OPTIMALITY OF A THRESHOLD POLICY IN THE M/M/1 QUEUE WITH REPEATED VACATIONS*

Eitan Altman and Philippe Nain
INRIA
2004, Route des Lucioles
06565 Valbonne Cedex
France

July 25, 1995

Abstract

Consider an M/M/1 queueing system with server vacations where the server is turned off as soon as the queue gets empty. We assume that the vacation durations form a sequence of i.i.d. random variables with exponential distribution. At the end of a vacation period, the server may either be turned on if the queue is non empty or take another vacation. The following costs are incurred: a holding cost of h per unit of time and per customer in the system and a fixed cost of γ each time the server is turned on. We show that there exists a threshold policy that minimizes the long-run average cost criterion. The approach we use was first proposed in Blanc et al. (1990) and enables us to determine explicitly the optimal threshold and the optimal long-run average cost in terms of the model parameters.

Subject classification: Dynamic programming; Markov decision processes; queueing control models; queues with vacations.

Queueing systems with vacations of the server have already received much attention in the literature and a comprehensive discussion can be found in the survey papers by Doshi (1986) and Teghem (1986). These models are commonly used for modeling and tuning various systems ranging from manufacturing systems to communication and computer systems (cf. Doshi (1986)). Vacation models can be classified into two categories: models with repeated (multiple) vacations of the server and models with a removable server. In the former case the length of a vacation period is driven by an external process (e.g., the vacation lengths are i.i.d. random variables, see Gelenbe and Mitrani (1980), Gelenbe et Iasnogorodski (1980), Kella (1989, 1990), Levy and Yechiali (1975),

*Submitted to *Operations Research*.

while in the latter case the length of a vacation period is driven by the arrival process (see Heyman and Sobel (1984, pp. 336-337), Yadin and Naor (1963)). In both cases, the server may be turned off at any service completion epoch of a customer and it may be turned on only at an arrival epoch of a customer in the case of a removable server, and only at the end of a vacation period in the case of repeated vacations. Of particular interest is the stationary policy that turns the server on only when the number of customers in the queue is equal to or larger than a given value, and turns the server off when the queue is empty. This policy will be referred to as the *threshold policy*.

A natural objective is to seek for an optimal vacation policy that optimizes a given cost function among certain classes of policies. For the M/G/1 queue with a removable server Heyman and Sobel (1984) and Talman (1979) have shown that a fairly general average cost criterion is minimized by a threshold policy among the class of all policies. For the more difficult problem of controlling the service process in a queue with repeated vacations, only a few optimization results have been reported so far. Recently, Kella (1989) has computed the best threshold policy for an M/G/1 queue with repeated vacations, and Lee and Srinivasan (1989) have carried out the same analysis in the case of batch arrivals. In the case where the decision to take another vacation is based on a random outcome depending on the number of consecutive vacations already taken, Kella (1990) has shown that a control policy of a limit type minimizes a long-run average cost criterion.

Our contribution is to establish the optimality of a threshold policy over the class of all policies, including policies that may depend on the history of the system (i.e., number of customers at any time and previous decisions made).

More precisely, consider an M/M/1 queue under the exhaustive service discipline and with repeated vacations, where the lengths of the vacation periods are i.i.d. random variables with an exponential distribution. Assume that a holding cost $h \geq 0$ is incurred per unit of time and per customer and that a fixed cost $\gamma \geq 0$ is incurred when the server is turned on. We show in this paper that there exists a threshold policy that minimizes the long-run average cost criterion.

The approach we use closely follows that proposed by Blanc et al. (1990) and enables us to *explicitly* compute both the optimal threshold and the optimal long-run average cost in terms of the model parameters.

Altman and Nain (1993) and Federgruen and So (1991) have shown independently that the optimality of a threshold policy actually extends to the M/G/1 queue with arbitrary repeated vacations. However, the optimal threshold cannot be determined explicitly in this case even by specializing the model to the model studied in the present paper, and a numerical procedure (see e.g., Kella (1990)) has to be used instead.

The paper is organized as follows. In Section 1, the problem is formulated as a Markov decision problem. In Sections 2 and 3 a related discounted cost problem is introduced and solved via dynamic programming and the technique proposed in Blanc et al. (1990). This finally enables us to derive the optimal policy for the long-run average cost criterion (Section 4).

1 The Model

We consider an $M/M/1$ queue where the server serves the customers according to an arbitrary work conserving service discipline until the queue gets empty, and then takes a vacation of random length. When returning from a vacation, the server may either be turned off (i.e., take a new vacation) or turned on (i.e., resume work), provided the queue is non empty, according to a *vacation policy* (a precise definition of a vacation policy is given below). The vacation lengths are i.i.d. random variables (r.v.'s) with an exponential distribution. We assume that the interarrival, service and vacation time processes are mutually independent. Let $1/\lambda$, $1/\mu$ and $1/\nu$ be the mean interarrival time, mean service time and mean vacation time, respectively.

We assume that a holding cost $h \geq 0$ is incurred per unit of time and per customer in the queue and that a restarting cost $\gamma \geq 0$ is incurred each time the server is turned on. Our objective is to find a vacation policy that minimizes the *long-run average cost* incurred over an infinite horizon.

Let $\mathbb{N}$ be the set of all nonnegative integers and $\mathbb{R}$ be the set of all real numbers.

In order to place our control problem in the Markov Decision Process (MDP) setting (see Bertsekas (1987), Ross (1970), Schäl (1975)), we need to introduce extra observation points. More specifically, we shall assume that the system (to be made more precise) is observed at every arrival time, service completion time and jump time of a Poisson process $\mathbf{V}$, independent of the arrival and service time processes, with intensity ν. The process $\mathbf{V}$ plays the role of a *virtual* vacation process in the sense that, if a jump occurs in $\mathbf{V}$ whenever the server is on vacation, then this jump may be taken as the next vacation completion time. This follows from the memoryless property of the exponential distribution, the independence of the arrival, service time and $\mathbf{V}$ processes, and the fact that the mean interevent time for the process $\mathbf{V}$ is the same as the mean vacation time. At every observation point that corresponds to a jump in $\mathbf{V}$, a decision will be made on whether or not to turn the server on (provided the queue is non-empty). This follows from the description of the control problem at hand and from the definition of the virtual process $\mathbf{V}$. At all other observation points, dummy actions will be taken that will not affect the system behavior.

With this definition of the observation (decision) points in mind, we shall construct an MDP such that the state of the process at the n-th decision epoch t_n, $n \geq 1$, is represented by the triple $(X_n, Y_n, Z_n) \in \mathbb{N} \times \{0,1\}^2$, where X_n is the queue-length at time t_n^+, $Y_n \in \{0,1\}$ describes the activity of the server at time t_n^- ($Y_n = 1$ if the server is working and $Y_n = 0$ if it is on vacation) and $Z_n = \mathbf{1}(t_n \in \mathbf{V})$, that is $Z_n = 1$ if t_n is a jump time of the process $\mathbf{V}$ and $Z_n = 0$, otherwise. It is readily seen from the definition of X_n, Y_n and Z_n, $n \geq 1$, that states $(0,0,0)$ and $(0,1,1)$ are not accessible.

