$M/G/1$ queue with exponential working vacation and gated service

Zsolt Saffer and Miklós Telek

Budapest University of Technology and Economics (BUTE), HUNGARY
{safferzs,telek}@hit.bme.hu

Abstract. In this paper we consider the analysis of $M/G/1$ queue with working vacation. In contrast to the previous literature where the working vacation starts when all customers are served (exhaustive discipline) we consider the case where the vacation period starts when the customers present at the system at beginning of the service period are served (gated discipline). The analysis of the model with gated discipline requires a different approach than the one with exhaustive discipline.
We present the probability generating function (PGF) and the mean of several interesting performance measures of this system.
Keywords: queueing theory, vacation model, working vacation

1 Introduction

Working vacation (WV) is an extension to the regular vacation model, where the server is switched off during the vacation. In WV systems, instead of completely stopping the service at the beginning of the vacation period, the server continuous serving the customers with a different service rate. In practice the service rate during WV is lower than the one during service period, but it is not a modeling restriction. WV models are suitable to describe practical system features like the effect of a lower intensity administrative task of the server, e.g. a safeguard service, following an active period. In the *multiple working vacation* (MWV) model a vacation period can be followed by another vacation period if the conditions of starting the vacation are observed.

The regular vacation model is a special case of the working vacation model with zero service rate during vacation. Thus the working vacation model is a generalization of the regular vacation model and it turned out that its analytical complexity is somewhat higher than the one of the regular vacation model.

Vacation models with Poisson arrival have been intensively studied in the last decades. For such models and for their solution we refer to the survey of Doshi [4] and to the fundamental book of Takagi [16].

The WV policy has been introduced by Servi and Finn [13] in 2002. They applied the M/M/1 queue with multiple working vacations to model a Wavelength-Division Multiplexing (WDM) optical access network. They derived the probability-generating function (PGF) of the number of customers in the system. Baba [1] generalized this model to a renewal input GI/M/1 queue with working vacations and derived the number of customers at customer arrival epochs and at arbitrary time. Wu and Takagi [18] generalized the model of Servi and Finn to the M/G/1 queue with working vacations. They provided a numerical solution for computing the distribution of the number of customers in the system. Liu, Xu and Tian [10] proved that the well-known *stochastic decomposition property* of the vacation queues [5] holds also for M/M/1 queue with working vacations. Recently Li and Tian with co-authors in [9] and [8] applied the matrix analytic approach subsequently to M/G/1 queue with exponential working vacations and to its extension to batch arrival, respectively. In all the above-mentioned working vacation queueing models the *exhaustive discipline* has been applied, i.e. the vacation starts, when the queue becomes empty.

In this paper we analyze the M/G/1 queue with exponential working vacations, but in contrast to the above references we consider the *gated discipline*. According to this discipline only those customers are served during the actual service period, which are already present at the beginning of the service period. Thus the customers arriving during a service period are present at the beginning of the next vacation period. A potential application of this model is the analysis of an Internet Protocol (IP) over WDM optical access network under heavily loaded traffic conditions (see in [11]). A motivation to apply the gated discipline instead of the exhaustive one is that under heavily loaded traffic condition the exhaustive discipline would monopolize the server in case of cyclic service queue, while the gated discipline weakens this effect.

The contribution of this paper is the queueing theoretic analysis and the results for the M/G/1 working vacation queueing model with gated discipline. In the first part of the analysis we establish a functional equation for the PGF of the stationary number of customers at the start of vacation, for which we utilize known results for the transient behavior of the $M/G/1$ queue. We solve this equation by using a convergence property of series of scalar probability functions (see [7] or [6]), which results in the expression of the PGF of the stationary number of customers at the beginning of vacation. In the second part of the analysis we establish a relation for the PGF of the stationary number of customers at an arbitrary epoch in terms of the PGFs of the stationary number of customers at the end and at the start of vacations. For the analysis we apply several elements of the methodology of [18]. The main results are the expressions of the PGF of the stationary number of customers at an arbitrary epoch and of the Laplace-Stieljes transform (LST) of the stationary waiting time.

The change of service rate is an important feature of the WV models. Both in case of exhaustive and gated disciplines the service period ends in a service completion epoch. I means that the last customer served during the service period is served with the regular service time and the first in the vacation period is served with the service time of the vacation period. In contrast the end of the vacation period is not synchronized with service completion. If a vacation period expires during the service of a customer the service period starts with the complete service of the same customer according to the service time of the service period. If the system is empty at the end of a vacation period then it immediately starts another vacation period.

The rest of this paper is organized as follows. In section 2 we introduce the model and the notations. The stationary number of customers at the start of vacation are derived in 3. The PGF of the stationary number of customers is given in section 4. The LST of the stationary waiting time is provided in section 5.

