OPTIMAL DYNAMIC ROUTING IN AN UNRELIABLE QUEUING SYSTEM

by

JOHN NIKOLAOS TSITSIKLIS

B. Sc., Massachusetts Institute of Technology

(1980)

SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE

DEGREE OF

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January, 1981

[redacted]

Signature of Author .....

Department of Electrical Engineering and Computer Science, January 22, 1981.

Certified by .....

Dr. Stanley B. Gershwin

Thesis Supervisor

Certified by .....

Professor Michael Athans

Thesis Supervisor

Accepted by .....

Chairman, Departmental Committee on Graduate Students

OPTIMAL DYNAMIC ROUTING IN AN UNRELIABLE QUEUING SYSTEM

by

JOHN NIKOLAOS TSITSIKLIS

Submitted to the Department of Electrical Engineering and Computer Science
on January 22, 1981 in partial fulfillment of the
requirements for the Degree of Master of Science in
Electrical Engineering and Computer Science

ABSTRACT

This thesis considers the optimal dynamic routing problem in a queuing system of three machines and two finite storage buffers. Machines are assumed to be failure prone and the material level in the buffers is assumed to be continuous. The objective is to optimize the long-run average performance of the system.

Cost-to-go functions of dynamic programming are defined and a set of partial differential equations for the cost-to-go functions is derived under any fixed dynamic routing strategy. Necessary and sufficient conditions for optimality involving the cost-to-go functions are also derived. Two iterative algorithms for optimizing the performance of the system ar proposed.

The case where the average production rate of the system is to be maximized is considered in more detail and the particular case where the lead machine is perfectly reliable is completely solved, theoretically and numerically.

Thesis Supervisor:	Stanley B. Gershwin
Title:	Lecturer, Department of Electrical Engineering
	and Computer Science
	Principal Research Scientist, Laboratory for
	Information and Decision Systems
Thesis Supervisor:	Michael Athans
Title:	Professor of Systems Science and Engineering
	Director, Laboratory for Information
	and Decision Systems

Acknowledgements

I would like to thank Professor Mike Athans and Dr. Stan Gershwin for suggesting the topic, their guidance, helpful discussions and for reading the manuscript. In addition, Professor Bernard Lévy and Dr. Dave Castanon for clarifying some fine theoretical points; Hellen Hahne, for helping my understanding of the problem. Finally, Panos Cavoulacos and John Bakopoulos for introducing me to the text editing program, TEX, that was used to typeset this thesis.

This research was carried out in the M.I.T. Laboratory for Information and Decision Systems (formerly the Electronic Systems Laboratory) with support extended by National Science Foundation Grants DAR78-17826.

Any opinions, findings and conclusions or recommendations expressed in this publication are those of the author, and do not necessarily reflect the views of the National Science Foundation.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

From an abstract point of view, manufacturing networks with storage spaces are just queuing networks. The main body of queuing theory has been concerned with the properties of queuing systems that are operated in a certain, fixed fashion. (Kleinrock, 1975) Recently however, some interest has arisen in comparing different ways to design (Schick and Gershwin, 1978) or operate (i.e. control) a queuing system. (Foschini, 1977; Foschini and Salz, 1978; Hahne, 1980; Ephremides, Varaiya and Warland, 1980; Olsder and Suri, 1980)

The main problem in controlling a queuing system appears when we are given a network of queues in which there is potential for a decision as to how material will be routed in the network. This thesis aims toward the development of a methodology for solving the routing problem for continuous queuing systems, i.e. systems in which the queue levels are continuous variables.

1.1. Problem Description

The subject of this thesis is the dynamic routing of material in flexible, unreliable networks with queues. Each node of the network is meant to represent a processor of some sort. The material flowing into the node is being processed at the node and then it is routed to a downstream node. We consider flexible networks, that is networks in which there is potential for a decision regarding the routing of material that has been processed by a node, or networks in which there are one or more processors that can operate on different kinds of material with no switchover time. These networks are also unreliable

in the sense that each of the processors (nodes) fails and is repaired in a random manner, but according to a known probabilistic law. Finally, we assume that each processor is preceeded by a finite storage space (buffer, queue) in which the material flowing into the node may be stored for an indefinite amount of time before it is processed.

Some examples of such networks are the following:

a) Manufacturing networks in which each node represents an unreliable machine. The simplest manufacturing network is a production line, in which some material is processed serially by a sequence of machines to produce a final good. Cleraly, there is no routing choice in a production line. However, in some more complex manufacturing networks, two machines (or two sequences of machines) can perform the same tasks. In this case, the routing problem appears naturally. (Hahne, 1980)

b) Communication networks in which each node represents a computer. The material flowing into the network consists of messages sent from one computer to another. Given an origin-destination pair, there are usually many alternative paths between which we can choose.

c) Hydraulic systems (Buzacott, 1971) or continuous chemical processes. Here fluids or chemicals flow through series of unreliable stages separated by holding tanks.

Although the methodology in this thesis is a general one, we use models and examples corresponding to manufacturing networks and occasionally comment on the modifications required to handle communication networks. Consequently, from now on, the nodes of the network are referred to as machines.

Given a flexible network one needs a rule for making a decision whenever there is a choice present. Any such rule will be called a routing strategy or a control law. The

problem which we aim to solve is to determine a routing strategy that will optimize the performance of the system with respect to some prespecified performance criterion. The optimal routing strategy clearly depends on the set of admissible routing strategies that has been chosen.

Traditionally, researchers have been interested in problems involving a relatively simple set of admissible strategies, namely the so-called static and quasi-static strategies. In a static strategy, the routing decision is the same at all times regardless of which machines are down and regardless of the buffer levels. In a quasi-static strategy the routing decision may depend on the availability of the machines but not on any information concerning the buffer levels. (Cantor and Gerla, 1974)

Such restricted classes of strategies are attractive because the set of admissible strategies is a subset of a vector space of reasonably small finite dimension. Traditional vector space optimization techniques can be used to find an optimal static or quasi-static strategy. For example, the routing problem in a reliable communication network has been formulated as a nonlinear minimum cost multicommodity network flow problem. (Bertsekas, 1979) However, a truly optimal strategy belongs, in general, to the class of dynamic strategies. These are strategies in which all information available is used whenever a decision is to be made. The optimal dynamic strategy has better performance than the optimal quasi-static strategy. On the other hand dynamic strategies have certain drawbacks: the set of strategies is a subset of a vector space of huge finite dimension, when the buffer levels are discrete variables, or of an infinite dimensional vector space (a function space) when the buffer levels are continuous variables. This leads to considerable computational complexity. Dynamic strategies are also harder to implement because they

presuppose a fast flow of significant amounts of information throughout the network, as well as the existence of some information processor. This is not a problem for manufacturing networks, although it may lead too increased costs in other situations. Despite these difficulties a theoretical investigation of dynamic strategies is of interest because it may suggest simple but good suboptimal heuristic rules for operating a network.

We study networks in which the material flow and the buffer levels are continuous variables. Such models can be considered as either true representations of a real network (like a chemical plant involving continuous flows of chemical materials) or, more importantly, as an approximation of a network of machines processing discrete but very small components. Indeed, as the component size becomes very small compared with the buffer capacities and as the processing time of a single component becomes very small compared with the time constants of the network (mean time to failure or repair of a machine, mean time to fill an empty buffer, etc.) we expect a continuous model to be a good approximation to the actual discrete system. (Gershwin and Schick, 1980)

An issue that must be considered is the choice of an appropriate continuous model. This model will not be unique but will depend on the sources of randomness inherent in the system.

There are, in general, four sources of randomness in a queuing network:

a) Random machine failures and repairs.

b) Random arrivals of components to be processed.

c) Random processing times of components.

d) Randomized routing strategies (i.e. routing strategies in which decisions are taken by flipping a biased coin).

If (a) is the only source of randomness, then the system behaves deterministically during any time interval in which no repair or failure occurs and it is relatively easy to specify an appropriate model.

If any of the sources of randomness (b), (c) or (d) is present, the system may behave in the limit like a diffusion process and the modelling problem is non-trivial. (Foschini, 1977; Foschini and Salz, 1978)

The failure and repair processes of the machines are assumed to be memoryless. This is a quite reasonable assumption which can often be physically justified and is adopted in many studies involving unreliable machines. (Koenigsberg, 1959; Sarma and Alam, 1975) Under this assumption, the system can be described by a simple Markov process. Stochastic control problems are almost always considered in a Markovian setting and this is why the above assumption is important.

It is known that any stochastic process becomes a Markov process if the state space is appropriately augmented. This usually results in untractable infinite dimensional state spaces. The memorylessness assumption keeps the state space finite–dimensional and with a small dimension. In many cases, with the augmentation of just a few states, we can come quite close to realistic distributions of repair and failure times. However, models with augmented state spaces are not considered in this thesis, because the machinery for the simple memoryless case should be developed first.

We study in detail one of the simplest possible routing problem, corresponding to the network in Figure 1.1. Here M_0, M_1 and M_2 are the nodes (machines, processors) of the network and B_1 and B_2 are finite buffers (storage spaces, queues).

Although one should be concerned with more general network topologies, this simple

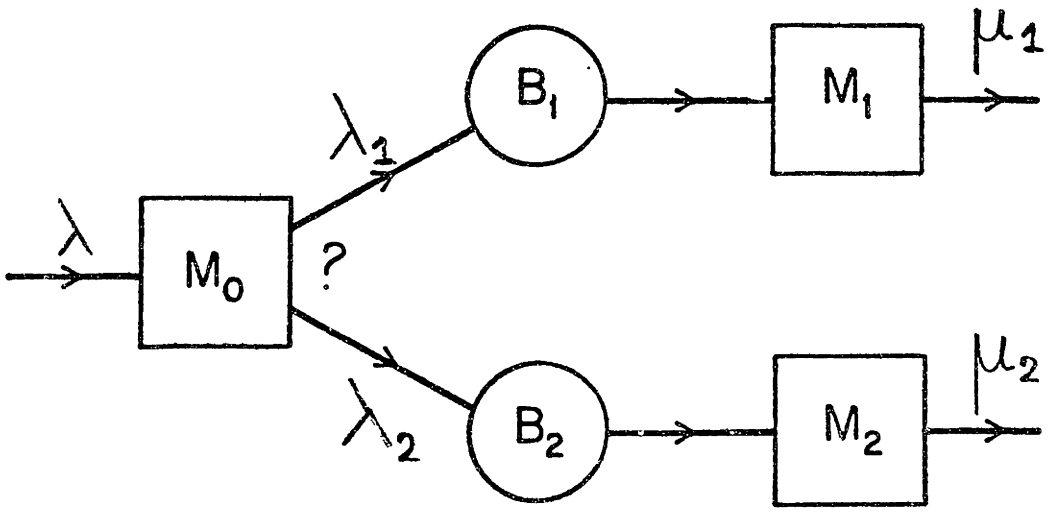


Figure 1.1: A Queuing System With a Single Decision

configuration is of interest because the methodology developed to handle this network is directly applicable to arbitrary network topologies. Moreover, the structure of the problem is simple enough so that we can get an intuitive feel for the results obtained. We also believe that the solution to this problem can be, possibly with some modifications, the building block for a decentralized strategy in a large scale network.

Finally, we should point out that we consider the infinite time horizon optimization problem. We are interested in the long-run average value of the performance criterion to be imposed; because of this, we restrict to stationary routing strategies, namely strategies that depend only on the state of the system (i.e. the machine repair states and the buffer levels) and not on the current time.

1.2. Description of Related Research.

In the field of data communication networks, researchers have dealt mainly with the problems of static and quasi-static routing and the stability properties of such algorithms. (Cantor and Gerla, 1974; Gafni, 1979; Bertsekas, 1979; Gallager,1977) The problem is usually formulated as a deterministic nonlinear minimum cost network flow problem to which various optimization techniques are applied.

A different approach was taken by Segall (1976) and Moss (1977) which led to a dynamic formulation of the problem but, unlike this thesis, a deterministic one.

In the field of manufacturing networks much work has been done in the analysis of the statistical behavior of such networks, in terms of steady-state probability distributions. Unreliable networks involving no decision (Schick and Gershwin, 1978; Gershwin

and Schick, 1979) as well as flexible networks have been considered. For example, the performance of alternative quasi-static strategies is analyzed by Gershwin and Ammar (1979).

The stochastic dynamic routing problem has also been studied by various researchers: in a discrete context (Ephremides, Varaiya and Warland, 1980; Hahne, 1980), in a continuous context (Santana and Platzman, 1979; Olsder and Suri, 1980; Deuermeyer and Pierskalla, 1978); and in a diffusion approximation context (Foschini, 1977; Foschini and Salz, 1978). Some of the researchers guess or derive heuristically an optimal solution and then proceed to prove that the proposed solution satisfies the conditions for optimality. While this approach works in simple problems with much symmetry, it does not give any insight as to how an optimal solution could be obtained in a general setting.

Two algorithmic approaches are given by Hahne (1980) who uses a successive approximation algorithm to solve a discrete Markovian decision problem and by Santana and Platzman (1979) who successively approximates cost-to-go functions by piecewise linear functions on the state space.

The only available optimality conditions for the routing problem follow from the Dynamic Programming Principle (Bellman and Dreyfus, 1962) and make use of cost-to-go functions. The theoretical foundation of the dynamic programming methodology for continuous state, continuous time processes has been recently developed. See for example the work of Boel and Varaiya (1977), Kushner (1965, 1967), Rishel (1975a, 1975b, 1977), Sworder (1969) for problems where total cost is minimized and of Wonham (1970), Stone (1973) and Kushner (1978) for problems where average cost is minimized.

1.3. Summary

In Chapter 2 we build a model for the simple queuing network of Figure 1.1 under a very general dynamic routing policy, in the case that machine failures and repairs are the only sources of randomness in the system. We model the system as a Markov process whose state consists of the buffer levels and a list of the repair conditions of the machines. We describe the evolution of the state variables as follows: Between any two failures or repairs of the machines the buffer levels change deterministically in a way determined by the routing strategy being used. A failed machine gets repaired at a predetermined repair rate; a working machine may fail at a rate proportional to the degree of its utilization.

We also consider the various performance criteria that can be imposed. Apart from the fact that we consider infinite time horizon, average performance, we allow the performance criterion to be as general as possible. However, we consider in more detail the case where weighted average production rate is maximized.

In Chapter 3 we assume that a certain routing strategy has been selected and we consider the behavior of the system under that assumption. We show that under any reasonable routing strategy, there exists a special state of the system which is visited an infinite number of times. This establishes certain ergodicity properties of the system and shows that long-run average quantities are well-defined. We discuss various methods for computing the value of the performance criterion corresponding to a fixed control law and we concentrate on the dynamic programming methodology and the concept of value (cost-to-go) functions. We derive a set of partial differential equations which, when solved, yields the value of the performance criterion.

In Chapter 4 we turn our attention to the optimizationn problem. We derive condi-

tions for optimality in terms of the value functions and then proceed to their interpretation. We show that they imply the following characterization of optimal routing strategies: The state space may be divided into three regions; in the first region, all material is routed to the first buffer; in the second region, all material is routed to the second buffer. A dividing line separates these two regions; if the state of the system ever reaches that line, it stays on it until a machine failure or repair occurs. (This can only be done by appropriately splitting the flow through the upstream machine). The third region is an indifference region and may appear in some special problems. The value of the performance criterion is not affected by the control law in that region.

Next, we formulate an iterative algorithm which, under some conditions, converges to an optimal control law. The algorithm finds at each stage the point at which conditions for optimality are not satisfied and makes an appropriate correction.

We have pointed out that the behavior of the system is deterministic until a machine failure or repair occurs. An optimal control law should certainly behave optimally during the time interval in which the system is deterministic. It follows that an optimal control law should simultaneously solve a certain deterministic optimal control problem. We formulate an iterative algorithm that converges to an optimal stochastic control law and uses as basis this deterministic problem. (Chapter 5)

In Chapter 6 we consider again the maximization of average production rate and we derive the following special results (under some inequality assumptions on the capacities of the machines): If the lead machine is perfectly reliable, or if the lead machine fails and is repaired very frequently, then any two routing strategies which coincide when both buffers are empty (but may be different everywhere else) have the same performance. We

then show how the dimension of the state space may be reduced and solve the optimal routing problem completely. Numerical results are also obtained for a few representative cases.

1.4. Contributions.

1. We demonstrate that under very general conditions there exists a special state of the state space of the system considered that is visited an infinite number of times and the inter–arrival times have bounded expecation. We then show that long–run average quantities are always well–defined and independent from the initial state.

2. We define rigorously cost-to-go functions and derive a set of partial differential equations satisfied by them. We show the uniqueness of solution of that set of equations.

3. We derive necessary and sufficient conditions for optimality for the routing problem in terms of the cost-to-go functions. We also derive a theorem that compares the performance of two different control laws.

4. We define, within the context of the routing problem, a deterministic control problem associated with the stochastic control problem and show the connection of the two.

5. We propose two iterative algorithms that solve the stochastic control problem.

6. We solve completely (including numerical results) the routing problem if the lead machine is perfectly reliable (or fails and gets repaired very frequently) under a few additional assumptions. We prove that the behavior of control laws when not both buffers are empty does not matter.

CHAPTER 2: A CONTINUOUS MODEL OF THE NETWORK

The behavior of a system in which machine failures and repairs are the only sources of randomness is qualitatively different from the general model when other sources of randomness are present, because, in this particular case, the evolution of the state variables is deterministic between any two consecutive failures or repairs.

In section 2.1 we develop a model for such systems and interpret the key variables.

In section 2.2 we complete the model by introducing a set of physical constraints (implied by conservation of flow and the finiteness of the queues) and a set of mathematical constraints which are necessary in order to obtain a mathematically well-defined stochastic process.

In section 2.3 we comment on various possible performance criteria and formulate a set of mathematical assumptions that a performance criterion should satisfy. We also introduce some additional constraints on the ways that the system may be operated which are quite reasonable when the performance criterion is average production rate.

2.1.1. Machine Behavior and Buffer Levels

Consider again the simple configuration introduced in Chapter 1 (Fig. 1.1). Machines M_0, M_1 and M_2 can each be in one of two states: up or down. We define an indicator variable α_i corresponding to machine M_i as follows:

$$\alpha_i = \begin{array}{ll} 0 & \text{if machine } M_i \text{ is down} \\ 1 & \text{if machine } M_i \text{ is up} \end{array} \qquad (2.1.1)$$

We let B_1 and B_2 be buffers (queues) with a finite storage capacity. Let x_i be the level of buffer B_i and let N_i be the maximum allowed level in the same buffer. Then,

$$0 \leq x_i \leq N_i \tag{2.1.2}$$

As we have said in the introduction, we consider only continuous models; so, x_i is a continuous variable.

We define the state of the system as

$$s \equiv ((x_1, x_2), (\alpha_0, \alpha_1, \alpha_2)) \equiv (\underline{x}, \underline{\alpha}) \tag{2.1.3}$$

The corresponding state space is:

$$S \equiv [0, N_1] \times [0, N_2] \times \{0, 1\}^3 \tag{2.1.4}$$

At any time, the state s contains all information relevant to the future evolution of the system. This is an informal way to say that the temporal evolution of s should be a Markov process.

Since we are building a continuous model, we can view the material being processed by the machines as a continuous flow. Let λ^* be the maximum possible flow rate through machine M_0, and μ_i^* be the maximum flow rate through machine M_i, for $i = 1, 2$. Let $\lambda(t)$, $\mu_1(t)$, $\mu_2(t)$ br the actual flow rates through machines M_0, M_1, M_2, respectively, at time t. Finally let $\lambda_1(t)$, $\lambda_2(t)$ be the actual flow rates from machine M_0 to buffers B_1 and B_2, respectively, at time t. Since machine M_0 has no inner storage, conservation of flow

implies:

$$\lambda(t) = \lambda_1(t) + \lambda_2(t) \tag{2.1.5}$$

Our assumptions on the sources of randomness of the system imply that all flows through the network can be set exactly to a desired level provided that no physical constraints are violated. An additional implication is that once the flow rates have been decided by the controller, the state of the system can be predicted with absolute certainty until the next time a working machine fails or a damaged machine is repaired.

Clearly, the assumption that input to the lead machine is always available (and in particular it is not random) is indispensable here. In contrast with discrete systems, *(Hahne, 1980)*, it makes no difference whether the routing decision is made after the material leaves, or before it enters machine M_0. This is because, in a continuous model, it takes an infinitesimal amount of time for the material to go through M_0.

We have defined a routing strategy as "a way to operate the system". In the present case, a way to operate the system is completely determined if, for any state in the state-space, we specify λ, λ_1, λ_2, μ_1 and μ_2. So, instead of viewing these variables as time functions, it is more useful to view them as functions on the state space. We define a routing strategy u as a set of functions on the state space.

$$u \equiv (\lambda(\underline{x}, \underline{a}), \lambda_1(\underline{x}, \underline{a}), \lambda_2(\underline{x}, \underline{a}), \mu_1(\underline{x}, \underline{a}), \mu_2(\underline{x}, \underline{a})) \tag{2.1.6}$$

Of course, this cannot be an arbitrary set of functions. There are constraints arising from both physical and mathematical considerations. These are treated in the next section.

We have said sofar enough to describe the law of evolution of the variables x_1, x_2. Namely, conservation of flow implies, for small Δt:

$$x_i(t+\Delta t) = x_i(t) + [\lambda_i(\underline{x}(t), \underline{\alpha}(t)) - \mu_i(\underline{x}(t), \underline{\alpha}(t))]\Delta t \quad i = 1, 2 \tag{2.1.7}$$

which is abbreviated as

$$dx_i = (\lambda_i - \mu_i)\, dt \tag{2.1.8}$$

We note that this would be a deterministic law, were it not for the random changes in $\underline{\alpha}(t)$. We must now describe the probabilistic law of evolution of the variables α_0, α_1, α_2.

If machine M_i $(i = 0, 1, 2)$ is down, we assume that the time until it is repaired is an exponential random variable with mean $1/r_i$, where r_i is the repair rate of machine M_i.

There are two alternatives, however, concerning the failures of the machines. We could assume that the time to failure is an exponentially distributed random variable with rate p_i. This would be a good model if failures were due to external causes (like power failures) and would imply that a machine may fail even at a time when it is not being operated.

However, in manufacturing networks, failures are caused by the utilization of the machines. It seems reasonable to scale the failure rate of a machine proportionately to the degree that it is being utilized *(Gershwin and Schick, 1980)*. We define $p_0{}^*$, $p_1{}^*$, $p_2{}^*$ as the failure rates of machines M_0, M_1, M_2, respectively, when these machines are operated in full capacity. Then, we let the actual failure rate be equal to:

$$p_0(\underline{x}(t), \underline{\alpha}(t)) = p_0{}^* \frac{\lambda(\underline{x}(t), \underline{\alpha}(t))}{\lambda^*} \tag{2.1.9}$$

$$p_i(\underline{x}(t), \underline{\alpha}(t)) = p_i{}^* \frac{\mu_i(\underline{x}(t), \underline{\alpha}(t))}{\mu_i{}^*} \quad i = 1, 2 \tag{2.1.10}$$

Failure rates can be viewed as functions on the state space or as functions of time, i.e. as stochastic processes. Let $R_i(t)$ be an integer valued, right continuous process which increases by one each time that machine M_i is repaired. Let $F_i(t)$ be an integer valued, right-continuous process which increases by one each time that machine M_i fails. $R_i(t)$ and $F_i(t)$ are not Poisson processes because the instantaneous jumping rate depends on the current state of the system which is itself a stochastic process depending on the past of $F_i(t)$ and $R_i(t)$. They belong to the class of processes sometimes called "randomly modulated jump processes" *(Segall and Kailath, 1975).*

Despite the apparent complication, R_i, F_i are still well-defined stochastic processes. One way of rigorously defining $F_i(t)$ is to require that $F_i(t)$ is a right-continuous jump process with unit jumps such that

$$F_i(t) - \int_0^t p_i(\underline{x}(\tau), \underline{\alpha}(\tau))\, d\tau \tag{2.1.11}$$

is a martingale with respect to a family of σ-fields $\mathfrak{B}_t$ such that $\underline{x}(t)$ and $\underline{\alpha}(t)$ are $\mathfrak{B}_t$ measurable for all t. In simpler words,

$$E\left[F_i(t) - \int_0^t p_i(s)\, ds \,|\, (\underline{x}(\tau), \underline{\alpha}(\tau)), \tau \in [0, t_0]\right] = F_i(t_0) - \int_0^{t_0} p_i(s)\, ds \tag{2.1.12}$$

Then *(Segall and Kailath, 1975)*, the process $F_i(t)$ is uniquely defined and has all desired propreties.

A second way of rigorously defining $F_i(t)$ is the following: Consider a standard Poisson process $F_i{}^*(t)$ with rate $p_i{}^*(t)$. Suppose that its first jump occurs at time $t_0{}^*$. Then we let

the process $F_i(t)$ have its first jump at a time t_0 such that $\int_0^t \mu_i(\tau)/\mu_i{}^* \, d\tau = t_0{}^*$, and so on. In other words, $F_i(t)$ can be obtained by a time transformation of a standard Poisson process.

The processes $R_i(t)$ may also be rigorously defined in a slightly different way.

The second definition leads to a very simple representation of the probability space on which the stochastic proces is defined. Let Ω_{F_i}, Ω_{R_i} be the sample spaces of standard Poisson processes with rates $p_i{}^*$ and r_i respectively. Let $P_{F_i}{}^*$, P_{R_i} be the corresponding probability measures. We define the probability space of our Markov process by:

$$\Omega = \Omega_{F_0} \times \Omega_{F_1} \times \Omega_{F_2} \times \Omega_{R_0} \times \Omega_{R_1} \times \Omega_{R_2} \tag{2.1.13}$$

and we endow Ω with the product measure. If a routing strategy (control law) and an initial state has been fixed, then, for any $\omega \in \Omega$, the time evolution of the system is completely determined.

