves the theorem.

It will now be shown that the feedback control $\overline{u}^M$ produces a cost that differs from that of the control u^M only in high order terms. In order to do this, first some preliminary results will be generated. Let

$$q^M \triangleq \sum_{m=0}^{M} \epsilon^m q_m \tag{7.120}$$

and

$$X^M \triangleq [A^T(u^M) + \epsilon B^T]q^M - A^T(u_0)q^M - \sum_{i=1}^{M} A_u^T(u_0)q_0\epsilon^i u_i - \sum_{m=0}^{M} \epsilon^m X_m \tag{7.121}$$

where we have defined

$$X_0 = 0 \tag{7.122}$$

and

$$X_1 = B^T q_0 \tag{7.123}$$

The following lemmas are then valid

Lemma 5 $\forall \alpha \in [1, \infty), \forall t$

$$[E \int_0^T |X^M(t)|^\alpha dt]^{\frac{1}{\alpha}} \leq C_{M,\alpha} \epsilon^{M+1} \tag{7.124}$$

Proof: We first verify that $\forall \beta \in [1, \infty)$,

$$E \int_0^T |q_i|^\beta dt < \infty \tag{7.125}$$

$$E \int_0^T |u_i|^\beta dt < \infty \tag{7.126}$$

From equation (7.61) q_1 is the solution to a linear stochastic differential equation. It therefore has finite moments of all orders. Since u_1 is an affine function of q_1, it has the same property. We have then the required result for X_2 (c.f. equation (7.42)). We can use the same argument for each $X_i, i \geq 2$. We will now prove the lemma.

Expanding the matrix $A^T(u^M)$ in a Taylor series about u_0 (c.f. equation (7.21))

$$A^T(u^M) = A(u_0) + \sum_{k=1}^{\infty} \epsilon^k \sum_{j=1}^{k} \sum_{\substack{l_1 + \cdots + l_j = k \\ l_1, \ldots, l_j \geq 1}} \frac{D^{(j)} A(u_0) u_{l_1} \ldots u_{l_j}}{j!} + O(\epsilon^{M+1}) \tag{7.127}$$

We now collect coefficients of like powers of ϵ in the term $[A^T(u^M) + \epsilon B^T] q^M$. The coefficient of ϵ^0 is

$$A^T(u_0) q_0 \tag{7.128}$$

The coefficients of ϵ^1 are

$$A^T(u_0)q_1 + A_u^T(u_0)q_0u_1 + B^Tq_0 \tag{7.129}$$

The coefficients of ϵ^m are

$$A^T(u_0)q_m + A_u^T(u_0)q_0u_m + X_m \tag{7.130}$$

We have then, by combining the last three equations,

$$\begin{aligned} [A^T(u^M) + \epsilon B^T]q^M &= A^T(u_0)q_0 + \epsilon\{A^T(u_0)q_1 + A_u^T(u_0)q_0u_1 + B^Tq_0\} \\ &\quad \sum_{i=2}^{M} \epsilon^i\{A^T(u_0)q_i + A_u^T(u_0)q_0u_i + X_i\} \\ &\quad +O(\epsilon^{M+1}) \qquad (7.131) \\ &= A^T(u_0)q^M + \sum_{i=1}^{M} A_u^T(u_0)q_0\epsilon^i u_i + \sum_{i=0}^{M} \epsilon^i X_i \\ &\quad +O(\epsilon^{M+1}) \qquad (7.132) \end{aligned}$$

Combining this result with equation (7.121) we have

$$X^M = O(\epsilon^{M+1}) \tag{7.133}$$

Thus the lemma is proven.

Let z^M be the state of the system in response to the control u^M. Thus

$$dz^M = [A^T(u^M) + \epsilon B^T]z^M dt + \frac{\epsilon}{\sqrt{N_0}} H z^M dy \tag{7.134}$$

$$z^M(0) = \overline{q} \tag{7.135}$$

Then the next lemma follows

Lemma 6 $\forall \alpha \in [1, \infty)$

$$E \sup_{0 \le t \le T} |z^M(t) - q^M(t)| \le C_{M,\alpha}\epsilon^{M+1} \tag{7.136}$$

Proof: Let

$$\tilde{z}^M = z^M - q^M \tag{7.137}$$

By differentiating equation (7.137) and substituting equations (7.27), (7.32), (7.37), and (7.134) into the resulting equation we have

$$\begin{aligned} d\tilde{z}^M &= [A^T(u^M) + \epsilon B^T]z^M dt + \frac{\epsilon}{\sqrt{N_0}} H z^M dy \\ &\quad \{-A^T(u_0)q_0 - \epsilon[A^T(u_0)q_1 + A_u^T(u_0)q_0u_1 + B^T q_0]\}dt + \\ &\quad -\frac{\epsilon}{\sqrt{N_0}} H q_0 dy - \sum_{i=2}^{M} \epsilon^i [A^T(u_0)q_i + A_u^T(u_0)q_0u_i + X_i]dt \\ &\quad + \frac{1}{\sqrt{N_0}} \sum_{i=2}^{M} \epsilon^i H q_{i-1} dy \end{aligned} \tag{7.138}$$

Using equation (7.121) and collecting terms yields

$$d\tilde{z}^M = \{[A^T(u^M) + \epsilon B^T]\tilde{z}^M + X^M\}dt + \frac{\epsilon}{\sqrt{N_0}} H(\tilde{z}^M + \epsilon^M q_M)dy \tag{7.139}$$

And applying the defined initial conditions,

$$\tilde{z}^M(0) = 0 \tag{7.140}$$

In order to apply theorem 7 we define

$$dz_c^M = \{[A^T(u^M(t)) + \epsilon B^T]z_c^M + X^M\}dt \tag{7.141}$$

$$z_c^M(0) = 0 \tag{7.142}$$

Using a Peano-Baker series to represent z_c^M and applying lemma 5 we see that

$$E|z_c^M| \leq C\epsilon^{M+1} \tag{7.143}$$

Now we define

$$\hat{z}^M = \tilde{z}^M - z_c^M \tag{7.144}$$

Then by substitution

$$d\hat{z}^M = [A^T(u^M) + \epsilon B^T]\hat{z}^M dt + \frac{\epsilon}{\sqrt{N_0}} H \hat{z}^M dy + \frac{\epsilon H}{\sqrt{N_0}}(z_c^M + \epsilon^M q_M)dy \tag{7.145}$$

$$\hat{z}^M(0) = 0 \tag{7.146}$$

Applying theorem 7 we see that

$$E|\hat{z}^M| \leq C\epsilon^{M+1} \tag{7.147}$$

Therefore

$$\begin{aligned} E|\tilde{z}^M| &= E|\hat{z}^M + z_c^M| & (7.148) \\ &\leq C\epsilon^{M+1} & (7.149) \end{aligned}$$

This is the desired result.

Now the theorem relating the cost of the feedback control to that of the control u^M will be derived.

Theorem 9 *The following estimate is valid*

$$|J(\overline{u}^M) - J(u^M)| \leq C_M\epsilon^{M+1} \tag{7.150}$$

Proof: Let z denote the state of the system in response to the feedback control $\overline{u}^M$.

$$dz = [A^T(\overline{u}^M) + \epsilon B^T]zdt + \frac{\epsilon}{\sqrt{N_0}}Hzdy \tag{7.151}$$

$$z(0) = \overline{q} \tag{7.152}$$

Let $\tilde{z}$ denote the difference between the state z and the estimate of the optimal trajectory, q^M.

$$\tilde{z} = z - q^M \tag{7.153}$$

Applying equations (7.96) and (7.95) to the difference between the controls $\overline{u}^M$ and u^M,

$$\begin{aligned} \tilde{u} &\triangleq \overline{u}^M - u^M & (7.154) \\ &= K[z - \sum_{m=0}^{M} \epsilon^m q_m] & (7.155) \\ &= K(t)\tilde{z}(t) & (7.156) \end{aligned}$$

Thus, $\overline{u}^M = u^M + K\tilde{z}$, and

$$d\tilde{z} = [A^T(u^M + K\tilde{z}) + \epsilon B^T](\tilde{z} + q^M)dt + \frac{\epsilon}{\sqrt{N_0}}H(\tilde{z} + q^M)dy - dq^M \quad (7.157)$$

$$\tilde{z}(0) = 0 \quad (7.158)$$

By direct substitution

$$dq^M = \{[A^T(u^M) + \epsilon B^T]q^M - X^M\}dt + \frac{\epsilon}{\sqrt{N_0}}Hq^{M-1}dy \quad (7.159)$$

Thus,

$$\begin{aligned} d\tilde{z} &= \{[A^T(u^M + K\tilde{z}) + \epsilon B^T](\tilde{z} + q^M) - [A^T(u^M) + \epsilon B^T]q^M + X^M\}dt \\ &\quad + \frac{\epsilon H}{\sqrt{N_0}}(\tilde{z} + \epsilon^M q_M)dy \end{aligned} \quad (7.160)$$

Now, by Taylor's Theorem,

$$A^T(u^M + K\tilde{z}) = A^T(u^M) + [\int_0^1 A_u^T(u^M + \lambda K\tilde{z})d\lambda]K\tilde{z} \quad (7.161)$$

Then, by substitution,

$$d\tilde{z} = [P(t)\tilde{z} + X^M]dt + \frac{\epsilon}{\sqrt{N_0}}H(\tilde{z} + \epsilon^M q_M)dy \quad (7.162)$$

where

$$P(t) = A^T(u^M(t) + K(t)\tilde{z}(t)) + [\int_0^1 A_u^T(u^M(t) + \lambda K(t)\tilde{z}(t))q^M d\lambda]K(t) + \epsilon B^T \quad (7.163)$$

In order to use Theorem 7, the variable z_c is introduced

$$dz_c = [P(t)z_c + X^M]dt \quad (7.164)$$

$$z_c(0) = 0 \quad (7.165)$$

Defining $\hat{z}$ by

$$\hat{z} = \tilde{z} - z_c \quad (7.166)$$

Then,

$$d\hat{z} = P\hat{z}dt + \frac{\epsilon}{\sqrt{N_0}}H\hat{z}dy + \frac{\epsilon}{\sqrt{N_0}}H(z_c + \epsilon^M q_M)dy \quad (7.167)$$

$$\hat{z}(0) = 0 \tag{7.168}$$

And, from equation (7.164),

$$z_c(t) = \int_0^t \Phi(t,\sigma) X^M d\sigma \tag{7.169}$$

where Φ is the fundamental transition matrix for the system matrix $P(t)$.

Using a Peano-Baker series to represent $z_c(t)$ and applying lemma 5 we have

$$E|z_c| \leq C\epsilon^{M+1} \tag{7.170}$$

Applying theorem 7 to equation (7.167),

$$\begin{aligned} \hat{z} &= \frac{\epsilon}{\sqrt{N_0}} \int_0^t \Phi(t,\sigma_1) H(z_c + \epsilon^M q_M) dy_{\sigma_1} + \\ &\quad + \frac{\epsilon^2}{N_0} \int_0^t \Phi(t,\sigma_1) H \int_0^{\sigma_1} \Phi(t,\sigma_2) H(z_c + \epsilon^M q_M) dy_{\sigma_1} dy_{\sigma_2} + \cdots \end{aligned} \tag{7.171}$$

Thus

$$\begin{aligned} E|\tilde{z}| &= E|\hat{z} + z_c| & (7.172) \\ &\leq C\epsilon^{M+1} & (7.173) \end{aligned}$$

Then

$$E|z - z^M| \leq E|z - q^M| + E|q^M - z^M| \tag{7.174}$$

Using the definition of $\tilde{z}$, then,

$$E|z - z^M| \leq E|\tilde{z}| + E|q^M - z^M| \tag{7.175}$$

Applying equations (7.136), (7.173) and (7.174) to the definition of the cost, the theorem follows.

Chapter 8

Singular perturbations

In this section the singular perturbation problem described by equations (4.9) through (4.11) will be examined. As is usually the case when one attempts to find time scale decompositions for such problems, the first step is to separate the variables into a set of variables representing a "slow" system, and a set of variables representing a "fast" system. That is, the variations in time of the slow variables are on the order of unity, while those of the fast variables are of the order $\frac{1}{\epsilon}$; where ϵ is a small parameter.

The term singular perturbations used here is contrasted to that of regular perturbations, which was described in the last section. The problem discussed in this section is of a fundamentally different nature to that of regular perturbations. In the case of regular perturbations, in the limit as $\epsilon \to 0$, the number of differential equations that describe the plant remain the same. In contrast, for the case of singular perturbations, in the limit as $\epsilon \to 0$, the differential equations that describe the fast interactions become algebraic. In fact, the initial conditions for those differential equations are not part of the limit problem.

The method of solution will be dynamic programming. The dynamic programming formulation permits the introduction of the concept of composite feedback [9]. The form of this feedback law is a summation of the feedback that corresponds to the limit problem (i.e. the control law for the slow system) and a corrector term that is a function of the difference between the actual state of the fast system and the limit state of the fast system.

The dynamics of the singular perturbation problem of concern in this paper consist of stochastic differential equations. It will be shown that the limit problem corresponds to the problem of optimizing a stochastic system (of reduced order) over a finite time horizon. In contrast, the fast problem, which produces the corrector term described above, corresponds to the optimization of a (reduced order) deterministic system over an infinite time horizon.

In the first part of this section the problem is separated into fast and slow systems of differential equations and the corresponding Hamilton Jacobi Bellman equation is written. The next part contains a statement of the limit problem. This is followed by a formal derivation of the fast dynamics. The composite control is then introduced. Finally, a proof of convergence is presented.

8.1 Separation into fast and slow variables

In order to obtain a useful decomposition of the singular perturbation problem described by equations (1.16) through (1.22) we must separate slowly changing effects from those effects that change quickly. We turn, therefore, to a consideration of the effects of time scales on the dynamics and on the corresponding probability distributions. The Fokker-Planck equation for the continuous time Markov chain approximation to the singular perturbation problem is

$$\frac{dp}{dt} = (\frac{A^T}{\epsilon} + B^T)p \tag{8.1}$$

where p is a column vector consisting of probabilities of being in a given state at time t; and A and B are matrices defined by equations (3.23) through (3.35) which together form the infinitesimal generator for the continuous time

Markov chain that approximates the diffusion processes x^1 and x^2 given in equations (1.16) and (1.17).