2 Model description

We consider a queue with multiple working vacation and gated service. The customer arrival process is Poisson with rate λ. Due to the gated service only those customers are served, which are present at the start of service period. The customer service times are independent and identically distributed. B, $B(t)$, $\widetilde{B}(s)$, b denote the service time r.v., its cumulated distribution function, its LST and its mean, respectively. The service during the service period is work conserving and non-preemptive. After finishing the service in the service period the server goes to vacation. The vacation period, V, is exponentially distributed r.v. with parameter μ. Note that the customers arriving during the service period are present at the start of vacation. According to the multiple working vacation policy the server serves the customers with a different service time distribution during the vacation period instead of completely stopping the service. The customer service times in the vacation period are independent and identically distributed. H, $\widetilde{H}(s)$, h denote the service time r.v. in the vacation period, its LST and its mean, respectively. If there is a customer under service at the end of vacation, the server changes to another service rate, i.e. it interrupts the service and starts a new service on that customer with the customer service time B. If there are no customers in the system at the end of the vacation, the server immediately takes another vacation. We define the *cycle time* as a service period and a vacation period together. On this vacation model we impose the following assumptions:

A.1 The arrival rate, the mean customer service time, the mean vacation time and the mean customer service time in the vacation period are positive and finite, i.e. $0 < \lambda < \infty$, $0 < b < \infty$, $0 < 1/\mu < \infty$ and $0 < h < \infty$.

A.2 The arrival process, the customer service times, the sequence of vacation periods and the customer service times in the vacation periods are mutually independent.

A.3 The customers are served in First-In-First-Out (FIFO) order.

We assume that the model is stable. It means that the arrival rate can not exceed the mean service rate ($\frac{1}{b}$). Therefore the necessary condition of the stability is $\lambda b < 1$.

The *Markov Regenerative Process* (MRP) framework (introduced in [12], Theorem 1) holds also for this model, since all the necessary assumptions are fulfilled. Thus the limiting distributions of the number of customers at different epoch and in different intervals are stationary distributions. Hence through this paper we use the term "stationary" instead of "limiting".

When $y(z)$ is a PGF, $y^{(k)}$ denotes its k-th derivative at $z = 1$ for $k \geq 1$, i.e., $y^{(k)} = \frac{d^k}{dz^k} y(z)|_{z=1}$.

3 The stationary number of customers at the start of vacation

In this section we derive the PGF of the stationary number of customers at the start of vacation. We describe the evolution of the system over the vacation period in terms of PGFs of the stationary number of customers. For doing this we take the idea of utilizing the transient behavior of the $M/G/1$ queue from [18], which we recall first. The descriptions of the evolutions of the system over the vacation period and over the service period lead to a functional equation for the PGF of the stationary number of customers at the start of vacation, which we call the governing equation of the system. Afterwards we solve this functional equation by applying a convergence property of series of scalar functions.

3.1 Transient behavior of the $M/G/1$ queue

Let $\mathit{\Gamma}(t)$ be the number of customers in the $M/G/1$ queue at time t for $t \geq 0$. Suppose that the queue starts to work at $t = 0$ and $\mathit{\Gamma}(0) = i$ for $i \geq 0$. The transition probability describing the changes of the number of customers up to t is defined as

$$p_{i,j}(t) = P\left\{\mathit{\Gamma}(t) = j | \mathit{\Gamma}(0) = i\right\}, \quad t \geq 0, \quad i, j \geq 0.$$

The corresponding Laplace transform (LT) and PGF are given as

$$\widetilde{T}_{i,j}(s) = \int_{t=0}^{\infty} e^{-st} p_{i,j}(t) dt, \quad Re(s) \geq 0.$$

$$\widehat{T}_i(z,t) = \sum_{j=0}^{\infty} p_{i,j}(t) z^j, \quad |z| \leq 1.$$

The LT of $\widehat{T}_i(z,t)$

$$\widetilde{\widehat{T}}_i(z,s) = \int_{t=0}^{\infty} e^{-st} \widehat{T}_i(z,t) dt, \quad i \geq 0.$$

can be explicitly expressed as ([14], p. 74, eq. (77))

$$\widetilde{\widehat{T}}_i(z,s) = \frac{z^{i+1}\left(1 - \widetilde{H}(s+\lambda-\lambda z)\right) + (z-1)(s+\lambda-\lambda z)\widetilde{H}(s+\lambda-\lambda z)\widetilde{T}_{i,0}(s)}{(s+\lambda-\lambda z)\left(z - \widetilde{H}(s+\lambda-\lambda z)\right)}. \quad (1)$$

Here $\widetilde{T}_{i,0}(s)$ can be explicitly given as

$$\widetilde{T}_{i,0}(s) = \frac{\theta(s)^i}{s+\lambda-\lambda\theta(s)}, \tag{2}$$

where $\theta(s)$ can be determined from the equation

$$\theta(s) = \widetilde{H}(s+\lambda-\lambda\theta(s)), \tag{3}$$

as its root with the smallest absolute value.