The MDP may now be constructed as follows. Let $\mathbf{S} := \mathbb{N} \times \{0,1\}^2 - \{(0,0,0),(0,1,1)\}$ be the state space, and let $\mathbf{A}_{x,y,z} \subset \{0,1\}$ be the set of all available actions when the system is in state $(x,y,z) \in \mathbf{S}$. We assume that

$$\mathbf{A}_{x,1,1} = \{1\}, \quad \text{for } x \geq 1;$$

$$\mathbf{A}_{x,1,0} = \begin{cases} \{1\}, & \text{for } x \geq 1; \\ \{0\}, & \text{for } x = 0; \end{cases}$$

$$\mathbf{A}_{x,0,1} = \begin{cases} \{0,1\}, & \text{for } x \geq 1; \\ \{0\}, & \text{for } x = 0; \end{cases}$$

$$\mathbf{A}_{x,0,0} = \{0\}, \quad \text{for } x \geq 1.$$

where by convention action 1 (resp. 0) is taken if the decision is to turn the server on (resp. off). Note that the only states when more that one action is available are states $(x, 0, 1)$ with $x \geq 1$, that is, when a jump occurs in $\mathbf{V}$ $(z = 1)$, that the server is on vacation $(y = 0)$ and that the queue is non-empty $(x \geq 1)$. This reflects the control problem at hand. In all other cases, the definition of the single available action is arbitrary, except for state $(0, 1, 0)$ when necessarily action 0 has to be taken to follow the definition of the vacation scheme.

Following Schäl (1975), a vacation policy U is then defined as a sequence of conditional probabilities $U_n : \mathbb{H}_n \to \mathcal{P}(\{0,1\})$, $\mathbb{H}_1 = \mathbf{S}$, $\mathbb{H}_{n+1} = \mathbb{H}_n \times \{0,1\}$ for all $n \in \mathbb{N}$, such that $U_n(h_n; \bullet)$, $h_n = (s_1, a_1, \ldots, s_{n-1}, a_{n-1}, s_n)$, assigns probability one to the set $\mathbf{A}_{s_n}$ (the notation $\mathcal{P}(\{0,1\})$ stands the set of all probability measures on $\{0,1\}$). As usual, a policy is *stationary* if $U_n(h_n; \bullet)$ is concentrated at the point $\alpha(s_n)$ for all $h_n = (s_1, a_1, \ldots, s_{n-1}, a_{n-1}, s_n) \in \mathbb{H}_n$, $n \geq 1$, where α is a measurable mapping from $\mathbf{S}$ to $\{0,1\}$. Let $\mathcal{U}$ be the set of all vacation policies.

Let $q(\bullet \,|\, s; a)$ be the probability distribution of the next state visited by the system if the system is in state $(x, y, z) \in \mathbf{S}$ and the action $a \in \mathbf{A}_{x,y,z}$ is chosen. The transition probabilities are given by (with $\beta := \lambda + \mu + \nu$)

$$\begin{aligned} q(x-1,1,0 \,|\, x,1,1;1) &= \mu/\beta, \quad \text{for } x \geq 1; \\ q(x+1,1,0 \,|\, x,1,1;1) &= \lambda/\beta, \quad \text{for } x \geq 1; \\ q(x,1,1 \,|\, x,1,1;1) &= \nu/\beta, \quad \text{for } x \geq 1; \end{aligned}$$

$$\begin{aligned} q(x-1,1,0 \,|\, x,1,0;1) &= \mu/\beta, \quad \text{for } x \geq 1; \\ q(x+1,1,0 \,|\, x,1,0;1) &= \lambda/\beta, \quad \text{for } x \geq 1; \\ q(x,1,1 \,|\, x,1,0;1) &= \nu/\beta, \quad \text{for } x \geq 1; \\ q(0,0,1 \,|\, 0,1,0;0) &= \nu/(\lambda+\nu); \\ q(1,0,0 \,|\, 0,1,0;0) &= \lambda/(\lambda+\nu); \end{aligned}$$

$$\begin{aligned} q(x-1,1,0 \,|\, x,0,1;1) &= \mu/\beta, \quad \text{for } x \geq 1; \\ q(x+1,1,0 \,|\, x,0,1;1) &= \lambda/\beta, \quad \text{for } x \geq 1; \\ q(x,1,1 \,|\, x,0,1;1) &= \nu/\beta, \quad \text{for } x \geq 1; \\ q(x+1,0,0 \,|\, x,0,1;0) &= \lambda/(\lambda+\nu), \quad \text{for } x \in \mathbb{N}; \\ q(x,0,1 \,|\, x,0,1;0) &= \nu/(\lambda+\nu), \quad \text{for } x \in \mathbb{N}; \end{aligned}$$

$$\begin{aligned} q(x+1,0,0 \mid x,0,0;0) &= \lambda/(\lambda+\nu), \;\; \text{for } x \geq 1; \\ q(x,0,1 \mid x,0,0;0) &= \nu/(\lambda+\nu), \;\; \text{for } x \geq 1. \end{aligned}$$

Given an initial distribution on $\mathbf{S}$ and the transition law q, any policy $U \in \mathcal{U}$ defines a probability measure on the product space $(\mathbf{X} \times \mathbf{A})^\infty$ endowed with the product σ-algebra. Let E^U be the expectation operator associated with this probability measure. On this probability space are defined the random vector (X_n, Y_n, Z_n) that describes the state of the system at the n-th decision epoch as well as the r.v. A_n that describes the action taken at the n-th decision epoch, $n \geq 1$. The reader can check from the above construction (in particular, from the definition of the transition law q) that the interpretation of the r.v.'s X_n, Y_n, Z_n and t_n, $n \geq 1$, is indeed the one given at the beginning of this section.

Our objective is to minimize

$$W^U(x,y,z) := \limsup_{T \uparrow \infty} T^{-1} W_T^U(x,y,z), \tag{1.1}$$

over the set $\mathcal{U}$ for all $(x,y,z) \in \mathbf{S}$, where

$$W_T^U(x,y,z) := \mathrm{E}_U \left[\sum_{0 \leq t_n \leq T} c(X_n, Y_n, Z_n; A_n) \mid (X_1, Y_1, Z_1) = (x,y,z) \right], \tag{1.2}$$

with

$$c(x,y,z;a) := (h/\nu)\, xz + \gamma\, (1-y) z\, \mathbf{1}(a=1), \tag{1.3}$$

for all $(x,y,z) \in \mathbf{S}$, $a \in \mathbf{A}_{x,y,z}$.

The first result shows that (1.1) is equal to the long-run average cost incurred in the M/M/1 queue when the vacation policy U is used.

Proposition 1.1 *Fix $U \in \mathcal{U}$. Let $X(t)$ and $Y(t)$ be the queue-length and the state of the server at time $t \geq 0$, respectively, when policy U is used. We assume that the sample paths of the processes $\{X(t),\, t \geq 0\}$ and $\{Y(t),\, t \geq 0\}$ are right-continuous and left-continuous, respectively. Then,*

$$W_T^U(x,y,z) = E_U \left[\int_0^T hX(t)\, dt + \gamma\, (1 - Y(t))\, dY(t) \mid X(0) = x, Y(0) = y, Z_1 = z \right], \tag{1.4}$$

for all $(x,y,z) \in \mathbf{S}$, where $X(0) = X_1$ and $Y(0) = Y_1$ under the assumption that $t_1 = 0$.

