JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol 44, No. 1, SEPTEMBER 1984

Convexity and Characterization of Optimal Policies in a Dynamic Routing Problem[1]

J. N. TSITSIKLIS[2]

Communicated by P. Varaiya

Abstract. An infinite horizon, expected average cost, dynamic routing problem is formulated for a simple failure-prone queueing system, modelled as a continuous time, continuous state controlled stochastic process. We prove that the optimal average cost is independent of the initial state and that the cost-to-go functions of dynamic programming are convex. These results, together with a set of optimality conditions, lead to the conclusion that optimal policies are switching policies, characterized by a set of switching curves (or regions), each curve corresponding to a particular state of the nodes (servers).

Key Words. Stochastic control, unreliable queueing systems, average cost, jump disturbances.

1. Introduction

Overview. The main body of queueing theory has been concerned with the properties of queueing systems that are operated in a certain, fixed fashion (Ref. 1). Considerable attention has also been given to optimal static (stationary) routing strategies in queueing networks (Refs. 2–4), which are often found from the solution of a nonlinear programming problem (flow assignment problem).

Concerning dynamic control strategies, most of the literature (Refs. 5 and 6 are good surveys) deals with the control of the queueing discipline (priority setting) or with the control of the arrival and/or service rate in an

[1] The author would like to thank Prof. M. Athans for his help; Dr. S. Gershwin for his suggestions and for reading the manuscript; finally, the Associate Editor, Prof. P. Varaiya, for his numerous suggestions for improving the style of the paper. This research was supported by the National Science Foundation, Grant No. DAR-78-17826.

[2] Doctoral Student, Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts.

0022-3239/84/0900-0105$03.50/0

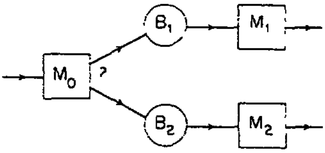

Fig. 1. Simple queueing system.

$M/M/1$ queue (Ref. 7) or $M/G/1$ queue (Ref. 8). Reference 9 considers the problem of controlling the service rate in a two-stage tandem queue.

Results for queueing systems where customers have the choice of selecting a server are fewer. Reference 10 considers multiserver queueing models with lane selection and derives mean waiting times but does not consider the optimization problem. Some problems with a high degree of symmetry have been solved (Refs. 11–13), leading to intuitively appealing strategies like, for example, "join the shortest queue." Results for systems without any particular symmetry are rare. Reference 14 contains a qualitative analysis of a dual-purpose system. In Ref. 15, a routing problem (very similar to ours), where the servers are allowed to be failure prone, is solved numerically. A simpler failure-prone system is studied in Ref. 16, and some analytical results are derived. Finally, the dynamic control problem for a class of flexible manufacturing systems, as defined in Ref. 17, has significant qualitative similarities with our problem.

In this paper, we consider an unreliable, failure-prone system (Fig. 1) with arrivals modelled as a continuous flow. Consequently, our model concentrates on the effects of failures, rather than the effects of random arrivals and service times. We prove convexity of the cost-to-go functions of dynamic programming (and hence, optimality of switching policies). Our method extends to more complex configurations.

Problem Description. We study a queueing control problem corresponding to the unreliable queueing system depicted in Fig. 1. We let M_0, M_1, M_2 be failure-prone nodes (servers, machines, processors) and B_1, B_2 be buffers (queues) with finite storage capacity. Machine M_0 receives external input (assumed to be always available), which it processes and sends to either of the buffers B_1, B_2. Machines M_1, M_2 then process the material in the buffers that precede them. We assume that each of the machines may fail and get repaired in a random manner. The failure and repair processes are modelled as memoryless stochastic processes (continuous-time Markov chains). We also assume that the maximum processing rate of a machine which is in working condition is finite.

With this system, there are two kinds of decisions to be made: (a) decide on the actual processing rate of each machine, at any time when it is in working condition and input to it is available; (b) decide, at any time, on how to route the output of machine M_0.

We consider a performance criterion which is linear in throughput and convex in storage.

The above configuration arises in certain manufacturing systems (Refs. 15, 18), from which our terminology is borrowed, and also in communication networks where the nodes may be thought of as being computers and the material being processed as messages (packets, Ref. 11). Note that the Markovian assumption on the failure and repair process of the nodes implies that a node may fail even at a time when it is not operating. This is a realistic assumption, in unreliable communication networks and in those manufacturing systems where failures may be ascribed to external causes (Refs. 18, 19). On the other hand, in some manufacturing systems, failure probabilities increase with the degree of utilization of the machines (Ref. 20). Such systems are not captured by our model and require a substantially different mathematical approach.

2. Dynamic Routing Problem

In this section, we formulate mathematically the dynamic routing problem and define the set of admissible control laws and the performance criterion to be minimized.

Consider the queueing system of Fig. 1, as described in Section 1. Let x_i be a continuous variable indicating the amount of material in buffer B_i, $i=1,2$, and let N_i be the maximum allowed level in that buffer. We denote (x_1, x_2) by x. Let

$$\alpha_i(t) \in \{0, 1\}, \qquad i=0,1,2,$$

be independent, right-continuous Markov chains with transition rates r_i from 0 to 1 and p_i from 1 to 0. We have

$$\alpha_i(t)=0 \text{ or } 1,$$

according to whether machine M_i is down or up. Let Ω denote the set of sample paths ω of

$$\alpha=(\alpha_0, \alpha_1, \alpha_2)$$

and $\mathscr{A}$, $\mathscr{A}_t$ the σ-field generated by

$$\{\alpha(\tau)|\tau \leq t\}, \qquad \{\alpha(\tau)|\tau \geq 0\},$$

respectively. Let $\mathscr{P}(\alpha(0))$ be the measure on $(\Omega, \mathscr{A})$ when the initial state of the Markov chain is $\alpha(0)$.

We define the state space S of the system by

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

The state $s(t) \in S$ of the system at time t is defined as

$$s(t) = ((x_1, x_2), (\alpha_0, \alpha_1, \alpha_2))(t) = (x, \alpha)(t). \tag{2}$$

Let λ^*, μ_1^*, μ_2^* be the maximum allowed flow rates through machines M_0, M_1, M_2, respectively; let $\lambda(t)$, $\mu_1(t)$, $\mu_2(t)$ be the actual flow rates, at time t, through machines M_0, M_1, M_2, respectively; finally, let $\lambda_1(t)$, $\lambda_2(t)$ be the flow rates, at time t, from machine M_0 to the buffers B_1, B_2, respectively. No flow may go through a machine which is down:

$$\alpha_i(t) = 0 \Rightarrow \begin{cases} \lambda(t) = \lambda_1(t) = \lambda_2(t) = 0, & i = 0, \\ \mu_i(t) = 0, & i = 1, 2. \end{cases} \tag{3}$$

Conservation of flow implies

$$\lambda(t) = \lambda_1(t) + \lambda_2(t), \tag{4}$$

$$x_i(t) = x_i(0) + \int_0^t (\lambda_i(\tau) - \mu_i(\tau))\, d\tau. \tag{5}$$

An admissible control law u is a mapping which to any initial state

$$s(0) = (x(0), \alpha(0))$$

assigns a right-continuous stochastic process

$$u(\omega, t) = (\lambda_1^u(\omega, t), \lambda_2^u(\omega, t), \mu_1^u(\omega, t), \mu_2^u(\omega, t)),$$

defined on the previously introduced probability space $(\Omega, \mathscr{A}, \mathscr{P}(\alpha(0)))$, with the following properties.

Property (S1). Each of the random variables $\lambda_i^u(t)$, $\mu_i^u(t)$ is $\mathscr{A}_t$-measurable.

Property (S2). The following relations hold:

$$0 \le \lambda_i^u(\omega, t), \qquad i = 1, 2, \tag{6}$$

$$\lambda_1^u(\omega, t) + \lambda_2^u(\omega, t) \le \alpha_0(\omega, t)\lambda^*, \tag{7}$$

$$0 \le \mu_i^u(\omega, t) \le \alpha_i(\omega, t)\mu_i^*, \qquad i = 1, 2. \tag{8}$$

Property (S3). The state

$$x_i^u(t) = x_i(0) + \int_0^t (\lambda_i^u(\tau) - \mu_i^u(\tau))\, d\tau$$

satisfies

$$0 \leq x_i^u(t) \leq N_i, \qquad i = 1, 2,$$

and $x_i^u(\omega, t)$ is a measurable function of the initial state $s(0)$, $\forall \omega \in \Omega$, $\forall r \geq 0$.

We let U be the set of all admissible control laws, and let $U_M \subset U$ be the set of those control laws such that $u(t)$ depends only on $s(t)$.

