Optimal control of the M/G/1/ queue with repeated vacations of the server

Eitan Altman, Philippe Nain

UNITÉ DE RECHERCHE
INRIA-SOPHIA ANTIPOLIS

Rapports de Recherche

N°1637

Programme 1
Architecture parallèles, Bases de données
Réseaux et Systèmes distribués

OPTIMAL CONTROL OF THE M/G/1 QUEUE WITH REPEATED VACATIONS OF THE SERVER

Eitan ALTMAN
Philippe NAIN

Mars 1992

Institut National
de Recherche
en Informatique
et en Automatique

Sophia Antipolis
B.P. 109
06561 Valbonne Cedex
France
Tél.: 93 65 77 77

OPTIMAL CONTROL OF THE M/G/1 QUEUE WITH REPEATED VACATIONS OF THE SERVER*

Eitan ALTMAN and Philippe NAIN
INRIA
Centre Sophia Antipolis
06565 Valbonne Cedex, France
altman@sophia.inria.fr nain@sophia.inria.fr

Abstract

We consider an M/G/1 queue where the server may take repeated vacations. Whenever a busy period terminates (i.e., when the queue empties) the server takes a vacation of random duration. At the end of each vacation the server may either take a new vacation or resume service provided that the system is nonempty. The decision to turn the server on/off may depend on all the history of the process (which includes the number of customers and all past decisions). The optimization problem typically arises when imposing a cost structure that involves a holding cost per unit time and per customer and a cost for turning the server on (a shut-down cost may also be included in the latter cost). One may wish to restrict to threshold policies where the server is turned on at the end of each vacation if and only if the queue-length is greater than or equal to a fixed threshold. A few recent papers address the problem of optimally choosing the threshold. The objective of this paper is to establish the optimality of threshold policies over all policies for two long-run average cost criteria.

*Work partially supported by the France-Israel Scientific Cooperation in Computer Science and Engineering under Grant 3321190.

1 The Optimization Problem

We consider an M/G/1 queue where the server may take repeated vacations. Whenever a busy period terminates (i.e., when the queue empties) the server takes a vacation whose duration is distributed like a generic random variable (r.v.) D with Laplace-Stieltjes Transform (LST) $d^*(\cdot)$, first moment $0 < d < \infty$ and second moment $d^{(2)} < \infty$. The durations of the vacation periods are assumed to be mutually independent r.v.'s, independent of the arrival and service processes. At the end of each vacation the server may either take a new vacation of random duration D or resume service provided that the system is nonempty. The cost structure includes a customer holding cost of rate 1 and a constant cost $\gamma \geq 0$ that is incurred each time the server is turned on (γ may also include a constant shut-down cost for turning the server off).

Let $X(t)$ be the queue length at time $t \geq 0$. The sample paths of the process $\{X(t), t \geq 0\}$ are assumed to be right-continuous. Define t_n ($n \geq 1$) to be the n-th decision epoch, namely the n-th vacation completion time. We shall assume without loss of generality that $t_1 = 0$. Also define $\mathbb{R}$ (resp. $\mathbb{N}$) to be the set of real numbers (resp. nonnegative integers).

Given the above cost structure, a natural objective is to solve the following optimization problem:

Problem P1: Minimize the long-run average operating cost

$$V(x,u) := \overline{\lim_{t\to\infty}} \frac{1}{t} E^u \left[\int_0^t X(\xi)\, d\xi + \gamma \sum_{n=1}^{N(t)} \mathbf{1}(A_n = s) \,|\, X(0) = x \right], \tag{1.1}$$

for all $x \in \mathbb{N}$ and for any admissible vacation policy u (the precise definition of an admissible vacation policy is given in Section 2), where A_n denotes the action chosen at time t_n ($A_n = s$ if the decision is to serve and $A_n = v$ if the decision is to take another vacation) and $N(t) := \sup\{n \geq 1 : t_n < t\}$ for all $t > 0$.

While the authors were polishing this paper, a similar study by Federgruen and So [6] appeared (see Remark 1.1). We have however made the decision to proceed with this paper for several reasons: Federgruen and So do not consider the problem **P1** but instead a different version of this problem (called problem **P2** in Section 2; see also Remark 2.1); the proof of Theorem 3.2 in [6] is not clearly established since it relies upon a result by Schweitzer [17] that can only be used in the case of finite state spaces (see Remark 4.1); our approach (value iteration) is completely different from the approach used by Federgruen and So (the method of proof in [6] is based on a systematic variation of the model parameter γ) and appears to be simpler.

The contributions of the present paper are the following: first, we (rigorously) solve the problem **P2** over the class of all policies; second, we solve the problem **P1** over the (broad) class of all *regenerative* policies. For both problems the optimality of a threshold policy, namely a policy that turns the server on if and only if the queue-length is greater than or equal to a fixed threshold, is established; in particular, we show that the optimal threshold is bounded from above by a known constant. We also believe that the way the value iteration algorithm is used in this paper is of interest since the properties to be propagated in the induction of dynamic programming appear to be non-trivial, and different from the standard monotonicity, convexity, concavity or supermodularity properties.

The literature on vacation queueing models is rapidly growing. This is because these models provide an ad hoc formalism for the study of various discrete event systems ranging from production systems to communication and computer systems (see Doshi [5] for a survey paper). Three types of server vacation schemes are commonly encountered in the literature: the scheme with repeated vacations of the server that has just been described above (see also Gelenbe and Iasnogorodski [7], Gelenbe and Mitrani [8], Kella [10, 11], Levy and Yechiali [13]); the scheme where the server may resume service upon the arrival of a new customer (the so-called "removable server", see Heyman and Sobel [9, pp. 336-337], Yadin and Naor [20]) and a mixture of those two schemes (Doğanata [4]).

Optimization issues for queueing models with server vacations have already received some attention. In [19] Talman solved a control problem for the M/G/1 queue with a removable server and proved the existence of an optimal threshold type. Lately, Kella [10] addressed the problem of computing the optimal threshold policy over the class of all threshold policies for the M/G/1 queue with repeated vacations of the server. Lee and Srinivasan [12] carried out the same analysis in the case of batch arrivals. In [1] Altman and Nain solved the problem **P1** for the M/M/1 queue with exponential repeated vacations.

The paper is organized as follows. Since the problem **P1** cannot be tackled directly, we instead introduce in Section 2, and solve in Sections 3 and 4, an intermediate problem using the value iteration approach. The solution to this problem will in turn enable us to solve the original problem **P1** in Section 5.

A last word about the notation: $\lambda > 0$ will denote the arrival intensity and $b < \infty$ (resp. $b^{(2)}$) the first (resp. second) moment of the service time distribution. We further assume that $\rho := \lambda b < 1$ (stability condition).