Recall that each block in the matrices A and B as described by equations (3.23) and (3.32) correspond to a fixed value of x^1. Furthermore, since A is block diagonal, the matrix A describes transition rates among the allowed values of the variable x^2 (with x^1 fixed). This is in contrast to the matrix B which describes transition rates among the allowed values of the variable x^1 (with x^2 fixed).

We consider here two time scales: (1) the time scale t used above in describing the evolution of the conditional probabilities, and (2) the time scale $\tau = \frac{t}{\epsilon}$. The τ time scale is also commonly referred to as the "fast" time scale since as $\epsilon \to 0$, $\tau \to \infty$. Behavior in this fast time scale will be described first. In the fast time scale, the effects of strong interactions (i.e. those of $\frac{A}{\epsilon}$) are apparent immediately; while those of the weak interactions (i.e. those of B) become significant only after a long period of time τ. This relationship can be made more apparent by introducing the infinitesimal generator of the Markov process in the τ time scale

$$\frac{dp}{d\tau} = (A^T + \epsilon B^T)p \tag{8.2}$$

From this equation it is clear that as $\epsilon \to 0$, the effects of ϵB become insignificant, leaving the interactions of the A matrix as the major contributor to changes in the conditional probabilities. The A matrix describes transitions in the variable x^2 only. That is, as $\epsilon \to 0$, equation (8.2) becomes

$$\frac{dp}{d\tau} = A^T p \tag{8.3}$$

which is a discrete state space approximation to transitions in the variable x^2 with x^1 fixed. It is, in fact, the discrete state space approximation to equations (1.23) and (1.24), in the limit as $\epsilon \to 0$. As such, it corresponds to the observation that, for the continuous state space problem, as $\epsilon \to 0$, x^1 can be treated as a constant parameter in the dynamics of the variable x^2.

It has thus been shown that, for the Markov chain approximation to the diffusions x^1 and x^2, transitions in the state of the chain that correspond to changes in the variable x^1 are "slowly" varying with respect to transitions of the chain that correspond to changes in the variable x^2. Since in the "fast" time scale y behaves as

$$dy_\tau = \epsilon h(x^1_\tau, x^2_\tau)d\tau + \sqrt{\epsilon N_0}dv_\tau \tag{8.4}$$

y is also "slowly" varying with respect to transitions in the Markov chain that correspond to changes of state of the variable x^2. Moreover, since the control, u, is a function of y, it may also be approximated by a constant in this time scale. Equation (8.3) describes this situation and is an approximation to equation (1.26), the continuous state space representation of this same result. We note that in equation (8.3), the parametric dependence on the control, u, is implicit (A is a function of u), as is the parametric dependence on the state x^1, which appears due to the block diagonal nature of A.

As in the continuous state space case, we require the variable x^2 to converge to a random variable as $\epsilon \to 0$. That is, we require the existence of an invariant measure $\overline{p}$ which satisfies the equation

$$0 = A^T \overline{p} \tag{8.5}$$

and allow as admissible controls only those controls for which there exists an invariant measure. Equation (8.5) is a discrete state space approximation to equation (1.28).

Due to the block diagonal structure of the matrix A, equation (8.5) is equivalent to the requirement that there exist vectors $p^i, i = 0, \ldots, m$, such that

$$A_i^T p^i = 0 \tag{8.6}$$

That is,

$$\frac{1}{\rho^2}\begin{pmatrix} \alpha^i_{11} & \alpha^i_{21} & & 0 \\ \alpha^i_{12} & \alpha^i_{22} & \ddots & \\ & \ddots & \ddots & \alpha^i_{k+1,k} \\ 0 & & \alpha^i_{k,k+1} & \alpha^i_{k+1,k+1} \end{pmatrix} p^i = 0 \tag{8.7}$$

But A_i is an infinitesimal generator of a continuous time Markov process. As such, its columns sum to zero (and hence, the rows of A_i^T sum to zero). The rank of A_i is therefore less than its dimension $k+1$. Moreover, in order to satisfy the ergodicity condition, its rank must be k. Therefore, summing over all the rows of A_i^T in order to obtain k linearly independent equations, we find the condition

$$\frac{1}{\rho^2}\begin{pmatrix} \alpha^i_{11} & \alpha^i_{21} & & & 0 \\ \alpha^i_{12} & \alpha^i_{22} & \ddots & & \\ & \ddots & \ddots & \ddots & \\ 0 & & \alpha^i_{k-1,k} & \alpha^i_{k,k} & \alpha^i_{k+1,k} \end{pmatrix} p^i = 0 \tag{8.8}$$

This equation can be readily solved recursively. From the first row we have

$$p^i_2 = -\frac{\alpha^i_{11}}{\alpha^i_{21}} p^i_1 \tag{8.9}$$

The second row yields

$$\alpha^i_{12} p^i_1 + \alpha^i_{22} p^i_2 + \alpha^i_{32} p^i_3 = 0 \tag{8.10}$$

And by substitution of equation (8.9),

$$p^i_3 = -\frac{1}{\alpha^i_{32}}(\alpha^i_{12} - \frac{\alpha^i_{11}\alpha^i_{22}}{\alpha^i_{21}}) p^i_1 \tag{8.11}$$

In general,

$$p^i_{j+1} = -\frac{1}{\alpha^i_{j+1,j}}(\alpha^i_{j-1,j} p^i_{j-1} + \alpha^i_{j,j} p^i_j) \tag{8.12}$$

So we can represent the kernel of A^T by

$$\overline{V} = \begin{pmatrix} c_0 v^0 & & & 0 \\ & c_1 v^1 & & \\ & & \ddots & \\ 0 & & & c_m v^m \end{pmatrix} \tag{8.13}$$

where v^i are column vectors of dimension $k+1$ such that

$$A_i^T v^i = 0 \tag{8.14}$$

and c_i are scalars. In order for the vectors v^i to represent probabilities, c_i will be chosen so that the sum of the elements of v^i times c_i equals one. That is, for $i = 0, \ldots, m$,

$$\sum_{j=0}^{k} c_i v_j^i = 1 \tag{8.15}$$

The matrix $\overline{V}$ is therefore equivalently defined by

$$\overline{V} \triangleq \begin{pmatrix} \overline{V}_0 & & & 0 \\ & \overline{V}_1 & & \\ & & \ddots & \\ 0 & & & \overline{V}_m \end{pmatrix} \tag{8.16}$$

where $\overline{V}_i$ are column vectors of dimension $k+1$ such that, for $i = 0, \ldots, m$,

$$A_i^T \overline{V}_i = 0 \tag{8.17}$$

$$\underline{1}\overline{V}_i = 1 \tag{8.18}$$

where $\underline{1}$ is the row vector of ones, $(1, \ldots, 1)$, of dimension $k+1$.

From equation (8.17) it is evident that each vector $\overline{V}_i$ is an invariant measure for the corresponding block of A. But we have previously shown that blocks of A correspond to variations in the variable x^2 in the original diffusion. We shall see that $\overline{V}_i$ plays a role that is analogous to the invariant measure over the ergodic set associated with the variable x^2.

We have thus described the behavior of the Markov chain approximation to the original diffusions in the fast time scale. Consider now the behavior of the state of the Zakai equation

$$dq(t) = (\frac{A^T(u)}{\epsilon} + B^T)q(t)dt + \frac{1}{\sqrt{N_0}} Hq(t)dy(t) \tag{8.19}$$

in the fast time scale. In the fast time scale equation (8.19) has the form

$$dq(\tau) = (A^T(u) + \epsilon B^T)q(\tau)d\tau + \frac{1}{\sqrt{N_0}}\sqrt{\epsilon}Hq(\tau)dy(\tau) \tag{8.20}$$

In the limit as $\epsilon \to 0$ one would expect equation (8.20) to approach

$$dq(\tau) = A^T(u)q(\tau)d\tau \tag{8.21}$$

Equation (8.21) is consistent with the continuous state space formulation of equations (1.26) and (1.25) inasmuch as in this time scale as $\epsilon \to 0$, $dy = 0$. That is, y is approximately constant. Since this implies no observations, the filter that determines the state of the system given observations must depend on the infinitesimal generator that describes the system dynamics only. Here again we note that in equation (8.21), the parametric dependence on the control, u, is implicit (A is a function of u), as is the parametric dependence on the state x^1, which appears due to the block diagonal nature of A. This function form is consistent with equation (1.26), describing the continuous state space situation.

Behavior of the system in the t time scale will now be examined. The t time scale may also be referred to as the slow time scale. In this time scale the effects of the weak interactions occur "sooner" than they would in the fast time scale. For example, if the units of τ are weeks and $\epsilon = \frac{1}{52}$, then the units of t are years.

Consider the limit problem in the t time scale as described by equations (1.29) through (1.31). The Zakai equation for this problem is given by

$$d\vartheta_t = \mathcal{L}^*\vartheta_t dt + \frac{1}{\sqrt{N_0}}h\vartheta_t dy_t \tag{8.22}$$

$$= f(x^1, \overline{x^2})\vartheta_t dt + \frac{1}{\sqrt{N_0}}h(x^1, \overline{x^2})\vartheta_t dy_t \tag{8.23}$$

and was described in the introduction. Once we have found a separation of the process which is a discrete state space approximation to the process ϑ_t, we will show that for the resulting slow process, in the limit as $\epsilon \to 0$ (i.e. for time horizons that are long with respect to the fast process, but short with respect to the slow process), the influence of x^2 is felt only in an average sense.

We now turn to the problem of finding a useful separation of time scales. That is, we seek a decomposition of the state space (i.e. the solution space of the Zakai equation) into rapidly varying processes and slowly varying processes. There are many feasible separations of time scales. However, only the one that directly reflects the underlying dynamics and observations of equations (1.16) through (1.21) will be of use in the optimization problem.

We will first find a process that describes the slowly varying process. As one might expect, that process will correspond to the filtered version of x^1. The solution to the Zakai equation for the dynamics of equations (1.16) through (1.21) is an unnormalized version of

$$\zeta_t = \Pr\{x_t^1 \cap x_t^2 | \mathcal{Y}_t\} \tag{8.24}$$

where $\mathcal{Y}_t$ is the $\sigma-$algebra of observations over the interval $[0, t]$. This process may equivalently be expressed as

$$\zeta_t = \Pr\{x_t^2 | x_t^1 \cap \mathcal{Y}_t\} \cdot \Pr\{x_t^1 | \mathcal{Y}_t\} \tag{8.25}$$

The term $\Pr\{x_t^1 | \mathcal{Y}_t\}$ represents the behavior of the filtered version of x^1. Due to the slowly varying nature of both x^1 and y_t, this term should be a slowly varying quantity. It can be related directly to the diffusion ζ_t as follows:

$$\Pr\{x_t^1 | \mathcal{Y}_t\} = \int_{\mathcal{R}} \Pr\{x_t^1 \cap x_t^2 | \mathcal{Y}_t\} dx^2 \tag{8.26}$$

This term is our slow process.

Turning now to the determination of the fast process, we see that the only term in ζ_t that is quickly varying with respect to both x^1 and y is $\Pr\{x_t^2 | x_t^1 \cap \mathcal{Y}_t\}$. In the limit as $\epsilon \to 0$, we will see that $\Pr\{x_t^2 | x_t^1 \cap \mathcal{Y}_t\}$ converges to the invariant measure of x^2, where x^1 enters parametrically. So we have the decomposition

$$\zeta_t = \Pr\{x_t^2 | x_t^1 \cap \mathcal{Y}_t\} \cdot \int_{\mathcal{R}} \Pr\{x_t^1 \cap x_t^2 | \mathcal{Y}_t\} dx^2 \tag{8.27}$$

A representation of the slow and fast variables for the discrete state space approximation will now be found. Consider first the slow variable given by equation (8.26). The integration can be performed for the discrete state space problem as a matrix multiplication. We first define the matrix T as

$$T \triangleq \begin{pmatrix} T_0 & & & 0 \\ & T_1 & & \\ & & \ddots & \\ 0 & & & T_m \end{pmatrix} \tag{8.28}$$

where $T_i = (1 \ldots 1), i = 0, \ldots m$ are row vectors of dimension $k+1$. Using this definition and that of $q(t)$, we have

$$\Pr\{x_t^1 | \mathcal{Y}_t\} = \frac{Tq(t)}{\tilde{e}Tq(t)} \tag{8.29}$$

where $\tilde{e} = (1 \ldots 1)$ is a row vector of length $m+1$. Thus $\tilde{e}Tq(t)$ is a normalizing scalar that arises due to the fact that $q(t)$ is an **unnormalized** probability. The analog of equation (8.27) is

$$\frac{q_t}{\tilde{e}Tq_t} = \Pr\{x_t^2 | x_t^1 \cap \mathcal{Y}_t\} \cdot \frac{Tq_t}{\tilde{e}Tq_t} \tag{8.30}$$

We now seek a definition of the fast process. First we define

$$V(t) \triangleq \begin{pmatrix} V_0(t) & & & 0 \\ & V_1(t) & & \\ & & \ddots & \\ 0 & & & V_m(t) \end{pmatrix} \tag{8.31}$$

where $V_i(t), i = 0, \ldots, m$ are column vectors of dimension $k+1$ that are described by $V_i(t) = \Pr\{x_t^2 | x_t^1 = \beta_i \cap \mathcal{Y}_t\}$. Equation (8.27) then becomes (for the discrete state space approximation):

$$\frac{q(t)}{\tilde{e}Tq(t)} = V(t)\frac{Tq(t)}{\tilde{e}Tq(t)} \tag{8.32}$$

Or, equivalently,

$$q(t) = V(t)Tq(t) \tag{8.33}$$

The slow variables will be denoted by $\eta(t)$ and are defined by

$$\eta(t) = Tq(t) \tag{8.34}$$

The fast process is, of course, $V(t)$.

