OPTIMAL CONTROL OF SERVICE IN TANDEM QUEUES

by

Z. Rosberg, P. Varaiya and J. Walrand

Memorandum No. UCB/ERL M80/42

2 September 1980

OPTIMAL CONTROL OF SERVICE IN TANDEM QUEUES

by

Z. Rosberg, P. Varaiya and J. Walrand

Memorandum No. UCB/ERL M80/42

2 September 1980

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

OPTIMAL CONTROL OF SERVICE IN TANDEM QUEUES[1]

Z. Rosberg[2], P. Varaiya, and J. Walrand[3],

Department of Electrical Engineering and Computer Science,
and Electronics Research Laboratory,
University of California, Berkeley, CA 94720

ABSTRACT

Customers arrive in a Poisson stream into a network consisting of two M/M/1 service stations in tandem. The service rate $u \in [0,a]$ at station 1 is to be selected as a function of the state (x_1, x_2) where x_i is the number of customers at station i, so as to minimize the expected total discounted or average cost corresponding to the instantaneous cost $c_1 x_1 + c_2 x_2$. The optimal policy is of the form $u = a$ or $u = 0$ according as $x_1 < S(x_2)$ or $x_1 \geq S(X_2)$ and S is a switching function. For the case of discounted cost the optimal process can be nonergodic, but is is ergodic for the case of average cost.

[1]Research sponsored in part by National Science Foundation under Grant ENG-7903879 A01.

[2]Current address: Department of Computer Science, Technion, Haifa, Israel.

[3]Current address: School of Electrical Engineering, Cornell University, Ithaca, N.Y. 14853.

1. Introduction

The search for the optimal control policy in queuing control problems can be simplified if it is known that the optimal policy has a special structure e.g. 'bang-bang' or 'switch-over'. Queuing models consisting of a single service station with controllable service times have been studied for various cost functions and shown to possess switch-over optimal policies.

An extensive bibliography on queuing control models can be found in Grabill et al (1977). The analysis of an M/M/1 model by Lippman (1973,1973a) extends earlier results of Grabill (1972). In his study of an M/G/1 model, Gallisch (1979) gives conditions on the cost function and service time distribution which imply switch-over optimal policies. He extends results obtained in Schassberger (1965) where the service cost is linear and waiting time cost is not allowed, and in Tijms (1975) where waiting time cost is linear. Formulas and properties of the long run average cost under switch-over policies are derived in Tijms et al (1978) and Weiss (1979). The former unifies several more specific models for controlling service and arrival rates in an M/G/1 model. The formulas derived there for the cost may be used to obtain operating characteristics of the system. Weiss considers batch service and gives computable upper and lower bounds for the optimal switch-over policy and an algorithm for finding the optimal policy.

There does not seem to be any study of optimal policies when there are two or more connected service stations. This paper analyzes the smiplest such case where customers in a Poisson stream enter a network consisting of two exponential servers in tandem. The service rate

$u \in [0,a]$ at station 1 is to be selected as a function of the state (x_1,x_2) where x_i is the number of customers in station i. The instantaneous cost is linear in the waiting times at the two stations, $c_1x_1 + c_2x_2$, and we show that the policies which minimize the total expected discounted or long run average cost are switch-over. That is, the optimal policy is characterized by a switching curve S: $u = a$ or $u = 0$ according as $x_1 \geq S(x_2)$ or $x_1 < S(x_2)$. The optimal S can be interpreted as the condition that $x_1 = S(x_2)$ iff in the state (x_1,x_2) the marginal increase in the expected cost is the same whether a marginal customer is added to the queue at 1 or at 2. This interpretation also explains the fact that the optimal S is monotonically increasing.

The analysis of the problem differs in two respects from the previously mentioned studies of a single server. The first concerns the convexity of the value function. In a single server model the state space is one dimensional and showing convexity of the value function using the optimality equations is usually sufficient for proving switch-over of the optimal policy. In our case the state space has dimension two, and the proof of convexity does not directly follow from the optimality equations. Our proof is based upon constructing an equivalent linear programming problem and deducing convexity from duality theory. The second difference concerns the analysis of the case of long run average cost. In single server models, it is usually trivial to show ergodicity of the optimal process for the discounted case. The average cost is then treated as a limit of the discounted case as in Lippman (1973). In our problem the optimal process for the discounted case can be nonergodic (in fact all states can be transient), and so the case of average cost cannot be treated as the limit of the discounted case. Instead we follow the unusual procedure of moving to the average cost from the finite horizon problem.

The paper is organized as follows. In the next section the continuous time problem is transformed into one with discrete time by looking at an embedded Markov chain, and the optimality conditions are written down. In section 3 the equivalent linear programming is introduced to prove convexity of the value function. The optimality conditions and convexity are used in section 4 to show that the optimal policy is switch-over for the finite horizon and discounted problems. The average cost case is treated in section 5 where a separate argument is given to show ergodicity.

2. The equivalent discrete time problem

Customers arrive into station 1 in a Poisson stream with constant rate λ . The rate at time t of the exponential server in station 1 can be selected to be any number u_t in $[0,a]$. Upon completing service at 1 a customer joins the queue at station 2 which is served by another exponential server with constant service rate μ. Let x_{it} be the number of customers at time t in station i, the customer in service included, and let $x_t = (x_{1t}, x_{2t})$ denote the state at t. u_t is to be selected knowing x_t.

The cost incurred per unit time in state x is $c'x = c_1x_1 + c_2x_2$ where $c_1 > 0$, $c_2 > 0$ are fixed. Let $\alpha \geq 0$ be the interest rate used for discounting future cost i.e. the present value of cost c incurred at time t is $ce^{-\alpha t}$. Let $V_t^\alpha(x)$ be the minimum acheivable cost when the time horizon is $t \geq 0$ and $x_0 = x$.

Dynamic programming considerations lead to the following optimality conditions for V_t^α.

$$V_{t+dt}^\alpha(x) = c'xdt + e^{-\alpha dt} \inf_{0\le u\le a} \{\lambda dt V_t^\alpha(Ax) + \mu dt V_t^\alpha(Dx) +$$

$$+ udt V_t^\alpha(Tx) + [1-(\lambda+\mu+u)dt]V_t^\alpha(x) + o(dt)\} \quad ,$$

$$= c'xdt + e^{-\alpha dt}\{\lambda dt V_t^\alpha(Ax) + \mu dt V_t^\alpha(Dx) + [1 - (\lambda+\mu+u)dt]V_t^\alpha(x)\}$$

$$+ e^{-\alpha dt} \inf_{0\le u\le a} udt[V_t^\alpha(Tx) - V_t^\alpha(x)] + o(dt). \tag{2.1}$$

In (2.1), A,D,T are functions representing an arrival at station 1, a departure from station 2 and a service completion at 1. That is, $A(x_1,x_2) = (x_1+1,x_2)$; $D(x_1,x_2) = (x_1,(x_2-1)^+)$; $T(x_1,x_2) = (x_1-1,x_2+1)$ or (x_1,x_2) according as $x_1 > 0$ or $x_1 = 0$.

Observe that by writing the optimality conditions in this way we are adopting the convention that an idle server (idle by choice or because of lack of customers) is serving a dummy customer who incurs no waiting cost and who never leaves the station. This is permissible because of the memoryless nature of the exponential service time.

From (2.1) it follows that the policy given by $u_s = 0$ if $V_{t-s}^\alpha(Tx_s) > V_{t-s}^\alpha(x_s)$ and $u_t = a$ if $V_{t-s}^\alpha(Tx_s) \le V_{t-s}^\alpha(x_s)$, is optimal. Hence we can limit attention to policies which are 'bang-bang' i.e. take values in {0,a}. To convert the problem into one in discrete time we henceforth limit ourselves further to those bang-bang policies which change values only at a transition epoch, including as transitions those due to service completions for dummy customers. Call this class of policies P. For the infinite horizon problem this is an inessential restriction since the optimal policy is stationary and hence it is in P. For the finite horizon this is a restriction.