Remark 1.1 The model investigated by Federgruen and So [6] is more general than the one described in this section. In particular, they allow for batch arrivals and system dependent holding cost rates. However, the analysis developed in the forthcoming sections extends to their model provided that Conditions 1 and 2 in [6, p. 391] are satisfied.

2 A Semi-Markov Decision Problem

Solving directly for problem **P1** is a difficult task, since the structure of the cost criterion (1.1) does not fit the standard Semi-Markov Decision Process (SMDP) setting (cf. Lippman [14], Ross

[16]). To see that, let us rewrite $V(x,u)$ as, cf. (1.1),

$$V(x,u) = \overline{\lim_{t\to\infty}} \frac{1}{t} E_x^u \left[\sum_{0\le t_n<t} \left(\int_{t_n}^{t_{n+1}} X(\xi)\, d\xi + \gamma\, \mathbf{1}(A_n = s) \right) - \int_t^{t_{N(t)+1}} X(\xi)\, d\xi \right], \quad x \in \mathbb{N}, \quad (2.1)$$

where the symbol E_x^u stands for the expectation operator given that $X(0) = x$ and that policy u is employed. We observe from (2.1) that the cost incurred in $[t_n, t_{n+1})$ $(n \ge 1)$ given $X(t_n) = x$ and $A_n = a$ is not a *deterministic* function of (x,a), which precludes the use of the SMDP theory.

So, instead of starting with the cost function (2.1), we shall first address the following optimization problem:

Problem P2: Minimize the long-run expected average operating cost

$$\Phi(x,u) := \overline{\lim_{t\to\infty}} \frac{1}{t} E_x^u \left[\sum_{0\le t_n<t} E\left[\int_{t_n}^{t_{n+1}} X(\xi)\, d\xi + \gamma\, \mathbf{1}(A_n = s) \,|\, X_n, A_n \right] \right], \quad (2.2)$$

for all $x \in \mathbb{N}$ and for any admissible vacation policy u (the precise definition of an admissible vacation policy is given at the end of this section). Observe now that the cost incurred in $[t_n, t_{n+1})$ $(n \ge 1)$ is a deterministic function of (x,a) when $X_n = x$ and $A_n = a$. The end of the section is devoted to showing that (2.2) is a long-run expected average cost associated with an SMDP.

This SMDP is defined as follows (Lippman [14]):

The state of the process The state of the process is $X(t)$ for all $t \ge 0$. Define $X_n := X(t_n)$ for all $n \ge 1$. Recall that $t_1 = 0$ so that $X_1 = X(0)$.

The action space. The action space is $\{s, v\}$ when the system is in state $x \ne 0$ and $\{v\}$ otherwise.

The expected cost. If action a is chosen when the system is in state x, then the immediate expected cost

$$z(x,a) := E\left[\int_0^{t_2} X(\xi)\, d\xi + \gamma\, \mathbf{1}(a = s), \,|\, X_1 = x, A_1 = s \right], \quad (2.3)$$

is incurred. It is a simple exercise (see Altman and Nain [2, Appendix A]) to show that

$$z(x,a) = \begin{cases} xd + \dfrac{\lambda d^{(2)}}{2}, & \text{if } a = v \text{ and } x \in \mathbb{N}; \\[2ex] \dfrac{b}{2(1-\rho)}\, x^2 + \dfrac{\lambda b^{(2)} + b(1-\rho)}{2(1-\rho)^2}\, x + \dfrac{\lambda d^{(2)}}{2} + \gamma, & \text{if } a = s \text{ and } x \ge 1. \end{cases} \quad (2.4)$$

The transition times. Let $\tau(x,a)$ be the conditional sojourn time in state x given action a is chosen. If $a = v$ then $\tau(x,a)$ is distributed like D for all $x \in \mathbb{N}$. If $a = s$ and $x \ge 1$ then $\tau(x,a)$ is distributed as the sum of x i.i.d. busy periods of an M/G/1 queue plus a vacation of length D. For

$\Re(\omega) \geq 0$, let $\tau^*_{x,a}(\omega) := E[e^{-\omega\tau(x,a)}]$ be the LST of $\tau(x,a)$. Let $T^*(\omega)$ denote the LST of a busy period duration in an M/G/1 queue. Clearly, for $\Re(\omega) \geq 0$,

$$\tau^*_{x,a}(\omega) = \begin{cases} d^*(\omega), & \text{if } a = v \text{ and } x \in \mathbb{N}; \\ d^*(\omega)\, [T^*(\omega)]^x, & \text{if } a = s \text{ and } x \geq 1. \end{cases} \tag{2.5}$$

The transition law. If the process is in state x and action a is chosen then the next state will be y with the probability

$$P_{xy}(a) := \begin{cases} P_d(y), & \text{if } a = s \text{ and } x \geq 1; \\ \\ P_d(y-x), & \text{if } a = v \text{ and } x \leq y; \\ \\ 0, & \text{if } a = v \text{ and } x > y, \end{cases} \tag{2.6}$$

where $P_d(y)$ stands for the probability of y arrivals during a vacation period.

It is seen that the stochastic process $\{X(t),\, t \geq 0\}$ endowed with the above structure is an SMDP with state space $\mathbb{N}$ (see Lippman [14] for instance). Further, cf. (2.2), (2.3),

$$\Phi(x,u) = \overline{\lim_{t\to\infty}} \frac{1}{t} E^u_x \left[\sum_{0 \leq t_n < t} z(X_n, A_n) \right], \tag{2.7}$$

for all $u \in \mathcal{U}$, $x \in \mathbb{N}$. In other words, as announced $\Phi(x,u)$ is a long-run expected average cost associated with the semi-Markov decision process $\{X(t),\, t \geq 0\}$.

We conclude this section by giving a precise definition of an admissible vacation policy. Given $(X_i = x_i, A_i = a_i)$, $1 \leq i \leq n-1$ and $X_n = x_n$, the n-th action A_n will depend (possibly in a random way) on the history $h_n := (x_1, \ldots, x_n, a_1, \ldots, a_{n-1})$ with $h_1 := x_1$. An admissible vacation policy u is then defined as any collection $\{u_n\}_1^\infty$ of mappings $u_n : \mathbb{N}^n \times \{s,v\}^{n-1} \to [0,1]$ with the interpretation that the n-th decision is to serve (resp. to take another vacation) with the probability $u_n(h_n)$ (resp. $1 - u_n(h_n)$) whenever the information h_n is available to the decision-maker. In the sequel, $\mathcal{U}$ will denote the collection of all admissible vacation policies. As usual, a policy $u \in \mathcal{U}$ is said to be stationary if u_n only depends on X_n and if it is nonrandomized (i.e., $u_n(\cdot) \in \{0,1\}$). For every stationary policy $u \in \mathcal{U}$, the notation $u(x)$ will stand for the action to be chosen when the system is in state $x \in \mathbb{N}$. In particular, the stationary policy u that satisfies $u(x) = v$ for all $x < l$ and $u(x) = s$ for all $s \geq l$, $l \geq 1$, will be called a threshold policy with threshold l and denoted by u_l.