We will now find the differential equation for $\eta(t)$ and $V(t)$. Premultiplying both sides of the Zakai equation (8.19) by T,

$$Tdq(t) = T(\frac{A^T}{\epsilon}(u) + B^T)q(t)dt + \frac{1}{\sqrt{N_0}}THq(t)dy(t) \tag{8.35}$$

But since A is the infinitesimal generator of a Markov Chain,

$$TA^T = 0 \tag{8.36}$$

By substitution of equations (8.36), (8.34), and (8.33) into equation (8.35),

$$d\eta(t) = TB^T V(t)\eta(t)dt + \frac{1}{\sqrt{N_0}}THV(t)\eta(t)dy(t) \tag{8.37}$$

The derivation for $dV(t)$ is somewhat more lengthy. Suppose that for $i = 0, \ldots, m$,

$$dV_i(t) = [\frac{A_i^T(u)V_i(t)}{\epsilon} + M_i(t)]dt + \sigma_i(t)dy(t) \tag{8.38}$$

$$d\eta(t) = a(t)dt + \Gamma(t)dy(t) \tag{8.39}$$

where $a(t) = TB^T V(t)\eta(t)$ and $\Gamma(t) = \frac{1}{\sqrt{N_0}}THV(t)\eta(t)$.

Using equations (8.33), (8.34), and (8.38) we have

$$dq(t) = d(V(t)\eta(t)) \tag{8.40}$$

$$= \begin{pmatrix} d(V_0(t)\eta_0(t)) \\ \vdots \\ d(V_m(t)\eta_m(t)) \end{pmatrix} \tag{8.41}$$

That is, for $i = 0, \ldots, m$,

$$dq_i(t) = [dV_i(t)\eta_i(t) + V_i(t)d\eta_i(t) + \frac{1}{2}(2\sigma_i(t)\Gamma_i(t))]dt \tag{8.42}$$

where $q_i(t)$ is a column vector of dimension $k+1$ corresponding to the i^{th} group of components in $q(t)$:

$$q(t) = \begin{pmatrix} q_0(t) \\ q_1(t) \\ \vdots \\ q_m(t) \end{pmatrix} \tag{8.43}$$

$\eta_i(t)$ and $\Gamma_i(t)$ are scalars, and $V_i(t)$ and $\sigma_i(t)$ are column vectors of dimension $k+1$.

In order to evaluate equation (8.38) we must determine the form of the terms $TB^T V(t)\eta(t)$ and $THV(t)\eta(t)$. We have

$$TB^T V(t)\eta(t) =$$

$$\begin{pmatrix} T_0 & & 0 \\ & \ddots & \\ 0 & & T_m \end{pmatrix} \begin{pmatrix} -B_0^+ & B_1^- & & 0 \\ B_0^+ & B_1 & \ddots & \\ & \ddots & \ddots & B_m^- \\ 0 & & B_{m-1}^+ & -B_m^- \end{pmatrix}$$

$$\begin{pmatrix} V_0(t) & & 0 \\ & \ddots & \\ 0 & & V_m(t) \end{pmatrix} \begin{pmatrix} \eta_0(t) \\ \vdots \\ \eta_m(t) \end{pmatrix}$$

That is,

$$TB^T V(t)\eta(t) = \begin{pmatrix} (TB^T V(t)\eta(t))_0 \\ (TB^T V(t)\eta(t))_1 \\ \vdots \\ (TB^T V(t)\eta(t))_{m-1} \\ (TB^T V(t)\eta(t))_m \end{pmatrix} \tag{8.44}$$

where

$$(TB^T V(t)\eta(t))_0 = -T_0 B_0^+ V_0(t)\eta_0(t) + T_0 B_1^- V_1(t)\eta_1(t) \tag{8.45}$$

$$(TB^T V(t)\eta(t))_1 = T_1 B_0^+ V_0(t)\eta_0(t) + T_1 B_1 V_1(t)\eta_1(t) + T_1 B_2^- V_2(t)\eta_2(t) \tag{8.46}$$

$$\vdots$$

$$\begin{aligned}(TB^TV(t)\eta(t))_{m-1} &= T_{m-1}B^+_{m-2}V_{m-2}(t)\eta_{m-2}(t) + \\ &\quad T_{m-1}B_{m-1}V_{m-1}(t)\eta_{m-1}(t) + \\ &\quad T_{m-1}B^-_mV_m(t)\eta_m(t) \qquad (8.47)\\ (TB^TV(t)\eta(t))_m &= -T_mB^+_{m-1}V_{m-1}(t)\eta_{m-1}(t) + \\ &\quad T_mB^-_mV_m(t)\eta_m(t) \qquad (8.48)\end{aligned}$$

We also have

$$THV(t)\eta(t) = \begin{pmatrix} T_0 & & 0 \\ & \ddots & \\ 0 & & T_m \end{pmatrix}\begin{pmatrix} H_0 & & 0 \\ & \ddots & \\ 0 & & H_m \end{pmatrix}\begin{pmatrix} V_0(t) & & 0 \\ & \ddots & \\ 0 & & V_m(t) \end{pmatrix}\begin{pmatrix} \eta_0(t) \\ \vdots \\ \eta_m(t) \end{pmatrix} \qquad (8.49)$$

where H_i are diagonal matices corresponding to the i^{th} blocks of the H matrix. We have, then

$$THV(t)\eta(t) = \begin{pmatrix} T_0H_0V_0(t)\eta_0(t) \\ \vdots \\ T_mH_mV_m(t)\eta_m(t) \end{pmatrix} \qquad (8.50)$$

From equations (8.42), (8.38), and (8.39) we have

$$\begin{aligned}dq_i(t) &= [\frac{A_i^TV_i(t)}{\epsilon} + M_i(t)]\eta_i(t)dt + \sigma_i(t)\eta_i(t)dy(t) + \\ &\quad V_i(t)[a_i(t)dt + \Gamma_i(t)dy(t)] + \sigma_i(t)\Gamma_i(t)dt \qquad (8.51)\end{aligned}$$

Regrouping terms,

$$\begin{aligned}dq_i(t) &= [\frac{A_i^TV_i(t)}{\epsilon}\eta_i(t) + M_i(t)\eta_i(t) + V_i(t)a_i(t) + \sigma_i(t)\Gamma_i(t)]dt + \\ &\quad [\sigma_i(t)\eta_i(t) + V_i(t)\Gamma_i(t)]dy(t) \qquad (8.52)\end{aligned}$$

But $q(t)$ is the solution to the Zakai equation (8.19). Therefore we must also have

$$dq(t) = (\frac{A^T}{\epsilon} + B^T)q(t)dt + \frac{1}{\sqrt{N_0}}Hq(t)dy(t) \qquad (8.53)$$

Equating the diffusion terms in equations (8.52) and (8.19),

$$\frac{1}{\sqrt{N_0}} H_i q_i(t) = \sigma_i(t)\eta_i(t) + V_i(t)\Gamma_i(t) \tag{8.54}$$

Using the definition of $\Gamma(t)$ and equation (8.50)

$$\frac{1}{\sqrt{N_0}} H_i q_i(t) = [\sigma_i(t) + \frac{1}{\sqrt{N_0}} V_i(t) T_i H_i V_i(t)]\eta_i(t) \tag{8.55}$$

Now applying equations (8.33) and (8.34) we have

$$\frac{1}{\sqrt{N_0}} H_i V_i(t)\eta_i(t) = (\sigma_i(t) + \frac{1}{\sqrt{N_0}} V_i(t) T_i H_i V_i(t))\eta_i(t) \tag{8.56}$$

Solving for $\sigma_i(t)$,

$$\sigma_i(t) = \frac{1}{\sqrt{N_0}} (I - V_i(t) T_i) H_i V_i(t) \tag{8.57}$$

Or, equivalently

$$\sigma(t) = \frac{1}{\sqrt{N_0}} [I - V(t)T] H V(t) \tag{8.58}$$

We need now only to find $M_i(t)$ in order to completely determine the form of the diffusion $V_i(t)$. First we note that for $i = 2, \ldots, m$,

$$(B^T q(t))_0 = (B^T V(t)\eta(t))_0 \tag{8.59}$$

$$(B^T q(t))_1 = (B^T V(t)\eta(t))_1 \tag{8.60}$$

$$\vdots$$

$$(B^T q(t))_i = (B^T V(t)\eta(t))_i \tag{8.61}$$

$$\vdots$$

$$(B^T q(t))_m = (B^T V(t)\eta(t))_m \tag{8.62}$$

And therefore,

$$(B^T q(t))_0 = -B_0^+ V_0(t)\eta_0(t) + B_1^- V_1(t)\eta_1(t) \tag{8.63}$$

$$(B^T q(t))_1 = B_0^+ V_0(t)\eta_0(t) + B_1 V_1(t)\eta_1(t) + B_2^- V_2(t)\eta_2(t) \tag{8.64}$$

$$\vdots$$

$$(B^T q(t))_i = B_{i-1}^+ V_{i-1}(t)\eta_{i-1}(t) + B_i V_i(t)\eta_i(t) +$$

$$B_{i+1}^{-}V_{i+1}(t)\eta_{i+1}(t) \tag{8.65}$$

$$\vdots$$

$$(B^Tq(t))_m = B_{m-1}^{+}V_{m-1}(t)\eta_{m-1}(t) + B_m^{-}V_m(t)\eta_m(t) \tag{8.66}$$

We also have

$$(Hq(t))_i = (HV(t)\eta(t))_i = H_iV_i(t)\eta_i(t) \tag{8.67}$$

Now equating the drift terms in equations (8.52) and (8.19), and substituting terms using equations (8.57) and (8.44), we have (for $i = 1, \ldots, m-1$),

$$\begin{aligned}
&\frac{A_0^TV_0(t)}{\epsilon}\eta_0(t) + M_0(t)\eta_0(t) + \\
&\quad V_0(t)[-T_0B_0^{+}V_0(t)\eta_0(t) + T_0B_1^{-}V_1(t)\eta_1(t)] + \\
&\quad \frac{1}{\sqrt{N_0}}[I - V_0(t)T_0]H_0V_0(t)\cdot\frac{1}{\sqrt{N_0}}T_0H_0V_0(t)\eta_0(t) \\
&= \frac{A_0^TV_0(t)}{\epsilon}\eta_0(t) + [-B_0^{+}V_0(t)\eta_0(t) + B_1^{-}V_1(t)\eta_1(t)]
\end{aligned} \tag{8.68}$$

$$\vdots$$

$$\begin{aligned}
&\frac{A_i^TV_i(t)}{\epsilon}\eta_i(t) + M_i(t)\eta_i(t) + \\
&\quad V_i(t)[T_iB_{i-1}^{+}V_{i-1}(t)\eta_{i-1}(t) + T_iB_iV_i(t)\eta_i(t) + T_iB_{i+1}^{-}V_{i+1}(t)\eta_{i+1}(t)] + \\
&\quad \frac{1}{\sqrt{N_0}}[I - V_i(t)T_i]H_iV_i(t)\cdot\frac{1}{\sqrt{N_0}}T_iH_iV_i(t)\eta_i(t) \\
&= \frac{A_i^TV_i(t)}{\epsilon}\eta_i(t) + \\
&\quad [-B_{i-1}^{+}V_{i-1}(t)\eta_{i-1}(t) + B_iV_i(t)\eta_i(t) + B_{i+1}^{-}V_{i+1}(t)\eta_{i+1}(t)]
\end{aligned} \tag{8.69}$$

$$\vdots$$

$$\begin{aligned}
&\frac{A_m^TV_m(t)}{\epsilon}\eta_m(t) + M_m(t)\eta_m(t) + \\
&\quad V_m(t)[T_mB_{m-1}^{+}V_{m-1}(t)\eta_{m-1}(t) - T_mB_m^{-}V_m(t)\eta_m(t)] + \\
&\quad \frac{1}{\sqrt{N_0}}[I - V_m(t)T_m]H_mV_m(t)\cdot\frac{1}{\sqrt{N_0}}T_mH_mV_m(t)\eta_m(t) \\
&= \frac{A_m^TV_m(t)}{\epsilon}\eta_m(t) + [-B_{m-1}^{+}V_{m-1}(t)\eta_{m-1}(t) - B_m^{-}V_m(t)\eta_m(t)]
\end{aligned} \tag{8.70}$$

Solving these equations for $M_j(t), j = 0, \ldots, m$,

$$
\begin{aligned}
M_0(t) = \\
& [I - V_0(t)T_0][-B_0^+ V_0(t)\eta_0(t) + B_1^- V_1(t)\eta_1(t)](\eta_0(t))^{-1} \\
& -\frac{1}{N_0}[I - V_0(t)T_0]H_0 V_0(t) T_0 H_0 V_0(t) \qquad (8.71)
\end{aligned}
$$

$\vdots$

$$
\begin{aligned}
M_i(t) = \\
& [I - V_i(t)T_i][B_{i-1}^+ V_{i-1}(t)\eta_{i-1}(t) + B_i V_i(t)\eta_i(t) + B_{i+1}^- V_{i+1}(t)\eta_{i+1}(t)](\eta_i(t))^{-1} \\
& -\frac{1}{N_0}[I - V_i(t)T_i]H_i V_i(t) T_i H_i V_i(t) \qquad (8.72)
\end{aligned}
$$

$\vdots$

$$
\begin{aligned}
M_m(t) = \\
& [I - V_m(t)T_m][B_{m-1}^+ V_{m-1}(t)\eta_{m-1}(t) - B_m^- V_m(t)\eta_m(t)](\eta_m(t))^{-1} \\
& -\frac{1}{N_0}[I - V_m(t)T_m]H_m V_m(t) T_m H_m V_m(t) \qquad (8.73)
\end{aligned}
$$

Equations (8.71) through (8.73) can be written in a matrix form. Let

$$
M(t) = \begin{pmatrix} M_0(t) & & 0 \\ & \ddots & \\ 0 & & M_m(t) \end{pmatrix} \qquad (8.74)
$$

Then we have

$$
M(t) = [I - V(t)T][E(t) - \frac{1}{N_0}(HV(t))(THV(t))] \qquad (8.75)
$$

where the matrix $E(t)$ is defined by

$$
E(t) = \begin{pmatrix} (B^T V(t)\eta(t))_0(\eta_0(t))^{-1} & & 0 \\ & \ddots & \\ 0 & & (B^T V(t)\eta(t))_m(\eta_m(t))^{-1} \end{pmatrix} \qquad (8.76)
$$

By substitution of equations (8.71) through (8.73) and equation (8.57) into equation (8.38) we have

$$
dV(t) \quad = \quad \{\frac{A^T}{\epsilon}V(t) + [I - V(t)T][E(t) - \frac{1}{N_0}(HV(t))(THV(t))]\}dt +
$$

$$\frac{1}{\sqrt{N_0}}(I - V(t)T)HV(t)dy(t) \tag{8.77}$$

In summary, then, the singular perturbations problem has the following decomposition:

Minimize

$$E\{\int_0^T C^T V(t)\eta(t)dt\} \tag{8.78}$$

subject to the dynamics

$$d\eta(t) = TB^T V(t)\eta(t)dt + \frac{1}{\sqrt{N_0}} THV(t)\eta(t)dy(t) \tag{8.79}$$

For $i = 0, \ldots, m$,

$$dV_i(t) = [\frac{A_i^T}{\epsilon} V_i(t) + M_i(t)]dt + \frac{1}{\sqrt{N_0}}[I - V_i(t)T_i]H_i V_i(t)dy(t) \tag{8.80}$$

where

$$V(t) = \begin{pmatrix} V_0(t) & & 0 \\ & \ddots & \\ 0 & & V_m(t) \end{pmatrix} \tag{8.81}$$

and $M_i(t)$ are defined by equations (8.71) through (8.73).