If $N_i(t)$ is a right-continuous jump process, we define $dN_i(t)$ by

$$dN(t) \equiv N(t) - \lim_{\tau \to t-} N(\tau) \equiv N(t) - N(t-) \tag{2.1.14}$$

Then, we can describe the evolution of the variables $\underline{\alpha}(t)$ by:

$$d\alpha_i = (1 - \alpha_i) dR_i - \alpha_i dF_i \tag{2.1.15}$$

Equations (2.1.8), (2.1.15) and the definition of the processes R_i and F_i provide a complete probabilistic description of a unique Markov process.

2.1.2. Interpretation of the Routing Variables.

If our model is considered as a continuous representation of a discrete system we have to solve the following problem: given a non-randomized routing strategy for a discrete system with deterministic processing times of components find piecewise continuous functions $\lambda_1(\underline{x}, \underline{a})$, $\lambda_2(\underline{x}, \underline{a})$, $\mu_1(\underline{x}, \underline{a})$, $\mu_2(\underline{x}, \underline{a})$ such that the behavior of the continuous system is close to that of the discrete system. Conversely, given the above set of functions, find a strategy for the discrete system that corresponds to these functions.

Given any initial state of the discrete system, and a fixed routing strategy, the time evolution of the state is deterministic until the first machine failure or repair and may be described by a path in the state space. This path can be approximated by a piecewise continuously differentiable path on a continuous state space. Then by evaluating the slope of the continuous path at each point we may obtain λ_1, λ_2, μ_1, μ_2 for those $\underline{x}$ that belong on the path. This procedure has to be repeated for all initial states. Then, these functions can be smoothly extrapolated and be defined on the entire state space. (Fig. 2.1)

Conversely, given λ_1, λ_2, μ_1, μ_2, we obtain a family of continuous paths on the continuous state space. We may discretize these paths (Fig. 2.1) and obtain a path for each possible initial state on the discrete state space. This will determine the routing strategy for the actual discrete system.

It should be clear from the above discussion that the model

$$dx_i = (\lambda_i - \mu_i)\,dt \tag{2.1.16}$$

cannot represent a discrete system in which a randomized strategy is used, because a

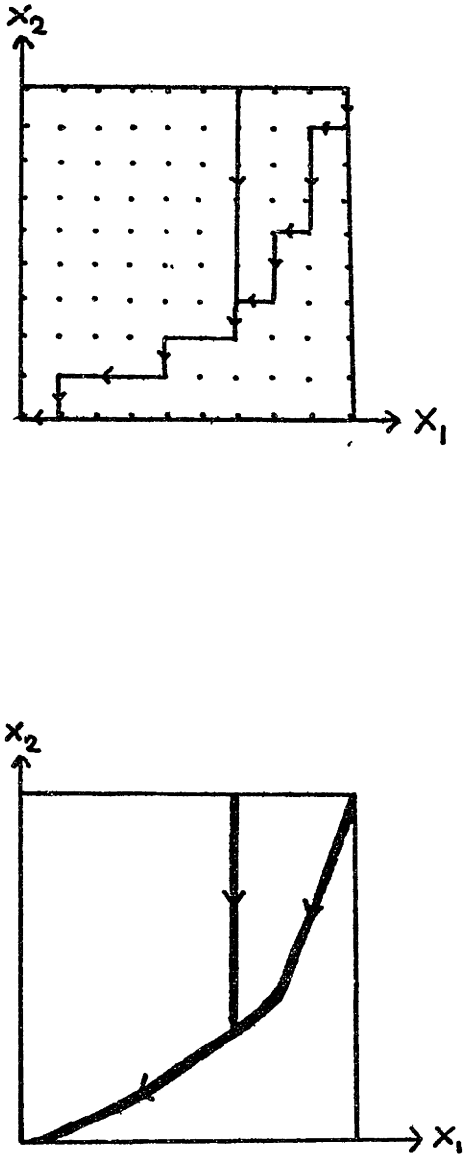


Figure 2.1: Relation between the Discrete and the Continuous Model

randomized strategy does not give rise to deterministic paths but to a random walk. This has no serious consequences, because a discrete Markovian decision theory problem always has an optimal solution which is not randomized. This is so, because any such decision problem is equivalent to a linear programming problem *(Shapiro, 1979)* and non-randomized strategies correspond to corners of the feasible set. Therefore, non-randomized strategies are at least as good as randomized ones.

2.2. Mathematical and Physical Constraints

In this section we complete the development of the model. We impose a set of constraints needed to ensure that a Markov process with the properties described in the last section exists. We start with the physical constraints which are relatively simple:

$$\lambda(\underline{x}, \underline{a}) = \lambda_1(\underline{x}, \underline{a}) + \lambda_2(\underline{x}, \underline{a}) \tag{2.2.1}$$

$$\text{if } x_i = 0 \text{ then } \mu_i(\underline{x}, \underline{a}) \leq \lambda_i(\underline{x}, \underline{a}) \quad i = 1, 2 \tag{2.2.2}$$

$$\text{if } x_i = N_i \text{ then } \mu_i(\underline{x}, \underline{a}) \geq \lambda_i(\underline{x}, \underline{a}) \quad i = 1, 2 \tag{2.2.3}$$

$$\text{if } a_0 = 0 \text{ then } \lambda(\underline{x}, \underline{a}) = 0 \tag{2.2.4}$$

$$\text{if } a_i = 0 \text{ then } \mu_i(\underline{x}, \underline{a}) = 0 \tag{2.2.5}$$

These constraints imply that buffer levels cannot be negative or exceed the buffer capacity and that there is no flow through machines which are down.

Now we turn to the mathematical constraints. We define a family of σ-algebras $(\mathfrak{B}_t,\ t \in [0,\infty))$ where $\mathfrak{B}_t$ is the smallest σ-algebra with respect to which the random vector $(\underline{x}(t), \underline{\alpha}(t))$ is measurable. We must determine what additional conditions have to be satisfied if $(\underline{x}(t), \underline{\alpha}(t))$ is a well-defined Markov process.

We require that λ, λ_1, λ_2, μ_1, μ_2 are Borel measurable functions on the state space. This is needed to ensure that $\lambda(t)$, $\lambda_1(t)$, $\lambda_2(t)$, $\mu_1(t)$, $\mu_2(t)$ are $\mathfrak{B}_t$-measurable random variables. We note that starting from an arbitrary point in time and until a change in $\underline{\alpha}$ occurs (i.e. until a repair or a failure occurs) $\underline{x}(t)$ evolves deterministically. So, we must require that the deterministic differential equation

$$dx_i = (\lambda_i(\underline{x}, \underline{\alpha}) - \mu_i(\underline{x}, \underline{\alpha}))dt \tag{2.2.6}$$

has a unique piecewise continuous differentiable solution for any fixed $\underline{\alpha}$. This will certainly be the case if λ_i and μ_i satisfy a Lipschitz condition: there exists some d such that

$$|\lambda_i(\underline{x}, \underline{\alpha}) - \lambda_i(\underline{x^*}, \underline{\alpha})| \leq d\|(\underline{x}, \underline{\alpha}) - (\underline{x^*}, \underline{\alpha})\| \tag{2.2.7}$$

$$|\mu_i(\underline{x}, \underline{\alpha}) - \mu_i(\underline{x^*}, \underline{\alpha})| \leq \|(\underline{x}, \underline{\alpha}) - (\underline{x^*}, \underline{\alpha})\| \tag{2.2.8}$$

where $\|\ \|$ is the Euclidean norm. However, this condition is unnecessarily strong. In fact, in Chapter 4, we focus on a subclass of routing strategies for which the λ_i's are discontinuous.

If a solution to equation (2.2.6) exists, it will be continuous in time. Since μ_i is a Borel measurable function of a continuous function of time, the integral $\int_{t_1}^{t_2} \mu_i(t)\,dt$ exists for any t_1, t_2 such that no failure or repair occurs in $[t_1, t_2]$. This implies that

$$\int_{t_1}^{t_2} p_i{}^* \frac{\mu_i(\underline{x}(t), \underline{\alpha}(t))}{\mu_i{}^*}\,dt \equiv \int_{t_1}^{t_2} p_i\,dt \tag{2.2.9}$$

exists. Therefore, the integral appearing in the definition of the process $F_i(t)$, in equation (2.1.9) is a meaningful quantity. The same argument is valid for the integral appearing in the definition of $F_0(t)$.

If these requirements are met, then it is easy to see that we can piece together parts of the process and obtain a process unambiguously defined for all times.

In order to summarize this discussion we recapitulate our assumptions:

a) λ_i and μ_i are Borel measurable functions on the state space.

b) The deterministic differential equation (2.2.6) has a unique piecewise continuously differentiable solution, for any fixed $\underline{\alpha}$, and any initial $\underline{x}$.

Then, the Markov process described by equations (2.1.8) and (2.1.15) is well-defined.

In this section we have made a distinction between physical and mathematical constraints. Physical constraints are those that follow from the structure of the system. Mathematical constraints are introduced for different reasons. If they are violated, the functions λ_i and μ_i would be so ill-behaved that they can have no interpretation in terms of a physical system. Moreover, these constraints are indispensable in the proof of any conceivable theorem and they are routinely assumed by researchers in the field of stochastic control.*(Boel and Varaiya, 1977; Rishel, 1975)*

2.3. Performance Criteria

We are interested in maximizing the long-run (infinite horizon) aaverage performance of the system. We let $k(\underline{x}, \underline{a})$ denote a function on the state space equal to the instantaneous benefit from the operation of the system when the current state is $(\underline{x}, \underline{a})$. The function to be maximized will then be:

$$g = \lim_{T \to \infty} E\left[\frac{1}{T}\int_0^T k(\underline{x}, \underline{a})\, dt\right] \tag{2.3.1}$$

There are many different possible choices of $k(\underline{x}, \underline{a})$ corresponding to different objectives and to different notions of good performance. In the present study we focus on maximization of average production rate. We also impose different weights on the outputs of the two downstream machines. This means that

$$k(\underline{x}, \underline{a}) = c_1\mu_1(\underline{x}, \underline{a}) + c_2\mu_2(\underline{x}, \underline{a}) \tag{2.3.2}$$

where c_1 and c_2 are constants reflecting the relative values of the outputs of machines M_1 and M_2.

In a communication network context one is usually interested in minimizing the average time that a message has to wait in a queue. In that case an appropriate performance functional would be

$$k(\underline{x}, \underline{a}) = -(x_1 + x_2) \tag{2.3.3}$$

Even in a manufacturing network, economic reasons might require that buffer levels be kept as low as possible. Then one can use a combination of the performance functionals

introduced above, like

$$k(\underline{x}, \underline{a}) = c_1\mu_1(\underline{x}, \underline{a}) + c_2\mu_2(\underline{x}, \underline{a}) - f(x_1, x_2) \tag{2.3.4}$$

where f is a positive function monotonically increasing in each variable. One could then relax the requiement that buffers have finite storage capacities but make sure that buffer levels do not become arbitrarily large by penalizing heavily large buffer levels. This can be done, for example, by letting $f(x_1, x_2) = x_1^2 + x_2^2$. However, this does not in general, make the problem easier to solve. Unbounded state spaces give rise to mathematical and computational difficulties, except for a few rare cases where a simple closed form solution exists.

In the sequel we derive some theoretical results valid for a general performance criterion $k(\underline{x}, \underline{a})$ and then proceed to obtain some more specific results for the special case where average production rate is maximized.

2.3.1. Assumptions on the Performance Criterion.

We assume that:

$k(\underline{x}, \underline{a})$ is a measurable function of the state space.

$k(\underline{x}, \underline{a})$ is bounded uniformly for all admissible control laws.

$$\lim_{t \to 0+} E[k(\underline{x}(t), \underline{a}(t))] = k(\underline{x}(0), \underline{a}(0)) \tag{2.3.5}$$

for all admissible control laws.

The last condition can be satisfied by requiring that $k(\underline{x}(t), \underline{\alpha}(t))$ has (almost surely) right-continuous sample functions. Equation (2.3.5) then follows from the boundedness of $k(\underline{x}, \underline{\alpha})$ and the dominated convergence theorem. (Appendix I)

2.3.2. Special Assumptions on Routing Strategies when Production Rate is Maximized.

If we seek to maximize average production rate, there are some reasonable requirements that may be imposed on the way that the system will be operated. These requirements are most often imposed in practice and some heuristic reasoning indicates that, under certain conditions, the production rate of the system is not worsened. Moreover, these assumptions reduce the size of the problem and make it easier to solve. The philosophy underlying these requirements is the following: If at any time we can increase the flow through some machine without decreasing the flow through another machine, we will do so. This statement can be translated mathematically to the following set of constraints:

$$\text{If } \alpha_i = 1 \text{ and } x_i \neq 0 \text{ then } \quad \mu_i(\underline{x}, \underline{\alpha}) = \mu_i^* \quad i = 1, 2 \tag{2.3.6}$$

$$\text{If } \alpha_i = 1 \text{ and } x_i = 0 \text{ then } \quad \mu_i(\underline{x}, \underline{\alpha}) = \min(\lambda_i(\underline{x}, \underline{\alpha}), \mu_i^*) \tag{2.3.7}$$

(Downstream machines are used as much as possible.)

$$\text{If } \alpha_0 = 1,\ x_1 \neq N_1 \text{ and } x_2 \neq N_2 \text{ then } \quad \lambda(\underline{x}, \underline{\alpha}) = \lambda^* \tag{2.3.8}$$

The above two constraints will be referred to as constraint (CT1). We should note that they are reasonable constraints not only for manufacturing networks maximizing production rate but for many other real world networks. It seems that (CT1) would be inappropriate only in the cases where heavy instantaneous utilization of the machines or large buffer levels are penalized by the cost functional $k(\underline{x}, \underline{\alpha})$.

We should point out that the assumptions introduced in this section are of a different nature than those introduced in earlier sections. Formally, they are just restrictions on the set of admissible control laws. Essentially, however, they are guesses about the optimal way that the system should be operated.

Constraints (CT1) are always assumed in the next chapters, unless the contrary is explicitly stated. We now introduce four additional assumptions that will be referred to as (CT2) and will only be assumed in Chapter 6.

$$\text{If } \alpha_0 = 1 \text{ and } (x_1, x_2) \neq (N_1, N_2) \text{ then } \quad \lambda(\underline{x}, \underline{\alpha}) = \lambda^* \tag{2.3.9}$$

$$\text{If } \alpha_0 = 1 \text{ and } (x_1, x_2) = (N_1, N_2) \text{ then } \quad \lambda(\underline{x}, \underline{\alpha}) = \min(\lambda^*, \alpha_1 \mu_1{}^* + \alpha_2 \mu_2{}^*) \tag{2.3.10}$$

$$\text{If } \alpha_0 = 1,\ x_1 = 0,\ x_2 > 0 \text{ then } \quad \lambda_1(\underline{x}, \underline{\alpha}) \geq \min(\lambda^*, \alpha_1 \mu_1{}^*) \tag{2.3.11}$$

$$\text{If } \alpha_0 = 1,\ x_1 > 0,\ x_2 = 0 \text{ then } \quad \lambda_2(\underline{x}, \underline{\alpha}) \geq \min(\lambda^*, \alpha_2 \mu_2{}^*) \tag{2.3.12}$$

If buffer B_1 is empty and buffer B_2 is not, machine M_2 may work at full capacity, irrespectively of λ_2. So, we require that we feed machine M_1 at least as much as it needs ($\alpha_1\mu_1^*$) unless this is not possible, i.e. if $\alpha_1\mu_1^* > \lambda^*$.

Constraints (CT2) are not always appropriate. For example, if the output of machine M_1 has almost no value compared with the output of machine M_2 we wouldn't have any reason to feed machine M_1, unless B_1 is almost full. However, if the output of each machine has the same value ($c_1 = c_2$) then, we conjecture that (CT2) does not deteriorate the achievable performance of the system.

CHAPTER 3: A POLICY EVALUATION METHODOLOGY

As discussed in earlier sections, for any choice of an admissible routing strategy we obtain a well–defined Markov process. Throughout this chapter we assume that an arbitrary routing strategy has been fixed.

In section 3.1 we establish that the Markov process corresponding to a fixed control law is semi–stationary. This is essentially a property closely related (but weaker than) ergodicity which will enable us to show, in section 3.2, that the performance functional is well–defined and independent of the initial state of the system. In section 3.3 we discuss the possible methodologies for evaluating the performance functional. Having decided that dynamic programming is the most appropriate methodology we define (in section 3.4) cost–to–go functions and derive some of their properties. In section 3.5 we derive, in a very general setting, a set of partial differential equations that have to be solved in order to obtain the value of the performance functional and the cost–to–go functions. We prove the uniqueness of the solution of these equations and comment on possible numerical methods of solution.

3.1. A Regeneration Property of the System.

Recall that the performance functional was defined by:

$$g = \lim_{T\to\infty} \frac{1}{T} E\left[\int_0^T k(s(\tau))d\tau\right] \tag{3.1.1}$$

The dynamic routing problem would be meaningless if g were not well–defined. It is also

desirable to have g independent of the initial state because otherwise we would have to partition the state space into a class of noncommunicating subsets.

The existence of g, and its independence from the initial state, can be established by showing the existence of a unique steady-state probability distribution, independent of the initial state. Because of the special structure of the system we are dealing with, we follow a simpler approach. We show some properties of the sample paths of the Markov process from which the existence of g will follow.

We show that there exists (at least) one point of the state space which is visited an infinite number of times by almost all sample functions and such that the Markov process starts afresh each time this point is visited.

It is not difficult to find one such point. Consider the point $s_0 = ((0,0),(0,1,1))$ where the upstream machine is down, the downstream machines are up and both buffers empty. For any initial state there is some probability that the machine configuration eventually becomes $(0,1,1)$ and that this configuration does not change until both buffers are empty. This is the main line of argument in Theorem 3.1 below.

Let s be any point of the state space S. Let $T_n(s)$ be the n-th time that the point s is reached. (If s is the initial state, then we let $T_1(s) = 0$. If the point s is not reached n times during $[0,\infty)$ we let $T_n(s) = \infty$).

Theorem 3.1: There exists a state $s_0 \in S$ and a constant M such that for every initial state and any admissible control law staisfying CT1:

a) $T_n(s_0)$ is a Markov time (see Appendix I).

$$b) \qquad E[T_n(s_0)] < \infty \tag{3.12}$$

$$c) \quad E[T_n(s_0) - T_{n-1}(s_0)] < M \tag{3.1.3}$$

$$d) \quad \lim_{n\to\infty} T_n(s_0) \quad \text{almost surely} \tag{3.1.4}$$

Proof: We need the following lemma:

Lemma 3.1: Let

$$s_0 == ((0,0),(0,1,1)) \tag{3.1.5}$$

$$T = 3\max\left(\frac{N_1}{\mu_1^*}, \frac{N_2}{\mu_2^*}\right) \tag{3.1.6}$$

(Here, $T/3$ is the time needed to empty full buffers). If $r_0 > 0$, then there exists an $\epsilon > 0$ such that for any initial state and any control law satisfying CT1,

$$Pr\big(s(T) = s_0,\ \exists t \in [0,T) s(t) \neq s_0\big) > \epsilon \tag{3.1.7}$$

Proof: See Appendix II.

We now define a sequence of random variables $\{R_n\}$ which may take as values only integer multiples of T:

$$R_0 = 0 \tag{3.1.8}$$

$$R_n = \min\{kT\colon k \in Z^+, kT > R_{n-1}, s(kT) = s_0, s(t) \neq s_0 \text{ for some } t \in [(k-1)T, kT]\} \tag{3.1.9}$$

As defined by equation (3.1.9), R_n is the n-th time such that: a) It is a multiple of T. b) The state of the system is s_0. c) The state of the system was not identically s_0 during the time interval $[R_{n-1}, R_n]$.

From the definition of R_n, we have

$$R_n \geq T_n(s_0) \tag{3.1.10}$$

Therefore, in order to prove (3.1.2) it is sufficient to show $E[R_n] < \infty$.

Lemma 3.1 implies that

$$Pr(R_n - R_{n-1} = T) > \epsilon \tag{3.1.11}$$

$$Pr(R_n - R_{n-1} > T) < (1 - \epsilon) \tag{3.1.12}$$

Suppose that we have established that

$$Pr(R_n - R_{n-1} > kT) < (i - \epsilon)^k \tag{3.1.13}$$

Clearly,

$$Pr(R_n - R_{n-1} > (k+1)T) = Pr(R_n - R_{n-1} > (k+1)T \mid R_n - R_{n-1} > kT)\, Pr(R_n - R_{n-1} > kT) \tag{3.1.14}$$

The first factor in the right hand side of equation (3.1.14) is less than $(1-\epsilon)$, by Lemma 3.1; the second factor is less than $(1-\epsilon)^k$ by the hypothesis (3.1.13). Therefore,

$$Pr(R_n - R_{n-1} > (k+1)T) < (1-\epsilon)^{k+1} \tag{3.1.15}$$

This is an inductive argument which proves hat equation (3.1.13) holds for all k. Moreover, equation (3.1.14) implies that

$$Pr(R_n - R_{n-1} = (k+1)T) < (1-\epsilon)^k \tag{3.1.16}$$

and

$$E[R_n - R_{n-1}] < \sum_{k=1}^{\infty} Tk(1-\epsilon)^{k-1} \tag{3.1.17}$$

The above series converges; consequently, there exists some M such that

$$E[R_n - R_{n-1}] < M \tag{3.1.18}$$

Since ϵ is independent of the control law, M can be chosen to be independent of the control law.

Clearly,

$$E[R_n] = E[R_0] + E[R_1 - R_0] + \cdots + E[R_n - R_{n-1}] \tag{3.1.19}$$

which implies that

$$E[R_n] < \infty \tag{3.1.20}$$

and by virtue of (3.1.10)

$$E[T_n(s_0)] < \infty \tag{3.1.21}$$

which proves (3.1.2).

The definitions of T_n and R_n imply that

$$E[T_n - T_{n-1}] \leq E[R_n - R_{n-1}] \tag{3.1.22}$$

Equations (3.1.18) and (3.1.22) establish (3.1.3).

The fact that $T_n(s_0)$ is a Markov time is well-known. (Meyer,1966)

Finally, if $\lim_{n\to\infty} T_n(s_0)$ were finite, the process $\underline{a}(t)$ would have an infinite number of jumps during a finite interval. Since the processes R_i and F_i (as defined in section 2.1) have bounded jumping rates, this is an event of zero probability and this proves statement d). ∎

Lemma 3.1 has been proved for the simple three machine network considered in this thesis. However, the same argument can be extended to the case of a general network under the following condition: "If all machines receiving input from outside are down and all downstream machines are up, then the system is operated in such a way that all buffers become empty, with finite probability, before a change in the machine configuration occurs".

Fortunately, these conditions are very reasonable and they should be certainly be satisfied an any real system. For the simple system considered in this thesis they are implied by CT1.

Note that the entire discussion has been based on the finiteness of buffers. If buffers are allowed to be infinite some stability conditions would be required to guarantee that buffer levels will not grow unbounded as $t \to \infty$.

An additional comment is in order here. The proof of Lemma 3.1 is based on the assumption that the lead machine has positive failure rate. Theorem 3.1 is however true even if the lead machine is perfectly reliable except that the recurrent point s_0 should be chosen differently.

From now on we will write T_n instead of $T_n(s_0)$.

We have shown that there exists a certain state which is positive recurrent. We now present a theorem of the same form that says that under any control law, every open ball is positive recurrent. We define the open ball of radius ϵ about a point $(\underline{x}^*, \underline{\alpha}^*)$ as the set of all points $(\underline{x}, \underline{\alpha}^*)$ such that $\|\underline{x}^* - \underline{x}\| < \epsilon$.

Theorem 3.2: For every $\epsilon > 0$, there exists a $\delta > 0$ and a deterministic time $T > 0$ such that for any $s^* \in S$ and for any admissible control law satisfying CT3 below

$$Pr(s(T) \in B_\epsilon(s^*)) > \delta \tag{3.1.23}$$

CT3: (i) There exists a deterministic time T^* such that if both downstream machines are down, the upstream machine is up, and no failure or repair occurs until time T^*, then both buffers are full at time T^*.

(ii) If the upstream machine is down, a downstream machine M_i is up and $x_i \neq 0$, then $\mu_i(\underline{x}, \underline{\alpha}) = \mu_i^*$.

Note that condition CT3 is implied by CT1 and equation (2.3.9).

This theorem can be proved in a way which is very similar to the proof of Lemma 3.1 and Theorem 3.1. We need again to identify, for every initial state, a set of paths that has probability greater than δ and which leads to the ball $B_\epsilon(s^*)$ at time T.

From an abstract point of view the Theorem 3.2 means that the Borel measure on the state space is absolutely continuous with respect to the invariant measure of the process (if it exists), *uniformly* for all admissible control laws.

3.2. Existence of the Performance Functional and Independence from the Initial State.

Theorem 3.1 essentially establishes the fact that the infinite time horizon can be broken down into a sequence of almost surely finite intervals $[T_n, T_{n-1}]$ such that the Markov process starts afresh at each time T_n. In fact the strong Markov property (Appendix I) implies that the sequence of random variables $\{s(T_n + t)\}, n = 1, 2, \cdots$ is a stationary sequence of random variables for any fixed $t > 0$. Such processes are called semi–stationary.

Semi–stationarity can be exploited to establish ergodicity properties of the system and the existence of long–run average quantities (Stidham, 1977; Serfozo and Stidham, 1978). We take such an approach in the following fashion: whenever we want to take a limit as time goes to infinity, we take this limit along the sequence of Markov times T_n and then exploit the fact that the state of the system at time T_n is known to be s_0.