Remark 2.1 The objective function in Federgruen and So [6] is not explicitly defined. However, the technique used by these authors and the definition of the one-step expected cost (see (2) in [6]) seem to indicate that their objective function is (2.2).

3 The Value Iteration Approach

In this section, we introduce and solve a discounted version of problem **P2**. As it is often the case in this context, the solution to this discounted cost problem will yield the solution to the average cost problem **P2** (see Section 4).

Define

$$W_\alpha(x,u) := E_x^u \left[\sum_{n=1}^{\infty} z(X_n, A_n) \exp\left(-\alpha \sum_{j=1}^{n-1} \tau(X_j, A_j) \right) \right], \tag{3.1}$$

for $\alpha > 0$, $x \in \mathbb{N}$, $u \in \mathcal{U}$, and let

$$W_\alpha(x) := \inf_{u \in U} W_\alpha(x,u). \tag{3.2}$$

A policy $u_\alpha \in \mathcal{U}$, $\alpha > 0$, is said to be α-discounted optimal if $W_\alpha(x) = W_\alpha(x, u_\alpha)$ for all $x \in \mathbb{N}$.

Let $\mathcal{K}$ be the set of functions $f : \mathbb{N} \to \mathbb{R}$ such that $||f|| := \sup_{x \in \mathbb{N}} |f(x)| \max(x,1)^{-2} < \infty$. Define the Dynamic Programming (DP) operator $T_\alpha : \mathcal{K} \to \mathcal{K}$ by

$$T_\alpha f(x) := \min_a \left\{ z(x,a) + \tau^*_{x,a}(\alpha) \sum_{y=0}^{\infty} P_{xy}(a)\, f(y) \right\}, \quad f \in \mathcal{K},\ x \in \mathbb{N}. \tag{3.3}$$

The proof that $T_\alpha(f) \in \mathcal{K}$ for every $f \in \mathcal{K}$ can be found in Lippman [14] (in our case the constant b that appears in Assumption 2 in [14] can be chosen to be $1 + \lambda d + \lambda^2 d^{(2)}$).

The following results hold:

Proposition 3.1 *For every $\alpha > 0$, $W_\alpha(x)$ satisfies the optimality equation*

$$W_\alpha(x) = T_\alpha W_\alpha(x), \quad x \in \mathbb{N}, \tag{3.4}$$

and further,

$$\lim_{n \to \infty} T_\alpha^n g(x) = W_\alpha(x), \quad x \in \mathbb{N}, \tag{3.5}$$

for every function $g \in \mathcal{K}$. In addition, any stationary policy that minimizes the right-hand side of (3.4) is α-discounted optimal.

Proof. See Lippman [14, Theorem 1]. ■

Define $\mathbb{N}^* := \mathbb{N} - \{0\}$. A set of functions $f_\alpha \in \mathcal{K}$, $\alpha > 0$, is said to satisfy condition

C1 if for any $x \in \mathbb{N}$, $k \in \mathbb{N}^*$, there exists $\alpha_{x,k} > 0$ such that

$$f_\alpha(x + r + k) - [T^*(\alpha)]^k f_\alpha(x + r) \geq z(x + k, s) - z(x, s), \tag{3.6}$$

for all $\alpha \in (0, \alpha_{x,k})$, $r \in \mathbb{N}$;

C2 if for any $\alpha > 0$, $x \in \mathbb{N}$, $k \in \mathbb{N}^*$,

$$f_\alpha(x+k) - [T^*(\alpha)]^k f_\alpha(x) \geq 0. \tag{3.7}$$

It is easily seen from (2.4) that the set of functions $f_\alpha(\cdot) := z(\cdot, s)$ satisfies **C1** and **C2**.

The following lemma holds:

Lemma 3.1 *W_α satisfies conditions* **C1** *and* **C2**.

Proof. Let $f_\alpha \in \mathcal{K}$ be an arbitrary set of functions that satisfies **C1** and **C2**. Let us show that $T_\alpha f_\alpha$ also satisfies these conditions. Throughout the proof $\alpha > 0$, $(x,r) \in \mathbb{N}^2$ and $k \in \mathbb{N}^*$ are fixed numbers.

We have, cf. (2.6) and (3.3),

$$\begin{aligned}
& T_\alpha f_\alpha(x+r+k) - [T^*(\alpha)]^k T_\alpha f_\alpha(x+r) = \\
& \qquad \min\left\{ z(x+r+k,v) + d^*(\alpha) \sum_{y=0}^{\infty} P_d(y) f_\alpha(x+r+k+y)\,, \right. \\
& \qquad \left. z(x+r+k,s) + d^*(\alpha)\, [T^*(\alpha)]^{x+r+k} \sum_{y=0}^{\infty} P_d(y) f_\alpha(y) \right\} \\
& \qquad - [T^*(\alpha)]^k \min\left\{ z(x+r,v) + d^*(\alpha) \sum_{y=0}^{\infty} P_d(y) f_\alpha(x+r+y)\,, \right. \\
& \qquad \left. z(x+r,s) + d^*(\alpha)\, [T^*(\alpha)]^{x+r} \sum_{y=0}^{\infty} P_d(y) f_\alpha(y) \right\}, \\
& \geq \quad \min\left\{ z(x+r+k,v) - [T^*(\alpha)]^k z(x+r,v) \right. \\
& \qquad + d^*(\alpha) \sum_{y=0}^{\infty} P_d(y) \Big(f_\alpha(x+r+k+y) - [T^*(\alpha)]^k f_\alpha(x+r+y) \Big)\,, \\
& \qquad \left. z(x+r+k,s) - [T^*(\alpha)]^k z(x+r,s) \right\}, \qquad (3.8) \\
& > \quad \min\left\{ kd + d^*(\alpha) \sum_{y=0}^{\infty} P_d(y) \Big(f_\alpha(x+r+k+y) - [T^*(\alpha)]^k f_\alpha(x+r+y) \Big)\,, \right. \\
& \qquad \left. z(x+k,s) - z(x,s) \right\}. \qquad (3.9)
\end{aligned}$$

The inequality (3.8) follows from the inequality $\min(a,b) - \min(c,d) \geq \min(a-c, b-d)$; the inequality (3.9) follows from the definition of $z(x,a)$ together with the fact that $T^*(\alpha) \leq 1$.

