*J. Appl. Prob.* **34**, 773–784 (1997)
*Printed in Israel*


# AN $M/G/1$ QUEUE WITH MULTIPLE TYPES OF FEEDBACK AND GATED VACATIONS

ONNO J. BOXMA,* *CWI*
URI YECHIALI,** *Tel Aviv University*

**Abstract**

This paper considers a single-server queue with Poisson arrivals and multiple customer feedbacks. If the first service attempt of a newly arriving customer is not successful, he returns to the end of the queue for another service attempt, with a different service time distribution. He keeps trying in this manner (as an 'old' customer) until his service is successful. The server operates according to the 'gated vacation' strategy; when it returns from a vacation to find $K$ (new and old) customers, it renders a single service attempt to each of them and takes another vacation, etc. We study the joint queue length process of new and old customers, as well as the waiting time distribution of customers. Some extensions are also discussed.

$M/G/1$ QUEUE; FEEDBACK; GATED VACATIONS; QUEUE LENGTH

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60K25
SECONDARY 90B22

## 1. Introduction

This paper is dedicated to Professor Joe Gani. The applied probability community is forever indebted to him for this vigorous and unselfish promotion of applied probability as a scientific discipline.

Consider a dedicated editor-in-chief who devotes sessions to handling tasks for his scientific journal. He distinguishes between new and old tasks. New tasks are papers/reports that he has received from his managing editor since the start of the previous working session. He decides on acceptance, or on the choice of a referee, etc. Some of the new tasks cannot yet be handled, as the editor-in-chief needs to 'sleep on it' or to get additional information from his managing editor. Such deferred tasks will be treated in the next session as old tasks. At the end of the session the editor sends a message to the managing editor and takes a vacation until a reply comes. This situation can be modeled by a feedback queue with so-called gated vacations — which is the subject of this paper.

Received 10 October 1996.
* Postal address: P.O. Box 94079, 1090 GB Amsterdam, The Netherlands.
** Postal address: Department of Statistics, School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel.
This paper is based on an invited presentation at the Applied Probability conference in honour of J. Gani and T. E. Hannan, held in Athens in March 1995.

Other situations modeled by such a feedback queue are that of a machine processing parts, X-ray taking, and the like. After a session, the processed parts (or X-rays) are inspected and it is decided which have to be redone. In the model studied in this paper we allow different service time distributions for new and old tasks, and for successful and unsuccessful tasks. This enables us, for example, to model a situation where a job can be processed either in one session, or be divided over two (or more) sessions. Our model contains the ordinary $M/G/1$ queue with gated vacations as a special case (cf. Takagi [8]); it also encompasses the $M/G/1$ queue with Bernoulli feedback (cf. Takács [7]), and its generalization to an $M/G/1$ queue with Bernoulli feedback and gated vacations (all service time distributions being the same), cf. Takine *et al.* [9].

The paper is organized as follows. The model is described in detail below. In Section 2 we determine the joint distribution of the number of old and new customers at session beginnings. The steady-state joint queue length distribution is studied in Section 3, and the waiting time distribution in Section 4.

*The model.* Customers arrive at a single-server $M/G/1$-type queue according to a Poisson process with rate $\lambda$. A newly arriving customer receives a 'successful' service with probability $p$, or an 'unsuccessful' service with probability $1-p$. The duration of a successful service attempt is $S$ (with mean $s$, p.d.f. $S(\cdot)$ and LST $\tilde{S}(\cdot)$), while the duration of an unsuccessful service attempt is $U$ (mean $u$, p.d.f. $U(\cdot)$ and LST $\tilde{U}(\cdot)$). If a service attempt is successful the customer leaves the system, whereas if the attempt is unsuccessful, the customer is immediately fed back to the end of the queue. From that moment on, such a customer is an 'old' one. If a customer has been fed back *at least once* then, on each of his following service attempts, with probability $p_1$ he receives a successful service with duration $S_1$ (mean $s_1$, p.d.f. $S_1(\cdot)$, LST $\tilde{S}_1(\cdot)$), and with probability $1-p_1$ he receives an unsuccessful service $U_1$ (mean $u_1$, p.d.f. $U_1(\cdot)$ and LST $\tilde{U}_1(\cdot)$).

A generalization of the above model to $M$ types of service re-attempts is straightforward to formulate, and its analysis is not fundamentally more complicated; it will not be discussed in this paper.