Let $K_i(t)$ denote the mean number of customers served in the time interval $(0,t]$ and suppose that $\Gamma(0) = i$ for $i \geq 0$. The LT of $K_i(t)$ is given by ([14], p. 78, eq. (91))

$$\widetilde{K}_i(s) = \int_{t=0}^{\infty} e^{-st} K_i(t) dt = \frac{\widetilde{H}(s)}{s(1-\widetilde{H}(s))} \left(1 - \frac{s\theta(s)^i}{s+\lambda-\lambda\theta(s)}\right). \tag{4}$$

3.2 Evolution of the system over the vacation period

Let $N(t)$ the number of customers in the system at time t for $t \geq 0$. Furthermore let t_k^f and t_k^m denote the end and the start of vacation in the k-th cycle for $k \geq 1$, respectively. The stationary PGF of the number of customers at the end of vacation, $\widehat{f}(z)$, and at the beginning of vacation, $\widehat{m}(z)$, are defined as

$$\widehat{f}(z) = \lim_{k\to\infty} \sum_{n=0}^{\infty} P\left\{N(t_k^f) = n\right\} z^n, \quad |z| \leq 1,$$

$$\widehat{m}(z) = \lim_{k\to\infty} \sum_{n=0}^{\infty} P\left\{N(t_k^m) = n\right\} z^n, \quad |z| \leq 1.$$

Let $f = f^{(1)}$ denote the mean of the stationary number of customers at the end of vacation. Similarly let $m = m^{(1)}$ denote the mean of the stationary number of customers at the start of vacation.

Theorem 1. *In the stable M/G/1 multiple working vacation model with gated discipline satisfying assumptions* **A.1** - **A.3** *the $m \to f$ transition can be described as*

$$\widehat{f}(z) = \frac{\mu z \widehat{m}(z)\left(1 - \widetilde{H}(\mu+\lambda-\lambda z)\right)}{(\mu+\lambda-\lambda z)\left(z - \widetilde{H}(\mu+\lambda-\lambda z)\right)} + \frac{\mu(z-1)\widetilde{H}(\mu+\lambda-\lambda z)}{z-\widetilde{H}(\mu+\lambda-\lambda z)} \frac{\widehat{m}(\theta(\mu))}{\mu+\lambda-\lambda\theta(\mu)}. \tag{5}$$

Proof. The evolution of the number of customers in the working vacation can be described by the transient behavior of the M/G/1 queue starting with the number of customers present at the start of vacation. Assuming that the number of customers present at the start of vacation in the k-th cycle is $i \geq 0$, the PGF of the number of customers present at the end of that vacation can be expressed as

$$\int_{t=0}^{\infty} \widehat{T}_i(z,t)\mu e^{-\mu t} dt.$$

Unconditioning on the number of customers at the start of vacation and letting m tend to ∞ gives the PGF of the number of customers at the end of vacation as

$$\widehat{f}(z) = \lim_{k\to\infty} \sum_{i=0}^{\infty} P\left\{N(t_k^m) = i\right\} \int_{t=0}^{\infty} \widehat{T}_i(z,t)\mu e^{-\mu t} dt. \tag{6}$$

Applying the definition of $\widetilde{\widehat{T}}_i(z,s)$ and (1) in (6) results in

$$\begin{aligned}\widehat{f}(z) &= \mu \lim_{k\to\infty} \sum_{i=0}^{\infty} P\left\{N(t_k^m) = i\right\} \widetilde{\widehat{T}}_i(z,\mu) \qquad (7)\\ &= \mu \lim_{k\to\infty} \sum_{i=0}^{\infty} P\left\{N(t_k^m) = i\right\} \\ &\quad \frac{z^{i+1}\left(1-\widetilde{H}(\mu+\lambda-\lambda z)\right) + (z-1)(\mu+\lambda-\lambda z)\widetilde{H}(\mu+\lambda-\lambda z)\widetilde{T}_{i,0}(\mu)}{(\mu+\lambda-\lambda z)\left(z-\widetilde{H}(\mu+\lambda-\lambda z)\right)}.\end{aligned}$$

Applying (2) and the definition of $\widehat{m}(z)$ in (7) gives the statement of the theorem. □

3.3 Computation of $\widehat{f}(z)$

Theorem 2. *The governing equation of the stable M/G/1 multiple working vacation model with gated discipline satisfying assumptions* **A.1** - **A.3** *is given in term of* $\widehat{f}(z)$ *as*

$$\begin{aligned}\widehat{f}(z) = &\frac{\widehat{f}(\widetilde{B}(\lambda-\lambda z))\mu z\left(1-\widetilde{H}(\mu+\lambda-\lambda z)\right)}{(\mu+\lambda-\lambda z)\left(z-\widetilde{H}(\mu+\lambda-\lambda z)\right)} \\ &+ \frac{\mu(z-1)\widetilde{H}(\mu+\lambda-\lambda z)}{z-\widetilde{H}(\mu+\lambda-\lambda z)} \frac{\widehat{f}(\widetilde{B}(\lambda-\lambda\theta(\mu)))}{\mu+\lambda-\lambda\theta(\mu)}.\end{aligned} \tag{8}$$

Proof. Under the gated discipline the number of customers at the end of the service period (= at the start of vacation) equals to the number of customers arrive during the service period. The PGF of the number of arriving customers during a customer service time is $\widetilde{B}(\lambda - \lambda z)$. Hence the PGF of the number of customers at the start of vacation is given by (see e.g. [15])

$$\widehat{m}(z) = \widehat{f}(\widetilde{B}(\lambda - \lambda z)) \tag{9}$$

This relation describes the $f \to m$ transition under the gated discipline. Substituting (9) into (5) results in the governing equation of the system. □