Proof. Define v_n to be the n-th jump time of the process $\mathbf{V}$, $n \geq 1$. Let $N(t) := \sum_{n \geq 1} \mathbf{1}(v_n \leq t)$ be the number of jumps of the process $\mathbf{V}$ in $[0,t]$, $t \geq 0$. It is seen from (1.2), (1.3) that

$$W_T^U(x,y,z) = \mathrm{E}_U \left[\sum_{0 \leq t_n \leq T} ((h/\nu)\, X_n + \gamma(1-Y_n)\, \mathbf{1}(a_n = 1))\, Z_n \mid (X_1, Y_1, Z_1) = (x,y,z) \right],$$

$$= \mathrm{E}_U \left[\int_0^T (h/\nu)\, X(t)\, dN(t) + \gamma\, (1 - Y(t))\, dY(t) \,|\, (X_1, Y_1, Z_1) = (x, y, z) \right],$$

$$= \mathrm{E}_U \left[\int_0^T (h/\nu)\, X(t-)\, dN(t) + \gamma\, (1 - Y(t))\, dY(t) \,|\, (X_1, Y_1, Z_1) = (x, y, z) \right], \tag{1.5}$$

$$= \mathrm{E}_U \left[\int_0^T h\, X(t-)\, dt + \gamma\, (1 - Y(t))\, dY(t) \,|\, (X_1, Y_1, Z_1) = (x, y, z) \right], \tag{1.6}$$

$$= \mathrm{E}_U \left[\int_0^T h\, X(t)\, dt + \gamma\, (1 - Y(t))\, dY(t) \,|\, (X_1, Y_1, Z_1) = (x, y, z) \right]. \tag{1.7}$$

Equality (1.5) follows from the fact that with probability one both processes $\{X(t),\, t \geq 0\}$ and $\mathbf{V}$ have no common jumps. Equality (1.6) follows from Brémaud (1981), formula 2.3, p. 24, with $C_t := X(t-)\, \mathbf{1}(t \leq T)$, $N_t := N(t)$, $\lambda_t := \nu$ and $\mathcal{F}_t = \sigma(X(s),\, 0 \leq s < t) \vee \sigma(N(s),\, 0 \leq s \leq t)$. ■

Remark 1.1 Additional costs/rewards could be considered. In particular, the system could receive a reward for each unit of time the server is on vacation, and a constant cost could be incurred each time the server is turned off. However, and as observed by Kella (1990, p. 116), these extensions can easily be captured by our model.

2 A Related Discounted Cost Problem

Minimizing directly $W^U(x, y, z)$, cf. (1.1), over the set of policies $\mathcal{U}$ turns out to be a difficult task. To achieve this goal, we shall first (partially) solve a related discounted cost problem. Then, the use of a Tauberian theorem (see Section 4) will allow us to determine an optimal (stationary) policy for the long-run average cost problem.

Fix $\alpha > 0$. Our first objective is to minimize over $\mathcal{U}$ the α-discounted cost function

$$V_\alpha^U(x, y, z) := \mathrm{E}_U \left[\sum_{n \geq 1} e^{-\alpha t_n}\, c(X_n, Y_n, Z_n; u_n) \,|(X_1, Y_1, Z_1) = (x, y, z) \right], \tag{2.1}$$

for all (x, y, z) in $\mathbf{S}$, where the cost c is defined in (1.3).

Let $V_\alpha(x, y, z) := \inf_{U \in \mathcal{U}} V_\alpha^U(x, y, z)$, $(x, y, z) \in \mathbf{S}$.

For each function $f : \mathbf{S} \to \mathbb{R}$, set

$$\| f \| := \sup_{\substack{(x,y,z) \in \mathbf{S} \\ x \neq 0}} |f(x, y, z)|\, x^{-1}, \tag{2.2}$$

and define $\mathcal{B}$ to be the Banach space (with norm given in (2.2)) of all such f for which $\| f \| < \infty$.

Unless otherwise mentioned, we will assume from now on that $h/\nu = 1$. The following basic result of Dynamic Programming (DP) holds:

Theorem 2.1 *There exists an optimal stationary policy for the α-discounted problem. In addition, V_α is the unique solution in the Banach $\mathcal{B}$ to the DP equation*

$$V_\alpha(x,y,z) = \min_{u \in \mathbf{A}_{x,y,z}} \left\{ c(x,y,z;u) + \frac{\theta_{x,y,z}(u)}{\alpha + \theta_{x,y,z}(u)} \sum_{(x',y',z') \in \mathbf{S}} V_\alpha(x',y',z')\, q(x',y',z' \,|\, x,y,z;u) \right\}, \tag{2.3}$$

for all $(x,y,z) \in \mathbf{S}$, where $\theta_{x,y,z}(u)$ is the transition rate out of state (x,y,z) given that action $u \in \mathbf{A}_{x,y,z}$ is chosen.

Furthermore, the stationary vacation policy which selects an action minimizing the right-hand side of (2.3) for all $(x,y,z) \in \mathbf{S}$ is optimal.

Proof. First note that the costs (1.3) are unbounded. However, one can easily establish that Assumptions 2 and 3 in Lippman (1975) are satisfied. Therefore, the proof follows from Theorem 1 in Lippman (1975). ∎

It is easily seen from Theorem 2.1 that

$$\begin{aligned}(\alpha+\beta)\,V_\alpha(x,1,1) &= x\,(\alpha+\beta) + \mu\,V_\alpha(x-1,1,0) + \lambda\,V_\alpha(x+1,1,0)\\ &+ \nu\,V_\alpha(x,1,1), \quad \text{for } x \geq 1;\end{aligned} \tag{2.4}$$

$$\begin{aligned}(\alpha+\beta)\,V_\alpha(x,1,0) &= \mu\,V_\alpha(x-1,1,0) + \lambda\,V_\alpha(x+1,1,0)\\ &+ \nu\,V_\alpha(x,1,1), \quad \text{for } x \geq 1;\end{aligned} \tag{2.5}$$

$$(\alpha+\lambda+\nu)\,V_\alpha(0,1,0) = \lambda\,V_\alpha(1,0,0) + \nu\,V_\alpha(0,0,1); \tag{2.6}$$

$$(\alpha+\lambda+\nu)\,V_\alpha(0,0,1) = \lambda\,V_\alpha(1,0,0) + \nu\,V_\alpha(0,0,1); \tag{2.7}$$

$$\begin{aligned}V_\alpha(x,0,1) &= \min\left\{ x + \frac{\lambda}{\alpha+\lambda+\nu}\,V_\alpha(x+1,0,0) + \frac{\nu}{\alpha+\lambda+\nu}\,V_\alpha(x,0,1);\right.\\ &\quad x + \gamma + \frac{\mu}{\alpha+\beta}\,V_\alpha(x-1,1,0) + \frac{\lambda}{\alpha+\beta}\,V_\alpha(x+1,1,0)\\ &\quad \left. + \frac{\nu}{\alpha+\beta}\,V_\alpha(x,1,1) \right\}, \quad \text{for } x \geq 1;\end{aligned} \tag{2.8}$$

$$(\alpha+\lambda+\nu)\,V_\alpha(x,0,0) = \lambda\,V_\alpha(x+1,0,0) + \nu\,V_\alpha(x,0,1), \quad \text{for } x \geq 1. \tag{2.9}$$

The end of this section is devoted to reducing the number of unknown quantities involved in the set of equations (2.4)-(2.9).

First, combining (2.6) and (2.7) gives us

$$V_\alpha(0,1,0) = V_\alpha(0,0,1) = \frac{\lambda}{\alpha+\lambda} V_\alpha(1,0,0). \tag{2.10}$$

Then, using (2.5) and (2.9) we see that (2.8) can be rewritten as

$$V_\alpha(x,0,1) = x + \min\{V_\alpha(x,0,0); \gamma + V_\alpha(x,1,0)\}, \quad \text{for } x \geq 1. \tag{2.11}$$

Introducing now (2.11) into (2.9) yields for $x \geq 1$

$$(\alpha+\lambda+\nu)\, V_\alpha(x,0,0) = \nu x + \lambda\, V_\alpha(x+1,0,0) + \nu\, \min\{V_\alpha(x,0,0); V_\alpha(x,1,0)+\gamma\}. \tag{2.12}$$

On the other hand, we observe from (2.4) and (2.5) that

$$V_\alpha(x,1,1) = x + V_\alpha(x,1,0), \quad \text{for } x \geq 1. \tag{2.13}$$

Introducing (2.13) into (2.5) finally yields for $x \geq 1$

$$(\alpha+\lambda+\mu)\, V_\alpha(x,1,0) = \nu x + \lambda\, V_\alpha(x+1,1,0) + \mu\, V_\alpha(x-1,1,0). \tag{2.14}$$

Relations (2.10), (2.12), (2.13) and (2.14) contain all the information carried by the DP equation (2.3). A glance at relation (2.14) indicates that it defines a difference equation. This yields the following result:

Lemma 2.1

$$V_\alpha(x,1,0) = \left(V_\alpha(1,0,0)\left(\frac{\lambda}{\alpha+\lambda}\right) + \nu\left(\frac{\mu-\lambda}{\alpha^2}\right)\right)\beta_{\alpha,1}^x + \frac{\nu}{\alpha}\,x + \nu\left(\frac{\lambda-\mu}{\alpha^2}\right), \tag{2.15}$$

for all $x \in \mathbb{N}$, where $\beta_{\alpha,1}$, $0 < \beta_{\alpha,1} < 1$, is the smallest root of the polynomial (in z) $\lambda z^2 - (\alpha + \lambda + \mu)\, z + \mu$.

