$M/PH/1$ QUEUE WITH DETERMINISTIC IMPATIENCE TIME

JERIM KIM AND JEONGSIM KIM

ABSTRACT. We consider an $M/PH/1$ queue with deterministic impatience time. An exact analytical expression for the stationary distribution of the workload is derived. By modifying the workload process and using Markovian structure of the phase-type distribution for service times, we are able to construct a new Markov process. The stationary distribution of the new Markov process allows us to find the stationary distribution of the workload. By using the stationary distribution of the workload, we obtain performance measures such as the loss probability, the waiting time distribution and the queue size distribution.

1. Introduction

In many service systems, customers wait for service for a limited time only and leave the system if not served during that time. Such customers with limited waiting time are usually referred to as impatient customers. Impatience phenomena are often encountered in real-time communication systems, inventory systems with storage of perishable goods, telecommunication networks, call centers, etc.

Systems with limited waiting times can be classified as follows:

- The limitation acts only on waiting time or on sojourn time.
- The customer can calculate his prospective waiting time at the arrival epoch and balks if this exceeds his patience or he joins the queue regardless, leaving the system if and when his patience expires.

Combining these two classifications, Baccelli et al. [3] described the following four queueing systems with impatient customers:

(a) limitation on sojourn time (impatience until the end of service), aware customers: The entering customer leaves immediately if he knows that his total sojourn time is beyond his patience.

(b) limitation on sojourn time, unaware customers: This is the case if customers do not know anything about the system and are unaware of the beginning of the service.

(c) limitation on waiting time (impatience until the beginning of service), aware customers: The same as (a) above with the impatience acting only on waiting time.

(d) limitation on waiting time, unaware customers: The same as (b) above with the impatience acting only on waiting time.

This paper deals with the case (d). There has been a large amount of literature on queues with impatient customers. We use the notation $G/G/c+G$: the first three symbols have the same meaning as in Kendall's notation and the last one specifies the impatience law. Barrer [5] and [6] analyzed the $M/M/1+D$ and $M/M/c + D$ queues with deterministic patience times and obtained equilibrium queue size distributions. Gnedenko and Kovalenko [12] (pp. 34 and 47) pointed out that the results of Barrer [6] are perfectly correct, but his derivation is not faultless, and then gave the correct derivation for the same models (pp. 33–47). Jurkevic [13] and [14] analyzed the $M/M/c$ queue where impatient time is the minimum of a constant and an exponentially distributed time, and the $M/M/c + G$ queue with general impatience time distribution. Independently, the $M/M/c+G$ queue was analyzed by Baccelli and Hebuterne [4]. Boxma and de Waal [7] developed several approximations for the overflow probability in the $M/M/c+G$ queue. The derivation of performance measures for $M/M/c+G$ queue continued in Brandt and Brandt [8, 9]. They considered the more general $M(n)/M(n)/c + G$ queue where arrival and service rates are allowed to depend on the number n of calls in the system. For the $M(n)/M/c+G$ queue, see Movaghar [16].

de Kok and Tijms [15] and Xiong et al. [21] studied the $M/G/1+D$ queue. de Kok and Tijms [15] obtained an expression for the distribution function of the workload in terms of the workload in the modified system. They obtained the exact expressions for the loss probability and the mean waiting time in the $M/M/1 + D$ queue. Xiong et al. [21] set up an integral equation for the distribution of the workload using level crossing analysis. An analytical solution for this equation was given only for $M/H_2/1 + D$ queue. de Kok and Tijms [15] and Xiong et al. [21] presented approximations for the loss probability and the mean waiting time in the $M/G/1 + D$ queue.

Finch [11] obtained the actual waiting time distribution for the $G/M/1+D$ queue. Bae and Kim [2] derived the stationary distribution of the workload in the $G/M/1 + D$ queue using level crossing analysis. Daley [10] studied the $GI/G/1 + G$ queue by setting up an integral equation for the waiting time distribution and focused on $M/G/1 + D$ and $M/G/1 + M$ (where $+M$ refers to the exponential impatience time) queues. The $GI/G/1 + G$ queue was also studied by Baccelli et al. [3] and Stanford [18, 19], and results for actual and virtual waiting times were obtained.

This paper is inspired by de Kok and Tijms [15] and Xiong et al. [21] who studied the $M/G/1+D$ queue. As mentioned above, de Kok and Tijms [15] and Xiong et al. [21] obtained the exact expressions for performance measures only for exponential and two-stage hyper-exponential service times. Therefore, we want to find the exact expressions for performance measures for a more general service time distribution than exponential and hyper-exponential distributions. This leads us to consider a phase-type distribution for service times. Note that every distribution can be approximated arbitrarily closely by a phase-type distribution.

In this paper we consider an $M/PH/1$ queue with deterministic impatience time in which customers have phase-type distributed service times. The purpose of this paper is to obtain the exact analytical expressions for performance measures. First we obtain an exact analytical expression for the stationary distribution of the workload. This is a generalization of the result of Xiong et al. [21] as mentioned above. To get the stationary distribution of the workload, we construct a new Markov process by modifying the workload process and using Markovian structure of the phase-type distribution for service times, and then obtain the stationary distribution of the Markov process. By using the stationary distribution of the workload, we derive the exact analytical expressions for performance measures such as the loss probability, the waiting time distribution and the queue size distribution.