The limit problem is now clear. As $\epsilon \to 0$ the problem becomes one of minimizing

$$E \int_0^T C^T \overline{V}(u)\eta(t)dt \tag{8.82}$$

subject to the dynamics

$$d\eta(t) = TB^T \overline{V}(u)\eta(t)dt + \frac{1}{\sqrt{N_0}} TH\overline{V}(u)\eta(t)dy(t) \tag{8.83}$$

and the constraints

$$A^T(u)\overline{V}(u) = 0 \tag{8.84}$$

$$T\overline{V}(u) = I \tag{8.85}$$

Equation (8.84) is obtained formally by taking the limit as $\epsilon \to 0$ in equation (8.80). It indicates that in the long term, the states $V(t)$ evolve towards the kernel of A^T. That is, in the limit they become the invariant measures over the states that correspond to the process x^2. Equation (8.85) is simply the requirement that the vectors $V_i(t)$ correspond to (normalized) probabilities.

Equation (8.83) is the Zakai equation that is the discrete analog of the Zakai equation for the continuous state space system of equations

$$dx_t^1 = f(x^1, \overline{x^2})dt \quad (8.86)$$

$$dy_t = h(x^1, \overline{x^2})dt + \sqrt{N_0}dv_t \quad (8.87)$$

That is, it is the analog of equation (8.23). In the case of equation (8.83) the averaging over the ergodic sets of x^2 is accomplished by premultiplication by T and by post multiplication by $\overline{V}$, the invariant measure. The functional dependence on f and h is captured by the presence of the matrices B and H in equation (8.83).

We note here that the dimension of the limit dynamics is considerably smaller than those of the original dynamics. The dimension of η is, in fact, $m+1$, which is consistent with the discretization of the slowly changing variable x^1.

The dimension of each vector V_i is $k+1$, which is consistent with the discretization of the "fast variable," x^2. However, in contrast to $\eta(t)$, $V_i(t)$ is a **normalized** probability. As such, its elements are not linearly independent. In the next section we will remove the linear dependence, resulting in the representation of the "fast variables" by $m+1$ vectors, each one of dimension k. This representation will yield a transformation of the original problem into one whose dimension is $m+1$ (the dimension of η) $+(m+1)k$ (the dimension of the linearly independent representation of V). That is, the dimension of the state variables in the transformed problem will be $(m+1)(k+1)$, which is the same as that of q, the solution of the Zakai equation.

8.2 Reduced Dimensional Representation

As noted at the end of the last section, equation (8.80) describes a set of linearly dependent equations. This is evident from the observation that $V_i(t)$ is a probability vector:

$$V_i(t) = \Pr\{x_t^2 | x_t^1 = \beta_i \cap \mathcal{Y}_t\} \tag{8.88}$$

Therefore, for any $i = 0, \ldots, m$, $T_i V_i = 1$ and so, $T_i dV_i(t) = 0$. That is, the sum of all the (scalar) equations represented by the (vector) equation (8.80) is zero.

We can readily verify by direct calculation that equation (8.80) represents a set of linearly dependent equations. From equations (8.36), (8.71), (8.72), (8.73), and (8.80) we see that equation (8.80) is of the form

$$dV_i(t) = [I - V_i(t)T_i]\varrho_i(t) \tag{8.89}$$

for some vector $\varrho_i(t)$. Summing over the elements of $dV_i(t)$ is equivalent to premultiplication by T_i:

$$\begin{aligned} T_i dV_i(t) &= T_i[I - V_i(t)T_i]\varrho_i(t) && (8.90)\\ & \quad [T_i - (T_i V_i(t))T_i]\varrho_i(t) && (8.91)\\ &= 0 && (8.92) \end{aligned}$$

since $T_i V_i(t) = 1$.

Equivalently, then, the last element of $V_i(t)$ is determined from the first k elements. In order to reduce the number of state variables representing the fast process, we introduce the following transformation:

$$\mu_i(t) = S_i V_i(t) \tag{8.93}$$

and

$$V_i(t) = e_i + W_i \mu_i(t) \tag{8.94}$$

where S_i is a matrix of dimension $(k) \times (k+1)$ defined by

$$S_i = (I_{k\times k}|0) \tag{8.95}$$

W_i is a matrix of dimension $(k+1) \times (k)$ defined by

$$W_i = \begin{pmatrix} & I_{k\times k} & \\ & & \\ -1 & \ldots & -1 \end{pmatrix} \tag{8.96}$$

and e_i is a column vector of length $(k+1)$ defined by

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{8.97}$$

Equation (8.93) defines a column vector, $\mu_i(t), i = 0, \ldots, m$, of length k with components equal to the first k components of $V_i(t)$. These vectors, $\mu_i(t)$, will be our linearly independent set of "fast states." We can recover $V_i(t)$ from $\mu_i(t)$ unambiguously by use of equation (8.94).

The above transformations can also be written in matrix form:

$$\mu(t) = SV(t) \tag{8.98}$$

$$V(t) = e + W\mu(t) \tag{8.99}$$

where we have defined the block diagonal matrices

$$S \triangleq \begin{pmatrix} S_0 & & 0 \\ & \ddots & \\ 0 & & S_m \end{pmatrix} \tag{8.100}$$

$$W \triangleq \begin{pmatrix} W_0 & & 0 \\ & \ddots & \\ 0 & & W_m \end{pmatrix} \tag{8.101}$$

$$e \triangleq \begin{pmatrix} e_0 & & 0 \\ & \ddots & \\ 0 & & e_m \end{pmatrix} \tag{8.102}$$

$$\mu(t) \triangleq \begin{pmatrix} \mu_0(t) & & 0 \\ & \ddots & \\ 0 & & \mu_m(t) \end{pmatrix} \tag{8.103}$$

The dynamics for $\mu_i(t)$ are obtained from equation (8.80) by first premultiplying by S_i

$$d\mu_i(t) = S_i[\frac{A_i^T}{\epsilon}V_i(t) + M_i(t)]dt + \frac{1}{\sqrt{N_0}}S_i[I - V_i(t)T_i]H_iV_i(t)dy(t) \tag{8.104}$$

and then by substituting equation (8.94)

$$\begin{aligned} d\mu_i(t) &= S_i\{[\frac{A_i^T}{\epsilon}][e_i + W_i\mu_i(t)] + M_i(t)\}dt + \\ &\quad \frac{1}{\sqrt{N_0}}S_i\{I - [e_i + W_i\mu_i(t)]T_i\}H_i[e_i + W_i\mu_i(t)]dy(t) \end{aligned} \tag{8.105}$$

Equation (8.105) can also be written in matrix form as follows:

$$\begin{aligned} d\mu(t) &= S\{[\frac{A^T}{\epsilon}][e + W\mu(t)] + M(t)\}dt + \\ &\quad \frac{1}{\sqrt{N_0}}S\{I - [e + W\mu(t)]T\}H[e + W\mu(t)]dy(t) \end{aligned} \tag{8.106}$$

But from the definitions of S, W, and e,

$$SW = I \tag{8.107}$$

and

$$Se = 0 \tag{8.108}$$

Thus,

$$\begin{aligned} d\mu(t) &= S\{[\frac{A^T}{\epsilon}][e + W\mu(t)] + M(t)\}dt + \\ &\quad \frac{1}{\sqrt{N_0}}\{S - \mu(t)T\}H[e + W\mu(t)]dy(t) \end{aligned} \tag{8.109}$$

where in the definition of $M(t)$ of equations (8.71) through (8.73) the substitution

$$V_i(t) = e_i + W_i \mu_i(t) \tag{8.110}$$

must be made.

Note here that due to the ergodicity requirements for the fast process, the blocks of A (i.e. the matrices $A_i, i = 0, \ldots, m$) have a semi-simple null structure; zero eigenvalues appearing only due to the linear dependence of the rows of A^T. But since the transformation from V_i to μ_i removes this linear dependence, the transformed matrix

$$SA^T W \tag{8.111}$$

is of full rank, and therefore, invertible.

Having thus described the fast process, $\mu(t)$, in such a way that its elements are linearly independent, we now turn to a description of the invariant measure in terms of this process. Using equation (8.94), the requirements of the last section that $\overline{V}(u) \in \ker(A^T)$ and that $T\overline{V}(u) = I$ are equivalent to

$$(A^T)(e + W\overline{\mu}(u)) = 0 \tag{8.112}$$

But due to the linear dependence of the rows of A^T (i.e. $TA^T = 0$) we can eliminate the equation associated with the last row of each block of A_i^T. That is, equation (8.112) is equivalent to

$$(SA^T)(e + W\overline{\mu}(u)) = 0 \tag{8.113}$$

Solving for $\overline{\mu}(u)$,

$$\overline{\mu}(u) = -(SA^T(u)W)^{-1} SA^T(u)e \tag{8.114}$$

which is also the condition one formally obtains in taking the limit as $\epsilon \to 0$ in equation (8.106).

By substitution of $\overline{\mu}(u)$ into equation (8.99) we have

$$\overline{V}(u) = [I - W(SA^T(u)W)^{-1} SA^T(u)]e \tag{8.115}$$

which is the expression for the invariant measure over blocks of the A matrix as a function of the control u.

In summary, then, the above state space decomposition gives the following equivalent statement of the singular perturbation problem

$$\min E\{\int_0^T C^T[e + W\mu(t)]\eta(t)dt\} \tag{8.116}$$

subject to

$$d\eta(t) = TB^T[e + W\mu(t)]\eta(t)dt + \frac{1}{\sqrt{N_0}}TH[e + W\mu(t)]\eta(t)dy(t) \tag{8.117}$$

$$\begin{aligned} d\mu(t) &= \{\frac{1}{\epsilon}[SA^T(u)][e + W\mu(t)] + SM(t)\}dt + \\ &\quad \frac{1}{\sqrt{N_0}}\{S - \mu(t)T\}H[e + W\mu(t)]dy(t) \end{aligned} \tag{8.118}$$

where $M(t)$ is defined by equations (8.71) through (8.73) with

$$V_i(t) = e_i + W_i\mu_i(t) \tag{8.119}$$

In the following sections a derivation of the expansion of the optimal cost will be provided, followed by a convergence proof.

8.3 Hamilton-Jacobi-Bellman Equation

The Hamilton-Jacobi-Bellman equation for the problem of minimizing the integral cost

$$\min E\{\int_0^T C^T[e + W\mu(t)]\eta(t)dt\} = \min_u J(u) \tag{8.120}$$

subject to the dynamics of equations (8.117) and (8.118) is given by

$$\begin{aligned} \frac{\partial\phi^\epsilon}{\partial t} + \inf_{u \in U_{ad}} &\{\frac{\partial\phi^\epsilon}{\partial\eta} \cdot TB^T[e + W\mu]\eta + \\ &\sum_{i=0}^m \frac{\partial\phi^\epsilon}{\partial\mu_i} \cdot \{\frac{S_iA_i^T(u)}{\epsilon}[e_i + W_i\mu_i] + S_iM_i\} + \\ &\frac{1}{2}\sum_{i,j}(\Sigma\Sigma^T)_{ij}\frac{\partial^2\phi^\epsilon}{\partial\nu_i\partial\nu_j} + C^T[e + W\mu]\eta\} = 0 \end{aligned} \tag{8.121}$$

$$\phi^{\epsilon}(\eta, \mu, T) = 0 \tag{8.122}$$

where

$$\nu = \begin{pmatrix} \eta \\ \mu_0 \\ \vdots \\ \mu_m \end{pmatrix} \tag{8.123}$$

and

$$\Sigma = \frac{1}{\sqrt{N_0}} \begin{pmatrix} TH[e + W\mu]\eta \\ (S_0 - \mu_0 T_0)H_0[e_0 + W_0\mu_0] \\ \vdots \\ (S_m - \mu_m T_m)H_m[e_m + W_m\mu_m] \end{pmatrix} \tag{8.124}$$

8.4 The limit problem

The limit problem is obtained by (formally) taking the limit as $\epsilon \to 0$ in equations (8.116) through (8.118). This operation results in the problem of minimizing

$$\min E\{\int_0^T C^T \overline{V}(u)\eta(t)dt\} = \min_u J(u) \tag{8.125}$$

subject to the dynamics

$$d\eta(t) = TB^T\overline{V}(u)\eta(t)dt + \frac{1}{\sqrt{N_0}} TH\overline{V}(u)\eta(t)dy(t) \tag{8.126}$$

where $\overline{V}(u)$ is given by

$$\overline{V}(u) = [I - W(SA^T(u)W)^{-1}SA^T(u)]e \tag{8.127}$$

Or equivalently by

$$A^T(u)\overline{V}(u) = 0 \tag{8.128}$$

and

$$T\overline{V}(u) = I \tag{8.129}$$

The Hamilton-Jacobi-Bellman Equation for the problem of equations (8.125) through (8.127) is

$$\frac{\partial \phi_0}{\partial t} = \inf_{u \in U_{ad}} \{ -\frac{\partial \phi_0}{\partial \eta} \cdot [TB^T\overline{V}(u)\eta] - C^T\overline{V}(u)\eta + -\frac{1}{2N_0}\sum_{i,j}(TH\overline{V}(u)\eta\eta^T\overline{V}^T(u)HT^T)_{ij}\frac{\partial^2 \phi_0}{\partial \eta_i \partial \eta_j}\} \quad (8.130)$$

$$\phi_0(\eta, T) = 0 \quad (8.131)$$

The control achieving the minimization on the right hand side will be denoted by $\overline{u}(\eta, t)$.