Proof. The general solution of the difference equation (2.14) is

$$V_\alpha(x,1,0) = a\,\beta_{\alpha,1}^x + b\,\beta_{\alpha,2}^x + c\,x + d, \tag{2.16}$$

for $x \in \mathbb{N}$, where $\beta_{\alpha,1}$ and $\beta_{\alpha,2}$ are the roots of the polynomial (in z) $\lambda z^2 - (\alpha+\lambda+\mu)\, z + \mu$, with $0 < \beta_{\alpha,1} < 1 < \beta_{\alpha,2}$.

By remembering that $V_\alpha(x,1,0)/x$ is uniformly bounded in $\mathbb{N}^*$, we see from (2.16) that necessarily $b=0$ since $\beta_{\alpha,2}>1$. The remaining coefficients a, c and d are easily identified by introducing (2.16) into (2.14) and by using (2.10). ■

In summary, we have shown in this section that all $V(x,y,z)$, $(x,y,z)\in\mathbf{S}$, only express in terms of $V(1,1,0)$, $V(x,0,1)$ and $V(x,0,0)$ for $x\geq 1$ through equations (2.10), (2.12), (2.13) and (2.15). In the following, (2.12) and (2.15) will turn out to be the key equations.

3 Properties of the Optimal α-Discounted Policy

The goal of this section is to further characterize the optimal α-discounted policy.

¿From now on we shall assume that $\lambda<\mu$ (condition of ergodicity). Let $U_\alpha(x,y,z)$ be the optimal stationary vacation policy when in state $(x,y,z)\in\mathbf{S}$. As already observed, we only need to focus on states $(x,y,z)\in\mathbf{S}$ such that $x\neq 0$, $y=0$ and $z=1$. With a slight abuse of notation, define $U_\alpha(x):=U_\alpha(x,0,1)$ for $x\neq 0$, with the interpretation that $U_\alpha(x)=1$ (resp. 0) if the decision is to turn the server on (resp. to take another vacation) when the system is in state $(x,0,1)$, $x\neq 0$.

As mentioned in the introduction, we follow the method proposed by Blanc et al. (1990) for determining the optimal α-discounted policy. The procedure goes as follows:

(1) Assume that the optimal policy is a threshold policy with threshold $L\geq 1$ (call U_L this policy);

(2) Construct a function that would be a solution of the DP equation if policy U_L were indeed optimal;

(3) Show the existence of an integer L such that the function constructed in (2) satisfies the DP equation.

Steps (1) and (2) in yield the following proposition:

Proposition 3.1 *Assume that there exists a finite integer $L\geq 1$ and a family of L numbers $\{Y_L(x)\}_{x=1}^{L}$ that satisfy*

$$(\alpha+\lambda)\,Y_L(x) \;=\; \nu x+\lambda\,Y_L(x+1),\ \ \textit{for}\ x=1,2,\ldots,L-1; \tag{3.1}$$

$$Y_L(L) \;=\; C_\alpha\left(Y_L(1)\left(\frac{\lambda}{\alpha+\lambda}\right)+\nu\left(\frac{\mu-\lambda}{\alpha^2}\right)\right)\beta_{\alpha,1}^{L}+\frac{\nu}{\alpha}\,L+K_\alpha, \tag{3.2}$$

and such that

$$Y_L(1) \;\geq\; 0; \tag{3.3}$$

$$\begin{aligned} Y_L(x) &< Z_L(x)+\gamma, \quad \textit{for } x=1,2,\ldots,L-1; && (3.4)\\ Y_L(L) &\geq Z_L(L)+\gamma, && (3.5) \end{aligned}$$

where

$$\begin{aligned} Z_L(x) &:= \left(Y_L(1)\left(\frac{\lambda}{\alpha+\lambda}\right)+\nu\left(\frac{\mu-\lambda}{\alpha^2}\right)\right)\beta_{\alpha,1}^x+\frac{\nu}{\alpha}x+\nu\left(\frac{\lambda-\mu}{\alpha^2}\right); && (3.6)\\ C_\alpha &:= \frac{\nu}{\alpha+\nu+\lambda\left(1-\beta_{\alpha,1}\right)}; && (3.7)\\ K_\alpha &:= \left(\frac{\nu}{\alpha+\nu}\right)\left(\frac{\lambda}{\alpha}+\nu\left(\frac{\lambda-\mu}{\alpha^2}\right)+\gamma\right). && (3.8) \end{aligned}$$

Then,

$$U_\alpha(x)=\mathbf{1}(x\geq L), \quad \textit{for } x\geq 1,$$

and further, $Y_L(x)=V_\alpha(x,0,0)$ *for* $x=1,2,\ldots,L$ *and* $Z_L(x)=V_\alpha(x,1,0)$ *for* $x\geq 1$.

Proof. Let $\{Y_L(x)\}_{x=1}^L$ be a family of numbers that satisfy (3.1)-(3.5).

Define (see Remark 3.1) for detailed comments on this definition):

$$\tilde{Y}_L(x) := \begin{cases} Y_L(x), & \text{for } x=1,2,\ldots,L;\\[2ex] C_\alpha\left(Y_L(1)\left(\frac{\lambda}{\alpha+\lambda}\right)+\nu\left(\frac{\mu-\lambda}{\alpha^2}\right)\right)\beta_{\alpha,1}^x+\frac{\nu}{\alpha}x+K_\alpha, & \text{for } x\geq L+1. \end{cases} \quad (3.9)$$

Let us show that $\tilde{Y}_L(x)\geq Z_L(x)+\gamma$ for $x\geq L$, where $Z_L(x)$ is defined in (3.6).

For $x\geq L$, we have from (3.6) and (3.9)

$$\tilde{Y}_L(x)-Z_L(x)-\gamma=(C_\alpha-1)\left(Y_L(1)\left(\frac{\lambda}{\alpha+\lambda}\right)+\nu\left(\frac{\mu-\lambda}{\alpha^2}\right)\right)\beta_{\alpha,1}^x+K_\alpha+\nu\left(\frac{\mu-\lambda}{\alpha^2}\right)-\gamma. \quad (3.10)$$

We know from (3.5) and the definition of $\tilde{Y}_L(L)$ that the right-hand side of (3.10) is non negative for $x=L$. Since the right-hand side of (3.10) is a nondecreasing function of x (because $\lambda<\mu$, $Y_L(1)\geq 0$, $0<\beta_{\alpha,1}<1$, and $C_\alpha<1$), we therefore deduce from (3.10) that

$$\tilde{Y}_L(x)\geq Z_L(x)+\gamma, \quad \text{for } x\geq L. \quad (3.11)$$

Using this result, (2.15), (3.1), (3.6), and (3.9), it is easily seen that $V_\alpha(x,0,0)=\tilde{Y}_L(x)$ and $V_\alpha(x,1,0)=Z_L(x))$ satisfy (2.12) and (2.14).