2. Analysis of the workload process

We consider the $M/PH/1$ queue with deterministic impatience time τ in which customers arrive according to a Poisson process with intensity λ. The service time has a phase-type distribution with representation $(\boldsymbol{\alpha}, T)$ of order m and mean μ^{-1}. The offered load ρ is defined as $\rho \equiv \frac{\lambda}{\mu}$.

We denote by $V(t)$ the workload (unfinished work, or virtual waiting time) at time t. Figure 1 shows a sample path of $\{V(t) : t \geq 0\}$. Note that an arriving customer who finds that the workload exceeds τ is lost, when his attained waiting time becomes τ. Clearly $\{V(t) : t \geq 0\}$ is a Markov process. Furthermore, it is a regenerative process with returning points to 0 as regeneration epochs. The mean of a regeneration cycle is finite since customers who arrive when the workload is larger than τ will be lost; and the mean service time is finite. In addition, the distribution of a regeneration cycle is nonlattice. Hence the workload process $\{V(t) : t \geq 0\}$ has a limiting distribution, which is also a stationary distribution, see Theorem 17 on p. 112 in [20]; Theorem 20 on p. 120 in [20]; and Theorem 1.2 on p. 170 in [1]. Suppose that $\{V(t) : t \geq 0\}$ is in the steady state, i.e., $\{V(t) : t \geq 0\}$ is stationary. Let

$$P(x) \equiv \mathbb{P}(V(t) \leq x), \ x \in \mathbb{R}.$$

Define

(1) $\qquad \sigma \quad \equiv \quad \inf\{t > 0 : V(t) = 0, V(u) \neq 0 \text{ for some } u \in (0, t)\},$

$$\text{(2)} \qquad \nu \;\equiv\; \mathbb{E}\big(\sigma|V(0)=0\big).$$

Then, by Theorem 1.2 on p. 170 in [1], for every Borel subset B of $\mathbb{R}$,

$$\text{(3)} \qquad \int_B dP(x) = \frac{1}{\nu}\mathbb{E}\Big(\int_0^\sigma \mathbb{1}_{\{V(t)\in B\}}dt \mid V(0)=0\Big).$$

Specifically if B is a Lebesgue null set of $(0,\infty)$, then the right-hand side of (3) is zero (see Figure 1). Hence $\int_B dP(x) = 0$ for all Lebesgue null sets B of $(0,\infty)$, which implies that $P(x)$ is absolutely continuous in x on $[0,\infty)$. From this we have the following lemma.

Lemma 1. *There exists a density function $p(x)$, $x \geq 0$, such that*

$$P(x) - P(0) = \int_0^x p(y)\ dy.$$

Define

$$p_0 \equiv P(0) = \mathbb{P}(V(t)=0).$$

The remainder of this section is devoted to the derivation of p_0 and $p(x)$, $x \geq 0$. First, we introduce a new process $\{\tilde{V}(s) : s \geq 0\}$ by modifying $\{V(t) : t \geq 0\}$. We explain how to obtain this by illustrations with a sample path. Let $0 < t_1 < t_2 < \cdots$ be the arrival epochs of customers who will be served, see Figure 1. Let S_k, $k = 1, 2, \ldots$ be the required service time of the customer who arrives at time t_k. Let $A(t)$ be the sum of service times for the customers who arrive until time t and will be served, i.e.,

$$A(t) \equiv \sum_{t_k \leq t} S_k,\ t \geq 0.$$

Figure 2 shows the plot of $(t + A(t), V(t))$. Now we obtain Figure 3 by adding line segments with slope 1 to Figure 2. Let $\{\tilde{V}(s) : s \geq 0\}$ be the process obtained in this way, as in Figure 3.

Let us describe the phase-type distribution for service times in detail. Consider a continuous time Markov process with state space $\{1, \ldots, m, m+1\}$ and an infinitesimal generator of the form

$$\text{(4)} \qquad \begin{bmatrix} T & \boldsymbol{T}^0 \\ \boldsymbol{0}^\top & 0 \end{bmatrix},$$

where $T = (T_{ij})_{1\leq i,j\leq m}$ is a nonsingular $m\times m$ matrix and $\boldsymbol{T}^0 = (T_1^0, \ldots, T_m^0)^\top$ is an m-dimensional column vector satisfying $T\boldsymbol{1} + \boldsymbol{T}^0 = 0$. Here and subsequently, $\boldsymbol{0}$ and $\boldsymbol{1}$ are m-dimensional column vectors with all components equal to zero and one, respectively. Let $(\boldsymbol{\alpha}, 0)$ be the initial distribution of the Markov process, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)$ with $\sum_{i=1}^m \alpha_i = 1$. Then the time until absorption into the state $m+1$ in the Markov process has the phase-type distribution with representation $(\boldsymbol{\alpha}, T)$. Without loss of generality, we may assume that the representation $(\boldsymbol{\alpha}, T)$ is irreducible in the sense that the Markov process with initial distribution $(\boldsymbol{\alpha}, 0)$ and infinitesimal generator (4) has no redundant states. Equivalently, the square matrix with nonnegative off-diagonal entries, $\boldsymbol{T}^0\boldsymbol{\alpha} + T$ is assumed to be irreducible.
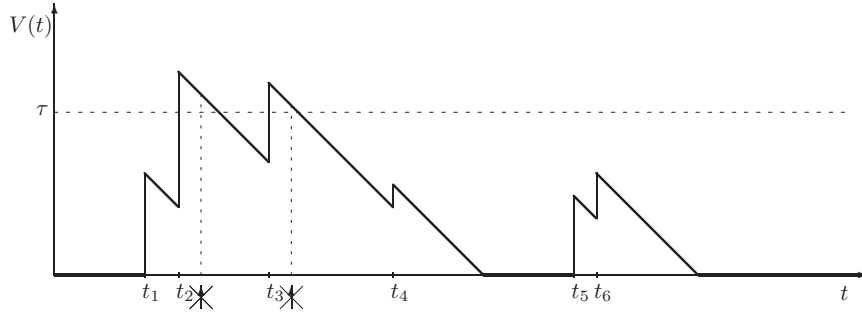
$t_1, t_2, \ldots$: arrival epochs of customers who will be served.

