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

Eitan ALTMAN and Philippe NAIN
INRIA
B.P. 93
06902 Sophia Antipolis Cédex, France
altman@sophia.inria.fr   nain@sophia.inria.fr

Published in: **IEEE Trans. Automatic Control 38, 12 (1993) 1766-1775**

Abstract

We consider a M/G/1 queue where the server may take repeated vacations. Whenever a busy period terminates 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; if the queue is found empty the server always takes a new vacation. The cost structure includes a holding cost per unit of time and per customer in the system and a cost each time the server is turned on. One discounted cost criterion and two average cost criteria are investigated. We show that the vacation policy that minimizes the discounted cost criterion over all policies (randomized, history dependent, etc.) converges to a threshold policy as the discount factor goes to zero. This result relies on a non standard use of the value iteration algorithm of dynamic programming and is used to prove that both average cost problems are minimized by a threshold policy

1 Introduction

We consider a M/G/1 queue where the server may take repeated vacations. Let $\lambda > 0$ be the intensity of the Poisson arrival process and let $0 < b < \infty$ and $b^{(2)} < \infty$ be the first and second moment of the service time distribution $B(\cdot)$, respectively. We shall assume throughout this paper that $\rho := \lambda b < 1$. 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) $\hat{d}(\cdot)$, first moment $0 < d < \infty$ and second moment $d^{(2)} < \infty$. The

[*]To appear in the IEEE *Transactions on Automatic Control* (Dec. 1993).
[†]Work partially supported by the France-Israel Scientific Cooperation in Computer Science and Engineering under Grant 3321190.

durations of the vacation periods are assumed to be mutually independent r.v.'s independent of the arrival and service processes. At each vacation completion time (or decision epoch) the control policy specifies whether the server should take a new vacation of random duration D (action v) or resume service (action s); if the queue is found empty the server always takes a new vacation. The service discipline is arbitrary as long as it is work conserving. The cost structure includes a holding cost of 1 per unit of time and per customer in the system and a constant cost $\gamma > 0$ each time the server is turned on.

In this paper we shall adopt the definition of a control policy used in the Semi-Markov Decision Process (SMDP) setting (see [14, 17] among others). A control policy u is a sequence of decision rules $u_1, u_2, \ldots$ where the nth decision rule u_n selects an action in $\{s, v\}$ after completion of the nth vacation. More precisely, u_n is a conditional probability on the set $\{s, v\}$ given the history $h_n = (x_1, a_1, \tau_2, \ldots, x_{n-1}, a_{n-1}, \tau_n, x_n)$ of the system up to and including the n-th decision epoch, where x_j, a_j and τ_j represent the queue-length at the jth decision epoch, the action made at the jth decision epoch and the time that elapses between the $j-1$st and the jth decision, respectively. Let $\mathcal{U}$ be the set of all control policies such that $u_n(v \,|\, h_n) = 1$ if $x_n = 0$ for all $n \geq 1$. As usual a policy in $\mathcal{U}$ is said to be a stationary policy if there exists a measurable mapping f from $\{0, 1, 2, \ldots\}$ into $\{s, v\}$ such that $u_n(a \,|\, h_n) = \mathbf{1}(f(x_n) = a)$ for all $a \in \{s, v\}$, $n \geq 1$. 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 particular, a stationary policy $u \in \mathcal{U}$ such that $u(x) = v$ for $0 \leq x < l$ and $u(x) = s$ for $x \geq l$, $l \geq 1$, is called a threshold policy with threshold l and is denoted by u_l.

Given that policy u is used and that x customers are present in the system at time 0 there exists a probability space $(\Omega, \mathcal{F}, P_x^u)$ that simultaneously carries the queue-length process $\{X(t),\, t \geq 0\}$, the action process $\{A_n,\, n \geq 1\}$ where $A_n \in \{s, v\}$ is the nth action taken, and the vacation completion time process $\{t_n, n \geq 1\}$, $0 \leq t_1 < t_2 < \cdots$ a.s., where t_n is the time when the nth vacation ends (nth decision epoch). In the following we shall assume that the first decision is made at time 0 (i.e., $t_1 = 0$). The construction of P_x^u is a standard exercise that will not be addressed in this paper (see [17]).

Let $X_n := X(t_n)$ be the queue-length at the nth decision epoch t_n, $n \geq 1$. We now introduce the three control problems that will be successively considered in this paper:

Problem $\mathbf{P0}(\alpha)$. Find $u \in \mathcal{U}$ that minimizes ($\alpha > 0$)

$$W_\alpha(x, u) := E_x^u \left[\sum_{n \geq 1} e^{-\alpha t_n}\, C(X_n, A_n) \right], \quad x = 0, 1, 2, \ldots; \tag{1.1}$$

Problem $\mathbf{P1}$. Find $u \in \mathcal{U}$ that minimizes

$$\Phi(x, u) := \overline{\lim_{t \to \infty}}\, \frac{1}{t}\, E_x^u \left[\sum_{n=1}^{N(t)} C(X_n, A_n) \right], \quad x = 0, 1, 2, \ldots; \tag{1.2}$$

Problem **P2**. Find $u \in \mathcal{U}$ that minimizes

$$V(x,u) := \overline{\lim_{t\to\infty}} \, \frac{1}{t} \, E_x^u \left[\int_0^t X(\xi)\, d\xi + \gamma \sum_{n=1}^{N(t)} \mathbf{1}(A_n = s) \right], \quad x = 0, 1, 2, \ldots \tag{1.3}$$

where E_x^u is the expectation operator associated with the probability measure P_x^u and $N(t) := \sup\{n \geq 1 : 0 \leq t_n < t\}$ is the number of decision epochs in $[0, t)$.

In (1.1)-(1.2) the quantity $C(x,a)$ is defined to be the cumulative expected cost incurred by the system between two consecutive decision epochs, say t_n and t_{n+1}, given that the queue-length is x at time t_n and that action a is chosen at time t_n. Since the arrival process is Poisson we observe that the evolution of the queue-length process in $[t_n, t_{n+1})$, given that $X_n = x$ and $A_n = a$, does not depend on n. So, letting $n = 1$ we see that

$$C(x,a) = E_x^u \left[\int_0^{t_2} X(\xi)\, d\xi \right] + \gamma\, \mathbf{1}(a = s), \quad x \in \mathbb{N}. \tag{1.4}$$

A closed-form expression for $C(x,a)$ is given in Section 2.

The cost functions (1.1) and (1.2) are the total expected α-discounted cost and the long-run expected average cost associated with the one-step cost function C, respectively. On the other hand, (1.3) gives the long-run average cost corresponding to the original system. So, in a way, (1.3) is the "natural" cost criterion associated with the cost struture introduced at the beginning of this section. However, solving directly for problem **P2** is a difficult task for this problem cannot be formulated as a SMDP control problem in contrast with problems $\mathbf{P0}(\alpha)$ and **P1**. Therefore, in a first step our approach will consist in solving successively problems $\mathbf{P0}(\alpha)$ and **P1** by using tools from SMDP theory. In a second step, we shall invoke renewal theory to show that both cost functions (1.2) and (1.3) agree over the broad class of so-called regenerative policies (see the definition in Section 4). This will finally yield the solution to problem **P2**.