Consequently, we have found a solution to the DP equation that belongs to $\mathcal{B}$ (because $\tilde{Y}_L(x)/x$ and $Z_L(x)/x$ are uniformly bounded in $\mathbb{N}^*$). Since such a solution is unique from Theorem 2.1, we deduce that necessarily (cf. also (2.10), (2.13)),

$$\begin{aligned} V_\alpha(x,0,0) &= \tilde{Y}_L(x), \text{ for } x \geq 1; && (3.12)\\ V_\alpha(0,1,0) &= (\lambda/(\alpha+\lambda))\, Y_L(1); && (3.13)\\ V_\alpha(x,1,0) &= Z_L(x), \text{ for } x \geq 1; && (3.14)\\ V_\alpha(x,1,1) &= x + Z_L(x), \text{ for } x \geq 1. && (3.15)\end{aligned}$$

Hence, cf. (3.4), (3.12), (3.14),

$$V_\alpha(x,0,0) < V_\alpha(x,1,0) + \gamma, \tag{3.16}$$

for $x = 1, 2, \ldots, L-1$, and, cf. (3.11), (3.12), (3.14),

$$V_\alpha(x,0,0) \geq V_\alpha(x,1,0) + \gamma, \tag{3.17}$$

for all $x \geq L$, or equivalently, $U_\alpha(x) = \mathbf{1}(x \geq L)$ for all $\in \mathbb{N}^*$, which concludes the proof. ■

We are now left with proving the existence of the integer L in Proposition 3.1. To do so, let us first introduce further notation. Let $\rho := \lambda/\mu$, $a := \lambda/\nu$ and define x_0 as the unique zero in $[0,\infty)$ of the polynomial (in w) $w^2 + (2a+1)\, w - 2a\gamma\,(1-\rho)$, that is

$$x_0 := \frac{-(2a+1) + \sqrt{(2a+1)^2 + 8a\gamma\,(1-\rho)}}{2}. \tag{3.18}$$

Last, let

l_0 be the smallest integer larger or equal to x_0 such that $l_0 \geq 1$. (3.19)

The symbol $O(\alpha)$ (resp. $O(1)$) will denote a function such that $\lim_{\alpha\downarrow 0} O(\alpha) = 0$ (resp. $= K$, $|K| < \infty$).

The following result holds:

Proposition 3.2 *Assume that $x_0 < l_0$. Then, there exists $\alpha_0 > 0$, such that for all $\alpha \in (0,\alpha_0)$, there exists a family of l_0 numbers $\{Y_{\alpha,l_0}(x)\}_{x=1}^{l_0}$ that satisfy (3.1)-(3.5).*

Proof. We first proceed with the solution $\{Y_{\alpha,L}(x)\}_{x=1}^{L}$ of the system of equations (3.1)-(3.2) when $L \geq 1$ and $\alpha > 0$ are both fixed.

¿From (3.1), we obtain that

$$Y_{\alpha,L}(x) = \left(\frac{\alpha+\lambda}{\lambda}\right)^{x-1} Y_{\alpha,L}(1) - \frac{\lambda\nu}{\alpha^2}\left[\left(\frac{\alpha+\lambda}{\lambda}\right)^{x} - 1\right] + \frac{\nu}{\alpha}\, x, \tag{3.20}$$

for $x = 2, \ldots, L$.

Combining (3.20) for $x = L$ and (3.2) yields $Y_{\alpha,L}(1)$, from which we deduce with (3.20) that

$$\begin{aligned} Y_{\alpha,L}(x) &= \left[\frac{\frac{\lambda\nu}{\alpha^2}\left(\left(\frac{\alpha+\lambda}{\lambda}\right)^L - 1\right) + \nu\left(\frac{\mu-\lambda}{\alpha^2}\right)\beta_{\alpha,1}^L\, C_\alpha + K_\alpha}{\left(\frac{\alpha+\lambda}{\lambda}\right)^L - \beta_{\alpha,1}^L\, C_\alpha} \right] \left(\frac{\alpha+\lambda}{\lambda}\right)^x \\ &+ \frac{\lambda\nu}{\alpha^2}\left[1 - \left(\frac{\alpha+\lambda}{\lambda}\right)^x\right] + \frac{\nu}{\alpha}\, x, \end{aligned} \tag{3.21}$$

for $x = 1, 2, \ldots, L$.

Finally, using (3.21) and the definition of $Z_{\alpha,L}(x)$ gives

$$Y_{\alpha,L}(x) - Z_{\alpha,L}(x) = \left[\frac{\frac{\nu\mu}{\alpha^2}\left(\beta_{\alpha,1}^L\, C_\alpha - \frac{\nu}{\alpha+\nu}\right) + \frac{\gamma\,\nu}{\alpha+\nu}}{\left(\frac{\alpha+\lambda}{\lambda}\right)^L - \beta_{\alpha,1}^L\, C_\alpha} \right] \left[\left(\frac{\alpha+\lambda}{\lambda}\right)^x - \beta_{\alpha,1}^x\right] + \frac{\nu\mu}{\alpha^2}\left(1 - \beta_{\alpha,1}^x\right), \tag{3.22}$$

for $x = 1, 2, \ldots, L$.

The first step is to prove the existence of $\alpha_1 > 0$ such that $Y_{\alpha,L}(1) > 0$ for $\alpha \in (0, \alpha_1)$. This property follows from the fact that (cf. Appendix A)

$$Y_{\alpha,L}(1) = \frac{A}{\alpha} + O(1), \tag{3.23}$$

where $A > 0$ does not depend on α.

The next step consists in evaluating the difference $Y_{\alpha,L}(x) - Z_{\alpha,L}(x)$ when α gets close to 0. After tedious algebra (cf. Appendix A), we obtain that

$$Y_{\alpha,L}(x) - Z_{\alpha,L}(x) = \left(\frac{L^2 + 2a\,L - x\,(L+a) + 2a^2 + 2a\,\gamma\,(1-\rho) + a}{L+a}\right)\left(\frac{x}{2a\,(1-\rho)}\right) + O(\alpha), \tag{3.24}$$

for $x = 1, 2, \ldots, L$.

On the other hand, it is seen from (3.24) that

$$Y_{\alpha,L}(L) - Z_{\alpha,L}(L) - \gamma = \frac{P(L)}{2\,(L+a)\,(1-\rho)} + O(\alpha), \tag{3.25}$$

where

$$P(w) := w^2 + (2a+1)\,w - 2a\,\gamma\,(1-\rho). \tag{3.26}$$

Recall the definitions of x_0 and l_0. Since $x_0 < l_0$ by assumption, we see that

$$P(L) \;<\; 0, \;\text{ for } L = 1, 2, \ldots, l_0 - 1; \tag{3.27}$$

$$P(l_0) \;>\; 0, \tag{3.28}$$

and so, cf. (3.25),

$$Y_{\alpha,L}(L) \;<\; Z_{\alpha,L}(L) + \gamma, \;\text{ for } L = 1, 2, \ldots, l_0 - 1; \tag{3.29}$$

$$Y_{\alpha,l_0}(l_0) \;>\; Z_{\alpha,l_0}(l_0) + \gamma, \tag{3.30}$$

for $\alpha \in (0, \alpha_2)$.