This approach enables us to prove that the performance functional g is well–defined and is independent of the initial state.

Recall that we assumed that the cost function $k(\underline{x}, \underline{a})$ is a Borel measurable, bounded function on the state space. We let M be such that $|k(s)| \leq M$, $\forall s \in S$.

This measurability assumption, together with the right continuity of the process $s(t)$ imply that for any random variables U, V such that $V \geq U$ almost surely, the integral $\int_U^V k(s(t))dt$, interpreted as an integral of sample functions is well-defined.

Consider the sequence of Markov processes $s_n(t) \equiv s(T_n + t), t \in [0, T_{n+1} - T_n]$. By the strong Markov property, all processes in the sequence have an identical behavior from a probabilistic point of view. We therefore expect that any quantity related to the average behavior of the initial process may be deduced by considering only one of the processes in this sequence, $s_1(t)$ for example. The following theorem shows that the average performance of the system may be evaluated by averaging only the process $s_1(t)$. One might suspect that one of the following two equations would be true:

$$g = E\left[\frac{\int_{T_1}^{T_2} k(s(t))dt}{T_2 - T_1}\right] = E\left[\frac{\int_0^{T_2 - T_1} k(s_1(t))dt}{T_2 - T_1}\right] \tag{3.2.1}$$

(which is the average performance of one piece of the process), or

$$g = \frac{E\left[\int_{T_1}^{T_2} k(s(t))dt\right]}{E[T_2 - T_1]} = \frac{E\left[\int_0^{T_2 - T_1} k(s_1(t))dt\right]}{E[T_2 - T_1]} \tag{3.2.2}$$

Theorem 3.3 says that equation (3.2.2) is true (and therefore (3.2.1) is false).

Theorem 3.3: For any initial state and any control law such that the conclusions of Theorem 3.1 hold,

$$\lim_{n\to\infty} E\left[\frac{1}{T_n}\int_0^{T_n} k(s(\tau))d\tau\right] = \lim_{t\to\infty} E\left[\frac{1}{t}\int_0^t k(s(\tau))d\tau\right] = \frac{E\left[\int_{T_n}^{T_{n+1}} k(s(\tau))d\tau\right]}{E[T_{n+1} - T_n]} = g \quad n = 1, 2, \cdots \tag{3.2.3}$$

(Note that the second equality implies that g is independent of the initial state).

Proof: We define a random variable $T(t)$ by

$$T(t) = \sup\{T_n : T_n \leq t\} \tag{3.2.4}$$

Here, $T(t)$ is the last time before time t state s_0 was reached.

Lemma 3.2: There exists $N : E[t - T(t)] < N$, for all initial states, for all control laws.

Proof: See Appendix II.

Lemma 3.3: $\lim_{t\to\infty} T(t) = \infty$, almost surely.

Proof: Follows trivially from the fact that $\lim_{n\to\infty} T_n = \infty$, almost surely.

Proof of the theorem: Let

$$W_n = T_n - T_{n-1} \tag{3.2.5}$$

$$U_n = \int_{T_{n-1}}^{T_n} k(s(\tau))d\tau \tag{3.2.6}$$

Semi-stationarity implies that the random vectors $(W_n, U_n), n = 2, 3, \ldots$ are independent,identically distributed. Then, an ergodic theorem, (Loeve, 1967, Vol. 2) implies that

$$\lim_{n\to\infty} \frac{\sum_{k=2}^{n} U_k}{\sum_{k=2}^{n} W_k} = \frac{E[U_m]}{E[W_m]} \quad \text{almost surely} \quad (m = 2, 3, \ldots) \tag{3.2.7}$$

$$lim_{n\to\infty} \frac{\sum_{k=2}^{n} U_k}{\sum_{k=2}^{n} W_k} = \lim_{n\to\infty} \frac{\int_{T_1}^{T_n} k(s(\tau))d\tau}{T_n - T_1} = \tag{3.2.7}$$

$$\lim_{n\to\infty}\left[\frac{1}{T_n}\int_0^{T_n}k(s)d\tau\right]+\lim_{n\to\infty}\left[\left(\frac{1}{T_n-T_1}-\frac{1}{T_n}\right)\int_0^{T_n}k(s)d\tau\right]-\lim_{n\to\infty}\frac{1}{T_n-T_1}\int_0^{T_1}k(s)d\tau \tag{3.2.8}$$

We claim that the second and third summands converge almost surely to zero.

Indeed:

$$\left|\left(\frac{1}{T_n-T_1}-\frac{1}{T_n}\right)\int_0^{T_n}k(s)d\tau\right|\leq M\frac{T_n-(T_n-T_1)}{T_n(T_n-T_1)T_n}=M\frac{T_1}{T_n-T_1} \tag{3.2.9}$$

Now, $T_1<\infty$ (almost surely) and $\lim_{n\to\infty}T_n=\infty$ (almost surely). Therefore, $MT_1/(T_n-T_1)$ converges to zero almost surely. Also,

$$\left|\frac{1}{T_n-T_1}\int_0^{T_1}k(s)d\tau\right|\leq\frac{MT_1}{T_n-T_1}\to 0,\quad a.s. \tag{3.2.10}$$

Therefore,

$$\lim_{n\to\infty}\frac{1}{T_n}\int_0^{T_n}k(s)d\tau=\frac{E[U_n]}{E[W_n]},\quad a.s. \tag{3.2.11}$$

The sequence of random variables $(1/T_n)\int_0^{T_n}k(s)d\tau$ is uniformly bounded by M. Therefore, by the dominated convergence theorem (Appendix I):

$$\lim_{n\to\infty}E\left[\frac{1}{T_n}\int_0^{T_n}k(s)d\tau\right]=\frac{E[U_n]}{E[W_n]} \tag{3.2.12}$$

We now need to show the first equality in the statement of the theorem.

$$\left|E\left[\left(\frac{1}{t}-\frac{1}{T(t)}\right)\int_0^{T(t)}k(s)d\tau\right]\right|\leq\left|E\left[\left(\frac{1}{t}-\frac{1}{T(t)}\right)T(t)M\right]\right|=M\frac{1}{t}E[t-T(t)]\leq\frac{MN}{t}\to_{t\to\infty}0 \tag{3.2.13}$$

Also,

$$\left|E\left[\frac{1}{t}\int_{T(t)}^{t} k(s)d\tau\right]\right| \leq \frac{M}{t}E[t - T(t)] \leq \frac{MN}{t} \to_{t\to\infty} 0 \tag{3.2.14}$$

Using (3.2.13) and (3.2.14)

$$\lim_{t\to\infty} E\left[\frac{1}{t}\int_{0}^{t} k(s)d\tau\right] = \lim_{t\to\infty} E\left[\frac{1}{T(t)}\int_{0}^{T(t)} k(s)d\tau\right] = \lim_{n\to\infty} E\left[\frac{1}{T_n}\int_{0}^{T_n} k(s)d\tau\right] \tag{3.2.15}$$

∎

This theorem establishes that the expected average cost (or benefit) is a well-defined quantity and is equal to $(EU_n)/(EW_n)$. In the next section we develop a method of calculating this average cost without having to obtain expressions for EU_n and EW_n.

3.3. Discussion of Possible Methodologies.

Theorem 3.3 of the last section indicates that given a control law and the corresponding Markov process we can evaluate the performance of the system by obtaining an expression for EU_n and EW_n. Another possible method is to establish the existence of the steady-state probability measure of the process $\mathcal{P}(s)$, evaluate it and then calculate the performance of the system by using the formula

$$\lim_{T\to\infty} E\frac{1}{t}\int_{0}^{T} k(s)d\tau = \int k(s)d\mathcal{P}(s) \tag{3.3.1}$$

the last integral being a Lebesgue integral over the state space. Both of these approaches work in principle. In fact, the second method has been often used to calculate quantities

such as the average throughput of the system or the average waiting time in a queue (Gershwin and Schick,1979).

The drawback of both methods is that they are not very useful when one is interested in the optimization problem. Suppose that the performance of a system under a specific control law has been evaluated using the steady-state distribution. Then, there is no way to check whether this specific control law has optimal performance or not. (The only exception is when the steady-state distribution can be obtained parametrically, the parameter being the routing strategy, but this is rarely the case).

In fact, the only available optimality conditions for stochastic optimal control problems are those given by Bellman's Optimality Principle of dynamic programming (Bellman and Dreyfus, 1962). In order to use these optimality conditions we need an approach which is based on cost-to-go functions. Such an approach is always followed in the case of finite-state, discrete time, Markovian decision problems. The evaluation of the cost-to-go functions, for a discrete problem, involves the solution of a system of algebraic equations (Howard, 1960). Many simple optimization algorithms based on cost-to-go functions exist. (For example, the Howard algorithm, the successive approximation algorithm, etc. (Bertsekas,1976)). In the continuous case, the evaluation of cost-to-go functions involves the solution of a system of partial differential equations which (except for a driving term) is dual to the system of partial differential equations that has to be solved in order to evaluate the steady-state distribution. Therefore, the dynamic programming approach does not have any increased computational requirements over the other approach. The only deficiency is that cost-to-go functions contain information relevant to the specific cost functional considered, whereas steady-state probability distributions contain all in-

formation relevant to the system.

3.4. Definition of Value (Cost–to–go) Functions.

We are still assuming that a routing strategy and a corresponding Markov process have been fixed. We have established that the average reward g is well-defined (Theorem 3.3). As is always done in infinite time horizon, average reward stochastic control problems we assume that

$$E\left[\int_0^t k(s(\tau))d\tau \mid s(0) = s^*\right] \approx V(s^*) + gt \tag{3.4.1}$$

for some function V of the state space. This is meant to be valid in some sense to be defined below as $t \to \infty$. The meaning of the function V is clear. If s_1 and s_2 are two points of the state space, $V(s_1) - V(s_2)$ is the expected increase in the reward $\int_0^\infty k(s)d\tau$ (not in the average reward!) if the initial conditions are moved from s_2 to s_1. This difference is equal to the maximum amount we would be willing to pay in order to have the initial conditions moved from s_2 to s_1. We now proceed with a rigorous definition and treatment of the function $V(s)$ which will be called "cost–to–go" or "value" function.

Definition: We define $V{:}S \mapsto R$ by

$$V(s) = E\left[\int_0^{T_1} (k(s(t)) - g)dt \mid s(0) = s\right] \tag{3.4.2}$$

where T_1 is again the first time that state s_0 is reached.

Proposition 3.1: V is uniformly bounded (for all control laws) and measurable.

Proof: By assumption, $|k(s)| \leq M, \forall s$; by Theorem 3.1 we have $E[T_1] \leq N$ for some N; from the definition of g (equation 3.1.1), $|g| \leq M$. Therefore,

$$|V(s)| = \left| E\left[\int_0^T (k(s(t)) - g)dt \mid s(0) = s \right] \right| \leq 2MN \tag{3.4.3}$$

Since M and N are independent of which control law is used, V is uniformly bounded for all control laws.

The fact that V is measurable follows directly from the rigorous definition of conditional expectations (Wong, 1971). ∎

Proposition 3.2:

$$V(s) = E\left[\int_0^{T_n} (k(s(t)) - g)dt \mid s(0) = s \right] = \lim_{n \to \infty} E\left[\int_0^{T_n} (k(s(t)) - g)dt \mid s(0) = s \right] \tag{3.4.4}$$

where T_n is as defined in section 3.1.

Proof: According to Theorem 3.3

$$g = \frac{E[\int_{T_n}^{T_n+1} k(s(t))dt]}{E[T_{n+1} - T_n]}, \quad n = 1, 2, \ldots \tag{3.4.5}$$

Therefore,

$$E\left[\int_{T_n}^{T_{n+1}} (k(s(t)) - g)\, dt \right] = E\left[\int_{T_n}^{T_{n+1}} k(s(t))\, dt \right] - E[T_{n+1} - T_n] \frac{E\left[\int_{T_n}^{T_{n+1}} k(s(t))\, dt \right]}{E[T_{n+1} - T_n]} = 0 \tag{3.4.6}$$

∎

We should note that Proposition 3.2 fails if we attempt to phrase it in a stronger way. The statements

$$V(s) = E\left[\lim_{n\to\infty} \int_0^{T_n} (k-g)\,dt \mid s(0) = 0\right] \tag{3.4.7}$$

$$V(s) = E\left[\lim_{n\to\infty} \int_0^{t} (k-g)\,d\tau \mid s(0) = s\right] \tag{3.4.8}$$

are false because the above limits of random variables do not exist.

Our definition of V implies that $V(s_0) = 0$. This is not essential. Any other function different from V by a constant would be equally suitable for the purposes of this study.

3.5. An Operator Equation for the Value Functions.

There are two reasons why value functions are of interest. First, they can be evaluated by solving a system of partial differential equations which yields at the same time the value of the performance criterion g. Second, there are very simple optimality conditions which are formulated in terms of value functions. In this section we explore the first issue and we defer optimality considerations to Chapter 4.

To every Markov process there corresponds an operator $\mathcal{L}$, on the set of functions of the state space, which characterizes the Markov process completely. For a deterministic process, $\mathcal{L}$ is equal to the time derivative operator. For a random process, $\mathcal{L}$ is still very much like a time derivative except that the uncertainty in the future is taken care by taking expectations. (See Appendix I)

As this operator summarizes the structure of the process, it appears in the differential equations that are satisfied by any of the interesting properties of the process (e.g. exit times, exit distributions etc.) *(Dynkin, 1965)*. This is also true for the value functions as the following theorem indicates.

Theorem 3.4: Let s(t) be a Markov process for which the conclusions of Theorems 3.1 and 3.3 hold and let $\mathcal{L}$ be the weak infinitesimal operator of the process. Let V be a function as defined by equation (3.4.2). Then V is in the domain of $\mathcal{L}$ and satisfies

$$\mathcal{L}V(s) = -k(s) + g \quad \forall s \in S \tag{3.5.1}$$

Proof: Fix the initial state $s(0)$ of the process. Let $T(t)$ be the first time after t that the recurrent state s_0 of Theorem 3.1 is reached. Then, by (3.4.2),

$$V(s(t)) = E\left[\int_t^{T(t)} (k(s(\tau)) - g)\, d\tau \mid s(t)\right] \tag{3.5.2}$$

$$E[V(s(t)) \mid s(0)] = E\left[E\left[\int_t^{T(t)} (k(s(\tau)) - g)\, d\tau \mid s(t)\right] \mid s(0)\right] \tag{3.5.3}$$

Using the properties of repeated conditional expectations *(Wong, 1971)* we obtain

$$E[V(s(t)) \mid s(0)] = E\left[\int_t^{T(t)} (k(s(\tau)) - g)\, d\tau \mid s(0)\right] \tag{3.5.4}$$

We also have, by Proposition 3.2

$$V(s(0)) = E\left[\int_0^{T(t)} (k(s(\tau)) - g)\, d\tau \mid s(0)\right] \tag{3.5.5}$$

By using (3.5.4) and (3.5.5) we obtain

$$\frac{E[V(s(t)) \mid s(0)] - V(s(0))}{t} = \frac{E\left[\int_0^t (k(s(\tau)) - g)\, d\tau \mid s(0)\right]}{t} \tag{3.5.6}$$

The assumptions of section (2.3.1) imply that

$$\lim_{t \to 0+} E[k(s(\tau)) - g \mid s(0)] = k(s(0)) - g \tag{3.5.7}$$

By the Fubini Theorem (Appendix I)

$$\lim_{t \to 0+} \frac{E[V(s(t)) \mid s(0)] - V(s(0))}{t} = \lim_{t \to 0+} \frac{\int_0^t \left(E[k(s(\tau)) \mid s(0)] - g\right) d\tau}{t} \tag{3.5.8}$$

Equations (3.5.7),(3.5.8) and the fundamental theorem of integral calculus imply

$$\lim_{t \to 0+} \frac{E[V(s(t)) \mid s(0)] - V(s(0))}{t} = k(s(0)) - g \tag{3.5.9}$$

Since $s(0)$ was arbitrarily chosen, the limit in (3.5.9) exists as a pointwise limit for all $s \in S$ and converges to the bounded and measurable function $k(s) - g$ of the state space. This fact together with the assumptions on $k(s)$ show that V belongs to the domain of the

weak infinitesimal operator $\mathcal{L}$ of the process (Appendix I) and satisfies equation (3.5.1).

■

3.6. Identification of the Infinitesimal Operator.

The operator equation in the last section is not very useful unless there is a way to identify and obtain explicitly the operator $\mathcal{L}$. Fortunately this is possible for the processes we are considering.

It is well known that for processes described in terms of stochastic differential equations, as is our case, the infinitesimal operator can be written down almost by inspection and is always a differential operator for those points of the state space at which the function to be operated satisfies several differentiability requirements. The usual procedures for obtaining $\mathcal{L}$ use either the Ito or the Kunita–Watanabe *(Meyer,1967)* change of variables formula. However, both of these formulas have been rigorously proved for twice continuously differentiable functions on the state space. In our case the expression for the weak infinitesimal operator of the process may be derived using a rigorous version of the approach followed in Section 3.7, but this is not done in this thesis.

We should point out that the interest in functions which are not continuously differentiable is not academic. There is no guarantee that the value functions arising from a discontinuous control law such as "send all flow to the buffer with the smallest level" are differentiable on the set $\{s \in S: a_0 = 1, x_1 = x_2\}$. The results of the next chapter indicate that we will be primarily interested in such discontinuous control laws.

We take the following approach for identifying the weak operator of the process. We

first state what the operator is when it is applied on twice continuously differentiable functions and then give rule as to how derivatives should be interpreted when we have a function with discontinuous derivatives. This approach leads, with some care, to correct results. A general and rigorous justification, however, falls outside the scope of this thesis. It will be supported, however, by an informal argument presented in the next section.

Let $\mathcal{L}$ be the weak infinitesimal operator of the process $s(t)$ obeying the stochastic differential equations (2.1.8) and (2.1.15) and let f be a continuously differentiable function on the state space which is in the domain of $\mathcal{L}$. Then, assuming that the instantaneous failure rates $p_i(\underline{x}(t), \underline{a}(t))$ are right continuous in t (which is our case because of the constraints CT1), we have *(Meyer, 1967)*

$$\mathcal{L}f = (\lambda_1 - \mu_1)\frac{\partial f}{\partial x_1} + (\lambda_2 - \mu_2)\frac{\partial f}{\partial x_2} + \sum_{i=0}^{2}(\langle 1 - a_i, f\rangle_i - f)p_i + \sum_{i=0}^{2}(\langle a_i, f\rangle_i - f)r_i \quad (3.6.1)$$

where,

$$\langle \phi(\underline{a}), \psi(\underline{a})\rangle_i \equiv \sum_{a_i=0}^{1} \phi(\underline{a})\psi(\underline{a}) \quad (3.6.2)$$

For example,

$$\begin{aligned}(\mathcal{L}f)(\underline{x},(1,1,1)) =&(\lambda_1 - \mu_1)\frac{\partial f}{\partial x_1}(\underline{x},(1,1,1)) + (\lambda_2 - \mu_2)\frac{\partial f}{\partial x_2}(\underline{x},(1,1,1)) + \\ &(f(\underline{x},(0,1,1)) - f(\underline{x},(1,1,1)))p_0 + (f(\underline{x},(1,0,1)) - f(\underline{x},(1,1,1)))p_1 \\ &+ (f(\underline{x},(1,1,0)) - f(\underline{x},(1,1,1)))p_2\end{aligned} \quad (3.6.3)$$

We conjecture that if f is not continuously differentiable, $\mathcal{L}f$ will be the same as in (3.6.2), except that all derivatives will be interpreted as "forward derivatives along the direction of the motion". Let us make this statement more precise.

We know that between machine transitions the system follows a deterministic path. Suppose that at time $t = 0$, $\underline{x}$ is equal to $(x_1(0), x_2(0))$ and until the first machine transiition the state evolution is given by the deterministic functions $(x_1(t), x_2(t))$ that satisfy equation (2.1.8).

If f is a continuously differentiable function, then, by the chain rule

$$\lim_{t\to 0+} \frac{f(x_1(t), x_2(t)) - f(x_1(0), x_2(0))}{t} = (\lambda_1 - \mu_1)\frac{\partial f}{\partial x_1} + (\lambda_2 - \mu_2)\frac{\partial f}{\partial x_2} \tag{3.6.4}$$

If f belongs to the domain of $\mathcal{L}$ but is not differentiable at some point, then we will interpret the right-hand side of (3.6.4) (which also appears in (3.6.1)) as meaning the expression in the left hand side of (3.6.4). In that case, equation 3.6.1 becomes:

$$\begin{aligned} \mathcal{L}f = \lim_{t\to 0+} \frac{f(x_1(t), x_2(t)) - f(x_1(0), x_2(0))}{t} + \\ \sum_{i=0}^{2}(\langle 1 - a_i, f\rangle_i - f)p_i + \sum_{i=0}^{2}(\langle a_i, f\rangle_i - f)r_i \end{aligned} \tag{3.6.5}$$

where $(x_1(t), x_2(t))$ is the deterministic path followed by the system if no machine transition occurs in the time-interval $[0, t]$.

3.7. Informal Derivation of the Differential Equations for the Value Functions.

In this section we derive again the differential equations for the value functions given by equations 3.5.1 and 3.6.1 using an informal argument. This derivation is useful in understanding what these differential equations mean as well as the way that one-sided derivatives arise.

For notational simplicity, let $V_{\alpha_0\alpha_1\alpha_2}(\underline{x}) = V(\underline{x}, \underline{\alpha})$. Consider some point $(x_1, x_2) \in [0, N_1] \times [0, N_2]$. We derive the differential equation for $V_{111}(\underline{x})$. The equations for different values of $\underline{\alpha}$ can be derived in an almost identical way.

By the definition of V:

$$V_{111}(x_1, x_2) + gT_1 = E\left[\int_0^{T_1} k(s(\tau))\, d\tau \mid s(0) = ((x_1, x_2), (1, 1, 1))\right] =$$

$$E\left[\int_0^{\Delta} k(s(\tau))\, d\tau \mid s(0) = ((x_1, x_2), (1, 1, 1))\right] + E\left[\int_{\Delta}^{T_1} k(s(\tau))\, d\tau \mid s(0) = ((x_1, x_2), (1, 1, 1))\right] \approx$$

(when Δ is small)

$$\Delta k((x_1, x_2), (1, 1, 1)) + (1 - p_0\Delta - p_1\Delta - p_2\Delta)V_{111}(x_1 + (\lambda_1 - \mu_1)\Delta, x_2 + (\lambda_2 - \mu_2)\Delta) + p_0\Delta V_{011}(x_1, x_2) + p_1\Delta V_{101}(x_1, x_2) + p_2\Delta V_{110}(x_1, x_2) + g(T_1 - \Delta) + O(\Delta^2) \tag{3.7.1}$$

where $O(\Delta^2)$ is such that $\lim_{\Delta \to 0}(1/\Delta)O(\Delta^2) = 0$. Then we substract gT_1 from both sides, divide by Δ and obtain:

$$\frac{V_{111}(x_1, x_2) - V_{111}(x_1 + (\lambda_1 - \mu_1)\Delta, x_2 + (\lambda_2 - \mu_2)\Delta)}{\Delta} \approx$$

$$p_0V_{011}(x_1, x_2) + p_1V_{101}(x_1, x_2) + p_2V_{110}(x_1, x_2) - (p_0 + p_1 + p_2)V_{111}(x_1, x_2) + k((x_1, x_2), (1, 1, 1)) - g$$

Now, let $\Delta \to 0$ to get:

$$-(\lambda_1 - \mu_1)\frac{\partial V_{111}}{\partial x_1} - (\lambda_2 - \mu_2)\frac{\partial V_{111}}{\partial x_2} = p_0V_{011} + p_1V_{101} + p_2V_{110} - (p_0 + p_1 + p_2)V_{111} + k - g \tag{3.7.2}$$

We repeat again that $\partial V_{111}/\partial x_1$ and $\partial V_{111}/\partial x_2$ need not exist. However,

$$\lim_{\Delta \to 0^+} \frac{V_{111}(x_1, x_2) - V_{111}(x_1 + (\lambda_1 - \mu_1)\Delta, x_2 + (\lambda_2 - \mu_2)\Delta)}{\Delta} \tag{3.7.3}$$

exists and this is how $-(\lambda_1 - \mu_1)(\partial V_{111}/\partial x_1) - (\lambda_2 - \mu_2)(\partial V_{111}/\partial x_2)$ should be interpreted. We note that this limit is just a one sided derivative along the direction of motion of the system, exactly as we said in Section 3.6.

3.8. Uniqueness of Solution of the Equation for the Value Function.

We have shown that the value function V as defined in section 3.2 satisfies the operator equation

$$\mathcal{L}V = -k + g \tag{3.8.1}$$

Suppose that a function U is also in the domain of the weak operator $\mathcal{L}$ and satisfies

$$\mathcal{L}U = -k + g \tag{3.8.2}$$

Then, by Dynkin's theorem (Appendix I),

$$E[U(s(T_1))] - U(s) = E\left[\int_0^{T_1} -[k(s(\tau)) - g]\, d\tau \,|\, s(0) = s\right] = -V(s) \tag{3.8.3}$$

where T_1 is the first passage time frorm state s_0. Therefore, $s(T_1) = s_0$ is a fixed unique state and we conclude that

$$U(s) = V(s) + c \tag{3.8.4}$$

for some constant c. This shows that the solution of equation 3.8.1 is unique within the domain of the weak infinitesimal operator.

