

THE MULTI-CLASS FIFO M/G/1 QUEUE WITH EXPONENTIAL WORKING VACATIONS

Yoshiaki Inoue Tetsuya Takine
Osaka University

(Received March 10, 2012; Revised January 1, 2013)

Abstract We consider a stationary multi-class FIFO M/G/1 queue with exponential working vacations, where a server works at two different processing rates. There are K classes of customers, and the arrival rates and the distributions of the amount of service requirements of arriving customers depend on both their customer classes and the server state. When the system becomes empty, the server takes a working vacation, during which customers are served at processing rate γ ($\gamma > 0$). If the system is empty at the end of the working vacation, the server takes another working vacation. On the other hand, if a customer is being served at the end of the working vacation, the server switches its processing rate to one and continues to serve customers in a preemptive-resume manner, until the system becomes empty. For this queue, we derive various quantities of interest, including the Laplace-Stieltjes transforms of the actual waiting time and sojourn time distributions, and the joint transform of the numbers of customers and the amounts of unfinished work in respective classes. As by-products, we also obtain various results of a stationary multi-class FIFO M/G/1 queue with Poisson disasters.

Keywords: Queue, M/G/1, multi-class, FIFO, working vacations, disasters

1. Introduction

This paper considers a single-server queue with working vacations. In queues with working vacations, the server takes a working vacation when the system becomes empty. Contrary to ordinary vacations models, customers are served at processing rate γ ($\gamma > 0$), which may differ from the normal processing rate of one. If a customer is being served at the end of the working vacation, the server switches its processing rate to one and continues to serve customers, until the system becomes empty.

The queueing model with working vacations was first introduced in [9], as a model of an access router in a reconfigurable wavelength division multiplexing (WDM) optical access network. While each access router has its own wavelength, there are some additional wavelengths that are shared among several access routers, and those additional wavelengths are assigned to those access routers cyclically. A working vacation period then corresponds to the situation that the access router has no additional wavelengths and the following period with the normal processing rate of one corresponds to the situation that the access router utilizes the additional wavelengths as well. In [9], an M/M/1 queue with exponential working vacations is studied. In [5, 6, 14], the model of [9] is generalized to the M/G/1 queue.

In current communication networks, input traffic is usually a superposition of several packet streams such as video, audio, and data traffic, which have different arrival rates and packet length distributions. We thus consider a model with several classes of customers so that such a feature can be incorporated.

Note that queues with working vacations are also applicable to modeling a class of traffic engineering schemes. For example, we consider the following scenario. The network system provides a fixed, primary route for each destination. When packet transmissions start on the primary path, the network system tries to find the lightly-loaded second path, and if such a path is found after some delay (and the sender node still transmits packets), some of packet streams served on the primary path will be re-routed to the second path. In this scenario, the working vacation period corresponds to the interval during which the system is seeking the second path, and the following normal service period corresponds to the interval after re-routing. To model this scenario, we set γ to be one, while the arrival rate in the normal service period is less than that in the working vacation. We thus generalize the conventional model with working vacations and assume that the arrival rate in the working vacation and normal service periods may be different.

Past studies on the M/G/1 queues with working vacations take an approach that the queue length process is analyzed first and then other performance measures of the model are derived from the result of the queue length. Furthermore, to make the analysis of the queue length simple, those studies assume the preemptive-repeat with resampling when working vacations end, i.e., the server always restarts the ongoing service at the beginning of normal service periods, where the new service time is resampled according to the service time distribution. On the other hand, in our model, the server continues the ongoing service at the beginning of a normal service period in a preemptive-resume manner.

In general, the queue length process in multi-class FIFO queues is not easy to analyze directly [1, 10]. Therefore, we first analyze the stationary amount of work in system and obtain its LST. Using this result, we derive the joint LST of the attained waiting time [8] and the remaining service requirement in terms of the LST of work in system. Because the server has two different processing rates, the analysis of the attained waiting time distribution in our model is not as simple as in [1, 10]. This also makes the joint LST of the attained waiting time and the remaining service requirement complicated. We classify the attained waiting time into several cases, so that the formula for the joint LST of the attained waiting time and the remaining service requirement is given in a comprehensible form.

Note that all *waiting* customers in the FIFO system arrived during the attained waiting time [1, 10]. Based on this observation, we obtain the joint transform of the queue lengths and the amounts of work in system in respective classes, which is the main result of this paper. We also derive the LSTs of the stationary distributions of waiting time and sojourn time and the joint transform of the length of a randomly chosen busy cycle and the number of customers served in the cycle.

Owing to the independent and stationary increment of Poisson arrival processes, the stationary system behavior conditioned that the server is on working vacation is equivalent to that in the corresponding queue with disasters. Therefore, as by-products, we also obtain various formulas for the multi-class FIFO M/G/1 queue with Poisson disasters.

The rest of this paper is organized as follows. In section 2, we describe the mathematical model. In section 3, the stationary amount of work in system is analyzed. In section 4, the actual waiting time and sojourn time distributions are analyzed. In section 5, we study the joint distribution of the numbers of customers and the amounts of work in system in respective classes. In section 6, we analyze the busy cycle. Finally, some concluding remarks are provided in section 7.

2. Model

We consider a stationary multi-class FIFO M/G/1 queue with exponential working vacations. When the system becomes empty, the server takes a working vacation, during which customers are served at processing rate γ ($\gamma > 0$). If the system is empty at the end of the working vacation, the server takes another working vacation. On the other hand, if a customer is being served at the end of the working vacation, the server switches its processing rate to one and continues to serve customers in a *preemptive-resume* manner, until the system becomes empty. In what follows, we call time intervals during which customers are served at processing rate one *normal service periods*. We assume that lengths of working vacations are independent and identically distributed (i.i.d.) according to an exponential distribution with parameter η ($\eta > 0$). Let V denote a random variable representing the length of a randomly chosen working vacation.

There are K classes of customers, labeled one to K . Let \mathcal{K} denote $\{1, 2, \dots, K\}$. During working vacation periods (resp. normal service periods), class k ($k \in \mathcal{K}$) customers arrive according to a Poisson process at rate $\lambda_{\text{WV},k}$ (resp. $\lambda_{\text{NP},k}$). Let λ_{WV} and λ_{NP} denote the total arrival rates during working vacation periods and during normal service periods, respectively.

$$\lambda_{\text{WV}} = \sum_{k \in \mathcal{K}} \lambda_{\text{WV},k}, \quad \lambda_{\text{NP}} = \sum_{k \in \mathcal{K}} \lambda_{\text{NP},k},$$

where we assume $\lambda_{\text{WV}} > 0$ to avoid trivialities. The amounts of service requirements of class k ($k \in \mathcal{K}$) customers who arrive during working vacation periods (resp. normal service periods) are assumed to be i.i.d. according to a general distribution function $H_{\text{WV},k}(x)$ (resp. $H_{\text{NP},k}(x)$). For each k ($k \in \mathcal{K}$), let $H_{\text{WV},k}$ (resp. $H_{\text{NP},k}$) denote a random variable representing the amount of the service requirement of a randomly chosen class k customer arriving in working vacation periods (resp. normal service periods). We denote the Laplace-Stieltjes transforms (LSTs) of $H_{\text{WV},k}$ and $H_{\text{NP},k}$ ($k \in \mathcal{K}$) by $h_{\text{WV},k}^*(s)$ and $h_{\text{NP},k}^*(s)$, respectively. Let H_{WV} (resp. H_{NP}) denote a random variable representing the amount of the service requirement brought by a customer randomly chosen among those arriving in working vacation periods (resp. normal service periods). We then define $h_{\text{WV}}^*(s)$ and $h_{\text{NP}}^*(s)$ as the LSTs of H_{WV} and H_{NP} , respectively.

$$h_{\text{WV}}^*(s) = \sum_{k \in \mathcal{K}} \frac{\lambda_{\text{WV},k}}{\lambda_{\text{WV}}} \cdot h_{\text{WV},k}^*(s), \quad h_{\text{NP}}^*(s) = \sum_{k \in \mathcal{K}} \frac{\lambda_{\text{NP},k}}{\lambda_{\text{NP}}} \cdot h_{\text{NP},k}^*(s).$$

We define $\rho_{\text{WV},k}$ and $\rho_{\text{NP},k}$ ($k \in \mathcal{K}$) as $\rho_{\text{WV},k} = \lambda_{\text{WV},k} \text{E}[H_{\text{WV},k}]$ and $\rho_{\text{NP},k} = \lambda_{\text{NP},k} \text{E}[H_{\text{NP},k}]$, respectively. Let $\rho_{\text{WV}} = \sum_{k \in \mathcal{K}} \rho_{\text{WV},k}$ and $\rho_{\text{NP}} = \sum_{k \in \mathcal{K}} \rho_{\text{NP},k}$. In what follows, we assume $\rho_{\text{NP}} < 1$. The service discipline is assumed to be FIFO, unless otherwise mentioned, and services are nonpreemptive.

Remark 1. When $\eta > 0$, the system is stable if and only if $\rho_{\text{WV}} < \infty$ and $\rho_{\text{NP}} < 1$. To see this, consider the length \bar{C} of an interval between successive starts of working vacations. Note that the system is stable if and only if $\text{E}[\bar{C}] < \infty$. By definition, \bar{C} can be divided into two parts, one of which is the length of a working vacation period \bar{C}_{WV} with mean $1/\eta$ and the other is the length of the following normal service period \bar{C}_{NP} . Let U_{WV}^{E} denote the total amount of work in system at the end of the working vacation period. If $\rho_{\text{WV}} < \infty$ and $\rho_{\text{NP}} < 1$, the stability of the system is ensured because $\text{E}[\bar{C}_{\text{NP}}] = \text{E}[U_{\text{WV}}^{\text{E}}]/(1 - \rho_{\text{NP}})$ and

$$\text{E}[\bar{C}] = \frac{1}{\eta} + \frac{\text{E}[U_{\text{WV}}^{\text{E}}]}{1 - \rho_{\text{NP}}} \leq \frac{1}{\eta} + \frac{\rho_{\text{WV}}/\eta}{1 - \rho_{\text{NP}}} < \infty,$$

where the first inequality comes from the fact that in every sample path, U_{WV}^{E} is bounded above by the total amount of work brought in the working vacation period.

Conversely, if the system is stable, $\text{E}[U_{\text{WV}}^{\text{E}}] < \infty$ holds, and therefore $\rho_{\text{WV}} < \infty$. Furthermore, in an ordinary M/G/1 queue, the first passage time to the idle state with finite initial workload is finite if and only if the traffic intensity is less than one. Therefore, we have $\rho_{\text{WV}} < \infty$ and $\rho_{\text{NP}} < 1$ if the system is stable.

Remark 2. If we ignore customer classes, the above model is reduced to a single-class M/G/1 with exponential working vacations characterized by arrival rates λ_{WV} and λ_{NP} , amounts of service requirements H_{WP} and H_{NP} , processing rate γ during working vacation periods, and exponential lengths of working vacation periods with mean $1/\eta$.

3. Total Work in System

In this section, we discuss the total amount of work in system in steady state. Let U denote the total amount of work in system. We define U_{WV} (resp. U_{NP}) as the conditional total amount of work in system given the server being on working vacation (resp. being in a normal service period). Let $u^*(s)$, $u_{\text{WV}}^*(s)$, and $u_{\text{NP}}^*(s)$ denote the LSTs of U , U_{WV} , and U_{NP} , respectively. We then have

$$u^*(s) = P_{\text{WV}} \cdot u_{\text{WV}}^*(s) + P_{\text{NP}} \cdot u_{\text{NP}}^*(s), \quad (1)$$

where P_{WV} (resp. P_{NP}) denotes the time-average probability of the server being on working vacation (resp. being in a normal service period).

Let U_{WV}^{E} denote the total amount of work in system at the end of a working vacation. We denote the LST of U_{WV}^{E} by $u_{\text{WV,E}}^*(s)$. Consider a censored workload process by removing all normal service periods. In the resulting process, the ends of working vacations occur according to a Poisson process with rate η . Therefore, owing to PASTA [13], we have

$$u_{\text{WV,E}}^*(s) = u_{\text{WV}}^*(s), \quad \text{E}[U_{\text{WV}}^{\text{E}}] = \text{E}[U_{\text{WV}}]. \quad (2)$$

We then have the following two lemmas, whose proofs are given in Appendices A and B, respectively.

Lemma 1. $u_{\text{NP}}^*(s)$ is given by

$$u_{\text{NP}}^*(s) = \frac{1 - u_{\text{WV}}^*(s)}{s\text{E}[U_{\text{WV}}]} \cdot u_{\text{M/G/1}}^*(s), \quad (3)$$

where $u_{\text{M/G/1}}^*(s)$ denotes the LST of the amount of work in system in an ordinary M/G/1 queue and it is given by

$$u_{\text{M/G/1}}^*(s) = \frac{(1 - \rho_{\text{NP}})s}{s - \lambda_{\text{NP}} + \lambda_{\text{NP}}h_{\text{NP}}^*(s)}. \quad (4)$$

Lemma 2. P_{WV} and P_{NP} are given by

$$P_{\text{WV}} = \frac{1 - \rho_{\text{NP}}}{1 - \rho_{\text{NP}} + \eta\text{E}[U_{\text{WV}}]}, \quad P_{\text{NP}} = \frac{\eta\text{E}[U_{\text{WV}}]}{1 - \rho_{\text{NP}} + \eta\text{E}[U_{\text{WV}}]}, \quad (5)$$

respectively.

With Lemma 1, $u^*(s)$ is given in terms of $u_{\text{WV}}^*(s)$ and $E[U_{\text{WV}}]$.

$$u^*(s) = P_{\text{WV}} \cdot u_{\text{WV}}^*(s) + P_{\text{NP}} \cdot \frac{1 - u_{\text{WV}}^*(s)}{sE[U_{\text{WV}}]} \cdot u_{\text{M/G/1}}^*(s), \tag{6}$$

where P_{WV} and P_{NP} are given in (5).

We now characterize $u_{\text{WV}}^*(s)$. Note here that the conditional total amount U_{WV} of work in system is equivalent to that in the corresponding M/G/1 queue with Poisson disasters [3, 15]. Therefore we can readily obtain $u_{\text{WV}}^*(s)$ using the results in [3, 15]. Note that a similar observation with respect to the queue length has been made in [5] for a single-class M/G/1 queue with exponential working vacations.