The server operates according to the gated vacation strategy. When the server returns from a vacation to find, say, $N$ new customers and $O$ old ones, it starts a 'session' and renders each one of those customers a single service attempt (which may be successful or not) and then leaves for another vacation whose duration $V$ has mean $v$, p.d.f. $V(\cdot)$ and LST $\tilde{V}(\cdot)$. If the queue is empty upon returning from a vacation, the server immediately leaves for another vacation, etc. (Note that we have assumed that a customer immediately leaves after a successful service. In some applications it may be more realistic that a successful customer is only released after an inspection at the beginning of the server vacation. The analysis to be presented below can easily handle that.)

We impose all the usual independence assumptions between the arrival process, successful and unsuccessful services, and vacations. Given that a customer's first service attempt is unsuccessful (with duration $U$), let $Y$ be the number of his additional unsuccessful attempts *before* he leaves the system. Clearly, $\boldsymbol{P}(Y=n)=(1-p_1)^n p_1$, $n=0, 1, 2, \cdots$. Thus, the expected total amount of service provided to a customer by the time he successfully leaves the system is

$$ps+(1-p)\left\{u+E\left[\sum_{j=1}^{Y} U_{1j}\right]+s_1\right\}$$

where $U_{1j}$ are independent, identically distributed as $U_1$.

Since $EY=(1-p_1)/p_1$, by Wald's theorem, $E[\Sigma_{j=1}^{Y} U_{1j}]=[(1/p_1)-1]u_1$. Thus, the system load is

$$\rho := \lambda\left[ps+(1-p)\left\{u+\left(\frac{1}{p_1}-1\right)u_1+s_1\right\}\right]. \tag{1.1}$$

An interpretation of $\rho$ is as follows. Let $B_N=pS+(1-p)U$ denote the duration of a service attempt of a *new* customer (mean $b_N$, p.d.f. $B_N(\cdot)$, LST $\tilde{B}_N(\cdot)$), and let $B_O=p_1S_1+(1-p_1)U_1$ denote the duration of a service attempt of an *old* customer (mean $b_O$, p.d.f. $B_O(\cdot)$, LST $\tilde{B}_O(\cdot)$). Clearly,

$$b_N=ps+(1-p)u, \qquad b_O=p_1s_1+(1-p_1)u_1.$$

Each customer makes one service attempt (of mean duration $b_N$) as a *new* customer. With probability $(1-p)$ it is unsuccessful, and then it obtains an average of $1/p_1$ additional service attempts as an *old* customer, each attempt with mean duration $b_O$. Thus, the mean load of each customer on the system is $b_N+((1-p)/p_1)b_O$, and the overall system load $\rho$ is as obtained in (1.1). We assume that $\rho<1$.

*Remark.* In [3] polling models with customer routing are considered; this includes feedback queues, with different service time distributions at successive visits. That paper sketches a framework for analyzing joint queue length distributions in such cases. The approach is based on branching processes, which are basically also at the heart of the present paper.

## 2. The joint queue length distribution at session beginnings

Define:

$T_n :=$ the time instant when the $n$th vacation ends;

$N_n :=$ number of 'new' customers at $T_n$ (those customers have arrived during the time interval $(T_{n-1}, T_n]$);

$O_n :=$ number of 'old' customers at $T_n$.

The joint probability generating function (p.g.f.) of $N_n$ and $O_n$ is

$$G_n(z, w) := E[z^{N_n}w^{O_n}], \qquad n=1,2,\cdots, \ |z|\leq 1, \ |w|\leq 1. \tag{2.1}$$

We shall derive a recursive formula for $G_n(z, w)$ that will lead to the limiting distribution of $N_n$ and $O_n$. Let $X(N_n)$ (respectively $X_1(O_n)$, $A(V_n)$) denote the number of new arrivals during the service time of the $N_n$ 'new' customers (respectively the service time of the $O_n$ 'old' customers, and the $n$th vacation $V_n$). Let $F(N_n)$ (respectively $F_1(O_n)$) denote the number of customers fed back out of $N_n$ (respectively, $O_n$). Then,

$$N_{n+1} = X(N_n) + X_1(O_n) + A(V_n),$$
$$O_{n+1} = F(N_n) + F_1(O_n). \tag{2.2}$$

We claim the following.

*Proposition* 1. For $|z| \leqq 1$, $|w| \leqq 1$:

$$G_{n+1}(z, w) = \tilde{V}(\lambda(1-z))G_n(p\tilde{S}[\lambda(1-z)] + (1-p)w\tilde{U}[\lambda(1-z)],$$
$$p_1\tilde{S}_1[\lambda(1-z)] + (1-p_1)w\tilde{U}_1[\lambda(1-z)]). \tag{2.3}$$

*Proof.* If $N_n = j$ and $O_n = k$ then, with probability

$$\binom{j}{m} p^m(1-p)^{j-m} \binom{k}{m_1} p_1^{m_1}(1-p_1)^{k-m_1},$$

there will be $m$ type $S$ ($m_1$ type $S_1$) successful services out of $N_n = j$ (out of $O_n = k$) attempts, and $j-m$ type $U$ ($k-m_1$ type $U_1$) unsuccessful services.

