Journal of the Operations Research
Society of Japan
Vol. 43, No. 4, December 2000

A REGENERATIVE CYCLE APPROACH TO AN $M/G/1$ QUEUE WITH EXCEPTIONAL SERVICE

Hironobu Kengaku Masakiyo Miyazawa
Science University of Tokyo

(Received November 4, 1999; Revised August 3, 2000)

Abstract We are concerned with an $M/G/1$ queue in which service time distributions in each busy period may depend on the number of customers who have been served in the same busy period. This model is called an exceptional service model. Our major interest is to see a general structure of this model through the stationary waiting time distribution and some other characteristics. To this end, we take a regenerative cycle approach with respect to a busy cycle. This approach enables us to get several characteristics in tractable forms, from which we get some interesting properties of the exceptional service model. Numerical examples are presented as well.

1. Introduction

We consider a single server queue with Poisson arrivals and an infinite buffer, where service discipline is FIFO, i.e., first in first out. Service times are assumed to be independent, but their distributions may depend on the number of customers who have been served in the current busy period. Thus, the queue is a modification of the standard $M/G/1$ model, and called an $M/G/1$ queue with *exceptional service.*

This class of models are motivated by the fact that a server may need more (or less) time when he resumes service after idling. For example, he may need an extra job to resume service for warming up. This also reflects a certain aspect of cashing in a computer system, in which transfer times of data correspond to service times (see, e.g., [10]). Furthermore, in a practical service system, a server may be refreshed at the service resumption or may be getting tired. In those situations, it is very natural for the service times of not only the first customer but also the subsequent several customers to be changed. We consider this case, while the traditional setting usually assumes a single exceptional service. A primary concern of the exceptional service model is to study the effect of such deviations of service times at the beginning of each busy period.

The model is also closely related to queues with vacations (see [4] for their survey). In the latter model, a server takes a vacation, and the customers who arrive during the server vacation experience longer waiting times. However, those delays largely depend on a policy of the server vacation, and only the first customer at each busy period can be considered to have exceptional service for a typical vacation policy such as exhaustive service. On the other hand, in the exceptional service model, service time distributions are specified, and more than one customers may have exceptional service. Thus, the exceptional service model has a different model structure from the vacation model, although certain special cases such as the single exceptional service are indeed closely related. The exceptional service model can be also used to analyze other models such as priority queues (see, e.g., [13]).

The $M/G/1$ queue with exceptional service was firstly studied by Welch [15], and recently further studied by Igaki et al. [5] and Baba [2] (see also [10] for the case of exponentially distributed service times). For the case of the renewal arrivals, the single exceptional service model was studied as a modification of $GI/G/1$ queue with server vacations (see [9, 11] and references in [4]). The papers [2, 5, 10] for the $M/G/1$ models generalizes the single exceptional service to multiple exceptional services. However, the approaches of those papers have some limitations. The paper of [10] assumes the exponential service times, while the other two papers use the Kolmogorov differential equations for stationary analysis, which requires unnecessary regularity conditions such as the existence of densities of the service time distributions. Furthermore, from the results of those papers, it is hard to see the effect of exceptional services. Moreover, closed form formulas for stationary characteristics are obtained only for two exceptional services in [5], and three exceptional services in [2].

The aim of this paper is twofold. First, we show usefulness of a regenerative cycle approach, which is only concerned with a single busy cycle, and uses a well known regenerative cycle formula to derive stationary distributions (e.g., see [3]). This approach well fits to exceptional service mechanism, and neither needs to assume the steady state as in [2] nor to use the Kolmogorov differential equations. The approach has been used for the single exceptional service model (see [11]), but has never been applied to the multiple exceptional service model, as long as the authors know.

Secondly, we study general structures of characteristics of interests such as the mean waiting time, in particular, how they are affected by the exceptional service. This is our major interest. For this, we derive the Laplace transform of the stationary waiting time using some unknown parameters, which can be determined by algorithmic computations. The result is presented as a certain decomposition formula (see Theorem 3.1), which is our main result. The formula enables us to see how the empty probability and the mean waiting time are affected by the distributions of exceptional service times as well as to get some other stationary characteristics such as queue length. As byproduct, the Laplace transform of the stationary waiting time of each customer in a given busy period is obtained. It is not our major interest to derive the closed form formulas, but we demonstrate the case of up to five exceptional services to show how our approach produces those formulas in a systematic way (see Appendix A). Note that those formulas not only extend the known results but also prevent possible errors found in the literature.

This paper is composed of six sections. Section 2 introduces the exceptional service model, and discusses general expressions for the waiting time distributions in transient as well as stationary cases. In this section, the number of exceptional services may be infinite, but the subsequent sections assumes that it is finite. In Section 3, the stationary distribution and its moments are computed for the waiting time, while they are computed for the sojourn time and the queue length in Section 4. The unknown parameters in those formulas are computed in Section 5 and Appendix A. We also give numerical examples in Section 6.

2. Exceptional Service Model and Waiting Time

Consider the $M/G/1$ queue with exceptional service. Denote the mean arrival rate by λ, which is assumed to be positive and finite. We are only concerned with one busy period starting with an empty state, i.e. the state that there is no customer in system, since the empty state is a regeneration epoch for both of the waiting time and queue length process. For an integer $n \geq 1$, let S_n be the service time of the n-th arriving customer in the first busy period. It is assumed that $\{S_n; n = 1, 2, \ldots\}$ is a sequence of independent random variables that are independent of the arrival process. The expectation $E(S_n)$ is assumed to be finite for all n. We denote the distribution of S_n by G_n. If only the first ℓ customers get exceptional service, the service time distribution of the n-th customer for $n \geq \ell+1$ is independent of n, and denoted by G. That is, $G_{\ell+1} = G_{\ell+2} = \ldots = G$. We are mainly concerned with the case of a finite ℓ, but, in this section, the ℓ may be infinite.

Let T_n be interarrival between n-th and $(n+1)$-th arriving customer, and let W_n be the waiting time of the n arriving customer for $n = 1, 2, \ldots$. It is assumed that the first arriving customer finds the system empty. Namely, the system starts with the empty state at time 0. Let N be the total number of customers that arrive in the first busy period. We do not assume that N is finite at this stage, although our main interest is in the case that $E(N) < \infty$.

For $n = 1, 2, \ldots$, define function f_n as

$$f_n(\theta) = E(e^{-\theta W_n}; N \geq n), \qquad \theta \geq 0,$$

where $E(X; A) = E(X 1_A)$ for an event A and the indicator function 1_A. Since $W_{n+1} = W_n + S_n - T_n \geq 0$ on the event that $N \geq n+1$, we have

$$\begin{aligned} f_{n+1}(\theta) &= E(e^{-\theta(W_n+S_n-T_n)}; N \geq n+1) \\ &= E(e^{-\theta(W_n+S_n-T_n)}; W_n + S_n \geq T_n, N \geq n) \\ &= \frac{\lambda}{\lambda-\theta} E(e^{-\theta(W_n+S_n)} - e^{-\lambda(W_n+S_n)}; N \geq n), \end{aligned} \tag{2.1}$$

where the second equality is obtained from the fact that T_n is subject to the exponential distribution with mean $1/\lambda$. For each $n = 1, 2, \ldots$, the Laplace transform of G_n is denoted by g_n, i.e.,

$$g_n(\theta) = E(e^{-\theta S_n}).$$

Since $W_1 = 0$ and W_n and S_n are independent, (2.1) leads to

$$f_1(\theta) = E(e^{-\theta W_1}; N \geq 1) \equiv 1 \tag{2.2}$$

$$f_{n+1}(\theta) = \frac{\lambda}{\lambda-\theta}(f_n(\theta) g_n(\theta) - f_n(\lambda) g_n(\lambda)), \qquad n \geq 1. \tag{2.3}$$

In principle, (2.2) and (2.3) successively determine f_n. Hence, if we can write characteristics of interest in terms of them, our problems are solved. However, the determination of f_n is not easy as it looks. In fact, from (2.2) and (2.3), we have

$$f_n(\theta) = \left(\frac{\lambda}{\lambda-\theta}\right)^{n-1} \prod_{i=1}^{n-1} g_i(\theta) - \sum_{i=1}^{n-1} \left(\frac{\lambda}{\lambda-\theta}\right)^{n-i} f_i(\lambda) g_i(\lambda) \prod_{i=i+1}^{n-1} g_i(\theta), \quad n \geq 1, \tag{2.4}$$

where the empty product is unity, and the empty sum is zero. Thus, $f_n(\theta)$ is determined if $f_i(\lambda)$'s are obtained. But, the evaluation of $f_n(\theta)$ at $\theta = \lambda$ is not easy, because of the singularity in the expression (2.4). We defer this problem to Section 5. Since (2.3) with $\theta = 0$ implies

$$f_n(\lambda) g_n(\lambda) = f_n(0) - f_{n+1}(0), \qquad n \geq 1. \tag{2.5}$$

$f_n(\theta)$ is also determined by $f_i(0)$ for $i = 1, 2, \ldots, n$.