The derivative of (9) at $z = 1$ gives

$$m = \lambda b f. \tag{10}$$

We define a series of functions recursively as

$$\begin{aligned} \beta_0(z) &= z, \quad |z| \le 1, \\ \beta_{k+1}(z) &= \widetilde{B}(\lambda - \lambda \beta_k(z)), \quad k \ge 0. \end{aligned} \tag{11}$$

Theorem 3. *In the stable M/G/1 multiple working vacation model with gated discipline satisfying assumptions* **A.1** - **A.3** *the stationary PGF of the number of customers at the start of vacation is given as*

$$\widehat{f}(z) = \phi(z) + \frac{\phi(\widetilde{B}(\lambda - \lambda\theta(\mu)))}{\mu + \lambda - \lambda\theta(\mu) - \psi(\widetilde{B}(\lambda - \lambda\theta(\mu)))}\, \psi(z). \tag{12}$$

where $\phi(z)$ and $\psi(z)$ are given as

$$\phi(z) = \prod_{k=1}^{\infty} \eta_k(z), \tag{13}$$

$$\psi(z) = \sum_{k=0}^{\infty} \left(\omega_k(z) \prod_{r=0}^{k-1} \eta_r(z) \right), \tag{14}$$

with

$$\eta_k(z) = \frac{\mu\beta_k(z)\left(1 - \widetilde{H}(\mu + \lambda - \lambda\beta_k(z))\right)}{(\mu + \lambda - \lambda\beta_k(z))\left(\beta_k(z) - \widetilde{H}(\mu + \lambda - \lambda\beta_k(z))\right)}, \tag{15}$$

$$\omega_k(z) = \frac{\mu(\beta_k(z) - 1)\widetilde{H}(\mu + \lambda - \lambda\beta_k(z))}{\left(\beta_k(z) - \widetilde{H}(\mu + \lambda - \lambda\beta_k(z))\right)} \tag{16}$$

and an empty product is 1.

Proof. Replacing z by $\beta_k(z)$ in (8) for $k \geq 0$ leads to

$$\widehat{f}(\beta_k(z)) = \widehat{f}(\beta_{k+1}(z)) \frac{\mu\beta_k(z)\left(1-\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))\right)}{(\mu+\lambda-\lambda\beta_k(z))\left(\beta_k(z)-\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))\right)} + \frac{\mu(\beta_k(z)-1)\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))}{\left(\beta_k(z)-\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))\right)} \frac{\widehat{f}(\widetilde{B}(\lambda-\lambda\theta(\mu)))}{(\mu+\lambda-\lambda\theta(\mu))}. \tag{17}$$

Solving (17) by recursive substitution for $k \geq 0$ yields

$$\widehat{f}(z) = \widehat{f}(\lim_{k\to\infty}\beta_k(z)) \prod_{k=1}^{\infty} \frac{\mu\beta_k(z)\left(1-\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))\right)}{(\mu+\lambda-\lambda\beta_k(z))\left(\beta_k(z)-\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))\right)} + \frac{\widehat{f}(\widetilde{B}(\lambda-\lambda\theta(\mu)))}{\mu+\lambda-\lambda\theta(\mu)} \sum_{k=0}^{\infty}\left(\frac{\mu(\beta_k(z)-1)\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))}{\left(\beta_k(z)-\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))\right)} \prod_{r=0}^{k-1} \frac{\mu\beta_r(z)\left(1-\widetilde{H}(\mu+\lambda-\lambda\beta_r(z))\right)}{(\mu+\lambda-\lambda\beta_r(z))\left(\beta_r(z)-\widetilde{H}(\mu+\lambda-\lambda\beta_r(z))\right)}\right). \tag{18}$$

It can be shown that for $\lambda b < 1$, which is the condition of stability, $\lim_{k\to\infty}\beta_k(z) = 1$ for any $|z| \leq 1$ [6]. Using this in (18) yields

$$\widehat{f}(z) = \prod_{k=1}^{\infty} \frac{\mu\beta_k(z)\left(1-\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))\right)}{(\mu+\lambda-\lambda\beta_k(z))\left(\beta_k(z)-\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))\right)} + \frac{\widehat{f}(\widetilde{B}(\lambda-\lambda\theta(\mu)))}{\mu+\lambda-\lambda\theta(\mu)} \sum_{k=0}^{\infty}\left(\frac{\mu(\beta_k(z)-1)\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))}{\left(\beta_k(z)-\widetilde{H}(\mu+\lambda-\lambda\beta_k(z))\right)} \prod_{r=0}^{k-1} \frac{\mu\beta_r(z)\left(1-\widetilde{H}(\mu+\lambda-\lambda\beta_r(z))\right)}{(\mu+\lambda-\lambda\beta_r(z))\left(\beta_r(z)-\widetilde{H}(\mu+\lambda-\lambda\beta_r(z))\right)}\right). \tag{19}$$

By the notations introduced in (13) and (14), (19) can be rewritten as

$$\widehat{f}(z) = \phi(z) + \frac{\widehat{f}(\widetilde{B}(\lambda-\lambda\theta(\mu)))}{\mu+\lambda-\lambda\theta(\mu)}\psi(z). \tag{20}$$