✳ : arrival epochs of customers who will be lost.

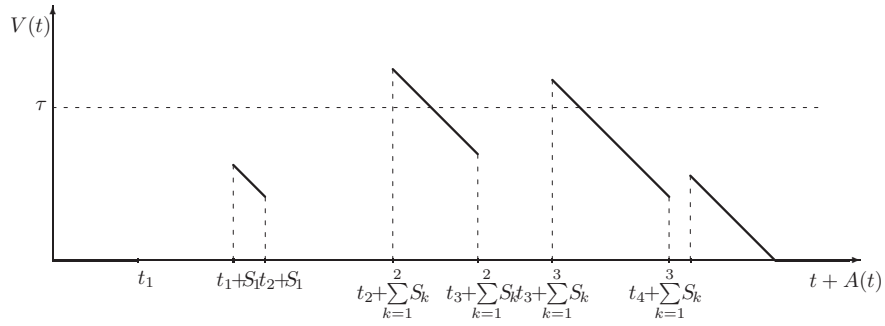FIGURE 1. A sample path of the workload $\{V(t) : t \geq 0\}$.
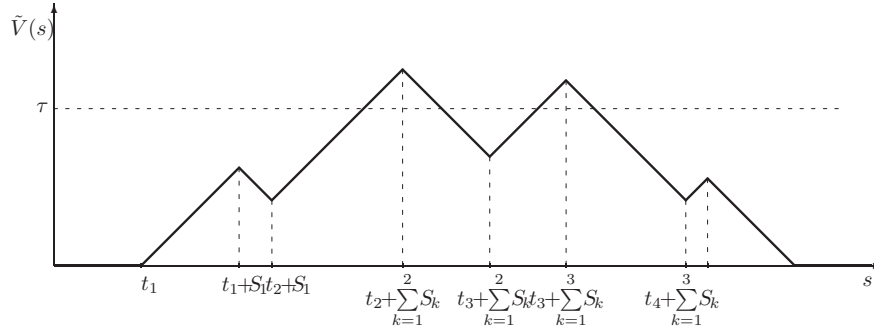

FIGURE 2. Plot $(t + A(t), V(t))$.

FIGURE 3. A sample path of $\{\tilde{V}(s) : s \geq 0\}$.

Let $\{J_k(s) : s \geq 0\}$, $k = 1, 2, \ldots$, be independent Markov processes on $\{1, 2, \ldots, m+1\}$ with initial distribution $(\boldsymbol{\alpha}, 0)$ and infinitesimal generator (4).

The service times S_k, $k = 1, 2, \ldots$, are obtained by

$$S_k = \inf\{s \geq 0 : J_k(s) = m + 1\}.$$

Define

$$\tilde{J}(s) = \begin{cases} J_k(s - (t_k + \sum_{l=1}^{k-1} S_l)) & \text{if } t_k + \sum_{l=1}^{k-1} S_l \leq s < t_k + \sum_{l=1}^{k} S_l \\ & \text{for some } k = 1, 2, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\{(\tilde{V}(s), \tilde{J}(s)) : s \geq 0\}$ is a Markov process. Observe that the Markov process $\{(\tilde{V}(s), \tilde{J}(s)) : s \geq 0\}$ is also a regenerative process with returning points to $(0, 0)$ as regeneration epochs. The mean of a regeneration cycle is finite since customers who arrive when the workload is larger than τ will be lost; and the mean service time is finite. In addition, the distribution of a regeneration cycle is nonlattice. Hence the process $\{(\tilde{V}(s), \tilde{J}(s)) : s \geq 0\}$ has a limiting distribution, which is also a stationary distribution, see Theorem 17 on p. 112 in [20]; Theorem 20 on p. 120 in [20]; and Theorem 1.2 on p. 170 in [1]. Suppose that $\{(\tilde{V}(s), \tilde{J}(s)) : s \geq 0\}$ is in the steady state, i.e., $\{(\tilde{V}(s), \tilde{J}(s)) : s \geq 0\}$ is stationary. We define

$$F_i(x) = \mathbb{P}(\tilde{V}(s) \leq x, \tilde{J}(s) = i), \ x \geq 0, \ i = 0, 1, \ldots, m,$$

and

$$\boldsymbol{F}(x) = \big(F_1(x), F_2(x), \ldots, F_m(x)\big).$$

We can express p_0 and $p(x)$ in terms of $F_0(0)$ and $F_0(x)$ as shown in the following lemma.

Lemma 2. *We have*

(i) $F_0(0) > 0$ *and* $F_i(0) = 0$ *for* $i = 1, 2, \ldots, m$;

(ii) $F_0(x) - F_0(0) = \boldsymbol{F}(x)\mathbf{1}$;

(iii) $p_0 = \frac{2F_0(0)}{1+F_0(0)}$ *and* $\int_0^x p(y)dy = \frac{2}{1+F_0(0)}(F_0(x) - F_0(0))$.