For $u \in U$ and f a bounded function of s, define

$$(\mathscr{L}^u f)(s) = \lim_{t \downarrow 0} \frac{E[f(s^u(t)) | s^u(0) = s] - f(s)}{t}, \tag{9}$$

whenever the limit exists. It can be verified that, for continuous functions f in the domain of $\mathscr{L}^u$,

$$(\mathscr{L}^u f)(x, \alpha) = \lim_{t \downarrow 0} \frac{f(x_\alpha^u(t), \alpha) - f(x, \alpha)}{t} + \sum_{\alpha^*} p_{\alpha\alpha^*}[f(x, \alpha^*) - f(x, \alpha)], \tag{10}$$

where $p_{\alpha\alpha^*}$ is the transition rate from α to α^* and where $x_\alpha^u(t)$ is the value of x at time t if control law u is used and no jump of α occurs until time t.

Performance Criterion. We are interested in minimizing the long-run (infinite horizon) average cost resulting from the operation of the system. Let $k(s, \lambda_1, \lambda_2, \mu_1, \mu_2)$ be a function of the state and control variables representing the instantaneous cost. For notational convenience, we define

$$k^u(s^u(\omega, t)) \equiv k(s^u(\omega, t), \lambda_1^u(\omega, t), \lambda_2^u(\omega, t), \mu_1^u(\omega, t), \mu_2^u(\omega, t)). \tag{11}$$

We introduce the following assumptions:

$$k^u(s^u(\omega, t)) = f(x^u(\omega, t), \alpha(\omega, t)) - c_1\mu_1^u(\omega, t) - c_2\mu_2^u(\omega, t), \tag{12}$$

where $c_1, c_2 > 0$ and, for any $\alpha \in \{0, 1\}^3$, $f(x, \alpha)$ is (i) nondecreasing in x_1 and x_2, (ii) convex, and (iii) Lipschitz continuous. Let $f_\alpha(x)$ be an alternative notation for $f(x, \alpha)$.

The function to be minimized is

$$g^u(s) = \limsup_{T \to \infty} E\left[\frac{1}{T}\int_0^T k^u(s^u(\omega, t))\, dt | s^u(0) = s\right]. \tag{13}$$

We define the optimal average cost by

$$g^*(s) = \inf_{u \in U} g^u(s). \tag{14}$$

In Section 3, we show that g^* is independent of s.

3. Reduction of the Set of Admissible Control Laws

Suppose that, at some time, the lead machine is down and both downstream machines are up. If this configuration does not change for a large enough time interval, we expect that any reasonable control law would eventually empty both buffers. Indeed, Theorem 3.1 shows that we may so restrict the set of admissible control laws without worsening the optimal performance of the system (i.e., without increasing the optimal value of the cost functional).

We then show that there exists a particular state which is recurrent under any control law that satisfies the above introduced constraint (Theorem 3.2). The existence of such a recurrent state permits a significant simplification of the mathematical issues typically associated with average cost problems.

We end this section by introducing the class of regenerative control laws. This is the class of control laws for which the stochastic process regenerates (i.e., forgets the past history and starts afresh) each time the particular recurrent state is reached. In that case, g^u admits a simple and useful representation and is independent of s (Theorem 3.3). We show that we may restrict ourselves to regenerative control laws without any loss of performance (Theorem 3.4).

Definition 3.1. Let U_A be the set of control laws in U with the following property. If, at some time t_0,

$$\alpha(t_0) = (0, 1, 1),$$

and $\alpha(t)$ does not change for a further time interval of $\max\{N_1/\mu_1^*, N_2/\mu_2^*\}$ time units, then

$$s^u(t = t_0 + \max\{N_1/\mu_1^*, N_2/\mu_2^*\}) = ((0, 0), (0, 1, 1)).$$

Remark 3.1. A sufficient (but not necessary) condition for a control law u to belong in U_A is that downstream machines operate at full capacity whenever

$$\alpha = (0, 1, 1).$$

However, we do not want to impose the latter condition, because in the course of the proofs in Section 5 we will use control laws that violate it.

Theorem 3.1. For any $u \in U$, $s(0) \in S$, there exists some $w \in U_A$ such that

$$\int_0^t k^w(s^w(\omega, \tau))\, d\tau \leq \int_0^t k^u(s^u(\omega, \tau))\, d\tau, \qquad \forall t \geq 0, \forall \omega \in \Omega. \tag{15}$$

Proof. Fix some initial state $s(0)$ and a control law $u \in U$. Let $w \in U$ be a control law such that, with the same initial state, we have

$$\lambda_i^w(\omega, t) = \lambda_i^u(\omega, t), \qquad i = 1, 2, \forall \omega, t, \tag{16}$$

$$\mu_i^w(\omega, t) = \begin{cases} \mu_i^*, & \text{if } x_i^w(\omega, t) \neq 0, \quad \alpha = (0, 1, 1), \\ 0, & \text{if } x_i^w(\omega, t) = 0, \quad \alpha = (0, 1, 1), \\ \mu_i^u(\omega, t), & \text{if } x_i^w(\omega, t) = x_i^u(\omega, t), \quad \alpha \neq (0, 1, 1), \\ 0, & \text{if } x_i^w(\omega, t) \neq x_i^u(\omega, t), \quad \alpha \neq (0, 1, 1), \end{cases} \tag{17}$$

where $x_i^w(\omega, t)$ is determined by

$$x_i^w(\omega, t) = x_i(0) + \int_0^t (\lambda_i^w(\omega, \tau) - \mu_i^w(\omega, \tau))\, d\tau. \tag{18}$$

It is easy to see that $w \in U_A$ and

$$0 \leq x_i^w(\omega, t) \leq x_i^u(\omega, t), \qquad \forall \omega, t, \quad i = 1, 2. \tag{19}$$

From (19) and the monotonicity of f_α, we have

$$\int_0^t f_{\alpha(\omega,\tau)}(x^w(\omega, \tau))\, d\tau \leq \int_0^t f_{\alpha(\omega,\tau)}(x^u(\omega, \tau))\, d\tau, \qquad \forall \omega, t. \tag{20}$$

Using (16), (18), (19), we have

$$\begin{aligned} \int_0^t \mu_i^w(\omega, \tau)\, d\tau &= x_i(0) - x_i^w(\omega, t) + \int_0^t \lambda_i^w(\omega, \tau)\, d\tau \\ &\geq x_i(0) - x_i^u(\omega, t) + \int_0^t \lambda_i^u(\omega, \tau)\, d\tau \\ &= \int_0^t \mu_i^u(\omega, \tau)\, d\tau, \qquad \forall \omega, t. \end{aligned} \tag{21}$$

Adding Ineqs. (20) and (21), for $i = 1, 2$, we obtain the desired result. □

Corollary 3.1. We may restrict to control laws in U_A without loss of optimality. Namely,

$$\inf_{u \in U_A} g^u(s) = \inf_{u \in U} g^u(s) = g^*(s), \qquad \forall s \in S. \tag{22}$$

We now proceed with the recurrence properties of control laws in U_A. For the rest of this paper, we let s_0 denote the special state

$$(x, \alpha) = ((0, 0), (0, 1, 1)).$$

Let $u \in U_A$. We define the stopping time T_n^u, $n \geq 1$, as the nth time, after time 0, that the state s_0 is reached, given that control law u is used. We also let

$$T_0^u = 0.$$

Let

$$s^u(0) = (x^u(0), \alpha), \; s^w(0) = (x^w(0), \alpha)$$

be elements of S with the same value of α; let $u, w \in U_A$. We define the stopping time T^{uw} by

$$T^{uw} = \inf\{t > 0: s^u(t) = s^w(t) = s_0\},$$

$$T^{uw} = \infty, \qquad \text{if the above set is empty.} \tag{23}$$

If we are given a third element of S,

$$s^v(0) = (x^v(0), \alpha),$$

with the same value of α, and a third control law $v \in U$, we may define T^{uvw} in a similar way, as the first time that

$$s^u(t) = s^w(t) = s^v(t) = s_0.$$

Theorem 3.2. Let u, v, $w \in U_A$, and let $s^u(0)$, $s^v(0)$, $s^w(0)$ be three initial states with the same value of α. Assume that

$$p_0 \neq 0, \qquad r_i \neq 0, \qquad i = 0, 1, 2.$$

Then,

$$\text{(a)} \quad E[T_{n+1}^u - T_n^u] \leq B, \tag{24}$$

$$\text{(b)} \quad E[T^{uvw}] \leq B \qquad E[T^{uw}] \leq B, \tag{25}$$

where B is a constant independent of u, v, w and the initial states $s^u(0)$, $s^v(0)$, $s^w(0)$.