8.5 Method of Solution

The determination of ϵ-optimal controls for the singular perturbations problem will be accomplished by establishing an expansion in ϵ for the solution ϕ^ϵ of the Hamilton Jacobi Bellman Equation. In this section we provide formal arguments (which will be made more precise in future sections) that outline the method of solution.

The Bellman equation (8.121) can be written in the following form

$$\frac{\partial \phi^\epsilon}{\partial t} + \inf_{u \in U_{ad}} \{C^T[e + W\mu]\eta + \mathcal{L}\phi^\epsilon\} = 0 \quad (8.132)$$

where $\mathcal{L}$ is the differential generator for the dynamics of equations (8.117) and (8.118) and is given by

$$\mathcal{L}\psi = \frac{\partial \psi}{\partial \eta} \cdot TB^T[e + W\mu]\eta + \sum_{i=0}^{m} \frac{\partial \psi}{\partial \mu_i} \cdot \{\frac{S_iA_i^T(u)}{\epsilon}[e_i + W_i\mu_i] + S_iM_i\} + \frac{1}{2}\sum_{i,j}(\Sigma\Sigma^T)_{ij}\frac{\partial^2 \psi}{\partial \nu_i \nu_j} \quad (8.133)$$

We seek an expansion of equation (8.132) in which

$$\phi^\epsilon(\eta, \mu, t) = \phi_0(\eta, t) + \epsilon\phi_1(\eta, \mu, t) + \dots \quad (8.134)$$

By substitution of equation (8.134) into equation (8.132) and taking the limit as $\epsilon \to 0$, we retain those terms that are independent of ϵ.

$$\frac{\partial \phi_0}{\partial t} + \inf_{u \in U_{ad}} \{C^T[e + W\mu]\eta + \mathcal{L}\phi_0 + \lim_{\epsilon \to 0} \epsilon \mathcal{L}\phi_1\} = 0 \tag{8.135}$$

where, since ϕ_0 is a function of η and t only,

$$\mathcal{L}\phi_0 = \frac{\partial \phi_0}{\partial \eta} \cdot TB^T[e + W\mu]\eta + \frac{1}{2}\sum_{i,j}(TH[e + W\mu]\eta\eta^T[e + W\mu]^T HT^T)_{ij} \frac{\partial^2 \phi_0}{\partial \eta_i \partial \eta_j} \tag{8.136}$$

and, from equation (8.133),

$$\lim_{\epsilon \to 0} \epsilon \mathcal{L}\phi_1 = \sum_{i=0}^{m} \frac{\partial \phi_1}{\partial \mu_i} \cdot [S_i A_i^T(u)][e_i + W_i \mu_i] \tag{8.137}$$

The important observation to make regarding equation (8.135) is that, using equation (8.137), it can be interpreted as the Bellman equation that corresponds to the deterministic optimal control problem of minimizing the cost

$$\min_{u} \int_0^\infty \frac{\partial \phi_0}{\partial t} + C^T[e + W\mu(\tau)]\eta + \mathcal{L}\phi_0 d\tau \tag{8.138}$$

subject to the dynamics

$$\frac{d\mu}{d\tau} = [SA^T(u(\tau))][e + W\mu(\tau)] \tag{8.139}$$

where η and t enter as parameters. The dynamics are deterministic since, in the limit as $\epsilon \to 0$, the dynamics of equation (8.118) are deterministic. Similarly, since we are optimizing with respect to the "fast variables," in the limit as $\epsilon \to 0$, $\frac{t}{\epsilon} \to \infty$; and we have an infinite time horizon problem. The resulting control will be of the form

$$u = u(\eta, t; \mu) \tag{8.140}$$

In order to evaluate equation (8.138) we must determine $\frac{\partial \phi_0}{\partial t}$, $\frac{\partial \phi_0}{\partial \eta}$, $\frac{\partial^2 \phi_0}{\partial \eta_i \partial \eta_j}$, and $\eta(t)$. One may think that a natural way to proceed is to obtain these functions by averaging equation (8.135) over the fast process and solving the resultant limit problem. That is, by multiplying equation (8.135) by the invariant measure of the fast process and integrating with respect to the fast process. This operation succeeds since the influence of the fast process in the slow time scale is only felt in an average sense.

We will denote the invariant measure of the fast process (μ) given the slow process (η) when the control v is applied by $m^v(\eta;\mu)$. The limit problem is given by

$$\frac{\partial \phi_0}{\partial t} + \inf_{u} \{ \int_{\mathcal{R}^{m+1}} [C^T(e + W\mu(t))\eta(t) + \mathcal{L}\phi_0] m^u(\eta;\mu) d\mu \} = 0 \qquad (8.141)$$

Note that the operator

$$\{ \int_{\mathcal{R}^{m+1}} (\cdot) m^u(\eta;\mu) d\mu \} \qquad (8.142)$$

annihilated the term $\lim_{\epsilon \to 0} \epsilon \mathcal{L}\phi_1$. In the case of a fast process that is stochastic in the limit as $\epsilon \to 0$ this annihilation follows readily after integrating by parts and applying the Fokker-Planck equation (c.f. [2] or [31]). On the other hand, for the system of equations (8.117) and (8.118), when one applies the control u, after long periods of time (in the fast time scale) the state $\mu(t)$ evolves towards $\overline{\mu}(u(t))$. One can think of this limit value for μ as a point at which all probability mass is located in a degenerate stochastic formulation. For the system of equations (8.117) and (8.118), then

$$m^u(\eta;\mu) = \delta(\mu(u) - \overline{\mu}(u)) \qquad (8.143)$$

where δ is the Dirac delta function and $\overline{\mu}(u)$ is determined by the long term behavior of μ. That is, formally taking the limit as $\epsilon \to 0$ in equation (8.118), $\overline{\mu}(u)$ is such that

$$e + W\overline{\mu}(u) \in \ker(A^T(u)) \qquad (8.144)$$

(see, for example, equation (8.114)).

Applying equations (8.143) and (8.144) to equation (8.137) we have

$$\int_{\mathcal{R}^{m+1}} [\lim_{\epsilon \to 0} \epsilon \mathcal{L}\phi_1] \delta(\mu(u) - \overline{\mu}(u)) d\mu = 0 \tag{8.145}$$

That is, the operator described by expression (8.142) is a projection onto the null space of the ergodic kernels (i.e. the null space of $\lim_{\epsilon \to 0} \epsilon \mathcal{L}(\cdot)$). It is precisely this operator that often plays an analogous role in the problem of the optimal control of **fully observed** Markov Chains. In [13] Delebecque and Quadrat consider such a problem in which the infinitesimal generator for the chain is given by

$$\frac{A(u)}{\epsilon} + B \tag{8.146}$$

The operator they use is $\overline{V}T$, the eigenprojection of the zero eigenvalue of the A matrix (where, of course, $\overline{V}$ is the invariant measure for blocks of the A matrix).

Combining equations (8.143) and (8.141), we have the limit problem for the dynamics of equations (8.117) and (8.118) given by

$$\frac{\partial \phi_0}{\partial t} + \inf_{u} \{C^T[e + W\overline{\mu}(u(\eta, t))]\eta(t) + \mathcal{L}\phi_0|_{\mu = \overline{\mu}(u)}\} = 0 \tag{8.147}$$

Equation (8.147) is the Bellman equation for the stochastic optimal control problem of minimizing the (averaged) cost

$$E\{\int_0^T C^T[e + W\overline{\mu}(u(\eta, t))]\eta(t)dt\} \tag{8.148}$$

subject to the (averaged) dynamics

$$\begin{aligned} d\eta(t) &= TB^T[e + W\overline{\mu}(u(\eta, t))]\eta(t)dt + \\ &\quad \frac{1}{\sqrt{N_0}} TH[e + W\overline{\mu}(u(\eta, t))]\eta(t)dy(t) \end{aligned} \tag{8.149}$$

A significant aspect of equation (8.149) is that the slow control, u, is a function of η and t only. This functional relationship is due to the deterministic nature of the fast variables over long time horizons (in the fast time scale).

In contrast, when in the limit as $\epsilon \to 0$, the fast variable (for us, fast variables refer to the Zakai equation) is stochastic, the slow control is a function of the fast variables, the slow variables, and time. This result is suggested by equation (8.141), in which, for a stochastically driven fast variable, the invariant measure converges to a non-impulsive distribution and we are left with the control as a function of all variables.

We note here that the control that minimizes the term in braces in equation (8.135) will also achieve the minimum in equation (8.147). Indeed, suppose $\hat{u}$ is the minimizing control in equation (8.135). Then after integration we obtain

$$\frac{\partial \phi_0}{\partial t} + C^T[e + W\overline{\mu}(\hat{u})]\eta(t) + \mathcal{L}\phi_0|_{\mu=\overline{\mu}(\hat{u})} = 0 \qquad (8.150)$$

Deterministic problems (and problems such as we have here in which the fast dynamics are, in the limit, deterministic) have ϵ-optimal controls that can be represented as the sum of two terms. The first term is the solution to the limit problem (i.e. the slow control described already). The second term is a corrector which is parameterized by the slow variable and time and is a function of the difference between the instantaneous value of the fast variable and its limiting value:

$$u_c(s) = \overline{u}(\eta(s), s) + \overline{\xi}(\eta(s), s; \mu(s) - \underline{\mu}(\eta, s)) \qquad (8.151)$$

This control is called the composite control [9]. In future sections we will demonstrate the ϵ-optimality of this control. We wish here only to show how this form follows from the preceding discussion.

If we rewrite equation (8.137) as

$$\lim_{\epsilon \to 0} \epsilon\mathcal{L}\phi_1 = \sum_{i=0}^{m}[\overline{\lambda}_i^T(\eta, t) + \frac{\partial \chi}{\partial \mu_i}(\eta, t; \mu - \overline{\mu})] \cdot [S_i A_i^T(u)][e_i + W_i\mu_i] \qquad (8.152)$$

Then, as is described in some detail in the section in which the asymptotic analysis is done, χ becomes the value function for an optimal control problem which is very similar to that of equations (8.138) and (8.139); except for the fact that the cost for this new problem includes the term

$$\int_0^\infty \sum_{i=0}^{m} \overline{\lambda}_i^T(\eta,t) S_i A_i^T(u)[e_i + W_i\mu_i] d\tau \tag{8.153}$$

The function $\overline{\lambda}(\eta,t)$ will be chosen to be $\frac{\partial\phi_1}{\partial\mu}(\eta,\mu,t)$ evaluated at $\mu = \overline{\mu}(\overline{u}(\eta,t))$ and $u = \overline{u}(\eta,t)$. And, of course, we have

$$\frac{\partial\phi_1}{\partial\mu}(\eta,\mu,t) = \overline{\lambda}(\eta,t) + \frac{\partial\chi}{\partial\mu}(\eta,t;\mu-\overline{\mu}) \tag{8.154}$$

8.6 Asymptotic analysis

We seek a solution of the form

$$\phi^\epsilon(\eta,\mu,t) = \phi_0(\eta,t) + \epsilon\phi_1(\eta,\mu,t) + \dots \tag{8.155}$$

Formally substituting equation (8.155) into the Hamilton-Jacobi-Bellman equation (8.121) and collecting terms that are independent of ϵ, we obtain

$$\begin{aligned}\frac{\partial\phi_0}{\partial t} + \inf_{u \in U_{ad}} \{ & \frac{\partial\phi_0}{\partial\eta} \cdot TB^T[e + W\mu]\eta + \\ & \sum_{i=0}^{m} \frac{\partial\phi_1}{\partial\mu_i} \cdot [S_i A_i^T(u)][e_i + W_i\mu_i] + \\ \frac{1}{2}\sum_{i,j}(TH(e+W\mu)\eta\eta^T(e+W\mu)^T HT^T)_{ij} & \frac{\partial^2\phi_0}{\partial\eta_i\partial\eta_j} + \\ & C^T[e + W\mu]\eta\} = 0\end{aligned} \tag{8.156}$$

It will now be shown that equation (8.156) can be resolved, where (η,t) enter as parameters. Let

$$\underline{\mu}(\eta,t) \triangleq \overline{\mu}(\overline{u}(\eta,t)) \tag{8.157}$$

Then introduce the functions

$$\begin{aligned}& F(\eta,t;\tilde{\mu},\xi) = \\ & \quad \frac{\partial\phi_0}{\partial t} + \frac{\partial\phi_0}{\partial\eta} \cdot TB^T\{e + W[\underline{\mu}(\eta,t) + \tilde{\mu}]\}\eta +\end{aligned}$$

$$+\frac{1}{2}\sum_{i,j}(TH[e+(W)(\underline{\mu}(\eta,t)+\tilde{\mu})]\eta$$

$$\cdot\eta^T[e+(W)(\underline{\mu}(\eta,t)+\tilde{\mu})]^T HT^T)_{ij}\frac{\partial^2\phi_0}{\partial\eta_i\partial\eta_j}+$$

$$+\sum_{i=0}^{m}\overline{\lambda}_i^T(\eta,t)[S_iA_i^T(\overline{u}(\eta,t)+\xi)][e_i+(W)(\underline{\mu}_i(\eta,t)+\tilde{\mu}_i)]$$

$$+C^T\{e+W[\underline{\mu}(\eta,t)+\tilde{\mu}]\}\eta \tag{8.158}$$

$$G(\eta,t;\tilde{\mu},\xi)=$$

$$[SA^T(\overline{u}(\eta,t)+\xi)][e+(W)(\underline{\mu}(\eta,t)+\tilde{\mu})] \tag{8.159}$$

where $\tilde{\mu}$ is defined by the matrix of k dimensional column vectors $\tilde{\mu}_i, i = 0,\ldots,m$ as follows

$$\tilde{\mu}=\begin{pmatrix}\tilde{\mu}_0 & & 0\\ & \ddots & \\ 0 & & \tilde{\mu}_m\end{pmatrix} \tag{8.160}$$

and $\overline{\lambda}(\eta,t)$ is defined by

$$\begin{aligned}\overline{\lambda}(\eta,t) \;&=\; -[W^TA(\overline{u}(\eta,t))S^T]^{-1}NW^T\{BT^T\frac{\partial\phi_0}{\partial\eta}(\eta,t)+C+\\ &\quad HT^T\{\frac{\partial}{\partial\eta}[\frac{\partial}{\partial\eta}\phi_0(\eta,t)]^T\}TH[e+W\overline{\mu}(\overline{u}(\eta,t))]\eta\}\end{aligned} \tag{8.161}$$

where

$$N\triangleq\begin{pmatrix}\eta_0 I & & 0\\ & \ddots & \\ 0 & & \eta_m I\end{pmatrix} \tag{8.162}$$

is an $(m+1)k\times(m+1)k$ matrix.