The total service duration of the $N_n = j$ and $O_n = k$ customers will have a convolution p.d.f. $[S^{m*} * U^{(j-m)*} * S_1^{m_1*} * U_1^{(k-m_1)*}](\cdot)$. If that convoluted service lasts $t$ units of time, there will be $l$ new arrivals during that time with probability $e^{-\lambda t}(\lambda t)^l/l!$ ($l = 0, 1, 2, \cdots$). Thus,

$$E[z^{N_{n+1}}w^{O_{n+1}} \mid N_n = j, O_n = k]$$
$$= E[z^{A(V_n)}] \sum_{m=0}^{j} \sum_{m_1=0}^{k} \binom{j}{m} p^m(1-p)^{j-m}w^{j-m} \binom{k}{m_1} p_1^{m_1}(1-p_1)^{k-m_1}w^{k-m_1}$$
$$\times \sum_{l=0}^{\infty} z^l \int_{t=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^l}{l!}\, d[S^{m*} * U^{(j-m)*} * S_1^{m_1*} * U_1^{(k-m_1)*}](t) \tag{2.4}$$
$$= \tilde{V}[\lambda(1-z)][p\tilde{S}[\lambda(1-z)] + (1-p)w\tilde{U}[\lambda(1-z)]]^j$$
$$\times [p_1\tilde{S}_1[\lambda(1-z)] + (1-p_1)w\tilde{U}_1[\lambda(1-z)]]^k.$$

Now, taking expectation with respect to $N_n$ and $O_n$, we obtain equation (2.3). This completes the proof.

Let $N$ and $O$ be the random variables with as joint distribution, the joint *limiting distribution* of $N_n$ and $O_n$. Then, for $|z| \leqq 1$, $|w| \leqq 1$,

$$G(z, w) = E[z^N w^O] = \tilde{V}[\lambda(1-z)]G(p\tilde{S}[\lambda(1-z)] + (1-p)w\tilde{U}[\lambda(1-z)],$$
$$p_1\tilde{S}_1[\lambda(1-z)] + (1-p_1)w\tilde{U}_1[\lambda(1-z)]). \tag{2.5}$$

*Remark.* A fast way to understand (2.5) is to use the concept of branching processes. According to this concept one can interpret the factor $p\tilde{S}[\lambda(1-z)] + (1-p)w\tilde{U}[\lambda(1-z)]$ as follows. Define the 'descendants' of a customer to be the new arrivals during his service time. Then the number of descendants of a 'new' customer has generating function $p\tilde{S}[\lambda(1-z)]+(1-p)w\tilde{U}[\lambda(1-z)]$. The extra factor $w$ indicates the fact that, with probability $1-p$, the 'new' customer also generates one 'old' customer, which is he himself, being fed back. In the same way one can easily handle the case of $M$ different customer types. Formula (2.5) would keep the same structure, the $(k+1)$th element of $G$ in the right-hand side becoming (with an obvious extension of notation):

$$p_k\tilde{S}_k(\lambda(1-z))+(1-p_k)w_{k+1}\tilde{U}_k(\lambda(1-z)), \qquad k=1,\cdots,M-2.$$

We now solve (2.5) by iteration. For $|z|\leqq 1$, $|w|\leqq 1$, define

$$\begin{aligned} &m_N^{(0)}(z,w):=z, \qquad m_O^{(0)}(z,w):=w,\\ &m_N^{(1)}(z,w):=p\tilde{S}[\lambda(1-z)]+(1-p)w\tilde{U}[\lambda(1-z)],\\ &m_O^{(1)}(z,w):=p_1\tilde{S}_1[\lambda(1-z)]+(1-p_1)w\tilde{U}_1[\lambda(1-z)].\end{aligned} \tag{2.6}$$

For $n\geqq 1$,

$$\begin{aligned} m_N^{(n)}(z,w)&=p\tilde{S}[\lambda(1-m_N^{(n-1)}(z,w))]+(1-p)m_O^{(n-1)}(z,w)\tilde{U}[\lambda(1-m_N^{(n-1)}(z,w))],\\ m_O^{(n)}(z,w)&=p_1\tilde{S}_1[\lambda(1-m_N^{(n-1)}(z,w))]+(1-p_1)m_O^{(n-1)}(z,w)\tilde{U}_1[\lambda(1-m_N^{(n-1)}(z,w))].\end{aligned} \tag{2.7}$$