Letting $r = 0$ in (3.9) it is seen that $T_\alpha f_\alpha$ satisfies **C2** because f_α satisfies **C2** (which implies that the first argument of the min is nonnegative) and that $x \to z(x,s)$ is nondecreasing (see (2.4)).

Hence, we may deduce by induction that $T_\alpha^n f_\alpha$ satisfies **C2** for all $n \geq 1$. Consequently, by Proposition 3.1

$$W_\alpha(x+k) - [T^*(\alpha)]^k W_\alpha(x) = \lim_{n\to\infty} \left[T_\alpha^n f_\alpha(x+k) - [T^*(\alpha)]^k T_\alpha^n f_\alpha(x)\right] \geq 0, \tag{3.10}$$

which shows that W_α satisfies **C2**.

On the other hand, since f_α satisfies **C1** by assumption and since $kd > 0$, it is seen from (3.9) that there exists $\alpha_{x,k} > 0$ such that

$$T_\alpha f_\alpha(x+r+k) - [T^*(\alpha)]^k \, T_\alpha f_\alpha(x+r) \geq z(x+k,s) - z(x,s), \tag{3.11}$$

for all $\alpha \in (0, \alpha_{x,k})$. The proof is concluded by using (3.10). ∎

Let u_α^* be a stationary policy that minimizes the right-hand side of the DP equation (3.4). Also define

$$N_0 := \inf_{x\in\mathbb{N}} \left\{ x > \gamma \left(\frac{d}{1-\rho}\right)^{-1} \right\}. \tag{3.12}$$

It is worth observing that the constant $d/(1-\rho)$ that appears in the definition of N_0 is the expected duration between two consecutive vacation completion times if one uses the vacation policy that always turns the server on when the queue is nonempty (i.e., the threshold policy with threshold 1).

The following lemma partially characterizes the α-discounted optimal policy for small discount factors.

Lemma 3.2 *There exists $\alpha^* > 0$ such that for every $\alpha \in (0, \alpha^*)$, there exist two integers L_α and N_α such that*

$$u_\alpha^*(x) = v \quad \textit{if} \quad 0 \leq x < L_\alpha; \tag{3.13}$$

$$u_\alpha^*(x) = s \quad \textit{if} \quad L_\alpha \leq x < N_\alpha, \tag{3.14}$$

for all $x \in \mathbb{N}$, where $0 \leq L_\alpha \leq N_0$ and

$$N_\alpha \to \infty \quad \textit{when} \quad \alpha \to 0. \tag{3.15}$$

Proof. Using (2.5), (2.6) and (3.3), the DP equation (3.4) becomes

$$W_\alpha(x) = \min\left\{ z(x,v) + d^*(\alpha) \sum_{y=0}^{\infty} P_d(y) W_\alpha(x+y),\, z(x,s) + d^*(\alpha)\, [T^*(\alpha)]^x \sum_{y=0}^{\infty} P_d(y) W_\alpha(y) \right\}, \tag{3.16}$$

for all $x \in \mathbb{N}$.

We first show that for any $x \geq N_0$, there exists $\alpha_x > 0$ such that for all $\alpha \in (0, \alpha_x)$, $u^*_\alpha(x) = s$, or equivalently from (3.16),

$$\Delta_\alpha(x) := z(x,v) - z(x,s) + d^*(\alpha) \sum_{y=0}^{\infty} P_d(y) \left(W_\alpha(x+y) - [T^*(\alpha)]^x W_\alpha(y)\right) > 0. \tag{3.17}$$

Fix $x \geq N_0$. Since $W_\alpha(x+y) - [T^*(\alpha)]^x W_\alpha(y)$ is nonnegative for all $\alpha > 0$, $y \in \mathbb{N}$ (for W_α satisfies **C2**), we may apply Fatou's lemma to (3.17) to obtain

$$\begin{aligned}
\varliminf_{\alpha \to 0} \Delta_\alpha(x) &\geq z(x,v) - z(x,s) + \sum_{y=0}^{\infty} P_d(y) \varliminf_{\alpha \to 0} \left(W_\alpha(x+y) - [T^*(\alpha)]^x W_\alpha(y)\right), \\
&\geq z(x,v) - z(x,s) + \sum_{y=0}^{\infty} P_d(y) \left(z(x+y,s) - z(y,s)\right), \text{ since } W_\alpha \text{ satisfies } \mathbf{C1}, \\
&= \frac{xd}{1-\rho} - \gamma > 0, \qquad (3.18)
\end{aligned}$$

by using (2.4), (3.12) and the identity $\sum_{y=0}^{\infty} y P_d(y) = \lambda d$. This proves (3.14) and (3.15).

Let us now turn to the proof of (3.13). Since we have already shown that $\varliminf_{\alpha \to 0} \Delta_\alpha(x) > 0$ for $x \geq N_0$, it suffices to show that, for α small enough, the mapping $x \to \Delta_\alpha(x)$ is nondecreasing for $0 \leq x \leq N_0 - 1$.

We have for $x \in \mathbb{N}$, cf. (2.4), (3.17),

$$\begin{aligned}
\Delta_\alpha(x+1) - T^*(\alpha)\Delta_\alpha(x) &= (1 - T^*(\alpha)) \left(dx + \frac{\lambda d^{(2)}}{2}\right) + d - \left(z(x+1,s) - T^*(\alpha)\, z(x,s)\right) \\
&\quad + d^*(\alpha) \sum_{y=0}^{\infty} P_d(y) \left(W_\alpha(x+1+y) - T^*(\alpha) W_\alpha(x+y)\right). \qquad (3.19)
\end{aligned}$$

Letting $\alpha \to 0$ in (3.19), we have again by Fatou's lemma

$$\begin{aligned}
\varliminf_{\alpha \to 0} \{\Delta_\alpha(x+1) - T^*(\alpha)\Delta_\alpha(x)\} &\geq d - \left(z(x+1,s) - z(x,s)\right) \\
&\quad + \sum_{y=0}^{\infty} P_d(y) \varliminf_{\alpha \to 0} \left(W_\alpha(x+1+y) - T^*(\alpha) W_\alpha(x+y)\right), \\
&\geq d, \qquad (3.20)
\end{aligned}$$

where (3.20) follows from the fact that W_α satisfies condition **C1**. Combining (3.20) together with the inequality

$$\begin{aligned}
\varliminf_{\alpha \to 0} \left(\Delta_\alpha(x+1) - T^*(\alpha)\, \Delta_\alpha(x)\right) &= \varliminf_{\alpha \to 0} \left(\Delta_\alpha(x+1) - \Delta_\alpha(x) + (1 - T^*(\alpha))\, \Delta_\alpha(x)\right), \\
&\leq \varliminf_{\alpha \to 0} \left(\Delta_\alpha(x+1) - \Delta_\alpha(x)\right) + \varlimsup_{\alpha \to 0} (1 - T^*(\alpha))\, \Delta_\alpha(x),
\end{aligned}$$

we see that the nondecreasingness property of the mapping $x \to \Delta_\alpha(x)$ for $x \leq N_0 - 1$ and for α small enough is established if one can show that

$$\overline{\lim_{\alpha\to 0}}\,(1 - T^*(\alpha))\,\Delta_\alpha(x) \leq 0, \tag{3.21}$$

for $x = 0, 1, \ldots, N_0 - 2$. The remainder of the proof is devoted to proving (3.21).