It remains to prove that for α small enough,

$$Y_{\alpha,l_0}(x) < Z_{\alpha,l_0}(x) + \gamma, \;\text{ for } x = 1, 2, \ldots, l_0 - 1. \tag{3.31}$$

To do so, rewrite $Y_{\alpha,L}(x) - Z_{\alpha,L}(x)$ as, cf. (3.24),

$$Y_{\alpha,L}(x) - Z_{\alpha,L}(x) = Y_{\alpha,L-1}(x) - Z_{\alpha,L-1}(x) + \left(\frac{x}{2a\,(1-\rho)}\right)(Q(L) - Q(L-1)) + O(\alpha), \tag{3.32}$$

for $x = 1, 2, \ldots, L - 1$, where

$$Q(w) := \frac{w^2 + 2a\,w + 2a^2 + 2a\,\gamma\,(1-\rho) + a}{w + a}. \tag{3.33}$$

Since

$$\frac{\partial Q(w)}{\partial w} = \frac{P(w) - (w+a)}{(w+a)^2} < 0 \;\text{ for } 0 \le w \le x_0,$$

we get that $Q(L) - Q(L-1) < 0$ for $L = 1, 2, \ldots, l_0 - 1$. Further, it is shown in Appendix B that $Q(l_0) - Q(l_0 - 1) < 0$ when $x_0 < l_0$. Consequently,

$$Q(L) - Q(L-1) < 0, \;\text{ for } l = 1, 2, \ldots l_0, \tag{3.34}$$

which implies from (3.32) that for α small enough,

$$Y_{\alpha,L}(x) - Z_{\alpha,L}(x) < Y_{\alpha,L-1}(x) - Z_{\alpha,L-1}(x), \tag{3.35}$$

for $x = 1, 2, \ldots, L - 1$, $L = 1, 2, \ldots, l_0$.

Combining (3.29) and (3.35), it is easily seen that there exists $\alpha_3 > 0$, such that for $\alpha \in (0, \alpha_3)$,

$$Y_{\alpha,l_0}(x) - Z_{\alpha,l_0}(x) < \gamma, \tag{3.36}$$

for $x = 1, 2, \ldots, l_0 - 1$. The proof is concluded by letting $\alpha_0 := \min(\alpha_1, \alpha_2, \alpha_3)$. ■

Propositions 3.1 and 3.2 yield the following

Proposition 3.3 *Assume that $x_0 < l_0$. Then, there exists $\alpha_0 > 0$ such that for all $\alpha \in (0, \alpha_0)$, $U_\alpha(x) = \mathbf{1}(x \geq l_0)$ for $x \geq 1$.*

We draw the reader's attention on the fact that Proposition (3.3) *does not* solve the discounted cost problem since (1) it does not cover the case when $l_0 = x_0$ and (2) it does not say how to compute α_0. However, Proposition (3.3) will turn out to be sufficient for solving the long-run average cost problem, as shown in the next section.

Remark 3.1 According to Step (1) of our "algorithm" at the beginning of Section 3, $\tilde{Y}_L(x)$ is obtained by assuming that the threshold policy U_L is optimal (i.e., $V_\alpha(x,0,0) = \tilde{Y}_L(x)$). This clearly motivates the definition of $\tilde{Y}_L(x)$ for $x = 1, 2, \ldots, L$ (see (2.12) and (3.1)). The expression of $\tilde{Y}_L(x)$ for $x \geq L+1$ is obtained by introducing (2.15) into (2.12) (still assuming that U_L is optimal), which gives for $x \geq L$,

$$(\alpha + \lambda + \nu)\, V_\alpha(x,0,0) = \nu x + \lambda\, V_\alpha(x+1,0,0) + \nu a_\alpha\, \beta_{\alpha,1}^x + \frac{\nu^2}{\alpha}\, x + \frac{\nu^2}{\alpha^2}\, (\lambda - \mu) + \nu\, \gamma, \tag{3.37}$$

where a_α is the coefficient of $\beta_{\alpha,1}^x$ in (2.15). Again, we have a difference equation, of which the general solution is given by

$$V_\alpha(x,0,0) = K_0\, \beta_{\alpha,1}^x + K_1 \left(\frac{\alpha+\lambda+\nu}{\lambda}\right)^x + K_2\, x + K_3, \ \text{ for } x \geq L. \tag{3.38}$$

Since $V_\alpha(x,0,0)/x$ must be uniformly bounded in x, we deduce from (3.38) that necessarily $K_1 = 0$. The remaining constants K_0, K_2 and K_3 are easily identified by pluging (3.38) into (3.37), which yields

$$V_\alpha(x,0,0) = a_\alpha\, C_\alpha\, \beta_{\alpha,1}^x + \frac{\nu}{\alpha}\, x + K_\alpha, \ \text{ for } x \geq L, \tag{3.39}$$

where C_α and K_α are defined in (3.7) and (3.8), respectively, which now motivates the definition of $\tilde{Y}_L(x)$ for $x \geq L$.

Remark 3.2 If

$$0 \leq \gamma < \frac{a+1}{a\,(1-\rho)}, \tag{3.40}$$

then for α small enough $U_\alpha(x) = 1$. This result follows from (3.25) by noting that the condition (3.40) is equivalent to $P(1) > 0$.

4 The Long-Run Average Cost Problem

In this section we shall discuss the long-run average cost problem (1.1) and we shall establish the optimality of a threshold policy.

Since $V_\alpha^U(x,y,z)$ is well defined for all $(x,y,z) \in \mathbf{S}$, $U \in \mathcal{U}$ (see (2.1)), we know from a Tauberian theorem (Widder (1941, pp. 181-182)) that

$$\text{limsup}_{\alpha\downarrow 0}\, \alpha\, V_\alpha^U(x,y,z) \leq \text{limsup}_{T\uparrow\infty}\, T^{-1}\, W_T^U(x,y,z), \tag{4.1}$$

for all $(x,y,z) \in \mathbf{S}$, $U \in \mathcal{U}$. Further, if $\lim_{T\uparrow\infty} T^{-1}\, W_T^U(x,y,z)$ exists then $\lim_{\alpha\downarrow 0} \alpha\, V_\alpha^U(x,y,z)$ exists as well, and

$$W^U(x,y,z) = \lim_{\alpha\downarrow 0} \alpha\, V_\alpha^U(x,y,z), \tag{4.2}$$

for all $(x,y,z) \in \mathbf{S}$, $U \in \mathcal{U}$.

We know from a standard result from Markov chain theory that (4.2) holds whenever $(X_n, Y_n, Z_n)_n$ is an ergodic Markov chain. On the other hand, $(X_n, Y_n, Z_n)_n$ is an ergodic Markov chain when a threshold policy with a finite threshold is used and $\rho < 1$. Hence,

$$\text{(4.2) holds when a threshold policy with a finite threshold is used.} \tag{4.3}$$

Fix $x \geq 1$ and let $U \in \mathcal{U}$ be an arbitrary policy. Two cases need be distinguished:

Case 1: $x_0 < l_0$.

¿From Proposition 3.3 we have

$$V_\alpha^{U_{l_0}}(x,0,1) \leq V_\alpha^U(x,0,1), \tag{4.4}$$

for all $\alpha \in (0, \alpha_0)$. Hence, cf. (4.1),

$$\text{limsup}_{\alpha\downarrow 0} \alpha\, V_\alpha^{U_{l_0}}(x,0,1) \leq W^U(x,0,1). \tag{4.5}$$

Combining now (4.3) and (4.5) yields

$$W^{U_{l_0}}(x,0,1) \leq W^U(x,0,1). \tag{4.6}$$

Case 2: $x_0 = l_0$.

¿From now on, $V_\alpha^U(x,0,1)$, $W^U(x,0,1)$, x_0 and l_0 are considered as functions of the parameter γ, and denoted as $V_\alpha^{U,\gamma}(x,0,1)$, $W^{U,\gamma}(x,0,1)$, x_0^γ and l_0^γ, respectively.