The transition epochs $0 = t_0 < t_1 < t_2 \ldots < t_n \ldots$ are the same for all policies in P. In fact the inter-epoch intervals are independent and have the same distribution, $\text{Prob}\{t_{k+1} - t_k > t\} = \exp -t(\lambda+\mu+a)$. The cost incurred by a policy (in P) over the random interval $[0,t_n]$ and with initial state x is

$$E\int_0^{t_n} e^{-\alpha t} c'x_t dt = E_x \sum_0^{n-1} \int_{t_k}^{t_{k+1}} e^{-\alpha t} c'x_{t_k} dt$$

$$= E_x \Sigma\{c'x_{t_k} E\int_{t_k}^{t_{k+1}} e^{-\alpha t} dt\} = E_x \Sigma\, c'x_{t_k} \frac{1}{\alpha} E[e^{-\alpha t_k} - e^{-\alpha t_{k+1}}]. \tag{2.2}$$

Since the $t_{k+1} - t_k$ are i.i.d., therefore $E\, e^{-\alpha t_k} = \beta^k$ where

$$\beta = E\, e^{-\alpha t_1} = \int_0^{\infty} e^{-\alpha t} (\lambda+\mu+a) e^{-(\lambda+\mu+a)t} dt = \frac{\lambda+\mu+a}{\alpha+\lambda+\mu+a} \tag{2.3}$$

Substituting we see that the cost (2.2) equals

$$\frac{1-\beta}{\alpha} E_x \sum_{k=0}^{n-1} \beta^k c'x_{t_k} \quad , \tag{2.4}$$

provided that $\beta < 1$, whereas if $\beta = 1$ then it equals $E_x\, \Sigma c'x_k$. It is convenient to ignore the constant factor and take as the cost

$$E_x \sum_{k=0}^{n-1} \beta^k c'x_{t_k} \tag{2.5}$$

which is valid for $0 \leq \beta \leq 1$. Writing $x_k = x_{t_k}$, we see that the search for the optimal policy in P over the random interval $[0,t_n]$ is equivalent to finding the best policy over the horizon $[0,n]$ for the following discrete time Markov decision problem.

The state space is N^2 where $N = \{0,1,2,..\}$, the action space is $\{0,1\}$ and the one-step transition probabilities are

$$p(x_{k+1}|x_k,z_k) = \begin{cases} \lambda \text{ if } x_{k+1} = Ax_k, \\ \mu \text{ if } x_{k+1} = Dx_k, \\ a \text{ if } x_{k+1} = Tx_k,\ z_k = 1 \\ a \text{ if } x_{k+1} = x_k,\ z_k = 0 \end{cases} \tag{2.5}$$

In (2.5) the (λ,μ,a) are proportional to the (λ,μ,a) in (2.1) and normalized so that $\lambda + \mu + a = 1$. $z_k = 1$, respectively 0, corresponds to $\mu_t \equiv a$, respectively 0, during the interval $[t_k,t_{k+1}]$. $1 - z_k$ can be interpreted as the probability of recycling a customer as in Figure 1.

The cost is given by $E_x \sum^{n-1} \beta^k c'x_k$. Let V_n^β denote the minimum achievable cost. It is characterized by these optimality conditions: $V_1^\beta(x) = c'x$ and

$$V_{n+1}^\beta(x) = c'x + \beta\{\lambda V_n^\beta(Ax) + \mu V_n^\beta(Dx) + a \min(V_n^\beta(x),V_n^\beta(Tx))\} \quad . \tag{2.6}$$

Furthermore, when there are k steps to go and the current state is x, the optimal action is $z = 0$ if $V_k^\beta(Tx) > V_k^\beta(x)$ and $z = 1$ if $V_k^\beta(Tx) \leq V_k^\beta(x)$.

Since $c'x \geq 0$, therefore $V_n^\beta(x) \leq V_{n+1}^\beta(x)$ and so this limit exists:

$$V_\infty^\beta(x) = \lim_{n\to\infty} V_n^\beta(x) \tag{2.7}$$

<u>Lemma 2.1</u> $V_\infty^\beta < \infty$ for $\beta < 1$.

<u>Proof</u> Let $|x| = |x_1| + |x_2|$ and $|c| = \max(c_1, c_2)$. For any initial state x and any policy, the state x_k at time k must satisfy $|x_k| \leq |x| + k$. Hence

$$V_n^\beta(x) \leq \sum_{k=0}^{n-1} \beta^k |c|(|x|+k) < \frac{|c||x|}{1-\beta} + \frac{|c|\beta}{(1-\beta)^2} < \infty \quad . \tag{2.8}$$

¤

It follows from Theorem 1 in Lippman (1973) that for $\beta < 1$ the minimum cost for the infinite horizon,

$$V^\beta(x) = \min \sum_{k=0}^{\infty} E_x c'x_k \quad ,$$

is achieved by a stationary policy. Moreover V^β is the unique solution to the optimality condition

$$V^\beta(x) = c'x + \beta\{\lambda V^\beta(Ax) + \mu V^\beta(Dx) + a \min(V^\beta(x), V^\beta(Tx))\} \quad . \tag{2.9}$$

From (2.6), (2.7) and the uniqueness of V^β it follows that

$$V^\beta = V_\infty^\beta \quad . \tag{2.10}$$

3. $V_n^\beta(x)$ is convex

We have been unable to show that $V_n^\beta(x)$ is convex by using the relation (2.6). Instead, we give a linear programming problem equivalent to the Markov decision problem and then deduce convexity by duality. This technique appears novel and may be useful in studying other queuing problems.

The basic sample space for the Markov decision problem is denoted Ω^n and it consists of all sequences $\omega^n = (\omega_1,..,\omega_n)$ where $\omega_k \in \{A,T,D\}$; $\omega_k = A$, respectively T or D, according as the transition at step k is an arrival, respectively a service completion at station 1 or 2, including service completion for dummy customers. Let p denote the probability distribution on Ω^n. Under p the ω_k are i.i.d. and $\omega_k = A$, T or D with probability λ, a or μ. Let $\underline{\underline{F}}_k$ be the field on Ω^n generated by $\omega^k = (\omega_1,..,\omega_k)$. A <u>process</u> is any sequence of random variables $f = (f_1,..,f_n)$ such that f_k is adapted to $\underline{\underline{F}}_k$. Hence we can and will regard f_k as a function on Ω^k. Let

$$\xi(\omega) = \begin{cases} (1,0) & \text{if} \quad \omega = A \\ (-1,1) & \text{if} \quad \omega = T \\ (0,-1) & \text{if} \quad \omega = D \end{cases}$$

and let $\xi = (\xi_1,..,\xi_n)$ be the process given by $\xi_k(\omega^k) = \xi(\omega_k)$.

A <u>policy</u> is any process $z = (z_1,..,z_n)$ satisfying

$$z_k(\omega^k) \begin{cases} = 1 & \text{if} \quad \omega_k = A \\ \in [0,1] & \text{if} \quad \omega_k = T \\ \in [0,1] & \text{if} \quad \omega_k = D \end{cases} \tag{3.1}$$

Let Z be the set of all policies and Z_I the subset of policies which are integer-valued.

queues under the policy z and $J_n(z,x)$ is the n-step cost. (If x and u are not integer-valued but if $x \geq 0$, then we can interpret $\hat{x}$ as the evolution of the queuing system in which there are "fractional" customers.) The value function for the Markov decision problem is also given by

$$V_n^\beta(x) = \mathrm{Min}\{J_n(z,x) \mid z \in Z_I,\ \hat{x} \geq 0\} \quad . \tag{3.5}$$

Define

$$W_n^\beta(x) = \mathrm{Min}\{J_n(z,x) \mid z \in Z,\ \hat{x} \geq 0\} \quad . \tag{3.6}$$

Theorem 3.1 an immediate consequence of Lemmas 3.1 and 3.2.

<u>Theorem 3.1</u> $V_n^\beta(x)$ is convex for $x \in N^2$, i.e. $x \geq 0$ and integer-valued.

<u>Lemma 3.1</u> $W_n^\beta(x)$ is a convex function for $x \geq 0$.

<u>Lemma 3.2</u> $V_n^\beta(x) = W_n^\beta(x)$ for $x \in N^2$.