Note that $m_N^{(n)}(z,w)$ (respectively, $m_O^{(n)}(z,w)$) may be interpreted as the $n$th generation descendants of one new (respectively, old) customer. Formula (2.5) can be rewritten as

$$G(z,w)=\tilde{V}(\lambda(1-z))G(m_N^{(1)}(z,w),m_O^{(1)}(z,w)). \tag{2.8}$$

Iterating this formula $k$ times results in

$$G(z,w)=\prod_{n=0}^{k}\tilde{V}[\lambda(1-m_N^{(n)}(z,w))]G(m_N^{(k+1)}(z,w),m_O^{(k+1)}(z,w)).$$

It can be shown that, if $\rho<1$, $\lim_{n\to\infty} m_N^{(n)}(z,w)=1$ and $\lim_{n\to\infty} m_O^{(n)}(z,w)=1$, convergence is geometrically fast, so that $\Pi_{n=0}^{\infty}\tilde{V}[\lambda(1-m_N^{(n)}(z,w))]$ is a convergent infinite product. Hence, for $\rho<1$, as $G(1,1)=1$,

$$G(z,w)=\prod_{n=0}^{\infty}\tilde{V}[\lambda(1-m_N^{(n)}(z,w))], \qquad |z|\leqq 1,\ |w|\leqq 1. \tag{2.9}$$

*Moments.* From (2.2), using the Poisson arrival property, we readily obtain

$$\begin{aligned} &\boldsymbol{E}[N]=\lambda\{b_N\boldsymbol{E}[N]+b_O\boldsymbol{E}[O]+v\},\\ &\boldsymbol{E}[O]=(1-p)\boldsymbol{E}[N]+(1-p_1)\boldsymbol{E}[O].\end{aligned} \tag{2.10}$$

Thus,

$$\boldsymbol{E}[O]=\frac{1-p}{p_1}\boldsymbol{E}[N], \tag{2.11}$$

and hence, using (1.1),

$$E[N]=\rho E[N]+\lambda v. \tag{2.12}$$

This implies

$$E[N]=\frac{\lambda v}{1-\rho}. \tag{2.13}$$

Clearly, (2.12) can be directly obtained by using balance arguments.

Finally, the mean *cycle time*, i.e. the time between two consecutive vacation beginnings, is given by

$$E[C]=E[N]/\lambda=\frac{v}{1-\rho}. \tag{2.14}$$

## 3. The steady-state joint queue length distribution

$G(z, w)$ is the p.g.f. of the steady-state joint queue length distribution at the *beginning* of a session (end of a vacation). Let $G^{(e)}(z, w)$ denote the corresponding p.g.f. at the *end* of a session (beginning of a vacation). Also, let $H_N(z, w)$ and $H_O(z, w)$ be the joint queue length p.g.f. at the *beginning* of a service attempt of a new (respectively, old) customer. Similarly, let $H_N^{(e)}(z, w)$ and $H_O^{(e)}(z, w)$ be the corresponding p.g.f.s right after the *end* of a service. Considering one cycle (service plus vacation), the p.g.f. of the state of the system at an arbitrary moment, $F^*(z, w)$, is given by

$$\begin{aligned} F^*(z, w) = \frac{1}{E[C]} & [b_N E[N] F^*(z,w \mid \text{a new customer is served}) \\ & + b_O E[O] F^*(z,w \mid \text{an old customer is served}) \\ & + v F^*(z, w \mid \text{the server is on vacation})]. \end{aligned} \tag{3.1}$$

Now,

$$F^*(z, w \mid \text{the server is on vacation}) = G^{(e)}(z, w)\frac{1-\tilde{V}[\lambda(1-z)]}{v(\lambda(1-z))}, \tag{3.2}$$

where $(1-\tilde{V}(s))/vs$ is the LST of the past part, $V_P$, of $V$;

$$F^*(z, w \mid \text{a new customer is served}) = H_N(z, w)\frac{1-\tilde{B}_N[\lambda(1-z)]}{b_N(\lambda(1-z))}, \tag{3.3}$$

$$F^*(z, w \mid \text{an old customer is served}) = H_O(z, w)\frac{1-\tilde{B}_O[\lambda(1-z)]}{b_O(\lambda(1-z))}. \tag{3.4}$$