Lemma 3. $u_{\text{WV}}^*(s)$ and $E[U_{\text{WV}}]$ are given by

$$u_{\text{WV}}^*(s) = \frac{(1 - \nu)s - \eta/\gamma}{s - \lambda_{\text{WV}}/\gamma + (\lambda_{\text{WV}}/\gamma)h_{\text{WV}}^*(s) - \eta/\gamma}, \quad E[U_{\text{WV}}] = \frac{\rho_{\text{WV}} - \gamma\nu}{\eta}, \tag{7}$$

respectively, where ν denotes the conditional steady state probability that the server is busy given that it is on working vacation. Note that ν is given by

$$\nu = \frac{(1 - r)\lambda_{\text{WV}}}{(1 - r)\lambda_{\text{WV}} + \eta}, \tag{8}$$

where r ($r > 0$) denotes the unique real root of the following equation.

$$z = h_{\text{WV}}^*(\eta/\gamma + \lambda_{\text{WV}}/\gamma - (\lambda_{\text{WV}}/\gamma)z), \quad |z| < 1. \tag{9}$$

The proof of Lemma 3 is given in Appendix C.

Remark 3 (Remark 2.2 in [15]). *The solution r of (9) represents the probability that a randomly chosen busy period starting in a working vacation ends within the working vacation. To see this, consider an M/G/1 queue with arrival rate λ_{WV} , the LST $h_{\text{WV}}^*(s)$ of service requirements of customers, and the processing rate γ . The LST $\theta^*(s)$ of the lengths of busy periods is then given by $\theta^*(s) = h_{\text{WV}}^*(s/\gamma + \lambda_{\text{WV}}/\gamma - (\lambda_{\text{WV}}/\gamma)\theta^*(s))$. Comparing this with (9), we have $r = \theta^*(\eta) > 0$.*

Rearranging terms on the right side of $u_{\text{WV}}^*(s)$ in (7) yields

$$u_{\text{WV}}^*(s) = \frac{1 - \nu}{1 - \nu \tilde{f}_{\text{WV}}^*(s)}, \tag{10}$$

where $\tilde{f}_{\text{WV}}^*(s)$ is given by

$$\tilde{f}_{\text{WV}}^*(s) = \frac{h_{\text{WV}}^*(s) - r}{(\gamma\nu/\lambda_{\text{WV}})\{\eta/\gamma + \lambda_{\text{WV}}/\gamma - (\lambda_{\text{WV}}/\gamma)r - s\}}. \tag{11}$$

Remark 4. *Theorem 2 in [3] shows that $\tilde{f}_{\text{WV}}^*(s)$ represents the LST of the remaining service requirement \tilde{F}_{WV} of a randomly chosen customer present in working vacation periods when customers are served on a LIFO preemptive resume basis. Note that (7) and (10) imply*

$$E[\tilde{F}_{\text{WV}}] = \frac{1 - \nu}{\nu} \cdot E[U_{\text{WV}}] = \frac{1 - \nu}{\nu} \cdot \frac{\rho_{\text{WV}} - \gamma\nu}{\eta}. \tag{12}$$

Theorem 1. $u^*(s)$ is given by

$$u^*(s) = u_{\text{WV}}^*(s) \cdot \left(P_{\text{WV}} + P_{\text{NP}} \cdot \frac{1 - \tilde{f}_{\text{WV}}^*(s)}{sE[\tilde{F}_{\text{WV}}]} \cdot u_{\text{M/G/1}}^*(s) \right), \quad (13)$$

where $u_{\text{M/G/1}}^*(s)$, $u_{\text{WV}}^*(s)$, $\tilde{f}_{\text{WV}}^*(s)$, and $E[\tilde{F}_{\text{WV}}]$ are given by (4), (7), (11), and (12), respectively, and P_{WV} and P_{NP} are given by

$$P_{\text{WV}} = \frac{1 - \rho_{\text{NP}}}{1 - \rho_{\text{NP}} + \rho_{\text{WV}} - \gamma\nu}, \quad P_{\text{NP}} = \frac{\rho_{\text{WV}} - \gamma\nu}{1 - \rho_{\text{NP}} + \rho_{\text{WV}} - \gamma\nu}, \quad (14)$$

respectively.

Proof. It follows from (10) and (12) that

$$\frac{1 - u_{\text{WV}}^*(s)}{sE[U_{\text{WV}}]} = \frac{1 - \nu}{1 - \nu \tilde{f}_{\text{WV}}^*(s)} \cdot \frac{1 - \tilde{f}_{\text{WV}}^*(s)}{sE[\tilde{F}_{\text{WV}}]} = u_{\text{WV}}^*(s) \cdot \frac{1 - \tilde{f}_{\text{WV}}^*(s)}{sE[\tilde{F}_{\text{WV}}]}, \quad (15)$$

Substituting (15) into (6) yields (13). Further (14) follows from (5) and (7). \square

Remark 5. Theorem 1 shows that U is stochastically decomposed into two independent nonnegative random variables, i.e., $U = U_{\text{WV}} + U_{\text{I}}$, where the LST of non-negative random variable U_{I} is given by

$$u_{\text{I}}^*(s) = P_{\text{WV}} + P_{\text{NP}} \cdot \frac{1 - \tilde{f}_{\text{WV}}^*(s)}{sE[\tilde{F}_{\text{WV}}]} \cdot u_{\text{M/G/1}}^*(s).$$

4. Waiting Time and Sojourn Time

In this section, we consider the actual waiting time and sojourn time distributions of class k ($k \in \mathcal{K}$) customers in steady state, assuming the FIFO service discipline. Let W_k ($k \in \mathcal{K}$) denote the waiting time of a randomly chosen class k customer. For each k ($k \in \mathcal{K}$), we define $W_{\text{WV},k}$ (resp. $W_{\text{NP},k}$) as the waiting time of a randomly chosen class k customer arriving in a working vacation period (resp. a normal service period). Let $w_k^*(s)$, $w_{\text{WV},k}^*(s)$, and $w_{\text{NP},k}^*(s)$ ($k \in \mathcal{K}$) denote the LSTs of W_k , $W_{\text{WV},k}$, and $W_{\text{NP},k}$, respectively. Similarly, let Q_k ($k \in \mathcal{K}$) denote the stationary sojourn time of class k customers. For each k ($k \in \mathcal{K}$), we define $Q_{\text{WV},k}$ (resp. $Q_{\text{NP},k}$) as the sojourn time of a randomly chosen class k customer arriving in a working vacation period (resp. a normal service period). Let $q_k^*(s)$, $q_{\text{WV},k}^*(s)$, and $q_{\text{NP},k}^*(s)$ ($k \in \mathcal{K}$) denote the LSTs of Q_k , $Q_{\text{WV},k}$, and $Q_{\text{NP},k}$, respectively.

For each k ($k \in \mathcal{K}$), we define $P_{\text{WV},k}^{\text{A}}$ (resp. $P_{\text{NP},k}^{\text{A}}$) as the probability that a randomly chosen class k customer finds the server being on working vacation (resp. being in a normal service period) upon arrival. By definition, $w_k^*(s)$ and $q_k^*(s)$ ($k \in \mathcal{K}$) are given by

$$w_k^*(s) = P_{\text{WV},k}^{\text{A}} \cdot w_{\text{WV},k}^*(s) + P_{\text{NP},k}^{\text{A}} \cdot w_{\text{NP},k}^*(s), \quad (16)$$

$$q_k^*(s) = P_{\text{WV},k}^{\text{A}} \cdot q_{\text{WV},k}^*(s) + P_{\text{NP},k}^{\text{A}} \cdot q_{\text{NP},k}^*(s), \quad (17)$$

respectively. Because class k customers arrive according to a Poisson process with rate $\lambda_{\text{WV},k}$ during working vacation periods and rate $\lambda_{\text{NP},k}$ during normal service periods, $P_{\text{WV},k}^{\text{A}}$ and $P_{\text{NP},k}^{\text{A}}$ satisfy

$$\frac{P_{\text{WV},k}^{\text{A}}}{P_{\text{NP},k}^{\text{A}}} = \frac{\lambda_{\text{WV},k} P_{\text{WV}}}{\lambda_{\text{NP},k} P_{\text{NP}}}.$$

Therefore, using $P_{WV,k}^A + P_{NP,k}^A = 1$, we obtain

$$P_{WV,k}^A = \frac{\lambda_{WV,k}P_{WV}}{\lambda_{WV,k}P_{WV} + \lambda_{NP,k}P_{NP}} = \frac{\lambda_{WV,k}(1 - \rho_{NP})}{\lambda_{WV,k}(1 - \rho_{NP}) + \lambda_{NP,k}(\rho_{WV} - \gamma\nu)}, \quad (18)$$

$$P_{NP,k}^A = \frac{\lambda_{NP,k}P_{NP}}{\lambda_{WV,k}P_{WV} + \lambda_{NP,k}P_{NP}} = \frac{\lambda_{NP,k}(\rho_{WV} - \gamma\nu)}{\lambda_{WV,k}(1 - \rho_{NP}) + \lambda_{NP,k}(\rho_{WV} - \gamma\nu)}. \quad (19)$$

Both W_k and Q_k ($k \in \mathcal{K}$) are considered as the processing time of a certain amount of work. More specifically, W_k ($k \in \mathcal{K}$) corresponds to the stationary processing time of work in system seen by an arriving customer of class k . On the other hand, Q_k ($k \in \mathcal{K}$) corresponds to the stationary processing time of the sum of work in system seen by an arriving customer of class k and his/her service requirement. To treat W_k and Q_k in a unified way, we define $T_{WV}(U_X)$ (resp. $T_{NP}(U_X)$) as the processing time of the amount U_X of work conditioned that the server is on working vacation (resp. in a normal service period) when its processing starts, where U_X is assumed to be a nonnegative random variable whose distribution function and LST are given by $U_X(x)$ and $u_X^*(s)$, respectively. Because the processing rate in $T_{WV}(U_X)$ may change from γ to one, we divide $T_{WV}(U_X)$ into two parts, $T_{WV}^{(\gamma)}(U_X)$ and $T_{WV}^{(1)}(U_X)$, where $T_{WV}^{(\gamma)}(U_X)$ (resp. $T_{WV}^{(1)}(U_X)$) is defined as the length of a subinterval in $T_{WV}(U_X)$, during which the processing rate is equal to γ (resp. one). By definition, $T_{WV}(U_X) = T_{WV}^{(\gamma)}(U_X) + T_{WV}^{(1)}(U_X)$, where $T_{WV}^{(\gamma)}(U_X) > 0$ for $U_X > 0$, and $T_{WV}^{(1)}(U_X) \geq 0$. We then define $\phi_{WV}^{**}(\omega, s | U_X)$ and $\phi_{NP}^*(s | U_X)$ as

$$\phi_{WV}^{**}(\omega, s | U_X) = E \left[e^{-\omega T_{WV}^{(\gamma)}(U_X)} e^{-s T_{WV}^{(1)}(U_X)} \right], \quad \phi_{NP}^*(s | U_X) = E \left[e^{-s T_{NP}(U_X)} \right],$$

respectively.

Lemma 4. $\phi_{WV}^{**}(\omega, s | U_X)$ and $\phi_{NP}^*(s | U_X)$ are given by

$$\phi_{WV}^{**}(\omega, s | U_X) = u_X^* \left(\frac{\omega + \eta}{\gamma} \right) + \frac{u_X^*(s) - u_X^* \left(\frac{\omega + \eta}{\gamma} \right)}{(\gamma/\eta) \{ (\omega + \eta)/\gamma - s \}}, \quad \phi_{NP}^*(s | U_X) = u_X^*(s),$$

respectively.

Proof. We first consider $\phi_{NP}^*(s | U_X)$. When the processing of U_X starts in a normal service period, the processing rate is fixed to one throughout its processing. We then have $T_{NP}(U_X) = U_X/1$, from which $\phi_{NP}^*(s | U_X) = u_X^*(s)$ follows. On the other hand, when the processing of U_X starts in a working vacation period, we have

$$(T_{WV}^{(\gamma)}(U_X), T_{WV}^{(1)}(U_X)) = \begin{cases} (\frac{U_X}{\gamma}, 0), & \tilde{V}_S > \frac{U_X}{\gamma}, \\ (\tilde{V}_S, U_X - \gamma\tilde{V}_S), & \tilde{V}_S \leq \frac{U_X}{\gamma}, \end{cases}$$

where \tilde{V}_S denotes the remaining length of the working vacation when the processing starts. Owing to the memoryless property of the exponential distribution, \tilde{V}_S is exponentially distributed with parameter η . We then have

$$\phi_{WV}^{**}(\omega, s | U_X) = \int_0^\infty dU_X(x) \left[e^{-\eta(x/\gamma)} \cdot e^{-\omega(x/\gamma)} \right]$$

$$+ \left\{ 1 - e^{-\eta(x/\gamma)} \right\} \cdot \int_0^{x/\gamma} \frac{\eta e^{-\eta\tau} d\tau}{1 - e^{-\eta(x/\gamma)}} \cdot e^{-\omega\tau} e^{-s(x-\gamma\tau)} \Bigg], \quad (20)$$

from which the expression of $\phi_{\text{WV}}^{**}(\omega, s \mid U_X)$ follows. □

Using Lemma 4, $E[T_{\text{WV}}^{(\gamma)}(U_X)]$, $E[T_{\text{WV}}^{(1)}(U_X)]$, and $E[T_{\text{NP}}(U_X)]$ are obtained to be

$$E \left[T_{\text{WV}}^{(\gamma)}(U_X) \right] = (-1) \cdot \lim_{\omega \rightarrow 0} \frac{d}{d\omega} [\phi_{\text{WV}}^{**}(\omega, 0 \mid U_X)] = \frac{1 - u_X^*(\eta/\gamma)}{\eta}, \quad (21)$$

$$E \left[T_{\text{WV}}^{(1)}(U_X) \right] = (-1) \cdot \lim_{s \rightarrow 0} \frac{d}{ds} [\phi_{\text{WV}}^{**}(0, s \mid U_X)] = E[U_X] - \gamma \cdot \frac{1 - u_X^*(\eta/\gamma)}{\eta}, \quad (22)$$

$$E[T_{\text{NP}}(U_X)] = E[U_X]. \quad (23)$$

We now turn our attention to the waiting time distribution. Consider the censored process obtained by removing all normal service periods. In the resulting process, class k customers arrive according to a Poisson process. Owing to PASTA, the conditional amount of work in system seen by a randomly chosen class k customer arriving in a working vacation period has the same distribution as U_{WV} . Therefore the conditional waiting time distributions are identical among classes. Similarly, the conditional amount of work in system seen by class k ($k \in \mathcal{K}$) customers arriving in normal service periods has the same distribution as U_{NP} . Thus, the conditional waiting time distributions are also identical among classes.