<u>Proof of Lemma 3.1</u> For $x \geq 0$ let $M(\hat{x}) = W_{n+1}^\beta(x) - (n+1)c'x$. From (3.1), (3.3), (3.4) and (3.6) we get

$$M(x) = \min \sum_{k=1}^{n} \sum_{\omega^k} \gamma_k(\omega^k) z_k(\omega^k)$$

$$\text{s.t.}\quad z_k(\omega^k) = 1 \qquad \text{if } \omega_k = A \quad , \tag{3.7}$$

$$0 \leq z_k(\omega^k) \leq 1 \text{ if } \omega_k = T \text{ or } D \quad , \tag{3.8}$$

$$x_k(\omega^k,z) := x + \sum_{j=1}^{k} z_j(\omega^j)\xi_j(\omega^j) \geq 0,\ \omega^k \in \Omega^k,\ k > 1. \tag{3.9}$$

The <u>trajectory</u> corresponding to a policy z (and initial state x) is the process $\hat{x} = (x_1,..x_n)$ with values in R^2, defined recursively by

$$x_0 = x,$$

$$x_k(\omega^k) = x_{k-1}(\omega^{k-1}) + z_k(\omega^k)\xi_k(\omega^k), \ k > 0 \quad . \tag{3.2}$$

The trajectory is said to <u>nonnegative</u>, and we write $z \geq 0$, if for all ω^k, $x_k(\omega^k) \geq 0$.

The cost incurred by policy z and initial state x is defined to be

$$J_{n+1}(z,x) = E_{\hat{x}} \sum_{k=0}^{n} \beta^k c' x_k \quad .$$

From (3.2),

$$x_k(\omega^k) = x + \sum_{j=1}^{k} z_j(\omega^j)\xi_j(\omega^j), \ k > 0 \ .. \tag{3.3}$$

Using this we obtain after some manipulations

$$J_{n+1}(z,x) = (n+1)c'x + \sum_{k=1}^{n} \sum_{\omega^k \in \Omega^k} \gamma_k(\omega^k) z_k(\omega^k) \quad ,$$

where $(\gamma_1,\ldots,\gamma_n)$ is the process

$$\gamma_k(\omega^k) = \sum_{j=k}^{n} \{ \sum_{\omega_{k+1},\ldots,\omega_j} p(\omega^k,\omega_{k+1},\ldots,\omega_j)\}\beta^k c'\xi(\omega^k) \quad . \tag{3.4}$$

When x and z are integer-valued then so is the trajectory $\hat{x}$; if, moreover, $\hat{x} \geq 0$, then it does describe the evolution of the two tandem

This is a linear programming (LP) problem in the finite array of variables $z = \{z_k(\omega^k) | \omega^k \in \Omega^k, 1 \leq k \leq n\}$. Since x enters linearly in the constraint equation, therefore M(x) is convex (see, e.g. Lemma 3, p. 93 of Varaiya (1972)), and Lemma 3.1 is proved.

Proof of Lemma 3.2 By duality theory (Theorem 2, p. 98 of Varaiya (1972)), z* is an optimal solution of the LP above if and only if there exist dual variables $\lambda_k^*(\omega^k) \in R^2$, $\lambda_k^*(\omega^k) \geq 0$, $\omega^k \in \Omega^k$, $k > 1$ (and corresponding to the nonnegativity constraints (3.9)) such that conditions (i), (ii) and (iii) below hold.

(i) z^* is an optimal solution of this LP:

$$\min \sum_{k=1}^{n} \sum_{\omega^k} \gamma_k(\omega^k) z_k(\omega^k) - \sum_{k=1}^{n} \lambda^*(\omega^k)'[x + \sum_{j=1}^{k} z_j(\omega^j)\xi_j(\omega^j)]$$

s.t. (3.7) and (3.8).

(ii) (Feasability) $x_k(\omega^k, z^*) \geq 0$, $\omega^k \in \Omega^k$, $k > 1$.

(iii) (Complementary slackness) If $\lambda_{ki}^*(\omega^k) > 0$, then $x_{ki}(\omega^k, z^*) = 0$, $i = 1, 2$, $k > 1$.

The cost function in (i) may be rewritten as

$$\sum_{k=1}^{n} \sum_{\omega^k} [\gamma_k(\omega^k) - (\sum_{j=k}^{n} \lambda_j^*(\omega^j)'\xi_j(\omega^j))] z_k(\omega^k) + C,$$

where C is a constant independent of z. Hence condition (i) can be more conveniently rewritten as (i').

(i') If $\omega_k = A$, then $z^*(\omega^k) = 1$;

if ω_k = T or D, then

$$z^*(\omega^k) \begin{cases} = 1 & \text{if} \quad \alpha_k := \gamma_k(\omega^k) - \sum_{j=k}^{n} \lambda_j^*(\omega^j)'\xi_j(\omega^j) < 0 \\ = 0 & \text{if} \quad \alpha_k > 0 \\ \in [0,1] & \text{if} \quad \alpha_k = 0 \end{cases} .$$

Henceforth suppose $x \in N^2$. Let z^*, λ^* satisfy (i'), (ii), (iii). We will use z^* to construct an integer-valued policy z which also satisfies these conditions.

<u>Proposition A</u> Let $X = \{z_1\xi(T) + z_2\xi(D) \mid -\frac{1}{2} < z_i \leq \frac{1}{2},\ i = 1,2\}$. Then

$$X + \{z\xi(\omega) \mid 0 \leq z \leq 1\} \subset X \cup \{X + \xi(\omega)\},\ \omega = T \text{ or } D. \tag{3.10}$$

<u>Proof</u> Simple verification. ¤

<u>Proposition B</u> There is an integer-valued policy $z = \{z_k(\omega^k)\}$ such that $z_k(\omega^k) = z_k^*(\omega^k)$ whenever the latter is integer-valued, and

$$\Delta_k := [x_k(\omega^k, z^*) - x_k(\omega^k, z)] \in X \text{ for all } \omega^k,\ k > 1 \quad .$$

<u>Proof</u> Suppose that for some $k \geq 0$, the $z_j(\omega^j)$, $j \geq k$, have been selected and $\Delta_k \in X$. Now

$$\Delta_{k+1} = \Delta_k + z_{k+1}^*(\omega^{k+1})\xi(\omega_{k+1}) - z_{k+1}(\omega^{k+1})\xi(\omega_{k+1}) \quad .$$

If $z_{k+1}^*(\omega^{k+1})$ is integer-valued, then take $z_{k+1}(\omega^{k+1}) = z_{k+1}^*(\omega^{k+1})$ and then certainly $\Delta_{k+1} \in X$. Otherwise, by (3.10), either $[\Delta_k + z^*(\omega^{k+1})\xi(\omega_{k+1})]$ is in X and then take $z_{k+1}(\omega^{k+1}) = 0$, or $[\Delta_k + z^*(\omega^{k+1})\xi(\omega_{k+1})] \in X + \xi(\omega_{k+1})$ and then take $z_{k+1}(\omega^{k+1}) = 1$. In either case $\Delta_{k+1} \in X$ and the proposition follows by induction. ¤

Lemma 3.2 follows from the next proposition.

<u>Proposition C</u> The integer-valued policy z constructed above satisfies (i'), (ii) and (iii).

<u>Proof</u> Since $z_k(\omega^k) = z_k^*(\omega^k)$ whenever the latter is integer valued, condition (i') is satisfied.

To prove (ii) it must be shown that $y := x_k(\omega^k, z) \geq 0$. Suppose to the contrary that $y \ngeq 0$. Since y is integer-valued, therefore for at least one i = 1 or 2, $y_i \leq -1$. Let $y^* := x_k(\omega^k, z^*)$. By Proposition B, and recalling the definition of $\xi(T)$ and $\xi(D)$,

$$y^* \in y + \{z_1(-1,1) + z_2(0,-1) \mid -\tfrac{1}{2} < z_i \leq \tfrac{1}{2}\}$$

$$= \{(y_1 - z_1, y_2 + z_1 - z_2) \mid -\tfrac{1}{2} < z_i \leq \tfrac{1}{2}\} \quad .$$

Now if $y_1 \leq -1$, then $y_1 - z_1 < 0$ if $-\frac{1}{2} < z_1$ and so $y_1^* < 0$; and if $y_2 \leq -1$, then $y_2 + z_1 - z_2 < 0$ if $z_1 \leq \frac{1}{2}$ and $-\frac{1}{2} < z_2$ and so $y_2^* < 0$. In either case if $y \ngeq 0$, then $y^* \ngeq 0$ which contradicts the hypothesis that z^* satisfies (ii). Thus z satisfies (ii) as well.