We now summarize the main contributions of this paper. The first set of results (Propositions 2.2-2.3) addresses the problem $\mathbf{P0}(\alpha)$ for small values of the discount factor α. We show that there exists $L \in \mathbb{N}$, $1 \leq L \leq N_0 := \inf\{k \geq 1 : k > \gamma\,(1-\rho)/d\}$, such that for every $k \in \mathbb{N}$ there exists $\gamma_k > 0$ such that the optimal action when the queue-length is x is to turn the server off if $0 \leq x < L$ and to turn it on if $L \leq x \leq N_0 + k$ for all $\alpha \in (0, \gamma_k)$. The proof relies on a non-standard use of the value iteration algorithm of the Dynamic Programming (DP) theory [16]. Indeed, the conditions to be propagated involve the entire range of the parameter α and require that they be propagated through the entire family of DP operators. This is in contrast with usual application of the value iteration algorithm where the conditions to be propagated through the DP operator are given for a single value of α.

The second contribution (Theorem 3.1) is to show that the threshold policy u_L solves the average cost problem **P1**. The proof of this result will require a generalization of one of Sennott's conditions for the existence of an average optimal policy [18, 19]. The third contribution (Theorem 4.1) is to prove that the threshold policy u_L also minimizes the cost function (1.3) over the set of all regenerative policies.

The literature on queueing models with vacations is rapidly growing. This is because these models provide an appealing formalism for the study of various discrete event systems ranging from production systems to communication and computer systems (see [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 [7, 8, 10, 11, 13]); the scheme where the server may resume service upon the arrival of a new customer (the so-called "removable server", see [9, pp. 336-337], [21]) and a mixture of those two schemes [4].

Dynamic control and optimization issues for queueing models with server vacations have already received some attention. In [20] a control problem for the M/G/1 queue with a removable server is solved and the optimality of a threshold policy is established. For the same model, the optimal threshold policy out of all threshold policies is computed in [10]. The same analysis is carried out in [12] when batch arrivals are allowed. In [1] the optimality of a threshold policy for the problem **P2** is obtained for the M/M/1 queue with exponential repeated vacations.

While we were completing this paper a related study appeared in [6]. The model investigated in [6] is more general than the present model since batch arrivals and system dependent holding cost rates are allowed. However, the analysis developed in the forthcoming sections extends to this model provided that Conditions 1 and 2 in [6, p. 391] are satisfied. In [6] the existence of a threshold policy that minimizes a cost function similar to the cost function (1.2) is shown. The approach relies on a systematic variation of the parameter γ and is completely different from our approach.

The discussion is organized as follows: Section 2 addresses the problem $\mathbf{P0}(\alpha)$ for small values of α. In Section 3 we solve the problem **P1** while the problem **P2** is solved in Section 4. A few words on the notation used in this paper: we denote the set of nonnegative integers by $\mathbb{N}$ and the set of all real numbers by $\mathbb{R}$

2 The Discounted Cost Problem

This section addresses the control problem $\mathbf{P0}(\alpha)$ introduced in Section 1. For every fixed $\alpha > 0$, a policy $u \in \mathcal{U}$ is said to be α-discounted optimal if $W_\alpha(x) = W_\alpha(x, u)$ for all $x \in \mathbb{N}$ where

$$W_\alpha(x) := \inf_{v \in \mathcal{U}} W_\alpha(x, v), \quad n \in \mathbb{N}. \tag{2.1}$$

Our analysis begins with the observation that the queue-length process embedded at the decision epochs $\{X_n,\, n \geq 1\}$ is a controlled Markov chain [16] with state space $\mathbb{N}$ and transition probabilities

$P_{xy}(a)$ given by

$$P_{xy}(a) := \begin{cases} P_V(y) & \text{if } a = s \text{ and } x \geq 1 \\ P_V(y-x) & \text{if } a = v \text{ and } x \leq y \\ 0 & \text{if } a = v \text{ and } x > y \end{cases} \tag{2.2}$$

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

The LST of the sojourn times of the process $\{X_n,\, n \geq 1\}$ in state x, given that action a has been chosen, is given by ($\alpha > 0$)

$$\hat{\tau}_{x,a}(\alpha) := \begin{cases} \hat{d}(\alpha) & \text{if } a = v \text{ and } x \in \mathbb{N} \\ \hat{d}(\alpha)\,\left[\hat{T}(\alpha)\right]^x & \text{if } a = s \text{ and } x \geq 1 \end{cases} \tag{2.3}$$

where $\hat{T}(\alpha)$ stands for the LST of a busy period duration in an M/G/1 queue with arrival intensity λ and service time distribution $B(\cdot)$ (observe that $P_{0y}(s)$ and $\hat{\tau}_{0,s}(\alpha)$ need not to be defined since the action s is not permitted when the system is found empty at a decision epoch).

Therefore (see [14, 16]) the cost functions $W_\alpha(x,u)$ and $\Phi(x,u)$ are seen to be the total expected α-discounted cost and the long-run expected average cost, respectively, associated with the SMDP $\{X(t),\, t \geq 0\}$ with state space $\mathbb{N}$, action spaces $\{A_x,\, x \in \mathbb{N}\}$ with $A_x := \{s, v\}$ if $x \neq 0$ and $A_x := \{v\}$ if $x = 0$, one-step transition probabilities $P_{xy}(a)$, LST of the conditional sojourn times $\hat{\tau}_{x,a}$ and one-step cost function C defined in (1.4), respectively.

We have shown in [2, Appendix A] that for $x \in \mathbb{N}$

$$C(x,a) = \frac{\lambda\, d^{(2)}}{2} + xd\,\mathbf{1}(a = v) + \left(\frac{b}{2(1-\rho)}x^2 + \frac{\lambda b^{(2)} + b(1-\rho)}{2(1-\rho)^2}x + \gamma\right)\mathbf{1}(a = s, x \neq 0). \tag{2.4}$$

We observe from (2.4) that for each $a \in \{s, v\}$ the cost $C(x,a)$ is not bounded on $\mathbb{N}$. One way of handling unbounded costs is to define a new norm for which the costs are bounded [14]. Adapting the results in [14] to our problem, we define $\mathcal{K}$ to 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

$$T_\alpha f(x) := \min_{a \in A_x} \left\{ C(x,a) + \hat{\tau}_{x,a}(\alpha) \sum_{y \in \mathbb{N}} P_{xy}(a)\, f(y) \right\}, \quad x \in \mathbb{N}. \tag{2.5}$$

The proof that $T_\alpha(f)$ belongs to the set $\mathcal{K}$ can be found in [14, p. 1228] (the proof that Assumptions 1-3 in [14, p. 1227] hold is straightforward by using the explicit form of C given in (2.4); in particular, the constant b that appears in Assumption 3 in [14] may be chosen to be $1 + \lambda d + \lambda^2\, d^{(2)}$).

We have the following results [14, Theorem 1]:

Proposition 2.1 *Fix $\alpha > 0$. For every stationary policy $u \in \mathcal{U}$, $W_\alpha(x,u)$ satisfies the equation*

$$W_\alpha(x,u) = C(x,u(x)) + \hat{\tau}_{x,u(x)}(\alpha) \sum_{y \in \mathbb{N}} P_{xy}(u(x))\, W_\alpha(y,u), \quad x \in \mathbb{N}. \tag{2.6}$$

On the other hand, $W_\alpha(x)$ is (the unique) solution in $\mathcal{K}$ of the DP equation

$$W_\alpha(x) = T_\alpha W_\alpha(x), \quad x \in \mathbb{N} \tag{2.7}$$

and can be obtained as

$$\lim_{n \to \infty} T_\alpha^n f(x) = W_\alpha(x), \quad x \in \mathbb{N} \tag{2.8}$$

for every function $f \in \mathcal{K}$, where $T_\alpha^{n+1} f := T_\alpha[T_\alpha^n f]$ for $n \geq 1$.