In the rest of this section and in the next section, we shall express various characteristics using $f_n(\theta)$ and $f_n(0)$ in addition to $g_n(\theta)$. We first note that

$$E(N) = \sum_{n=1}^{+\infty} P(N \geq n) = \sum_{n=1}^{+\infty} f_n(0). \tag{2.6}$$

Then, from (2.5), we have

Lemma 2.1 The total number N is finite, i.e., $P(N < +\infty) = 1$, if and only if

$$\sum_{n=1}^{\infty} f_n(\lambda) g_n(\lambda) = 1.$$

Lemma 2.2 The queue is stable, i.e., there exists a stationary waiting time distribution for all customers over all busy periods if and only if $E(N)$ is finite, i.e.

$$E(N) = \sum_{n=1}^{\infty} f_n(0) < \infty. \tag{2.7}$$

If this is the case, Laplace transform ϕ of the stationary waiting time distribution is given by

$$\phi(\theta) = \frac{1}{E(N)} \sum_{n=1}^{+\infty} f_n(\theta). \tag{2.8}$$

PROOF. Since $\{W_n; n = 1, 2, \ldots\}$ is a discrete-time regenerative process, it has the stationary distribution only if each regenerative cycle has a finite mean, i.e., $E(N) < \infty$. (2.8) is a direct consequence of the so called cycle formula, i.e.,

$$\begin{aligned} \phi(\theta) &= \frac{1}{E(N)} E\left(\sum_{n=1}^{N} e^{-\theta W_n}\right) \\ &= \frac{1}{E(N)} E\left(\sum_{n=1}^{\infty} e^{-\theta W_n} 1_{\{N \geq n\}}\right). \end{aligned}$$

□

Another interesting characteristic is the conditional waiting time distribution of the n-th arriving customer, given that he arrived in the first busy period. That is,

$$E(e^{-\theta W_n} | N \geq n) = \frac{f_n(\theta)}{f_n(0)}.$$

3. Stationary Waiting Time for Finite Exceptional Services

In this section, we assume that only the first ℓ customers get exceptional service, and $E(N) < \infty$, where ℓ is finite. It is easy to see that the latter holds if and only if

$$\rho \equiv \lambda E(S) < 1, \tag{3.1}$$

where S is subject to the distribution G. For instance, one can verify this fact by formulating the waiting time process as a random walk with jumps $S_n - T_n$ and reflecting barriers at the origin. Then, the waiting time distribution weakly converges to a probability distribution if and only if $\sum_{i=1}^{n}(S_i - T_i)$ goes to $-\infty$ with probability one as n tends to infinity. The latter is only possible for the case that $E(S_{\ell+1} - T_{\ell+1}) = E(S) - 1/\lambda < 0$.

Our aim of this section is to compute stationary characteristics such as the stationary waiting time distribution, using $f_1, f_2, \ldots, f_\ell$. Even though they are not completely known, informative expressions will be found.

For each $n = 0, 1, \ldots$, let

$$\psi_n(\theta) = \sum_{i=n}^{+\infty} f_i(\theta).$$

Then, from Lemma 2.2, we have

$$\phi(\theta) = \frac{\sum_{n=1}^{+\infty} f_n(\theta)}{E(N)} = \frac{\psi_1(\theta)}{\psi_1(0)}. \tag{3.2}$$

We first compute ψ_1.

Lemma 3.1

$$\psi_1(\theta) = \frac{\theta + \lambda \sum_{n=1}^{\ell} f_n(\theta)(g(\theta) - g_n(\theta))}{\theta - \lambda(1 - g(\theta))}. \tag{3.3}$$

PROOF. Using (2.5), the inductive formula (2.3) can be written as, for $1 \le n \le \ell$,

$$\begin{aligned}(\lambda - \theta) f_{n+1}(\theta) &= \lambda f_n(\theta) g_n(\theta) - \lambda f_n(\lambda) g_n(\lambda), \\ &= \lambda f_n(\theta) g_n(\theta) + \lambda(f_{n+1}(0) - f_n(0)),\end{aligned} \tag{3.4}$$

and, for $n \ge \ell + 1$,

$$(\lambda - \theta) f_{n+1}(\theta) = \lambda f_n(\theta) g(\theta) + \lambda(f_{n+1}(0) - f_n(0)). \tag{3.5}$$

Summing (3.5) over $n = \ell+1, \ell+2, \dots$ yields

$$(\lambda - \theta)\psi_{\ell+2}(\theta) = \lambda \psi_{\ell+1}(\theta) g(\theta) - \lambda f_{\ell+1}(0).$$

Substituting $\psi_{\ell+1}(\theta) = \psi_{\ell+2}(\theta) + f_{\ell+1}(\theta)$ to this formula, we have

$$(\lambda - \theta - \lambda g(\theta))\psi_{\ell+2}(\theta) = \lambda f_{\ell+1}(\theta) g(\theta) - \lambda f_{\ell+1}(0). \tag{3.6}$$

On the other hand, from (3.4), we have, for $n \le \ell$

$$(\lambda - \theta - \lambda g(\theta)) f_{n+1}(\theta) = \lambda(f_n(\theta) g_n(\theta) - f_{n+1}(\theta) g(\theta)) + \lambda(f_{n+1}(0) - f_n(0)). \tag{3.7}$$

Summing (3.6) and (3.7) for $n = 1, 2, \dots, \ell$ and the fact that $f_1(\theta) = 1$ lead to

$$(\lambda - \theta - \lambda g(\theta))(\psi_1(\theta) - 1) = \lambda \sum_{n=2}^{\ell} f_n(\theta)(g_n(\theta) - g(\theta)) + \lambda g_1(\theta) - \lambda.$$

This yields (3.3). □

To get the Laplace transform of the stationary waiting time, we need to compute $\psi_1(0)$. From Lemma 3.1, we have

$$\psi_1(0) = \frac{1 + \sum_{n=1}^{\ell} f_n(0)(\rho_n - \rho)}{1 - \rho} \tag{3.8}$$

Hence, from (3.3) and (3.8), we have

$$E(e^{-\theta W}) = \frac{(1-\rho)\Big(\theta + \lambda \sum_{n=1}^{\ell} f_n(\theta)(g(\theta) - g_n(\theta))\Big)}{\Big(1 + \sum_{n=1}^{\ell} f_n(0)(\rho_n - \rho)\Big)\Big(\theta - \lambda(1 - g(\theta))\Big)}. \tag{3.9}$$

Let $W_{M/G/1}$ be a random variable subject to the stationary waiting time distribution of the standard $M/G/1$ queue that corresponds with the exceptional model, i.e., both models has the same arrival rate and the same distribution for the non-exceptional service times. That is,

$$E(e^{-\theta W_{M/G/1}}) = \frac{(1-\rho)\theta}{\theta - \lambda(1-g(\theta))}.$$

We are now ready to present the result (3.9) in the following form.