Finally to prove (iii) it is evidently sufficient to show that $y_i = 0$ whenever $y_i^* = 0$ where these are defined as above. By Proposition B again,

$$y \in \{(y_1^* + z_1, y_2^* - z_1 + z_2) \mid -\tfrac{1}{2} < z_i \leq \tfrac{1}{2}\} \quad .$$

If $y_1^* = 0$ then $y_1 \in (-\frac{1}{2}, \frac{1}{2}]$ and so being integer-valued $y_1 = 0$; and if $y_2^* = 0$ then $y_2 \in (-1,1)$ and so being integer-valued $y_2 = 0$.

$\blacksquare$

Corollary 3.1 For $\beta < 1$, $V^\beta = V^\beta_\infty$ is convex.

Proof Immediate from Theorem 3.1 and (2.7). ⌑

4. The optimal switching function

In this section conditions (2.6), (2.9) and the convexity of V^β_n are used to show that the optimal policy is switch-over. For the infinite horizon problem only the discounted case, $\beta < 1$, is considered here.

Fix $\beta \leq 1$. Let $V_n = V^\beta_n$. From (2.6)

$$V_1(x) = c'x, \tag{4.1}$$

$$V_2(x) = (1+\beta)c'x + \beta\lambda c_1 - \beta\mu c_2\, 1(x_2>0) + \beta a(c_2-c_1)1(c_2\leq c_1, x_1>0), \tag{4.2}$$

where $1(\cdot)$ denotes the indicator function. To evaluate V_n for $n \geq 2$, define

$$U_n(x) = V_n(Tx) - V_n(x). \tag{4.3}$$

Observe, with the aid of (2.6), that when there are n steps to go the optimal action is

$$z = \begin{cases} 0 & \text{if } U_n(x) > 0 \\ 1 & \text{if } U_n(x) \leq 0 \end{cases} . \tag{4.4}$$

Also from (2.6),

$$U_{n+1}(x) = (c_2-c_1) + \beta\lambda\, U_n(Ax) + \beta\mu[V_n(DTx) - V_n(Dx)] + \beta a\, \phi_n(x) \tag{4.5}$$

where

$$\phi_n(x) = V_n(Tx) \wedge V_n(T^2x) - V_n(x) \wedge V_n(Tx) \tag{4.6}$$

and $f \wedge g = \min(f,g)$ for any two numbers f and g.

It is clear, and can be proved easily by an argument based on stochastic dominance, that the total cost is an increasing function of the initial state.

<u>Lemma 4.1</u> $V_n^\beta(x_1,x_2)$ is non-decreasing in x_i, i = 1, 2.

If $c_1 \geq c_2$, that is the waiting cost at station 2 is not greater than at station 1, then evidently the optimum policy is $z_k \equiv 1$.

<u>Lemma 4.2</u> If $c_1 \geq c_2$, then $z_k \equiv 1$ is an optimum policy

Henceforth it is assumed that $c_2 > c_1$.

<u>Lemma 4.3</u> $U_n(x_1,x_2)$ is increasing in x_2 and decreasing in x_1, for $x_1 > 0$, $x_2 \geq 0$ (Increasing means non-decreasing, similarly for decreasing).

<u>Proof</u> From (4.1), (4.2), (4.3)

$$U_1(x) = c_2 - c_1 \tag{4.7}$$

$$U_2(x) = (1+\beta)(c_2-c_1) - \beta\mu c_2 1(x_2=0) \tag{4.8}$$

so the assertion is true for n = 1, 2. Suppose it is true for n. The first two terms in the formula (4.5) clearly have the indicated property. The third term is

$$\beta\mu[V_nDTx) - V_n(Dx)] = \begin{cases} U_n(Dx) & \text{if } x_1 > 0,\ x_2 > 0 \\ V_n(x_1-1,0) - V_n(x_1,0) & \text{if } x_1 > 0,\ x_2 = 0 \end{cases} \tag{4.9}$$

By the induction hypothesis, $U_n(Dx)$ has the indicated property for $x_1 > 0$, $x_2 > 0$. By Theorem 3.1 and Lemma 4.1 $V_n(x_1,0)$ is an increasing convex function of x_1, and so $V_n(x_1-1,0) - V_n(x_1,0)$ is decreasing in x_1, for $x_1 > 0$. It only remains to check that (4.9) is larger for $(x_1,1)$ than for $(x_1,0)$ i.e. that

$$U_n(D(x_1,1)) = V_n(x_1-1,1) - V_n(x_1,0) \geq V_n(x_1-1,0) - V_n(x_1,0) \quad .$$

But this inequality follows from Lemma 4.1. The last term in (4.5) can be rewritten as

$$\beta a\{[V_n(T^2x) - V_n(Tx)] \wedge 0 + [V_n(Tx) - V_n(x)] \vee 0\} \quad .$$

$$= \beta a[U_n(Tx) \wedge 0 + U_n(x) \vee 0],$$

and by the hypothesis this expression also has the indicated property. The lemma follows by induction. ⌑

Define the <u>switching function</u> $S_n(x_1)$ by

$$S_n(x_1) = \min\{x_2 \geq 0 | U_n(x_1,x_2) > 0\}$$

$$= \infty \text{ if } U_n(x_1,x_2) \leq 0 \text{ for all } x_2 \quad .$$

It will be shown in Lemma 4.5 that S_n is always finite.

<u>Corollary 4.1</u> $S_n(x_1)$ is increasing in x_1. When there are n steps to go the optimal action is z = 0 or 1 according as $x_2 \geq S_n(x_1)$ or $x_2 < S_n(x_1)$.

<u>Proof</u> The first assertion follows from the fact that $U_n(x_1,x_2)$ is increasing in x_2, and the second from (4.4) and the fact that $U_n(x_1,x_2)$ is decreasing in x_1. ⌑

Next we study the behavior of S_n as n increases.

Lemma 4.4 For $n \geq 1$,

$$U_n(x) \leq 0 \Rightarrow U_{n+1}(x) \leq U_n(x), \tag{4.10}$$

$$V_n(Ax) - V_n(x) \leq V_{n+1}(Ax) - V_{n+1}(x), \tag{4.11}$$

$$V_n(x) - V_n(Dx) \leq V_{n+1}(x) - V_{n+1}(Dx) \quad . \tag{4.12}$$

Proof Consider $n = 1$. By (4.7), $U_1(x) = c_2 - c_1 > 0$, so (4.10) holds trivially. By (4.1) and (4.2),

$$V_1(Ax) - V_1(x) = c_1 \leq (1+\beta)c_1 = V_2(Ax) - V_2(x),$$

$$V_1(x) - V_1(Dx) = c_2 1(x_2>0) \leq (1+\beta)c_2 1(x_2>0) - \beta\mu c_2 1(x_2=1) = V_2(x)-V(Dx),$$

and so the assertion is true for $n = 1$. Suppose the assertion is true for $n - 1$. Will prove it for n.

Consider (4.10), and suppose $U_n(x) \leq 0$. By (4.5)

$$U_n(x) - U_{n+1}(x) = \beta\lambda[U_{n-1}(Ax) - U_n(Ax)] + \beta\mu\{[V_{n-1}(DTx) - V_{n-1}(Dx)]$$
$$- [V_n(DTx) - V_n(Dx)]\} + \beta a[\phi_{n-1}(x) - \phi_n(x)] \quad . \tag{4.13}$$

Now $[U_{n-1}(Ax) - U_n(Ax)] \geq 0$. To see this, note that $U_n(x) \leq 0$ implies $U_n(Ax) \leq 0$ since U_n is decreasing in x_1. Therefore, if $U_{n-1}(Ax) > 0$, then certainly $U_{n-1}(Ax) - U_n(Ax) \geq 0$; whereas if $U_{n-1}(Ax) \leq 0$, then the same conclusion follows from the induction hypothesis. The coefficient of $\beta\mu$ in (4.13) is also nonnegative. Indeed, if $x_2 > 0$, then this coefficient is $[U_{n-1}(Dx) - U_n(Dx)]$ and the same argument applies; whereas if $x_2 = 0$, this coefficient is $[V_{n-1}(x_1-1,0) - V_{n-1}(x_1,0)]$ $- [V_n(x_1-1,0) - V_n(x_1,0)]$ and this is nonnegative by the induction hypothesis. The coefficient of βa in (4.13) can be rewritten as

$$[U_{n-1}(Tx) \wedge 0 + U_{n-1}(x) \vee 0] - [U_n(Tx) \wedge 0 + U_n(x) \vee 0]$$

$$= U_{n-1}(Tx) \wedge 0 + U_{n-1}(x) \vee 0 - U_n(Tx) \wedge 0 \quad , \tag{4.14}$$

since $U_n(x) \leq 0$. The induction hypothesis implies $U_{n-1}(Tx) \wedge 0 \geq U_n(Tx) \wedge 0$, and so the coefficient of βa is nonnegative. Thus (4.10) is true for n.