3.9. Comments on the Solution of the Differential Equations.

We note that for any admissible control law, which is not quasi-static, equation 3.6.1 is a set of coupled linear partial differential equations with non-constant coefficients. (In particular, λ_1, λ_2 vary with x_1, x_2.) Because coefficients are not constant, separation of variables cannot be used as in classical problems involving PDE's. Even if we consider a control law which is characterized by a set of dividing lines (See Chapter 4) this difficulty persists. We may write V as a superposition of eigenfunctions, for each side of the line, but the two sets of eigenfunctions obtained at each side are not the same. These considerations render an analytical solution by this technique unlikely.

One has to resort to numerical techniques. Any such technique will involve a discretizaton of the state space and of the equations and we end up with a system of linear algebraic equations to be solved. The simplest discretization would be to substitute the state space with a mesh of points and discretize all derivatives in a standard way.

The structure of the system of algebraic equations so obtained is similar to the structure of the system of equations that would have to be solved if we were dealing with a finite state space Markovian decision problem (Hahne, 1980). However, the numerical results of Hahne indicate that the variation of value functions is very slow at some regions of the state space (for example far away from the origin). We may then discretize the continuous equations using a less dense mesh of points in those regions. This approach

may lead to significant computational savings over the discrete formulation. However, the computational complexity of the problem will still be of the same order of magnitude.

CHAPTER 4: THE OPTIMIZATION PROBLEM.

In Chapter 3 we had assumed that a control law had been fixed; then, the performance functional and the value functions associated with the specific control law were examined. In this Chapter, we drop this assumption and we compare different control laws with the ultimate aim of characterizing and finding an optimal admissible control law.

The set of admissible control laws does not have enough of a topological or analytical structure to ascertain whether an optimal control law exists with respect to a given performance functional. Nevertheless, we derive necessary and sufficient conditions of optimality. From a conceptual point of view, these conditions are just particular ways of interpreting the Optimality Principle of Dynamic Programming and they are, invariably, phrased in terms of value functions. (Sections 4.1, 4.2.)

These optimality conditions lead to intuitively appealing conclusions. They imply the following property of an optimal control law: the state space is divided into three regions. In the first one, all material is routed towards buffer B_1; in the second region, all material is routed towards buffer B_2. In the third region either the routing decision has no effect on performance or there is no freedom in making the decision. Sections 4.4 to 4.7 deal with the complete interpretation of the optimality conditions, with various technical problems that arise and a description of some particularly interesting sets of admissible control laws.

Sections 4.8 and 4.9 contain an iterative algorithm for obtaining an optimal control law, together with a proof that it converges. Finally, Section 4.10 contains a brief discussion of the mathematical difficulties encountered in this chapter.

4.0. Assumptions and Notation for Reward Fuctions.

In Chapter 3 we focused on a single control law and therefore $k(s)$ was only considered as a function of the state and (indirectly) as a function of time.

In general, k depends both on the state and the control. The dependence on the control may be viewed in two ways:

(i) Let every admissible control law belong to some abstract space U of control laws and let a reward function $k^u(s)$ correspond to each $u \in U$.

(ii) Let the reward function $k(s)$ depend explicitly on the state and on the value of the control currently applied. Then we may write

$$k = k(\underline{x}, \underline{a}, \lambda_1, \lambda_2, \mu_1, \mu_2)$$

For example, under production rate maximization

$$k(\underline{x}, \underline{a}, \lambda_1, \lambda_2, \mu_1, \mu_2) = \mu_1 + \mu_2$$

The first approach will be more convenient to use in most proofs but not always.

For the rest of this thesis we assume that $k(\underline{x}, \underline{a}, \lambda_1, \lambda_2, \mu_1, \mu_2)$ is a continuous function of each of its arguments.

4.1. Informal Derivation of Optimality Conditions.

Let $k^u(s)$ be the reward function when control law u is applied. In this section we examine what the optimality conditions would be in the interior and in a region in which

$k^u(s)$, the failure probabilities, and the repair probabilities are the same for all control laws. This is done only with the purpose of illustrating in a simplified manner the meaning of the optimality conditions to be derived in the next secttions. One may construct a problem for which $k^u(s)$ depends on u everywhere. Let, for example,

$$k_u(\underline{x}, \underline{a}) = \lambda_1^u(\underline{x}, \underline{a})\, \lambda_2^u(\underline{x}, \underline{a}). \tag{4.1.1}$$

Even in the case where the production rate is maximized, there may be regions in which $k^u(s)$ is not independent of u. For example,

$$\lambda^* < \mu_1^* + \mu_2^* \quad \Rightarrow \quad k^u((0,0),(1,1,1)) = c_1\lambda_1^u + c_2\lambda_2^u \tag{4.1.2}$$

for some λ_1^u, λ_2^u depending on u. However, our assumption is valid in the interior of the state space when production rate is maximized and constraints CT1 hold. This is because

$$k^u(\underline{x}, \underline{a}) = c_1\mu_1^u + c_2\mu_2^u = c_1\alpha_1\mu_1^* + c_2\alpha_2\mu_2^* \tag{4.1.3}$$

in the interior.

Our informal derivation uses Bellman's Optimality Principle of Dynamic Programming *(Bellman and Dreyfus, 1962)*. Suppose that a control law associated with a routing function $\lambda(\underline{x}, \underline{a})$ is an optimal control law and that the corresponding value functions and average performance are $V_{\underline{a}}(\underline{x})$. (Here, $V_{\underline{a}}(\underline{x})$ is a different notation for $V(\underline{x}, \underline{a})$.) We only consider the case where $\underline{a} = (1, 1, 1)$ and (x_1, x_2) is in the interior of the state space and constraints CT1 hold. Optimality conditions for other choices of $\underline{a}$ can be derived in a similar way.

As T becomes very large, the expected reward during $[0, T]$ is roughly $V_{111}(x_1, x_2)+gT$. This reward is equal to the sum of the immediate reward to be obtained during a very small time interval $[0, \Delta]$ and the expected reward during $[\Delta, T]$. Under several continuity assumptions we have:

$$V_{111}(x_1, x_2) + gT = (1 - p_0\Delta - p_1\Delta - p_2\Delta)\, V_{111}(x_1 + (\lambda_1 - \mu_1)\Delta,\, x_2 + (\lambda_2 - \mu_2)\Delta) + p_0\Delta V_{011}(x_1, x_2) + p_1\Delta V_{101}(x_1, x_2) + p_2\Delta V_{110} + g(T - \Delta) + k(\underline{x}, \underline{a})\Delta \tag{4.1.4}$$

Since the control law considered is optimal, we have:

$$V_{111}(x_1, x_2) + gT = \max_{\lambda_1, \lambda_2}[(1 - p_0\Delta - p_1\Delta - p_2\Delta)V_{111}(x_1 + (\lambda_1 - \mu_1)\Delta, x_2 + (\lambda_2 - \mu_2)\Delta) + p_0\Delta V_{011} + p_1\Delta V_{101} + p_2\Delta V_{110} + g(T - \Delta) + k(\underline{x}, \underline{a})\Delta] \tag{4.1.5}$$

We now assume that V_{111} is differentiable and write

$$V_{111}(x_1 + (\lambda_1 - \mu_1)\Delta,\, x_2 + (\lambda_2 - \mu_2)\Delta) \approx V_{111}(x_1, x_2) + (\lambda_1 - \mu_1)\frac{\partial V_{111}}{\partial x_1} + (\lambda_2 - \mu_2)\frac{\partial V_{111}}{\partial x_2} \tag{4.1.6}$$

Equations (4.1.5) and (4.1.6) yield

$$V_{111}(x_1, x_2) + gT = \max_{\lambda_1, \lambda_2}\left[(1 - p_0\Delta - p_1\Delta - p_2\Delta)\left(V_{111} + (\lambda_1 - \mu_1)\frac{\partial V_{111}}{\partial x_1} + (\lambda_2 - \mu_2)\frac{\partial V_{111}}{\partial x_2}\right)\right] +$$

$$p_0\Delta V_{011} + p_1\Delta V_{101} + p_2\Delta V_{110} + g(T - \Delta) + k\Delta \tag{4.1.7}$$

We discard $O(\Delta^2)$ terms and separate terms that do not depend on λ_1 and λ_2, to obtain,

$$0 = \max_{\lambda_1, \lambda_2}\left[\lambda_1\frac{\partial V_{111}}{\partial x_1} + \lambda_2\frac{\partial V_{111}}{\partial x_2}\right] - (p_0 + p_1 + p_2)\Delta V_{111} +$$

$$p_0\Delta V_{011} + p_1\Delta V_{101} + p_2\Delta V_{110} - g\Delta + k\Delta \tag{4.1.8}$$

The Dynamic Programming conditions for Optimality then reduce to the requirement that the expression

$$\lambda_1\frac{\partial V_{111}}{x_1} + \lambda_2\frac{\partial V_{111}}{\partial x_2} \equiv \lambda_1\left(\frac{\partial V_{111}}{x_1} - \frac{\partial V_{111}}{\partial x_2}\right) + (\lambda_1 + \lambda_2)\frac{\partial V_{111}}{\partial x_2} \tag{4.1.9}$$

is maximized at very point. Because of CT1, we have $\lambda_1 + \lambda_2 = \lambda^*$ and we come down to the maximization of

$$\lambda_1\left(\frac{\partial V_{111}}{\partial x_1} - \frac{\partial V_{111}}{\partial x_2}\right)$$

subject to $0 \leq \lambda_1 \leq \lambda^*$. The solution is then

$$\begin{aligned} \lambda_1 = \lambda^* \quad & \text{if} \quad \frac{\partial V_{111}}{\partial x_1} > \frac{\partial V_{111}}{\partial x_2} \\ 0 \quad & \text{if} \quad \frac{\partial V_{111}}{\partial x_1} < \frac{\partial V_{111}}{\partial x_2} \end{aligned} \tag{4.1.10}$$

Therefore, an optimal routing strategy always tries to direct the system towards higher values of the value functions and this is done by evaluating and comparing their partial derivatives. However, the situation may become considerably more complex at points where the partial derivatives are equal or undefined.

At all points where the partial derivatives are not equal, the optimal policy consists of routing all material to one of the two buffers, instead of splitting the flow. This property is valid whenever the reward function is independent or a linear function of the control law that is being used. This is the case for most physically meaningful routing problems,

in the interior of the state space, and it is used in Sections 4.4 and 4.5 in order to reduce the set of control laws that are likely to be optimal.

4.2. Optimality and Comparison Theorems.

We note that the right-hand side of equation (4.1.7) is very closely related to the infinitesimal operator of the process. We therefore expect that optimality conditions can be expressed in terms of that operator. In fact, this is the case. The result obtained is very similar to the Hamiltonian minimization of deterministic optimal control except that now the quantity minimized is $\mathcal{L}V + k$.

Let U_0 be a set of admissible control laws for which the conclusions of the theorems of Chapter 3 are true. Let $s^u(t)$ be the Markov process on the state space S resulting from the application of a control law $u \in U_0$. Let k^u, $\mathcal{L}^u$, V^u, g^u be the corresponding reward function, weak infinitesimal operator, value function and expected average reward, respectively.

Theorem 4.1: (Sufficiency)

If for any control law $w \in U_0$, V^u belongs to the domain of $\mathcal{L}^w$ and if for all $s \in S$

$$\max_{w \in U_0} [(\mathcal{L}^w V^u + k^w)(s)] = (\mathcal{L}^u V^u + k^u)(s) \tag{4.2.1}$$

then control law u is optimal.

Proof:

$$\max_{w \in U_0} [(\mathcal{L}^w V^u + k^w)(s)] = (\mathcal{L}^u V^u + k^u)(s), \quad \forall s \in S \quad \Rightarrow$$

$$(\mathcal{L}^w V^u + k^w)(s) \leq (\mathcal{L}^u V^u + k^u)(s), \quad \forall w \in U_0, \forall s \in S \tag{4.2.2}$$

From Theorem 3.4, section 3.5 we have:

$$(\mathcal{L}^u V^u)(s) = -k^u(s) + g^u \tag{4.2.3}$$

Combining (4.2.2) and (4.2.3) we obtain:

$$-k^w(s) + g^u \geq (\mathcal{L}^w V^u)(s), \quad \forall s \in S, \forall w \in U_0 \tag{4.2.4}$$

Inequality (4.2.4) holds in particular for the sample functions of the process $s^w(t)$, so

$$-k^w(s^w(t)) + g^u \geq (\mathcal{L}^w V^u)(s), \quad \forall w \in U_0, \ \forall t \geq 0 \tag{4.2.5}$$

Let T_1^w, T_2^w be two consecutive regeneration times of the process $s^w(t)$ as defined in section 3.1. We integrate both sides of inequality (4.2.5) over the random interval $[T_1^w, T_{2,}{}^w]$ and take expectations:

$$E\left[\int_{T_1^w}^{T_2^w} (-k^w(s^w(t)) + g^u)\, dt\right] \geq E\left[\int_{T_1^w}^{T_2^w} (\mathcal{L}^w V^u)(s^w(t))\, dt\right]$$

By Dynkin's theorem (Appendix I):

$$E\left[\int_{T_1^w}^{T_2^w} (\mathcal{L}^w V^u)(s^w(t))\, dt\right] = E[V^u(s^w(T_2^w)) - V^u(s^w(T_1^w))]$$

By definition, $s^w(T_2^w) = s^w(T_1^w)$. Therefore,

$$E\left[\int_{T_1^w}^{T_2^w} k^w(s^w(t))\,dt\right] \leq E\left[\int_{T_1^w}^{T_2^w} g^u\,dt\right] = E[T_2^w - T_1^w]g^u \quad \Rightarrow$$

$$g^u \geq \frac{E\left[\int_{T_1^w}^{T_2^w} k^w(s^w(t))\,dt\right]}{E[T_2^w - T_1^w]} \tag{4.2.6}$$

The quantity on the right-hand side of (4.2.6) is equal to g^w by Theorem 3.3, section 3.2. Therefore, $g^u \geq g^w$, $\forall w \in U_0$. ∎

Theorems similar to the above have appeared in the literature for various classes of stochastic processes and they have been proved using a wide range of approaches. Of those that are similar to Theorem 4.1, Kushner (1967) and Rishel (1970) seem to be the first that have formulated a stochastic maximum principle by making use of infinitesimal operators. Wohnam (1970), Stone (1973) and Kushner (1978) have addressed the infinite time horizon, average cost problem. Stone considers pure jump processes, while the other two authors consider Ito processes.

Our proof of Theorem 4.1, as well as of the next theorems (4.2, 4.3, 4.4) is quite different than the proofs appearing in the literature because it exploits in a specific way the existence of the regeneration times $(T_1^w, T_2^w, \ldots)$ and does not use the invariant, steady-state probability measure of the process. The major exception is the work of Kushner (1978) which has striking similarities with our work, except that it deals with Ito processes.

By exploiting the special structure of the problems we are considering, one can prove under suitable assumptions a necessity theorem of the following form: If a control law u

is optimal, then

$$(\mathcal{L}^u V^u + k^u)(s) \geq (\mathcal{L}^w V^u + k^w)(s), \quad \forall s \in S, \forall w \in U_0 \tag{4.2.7}$$

We derive, however, a different necessity theorem which will be more helpful in the development of an iterative algorithm. Before that, we prove a theorem which compares the performance of two different control laws. The only similar result in the literature we are aware of is in Kushner (1978).

Consider two control laws u, $w \in U_0$ and suppose that V^u belongs to the domain of $\mathcal{L}^w$. Let

$$Q_\epsilon = \{s \in S: (\mathcal{L}^w V^u + k^w - \mathcal{L}^u V^u - k^u)(s) = \epsilon\} \tag{4.2.8}$$

Let T_1^w, T_2^w be two consecutive regeneration times of the process $s^w(t)$, as defined in Section 3.1. Let

$$P_\epsilon = \{t \in [T_1^w, T_2^w]: s^w(t) \in \bigcup_{\delta \leq \epsilon} Q_\delta\} \tag{4.2.9}$$

Clearly, P_ϵ is a random set, because it depends on the specific sample function $s^w(t)$ that is being observed.

Lemma 4.1: $\forall \omega \in \Omega$, P_ϵ is a measurable set.

Proof: Given some $\omega_o \in \Omega$, T_1 and T_2 are completely determined. Let t_1, t_2 be defined by

$$t_1 = T_1^w(\omega_o), \qquad t_2 = T_2^w(\omega_o) \tag{4.2.10}$$

The Markov process $s^w(t)$ is right-continuous. Therefore it has the following property *(Meyer, 1966, p.70, theorem T47)*: The mapping $s^w(\omega, \tau)$:$\Omega \times [0, t] \mapsto S$ is a measurable function, $\forall t$.

Therefore, for every measurable subset A of S, the set

$$(s^w)^{-1}(A) = \{(\omega, t) \in \Omega \times [t_1, t_2]: s^w(\omega, t) \in A\} \tag{4.2.11}$$

is a measurable subset of $\Omega \times [0, t_2]$. The set $\bigcup_{\delta \leq \epsilon} Q_\delta$ is a measurable subset of S because $\mathcal{L}^w V^u$, k^w, $\mathcal{L}^u V^u$, k^u are all measurable functions on S.

Therefore, if $A = \bigcup_{\delta \leq \epsilon} Q_\delta$, $(s^w)^{-1}(A)$ is measurable. Then,*(Loeve, 1977, Vol. 1, p.61)* the following "section" of the above set is also measurable:

$$P_\epsilon{}^* = \{(\omega, t) \in \Omega \times [t_1, t_2]: \omega = \omega_o,\ s^w(\omega, t) \in (\bigcup_{\delta \leq \epsilon} Q_\delta)\} \tag{4.2.12}$$

This shows that P_ϵ is also measurable, because $P_\epsilon = \{\omega_o\} \times P_\epsilon{}^*$.

Let $l()$ be the Borel measure on $(-\infty, +\infty)$. Then $l(P_\epsilon)$ can be interpreted as the time spent in $\bigcup_{\delta \leq \epsilon} Q_\delta$ and is a random variable itself because it depends on which sample function is being observed. We define a set function ϕ on the family of left-open, right-closed intervals by

$$\phi((a, b]) = E[l(P_b) - l(P_a)] \tag{4.2.13}$$

Then $\phi((a, b])$ can be interpreted as the expected time spent by the process in the set $\bigcup_{a < \delta \leq b} Q_\delta$.

The set function ϕ is σ-additive. Therefore, *(Halmos, 1974)* it can be extended uniquely to a measure ϕ on $(-\infty, +\infty)$.

We now motivate the next theorem. The quantity

$$(\mathcal{L}^w V^u + k^w - \mathcal{L}^u V^u - k^u)(s) \tag{4.2.14}$$

can be interpreted as the benefit to be obtained if we use control law w when the state of the system is s and then use control law u afterwards. By averaging this expression over the entire space we obtain the difference in performance $g^w - g^u$ of the two control laws. Kushner (1978) shows that (4.2.14) has to be weighted by the invariant measure of the process resulting from control law u. We use instead the measure ϕ previously defined.

Theorem 4.2:(comparison Theorem)

If for some $u \in U_o$, $w \in U_o$, V^u is in the domain of $\mathcal{L}^w$ then

$$g^w = g^u + \frac{\int \epsilon \, d\phi(\epsilon)}{E[T_2^w - T_1^w]} \tag{4.2.15}$$

Proof: Let

$$h(s) = (\mathcal{L}^w V^u + k^w - \mathcal{L}^u V^u - k^u)(s) \tag{4.2.16}$$

$$E\left[\int_{T_1^w}^{T_2^w} -k^w(s^w(t)) + g^u + h(s^w(t))\, dt\right] =$$

$$E\left[\int_{T_1^w}^{T_2^w} -k^w(s^w(t)) + (\mathcal{L}^u V^u + k^u)(s^w(t)) + (\mathcal{L}^w V^u + k^w - \mathcal{L}^u V^u - k^u)(s^w(t))\, dt\right] = \tag{4.2.18}$$

$$E\left[\int_{T_1^w}^{T_2^w} (\mathcal{L}^w V^u)(s^w(t))\, dt\right] = \tag{4.2.19}$$

(by Dynkin's theorem, Appendix I)

$$E[V^u(s^w(T_2^w)) - V^u(s^w(T_1^w))] = 0 \tag{4.2.20}$$

because $s^w(T_2^w) = s^w(T_1^w)$.

Therefore,

$$E\left[\int_{T_1^w}^{T_2^w} k^w(s^w(t))\,dt\right] = g^u E[T_2^w - T_1^w] + E\left[\int_{T_1^w}^{T_2^w} h^w(s^w(t))\,dt\right] \quad \Rightarrow \tag{4.2.21}$$

$$g^w = g^u + \frac{E\left[\int_{T_1^w}^{T_2^w} h^w(s^w(t))\,dt\right]}{E[T_2^w - T_1^w]}. \tag{4.2.22}$$

For every $\omega \in \Omega$, the set function ψ defined by $\psi((a,b]) = l(P_b) - l(P_a)$ is σ-additive and can be extended to a unique measure on $(-\infty, +\infty)$. Moreover, since all set functions considered are bounded, we have (by dominated convergence)

$$E[\psi(A)] = \phi(A) \tag{4.2.23}$$

for any Borel measurable $A \subseteq (-\infty, +\infty)$. Then, the Fubini theorem (Appendix I) implies

$$\int \epsilon\, d\phi = E\left[\int \epsilon\, d\psi\right] \tag{4.2.24}$$

To complete the proof, we only need to show that

$$\int_{T_1^w}^{T_2^w} h(s^w(t))\,dt = \int \epsilon\, d\psi, \qquad \forall \omega \in \Omega \tag{4.2.25}$$

But this is just a special case of the theorem on induced measures in Appendix I. ∎

The above theorem has two sorts of implications. On one hand it implies the necessity theorem presented below. On the other hand it provides us with a measure of the performance improvement when control law w is substituted for control law u. It is practically impossible to obtain the measure ϕ and apply Theorem 4.2 in an exact way. However, in many cases it is possible to get some useful inequality bounds. For example if the set

$$\{s \in S: h(s) \in [a, b]\} \tag{4.2.26}$$

contains an open ball of radius r and the assumptions of Theorem 3.2 hold, we may conclude that $\phi([a, b]) \geq f(r)$ where f is a real function, independent of the specific control law that is being used. This can be done as follows.

Theorem 3.2 shows that for any control law and any open ball of radius r, there is probability greater than some p_r that this ball is reached between two consecutive regeneration times of the process. The speed of motion of the state variables $(dx_1/dt,\ dx_2/dt)$ is bounded. Let B be such that $|dx_i/dt| \leq B$, for all control laws and all states. Then, the expected time that the process spends in a ball of radius $2r$ is greater than $p_r r/B$.

The following necessary conditions for optimality is a rewording of the comparison theorem:

Theorem 4.3: (Necessity) If $u \in U \subseteq U_o$, $w \in U \subseteq U_o$, V^u is in the domain of $\mathcal{L}^w$, u is optimal among all $w \in U$,

$$(\mathcal{L}^w V^u + k^w)(s) \geq (\mathcal{L}^u V^u + k^u)(s), \quad \forall s \in S \tag{4.2.27}$$

$$(\mathcal{L}^w V^u + k^w)(s) \geq (\mathcal{L}^u V^u + k^u)(s) + \delta, \quad \forall s \in Q, \delta > 0 \tag{4.2.28}$$

for some $Q \subseteq S$, then the set $\{t \in [T_1^w, T_2^w],\ s^w(t) \in Q\}$ has measure zero, almost surely. Moreover, if the assmptions of Theorem 3.2 are satisfied, the set Q contains no open set.

Proof: By Theorem 4.2,

$$g^w = g^u + \frac{\int \epsilon \, d\phi(\epsilon)}{E[T_2^w - T_1^w]} \tag{4.2.29}$$

Equation (4.2.27) implies that $d\phi(\epsilon) = 0$ whenever $\epsilon < 0$. Therefore,

$$\int \epsilon \, d\phi(\epsilon) \geq 0 \tag{4.2.30}$$

Since u is optimal, $g^w \leq g^u$ and this implies that

$$\int \epsilon \, d\phi(\epsilon) = 0 \tag{4.2.31}$$

This implies that $\delta\phi([\delta, \infty)) = 0$. Recall that

$$\phi([\delta, \infty)) = E[\psi([\delta, \infty))] \tag{4.2.32}$$

Since $\psi([\delta, \infty))$ is a non-negative random variable with zero expectation we have

$$\psi([\delta, \infty)) = 0, \quad \text{almost surely} \tag{4.2.33}$$

But

$$\psi([\delta, \infty)) \geq l(\{t \in [T_1^w, T_2^w],\ s^w(t) \in Q\}) \tag{4.2.34}$$

where l is again the Borel measure.

The last assertion of the Theorem follows immediately from Theorem 3.2. (See also the comments preceding this theorem.) ∎

4.3. Description of Various Classes of Admissible Strategies.

Recall that the operator equation $\mathcal{L}V = -k + g$ was derived in Chapter 3 with practically no restriction on admissible control laws except for the requirement that the resulting sample paths should be piecewise continuously differentiable. This was because, no matter how irregular a control law is, the corresponding value function is guaranteed to belong to the domain of the weak infinitesimal operator. However, in order to compare two control laws u and w we need to consider expressions of the form $\mathcal{L}^u V^w$ where V^w is the value function corresponding to control law w and $\mathcal{L}^u$ is the infinitesimal operator corresponding to control law u. There is no guarantee, in general, that V^w will be in the domain of $\mathcal{L}^u$. In this section we introduce several assumptions and we define several classes of admissible control laws, in a way that enables us to study the optimization problem. These assumptions are of two different kinds:

1) Stability assumptions.

2) Smoothness assumptions.

Stability assumptions correspond to an intuitively appealing class of control laws and can be justified on physical grounds.

Smoothness assumptions, on the other hand are justified on mathematical grounds: the set of smooth functions or the set of smooth control laws, is, more often than not, dense in any arbitrary set of functions or of control laws. So, even if the optimal control

law does not obey these assumptions, there exist smooth control laws whose performance is as close as desired to the optimal. Also, any law that is physically implemented is likely to be a a smooth one.