Let $W_{\text{WV}}^{(\gamma)}$ (resp. $W_{\text{WV}}^{(1)}$) denote the length of an interval during which a randomly chosen customer waits for his/her service in a working vacation period (resp. normal service period), given that the customer arrived in the working vacation period. By definition, $W_{\text{WV},k} = W_{\text{WV}}^{(\gamma)} + W_{\text{WV}}^{(1)}$ for all k ($k \in \mathcal{K}$). Also, let W_{NP} denote the conditional waiting time of a randomly chosen customer given that the customer arrives in a normal service period. We then define $w_{\text{WV}}^{**}(\omega, s)$ as the joint LST $E[\exp(-\omega W_{\text{WV}}^{(\gamma)}) \exp(-s W_{\text{WV}}^{(1)})]$ of $W_{\text{WV}}^{(\gamma)}$ and $W_{\text{WV}}^{(1)}$, and $w_{\text{NP}}^*(s)$ as the LST of W_{NP} .

Theorem 2. $w_{\text{WV}}^{**}(\omega, s)$ and $w_{\text{NP}}^*(s)$ are given by

$$w_{\text{WV}}^{**}(\omega, s) = u_{\text{WV}}^* \left(\frac{\omega + \eta}{\gamma} \right) + \frac{u_{\text{WV}}^*(s) - u_{\text{WV}}^* \left(\frac{\omega + \eta}{\gamma} \right)}{(\gamma/\eta)\{(\omega + \eta)/\gamma - s\}}, \quad w_{\text{NP}}^*(s) = u_{\text{NP}}^*(s),$$

respectively.

Proof. By definition, $w_{\text{WV}}^{**}(\omega, s) = \phi_{\text{WV}}^{**}(\omega, s \mid U_{\text{WV}})$ and $w_{\text{NP}}^*(s) = \phi_{\text{NP}}^*(s \mid U_{\text{NP}})$, so that Theorem 2 immediately follows from Lemma 4. □

Because $W_{\text{WV},k} = W_{\text{WV}}^{(\gamma)} + W_{\text{WV}}^{(1)}$ and $W_{\text{NP},k} = W_{\text{NP}}$ for all k ($k \in \mathcal{K}$),

$$w_{\text{WV},k}^*(s) = w_{\text{WV}}^{**}(s, s), \quad w_{\text{NP},k}^*(s) = w_{\text{NP}}^*(s), \quad \forall k \in \mathcal{K}.$$

Thus the LST $w_k^*(s)$ ($k \in \mathcal{K}$) of the waiting time distribution of class k customers is obtained by (16). In particular, the mean waiting time is given by

$$E[W_k] = P_{\text{WV},k}^A \cdot \left[\frac{(1 - \gamma)(1 - u_{\text{WV}}^*(\eta/\gamma))}{\eta} + E[U_{\text{WV}}] \right] + P_{\text{NP},k}^A \cdot E[U_{\text{NP}}],$$

where $E[U_{NP}]$ denotes the mean amount of conditional work in system given the system being in a normal service period and it is obtained from (3) and Lemma 3.

$$E[U_{NP}] = \frac{\rho_{WV} - \gamma}{\eta} + \frac{\lambda_{WV}E[H_{WV}^2]}{2(\rho_{WV} - \gamma\nu)} + \frac{\lambda_{NP}E[H_{NP}^2]}{2(1 - \rho_{NP})}.$$

Next we consider the sojourn time distribution. For each k ($k \in \mathcal{K}$), let $Q_{WV,k}^{(\gamma)}$ (resp. $Q_{WV,k}^{(1)}$) denote the length of time during which a randomly chosen class k customer spends in a working vacation period (resp. a normal service period), given that the customer arrives in the working vacation period. By definition, $Q_{WV,k}^{(\gamma)} > 0$, $Q_{WV,k}^{(1)} \geq 0$, and $Q_{WV,k} = Q_{WV,k}^{(\gamma)} + Q_{WV,k}^{(1)}$. We define $q_{WV,k}^{**}(\omega, s)$ ($k \in \mathcal{K}$) as the joint LST $E[\exp(-\omega Q_{WV,k}^{(\gamma)}) \exp(-s Q_{WV,k}^{(1)})]$ of $Q_{WV,k}^{(\gamma)}$ and $Q_{WV,k}^{(1)}$, and $q_{NP,k}^*(s)$ ($k \in \mathcal{K}$) as the LST of $Q_{NP,k}$.

Theorem 3. $q_{WV,k}^{**}(\omega, s)$ and $q_{NP,k}^*(s)$ ($k \in \mathcal{K}$) are given by

$$\begin{aligned} q_{WV,k}^{**}(\omega, s) &= u_{WV}^* \left(\frac{\omega + \eta}{\gamma} \right) \cdot h_{WV,k}^* \left(\frac{\omega + \eta}{\gamma} \right) \\ &\quad + \frac{u_{WV}^*(s) \cdot h_{WV,k}^*(s) - u_{WV}^* \left(\frac{\omega + \eta}{\gamma} \right) \cdot h_{WV,k}^* \left(\frac{\omega + \eta}{\gamma} \right)}{(\gamma/\eta)\{(\omega + \eta)/\gamma - s\}}, \\ q_{NP,k}^*(s) &= u_{NP}^*(s) \cdot h_{NP,k}^*(s), \end{aligned}$$

respectively.

Proof. By definition, $q_{WV,k}^{**}(\omega, s) = \phi_{WV}^{**}(\omega, s \mid U_{WV} + H_{WV,k})$, and $q_{NP,k}^*(s) = \phi_{NP}^*(s \mid U_{NP} + H_{NP,k})$. Theorem 3 then follows from Lemma 4. \square

Note that $q_{WV,k}^*(s) = q_{WV,k}^{**}(s, s)$ ($k \in \mathcal{K}$). Thus the LST $q_k^*(s)$ ($k \in \mathcal{K}$) of the sojourn time distribution of class k customers is obtained by (17). In particular, the mean sojourn time is given by

$$E[Q_k] = P_{WV,k}^A \cdot E[Q_{WV,k}] + P_{NP,k}^A \cdot E[Q_{NP,k}],$$

where

$$\begin{aligned} E[Q_{WV,k}] &= \frac{(1 - \gamma)(1 - u_{WV}^*(\eta/\gamma)h_{WV,k}^*(\eta/\gamma))}{\eta} + E[U_{WV}] + E[H_{WV,k}], \\ E[Q_{NP,k}] &= E[U_{NP}] + E[H_{NP,k}]. \end{aligned}$$

5. Joint Distribution of Queue Lengths and Work in System

In this section, we consider the joint distribution of the numbers of customers and the amounts of work in system in respective classes. To do so, we first derive the joint LST of the attained waiting time and the remaining amount of service requirement of a class k customer being served. With this result, the joint distributions are derived.

For each k ($k \in \mathcal{K}$), let $\sigma_{WV,k}^{(\gamma)}$ (resp. $\sigma_{WV,k}^{(1)}$) denote the time-average probability that class k customers, who arrived in working vacation periods, are being served in working vacation periods (resp. in normal service periods). Also, let $\sigma_{NP,k}^{(1)}$ ($k \in \mathcal{K}$) denote the time-average probability that class k customers arriving in normal service periods are being served.

Lemma 5. $\sigma_{\text{WV},k}^{(\gamma)}$, $\sigma_{\text{WV},k}^{(1)}$, and $\sigma_{\text{NP},k}^{(1)}$ ($k \in \mathcal{K}$) are given by

$$\sigma_{\text{WV},k}^{(\gamma)} = P_{\text{WV}} \cdot \nu \cdot \frac{\lambda_{\text{WV},k}(1 - h_{\text{WV},k}^*(\eta/\gamma))}{\lambda_{\text{WV}}(1 - h_{\text{WV}}^*(\eta/\gamma))}, \quad (24)$$

$$\sigma_{\text{WV},k}^{(1)} = P_{\text{WV}} \cdot \left[\rho_{\text{WV},k} - \gamma \nu \cdot \frac{\lambda_{\text{WV},k}(1 - h_{\text{WV},k}^*(\eta/\gamma))}{\lambda_{\text{WV}}(1 - h_{\text{WV}}^*(\eta/\gamma))} \right], \quad (25)$$

$$\sigma_{\text{NP},k}^{(1)} = P_{\text{NP}} \cdot \rho_{\text{NP},k}, \quad (26)$$

respectively.

The proof of Lemma 5 is given in Appendix D.

Remark 6. Let σ denote the utilization factor, i.e., the time-average probability that customers are being served. Recall that ν in (8) represents the conditional probability of the server being busy given that the server is on working vacation. We then have

$$\sigma = 1 - P_{\text{WV}} \cdot (1 - \nu) = P_{\text{WV}} \cdot \nu + P_{\text{NP}} = \frac{(1 - \rho_{\text{NP}})\nu + \rho_{\text{WV}} - \gamma \nu}{1 - \rho_{\text{NP}} + \rho_{\text{WV}} - \gamma \nu}.$$

Furthermore, using Lemma 5, we can verify

$$\sum_{k \in \mathcal{K}} \sigma_{\text{WV},k}^{(\gamma)} = P_{\text{WV}} \cdot \nu, \quad \sum_{k \in \mathcal{K}} (\sigma_{\text{WV},k}^{(1)} + \sigma_{\text{NP},k}^{(1)}) = P_{\text{NP}}.$$

We now consider the attained waiting time [8], which is defined as the length of time spent by a customer being served (if any) in the system. When the system is empty, the attained waiting time is defined to be zero. Note that under the FIFO service discipline, all waiting customers in the system arrived during the attained waiting time.

For later use, we divide the attained waiting time into two parts: One is the (sub)interval in working vacation periods and the other is the (sub)interval in normal service periods. Let $A_{\text{WV},k}^{(\gamma)}$ ($k \in \mathcal{K}$) denote the length of time in the attained waiting time, during which the server was on working vacation, given that a class k customer is being served. Furthermore, for each k ($k \in \mathcal{K}$), let $A_{\text{WV},k}^{(1)}$ (resp. $A_{\text{NP},k}^{(1)}$) denote the length of time in the attained waiting time, during which the server worked in a normal service period, given that a class k customer, who arrived in a working vacation period (resp. a normal service period), is being served. For a class k ($k \in \mathcal{K}$) customer being served, let \tilde{H}_k denote the remaining amount of his/her service requirement. We then define the following joint LSTs:

$$a_{\text{WV},\text{WV},k}^{**}(\omega_k, \alpha_k) = \mathbb{E} \left[e^{-\omega_k A_{\text{WV},k}^{(\gamma)}} e^{-\alpha_k \tilde{H}_k} \left| \begin{array}{l} \text{a class } k \text{ customer is being served} \\ \text{at processing rate } \gamma \end{array} \right. \right],$$

$$a_{\text{WV},\text{NP},k}^{***}(\omega_k, s_k, \alpha_k) = \mathbb{E} \left[e^{-\omega_k A_{\text{WV},k}^{(\gamma)}} e^{-s_k A_{\text{WV},k}^{(1)}} e^{-\alpha_k \tilde{H}_k} \left| \begin{array}{l} \text{a class } k \text{ customer, who arrived} \\ \text{in a working vacation period, is} \\ \text{being served at processing rate one} \end{array} \right. \right],$$

$$a_{\text{NP},k}^{**}(s_k, \alpha_k) = \mathbb{E} \left[e^{-s_k A_{\text{NP},k}^{(1)}} e^{-\alpha_k \tilde{H}_k} \left| \begin{array}{l} \text{a class } k \text{ customer, who arrived in a} \\ \text{normal service period, is being served} \end{array} \right. \right].$$

See Figures 1–4, where Figure 1 corresponds to $a_{\text{WV},\text{WV},k}^{**}(\omega_k, \alpha_k)$, Figures 2 and 3 correspond to $a_{\text{WV},\text{NP},k}^{***}(\omega_k, s_k, \alpha_k)$, and Figure 4 corresponds to $a_{\text{NP},k}^{**}(s_k, \alpha_k)$.

Moreover, for each k ($k \in \mathcal{K}$), let $H_{\text{WV},k}^{(\gamma)}$ (resp. $H_{\text{WV},k}^{(1)}$) denote the lengths of time during which a class k customer, who started his/her service in a working vacation period, is served

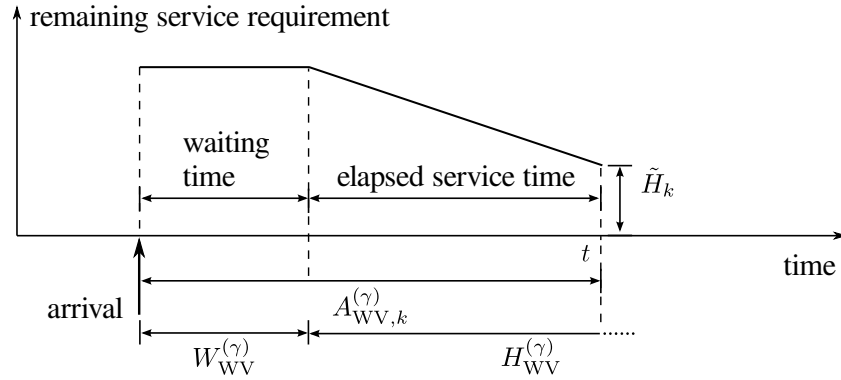


Figure 1: Attained waiting time of a class k customer in a working vacation period

during the working vacation period (resp. the subsequent normal service period). We then define $\hat{h}_{WV,k}^{**}(\omega, s)$ as the joint LST of $H_{WV,k}^{(\gamma)}$ and $H_{WV,k}^{(1)}$. Using Lemma 4, we obtain

$$\begin{aligned} \hat{h}_{WV,k}^{**}(\omega, s) &= \mathbb{E} \left[e^{-\omega H_{WV,k}^{(\gamma)}} e^{-s H_{WV,k}^{(1)}} \right] = \phi_{WV}^{**}(\omega, s \mid H_{WV,k}) \\ &= h_{WV,k}^* \left(\frac{\omega + \eta}{\gamma} \right) + \frac{h_{WV,k}^*(s) - h_{WV,k}^* \left(\frac{\omega + \eta}{\gamma} \right)}{(\gamma/\eta) \{ (\omega + \eta)/\gamma - s \}}. \end{aligned}$$

We then have the following theorem, whose proof is provided in Appendix E.