Next consider (4.11). Using (2.6),

$$V_{n+1}(Ax) - V_{n+1}(x) = c_1 + \beta\lambda[V_n(A^2x) - V_n(Ax)] + \beta\mu[V_n(DAx) - V_n(Dx)]$$

$$+ \beta a[V_n(Ax) \wedge V_n(TAx) - V_n(x) \wedge V_n(Tx)] \; .$$

A similar expression is valid for $V_n(Ax) - V_n(x)$. By the induction hypothesis

$$V_n(A^2x) - V_n(Ax) \geq V_{n-1}(A^2x) - V_{n-1}(Ax) \quad ,$$

$$V_n(DA_x) - V_n(Dx) \geq V_{n-1}(DAx) - V_{n-1}(Dx) \quad .$$

Therefore to show (4.11), it only remains to show that

$$v_n(x) \geq v_{n-1}(x) \quad ,$$

where $v_n(x) = V_n(Ax) \wedge V_n(TAx) - V_n(x) \wedge V_n(Tx)$. Since for real numbers a,b,c,d we have $a \wedge b - c \wedge d = a - d - (c-b) \wedge 0$, we may write

$$v_n(x) = V_n(TAx) - V_n(x) - U_n(Ax) \vee 0 - U_n(x) \wedge 0 \quad .$$

Since x = DTAx, the induction hypothesis yields

$$V_n(TAx) - V_n(x) \geq V_{n-1}(TAx) - V_n(x) \quad , \tag{4.16}$$

$$U_n(x) \wedge 0 \leq U_{n-1}(x) \wedge 0 \quad . \tag{4.17}$$

Now, if $U_{n-1}(Ax) \leq 0$ then by the induction hypothesis $U_n(Ax) \leq 0$ also, which together with (4.16), (4.17) gives (4.15); if $U_{n-1}(Ax) > 0$ and $U_n(Ax) \leq 0$ then again we get (4.15); finally if $U_{n-1}(Ax) > 0$ and $U_n(Ax) > 0$ then $U_{n-1}(x) > 0$ and $U_n(x) > 0$ (since U_n and U_{n-1} are decreasing in x_1), and so again

$$v_n(x) = V_n(TAx) - V_n(x) - U_n(Ax) = V_n(Ax) - V_n(x) \geq V_{n-1}(Ax) - V_{n-1}(x) = v_{n-1}(x).$$

Finally consider (4.12). Using (2.6) it can be seen that (4.12) holds if

$$w_n(x) \geq w_{n-1}(x) \tag{4.18}$$

where $w_n(x) = V_n(x) \wedge V_n(Tx) - V_n(Dx) \wedge V_n(TDx)$. Now

$$w_n(x) = V_n(x) - V_n(Dx) + [V_n(Tx) - V_n(x)] \wedge 0 - [V_n(TDx) - V_n(Dx)] \wedge 0.$$

By the induction hypothesis,

$$V_n(x) - V_n(Dx) \geq V_{n-1}(x) - V_{n-1}(Dx) \quad ,$$

$$[V_n(TDx) - V_n(Dx)] \wedge 0 \leq [V_{n-1}(TDx) - V_{n-1}(Dx)] \wedge 0 \quad .$$

Hence (4.18) certainly holds if $U_{n-1}(x) \leq 0$ or $U_n(x) \geq 0$. Whereas if $U_{n-1}(x) > 0$ and $U_n(x) < 0$, then $U_n(Dx) < 0$ (since U_n is increasing in x_2), and so

$$w_n(x) - V_n(x) - V_n(TDx) \quad ,$$

$$w_{n-1}(x) = V_{n-1}(x) - V_{n-1}(Dx) \wedge V_{n-1}(TDx) \quad .$$

Now if $U_{n-1}(Dx) < 0$, then

$$w_{n-1}(x) = V_{n-1}(x) - V_{n-1}(TDx) = V_{n-1}(Tx) - V_{n-1}(TDx) - U_{n-1}(x)$$

$$\leq V_{n-1}(Tx) - V_{n-1}(TDx), \text{ since } U_{n-1}(x) > 0,$$

$$\leq V_n(x) - V_n(TDx), \text{ by induction hypothesis,}$$

$$= w_n(x);$$

and similarly, if $U_{n-1}(Dx) > 0$, then

$$w_{n-1}(x) = V_{n-1}(x) - V_{n-1}(Dx)$$

$$\leq V_n(x) - V_n(Dx), \text{ by induction hypothesis,}$$

$$\leq V_n(Tx) - V_n(TDx) = w_n(x) \quad .$$

The last inequality comes from the fact that $U_n(Dx) \leq U_n(x)$ since U_n is increasing in x_2.

The lemma now follows by induction. ¤

<u>Corollary 4.2</u> The switching curve $S_n(x_1)$ is increasing in n.

<u>Proof</u> If $x_2 \leq S_n(x_1)$, then $U_n(x_1,x_2) \leq 0$, and so $U_{n+1}(x_1,x_2) \leq 0$ which implies $x_2 \leq S_{n+1}(x_1)$. Hence $S_n(x_1) \leq S_{n+1}(x_1)$. ¤

<u>Lemma 4.5</u> $S_n(x_1) \leq (n-2) \vee 0$ for all x_1.

<u>Proof</u> Because of Corollary 4.1 we may limit ourselves to $x_1 \geq 1$.

By Lemma 4.3 it is enough to show $U_n(x_1,(n-2) \vee 0) > 0$. For n = 1,2 this follows from (4.7), (4.8). Suppose the assertion holds for some $n \geq 2$. Using (4.5),

$$U_{n+1}(x_1,n-1) = c_2 - c_1 + \beta\lambda U_n(x_1+1,n-1) + \beta\mu[V_n(x_1-1,n-1) - V_n(x_1,n-2)]$$

$$+ \beta a\phi_n(x_1,n-1)$$

$$= c_2 - c_1 + \beta\lambda U_n(x_1+1,n-1) + \beta\lambda U_n(x_1,n-2) + \beta a\phi_n(x_1,n-1)$$

and so it is enough to prove that $\phi_n(x_1,n-2) \geq 0$. By (4.6)

$$\phi_n(x_1,n-1) = [V_n(T^2x) - V_n(Tx)]\wedge 0 + U_n(x)\vee 0$$

$$\geq [V_n(T^2x) - V_n(Tx)]\wedge 0 + U_n(x)\vee 0, \text{ with } x = (x_1,n-1)$$

$$= U_n(x_1-1, n)$$

which is positive by the induction hypothesis if $x_1-1 \geq 1$, and also if $x_1 = 1$, because $U_n(0,n) = 0$. ¤

Since we wish to consider different discount factors we reintroduce the superscript β and use the notation V_n^β, U_n^β, S_n^β. Because of Corollary 4.2 the following limit exists.

$$S^\beta(x_1) = S_\infty^\beta(x_1) = \lim_{n\to\infty} S_n^\beta(x_1) \quad ,$$

and by Corollary 4.2 it is increasing in x_1.

Theorem 4.1 An optimal policy for the infinite horizon discounted case, $\beta < 1$, is given by the stationary switch-over policy $z(x) = 0$ or 1 according as $x_2 \geq S^\beta(x_1)$ or $x_2 < S^\beta(x_1)$.