Recall that all control laws considered in this thesis have been assumed to satisfy CT1 (section 2.3). Therefore any class of admissible control laws may be viewed as a set of functions $(\lambda_1(\underline{x}, \underline{a}), \lambda_2(\underline{x}, \underline{a}))$ on the state space.

Definition 4.1: Let U_o be the set of functions $(\lambda_1(\underline{x}, \underline{a}), \lambda_2(\underline{x}, \underline{a}))$ on the state space which are Borel measurable, satisfy CT1, and are such that the resulting deterministic differential equations

$$dx_i = (\lambda_i(\underline{x}, \underline{a}) - \mu_i(\underline{x}, \underline{a}))\, dt \tag{4.3.1}$$

(for any fixed $\underline{a}$ and any initial state) a unique solution $\underline{x}(t)$ which is piecewise continuously differentiable.

Definition 4.2: Let U_S be the subset of U_o such that the solution $\underline{x}(t)$ of the deterministic state equations is continuous in the initial state, for any t and for any $\underline{a}$. We will say that a control law is stable if it belongs in U_S.

Stability here means that if we change the initial state by a little bit, the future evolution of the system is not changed significantly. Figure 4.1 illustrates a control law which belongs in U_o but not in U_S. Namely, for $\underline{a} = (1, 0, 0)$:

$$\text{If } x_1 \geq x_2,\ x_1 \neq N\text{:} \quad \lambda_1(\underline{x}, \underline{a}) = \lambda^* \quad l_2(\underline{x}, \underline{a}) = 0 \tag{4.3.2}$$

$$\text{If } x_1 < x_2,\ x_2 \neq N\text{:} \quad \lambda_1(\underline{x}, \underline{a}) = 0 \quad \lambda_2(\underline{x}, \underline{a}) = \lambda^* \tag{4.3.3}$$

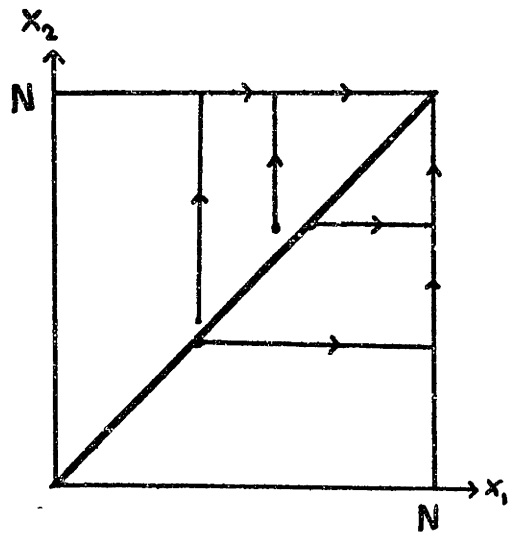


Figure 4.1: An Unstable Control Law

Let the initial state be $(x_1, x_2) = (0, 0)$. Then, $(x_1(t), x_2(t)) = (\lambda^* t, 0)$. Then, let the initial state be $(x_1^*, x_2^*) = (0, \epsilon)$. Then, $(x_1^*(t), x_2^*(t)) = (0, \epsilon + \lambda^* t)$. As $\epsilon \rightarrow 0$, $(x_1^*(0), x_2^*(0))$ converges to $(x_1(0), x_2(0))$ but $(x_1^*(t), x_2^*(t))$ does not converge to $(x_1(t), x_2(t))$, for $t > 0$. Therefore, this is not a stable control law.

Such a control law might be optimal if the performance criterion rewarded the states on the boundary heavily. However, such a performance criterion makes little physical sense, especially for manufacturing systems. We expect that only stable control laws need to be considered for optimality in a physical system.

Definition 4.3: Let U_V be the set of control laws such that the state space can be partitioned to a finite collection of sets of two types:

a) Sets of type I in which the value function is continuously differentiable and its derivatives are bounded.

b) Sets of type II in which V possesses one-sided directional derivatives, along all directions. Moreover, we require that sets of type II are either single lines or piecewise continuously differentiable lines.

Definition 4.4: Let U_c be the set of control laws such that the value function is continuous.

Remark: $U_c \subseteq U_V$

Definitions 4.3 and 4.4 are only indirect characterizations of the classes U_c and U_V. We would like to have an alternative characterization, in terms of $\lambda_1(\underline{x}, \underline{\alpha})$. We make the following conjectures that link the various classes of admissible control laws:

Conjecture 4.1: Let $u \in U_o$. Let $\lambda_1^u(\underline{x}, \underline{\alpha})$ and $\lambda_2^u(\underline{x}, \underline{\alpha})$ be the corresponding functions

that determine u and let $k^u(\underline{x}, \underline{a})$ be the corresponding reward function. If (1) $u \in U_S$, (2) $\lambda_1^u(\underline{x}, \underline{a})$, $\lambda_2^u(\underline{x}, \underline{a})$ and $k^u(\underline{x}, \underline{a})$ are continuous except on a set A on which they have one-sided limits along any direction, (3) the set A consists of isolated points or piecewise continuously differentiable lines, then $\lambda_1(\underline{x}, \underline{a}) \in U_V$.

Conjecture 4.1 is very similar to a theorem proved in *(Rishel, 1975)* in a slightly different setting.

Conjecture 4.2: $U_S \subseteq U_c$ (The value functions associated with stable control laws are continuous.)

The next conjecture concerns the possible deterioration of attainable performance if we restrict the admissible controls to the class U_V.

Conjecture 4.3:

$$\sup_{u \in U_o} g^u = \sup_{w \in U_V} g^w \tag{4.3.4}$$

$$\sup_{u \in U_S} g^u = \sup_{w \in U_V \bigcap U_S} g^w \tag{4.3.5}$$

However, the main reason that class U_V is of interest is the following proposition:

Proposition 4.1: If $u \in U_V$ and $w \in U_V$ then V^u is in the domain of $\mathcal{L}^w$.

Proof: In Chapter 3 we showed that the infinitesimal operator involves one-sided derivatives only. Moreover, the definition of U_V implies (indirectly) that these derivatives are bounded; so $\mathcal{L}^w V^u$ exists. ∎

The statements of Theorems 4.1, 4.2 and 4.3 show that U_V is the natural setting in which the optimization problem should be solved.

4.4. Definition of Various Classes of Control Laws Relevant to Production Maximization.

The introduction of the classes U_o, U_S, U_V, U_c of admissible control laws was very much motivated by a desire to determine the mathematical framework within which the optimization problem can be addressed.

On the other hand there are certain classes of control laws that are especially relevant to the routing problem when production rate is maximized, or, more generally, when the reward function $k(\underline{x}, \underline{a})$ is independent of the control in the interior and constraints CT1 have been imposed.

Definition 4.5: We will say that a control law $u \in U_o$ belongs to U_R if the following conditions are met: For every $\underline{a}$ such that $a_0 = 1$, the interior of the state space can be partitioned in four regions R_1, R_2, R_3, R_4 such that:

$\lambda_1(\underline{x}, \underline{a}) = \lambda^* \quad \text{in } R_1$

$\lambda_1(\underline{x}, \underline{a}) = 0 \quad \text{in } R_2$

In R_3 the evolution of the state variables $\underline{x}$ (and therefore $\lambda_1(\underline{x}, \underline{a})$) is uniquely determined by R_1 and R_2.

The performance of the system is unchanged if $\lambda_1(\underline{x}, \underline{a})$ is modified only on R_4.

Figure 4.2 provides an example of a control law in U_R in which the fourth region is empty and which illustrates how it is possible that the shape of the first and the second regions may determine the control law on a third region. We assume that Figure 4.2 shows a control law when $a = (1, 1, 1)$. We also assume that $\mu_i^* > \lambda^*$, $i = 1, 2$ so that state trajectories move towards the origin. We have

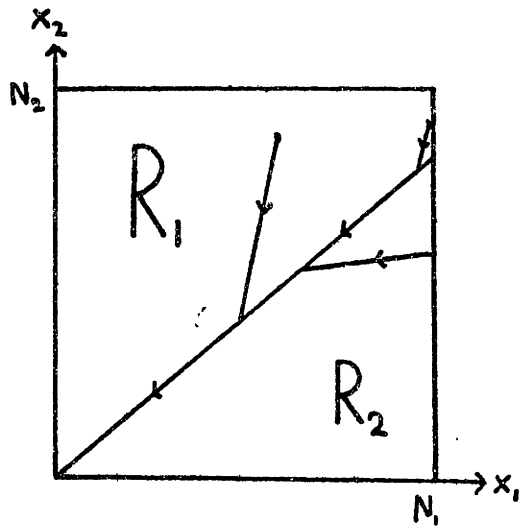


Figure 4.2: A Control Law in U_R

$$R_1 = \{\underline{x}: x_2 > x_1\}$$

$$R_2 = \{\underline{x}: x_1 > x_2\}$$

Consider the behavior of the system during a time interval in which no machine fails. Whenever the state of the system is in region R_1 or R_2, the control law is such that the state moves towards the line $x_1 = x_2$. Once the set $\{\underline{x}: x_1 = x_2\}$ is reached a trajectory cannot leave this set. This is because the control law exerts an infinitely fast stabilizing action. We then have,

$$\lambda_1 - \mu_1^* = \lambda_2 - \mu_2^* \quad \Rightarrow$$

$$\lambda_1 - \mu_1^* = \lambda^* - \lambda_1 - \mu_2^* \quad \Rightarrow$$

$$\lambda_1 = \frac{1}{2}(\mu_1^* - \mu_2^*)$$

on the set $\{\underline{x}: x_1 = x_2\}$. Therefore, $R_3 = \{\underline{x}: x_1 = x_2\}$, because λ_1 on R_3 is uniquely determined by the shape of R_1 and R_2. Such will always be the case when regions R_1 and R_2 are separated by a line and the control law in these regions forces trajectories to move towards the dividing line.

Region R_3 could not have nonempty interior and still satisfy the conditions appearing in the definition of U_R. Forgetting about mathematical technicalities, we may assume that R_3 will only consist of lines.

Recall that the definition of U_R referred exclusively to the interior. Additional information (other than the nature of the regions R_i) is needed to determine completely any control law $u \in U_R$.

Definition 4.6: We let U_L be the subset of $U_R \bigcap U_S$ such that the interior of the state space may be partitioned in two parts R_1 and R_2 (see Definition 4.5) and, for each value of $\underline{a}$ with $a_0 = 1$, a single piecewise continuously differentiable line R_L separates R_1 and R_2.

The control law in Figure 4.2 is an example of a member of U_L.

The sets U_R and U_L are rather restricted classes of control laws. However, we argue in the next section that U_R contains all candidates for optimality. In fact the empirical evidence (Hahne, 1980) indicates that optimal control laws belong to U_L. Unfortunately, from a mathematically rigorous point of view, the optimality conditions do not suggest directly that we may restrict our attention to U_L.

4.5. Interpretation of Optimality Conditions.

Theorems 4.1 and 4.3 are only a formal way to state and generalize the conclusions of the informal argument of section 4.1. Recall that the main conclusion was that for a control law to be optimal we should have:

$$\frac{\partial V^u}{\partial x_1} > \frac{\partial V^u}{\partial x_2} \quad \Rightarrow \quad \lambda_2 = 0 \tag{4.5.1}$$

$$\frac{\partial V^u}{\partial x_1} < \frac{\partial V^u}{\partial x_2} \quad \Rightarrow \quad \lambda_1 = 0 \tag{4.5.2}$$

for those $(\underline{x}, \underline{a})$ for which the above derivatives are well-defined, $k^u(\underline{x}, \underline{a})$ is independent of the control law and the failure rates are also independent of the control law.

Theorem 4.2 implies that the value of the performance criterion is insensitive to what the control law is on the set

$$R_4 = \{x : \frac{\partial V}{\partial x_1} = \frac{\partial V}{\partial x_2}\}$$

This establishes the existence of regions R_1, R_2, R_4 as they appeared in the definition of class U_R in section 4.4.

We now demonstrate that if a stable control law u belongs in U_V (definition 4.3, section 4.3) and satisfies the optimality conditions then it belongs in U_R.

Suppose that $u \in U_V$. Then, for all points $(\underline{x}, \underline{a})$ belonging in a set of type I, the partial derivatives of the value functions exist. Depending on the sign of $\partial V/\partial x_1 - \partial V/\partial x_2$, such points belong in R_1, R_2, or R_4.

We also need to examine the points where V is not differentiable. These points belong (by definition) to sets of type II. Fix $\underline{a}$ such that $a_0 = 1$ and suppose that $(\underline{x}, \underline{a}) \in A \subseteq S$ and A is a line of type II. Then V possesses one-sided directional derivatives along any direction at x. Consider the set of "feasible" directions, that is the set of directions along which the system is allowed to move. For example, if $\underline{a} = (1, 0, 0)$,feasible directions are those along which both x_1 and x_2 are non-decreasing. Then we amy have one of the three cases in Figure 4.4. (Arrows B and C indicate the range of feasible directions.)

The direction in which the system will move is determined by λ_1. The optimal direction is the direction in which $(\partial V/\partial x_1 - \partial V/\partial x_2)\lambda_1$ is maximized. On each side of the line, $\partial V/\partial x_1 - \partial V/\partial x_2$ is uniquely defined. Let

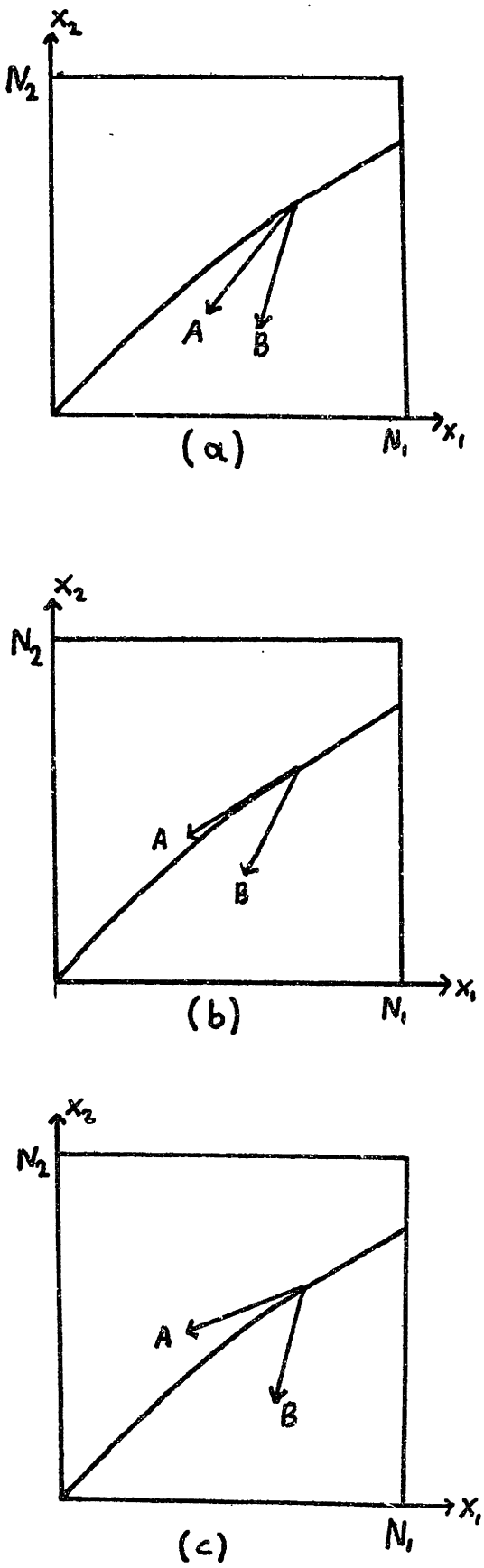


Figure 4.3: The Set of Feasible Directions of Motion

$$I_L = \frac{\partial V}{\partial x_1} - \frac{\partial V}{\partial x_2} \qquad \text{on the left half}$$

$$I_R = \frac{\partial V}{\partial x_1} - \frac{\partial V}{\partial x_2} \qquad \text{on the right half}$$

In cases (a) and (b), in Figure 4.4, the optimal value of λ_1 is

$$\lambda_1 = \begin{array}{ll} 0 & \text{if } I_L < 0 \\ \lambda^* & \text{if } I_L \geq 0 \end{array}$$

and either $x \in R_1$ or $x \in R_2$.

In case (c) we distinguish several possibilities:

(i) $I_L \geq 0$, $I_R \geq 0$. Then, $\lambda_1 = \lambda^*$, $x \in R_1$.

(ii) $I_L \leq 0$, $I_R \leq 0$. Then, $\lambda_1 = 0$, $x \in R_2$.

(iii) $I_L \geq 0$, $I_R \leq 0$. Then, we have $\lambda_1 = 0$ to the right of the line and $\lambda_2 = 0$ to the left of the line. So, the system can only stay on the line. (Hence, $x \in R_3$.)

(iv) $I_L < 0$, $I_R > 0$. This is impossible because we would have $\lambda_2 = 0$ to the right, $\lambda_1 = 0$ to the left and the system would violate our stability assumption.

4.6. Interpretation of Optimality Conditions under Production Rate Maximization.

We saw in the previous sections that if $k(\underline{x}, \underline{a})$ is control independent then a control law is optimal if and only if it always tries to move the state of the system towards regions where value functions have higher values. However, when $k(\underline{x}, \underline{a})$ is control dependent, a tradeoff appears between immediate and future rewards and, usually, one has to do more

than just calculate the partials $\partial V/\partial x_1$, $\partial V/\partial x_2$. Moreover, under different control laws the failure rates of some machines may be different and this effect has to be taken into account directly.

In this section we derive the optimality conditions for some illustrative examples. In Appendix IV we give a complete list of the optimality conditions for the particular case $\lambda^* = \mu_1^* + \mu_2^*$.

Fix again $a = (1, 1, 1)$ and

$$k(\underline{x}, \underline{a}) = c_1\mu_1(\underline{x}, \underline{a}) + c_2\mu_2(\underline{x}, \underline{a}) \tag{4.6.1}$$

Theorems 4.1 and 4.3 imply that

$$\begin{aligned} H \equiv & c_1\mu_1 + c_2\mu_2 + (\lambda_1 - \mu_1)\frac{\partial V_{111}}{\partial x_1} + (\lambda_2 - \mu_2)\frac{\partial V_{111}}{\partial x_2} \\ & p_0^*\frac{\lambda}{\lambda^*}(V_{011} - V_{111}) + p_1^*\frac{\mu_1}{\mu_1^*}(V_{101} - V_{111}) + p_2^*\frac{\mu_2}{\mu_2^*}(V_{110} - V_{111}) \end{aligned} \tag{4.6.2}$$

has to be maximized at every point. We now interpret this condition at the various points of the state space.

A. $x_1 > 0$, $x_2 > 0$, $x_1 \neq N_1$, $x_2 \neq N_2$).

From constraints CT1, $\lambda = \lambda^*$, $\mu_1 = \mu_1^*$, $\mu_2 = \mu_2^*$. Therefore, we only need to maximize $\lambda_1\partial V_{111}/\partial x_1 + \lambda_2\partial V_{111}/\partial x_2$. This is the case that was examined in section 4.5.

B. $x_1 = 0$, $x_2 > 0$

Assume $\lambda^* > \mu_1^*$. From constraints CT1, $\lambda = \lambda^*$, $\mu_1 = \min\{\mu_1^*, \lambda_1\}$, $\mu_2 = \mu_2^*$. When $\lambda_1 \leq \mu_1^*$, H (equation (4.6.2)) becomes

$$c_1\lambda_1 + c_2\mu_2^* + (\lambda^* - \lambda_1 - \mu_2^*)\frac{\partial V_{111}}{\partial x_2} + p_0^*(V_{011} - V_{111}) + p_1^*\frac{\lambda_1}{\mu_1^*}(V_{101} - V_{111}) + p_2^*(V_{110} - V_{111}) \tag{4.6.3}$$

When $\lambda_1 \geq \mu_1^*$, H becomes

$$c_1\mu_1^* + c_2\mu_2^* + (\lambda_1 - \mu_1^*)\partial V_{111}\partial x_1 + (\lambda^* - \lambda_1 - \mu_2^*)\frac{\partial V_{111}}{\partial x_2} + p_0^*(V_{011} - V_{111}) + p_1^*(V_{101} - V_{111}) + p_2^*(V_{110} - V_{111}) \tag{4.6.4}$$

We note that H, as a function of λ_1 is piecewise linear. We therefore need to evaluate $H(\lambda_1)$ only at the extreme values of λ_1 ($\lambda_1 = 0$, $\lambda_1 = \lambda^*$) and at the value of λ_1 where $H(\lambda_1)$ changes slope ($\lambda_1 = \mu_1^*$). We then have:

$$H(0) = (\lambda^* - \mu_2^*)\frac{\partial V_{111}}{\partial x_2} + C \tag{4.6.5}$$

$$H(\mu_1^*) = (\lambda^* - \mu_1^* - \mu_2^*)\frac{\partial V_{111}}{\partial x_2} + p_1^*(V_{101} - V_{111}) + c_1\mu_1^* + C \tag{4.6.6}$$

$$H(\lambda^*) = (\lambda^* - \mu_1^*)\frac{\partial V_{111}}{\partial x_1} + p_1^*(V_{101} - V_{111}) + c_1\mu_1^* + C \tag{4.6.7}$$

where

$$C = p_0^*(V_{011} - V_{111}) + p_2^*(V_{110} - V_{111}) + c_2\mu_2^* \tag{4.6.8}$$

The optimal control law is

$$\lambda_1 = \text{argmax}\{H(0),\ H(\mu_1^*),\ H(\lambda^*)\} \tag{4.6.9}$$

A similar condition may be easily obtained when $\lambda^* < \mu_1^*$.

C. $x_1 = x_2 = 0$

Assume $\lambda^* < \mu_i^*$, $i = 1, 2$. Constraints CT1 imply, $\mu_i = \lambda_i$, $\lambda = \lambda^*$. Then,

$$H(\lambda_1) = p_0^*(V_{011} - V_{111}) + p_1^* \frac{\lambda_1}{\mu_1^*}(V_{101} - V_{111}) + p_2^* \frac{(\lambda^* - \lambda_1)}{\mu_2^*}(V_{110} - V_{111}) + c_1\lambda_1 + c_2(\lambda^* - \lambda_1) \tag{4.6.9}$$

We see that $H(\lambda_1)$ is linear, so the optimal value of λ_1 is either $\lambda_1 = 0$, or $\lambda_1 = \lambda^*$ depending on which one one of the following two expressions is larger:

$$\frac{p_2^*}{\mu_2^*}(V_{110} - V_{111}) + c_2 \tag{4.6.10}$$

or

$$\frac{p_1^*}{\mu_1^*}(V_{101} - V_{111}) + c_1 \tag{4.6.11}$$

Similar optimality conditions may be obtained when the assumption $\lambda^* < \mu_1^*$ is not true.

4.7. Differential Equations when a Control Law is Characterized by Dividing Lines.

In this section we present the partial differential equations for the value functions as they are simplified in the case of a control law belonging in U_L. In this case the control law is characterized, in the interior, by a single dividing line for each value of $\underline{a}$ such that $a_0 = 1$ and is stable (in the sense of the definition in Section 4.3.). In order to avoid considering many different possibilities we assume that:

(i) For any fixed $\underline{a}$, the slope of the dividing line is such that this line is an admissible path for the state trajectories.

(ii) We let $k(\underline{x}, \underline{a}) = c_1\mu_1(\underline{x}, \underline{a}) + c_2\mu_2(\underline{x}, \underline{a})$.