Theorem 4. $a_{WV,WV,k}^{**}(\omega_k, \alpha_k)$, $a_{WV,NP,k}^{***}(\omega_k, s_k, \alpha_k)$, and $a_{NP,k}^{**}(s_k, \alpha_k)$ are given by

$$a_{WV,WV,k}^{**}(\omega_k, \alpha_k) = \frac{(1/\eta) u_{WV}^* \left(\frac{\omega_k + \eta}{\gamma} \right) \cdot h_{WV,k}^*(\alpha_k) - h_{WV,k}^* \left(\frac{\omega_k + \eta}{\gamma} \right)}{\mathbb{E} \left[Q_{WV,k}^{(\gamma)} - W_{WV,k}^{(\gamma)} \right] \cdot (\gamma/\eta) \{ (\omega_k + \eta)/\gamma - \alpha_k \}}, \quad (27)$$

$$\begin{aligned} a_{WV,NP,k}^{***}(\omega_k, s_k, \alpha_k) &= \frac{1}{\mathbb{E} \left[Q_{WV,k}^{(1)} - W_{WV,k}^{(1)} \right]} \left[u_{WV}^* \left(\frac{\omega_k + \eta}{\gamma} \right) \frac{\hat{h}_{WV,k}^{**}(\omega_k, \alpha_k) - \hat{h}_{WV,k}^{**}(\omega_k, s_k)}{s_k - \alpha_k} \right. \\ &\quad \left. + \frac{u_{WV}^*(s_k) - u_{WV}^* \left(\frac{\omega_k + \eta}{\gamma} \right)}{(\gamma/\eta) \{ (\omega_k + \eta)/\gamma - s_k \}} \cdot \frac{h_{WV,k}^*(\alpha_k) - h_{WV,k}^*(s_k)}{s_k - \alpha_k} \right], \quad (28) \end{aligned}$$

$$a_{NP,k}^{**}(s_k, \alpha_k) = u_{NP}^*(s_k) \cdot \frac{h_{NP,k}^*(\alpha_k) - h_{NP,k}^*(s_k)}{\mathbb{E} [H_{NP,k}](s_k - \alpha_k)}, \quad (29)$$

respectively, where $\mathbb{E}[Q_{WV,k}^{(\gamma)} - W_{WV,k}^{(\gamma)}]$ and $\mathbb{E}[Q_{WV,k}^{(1)} - W_{WV,k}^{(1)}]$ are given in (47) and (48), respectively.

With Theorems 2, 3, and 4, we can verify that $a_{WV,WV,k}^{**}(\omega_k, \alpha_k)$ and $a_{WV,NP,k}^{***}(\omega_k, s_k, \alpha_k)$ are represented in terms of $w_{WV}^{**}(\omega, s)$ and $q_{WV,k}^{**}(\omega, s)$.

Corollary 1. $a_{WV,WV,k}^{**}(\omega_k, \alpha_k)$ and $a_{WV,NP,k}^{***}(\omega_k, s_k, \alpha_k)$ are given by

$$a_{WV,WV,k}^{**}(\omega_k, \alpha_k) = \frac{w_{WV}^{**}(\omega_k, \alpha_k) h_{WV,k}^*(\alpha_k) - q_{WV,k}^{**}(\omega_k, \alpha_k)}{(\omega_k - \gamma \alpha_k) \mathbb{E} \left[Q_{WV,k}^{(\gamma)} - W_{WV,k}^{(\gamma)} \right]},$$

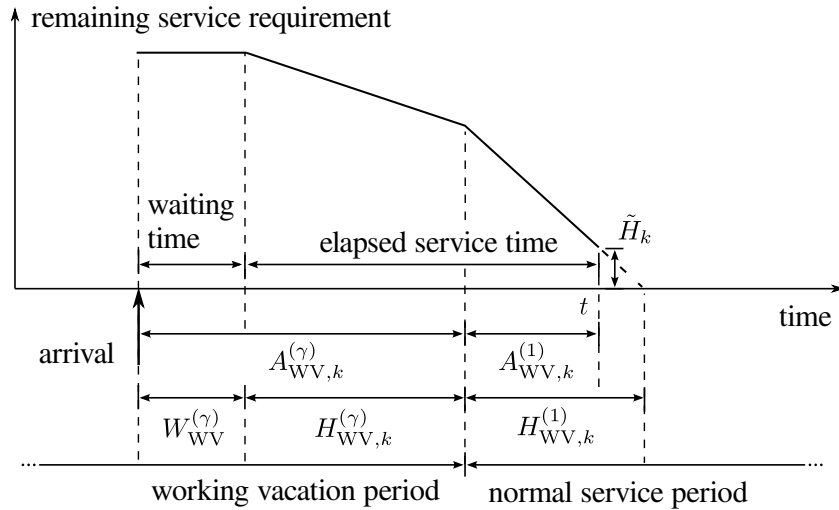


Figure 2: Attained waiting time of a class k customer who started his/her service in a working vacation period and will end his/her service in a normal service period

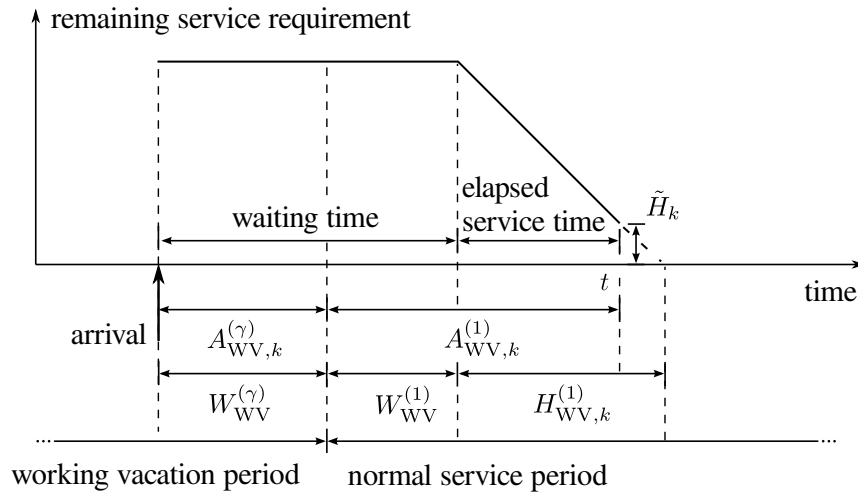


Figure 3: Attained waiting time of a class k customer who arrived in a working vacation period and started his/her service in a normal service period

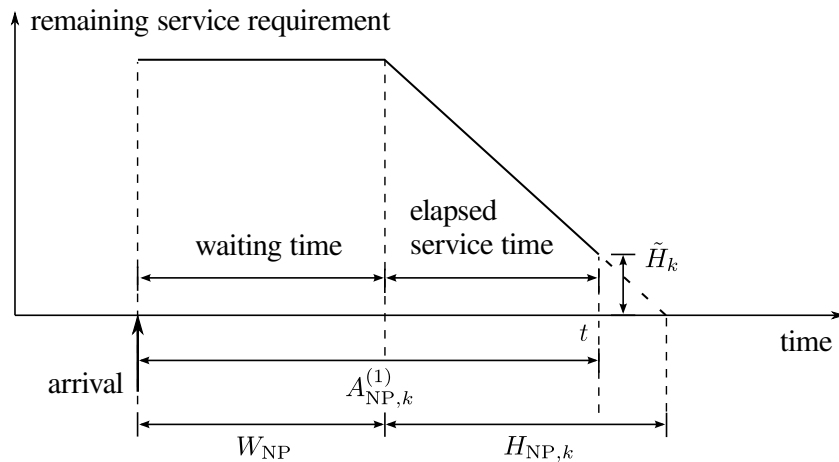


Figure 4: Attained waiting time of a class k customer who arrived in a normal service period

$$a_{\text{WV, NP}, k}^{***}(\omega_k, s_k, \alpha_k) = \frac{(w_{\text{WV}}^{**}(\omega_k, s_k) - w_{\text{WV}}^{**}(\omega_k, \alpha_k))h_{\text{WV}, k}^*(\alpha_k) - (q_{\text{WV}, k}^{**}(\omega_k, s_k) - q_{\text{WV}, k}^{**}(\omega_k, \alpha_k))}{(s_k - \alpha_k)\text{E}\left[Q_{\text{WV}, k}^{(1)} - W_{\text{WV}, k}^{(1)}\right]},$$

respectively.

Let $\bar{L}_{\text{WV}, k}$ (resp. $\bar{L}_{\text{NP}, k}$) ($k \in \mathcal{K}$) denote the number of class k customers in the system, who arrived during working vacation periods (resp. normal service periods). Also, let $\bar{U}_{\text{WV}, k}$ (resp. $\bar{U}_{\text{NP}, k}$) ($k \in \mathcal{K}$) denote the amount of work in system, which is brought by class k customers who arrived during working vacation periods (resp. normal service periods). We then define the joint transform $\psi(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \mathbf{s}_{\text{WV}}, \mathbf{s}_{\text{NP}})$ as

$$\psi(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \mathbf{s}_{\text{WV}}, \mathbf{s}_{\text{NP}}) = \text{E}\left[\prod_{k \in \mathcal{K}} \left(z_{\text{WV}, k}^{\bar{L}_{\text{WV}, k}} \cdot z_{\text{NP}, k}^{\bar{L}_{\text{NP}, k}} \cdot e^{-s_{\text{WV}, k}\bar{U}_{\text{WV}, k}} \cdot e^{-s_{\text{NP}, k}\bar{U}_{\text{NP}, k}}\right)\right],$$

where $\mathbf{z}_{\text{WV}} = (z_{\text{WV}, 1}, z_{\text{WV}, 2}, \dots, z_{\text{WV}, K})$, $\mathbf{z}_{\text{NP}} = (z_{\text{NP}, 1}, z_{\text{NP}, 2}, \dots, z_{\text{NP}, K})$, $\mathbf{s}_{\text{WV}} = (s_{\text{WV}, 1}, s_{\text{WV}, 2}, \dots, s_{\text{WV}, K})$, and $\mathbf{s}_{\text{NP}} = (s_{\text{NP}, 1}, s_{\text{NP}, 2}, \dots, s_{\text{NP}, K})$.

Theorem 5. $\psi(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \mathbf{s}_{\text{WV}}, \mathbf{s}_{\text{NP}})$ is given by

$$\begin{aligned} \psi(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \mathbf{s}_{\text{WV}}, \mathbf{s}_{\text{NP}}) &= (1 - \nu)P_{\text{WV}} + \sum_{k \in \mathcal{K}} z_{\text{WV}, k} \sigma_{\text{WV}, k}^{(\gamma)} a_{\text{WV}, \text{WV}, k}^{**} \left(\sum_{i \in \mathcal{K}} [\lambda_{\text{WV}, i} - \lambda_{\text{WV}, i} z_{\text{WV}, i} h_{\text{WV}, i}^*(s_{\text{WV}, i})], s_{\text{WV}, k} \right) \\ &+ \sum_{k \in \mathcal{K}} z_{\text{WV}, k} \sigma_{\text{WV}, k}^{(1)} a_{\text{WV}, \text{NP}, k}^{***} \left(\sum_{i \in \mathcal{K}} [\lambda_{\text{WV}, i} - \lambda_{\text{WV}, i} z_{\text{WV}, i} h_{\text{WV}, i}^*(s_{\text{WV}, i})], \right. \\ &\qquad \qquad \qquad \left. \sum_{i \in \mathcal{K}} [\lambda_{\text{NP}, i} - \lambda_{\text{NP}, i} z_{\text{NP}, i} h_{\text{NP}, i}^*(s_{\text{NP}, i})], s_{\text{WV}, k} \right) \\ &+ \sum_{k \in \mathcal{K}} z_{\text{NP}, k} \sigma_{\text{NP}, k}^{(1)} a_{\text{NP}, k}^{**} \left(\sum_{i \in \mathcal{K}} [\lambda_{\text{NP}, i} - \lambda_{\text{NP}, i} z_{\text{NP}, i} h_{\text{NP}, i}^*(s_{\text{NP}, i})], s_{\text{NP}, k} \right). \end{aligned}$$

Proof. Note first that the system is empty with probability $1 - \sigma = (1 - \nu)P_{\text{WV}}$ (see Remark 6). Furthermore, when a customer is being served, all waiting customers arrived during the attained waiting time, as noted at the beginning of this section. Theorem 5 immediately follows from those observations. \square

Remark 7. Let \bar{L}_{WV} (resp. \bar{L}_{NP}) denote the total number of customers in the system, who arrived during working vacation periods (resp. normal service periods). Also, let \bar{U}_{WV} (resp. \bar{U}_{NP}) denote the total amount of work in system, which was brought by customers who arrived during working vacation periods (resp. normal service periods). As stated in Remark 2, we can obtain those by considering the single-class system with λ_{WV} , $h_{\text{WV}}^*(s)$, λ_{NP} , and $h_{\text{NP}}^*(s)$. Therefore Theorem 5 also provides the formula for the joint transform of \bar{L}_{WV} , \bar{L}_{NP} , \bar{U}_{WV} , and \bar{U}_{NP} implicitly, because it corresponds the case of $K = 1$.