Proof Since $V_n^\beta \to V^\beta$, therefore $U^\beta(x) = \lim U_n^\beta(x)$. By (2.9) it is enough to show that

$$x_2 \geq S^\beta(x_1) \Rightarrow U^\beta(x) \geq 0 \quad ,$$

$$x_2 < S^\beta(x_1) \Rightarrow U^\beta(x) \leq 0 \quad .$$

Now if $x_2 \geq S^\beta(x_1)$ then $x_2 \geq S_n^\beta(x_1)$ for all n and so $U_n^\beta(x) \geq 0$; hence $U^\beta(x) \geq 0$. Suppose $x_2 < S^\beta(x_1)$. Since the switching curves are integer-valued and $S^\beta(x_1) = \lim S_n^\beta(x_1)$, therefore $x_2 < S_n^\beta(x_1)$ for large n; hence $U_n^\beta(x) \leq 0$ for large n and so $U^\beta(x) \leq 0$. ¤

A surprising feature of the discounted case is that it is optimal to provide no service at station 1 when the queue length at station 2 exceeds a threshold.

<u>Theorem 4.2</u> If $\beta < 1$, then

$$S^\beta(x_1) \leq \min\{x_2 | \beta^{x_2} \leq \frac{c_2-c_1}{c_2}\}$$

<u>Proof</u> Fix x with $x_1 > 0$. We claim that

$$U_n^\beta(x) = V_n^\beta(Tx) - V_n^\beta(x) \geq (c_2-c_1) \frac{1-\beta^{x_2 \wedge n}}{1-\beta} - c_1 \frac{\beta^{x_2 \wedge n} - \beta^n}{1-\beta} = v_n(x_2) \text{say.} \tag{4.19}$$

To see this, and using the notation of Section 3, along a sample path ω^n, let $x_k(\omega^k)$ be the trajectory corresponding to the optimal policy and initial state $x_0 = Tx$, and let $z_k(\omega^k)$ be the optimal control policy. Let $y_k(\omega^k)$ be the trajectory along the same sample path and corresponding to the same policy, but with initial condition $y_0 = x$. Then

$$V_n^\beta(Tx) = E \sum_0^{n-1} \beta^k c' x_k \text{ and } V_n^\beta(x) \leq E \sum_0^{n-1} \beta^k c' y_k \quad ,$$

and so to prove (4.19) it is enough to show that

$$E \sum^{n-1} \beta^k c' x_k - E \sum^{n-1} \beta^k c' y_k \geq v_n(x_2) \quad . \tag{4.20}$$

To do this define the stopping time

$$K = K(\omega^n) = \min\{k \geq 0 | x_{k2}(\omega^k) = 0\} \wedge n \quad .$$

Then $x_{k2} > 0$ for $k \leq K-1$ and a comparison of the two trajectories gives

$$\sum_0^{n-1} \beta^k c' x_k - \sum_0^{n-1} \beta^k c' y_k \geq (c_2 - c_1) \sum_0^{K-1} \beta^k - c_1 \sum_K^{n-1} \beta^k \quad .$$

Since x_{k2} can decrease by atmost one per step, therefore $K \geq x_2 \wedge n$, and so

$$\sum_0^{n-1} \beta^k c' \; x_k - \sum_0^{n-1} \beta^k c' y_k \geq (c_2 - c_1) \sum_0^{(x_2 \wedge n)-1} \beta^k - c_1 \sum_{x_2 \wedge n}^{n-1} \beta^k$$

from which follows (4.20) by taking expectations. Hence (4.19) is true. But then if $x_2 \geq 0$ is such that $\beta^{x_2} \leq \frac{c_2 - c_1}{c_2}$, then $U_n^\beta(x) > 0$ and so $x_2 \geq S_n^\beta(x_1)$. ¤

We have seen that if $\beta < 1$, then the switching curves S_n^β increase to S^β and are all uniformly bounded. (See Figure 2). Let x_k^β, $k = 0,1,..$ be the Markov process corresponding to the optimal switching curve S^β. Since the service rate at station 1 is nonzero only when $x_2 < S^\beta(x_1)$, therefore all states in $\{x | x_2 \geq S^\beta(x_1) + 1\}$ are transient. The process $\{x_{k2}^\beta\}$ may be ergodic nonetheless. But observe from Theorem 4.2 that the service at station 1 stops whenever $x_{k2}^\beta \geq s$ where $s > \infty$ is independent of λ, μ. It is not difficult to see that then there always exists μ sufficiently small, but with $\mu > \lambda$, such that the queue length in station 1 $x_{k_1}^\beta \to \infty$ with probability one. Thus in the discounted case, the optimal process may be non-ergodic, in fact all states may be transient!

The surprising result also suggests that the case of long run average cost, whose study requires ergodicity, cannot be approached by the standard approach of taking limits as β increases to 1.

5. Long run average cost

Throughout this section we consider the undiscounted case $\beta = 1$, and so we write V_n, U_n, S_n instead of V_n^1, U_n^1, S_n^1. The next result contrasts sharply with Theorem 4.2.

Lemma 5.1 For every x_2, $S_n(x_1) > x_2$ for all (x_1,n) sufficiently large.

Proof Suppose in contradiction that $S_n(x_1) < b$ for all x_1,n. We use the notation of Section 3. Fix the horizon n and an initial state x. Let $z_k(\omega^k)$ be the optimal control policy and $x_k(\omega^k)$ the corresponding trajectory. Also let $y_k(\omega^k)$ be the trajectory corresponding to the same policy but with initial state Tx. Then

$$E \sum_0^{n-1} c'x_k = V_n(x), \; E \sum_0^{n-1} c'y_k \geq V_n(Tx) \quad . \tag{5.1}$$

We wish to compare the two trajectories $\{x_k\}$ and $\{y_k\}$. Let K_i be the first time that queue i is empty,

$$K_i(\omega^n) = \min\{k | x_{k_i}(\omega^k) = 0\} \wedge n \quad .$$

If $K_1 \leq K_2$, one can verify that

$$y_k = \begin{cases} Tx_k , & k = 0,\ldots,K_1-1 \\ x_k , & k \geq K_1 \end{cases}$$

and therefore

$$\sum_{0}^{n-1} c'y_k - \sum_{0}^{n-1} c'x_k = (c_2-c_1)K_1 \quad ;$$

whereas if $K_1 > K_2$, then

$$y_k = \begin{cases} Tx_k & , k = 0,..,K_2-1 \\ (x_{k1}-1,x_{k2}) & , k = K_2,..,K_1-1 \\ (x_{k1},x_{k2}-1) \text{ or } x_k, k \geq K_1 \end{cases}$$

and therefore

$$\sum_{0}^{n-1} c'y_k - \sum_{0}^{n-1} c'x_k \leq (c_2-c_1)K_2 - c_1(K_1-K_2) \quad .$$

Combining these two estimates gives

$$E \sum_{0}^{n-1} c'y_k - E \sum_{0}^{n-1} c'x_k \leq c_2 \, E(K_1 \wedge K_2) - c_1 \, EK_1 = \delta \text{ say.} \tag{5.2}$$

We want to show that $\delta < 0$ for some (x_1,n). It will be first shown that $E(K_2) < \overline{K}_2$ for some constant $\overline{K}_2$ and for all initial states $x = (x_1,b)$, and n. To see this observe that, by assumption, there are no arrivals into station 2 whenever $x_{k2} \geq b$; and when there are arrivals these occur in a Poisson stream with rate a. Hence the process x_{k2}, $k = 0,1,..$ is stochastically dominated by the queue size process of a M/M/1/b queue with arrival rate a and service rate μ. Let $\overline{K}_2$ be the expected time (measured in number of transitions) to empty this queue if it initially starts with b customers. Clearly $E(K_2) < \overline{K}_2$. On the other hand since x_{k1} can decrease by at most one per step, therefore $E\,K_1 \geq x_1 \wedge n$.