Substituting (3.2), (3.3) and (3.4) in (3.1), and then using

$$G(z, w) = G^{(e)}(z, w)\tilde{V}[\lambda(1-z)], \tag{3.5}$$

we obtain

$$F^*(z,w)=\frac{1}{E[C]\lambda(1-z)}[[E[N](1-\tilde{B}_N[\lambda(1-z)])H_N(z,w)$$

$$+E[O](1-\tilde{B}_O[\lambda(1-z)])H_O(z,w)]$$

$$+(1-\tilde{V}[\lambda(1-z)])G(z,w)/\tilde{V}[\lambda(1-z)]]. \tag{3.6}$$

The steady-state joint queue length distribution of old and new customers depends on the order in which customers are served during a session. We now assume that during each session of the server, new customers have priority over old ones, i.e. they are served first. Thus, if $G(z, w)$ is the p.g.f. of the number of customers at the beginning of a session, let $G_N^{(e)}(z, w)$ be the p.g.f. of the number of customers at the moment where the server ends serving *new* customers. We now use a nice idea that seems to be due to Eisenberg [5] in a polling context; see also Altman and Yechiali [1], Borst and Boxma [2] and Sidi *et al.* [6]. It is the observation that the union of all session beginnings and all service endings of new customers coincides with the union of all service beginnings of new customers and endings of whole service periods of new customers. In terms of generating functions:

$$G(z,w)+E[N]H_N^{(e)}(z,w)=G_N^{(e)}(z,w)+E[N]H_N(z,w). \tag{3.7}$$

Similarly, with $G_O(z, w)$ denoting the p.g.f. of the number of customers at the moment where the server starts serving *old* customers:

$$G_O(z,w)+E[O]H_O^{(e)}(z,w)=G^{(e)}(z,w)+E[O]H_O(z,w). \tag{3.8}$$

Above we might as well have written $G_N^{(e)}(z, w)$ instead of $G_O(z, w)$; note that during a session the period in which the server serves new (or old) customers may have length zero, but this does not cause complications. Now

$$G_O(z,w)=G_N^{(e)}(z,w)=E[(m_N^{(1)}(z,w))^N w^O]=G(m_N^{(1)}(z,w),w). \tag{3.9}$$

Using (3.5), (3.9) and the obvious relations

$$H_N^{(e)}(z,w)=H_N(z,w)m_N^{(1)}(z,w)/z, \tag{3.10}$$

$$H_O^{(e)}(z,w)=H_O(z,w)m_O^{(1)}(z,w)/w, \tag{3.11}$$

equations (3.7) and (3.8) are written as

$$E[N]H_N(z,w)[1-m_N^{(1)}(z,w)/z]=G(z,w)-G(m_N^{(1)}(z,w),w), \tag{3.12}$$

$$E[O]H_O(z,w)[1-m_O^{(1)}(z,w)/w]=G(m_N^{(1)}(z,w),w)-G(z,w)/\tilde{V}(\lambda(1-z)). \tag{3.13}$$

Substituting $E[N]H_N(z, w)$ from (3.12) and $E[O]H_O(z, w)$ from (3.13) in equation (3.6) gives the result we were looking for: the p.g.f. $F^*(z, w)$ of the joint steady-state queue length distribution, expressed in terms of the known $G(\cdot,\cdot)$.

Briefly consider the case where old customers have priority over new ones. Let $G_O^{(e)}(z, w)=G_N(z, w)$ be the p.g.f. of the joint queue length distribution at the moment where the server ends serving old customers and starts serving new customers. Then (3.7) and (3.8) are modified into

$$G(z, w) + E[O]H_O^{(e)}(z, w) = G_O^{(e)}(z, w) + E[O]H_O(z, w), \tag{3.14}$$

$$G_N(z, w) + E[N]H_N^{(e)}(z, w) = G^{(e)}(z, w) + E[N]H_N(z, w). \tag{3.15}$$

Instead of (3.9) we have

$$G_O^{(e)}(z, w) = G_N(z, w) = E[z^N(m_O^{(1)}(z, w))^O] = G(z, m_O^{(1)}(z, w)). \tag{3.16}$$

Similar manipulations as above and finally substitution in (3.6) gives $F^*(\cdot, \cdot)$ in terms of $G(\cdot, \cdot)$.

## 4. Waiting times

The waiting time distribution obviously depends on the order of service. We assume that the order of service follows the 'longest present first' rule. That is, old customers are served before new customers and within each group the order is FCFS. We shall express the LST of the waiting time into the just obtained function $G(\cdot, \cdot)$.

For the calculation of the LST of the waiting time, $W$, of an arbitrary customer, we consider here the case where $p_1 = 1$. That is, a customer is either successful on his first service attempt or, if fed back after the unsuccessful service, his second service attempt is *always* successful (having duration $S_1$). More involved cases can in principle be handled in a similar manner, but the analysis may become quite complicated.