It is seen from (3.20) that there exists $\beta > 0$, such that for $0 < \alpha < \beta$,

$$\Delta_\alpha(x) \leq \frac{\Delta_\alpha(N_0)}{[T^*(\alpha)]^{N_0 - x}}, \tag{3.22}$$

for $x = 0, 1, \ldots, N_0 - 1$.

On the other hand, since $1 - T^*(\alpha) \leq \alpha\, b/(1-\rho)$ for all $\alpha \geq 0$, we have for $x \leq N_0 - 1$

$$\begin{aligned}
\overline{\lim_{\alpha\to 0}}\,(1 - T^*(\alpha))\,\Delta_\alpha(x) &\leq \overline{\lim_{\alpha\to 0}}\,(1 - T^*(\alpha))\,\frac{\Delta_\alpha(N_0)}{[T^*(\alpha)]^{N_0 - x}}, \ \text{ from (3.22)}, \\
&\leq \left(\frac{b}{1-\rho}\right) \overline{\lim_{\alpha\to 0}}\,\alpha\,\Delta_\alpha(N_0).
\end{aligned} \tag{3.23}$$

Let us prove that $\Delta_\alpha(N_0)$ is bounded from above by a number that does not depend on α, which will prove (3.21).

From (3.17) and the fact that W_α satisfies **C2**, we have for all $\alpha > 0$

$$\begin{aligned}
\Delta_\alpha(N_0) &\leq z(N_0, v) + \sum_{y=0}^{\infty} P_d(y)\left(W_\alpha(N_0 + y) - [T^*(\alpha)]^{N_0}\, W_\alpha(y)\right), \\
&= z(N_0, v) + \sum_{y=0}^{\infty} P_d(y)\,(W_\alpha(N_0 + y) - [T^*(\alpha)]^{y}\, W_\alpha(N_0)) \\
&\quad + \sum_{y=0}^{\infty} P_d(y)\,[T^*(\alpha)]^{y}\left(W_\alpha(N_0) - [T^*(\alpha)]^{N_0 - y}\, W_\alpha(y)\right).
\end{aligned} \tag{3.24}$$

We first compute an upper bound for the first summation in the right-hand side of (3.24). Since $\underline{\lim}_{\alpha\to 0}\,\Delta_\alpha(N_0) > 0$ (see (3.18)), it follows that there exists $\alpha_{N_0} > 0$ such that $\Delta_\alpha(N_0) > 0$ for $0 < \alpha < \alpha_{N_0}$, that is the α-optimal action in state N_0 is to serve whenever α is small enough. Let w_α be the policy that serves at the first decision epoch, and then follows the α-optimal policy.

For $0 < \alpha < \alpha_{N_0}$, we have

$$\begin{aligned}
&\sum_{y=0}^{\infty} P_d(y)\,(W_\alpha(N_0 + y) - [T^*(\alpha)]^{y}\, W_\alpha(N_0)) \leq \\
&\qquad \sum_{y=0}^{\infty} P_d(y)\,(W_\alpha(N_0 + y, w_\alpha) - [T^*(\alpha)]^{y}\, W_\alpha(N_0)),
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{y=0}^{\infty} P_d(y)\left(z(N_0+y,s) - [T^*(\alpha)]^y\, z(N_0,s)\right), \ \ \text{(hint: use (3.1))}, \\
&\le \sum_{y=0}^{\infty} P_d(y)\, z(N_0+y,s) = z(N_0,s) + \frac{N_0 \rho\, d}{1-\rho} - \frac{\lambda d^{(2)}}{2} - \gamma. \qquad (3.25)
\end{aligned}
$$

Next, we compute an upper bound for the term in the second summation in the right-hand side of (3.24). For $y \geq N_0$, we have by condition **C2**,

$$
W_\alpha(N_0) - [T^*(\alpha)]^{N_0-y}\ W_\alpha(y) = -\,[T^*(\alpha)]^{N_0-y}\left(W_\alpha(y) - [T^*(\alpha)]^{y-N_0}\ W_\alpha(N_0)\right) \le 0. \qquad (3.26)
$$

Consider now y such that $y < N_0$. If the α-optimal policy serves in state y, then we have for $0 < \alpha < \alpha_{N_0}$ (use (3.1) again),

$$
W_\alpha(N_0) - [T^*(\alpha)]^{N_0-y}\ W_\alpha(y) \le z(N_0,s). \qquad (3.27)
$$

If the α-optimal policy does not serve in state y, then for $0 < \alpha < \alpha_{N_0}$,

$$
\begin{aligned}
& W_\alpha(N_0) - [T^*(\alpha)]^{N_0-y}\ W_\alpha(y) \\
&\le\ z(N_0,s) + d^*(\alpha)\ [T^*(\alpha)]^{N_0-y} \sum_{y'=0}^{\infty} P_d(y')\left(W_\alpha(y')\,[T^*(\alpha)]^y - W_\alpha(y+y')\right), \\
&\le\ z(N_0,s), \qquad (3.28)
\end{aligned}
$$

since $W_\alpha(y+y') - [T^*(\alpha)]^y\, W_\alpha(y') \geq 0$ for all $(y,y') \in \mathbb{N}^2$, $\alpha > 0$ from Lemma 3.1.

Combining (3.24)-(3.28), we get that for all $\alpha \in (0, \alpha_{N_0})$

$$
\Delta_\alpha(N_0) \le z(N_0,s) + z(N_0,v) + \frac{N_0 \rho d}{1-\rho} < \infty,
$$

which completes the proof (with $\alpha^* := \min(\beta, \alpha_{N_0})$). ∎

What Lemma 3.2 says is that for small $\alpha > 0$ there exists an integer N_α that goes to ∞ when α goes to 0 such that the optimal α-discounted policy is a threshold policy in $\{0,1,\ldots,N_\alpha - 1\}$. Observe that Lemma 4.2 does not say anything about the optimal decision when the system is in state $x \geq N_\alpha$.

From Lemma 3.2, we immediately deduce the following

Corollary 3.1 *The limit as α converges to 0 of the α-discounted optimal policy u^*_α is a threshold policy u_L with threshold $L \leq N_0$.*

4 The Long-Run Expected Average Cost Problem P2

The threshold policy u_L given in Corollary 3.1 is our candidate for the solution to the problem **P2**. One way of proving that u_L is **P2**-average optimal is to apply Theorem 2 in Sennott [18].