Taking the partial derivatives of $\psi(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \mathbf{s}_{\text{WV}}, \mathbf{s}_{\text{NP}})$, we can obtain the moments of $\bar{L}_{\text{WV}, k}$, $\bar{L}_{\text{NP}, k}$, $\bar{U}_{\text{WV}, k}$, and $\bar{U}_{\text{NP}, k}$ ($k \in \mathcal{K}$). In particular, we have

$$\begin{aligned} \text{E}[\bar{L}_{\text{WV}, k}] &= \lambda_{\text{WV}, k} P_{\text{WV}} \cdot \text{E}[Q_{\text{WV}, k}], & \text{E}[\bar{L}_{\text{NP}, k}] &= \lambda_{\text{NP}, k} P_{\text{NP}} \cdot \text{E}[Q_{\text{NP}, k}], \\ \text{E}[\bar{U}_{\text{WV}, k}] &= P_{\text{WV}} \rho_{\text{WV}, k} \left(\text{E}[U_{\text{WV}}] + \frac{\text{E}[H_{\text{WV}, k}^2]}{2\text{E}[H_{\text{WV}, k}]} + \frac{1}{\eta} \right) - \frac{\gamma}{\eta} \left(\sigma_{\text{WV}, k}^{(\gamma)} + \sigma_{\text{WV}, k}^{(1)} \right), \\ \text{E}[\bar{U}_{\text{NP}, k}] &= P_{\text{NP}} \rho_{\text{NP}, k} \left(\frac{\text{E}[H_{\text{NP}, k}^2]}{2\text{E}[H_{\text{NP}, k}]} + \text{E}[U_{\text{NP}}] \right). \end{aligned}$$

6. Busy Cycle

The busy cycle is defined as the interval between ends of successive busy periods. In order to analyze the busy cycle and related quantities, we first consider the first passage time to the empty system. More specifically, we define F_{WV} (resp. F_{NP}) as the first passage time to the empty system given that the server is on working vacation (resp. in a normal service period) at time 0. We divide F_{WV} into two parts: $F_{\text{WV}}^{(\gamma)}$ and $F_{\text{WV}}^{(1)}$, where $F_{\text{WV}}^{(\gamma)}$ (resp. $F_{\text{WV}}^{(1)}$) denotes the length of a subinterval during which the server is on working vacation (resp. in a normal service period). By definition, $F_{\text{WV}} = F_{\text{WV}}^{(\gamma)} + F_{\text{WV}}^{(1)}$. Furthermore, for each k ($k \in \mathcal{K}$), we define $N_{\text{WV},k}^{(\gamma)}$ (resp. $N_{\text{WV},k}^{(1)}$) as the number of class k customers arriving in $F_{\text{WV}}^{(\gamma)}$ (resp. $F_{\text{WV}}^{(1)}$). Similarly, we define $N_{\text{NP},k}$ ($k \in \mathcal{K}$) as the number of class k customers arriving in F_{NP} . Let $S(t)$ ($t \geq 0$) denote the state of the server at time t , i.e., and $S(t) = \text{WV}$ if the server is on working vacation at time t , and otherwise $S(t) = \text{NP}$. We define $U(t)$ ($t \geq 0$) as the total amount of unfinished work at time t . We are interested in the following joint transforms.

$$\begin{aligned} \zeta_{\text{WV}}^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, s_{\text{WV}}, s_{\text{NP}} \mid x) &= \mathbb{E} \left[\left(\prod_{k \in \mathcal{K}} z_{\text{WV},k}^{N_{\text{WV},k}^{(\gamma)}} \cdot z_{\text{NP},k}^{N_{\text{WV},k}^{(1)}} \right) \cdot e^{-s_{\text{WV}} F_{\text{WV}}^{(\gamma)}} \cdot e^{-s_{\text{NP}} F_{\text{WV}}^{(1)}} \mid U(0) = x, S(0) = \text{WV} \right], \\ \zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid x) &= \mathbb{E} \left[\left(\prod_{k \in \mathcal{K}} z_{\text{NP},k}^{N_{\text{NP},k}} \right) \cdot e^{-s_{\text{NP}} F_{\text{NP}}} \mid U(0) = x, S(0) = \text{NP} \right], \end{aligned}$$

where $\mathbf{z}_{\text{WV}} = (z_{\text{WV},1}, z_{\text{WV},2}, \dots, z_{\text{WV},K})$ and $\mathbf{z}_{\text{NP}} = (z_{\text{NP},1}, z_{\text{NP},2}, \dots, z_{\text{NP},K})$.

Lemma 6. $\zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid x)$ is given by

$$\zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid x) = e^{-\beta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}})x}, \quad (30)$$

where $\beta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}})$ is defined as

$$\beta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}}) = s_{\text{NP}} + \lambda_{\text{NP}} - \sum_{k \in \mathcal{K}} z_{\text{NP},k} \lambda_{\text{NP},k} \int_0^\infty \zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid y) dH_{\text{NP},k}(y), \quad (31)$$

and it is given by

$$\beta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}}) = s_{\text{NP}} + \lambda_{\text{NP}} - \sum_{k \in \mathcal{K}} z_{\text{NP},k} \lambda_{\text{NP},k} h_{\text{NP},k}^*(\beta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}})). \quad (32)$$

The proof of Lemma 6 is given in Appendix F.

Next, we consider the joint transform $\zeta_{\text{WV}}^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, s_{\text{WV}}, s_{\text{NP}} \mid x)$. Given $S(0) = \text{WV}$, let T_V denote the time instant when the server ends the current working vacation for the first time after time 0. Because of the memoryless property, T_V is exponentially distributed with parameter η . We classify the first passage time F_{WV} to the empty system into two cases, $F_{\text{WV}} \leq T_V$ and $F_{\text{WV}} > T_V$, and we define $\zeta_{\text{WV},\text{C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} \mid x)$ and $\zeta_{\text{WV},\text{E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha \mid x)$ as

$$\begin{aligned} \zeta_{\text{WV},\text{C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} \mid x) &= \mathbb{E} \left[\left(\prod_{k \in \mathcal{K}} z_{\text{WV},k}^{N_{\text{WV},k}^{(\gamma)}} \right) \cdot e^{-s_{\text{WV}} F_{\text{WV}}^{(\gamma)}} \mid U(0) = x, S(0) = \text{WV}, F_{\text{WV}} \leq T_V \right], \\ \zeta_{\text{WV},\text{E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha \mid x) & \end{aligned}$$

$$= \mathbb{E} \left[\left(\prod_{k \in \mathcal{K}} z_{\text{WV},k}^{N_{\text{WV},k}^{(\gamma)}} \right) \cdot e^{-s_{\text{WV}}T_V} \cdot e^{-\alpha U(T_V)} \mid U(0) = x, S(0) = \text{WV}, F_{\text{WV}} > T_V \right],$$

respectively. Note here that (30) implies

$$\begin{aligned} & \mathbb{E} \left[\left(\prod_{k \in \mathcal{K}} z_{\text{NP},k}^{N_{\text{NP},k}^{(1)}} \right) \cdot e^{-s_{\text{NP}}F_{\text{WV}}^{(1)}} \mid U(0) = x, S(0) = \text{WV}, F_{\text{WV}} > T_V, U(T_V) = y \right] \\ &= \mathbb{E} \left[\left(\prod_{k \in \mathcal{K}} z_{\text{NP},k}^{N_{\text{NP},k}} \right) \cdot e^{-s_{\text{NP}}F_{\text{NP}}} \mid U(0) = y, S(0) = \text{NP} \right] \\ &= \zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid y) = e^{-\beta_{\text{NP}}^*(z_{\text{NP}}, s_{\text{NP}})y} \end{aligned}$$

We then have

$$\begin{aligned} & \zeta_{\text{WV}}^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, s_{\text{WV}}, s_{\text{NP}} \mid x) \\ &= P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} \mid x) + P_{\text{E}|x} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \beta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}}) \mid x), \end{aligned} \quad (33)$$

where $P_{\text{C}|x}$ and $P_{\text{E}|x}$ are defined as

$$\begin{aligned} P_{\text{C}|x} &= \Pr(F_{\text{WV}} \leq T_V \mid U(0) = x, S(0) = \text{WV}), \\ P_{\text{E}|x} &= \Pr(F_{\text{WV}} > T_V \mid U(0) = x, S(0) = \text{WV}). \end{aligned}$$

Lemma 7. *The following equations hold.*

$$P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} \mid x) = e^{-\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}})x}, \quad (34)$$

$$P_{\text{E}|x} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha \mid x) = \frac{e^{-\alpha x} - e^{-\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}})x}}{\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}}) - \alpha} \cdot \beta_{\text{WV,E}}^*(z_{\text{WV}}, s_{\text{WV}}, \alpha), \quad (35)$$

where $\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}})$ and $\beta_{\text{WV,E}}^*(z_{\text{WV}}, s_{\text{WV}}, \alpha)$ are defined as

$$\begin{aligned} \beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}}) &= \frac{s_{\text{WV}}}{\gamma} + \frac{\eta}{\gamma} + \frac{\lambda_{\text{WV}}}{\gamma} \\ &\quad - \sum_{k \in \mathcal{K}} \frac{z_{\text{WV},k} \lambda_{\text{WV},k}}{\gamma} \int_0^\infty P_{\text{C}|y} \cdot \zeta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}} \mid y) dH_{\text{WV},k}(y), \end{aligned} \quad (36)$$

$$\begin{aligned} \beta_{\text{WV,E}}^*(z_{\text{WV}}, s_{\text{WV}}, \alpha) &= \eta/\gamma \\ &\quad + \sum_{k \in \mathcal{K}} \frac{z_{\text{WV},k} \lambda_{\text{WV},k}}{\gamma} \int_0^\infty P_{\text{E}|y} \cdot \zeta_{\text{WV,E}}^*(z_{\text{WV}}, s_{\text{WV}}, \alpha \mid y) dH_{\text{WV},k}(y), \end{aligned} \quad (37)$$

and they satisfy

$$\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}}) = \frac{s_{\text{WV}}}{\gamma} + \frac{\eta}{\gamma} + \frac{\lambda_{\text{WV}}}{\gamma} - \sum_{k \in \mathcal{K}} \frac{z_{\text{WV},k} \lambda_{\text{WV},k}}{\gamma} \cdot h_{\text{WV},k}^*(\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}})), \quad (38)$$

$$\begin{aligned} \beta_{\text{WV,E}}^*(z_{\text{WV}}, s_{\text{WV}}, \alpha) &= \frac{\eta}{\gamma} + \sum_{k \in \mathcal{K}} \frac{z_{\text{WV},k} \lambda_{\text{WV},k}}{\gamma} \cdot \frac{h_{\text{WV},k}^*(\alpha) - h_{\text{WV},k}^*(\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}}))}{\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}}) - \alpha} \\ &\quad \cdot \beta_{\text{WV,E}}^*(z_{\text{WV}}, s_{\text{WV}}, \alpha). \end{aligned} \quad (39)$$

The proof of Lemma 7 is given in Appendix G.

It follows from (33), (34), and (35) that $\zeta_{\text{WV}}^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, s_{\text{WV}}, s_{\text{NP}} \mid x)$ is given by

$$\begin{aligned} \zeta_{\text{WV}}^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, s_{\text{WV}}, s_{\text{NP}} \mid x) &= e^{-\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}})x} \\ &+ \frac{e^{-\beta_{\text{NP}}^*(z_{\text{NP}}, s_{\text{NP}})x} - e^{-\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}})x}}{\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}}) - \beta_{\text{NP}}^*(z_{\text{NP}}, s_{\text{NP}})} \cdot \beta_{\text{WV,E}}^*(z_{\text{WV}}, s_{\text{WV}}, \beta_{\text{NP}}^*(z_{\text{NP}}, s_{\text{NP}})). \end{aligned} \quad (40)$$

With (36) and (37), we define $\beta_{\text{WV,C}}$ and $\beta_{\text{WV,E}}$ as

$$\beta_{\text{WV,C}} = \beta_{\text{WV,C}}^*(\mathbf{1}, 0), \quad \beta_{\text{WV,E}} = \beta_{\text{WV,E}}^*(\mathbf{1}, 0, 0),$$

where $\mathbf{1}$ denotes a vector whose elements are all equal to one.

Lemma 8. $\beta_{\text{WV,C}}$ and $\beta_{\text{WV,E}}$ are given by

$$\beta_{\text{WV,C}} = \beta_{\text{WV,E}} = \frac{\eta/\gamma}{1 - \nu}, \quad (41)$$

and $P_{\text{C}|x}$ and $P_{\text{E}|x}$ are given by

$$P_{\text{C}|x} = e^{-\beta_{\text{WV,C}} \cdot x}, \quad P_{\text{E}|x} = 1 - e^{-\beta_{\text{WV,C}} \cdot x}. \quad (42)$$

The proof of Lemma 8 is given in Appendix H.

We now consider the busy cycle. Recall that the server is always on working vacation at the beginning of busy cycle. Let Θ denote the length of a randomly chosen busy cycle. We divide Θ into two parts, and let $\Theta^{(\gamma)}$ (resp. $\Theta^{(1)}$) denote the length of the subinterval during which the server is on working vacation (resp. in a normal service period). Furthermore, we divide $\Theta^{(\gamma)}$ into two parts, and let $\Theta_{\text{E}}^{(\gamma)}$ (resp. $\Theta_{\text{B}}^{(\gamma)}$) denote the length of the subinterval during which the server is idle (resp. busy). By definition, $\Theta = \Theta_{\text{E}}^{(\gamma)} + \Theta_{\text{B}}^{(\gamma)} + \Theta^{(1)}$. For each k ($k \in \mathcal{K}$), let $\bar{N}_k^{(\gamma)}$ (resp. $\bar{N}_k^{(1)}$) denote the number of class k customers arriving during $\Theta^{(\gamma)}$ (resp. $\Theta^{(1)}$). We then define the joint transform of those quantities as follows.

$$\theta^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \omega, s_{\text{WV}}, s_{\text{NP}}) = \mathbb{E} \left[\left(\prod_{k \in \mathcal{K}} z_{\text{WV},k}^{\bar{N}_k^{(\gamma)}} \cdot z_{\text{NP},k}^{\bar{N}_k^{(1)}} \right) \cdot e^{-\omega \Theta_{\text{E}}^{(\gamma)}} \cdot e^{-s_{\text{WV}} \Theta_{\text{B}}^{(\gamma)}} \cdot e^{-s_{\text{NP}} \Theta^{(1)}} \right].$$

By definition, $\theta^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \omega, s_{\text{WV}}, s_{\text{NP}})$ satisfies

$$\begin{aligned} &\theta^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \omega, s_{\text{WV}}, s_{\text{NP}}) \\ &= \frac{\lambda_{\text{WV}}}{\omega + \lambda_{\text{WV}}} \sum_{k \in \mathcal{K}} \frac{z_{\text{WV},k} \lambda_{\text{WV},k}}{\lambda_{\text{WV}}} \int_0^\infty \zeta_{\text{WV}}^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, s_{\text{WV}}, s_{\text{NP}} \mid y) dH_{\text{WV},k}(y). \end{aligned}$$

Therefore, with (40), we obtain the following theorem.