Theorem 3.1 Under the stability condition (3.1), the Laplace transform of the stationary waiting time is

$$E(e^{-\theta W}) = E(e^{-\theta W_{M/G/1}}) \frac{1 + \lambda \sum_{n=1}^{\ell} f_n(\theta) \frac{g(\theta) - g_n(\theta)}{\theta}}{1 + \sum_{n=1}^{\ell} f_n(0)(\rho_n - \rho)}. \tag{3.10}$$

Remark 3.1 From (3.7), we have

$$\sum_{n=1}^{\ell+1} f_n(\theta) = \frac{\theta + \lambda(f_{\ell+1}(\theta)g(\theta) - f_{\ell+1}(0)) + \lambda \sum_{n=1}^{\ell} f_n(\theta)(g(\theta) - g_n(\theta))}{\theta - \lambda(1-g(\theta))},$$

for the Laplace transform g of an arbitrary distribution of a positive random variable. Hence, under the stability condition (2.7), equation (3.10) is valid also for $\ell = \infty$, if ρ_n is bound in n. If $g_n(\theta)$ converges to $g(\theta)$ for all θ as n goes to infinity, it is not difficult to see that $\rho \equiv \lim_{n\to\infty} \rho_n < 1$ is sufficient for (2.7). However, (3.10) with $\ell = \infty$ may not be useful since it includes f_n for all n.

Remark 3.2 Equation (3.10) may be interpreted as a decomposition formula similar to the one for the vacation models (see, e.g. [4]). However, it is not the decomposition of a distribution. The busy period of the exceptional service model may be shorter than the corresponding period of the non-exceptional model, so we may not be able to create any server vacation. This is the reason that the second component of the decomposition may not correspond with any distribution. Exceptionally, it indeed corresponds with a distribution if S_n is the sum of independent nonnegative random variables $S^{(n)}$ and V_n such that $S^{(n)}$ has the same distribution as S, because

$$\frac{g(\theta) - g_n(\theta)}{\theta} = \frac{1 - v_n(\theta)}{\theta} g(\theta),$$

where $v_n(\theta)$ is the Laplace transform of V_n. In this case, the server actually takes a vacation for each exceptional service.

Remark 3.3 If we only consider customers who have arrived after the exceptional service periods, we have the conventional decomposition. That is, from (3.6) and (3.7) with $n = \ell$,

$$\frac{\psi_{\ell+1}(\theta)}{\psi_{\ell+1}(0)} = E(e^{-\theta W_{M/G/1}}) \frac{\lambda\big(f_\ell(0) - f_\ell(\theta)g_\ell(\theta)\big)}{\theta\left(1 - \sum_{n=1}^{\ell} f_n(0)(1-\rho_n)\right)}.$$

Remark 3.4 To fully determine $E(e^{-\theta W})$, we need to compute $f_n(\theta)$ and $f_n(0)$. This will be considered in Section 5 and Appendix A. Note that, if their closed forms are obtained, one may numerically invert the Laplace transform using techniques developed in [1].

Using Theorem 3.1, we consider some properties of the exceptional service model.

Corollary 3.1 Under the assumption of Theorem 3.1, the stationary probability that an arriving customer finds the system empty is

$$P(W=0) = \frac{1}{E(N)} = \frac{1-\rho}{1+\sum_{n=1}^{\ell} f_n(0)(\rho_n - \rho)}. \tag{3.11}$$

By PASTA property (see, e.g., [3]), this probability equals the probability that the system is empty in the steady state.

PROOF. The first equality is a direct consequence of (2.8), since (2.3) implies that $f_n(\theta)$ goes to zero for $n \geq 2$ as θ goes to infinity, while the second equality follows from (2.6) and (3.8). □

From the sample path comparisons, it is easy to see that the empty probability is decreased if S_n is stochastically increased, i.e., the tail probability $P(S_n > x)$ is increased for all $x \geq 0$. Corollary 3.1 also implies this, since the denominator of the right-hand side of (3.11) is rewritten as

$$1 - \rho + \sum_{n=1}^{\ell} f_n(0)\rho_n + \sum_{n=\ell+1}^{\infty} f_n(0)\rho,$$

and $f_n(0)$ is increased as S_n is stochastically increased. Corollary 3.1 tells more about the empty probability.

For instance, if $\ell = 1$, then S_1 affects $P(W=0)$ only through its mean. In general, the last exceptional service time S_ℓ affects $P(W=0)$ only through its mean, since (2.4) implies that $f_n(0)$ is determined by g_i for $i \leq n-1$. If $\ell = 2$, the distribution of S_1 does affect $P(W=0)$. This affection is not obvious. For instance, compare the two exceptional service models one of which has an exponential distribution for S_1, and the other has a deterministic distribution with the same mean for S_1. We distinguish characteristics of those models by upper suffixes (M) and (D), respectively, if necessary. Then, since $f_2(0) = 1 - g_1(\lambda)$ from (2.3), we have

$$\begin{aligned} f_2^{(M)}(0) &= 1 - \frac{1}{1+\rho_1} \\ &\leq 1 - e^{-\rho_1} = f_2^{(D)}(0). \end{aligned}$$

Hence, if $E(S_2^{(D)}) = E(S_2^{(M)})$ and if $\rho_2 > (<) \rho$, the deterministic $S_1^{(D)}$ decreases (increases) the empty probability compared with the exponential $S_1^{(M)}$ (respectively). Thus, less random S_1 may decrease the empty probability, contrary to intuition. Since

$$1 - g_1(\lambda) = E(1 - e^{-\lambda S_1}),$$

is the expectation of the increasing concave function of S_1, the same property holds for the two systems if the S_1's are concave ordered (see [14] for the concave order). This may be interpreted that, for the system with $\ell = 2$, less random S_1 effectively increases N, if $\rho_2 > \rho$. For the case of $\ell = 3$, a similar tendency can be found in numerical examples of [2] (see Tables 1 and 2 of the paper). However, this is not always the case (see Appendix B). Thus, we can not expect such a concave order property for the $\ell \geq 3$.

We next consider the moments of W. Let k be an arbitrary positive integer. If $E(S^{n+1})$ and $E(S_k^{n+1})$ for $n = 1, 2, \ldots, \ell$ are finite, then $f_n^{(j)}(0)$ is finite for $n \leq \ell$ and $j \leq k$, where $f_n^{(j)}$ denotes the j-th derivative. This can be seen from the inductive relations,

$$\lambda\left(f_n^{(j)}(0) - f_{n-1}^{(j)}(0)\right) = j f_n^{(j-1)}(0) + \lambda \sum_{i=0}^{j-1} \binom{j}{i} f_{n-1}^{(i)}(0)(-1)^{j-i} E(S_{n-1}^{j-i}), \tag{3.12}$$

which are obtained from (2.3) by differentiating it after multiplying $(\lambda - \theta)$. Hence, taking the k-th order derivative of (3.10) at $\theta = 0$ and noting the fact that

$$(-1)^i \frac{d^i}{d\theta^i} \frac{g(\theta) - g_n(\theta)}{\theta}\bigg|_{\theta=0} = \frac{E(S_n^{i+1}) - E(S^{i+1})}{i+1},$$

we have the following result.

Corollary 3.2 Under the assumption of Theorem 3.1, for an arbitrary positive integer k, if $E(S^{n+1})$ and $E(S_k^{n+1})$ for $n = 1, 2, \ldots, \ell$ are finite, then we have

$$\begin{aligned} E(W^k) &= E(W_{M/G/1}^k) \\ &+ \sum_{j=0}^{k-1} \binom{k}{j} E(W_{M/G/1}^j) \frac{\lambda \sum_{n=1}^{\ell} \sum_{i=0}^{k-j} \binom{k-j}{i} (-1)^{k-j-i} f_n^{(k-j-i)}(0) \frac{E(S_n^{i+1}) - E(S^{i+1})}{i+1}}{1 + \sum_{n=1}^{\ell} f_n(0)(\rho_n - \rho)}. \end{aligned} \tag{3.13}$$

By Corollary 3.2, the moments of the stationary waiting time is evaluated only through $f_n^{(j)}(0)$ for $n \leq \ell$ and $j \leq k$. By (3.12), they are eventually determined by $f_n(0)$ for $n \leq \ell$. We defer computations of $f_n(0)$ to Section 5. We here keep them as unknown parameters. Nevertheless, we can conclude some general properties from Corollary 3.2. We consider them for the mean waiting time, which is given by

$$E(W) = \frac{\lambda E(S^2)}{2(1-\rho)} + \frac{\sum_{n=1}^{\ell} \left((-f_n'(0))(\rho_n - \rho) + \frac{\lambda}{2} f_n(0)(E(S_n^2) - E(S^2))\right)}{1 + \sum_{n=1}^{\ell} f_n(0)(\rho_n - \rho)}, \tag{3.14}$$

where $-f_n'(0)$ is nonnegative, and calculated as

$$-f_n'(0) = \frac{1}{\lambda} \sum_{i=2}^{n} \left(f_{i-1}(0)\rho_{i-1} - f_i(0)\right), \qquad n = 1, 2, \ldots, \ell. \tag{3.15}$$

Corollary 3.3 Under the assumption of Theorem 3.1, the following properties hold.