To this end, we first need to introduce the long-run expected average cost (see Ross [16, p. 159] and Sennott [18, p. 249] where this cost is also used)

$$\Psi(x,u) := \overline{\lim_{n\to\infty}} \frac{E_x^u[Z_n]}{E_x^u[t_n]}, \tag{4.1}$$

for $x \in \mathbb{N}$, $u \in \mathcal{U}$, where

$$Z_n := E_x^u \left[\sum_{i=1}^{n} z(X_i, A_i) \right], \tag{4.2}$$

is the total expected cost incurred by the system in $[0, t_{n+1})$, $n \geq 1$.

Then, the following result is needed:

Lemma 4.1 *For every $l \geq 1$, $x \in \mathbb{N}$,*

$$\Psi(x, u_l) = \Phi(x, u_l), \tag{4.3}$$

where $\Psi(x,u)$ has been defined in (4.1). Moreover, $\Phi(x, u_l)$ does not depend on x and is finite.

Proof. Assume that the system evolves under a threshold policy u_l and define

$$S_x := \inf\{t_n, n \geq 2 : X(t_n) = x\}, \tag{4.4}$$

given $X(0) = x$, $x \in \mathbb{N}$. Clearly, for every $x \in \mathbb{N}$, the process $\{X(t), t \geq 0\}$ is a regenerative process with regeneration point S_x. If $E_x^{u_l}[S_x] < \infty$ for every $x \in \mathbb{N}$, then (4.3) follows from Theorem 7.5 in [16]. Let us show that $E_x^{u_l}[S_x] < \infty$ for every $x \in \mathbb{N}$.

It is easily seen that $\mathbf{X} := \{X_n\}_{n\geq 1}$ is an irreducible, aperiodic Markov chain. Let us show that all the states of this chain are recurrent non-null. Clearly, under policy u_l

$$\begin{aligned} X_{n+1} &= U_n + X_n \mathbf{1}(X_n < l), \\ &\leq U_n + l, \end{aligned} \tag{4.5}$$

for all $n \geq 1$, where U_n denotes the number of arrivals during the n-th vacation period. Observe from Section 2 that the U_n's are i.i.d. random variables with probability distribution function $P_d(\cdot)$.

On the other hand, we have using (4.5)

$$P_l^{u_l}(X_{n+1} = l) = \sum_{j\in\mathbb{N}} P_l^{u_l}(U_n + X_n \mathbf{1}(X_n < l) = l, X_n = j),$$

$$\geq \quad P_l^{u_l}(X_n = l)\, P_d(l) + \sum_{j=0}^{l-1} P_l^{u_l}(X_n = j)\, P_d(l-j), \tag{4.6}$$

$$\geq \quad c\, P_l^{u_l}(X_n \leq l),$$

$$\geq \quad c\, P_d(0) > 0, \tag{4.7}$$

with $c := \min_{1\leq j\leq l} P_d(j)$. To derive (4.6) we have used the fact that X_n and U_n are independent r.v.'s, whereas (4.7) follows from (4.5). Consequently, the state l is recurrent non-null, which in turn entails that all the states are recurrent non-null since the Markov chain $\mathbf{X}$ is irreducible.

Let $\{\nu(y)\}_{y\in\mathbb{N}}$ denote the limiting probabilities of the irreducible, aperiodic and recurrent non-null Markov chain $\mathbf{X}$. Then, a standard result from the theory of regenerative processes (see Theorem 10.4.3 in Cinlar [3]) ensures that

$$\begin{aligned}
E_x^{u_l}[S_x] &= \frac{\sum_{y\in\mathbb{N}} E_y^{u_l}[t_2]\,\nu(y)}{\nu(x)}, \\
&= \frac{d\sum_{y=0}^{l-1}\nu(y) + (b/(1-\rho))\sum_{y\geq l} y\,\nu(y)}{\nu(x)}, \\
&\leq \frac{d + (b/(1-\rho))\,(l+\lambda d)}{\nu(x)} < \infty,
\end{aligned}$$

for every $x \in \mathbb{N}$, since $\sum_{y\in\mathbb{N}} y\,\nu(y) \leq E[U_n + l] = l + \lambda d$ from (4.5), which proves (4.3).

We now prove the second part of the lemma. Since the Markov chain $\mathbf{X}$ is recurrent non-null, this implies that $\{X(t),\, t\geq 0\}$ is a delayed regenerative process. Therefore (see Ross [16, Remark, p. 161]),

$$\Phi\,(x, u_l) = \Phi\,(y, u_l) := \Phi_l,$$

for all $x \neq y$. It remains to show that $\Phi_l < \infty$. Fix $x > l$.

Using again Theorem 7.5 in [16] we get that $\Phi_l = E_x^{u_l}[Z(S_x)]/E_x^{u_l}[S_x]$, where $Z(t)$ is the total cost incurred in $[0,t)$. Since $E_x^{u_l}[S_x] \geq d > 0$, we are left to prove that $E_x^{u_l}[Z(S_x)]$ is finite. It is easily seen by definition of $z(x,a)$ that there exist three constants $a_i < \infty$, $i = 1,2,3$, such that

$$z(x,a) \leq a_1 + a_2 x + a_3 x^2,$$

for all $x \in \mathbb{N}$, $a \in \{s, v\}$. Hence,

$$\begin{aligned}
E_x^{u_l}[Z(S_x)] &= E_x^{u_l}\left[\sum_{n=1}^{N_x} z(X_n, A_n)\right], \qquad (4.8)\\
&= z(x,s) + E_x^{u_l}\left[\sum_{n=1}^{N_x-1} z(X_{n+1}, A_{n+1})\right], \\
&\leq z(x,s) + a_1\, E_x^{u_l}[N_x] + a_2\, E_x^{u_l}\left[\sum_{n=1}^{N_x} X_{n+1}\right] + a_3\, E_x^{u_l}\left[\sum_{n=1}^{N_x} X_{n+1}^2\right],
\end{aligned}$$

$$\leq \quad z(x,s) + b_1\, E_x^{u_l}[N_x] + b_2\, E_x^{u_l}\left[\sum_{n=1}^{N_x} U_n\right] + a_3\ E_x^{u_l}\left[\sum_{n=1}^{N_x} U_n^2\right], \tag{4.9}$$