Hence $\delta \le c_2\bar{K}_2 - c_1(x_1 \wedge n) < 0$ for x_1, n large. From (5.1), (5.2) we conclude that

$$V_n(Tx) - V_n(x) = U_n(x) < 0 \tag{5.3}$$

for $x = (x_1,b)$ and x_1,n large. But if (5.3) holds then $S_n(x_1) > b$ contradicting the assumption. ¤

Let P denote the set of all policies over the infinite horizon. For $\pi \in P$ let

$$V_n(\pi,x) = E_x \sum_0^{n-1} c'x_k^\pi \quad , \tag{5.4}$$

where x_k^π, k = 0,1,.. is the state process corresponding to π and x is the initial state. Then

$$V_n(x) = \min_P V_n(\pi,x) \quad . \tag{5.5}$$

Let

$$\bar{V}(\pi,x) = \overline{\lim_n} \frac{1}{n} V_n(\pi,x) \quad , \tag{5.6}$$

$$\bar{V}(x) = \inf_P \bar{V}(\pi,x) \quad . \tag{5.7}$$

From (5.5) and (5.6), $\bar{V}(\pi,x) \ge \overline{\lim} \frac{1}{n} V_n(x)$ and so

$$\bar{V}(x) \ge \overline{\lim} \frac{1}{n} V_n(x) \quad . \tag{5.8}$$

Recall that the switching curves $S_n(x_1)$ are increasing in n and so the following limit exists.

$$S(x_1) = \lim_n S_n(x_1) \quad .$$

Abusing notation slightly, let S denote the stationary policy defined by the switching curve S. Our objective is to show that (i) under S the Markov process $\{x_k^S\}$ is ergodic and so $\overline{V}(S,x) = \overline{V}(S)$ is independent of x and (ii) $\overline{V}(S) = \overline{\lim} \frac{1}{n} V_n(x)$ so that S minimizes the long run average cost. It is assumed henceforth that $\lambda < \min(a,\mu)$.

Let σ denote the stationary policy under which $z_k \equiv 1$. In this case the network consists of two M/M/1 queues in tandem with arrival λ and service rates a and μ in station 1 and 2 respectively. Since $\lambda < a \wedge \mu$, therefore $\{x_k^\sigma\}$ is ergodic and $\overline{V}(\sigma,x) = \overline{V}(\sigma) < \infty$ is independent of x. From (5.7) and (5.8) we get

$$\overline{\lim} \ \frac{1}{n} V_n(x) \leq \overline{V}(\sigma) \quad . \tag{5.9}$$

<u>Lemma 5.2</u> The Markov process $\{x_k^S\}$ is ergodic.

<u>Proof</u> Let $|x| = x_1 + x_2$ denote total number of customers in the system. Let $\varepsilon > 0$. By (5.9), for all n sufficiently large,

$$\frac{1}{n} V_n(x) = \frac{1}{n} \sum^{n-1} E_x \, c' x_k^{\sigma^n} \leq \overline{V}(\sigma) + \varepsilon \tag{5.10}$$

where σ^n is the policy which achieves the minimum in (5.5). Since $0 < c_1 \leq c_2$, (5.10) implies

$$\frac{1}{n} \sum^{n-1} E_x |x_k^{\sigma^n}| \leq \frac{1}{c_1} [\overline{V}(\sigma) + \varepsilon] \quad .$$

The policy σ^n is defined by the switching curve S_k when there are k steps to go, and $S_k \leq S$. Hence, if $x_k^S = x_k^{\sigma^n}$ for some k, then the service rate at station 1 under S is always greater than under σ^n.

Considerations based on stochastic dominance then lead to

$$E_x|x_k^S| \leq E_x|x_k^{\sigma^n}| \quad ,$$

and so

$$\frac{1}{n}\sum^{n-1} E_x|x_k^S| \leq \frac{1}{c_1}[\overline{V}(\sigma) + \varepsilon] \quad .$$

Hence

$$\overline{\lim}\, \frac{1}{n}\sum^{n-1} E_x|x_k^S| \leq \frac{\overline{V}(\sigma)}{c_1} \quad , \tag{5.11}$$

which clearly implies ergodicity. ¤

The following consequence of (5.11) will be useful later.

Corollary 5.1 For any $\varepsilon > 0$ there is a finite set of states $A \subset N^2$ such that

$$\overline{\lim}\, \frac{1}{n}\sum^{n-1} c'x_k^S\, 1(x_k^S \notin A) \leq \varepsilon \quad . \tag{5.12}$$

Corollary 5.2 For $(m_1,m_2) \in N^2$ define

$$A(m_1,m_2) = \{x|x_1 \leq m_1 - 1 \text{ and } x_2 \leq S(x_1) \wedge (m_2-1)\}$$

Then for any $\varepsilon > 0$, there exists (m_1,m_2) such that $A(m_1,m_2)$ satisfies (5.12).

Proof This follows from Corollary (5.1) and the fact that the set $\{x|x_1 \geq S(x_2) + 1\}$ is transient under the policy S. ¤

By Lemma 5.2 $\overline{V}(S,x) = \overline{V}(S)$ is independent of x. To show $\overline{V}(S) = \overline{\lim}\, \frac{1}{n} V_n(x)$ we compare $V_n(S,x)$ and $V_n(x) = V_n(\sigma^n,x)$. The next lemma will be used in estimating $|V_n(S,x) - V_n(\sigma^n,x)|$.

<u>Lemma 5.3</u> Let $V_k(\cdot)$, $k = 0,1,..$ be a sequence of functions and x_k, $k = 0,1,..$ a process with values in N^2 such that

$$E\{V_k(x_k) - V_{k+1}(x_{k+1}) | x_k = x\} = c'x \quad .$$

Let K be a stopping time of the process $\{x_k\}$ and let Y_K be a $\underline{\underline{F}}_K^x$ - measurable variable such that $V_K(Y_K) \le V_K(x_K)$ a.s. Let $\{y_k\}$ be another process such that

$$y_k = x_k,\ k \le K - 1$$

$$y_K = Y_K \quad ,$$

$$E\{V_k(y_k) - V_{k+1}(y_{k+1}) | y_k = y,\ k > K\} = c'y \quad .$$

Then

$$E_x \sum^{n-1} c'x_k \ge E_x \sum^{n-1} c'y_k \quad .$$

(Here $\underline{\underline{F}}_k^x$ is the σ-field generated by $x_0,\dots,x_k$).

<u>Proof</u> $E_x \sum^{n-1} c'y_k = E_x\{\sum^{K-1} c'y_k + V_K(Y_K)\} = E_x\{\sum^{K-1} c'x_k + V_K(Y_K)\}$

$$\le E_x\{\sum^{K-1} c'x_k + V_K(x_K)\} = E_x \sum^{n-1} c'x_k. \qquad \square$$

<u>Lemma 5.4</u> $\overline{V}(S) = \overline{\lim}\, \frac{1}{n} V_n(x)$.

Proof Fix an initial state x. Fix $\varepsilon > 0$ and select $A = A(m_1, m_2)$ as in Corollary 5.2. Since the switching curves are integer-valued, there exists m such that (see Fig. 3)

$$S_n(x_1) \wedge m_2 = S(x_1) \wedge m_2, \text{ for } x_1 \leq m_1 \text{ and } n > m \quad .$$

Select $n > m$ such that

$$\frac{1}{n} E_x \sum_{n-m+1}^{n-1} c' x_k^S < \varepsilon \quad , \tag{5.13}$$

$$\frac{1}{n} E_x \sum_{0}^{n-1} c' x_k^S \, 1(x_k^S \notin A) < \varepsilon \quad . \tag{5.14}$$

Define the return times to A by $K_0 \equiv 0$,

$$K_r = \min\{k > K_{r-1} | x_k^S \in A,\ x_{k-1}^S \notin A\} \wedge n,\ r \geq 1 \quad .$$

For $r \geq 1$, let

$$y_k^r = \begin{cases} x_k^S, & k \leq K_{r-1} \\ \phi(K_r, k; x_{K_r}^S), & k \geq K_r \end{cases}$$

where $\phi(k_1, k_2, x)$ is the state process at time k_2 if the state at time k_1 is x and policy σ^n is used i.e. $\phi(k_2, k_1, x) = x_{k_2}^{\sigma^n}$ given that $x_{k_1}^{\sigma^n} = x$.