Proof. We define

$$\tilde{\sigma} = \inf\{s > 0 : (\tilde{V}(s), \tilde{J}(s)) = (0, 0), (\tilde{V}(u), \tilde{J}(u)) \neq (0, 0) \text{ for some } u \in (0, s)\},$$

$$\tilde{\nu} = \mathbb{E}\big(\tilde{\sigma}|(\tilde{V}(0), \tilde{J}(0)) = (0, 0)\big).$$

By Theorem 1.2 on page 170 in [1], we have

$$F_i(0) = \frac{1}{\tilde{\nu}}\mathbb{E}\left(\int_0^{\tilde{\sigma}} \mathbb{1}_{\{(\tilde{V}(s), \tilde{J}(s))=(0,i)\}}ds|(\tilde{V}(0), \tilde{J}(0)) = (0, 0)\right).$$

On $\{(\tilde{V}(0), \tilde{J}(0)) = (0, 0)\}$, it is observed that

$$\int_0^{\tilde{\sigma}} \mathbb{1}_{\{(\tilde{V}(s), \tilde{J}(s))=(0,i)\}}ds = 0 \ \text{ if } i = 1, \ldots, m,$$

$$\int_0^{\tilde{\sigma}} \mathbb{1}_{\{(\tilde{V}(s), \tilde{J}(s))=(0,i)\}}ds > 0 \ \text{ if } i = 0.$$

Hence $F_i(0) > 0$ if and only if $i = 0$, and so (i) is proved.

Next we prove (ii). We observe that

$$\int_0^{\tilde{\sigma}} \mathbb{1}_{\{\tilde{V}(s)\in(0,x],\tilde{J}(s)=0\}} ds$$

$$(5) \qquad = \int_0^{\tilde{\sigma}} \mathbb{1}_{\{\tilde{V}(s)\in(0,x],1\le\tilde{J}(s)\le m\}} ds \quad \text{on } \{(\tilde{V}(0),\tilde{J}(0))=(0,0)\}.$$

By Theorem 1.2 on page 170 in [1], we have

$$F_0(x) - F_0(0) = \frac{1}{\tilde{\nu}}\mathbb{E}\left(\int_0^{\tilde{\sigma}} \mathbb{1}_{\{\tilde{V}(s)\in(0,x],\tilde{J}(s)=0\}} ds | (\tilde{V}(0),\tilde{J}(0))=(0,0)\right),$$

$$\boldsymbol{F}(x)\boldsymbol{1} = \frac{1}{\tilde{\nu}}\mathbb{E}\left(\int_0^{\tilde{\sigma}} \mathbb{1}_{\{\tilde{V}(s)\in(0,x],1\le\tilde{J}(s)\le m\}} ds | (\tilde{V}(0),\tilde{J}(0))=(0,0)\right),$$

which together with (5) leads to (ii).

Now we prove (iii). Recall σ and ν in (1) and (2). On $\{V(0)=0\}$, hence on $\{(\tilde{V}(0),\tilde{J}(0))=(0,0)\}$, it is observed that

$$\int_0^{\sigma} \mathbb{1}_{\{V(t)\in B\}} dt = \int_0^{\tilde{\sigma}} \mathbb{1}_{\{\tilde{V}(s)\in B,\tilde{J}(s)=0\}} ds$$

for all Borel sets B in $\mathbb{R}$. By Theorem 1.2 on page 170 in [1] again, we have

$$p_0 = \frac{1}{\nu}\mathbb{E}\left(\int_0^{\sigma} \mathbb{1}_{\{V(t)=0\}} dt | V(0)=0\right)$$

$$= \frac{\tilde{\nu}}{\nu}\frac{1}{\tilde{\nu}}\mathbb{E}\left(\int_0^{\tilde{\sigma}} \mathbb{1}_{\{(\tilde{V}(s),\tilde{J}(s))=(0,0)\}} ds | (\tilde{V}(0),\tilde{J}(0))=(0,0)\right)$$

$$(6) \qquad = \frac{\tilde{\nu}}{\nu}F_0(0),$$

and

$$\int_0^x p(y)dy = \frac{1}{\nu}\mathbb{E}\left(\int_0^{\sigma} \mathbb{1}_{\{V(t)\in(0,x]\}} dt | V(0)=0\right)$$

$$= \frac{\tilde{\nu}}{\nu}\frac{1}{\tilde{\nu}}\mathbb{E}\left(\int_0^{\tilde{\sigma}} \mathbb{1}_{\{\tilde{V}(s)\in(0,x],\tilde{J}(s)=0\}} ds | (\tilde{V}(0),\tilde{J}(0))=(0,0)\right)$$

$$(7) \qquad = \frac{\tilde{\nu}}{\nu}(F_0(x) - F_0(0)).$$

Further,

$$\nu = \mathbb{E}\left(\int_0^{\sigma} ds | V(0)=0\right)$$

$$= \tilde{\nu}\frac{1}{\tilde{\nu}}\mathbb{E}\left(\int_0^{\tilde{\sigma}} \mathbb{1}_{\{\tilde{J}(s)=0\}} ds | (\tilde{V}(0),\tilde{J}(0))=(0,0)\right)$$