with $N_x := \inf\{n \geq 1 : X_{n+1} = x\}$ given $X_1 = x$, $b_1 := a_1 + a_2 l + a_3 l^2$ and $b_2 := a_2 + 2a_3 l$. Note that (4.9) follows from (4.5). Since N_x is a stopping time for the renewal sequence $\{U_n\}_{n\geq 1}$ (i.e., $\{N_x \leq i\} \subset \sigma(U_1, \ldots, U_i)$ for all $i \geq 1$), Walds' relation (Loève [15, p. 377]) applies to the summations $E_x^{u_l}\left[\sum_{n=1}^{N_x} U_n^j\right]$, $j = 1, 2$, which gives

$$E_x^{u_l}\left[\sum_{n=1}^{N_x} U_n^j\right] = E_x^{u_l}[N_x]\, E[U_n^j] < \infty, \quad j = 1, 2,$$

where the boundedness follows from the fact that $E_x^{u_l}[N_x] < \infty$ (cf. Ross [16, Lemma 7.4], where the validity of Condition 1 in [16, p 157] is established in Appendix A), $E[U_n] = \lambda d < \infty$, and $E[U_n^2] = \lambda^2 d^{(2)} + \lambda d < \infty$. Combining the above results with (4.9) yields $E_x^{u_l}[Z(S_x)] < \infty$, which completes the proof. ∎

We are now in position to solve the problem **P2**.

Theorem 4.1 *The threshold policy u_L given in Corollary 3.1 is optimal for problem* **P2**.

Proof. Assume that Assumptions 1-5 in Sennott [18, p. 250] hold. Then, the proof follows from Theorem 2 in [18] since we have shown in Lemma 4.1 that $\Psi(x, u_L) = \Phi(x, u_L)$ for all $x \in \mathbb{N}$.

It has been shown by Federgruen and So [6, Theorem 3.1] that Assumptions 3, 4 and 5 in Sennott [18] hold. It is shown in the Appendix that Assumption 1 in [18] also holds. Assumption 2 in [18] does not hold. However, Sennott only uses Assumption 2 to ensure that $E_x^{u_L}(\tau(X(t), a)) < \infty$ for every $x \in \mathbb{N}$, $t > 0$, $a \in \{s, v\}$ (see Sennott's comment in the proof of Theorem 2). It is shown in the Appendix that this expectation is finite in our case, which concludes the proof. ∎

Remark 4.1 As mentioned in the introduction, the proof of Theorem 3.2 in [6] is not complete. This theorem can be proved by using the proof of Theorem 4.1. If the approach used in [6] were correct then the problem of the existence of an expected average cost optimal stationary policy for a SMDP would reduce to the problem of the existence of an expected average cost optimal stationary policy for a Markov decision process. To the best of the authors' knowledge such a result has not been proved.

5 The Long-Run Average Cost Problem P1

We now address the optimization problem **P1** introduced in Section 1. We shall show that the threshold policy u_L solves **P1** over a subset $\mathcal{V}$ of the set $\mathcal{U}$ of all admissible vacation policies.

The set $\mathcal{V}$ is defined as follows: $u \in \mathcal{V}$ if $u \in \mathcal{U}$ and if there exists a state $x_0 \in \mathbb{N}$ (that may depend on u) and a subsequence $\{S_n\}_{n\geq 1}$ of $\{t_n\}_{n\geq 1}$ such that

- $X(S_n) = x_0$ for all $n \geq 1$;
- $\{S_n\}_{n\geq 1}$ is a renewal process with finite expected cycle length;
- $\{X(t),\, t \geq 0\}$ is a delayed regenerative process with respect to the renewal process $\{S_n\}_{n\geq 1}$.

Note that $\mathcal{V} \neq \emptyset$ since $u_l \in \mathcal{V}$ for every $1 \leq l < \infty$ (cf. the proof of Lemma 4.1).

The following result holds:

Theorem 5.1 *The threshold policy u_L given in Corollary 3.1 solves the problem* **P1** *over the set of policies $\mathcal{V}$.*

Proof. Fix $u \in \mathcal{V}$ and assume without loss of generality that $X(0) = x_0$ (that is $S_1 = 0$). Define $S := S_2$, $N := \inf\{n \geq 1 : X_{n+1} = x_0\}$, and observe from Ross [16, Lemma 7.4] that $E^u_{x_0}[N] < \infty$ since $E^u_{x_0}[S] < \infty$ by assumption.

Since $E^u_{x_0}[S] < \infty$, Theorem 7.5 in Ross [16] applies to the cost $\Phi(x_0, u)$ to give

$$\Phi(x_0, u) = \frac{E^u_{x_0}[Z_N]}{E^u_x[S]}, \tag{5.1}$$

where we recall that $Z_n = \sum_{i=1}^n z(X_i, A_i)$. Now, it is easy to see that $\{X_n,\, n \geq 1\}$ is a discrete regenerative process with regeneration time N. Hence, by regarding Z_N as the reward earned during the first cycle, it follows from Theorem 3.16 in [16] that

$$\lim_{n\to\infty} E^u_{x_0}\left[\frac{Z_n}{n}\right] = \frac{E^u_{x_0}[Z_N]}{E^u_{x_0}[N]}. \tag{5.2}$$

Combining (5.1) and (5.2) gives

$$\Phi(x_0, u) = \frac{E^u_{x_0}[N]}{E^u_{x_0}[S]} \lim_{n\to\infty} E^u_{x_0}\left[\frac{Z_n}{n}\right]. \tag{5.3}$$

Let us show that $V(x_0, u)$ is equal to the right-hand side of (5.3). Consider the renewal reward process $\{(S_n, Y_n),\, n \geq 1\}$ where

$$Y_n := \int_{S_n}^{S_{n+1}} X(\xi)\, d\xi + \gamma \sum_{i=M_n}^{M_{n+1}-1} \mathbf{1}(A_i = s), \quad n \geq 1,$$

is the reward earned during the n-th renewal cycle $[S_n, S_{n+1})$, and where M_n is such that $t_{M_n} = S_n$. Since, cf. (1.1),

$$\lim_{t\to\infty} E^u_{x_0}\left[\frac{1}{t}\sum_{n=1}^{M(t)-1} Y_n\right] \le V(x_0,u) \le \lim_{t\to\infty} E^u_{x_0}\left[\frac{1}{t}\sum_{n=1}^{M(t)-1} Y_n\right] + \lim_{t\to\infty} E^u_{x_0}\left[\frac{Y_{M(t)}}{t}\right], \tag{5.4}$$

with $M(t) := \sup\{n \ge 1 : S_n < t\}$ for $t > 0$, it follows again by Theorem 3.16 in Ross [16] (see also the bottom of p. 53 in [16]) that