Let $W_S$ ($W_U$) denote the waiting time of a customer who is successful (unsuccessful) on his initial service attempt. Then,

$$P(W \leqq t) = pP(W_S \leqq t) + (1-p)P(W_U \leqq t). \tag{4.1}$$

4.1. *Calculation of the distribution of* $W_S$. Consider a tagged successful customer $K$. We can write $W_S = C_R + D_1 + D_2$. Here $C_R$ denotes the residual cycle time, measured from the moment of the arrival of $K$. $D_1$ is the sum of the service times of those $F(N)$ customers, out of $N$ present at the *start* of the cycle, who were fed back during the cycle (the old customers were successful and left the system). Each of those customers now receives a successful service $S_1$. $D_2$ is the sum of service times of *new* arrivals during the *past* part, $C_P$, of the cycle (before the arrival of $K$). Each of those new customers requires a service duration $B_N$. Thus,

$$\begin{aligned} E[e^{-\omega D_2} \mid C_P = t] &= E[E[e^{-\omega D_2} \mid A(t)] \mid C_P = t] \\ &= E[(\tilde{B}_N(\omega))^{A(t)} \mid C_P = t] = e^{-\lambda(1-\tilde{B}_N(\omega))t}, \end{aligned} \tag{4.2}$$

where $\tilde{B}_N(\omega) = p\tilde{S}(\omega) + (1-p)\tilde{U}(\omega)$, and $A(t)$ is the number of Poisson arrivals during a time interval of length $t$.

Also,

$$E[e^{-\omega D_1} \mid F(N)] = [\tilde{S}_1(\omega)]^{F(N)}. \tag{4.3}$$

Combining the above, we have

$$E[e^{-\omega W_S}] = \int_{t_P=0}^{\infty}\int_{t_R=0}^{\infty} dP\{C_P \leqq t_P,\, C_R \leqq t_R\} e^{-\omega t_R} e^{-\lambda(1-\tilde{B}_N(\omega))t_P}$$

(4.4)
$$\times \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{k} P(N=k,\, O=j,\, F(N)=k-l \mid C_P=t_P,\, C_R=t_R)\tilde{S}_1(\omega)^{k-l}.$$

It follows from renewal theory (cf. formula (I.6.23) of [4]) that one can write

$$\int_{t_P=0}^{\infty}\int_{t_R=0}^{\infty} g(t_P, t_R)dP\{C_P \leqq t_P,\, C_R \leqq t_R\}$$
$$= \int_{x=0}^{\infty}\int_{t_R=0}^{x} g(x-t_R, t_R)\frac{dt_R}{x}\frac{xdP(C \leqq x)}{E[C]}.$$

Hence,

$$E[e^{-\omega W_S}] = \frac{1}{E[C]}\int_{x=0}^{\infty}\int_{t_R=0}^{x} dP(C \leqq x) e^{-\omega t_R}\exp[-\lambda(1-\tilde{B}_N(\omega))(x-t_R)]dt_R$$

(4.5)
$$\times \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{k} P(N=k,\, O=j,\, F(N)=k-l \mid C=x)\tilde{S}_1(\omega)^{k-l}.$$

Using $P(F(N)=k-l \mid N=k,\, O=j)=\binom{k}{l}p^l(1-p)^{k-l}$ and performing the integration in (4.5) with respect to $t_R$ we have, letting $\xi(\omega) \equiv \lambda(1-\tilde{B}_N(\omega))$,

$$E[e^{-\omega W_S}] = \frac{1}{E[C]}\frac{1}{\omega-\xi(\omega)}\int_{x=0}^{\infty}(e^{-\xi(\omega)x}-e^{-\omega x})$$
$$\times \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{k} P(N=k,\, O=j)\binom{k}{l}p^l(1-p)^{k-l}\tilde{S}_1(\omega)^{k-l}$$
$$\times dP(C \leqq x \mid N=k,\, O=j,\, F(N)=k-l)$$
$$= \frac{1}{E[C]}\frac{1}{\omega-\xi(\omega)}\int_{x=0}^{\infty}(e^{-\xi(\omega)x}-e^{-\omega x})\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{k} P(N=k,\, O=j)$$
$$\times \tilde{S}_1(\omega)^{k-l}\binom{k}{l}p^l(1-p)^{k-l}d[S^{l*} * U^{(k-l)*} * S_1^{j*} * V](x)$$
$$= \frac{1}{E[C]}\frac{1}{\omega-\xi(\omega)}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} P(N=k,\, O=j)$$
$$\times \sum_{l=0}^{k}\{[\tilde{S}(\xi(\omega))]^l[\tilde{U}(\xi(\omega))]^{k-l}[\tilde{S}_1(\xi(\omega))]^j\tilde{V}(\xi(\omega))$$
$$-[\tilde{S}(\omega)]^l[\tilde{U}(\omega)]^{k-l}[\tilde{S}_1(\omega)]^j\tilde{V}(\omega)\}\binom{k}{l}p^l(1-p)^{k-l}\tilde{S}_1(\omega)^{k-l}.$$