Theorem 6. $\theta^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \omega, s_{\text{WV}}, s_{\text{NP}})$ is given by

$$\begin{aligned} \theta^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \omega, s_{\text{WV}}, s_{\text{NP}}) &= \frac{\lambda_{\text{WV}}}{\omega + \lambda_{\text{WV}}} \sum_{k \in \mathcal{K}} \frac{z_{\text{WV},k} \lambda_{\text{WV},k}}{\lambda_{\text{WV}}} \left[h_{\text{WV},k}^*(\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}})) \right. \\ &+ \left. \frac{h_{\text{WV},k}^*(\beta_{\text{NP}}^*(z_{\text{NP}}, s_{\text{NP}})) - h_{\text{WV},k}^*(\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}}))}{\beta_{\text{WV,C}}^*(z_{\text{WV}}, s_{\text{WV}}) - \beta_{\text{NP}}^*(z_{\text{NP}}, s_{\text{NP}})} \cdot \beta_{\text{WV,E}}^*(z_{\text{WV}}, s_{\text{WV}}, \beta_{\text{NP}}^*(z_{\text{NP}}, s_{\text{NP}})) \right]. \end{aligned}$$

Remark 8. It is clear from the derivation of Theorem 6 that

$$\begin{aligned} & \Pr(\text{A randomly chosen busy period ends while the server is on working vacation}) \\ &= \lim_{\omega \rightarrow 0} \frac{\lambda_{\text{WV}}}{\omega + \lambda_{\text{WV}}} \sum_{k \in \mathcal{K}} \frac{\lambda_{\text{WV},k}}{\lambda_{\text{WV}}} \cdot h_{\text{WV},k}^*(\beta_{\text{WV},\text{C}}^*(\mathbf{0}, 0)) = h_{\text{WV}}^*(\beta_{\text{WV},\text{C}}) = r, \end{aligned}$$

where we use (53). This result is consistent with Remark 3. Furthermore, using (38) and (39), we obtain an alternative expression for $\theta^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \omega, s_{\text{WV}}, s_{\text{NP}})$.

$$\begin{aligned} \theta^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \omega, s_{\text{WV}}, s_{\text{NP}}) &= \frac{\lambda_{\text{WV}}}{\omega + \lambda_{\text{WV}}} \left[\frac{1}{\lambda_{\text{WV}}} \left\{ s_{\text{WV}} + \lambda_{\text{WV}} - \gamma \beta_{\text{WV},\text{C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}) \right. \right. \\ &\quad \left. \left. + \gamma \beta_{\text{WV},\text{E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \beta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}})) \right\} \right]. \end{aligned}$$

Taking the partial derivatives of $\theta^*(\mathbf{z}_{\text{WV}}, \mathbf{z}_{\text{NP}}, \omega, s_{\text{WV}}, s_{\text{NP}})$, we can obtain the moments of $\bar{N}_k^{(\gamma)}$, $\bar{N}_k^{(1)}$, $\Theta_{\text{B}}^{(\gamma)}$, and $\Theta^{(1)}$. In particular,

$$\begin{aligned} \mathbb{E}[\bar{N}_k^{(\gamma)}] &= \lambda_{\text{WV},k} \cdot \left(\frac{1}{\lambda_{\text{WV}}} + \mathbb{E}[\Theta_{\text{B}}^{(\gamma)}] \right), & \mathbb{E}[\bar{N}_k^{(1)}] &= \lambda_{\text{NP},k} \cdot \mathbb{E}[\Theta^{(1)}], \\ \mathbb{E}[\Theta_{\text{B}}^{(\gamma)}] &= \frac{\gamma \beta_{\text{WV},\text{C}}}{\eta \lambda_{\text{WV}}} - \frac{1}{\lambda_{\text{WV}}}, & \mathbb{E}[\Theta^{(1)}] &= (1 - r) \cdot \frac{\mathbb{E}[U_{\text{WV}}]/\nu}{1 - \rho_{\text{NP}}}. \end{aligned}$$

7. Concluding Remarks

We considered the stationary multi-class FIFO M/G/1 queue with exponential working vacations. We derived the LST of the stationary work in system, and the LSTs of the stationary waiting time and sojourn time in each class. We also obtained the joint transform for the queue lengths and the amounts of work in system in respective classes and the joint transform associated with the busy cycle. Before closing this paper, we provide some remarks.

As stated in section 1, if we delete time intervals in normal service periods from the time axis, the resulting process can be viewed as a multi-class FIFO M/G/1 queue with Poisson disasters, where the processing rate is equal to γ . Because queues with disasters are of independent interest, Appendix I summarizes the analytical results for the multi-class FIFO M/G/1 queue with Poisson disasters, all of which are immediately obtained from the results in this paper.

In queueing models with working vacations, the processing rate is always equal to γ when the system becomes empty. In other words, the queue length process directly affects the processing rate. From this point of view, the queueing model with working vacations differs from the queueing model embedded in a random environment (i.e., the processing rate is assumed to change according to an underlying environmental process). More specifically, we can see the difference between these two models by considering a special case of our model, where the processing rate is proportional to the arrival rate and comparing it to the corresponding queue embedded in a random environment of a two-state Markov chain. In the latter, the stationary number of customers in the system is independent of the underlying Markov chain and its conditional distribution given a specific state of the Markov chain is the same as that of the ordinary M/G/1 queue (Section 6 in [11]). On the other hand, it is verified that the model we considered does not have such a property. Thus, the queueing model with working vacation is essentially different from the queueing model embedded in a random environmental process.

Appendix

A. Proof of Lemma 1

We define U_{NP}^{B} as the total amount of work in system at the beginning of a normal service period. Note that U_{NP}^{B} is a conditional random variable of U_{WV}^{E} given that the server is busy at the end of a working vacation. Let $u_{\text{NP},\text{B}}^*(s)$ denote the LST of U_{NP}^{B} . We then have

$$u_{\text{NP},\text{B}}^*(s) = \mathbb{E} \left[e^{-sU_{\text{WV}}^{\text{E}}} \mid U_{\text{WV}}^{\text{E}} > 0 \right] = \frac{u_{\text{WV},\text{E}}^*(s) - \Pr(U_{\text{WV}}^{\text{E}} = 0)}{1 - \Pr(U_{\text{WV}}^{\text{E}} = 0)}, \quad (43)$$

$$\mathbb{E} [U_{\text{NP}}^{\text{B}}] = \frac{\mathbb{E} [U_{\text{WV}}^{\text{E}}]}{1 - \Pr(U_{\text{WV}}^{\text{E}} = 0)}. \quad (44)$$

Consider a censored workload process by removing all working vacation periods from the time axis. In steady state, the censored process has the same distribution as U_{NP} . Also, the censored process can be viewed as the conditional workload process of the M/G/1 vacation queue with exhaustive services, given that the server is busy. Therefore, it follows from (5.6) in [2] that $u_{\text{NP}}^*(s)$ is given by

$$u_{\text{NP}}^*(s) = \frac{1 - u_{\text{NP},\text{B}}^*(s)}{s\mathbb{E} [U_{\text{NP}}^{\text{B}}]} \cdot u_{\text{M/G/1}}^*(s).$$

Note here that (2), (43), and (44) imply

$$\frac{1 - u_{\text{NP},\text{B}}^*(s)}{s\mathbb{E} [U_{\text{NP}}^{\text{B}}]} = \frac{1 - u_{\text{WV},\text{E}}^*(s)}{s\mathbb{E} [U_{\text{WV}}^{\text{E}}]} = \frac{1 - u_{\text{WV}}^*(s)}{s\mathbb{E} [U_{\text{WV}}]},$$

which completes the proof.

B. Proof of Lemma 2

We regard an interval between successive ends of working vacations as a cycle. Let C_{WV} (resp. C_{NP}) denote the length of an interval during which the server is on working vacation (resp. in a normal service period) in a randomly chosen cycle. Owing to the renewal reward theorem, we have

$$P_{\text{WV}} = \frac{\mathbb{E}[C_{\text{WV}}]}{\mathbb{E}[C_{\text{WV}}] + \mathbb{E}[C_{\text{NP}}]}, \quad P_{\text{NP}} = \frac{\mathbb{E}[C_{\text{NP}}]}{\mathbb{E}[C_{\text{WV}}] + \mathbb{E}[C_{\text{NP}}]}. \quad (45)$$

Because C_{WV} is equivalent to the working vacation length V , we have $\mathbb{E}[C_{\text{WV}}] = \mathbb{E}[V]$. On the other hand, $\mathbb{E}[C_{\text{NP}}]$ equals to the mean first passage time to the empty system in the corresponding ordinary M/G/1 queue with initial workload of U_{NP}^{B} . Noting that $C_{\text{NP}} = 0$ if the system is empty at the end of the working vacation, we have

$$\mathbb{E}[C_{\text{NP}}] = \Pr(U_{\text{WV}}^{\text{E}} = 0) \cdot 0 + \{1 - \Pr(U_{\text{WV}}^{\text{E}} = 0)\} \cdot \frac{\mathbb{E}[U_{\text{NP}}^{\text{B}}]}{1 - \rho_{\text{NP}}} = \frac{\mathbb{E}[U_{\text{WV}}]}{1 - \rho_{\text{NP}}}, \quad (46)$$

where we use (2) and (44). (5) now follows from (45), (46), and $\mathbb{E}[C_{\text{WV}}] = \mathbb{E}[V] = 1/\eta$.

C. Proof of Lemma 3

The censored process obtained by removing all normal service periods is considered as an M/G/1 queue with Poisson disasters with rate η , where the system becomes empty when

disasters occur. The M/G/1 queue with Poisson disasters has already been studied in [3, 15], where the processing rate is assumed to be one. In order to apply the results in [3, 15] to our system, we consider the new process created by extending the time axis of the workload process in working vacation periods γ times so that the processing rate becomes one. Note that the time-average quantities of the new censored process are identical to those of the original process. In the new process, the arrival rate of customers is equal to $\lambda_{\text{WV}}/\gamma$ and lengths of working vacations are exponentially distributed with parameter η/γ . $u_{\text{WV}}^*(s)$ in (7) then immediately follows from Proposition 1 in [3]. We also obtain (8) by substituting 0 to the repair time in (2.1a) in [15]. The existence of the unique real root of (9) is shown in Remark 2.2 in [15]. See Remark 3 for the positivity of r . Furthermore, taking the derivative of $u_{\text{WV}}^*(s)$ in (7) and evaluating at $s = 0$ yields $E[U_{\text{WV}}]$ in (7).

D. Proof of Lemma 5

We first consider $\sigma_{\text{WV},k}^{(\gamma)}$. Note that all customers being served in working vacation periods arrived during working vacation periods. Thus, from Little’s law, we have $\sigma_{\text{WV},k}^{(\gamma)} = \lambda_{\text{WV},k} P_{\text{WV}} \cdot E[Q_{\text{WV},k}^{(\gamma)} - W_{\text{WV}}^{(\gamma)}]$. Furthermore, with Lemma 3 and (21), $E[Q_{\text{WV},k}^{(\gamma)} - W_{\text{WV}}^{(\gamma)}]$ is obtained to be

$$\begin{aligned} E \left[Q_{\text{WV},k}^{(\gamma)} - W_{\text{WV}}^{(\gamma)} \right] &= E \left[T_{\text{WV}}^{(\gamma)}(U_{\text{WV}} + H_{\text{WV},k}) \right] - E \left[T_{\text{WV}}^{(\gamma)}(U_{\text{WV}}) \right] \\ &= \frac{u_{\text{WV}}^*(\eta/\gamma)(1 - h_{\text{WV},k}^*(\eta/\gamma))}{\eta} = \frac{\nu}{\lambda_{\text{WV}}} \cdot \frac{1 - h_{\text{WV},k}^*(\eta/\gamma)}{1 - h_{\text{WV}}^*(\eta/\gamma)}, \end{aligned} \tag{47}$$

from which (24) follows.

Similarly, $\sigma_{\text{WV},k}^{(1)}$ follows from $\sigma_{\text{WV},k}^{(1)} = \lambda_{\text{WV},k} P_{\text{WV}} \cdot E[Q_{\text{WV},k}^{(1)} - W_{\text{WV}}^{(1)}]$ and

$$\begin{aligned} E \left[Q_{\text{WV},k}^{(1)} - W_{\text{WV}}^{(1)} \right] &= E \left[T_{\text{WV}}^{(1)}(U_{\text{WV}} + H_{\text{WV},k}) \right] - E \left[T_{\text{WV}}^{(1)}(U_{\text{WV}}) \right] \\ &= E[H_{\text{WV},k}] - \frac{\gamma\nu}{\lambda_{\text{WV}}} \cdot \frac{1 - h_{\text{WV},k}^*(\eta/\gamma)}{1 - h_{\text{WV}}^*(\eta/\gamma)}. \end{aligned} \tag{48}$$

Finally, we consider $\sigma_{\text{NP},k}^{(1)}$. Note that all customers arriving in normal service periods are served in normal service periods. Therefore $\sigma_{\text{NP},k}^{(1)} = \lambda_{\text{NP},k} P_{\text{NP}} \cdot E[H_{\text{NP},k}] = P_{\text{NP}} \cdot \rho_{\text{NP},k}$, from which (26) follows.