Proof. Let Q_n be the nth time that the continuous-time Markov chain $\alpha(t)$ reaches the state

$$\alpha = (0, 1, 1).$$

Since p_0, r_i are nonzero, there exists a constant A such that

$$E[Q_n] \leq nA,$$

for all initial states $\alpha(0)$, and

$$E[Q_n - Q_m] \leq (n - m)A.$$

If

$$\alpha(t)=(0, 1, 1)$$

and if no jumps of α occur for a further time interval of

$$T\equiv \max\{N_1/\mu_1^*, N_2/\mu_2^*\}$$

time units, which is the case with probability equal to or larger than

$$q\equiv \exp(-(r_0+p_1+p_2)T),$$

the state becomes

$$s^u(t+T)=s_0,$$

for any $u\in U_A$ and regardless of the initial state. It follows that

$$E[T_{n+1}^u-T_n^u]\le \sum_{k=1}^{\infty}(kA+T)q(1-q)^{k-1}=B.$$

Similar inequalities hold for $E[T^{uw}]$, $E[T^{uwv}]$. □

It will be assumed throughout this paper that

$$p_0\neq 0, \qquad r_i\neq 0, \qquad i=0, 1, 2.$$

If we allowed

$$p_0=0,$$

all subsequent results would be still valid, but the recurrent state s_0 should be differently chosen.

Theorem 3.2 allows us to break down the infinite horizon into a sequence of almost surely finite and disjoint time intervals $[T_n^u, T_{n+1}^u)$. If, in addition, the stochastic process $s^u(t)$ regenerates at the times T_n^u, the infinite horizon average cost admits a simple and useful representation.

We define the set U_R of regenerative control laws to consist of those elements of U_A which forget the past each time that the state is equal to s_0 and start afresh. To make these requirements formal, we first define a regeneration time to mean an almost surely finite stopping time T, such that

$$s^u(T)=s_0,$$

with probability 1. Our first condition on regenerative control laws is that the past is forgotten at regeneration times. See the property below.

Property (S4). The stochastic process

$$\{(s^u(T+t), u(T+t)), t\ge 0\}$$

is independent of $\mathscr{A}_T$, for any regeneration time T.

The second requirement is that the stochastic process in (S4) be the same for all regeneration times T. See the property below.

Property (S5). For any two regeneration times S, T, the stochastic processes

$$\{(s^u(T+t), u(T+t)), t \geq 0\} \text{ and } \{(s^u(S+t), u(S+t)), t \geq 0\}$$

are identically distributed.

Markovian control laws in U_A certainly belong to U_R. However, the proofs of the results of Section 5 require us to consider non-Markovian control laws as well. It turns out that U_R is a suitable framework.

Theorem 3.3. Let $u \in U_R$. Then,

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

$$= \frac{E\left[\int_{T_n^u}^{T_{n+1}^u} k^u(s^u(\tau))\, d\tau\right]}{E[T_{n+1}^u - T_n^u]}, \qquad n = 1, 2, \ldots. \tag{26}$$

Note that the first equality implies that the limit exists and is independent of the initial state.

Proof. Define

$$W_m \equiv T_{m+1}^u - T_m^u, \qquad m = 1, 2, \ldots, \tag{27}$$

$$U_m \equiv \int_{T_m^u}^{T_{m+1}^u} k^u(s^u(\tau))\, d\tau, \qquad m = 1, 2, \ldots. \tag{28}$$

The random vectors (W_m, U_m), $m = 1, 2, \ldots$, are independent (by S4), identically distributed (by S5). Then, an ergodic theorem (Ref. 22, Vol. 2) implies that

$$\lim_{m\to\infty} \frac{\sum_{k=1}^{m} U_k}{\sum_{k=1}^{m} W_k} = \frac{E[U_n]}{E[W_n]}, \qquad \text{a.s.} \tag{29}$$

Now,

$$\lim_{m\to\infty} \frac{\sum_{k=1}^{m} U_k}{\sum_{k=1}^{m} W_k} = \lim_{m\to\infty} \frac{\int_{T_1^u}^{T_m^u} k^u(s^u(\tau))\, d\tau}{T_m^u - T_1^u}$$

$$= \lim_{m\to\infty} \left[\frac{1}{T_m^u} \int_0^{T_m^u} k^u(s^u(\tau))\, d\tau\right]$$

$$+ \lim_{m\to\infty} \left[\left(\frac{1}{T_m^u - T_1^u} - \frac{1}{T_m^u}\right) \int_0^{T_m^u} k^u(s^u(\tau))\, d\tau\right]$$

$$- \lim_{m\to\infty} \left[\frac{1}{T_m^u - T_1^u} \int_0^{T_1^u} k^u(s^u(\tau))\, d\tau\right]. \tag{30}$$

We claim that the second and third terms are almost surely equal to zero. Let M be a bound on $|k^u|$. Then,.

$$\left|\left(\frac{1}{T_m^u - T_1^u} - \frac{1}{T_m^u}\right) \int_0^{T_m^u} k^u(s^u(\tau))\, d\tau\right| \leq M T_m^u \frac{T_1^u}{T_m^u(T_m^u - T_1^u)}. \tag{31}$$

Now,

$$T_1^u < \infty, \qquad \text{a.s.},$$

$$\lim_{m\to\infty} T_m^u = \infty, \qquad \text{a.s.}$$

Also,

$$\left|\frac{1}{T_m^u - T_1^u} \int_0^{T_1^u} k^u(s^u(\tau))\, d\tau\right| \leq \frac{M T_1^u}{T_m^u - T_1^u} \to 0, \qquad \text{a.s.}, \tag{32}$$

for the same reasons. We now take expectations in (30) and invoke (29) to obtain

$$E \lim_{m\to\infty} \left[\frac{1}{T_m^u} \int_0^{T_m^u} k^u(s^u(\tau))\, d\tau\right] = \frac{E[U_n]}{E[W_n]}. \tag{33}$$

Let

$$T^u(t) = \inf\{\tau \geq t \colon \exists n \text{ such that } \tau = T_n^u\},$$

and observe that the sequence

$$(1/T_m^u) \int_0^{T_m^u} k^u(s^u(\tau))\, d\tau$$

and the function

$$(1/T^u(t)) \int_0^{T^u(t)} k^u(s^u(\tau))\, d\tau$$

take the same values in the same order; therefore, they have the same limit and may be interchanged in (33). We then use the dominated convergence theorem to interchange the limit and the expectation on the left-hand side to obtain

$$\lim_{t\to\infty} E\left[\frac{1}{T^u(t)} \int_0^{T^u(t)} k^u(s^u(\tau))\, d\tau\right] = \frac{E[U_n]}{E[W_n]}. \tag{34}$$

Finally,

$$\begin{aligned}
&\lim_{t\to\infty} \left| E\left[\frac{1}{t}\int_0^t k^u(s^u(\tau))\, d\tau\right] - E\left[\frac{1}{T^u(t)} \int_0^{T^u(t)} k^u(s^u(\tau))\, d\tau\right]\right| \\
&\quad \le \lim_{t\to\infty} \left| E\left[\frac{1}{T^u(t)} \int_t^{T^u(t)} k^u(s^u(\tau))\, d\tau\right]\right| \\
&\quad + \lim_{t\to\infty} \left| E\left[\left(\frac{1}{t} - \frac{1}{T^u(t)}\right) \int_0^t k^u(s^u(\tau))\, d\tau\right]\right| = 0.
\end{aligned} \tag{35}$$

The two summands on the right-hand side of (35) converge to zero, because they are bounded above by $E[T^u(t) - t]M/t$, which is bounded by BM/t (Theorem 3.2). Equations (34), (35) complete the proof of (26). □

Remark 3.2. If $s(0) = s_0$, then (26) is obviously true for $n = 0$ as well.

The last result of this section shows that we may restrict ourselves to control laws in U_R without increasing the optimal value of the cost functional.

Theorem 3.4. The following result holds:

$$\inf_{u\in U_R} g^u = \inf_{u\in U_A} g^u(s) = g^*(s), \qquad \forall s \in S. \tag{36}$$

Proof. *Outline.* View our control problem as follows. Each time T_n the state s_0 is reached, a policy $u_n \in U_A$ to be followed in $[T_n, T_{n+1})$ is chosen. We then have a single-state semi-Markov renewal programming problem with an infinite action space and bounded costs per stage; regenerative control laws correspond to stationary policies of the semi-Markov problem. Moreover, $T_n - T_{n-1}$ is uniformly bounded, in expected value, for all policies of the semi-Markov problem. It follows that stationary policies exist that come arbitrarily close to being optimal. By translating this statement to the original problem, we obtain (36). □

4. Value Function of Dynamic Programming

Using the recurrence properties of control laws in U_R, we may now define value (cost-to-go) functions of dynamic programming. This is done by using the recurrent state s_0 as a reference state. Moreover, we exploit Theorem 3.3 to convert the average cost problem to a total cost problem.

Following Ref. 21, we define the value function $V^u: S \to R$, corresponding to $u \in U_R$, by

$$V^u(s) = E\left[\int_0^{T_1^u} (k^u(s^u(\tau)) - g^u)\, d\tau | s^u(0) = s\right]. \tag{37}$$

In view of Theorem 3.3, we have

$$V^u(s_0) = 0, \qquad \text{for all } u \in U_R.$$

We also define an auxiliary value function $\hat{V}^u(s)$ by

$$\hat{V}^u(s) = E\left[\cdot\int_0^{T_1^u} (k^u(s^u(\tau)) - g^*)\, d\tau | s^u(0) = s\right] \tag{38}$$

and the optimal value function $V^*(s)$ by

$$V^*(s) = \inf_{u \in U_R} \hat{V}^u(s). \tag{39}$$

The above defined functions are all bounded by $2BM$, where M is a constant bounding $|k^u(s)|$ and B is the constant of Theorem 3.2.