(8) $$= \tilde{\nu} F_0(\infty).$$

Noting from (ii) that $F_0(\infty) - F_0(0) = \boldsymbol{F}(\infty)\boldsymbol{1}$ and using $F_0(\infty) + \boldsymbol{F}(\infty)\boldsymbol{1} = 1$, we have

(9) $$F_0(\infty) = \frac{1 + F_0(0)}{2}.$$

Substituting (9) into (8) yields

$$\frac{\tilde{\nu}}{\nu} = \frac{2}{1 + F_0(0)}.$$

Finally, substituting the above into (6) and (7) leads to the assertion (iii). $\square$

Corollary 1. *For each $i = 0, 1, \ldots, m$, there exists a density function $f_i(x)$, $x \geq 0$ such that*

(10) $$F_i(x) - F_i(0) = \int_0^x f_i(y) dy.$$

Proof. From Lemma 2(iii) it follows that $f_0(x) = \frac{1 + F_0(0)}{2} p(x)$ satisfies (10) for $i = 0$. By Lemma 2(ii), $\boldsymbol{F}(x)\boldsymbol{1} = F_0(x) - F_0(0)$. Since $F_0(x)$, $x \geq 0$ is absolutely continuous, so are $F_i(x)$ for all $i = 1, \ldots, m$. This implies that there are density functions $f_i(x)$, $i = 1, \ldots, m$, satisfying (10). $\square$

The density functions $f_i(x)$, $x \geq 0$, $i = 1, 2, \ldots, m$, satisfy the following lemma.

Lemma 3. *Let $\boldsymbol{f}(x) = (f_1(x), f_2(x), \ldots, f_m(x))$. Then $\boldsymbol{f}(x)$ can be chosen so that it is continuous in $x \geq 0$. Moreover, it satisfies*

(11) $$\boldsymbol{f}(x) = \begin{cases} \boldsymbol{F}(x)M + \lambda F_0(0)\boldsymbol{\alpha}, & 0 \leq x \leq \tau, \\ (\boldsymbol{F}(x) - \boldsymbol{F}(\infty))T, & x > \tau, \end{cases}$$

where $M \equiv \lambda \boldsymbol{1}\boldsymbol{\alpha} + T$.

Proof. By Theorem 7.20 in [17], $\boldsymbol{F}(x)$ is differentiable in x almost everywhere with respect to the Lebesgue measure (Lebesgue-a.e) and

$$\boldsymbol{f}(x) = \frac{d}{dx}\boldsymbol{F}(x), \ x > 0, \text{ Lebesgue-a.e.}$$

For $x \geq 0$, $h > 0$ and $i = 1, 2, \ldots, m$,

$$F_i(x + h) = \mathbb{P}(\tilde{V}(h) \leq x + h, \tilde{J}(h) = i)$$

$$= \mathbb{P}(\tilde{V}(h) \leq x + h, \tilde{J}(h) = i, \tilde{J}(0) = i)$$

$$+ \sum_{1 \leq j \leq m, j \neq i} \mathbb{P}(\tilde{V}(h) \leq x + h, \tilde{J}(h) = i, \tilde{J}(0) = j)$$

(12) $$+ \mathbb{P}(\tilde{V}(h) \leq x + h, \tilde{J}(h) = i, \tilde{J}(0) = 0).$$

It is observed that

- For $x \geq 0$, $h > 0$ and $1 \leq i \leq m$,

$$\mathbb{P}(\tilde{V}(h) \leq x + h, \tilde{J}(h) = i, \tilde{J}(0) = i)$$
$$= \mathbb{P}(\tilde{V}(0) \leq x, \tilde{J}(0) = i)(1 + T_{ii}h) + o(h)$$
(13) $$= F_i(x) + F_i(x)T_{ii}h + o(h).$$

- For $x \geq 0$, $h > 0$ and $1 \leq i \neq j \leq m$,

$$\mathbb{P}(\tilde{V}(h) \leq x + h, \tilde{J}(h) = i, \tilde{J}(0) = j)$$
$$= \mathbb{P}(\tilde{V}(0) \leq x, \tilde{J}(0) = j)T_{ji}h + o(h)$$
(14) $$= F_j(x)T_{ji}h + o(h).$$

- For $x \geq 0$, $h > 0$ and $1 \leq i \leq m$,

$$\mathbb{P}(\tilde{V}(h) \leq x + h, \tilde{J}(h) = i, \tilde{J}(0) = 0)$$
$$= \begin{cases} \mathbb{P}(\tilde{V}(0) \leq x, \tilde{J}(0) = 0)\lambda h \alpha_i + o(h), & \text{if } x \leq \tau \\ \mathbb{P}(\tilde{V}(0) \leq \tau, \tilde{J}(0) = 0)\lambda h \alpha_i + o(h), & \text{if } x > \tau \end{cases}$$
$$= F_0(\min\{x, \tau\})\lambda h \alpha_i + o(h)$$
(15) $$= (F_0(0) + \boldsymbol{F}(\min\{x, \tau\})\boldsymbol{1})\lambda h \alpha_i + o(h),$$

where the last equality follows from Lemma 2(ii).