Thus,

$$E[\mathrm{e}^{-\omega W_S}] = \frac{1}{E[C]}\frac{1}{\omega-\xi(\omega)}$$

$$\times [\tilde{V}(\xi(\omega))G(p\tilde{S}(\xi(\omega))+(1-p)\tilde{U}(\xi(\omega))\tilde{S}_1(\omega),\ \tilde{S}_1(\xi(\omega))) \tag{4.6}$$

$$-\tilde{V}(\omega)G(p\tilde{S}(\omega)+(1-p)\tilde{U}(\omega)\tilde{S}_1(\omega),\ \tilde{S}_1(\omega))].$$

*Remark.* For $p=1$, i.e. the case without feedback, the expression in (4.6) reduces to the LST of the waiting distribution in the $M/G/1$ queue with multiple gated vacations, cf. [8], formula (5.21b) on p. 208. The latter formula can be decomposed into the waiting time LST of the ordinary $M/G/1$ queue and an additional term that is due to the occurrence of vacations. In (4.6) one also sees a hint of a decomposition; observe that the waiting time LST in the $M/G/1$ queue with service time LST $\tilde{B}_N(\omega)$ is given by $[1-\lambda(ps+(1-p)u)]\omega/[\omega-\xi(\omega)]$.

4.2. *Calculation of the distribution of* $W_U$. Consider the waiting time of an unsuccessful tagged customer $K$, arriving during a cycle $C$. Observe that, because old customers are served before new ones in a session, the first service time of $K$ does not influence his waiting time between both services; hence we can obtain his sojourn time distribution by convoluting his waiting time distribution with his two service time distributions.

We have

$$W_U = C_R + \text{service of } F(N) \text{ unlucky customers who were}$$
$$\text{present at the beginning of the cycle } C$$
$$+\text{service of all arrivals in } C_P \text{ (possibly two services)} \tag{4.7}$$
$$+\text{vacation at end of 2nd cycle}$$
$$+\text{first service of all (new) arrivals in } C_R.$$

Thus,

$$E[\mathrm{e}^{-\omega W_U}] = \tilde{V}(\omega)\frac{1}{E[C]}\int_{x=0}^{\infty}\int_{t_R=0}^{x}\mathrm{e}^{-\omega t_R}dP(C\leqq x)dt_R$$

$$\times\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{k}P(N=k,\,O=j,\,F(N)=k-l\mid C=x)\tilde{S}_1(\omega)^{k-l}$$

$$\times\sum_{m=0}^{\infty}\mathrm{e}^{-\lambda(x-t_R)}\frac{[\lambda(x-t_R)]^m}{m!}\sum_{n=0}^{\infty}\mathrm{e}^{-\lambda t_R}\frac{(\lambda t_R)^n}{n!} \tag{4.8}$$

$$\times\sum_{i=0}^{m}\binom{m}{i}p^i(1-p)^{m-i}\tilde{S}(\omega)^i(\tilde{U}(\omega)\tilde{S}_1(\omega))^{m-i}$$

$$\times\sum_{r=0}^{n}\binom{n}{r}p^r(1-p)^{n-r}\tilde{S}(\omega)^r\tilde{U}(\omega)^{n-r}.$$

Setting $\tilde{B}_2(\omega)=p\tilde{S}(\omega)+(1-p)\tilde{U}(\omega)\tilde{S}_1(\omega)$ we have

$$E[e^{-\omega W_U}] = \tilde{V}(\omega)\frac{1}{E[C]}\int_{x=0}^{\infty}\int_{t_R=0}^{x} e^{-\omega t_R} dt_R$$

$$\times \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{k} dP(C \leqq x \mid N=k, O=j, F(N)=k-l) P(N=k, O=j)$$

$$\times P(F(N)=k-l \mid N=k, O=j)\tilde{S}_1(\omega)^{k-l}$$

$$\times \exp\{-\lambda t_R[1-\tilde{B}_N(\omega)]\}\exp\{-\lambda(x-t_R)[1-\tilde{B}_2(\omega)]\} \tag{4.9}$$

$$= \tilde{V}(\omega)\frac{1}{E[C]}\int_{x=0}^{\infty}\int_{t_R=0}^{x} e^{-\omega t_R} dt_R \sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{k} d[S^{l*} * U^{(k-l)*} * S_1^{j*} * V](x)$$

$$\times P(N=k, O=j)\binom{k}{l} p^l (1-p)^{k-l}\tilde{S}_1(\omega)^{k-l}$$

$$\times \exp\{-\lambda t_R[1-\tilde{B}_N(\omega)]\}\exp\{-\lambda(x-t_R)[1-\tilde{B}_2(\omega)]\}.$$