(i) Suppose $E(S_n^2)$ for $n = 1, 2, \ldots, \ell - 1$ are finite. If $E(S)$, $E(S_n)$ for $n = 1, 2, \ldots, \ell$ and $E(S_n^2)$ for $n = 1, 2, \ldots, \ell - 1$ are fixed, $E(W)$ is a linear and increasing function of $E(S^2)$ and $E(S_\ell^2)$. In particular, $E(W)$ is finite only if $E(S^2)$ and $E(S_\ell^2)$ are finite.

(ii) The mean $E(W)$ is finite if and only if $E(S^2)$ and $E(S_n^2)$ for $n = 1, 2, \ldots, \ell$ are finite.
(iii) Under the condition in (ii), if $E(S) \leq (\geq) E(S_\ell)$ for $n = 1, 2, \ldots, \ell$ and if $E(S^2) \leq (\geq)$ $E(S_n^2)$ for $n = 1, 2, \ldots, \ell$, then $E(W)$ is not less (greater) than the mean waiting time of the corresponding standard $M/G/1$ queue, i.e., the first term in the right hand side of (3.14).

PROOF. From (2.5), $f_n(0)$ is determined by g_i for $i \leq n-1$. Hence, $E(W)$ of (3.14) is a function of ρ, ρ_n for $n = 1, 2, \ldots, \ell$, $E(S^2)$, $E(S_n^2)$ for $n = 1, 2, \ldots, \ell$, and g_n for $n = 1, 2, \ldots, \ell - 1$. Furthermore, the coefficients of $E(S^2)$ are summed up to

$$\frac{\lambda}{2(1-\rho)} + \frac{-\sum_{n=1}^{\ell} \frac{\lambda}{2} f_n(0)}{1 + \sum_{n=1}^{\ell} f_n(0)(\rho_n - \rho)} = \frac{\lambda\Big(1 - \sum_{n=1}^{\ell}(1-\rho_n) f_n(0)\Big)}{2(1-\rho)\Big(1 + \sum_{n=1}^{\ell} f_n(0)(\rho_n - \rho)\Big)} > 0,$$

since $\sum_{n=1}^{\ell} f_n(0) \leq 1$. Hence, the first part of (i) is obtained from (3.14), if $E(S^2)$ and $E(S_\ell^2)$ are finite. If either one of these second moments is infinite, we truncate S and S_ℓ by $K > 0$, i.e., they are replaced by $S^{(K)} = \min(K, S)$ and $S_\ell^{(K)} = \min(K, S_\ell)$, respectively. Since each sample of W_n is a nondecreasing function of the service times, the truncated system has stochastically not greater waiting times. Hence, for a random variable $W^{(K)}$ subject to the stationary waiting time of the truncated system, we have

$$E(W^{(K)}) \leq E(W).$$

Since $E(W^{(K)})$ is an increasing function of $E((S^{(K)})^2)$ and $E((S_\ell^{(K)})^2)$ and either one of $E(S^2)$ and $E(S_\ell^2)$ is infinite, letting K tend to infinite yields that $E(W) = \infty$. This concludes the second part of (i). (ii) is obviously obtained from (i) and (3.14), since $0 < f_n(0) < 1$ for $n = 1, 2, \ldots, \ell$. (iii) is also a direct consequence of (3.14). □

The property (iii) also follows from the rate conservation law (e.g., see [12]),

$$E(W) = \lambda\Big(E(WS) + \frac{1}{2}E(S^2)\Big),$$

since W_n and S_n are independent. So (iii) just verifies the compatibility of Corollary 3.3.

4. Sojourn Time and Queue Length

In this section, we consider the sojourn time of a customer and the queue length in the steady state, where the queue length is meant the total number of customers including a customer being served. We assume the finiteness of the number of the exceptional services and the stability condition $\rho < 1$ as in the previous section.

Let U_n be the sojourn time of the n-th arriving customer, and let U be a random variable subject to the stationary sojourn time distribution. Since S_n is independent of W_n and the event $\{N \geq n\}$, we get

$$E[e^{-\theta U_n}; N \geq n] = f_n(\theta) g_n(\theta).$$

Hence, similar to Lemma 2.2, we have

$$E[e^{-\theta U}] = \frac{1}{\psi_1(0)} \sum_{n=1}^{\infty} f_n(\theta) g_n(\theta).$$

Since (3.4) and (3.5) imply

$$(\lambda-\theta)\psi_1(\theta)+\theta=\lambda\sum_{n=1}^{\infty}f_n(\theta)g_n(\theta),$$

we have

$$\begin{aligned}E[e^{-\theta U}] &= \frac{1}{\lambda\psi_1(0)}(\theta+(\lambda-\theta)\psi_1(\theta)) \\ &= \frac{(1-\rho)\Big(\theta g(\theta)+(\lambda-\theta)\sum_{n=1}^{\ell}f_n(\theta)(g(\theta)-g_n(\theta))\Big)}{\Big(1+\sum_{n=1}^{\ell}f_n(0)(\rho_n-\rho)\Big)\Big(\theta-\lambda(1-g(\theta))\Big)}. \end{aligned} \tag{4.1}$$

Using $W_{M/G/1}$ in the corresponding $M/G/1$ queue, (4.1) is written as the following form.

Theorem 4.1 Under the assumptions of Theorem 3.1, we have

$$E[e^{-\theta U}]=E[e^{-\theta W_{M/G/1}}]\frac{g(\theta)+(\lambda-\theta)\sum_{n=1}^{\ell}f_n(\theta)\dfrac{g(\theta)-g_n(\theta)}{\theta}}{1+\sum_{n=1}^{\ell}f_n(0)(\rho_n-\rho)}. \tag{4.2}$$

Let L_n^+ be the queue length just after the n-th customer completes his service. Then, L_n^+ equals the number of customers who arrived during the n-th customer being in the system. Hence, we have

$$\begin{aligned}E[z^{L_n^+};N\geq n] &= \sum_{i=0}^{\infty}E[e^{-\lambda(W_n+S_n)}\frac{(\lambda(W_n+S_n))^i}{i!}z^k;N\geq n] \\ &= E[e^{-\lambda(W_n+S_n)(1-z)};N\geq n] \\ &= f_n(\lambda(1-z))g_n(\lambda(1-z)). \end{aligned} \tag{4.3}$$

Thus we can calculate the condition generating function $E[z^{L_n^+}|N\geq n]$ through f_n.

Let L be a random variable subject to the stationary queue length distribution at an arbitrary time instant, and let $Q_{M/G/1}$ be the queue length not including a customer being served in the corresponding $M/G/1$ queue, similar to $W_{M/G/1}$. That is,

$$E[z^{Q_{M/G/1}}]=\frac{(1-\rho)(1-z)}{g(\lambda(1-z))-z}.$$

Then, from (4.2) and the distributional Little's law, i.e.,

$$\begin{aligned}E(z^L) &= E(e^{-\lambda(1-z)U}), \\ E(z^{Q_{M/G/1}}) &= E(e^{-\lambda(1-z)W_{M/G/1}}),\end{aligned}$$

we have following result.

Corollary 4.1 Under the assumptions of Theorem 3.1, we have

$$E[z^L] = E(z^{Q_{M/G/1}}) \frac{g(\lambda(1-z)) + z\sum_{n=1}^{\ell} f_n(\lambda(1-z)) \dfrac{g(\lambda(1-z)) - g_n(\lambda(1-z))}{1-z}}{1+\sum_{n=1}^{\ell} f_n(0)(\rho_n - \rho)}. \tag{4.4}$$

For the moments of the stationary sojourn time and the stationary queue length, we can get similar formulas to those in Corollary 3.2 from Theorems 4.1 and 4.1. However, they are tedious and the resulted formulas are not tractable, so we here only give the first moment of the queue length.