Moreover, any stationary policy that minimizes the right-hand side of (2.7) is α-discounted optimal.

Specializing the DP equation (2.7) to our model we see from (2.2), (2.3) and (2.5) that

$$W_\alpha(x) = \min\Big\{C(x,v) + \hat{d}(\alpha) \sum_{y \in \mathbb{N}} P_V(y) W_\alpha(x+y), C(x,s) + \hat{d}(\alpha) \left[\hat{T}(\alpha)\right]^x \sum_{y \in \mathbb{N}} P_V(y) W_\alpha(y)\Big\} \tag{2.9}$$

for every $x \in \mathbb{N}$. For later use it is convenient to rewrite (2.9) in the form

$$W_\alpha(x) = C(x,s) + \hat{d}(\alpha) \left[\hat{T}(\alpha)\right]^x \sum_{y \in \mathbb{N}} P_V(y) W_\alpha(y) + \min\{\Delta_\alpha(x), 0\}, \quad x \in \mathbb{N} \tag{2.10}$$

where

$$\Delta_\alpha(x) := C(x,v) - C(x,s) + \hat{d}(\alpha) \sum_{y \in \mathbb{N}} P_V(y) \left(W_\alpha(x+y) - \left[\hat{T}(\alpha)\right]^x W_\alpha(y)\right). \tag{2.11}$$

By Proposition 2.1 the optimal action in state x is to take another vacation if $\Delta_\alpha(x) < 0$ and to resume service if $\Delta_\alpha(x) > 0$, while both actions are optimal when $\Delta_\alpha(x) = 0$. Denote by u_α^* the unique stationary policy that minimizes the right-hand side of (2.1) and that chooses action v in state x if $\Delta_\alpha(x) = 0$ for $x \geq 1$, $\alpha > 0$. The aim of this section is to characterize u_α^* for small values of the discount factor.

We now introduce some properties to be enjoyed by the set of optimal value functions $\{W_\alpha,\ \alpha > 0\}$ (see Lemma 2.1). We shall say that a set of functions $\{f_\alpha, \alpha > 0\}$ in $\mathcal{K}$ satisfies condition

C1 if for every $x \in \mathbb{N}$, $k \geq 1$, there exists $\alpha_{x,k} > 0$ such that

$$f_\alpha(x+r+k) - \left[\hat{T}(\alpha)\right]^k f_\alpha(x+r) \geq C(x+k,s) - C(x,s) \tag{2.12}$$

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

C2 if for all $\alpha > 0$, $x \in \mathbb{N}$, $k \geq 1$,

$$f_\alpha(x+k) - \left[\hat{T}(\alpha)\right]^k f_\alpha(x) \geq 0. \tag{2.13}$$

We observe from (2.4) that the set of functions $\{f_\alpha,\ \alpha > 0\}$ with $f_\alpha(\cdot) := C(\cdot, s)$ satisfies **C1** and **C2**.

Lemma 2.1 *The set of optimal value functions* $\{W_\alpha,\ \alpha > 0\}$ *satisfies conditions* **C1** *and* **C2**.

Proof. Unless otherwise mentioned $\alpha > 0$, $x \in \mathbb{N}$ and $k \geq 1$ are fixed numbers.

Let $\{f_\alpha,\ \alpha > 0\}$ be an arbitrary set of functions in $\mathcal{K}$ that satisfies **C1** and **C2**. We show that $\{T_\alpha f_\alpha,\ \alpha > 0\}$ also satisfies these conditions.

We have from (2.2), (2.3) and (2.5) that for every $r \in \mathbb{N}$

$$\begin{aligned}
& T_\alpha f_\alpha(x+r+k) - \left[\hat{T}(\alpha)\right]^k T_\alpha f_\alpha(x+r) \\
= \ & \min\Bigg\{ C(x+r+k, v) + \hat{d}(\alpha) \sum_{y \in \mathbb{N}} P_V(y) f_\alpha(x+r+k+y)\ , \\
& C(x+r+k, s) + \hat{d}(\alpha) \left[\hat{T}(\alpha)\right]^{x+r+k} \sum_{y \in \mathbb{N}} P_V(y) f_\alpha(y) \Bigg\} \\
& - \left[\hat{T}(\alpha)\right]^k \min\Bigg\{ C(x+r, v) + \hat{d}(\alpha) \sum_{y \in \mathbb{N}} P_V(y) f_\alpha(x+r+y)\ , \\
& C(x+r, s) + \hat{d}(\alpha) \left[\hat{T}(\alpha)\right]^{x+r} \sum_{y \in \mathbb{N}} P_V(y) f_\alpha(y) \Bigg\} \\
\geq \ & \min\Bigg\{ C(x+r+k, v) - \left[\hat{T}(\alpha)\right]^k C(x+r, v) \\
& + \hat{d}(\alpha) \sum_{y \in \mathbb{N}} P_V(y) \left(f_\alpha(x+r+k+y) - \left[\hat{T}(\alpha)\right]^k f_\alpha(x+r+y) \right), \\
& C(x+r+k, s) - \left[\hat{T}(\alpha)\right]^k C(x+r, s) \Bigg\} \qquad (2.14) \\
\geq \ & \min\Bigg\{ kd + \hat{d}(\alpha) \sum_{y \in \mathbb{N}} P_V(y) \left(f_\alpha(x+r+k+y) - \left[\hat{T}(\alpha)\right]^k f_\alpha(x+r+y) \right), \\
& C(x+k, s) - C(x, s) \Bigg\}. \qquad (2.15)
\end{aligned}$$

The inequality (2.14) follows from the inequality $\min(a,b) - \min(c,d) \geq \min(a-c, b-d)$. The first argument in the min in (2.15) follows from

$$C(x+r+k,v) - \left[\hat{T}(\alpha)\right]^k C(x+r,v) \geq C(x+r+k,v) - C(x+r,v) = kd \qquad (2.16)$$

where the last equality is obtained from (2.4). Finally, the second argument in the min is a consequence of the fact that $C(\cdot,s)$ satisfies **C1**.

Letting $r=0$ in (2.15) it is seen that $\{T_\alpha f_\alpha,\ \alpha>0\}$ satisfies **C2** since $x \to C(x,s)$ is nondecreasing by (2.4) and since $\{f_\alpha,\ \alpha>0\}$ satisfies **C2** (which implies that the first argument in the min is nonnegative).

Hence, we may deduce by induction that $\{T_\alpha^n f_\alpha,\ \alpha>0\}$ satisfies **C2** for every $n \geq 1$. Consequently, we have from (2.8)

$$W_\alpha(x+k) - [\hat{T}(\alpha)]^k W_\alpha(x) = \lim_{n\to\infty} \left[T_\alpha^n f_\alpha(x+k) - [\hat{T}(\alpha)]^k T_\alpha^n f_\alpha(x)\right] \geq 0 \qquad (2.17)$$

which shows that $\{W_\alpha,\ \alpha>0\}$ satisfies **C2**.