Lemma 4.1. The above value functions satisfy the relations:

(a) $0 \le \hat{V}^u(s) - V^u(s) \le (g^u - g^*)B, \forall s \in S$;
(b) $\hat{V}^u(s_0) = (g^u - g^*)E[T_1^u | s^u(0) = s_0]$;
(c) $g^u = g^*$, iff $\hat{V}^u(s_0) = 0$;
(d) $\hat{V}^u(s_0) \ge 0$, $V^*(s_0) = 0$.

Proof. It follows directly from the definitions and the inequality $E[T_1^u] \le B$. □

We will say that a control law $u \in U_R$ is everywhere optimal if

$$\hat{V}^u(s) = V^*(s), \qquad \forall s \in S;$$

it is optimal if

$$g^u = g^*.$$

Lemma 4.1 implies that an everywhere optimal control law is optimal.

We conclude this section with a few properties of $\hat{V}^u$, V^u that will be needed in the next section.

Lemma 4.2. (a) For any positive integers m and n such that $n \geq m$ and any $u \in U_R$,

$$E\left[\int_{T_m^u}^{T_n^u} (k^u(s^u(\tau)) - g^u)\, d\tau\right] = 0. \tag{40}$$

(b) For any positive integer n and any $u \in U_R$,

$$V^u(s) = E\left[\int_0^{T_u^n} (k^u(s^u(\tau)) - g^u)\, d\tau | s^u(0) = s\right]. \tag{41}$$

Proof. Both parts follow immediately from Theorem 3.3 □

The following result is essentially a version of Eq. (41).

Lemma 4.3. Let u, v, $w \in U_R$. Let $s^u(0)$, $s^v(0)$, $s^w(0)$ be three states with the same value of α. Let T^{uvw} be as defined in Section 3. Then,

$$V^u(s) = E\left[\int_0^{T^{uvw}} (k^u(s^u(\tau)) - g^u)\, d\tau | s^u(0) = s\right]. \tag{42}$$

Proof. Let

$$T = \min\{T_n^u \colon T_n^u \geq T^{uvw}\},$$

and let χ_n be the characteristic (indicator) function of the set of those $\omega \in \Omega$ such that

$$T > T_n^u.$$

We then have

$$E\left[\int_0^T (k^u(s^u(\tau)) - g^u)\, d\tau\right]$$

$$= E\left[\int_0^{T_1^u} (k^u(s^u(\tau)) - g^u)\, d\tau\right]$$

$$+ \sum_{n=1}^{\infty} E\left[\chi_n \int_{T_n^u}^{T_{n+1}^u} (k^u(s^u(\tau)) - g^u)\, d\tau\right]. \tag{43}$$

The random variable χ_n is $\mathscr{A}_{T_n^u}$-measurable. Therefore,

$$E\left[\chi_n \int_{T_n^u}^{T_{n+1}^u} (k^u(s^u(\tau))-g^u)\,d\tau\right]$$
$$= E\left[\chi_n E\left[\int_{T_n^u}^{T_{n+1}^u} (k^u(s^u(\tau))-g^u)\,d\tau \big| \mathscr{A}_{T_n^u}\right]\right] = 0. \tag{44}$$

The second equality in (44) follows from Lemma 4.2(a) and the assumption that u regenerates at time T_n^u. For the same reasons, we obtain

$$E\left[\int_{T^{uvw}}^{T} (k^u(s^u(\tau))-g^u)\,d\tau\right] = E\left[\int_{T_1^u}^{T_2^u} (k^u(s^u(\tau))-g^u)\,d\tau\right] = 0. \tag{45}$$

Combining (43), (44), (45), and using the definition of V^u, we obtain

$$E\left[\int_0^{T^{uvw}} (k^u(s^u(\tau))-g^u)\,d\tau\right] = E\left[\int_0^{T_1^u} (k^u(s^u(\tau))-g^u)\,d\tau\right]$$
$$= V^u(s^u(0)). \tag{46}$$

□

The last lemma is an elementary consequence of our definitions.

Lemma 4.4. Given some $s \in S$ and $\epsilon > 0$, $\exists u \in U_R$, such that

$$\hat{V}^u(s) \leq V^*(s) + \epsilon \quad \text{and} \quad g^u \leq g^* + \epsilon.$$

Proof. *Outline.* Assume $s \neq s_0$. Then, $\hat{V}^u$ depends on the choice of the control variables up to time T_1^u and g^u depends on the choice after that time. The control variables before and after T_1^u may be independently chosen so as to satisfy both inequalities. If $s = s_0$, choose u such that

$$g^u \leq g^* + \min\{\epsilon, \epsilon/B\}.$$

Then,

$$\hat{V}^u(s_0) \leq V^*(s_0) + \epsilon.$$

□

5. Convexity and Other Properties of V^*

In this section, we exploit the structure of our system to obtain certain basic properties of V^*. These properties, together with the optimality conditions, to be derived in Section 6, lead directly to the characterization of optimal control laws. For the rest of the paper, let $V_\alpha^*(x)$ denote $V^*(x, \alpha)$.

Theorem 5.1. $V^*(x, \alpha)$ is convex, $\forall \alpha$.

Proof. Let

$$s^u(0)=(x^u, \alpha) \quad \text{and} \quad s^v(0)=(x^v, \alpha)$$

be two states in S with the same value of α. Let

$$c \in (0, 1) \quad \text{and} \quad s^w(0)=(cx^u+(1-c)x^v, \alpha).$$

Then, $s^w(0) \in S$, because $[0, N_1] \times [0, N_2]$ is a convex set, and we need to show that

$$V^*(s^w(0)) \leq cV^*(s^u(0))+(1-c)V^*(s^v(0)). \tag{47}$$

Fix some $\epsilon > 0$, and let u, v be control laws in U_R such that

$$g^u \leq g^*+\epsilon, \qquad g^v \leq g^*+\epsilon, \tag{48}$$

$$\hat{V}^u(s^u(0)) \leq V^*(s^u(0))+\epsilon, \qquad \hat{V}^v(s^v(0)) \leq V^*(s^v(0))+\epsilon. \tag{49}$$

Let $s^u(\omega, t)$ and $s^v(\omega, t)$ be the corresponding sample paths. We now define a control law w to be used starting from the initial state $s^w(0)$. Let, for $i=1, 2$,

$$\lambda_i^w(\omega, t)=c\lambda_i^u(\omega, t)+(1-c)\lambda_i^v(\omega, t), \tag{50}$$

$$\mu_i^w(\omega, t)=c\mu_i^u(\omega, t)+(1-c)\mu_i^v(\omega, t). \tag{51}$$

With w defined by (50), (51), Assumptions (S1)–(S3) are satisfied, because these assumptions are satisfied by u and v. Moreover, by linearity of the dynamics,

$$x_i^w(\omega, t)=cx_i^u(\omega, t)+(1-c)x_i^v(\omega, t). \tag{52}$$

Since $x=(0, 0)$ is an extreme point of $[0, N_1] \times [0, N_2]$, Eq. (52) implies that, whenever

$$s^w(t)=s_0,$$

we also have

$$s^u(t)=s^v(t)=s_0.$$

Therefore,

$$T_1^w=T^{uvw}$$

and, consequently, $w \in U_A$. Moreover, u and v regenerate whenever

$$s^w(t)=s_0$$

and, therefore, $w \in U_R$. Using (52) and the convexity of the cost function, we obtain

$$\int_0^{T_1^w} (k^w(s^w(\omega,\tau))-g^*)\,d\tau \le c\int_0^{T_1^w} (k^u(s^u(\omega,\tau))-g^*)\,d\tau$$
$$+(1-c)\int_0^{T_1^w} (k^v(s^v(\omega,\tau))-g^*)\,d\tau. \tag{53}$$

We take expectations of both sides of Ineq. (53) and rearrange it to obtain

$$V^*(s^w(0)) \le \hat{V}^w(s^w(0)) \le cE\left[\int_0^{T_1^w} (k^u(s^u(\omega,\tau))-g^u)\,d\tau | s^u(0)\right]$$
$$+(1-c)E\left[\int_0^{T_1^w} (k^v(s^v(\omega,\tau))-g^v)\,d\tau | s^v(0)\right]$$
$$+(cg^u+(1-c)g^v-g^*)E[T_1^w]. \tag{54}$$

Since

$$T^{uvw}=T_1^w,$$

Lemma 4.3 applies. Using also Ineqs. (48), (49) and Lemma 4.1(a), we obtain

$$V^*(s^w(0)) \le cV^u(s^u(0))+(1-c)V^v(s^v(0))+\epsilon B$$
$$\le c\hat{V}^u(s^u(0))+(1-c)\hat{V}^v(s^v(0))+\epsilon B$$
$$\le cV^*(s^u(0))+(1-c)V^*(s^v(0))+\epsilon(1+B). \tag{55}$$

Since ϵ was arbitrary, we may let $\epsilon\downarrow 0$ in (55) to obtain Ineq. (47). □

It is not hard to show that, if the storage cost f_α [defined by Eq. (12)] is strictly convex, then Ineq. (47) is strict. In fact, it is also true that (47) is a strict inequality even if f_α is linear. A detailed proof would be fairly involved, and we only give here an outline.