Corollary 4.2 Under the same assumptions of Corollary 3.2 for $k = 1$,

$$E[L] = \frac{\lambda^2 E(S^2)}{2(1-\rho)} + \frac{\rho + \sum_{n=1}^{\ell} \left((\rho_n - \rho)\left(1 - \sum_{i=1}^{n-1} f_i(0)(1-\rho_i)\right) + \frac{\lambda^2}{2} f_n(0)(E(S_n^2) - E(S^2))\right)}{1+\sum_{n=1}^{\ell} f_n(0)(\rho_n - \rho)}. \tag{4.5}$$

PROOF. We differentiate (4.4) at $z = 1$. Then, we have (4.5) using the fact that

$$\lambda f_n'(0) - f_n(0) = -1 + \sum_{i=1}^{n-1} f_i(0)(1-\rho_i), \qquad n = 1, 2, \ldots, \ell,$$

which is obtained from (3.15). □

5. Computations of $f_n(\lambda)$'s

As we discussed, the computations of $f_n(0)$'s or equivalently those of $f_n(\lambda)$'s are essential to determine $f_n(\theta)$ as well as to compute the means of the waiting time and the queue length. We here give an algorithm to compute $f_n(\lambda)$'s. Then, $f_n(0)$'s are computed by

$$f_n(0) = 1 - \sum_{i=1}^{n-1} f_i(\lambda) g_i(\lambda), \qquad n \geq 2, \tag{5.1}$$

which is obtained from (2.5). Since these computations are irrelevant to the number of exceptional services ℓ, we mainly use g_n instead of g for $n \geq \ell$ as in Section 2.

In the following computations, we need to evaluate derivatives of Laplace transforms at λ. To this end, we introduce operator notation for differentiations. Let $h(\theta)$ be the Laplace transform of a distribution on $[0, \infty)$. For nonnegative integer n, define the operator D_λ^n and $\tilde{D}_\lambda^n$ as

$$\begin{aligned} D_\lambda^n h &= h^{(n)}(\theta)|_{\theta=\lambda}, \\ \tilde{D}_\lambda^n h &= \frac{(-\lambda)^n}{n!} h^{(n)}(\theta)|_{\theta=\lambda}. \end{aligned}$$

The following result recursively determines $f_n(\lambda)'s$.

Lemma 5.1 For $n = 0, 1, \ldots$, we have

$$f_{n+1}(\lambda) = \tilde{D}_\lambda^n\Big(\prod_{i=1}^n g_i\Big) - \sum_{j=1}^{n-1} f_j(\lambda)g_j(\lambda)\tilde{D}_\lambda^{n-j+1}\Big(\prod_{i=j+1}^n g_i\Big). \tag{5.2}$$

PROOF. Multiplying both sides of (2.4) for $n+1$ instead of n by $(\lambda-\theta)^n$ yields

$$(\lambda-\theta)^n f_{n+1}(\theta) = \lambda^n \prod_{i=1}^n g_i(\theta) - \sum_{j=1}^n f_j(\lambda)g_j(\lambda)(\lambda-\theta)^{j-1}\lambda^{n+1-j}\prod_{i=j+1}^n g_i(\theta). \tag{5.3}$$

Since, for $n \geq 1$ and $k \leq n$,

$$D_\lambda^n(\lambda-\theta)^k = \begin{cases} (-1)^k k!, & k = n, \\ 0, & k \neq n, \end{cases}$$

applying D_λ^n to both sides of (5.3) leads to

$$\begin{aligned}(-1)^n n! f_{n+1}(\lambda) &= \lambda^n D_\lambda^n\Big(\prod_{i=1}^n g_i\Big) \\ &\quad - \sum_{j=1}^{n-1} f_j(\lambda)g_j(\lambda)\binom{n}{j-1}(-1)^{j-1}(j-1)!\lambda^{n-j+1}D_\lambda^{n-j+1}\Big(\prod_{i=j+1}^n g_i\Big),\end{aligned}$$

where $n = j$ is dropped in the summation, since the empty product is unity and its derivative is zero. Dividing both sides of this equation by $(-1)^n n!$ yields (5.2). ∎

Remark 5.1 Note that, for a Laplace transform h of a nonnegative random variable X,

$$\tilde{D}_\lambda^n h = \frac{(-\lambda)^n}{n!}D_\lambda^n h = E\Big(\frac{(\lambda X)^n}{n!}e^{-\lambda X}\Big) \leq 1.$$

Let $Y_{j,n} = S_j + S_{j+1} + \cdots + S_n$. Then, $Y_{j,n}$ has the Laplace transform $\prod_{i=j}^n g_i(\theta)$, so (5.2) can be written as

$$f_{n+1}(\lambda) = E\Big(\frac{(\lambda Y_{1,n})^n}{n!}e^{-\lambda Y_{1,n}}\Big) - \sum_{j=1}^{n-1} f_j(\lambda)g_j(\lambda)E\Big(\frac{(\lambda Y_{j+1,n})^{n-j+1}}{(n-j+1)!}e^{-\lambda Y_{j+1,n}}\Big).$$

Thus, the recursive computation by (5.2) is expected to be numerically stable, since $f_j(\lambda)$'s are all positive (also see Lemma 2.1).

Remark 5.2 To get algebraic formulas for $f_n(\lambda)$ in terms of the derivatives of g_n at λ, (5.2) is not so convenient. To this end, we rewrite it in the following form by multiplying both sides of (5.2) with $n!/(-\lambda)^n$.

$$\frac{n!}{(-\lambda)^n}f_{n+1}(\lambda) = D_\lambda^n\Big(\prod_{i=1}^n g_i\Big) - \sum_{j=1}^{n-1}\binom{n}{j-1}\frac{(j-1)!}{(-\lambda)^{j-1}}f_j(\lambda)g_j(\lambda)D_\lambda^{n-j+1}\Big(\prod_{i=j+1}^n g_i\Big). \tag{5.4}$$

Then, we can recursively calculate $n!(-\lambda)^n f_{n+1}(\lambda)$ using either $g_n^{(j)}$ or D_λ^n.

Using (5.4), we exemplify the first six terms of $f_n(\lambda)$ below, where $D^0_\lambda g_j$ denotes $g_j(\lambda)$.

$$\begin{aligned}
f_1(\lambda) &= 1, \\
f_2(\lambda) &= -\lambda D^1_\lambda g_1 = -\lambda g_1^{(1)}(\lambda), \\
f_3(\lambda) &= \frac{(-\lambda)^2}{2!}\left(D^2_\lambda(g_1g_2) - D^0_\lambda g_1 D^2_\lambda g_2\right) \\
&= \frac{(-\lambda)^2}{2}\left(g_1^{(2)}(\lambda)g_2(\lambda) + 2g_1^{(1)}(\lambda)g_2^{(1)}(\lambda)\right), \\
f_4(\lambda) &= \frac{(-\lambda)^3}{3!}\left(D^3_\lambda(g_1g_2g_3) - D^0_\lambda g_1 D^3_\lambda(g_2g_3) - 3D^1_\lambda(g_1)D^0_\lambda g_2 D^2_\lambda g_3\right) \\
&= \frac{(-\lambda)^3}{3!}\left(g_1^{(3)}(\lambda)g_2(\lambda)g_3(\lambda) + 3g_1^{(2)}(\lambda)(g_2^{(1)}(\lambda)g_3(\lambda) + g_2(\lambda)g_3^{(1)}(\lambda))\right. \\
&\qquad \left. +3g_1^{(1)}(\lambda)(g_2^{(2)}(\lambda)g_3(\lambda) + 2g_2^{(1)}(\lambda)g_3^{(1)}(\lambda))\right), \\
f_5(\lambda) &= \frac{(-\lambda)^4}{4!}\left(D^4_\lambda(g_1g_2g_3g_4) - D^0_\lambda g_1 D^4_\lambda(g_2g_3g_4) - 4D^1_\lambda g_1 D^0_\lambda g_2 D^3_\lambda(g_3g_4)\right. \\
&\qquad \left. -6(D^2_\lambda(g_1g_2) - D^0_\lambda g_1 D^2_\lambda g_2)D^0_\lambda g_3 D^2_\lambda g_4\right), \\
f_6(\lambda) &= \frac{(-\lambda)^5}{5!}\left(D^5_\lambda(g_1g_2g_3g_4g_5) - D^0_\lambda g_1 D^5_\lambda(g_2g_3g_4g_5) - 5D^1_\lambda g_1 D^0_\lambda g_2 D^4_\lambda(g_3g_4g_5)\right. \\
&\qquad -10\left(D^2_\lambda(g_1g_2) - D^0_\lambda g_1 D^2_\lambda g_2\right)D^0_\lambda g_3 D^3_\lambda(g_4g_5) \\
&\qquad \left. -10\left(D^3_\lambda(g_1g_2g_3) - D^0_\lambda g_1 D^3_\lambda(g_2g_3) - 3D^1_\lambda g_1 D^0_\lambda g_2 D^2_\lambda g_3\right)D^0_\lambda g_4 D^2_\lambda g_5\right).
\end{aligned}$$