It remains to show that the set $\{W_\alpha,\ \alpha>0\}$ satisfies **C1**. Since the mapping $x \to C(x,s)$ is nondecreasing, $\hat{d}(\alpha) \to 1$ as $\alpha \to 0$ and since $kd>0$, there exists a real number $\beta_{x,k}>0$ such that for $\alpha \in (0, \beta_{x,k})$

$$kd + \hat{d}(\alpha)\left(C(x+k,s) - C(x,s)\right) \geq C(x+k,s) - C(x,s). \qquad (2.18)$$

On the other hand, since the set of functions $\{f_\alpha,\ \alpha>0\}$ satisfies **C1** by assumption, there exists $\alpha_{x,k}>0$ such that (2.12) holds for $\alpha \in (0, \alpha_{x,k})$. Combining (2.12) and (2.18) we deduce that the right-hand side of inequality (2.15) is equal to $C(x+k,s) - C(x,k)$ for $\alpha \in (0, \gamma_{x,k})$ where $\gamma_{x,k} := \min\{\alpha_{x,k}, \beta_{x,k}\}$. This in turn shows that $\{T_\alpha f_\alpha,\ \alpha>0\}$ satisfies **C1**.

Iterating this procedure we obtain that for every $n \geq 1$,

$$T_\alpha^n f_\alpha(x+r+k) - \left[\hat{T}(\alpha)\right]^k T_\alpha^n f_\alpha(x+r) \geq C(x+k,s) - C(x,s)$$

for $\alpha \in (0, \gamma_{x,k})$, $r \in \mathbb{N}$ (note that it is crucial here that $\gamma_{x,k}$ does not depend on n) and the proof is concluded as in (2.17). ■

Define

$$N_0 := \inf_{k\in\mathbb{N}} \left\{ k > \gamma \left(\frac{d}{1-\rho}\right)^{-1} \right\}. \qquad (2.19)$$

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 at the end of a vacation.

We now address the minimization of the right-hand side of the DP equation (2.10) for $x \geq N_0$ and for small values of α.

Proposition 2.2 *There exists a nonincreasing sequence $\{\gamma_k, k \in \mathbb{N}\}$ in $(0, \infty)$ such that for each $k \in \mathbb{N}$, whenever $\alpha \in (0, \gamma_k)$, $u^*_\alpha(x) = s$ for $x = N_0, N_0+1, \ldots, N_0+k$.*

Proof. Fix $x \geq N_0$. Owing to (2.10) it suffices to prove that $\Delta_\alpha(x) > 0$ for α small enough.

Since $W_\alpha(x+y) - [\hat{T}(\alpha)]^x\, W_\alpha(y)$ is nonnegative for all $\alpha > 0$, $y \in \mathbb{N}$ because $\{W_\alpha,\ \alpha > 0\}$ satisfies **C2** by Lemma 2.1, Fatou's lemma applies to (2.11) to yield

$$\begin{aligned}
\varliminf_{\alpha \to 0} \Delta_\alpha(x) &\geq C(x,v) - C(x,s) + \sum_{y \in \mathbb{N}} P_V(y) \varliminf_{\alpha \to 0} \Big(W_\alpha(x+y) - [\hat{T}(\alpha)]^x\, W_\alpha(y) \Big) \\
&\geq C(x,v) - C(x,s) + \sum_{y \in \mathbb{N}} P_V(y) \left(C(x+y,s) - C(y,s) \right) \qquad (2.20)
\end{aligned}$$

where the last inequality holds because $\{W_\alpha,\ \alpha > 0\}$ satisfies **C1** by Lemma 2.1. Using now the expression (2.4) of the cost C together with the identity $\sum_{y \in \mathbb{N}} y\, P_V(y) = \lambda d$ (see the definition of P_V in Section 2) we readily deduce that the right-hand side of (2.20) is equal to $xd/(1-\rho) - \gamma$ which is a strictly positive quantity for all $x \geq N_0$. This concludes the proof. ■

The next result addresses the minimization of the right-hand side of (2.10) for $x \leq N_0$ and for small values of α.

Proposition 2.3 *There exists an integer L such that $1 \leq L \leq N_0$ with the property that for α small enough*

$$u^*_\alpha(x) = \begin{cases} v & \text{for } x = 0, 1, \ldots, L-1 \\ s & \text{for } x = L, L+1, \ldots, N_0. \end{cases} \qquad (2.21)$$

The proof of Proposition 2.3 is given in Appendix A.

Combining Propositions 2.2 and 2.3 we obtain the following

Corollary 2.1 *The sequence of α-discounted optimal policies $\{u^*_\alpha,\ \alpha > 0\}$ converges to a threshold policy u_L as α goes to zero with $1 \leq L \leq N_0$.*

3 The Problem P1

We show in this section that the threshold policy u_L identified in Section 2 solves the problem **P1**. The proof of this result will rely on Theorem 2 in [18].

To apply this theorem (see [18, p. 249]) we need to introduce the long-run expected average cost

$$\Psi(x,u) := \overline{\lim_{n\to\infty}} \frac{E_x^u \left[\sum_{i=1}^n C(X_i,A_i)\right]}{E_x^u\left[t_n\right]}, \quad x \in \mathbb{N}, u \in \mathcal{U}. \tag{3.1}$$

The following result is also needed:

Lemma 3.1 *For every $l \geq 1$ we have*

$$\Psi\left(x,u_l\right) = \Phi\left(x,u_l\right), \quad x \in \mathbb{N} \tag{3.2}$$

where the cost functions $\Phi(x,u)$ and $\Psi(x,u)$ are defined in (1.2) and (3.1), respectively. Moreover, $\Phi\left(x,u_l\right)$ does not depend on x and is finite.

Proof. Let $x \in \mathbb{N}$, $l \geq 1$ be fixed integers. Throughout the proof we shall assume that the threshold policy u_l is used. Under policy u_l the queue-length process $\{X(t),\, t \geq 0\}$ is a delayed regenerative process with regeneration point $S_x := \inf\{n \geq 2 : X_n = x\}$. Hence, we know by Theorem 7.5 in [16] that (3.2) will hold if $E_x^u\left[S_x\right] < \infty$.

We have shown in [2] that $E_x^{u_l}[S_x] = \sum_{y\in\mathbb{N}} E_y^{u_l}[t_2]\,\nu(y)/\nu(x)$ where $\nu(\cdot)$ is the invariant measure of the irreducible, aperiodic, non-null recurrent Markov chain $\{X_n,\, n \geq 1\}$. Let U_n denote the number of arrivals during the n-th vacation period. Because

$$X_{n+1} = U_n + X_n\,\mathbf{1}(X_n < l) \leq U_n + l \tag{3.3}$$

for all $n \geq 1$, it is readily seen that $\sum_{y\in\mathbb{N}} E_y^{u_l}[t_2]\,\nu(y) \leq d + (b/(1-\rho))(l+\lambda d)$, which proves that $E_x^u\left[S_x\right] < \infty$. As a corollary of this result we know from [16, Remark, p. 161] that $\Phi(x,u_l)$ does not depend on x.

It remains to show that $\Phi_l := \Phi(\cdot,u_l) < \infty$. Until the end of the proof we shall assume that $x \geq l$. By applying 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 must prove that $E_x^{u_l}[Z(S_x)]$ is finite.