Hence, integrating with respect to $t_R$, we get

$$E[e^{-\omega W_U}] = \frac{\tilde{V}(\omega)}{\omega+\lambda(\tilde{B}_2(\omega)-\tilde{B}_N(\omega))}\frac{1}{E[C]}\int_{x=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} P(N=k, O=j)$$

$$\times \sum_{l=0}^{k}\binom{k}{l} p^l(1-p)^{k-l}\tilde{S}_1(\omega)^{k-l}\exp\{-\lambda x[1-\tilde{B}_2(\omega)]\}$$

$$\times [1-\exp\{-x(\omega+\lambda(\tilde{B}_2(\omega)-\tilde{B}_N(\omega)))\}] d[S^{l*} * U^{(k-l)*} * S_1^{j*} * V](x)$$

$$= \frac{\tilde{V}(\omega)}{\omega+\lambda(\tilde{B}_2(\omega)-\tilde{B}_N(\omega))}\frac{1}{E[C]}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} P(N=k, O=j)$$

$$\times \sum_{l=0}^{k}\binom{k}{l} p^l(1-p)^{k-l}\tilde{S}_1(\omega)^{k-l}\{\tilde{S}(\xi_2(\omega))^l\tilde{U}(\xi_2(\omega))^{k-l}\tilde{S}_1(\xi_2(\omega))^j\tilde{V}(\xi_2(\omega))$$

$$-\tilde{S}(\omega+\xi(\omega))^l\tilde{U}(\omega+\xi(\omega))^{k-l}\tilde{S}_1(\omega+\xi(\omega))^j\tilde{V}(\omega+\xi(\omega))\},$$

where $\xi_2(\omega) \equiv \lambda(1-\tilde{B}_2(\omega))$. Finally,

$$E[e^{-\omega W_U}] = \frac{\tilde{V}(\omega)}{\omega+\lambda(\tilde{B}_2(\omega)-\tilde{B}_N(\omega))}\frac{1}{E[C]}$$

$$\times [\tilde{V}(\xi_2(\omega))G(p\tilde{S}(\xi_2(\omega))+(1-p)\tilde{U}(\xi_2(\omega))\tilde{S}_1(\omega), \tilde{S}_1(\xi_2(\omega))) \tag{4.10}$$

$$-\tilde{V}(\omega+\xi(\omega))G(p\tilde{S}(\omega+\xi(\omega))+(1-p)\tilde{U}(\omega+\xi(\omega))\tilde{S}_1(\omega), \tilde{S}(\omega+\xi(\omega)))].$$

The waiting time LST of an arbitrary customer follows from (4.1), (4.6) and (4.10). In particular, moments can be obtained straightforwardly.

## Acknowledgement

The first author's research has been supported by the European Grant BRA-QMIPS of CEC DG XIII.

## References

[1] ALTMAN, E. AND YECHIALI, U. (1994) Polling in a closed network. *Prob. Eng. Inf. Sci.* **8**, 327–343.

[2] BORST, S. C. AND BOXMA, O. J. (1997) Polling models with and without switchover times. *Operat. Res.* To appear.

[3] BOXMA, O. J. (1994) Polling systems. In *From Universal Morphisms to Megabytes: A Baayen Space Odyssey.* (*Liber Amicorum for P. C. Baayen*). ed. K. R. Apt, A. Schrijver and N. M. Temme. CWI, Amsterdam. pp. 215–230.

[4] COHEN, J. W. (1982) *The Single Server Queue.* North-Holland, Amsterdam.

[5] EISENBERG, M. (1972) Queues with periodic service and changover time. *Operat. Res.* **20**, 440–451.

[6] SIDI, M., LEVY, H. AND FUHRMANN, S. W. (1992) A queueing network with a single cyclically roving server. *Queueing Systems* **11**, 121–144.

[7] TAKÁCS, L. (1963) A queueing model with feedback. *Operat. Res.* **11**, 345–354.

[8] TAKAGI, H. (1991) *Queueing Analysis. Vol. 1: Vacation and Priority Systems.* North-Holland, Amsterdam.

[9] TAKINE, T., TAKAGI, H. AND HASEGAWA, T. (1991) Sojourn times in vacation and polling systems with Bernoulli feedback. *J. Appl. Prob.* **28**, 422–432.