Let λ, μ, ν and $\gamma = \gamma_0$ be such that $x_0^{\gamma_0} = l_0^{\gamma_0}$. Since the mapping $\gamma \to x_0^\gamma$ is strictly increasing and continuous in $[0,\infty)$, cf. (3.18), we see that there exists $H > 0$ such that

$$l_0^{\gamma_0} - 1 < x_0^{\gamma_0 - h} < l_0^{\gamma_0}, \tag{4.7}$$

for $0 < h < H$. Therefore, Proposition 3.3 applies to the parameters λ, μ, ν and $\gamma_0 - h$ for $0 < h < H$, which implies that for $h \in (0, H)$,

$$V_\alpha^{U_{l_0^{\gamma_0}}, \gamma_0 - h}(x,0,1) \leq V_\alpha^{U,\gamma_0 - h}(x,0,1) \leq V_\alpha^{U,\gamma_0}(x,0,1), \tag{4.8}$$

for $\alpha \in (0, \alpha_0(h))$, where the second inequality follows from the fact that the mapping $\gamma \to V_\alpha^{U,\gamma}(x,0,1)$ is nondecreasing in $[0,\infty)$ for $\alpha > 0$. Using again (4.3) together with (4.8) gives

$$W^{U_{l_0^{\gamma_0}},\gamma_0-h}(x,0,1) \le W^{U,\gamma_0}(x,0,1), \tag{4.9}$$

for $0 < h < H$, which in turn yields

$$W^{U_{l_0^{\gamma_0}},\gamma_0}(x,0,1) \le W^{U,\gamma_0}(x,0,1) \tag{4.10}$$

by using the continuity of the mapping $\gamma \to W^{U_{l_0^{\gamma_0}},\gamma}(x,0,1)$.

This shows the optimality of a threshold policy.

Let us now compute the optimal long-run average cost. Assume first that $x_0 < l_0$. Because $\lim_{\alpha\downarrow 0} \alpha V_\alpha(x,0,1)$ does not depend on x, this limit can be obtained (in particular) from (3.23) and (A.14). Therefore, the optimal long-run average cost $W^{U_{l_0}}$ is given by

$$W^{U_{l_0}} = \left(\frac{(l_0-1)\,l_0}{2a} + \left(\frac{\rho}{a\,(1-\rho)} + 1\right)(l_0+a) + \gamma\,(1-\rho)\right)\left(\frac{\lambda}{l_0+a}\right). \tag{4.11}$$

By using now the continuity of the mapping $\gamma \to W^{U_{l_0^{\gamma}},\gamma}(x,y,z)$ (cf. Appendix C) we deduce that (4.11) also gives the optimal long-run average cost when $x_0 = l_0$.

The results of this paper are collected in the following proposition (we relax the assumption that $h/\nu = 1$):

Proposition 4.1 *Assume that $\lambda < \mu$. Then, there exists a threshold policy that solves the long-run average cost problem (1.1). The long-run average cost corresponding to the optimal policy is given by*

$$W^{U_{l_0}} = \left(\frac{l_0\,(l_0-1)}{2(l_0+a)} + \frac{\rho}{(1-\rho)} + a\right)h + \gamma\,\frac{\lambda\,(1-\rho)}{l_0+a}, \tag{4.12}$$

where the optimal threshold l_0 is given in (3.19) once γ is substituted for $\gamma\nu/h$ in (3.18).

Remark 4.1 It may be checked from (4.12) that the policy U_{l_0+1} is also optimal whenever $x_0 = l_0$ (i.e., $W^{U_{l_0}} = W^{U_{l_0+1}}$). Note that this property can directly be obtained by considering $\gamma + h$ instead of $\gamma - h$ in the proof of Proposition 4.1.

Remark 4.2 ¿From (1.2) and the definition of the cost c (see Section 2) one immediately deduces from (4.12) that $(L+a)/(\lambda\,(1-\rho))$ is the average return time to an empty queue when the threshold policy U_L is used, for any $L \ge 1$. Similarly, one observes that $L\,(L-1)/(2\,(L+a)) + \rho/(1-\rho) + a$ is the expected queue length under policy U_L, for any $L \ge 1$.

A Appendix

In this appendix we establish formulas (3.23) and (3.24).

To do so, the following power series expansions are needed:

$$\frac{\nu}{\alpha+\nu} = 1-\frac{\alpha}{\nu}+\frac{\alpha^2}{\nu^2}+o\left(\alpha^2\right); \tag{A.1}$$

$$\beta^x_{\alpha,1} = 1+\frac{x\alpha}{\lambda-\mu}+\left(\frac{(x-1)x}{2(\lambda-\mu)^2}-\frac{x\mu}{(\lambda-\mu)^3}\right)\alpha^2+o\left(\alpha^2\right); \tag{A.2}$$

$$\left(\frac{\alpha+\lambda}{\lambda}\right)^x = 1+\frac{x}{\lambda}\,\alpha+\frac{(x-1)x}{2\,\lambda^2}\,\alpha^2+o\left(\alpha^2\right); \tag{A.3}$$

$$C_\alpha = 1+\frac{\mu}{(\lambda-\mu)\nu}\,\alpha+\left(\frac{(\lambda\mu-\lambda\nu-\mu^2)\,\mu}{\nu^2\,(\lambda-\mu)^3}\right)\alpha^2+o\left(\alpha^2\right), \tag{A.4}$$

where $\lim_{\alpha\downarrow 0} o(\alpha^i)/\alpha^i = 0$ for any $i \in \mathbb{N}^*$.

¿From (A.1)-(A.4), we readily obtain that

$$\frac{\nu\mu\,(1-\beta^x_{\alpha,1})}{\alpha^2} = \frac{\nu\lambda\,a_1}{\alpha}-\frac{\nu\mu\,x}{(\lambda-\mu)^2}\left(\frac{x-1}{2}-\frac{\mu}{\lambda-\mu}\right)+O(\alpha); \tag{A.5}$$

$$\left(\frac{\alpha+\lambda}{\lambda}\right)^x-\beta^x_{\alpha,1} = a_1\,\alpha+a_2\,\alpha^2+o\left(\alpha^2\right); \tag{A.6}$$

$$\frac{\nu\mu}{\alpha^2}\left(\beta^L_{\alpha,1}\,C_\alpha-\frac{\nu}{\alpha+\nu}\right)+\frac{\nu\,\gamma}{\alpha+\nu} = -\frac{\nu\lambda\,b_1}{\alpha}-\nu\mu\left(b_2+\frac{1}{\nu^2}-\frac{(L-1)L}{2\lambda^2}\right)+\gamma+O(\alpha); \tag{A.7}$$

$$\left(\frac{\alpha+\lambda}{\lambda}\right)^L-\beta^L_{\alpha,1}\,C_\alpha = b_1\,\alpha+b_2\,\alpha^2+o\left(\alpha^2\right); \tag{A.8}$$

$$\frac{\left(\frac{\alpha+\lambda}{\lambda}\right)^x-\beta^x_{\alpha,1}}{\left(\frac{\alpha+\lambda}{\lambda}\right)^L-\beta^L_{\alpha,1}\,C_\alpha} = \frac{a_1}{b_1}+\left(\frac{a_2}{b_1}-\frac{a_1\,b_2}{b_1^2}\right)\alpha+o\left(\alpha\right), \tag{A.9}$$

with

$$a_1 := \frac{\mu\,x}{\lambda\,(\mu-\lambda)}; \tag{A.10}$$

$$a_2 := -\left(\frac{x-1}{2\,(\lambda-\mu)^2}-\frac{\mu}{(\lambda-\mu)^3}-\frac{x-1}{2\,\lambda^2}\right)x; \tag{A.11}$$

$$b_1 := \left(\frac{\mu}{\mu-\lambda}\right)\left(\frac{L}{\lambda}+\frac{1}{\nu}\right); \tag{A.12}$$

$$b_2 := -\left(\frac{(L-1)L}{2\,(\lambda-\mu)^2}-\frac{L\mu}{(\lambda-\mu)^3}+\frac{\mu(\lambda\mu-\lambda\nu-\mu^2)}{\nu^2\,(\lambda-\mu)^3}+\frac{L\mu}{\nu\,(\lambda-\mu)^2}-\frac{(L-1)L}{2\lambda^2}\right). \tag{A.13}$$

Using (3.22), (A.5)-(A.13) we obtain after straightforward algebra

$$
\begin{aligned}
Y_{\alpha,L}(x) - Z_{\alpha,L}(x) &= \left(\frac{\nu\, x}{L+\dfrac{\lambda}{\nu}}\right)\left(\frac{(L-1)\,L\mu}{2\lambda\,(\mu-\lambda)}+\frac{\mu\, L}{\nu\,(\mu-\lambda)}+\frac{\mu\,\lambda}{\nu^2\,(\mu-\lambda)}+\frac{\gamma}{\nu}\right)\\
&- \frac{\nu\mu\,(x-1)\,x}{2\lambda\,(\mu-\lambda)}+O(\alpha),\\
&= \left(\frac{x}{L+a}\right)\left(\frac{(L-1)\,L}{2a\,(1-\rho)}+\frac{L}{1-\rho}+\frac{a}{1-\rho}+\gamma\right)-\frac{(x-1)\,x}{2a\,(1-\rho)}+O(\alpha),\\
&= \left(\frac{L^2+2a\,L-x\,(L+a)+2a^2+2a\,\gamma\,(1-\rho)+a}{L+a}\right)\left(\frac{x}{2a\,(1-\rho)}\right)+O(\alpha),
\end{aligned}
$$

which proves (3.24).