Interior Equations

$\underline{a} = (1, 1, 1)$, Upper Half:

$$(\mu_1{}^* - \lambda^*)\frac{\partial V_{111}}{\partial x_1} + \mu_2{}^*\frac{\partial V_{111}}{\partial x_2} = -(p_0{}^* + p_1{}^* + p_2{}^*)V_{111} + p_0{}^*V_{011} + p_1{}^*V_{101} + p_2{}^*V_{110} + c_1\mu_1{}^* + c_2\mu_2{}^* - g \tag{4.7.1}$$

$\underline{a} = (1, 1, 1)$, Lower Half:

$$\mu_1{}^*\frac{\partial V_{111}}{\partial x_1} + (\mu_2{}^* - \lambda^*)\frac{\partial V_{111}}{\partial x_2} = -(p_0{}^* + p_1{}^* + p_2{}^*)V_{111} + p_0{}^*V_{011} + p_1{}^*V_{101} + p_2{}^*V_{110} + c_1\mu_1{}^* + c_2\mu_2{}^* - g \tag{4.7.2}$$

$\underline{a} = (1, 1, 0)$, Upper Half:

$$(\mu_1{}^* - \lambda^*)\frac{\partial V_{110}}{\partial x_1} = -(p_0{}^* + p_1{}^* + r_2)V_{110} + p_0{}^*V_{010} + p_1{}^*V_{100} + r_2V_{111} + c_1\mu_1{}^* - g \tag{4.7.3}$$

$\underline{a} = (1, 1, 0)$, Lower Half:

$$\mu_1{}^*\frac{\partial V_{110}}{\partial x_1} - \lambda^*\frac{\partial V_{110}}{\partial x_2} = -(p_0{}^* + p_1{}^* + r_2)V_{110} + p_0{}^*V_{010} + p_1{}^*V_{100} + r_2V_{111} + c_1\mu_1{}^* - g \tag{4.7.4}$$

$\underline{a} = (1, 0, 1)$, Upper Half:

$$-\lambda^*\frac{\partial V_{101}}{\partial x_1} + \mu_2{}^*\frac{\partial V_{101}}{\partial x_2} = -(p_0{}^* + r_1 + p_2{}^*)V_{101} + p_0{}^*V_{001} + r_1V_{111} + p_2{}^*V_{100} + c_2\mu_2{}^* - g \tag{4.7.5}$$

$\underline{a} = (1, 0, 1)$, Lower Half:

$$(\mu_2{}^* - \lambda^*)\frac{\partial V_{101}}{\partial x_2} = -(p_0{}^* + r_1 + p_2{}^*)V_{101} + p_0{}^* V_{001} + r_1 V_{111} + p_2{}^* V_{100} + c_2\mu_2{}^* - g \quad (4.7.6)$$

$\underline{a} = (1, 0, 0)$, Upper Half:

$$-\lambda^*\frac{\partial V_{100}}{\partial x_1} = -(p_0{}^* + r_1 + r_2)V_{100} + p_0{}^* V_{000} + r_1 V_{110} + r_2 V_{101} - g \quad (4.7.7)$$

$\underline{a} = (1, 0, 0)$, Lower Half:

$$-\lambda^*\frac{\partial V_{100}}{\partial x_2} = -(p_0{}^* + r_1 + r_2)V_{100} + p_0{}^* V_{000} + r_1 V_{110} + r_2 V_{101} - g \quad (4.7.8)$$

$\underline{a} = (0, 1, 1)$:

$$\mu_1{}^*\frac{\partial V_{011}}{\partial x_1} + \mu_2{}^*\frac{\partial V_{011}}{\partial x_2} = -(r_0 + p_1{}^* + p_2{}^*)V_{011} + r_0 V_{111} + p_1{}^* V_{001} + p_2{}^* V_{010} + c_1\mu_1{}^* + c_2\mu_2{}^* - g \quad (4.7.9)$$

$\underline{a} = (0, 1, 0)$:

$$\mu_1{}^*\frac{\partial V_{010}}{\partial x_1} = -(r_0 + p_1{}^* + r_2)V_{010} + r_0 V_{110} + p_1{}^* V_{000} + r_2 V_{011} + c_1\mu_1{}^* - g \quad (4.7.10)$$

$\underline{a} = (0, 0, 1)$:

$$\mu_2{}^*\frac{\partial V_{001}}{\partial x_2} = -(r_0 + r_1 + p_2{}^*)V_{001} + r_0 V_{101} + r_1 V_{011} + p_2{}^* V_{000} + c_2\mu_2{}^* - g \quad (4.7.11)$$

$\underline{a} = (0, 0, 0)$:

$$0 = -(r_0 + r_1 + r_2)V_{000} + r_0V_{100} + r_1V_{010} + r_2V_{001} - g \qquad (4.7.12)$$

Equations on the Dividing Lines

On the dividing lines λ_1 and λ_2 are uniquely determined by the slope of the dividing line. More precisely, if at some point the slope of the the dividing line is l, λ_1 and λ_2 are uniquely determined by:

$$\lambda^* = \lambda_1 + \lambda_2$$

and

$$\frac{\lambda_2 - \mu_2^*}{\lambda_1 - \mu_1^*} = l$$

We may therefore assume that λ_1 and λ_2 are known.

$\underline{a} = (1, 1, 1)$:

$$(\mu_1^* - \lambda_1)\frac{\partial V_{111}}{\partial x_1} + (\mu_2^* - \lambda_2)\frac{\partial V_{111}}{\partial x_2} = -(p_0^* + p_1^* + p_2^*)V_{111} + p_0^*V_{011} + p_1^*V_{101} + p_2^*V_{110} + \mu_1^*c_1 + \mu_2^*c_2 - g \qquad (4.7.13)$$

$\underline{a} = (1, 1, 0)$:

$$(\mu_1^* - \lambda_1)\frac{\partial V_{110}}{\partial x_1} - \lambda_2\frac{\partial V_{110}}{\partial x_2} = -(p_0^* + p_1^* + r_2)V_{110} + p_0^*V_{010} + p_1^*V_{100} + r_2V_{111} + \mu_1^*c_1 - g \qquad (4.7.14)$$

$\underline{a} = (1, 0, 1)$:

$$-\lambda_1 \frac{\partial V_{101}}{\partial x_1} + (\mu_2{}^* - \lambda_2)\frac{\partial V_{101}}{\partial x_2} = -(p_0{}^* + r_1 + p_2{}^*)V_{101} + p_0{}^* V_{001} + r_1 V_{111} + p_2{}^* V_{100} + c_2\mu_2{}^* - g \tag{4.7.15}$$

$\underline{a} = (1, 0, 0)$:

$$-\lambda_1 \frac{\partial V_{100}}{\partial x_1} - \lambda_2 \frac{\partial V_{100}}{\partial x_1} = -(p_0{}^* + r_1 + r_2)V_{100} + p_0{}^* V_{000} + r_1 V_{110} + r_2 V_{101} - g \tag{4.7.16}$$

Boundary Conditions

The equations for the value functions may be slightly different on some points of the boundary because constraints on λ_i and μ_i may become effective. Then, one has to go back to the general form of the equations for the value functions as they have been derived in Chapter 3. Essentially, the equations valid at those points are the same as the interior equations, except that failure rates may have to be adjusted according to the degree of utilization of the machines. Namely, if one (or more) of the machines is utilized below its maximum capacity $p_i{}^*$ has to be replaced by p_i, the adjusted failure rate (equation 2.1.9). In Appendix IV we give a complete list of the boundary conditions when $\lambda^* = \mu_1{}^* + \mu_2{}^*$.

Despite the fact that this is a large number of equations, we should emphasize again that they are all special cases of the operator equation

$$\mathcal{L}^u V^u + k^u = g^u \tag{4.7.18}$$

4.8. Conjectures on Optimal Control Laws.

In the work of Hahne (1980) on the discrete version of our problem, optimal control laws were observed to follow certain patterns. We conjecture that optimal control laws for the continuous problem will also follow the same pattern. The main reason, however, for making these conjectures is that they are perfectly reasonable from an intuitive point of view:

Let

$$k(\underline{x}, \underline{a}) = c_1\mu_1(\underline{x}, \underline{a}) + c_2\mu_2(\underline{x}, \underline{a}) \tag{4.8.1}$$

Let u be an optimal control law and V the corresponding value function. We then conjecture that the following will be true:

C1: Control law u satisfies conditions CT1.

C2: If $c_1 = c_2$, then control law u satisfies conditions CT2.

C3: Control law u belongs to U_L.

C4: All four dividing lines (corresponding to the four values of $\underline{a}$ with $a_0 = 1$) have positive slopes. Equivalently,

$$\frac{\partial V_{\underline{a}}}{\partial x_1}(x_1, x_2) > \frac{\partial V_{\underline{a}}}{\partial x_2}(x_1, x_2) \quad \Rightarrow \quad \frac{\partial V_{\underline{a}}}{\partial x_1}(x_1, x_2 + \Delta) > \frac{\partial V_{\underline{a}}}{\partial x_2}(x_1, x_2 + \Delta), \quad \Delta > 0 \tag{4.8.2}$$

and

$$\frac{\partial V_{\underline{a}}}{\partial x_1}(x_1, x_2) < \frac{\partial V_{\underline{a}}}{\partial x_2}(x_1, x_2) \quad \Rightarrow \quad \frac{\partial V_{\underline{a}}}{\partial x_1}(x_1 + \Delta, x_2) < \frac{\partial V_{\underline{a}}}{\partial x_2}(x_1 + \Delta, x_2), \quad \Delta > 0 \tag{4.8.3}$$

C5: The line corresponding to $\underline{a} = (1, 1, 0)$ is to the left of the lines corresponding to $\underline{a} = (1, 1, 1)$ and $\underline{a} = (1, 0, 0)$. Equivalently,

$$\frac{\partial V_{111}}{\partial x_1}(x_1, x_2) > \frac{\partial V_{111}}{\partial x_2}(x_1, x_2) \quad \Rightarrow \quad \frac{\partial V_{110}}{\partial x_1}(x_1, x_2) > \frac{\partial V_{110}}{\partial x_2}(x_1, x_2) \tag{4.8.4}$$

$$\frac{\partial V_{100}}{\partial x_1}(x_1, x_2) > \frac{\partial V_{100}}{\partial x_2}(x_1, x_2) \quad \Rightarrow \quad \frac{\partial V_{110}}{\partial x_1}(x_1, x_2) > \frac{\partial V_{110}}{\partial x_2}(x_1, x_2) \tag{4.8.5}$$

C6: The line corresponding to $\underline{a} = (1, 0, 1)$ is to the right of the lines corresponding to $\underline{a} = (1, 1, 1)$ and $\underline{a} = (1, 0, 0)$. Equivalently,

$$\frac{\partial V_{111}}{\partial x_1}(x_1, x_2) < \frac{\partial V_{111}}{\partial x_2}(x_1, x_2) \quad \Rightarrow \quad \frac{\partial V_{101}}{\partial x_1}(x_1, x_2) < \frac{\partial V_{101}}{\partial x_2}(x_1, x_2) \tag{4.8.6}$$

$$\frac{\partial V_{111}}{\partial x_1}(x_1, x_2) > \frac{\partial V_{111}}{\partial x_2}(x_1, x_2) \quad \Rightarrow \quad \frac{\partial V_{101}}{\partial x_1}(x_1, x_2) > \frac{\partial V_{101}}{\partial x_2}(x_1, x_2) \tag{4.8.7}$$

4.9. An Iterative Optimizing Algorithm.

It should have become clear by now that no analytical metthod is likely to yield an optimal control law in a straightforward manner. The optimality conditions that we have derived concern the maximization over a space of admissible controll laws, which is a space without much of an analytical structure. Therefore, we have to use some iterative algorithm.

The main idea behind many algorithms used in optimization problems is the following: Consider a set of optimality conditions. Consider, then, some control law; find at which points the optimality conditions are violated and, at those points, modify the control law so that they are violated less. We propose an algorithm following these lines and relying on the optimality conditions of section 4.2. This algorithm is the stochastic version of an algorithm proposed by Kelley (1962).

Algorithm:

Start with an arbitrary admissible control law u_0. Given a control law u_n and the resulting Markov process $s_n(t)$:

a) Evaluate the operator of the process $\mathcal{L}^n$.

b) Solve the equation

$$\mathcal{L}^n V^n = -k^{u_n} + g^n$$

for the unknowns V^n and g^n.

c) Define a new control law u_{n+1}, as the control law which satisfies, at all points of the state space

$$(\mathcal{L}^{n+1} V^n + k^{u_{n+1}})(s) = \max(\mathcal{L} V^n + k) \qquad (4.9.1)$$

the maximum taken over all $(\mathcal{L}, k)$ corresponding to an admissible control law (such that V_n is in the domain of $\mathcal{L}$).

The last step of the algorithm (step c) is not so hard as it may seem. When $k(\underline{x}, \underline{a})$ is control independent in the interior, step (c) reduces (in the interior) to the evaluation and comparison of $\partial V/\partial x_1$ and $\partial V/\partial x_2$. Some more care is needed at boundary points where optimality conditions assume slightly different forms (section 4.6) but the computation

needed is still trivial. However, the overall computational requirements of the algorithm are significant because each iteration involves the solution of the P.D.E.'s for the value functions in step (b).

A problem with this algorithm arises if the control law u^{n+1} is not well-defined, i.e. if it does not give rise to piecewise continuously differentiable paths. We do not think that such a situation will ever arise in a practical implementation of the algorithm. However, if we attempt to prove the convergence of the algorithm, this issue has to be taken into account. We circumvent this difficulty by just assuming it will not occur.

4.10. Convergence of the Algorithm.

The algorithm of the last section can be seen as a continuous time analog of the Howard algorithm for discrete problems. *(Howard, 1960)* Unlike the discrete case, a proof of the convergence of the algorithm encounters many mathematical difficulties. Even if we initialize the algorithm with a "smooth" control law, there is no guarantee that the algorithm will generate at subsequent stages equally "smooth" control laws.

We prove the convergence of the algorithm conditionally on the assumption that there is a class of smooth initial control laws such that all control laws generated at some stage of the algorithm belong to that class. This result will become useful provided that one is able to:

(i) identify such a class of control laws (in particular, show that such a class exists)

(ii) show that this class is, in some sense, dense in the set of control laws. Then, no performance is lost by restricting our attention to that class.

We view a single stage of the algorithm as an operator $\mathcal{F}:U_0 \mapsto U_0$. If $u \in U_0$, then $\mathcal{F}(u)$ is the control law generated after one step of the algorithm, provided that u was the control law that initialized the algorithm.

Theorem 4.4: Suppose that there exists a set of admissible control laws $U \subseteq U_0$ such that:

$$(i) \qquad \sup_{u \in U} g^u = \sup_{u \in U_0} g^u \tag{4.10.1}$$

(ii) $\mathcal{F}(u)$ exists and belongs in U, $\forall u \in U$

(iii) V^u is in the domain of $\mathcal{L}^w$, $\forall u, w \in U$.

Let

$$u^* \equiv \mathcal{F}(u) \tag{4.10.2}$$

and

$$h^u(s) \equiv (\mathcal{L}^{u^*}V^u + k^{u^*} - \mathcal{L}^u V^u - k^u)(s) \tag{4.10.3}$$

and suppose that

$$(iv) \qquad \sup_{s \in S} h^u(s) = H^u < \infty, \quad \forall u \in U \tag{4.10.4}$$

(v) There exists a continuous function $f:R \mapsto R^+$ such that $f(d) \neq 0$, if $d \neq 0$ and

$$E[l(\{t \in [T_1, T_2]: h^u(s^{u^*}(t)) > \frac{H^u}{2}\})] > f(H^u) \tag{4.10.5}$$

for all $u \in U$. ($l(.)$ is the Borel measure)

Then, if $u_0 \in U$

$$\lim_{n \to \infty} g^{\mathcal{F}^n(u_0)} = \sup_{u \in U_0} g^u \tag{4.10.6}$$

Before proceeding to the proof of the theorem we discuss assumptions (i) to (iv).

Assumption (i) merely states that the best achievable performance is not deteriorated if we restrict our attention to U, some restricted set of smooth admissible control laws.

Assumption (ii) states that if the algorithm is initialized in U, then it produces meaningful control laws which also belong in U. We need this assumption in order to be able to exploit the nice properties of U at all stages of the algorithm.

We need assumption (iii) in order to consider expressions involving $\mathcal{L}^w V^u$. The discussion of section 4.3 implies that if we take $U \subseteq U_V$, then this assumption is automatically satisfied.

Assumption (iv) essentially says that the partial derivatives of V^u are bounded. Again, if we take $U \subseteq U_V$, this condition is automatically satisfied.

Assumption (v) is the most esential, and together with assumption (ii), the most restrictive one. Recall the discussion in section 4.2 that precedes Theorem 4.2. The value of the function $h^u(s)$ at some point $s \in S$, is equal to the instantaneous improvement to control law u, if control law u^* is used at that point. The expression in the left-hand side of (4.10.5) measures for how much time the instantaneous improvement is significant. Assumption (v) means that the time during which a significant improvement is incurred is bounded below uniformly for all control laws which have the same value of H^u. If the assumptions of Theorem 3.2 hold, assumption (v) is satisfied if V^u has uniformly bounded second-order partial derivatives, for all $u \in U$ but this condition is not necessary.

Proof: By assumption (ii) $\mathcal{F}^n(u_o) \in U$, $\forall n$. By Theorem 4.2

$$g^{\mathcal{F}^n(u_0)} - g^{\mathcal{F}^{n-1}(u_o)} > \frac{H^{\mathcal{F}^n(u_o)}}{2} f(H^{\mathcal{F}^n(u_o)}) \qquad (4.10.7)$$

The sequence $g^{\mathcal{F}^n(u_o)}$ is increasing and bounded, therefore convergent. So,

$$\lim_{n\to\infty} H^{\mathcal{F}^n(u_o)} f(H^{\mathcal{F}^n(u_o)}) = 0 \quad \Rightarrow \tag{4.10.8}$$

$$\lim_{n\to\infty} H^{\mathcal{F}^n(u_o)} = 0 \tag{4.10.9}$$

From the definition of the algorithm,

$$\sup_{s\in S} \sup_{w\in U} \left[\mathcal{L}^w V^{\mathcal{F}^n(u_o)} + k^w - \mathcal{L}^{\mathcal{F}^n(u_o)} V^{\mathcal{F}^n(u_o)} - k^{\mathcal{F}^n(u_o)} \right](s) = H^{\mathcal{F}^n(u_o)} \tag{4.10.10}$$

Therefore,

$$\sup_{s\in S} \left[\mathcal{L}^w V^{\mathcal{F}^n(u_o)} + k^w - \mathcal{L}^{\mathcal{F}^n(u_o)} V^{\mathcal{F}^n(u_o)} - k^{\mathcal{F}^n(u_o)} \right](s) \leq H^{\mathcal{F}^n(u_o)} \quad \forall w \in U \tag{4.10.11}$$

Using the notation of Theorem 4.2, we have:

$$\phi((H^{\mathcal{F}^n(u_o)}, \infty)) = 0 \tag{4.10.12}$$

Then, Theorem 4.2 implies

$$g^w \leq g^{\mathcal{F}^n(u_o)} + \frac{H^{\mathcal{F}^n(u_o)}}{E[T_2^w - T_1^w]} \quad \forall w \in U \tag{4.10.13}$$

Now, let $n\to\infty$. Using (4.10.9) and (4.10.13) we obtain,

$$g^w \leq \lim_{n\to\infty} g^{\mathcal{F}^n(u_o)} \tag{4.10.15}$$

Then, by assumption (i),

$$\sup_{w \in U_o} g^w = \lim_{n \to \infty} g^{\mathcal{F}^n(u_o)} \tag{4.10.16}$$

■

4.11 Mathematical Difficulties and the Modelling Issue.

In the previous sections various mathematical difficulties are encountered which are related to the smoothness of the value functions. Special assumptions were introduced, wherever needed, to circumvent these difficulties.

On the other hand it is well-known that the optimality theorems and the convergence of the algorithm have been proved in a discrete time framework. (The algorithm of section 4.9 is just a continuous version of the Howard algorithm.) Since our continuous model is primarily meant to represent a discrete system, all techniques valid for discrete processes should be appliccable (with minor adjustments) to the continuous model.

This discrepancy between the discrete and the continuous formulation has the following explanantion. All difficulties that have appeared have no analogs in discrete systems. So, as long as the control laws considered are good representations of discrete control laws these difficulties will not arise. Very singular and highly discontinuous behaviors of value functions can only be associated with nonphysical systems. Therefore we should not have to worry about these problems.

However, the set of physically meaningful control laws cannot be easily characterized nor is it a closed or complete set. Therefore, a mathematical treatment has to take explicitly into account pathological cases.

CHAPTER 5: RELATIONSHIP BETWEEN THE ROUTING PROBLEM AND DETERMINISTIC OPTIMAL CONTROL

The optimality conditions of the previous chapter were in some sense local. That is, they checked whether a control law is optimal at some point by considering the value function in the vicinity of that point. They were based on the following formulation of the Dynamic Programming Principle: "For a control law to be optimal, it has to be optimal during a very small time increment, assuming it will be optimal later on."

In this chapter we use the following form of the Dynamic Programming Principle: "For a control law to be optimal, it has to be optimal until the first machine failure or repair, assuming it will be optimal later on." Because the motion of the system is deterministic until the first machine repair or failure we may conclude that an optimal control law is also the solution of a certain *deterministic* optimal control problem. We present an iterative algorithm that solves the stochastic control problem by solving the deterministic problem at every iteration.

The fact that a certain deterministic optimal control problem is associated with the stochastic control problem considered in this thesis has been observed in *(Rishel, 1975)*.

5.1. The Associated Deterministic Problem.

We present the deterministic optimal control problem associated with our routing problem. There is one such deterministic problem for each value of $\underline{\alpha}$ for which $\alpha_0 = 1$. For the sake of simplicity we only present the problem corresponding to the choice $\underline{\alpha} = (1, 1, 1)$.

Let T be the first time at which a machine failure occurs. We may break down the expected future reward in two parts:

a) Expected reward until time T.

b) Expected reward after time T.

Suppose that a control law has been fixed. The process is then detreministic until time T. Therefore,

$$\int_0^T k(s(\tau))\,d\tau = f(T) \tag{5.1.1}$$

where f is a deterministic function of T.

The reward after time T and until some time t, which is assumed to be very large, is approximately equal to $V(s(T)) + g(t-T)$. The dynamic programming principle implies that

$$E[f(T) + V(s(T)) - gT] \tag{5.1.2}$$

has to be maximized.

We now identify the quantities appearing in the expression to be maximized. We need first the probability distribution of the random variable T. Let

$$Q(s(t)) \equiv p_0{}^*\frac{\lambda(s(t))}{\lambda^*} + p_1{}^*\frac{\mu_1(s(t))}{\mu_1{}^*} + p_2{}^*\frac{\mu_2(s(t))}{\mu_2{}^*} \tag{5.1.3}$$

Then, $Q(s(t))\Delta$ is the probability of a machine failure during the time interval $[t, t+\Delta]$ and

$$Pr(T \geq \tau) = \exp(-\int_0^\tau Q(s(t))\,dt) \tag{5.1.4}$$

where T is the time of the first machine failure. Moreover,

$$Pr(\underline{a}(T) = (0,1,1) \mid T = t) = \frac{p_0{}^*}{Q(t)} \frac{\lambda(s(t))}{\lambda^*} \tag{5.1.5}$$

$$Pr(\underline{a}(T) = (1,0,1) \mid T = t) = \frac{p_1{}^*}{Q(t)} \frac{\mu_1(s(t))}{\mu_1{}^*} \tag{5.1.6}$$

$$Pr(\underline{a}(T) = (1,1,0) \mid T = t) = \frac{p_1{}^*}{Q(t)} \frac{\mu_2(s(t))}{\mu_2{}^*} \tag{5.1.7}$$

Finally, the probability density function of T is

$$Q(s(\tau)) \exp(-\int_0^\tau Q(s(t))\, dt \tag{5.1.8}$$

Combining equations (5.1.5) to (5.1.8) $E[V(s(T))]$ can be written as

$$\int_0^\infty \exp\left(-\int_0^t Q(s(\tau))\, d\tau\right) \left[p_0{}^* \frac{\lambda(s(t))}{\lambda^*} V_{0111}(s(t)) + p_1{}^* \frac{\mu_1(s(t))}{\mu_1{}^*} V_{101}(s(t)) + p_2{}^* \frac{\mu_2(s(t))}{\mu_2{}^*} V_{110}(s(t))\right] dt$$

Then, recalling equation (5.1.2), we must maximize

$$J \equiv E[f(T) + V(s(T)) - gT] = \int_0^\infty [k(s(t)) - g]\, dt + \int_0^\infty Q(s(t)) \exp(-\int_0^\tau Q(s(u))\, du) +$$

$$\int_0^\infty \exp\left(-\int_0^t Q(s(\tau))\, d\tau\right) \left[p_0{}^* \frac{\lambda(s(t))}{\lambda^*} V_{0111}(s(t)) + p_1{}^* \frac{\mu_1(s(t))}{\mu_1{}^*} V_{101}(s(t)) + p_2{}^* \frac{\mu_2(s(t))}{\mu_2{}^*} V_{110}(s(t))\right] dt \tag{5.1.9}$$

subject to:

$$s(t) = (x_1(t), x_2(t)) \tag{5.1.10}$$

$$0 \leq \lambda \leq \lambda^*, \quad 0 \leq \mu_i \leq \mu_i^* \quad i = 1, 2 \tag{5.1.11}$$

$$\lambda = \lambda_1 + \lambda_2, \quad \lambda_i \geq 0 \quad i = 1, 2 \tag{5.1.12}$$

$$dx_i = (\lambda_i - \mu_i)dt, \quad 0 \leq x_i \leq N_i, \quad i = 1, 2 \tag{5.1.13}$$

If u is an optimal admissible control law and V is the corresponding value function, the discussion in the beginning of this section shows that u maximizes J for every initial state. Conversely, if u maximizes J and also solves the deterministic control problems corresponding to other choices of $\underline{\alpha}$ then u is optimal.

Recall that an optimal control problem has to be solved for every initial state. Let $J^*(s)$ be the optimal value of J as a function of the initial state. If V is the value function corresponding to an optimal control law then it is not hard to show that $J^*(s) = V_{111}(s) + \text{constant}$.

5.2. An Iterative Algorithm.

We now present an optimizing algorithm that is based on the discussion of the previous section:

a) Given (at the n-th stage) a control law u^n, evaluate the corresponding value function, to be denoted by V^n.

b) Using V^n as an argument in the expression (5.1.9) for J, maximize J for all initial states and obtain a new control law u^{n+1}. This is to be done for all four choices of $\underline{a}$ in which $a_0 = 1$.

The control law u_{n+1} so obtained has an appealling interpretation. If we were restricted to use control law u_n after the first machine failure or repair, but were free to use any control law until that time, then u^{n+1} would be an optimal choice of a control law during this initial (random) interval. In this respect this algorithm is very similar to the successive approximation algorithm for discrete problems *(Bertsekas, 1976)* except that here the "time step" is a random interval.

We note that this algorithm has more computational requirements than the algorithm of Section 4.8. Not only do the value functions have to be computed at each iteratiion, but a nontrivial infinite time horizon deterministic optimal control problem has to be solved as well.

Fortunately, the second part can be considerably simplified. For example, the structure of our system, together with constraints CT1, implies that if no failure or repair occurs until some finite time t_0, then the system will reach a known state and stay there. In this way, the infinite time horizon problem can be very easily transformed to a finite time horizon problem. If in addition this terminal state is known beforehand, than the problem is equivalent to finding a minimum cost path from a an initial to a terminal state. This latter problem can be solved quite efficiently by discretizing the state space to a mesh of points and applying any of the standard algorithms for solving minimum

cost (shortest path) problems.