We will now formally show that this definition of $\overline{\lambda}(\eta,t)$ results in $\overline{\lambda}(\eta,t)$ being $\frac{\partial\phi_1}{\partial\mu}$ evaluated at $\mu=\overline{\mu}(\overline{u}(\eta,t))$ and $u=\overline{u}(\eta,t)$.

First we define the following vectors

$$\rho=\begin{pmatrix}\hat{\rho}\\ \rho_0\\ \vdots\\ \rho_m\end{pmatrix} \tag{8.163}$$

where $\hat{\rho}$ is a vector of length $m+1$, and, for $i = 0, \ldots, m$, ρ_i is a vector of length k.

$$r = \begin{pmatrix} \hat{r} \\ r_0 \\ \vdots \\ r_m \end{pmatrix} \tag{8.164}$$

where $\hat{r}$ is a vector of length $m+1$, and, for $i = 0, \ldots, m$, r is a vector of length k.

Applying the Stochastic Maximum Principle (Theorem 3, equation (5.35)) to the problem of equations (8.116) through (8.118) we have

$$\begin{aligned} -d\rho_i &= \{\frac{\partial}{\partial \mu_i}\hat{\rho}^T TB^T(e + W\mu)\eta + \\ &\quad \frac{\partial}{\partial \mu_i}\sum_{j=0}^{m}\rho_j^T[\frac{1}{\epsilon}(S_j A_j^T)(e_j + W_j\mu_j) + S_j M_j] \\ &\quad \frac{\partial}{\partial \mu_i}C^T(e + W\mu)\eta + \frac{1}{\sqrt{N_0}}\frac{\partial}{\partial \mu_i}\hat{r}^T TH(e + W\mu)\eta + \\ &\quad \frac{1}{\sqrt{N_0}}\frac{\partial}{\partial \mu_i}\sum_{j=0}^{m} r_j^T[(S_j - \mu_j T_j)H_j(e_j + W_j\mu_j)]\}dt - r_i^T dy \end{aligned} \tag{8.165}$$

But from equation (5.71)

$$r_i = \{\frac{\partial}{\partial \mu_i}[\frac{\partial}{\partial \nu}\phi^\epsilon]^T\}\Gamma \tag{8.166}$$

and

$$\hat{r} = \{\frac{\partial}{\partial \eta}[\frac{\partial}{\partial \nu}\phi^\epsilon]^T\}\Gamma \tag{8.167}$$

where

$$\Gamma = \begin{pmatrix} TH(e + W\mu)\eta \\ (S_0 - \mu_0^T)H_0(e_0 + W_0\mu_0) \\ \vdots \\ (S_m - \mu_m^T)H_m(e_m + W_m\mu_m) \end{pmatrix} \tag{8.168}$$

is a vector of length $m(k+1)$,

$$\nu = \begin{pmatrix} \eta \\ \mu_0 \\ \vdots \\ \mu_m \end{pmatrix} \tag{8.169}$$

is a column vector of length $(m+1)(k+1)$, and

$$\frac{\partial}{\partial \mu_i}[\frac{\partial}{\partial \nu}\phi^\epsilon]^T \tag{8.170}$$

is a matrix of dimension $k \times (m+1)(k+1)$.

Now let

$$\phi^\epsilon(\eta, \mu, t) = \phi_0(\eta, t) + \epsilon\phi_1(\eta, t) + \cdots \tag{8.171}$$

Substitution of equation (8.171) into equations (8.166) and (8.167) yield ϵ-expansions for r_i and $\hat{r}_i$. Let

$$r = R + \epsilon\Theta + \cdots \tag{8.172}$$

where

$$R = \begin{pmatrix} \hat{R} \\ R_0 \\ \vdots \\ R_m \end{pmatrix} \tag{8.173}$$

and

$$\Theta = \begin{pmatrix} \hat{\Theta} \\ \Theta_0 \\ \vdots \\ \Theta_m \end{pmatrix} \tag{8.174}$$

Since ϕ_0 is a function of η and t only, by substitution of equation (8.171) into equations (8.166) and (8.167) we have, for $i = 0, \ldots, m$,

$$R_i = 0 \tag{8.175}$$

and

$$\hat{R} = \{\frac{\partial}{\partial \eta}[\frac{\partial}{\partial \eta}\phi_0(\eta, t)]^T\}TH(e + W\mu)\eta \tag{8.176}$$

We also know from equation (5.66) that

$$\rho = \frac{\partial \phi^\epsilon}{\partial \nu} \tag{8.177}$$

So, by substitution of equation (8.171) into equation (8.177), for $i = 1, \ldots, m-1$,

$$\hat{\rho} = \frac{\partial \phi_0}{\partial \eta} + \epsilon \frac{\partial \phi_1}{\partial \eta} + \cdots \tag{8.178}$$

$$\rho_0 = \epsilon \frac{\partial \phi_1}{\partial \mu_0} + \cdots \tag{8.179}$$

$$\vdots$$

$$\rho_i = \epsilon \frac{\partial \phi_1}{\partial \mu_i} + \cdots \tag{8.180}$$

$$\vdots$$

$$\rho_m = \epsilon \frac{\partial \phi_1}{\partial \mu_m} + \cdots \tag{8.181}$$

We will denote by λ the column vector of length $(m+1)k$ given by

$$\lambda = \frac{\partial \phi_1}{\partial \mu} \tag{8.182}$$

which we will partition into $m+1$ subvectors, each of length k as

$$\lambda = \begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_m \end{pmatrix} \tag{8.183}$$

We will also define the column vector p (of length $m+1$) by

$$p = \frac{\partial \phi_0}{\partial \eta} \tag{8.184}$$

Substituting equation (8.171) into equation (8.165), collecting terms independent of ϵ, and using equation (8.175) and equations (8.178) through (8.181), we have

$$0 = \frac{\partial}{\partial \mu_i} p^T T B^T (e + W\mu)\eta + \frac{\partial}{\partial \mu_i} \sum_{i=0}^{m} \lambda_i^T (S_i A_i^T)(e_i + W_i \mu_i) + \frac{\partial}{\partial \mu_i} C^T (e + W\mu)\eta + \frac{1}{\sqrt{N_0}} \frac{\partial}{\partial \mu_i} \hat{R}^T T H (e + W\mu)\eta \quad (8.185)$$

Performing the indicated partial differentiations,

$$0 = NW^T BT^T p + W^T A(u) S^T \lambda + NW^T C + \frac{1}{\sqrt{N_0}} NW^T HT^T \hat{R} \quad (8.186)$$

Due to the invertibility of

$$SA^T(u)W \quad (8.187)$$

for all u, equation (8.186) has the unique solution

$$\lambda = -(W^T A(u) S^T)^{-1} NW^T [BT^T p + C + HT^T \hat{R}] \quad (8.188)$$

And so, using equations (8.176) and (8.184),

$$\overline{\lambda}(\eta, t) = -[W^T A(\overline{u}(\eta, t)) S^T]^{-1} NW^T \{BT^T \frac{\partial \phi_0}{\partial \eta}(\eta, t) + C + HT^T \{\frac{\partial}{\partial \eta}[\frac{\partial}{\partial \eta}\phi_0(\eta, t)]^T\} TH[e + W\overline{\mu}(\overline{u}(\eta, t))]\eta\} \quad (8.189)$$

which is the expression for $\overline{\lambda}(\eta, t)$ given in equation (8.161).

Returning now to the functions F and G, we will partition $G(\eta, t; \tilde{\mu}, \xi)$ into column vectors of dimension k as

$$\begin{pmatrix} G_0 & & 0 \\ & \ddots & \\ 0 & & G_m \end{pmatrix} \quad (8.190)$$

where $G_i, i = 0, \ldots, m$ is given by

$$G_i(\eta, t; \tilde{\mu}, \xi) = [S_i A_i^T (\overline{u}(\eta, t) + \xi)][e_i + (W_i)(\underline{\mu}_i(\eta, t) + \tilde{\mu}_i)] \quad (8.191)$$

And similarly we will partition $\overline{\lambda}(\eta, t)$ as follows

$$\overline{\lambda}(\eta, t) = \begin{pmatrix} \overline{\lambda}_0(\eta, t) \\ \vdots \\ \overline{\lambda}_m(\eta, t) \end{pmatrix} \quad (8.192)$$

where $\overline{\lambda}_i$ are column vectors of dimension k.

Consider now the equation, which determines χ:

$$\inf_{\xi} \{F(\eta, t; \tilde{\mu}, \xi) + \sum_{i=0}^{m} D_{\tilde{\mu}_i}\chi \cdot G_i(\eta, t; \tilde{\mu}, \xi)\} = 0 \tag{8.193}$$

Equation (8.193) corresponds to the optimal control problem of minimizing

$$J(U) = \int_0^\infty F(\eta, t; \tilde{\mu}(\tau), U(\tau))d\tau \tag{8.194}$$

subject to the dynamics (for $i = 0, \ldots, m$),

$$\frac{d\tilde{\mu}_i(\tau)}{d\tau} = G_i(\eta, t; \tilde{\mu}(\tau), U(\tau)) \tag{8.195}$$

$$\tilde{\mu}(0) = \tilde{\mu} \tag{8.196}$$

where the optimal value function is

$$\chi(\eta, t; \tilde{\mu}) = \inf_{U} J(U) \tag{8.197}$$

It is easily seen from the form of F and G in equations (8.159) and (8.158) that this optimal control problem has a solution. Thus χ exists.

The control that minimizes the value function will be denoted $\overline{\xi}$. It will depend on $\tilde{\mu}$ (and on τ through $\tilde{\mu}$) and parametrically on η and t (i.e. $\overline{\xi}(\eta(t), t; \tilde{\mu}(\tau))$).

Having thus defined χ, we may now construct a solution to equation (8.156). Let

$$\phi_1(\eta, \mu, t) = \sum_{i=0}^{m} \overline{\lambda}_i^T(\eta, t)(\mu_i - \underline{\mu}_i(\eta, t)) + \chi(\eta, t; \mu - \underline{\mu}(\eta, t)) \tag{8.198}$$

The verification of ϕ_1 as a solution to equation (8.156) is straightforward. By direct substitution of equation (8.198) into equation (8.156) we have

$$\begin{aligned} \frac{\partial \phi_0}{\partial t} + \inf_{u \in U_{ad}} \{ & \frac{\partial \phi_0}{\partial \eta} \cdot TB^T[e + W\mu]\eta + \\ & \sum_{i=0}^{m} [\overline{\lambda}_i^T(\eta, t) + \frac{\partial \chi}{\partial \mu_i}(\eta, t; \mu - \underline{\mu}(\eta, t))] \cdot S_i A_i^T(u)[e_i + W_i \mu_i] + \\ & \frac{1}{2} \sum_{i,j} (TH(e + W\mu)\eta\eta^T(e + W\mu)^T HT^T)_{ij} \frac{\partial^2 \phi_0}{\partial \eta_i \partial \eta_j} + \\ & C^T[e + W\mu]\eta \} = 0 \end{aligned} \tag{8.199}$$

Now let

$$\tilde{\mu} = \mu - \underline{\mu}(\eta, t) \tag{8.200}$$

$$\xi = u - \overline{u}(\eta, t) \tag{8.201}$$

Then by substitution of equations (8.200) and (8.201) into equation (8.199) and applying the definitions of F and G, we obtain

$$\inf_{\xi} \; \{F(\eta, t; \tilde{\mu}, \xi) + \sum_{i=0}^{m} D_{\tilde{\mu}_i}\chi \cdot G_i(\eta, t; \tilde{\mu}, \xi)\} = 0 \tag{8.202}$$

Thus it has been formally shown that the optimal control problems given by equations (8.125) through (8.129) and equations (8.194) through (8.196) provide the first and second order terms in the expansion of the optimal cost for the original singular perturbation problem, with the help of equations (8.157), (8.161), and (8.198).

8.7 The composite control

The composite feedback control is defined by

$$u_c(\eta(t), \mu(t), t) = \overline{u}(\eta(t), t) + \overline{\xi}(\eta(t), t; \mu(t) - \underline{\mu}(\eta(t), t)) \tag{8.203}$$

The form of this equation suggests the name composite feedback. The feedback is the sum of two terms. The first term corresponds to the optimal feedback for the limit system. This term resembles the result one obtains from deterministic systems in that it is a function of only the slow variable and time. This result is in contrast to the stochastic systems Bensoussan considers in [2], in which the solution to the limit problem is a function of both the fast and slow variables.

The second term is a corrector term that is a function of the difference between the actual value of the fast variables, μ, and the limiting value of these variables, $\underline{\mu}$, under the optimal slow control. Chow and Kokotovic demonstrate that such a control arises in the situation in which the fast system is linear. Bensoussan considers a deterministic problem with nonlinear dynamics in which he shows that in spite of the nonlinear dynamics, the composite control is a near optimal solution to the problem. [2]. Here, although the dynamics are stochastic, in the fast time scale the problem is essentially deterministic. As a result, it is quite similar to the problem that Bensoussan considers; and the composite control is again near optimal.