Here, $f_2(\lambda)$, $f_3(\lambda)$ and $f_4(\lambda)$ are given also in terms of $g_i^{(n)}(\lambda)$ for convenience. These computation indicates how much $f_n(\lambda)$ become complicated as n increases. Thus, it seems not so meaningful to present closed form formulas for them except for small n. See Appendix A for the corresponding formulas of $f_n(\theta)$, $f_n(0)$ and $f_n'(0)$ for $n = 1, 2, \ldots, 5$. These formulas together with (4.1) and (4.4) give closed form formulas for the case of ℓ exceptional services for $\ell \leq 5$.

Although $f_n(\lambda)$ and therefore $f_n(0)$ are complicated for large n, it is expected that $f_n(\theta)/f_n(0)$ converges as n goes to infinity, if the number of exceptional services ℓ is finite. This limit corresponds with the so called quasi-stationary distribution on $(0, \infty)$, which is independent of ℓ. The following result is known for the standard $M/G/1$ queue, but it is intuitively clear that they also hold for the exceptional service model, as long as the ℓ is finite.

Suppose that the queue is stable, i.e. $\rho < 1$, and there is a $\theta_0 > 0$ such that $g(-\theta_0) < \infty$. Then, from Theorem 2.3 of [6] (see also [8]), we have

$$\lim_{n\to\infty} \frac{f_n(\theta)}{f_n(0)} = \frac{\lambda(1-\gamma)}{(\theta-\lambda)\gamma + \lambda g(\theta)}, \tag{5.5}$$

where $\gamma = \lambda g(\tau)/(\lambda - \tau) < 1$, and τ is a unique solution of the equation of $\theta < 0$,

$$g'(\theta) + \frac{g(\theta)}{\lambda - \theta} = 0.$$

That is, $E(e^{-(S-T)\theta}) = g(\theta)\lambda/(\lambda - \theta)$ attains a minimum value at $\theta = \tau$. For instance, if the service time distribution is the exponential with mean $1/\mu$, we have $\tau = (\lambda - \mu)/2$, so we get

$$\gamma = \frac{4\rho}{(1+\rho)^2}. \tag{5.6}$$

Furthermore, the asymptotic behavior of $f_n(0)$ is also obtained in [6], i.e.,

$$\lim_{n\to\infty} \frac{f_n(0)}{n^{-3/2}\gamma^n} = c, \tag{5.7}$$

where c is a constant determined by the distribution of S (see Theorem 2.2 of [6]). Note that these results are only useful for f_n with sufficient large n compared with the ℓ.

6. Numerical Examples

In the computations of $f_n(\lambda)$'s through (5.2), we need to evaluate $D_\lambda^n(\prod_{i=1}^n g_i)$'s. It is possible to give analytic expression for an arbitrary n, but the expression as well as its numerical evaluation rapidly becomes complicated as n is increased. So, we evaluate their numerical values inductively. We first note the following equation.

$$\begin{aligned}\tilde{D}_\lambda^n\left(\prod_{i=1}^k g_i\right) &= \frac{(-\lambda)^n}{n!}\sum_{j=0}^n \binom{n}{j} D_\lambda^j\left(\prod_{i=1}^{k-1} g_i\right) D_\lambda^{n-j} g_k \\ &= \sum_{j=0}^n \tilde{D}_\lambda^j\left(\prod_{i=1}^{k-1} g_i\right)\frac{(-\lambda)^{n-j}}{(n-j)!} g_k^{(n-j)}(\lambda), \qquad 1 \le k \le n.\end{aligned}$$

Hence, if $g_j^{(i)}(\lambda)$ is evaluated for $0 \le i \le j$ and $1 \le j \le n$, $\tilde{D}_\lambda^n(\prod_{i=1}^j g_i)$ is evaluated inductively. Thus, if we know the Laplace transform g_j of the j-th service time S_j, we can numerically compute $f_n(\lambda)$'s. For instance, if S_j's are all exponentially distributed with mean $1/\mu_j$'s, we have

$$\frac{(-\lambda)^i}{(i)!} g_j^{(i)}(\lambda) = \frac{\lambda^i \mu_j}{(\mu_j+\lambda)^{i+1}} = \frac{\rho_j^i}{(1+\rho_j)^{i+1}},$$

where $\rho_j = \lambda/\mu_j$.

Using the above algorithm, for the case of the exponentially distributed service times, we compute $E(W)$ for $\ell = 1,2,3,4$, and $E(W_n|N \ge n)$ for $\ell = 4$ and $n \le 5$. For the latter, we also compute the limiting value of $E(W_n|N \ge n)$ when n goes to infinity, which is denoted by $E(W_\infty|N = \infty)$. From (5.5), this is computed as

$$E(W_\infty|N=\infty) = \frac{\gamma-\rho}{\lambda(1-\gamma)},$$

if the total number ℓ of exceptional services is finite. On the other hand, from (3.15), $E(W_n|N \ge n)$ is computed as

$$\begin{aligned}E(W_n|N \ge n) &= \frac{-f_n'(0)}{f_n(0)} \\ &= \frac{1}{\lambda f_n(0)}\sum_{i=2}^n \big(f_{i-1}(0)\rho_{i-1} - f_i(0)\big).\end{aligned}$$

In Tables 1 and 2, the expected values of the conditional waiting times are presented for the first $n \le 15$ customers when all the service times are exponentially distributed. In

Tables 3 to 6, the expected values of the stationary waiting times are presented for the case that S is exponentially distributed. In Tables 3 and 5, all S_i's are exponentially distributed, while, in Tables 4 and 6, all S_i's are deterministic. In all the tables, case 1 assumes that

$$E(S) = 1, \qquad E(S_i) = \frac{10}{15 - i}, \quad i = 1, 2, 3, 4,$$

while case 2 assumes that

$$E(S) = 1, \qquad E(S_i) = \frac{10}{5 + i}, \quad i = 1, 2, 3, 4.$$

Note that $\lambda = \rho$, since $E(S) = 1$.

From those tables, we can see that the variability of S_i's as well as their means increases the mean waiting times, as it is expected. One interesting feature is that the mean stationary waiting times are maximized for $\ell = 3$ in Table 6. This shows a trade off between the means and variability of S_i.