E. Proof of Theorem 4

We first consider (27). Suppose a class k ($k \in \mathcal{K}$) customer is being served at processing rate γ (i.e., in a working vacation period). Note here that

$$E[Q_{\text{WV},k}^{(\gamma)} - W_{\text{WV},k}^{(\gamma)} \mid Q_{\text{WV},k}^{(\gamma)} - W_{\text{WV},k}^{(\gamma)} > 0] = \frac{E[Q_{\text{WV},k}^{(\gamma)} - W_{\text{WV},k}^{(\gamma)}]}{u_{\text{WV}}^*(\eta/\gamma)}.$$

We thus have

$$\begin{aligned} &a_{\text{WV},\text{WV},k}^{**}(\omega_k, \alpha_k) \\ &= \frac{1}{\frac{E[Q_{\text{WV},k}^{(\gamma)} - W_{\text{WV},k}^{(\gamma)}]}{u_{\text{WV}}^*(\eta/\gamma)}} \cdot \frac{u_{\text{WV}}^* \left(\frac{\omega_k + \eta}{\gamma} \right)}{u_{\text{WV}}^*(\eta/\gamma)} \end{aligned}$$

$$\cdot \int_0^\infty dH_{\text{WV},k}(x) \left[e^{-\eta(x/\gamma)} \int_0^{x/\gamma} e^{-\omega_k t} e^{-\alpha_k(x-\gamma t)} dt + \int_0^{x/\gamma} \eta e^{-\eta\tau} d\tau \int_0^\tau e^{-\omega_k t} e^{-\alpha_k(x-\gamma t)} dt \right],$$

from which (27) follows.

Next we consider (28). Suppose a class k ($k \in \mathcal{K}$) customer, who arrived in a working vacation period, is being served at processing rate one (i.e., in a normal service period). We then have

$$\mathbb{E}[Q_{\text{WV},k}^{(1)} - W_{\text{WV},k}^{(1)} \mid Q_{\text{WV},k}^{(1)} - W_{\text{WV},k}^{(1)} > 0] = \frac{\mathbb{E}[Q_{\text{WV},k}^{(1)} - W_{\text{WV},k}^{(1)}]}{1 - u_{\text{WV}}^*(\eta/\gamma)h_{\text{WV},k}^*(\eta/\gamma)}.$$

Therefore

$$\begin{aligned} & a_{\text{WV},\text{NP},k}^{***}(\omega_k, s_k, \alpha_k) \\ &= \frac{1}{\frac{\mathbb{E}[Q_{\text{WV},k}^{(1)} - W_{\text{WV},k}^{(1)}]}{1 - u_{\text{WV}}^*(\eta/\gamma)h_{\text{WV},k}^*(\eta/\gamma)}} \\ & \cdot \left[\frac{u_{\text{WV}}^*(\eta/\gamma)(1 - h_{\text{WV},k}^*(\eta/\gamma))}{1 - u_{\text{WV}}^*(\eta/\gamma)h_{\text{WV},k}^*(\eta/\gamma)} \cdot \frac{u_{\text{WV}}^*\left(\frac{\omega_k + \eta}{\gamma}\right)}{u_{\text{WV}}^*(\eta/\gamma)} \right. \\ & \cdot \frac{1}{1 - h_{\text{WV},k}^*(\eta/\gamma)} \int_0^\infty dH_{\text{WV},k}(x) \int_0^{x/\gamma} \eta e^{-\eta\tau} d\tau \int_\tau^{\tau+x-\gamma\tau} e^{-\omega_k\tau} e^{-s_k(t-\tau)} e^{-\alpha_k(x-\gamma\tau-(t-\tau))} dt \\ & \left. + \frac{1 - u_{\text{WV}}^*(\eta/\gamma)}{1 - u_{\text{WV}}^*(\eta/\gamma)h_{\text{WV},k}^*(\eta/\gamma)} \cdot \frac{1}{1 - u_{\text{WV}}^*(\eta/\gamma)} \cdot \frac{u_{\text{WV}}^*(s_k) - u_{\text{WV}}^*\left(\frac{\omega_k + \eta}{\gamma}\right)}{(\gamma/\eta)\{(\omega_k + \eta)/\gamma - s_k\}} \right. \\ & \left. \cdot \int_0^\infty dH_{\text{WV},k}(x) \int_0^x e^{-s_k t} e^{-\alpha_k(x-t)} dt \right], \end{aligned}$$

from which (28) follows.

Finally, $a_{\text{NP},k}^{**}(s_k, \alpha_k)$ is given by

$$a_{\text{NP},k}^{**}(s_k, \alpha_k) = \frac{1}{\mathbb{E}[H_{\text{NP},k}]} \cdot u_{\text{NP},k}^*(s) \int_0^\infty dH_{\text{NP},k}(x) \int_0^x e^{-s_k t} e^{-\alpha_k(x-t)} dt,$$

from which (29) follows.

F. Proof of Lemma 6

For $x \geq 0$, $y \geq 0$, we have $\zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid x + y) = \zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid x) \cdot \zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid y)$. Therefore

$$\begin{aligned} \zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid x + \Delta x) &= \zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid x) \cdot \zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid \Delta x) \\ &= \zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid x) \left[1 - s_{\text{NP}}\Delta x - \lambda_{\text{NP}}\Delta x + \lambda_{\text{NP}}\Delta x \sum_{k \in \mathcal{K}} z_{\text{NP},k} \right. \\ & \left. \cdot \frac{\lambda_{\text{NP},k}}{\lambda_{\text{NP}}} \int_0^\infty \zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} \mid y) dH_{\text{NP},k}(y) + o(\Delta x) \right], \end{aligned}$$

from which it follows that

$$\frac{\partial}{\partial x} \left[\zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} | x) \right] = -\zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} | x) \beta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}}).$$

Noting $\zeta_{\text{NP}}^*(\mathbf{z}_{\text{NP}}, s_{\text{NP}} | 0) = 1$, we obtain (30). Also, substituting (30) into (31), we obtain (32).

G. Proof of Lemma 7

It is easy to see that for $x \geq 0, y \geq 0$,

$$\begin{aligned} P_{\text{C}|x+y} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x+y) &= P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x) \cdot P_{\text{C}|y} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | y), \\ P_{\text{E}|x+y} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | x+y) &= P_{\text{E}|x} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | x) \cdot e^{-\alpha y} \\ &\quad + P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x) \cdot P_{\text{E}|y} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | y). \end{aligned}$$

Therefore we have

$$\begin{aligned} &P_{\text{C}|x+\Delta x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x+\Delta x) \\ &= P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x) \cdot P_{\text{C}|\Delta x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | \Delta x) \\ &= P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x) \left[1 - s_{\text{WV}} \frac{\Delta x}{\gamma} - \eta \frac{\Delta x}{\gamma} - \lambda_{\text{WV}} \frac{\Delta x}{\gamma} \right. \\ &\quad \left. + \lambda_{\text{WV}} \frac{\Delta x}{\gamma} \sum_{k \in \mathcal{K}} \frac{z_{\text{WV},k} \lambda_{\text{WV},k}}{\lambda_{\text{WV}}} \int_0^\infty P_{\text{C}|y} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | y) dH_{\text{WV},k}(y) + o(\Delta x) \right], \end{aligned}$$

from which it follows that

$$\frac{\partial}{\partial x} \left[P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x) \right] = -P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x) \cdot \beta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}). \quad (49)$$

(34) now follows from (49) with $P_{\text{C}|0} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | 0) = 1$.

Similarly,

$$\begin{aligned} &P_{\text{E}|x+\Delta x} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | x+\Delta x) \\ &= P_{\text{E}|x} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | x) \cdot e^{-\alpha \Delta x} \\ &\quad + P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x) \cdot P_{\text{E}|\Delta x} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | \Delta x) \\ &= P_{\text{E}|x} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | x) \cdot (1 - \alpha \Delta x) + o(\Delta x) \\ &\quad + P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x) \left[\eta \frac{\Delta x}{\gamma} \right. \\ &\quad \left. + \lambda_{\text{WV}} \frac{\Delta x}{\gamma} \sum_{k \in \mathcal{K}} \frac{z_{\text{WV},k} \lambda_{\text{WV},k}}{\lambda_{\text{WV}}} \int_0^\infty P_{\text{E}|y} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | y) dH_{\text{WV},k}(y) + o(\Delta x) \right], \end{aligned}$$

and therefore

$$\begin{aligned} &\frac{\partial}{\partial x} \left[P_{\text{E}|x} \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | x) \right] \\ &= -\alpha P_{\text{E}|x} \cdot \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | x) + P_{\text{C}|x} \cdot \zeta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}} | x) \beta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha). \quad (50) \end{aligned}$$

Multiplying both sides of (50) by $e^{\alpha x}$ and using (34) yield

$$\frac{\partial}{\partial x} \left[P_{\text{E}|x} \zeta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha | x) \cdot e^{\alpha x} \right] = e^{-\beta_{\text{WV,C}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}})x} \cdot \beta_{\text{WV,E}}^*(\mathbf{z}_{\text{WV}}, s_{\text{WV}}, \alpha) \cdot e^{\alpha x}.$$

Because $P_{E|0} = 0$, we obtain

$$P_{E|x} \cdot \zeta_{\text{WV},E}^*(z_{\text{WV}}, s_{\text{WV}}, \alpha | x) \cdot e^{\alpha x} = \int_0^x e^{-\{\beta_{\text{WV},C}^*(z_{\text{WV}}, s_{\text{WV}}) - \alpha\}y} \cdot \beta_{\text{WV},E}^*(z_{\text{WV}}, s_{\text{WV}}, \alpha) dy,$$

from which (35) follows. Substituting (34) into (36) yields (38), and substituting (35) into (37) yields (39).

H. Proof of Lemma 8

(42) follows from $\zeta_{\text{WV},C}^*(\mathbf{1}, 0 | x) = 1$, $P_{C|x} + P_{E|x} = 1$, and (34). We thus consider (41) below. Note that $\zeta_{\text{WV},E}^*(\mathbf{1}, 0, 0 | x) = 1$. Therefore, taking the limits $\alpha \rightarrow 0$ and $s \rightarrow 0$ in (35), we obtain

$$P_{E|x} = \frac{\beta_{\text{WV},E}}{\beta_{\text{WV},C}} \cdot (1 - e^{-\beta_{\text{WV},C} \cdot x}),$$

from which and (42), we have $\beta_{\text{WV},C} = \beta_{\text{WV},E}$.

It is readily seen from (38) that $\beta_{\text{WV},C}$ satisfies

$$\beta_{\text{WV},C} = \eta/\gamma + \lambda_{\text{WV}}/\gamma - (\lambda_{\text{WV}}/\gamma)h_{\text{WV}}^*(\beta_{\text{WV},C}), \quad (51)$$

and $h_{\text{WV}}^*(\beta_{\text{WV},C}) = h_{\text{WV}}^*(\eta/\gamma + \lambda_{\text{WV}}/\gamma - (\lambda_{\text{WV}}/\gamma)h_{\text{WV}}^*(\beta_{\text{WV},C}))$. Furthermore, we have from (36)

$$\begin{aligned} \beta_{\text{WV},C} &= \eta/\gamma + \lambda_{\text{WV}}/\gamma - \sum_{k \in K} (\lambda_{\text{WV},k}/\gamma) \int_0^\infty P_{C|y} dH_{\text{WV},k}(y) \\ &= \eta/\gamma + \lambda_{\text{WV}}/\gamma - (\lambda_{\text{WV}}/\gamma) \int_0^\infty P_{C|y} dH_{\text{WV}}(y) \geq \eta/\gamma > 0, \end{aligned} \quad (52)$$

so that $|h_{\text{WV}}^*(\beta_{\text{WV},C})| < 1$. As a result, $h_{\text{WV}}^*(\beta_{\text{WV},C})$ is identical to the minimum nonnegative root r of (9). Finally, from (8) and (51), we obtain

$$\eta/\gamma + \lambda_{\text{WV}}/\gamma - (\lambda_{\text{WV}}/\gamma)r = \frac{\eta/\gamma}{1 - \nu}, \quad (53)$$

which completes the proof.

I. The Multi-Class FIFO M/G/1 Queue with Poisson Disasters

In this Appendix, we summarize the results of the stationary multi-class FIFO M/G/1 queue with Poisson disasters, where the processing rate is equal to one. We can readily obtain those results by considering the conditional counterparts in the multi-class FIFO M/G/1 with exponential working vacations and $\gamma = 1$, given that the server is on working vacation.

I.1. Model

Consider a stationary multi-class FIFO M/G/1 queue with Poisson disasters. Class k ($k \in \mathcal{K}$) customers arrive according to a Poisson process with rate λ_k . Let $h_k(x)$ and $h_k^*(s)$ ($k \in \mathcal{K}$) denote the distribution function of service times H_k of class k customers and its LST, respectively. Disasters occurs according to a Poisson process with rate η ($\eta > 0$), and the system becomes empty when disasters occur. We define λ and $h(x)$ as

$$\lambda = \sum_{k \in \mathcal{K}} \lambda_k, \quad h^*(s) = \sum_{k \in \mathcal{K}} \frac{\lambda_k}{\lambda} \cdot h_k^*(s).$$

Note that if we ignore customer classes, the system can be regarded as a single-class FIFO M/G/1 queue with Poisson disasters. Note also that the system is stable regardless of values of system parameters.