Substituting (13)-(15) into (12), we have for $x \geq 0$ and $1 \leq i \leq m$,

$$\frac{1}{h}(F_i(x + h) - F_i(x))$$
$$= \sum_{j=1}^{m} F_j(x)T_{ji} + \lambda(F_0(0) + \boldsymbol{F}(\min\{x, \tau\})\boldsymbol{1})\alpha_i + o(1) \text{ as } h \to 0+ .$$

Letting $h \to 0+$ yields

$$\boldsymbol{f}(x) = \boldsymbol{F}(x)T + \lambda(F_0(0) + \boldsymbol{F}(\min\{x, \tau\})\boldsymbol{1})\boldsymbol{\alpha}, \; x \geq 0, \text{ Lebesgue-a.e.}$$

Modifying $\boldsymbol{f}$ on a Lebesgue null set, we may assert

(16) $$\boldsymbol{f}(x) = \boldsymbol{F}(x)T + \lambda(F_0(0) + \boldsymbol{F}(\min\{x, \tau\})\boldsymbol{1})\boldsymbol{\alpha} \text{ for all } x \geq 0.$$

This implies that $\boldsymbol{f}$ is continuous. By (16), $\boldsymbol{f}(\infty) \equiv \lim_{x \to \infty} \boldsymbol{f}(x)$ exists and $\boldsymbol{f}(\infty) = \boldsymbol{F}(\infty)T + \lambda(F_0(0) + \boldsymbol{F}(\tau)\boldsymbol{1})\boldsymbol{\alpha}$. Since $\boldsymbol{F}(\infty) = \int_0^\infty \boldsymbol{f}(x)dx$ is finite, we have $\boldsymbol{f}(\infty) = \boldsymbol{0}$ and so

(17) $$\lambda(F_0(0) + \boldsymbol{F}(\tau)\boldsymbol{1})\boldsymbol{\alpha} = -\boldsymbol{F}(\infty)T.$$

When $x > \tau$, (16) together with (17) yields

$$\boldsymbol{f}(x) = (\boldsymbol{F}(x) - \boldsymbol{F}(\infty))T, \; x > \tau.$$

This and (16) with $0 \leq x \leq \tau$ yield (11). $\square$

Solving the equation (11) gives

$$
(18) \qquad \boldsymbol{f}(x) = \begin{cases} \lambda F_0(0)\boldsymbol{\alpha}e^{Mx}, & 0 \le x \le \tau, \\ (\boldsymbol{F}(\tau) - \boldsymbol{F}(\infty))Te^{T(x-\tau)}, & x > \tau. \end{cases}
$$

Since $\boldsymbol{f}(x)$ is continuous at $x = \tau$, we have $(\boldsymbol{F}(\tau) - \boldsymbol{F}(\infty))T = \lambda F_0(0)\boldsymbol{\alpha}e^{M\tau}$ and (18) becomes

$$
(19) \qquad \boldsymbol{f}(x) = \begin{cases} \lambda F_0(0)\boldsymbol{\alpha}e^{Mx}, & 0 \le x \le \tau, \\ \lambda F_0(0)\boldsymbol{\alpha}e^{M\tau}e^{T(x-\tau)}, & x > \tau. \end{cases}
$$

Now we determine $F_0(0)$. From $F_0(\infty) + \boldsymbol{F}(\infty)\boldsymbol{1} = 1$ and Lemma 2(ii), we have

$$
1 = F_0(0) + 2\int_0^\infty \boldsymbol{f}(x)dx\boldsymbol{1},
$$

which leads to

$$
1 = F_0(0) + 2\lambda F_0(0)\boldsymbol{\alpha}\Big(\int_0^\tau e^{Mx}dx + e^{M\tau}(-T)^{-1}\Big)\boldsymbol{1}.
$$

Therefore

$$
(20) \qquad F_0(0) = \Big(1 + 2\lambda\boldsymbol{\alpha}\{\int_0^\tau e^{Mx}dx + e^{M\tau}(-T)^{-1}\}\boldsymbol{1}\Big)^{-1}.
$$

Finally, using (19), (20), and Lemma 2(ii) and (iii), we have the following theorem.

Theorem 1. *For the $M/PH/1$ queue with deterministic impatience time, the density function $p(x)$ of the workload is given by*

$$
p(x) = \begin{cases} p_0\lambda\boldsymbol{\alpha}e^{Mx}\boldsymbol{1}, & 0 \le x \le \tau, \\ p_0\lambda\boldsymbol{\alpha}e^{M\tau}e^{T(x-\tau)}\boldsymbol{1}, & x > \tau, \end{cases}
$$

where

$$
(21) \qquad p_0 = \Big(1 + \lambda\boldsymbol{\alpha}\{\int_0^\tau e^{My}dy + e^{M\tau}(-T)^{-1}\}\boldsymbol{1}\Big)^{-1}.
$$

We remark that for exponential service times, Theorem 1 is reduced to

$$
p(x) = \begin{cases} \frac{(1-\rho)\lambda e^{(\rho-1)\mu x}}{1-\rho^2 e^{(\rho-1)\mu\tau}}, & 0 \le x \le \tau, \\ \frac{(1-\rho)\lambda e^{\lambda\tau-\mu x}}{1-\rho^2 e^{(\rho-1)\mu\tau}}, & x > \tau, \end{cases}
$$

which is identical to the result of de Kok and Tijms [15] for $0 \le x \le \tau$.