Assume, without loss of generality, that

$$x_1^u(0)\neq x_1^v(0).$$

With control law w, defined by (50), (51), there is positive probability that

$$\alpha(t)=(0,1,1), \qquad x_1^w(t)\neq 0, \qquad \mu_1^w(t)<\mu_1^*,$$

for all t belonging to a time interval of positive length. We can then show (in a way similar to the proof of Theorem 3.1) that any control law with the above property does not minimize $\hat{V}^w(s^w(0))$ and that

$$V^*(s^w(0))<\hat{V}^w(s^w(0))-\delta,$$

for some δ independent of ϵ. Using this inequality in (54) and (55), (47) becomes a strict inequality.

Let M be such that

$$|f_\alpha(x_1, x_2) - f_\alpha(x_1 + \Delta_1, x_2 + \Delta_2)| \leq M(|\Delta_1| + |\Delta_2|). \tag{56}$$

Such an M exists, since f_α is Lipschitz continuous. We then have the following result.

Theorem 5.2. Let

$$0 \leq x_i \leq x_i + \Delta_i \leq N_i, \qquad i = 1, 2,$$

$$\Delta_1 + \Delta_2 > 0.$$

Then,

$$-c_1\Delta_1 - c_2\Delta_2 < V_\alpha^*(x_1 + \Delta_1, x_2 + \Delta_2) - V_\alpha^*(x_1, x_2) < MB(\Delta_1 + \Delta_2). \tag{57}$$

In particular, V_α^* is Lipschitz continuous and, if $f_\alpha \equiv 0$, then $M = 0$ and V_α^* is strictly decreasing in each variable.

Proof. The two inequalities in (57) will be proved separately. Without loss of generality, we assume that $\Delta_2 = 0$ and we start by proving the second inequality.

(a) Fix two initial states

$$s^u(0) = (x_1, x_2, \alpha) \quad \text{and} \quad s^w(0) = (x_1 + \Delta, x_2, \alpha), \qquad \Delta > 0,$$

with the same value of α. Let $u \in U_R$ be such that (Lemma 4.4)

$$\hat{V}^u(s^u(0)) \leq V^*(s^u(0)) + \epsilon, \qquad g^u \leq g^* + \epsilon. \tag{58}$$

We now define a new control law $w \in U_R$ to be used starting from $s^w(0)$ as follows:

$$\lambda_1^w(\omega, t) = \begin{cases} 0, & \text{if } x_1^w(\omega, t) \neq x_1^u(\omega, t), \\ \lambda_1^u(\omega, t), & \text{if } x_1^w(\omega, t) = x_1^u(\omega, t), \end{cases} \tag{59}$$

$$\mu_1^w(\omega, t) = \begin{cases} \alpha_1(\omega, t)\mu_1^*, & \text{if } x_1^w(\omega, t) \neq x_1^u(\omega, t), \\ \mu_1^u(\omega, t), & \text{if } x_1^w(\omega, t) = x_1^u(\omega, t), \end{cases} \tag{60}$$

$$\lambda_2^w(\omega, t) = \lambda_2^u(\omega, t), \qquad \mu_2^w(\omega, t) = \mu_2^u(\omega, t). \tag{61}$$

Then, $w \in U_R$ and

$$x_1^u(\omega, t) \leq x_1^w(\omega, t) \leq x_1^u(\omega, t) + \Delta, \tag{62}$$

$$\mu_1^w(\omega, t) \geq \mu_1^u(\omega, t). \tag{63}$$

Moreover,

$$T^{uw} = T_1^w$$

and

$$\int_0^{T_1^w} k^w(s^w(\omega, \tau))\, d\tau = \int_0^{T_1^w} (f_\alpha(x^w(\omega, \tau)) - c_1\mu_1^w(\omega, \tau) - c_2\mu_2^w(\omega, \tau))\, d\tau$$

$$\leq \int_0^{T_1^w} (f_\alpha(x^u(\omega, \tau)) + M\Delta - c_1\mu_1^u(\omega, \tau) - c_2\mu_2^u(\omega, \tau))\, d\tau$$

$$= \int_0^{T_1^w} k^u(s^u(\omega, \tau))\, d\tau + M\Delta T_1^w. \tag{64}$$

We claim that there exists a set $A \subset \Omega$ of positive probability measure such that Ineq. (64) is strict for all $\omega \in A$. Namely, consider all those ω for which $\alpha(\omega, t)$ becomes $(0, 1, 1)$ before time $\Delta/2(\lambda^* + \mu_1^*)$ and stays equal to $(0, 1, 1)$ until time T_1^w. Let $\delta > 0$ be such that

$$\Pr(\omega \in A) > \delta.$$

For all $\omega \in A$, we have

$$\int_0^{T_1^w} c_1\mu_1^w(\omega, \tau)\, d\tau > \int_0^{T_1^w} c_1\mu_1^u(\omega, \tau)\, d\tau + c_1\Delta/2 \tag{65}$$

and, consequently,

$$\int_0^{T_1^w} k^w(s^w(\omega, \tau))\, d\tau < \int_0^{T_1^w} k^u(s^u(\omega, \tau))\, d\tau + M\Delta T_1^w - c_1\Delta/2, \qquad \omega \in A. \tag{66}$$

Taking expectations in (64) and using (66) for $\omega \in A$, we obtain

$$E\left[\int_0^{T_1^w} (k^w(s^w(\omega, \tau)) - g^*)\, d\tau\right] \leq E\left[\int_0^{T_1^w} (k^u(s^u(\omega, \tau)) - g^*)\, d\tau\right] + M\Delta E[T_1^w] - c_1\delta\Delta/2. \tag{67}$$

Using Lemma 4.3 and following the same steps as in the proof of Theorem 5.1, we obtain

$$V^*(s^w(0)) \leq \hat{V}^w(s^w(0)) \leq V^*(s^u(0)) + \epsilon(1 + B) + M\Delta B - \delta c_1\Delta/2. \tag{68}$$

Since ϵ was arbitrary, we may let ϵ decrease to zero to obtain the second inequality in (57).

(b) For the proof of the left-hand-side of (57), let $s^u(0)$ and $s^w(0)$ be as before; let $w \in U_R$ be such that

$$\hat{V}^w(s^w(0)) \leq V^*(s^w(0)) + \epsilon, \qquad g^w \leq g^* + \epsilon; \tag{69}$$

and define $u \in U_R$ [to be used starting from $s^u(0)$] as follows:

$$\lambda_1^u(\omega, t) = \begin{cases} \alpha_0(\omega, t)(\lambda^* - \lambda_2^w(\omega, t)), & \text{if } x_1^u(\omega, t) \neq x_1^w(\omega, t), \\ \lambda_1^w(\omega, t), & \text{if } x_1^u(\omega, t) = x_1^w(\omega, t), \end{cases} \tag{70}$$

$$\mu_1^u(\omega, t) = \begin{cases} 0, & \text{if } x_1^u(\omega, t) \neq x_1^w(\omega, t), \\ \mu_1^w(\omega, t), & \text{if } x_1^u(\omega, t) = x_1^w(\omega, t), \end{cases} \tag{71}$$

$$\lambda_2^u(\omega, t) = \lambda_2^w(\omega, t), \qquad \mu_2^u(\omega, t) = \mu_2^w(\omega, t). \tag{72}$$

We now assume that

$$x_1^u(0) > 0.$$

Then, it is easy to check that (62) holds, that

$$T^{uw} = T_1^u = T_1^w,$$

and that

$$\int_0^{T^{uw}} k^u(s^u(\omega, \tau))\, d\tau \leq \int_0^{T^{uw}} k^w(s^w(\omega, \tau))\, d\tau + c_1\Delta. \tag{73}$$

Consider the set $A \subset \Omega$ of those ω such that α becomes $(1, 0, 0)$ within $\Delta/2(\lambda^* + \mu_1^*)$ time units and stays equal to $(1, 0, 0)$ for at least $(N_1 + N_2)/\lambda^*$ additional time units. For any $\omega \in A$, we will have

$$\int_0^{T^{uw}} k^u(s^u(\omega, \tau))\, d\tau \leq \int_0^{T^{uw}} k^w(s^w(\omega, \tau))\, d\tau + c_1\Delta/2. \tag{74}$$

Taking expectations and following the same procedure as in Part (a), we establish the desired result for

$$x_1^u(0) > 0.$$

Now, if

$$x_1^u(0) = 0,$$

the statements

$$T_1^u = T_1^w$$

and $u \in U_R$ are not necessarily true, and the above argument fails. However, a sample path argument of the same flavor easily shows that V^* is continuous at

$$x_1 = 0.$$

Since (57) has been proved for

$$x_1 \neq 0,$$

it follows, by continuity, that (57) is also true at

$$x_1 = 0. \qquad \square$$

Corollary 5.1. Let $0 < x_i < N_i$. Then,

$$-c_1 < \lim_{\Delta \downarrow 0} \frac{V_\alpha^*(x_1 + \Delta_1, x_2) - V_\alpha^*(x_1, x_2)}{\Delta} < M. \tag{75}$$

The right-hand-side inequality also holds if $x_1 = 0$. Inequalities (75) also hold with $\Delta \uparrow 0$, $0 < x_i < N_i$. In that case, the left-hand-side inequality also holds for $x_i = N_i$. Finally, the same results are obtained if we consider the slopes with respect to x_2.