The unknown term $\frac{\widehat{f}(\widetilde{B}(\lambda-\lambda\theta(\mu)))}{\mu+\lambda-\lambda\theta(\mu)}$ is obtained by setting $z = \widetilde{B}(\lambda-\lambda\theta(\mu))$ in (20), therefore

$$\widehat{f}(\widetilde{B}(\lambda-\lambda\theta(\mu))) = \frac{\phi(\widetilde{B}(\lambda-\lambda\theta(\mu)))}{1-\dfrac{\psi(\widetilde{B}(\lambda-\lambda\theta(\mu)))}{\mu+\lambda-\lambda\theta(\mu)}}. \tag{21}$$

Substituting it into (20) gives the theorem. □

Remark 1. $\widehat{m}(z)$ can be also determined by applying (12) in the gated discipline specific $f \to m$ transition (9).

From the definition of $\beta_k(z)$, $\eta_k(z)$ and $\omega_k(z)$ we have $\beta_k(1) = 1$, $\beta_k^{(1)} = (b\lambda)^k$, $\eta_k(1) = 1$, $\eta_k^{(1)} = \frac{(\lambda+\mu)\widetilde{H}(\mu)-\lambda}{\mu(\widetilde{H}(\mu)-1)}(b\lambda)^k$, $\omega_k(1) = 0$, $\omega_k^{(1)} = \frac{\mu\widetilde{H}(\mu)}{1+\widetilde{H}(\mu)}(b\lambda)^k$, for $k \geq 0$. Using these properties we obtain

$$\phi^{(1)} = \sum_{k=1}^{\infty}\eta_k^{(1)} = \frac{(\lambda+\mu)\widetilde{H}(\mu)-\lambda}{\mu(\widetilde{H}(\mu)-1)}\ \frac{b\lambda}{1-b\lambda}, \tag{22}$$

$$\psi^{(1)} = \sum_{k=0}^{\infty}\omega_k^{(1)} = \frac{\mu\widetilde{H}(\mu)}{1+\widetilde{H}(\mu)}\ \frac{1}{1-b\lambda}, \tag{23}$$

and from (12) we have

$$f^{(k)} = \phi^{(k)} + \frac{\phi(\widetilde{B}(\lambda-\lambda\theta(\mu)))}{\mu+\lambda-\lambda\theta(\mu)-\psi(\widetilde{B}(\lambda-\lambda\theta(\mu)))}\ \psi^{(k)}. \tag{24}$$

The computation of the second moments is more involved.

$$\beta_k^{(2)} = \lambda^2 E(B^2)\sum_{i=k-1}^{2k-2}(b\lambda)^i, \tag{25}$$

$$\eta_k^{(2)} = -\frac{\left(4\lambda^2(b\lambda)^{2k}+\mu^2\left(\beta_k^{(2)}-2(b\lambda)^{2k}\right)+2\lambda\mu\left((b\lambda)^{2k}+\beta_k^{(2)}\right)\right)\widetilde{H}(\mu)}{\mu^2(\widetilde{H}(\mu)-1)^2} + \tag{26}$$

$$+\frac{(\lambda+\mu)\left(2\lambda(b\lambda)^{2k}+\mu\beta_k^{(2)}\right)\widetilde{H}(\mu)^2}{\mu^2(\widetilde{H}(\mu)-1)^2} + \frac{\lambda\left(2\lambda(b\lambda)^{2k}+2\mu^2\widetilde{H}'(\mu)(b\lambda)^{2k}+\mu\beta_k^{(2)}\right)}{\mu^2(\widetilde{H}(\mu)-1)^2},$$

$$\omega_k^{(2)} = \frac{2\lambda\mu\widetilde{H}'(\mu)(b\lambda)^{2k}+\mu\widetilde{H}(\mu)\left(2(b\lambda)^{2k}-\beta_k^{(2)}\right)+\mu\widetilde{H}(\mu)^2\beta_k^{(2)}}{-(\widetilde{H}(\mu)-1)^2}, \tag{27}$$

and

$$\phi^{(2)} = \sum_{k=1}^{\infty}\eta_k^{(2)} + 2\sum_{k=1}^{\infty}\eta_k^{(1)}\sum_{i=k+1}^{\infty}\eta_i^{(1)}, \tag{28}$$

$$\psi^{(2)} = \sum_{k=0}^{\infty}\omega_k^{(2)} + 2\sum_{k=0}^{\infty}\omega_k^{(1)}\sum_{i=0}^{k-1}\eta_i^{(1)}. \tag{29}$$

4 The stationary number of customers at an arbitrary epoch

Following the same line of argument as in [18] we determine the PGF of the stationary number of customers at an arbitrary epoch from the PGFs of the stationary number of customers in the service period and in the vacation period. These PGFs are determined in terms of $\widehat{f}(z)$ and $\widehat{m}(z)$. Putting all these together and applying the formulas for $\widehat{f}(z)$ and $\widehat{m}(z)$ derived in the previous section gives the PGF of the stationary number of customers at an arbitrary epoch.