It is easily seen from (2.4) that there exist three constants a_1, a_2, a_3 such that $C(y,a) \leq a_1 + a_2 y + a_3 y^2$ for every $y \in \mathbb{N}$, $a \in \{s,v\}$. Hence, with $N_x := \inf\{n \geq 1 \,:\, X_{n+1} = x\}$, we have

$$\begin{aligned}
E_x^{u_l}[Z(S_x)] &= E_x^{u_l}\left[\sum_{n=1}^{N_x} C(X_n,A_n)\right] && (3.4)\\
&= C(x,s) + E_x^{u_l}\left[\sum_{n=1}^{N_x-1} C(X_{n+1},A_{n+1})\right] \\
&\leq C(x,s) + a_1\,E_x^{u_l}[N_x] + a_2\,E_x^{u_l}\left[\sum_{n=1}^{N_x} X_{n+1}\right] + a_3\,E_x^{u_l}\left[\sum_{n=1}^{N_x} X_{n+1}^2\right] \\
&\leq C(x,s) + b_1\,E_x^{u_l}[N_x] + b_2\,E_x^{u_l}\left[\sum_{n=1}^{N_x} U_n\right] + a_3\;E_x^{u_l}\left[\sum_{n=1}^{N_x} U_n^2\right] && (3.5)
\end{aligned}$$

with $b_1 := a_1 + a_2 l + a_3 l^2$, $b_2 := a_2 + 2a_3 l$, in view of (3.3). Since $x \geq l$ by assumption, we observe that under policy u_l the r.v. N_x is a stopping time for the renewal sequence $\{U_n,\, n \geq 1\}$. Therefore, Walds' formula [15, p. 377] applies to yield $E_x^{u_l}\left[\sum_{n=1}^{N_x} U_n^j\right] = E_x^{u_l}[N_x]\, E[U_n^j] < \infty$ for $j = 1, 2$, where the boundedness follows from the fact that $E_x^{u_l}[N_x] < \infty$ (cf. [16, Lemma 7.4], where the validity of Condition 1 in [16, p 157] is established in Appendix B), $E[U_n] = \lambda d < \infty$, and $E[U_n^2] = \lambda^2 d^{(2)} + \lambda d < \infty$. Combining these results with (3.5) yields $E_x^{u_l}[Z(S_x)] < \infty$ and the proof is complete. ■

We are now in position to solve problem **P1**.

Theorem 3.1 *There exists an integer L, $1 \leq L \leq N_0$, such that $\Phi(x, u_L) \leq \Phi(x, u)$ for every $u \in \mathcal{U}$, $x \in \mathbb{N}$.*

Proof. If Assumptions 1-5 in [18, p. 250] were to hold then the proof would follow from Theorem 2 in [18] since we have shown in Lemma 3.1 that $\Psi(x, u_L) = \Phi(x, u_L)$ for all $x \in \mathbb{N}$.

We show in Appendix B that Assumptions 1, 3 and 4 in [18] hold for our model. However, Assumption 2 does not hold and it is quite some work to prove that Assumption 5 holds. We shall instead show that Assumption 5 can be replaced by a weaker assumption which turns out to be satisfied by our model.

Assumption 2 is only used in [18] to ensure that $E_x^{u_L}(\tau(X_n, u_L(X_n))) < \infty$ for every $n \geq 1$, $x \in \mathbb{N}$, where $\tau(x, a)$ is the sojourn time in state x given action a is chosen (see Sennott's comment in the proof of Theorem 2). This result is true in our case. Indeed, we have for every $n \geq 1$, $x \in \mathbb{N}$, $a \in A_x$,

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

since $E_x^{u}(X_n)$ is maximized when using the policy that never turns the server on.

We shall replace Assumption 5 in [18] by the following assumption (we use Sennott's notation):

Assumption 5*: There exists a nonnegative number M, such that for every $i \in \mathbb{N}$ there exists $\beta_i > 0$ such that $-M \leq h_\alpha(i) := W_\alpha(i) - W_\alpha(N_0)$ for every i and $0 < \alpha < \beta_i$.

We show in Appendix B that Assumption 5* holds. Note that this assumption reduces to Assumption 5 in [18] if β_i does not depend on i.

Assumption 5 is used by Sennott (see [18], [19, pp. 632-633]) to prove the existence of a sequence

$\{\alpha_n\}$ in $(0,\infty)$ converging to 0 as n goes to ∞ such that for all $i \in \mathbb{N}$, $\lim_{n\to\infty} h_{\alpha_n}(i) \geq -M$. We now show that this result is still valid under Assumption 5*.

We know from Assumption 4 in [18] and Assumption 5* above that for every $i \in \mathbb{N}$, $h_\alpha(i) \in [-M, M_i]$ for $\alpha \in (0, \min(\alpha_0, \beta_i))$. Therefore, for any sequence $\{\alpha_n\}$ in $(0,\infty)$ converging to 0, there exists a subsequence $\{\alpha_n(0)\}$ of the sequence $\{\alpha_n\}$ such that $\lim_{n\to\infty} h_{\alpha_n(0)}(0)$ exists and is greater than or equal to $-M$. Similarly, there exists a subsequence of $\{\alpha_n(0)\}$ (call it $\{\alpha_n(1)\}$) such that $\lim_{n\to\infty} h_{\alpha_n(1)}(1)$ exists and is greater than or equal to $-M$. Iterating this procedure, we see that for every $i \in \mathbb{N}$ there exists a sequence $\{\alpha_n(i)\} \in (0,\infty)$ converging to 0 with the property that $\{\alpha_n(i)\} \subset \{\alpha_n(i-1)\} \subset \cdots \subset \{\alpha_n(0)\}$ and such that $\lim_{n\to\infty} h_{\alpha_n(i)}(i)$ exists and is greater than or equal to $-M$. Hence, $\lim_{n\to\infty} h_{\alpha_n(n)}(i) \geq -M$ for every $i \in \mathbb{N}$, and Theorem 2 in [18] applies, which concludes the proof. ■

4 The Problem P2

We now address the problem **P2** introduced in Section 1. We shall show that the threshold policy u_L solves **P2** over a subset $\mathcal{V}$ of the set $\mathcal{U}$ of all vacation policies.

The set $\mathcal{V}$ is defined as follows: $u \in \mathcal{V}$ if $u \in \mathcal{U}$ and if there exists a distinguished state $x_0 \in \mathbb{N}$ and a sequence $\{t'_n\} \subset \{t_n\}$ such that (1) $X(t'_n) = x_0$ for all $n \geq 1$, (2) $\{t'_n\}$ is a renewal process with finite expected cycle length and (3) $\{X(t), t \geq 0\}$ is a delayed regenerative process with respect to the renewal process $\{t'_n\}$. Note that $\mathcal{V} \neq \emptyset$ since $u_l \in \mathcal{V}$ for every $l \geq 1$ as shown in the proof of Lemma 3.1.

The following result holds:

Theorem 4.1 *There exists an integer L, $1 \leq L \leq N_0$, such that $V(x, u_L) \leq V(x, u)$ for every $u \in \mathcal{V}$, $x \in \mathbb{N}$.*

Proof. Fix $u \in \mathcal{V}$. We shall first assume that $X(0) = x_0$ (i.e., $t'_1 = 0$) where x_0 is the distinguished state associated with the policy u. Let S be a generic r.v. distributed as the r.v.'s $t'_{n+1} - t'_n$ and define $N := \inf\{n \geq 1 : X_{n+1} = x_0\}$. We have by Lemma 7.4 in [16] that $E^u_{x_0}[N] < \infty$ since $E^u_{x_0}[S] < \infty$ by definition of the set $\mathcal{V}$.