On the other hand, from (3.21) and the series expansions (cf. (A.1)-(A.4))

$$
\begin{aligned}
\frac{1}{\left(\dfrac{\alpha+\lambda}{\lambda}\right)^L-\beta^L_{\alpha,1}\,C_\alpha} &= \frac{1}{b_1\,\alpha}-\frac{b_2}{b_1^2}+O(\alpha);\\
\beta^L_{\alpha,1}\,C_\alpha-\frac{\nu}{\alpha+\nu} &= -\frac{\lambda b_1\,\alpha}{\mu}-\left(b_2+\frac{1}{\nu^2}-\frac{(L-1)\,L}{2\lambda^2}\right)\alpha^2+o(\alpha^2),
\end{aligned}
$$

we get that

$$
Y_{\alpha,L}(1)=\left(\frac{(L-1)\,L}{2a}+\left(\frac{\rho}{a\,(1-\rho)}+1\right)(L+a)+\gamma\,(1-\rho)\right)\left(\frac{\lambda}{L+a}\right)\alpha^{-1}+O(1), \tag{A.14}
$$

which establishes (3.23).

B Appendix

We show that

$$
Q(l_0)-Q(l_0-1)<0, \tag{B.1}
$$

when $x_0 < l_0$.

¿From the definition of $P(w)$ and $Q(w)$ (cf. (3.26) and (3.33), respectively), it is easily seen that (B.1) holds if and only if

$$
P(l_0)<2\,(l_0+a). \tag{B.2}
$$

Let us prove (B.2). We have:

$$
\begin{aligned}
P(l_0) &= P(l_0)-P(x_0),\\
&= (l_0-x_0)(l_0+x_0+2a+1),\\
&= \Delta\,(2l_0-\Delta+2a+1):=f(\Delta),
\end{aligned}
$$

where $\Delta := l_0 - x_0$ with $0 < \Delta < 1$.

The proof is now completed by observing that the mapping $\Delta \to f(\Delta)$ is nondecreasing in $[0,1]$ (since $l_0 \geq 1$) and by noting that $f(1) = 2\,(l_0 + a)$.

C Appendix

We show in this appendix that the mapping $\gamma \to W^{U_{l_0^\gamma},\gamma}(x,y,z)$ is continuous in $[0,\infty)$.

Observe first from (4.3) and (1.3) that the mapping $\gamma \to W^{U_L,\gamma}(x,y,z)$ is nondecreasing (referred below to as property (P1)) and continuous (referred below to as property (P2)) in $[0,\infty)$ for $L < \infty$.

Let $0 \leq \gamma_1 < \gamma_2 < \infty$. Then,

$$W^{U_{l_0^{\gamma_1}},\gamma_1}(x,y,z) \leq W^{U_{l_0^{\gamma_2}},\gamma_1}(x,y,z) \leq W^{U_{l_0^{\gamma_2}},\gamma_2}(x,y,z) \leq W^{U_{l_0^{\gamma_1}},\gamma_2}(x,y,z), \tag{C.1}$$

where we have used (P1) to establish the second inequality.

Using now (P2) and (C.1) we obtain that

$$\lim_{\gamma_2 \to \gamma_1} W^{U_{l_0^{\gamma_2}},\gamma_2}(x,y,z) = W^{U_{l_0^{\gamma_1}},\gamma_1}(x,y,z),$$

which concludes the proof.

References

Altman, E., and P. Nain. 1993. Optimal Control of an M/G/1 Queue with Repeated Vacations. *IEEE Trans. Aut. Cont.*, Vol. 38, No. 12, Dec. 1993, 1766-1775.

Bertsekas, D. P. 1987. *Dynamic Programming. Deterministic and Stochastic Models*. Prentice Hall, Inc. Englewoods Cliff.

Brémaud, P. 1981. *Point Processes and Queues*. Springer Verlag, New York.

Blanc, J. P. C., de Waal, P., Nain, P., and D. Towsley. 1992. Optimal control of admission to a multi-server queue with two arrival streams. *IEEE Trans. Aut. Cont.*, Vol. 37, No. 6, Jun. 1992, 785-797.

Chung, K. L. 1967. *Markov Chains with Stationary Transition Probabilities*. 2nd. Ed., Springer Verlag, New York.

Doshi, B. 1986. Queueing Systems with Vacations - A Survey. *Queueing Syst.* **1**, 29-66.

Doshi, B. 1986. "Conditional and Unconditional Distributions for M/G/1 Type Queues with Server Vacations," *Queueing Syst.* **7**, 229-252.

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

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

Gelenbe, E., and R. Iasnogorodski. 1980. "A Queue with Server of Walking Type (Autonomous Service). *Ann. Inst. Henry Poincaré* **XVI**, 63-73.

Heyman, D. P., and M. J. Sobel. 1984. *Stochastic Models in Operations Research*, Volume II, Stochastic Optimization. McGraw-Hill Book Company, New York.

Kella, O. 1989. The Threshold Policy in the M/G/1 Queue with Server Vacations. *Naval Res. Quat.* **33**, 111-123.

Kella, O. 1990. Optimal Control of the Vacation Scheme in an M/G/1 Queue. *Opns. Res.* **38**, 724-728.

Lee, H.-S., and M. M. Srinivasan. 1989. Control Policies for the M^x/G/1 Queueing System. *Mgmt Sci.* **35**, 708-721.

Levy, H., and U. Yechiali. 1975. Utilization of Idle Time in an M/G/1 Queueing System. *Mgmt Sci.* **22**, 202-211.

Lippman, S. A. 1975. On Dynamic Programming with Unbounded Rewards. *Mgmt Sci.* **21**, 1225-233.

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

Schäl, M. 1975. Conditions for optimality in dynamic programming and for the limit of the n-stage optimal policies to be optimal. *Z. Wahrscheinlichkeitsth.* **32**, 179-196.

Sennott, L. I. 1989. Average Cost Semi-Markov Decision Processes and the Control of Queueing System. *Prob. in the Engineering and the Informational Sci.* **3**, 247-272.

Talman, A. J. J. 1979. A Simple Proof of the Optimality of the Best N-Policy in the M/G/1 Queueing Control Problem with Removable Server. *Statistica Neerlandica* **33**, 143-150.

Takagi, H. 1986. *Analysis of Polling Systems.* MIT Press, Cambridge (Mass.).

Teghem J. Jr. 1986. Control of the Service Process in a Queueing System. *European J. of Opns. Res.* **23**, 141-158, 1986.

Widder, D. V. 1941. *The Laplace Transform.* Princeton University Press, Princeton.

Yadin, M., and P. Naor. 1963. Queueing Systems with a Removable Service Station. *Opns. Res. Quat.* **14**, 393-405.