We define $\widehat{q}(z)$ as the PGF of the stationary number of customers in an arbitrary epoch as

$$\widehat{q}(z) = \lim_{t\to\infty} \sum_{n=0}^{\infty} P\left\{N\left(t\right) = n\right\} z^n, \quad |z| \leq 1.$$

4.1 The stationary number of customers in the service period

We define $\widehat{q}^b(z)$ as the PGF of the stationary number of customers in the service period as

$$\widehat{q}^b(z) = \lim_{t\to\infty} \sum_{n=0}^{\infty} P\left\{N\left(t\right) = n \mid t \in \text{service period}\right\} z^n, \quad |z| \leq 1.$$

Let $G(k)$ be the number of customers served in the service period in the k-th cycle for $k \geq 1$. Let $t^s_{k,\ell}$ and $t^d_{k,\ell}$ denotes the start and the end of the ℓ-th customer service in the service period in the k-th cycle for $\ell = 1, ...G(k)$ and $k \geq 1$, respectively.

Furthermore we define $\widehat{q}^s(z)$ as the PGF of the stationary number of customers at the customer service start epochs in the service period as

$$\widehat{q}^s(z) = \lim_{m\to\infty} \sum_{n=0}^{\infty} \frac{\sum_{k=1}^{m} \sum_{\ell=1}^{G(k)} P\left\{N(t^s_{k,\ell}) = n\right\}}{E\left(\sum_{k=1}^{m} G(k)\right)} z^n, \quad |z| \leq 1.$$

Similarly we define $\widehat{q}^d(z)$ as the PGF of the stationary number of customers at the customer departure epochs in the service period as

$$\widehat{q}^d(z) = \lim_{m\to\infty} \sum_{n=0}^{\infty} \frac{\sum_{k=1}^{m} \sum_{\ell=1}^{G(k)} P\left\{N(t^d_{k,\ell}) = n\right\}}{E\left(\sum_{k=1}^{m} G(k)\right)} z^n, \quad |z| \leq 1.$$

Proposition 1. *In the stable M/G/1 multiple working vacation model with gated discipline satisfying assumptions* **A.1** - **A.3** *the stationary PGF of the number of customers in the vacation period is given as*

$$\widehat{q}^b(z) = \frac{(1-\lambda b)(1-\widetilde{B}(\lambda-\lambda z))}{\lambda b(1-z)(\widetilde{B}(\lambda-\lambda z)-z)} \; \frac{\widehat{m}(z)-\widehat{f}(z)}{f-m}. \tag{30}$$

Proof. The service process in the models with and without working vacation are completely identical. Therefore the expression of $\widehat{q}^d(z)$ derived from the service process in [3] is valid also in our working vacation model. In fact in [3] this expression has been derived for $\widehat{q}(z)$, but a standard up-and down-crossing argument combined with PASTA [17] shows that the stationary number of customers at customer departure, at customer arrival and at arbitrary epochs are all the same. According to this $\widehat{q}^d(z)$ can be expressed as

$$\widehat{q}^d(z) = \frac{(1-\lambda b)\widetilde{B}(\lambda-\lambda z)}{\widetilde{B}(\lambda-\lambda z)-z} \; \frac{\widehat{m}(z)-\widehat{f}(z)}{f-m}. \tag{31}$$

Furthermore the number of customers at the customer departure epoch is the sum of the number of customers at the previous customer departure epoch and the number of customers arriving during the last customer service, which are independent. Therefore we have

$$\widehat{q}^d(z) = \widehat{q}^s(z)\widetilde{B}(\lambda-\lambda z). \tag{32}$$

Applying (31) in (32), $\widehat{q}^s(z)$ can be expressed as

$$\widehat{q}^s(z) = \frac{(1-\lambda b)}{\widetilde{B}(\lambda-\lambda z)-z} \; \frac{\widehat{m}(z)-\widehat{f}(z)}{f-m}. \tag{33}$$

The interval between the starting time of a customer service in the service period and an arbitrary epoch in that service time is the backward recurrence customer service time, whose probability density function (pdf) is given by

$$b^*(t) = \frac{1-B(t)}{b}. \tag{34}$$

The number of customers at an arbitrary epoch in any customer service time is the number of customers at the start of that customer service and the number of customers arriving in between. The later can be obtained by integrating the customer arrivals with the backward recurrence customer service time. Thus for $\widehat{q}^b(z)$ we obtain

$$\widehat{q}^b(z) = \lim_{m\to\infty} \sum_{j=1}^{\infty} z^j \frac{\sum_{k=1}^{m} \sum_{\ell=1}^{G(k)} \sum_{i=1}^{j} P\left\{N(t^s_{k,\ell}) = i\right\}}{E\left(\sum_{k=1}^{m} G(k)\right)} \int_{t=0}^{\infty} \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t} b^*(t)dt. \quad (35)$$