I.2. Results

The LST $u^*(s)$ of the amount of work in system is given by [3, 15] (cf. Lemma 3 and its proof)

$$u^*(s) = \frac{(1 - \nu)s - \eta}{s - \lambda + \lambda h^*(s) - \eta}.$$

Note that ν denotes the stationary probability of the server being busy.

$$\nu = \frac{(1 - r)\lambda}{(1 - r)\lambda + \eta}, \tag{54}$$

where r denotes the minimum nonnegative root of the following equation.

$$z = h^*(\eta + \lambda - \lambda z), \quad |z| < 1. \tag{55}$$

We denote the amount of work in system seen by a randomly chosen customer on arrival by U_A , and the length of the interval from the arrival of this customer to the occurrence of the next disaster by \tilde{D}_A . Owing to the memoryless property, \tilde{D}_A is exponentially distributed with parameter η . We define W_k and Q_k ($k \in \mathcal{K}$) as the waiting time and sojourn time, respectively, of class k customers, i.e., $W_k = \min(U_A, \tilde{D}_A)$ and $Q_k = \min(U_A + H_k, \tilde{D}_A)$. Note that owing to PASTA, W_k ($k \in \mathcal{K}$) is identical to the waiting time W of a randomly chosen customer. Furthermore, we define

$$\begin{aligned} P_N^W &= \Pr(U_A \leq \tilde{D}_A), & P_D^W &= \Pr(U_A > \tilde{D}_A), \\ w_N^*(s) &= \mathbb{E}[e^{-sW} \mid U_A \leq \tilde{D}_A], & w_D^*(s) &= \mathbb{E}[e^{-sW} \mid U_A > \tilde{D}_A], \end{aligned}$$

and for each k ($k \in \mathcal{K}$)

$$\begin{aligned} P_{N,k}^Q &= \Pr(U_A + H_k \leq \tilde{D}_A), & P_{D,k}^Q &= \Pr(U_A + H_k > \tilde{D}_A), \\ q_{N,k}^*(s) &= \mathbb{E}[e^{-sQ_k} \mid U_A + H_k \leq \tilde{D}_A], & q_{D,k}^*(s) &= \mathbb{E}[e^{-sQ_k} \mid U_A + H_k > \tilde{D}_A]. \end{aligned}$$

By definition, we have

$$w^*(s) = \mathbb{E}[e^{-sW}] = P_N^W w_N^*(s) + P_D^W w_D^*(s), \quad q_k^*(s) = \mathbb{E}[e^{-sQ_k}] = P_{N,k}^Q q_{N,k}^*(s) + P_{D,k}^Q q_{D,k}^*(s).$$

Because W corresponds to $W_{\text{WV}}^{(\gamma)}$ in the queue with working vacations, we obtain from Theorem 2

$$w^*(s) = u^*(s + \eta) + \frac{1 - u^*(s + \eta)}{(1/\eta)(s + \eta)}. \tag{56}$$

Note here that $u^*(s + \eta) = P_N^W w_N^*(s)$. Therefore the second term on the right hand side of (56) represents $P_D^W w_D^*(s)$. It then follows that

$$\begin{aligned} w_N^*(s) &= \frac{u^*(s + \eta)}{u^*(\eta)}, & w_D^*(s) &= \frac{1}{1 - u^*(\eta)} \cdot \frac{1 - u^*(s + \eta)}{(1/\eta)(s + \eta)}, \\ P_N^W &= u^*(\eta), & P_D^W &= 1 - u^*(\eta). \end{aligned}$$

Similarly, it follows from Theorem 3 that

$$q_k^*(s) = u^*(s + \eta)h_k^*(s + \eta) + \frac{1 - u^*(s + \eta)h_k^*(s + \eta)}{(1/\eta)(s + \eta)}, \quad k \in \mathcal{K},$$

and therefore for each k ($k \in \mathcal{K}$)

$$q_{N,k}^*(s) = \frac{u^*(s+\eta)h_k^*(s+\eta)}{u^*(\eta)h_k^*(\eta)}, \quad q_{D,k}^*(s) = \frac{1}{1-u^*(\eta)h_k^*(\eta)} \cdot \frac{1-u^*(s+\eta)h_k^*(s+\eta)}{(1/\eta)(s+\eta)},$$

$$P_{N,k}^Q = u^*(\eta)h_k^*(\eta), \quad P_{D,k}^Q = 1-u^*(\eta)h_k^*(\eta).$$

Let σ_k ($k \in \mathcal{K}$) denote the probability of a class k customer being served, which corresponds to $\sigma_{WV,k}^{(\gamma)}/P_{WV}$ in the queue with working vacations. It follows from (24) that

$$\sigma_k = \nu \cdot \frac{\lambda_k(1-h_k^*(\eta))}{\lambda(1-h^*(\eta))},$$

where ν is given in (54).

Let A_k ($k \in \mathcal{K}$) denote the conditional attained waiting time given that a class k customer is being served and let \tilde{H}_k ($k \in \mathcal{K}$) denote the remaining service time of class k customer being served. We then define $a_k^{**}(s_k, \alpha_k)$ ($k \in \mathcal{K}$) as

$$a_k^{**}(s_k, \alpha_k) = \mathbb{E}[e^{-s_k A_k} \cdot e^{-\alpha_k \tilde{H}_k} \mid \text{a class } k \text{ customer is being served}].$$

Note that $a_k^{**}(s_k, \alpha_k)$ corresponds to $a_{WV,WV,k}^{**}(\omega_k, \alpha_k)$ in the queue with working vacations. Moreover, Q_k and W_k corresponds to $Q_{WV,k}^{(\gamma)}$ and $W_{WV,k}^{(\gamma)}$, respectively. It then follows from (27) that

$$a_k^{**}(s_k, \alpha_k) = \frac{u^*(s_k+\eta)}{\mathbb{E}[Q_k - W_k]} \cdot \frac{h_k^*(\alpha_k) - h_k^*(s_k+\eta)}{s_k+\eta - \alpha_k},$$

where $\mathbb{E}[Q_k - W_k]$ is obtained from (47).

$$\mathbb{E}[Q_k - W_k] = \frac{\nu}{\lambda} \cdot \frac{1-h_k^*(\eta)}{1-h^*(\eta)}.$$

Let L_k ($k \in \mathcal{K}$) denote the number of class k customers in the system and let U_k ($k \in \mathcal{K}$) denote the total amount of work in system belonging to class k . We then define the joint transform $\psi(\mathbf{z}, \mathbf{s})$ as

$$\psi(\mathbf{z}, \mathbf{s}) = \mathbb{E} \left[\prod_{k \in \mathcal{K}} z_k^{L_k} \cdot e^{-s_k U_k} \right],$$

where $\mathbf{z} = (z_1, z_2, \dots, z_K)$ and $\mathbf{s} = (s_1, s_2, \dots, s_K)$. We then have

$$\psi(\mathbf{z}, \mathbf{s}) = 1 - \nu + \sum_{k \in \mathcal{K}} z_k \sigma_k a_k^{**} \left(\sum_{i \in \mathcal{K}} [\lambda_i - \lambda_i z_i h_i^*(s_i)], s_k \right),$$

which corresponds to Theorem 5.

Finally, we consider the busy cycle, which is defined as the interval between successive ends of busy periods. Let Θ denote the length of a randomly chosen busy cycle. We divide Θ into two parts, and let Θ_E (resp. Θ_B) denote the length of the subinterval during which the server is idle (resp. busy). We define \bar{N}_k ($k \in \mathcal{K}$) as the number of class k customers arriving during Θ . Let \tilde{U}_L denote the amount of work in system that is lost due to disasters. We then define joint transforms $\theta_N^*(\mathbf{z}, s)$ and $\theta_D^*(\mathbf{z}, s, \alpha)$ as follows.

$$\theta_N^*(\mathbf{z}, \omega, s) = \mathbb{E} \left[\left(\prod_{k \in \mathcal{K}} z_k^{\bar{N}_k} \right) \cdot e^{-\omega \Theta_E} \cdot e^{-s \Theta_B} \mid \text{a busy period ends without disasters} \right],$$

$$\theta_D^*(\mathbf{z}, \omega, s, \alpha) = \mathbb{E} \left[\left(\prod_{k \in \mathcal{K}} z_k^{\bar{N}_k} \right) \cdot e^{-\omega \Theta_E} \cdot e^{-s \Theta_B} \cdot e^{-\alpha \bar{U}_L} \mid \text{a busy period ends with disasters} \right].$$

We also define P_N^B as

$$P_N^B = \Pr(\text{a busy period ends without disasters}),$$

and let $P_D^B = 1 - P_N^B$. It then follows from Lemma 7 that

$$\begin{aligned} P_N^B \cdot \theta_N^*(\mathbf{z}, \omega, s) &= \frac{\lambda}{\lambda + \omega} \sum_{k \in \mathcal{K}} z_k \cdot \frac{\lambda_k}{\lambda} \int_0^\infty e^{-\beta_N^*(\mathbf{z}, s)y} dH_k(y) \\ &= \frac{\lambda}{\lambda + \omega} \sum_{k \in \mathcal{K}} z_k \cdot \frac{\lambda_k h_k^*(\beta_N^*(\mathbf{z}, s))}{\lambda} \end{aligned} \tag{57}$$

$$= \frac{\lambda}{\lambda + \omega} \cdot \frac{s + \eta + \lambda - \beta_N^*(\mathbf{z}, s)}{\lambda}, \tag{58}$$

$$\begin{aligned} P_D^B \cdot \theta_N^*(\mathbf{z}, \omega, s, \alpha) &= \frac{\lambda}{\lambda + \omega} \sum_{k \in \mathcal{K}} z_k \cdot \frac{\lambda_k}{\lambda} \int_0^\infty \frac{e^{-\alpha y} - e^{-\beta_N^*(\mathbf{z}, s)y}}{\beta_N^*(\mathbf{z}, s) - \alpha} \cdot \beta_D^*(\mathbf{z}, s, \alpha) dH_k(y) \\ &= \frac{\lambda}{\lambda + \omega} \sum_{k \in \mathcal{K}} z_k \cdot \frac{\lambda_k}{\lambda} \cdot \frac{h_k^*(\alpha) - h_k^*(\beta_N^*(\mathbf{z}, s))}{\beta_N^*(\mathbf{z}, s) - \alpha} \cdot \beta_D^*(\mathbf{z}, s, \alpha) \\ &= \frac{\lambda}{\lambda + \omega} \cdot \frac{\beta_D^*(\mathbf{z}, s, \alpha) - \eta}{\lambda}, \end{aligned} \tag{59}$$

where $\beta_N^*(\mathbf{z}, s)$ and $\beta_D^*(\mathbf{z}, s, \alpha)$ satisfy

$$\begin{aligned} \beta_N^*(\mathbf{z}, s) &= s + \eta + \lambda - \sum_{k \in \mathcal{K}} z_k \lambda_k h_k^*(\beta_N^*(\mathbf{z}, s)), \\ \beta_D^*(\mathbf{z}, s, \alpha) &= \eta + \sum_{k \in \mathcal{K}} z_k \lambda_k \cdot \frac{h_k^*(\alpha) - h_k^*(\beta_N^*(\mathbf{z}, s))}{\beta_N^*(\mathbf{z}, s) - \alpha} \cdot \beta_D^*(\mathbf{z}, s, \alpha), \end{aligned}$$

which correspond to $\beta_{WV,C}^*(\mathbf{z}_{WV}, s_{WV})$ in (36) and $\beta_{WV,E}^*(\mathbf{z}_{WV}, s_{WV}, \alpha)$ in (37), respectively.

We define β_N and β_D as

$$\beta_N = \beta_N^*(\mathbf{1}, 0), \quad \beta_D = \beta_D^*(\mathbf{1}, 0, 0).$$

We then have (cf. Lemma 8 and its proof)

$$\beta_N = \beta_D = \frac{\eta}{1 - \nu}, \quad h^*(\beta_N) = r,$$

where r is the minimum nonnegative root of (55). It then follows from (57) that

$$P_N^B = r, \quad P_D^B = 1 - r,$$

and from (58) and (59) that

$$\theta_N^*(\mathbf{z}, \omega, s) = \frac{\lambda}{\lambda + \omega} \cdot \frac{s + \eta + \lambda - \beta_N^*(\mathbf{z}, s)}{\eta + \lambda - \beta_N}, \quad \theta_D^*(\mathbf{z}, \omega, s, \alpha) = \frac{\lambda}{\lambda + \omega} \cdot \frac{\beta_D^*(\mathbf{z}, s, \alpha) - \eta}{\beta_D - \eta}.$$

References

- [1] O.J. Boxma and T. Takine: The M/G/1 FIFO queue with several customer classes. *Queueing Systems*, **45** (2003), 185–189.
- [2] B. Doshi: Conditional and unconditional distributions for M/G/1 type queues with server vacations. *Queueing Systems*, **7** (1990), 229–252.
- [3] G. Jain and K. Sigman: A Pollaczek-Khintchine formula for M/G/1 queues with disasters. *Journal of Applied Probability*, **33** (1996), 1191–1200.
- [4] J. Keilson and L.D. Servi: A distributional form of Little’s law. *Operations Research Letters*, **7** (1988), 223–227.
- [5] J.D. Kim, D.W. Choi, and K.C. Chae: Analysis of queue-length distribution of the M/G/1 queue with working vacations (M/G/1/WV). *Proceedings of Hawaii International Conference on Statistics and Related Fields* (2003), 1191–1200.
- [6] J.-H. Li, N.-S. Tian, A.G. Zhang, and H.P. Luh: Analysis of the M/G/1 queue with exponentially working vacations — A matrix analytic approach. *Queueing Systems*, **61** (2009), 139–166.
- [7] W.-Y. Liu, X.-L. Xu, and N.-S. Tian: Stochastic decompositions in the M/M/1 queue with working vacations. *Operations Research Letters*, **35** (2007), 595–600.
- [8] B. Sengupta: An invariance relationship for the G/G/1 queue. *Advances in Applied Probability*, **21** (1989), 956–957.
- [9] L.D. Servi and S.G. Finn: M/M/1 queues with working vacations (M/M/1/WV). *Performance Evaluation*, **50** (2002), 41–52.
- [10] T. Takine: Queue length distribution in a FIFO single-server queue with multiple arrival streams having different service time distributions. *Queueing Systems*, **39** (2001), 349–375.
- [11] T. Takine: Single-server queues with Markov-modulated arrivals and service speed. *Queueing Systems*, **49** (2005), 7–22.
- [12] T. Takine and T. Hasegawa: A generalization of the decomposition property in the M/G/1 queue with server vacations. *Operations Research Letters*, **12** (1992), 97–99.
- [13] R.W. Wolff: Poisson arrivals see time averages. *Operations Research*, **30** (1982), 223–231.
- [14] D.-A. Wu and H. Takagi: M/G/1 queue with multiple working vacations. *Performance Evaluation*, **63** (2006), 654–681.
- [15] W.S. Yang, J.D. Kim, and K.C. Chae: Analysis of M/G/1 stochastic clearing systems. *Stochastic Analysis and Applications*, **20** (2002), 1083–1100.

Yoshiaki Inoue
Graduate School of Engineering
Osaka University
2-1 Yamadaoka, Suita
Osaka 565-0871, Japan
E-mail: y-inoue@post.comm.eng.osaka-u.ac.jp