$$\begin{aligned}
V(x_0,u) &= E^u_{x_0}\left[\int_0^S X(\xi)\,d\xi + \gamma \sum_{n=1}^N \mathbf{1}(A_n = s)\right] / E^u_{x_0}[S], \\
&= E^u_{x_0}\left[\sum_{i=1}^N \left(\int_{t_i}^{t_{i+1}} X(\xi)\,d\xi + \gamma\,\mathbf{1}(A_i = s)\right)\right] / E^u_{x_0}[S] \quad \text{by definition of } S,\ N,\ t_i, \\
&= \frac{E^u_{x_0}[N]}{E^u_{x_0}[S]} \lim_{n\to\infty} E^u_{x_0}\left[\frac{1}{n}\sum_{i=1}^n \left(\int_{t_i}^{t_{i+1}} X(\xi)\,d\xi + \gamma\,\mathbf{1}(A_i = s)\right)\right], \qquad (5.5) \\
&= \frac{E^u_{x_0}[N]}{E^u_{x_0}[S]} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n E^u_{x_0}\left[E^u_{x_0}\left[\int_{t_i}^{t_{i+1}} X(\xi)\,d\xi + \gamma\,\mathbf{1}(A_i = s)\,|\, X_i, A_i\right]\right], \\
&= \frac{E^u_{x_0}[N]}{E^u_{x_0}[S]} \lim_{n\to\infty} E^u_{x_0}\left[\frac{Z_n}{n}\right], \quad \text{from (2.3), (4.2),}
\end{aligned}$$

where the derivation of (5.5) is analogous to the derivation of (5.2) by replacing the reward Z_N by the reward $\sum_{i=1}^N \left(\int_{t_i}^{t_{i+1}} X(\xi)d\xi + \gamma\mathbf{1}(A_i = s)\right)$.

This shows that $V(x_0,u) = \Psi(x_0,u)$. That $V(x,u) = \Psi(x,u)$ for all $x \in \mathbb{N}$ is a consequence of the assumption that $\{X(t),\, t \ge 0\}$ is a delayed regenerative process (see Ross [16, Remark p. 161]). Hence, by Theorem 4.1, $V(x,u_L) \le V(x,u)$ for all $x \in \mathbb{N}$, $u \in \mathcal{V}$, which completes the proof. ■

A Appendix

(a) Proof that Assumption 1 in Sennott [18, p. 250] holds.

Assumption 1: There exists $\delta > 0$ and $\epsilon > 0$ such that $P(t_2 > \delta \,|\, X_1 = x, A_1 = a) \ge \epsilon$ for all $x \in \mathbb{N}$, $a \in \{s, v\}$.

For all $x \in \mathbb{N}$, $a \in \{s, v\}$, $\delta > 0$,

$$P(t_2 > \delta \,|\, X_1 = x, A_1 = a) \ge P(D > \delta). \tag{A.1}$$

Since $E[D] = d > 0$ by assumption, there necessarily exists $\delta > 0$ such that the right-and side of (A.1) is strictly positive, which shows that Assumption 1 holds. ■

(b) Proof that $E_x^u(\tau(X(t),a)) < \infty$ for every $x \in \mathbb{N}$, $a \in \{s,v\}$, $t > 0$.

We have for every $x \in \mathbb{N}$, $a \in \{s,v\}$, $t > 0$,

$$\begin{aligned} E_x^{u_L}(\tau(X(t),a)) &= E_x^{u_L}\left[E_x^{u_L}(\tau(X(t),a) \,|\, X(t))\right], \\ &\leq \max\left\{d, \frac{E_x^{u_L}(X(t))}{1-\rho}\right\}, \\ &\leq \max\left\{d, \frac{x+\lambda t}{1-\rho}\right\} < \infty, \end{aligned}$$

since $E_x^u(X(t))$ is maximized by the policy that never turns the server on. This concludes the proof. ■

References

[1] E. Altman and P. Nain, "Optimality of a threshold policy in the M/M/1 queue with repeated vacations," INRIA Report no. 1419, Apr. 1991. Submitted to *Operations Research*.

[2] E. Altman and P. Nain, "Optimal control of the M/G/1 queue with repeated vacations." INRIA Report (to appear).

[3] E. Cinlar, *Introduction to Stochastic Processes*, Prentice Hill, Englewood Cliffs, N. J., 1975.

[4] Y. N. Doğanata, "NT-vacation policy for M/G/1 queue with starter," *Communication, Control, and Signal Processing*, Elsevier Science Pub., 1663-1669, 1990.

[5] B. T. Doshi, "Queueing systems with vacations – A survey," *Queueing Systems*, vol. 1, no. 1, 29-66, 1986.

[6] A. Federgruen and K. C. So, "Optimality of threshold policies in single-server queueing systems with server vacations," *Adv. Appl. Prob.*, vol. 23, 388-405, 1991.

[7] E. Gelenbe and R. Iasnogorodski, "A queue with server of walking type (Autonomous service)," *Ann. Inst. Henry Poincaré*, vol. **XVI**, 63-73, 1980.

[8] E. Gelenbe and I. Mitrani, *Analysis and Synthesis of Computer Systems*, Academic Press, London, 1980.

[9] D. P. Heyman and M. J. Sobel, *Stochastic Models in Operations Research*, vol. II, McGraw-Hill, New York, 1982.

[10] O. Kella, "The threshold policy in the M/G/1 queue with server vacations," *Nav. Res. Logist.*, vol. 36, 111-123, 1989.

[11] O. Kella, "Optimal control of the vacation scheme in an M/G/1 queue," *Opns. Res.*, vol. 38, no. 4, 724-728, 1990.

[12] H. Lee and M. M. Srinivasan, "Control policies for the M^X/G/1 queueing system," *Management Sci.*, vol. 35, no. 6, 708-721, 1989.

[13] H. Levy and U. Yechiali, "Utilization of idle time in an M/G/1 queueing system," *Management Sci.*, vol. 22, 202-211, 1975.

[14] S. A. Lippman, "On dynamic programming wth unbounded rewards," *Management Sci.*, vol. 21, no. 11, 1225-1233, 1975.

[15] M. Loève, *Probability Theory I*, 4th Ed., Springer-Verlag, New York, 1977.

[16] S. M. Ross, *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco, 1970.

[17] P. J. Schweitzer, "Iterative solution of the functional equations for undiscounted Markov renewal programming," *J. Math. Anal. Appl.*, vol. 34, 495-501, 1971.

[18] L. I. Sennott, "Average cost semi-Markov decision processes and the control of queueing systems," *Prob. in the Engineering and Informational Sci.*, vol. 3, 247-272, 1989.

[19] A. J. J. Talman, "A simple proof of the optimality of the best N-policy in the M/G/1 queueing control problem with removable server," *Statistica Neerlandica*, vol. 33, 143-150, 1979.

[20] M. Yadin and P. Naor, "Queueing systems with a removable service station," *Opns. Res. Quat.*, vol. 14, 393-405, 1963.