For $r = 0$, let $w_k^0 = x_k^{\sigma^n}$ and for $r \geq 1$ let

$$w_k^r = \begin{cases} w_k^{r-1}, & k \leq K_r - 1 \\ \phi(K_r, k; x_{K_r}^S), & k \geq K_r \quad . \end{cases}$$

We claim first of all that

$$E_x \sum_0^{n-m} c'w_k^r \leq E_x \sum_0^{n-m} c'w_k^{r-1} \quad . \tag{5.15}$$

Indeed, observe that on $\{K_r \leq n - m\}$,

$$|x_{K_r}^S| \leq |w_{K_r}^{r-1}| \text{ and } x_{K_r2}^S \geq w_{K_r2}^{r-1} \quad .$$

This implies that

$$V_{n-K_r}(x_{K_r}^S) \leq V_{n-K_r}(w_{K_r}^{r-1}) \tag{5.16}$$

since $x_{K_r}^S$ lies 'below' the switching curve S_{n-K_r}, and since in this region $V_{n-K_r}(x_{K_r}^S + (\ell,-\ell) + (\ell',0))$ is increasing in ℓ and ℓ'. Use of (5.16) in Lemma 5.3 gives (5.15).

Second, by definition of y_k^r,

$$0 \leq E_x \sum_0^{n-m} c'y_k^r - E_x \sum_0^{n-m} c'w_k^r \leq E_x \sum_0^{K_r} c'x_k^S \, 1(x_k^S \notin A) \quad ;$$

and so, using (5.15),

$$E_x \sum_0^{n-m} c'y_k^r \leq E_x \sum_0^{n-m} c'w_k^0 + E_x \sum_0^{K_r} c'x_k^S \, 1(x_k^S \notin A)$$

$$\leq V_n(x) + n\varepsilon \tag{5.17}$$

since $w_k^0 = x_k^{\sigma^n}$ and by (5.14).

Finally, since $y_k^r \to x_k^S$ as $r \to \infty$, (5.17) implies

$$E_x \sum_{0}^{n-m} c' x_k^S \le V_n(x) + n\varepsilon \quad ,$$

which, together with (5.13) gives

$$\overline{V}(S) \le \overline{\lim} \frac{1}{n} V_n(x) + 2\,\varepsilon \quad . \qquad ¤$$

Lemmas 5.2, 5.4 give the main result.

Theorem 5.1 The stationary policy defined by the switching curve S minimizes the long run average cost. Moreover the resulting Markov process is ergodic.

6. Conclusions

We have shown that the optimal policy which minimizes the cost over the infinite horizon is 'bang-bang' and characterized by a monotonically increasing switching curve $x_2 = S^{\beta}(x_1)$ where $\beta \le 1$ is the discount factor. The result is intuitively obvious: the 'bang-bang' nature is a consequence of the fact that there is no service cost, and the monotonicity of the switching curve can be anticipated once it is surmised that the difference in the total cost incurred by adding a marginal customer to the second queue rather than adding him to the first queue must increase with the number of customers in the second queue. This intuition also suggests the form of the optimal policy in more general networks. However, substantiation of this intuition even in the simple case treated here is a non-trivial exercise and differs from the case of a single queue in two important respects. First, the optimal processes for the discounted problem $\beta < 1$ may not be ergodic so that the average cost case cannot be studied by taking a limit as $\beta \to 1$. Second, the convexity of the value function, which is critical for further analysis and which is implicit in the intuitive argument given above, is difficult to establish. The argument for convexity put forward here extends to more complex networks.

References

Gallisch, E.,"On monotone optimal policies in a queuing model of M/G/1 type with controllable service time distribution," Advances in Applied Probability 11, 1979, 870-887.

Grabill, T., "Optimal control of service facility with variable exponential service time and constant arrival rate," Management Science 18, 1972, 560-566.

Grabill, T. B., Gross, D. and Magazine, M. J., "A classified bibliography of research on optimal design and control of queues," Operations Research 25(2), 1977, 219-232.

Lippman, S. A., "Semi-Markov decision processes with unbounded rewards, Management Science 19(7), 1973, 717-731.

Lippman, S. A., "A new technique in the optimization of exponential queuing systems," working paper No. 211, University of California, Los Angeles, 1973a.

Schassberger, R. "A note on optimal service selection in a single server queue," Management Science 21, 1975, 1326-1331.

Tijms, H., "On the optimality of a switch over policy for controlling the queue size in a M/G/1 queue with variable service rate," Report BW 25/75, Mathematisch Centrum, Amsterdam, 1975.

Tijms, H. C. and Van der Duyn Schouten, F. A., "Inventory control with two switch-over levels for a class of M/G/1 queuing systems with variable arrival and service rate," Stochastic Processes and their Applications 6, 1978, 213-222.

Weiss, H. J., "The computation of optimal control limits for a queue with batch services," Management Science 25(4), 1979, 320-328.

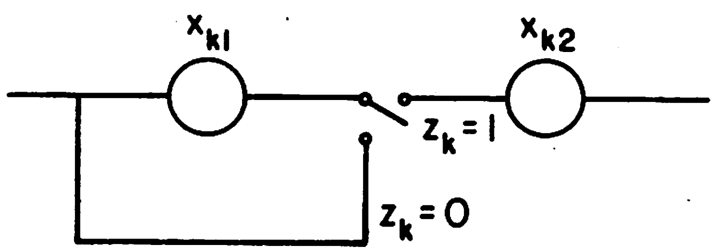
Varaiya, P., Notes on Optimization, Van Nostrand Reinhold, New York, 1972.

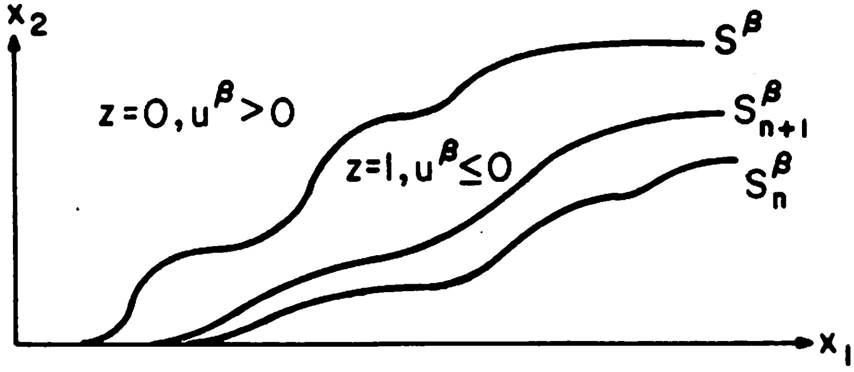
Figure Captions

Fig. 1. The Markov decision problem.

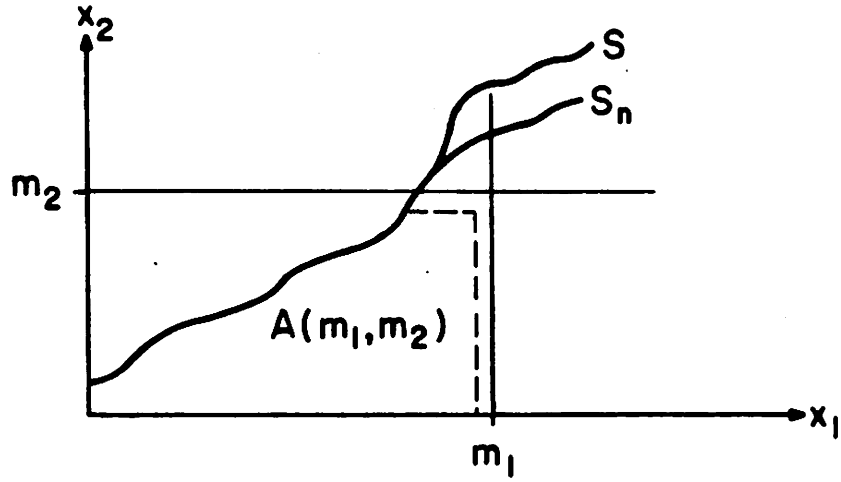
Fig. 2. Switching curves, $\beta < 1$.

Fig. 3. Illustration for Lemma 5.4.


x_{k1}
x_{k2}
$z_k = 1$
$z_k = 0$



x_2
S^{β}
$z=0, u^{\beta}>0$
S^{β}_{n+1}
$z=1, u^{\beta}\leq 0$
S^{β}_{n}
x_1


x2
S
Sn
m2
A(m1,m2)
x1
m1