We have shown (Theorem 3.3) that the optimal cost g^* is independent of the initial state. We now view g^* as a function of the parameters of the system and examine the form of the functional dependence. In particular, we consider the dependence of g^* on the buffer sizes N_1, N_2 as well as the machine capacities λ^*, μ_1^*, μ_2^*. To illustrate this dependence, we write $g^*(N_1, N_2, \lambda^*, \mu_1^*, \mu_2^*)$. Our result states that g^* is a convex function of its parameters. The proof is similar to the proof of Theorem 5.1, and hence it is omitted.

Theorem 5.3. $g^*(N_1, N_2, \lambda^*, \mu_1^*, \mu_2^*)$ is a convex function of its arguments.

6. Necessary Conditions for Optimality

In this section, we prove the necessary conditions for optimality that will be used in the next section. We start by demonstrating that V^* is in the domain of $\mathscr{L}^u$ (defined in Section 2), for any admissible control law u. Using the convexity and Lipschitz continuity of V_α^*, we get the following lemma.

Lemma 6.1. Let $x(t)$ be a trajectory in $[0, N_1] \times [0, N_2]$, and suppose that

$$d \equiv \lim_{t \downarrow 0} (x(t) - x(0))/t$$

exists. Then,

$$\lim_{t\downarrow 0} \frac{V^*_\alpha(x(t)) - V^*_\alpha(x(0))}{t} = \lim_{t\downarrow 0} \frac{V^*_\alpha(x(0)+td) - V^*_\alpha(x(0))}{t}. \tag{76}$$

The existence of the limits is part of the result.

Let $u \in U_R$. As in Section 2, for any fixed α, let $x^u_\alpha(t)$ be the value of x at time t if no jump occurs until time t. By right-continuity and boundedness of the control variables, the trajectory $x^u_\alpha(t)$ possesses right-hand-side derivatives. Then, Lemma 6.1 and (10) imply the following theorem.

Theorem 6.1. V^* belongs to the domain of $\mathscr{L}^u$, for any $u \in U_R$.

Lemma 6.2. For any $\epsilon > 0$, there exists some $w \in U_R$ such that

$$\hat{V}^w(s) \le V^*(s) + \epsilon, \qquad \forall s \in S. \tag{77}$$

Proof. *Outline.* Partition the state space S into a finite collection of disjoint and small enough rectangles $R_1, \ldots, R_k$. Choose a state $s_j \in R_j$ and a control law $w_j \in U_R$ such that

$$\hat{V}^{w_j}(s_j) \le V^*(s_j) + \epsilon_1,$$

where ϵ_1 is small enough. Define w_j for all initial states in the rest of the rectangle R_j, so that all sample paths $s^{w_j}(\omega, t)$, $w_j(\omega, t)$ starting from R_j stay close enough. In particular, choose w_j in such a way that $s^{w_j}(\omega, t)$ and $\mu_i^{w_j}(\omega, t)$ are continuous functions of the initial state, for any ω, t. In that case,

$$|\hat{V}^{w_j}(s_j) - \hat{V}^{w_j}(s)| \le \epsilon_2, \qquad \forall s \in R_j,$$

for some small enough ϵ_2. Then, define a control law w by lumping together control laws w_j, $j = 1, \ldots, k$. Given that V^* is Lipschitz continuous and since ϵ_1, ϵ_2 may be chosen as small as desired, w satisfies (77). □

Lemma 6.3. The following result holds:

$$\lim_{t\downarrow 0} (1/t) E[V^*(s^u(t))\chi] = 0,$$

where χ is the indicator function of the event $T^u_1 \le t$.

Proof. Observe that

$$\lim_{t\downarrow 0} \frac{E[V^*(s^u(t))\chi]}{t} = \lim_{t\downarrow 0} \frac{\Pr(T^u_1 \le t)}{t} \lim_{t\downarrow 0} E[V^*(s^u(t)) | T^u_1 \le t]. \tag{78}$$

The first limit in the r.h.s. of (78) is bounded by the transition rates p_i, r_i; the second limit is equal to

$$V^*(s_0)=0,$$

unless a jump occurs in $[T_1^u, t]$, which is an event whose probability goes to zero, as t goes to zero. □

Lemma 6.4. The following result holds:

$$\mathscr{L}^u V^* + k^u \geq g^*, \qquad \forall s \in S, \qquad \forall u \in U_R.$$

Proof. Let $u \in U_R$, $t>0$, $s \in S$ be fixed and let w be the control law of Lemma 6.2. Consider a new control law v with the following properties: v coincides with u up to time t; at that time the past is forgotten, and the process is restarted using control law w. Then,

$$\begin{aligned} V^*(s) \leq \hat{V}^v(s) &= E\left[\int_0^{\min\{t,T_1^u\}} (k^u(s^u(\tau))-g^*)\,d\tau\right] \\ &\quad + E[\hat{V}^w(s^u(t))(1-\chi)] \\ &\leq E\left[\int_0^{\min\{t,T_1^u\}} (k^u(s^u(\tau))-g^*)\,d\tau\right] + E[V^*(s^u(t))(1-\chi)]+\epsilon. \end{aligned} \tag{79}$$

Since ϵ was arbitrary, we may let $\epsilon\downarrow 0$, then divide by t, take the limit, as $t\downarrow 0$, and invoke Lemma 6.3 to obtain

$$\begin{aligned} \mathscr{L}^u V^*(s) &= \lim_{t\downarrow 0} \frac{E[V^*(s^u(t))]-V^*(s)}{t} \\ &\geq -\lim_{t\downarrow 0} \frac{1}{t} E\left[\int_0^{\min\{t,T_1^u\}} (k^u(s^u(\tau))-g^*)\,d\tau\right] \\ &= -k^u(s)+g^*. \end{aligned} \tag{80}$$

The last equality follows from the right-continuity of k^u and the dominated convergence theorem. □

Theorem 6.2. If $u \in U_R$ is everywhere optimal, then

$$\mathscr{L}^u V^* + k^u \leq \mathscr{L}^w V^* + k^w, \qquad \forall w \in U_R, \forall s \in S. \tag{81}$$

Proof. We note that

$$V^* = \hat{V}^u,$$

and we start with the equation

$$\mathscr{L}^u V^* + k^u = g^*,$$

which is derived in the same way as Lemma 6.4, except that inequalities become equalities. We then use Lemma 6.4 to get

$$\mathscr{L}^u V^* + k^u = g^* \leq \mathscr{L}^w V^* + k^w, \qquad \text{for all } w \in U_R. \qquad \square$$

7. Characterization of Optimal Control Laws

In this section, we use the optimality conditions (Theorem 6.2) together with the properties of V^* (Theorems 5.1 and 5.2) to characterize everywhere optimal control laws. We mainly consider Markovian control laws for which the control variables λ_i, μ_i (and hence the cost function k) can be viewed as functions of the state. The first two theorems (Theorems 7.1 and 7.2) state that the machines should be always operated at the maximum rate allowed, as it should be expected. Theorem 7.4 is much more substantial, as it characterizes the way that the flow through machine M_0 should be split.

Theorem 7.1. If $u \in U_M \cap U_R$ is everywhere optimal, then

(a) $\mu_i^u(x, \alpha) = \alpha_i \mu_i^*, \qquad \text{if } x_i \neq 0,\ i = 1, 2,$ (82)

(b) $\mu_i^u(x, \alpha) = \alpha_i \min\{\mu_i^*, \lambda_i^u(x, \alpha)\}, \qquad \text{if } x_i = 0,\ i = 1, 2.$ (83)

Proof. Let $u \in U_M \cap U_R$ be everywhere optimal. Then, u must minimize

$$(\mathscr{L}^u V^* + k^u)(s), \qquad \forall s \in S.$$

Using (10), Lemma 6.1, and dropping those terms that do not depend on u, we conclude that μ_1^u, μ_2^u must be chosen so as to minimize

$$\lim_{\Delta \downarrow 0} \frac{V_\alpha^*(x_1 + (\lambda_1^u - \mu_1^u)\Delta, x_2 + (\lambda_2^u - \mu_2^u)\Delta) - V_\alpha^*(x_1, x_2)}{\Delta} - c_1\mu_1^u - c_2\mu_2^u. \tag{84}$$

Let

$$x_i \neq 0 \quad \text{and} \quad \alpha_i \neq 0.$$

By Corollary 5.1, the slopes of V_α^* are strictly larger than $-c_1$, $-c_2$ and, as a result, μ_i^u must be set to its highest admissible value, which is $\alpha_i \mu_i^*$, thus proving (82).