We have not completed the proof of the convergence of this algorithm because we feel that the main issue of interest is whether the algorithm is amenable to computational implementation or its computational complexity is prohibitive. The proof encounters a lot of technical points but the general steps are the same as in the proof of the algorithm of Section 4.8 (although the assumptions needed are much less restrictive than those introduced in section 4.10). Namely, we have to show that in each stage of the algorithm the average performance of the system is improved by an amount proportional to the distance from the optimal performance. This guarantees an exponential rate of convergence.

5.3. The Maximum Principle Applied to the Deterministic Problem.

In this section we attempt to rederive the characterization of optimal control laws derived in Chapter 4 using Pontryagin's Maximum Principle *Athans and Falb,1966*

By observing the expression (5.1.9) for J, we note that we have to minimize

$$\int_0^\infty L(s,u,V(s))\,dt \tag{5.3.1}$$

where L is a function of the state s, of the controls $u = (\lambda(s), \mu_1(s), \mu_2(s))$, and of the value function $V(s)$. Under constraints CT1, there is no freedom in choosing λ,μ_1 and μ_2 and therefore L does not depend explicitly on the control. We form the Hamiltonian:

$$H = L(s,V(s)) - h^T\,ds/dt \tag{5.3.2}$$

where h is the costate vector and ^{T} denotes transposition. We note that $ds/dt = ((\lambda_1 - \mu_1), (\lambda_2 - \mu_2))$, so:

$$h^T\, ds/dt = h_1(\lambda_1 - \mu_1) + h_2(\lambda_2 - \mu_2) \tag{5.3.3}$$

Since L is independent of the control, maximization of the Hamiltonian implies that $h_1\lambda_1 + h_2\lambda_2$ is maximized in the interior, because there we have $\mu_i = a_i\mu_i^*$ =fixed. Therefore, wherever the costates h_1 and h_2 are well-defined and unequal either $(\lambda_1, \lambda_2) = (\lambda^*, 0)$ or $(\lambda_1, \lambda_2) = (0, \lambda^*)$ which is exactly the result of Chapter 4.

Unfortunately, however, the full power of the maximum principle cannot be used in this problem. This is because on the dividing lines $\partial V/\partial x_i$ need not be continuous (and as a result $\partial L_i/\partial x_i$ need not be continuous) and so, a major assumption of the maximum principle breaks down. Nevertheless, we have verified our previous result that in the regions where $\partial V_i/\partial x_i$ exist, an optimal control law is a bang-bang control. Moreover, the costate varibles h_i may be identified with $\partial V_i/\partial x_i$, at all points

CHAPTER 6: PERFECTLY RELIABLE LEAD MACHINE.

In this chapter we consider the problem of maximizing the average production rate

$$k(\underline{x}, \underline{\alpha}) = c_1\mu_1 + c_2\mu_2 \tag{6.0.1}$$

when the lead machine is perfectly reliable. The main result of this section is the following: Suppose that $\lambda^* < \mu_1^* + \mu_2^*$, $\lambda^* > \mu_1^*$ and $\lambda^* > \mu_2^*$ and that constraints CT2 have been imposed. Then, any two control laws which are the same when both buffers are empty and downstream machines are up, have the same performance. So, the optimization problem is reduced to finding the optimal policy just at the origin of the state space, which is very easy.

This result is of interest for two reasons:

a) It indicates that as the lead machine becomes more and more reliable, less care needs to be taken about finding a good dynamic control law. In the limit almost all reasonable dynamic control laws are equivalent.

b) It displays some of the merits of a continuous formulation. The result is obtained by means of a rather simple proof which would be much harder, if not impossible, in a discrete framework. In fact, this result may only be approximately true in the case studied by Hahne (1980).

6.1. Simplification of the Model

Throughout this chapter we assume

a) The lead machine never fails.

b) Constraints CT1 and CT2.

c) $\lambda^* < \mu_1^* + \mu_2^*$ Otherwise, both buffers would be always full and the control problem would be trivial.

d) $\lambda^* \geq \mu_1^*$, $\lambda^* \geq \mu_2^*$. This assumption says that the lead machine is faster than each of the downstream machines. Although this is usually the case in real world systems, it does not always have to be so. However, this assumption is indispensable for the validity of our result.

Consider the stochastic process

$$((x_1 + x_2)(t), \underline{\alpha}(t)) \equiv (y(t), \underline{\alpha}(t)) \tag{6.1.1}$$

We have

$$dx_i = (\lambda_i - \mu_i)\,dt \tag{6.1.2}$$

so that

$$dy = (\lambda - \mu_1 - \mu_2)\,dt \tag{6.1.3}$$

It follows from assumptions (c) and (d) that λ, μ_1, and μ_2 are uniquely determined by $\underline{\alpha}$ at all points other than $(y, \underline{\alpha}) = (0, (1, 1))$. Namely,

$$\mu_i(\underline{x}, \underline{\alpha}) = \alpha_i \mu_i^* \tag{6.1.4}$$

$$\lambda(\underline{x}, \underline{\alpha}) = \lambda^* \qquad \text{if } \underline{x} \neq (N_1, N_2) \tag{6.1.5}$$

$$\lambda(\underline{x}, \underline{a}) = \min\{\lambda^*, a_1\mu_1^* + a_2\mu_2^*\} \qquad \text{if } \underline{x} = (N_1, N_2) \tag{6.1.5}$$

There is only freedom in choosing μ_1, μ_2 at the point $(\underline{x}, \underline{a}) = ((0,0),(1,1))$ which is the same as the point $(y, \underline{a}) = (0,(1,1))$. So, μ_1, μ_2 and λ can be viewed as functions of $(y, \underline{a})$ instead of functions of $(\underline{x}, \underline{a})$. Moreover, the transition rates of the processes $a_1(t)$, $a_2(t)$ depend only on μ_1, μ_2, λ, therefore, these rates are also functions $(y, \underline{a})$. The conclusion is that the statistical behavior of the process $(y, \underline{a})(t)$ is completely determined by the value of μ_1 at the point $(y, \underline{a}) = (0,(1,1))$ (which is the same as the value of λ_1 at that point). So, the reduction of the state variables from $(\underline{x}, \underline{a})$ to $(y, \underline{a})$ leads to no loss of information concerning the stochastic process $(y, \underline{a})(t)$. This is another way of saying that $(y, \underline{a})(t)$ is itself a Markov process. Since μ_1 and μ_2 are also functions of $(y, \underline{a})$ we deduce that the statistics of the performance functional are also completely determined by the value of λ_1 at the point $(y, \underline{a}) = (0,(1,1))$. Therefore, any two control laws such that

$$\lambda_1^{(1)}((0,0),(1,1)) = \lambda_1^{(2)}((0,0),(1,1)) \tag{6.1.6}$$

have the same expected performance. So, the optimization problem is reduced to that of an optimal choice of $\lambda_1((0,0),(1,1))$.

We note that this demonstration is independent of the specific choice of the performance function in equation 6.0.6. Any performance function $k(\underline{x}, \underline{a})$ which depends on $\underline{x}$ only through $x_1 + x_2$ would lead to the same conclusion.

The optimality conditions of section 4.5 imply that the optimal choice of $\lambda_1((0,0),(1,1))$ will be either $\lambda_1 = \mu_1^*$ or $\lambda_1 = \lambda^* - \mu_2^*$. Then, in order to find the optimal control law we only need to solve the differential equations for the value functions for these two

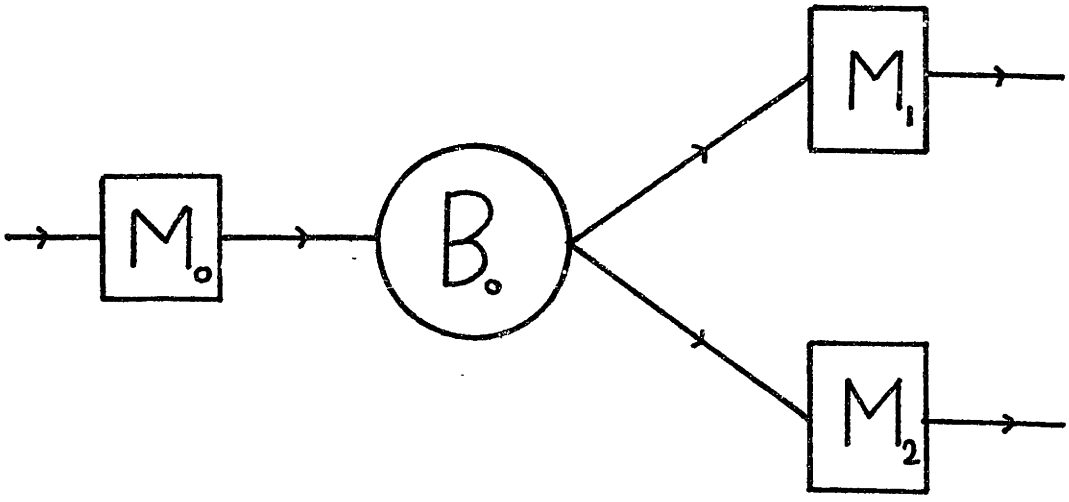


Figure 6.1: Reliable Lead Machine: Equivalent Network

control laws, obtain the values of the performance functional for each of them and simply compare them.

In the next section we exploit the fact that the performance functional depends only on the statistics of the one-dimensional Markov process $(y, \underline{\alpha})(t)$ and reduce the PDE's to ordinary differential equations. This reduction of the dimension of the state space has the following interpretation: under the assumptions made in the beginning of the chapter the network studied is equivalent to the network in Figure 6.1 which has only one shared buffer. We emphasize that this result depends crucially on the assumption that constraints CT2 have been imposed.

6.2. Solution of the Optimization Problem.

One of the conclusions of the last section is that that the value functions can also be viewed as functionsof $(y, \underline{\alpha})$ instead of functions of $(\underline{x}, \underline{\alpha})$. The infinitesimal operator corresponding to the process $(y, \underline{\alpha})(t)$ is an ordinary (not partial) differential operator. So, the solution of the equation

$$\mathcal{L}V = -k + g \tag{6.2.1}$$

is very easy. We can solve this equation for each of the two candidate optimal control laws and then compare the values of g obtained.

We now write down the equation 6.2.1 explicitly:

$$(\mu_1{}^* + \mu_2{}^* - \lambda^*)\frac{dV_{11}}{dy} = -(p_1 + p_2)V_{11} + p_1V_{01} + p_2V_{10} + c_1\mu_1{}^* + c_2\mu_2{}^* - g \tag{6.2.2}$$

$$(\mu_1{}^* - \lambda^*)\frac{dV_{10}}{dy} = -(p_1 + r_2)V_{10} + p_1V_{00} + r_2V_{11} + c_1\mu_1{}^* - g \tag{6.2.3}$$

$$(\mu_2{}^* - \lambda^*)\frac{dV_{01}}{dy} = -(r_1 + p_2)V_{01} + r_1 V_{11} + p_2 V_{00} + c_2\mu_2{}^* - g \qquad (6.2.4)$$

$$-\lambda^*\frac{dV_{00}}{dy} = -(r_1 + r_2)V_{00} + r_1 V_{10} + r_2 V_{10} - g \qquad (6.2.5)$$

Boundary Conditions:

At $y = 0$:

$$p_1\frac{\lambda_1(0)}{\mu_1{}^*}V_{01} + p_2\frac{\lambda_2(0)}{\mu_2{}^*}V_{10} - \left(p_1\frac{\lambda_1}{\mu_1{}^*} + p_2\frac{\lambda_2}{\mu_2{}^*}\right)V_{11} + c_1\lambda_1(0) + c_2\lambda_2(0) - g = 0 \qquad (6.2.6)$$

At $y = N_1 + N_2$:

$$p_1 V_{00} + r_2 V_{11} - (p_1 + r_2)V_{10} + c_1\mu_1{}^* - g = 0 \qquad (6.2.7)$$

$$p_2 V_{00} + r_1 V_{11} - (p_2 + r_1)V_{10} + c_2\mu_2{}^* - g = 0 \qquad (6.2.8)$$

$$r_1 V_{10} + r_2 V_{01} - (r_1 + r_2)V_{00} - g = 0 \qquad (6.2.9)$$

These equations determine V up to an arbitrary constant. We can set a reference point ($V_{11}(0) = 0$ for example) and proceed to the solution.

6.3. Numerical Results.

The differential equations in section 6.2 were solved numerically for a few particular cases. The results are reported in Appendix III. The conclusion that can be drawn (if any) is that the difference in performance of the two alternative control laws we are considering is very small (of the order of two to five percent) even when the availabilities $r_i/(r_i + p_i)$ of the machines are very different and even when the output of one machine is weighted twice as much as the output of the other machine in the cost functional.

6.4. Extension to Lead Machines that Change State Very Fast.

Suppose that the lead machine, instead of being reliable, does fail but its failure and repair rates are very large. When the expected time between two state transitions is much smaller than all the other time constants characterizing the system (average times to repair or failure of a downstream machine, average tiem to fill an empty buffer etc.) then the lead machine may be modelled as a perfectly reliable lead machine with maximum allowed flow rate equal to $\lambda^* r_0/(r_0 + p_0)$ where r_0 is the repair rate and p_0 is the failure rate of the lead machine.

In that case the results of the previous section are appliccable provided that the assumptions stated in section 6.1 are satisfied. Suppose that CT2 has been imposed and that

$$\mu_i^* < \frac{r_0}{r_0 + p_0}\lambda^* < \mu_1^* + \mu_2^* \qquad i = 1, 2. \tag{6.4.1}$$

Then, the average production rate of the system depends again only on the control law that is being applied when both buffers are empty.

CHAPTER 7: CONCLUSIONS, SUGGESTIONS FOR FURTHER RESEARCH.

In the first three chapters we examined the behavior of a continuous, unreliable queuing system in which a control law has been fixed. We have modelled such a system and developed a methodology for evaluating the average value of the performance criterion. We showed that this requires the solution of a system of linear partial differential equations with non-constant coefficients.

The next chapters dealt with the optimization issue. It became apparent that mathematical intricacies, not encountered in discrete problems, arise in the continuous framework. It is our belief that a thorough understanding of the mathematical issues involved is necessary before any practical solution of the problem or the design of an optimizing algorithm is attempted. One possible direction of further research is therefore to complete the investigation of the mathematical properties of the value functions. One should try to understand the effect of various restrictions on the set of admissible control laws on the properties of the value functions.

On the applied side there are both qualitative and numerical issues that have to be investigated. Qualitative issues concern the additional characterization of optimal control laws. The work of Hahne (1980) has shown that optimal control laws follow certain patterns. (See also section 4.8) One should attempt to prove as many as possible such results.

The main numerical issue that has to be considered concerns the develoment of efficient computational procedures that solve the partial differential equations for the value functions, so that any proposed algorithm can be implemented and tested.

A final direction of future research is the extension of the methodology and the results

of this thesis to other problems. We suggest some possible directions:

1) Extend the methodology to larger networks. The number of differential equations to be solved increases exponentially with the number of machines in the system, so it is not likely that one would ever try to solve these equations for large networks. However, it would be quite interesting to examine the generalization of the optimality conditions to larger networks. These conditions may suggest that the decisions at different points of the network may be decoupled.

2) Since an exact complete solution is unrealistic for a large network, one should examine the performance of various forms of decentralized strategies. It is our belief that any realistic decentralized strategy should be based on a decentralized model in the following sense: Let one decision maker correspond to each point of the system where a decision may be taken. Each such decision maker has an accurate model of the neighboring buffers and machines but a very aggregated amount of information. This decision maker will then try to optimize his behavior subject to his own model. That is, he will behave as if he believed that his model is accurate.

3) Consider the effect of more general distribution of the time to repair and the time to failure. Gamma distributions (that result to finite dimensional Markovian relizations) should be considered at first. Again the primary interest is for the optimality conditions that can be relatively easily obtained, rather than expecting any simple computational procedures. Such an investigation may also answer the question whether the results obtained under the memorylessness assumption are very sensitive to that assumption.

4) The model considered in this thesis took into account only one of the possible sources of randomness in a queuing network: the machine failures and repairs. We should

investigate the possibility of a continuous approximation to a discrete system with random component arrivals, random component processing times and randomized routing strategies. Some preliminary results indicate that as the component size, the interarrival times and the processing times of a single component become arbitrarily small, we will recover, in the limit, the same model that was studied in this thesis. In that case, the continuous approximation is associated with a loss of the "fine structure" of the process which may or may not be desirable. If not, ad-hoc diffusion models may also be considered.

REFERENCES

M. Athans, P. Falb, (1966), *Optimal Control* , N.Y., McGraw Hill, 1966.

R. Bellman, S. Dreyfus, (1962), *Applied Dynamic Programming* , Princeton, 1962.

D.P. Bertsekas, (1976), *Dynamic Programming and Stochastic Control* , Academic, 1976.

D.P. Bertsekas, (1979), Algorithms for Nonlinear Multicommodity Network Flow Problems, Intl. Symposium on Systems Optimization and Analysis, A. Bensoussan and J.T. Lions (eds.), Springer-Verlag, 1979, pp. 210-224.

R. Boel, P. Varaiya, (1977), Optimal Control of Jumap Processes, *SIAM J. Control and Optimization* , Vol. 15, No. 1, Jan. 1977.

J.A. Buzacott, (1971), The Role of Inventory Banks in Flow-Line Production Systems, *Int. J. Prod. Res.* 9, 4, pp. 425-436.

M. Cantor, M. Gerla, (1974), Optimal Routing in a Packet Switched Computer Network, *IEEE Trans. Computers* , Vol. C-23, No. 10, October 1974.

B.L. Deuermeyer, W.P. Pierskalla, (1978), A By-Product Production System with an Alternative, *Management Science* , Vol. 24, No. 13, September 1978.

E.B. Dynkin, (1965), *Markov Processes* , Academic, N.Y., 1965

A. Ephremides, P. Varaiya, J. Warland (1980), A Simple Dynamic Routing Problem, *IEEE Trans. on Automatic Control* , Vol. AC-25, No. 4, August 1980.

G.J. Foschini, (1977), On Heavy Traffic Diffusion Analysis and Dynamic Routing in Packet Switched Networks, *Computer Performance* , K.M. Chandy and M. Reiser (eds.), North-Holland, 1977.

G.J. Foschini, J. Salz, (1978), A Basic Dynamic Routing Problem and Diffusion, *IEEE Trans. on Communications* , Vol. COM-26, No. 3, March 1978.

R.G. Gallager, (1977), A Minimum Delay Routing Algorithm Using Distributed Computation, *IEEE Trans. on Communications* , Vol. COM-25, No.1, Jan. 1977

E.M. Gafni, (1979), Convergence of a Routing Algorithm, Massachusetts Institute of Technology, Laboratory for Information and Decision Systems, Report LIDS-R-907, May 1979.

S.B. Gershwin, M. Ammar, (1979), "Reliability in Flexible Manufacturing Systems", *Proceedings of the 18th IEEE Conference on Decision and Control* , Vol. 1, p. 540, December 1979.

S.B. Gershwin, I.C. Schick, (1979), "Analysis of Transfer Lines Consisting of Three Unreliable Machines and Two Finite Storage Buffers", Massachusetts Institute of Technology, Electronic Systems Laboratory, Report ESL-FR-834-9, August 1979.

S.B. Gershwin, I.C. Schick, (1980), "Continuous Model of an Unreliable Two-Stage Material Flow System with a Finite Interstage Buffer", Massachusetts Institute of Technology, Laboratory for Information and Decision Systems, Report LIDS-R-1039, September 1980.

E. Hahne, (1980), "Dynamic Routing in an Unreliable Manufacturing Network with Limited Storage", Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, M.Sc. Thesis (in preparation).

P.R. Halmos, (1974), *"Measure Theory"* , Springer Verlag, 1974.

R.A. Howard, (1960), *"Dynamic Programming and Markov Processes"* , MIT Press, Cambridge, Mass., 1960.

H. Kelley, (1962), "Method of Gradients", in *Optimization Techniques with Applications to Aerospace Systems* , G. Leitman (ed.), Academic, N.Y., 1962.

L. Kleinrock, (1975), *"Queuing Systems"* , J. Wiley, 1975.

E. Koenigsberg, (1959), "Production Lines and Internal Storage – A Review", *Management Science* , 5, pp. 410–433, 1959.

H. Kushner, (1965), "On the Stochastic Maximum Principle: Fixed Time Control", *J. Math. Anal. Appl.* , 11, pp.78–92.

H. Kushner, (1967), *"Stochastic Stability and Control"* , Academic, N.Y., 1967.

H. Kushner, (1978), "Optimality Conditions for the Average Cost per Unit Time Problem with a Diffusion Model", *SIAM J. on Cont. and Opt.*, Vol. 16, No. 2, March 1978.

M. Loeve, (1977), *"Probability Theory"* , Springer Verlag, 1977.

P.A. Meyer, (1966), "*Probability and Potentials* ", Blaisdell, Waltham, Mass., 1966.

P.A. Meyer, (1967), "Intégrales Stochastiques" (4 Exposés) *Séminaire de Probabilités I*, Lecture Notes in Mathematics No. 39, Springer Verlag, pp. 72–162, 1967

F. Moss, (1977), "The Application of Optimal Control Theory to Dynamic Routing in Data Communication Networks", Massachusetts Institute of Technology, Electronic Systems Laboratory, Report ESL–R–721, February 1977.

G.J. Olsder, R. Suri, (1980), "Time–Optimal Control of Parts Routing in a Manufacturing System with Failure Prone Machines", *Proceedings of the 19th Conference on Decision and Control* , 1980

R. Rishel, (1970), "Necessary and Sufficient Dynamic Programming Conditions for Continuous Time Optimal Control", *SIAM J. of Control* , 8 (1970), pp. 559–571.

R. Rishel, (1975a), "Dynamic Programming and Minimum Principles for Systems with Jump Markov Disturbances", *SIAM J. Control*, Vol. 13, No. 2, Feb. 1975.

R. Rishel, (1975b), "Control of Systems with Jump Markov Disturbances", *IEEE Trans. Automatic Control*, Vol. AC–20, pp. 241–244, 1975.

R. Rishel, (1977), "Optimality for Completely Observed Controlled Jump Processes", *IEEE Trans. Automatic Control*, No. 6, pp. 906–908, Dec. 1977.

W. Rudin, (1974), *"Real and Complex Analysis"* , McGraw Hill, 1974.

J.L. Santana, L.K. Platzman, (1979), "Real Time Scheduling of an "Open Warehouse" Sorter–Palletizer System", *Proceedings of the 18th IEEE Conference on Decision and Control*, December 1979.

V.V.S. Sarma, M. Alam, (1975), "Optimal Maintenance Policies for Machines Subject to Deterioration and Intermittent Breakdowns", *IEEE Trans. Systems, Man and Cybernetics* , Vol. SMC–4, pp. 396–398, 1975.

I.C. Schick, S.B. Gershwin, (1978), "Modelling and Analysis of Unreliable Transfer Lines with Finite Interstage Buffers", Massachusetts Institute of Technology, Electronic Systems Laboratory, Report ESL–FR–834–6, September 1978.

A. Segall, (1976), "The Modelling of Adaptive Routing in Data–Communication Networks", *IEEE Trans. on Comm., Special Issue on Comp. Comm.*, 1976.

A. Segall, T. Kailath, (1975), "The Modelling of Randomly Modulated Jump Processes, *IEEE Trans. on Information Theory*, Vol. IT–21, No. 2, pp. 135–143, March 1975.

R. Serfozo, S. Stidham, (1978), "Semi–Stationary Clearing Processes", in *Stochastic Processes and their Applications* , 6, North Holland, pp. 165-178.

J.F. Shapiro, (1979), "*Mathematical Programming: Structures and Algorithms* ",J. Wiley, N.Y., 1979.

S. Stidham, (1977), "Cost Models for Stochastic Clearing Systems", *Operations Research* , Vol. 25, No. 1, Jan.–Feb. 1977.

L.D. Stone, (1973), "Necessary and Sufficient Conditions for Optimal Control of Semi–Markov Jump Processes", *SIAM J. of Control* , Vol. 11, No. 2, May 1973.

D.D. Sworder, (1969), "Feedback Control of a Class of Linear Systems with Jump Parameters", *IEEE Trans. on Automatic Control* , Vol. AC–14, No. 1, Feb. 1969.

E. Wong, (1971), "*Stochastic Processes in Information and Dynamical Systems* ", McGraw Hill, 1971.

W.M. Wonham, (1970), "Random Differential Equations in Control Theory", in *Probabilistic Methods in Applied Mathematics* , A.T. Bharucha–Reid (ed.), Academic, N.Y., 1970, pp. 131–212.

APPENDIX I: MATHEMATICAL BACKGROUND

We present here some mathematical definitions and results that are repeatedly used.

Measure Theory:

Definition: A pair $(\Omega, \mathcal{A})$ is called a measurable space if $\mathcal{A}$ is a σ-algebra of subsets of Ω.

Definition: A function f from a measurable space $(\Omega, \mathcal{A})$ to another measurable space $(S, \mathcal{B})$ is called $\mathcal{A}$-measurable if $f^{-1}(B) \in \mathcal{A}$, $\forall B \in \mathcal{B}$.