Remark 1. When $\rho \le 1$, we have explicit expressions for the integral $\int_0^\tau e^{My}dy$ in (21). If $\rho < 1$, then it can be shown that all eigenvalues of M have negative real parts, and so M is invertible and

$$
\int_0^\tau e^{Mx}dx = (-M)^{-1}\big(I - e^{M\tau}\big).
$$

When $\rho = 1$, let $\boldsymbol{\xi} = (-T)^{-1}\mathbf{1}$ and $\boldsymbol{\eta} = (\boldsymbol{\alpha}(-T)^{-1}\boldsymbol{\xi})^{-1}\boldsymbol{\alpha}(-T)^{-1}$. It can be shown that all eigenvalues of $M - \boldsymbol{\xi}\boldsymbol{\eta}$ have negative real parts and

$$\int_0^\tau e^{Mx}dx = (\boldsymbol{\xi}\boldsymbol{\eta} - M)^{-1}(I - e^{(M-\boldsymbol{\xi}\boldsymbol{\eta})\tau}) + (\tau - 1 + e^{-\tau})\boldsymbol{\xi}\boldsymbol{\eta}.$$

3. Performance measures

In this section we obtain performance measures such as the loss probability, the waiting time distribution and the queue size distribution, by using Theorem 1.

Loss probability. The loss probability, denoted by p_{loss}, is the probability that the workload immediately before an arbitrary arrival is larger than τ, which is $1 - P(\tau)$ by the Poisson arrivals see time averages (PASTA) property. Therefore,

$$(22) \qquad p_{loss} = \int_\tau^\infty p(x)dx = p_0 \lambda \boldsymbol{\alpha} e^{M\tau}(-T)^{-1}\mathbf{1}.$$

We remark that for exponential service times, (22) is reduced to

$$p_{loss} = \frac{(1-\rho)\rho e^{(\rho-1)\mu\tau}}{1 - \rho^2 e^{(\rho-1)\mu\tau}},$$

which is identical to equation (4) in de Kok and Tijms [15].

Waiting time distribution. Let W denote the waiting time of an arbitrary customer among all customers who are served or reneged. By the PASTA property,

$$\mathbb{P}(W = 0) = p_0,$$

$$\frac{d}{dx}\mathbb{P}(W \le x) = p(x), \ 0 < x < \tau,$$

$$\mathbb{P}(W = \tau) = p_{loss}.$$

From this we can obtain the moments of W. Specifically,

$$(23) \qquad \mathbb{E}[W] = p_0 \lambda \boldsymbol{\alpha} \int_0^\tau x e^{Mx}dx\mathbf{1} + p_{loss}\tau,$$

$$(24) \qquad \mathbb{E}[W^2] = p_0 \lambda \boldsymbol{\alpha} \int_0^\tau x^2 e^{Mx}dx\mathbf{1} + p_{loss}\tau^2.$$

We remark that for exponential service times, (23) is reduced to

$$\mathbb{E}[W] = \frac{\frac{\rho}{1-\rho}\frac{1}{\mu} - (\frac{1}{1-\rho}\frac{1}{\mu} + \rho\tau)\rho e^{(\rho-1)\mu\tau}}{1 - \rho^2 e^{(\rho-1)\mu\tau}},$$

which is identical to equation (5) in de Kok and Tijms [15].

Let W_{served} denote the waiting time of an arbitrary customer who is served. Then the distribution of W_{served} is given by

$$\mathbb{P}(W_{served} = 0) = \frac{p_0}{1 - p_{loss}},$$

$$\frac{d}{dx}\mathbb{P}(W_{served} \leq x) = \frac{p(x)}{1 - p_{loss}}, \ 0 < x < \tau,$$

and

$$(25) \qquad \mathbb{E}[W_{served}] = \frac{p_0}{1 - p_{loss}}\lambda\boldsymbol{\alpha}\int_0^\tau xe^{Mx}dx\mathbf{1},$$

$$(26) \qquad \mathbb{E}[W_{served}^2] = \frac{p_0}{1 - p_{loss}}\lambda\boldsymbol{\alpha}\int_0^\tau x^2e^{Mx}dx\mathbf{1}.$$

Remark 2. When $\rho \leq 1$, we can obtain explicit expressions for the integrals $\int_0^\tau xe^{Mx}dx$ and $\int_0^\tau x^2e^{Mx}dx$ in (23)-(26) by the same argument as in Remark 1.

Queue size distribution. Let Q be the number of customers waiting in the queue at an arbitrary time, excluding the one who may be in service. By the PASTA property, Q has the same distribution as the number of customers in the queue immediately before an arbitrary arrival, which has the same distribution as the number of customers in the queue immediately after an arbitrary departure by Burke's theorem. Note that the departures from the queue consist of service initiations and reneging. The number of customers in the queue immediately after arbitrary service initiation or reneging is the number of customers arriving during the waiting time W of an arbitrary customer. Therefore we have

$$Q = N(W) \text{ in distribution,}$$

where $N(\cdot)$ is a Poisson process with intensity λ that is independent of W. Hence the probability generating function, $\mathbb{E}[z^Q]$, $|z| \leq 1$, of Q is given by

$$\mathbb{E}[z^Q] = \int_{0-}^\infty e^{\lambda x(z-1)}d\mathbb{P}(W \leq x)$$

$$(27) \qquad = p_0 + p_0\lambda\boldsymbol{\alpha}\int_0^\tau e^{\lambda x(z-1)}e^{Mx}dx\mathbf{1} + p_{loss}e^{\lambda\tau(z-1)}.$$