We consider the evolution of the system driven by the composite control. The dynamics are given by

$$\begin{aligned} d\eta^c(s) &= TB^T[e + W\mu^c(s)]\eta^c(s)ds + \\ &\quad \frac{1}{\sqrt{N_0}} TH[e + W\mu^c(s)]\eta^c(s)dy(s) \quad (8.204) \\ d\mu_i^c(s) &= \{\frac{1}{\epsilon}[S_iA_i^T(u_c)][e_i + W_i\mu_i^c(s)] + S_iM_i^c(s)\}ds \\ &\quad + \frac{1}{\sqrt{N_0}}[S_i - \mu_i^c(s)T_i]H_i[e_i + W_i\mu_i^c(s)]dy(s) \quad (8.205) \\ \eta^c(t) &= \eta \quad (8.206) \\ \mu_i^c(t) &= \mu_i \quad (8.207) \end{aligned}$$

where

$$u_c(s) = \overline{u}(\eta^c(s), s) + \overline{\xi}(\eta^c(s), s; \mu^c(s) - \overline{\mu}(\eta^c(s), s)) \quad (8.208)$$

8.8 Proof of convergence

Theorem 10 *If*

$$\frac{\partial^2\phi_1}{\partial\mu\partial\mu}, \frac{\partial^2\phi_1}{\partial\mu\partial\eta}, \frac{\partial^2\phi_1}{\partial\eta\partial\eta}, \frac{\partial\phi_1}{\partial\eta}, \frac{\partial\phi_1}{\partial\mu}, \frac{\partial\phi_1}{\partial s},$$
$$E\{TB^T[e + W\mu^c]\eta^c\}, E\{SM^c\}, E\{\Sigma_c\Sigma_c^T\}, \text{ and } E\{\phi_1(\eta^c(T), \mu^c(T), T)\}$$

are bounded by a constant $C(\eta, \mu)$ (8.209)

Then the following estimate is valid

$$|\phi^\epsilon(\eta, \mu, t) - \phi_0(\eta, t) - \epsilon\phi_1(\eta, \mu, t)| \leq \epsilon C(\eta, \mu, t) \quad (8.210)$$

Proof: By Ito's lemma,

$$\begin{aligned}
&\frac{d}{ds}[\phi_0(\eta^c,s)+\epsilon\phi_1(\eta^c,\mu^c,s)] = \\
&\quad \frac{\partial}{\partial s}[\phi_0(\eta^c,s)+\epsilon\phi_1(\eta^c,\mu^c,s)] + \\
&\quad \frac{\partial}{\partial \eta^c}[\phi_0(\eta^c,s)+\epsilon\phi_1(\eta^c,\mu^c,s)]\cdot TB^T[e+W\mu^c]\eta^c + \\
&\quad \frac{1}{\epsilon}\sum_{i=0}^{m}\{\frac{\partial}{\partial \mu_i^c}[\phi_0(\eta^c,s)+\epsilon\phi_1(\eta^c,\mu^c,s)]\}\cdot\{[S_iA_i^T(u_c)][e_i+W_i\mu_i^c]+\epsilon S_iM_i^c\} + \\
&\quad \frac{1}{2}\sum_{i,j}(\Sigma_c\Sigma_c^T)_{ij}\frac{\partial^2}{\partial\nu_i^c\partial\nu_j^c}[\phi_0(\eta^c,s)+\epsilon\phi_1(\eta^c,\mu^c,s)]
\end{aligned} \tag{8.211}$$

where ν^c is defined by equation (8.123), evaluated along the trajectory that results from the application of the composite control

$$\nu^c = \begin{pmatrix} \eta^c \\ \mu_0^c \\ \vdots \\ \mu_m^c \end{pmatrix} \tag{8.212}$$

Σ is defined by equation (8.124) also evaluated along the trajectory that results from an application of the composite control

$$\Sigma_c = \frac{1}{\sqrt{N_0}}\begin{pmatrix} TH[e+W\mu^c]\eta^c \\ (S_0-\mu_0^cT_0)H_0[e_0+W_0\mu_0^c] \\ \vdots \\ (S_m-\mu_m^cT_m)H_m[e_m+W_m\mu_m^c] \end{pmatrix} \tag{8.213}$$

and composite control, u_c is defined by equation (8.208). Applying these definitions, observing that ϕ_0 is a function of η^c and s only, and regrouping terms, equation (8.211) may be rewritten as

$$\begin{aligned}
&\frac{d}{ds}[\phi_0(\eta^c,s)+\epsilon\phi_1(\eta^c,\mu^c,s)] = \\
&\quad \frac{\partial}{\partial s}\phi_0(\eta^c,s)+[\frac{\partial}{\partial\eta^c}\phi_0(\eta^c,s)]\cdot TB^T[e+W\mu^c]\eta^c + \\
&\quad \sum_{i=0}^{m}[\frac{\partial}{\partial\mu_i^c}\phi_1(\eta^c,\mu^c,s)]\cdot[S_iA_i^T(u_c)][e_i+W_i\mu_i^c] +
\end{aligned}$$

$$\frac{1}{2}\sum_{i,j}(TH[e+W\mu^c]\eta^c(\eta^c)^T[e+W\mu^c]^T HT^T)_{ij}\frac{\partial^2}{\partial\eta_i^c\partial\eta_j^c}\phi_0(\eta^c,s)+$$
$$\epsilon\{\Psi(\eta^c,\mu^c,s)\} \tag{8.214}$$

where Ψ is defined by

$$\begin{aligned}\Psi(\eta^c,\mu^c,s) \quad = \quad & \frac{\partial\phi_1}{\partial s}(\eta^c,\mu^c,s)+[\frac{\partial}{\partial\eta^c}\phi_1(\eta^c,\mu^c,s)]\cdot TB^T[e+W\mu^c]\eta^c+ \\ & \sum_{i=0}^{m}[\frac{\partial}{\partial\mu_i^c}\phi_1(\eta^c,\mu^c,s)]\cdot S_iM_i^c+ \\ & \sum_{i,j}(\Sigma_c\Sigma_c^T)_{ij}\frac{\partial^2}{\partial\nu_i\partial\nu_j}\phi_1(\eta^c,\mu^c,s)\end{aligned} \tag{8.215}$$

Now let

$$\tilde{\mu}=\mu^c-\underline{\mu}(\eta^c,s) \tag{8.216}$$

Applying this relationship to the definition of ϕ_1 given by equation (8.198).

$$\phi_1(\eta^c,\mu^c,s)=\sum_{i=0}^{m}\overline{\lambda}_i^T(\eta^c,s)(\mu_i^c-\underline{\mu}_i(\eta^c,s))+\chi(\eta^c,s;\mu^c-\underline{\mu}(\eta^c,s)) \tag{8.217}$$

we obtain

$$\frac{\partial}{\partial\mu_i^c}\phi_1(\eta^c,\mu^c,s) \quad = \quad \overline{\lambda}_i^T(\eta^c,s)+\frac{\partial}{\partial\mu_i^c}\chi(\eta^c,s;\mu_c-\underline{\mu}(\eta^c,s)) \tag{8.218}$$
$$= \quad \overline{\lambda}_i^T(\eta^c,s)+\frac{\partial}{\partial\tilde{\mu}_i}\chi(\eta^c,s;\tilde{\mu}) \tag{8.219}$$

By substitution of equation (8.219) into equation (8.214),

$$\begin{aligned}&\frac{d}{ds}[\phi_0(\eta^c,s)+\epsilon\phi_1(\eta^c,\mu^c,s)]= \\ &\quad\frac{\partial}{\partial s}\phi_0(\eta^c,s)+[\frac{\partial}{\partial\eta^c}\phi_0(\eta^c,s)]\cdot TB^T[e+W(\tilde{\mu}+\underline{\mu}(\eta^c,s))]\eta^c+ \\ &\quad\sum_{i=0}^{m}[\overline{\lambda}_i^T(\eta^c,s)+\frac{\partial}{\partial\tilde{\mu}_i}\chi(\eta^c,s;\tilde{\mu})]\cdot \\ &\quad[S_iA_i^T(\overline{u}(\eta^c,s)+\overline{\xi}(\eta^c,s;\tilde{\mu}))][e_i+W_i(\tilde{\mu}_i+\underline{\mu}_i(\eta^c,s))]+ \\ &\quad\frac{1}{2}\sum_{i,j}(TH[e+W(\tilde{\mu}+\underline{\mu}(\eta^c,s))]\eta^c(\eta^c)^T[e+W(\tilde{\mu}+\underline{\mu}(\eta^c,t))]^T H^T T^T)_{ij}\cdot \\ &\quad\frac{\partial^2}{\partial\eta_i^c\partial\eta_j^c}\phi_0(\eta^c,s)+ \\ &\quad\epsilon\{\Psi(\eta^c,\mu^c,s)\}\end{aligned} \tag{8.220}$$

But by the definition of F and G (equations (8.158) and (8.159)) we obtain

$$\frac{d}{ds}[\phi_0(\eta^c, s) + \epsilon\phi_1(\eta^c, \mu^c, s)] = F(\eta^c, s; \tilde{\mu}, \overline{\xi}) + \sum_{i=0}^{m} \frac{\partial\chi}{\partial\tilde{\mu}_i} \cdot G_i(\eta^c, s; \tilde{\mu}, \overline{\xi}) - C^T[e + W(\tilde{\mu} + \underline{\mu}(\eta^c, s))]\eta^c + \epsilon\{\Psi(\eta^c, \mu^c, s)\} \tag{8.221}$$

But since $\overline{\xi}$ is the optimal control for the problem of equations (8.194) through (8.196) (in which η and s are parameters), it satisfies equation (8.193) and therefore

$$\frac{d}{ds}[\phi_0(\eta^c, s) + \epsilon\phi_1(\eta^c, \mu^c, s)] = -C^T[e + W(\tilde{\mu} + \underline{\mu}(\eta^c, s))]\eta^c + \epsilon\{\Psi(\eta^c, \mu^c, s)\} \tag{8.222}$$

Integrating both sides of equation (8.222)

$$\int_t^T \frac{d}{ds}[\phi_0(\eta^c, s) + \epsilon\phi_1(\eta^c, \mu^c, s)]ds = -\int_t^T C^T[e + W(\tilde{\mu} + \underline{\mu}(\eta^c, s))]\eta^c ds + \epsilon\int_t^T \Psi(\eta^c, \mu^c, s)ds \tag{8.223}$$

Performing the indicated integrations and using the equations for the initial conditions on η^c and μ^c (equations (8.206) and (8.207)) we obtain

$$\begin{aligned} \phi_0(\eta, t) + \epsilon\phi_1(\eta, \mu, t) &= J^\epsilon_{\eta,\mu,t}(u_c) \\ &\quad + \epsilon\{E[\phi_1(\eta^c(T), \mu^c(T), T)] \\ &\quad - \int_t^T \Psi(\eta^c, \mu^c, s)ds\} \end{aligned} \tag{8.224}$$

where we have defined the cost

$$J^\epsilon_{\eta,\mu,t}(v(\cdot)) = E\{\int_t^T C^T[e + W\mu^v(s)]\eta^v(s)ds\} \tag{8.225}$$

and $\mu^v(s)$ and $\eta^v(s)$ denote the trajectories that result when the control $v(\cdot)$ is applied to equations (8.77) and (8.79) respectively. That is,

$$\begin{aligned} d\eta^v(s) &= TB^T[e + W\mu^v(s)]\eta^v(s)ds + \\ &\quad \frac{1}{\sqrt{N_0}}TH[e + W\mu^v(s)]\eta^v(s)dy(s) \end{aligned} \tag{8.226}$$

$$\begin{aligned} d\mu_i^v(s) &= \{\frac{1}{\epsilon}[S_i A_i^T(v)][e_i + W_i \mu_i^v(s)] + S_i M_i^v(s)\} ds \\ &+ \frac{1}{\sqrt{N_0}}[S_i - \mu_i^v(s) T_i] H_i [e_i + W_i \mu_i^v(s)] dy(s) \end{aligned} \tag{8.227}$$

$$\eta^v(t) = \eta \tag{8.228}$$

$$\mu_i^v(t) = \mu_i \tag{8.229}$$

Since ϕ^ϵ is optimal,

$$J_{\eta,\mu,t}^\epsilon(u_c) \geq \phi^\epsilon \tag{8.230}$$

Thus, using the hypotheses of the theorem,

$$\phi_0(\eta, t) + \epsilon \phi_1(\eta, \mu, t) \geq \phi^\epsilon(\eta, \mu, t) - \epsilon C(\eta, \mu) \tag{8.231}$$

We will now find the reverse inequality by applying an arbitrary control, v where

$$v(s) = \overline{u}(\eta^v(s), s) + \xi(\eta^v(s), \mu^v(s), s; \mu^v(s) - \underline{\mu}(\eta^v(s), s)) \tag{8.232}$$

Then following the above steps for this arbitrary control we obtain

$$\begin{aligned} &\frac{d}{ds}[\phi_0(\eta^v, s) + \epsilon \phi_1(\eta^v, \mu^v, s)] = \\ &\quad F(\eta^v, s; \tilde{\mu}^v, \xi) + \sum_{i=0}^m \frac{\partial \chi}{\partial \tilde{\mu}_i^v} \cdot G_i(\eta^v, s; \tilde{\mu}^v, \xi) \\ &\quad - C^T[e + W(\tilde{\mu}^v + \underline{\mu}(\eta^v, s))]\eta^v + \epsilon\{\Psi(\eta^v, \mu^v, s)\} \end{aligned} \tag{8.233}$$

But from equation (8.193), the expression

$$\{F(\eta, s; \tilde{\mu}, \xi) + \sum_{i=0}^m D_{\tilde{\mu}_i}\chi \cdot G_i(\eta, s; \tilde{\mu}, \xi)\} \tag{8.234}$$

attains its minimum value, 0, when the control $\xi = \overline{\xi}$ is applied. We have, therefore,

$$\frac{d}{ds}[\phi_0(\eta^v, s) + \epsilon \phi_1(\eta^v, \mu^v, s)] \geq -C^T[e + W(\tilde{\mu}^v + \underline{\mu}(\eta^v, s))]\eta^v + \epsilon\{\Psi\} \tag{8.235}$$

Integrating with respect to s over the interval $[t, T]$, we have

$$\phi_0(\eta, t) + \epsilon \phi_1(\eta, \mu, t) \leq J_{\eta,\mu,t}^\epsilon(v(\cdot)) + \epsilon C(\eta, \mu) \tag{8.236}$$

Recall that up to this point the control v has been arbitrary. Suppose now it is chosen to be u^ϵ, the optimal control for the original problem. Then

$$\phi_0(\eta,t) + \epsilon\phi_1(\eta,\mu,t) \leq \phi^\epsilon(\eta,\mu,t) + \epsilon C(\eta,\mu) \tag{8.237}$$

Combining this equation with equation (8.231) completes the proof.