Table 1: $E(W_n|N \geq n)$ for the exponential S_i with $\ell = 4$: case 1

ρ	$n=1$	$n=2$	$n=3$	$n=10$	$n=15$	$+\infty$
0.3	0	0.714286	1.150183	2.968466	3.387836	4.714285
0.6	0	0.714286	1.180486	3.549099	4.302877	9.000000
0.9	0	0.714286	1.205739	4.113080	5.280517	39.00000

Table 2: $E(W_n|N \geq n)$ for the exponential S_i with $\ell = 4$: case 2

ρ	$n=1$	$n=2$	$n=3$	$n=10$	$n=15$	$+\infty$
0.3	0	1.666667	2.470238	3.571587	3.777754	4.714285
0.6	0	1.666667	2.568922	4.432692	4.952068	9.000000
0.9	0	1.666667	2.640693	5.275218	6.237180	39.00000

Table 3: $E(W)$ for the exponential S_i with $\ell \leq 4$: case 1

ρ	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
0.3	0.428571	0.267857	0.232045	0.221719	0.219091
0.6	1.500000	1.145320	0.995524	0.924230	0.896147
0.9	9.000000	8.406593	8.028617	7.787066	7.668081

Table 4: $E(W)$ for the deterministic S_i with $\ell \leq 4$: case 1

ρ	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
0.3	0.428571	0.184152	0.129821	0.114861	0.110378
0.6	1.500000	0.960591	0.714844	0.593581	0.533846
0.9	9.000000	8.097527	7.432807	6.960174	6.661003

Table 5: $E(W)$ for the exponential S_i with $\ell \leq 4$: case 2

ρ	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
0.3	0.428571	0.873016	0.998905	1.042953	1.055296
0.6	1.500000	2.261905	2.542279	2.673005	2.722213
0.9	9.000000	10.00000	10.39113	10.59174	10.67566

Table 6: $E(W)$ for the deterministic S_i with $\ell \leq 4$: case 2

ρ	$\ell=0$	$\ell=1$	$\ell=2$	$\ell=3$	$\ell=4$
0.3	0.428571	0.525794	0.560295	0.564263	0.558788
0.6	1.500000	1.666667	1.755184	1.772154	1.743789
0.9	9.000000	9.218750	9.354898	9.388998	9.338285

Acknowledgements

The decomposition in Remark 3.3 was suggested by Tetsuya Takine of Kyoto University. We are grateful to his comments. We also thanks for anonymous referees for helpful comments and suggesting some related references.

References

[1] J. Abate, G. L. Choudhury and Ward Whitt: An introduction to numerical transform inversion and its application to probability models. In Grassmann (ed.): *Computational Probability* (Kluwer Academic Publishers, New York, 1999).

[2] Y. Baba: On $M/G/1$ Queues with the first N customers of each busy period receiving exceptional services. *Journal of Operations Research Society of Japan*, **42** (1999) 490-500.

[3] F. Baccelli and P. Bremaud: *Elements of Queueing Theory* (Springer-Verlag, Berlin, 1991).

[4] B. T. Doshi: Queueing Systems with vacations – A survey. *Queueing Systems*, **6** (1986) 29-66.

[5] N. Igaki, U. Sumita and M. Kowada: On a generalized $M/G/1$ queue with service degradation/enforcement. *Journal of Operations Research Society of Japan*, **41** (1998) 415-429.

[6] D. L. Iglehart: Random walks with negative drift conditioned to stay positive. *Journal of Applied Probability*, **11** (1974) 742-751.

[7] J. Keilson and L. D. Servi: , A distributional form of Little's law. *Operations Research Letters*, **7** (1998) 223-227.

[8] E. K. Kyprianou: On the quasi-stationary distribution of the virtual waiting time in queues with Poisson arrivals. *Journal of Applied Probability*, **11** (1974) 742-751.

[9] A. J. Lemoine: Limit theorems for generalized single server queues: the exceptional system. *SIAM Journal on Applied Mathematics*, **28** (1975) 596-606.

[10] H. Li, Y. Zhu, P. Yang and S. Madhavapedy: On $M/M/1$ queues with a smart machine. *Queueing Systems*, **24** (1996) 23-36.

[11] D. L. Minh: Analysis of the exceptional queueing system by the use of regenerative processes and analytical methods. *Mathematics of Operations Research*, **5** (1980) 147-159.

[12] M. Miyazawa: Rate conservation laws: a survey. *Queueing Systems*, **15** (1994) 1-58.
[13] N. U. Prabhu: *Foundations to Queueing Theory* (Kluwer Academic Publishers, Boston, 1997).
[14] D. Stoyan: *Comparison Methods for Queues and Other Stochastic Models* (English edition, John Wiley & Sons, New York, 1983).
[15] P. D. Welch: On a generalized $M/G/1$ queueing process in which the first customer of each busy period receives exceptional service, *Operations Research*, **12** (1964) 736-752.

Appendix A

We calculate $f_n(\theta)$, $f_n(0)$ and $f_n'(0)$ for $n = 1, 2, \ldots, 5$. From $f_n(\lambda)$ of Section 5 and (2.4),

$$\begin{aligned}
f_1(\theta) &= 1, \\
f_2(\theta) &= \frac{\lambda}{\lambda-\theta}\big(g_1(\theta) - g_1(\lambda)\big), \\
f_3(\theta) &= \left(\frac{\lambda}{\lambda-\theta}\right)^2 \big(g_1(\theta)g_2(\theta) - g_1(\lambda)g_2(\theta) + (\lambda-\theta)g_1^{(1)}(\lambda)g_2(\lambda)\big), \\
f_4(\theta) &= \left(\frac{\lambda}{\lambda-\theta}\right)^3 \Big(g_1(\theta)g_2(\theta)g_3(\theta) - g_1(\lambda)g_2(\theta)g_3(\theta) + (\lambda-\theta)g_1^{(1)}(\lambda)g_2(\lambda)g_3(\theta) \\
&\qquad -\frac{(\lambda-\theta)^2}{2}\big(g_1^{(2)}(\lambda)g_2(\lambda) + 2g_1^{(1)}(\lambda)g_2^{(1)}(\lambda)\big)g_3(\lambda)\Big), \\
f_5(\theta) &= \left(\frac{\lambda}{\lambda-\theta}\right)^4 \Big(g_1(\theta)g_2(\theta)g_3(\theta)g_4(\theta) - g_1(\lambda)g_2(\theta)g_3(\theta)g_4(\theta) \\
&\qquad +(\lambda-\theta)g_1^{(1)}(\lambda)g_2(\lambda)g_3(\theta)g_4(\theta) \\
&\qquad -\frac{(\lambda-\theta)^2}{2}\big(g_1^{(2)}(\lambda)g_2(\lambda) + 2g_1^{(1)}(\lambda)g_2^{(1)}(\lambda)\big)g_3(\lambda)g_4(\theta) \\
&\qquad +\frac{(\lambda-\theta)^3}{6}\big(g_1^{(3)}(\lambda)g_2(\lambda)g_3(\lambda) + 3g_1^{(2)}(\lambda)(g_2^{(1)}(\lambda)g_3(\lambda) + g_2(\lambda)g_3^{(1)}(\lambda)) \\
&\qquad\qquad +3g_1^{(1)}(\lambda)(g_2^{(2)}(\lambda)g_3(\lambda) + 2g_2^{(1)}(\lambda)g_3^{(1)}(\lambda))\big)g_4(\lambda)\Big).
\end{aligned}$$

Substituting $\theta = 0$ in the above formulas or directly from (5.1), we have

$$\begin{aligned}
f_1(0) &= 1, \\
f_2(0) &= 1 - g_1(\lambda), \\
f_3(0) &= 1 - g_1(\lambda) + \lambda g_1^{(1)}(\lambda)g_2(\lambda), \\
f_4(0) &= 1 - g_1(\lambda) + \lambda g_1^{(1)}(\lambda)g_2(\lambda) - \frac{\lambda^2}{2}\big(g_1^{(2)}(\lambda)g_2(\lambda) + 2g_1^{(1)}(\lambda)g_2^{(1)}(\lambda)\big)g_3(\lambda), \\
f_5(0) &= 1 - g_1(\lambda) + \lambda g_1^{(1)}(\lambda)g_2(\lambda) - \frac{\lambda^2}{2}\big(g_1^{(2)}(\lambda)g_2(\lambda) + 2g_1^{(1)}(\lambda)g_2^{(1)}(\lambda)\big)g_3(\lambda) \\
&\quad +\frac{\lambda^3}{6}\big(g_1^{(3)}(\lambda)g_2(\lambda)g_3(\lambda) + 3g_1^{(2)}(\lambda)(g_2^{(1)}(\lambda)g_3(\lambda) + g_2(\lambda)g_3^{(1)}(\lambda)) \\
&\qquad +3g_1^{(1)}(\lambda)(g_2^{(2)}(\lambda)g_3(\lambda) + 2g_2^{(1)}(\lambda)g_3^{(1)}(\lambda))\big)g_4(\lambda).
\end{aligned}$$

These $f_n(\theta)$ and $f_n(0)$ explicitly determine the Laplace transforms of the stationary distributions of the waiting time, the sojourn time, and the queue length through (3.9), (4.1) and

(4.4), respectively, for $\ell \leq 5$. The results agrees with those in [15] for $\ell = 1$ and in [2] for $\ell = 2$. They also correct some errors in the corresponding formula for $\ell = 3$ in [2], in which apparently λ is dropped at $B_0^{(1)}(\lambda)$ of [2] which is g_1 in our notation. We here systematically derive the Laplace transforms.