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

$$\Phi(x_0, u) = \frac{E^u_{x_0} Z_N}{E^u_x [S]}, \tag{4.1}$$

with $Z_n := \sum_{i=1}^n C(X_i, A_i)$, $n \geq 1$. Now, it is easy to see that $\{X_n\}$ 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

$$\lim_{n\to\infty} E^u_{x_0}\left[\frac{Z_n}{n}\right] = \frac{E^u_{x_0}[Z_N]}{E^u_{x_0}[N]}. \tag{4.2}$$

Combining (4.1) and (4.2) gives

$$\Phi(x_0,u) = \frac{E^u_{x_0}[N]}{E^u_{x_0}[S]} \lim_{n\to\infty} E^u_{x_0}\left[\frac{Z_n}{n}\right]. \tag{4.3}$$

Let us show that $V(x_0,u)$ (cf. (1.3)) is equal to the right-hand side of (4.3). Consider the renewal reward process $\{(ft'_n, Y_n)\}$ where

$$Y_n := \int_{t'_n}^{t'_{n+1}} X(\xi)\,d\xi + \gamma \sum_{i=M_n}^{M_{n+1}-1} \mathbf{1}(A_i = s), \quad n \geq 1$$

is the reward earned during the n-th renewal cycle $[t'_n, t'_{n+1})$, and where M_n is such that $t_{M_n} = t'_n$. With $M(t) := \sup\{n \geq 1 \,:\, t'_n < t\}$ for $t > 0$ we have, cf. (1.3),

$$\lim_{t\to\infty} E^u_{x_0}\left[\frac{1}{t}\sum_{n=1}^{M(t)-1} Y_n\right] \leq V(x_0,u) \leq \lim_{t\to\infty} E^u_{x_0}\left[\frac{1}{t}\sum_{n=1}^{M(t)-1} Y_n\right] + \lim_{t\to\infty} E^u_{x_0}\left[\frac{Y_{M(t)}}{t}\right] \tag{4.4}$$

and it follows again by Theorem 3.16 in [16] (see also the bottom of p. 53 in [16]) that

$$\begin{aligned}
V(x_0,u) &= E^u_{x_0}\left[\int_0^S X(\xi)\,d\xi + \gamma\sum_{n=1}^N \mathbf{1}(A_n = s)\right] / E^u_{x_0}[S] \\
&= E^u_{x_0}\left[\sum_{i=1}^N \left(\int_{t_i}^{t_{i+1}} X(\xi)\,d\xi + \gamma\,\mathbf{1}(A_i = s)\right)\right] / E^u_{x_0}[S] \quad \text{by definition of } S,\ N,\ t_i \\
&= \frac{E^u_{x_0}[N]}{E^u_{x_0}[S]} \lim_{n\to\infty} E^u_{x_0}\left[\frac{1}{n}\sum_{i=1}^n \left(\int_{t_i}^{t_{i+1}} X(\xi)\,d\xi + \gamma\,\mathbf{1}(A_i = s)\right)\right] \qquad (4.5) \\
&= \frac{E^u_{x_0}[N]}{E^u_{x_0}[S]} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n E^u_{x_0}\left[E^u_{x_0}\left[\int_{t_i}^{t_{i+1}} X(\xi)\,d\xi + \gamma\,\mathbf{1}(A_i = s) \,|\, X_i, A_i\right]\right] \\
&= \frac{E^u_{x_0}[N]}{E^u_{x_0}[S]} \lim_{n\to\infty} E^u_{x_0}\left[\frac{Z_n}{n}\right] \\
&= \Phi(x_0,u) \qquad (4.6)
\end{aligned}$$

where the derivation of (4.5) is analogous to the derivation of (4.2) by replacing the reward Z_N by the reward $\sum_{i=1}^N \left(\int_{t_i}^{t_{i+1}} X(\xi)d\xi + \gamma\mathbf{1}(A_i = s)\right)$ and where (4.6) follows from (4.3).

Since $\{X(t),\, t \geq 0\}$ is a delayed regenerative process it follows from (4.6) (see [16, Remark p. 161]) that $V(x,u) = \Psi(x,u)$ for all $x \in \mathbb{N}$. Therefore, by Theorem 3.1, $V(x,u_L) = \Phi(x,u_L) \leq \Phi(x,u) =$

$V(x,u)$ for all $x \in \mathbb{N}$ and $u \in \mathcal{V}$. ■

Acknowledgments. We are very grateful to the referees for their thorough work and their constructive comments.

A Appendix

Proof of Proposition 2.3. It suffices to prove that for α small enough

$$\Delta_\alpha(x) \le \Delta_\alpha(x+1) \quad \text{for } x = 0, 1, \ldots, N_0 - 1. \tag{A.1}$$

Indeed, since $\Delta_\alpha(0) = 0$ for all $\alpha > 0$ (cf. (2.11)) and since $\Delta_\alpha(N_0) > 0$ for α small enough (cf. the proof of Proposition 2.2) we may then conclude from (A.1) that for α small enough, say $\alpha \in (0, \beta)$, there exists $L_\alpha \in \mathbb{N}$, $1 \le L_\alpha \le N_0$, such that $\Delta_\alpha(x) \le 0$ for $x = 0, 1, \ldots, L_\alpha - 1$ and $\Delta_\alpha(x) > 0$ for $x = L_\alpha, L_\alpha + 1, \ldots, N_0$. Let $\{\alpha_n\}$ be a sequence in $(0, \beta)$ that converges to 0 as $n \to \infty$. Because L_α lies in the compact set $[1, N_0]$ there exists a subsequence of $\{\alpha_n\}$ (also denoted by $\{\alpha_n\}$) such that $\lim_{n\to\infty} L_{\alpha_n} = L$ with $1 \le L \le N_0$. Because L is an integer we see that $L_\alpha = L$ for α small enough. Remembering now our convention that $u^*_\alpha(x) = v$ if $\Delta_\alpha(x) = 0$ we immediately deduce from the above that (2.21) holds for α small enough.

We now turn to the proof of (A.1)). We have for $x \in \mathbb{N}$, cf. (2.4), (2.11),

$$\begin{aligned} \Delta_\alpha(x+1) - \hat{T}(\alpha)\Delta_\alpha(x) &= \left(1 - \hat{T}(\alpha)\right)\left(dx + \frac{\lambda d^{(2)}}{2}\right) + d - \left(C(x+1,s) - \hat{T}(\alpha)\, C(x,s)\right) \\ &\quad + \hat{d}(\alpha) \sum_{y \in \mathbb{N}} P_V(y) \left(W_\alpha(x+1+y) - \hat{T}(\alpha) W_\alpha(x+y)\right). \end{aligned} \tag{A.2}$$

Letting α go to 0 in (A.2), we have again by Fatou's lemma that

$$\begin{aligned} \underline{\lim}_{\alpha\to 0} \left\{\Delta_\alpha(x+1) - \hat{T}(\alpha)\Delta_\alpha(x)\right\} &\ge d - (C(x+1,s) - C(x,s)) \\ &\quad + \sum_{y\in\mathbb{N}} P_V(y)\, \underline{\lim}_{\alpha\to 0} \left(W_\alpha(x+1+y) - \hat{T}(\alpha) W_\alpha(x+y)\right) \\ &= d \end{aligned} \tag{A.3}$$

where (A.3) follows from the fact that $\{W_\alpha,\ \alpha > 0\}$ satisfies condition **C1** by Proposition 2.1 We deduce from (A.3) that for α small enough

$$\Delta_\alpha(x+1) - \hat{T}(\alpha)\Delta_\alpha(x) \ge d/2, \quad x \in \mathbb{N}. \tag{A.4}$$

Sifrom (A.4) that (A.1) will hold for α small enough if we show that for α small enough

$$\left(1 - \hat{T}(\alpha)\right) \Delta_\alpha(x) \le d/2, \quad x = 1, 2, \ldots, N_0 - 1. \tag{A.5}$$

This is done in the remainder of the proof.