Differentiating the above equation with respect to z and letting $z \to 1$, we obtain the moments of Q. Specifically,

$$\mathbb{E}[Q] = \lambda\big(p_0\lambda\boldsymbol{\alpha}\int_0^\tau xe^{Mx}dx\mathbf{1} + p_{loss}\tau\big) = \lambda\mathbb{E}[W],$$

$$\mathbb{E}[Q^2] = \lambda^2\big(p_0\lambda\boldsymbol{\alpha}\int_0^\tau x^2e^{Mx}dx\mathbf{1} + p_{loss}\tau^2\big) + \lambda\mathbb{E}[W] = \lambda^2\mathbb{E}[W^2] + \lambda\mathbb{E}[W].$$

Now, let L be the number of customers in the system, including the one who may be in service. Then

$$L = \begin{cases} Q + 1 & \text{if the server is busy,} \\ 0 & \text{if the server is idle.} \end{cases}$$

Since $\mathbb{P}(L = 0) = p_0$, the above relation yields the probability generating function of L

$$\mathbb{E}[z^L] = p_0 + z(\mathbb{E}[z^Q] - p_0)$$

$$(28) \qquad = p_0 + zp_0\lambda\boldsymbol{\alpha}\int_0^\tau e^{\lambda x(z-1)}e^{Mx}dx\mathbf{1} + p_{loss}ze^{\lambda\tau(z-1)}.$$

From this we can obtain the moments of L. Specifically

$$\mathbb{E}[L] = \mathbb{E}[Q] + 1 - p_0,$$
$$\mathbb{E}[L^2] = E[Q^2] + 2\mathbb{E}[Q] + 1 - p_0.$$

Remark 3. When $\rho \le 1$, we can obtain explicit expressions for the integral $\int_0^\tau e^{\lambda x(z-1)}e^{Mx}dx$ in (27) and (28) by the same argument as in Remark 1.

References

[1] S. Asmussen, *Applied Probability and Queues*, Second Edition, Springer, 2003.

[2] J. Bae and S. Kim, *The stationary workload of the $G/M/1$ queue with impatient customers*, Queueing Syst. **64** (2010), no. 3, 253–265.

[3] F. Baccelli, P. Boyer, and G. Hebuterne, *Single server queues with impatient customers*, Adv. in Appl. Probab. **16** (1984), no. 4, 887–905.

[4] F. Baccelli and G. Hebuterne, *On queues with impatient customers*, Performance '81 (Amsterdam, 1981), 159–179, North-Holland, Amsterdam-New York, 1981.

[5] D. Y. Barrer, *Queueing with impatient customers and indifferent clerks*, Oper. Res. **5** (1957), 644–649.

[6] ———, *Queueing with impatient customers and ordered service*, Oper. Res. **5** (1957), 650–656.

[7] O. J. Boxma and P. R. de Waal, *Multiserver queues with impatient customers*, In: The Fundamental Role of Teletraffic in the Evolution of Telecommunications Networks (Proc. ITC 14), 743–756, North-Holland, Amsterdam, 1994.

[8] A. Brandt and M. Brandt, *On the $M(n)/M(n)/s$ queues with impatient calls*, Performance Evaluation **35** (1999), no. 1-2, 1–18.

[9] ———, *Asymptotic results and a Markovian approximation for the $M(n)/M(n)/s+GI$ system*, Queueing Syst. **41** (2002), no. 1-2, 73–94.

[10] D. J. Daley, *General customer impatience in the queue $GI/G/1$*, J. Appl. Probability **2** (1965), 186–205.

[11] P. D. Finch, *Deterministic customer impatience in the queueing system $GI/M/1$*, Biometrika **47** (1960), 45–52.

[12] B. V. Gnedenko and I. N. Kovalenko, *Introduction to Queueing Theory*, Israel Program for Scientific Translations, Jerusalem, 1968.

[13] O. M. Jurkevic, *On the investigation of many-server queueing systems with bounded waiting time* (*in Russian*), Izv. Akad. Nauk SSSR Techniceskaja Kibernetika **5** (1970), 50–58.

[14] ———, *On many-server systems with stochastic bounds for the waiting time* (*in Russian*), Izv. Akad. Nauk SSSR Techniceskaja Kibernetika **4** (1971), 39–46.

[15] A. G. de Kok and H. C. Tijms, *A queueing system with impatient customers*, J. Appl. Probab. **22** (1985), no. 3, 688–696.

[16] A. Movaghar, *On queueing with customer impatience until the beginning of service*, Queueing Syst. **29** (1998), no. 2-4, 337–350.

[17] W. Rudin, *Real and Complex Analysis*, Third Edition, McGraw-Hill, 1987.

[18] R. E. Stanford, *Reneging phenomena in single channel queues*, Math. Oper. Res. **4** (1979), no. 2, 162–178.

[19] R. E. Stanford, *On queues with impatience*, Adv. in Appl. Probab. **22** (1990), no. 3, 768–769.

[20] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*, Prentice Hall, Englewood Cliffs, NJ, 1989.

[21] W. Xiong, D. Jagerman, and T. Altiok, *M/G/1 queue with deterministic reneging times*, Performance Evaluation **65** (2008), no. 3-4, 308–316.

JERIM KIM
DEPARTMENT OF BUSINESS ADMINISTRATION
YONGIN UNIVERSITY
GYEONGGI 449-714, KOREA
E-mail address: jerimkim@yongin.ac.kr

JEONGSIM KIM
DEPARTMENT OF MATHEMATICS EDUCATION
CHUNGBUK NATIONAL UNIVERSITY
CHUNGBUK 361-763, KOREA
E-mail address: jeongsimkim@chungbuk.ac.kr