Now, let

$x_i = 0,$

and suppose that (83) is violated. In that case,

$\mu_i^u(x, \alpha) < \mu^*;$

also, the trajectory $s^u(t)$ enters immediately the region in which

$x_i > 0,$

and therefore

$\mu_i^u(t) = \mu^*,$

for all small enough positive times. Hence, the sample path $\mu_i^u(t)$ is not right-continuous and u is not an admissible control law. □

Theorem 7.2. If V_α^* is strictly decreasing in each variable (or in particular, by Theorem 5.2, if $k^u = -c_1\mu_1^u - c_2\mu_2^u$) and if $u \in U_M \cap U_R$ is everywhere optimal, then

$$\text{(a)} \quad \lambda^u(x, \alpha) = \alpha_0 \lambda^*, \qquad \text{if } x \neq (N_1, N_2), \tag{85}$$

$$\text{(b)} \quad \lambda^u(x, \alpha) = \alpha_o \min\{\lambda^*, \alpha_1\mu_1^* + \alpha_2\mu_2^*\}, \qquad \text{if } x = (N_1, N_2). \tag{86}$$

Proof. Let $u \in U_M \cap U_R$ be everywhere optimal. Then, u must minimize

$(\mathscr{L}^u V^* + k^u)(s), \qquad \forall s \in S.$

Theorem 7.1 determines μ_i^u uniquely and k^u is no more dependent on u. By dropping those terms that do not depend on u, we conclude that λ_1^u, λ_2^u must be chosen so as to minimize

$$\lim_{\Delta \downarrow 0} \frac{V_\alpha^*(x_1 + (\lambda_1^u - \mu_1^u)\Delta, x_2 + (\lambda_2^u - \mu_2^u)\Delta) - V_\alpha^*(x_1, x_2)}{\Delta}. \tag{87}$$

Let

$(x_1, x_2) \neq (N_1, N_2).$

Since V_α^* is strictly decreasing,

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

must be set equal to its highest admissible value, which is $\alpha_0\lambda^*$. If

$(x_1, x_2) = (N_1, N_2),$

Eq. (86) follows by the right-continuity requirement on admissible control laws, in the same way as in the proof of Theorem 7.1. □

If $f_\alpha(x)$ is nonzero and large enough, compared to c_1, c_2, then V_α^* need not be decreasing. Equivalently, the penalty for having large buffer levels will be larger than the future payoff in terms of increased production. In that case, for any optimal control law, λ^u should be set to zero whenever the buffer levels exceed some threshold.

From now on, we assume that V_α^* is strictly decreasing, for all α. Theorems 7.1 and 7.2 define μ_1^u, μ_2^u, λ^u uniquely. It only remains to decide on how λ^u is going to be split. It is here that convexity of V_α^* plays a major role. Note that there is no such decision to be made whenever $\alpha_0=0$. We will therefore assume that $\alpha_0 \neq 0$.

Let

$$h_\alpha(\rho)=\inf\{V_\alpha^*(x_1,x_2): x_1+x_2=\rho\}, \qquad \rho\in[0, N_1+N_2]. \tag{88}$$

Because of the continuity of V_α^*, the infimum is attained, for each ρ, and the set

$$H_\alpha(\rho)=\{(x_1,x_2): x_1+x_2=\rho,\ V_\alpha^*(x_1,x_2)=h_\alpha(\rho)\} \tag{89}$$

is nonempty. Finally, let (see Fig. 2)

$$H_\alpha=\bigcup_{\rho\in[0,N_1+N_2]} H_\alpha(\rho). \tag{90}$$

We should point out that the points on the x_1-axis to the left of point B (point C in particular) belong to H_α.

From the convexity and continuity of V^*, we can easily obtain the following theorem.

Theorem 7.3. (a) $H_\alpha(\rho)$ is connected for any ρ, α.

(b) If V^* is strictly convex, then $H_\alpha(\rho)$ is a singleton.

(c) H_α is closed and connected.

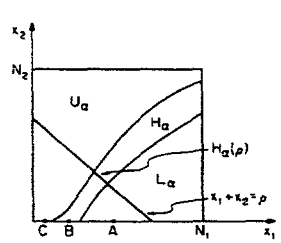

Fig. 2. Regions related to the optimality conditions.

Now, let

$$U_\alpha(\rho)=\{(x_1,x_2): x_1+x_2=\rho, x_1-x_2<y_1-y_2, \quad \forall(y_1,y_2)\in H_\alpha(\rho)\}, \tag{91}$$

$$L_\alpha(\rho)=\{(x_1,x_2): x_1+x_2=\rho, x_1-x_2>y_1-y_2, \quad \forall(y_1,y_2)\in H_\alpha(\rho)\}, \tag{92}$$

and (see Fig. 2)

$$U_\alpha=\bigcup_{\rho\in[0,N_1+N_2]} U_\alpha(\rho), \qquad L_\alpha=\bigcup_{\rho\in[0,N_1+N_2]} L_\alpha(\rho). \tag{93}$$

Since $H_\alpha(\rho)$ is connected, it follows that

$$U_\alpha(\rho)\cup L_\alpha(\rho)\cup H_\alpha(\rho)=\{(x_1,x_2): x_1+x_2=\rho\}; \tag{94}$$

consequently,

$$U_\alpha\cup L_\alpha\cup H_\alpha=[0,N_1]\times[0,N_2]. \tag{95}$$

Finally, note that, keeping $(x_1,x_2)\in U_\alpha$ fixed, the function $V_\alpha^*(x_1+\Delta, x_2-\Delta)$ is a strictly decreasing function of Δ (for small enough Δ), because of the convexity of V_α^* and the definition of U_α. With this remark, we have the following characterization of the optimal values of λ_1^u, λ_2^u in the interior of the state space.

Theorem 7.4. If V_α^* satisfies (57) with $M=0$, $u\in U_M\cap U_R$ is everywhere optimal, and x is in the interior of $[0,N_1]\times[0,N_2]$, the following results hold:

(a) if $x\in U_\alpha$, then $\lambda_1^u(x,\alpha)=\lambda^*\alpha_0$;
(b) if $x\in L_\alpha$, then $\lambda_2^u(x,\alpha)=\lambda^*\alpha_0$.

Proof. Let x belong to the interior of U_α. We must again minimize the expression (87). Because of the monotonicity property mentioned in the last remark, it follows that λ_1^u has to be set equal to its maximum value $\alpha_0\lambda^*$. Part (b) follows from a symmetrical argument. □

We now discuss the optimality conditions on the separating set H_α. We assume that V_α^* is strictly convex and (by Theorem 7.3) $H_\alpha(\rho)$ is a singleton, for any fixed ρ. Equivalently, H_α is a continuous curve. According to the remarks following Theorem 5.1, V_α^* is always strictly convex; but, since we have not given a proof of this fact, we introduce it as an assumption.

Fix $(x_1,x_2)\in H_\alpha$, and suppose that

$$0<x_i<N_i, \qquad i=1,2,$$

(interior point). Given a control law u, let

$$A(u)=\{\tau>0: x_\alpha^u(\tau)\in U_\alpha\},$$

$$B(u)=\{\tau>0: x_\alpha^u(\tau)\in L_\alpha\},$$

where $x_\alpha^u(\tau)$ is the path followed starting from $((x_1, x_2), \alpha)$ if no jump of α occurs. We distinguish four cases.

(a) Suppose that, for all $u\in U_M\cap U_R$, time $t=0$ is a limit point of $A(u)$. For all $\tau\in A(u)$, we have

$$\lambda_1^u(\tau)=\lambda^*,$$

by Theorem 7.4. Then, by right continuity of $\lambda_1^u(t)$, we must have

$$\lambda_1^u(0)=\lambda^*.$$

(b) Similarly, if for all $u\in U_M\cap U_R$, $t=0$ is a limit point of $B(u)$, we must have

$$\lambda_2^u(0)=\lambda^*.$$

(c) If $t=0$ is a limit point of both $A(u)$ and $B(u)$, for all $u\in U_M\cap U_R$, then no everywhere optimal control law exists. Fortunately, this will never be the case, if H_α is a sufficiently smooth curve.