It is seen by repeatedly applying (A.4) that for α small enough

$$\Delta_\alpha(x) \leq \frac{\Delta_\alpha(N_0)}{\left[\hat{T}(\alpha)\right]^{N_0-x}} \leq \frac{\Delta_\alpha(N_0)}{\left[\hat{T}(\alpha)\right]^{N_0}}, \quad x = 1, 2, \ldots, N_0 - 1. \tag{A.6}$$

Therefore, for α small enough,

$$\left(1 - \hat{T}(\alpha)\right) \Delta_\alpha(x) \leq \left(\frac{1 - \hat{T}(\alpha)}{[\hat{T}(\alpha)]^{N_0}}\right) \Delta_\alpha(N_0), \quad x = 1, 2, \ldots, N_0 - 1. \tag{A.7}$$

Since the coefficient of $\Delta_\alpha(N_0)$ in the right-hand side of (A.7) approaches 0 when α goes to 0, we see from (A.7) that (A.5) will hold for α small enough if $\Delta_\alpha(N_0)$ is bounded from above by a number that does not depend on α. Let us prove that $\Delta_\alpha(N_0)$ is bounded from above by a number that does not depend on α.

Because $\{W_\alpha,\, \alpha > 0\}$ satisfies **C2** by Lemma 2.1 we have for all $\alpha > 0$, cf. (2.11),

$$\begin{aligned}
\Delta_\alpha(N_0) &\leq C(N_0, v) + \sum_{y \in \mathbb{N}} P_V(y) \left(W_\alpha(N_0 + y) - \left[\hat{T}(\alpha)\right]^{N_0} W_\alpha(y)\right), \\
&= C(N_0, v) + \sum_{y \in \mathbb{N}} P_V(y) \left(W_\alpha(N_0 + y) - \left[\hat{T}(\alpha)\right]^{y} W_\alpha(N_0)\right) \\
&\quad + \sum_{y \in \mathbb{N}} P_V(y) \left[\hat{T}(\alpha)\right]^{y} \left(W_\alpha(N_0) - \left[\hat{T}(\alpha)\right]^{N_0 - y} W_\alpha(y)\right).
\end{aligned} \tag{A.8}$$

We first compute an upper bound for the first summation in the right-hand side of (A.8). Let $v_\alpha \in \mathcal{U}$ be the policy that serves at the first decision epoch whenever the queue is nonempty and then follows the α-optimal policy.

Hence, by definition of $W_\alpha(\cdot)$ (cf. (2.1)) we find for $\alpha > 0$

$$\begin{aligned}
&\sum_{y \in \mathbb{N}} P_V(y) \left(W_\alpha(N_0 + y) - [\hat{T}(\alpha)]^y\, W_\alpha(N_0)\right) \\
&\quad \leq \sum_{y \in \mathbb{N}} P_V(y) \left(W_\alpha(N_0 + y, v_\alpha) - [\hat{T}(\alpha)]^y\, W_\alpha(N_0)\right).
\end{aligned} \tag{A.9}$$

For every $y \in \mathbb{N}$, we have by (1.1) that

$$\begin{aligned}
&W_\alpha(N_0 + y, v_\alpha) \\
&= C(N_0 + y, s) + \sum_{z \in \mathbb{N}} P_{N_0+y,z}(s)\, E^{v_\alpha}_{N_0+y} \left[e^{-\alpha t_2} \sum_{n \geq 2} e^{-\alpha(t_n - t_2)}\, C(X_n, A_n) \,|\, X_2 = z\right] \\
&= C(N_0 + y, s) + E^{v_\alpha}_{N_0+y} \left[e^{-\alpha t_2}\right] \sum_{z \in \mathbb{N}} P_{N_0+y,z}(s)\, W_\alpha(z) \\
&= C(N_0 + y, s) + \hat{d}(\alpha)\, \hat{T}(\alpha)^{N_0+y} \sum_{z \in \mathbb{N}} P_V(z)\, W_\alpha(z).
\end{aligned} \tag{A.10}$$

On the other hand because $\Delta_\alpha(N_0) > 0$ for α small enough (see the proof of Proposition 2.2) we may deduce from Proposition 2.1 that the optimal action is to serve in state N_0 for α small enough. Hence, cf. (2.10), we have for small α

$$W_\alpha(N_0) = C(N_0, s) + \hat{d}(\alpha)\, \hat{T}(\alpha)^{N_0} \sum_{y\in\mathbb{N}} P_V(y)\, W_\alpha(y) \tag{A.11}$$

so that, cf. (A.10), (A.11),

$$\begin{aligned} W_\alpha(N_0 + y, v_\alpha) - \left[\hat{T}(\alpha)\right]^y W_\alpha(N_0) &= C(N_0 + y, s) - \left[\hat{T}(\alpha)\right]^y C(N_0, s) \\ &\leq C(N_0 + y, s) \end{aligned} \tag{A.12}$$

for any $y \in \mathbb{N}$ and for α small enough. Consequently, cf. (A.9), (A.12), (2.4),

$$\begin{aligned} &\sum_{y\in\mathbb{N}} P_V(y) \left(W_\alpha(N_0 + y) - \left[\hat{T}(\alpha)\right]^y W_\alpha(N_0)\right) \\ &\leq \sum_{y\in\mathbb{N}} P_V(y)\, C(N_0 + y, s) \\ &= C(N_0, s) + \left(\frac{\rho d}{1-\rho}\right)(N_0 + 1) + \left(\frac{\lambda^2}{2(1-\rho)}\right)\left(bd^{(2)} + \frac{db^{(2)}}{1-\rho}\right) \end{aligned} \tag{A.13}$$

by using the identities $\sum_{y\in\mathbb{N}} y\, P_V(y) = \lambda d$ and $\sum_{y\in\mathbb{N}} y^2\, P_V(y) = \lambda d + \lambda^2 d^{(2)}$.

Next, we compute an upper bound for the term in the second summation in the right-hand side of (A.8). We shall distinguish between the cases $y > N_0$ and $0 \leq y \leq N_0$.

Fix $y > N_0$. Since $\{W_\alpha,\, \alpha > 0\}$ satisfies condition **C2** we have for all $\alpha > 0$ that

$$W_\alpha(N_0) - \left[\hat{T}(\alpha)\right]^{N_0-y} W_\alpha(y) = -\left[\hat{T}(\alpha)\right]^{N_0-y}\left(W_\alpha(y) - \left[\hat{T}(\alpha)\right]^{y-N_0} W_\alpha(N_0)\right) \leq 0. \tag{A.14}$$

Fix now $0 \leq y \leq N_0$. If the α-optimal policy serves in state y, then (use (2.6) again and (A.11))

$$W_\alpha(N_0) - \left[\hat{T}(\alpha)\right]^{N_0-y} W_\alpha(y) = C(N_0, s) - \hat{T}(\alpha)^{N_0-y}\, C(y, s) \leq C(N_0, s). \tag{A.15}$$

for α small enough. If the α-optimal policy does not serve in state y, then

$$\begin{aligned} &W_\alpha(N_0) - \left[\hat{T}(\alpha)\right]^{N_0-y} W_\alpha(y) \\ &\leq C(N_0, s) + \hat{d}(\alpha) \left[\hat{T}(\alpha)\right]^{N_0-y} \sum_{y'=0}^{\infty} P_v(y') \left(W_\alpha(y') \left[\hat{T}(\alpha)\right]^y - W_\alpha(y + y')\right) \\ &\leq C(N_0, s) \end{aligned} \tag{A.16}$$

for α small enough, since by Lemma 2.1 $W_\alpha(y+y') - \left[\hat{T}(\alpha)\right]^y W_\alpha(y') \geq 0$ for all $(y, y') \in \mathbb{N}^2$, $\alpha > 0$.