Rearrangement of (35) and using the definition of $\widehat{q}^s(z)$ leads to

$$\begin{aligned}\widehat{q}^b(z) &= \lim_{m\to\infty} \sum_{i=1}^{\infty} \frac{\sum_{k=1}^{m} \sum_{\ell=1}^{G(k)} P\left\{N(t^s_{k,\ell}) = i\right\}}{E\left(\sum_{k=1}^{m} G(k)\right)} z^i \sum_{j-i=0}^{\infty} \int_{t=0}^{\infty} \frac{(\lambda t z)^{j-i}}{(j-i)!} e^{-\lambda t} b^*(t)dt \\ &= \widehat{q}^s(z) \int_{t=0}^{\infty} e^{-\lambda t(1-z)} b^*(t)dt. \end{aligned} \quad (36)$$

Applying (34) and the integral property of LT, that is $\int_{t=0}^{\infty} e^{-st}B(t)dt = \frac{\widetilde{B}(s)}{s}$, (36) yields

$$\begin{aligned}\widehat{q}^b(z) &= \frac{\widehat{q}^s(z)}{b} \left(\int_{t=0}^{\infty} e^{-\lambda t(1-z)}dt - \int_{t=0}^{\infty} e^{-\lambda t(1-z)}B(t)dt \right) \\ &= \widehat{q}^s(z) \frac{1-\widetilde{B}(\lambda - \lambda z)}{\lambda b(1-z)}. \end{aligned} \quad (37)$$

Applying (33) in (37) results in the statement of the proposition. □

4.2 The stationary number of customers in the vacation period

We define $\widehat{q}^v(z)$ as the PGF of the stationary number of customers in the vacation period as

$$\widehat{q}^v(z) = \lim_{t\to\infty} \sum_{n=0}^{\infty} P\left\{N(t) = n \mid t \in \text{vacation period}\right\} z^n, \quad |z| \le 1.$$

Proposition 2. *In the stable M/G/1 multiple working vacation model with gated discipline satisfying assumptions* **A.1** - **A.3** *the stationary PGF of the number of customers in the vacation period is given as*

$$\widehat{q}^v(z) = \widehat{f}(z). \quad (38)$$

Proof. The interval between the starting time of a vacation and an arbitrary epoch in that vacation is the backward recurrence vacation time. Let $v^*(t)$ denote the probability density function (pdf) of the backward recurrence vacation time.

Following the same line of argument as in theorem 1 the PGF of the number of customers present at an arbitrary epoch in the vacation can be expressed as

$$\widehat{q}^v(z) = \lim_{k\to\infty} \sum_{i=0}^{\infty} P\left\{N(t_k^m) = i\right\} \int_{t=0}^{\infty} \widehat{T}_i(z,t) v^*(t) dt. \qquad (39)$$

Due to the exponential distribution of the vacation time the backward recurrence vacation time is also exponentially distributed. Thus the right side of (39) is the same as the right side of (6), from which the statement of the proposition follows. □

4.3 Computation of $\widehat{q}(z)$

Theorem 4. *In the stable M/G/1 multiple working vacation model with gated discipline satisfying assumptions* **A.1** - **A.3** *the stationary PGF of the number of customers at an arbitrary epoch is given as*

$$\widehat{q}(z) = \frac{1}{\mu b f + 1} \left(\frac{\mu(1-\widetilde{B}(\lambda-\lambda z))\left(\widehat{f}(\widetilde{B}(\lambda-\lambda z)) - \widehat{f}(z)\right)}{\lambda(1-z)(\widetilde{B}(\lambda-\lambda z) - z)} + \widehat{f}(z) \right) \qquad (40)$$

where $\widehat{f}(z)$ and f are given by (12) and (24), respectively.

Proof. Let σ be the mean stationary length of the service period. Under gated discipline each customer present at the start of service period generates a service with mean length b. Applying Wald's lemma leads to

$$\sigma = fb. \qquad (41)$$

The state of the working vacation model alternates between service periods and vacation periods. According to the renewal theory the probabilities that the random epoch τ finds the model in service period or in vacation are given as

$$p_b = \{\tau \in \text{service period}\} = \frac{\sigma}{\sigma + 1/\mu},$$
$$p_v = \{\tau \in \text{vacation}\} = \frac{1/\mu}{\sigma + 1/\mu}. \qquad (42)$$

It follows from the theorem of total probability that

$$\widehat{q}(z) = p_b \widehat{q}^b(z) + p_v \widehat{q}^v(z). \qquad (43)$$

Applying (41), (42), and propositions 2 and 1 in (43) leads to

$$\widehat{q}(z) = \frac{\mu b f}{\mu b f + 1} \frac{(1-\lambda b)(1-\widetilde{B}(\lambda-\lambda z))}{\lambda b(1-z)(\widetilde{B}(\lambda-\lambda z) - z)} \frac{\widehat{m}(z) - \widehat{f}(z)}{f - m} + \frac{1}{\mu b f + 1} \widehat{f}(z). \qquad (44)$$

Applying (9) and (10) results in the statement of the theorem. □

Corollary 1. *Based on (40) the mean number of customers is*

$$q^{(1)} = \frac{2f(\lambda b - 1) + \mu f \lambda E(B^2)(3\lambda b - 1) + \mu b f^{(2)}(\lambda^2 b^2 - 1)}{2(\lambda b - 1)(\mu b f + 1)}. \quad (45)$$