Dominated Convergence Theorem: Let μ be a finite measure on a measurable space $(\Omega, \mathcal{A})$ and $\{f_n\}$ a sequence of real-valued measurable functions on Ω converging to a measurable function f. Suppose that there exists a measurable function g such that $|f_n| \leq |g|$, $\forall n$ and $\int g\, d\mu < \infty$. Then,

$$\lim_{n\to\infty} \int f_n\, d\mu = \int f\, d\mu \qquad (A.1.1)$$

Fubini Theorem: Let μ and ν be finite measures on the measurable spaces $(X, \mathcal{A})$ and $(Y, \mathcal{B})$, respectively, and let ρ be the product measure. Let $f: X \times Y \mapsto (-\infty, +\infty)$ be absolutely integrable (i.e. $\int_{X\times Y} |f|\, d\rho < \infty$). Then,

$$\int_{X\times Y} f\, d\rho = \int_X \int_Y f\, d\nu\, d\mu = \int_Y \int_X f\, d\mu\, d\nu \qquad (A.1.2)$$

Induced measures: Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces and let f be a measurable function from X into Y. Let μ be a finite measure on $(X, \mathcal{A})$. Then, f induces a finite

measure ν on $(Y, \mathfrak{B})$ such that

$$\nu(B) = \mu(f^{-1}(B)), \quad \forall B \in \mathfrak{B} \tag{A.1.3}$$

Moreover, if g is a real-valued measurable function on $(Y, \mathfrak{B})$, we have

$$\int g\, d\nu = \int g \circ f\, d\mu \tag{A.1.4}$$

where $(g \circ f)(x) \equiv g(f(x))$.

Good references on measure theory are Halmos (1974) and Rudin (1974).

Markov Processes and Associated Operators.

Let $(s(t), t \in [0, \infty))$ be a Markov process defined on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$. ($\Omega$ is the sample space, $\mathcal{A}$ is the σ-algebra of events and $\mathcal{P}$ a probability measure). We assume that the Markov process takes values in some metrizable topological space S. We let the σ-algebra $\mathfrak{B}$ of measurable sets of S be the smallest σ-algebra that contains all open subsets of S. Let $\mathcal{A}_t \subseteq \mathcal{A}$ be the smallest σ-algebra such that the random variable $s(\tau)$ is $\mathcal{A}_t$-measurable $\forall \tau \leq t$. A random variable T is called a stopping time or a Markov time with respect to $\{\mathcal{A}_t\}$ if

$$\{\omega: T(\omega) \leq t\} \in \mathcal{A}_t, \quad \forall t \geq 0 \tag{A.1.5}$$

This means that if we observe $(s(\tau),\ \tau \leq t)$ we can always determine whether $T(\omega) \leq t$ or not.

To every Markov process $s(t)$, with values in a state space S, there corresponds an operator mapping a subset of the Banach space of real, bounded, measurable functions on the space into itself. This operator, together with its domain, characterizes the Markov process completely. (Hille-Yosida theorem.) There are a few different versions of this operator which differ because they do not have the same domain of definition. Each one is an extension of another.

Let $f(s)$ be a real-valued, bounded, measurable function on the state space. The strong infinitesimal operator A of the process is defined by:

$$(Af)(s_o) = \lim_{t \to 0^+} \frac{E[f(s(t)) \mid s(0) = s_o] - f(s_o)}{t} \qquad (A.1.6)$$

The domain of definition of the operator consists of all functions f, whose limit on the above expression exists uniformly with respect to $s_o \in S$.

The weak infinitesimal operator $\mathcal{L}$ is also defined by A.1.6 but on a wider domain. The latter consists of all functions $f(s)$ for which: a) the ratio on the right-hand side of A.1.6 is bounded for all $s \in S$ and all t in some neighborhood of zero; b)the limit of this ratio exists for each $s \in S$ and determines a function h such that $E[h(s(t)) \mid s(0) = s_o]$ converges to $h(s_o)$ for any $s_o \in S$ as t decreases to 0.

A Markov process is called a strong Markov process if

$$E[f(s(t_1))g(s(t_3)) \mid s(t_2)] = E[f(s(t_1)) \mid s(t_2)]\, E[g(s(t_3)) \mid s(t_2)] \qquad (A.1.7)$$

for any three stopping times t_1, t_2, t_3 such that $t_1 \leq t_2 \leq t_3$ almost surely and any real-valued measurable functions f and g. All processes considered in this thesis are strong Markov processes.

The stochastic equivalent of the fundamental theorem of integral calculus is the following theorem:

Dynkin's Theorem: (Dynkin, 1965, p.133) Let $\mathcal{L}$ be the weak infinitesimal operator of a right-continuous strong Markov process $s(t)$ on a topological space and let τ be a stopping time. If $\mathcal{L}f = g$ and $E\tau < \infty$ then

$$E[f(s(\tau))] - f(s(0)) = E\left[\int_0^\tau g(s(t))\,dt \mid s(0)\right] \qquad (A.1.8)$$

We should point out the existence of an extension of the weak operator called "characteristic operator". This operator has many of the useful properties of the weak operator but both its domain and its range contain unbounded functions. This operator is more applicable for stochastic control problems with unbounded state spaces.

The most complete exposition of the modern theory of Markov processes is given by Dynkin (1965).

APPENDIX II

Lemma 3.1: Let

$$s_0 = ((0,0),(0,1,1)) \tag{A.2.1}$$

$$T = 3\max\left(\frac{N_1}{\mu_1^*}, \frac{N_2}{\mu_2^*}\right) \tag{A.2.2}$$

There exists an $\epsilon > 0$ such that for any initial state and any control law satisfying CT1,

$$Pr\big(s(T) = s_0,\ \exists t \in [0,T) s(t) \neq s_0\big) > \epsilon \tag{A.2.3}$$

Proof: We will only sketch a proof of the lemma by presenting, for each initial state, a chain of events that leads to state s_o at time T. This chain of events will have probability bounded away from zero, uniformly for all initial states and all control laws.

If the initial state has $\underline{\alpha}(0) = (1,1,1)$, the lead machine will be operated at a positive rate. Therefore, there is a probability larger than some positive number ϵ_o that it will fail and it will not be repaired until time $T/2$. Then, there is probability larger than $(1-\exp(-p_1 T))(1-\exp(-p_2 T))$ that none of the downstream machines will fail during $[0,T]$. Because of the way that T was selected, both buffers will be empty at time T, if the above described sequence of events occurs. So, there exists a positive number ϵ_{111} such that

$$Pr(((\underline{x},\underline{\alpha})(T) = s_o \,|\, s(0) = (\underline{x},(1,1,1))) > \epsilon_{111} \tag{A.2.4}$$

For all other choices of $\underline{\alpha}(0)$ we only state the corresponding sequences of events and omit the details.

If $\underline{\alpha}(0) = (0,1,1)$ consider the case where M_o is repaired before time $T/3$ then fails again during $[T/3, 2T/3]$ and is not repaired again until time T.

If $\underline{\alpha}(0) = (1,0,1)$ or $\underline{\alpha}(1,0,0)$ or $\underline{\alpha}(0) = (1,1,0)$, consider the case in which downstream machines are repaired until time $T/3$ and do not fail until time T. Then let machine M_o fail during the interval $[T/3, 2T/3]$ without getting repaired until time T.

If $\underline{\alpha}(0) = (0,0,0)$ or $(0,1,0)$ or $(0,0,1)$ consider the case where failed downstream machines are repaired until time $T/3$ and do not fail until time T, and machine M_o is not repaired until time T.

For each possible initial $\underline{\alpha}(0)$, there exists consequently a positive number $\epsilon_{\underline{\alpha}(0)}$ such that

$$Pr(((\underline{x}, \underline{\alpha})(T) = s_o \,|\, \underline{\alpha}(0)) > \epsilon_{\underline{\alpha}(0)} \tag{A.2.5}$$

If we let $\epsilon = \min_{\underline{\alpha}(0)}\{\epsilon_{\underline{\alpha}(0)}\}$, we get the desired result.

Lemma 3.2: Let $U(t)$ be the last time before t that state s_o was reached. Then,

$$\exists N{:}E[t - U(t)] < N \tag{A.2.6}$$

Proof: This proof is just a backward version of the proof of Theorem 3.1, so we only give a summary.

Let T and ϵ be defined as in Lemma 3.1. Lemma 3.1 shows that for any control law and any initial state there is a finite probability that the state of the system at time kT, ($k \in Z^+$, T as defined in Lemma 3.1) is equal to s_o. In particular, this is true for time

k_oT where

$$k_o = \max\{k: kT < t\} \tag{A.2.7}$$

Therefore,

$$Pr((\underline{x}, \underline{\alpha})(k_oT) \neq s_o) < 1 - \epsilon \tag{A.2.8}$$

Similarly,

$$Pr((\underline{x}, \underline{\alpha})(k_oT) \neq s_o, (\underline{x}, \underline{\alpha})((k_o - 1)T)) < (1 - \epsilon)^2 \tag{A.2.9}$$

This process may be continued and a geometric progression is obtained. Let $W(t)$ be the last time before t that the state of the system was s_o. By summing the terms of the geometric progression mentioned above, we have:

$$E[t - W(t)] < M_1 \tag{A.2.10}$$

for some finite M_1. Moreover, since ϵ can be selected independent of the initial state,of t and of the control law, M_1 is also independent of these.

Because of the way that $U(t)$ and $W(t)$ were defined, $U(t) < W(t)$ and the state of the system is identically equal to s_o during the time interval $[U(t), W(t)]$. The time that the state of the system stays equal to s_o is less or equal to the time to repair of one of downstream machines which has a finite expected value. Therefore, there exists some M_2, finite, such that

$$E[W(t) - U(t)] < M_2 \tag{A.2.11}$$

The proof is completed by letting $N = M_1 + M_2$.

APPENDIX III: NUMERICAL RESULTS

As mentioned in Chapter 6, the set of partial differential equations for the value functions becomes a system of ordinary differential equations when the lead machine is perfectly reliable, constraints CT2 have been imposed and $\mu_i^* \leq \lambda^* \leq \mu_1^* + \mu_2^*$. A FORTRAN program was written that solved this system of equations by using the method of undetermined coefficients. We present here the numerical results of a few representative sample runs.

For any given set of numerical values of the parameters of the system there are only two control laws that need to be considered as candidates for optimality (Recall Chapter 6). More precisely, the question is whether we will have $\lambda_1 = \mu_1^*$ or $\lambda_1 = \lambda^* - \mu_2^*$ when the state of the system is $((0,0),(1,1))$. Let g_1 and g_2 be the respective values of the performance criterion under each of the above two control laws. Table A.1 below presents and compares the values of g_1 and g_2 for several values of the parameters of the system. In these examples we have let $c_1 = r_1 = r_2 = 1.0$ and we have varied the other parameters. We also let N stand for $N_1 + N_2$, the sum of the capacities of the two buffers.

The main conclusion that can be drawn from the results in Table A.1 is that the difference in performance between the two control laws considered is negligible except if strong unbalances are introduced, like doubling the worth of the output of one of the machines or increasing the reliability of one of the machines.

Table A.1

Numerical Results: Reliable Lead Machine

p_1	p_2	μ_1*	μ_2*	λ*	c_1	c_2	N	g_1	g_2
.6	.5	3	4	5	1	1	10	4.47	4.47
.6	.5	3	4	5	1	1	1	4.10	4.08
.6	.5	3	4	5	1	1	.2	3.96	3.94
.6	.5	3	4	5	1	2	1	6.32	6.75
.6	.5	3	4	5	2	1	1	5.97	5.50
.2	.1	3	4	5	1	1	10	4.98	4.99
.2	.1	3	4	5	1	1	10	4.82	4.82
3	2	3	4	5	1	1	1	2.08	2.08
2	.2	3	4	5	1	1	1	4.11	4.16
.6	.5	3	4	6.8	1	1	1	4.54	4.54
.6	.5	3	4	4.3	1	1	1	4.02	4.01
.6	.5	1	5	5.5	1	1	1	3.93	3.93
.6	.5	1	5	5.5	1	2	1	7.24	7.26
.6	.5	1	5	5.5	2	1	1	4.56	4.53

APPENDIX IV: COMPLETE SET OF PARTIAL DIFFERENTIAL EQUATIONS AND OPTIMALITY CONDITIONS: $\lambda^* = \mu_1^* + \mu_2^*$

We have presented, in Chapter 4, the partial differential equations for the value functions and the optimality conditions for a few illustrative cases. In this Appendix we present the complete set of equations and optimality conditions for the special case where $\lambda^* = \mu_1^* + \mu_2^*$. This task becomes a little complicated because, not knowing exactly how an optimal control law looks like, we have to consider several different possibilities.

A. $\alpha = (1, 1, 1)$

Upper Half:

$$(\mu_1^* - \lambda^*)\frac{\partial V_{111}}{\partial x_1} + \mu_2^*\frac{\partial V_{111}}{\partial x_2} = -(p_0^* + p_1^* + p_2^*)V_{111} + p_0^* V_{011} + p_1^* V_{101} + p_2^* V_{110} + c_1\mu_1^* + c_2\mu_2^* - g \tag{A.4.1}$$

$$\frac{\partial V_{111}}{\partial x_1} \geq \frac{\partial V_{111}}{\partial x_2} \tag{A.4.2}$$

Lower Half:

$$\mu_1^*\frac{\partial V_{111}}{\partial x_1} + (\mu_2^* - \lambda^*)\frac{\partial V_{111}}{\partial x_2} = -(p_0^* + p_1^* + p_2^*)V_{111} + p_0^* V_{011} + p_1^* V_{101} + p_2^* V_{110} + c_1\mu_1^* + c_2\mu_2^* - g \tag{A.4.3}$$

$$\frac{\partial V_{111}}{\partial x_1} \leq \frac{\partial V_{111}}{\partial x_2} \tag{A.4.4}$$

Dividing Line:

$$0 = -(p_0{}^* + p_1{}^* + p_2{}^*)V_{111} + p_0{}^*V_{011} + p_1{}^*V_{101} + p_2{}^*V_{110} + c_1\mu_1{}^* + c_2\mu_2{}^* - g \quad (A.4.5)$$

No optimality conditions are needed on the dividing line. This is because if optimality conditions are satisfied on the upper and lower half then they are automatically satisfied on the dividing line.

Boundary Equations:

Suppose that the shape of the dividing line is as in Figure A.1. Then:

1) On segments AF and FE, equations (A.4.1) and (A.4.2) hold.

2) On segments BC and CD, equations (A.4.3) and (A.4.4) hold.

3) For any point on segment AB there are three possibilities:

3a) $\lambda_1 = \mu_1{}^*$ and $\lambda_2 = \mu_2{}^*$ In that case equation (A.4.5) holds.

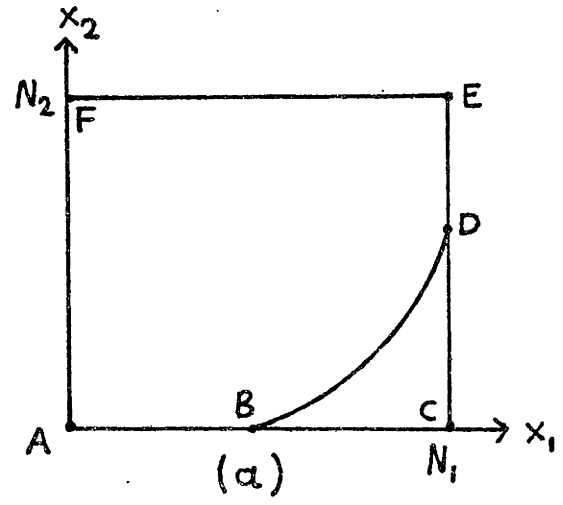
3b) $\lambda_1 = 0$ and $\lambda_2 = \mu_2{}^*$ In that case:

$$\mu_1{}^*\frac{\partial V_{111}}{\partial x_1} = -(\frac{\mu_2{}^*}{\lambda^*}p_o{}^* + p_1{}^* + p_2{}^*)V_{111} + \frac{\mu_2{}^*}{\lambda^*}p_o{}^*V_{011} + p_1{}^*V_{101} + p_2V_{110} + c_1\mu_1{}^* + c_2\mu_2{}^* - g \quad (A.4.6)$$

3c) $\lambda_1 = \lambda^*$ and $\lambda_2 = 0$ In that case:

$$-\mu_2{}^*\frac{\partial V_{111}}{\partial x_1} = -(p_o{}^* + p_1{}^*)V_{111} + p_o{}^*V_{011} + p_1{}^*V_{101} + c_1\mu_1{}^* - g \quad (A.4.7)$$

Which one of the above cases will be the optimal is determined by the optimality conditions. We do not write them explicitly, but they are the following: Gather all terms of



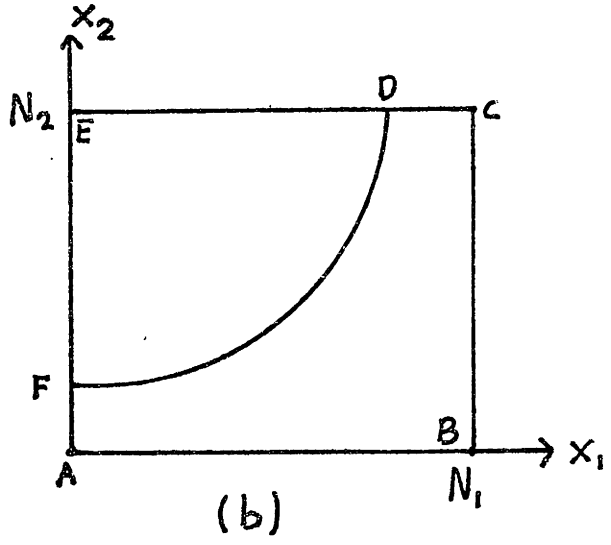


Figure A.1: Possible Dividing Lines

each one of the equations A.4.5, A.4.6, A.4.7 at the right–hand side. An optimal control law corresponds to the largest of the three right–hand sides considered.

4) At point A we have exactly the same three alternatives as in (3) above, except that the corresponding equations are slightly modified.

4a) $\lambda_1 = \mu_1^*$ and $\lambda_2 = \mu_2^*$ In tha case equation (A.4.5) holds.

4b) $\lambda_1 = 0$ and $\lambda_2 = \mu_2^*$ In that case:

$$0 = -(\frac{\mu_2^*}{\lambda^*} p_o^* + p_2^*)V_{111} + \frac{\mu_2^*}{\lambda^*} p_o^* V_{011} + p_2 V_{110} + c_2\mu_2^* - g \qquad (A.4.8)$$

4c) $\lambda_1 = \lambda^*$ and $\lambda_2 = 0$ In that case equation A.4.7 holds.

The optimality conditions at point A can be found in the same way as described in (3) above.

5) At any point on the segment DE we have two alternatives:

5a) $\lambda_1 = \mu_1^*$ and $\lambda_2 = \mu_2^*$ In that case equation (A.4.5) holds.

5b) $\lambda_1 = \mu_1^*$ and $\lambda_2 = 0$ In that case

$$0 = -(\frac{\mu_1^*}{\lambda^*} p_o^* + p_1^*)V_{111} + \frac{\mu_1^*}{\lambda^*} p_o^* V_{011} + p_1 V_{101} + c_1\mu_1^* + c_2\mu_2^* - g \qquad (A.4.9)$$

6) At point D we have exactly the same two alternatives as in (5) above and no change in the equations is needed.

If the dividing line BD had a different shape than that in Figure A.1, only minimal changes would be required to the above discussion. For example, if the dividing line

intercepted the $x_1 = 0$ axis instead of the $x_2 = 0$ axis, everything would remain true except that the subscripts 1 and 2 would have to be interchanged in parts (3), (4) and (5) above.

B $\alpha = (1, 0, 0)$ Upper Half:

$$-\lambda^* \frac{\partial V_{100}}{\partial x_1} = -(p_o^* + r_1 + r_2)V_{100} + p_o^* V_{000} + r_1 V_{110} + r_2 V_{101} - g \qquad (A.4.10)$$

Lower Half:

$$-\lambda^* \frac{\partial V_{100}}{\partial x_2} = -(p_o^* + r_1 + r_2)V_{100} + p_o^* V_{000} + r_1 V_{110} + r_2 V_{101} - g \qquad (A.4.11)$$

Dividing Line:

As shown in Section 4.7, λ_1^* and λ_2^* may be assumed to be known on the dividing line because they are uniquely determined by the slope of the dividing line.

$$-\lambda_1 \frac{\partial V_{100}}{\partial x_1} - \lambda_2 \frac{\partial V_{100}}{\partial x_2} = -(p_o^* + r_1 + r_2)V_{100} + p_o^* V_{000} + r_1 V_{110} + r_2 V_{101} - g \qquad (A.4.12)$$

Boundary Equations:

Assume that the dividing line has the shape shown in Figure 4.1.

1) On the segments AB, AF and FE, equation A.4.11 holds

2) On the segments BC and CD, equation A.4.12 holds.

3) At any point on the segment DE we have two choices:

3a) $\lambda_2 = \lambda^*$ In that case equation A.4.12 still holds.

3b) $\lambda_2 = 0$ (This will certainly be the case at point E.) Then,

$$0 = -(r_1 + r_2)V_{100} + r_1 V_{110} + r_2 V_{101} - g \qquad (A.4.13)$$

Optimality conditions determine whether (3a) or (3b) will be the case. We can again determine the optimality conditions by sending all terms of equations A.4.12 and A.4.13 to the right–hand side and compare them.

C: $a = (1, 1, 0)$

Upper Half:

$$(\mu_1{}^* - \lambda^*)\frac{\partial V_{110}}{\partial x_1} = -(p_0{}^* + p_1{}^* + r_2)V_{110} + p_0{}^* V_{010} + p_1{}^* V_{100} + r_2 V_{111} + c_1\mu_1{}^* - g \qquad (A.4.14)$$

Optimality Condition: Equation (A.4.2) except that the subscripts are changed to 110.

Lower Half:

$$\mu_1{}^*\frac{\partial V_{110}}{\partial x_1} - \lambda^*\frac{\partial V_{110}}{\partial x_2} = -(p_0{}^* + p_1{}^* + r_2)V_{110} + p_0{}^* V_{010} + p_1{}^* V_{100} + r_2 V_{111} + c_1\mu_1{}^* - g \qquad (A.4.15)$$

Optimality Condition: Equation (A.4.4) except that the subscripts are changed to 110.

Dividing Line:

$$(\mu_1{}^* - \lambda_1)\frac{\partial V_{110}}{\partial x_1} - \lambda_2\frac{\partial V_{110}}{\partial x_2} = -(p_0{}^* + p_1{}^* + r_2)V_{110} + p_0{}^* V_{010} + p_1{}^* V_{100} + r_2 V_{111} + \mu_1{}^* c_1 - g \qquad (A.4.16)$$

Boundary Equations:

These may be different depending on whether the dividing line intercepts the $x_1 = 0$ or the $x_2 = 0$ axis and whether it intercepts the $x_1 = N_1$ axis or the $x_2 = N_2$ axis. (Compare figures A.1a and A.1b)

A. 1: Assume that the dividing line intercepts the $x_2 = 0$ axis, as in Figure A.1a.

1) On the segments AF and AB, equation (A.4.14) holds.

2) On the segment BC, equation (A.4.15) holds.

A. 2: Assume that the dividing line intercepts the $x_1 = 0$ axis, as in Figure A.1b.

1) On the segment FE, equation (A.4.14) holds.

2) On the segment AB, equation (A.4.15) holds.

3) On the segment AF we have two choices:

3a) $\lambda_1 = \mu_1^*$ and $\lambda_2 = \mu_2^*$ In that case equation (A.4.16) holds, provided that we use the above values for λ_1 and λ_2.

3b) $\lambda_1 = 0$ and $\lambda_2 = \mu_2^*$ In that case we have:

$$0 = -(\frac{\mu_2^*}{\lambda^*} p_o^* + r_2)V_{110} + \frac{\mu_2^*}{\lambda^*} p_o V_{010} + r_2 V_{111} - g \qquad (A.4.17)$$

B. 1: Assume that the dividing line intercepts the $x_1 = N_1$ axis, as in Figure A.1a.

1) On the segment FE, equation (A.4.14) holds.

2) On the segment CD, equation (A.4.15) holds.

3) On the segment DE, we have two choices:

3a) $\lambda_1 = \mu_1^*$ and $\lambda_2 = 0$ In that case

$$0 = -(\frac{\mu_1^*}{\lambda^*} p_o^* + p_1^* + r_2)V_{110} + \frac{\mu_2^*}{\lambda^*} p_o V_{010} + p_1^* V_{100} + r_2 V_{111} + c_1\mu_1^* - g \qquad (A.4.18)$$

3b) $\lambda_1 = \mu_1^*$ and $\lambda_2 = \mu_2^*$ In that case, equation (A.4.16) holds, provided that we use the above values for λ_1 and λ_2.

B. 2: Assume that the dividing line intercepts the $x_2 = N_2$ axis, as in Figure A.1b.

1) On the segment ED, equation (A.4.14) holds.

2) On the segment BC, equation (A.4.15) holds.

3) On the segment CD, $\lambda_2 = 0$ but we have the following choices for λ_1:

3a) $\lambda_1 = 0$ In that case

$$\mu_1^* \frac{\partial V_{110}}{\partial x_1} = -(p_1^* + r_2)V_{110} + p_1^* V_{010} + r_2 V_{100} + c_1\mu_1^* - g \qquad (A.4.19)$$

3b) $\lambda_1 = \mu_1^*$ In that case

$$0 = -(\frac{\mu_1^*}{\lambda^*} p_0^* p_1^* + r_2)V_{110} + \frac{\mu_1^*}{\lambda^*} p_o^* V_{010} + p_1^* V_{010} + r_2 V_{100} + c_1\mu_1^* - g \qquad (A.4.20)$$

3c) $\lambda_1 = \lambda^*$ In that case equation (A.4.16) holds, provided that we use the above values for λ_1 and λ_2.

In the cases where we had to consider more than one possibility, we can again find the optimality conditions as follows: gather all terms of the equation corresponding to each possibility at the right hand side. The right-hand-side corresponding to an optimal control law should be the largest one.

The case $a = (1, 0, 1)$ is completely symmetrical to the case $a = (1, 1, 0)$, so we do not deal with separately. When $a_o = 0$, there is no decision to be made and the corresponding equations may be written in a similar but easier way.