Combining (A.8)-(A.16) we see that for α small enough $\Delta_\alpha(N_0)$ is bounded from above by a finite number that does not depend on α. This concludes the proof of Proposition 2.3.

B Appendix

We prove in this appendix that Assumptions 1, 3, 4 in [18] hold as well as Assumption 5* introduced in the proof of Theorem 3.1.

Assumption 1 in [18] holds because $P(t_2 > \delta \,|\, X_1 = x, A_1 = a) \geq P(D > \delta) > 0$ for some $\delta > 0$ since $d > 0$.

Assumption 3 holds because $W_\alpha \in \mathcal{K}$ by Proposition 2.1 which in turn implies by definition of the norm $||\cdot||$ that $W_\alpha(x)$ is finite for every $x \in \mathbb{N}$, $\alpha > 0$.

To prove that Assumption 4 holds we use again the policy v_α that resumes service at the first decision epoch (provided the queue is non-empty) and then follows the α-discounted optimal policy. Since the policy v_α cannot perform better than the optimal policy, we have for $y \in \mathbb{N}$

$$\begin{aligned} W_\alpha(y) - W_\alpha(N_0) &\leq W_\alpha(y, v_\alpha) - W_\alpha(N_0), \\ &= W_\alpha(y, v_\alpha) - [\hat{T}(\alpha)]^{y-N_0}\, W_\alpha(N_0) - \left(1 - [\hat{T}(\alpha)]^{y-N_0}\right) W_\alpha(N_0). \quad \text{(B.1)} \end{aligned}$$

We have shown in Proposition 2.3 that the optimal action in state N_0 is to resume service for α small enough. Hence, by (1.1) and (A.11) we obtain that for α small enough

$$\begin{aligned} W_\alpha(y, v_\alpha) - [\hat{T}(\alpha)]^{y-N_0}\, W_\alpha(N_0) &= C(y,s)\mathbf{1}(y \geq 1) + C(0,v)\mathbf{1}(y=0) - C(N_0,s)[\hat{T}(\alpha)]^{y-N_0} \\ &\leq N(y), \quad y \in \mathbb{N} \quad \text{(B.2)} \end{aligned}$$

where $N(y) := C(y,s)\mathbf{1}(y \geq 1) + C(0,v)\mathbf{1}(y=0)$. Because $1 - [\hat{T}(\alpha)]^{y-N_0}$ is nonnegative for $y \geq N_0$ it is seen from (B.1) and (B.2) that for α small enough $W_\alpha(y) - W_\alpha(N_0) \leq N(y)$ for $y \geq N_0$.

Consider now the case when $y < N_0$. We have

$$\begin{aligned} -\left(1 - [\hat{T}(\alpha)]^{y-N_0}\right) W_\alpha(N_0) &= [\hat{T}(\alpha)]^{y-N_0}\left(1 - [\hat{T}(\alpha)]^{N_0-y}\right) W_\alpha(N_0) \\ &\leq \left(\frac{bN_0}{1-\rho}\right) \alpha W_\alpha(N_0) \quad \text{(B.3)} \end{aligned}$$

by using the inequality $1 - [\hat{T}(\alpha)]^k \leq kb\alpha/(1-\rho)$ for all $\alpha \geq 0$, $k \in \mathbb{N}$. We now show that $\alpha W_\alpha(N_0)$ is uniformly bounded for small values of α.

Recall that $\overline{\lim}_{\alpha\to 0}\, \alpha W_\alpha(x,u) \leq \Phi(x,u)$ for every $x \in \mathbb{N}$, $u \in \mathcal{U}$ (see [18, Proposition 1]). On the other hand, $W_\alpha(x) \leq W_\alpha(x, u_l)$ for every $l \geq 1$, $x \in \mathbb{N}$ by definition of W_α. Hence, for every $l \geq 1$,

$$\overline{\lim_{\alpha\to 0}}\, \alpha W_\alpha(x) \leq \overline{\lim_{\alpha\to 0}}\, \alpha W_\alpha(x, u_l) \leq \Phi(x, u_l) < \infty, \quad x \in \mathbb{N}$$

where the last equality follows from Lemma 3.1. Therefore, there exist $M > 0$ and $\alpha_0 > 0$ such that

$$\alpha W_\alpha(N_0) \leq M \quad \text{for } \alpha \in (0, \alpha_0). \quad \text{(B.4)}$$

In summary, we have shown that for α small enough $W_\alpha(y) - W_\alpha(N_0) \le N(y) + (bN_0/(1-\rho))\, M :=$ $M(y)$ for every $y \in \mathbb{N}$.

The second condition in Assumption 4 requires that $\sum_{y\in\mathbb{N}} P_V(y)\, M(y) < \infty$. This follows from the definition of C together with the property that $\sum_{y\in\mathbb{N}} y^j\, P_V(y) < \infty$ for $j = 1, 2$.

We now examine the validity of Assumption 5*. For α small enough and $y \le N_0$ we see from (A.15) and (A.16) that

$$W_\alpha(y) - W_\alpha(N_0) \ge -C(N_0, s). \tag{B.5}$$

We now address the case when $y > N_0$. Recall that $\{W_\alpha,\ \alpha > 0\}$ satisfies condition **C1** by Lemma 2.1. So, for every $y > N_0$ there exists $\beta_y > 0$ such that whenever $\alpha \in (0, \beta_y)$

$$\begin{aligned}
W_\alpha(y) - W_\alpha(N_0) &= W_\alpha(y) - [\hat{T}(\alpha)]^{y-N_0}\, W_\alpha(N_0) + \left([\hat{T}(\alpha)]^{y-N_0} - 1\right) W_\alpha(N_0) \\
&\ge C(y, s) - C(N_0, s) + \left([\hat{T}(\alpha)]^{y-N_0} - 1\right) W_\alpha(N_0) \\
&\ge C(y, s) - C(N_0, s) - \frac{bM\, y}{1-\rho} \qquad \text{(B.6)} \\
&\ge -\frac{bM^2}{2(1-\rho)} - C(N_0, s) \qquad \text{(B.7)}
\end{aligned}$$

where (B.6) follows from (B.4) together with the inequality $1 - [\hat{T}(\alpha)]^k \le kb\alpha/(1-\rho)$ and where the lower bound in (B.7) is obtained from (2.4) (more precisely, $C(y, s) \ge (b/(2(1-\rho)))\, y^2$ for all $y \in \mathbb{N}$ so that $C(y, s) - (b/(1-\rho))\, My \ge -(b/(2(1-\rho)))\, M^2$ for all $y \in \mathbb{N}$). The validity of Assumption 5* follows from (B.5) and (B.7).

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 for publication.

[2] E. Altman and P. Nain, "Optimal control of the M/G/1 queue with repeated vacations," INRIA Report no. 1637, Mar. 1992.

[3] E. Cinlar, *Introduction to Stochastic Processes*, Prentice Hall, 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, *S<tochastic 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] M. Schäl, "Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal," *Z. Wahrscheinlichkeitsth.*, vol. 32, 179-196, 1975.

[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] L. I. Sennott, "Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs," *Opns. Res.*, vol. 37, no. 4, 626-636, 1989.

[20] 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.

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