5 The stationary waiting time

Let W_τ be the waiting time in the system at time τ. We define the distribution function of the stationary waiting time, $W(t)$, as

$$W(t) = \lim_{\tau \to \infty} P\left\{W_\tau \le t\right\}.$$

The LST of the stationary waiting time is defined as

$$\widetilde{w}\left(s\right) = \int_{t=0}^{\infty} e^{-st} dW(t), \quad Re(s) \ge 0.$$

Theorem 5. *In the stable M/G/1 multiple working vacation model with gated discipline satisfying assumptions* **A.1** - **A.3** *the LST of the stationary waiting time is given as*

$$\widetilde{w}\left(s\right) = \frac{1}{\mu b f + 1} \frac{1}{\widetilde{B}\left(s\right)} \left(\frac{\mu(1 - \widetilde{B}(s)) \left(\widehat{f}(\widetilde{B}(s)) - \widehat{f}(1 - \frac{s}{\lambda})\right)}{s(\widetilde{B}(s) - (1 - \frac{s}{\lambda}))} + \widehat{f}\left(1 - \frac{s}{\lambda}\right) \right) \quad (46)$$

where $\widehat{f}(z)$ and f are given by (12) and (24), respectively.

Proof. The model assumptions imply that a new arriving customer do not affect the time in the system of any previously arrived customers. This ensures the applicability of the distributional Little's law [2]. Furthermore the time in the system of an arbitrary customer is the sum of its waiting time and its service time, which are independent due to the model assumptions. Taking it also into account the distributional Little's law can be given to our model as

$$\widehat{q}(z) = \widetilde{w}\left(\lambda - \lambda z\right) \widetilde{B}\left(\lambda - \lambda z\right). \quad (47)$$

Substituting $z = 1 - \frac{s}{\lambda}$ into (47) and rearranging yields

$$\widetilde{w}\left(s\right) = \frac{\widehat{q}(1 - \frac{s}{\lambda})}{\widetilde{B}\left(s\right)}. \quad (48)$$

The statement comes from (48) and (40). □

From (48) the mean waiting time, $E(W) = -\frac{d}{ds}\widetilde{w}(s)|_{s\to 0}$, is

$$E(W) = \frac{q^{(1)}}{\lambda} - b, \tag{49}$$

which is the regular Little law obtained from the distributional Little law.

The applied analysis method can be used to analyze several further non exhaustive service disciplines, like the binomial-gated and the binomial-exhaustive disciplines.

References

1. Baba, Y.: Analysis of a GI/M/1 queue with multiple working vacations. Operation Research Letters 33, 201–209 (2005)
2. Bertsimas, D., Nakazato, D.: The distributional Little's law and its application. Operations Research 43, 298–310 (1995)
3. Borst, S.C., Boxma, O.J.: Polling models with and without switchover times. Operations Research 45, 536–543 (1997)
4. Doshi, B.T.: Queueing systems with vacations - a survey. Queueing Systems 1, 29–66 (1986)
5. Fuhrmann, S.W., Cooper, R.B.: Stochastic Decompositions in the M/G/1 Queue with Generalized Vacations. Operations Research 33, 1117–1129 (1985)
6. Kim, T.S., Chang, S.H., Chae, K.C.: Performance Analysis of a Discrete-Time Two-Phase Queueing System. ETRI Journal 25, 238–246 (2003)
7. Kuczma, M.: Functional equations in a single variable. PWN - Polish Scientific Publishers, Warsaw (1968)
8. Li, J., Jia, D., Tian, N.: A batch arrival queue with exponential working vacations. In: 5th International Conference on Queueing Theory and Network Applications (QTNA 2010). Beijing. pp. 169–173 (jul 2010)
9. Li, J., Tian, N., Zhang, Z.G., Luh, H.P.: Analysis of the M/G/1 queue with exponential working vacations - a matrix analytic approach. Queueing Systems 61, 139–166 (2009)
10. Liu, W., Xu, X., Tian, N.: Stochastic decompositions in the M/M/1 queue with working vacations. Operation Research Letters 35, 595–600 (2007)
11. Narula-Tam, A., Finn, S.G., Medard, M.: Analysis of reconfiguration in IP over WDM access networks. In: Proceedings of the Optical Fiber Communication Conference (OFC). p. MN4.1MN4.3 (2001)
12. Saffer, Z.: An introduction to classical cyclic polling model. In: Proc. of the 14th Int. Conf. on Analytical and Stochastic Modelling Techniques and Applications (ASMTA'07). pp. 59–64 (2007)
13. Servi, L.D., Finn, S.G.: M/M/1 queue with working vacations(M/M/1/WV). Performance Evaluation 50, 41–52 (2002)
14. Takács, L.: Introduction to the Theory of Queues. Oxford University Press, New York (1962)
15. Takagi, H.: Analysis of Polling Systems. MIT Press (1986)
16. Takagi, H.: Queueing Analysis - A Foundation of Performance Evaluation, Vacation and Prority Systems, vol.1. North-Holland, New York (1991)
17. Wolff, R.W.: Poisson Arrivals See Times Averages. Operations Research 30, 223–231 (1982)
18. Wu, D., Takagi, H.: M/G/1 queue with multiple working vacations. Performance Evaluation 63, 654–681 (2006)