(d) Finally, suppose that there exists some u such that $t=0$ is not a limit point of either $A(u)$ or $B(u)$. In that case,

$$x_\alpha^u(t)\in H_\alpha, \qquad \forall t\in[0,\Delta],$$

for some small enough $\Delta>0$. An argument similar to that in Theorem 7.4 will show that this control law satisfies the optimality conditions at (x_1, x_2). Such a control law travels on H_α, i.e., stays on the deepest part of the valley-like convex function V_α^*.

The optimality conditions on the boundaries are slightly more complicated, because the constraints on λ_i, μ_i are interrelated through the requirement that x_i stays in $[0, N_i]$. The exact form of these conditions depends, in general, on the relative magnitudes of the parameters λ^*, μ_1^*, μ_2^*. However, for any particular problem, Theorem 6.2 leads to an unambiguous selection of the values of the control variables.

8. Conclusions and Generalizations

Let us start by pointing out the main properties of our queueing system on which our development has been based:

(i) We first have the existence of a special state, which is recurrent when we restrict ourselves to a class of control laws that have equally good performance as the original set of admissible control laws.

(ii) We have the convexity of the optimal cost-to-go function, which only depends on the following facts: (a) the state space is convex; (b) the set of admissible values of the control variables is convex; and (c) the cost function is convex.

Our methodology is therefore applicable, with minor adjustments, to the large class of linear dynamical systems in which the above-enumerated properties are present.

We now indicate a few alternative configurations for which all steps of our development would remain valid. We may let the buffer capacities be infinite. Then, provided that storage costs increase fast enough with x_i, it is still possible to obtain a recurrence result. The convexity theorem would be still valid. A few derivations would need some more care, because V^* and f will no more be bounded functions of the state space, but the main results of Section 7 would remain unchanged.

We may also have three (instead of two) downstream buffers and machines, in which case the state space is three-dimensional. Convexity of V^* and the optimality conditions then imply that, for any fixed α, the three-dimensional state space is divided into three regions, separated by three two-dimensional surfaces that intersect on a one-dimensional curve. In each of the three regions, all material is to be routed to a unique buffer. The switching surfaces have interpretations similar to the switching curves H_α of Section 7.

As pointed out earlier, our recurrence results (Theorem 3.2) have been based on the assumption that the lead machine is unreliable, $p_0 \neq 0$. While this is a convenient assumption, it is not a necessary one, except that, if $p_0 = 0$, the reference state s_0 should be differently chosen. This choice should be problem specific and would not present any difficulties for most interesting cases. The only difference that arises when $p_0 = 0$ is that V^* need not be strictly convex and the separating set H_α could even be the entire state space (Ref. 23, Chapter 6).

As another variation of our problem, we could include a nonlinear, convex, and increasing cost on the utilization rates of the machines, to penalize utilization at or near capacity limits. The rationale behind this cost criterion is that high utilization rates are generally undesirable (in the long run). In that case, V^* would still be convex, but Theorem 7.1 would no longer hold. Rather, the optimal utilization rates μ_i^u of the downstream machines would be an increasing function of the buffer levels.

The next issue of concern is the computation of V^* and the generation of an optimal control law. One conceivable procedure (resembling the

Howard algorithm) is to evaluate V^u, for a fixed Markovian u, by solving the equation

$$\mathscr{L}^u V^u + k^u = g^u$$

for V^u and g^u. This equation has a unique solution within an additive constant for V^u. It really consists of eight coupled first-order, linear, partial differential equations with nonconstant coefficients and can only be solved numerically. Based on V^u, we may generate a control law w which improves performance by minimizing $\mathscr{L}^w V^u + k^w$, and so on. In practice, any such algorithm would involve a discretization procedure, so it might be preferable to formulate the problem on a discrete state space. In that case, the successive approximation algorithm (or accelerated versions of it) could yield a solution relatively efficiently.

An alternative iterative optimizing algorithm, based on an equivalent deterministic optimal control problem, has been also suggested in Ref. 24 (see also Refs. 17 and 23 for related ideas).

The drawback of any numerical procedure is that the computational requirements become immense, even for moderate sizes of the state space (e.g., $N_1 = N_2 = 20$, see Ref. 15). Fortunately, the existing numerical evidence shows that the performance functional is not very sensitive to variations of the dividing curve, so that rough approximations may be particularly useful. Estimates of the asymptotic slope of H_α, as N_1, N_2 increase, as well as of the intercepts of H_α with the axes $x_i = 0$ would be very helpful for obtaining an acceptable suboptimal control law.

References

1. KLEINROCK, L., *Queueing Systems, Vol. 1*, John Wiley, New York, New York, 1975.
2. KLEINROCK, L., *Queueing Systems, Vol. 2*, John Wiley, New York, New York, 1976.
3. GALLAGER, R. G., *A Minimum Delay Routing Algorithm Using Distributed Computation*, IEEE Transactions on Communications, Vol. COM-25, pp. 73–85, 1977.
4. KIMEMIA, J. G., and GERSHWIN, S. B., *Multicommodity Network Flow Optimization in Flexible Manufacturing Systems*, MIT, Electronic Systems Laboratory, Report No. ESL-FR-834-2, 1980.
5. CRABILL, T., GROSS, D., and MAGAZINE, M. J., *A Classified Bibliography of Research on Optimal Design and Control of Queues*, Operations Research, Vol. 25, pp. 219–232, 1977.

6. SOBEL, M. J., *Optimal Operation of Queues*, Mathematical Methods in Queueing Theory, Edited by A. B. Clarke, Springer-Verlag, New York, New York, 1974.
7. CRABILL, T., *Optimal Control of Service Facility with Variable Exponential Service Time and Constant Arrival Rate*, Management Science, Vol. 18, pp. 560–566, 1977.
8. GALLISH, E., *On Monotonic Optimal Policies in a Queueing Model of M/G/1 Type with Controllable Service Time Distribution*, Advances in Applied Probability, Vol. 11, pp. 870–887, 1979.
9. ROSBERG, Z., VARAIYA, P., and WALRAND, J., *Optimal Control of Service in Tandem Queues*, IEEE Transactions on Automatic Control, Vol. AC-27, pp. 600–610, 1982.
10. SCHWARTZ, B., *Queueing Models with Lane Selection: A New Class of Problems*, Operations Research, Vol. 22, pp. 331–339, 1974.
11. FOSCHINI, G. J., *On Heavy Traffic Diffusion Analysis and Dynamic Routing in Packet Switched Networks*, Computer Performance, Edited by K. M. Chandy and M. Reiser, North-Holland, Amsterdam, Holland, 1977.
12. FOSCHINI, G. J., and SALZ, J., *A Basic Dynamic Routing Problem and Diffusion*, IEEE Transactions on Communications, Vol. COM-26, pp. 320–327, 1978.
13. EPHREMIDES, A., VARAIYA, P., and WALRAND, J., *A Simple Dynamic Routing Problem*, IEEE Transactions on Automatic Control, Vol. AC-25, pp. 690–693, 1980.
14. DEUERMEYER, B. L., and PIERSKALLA, W. P., *A Byproduct Production System with an Alternative*, Management Science, Vol. 24, pp. 1373–1383, 1978.
15. HAHNE, E., *Dynamic Routing in an Unreliable Manufacturing Network with Limited Storage*, MIT, Laboratory for Information and Decision Systems, Report No. LIDS-TH-1063, 1981.
16. OLSDER, G. J., and SURI, R., *Time-Optimal Control of Parts Routing in a Manufacturing System with Failure Prone Machines*, Proceedings of the 19th IEEE Conference on Decision and Control, Albuquerque, New Mexico, 1980.
17. KIMEMIA, J. G., and GERSHWIN, S. B., *An Algorithm for the Computer Control of Production in a Flexible Manufacturing System*, Proceedings of the 20th IEEE Conference on Decision and Control, San Diego, California, 1981.
18. KOENIGSBERG, E., *Production Lines and Internal Storage: A Review*, Management Science, Vol. 5, pp. 410–433, 1959.
19. SARMA, V. V. S., and ALAM, M., *Optimal Maintenance Policies for Machines Subject to Deterioration and Intermittent Breakdowns*, IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-4, pp. 396–398, 1975.
20. GERSHWIN, S. B., and SCHICK, I. C., *Modelling and Analysis of Two-Stage and Three-Stage Transfer Lines with Unreliable Machines and Finite Buffers*, MIT, Laboratory for Information and Decision Systems, Report No. LIDS-R-979, 1980.
21. KUSHNER, H. J., *Optimality Conditions for the Average Cost per Unit Time Problem with a Diffusion Model*, SIAM Journal on Control and Optimization, Vol. 16, pp. 330–346, 1978.
22. LOEVE, M., *Probability Theory*, Springer-Verlag, New York, New York, 1977.

23. TSITSIKLIS, J. N., *Optimal Dynamic Routing in an Unreliable Manufacturing System*, MIT, Laboratory for Information and Decision Systems, Report No. LIDS-TH-1069, 1981.
24. RISHEL, R., *Dynamic Programming and Minimum Principles for Systems with Jump Markov Disturbances*, SIAM Journal on Control, Vol. 13, pp. 338–371, 1975.