Finally, we calculate $f_n'(0)$ for $n = 1, 2, \ldots, 5$, using (3.15).

$$
\begin{aligned}
\lambda f_1'(0) &= 0, \\
\lambda f_2'(0) &= 1 - g_1(\lambda) - \rho_1, \\
\lambda f_3'(0) &= 1 - g_1(\lambda) - \rho_1 + (1 - g_1(\lambda))(1 - \rho_2) + \lambda g_1^{(1)}(\lambda) g_2(\lambda) \\
&= (2 - \rho_2)(1 - g_1(\lambda)) - \rho_1 + \lambda g_1^{(1)}(\lambda) g_2(\lambda), \\
\lambda f_4'(0) &= (2 - \rho_2)(1 - g_1(\lambda)) - \rho_1 + \lambda g_1^{(1)}(\lambda) g_2(\lambda) + \big(1 - g_1(\lambda) + \lambda g_1^{(1)}(\lambda) g_2(\lambda)\big)(1 - \rho_3) \\
&\quad - \frac{\lambda^2}{2}\big(g_1^{(2)}(\lambda) g_2(\lambda) + 2 g_1^{(1)}(\lambda) g_2^{(1)}(\lambda)\big) g_3(\lambda) \\
&= -\rho_1 + (3 - \rho_2 - \rho_3)(1 - g_1(\lambda)) + \lambda(2 - \rho_3) g_1^{(1)}(\lambda) g_2(\lambda) \\
&\quad - \frac{\lambda^2}{2}\big(g_1^{(2)}(\lambda) g_2(\lambda) + 2 g_1^{(1)}(\lambda) g_2^{(1)}(\lambda)\big) g_3(\lambda), \\
\lambda f_5'(0) &= -\rho_1 + (3 - \rho_2 - \rho_3)(1 - g_1(\lambda)) + \lambda(2 - \rho_3) g_1^{(1)}(\lambda) g_2(\lambda) \\
&\quad - \frac{\lambda^2}{2}\big(g_1^{(2)}(\lambda) g_2(\lambda) + 2 g_1^{(1)}(\lambda) g_2^{(1)}(\lambda)\big) g_3(\lambda) \\
&\quad + (1 - \rho_4)\Big(1 - g_1(\lambda) + \lambda g_1^{(1)}(\lambda) g_2(\lambda) \\
&\qquad\qquad - \frac{\lambda^2}{2}\big(g_1^{(2)}(\lambda) g_2(\lambda) + 2 g_1^{(1)}(\lambda) g_2^{(1)}(\lambda)\big) g_3(\lambda)\Big) \\
&\quad + \frac{\lambda^3}{6}\Big(g_1^{(3)}(\lambda) g_2(\lambda) g_3(\lambda) + 3 g_1^{(2)}(\lambda)(g_2^{(1)}(\lambda) g_3(\lambda) + g_2(\lambda) g_3^{(1)}(\lambda)) \\
&\qquad\qquad + 3 g_1^{(1)}(\lambda)(g_2^{(2)}(\lambda) g_3(\lambda) + 2 g_2^{(1)}(\lambda) g_3^{(1)}(\lambda))\Big) g_4(\lambda) \\
&= -\rho_1 + (4 - \rho_2 - \rho_3 - \rho_4)(1 - g_1(\lambda)) + \lambda(3 - \rho_3 - \rho_4) g_1^{(1)}(\lambda) g_2(\lambda) \\
&\quad - \frac{\lambda^2}{2}(2 - \rho_4)\big(g_1^{(2)}(\lambda) g_2(\lambda) + 2 g_1^{(1)}(\lambda) g_2^{(1)}(\lambda)\big) g_3(\lambda) \\
&\quad + \frac{\lambda^3}{6}\Big(g_1^{(3)}(\lambda) g_2(\lambda) g_3(\lambda) + 3 g_1^{(2)}(\lambda)(g_2^{(1)}(\lambda) g_3(\lambda) + g_2(\lambda) g_3^{(1)}(\lambda)) \\
&\qquad\qquad + 3 g_1^{(1)}(\lambda)(g_2^{(2)}(\lambda) g_3(\lambda) + 2 g_2^{(1)}(\lambda) g_3^{(1)}(\lambda))\Big) g_4(\lambda).
\end{aligned}
$$

These together with $f_n(0)$'s yield closed form formulas for the means waiting time.

Appendix B

We compare the stationary empty probability $P(W = 0)$ for the two exceptional service models with $\ell = 3$, one of which exceptional service times are exponentially distributed, while those of the other model are deterministic. The other model assumptions are the same for both models. As in Section 3, we distinguish their characteristics by super script (M) and (D), respectively, if necessary. We assume that $\rho < 1$ and $E(S_n^{(M)}) = E(S_n^{(D)})$ for $n = 1, 2, 3$.

Since $P(W = 0) = 1/E(N)$, we compare $(1 - \rho)E(N)$. Form Corollary 3.1 and $f_n(0)$ in

Appendix A,

$$\begin{aligned}(1-\rho)E(N) &= 1+\sum_{n=1}^{3} f_n(0)(\rho_n-\rho)\\ &= 1+(\rho_1-\rho)+(\rho_2-\rho)\big(1-g_1(\lambda)\big)\\ &\quad +(\rho_3-\rho)\big(1-g_1(\lambda)+\lambda g_1^{(1)}(\lambda)g_2(\lambda)\big).\end{aligned}$$

Hence, we have

$$\begin{aligned}(1-\rho)\big(E(N^{(M)})-E(N^{(D)})\big) &= (\rho_2+\rho_3-2\rho)\big(e^{-\rho_1}-\frac{1}{1+\rho_1}\big)\\ &\quad +(\rho_3-\rho)\rho_1\big(e^{-\rho_1}e^{-\rho_2}-\frac{1}{(1+\rho_1)^2}\frac{1}{1+\rho_2}\big). \quad \text{(B.1)}\end{aligned}$$

Since, for small $x>0$,

$$e^{-x}=\frac{1}{1+x+\frac{1}{2}x^2+o(x^2)},$$

where $o(x)$ denote the small order, i.e., $o(x)/x$ goes to 0 as x goes to 0, we have

$$\begin{aligned}e^{-x} &= \frac{1}{1+x}\big(1-\frac{1}{2}x^2+o(x^2)\big),\\ e^{-x} &= \frac{1}{(1+x)^2}\big(1+x-\frac{1}{2}x^2+o(x^2)\big).\end{aligned}$$

Substituting these expressions into (B.1) yields

$$\begin{aligned}(1-\rho)\big(E(N^{(M)})-E(N^{(D)})\big) &= (\rho_2+\rho_3-2\rho)\frac{1}{1+\rho_1}\big(-\frac{1}{2}\rho_1^2+o(\rho_1^2)\big)\\ &\quad +(\rho_3-\rho)\frac{\rho_1}{(1+\rho_1)^2}\frac{1}{1+\rho_2}\Big(\big(1+\rho_1-\frac{1}{2}\rho_1^2+o(\rho_1^2)\big)\big(1-\frac{1}{2}\rho_2^2+o(\rho_2^2)\big)-1\Big).\end{aligned}$$

Thus, if we choose sufficiently small ρ_1 and ρ_2 in such a way that $\rho_2>\rho$ and $\rho_3>\rho$, this equation is positive, so we have

$$P(W^{(M)}=0)=\frac{1}{E(N^{(M)})}<\frac{1}{E(N^{(D)})}=P(W^{(D)}=0).$$

This is contrary to the case of $\ell=2$ discussed in Section 3. However, if ρ_1 and ρ_2 become large, for instance, if $\rho_1=\rho_2=2$, then the order is reversed. Thus, $P(W=0)$ can not be ordered by the concave orders of S_1 and S_2 for the $\ell=3$.

Masakiyo Miyazawa
Department of Information Sciences
Science University of Tokyo
Noda, Chiba 278-8510, Japan
E-mail: miyazawa@is.noda.sut.ac.jp