Lemma 7 *The composite control is ϵ-optimal. That is,*

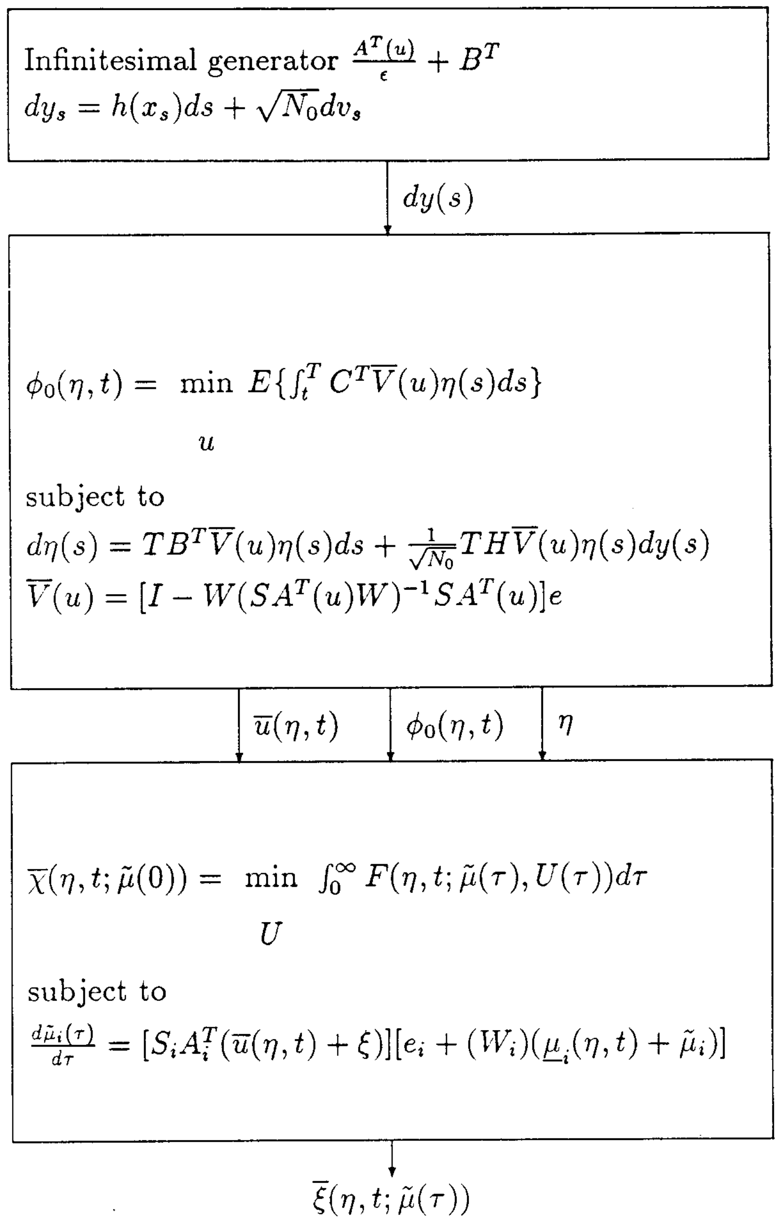
$$|\phi^\epsilon(\eta,\mu,t) - J^\epsilon_{\eta,\mu,t}(u_c)| \leq \epsilon C(\eta,\mu) \tag{8.238}$$

Proof: Immediate from the result of the last lemma and equation (8.224).

8.9 Realization of ϵ-Optimal Controls

The block diagram in Figure 8.1 demonstrates the relationships among the plant, the limit problem, the "fast problem," and the composite control. As noted above, the state vector of the limit problem is of dimension $k+1$, while the dimension of the state vector of the original problem is $(k+1)(m+1)$. The limit problem is thus of significantly lower order than the original optimization problem.

The solution of the limit problem, its state vector, and the optimizing control are used as parameters in the "fast time scale problem" to determine ϕ_1, the next term in the expansion of the optimal value function, ϕ^ϵ. It is noteworthy that the dynamics of the fast time scale problem decompose according to blocks in the A matrix. However, the dynamics are still coupled through the control, u. Moreover, the cost is a quadratic function of the fast states. As a result, although we have obtained separate fast and slow problems, the fast problem cannot readily be further decomposed along states that correspond to blocks of the A matrix.



$$u_c(s) = \overline{u}(\eta^c(s), s) + \overline{\xi}(\eta^c(s), s; \mu^c(s) - \underline{\mu}(\eta^c(s), s))$$

Figure 8.1: System block diagram for the composite control u_c

8.10 Extensions to Results

In this thesis ϵ-optimal controls were found for the problem of minimizing the cost

$$E\{\int_0^T l(x_t^1, x_t^2)dt\} \tag{8.239}$$

subject to the dynamics

$$\begin{aligned} dx_t^1 &= f(x_t^1, x_t^2)dt && (8.240)\\ dx_t^2 &= \frac{1}{\epsilon}b(x_t^1, x_t^2, u_t)dt + \frac{1}{\sqrt{\epsilon}}dw_t && (8.241)\end{aligned}$$

and the observation equation

$$dy_t = h(x_t^1, x_t^2)dt + \sqrt{N_0}dv_t \tag{8.242}$$

Near optimal controls for some more general models can be found with only minor modifications to the presented results. We consider here two such generalizations: (1) dynamical equations that are all stochastically driven, and (2) dynamical equations that are vector valued.

The problem of minimizing the cost of equation (8.239) subhect to the dynamics

$$\begin{aligned} dx_t^1 &= f(x_t^1, x_t^2)dt + dw_t^1 && (8.243)\\ dx_t^2 &= \frac{1}{\epsilon}b(x_t^1, x_t^2, u_t)dt + \frac{1}{\sqrt{\epsilon}}dw_t^2 && (8.244)\end{aligned}$$

and the observations of equation (8.242) requires only a modification to the Markov Chain approximating the dynamics. In particular, the B matrix becomes ([24] and [32])

$$B = \begin{pmatrix} -B_0^+ & B_0^+ & & 0 \\ B_1^- & B_1 & \ddots & \\ & \ddots & \ddots & B_{m-1}^+ \\ 0 & & B_m^- & -B_m^- \end{pmatrix} \tag{8.245}$$

where the diagonal matrices B_i^+, B_i^-, and B_i are given by

$$B_i^+ = \frac{1}{\rho^2}\begin{pmatrix} \frac{1}{2}+\rho f^+(z_{im}) & & 0 \\ & \ddots & \\ 0 & & \frac{1}{2}+\rho f^+(z_{im+k}) \end{pmatrix} \tag{8.246}$$

$$B_i^- = \frac{1}{\rho^2}\begin{pmatrix} \frac{1}{2}+\rho f^-(z_{im}) & & 0 \\ & \ddots & \\ 0 & & \frac{1}{2}+\rho f^-(z_{im+k}) \end{pmatrix} \tag{8.247}$$

$$B_i = \frac{1}{\rho^2}\begin{pmatrix} -1-\rho|f(z_{im})| & & 0 \\ & \ddots & \\ 0 & & -1-\rho|f(z_{im+k})| \end{pmatrix} \tag{8.248}$$

and the definition of the A matrix is unchanged (i.e. it is given by equation (3.23)).

Since only the block structure of the B matrix was used in the derived equations, once the substitution for the$B$ matix is made, all derived results also apply to the system of equations (8.243) and (8.244).

The problem of minimizing the cost of equation (8.239) subject to the dynamics of equations (8.243) and (8.244) (in which x_t^1 and x_t^2 are now vector valued processes) and the observations of equation (8.242) has not been solved in detail. However, there appears to be no conceptual problem in applying the procedure used in this thesis to solve it.

Bibliography

[1] Václav E. Beneš and Ioannis Karatzas, "On the Relation of Zakai's and Mortensen's Equations," SIAM Journal on Control and Optimization, Vol. 21, No. 3, May 1983. pp. 472–489.

[2] A. Bensoussan, *Methodes de Perturbations en Controle Optimal*, Preprint, 1986.

[3] A. Bensoussan, *Stochastic Control by Functional Analysis Methods*, New York: North-Holland Publishing Company, 1982.

[4] A. Bensoussan, "Maximum Principle and Dynamic Programming Approaches of the Optimal Control of Partially Observed Diffusions," *Stochastics*, Vol. 9, 1983. pp. 169–222.

[5] A. Bensoussan, "Stochastic Maximum Principle for Distributed Parameter Systems," *Journal of the Franklin Institute*, Vol. 315, No. 5/6, May/June 1983, pp. 387–406

[6] D. Bertsekas, *Dynamic Programming and Stochastic Control*, New York: Academic Press, 1976.

[7] Roger Brockett, *Finite Dimensional Linear Systems*, New York: John Wiley and Sons, Inc., 1970.

[8] A. Bryson, Jr., and Y. Ho, *Applied Optimal Control*, Waltham, Massachusetts: Ginn and Company, 1969.

[9] J. H. Chow and P. V. Kokotovic, "A two stage Lypunov-Bellman feedback design of a class of nonlinear systems," *IEEE Transactions on Automatic Control*, AC-26, pp. 656–663.

[10] M. Coderch, A. S. Willsky, S. S. Sastry, and D. A. Castanon, "Hierarchical Aggregation of Singularly Perturbed Finite State Markov Processes," *Stochastics*, Vol. 8, 1983. pp. 259–289.

[11] M. Coderch, A. S. Willsky, S. S. Sastry, and D. A. Castanon, "Hierarchical Aggregation of Linear Systems with Multiple Time Scales," *IEEE Transations on Automatic Control*, Vol. AC-28, No. 11, November 1983. pp. 1017–1029.

[12] P. J. Courtois, *Decomposability Queueing and Computer Systems Applications*, New York: Academic Press, 1977.

[13] François Delebecque and Jean-Pierre Quadrat, "Optimal Control of Markov Chains Admitting Strong and Weak Interactions," *Automatica*, Vol. 17, No. 2, 1981. pp. 281–196.

[14] F. Delebecque, "A Reduction Process for Perturbed Markov Chains," *SIAM Journal on Applied Mathematics*, Vol. 43, No. 2, April 1983. pp. 325–350.

[15] J. L. Doob, *Stochastic Processes*, New York: John Wiley & Sons, Inc., 1953.

[16] Robert J. Elliot, *Stochastic Calculus and Applications*, New York: Springer-Verlag, 1982,

[17] Wendell H. Fleming and Etienne Pardoux, "Optimal Control for Partially Observed Diffusions," *SIAM Journal on Control and Optimization*, Vol. 20, No. 2, March 1982. pp. 261–285.

[18] Wendell H Flemming, "Nonlinear Semigroup for Controlled Partially Observed Diffusions," *SIAM Journal on Control and Optimization*, Vol. 20, No. 2, March 1982. pp. 286–301.

[19] Avner Friedman, *Partial Differential Equations of Parabolic Type*, Englewood Cliffs, N.J.: Prentice Hall, Inc, 1964.

[20] Daniel T. Gillespie, "The Mathematics of Simple Random Walks," *Naval Research Reviews*, Vol. 3, 1983. pp. 46–52.

[21] R. Z. Has'minskiĭ, *Stochastic stability of differential equations*, Rockville, Maryland: Sijthoff & Noordhoff, 1980.

[22] T. Kato, *Perturbation Theory for Linear Operators*, New York: Springer-Verlag, 1984.

[23] Harold J. Kushner, *Introduction to Stochastic Control*, New York: Hold, Rinehart and Winston, Inc., 1971.

[24] Harold J. Kushner, *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, New York: Academic Press, 1977.

[25] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, New York: Springer-Verlag, 1971.

[26] M. Loeve, *Probability Theory*, New York: Springer-Verlag, 1978

[27] Charles–Michel Marchetti, "Methodes de Perturbations Singulieres en Filtrage Nonlineaire," These pour Doctorat de Troisieme Cycle de Mathematiques Appliquees, Universite de Provence Centre Saint Charles, Juillet 1986.

[28] G. B. Di Masi, M. Pratelli, and W. J. Runggaldier, "An Approximation for the Nonlinear Filtering Problem, with Error Bound," *Stochastics*, Vol. 14, 1985. pp. 247-271.

[29] R. E. Mortensen, "*A Priori* Open Loop Optimal Control of Continuous Time Stochastic Systems," *International Journal of Control*, Vol. 3, No. 2, 1966. pp. 113–127.

[30] R. E. Mortensen, "Stochastic Optimal Control with Noisy Observations," *International Journal of Control*, Vol. 4, No. 5, 1966. pp. 455–464.

[31] Randolph G. Phillips and Petar V. Kokotovic, "A Singular Perturbation Approach to Modeling and Control of Markov Chains," *IEEE Transactions on Automatic Control*, Vol. AC–26, No. 5, October 1981. pp. 1087–1094.

[32] M. Piccioni, "Convergence of Implicit Discretization Schemes for Linear Differential Equations with Applications to Filtering," Systems Research Center Technical Report TR-86-25.

[33] Jan Robin Rohlicek, "Aggregation and Time Scale Analysis of Perturbed Markov Systems," Ph.D. dissertation, Massachusetts Institute of Technology, January 1987.

[34] Wilson J. Rugh, *Nonlinear System Theory*, Baltimore: The Johns Hopkins University Press, 1981.

[35] Zeev Schuss, *Theory and Applications of Stochastic Differential Equations*, New York: John Wiley & Sons, 1980.

[36] Richard D. Smallwood and Edward J. Sondik, "The Optimal Control of Partially Observable Markov Processes over a Finite Horizon," *Operations Research*, No. 21, 1973, pp. 1071-1088.

[37] Eugene Wong, *Stochastic Processes in Information and Dynamical Systems*, Huntington, New York: Robert E. Krieger Publishing Company, 1979.

[38] W. M. Wonham, "Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering," *J. SIAM Control*, Ser. A, Vol. 2, No. 3, 1965, pp. 347-369.

[39] Kôsaku Yosida, *Functional Analysis*, New York: Springer-Verlag, 1980.