J. Appl. Prob. **23**, 443–458 (1986)
Printed in Israel

APPROXIMATION OF THE QUEUE-LENGTH DISTRIBUTION OF AN $M/GI/s$ QUEUE BY THE BASIC EQUATIONS

MASAKIYO MIYAZAWA,* *Science University of Tokyo*

Abstract

We give a unified way of obtaining approximation formulas for the steady-state distribution of the queue length in the $M/GI/s$ queue. The approximations of Hokstad (1978) and Case A of Tijms et al. (1981) are derived again. The main interest of this paper is in considering the theoretical meaning of the assumptions given in the literature. Having done this, we derive new approximation formulas. Our discussion is based on one version of the steady-state equations, called the basic equations in this paper. The basic equations are derived for $M/GI/s/k$ with finite and infinite k. Similar approximations are possible for $M/GI/s/k$ $(k < +\infty)$.

APPROXIMATION FORMULA; QUEUE LENGTH; STEADY-STATE DISTRIBUTION; POISSON ARRIVAL

1. Introduction

This paper deals with approximations for the steady-state distribution of the queue length in the $M/GI/s$ queue. The same method can be used for $M/GI/s/k$, i.e. $M/GI/s$ with k waiting places, and so this case is also discussed.

Much has been written on the approximation formulas of $M/GI/s$. We can classify the methods of obtaining these formulas into two main types. One is to obtain approximations by assuming a parametric formula, or the formula itself, beforehand. In the former case, these parameters are chosen to satisfy suitable conditions. The merit of this method is its simplicity. For example, Takahashi (1977) and Boxma et al. (1980) obtained useful formulas for the mean waiting time and the mean queue length by this method. The other type of method is to get approximation formulas by solving exact or approximated equations using some additional assumptions. This method is particularly useful when we discuss approximations of distributions such as the queue length, since it is very difficult to infer their approximation formulas intuitively. In fact, most approximations

Received 15 January 1985; revision received 10 June 1985.
* Postal address: Department of Information Sciences, Faculty of Science and Technology, Science University of Tokyo, Noda City, Chiba 278, Japan.

for queue-length distributions have been obtained by this method. For example, Tijms et al. (1981) obtained an excellent formula (cf. also Hokstad (1978), Stoyan (1976) and Kimura (1983)).

The purpose of this paper is to give theoretical insight into the latter type of approximations to the queue-length distribution. In particular, we consider why certain assumptions are necessary. This makes it possible for us to give a unified view of the approximation formulas. For example, we can provide interesting interpretations of the results by Nozaki and Ross (1978) Case B (equivalent to Hokstad (1978)) and Case A of Tijms et al. (1981)). We also give new approximation formulas, one of which is of a similar kind to the Tijms et al. approximation. All approximations of the queue-length distribution are given in terms of generating functions.

Our discussion is based on equations which hold exactly in the steady state, called the *basic equations* in this paper. In Section 2, we are concerned with the derivation of these basic equations, and, from them, we derive the parametric formula which holds exactly for $M/GI/s/k$; it contains $M/GI/s$ as a special case, $k = +\infty$. This formula, which is for the queue-length distribution, is the key point of our discussion. We also present some asymptotic results in light and heavy traffic. In Section 3, we examine the meaning of the assumptions often used to obtain the approximations, and using these we can obtain several approximations for $M/GI/s$ in a unified way. In the final section, we briefly discuss the approximation for $M/GI/s/k$.

2. Derivation of the basic equations for the steady-state distribution

In this section, both $M/GI/s$ and $M/GI/s/k$ queues are considered. $M/GI/s$ is an s-server queue with a stationary Poisson arrival process and i.i.d. service times which are independent of the arrival process. For convenience, we number the servers of the queue from 1 to s, and assume that an arriving customer chooses one of the idle servers with equal probability when fewer than s customers are in the system. This assumption has no effect on the distribution of the queue length, and ensures that the residual service times of s servers are symmetrically distributed, which is desirable for our analysis. For $M/GI/s/k$, it is assumed that arriving customers are rejected when k customers are waiting for service. In our model, those rejected customers are counted as immediately departing customers.

Now we introduce some notation which is valid for $M/GI/s$ and $M/GI/s/k$. Let λ be the mean arrival rate of customers. Let S, $G(x)$ and $\tilde{G}(\theta)$ denote the service time, its distribution and its Laplace transform, respectively. That is, $\tilde{G}(\theta) = \int \exp(-\theta x)G(dx)$. We always assume that ES (the mean of S) is finite. For $M/GI/s$, it is assumed that the traffic intensity $\rho = \lambda ES/s$ is less than 1, which implies the existence of the steady state. For $M/GI/s/k$, its existence is ensured under no additional condition. At time t in the steady state, let $l(t)$ and $r_i(t)$ be the number of customers in the system and the residual service time of the ith server, respectively, for $i = 1, \cdots, s$, where $r_i(t) = 0$ when the ith server is idle at time t. Also, $q(t)$ denotes the queue length at time t. From the existence of the steady state, there exists a probability measure P such that $\{(l(t), r_1(t), \cdots, r_s(t))\}$ is a stationary process with respect to P.

To derive the basic equations, we need some conditional distributions. Let P_0 (P_1) be the conditional distribution of P under the condition that a customer arrives at (leaves) the system at time 0. These distributions are called Palm distributions in the theory of point processes; see Miyazawa (1983) for precise definitions. We use the following notation:

$E, E_0, E_1, \cdots$ are expectations with respect to P, P_0, P_1, respectively,

$$l = l(0), \quad l^+ = l(0+), \quad l^- = l(0-),$$

$$q = q(0) \qquad (= (l-s)^+),$$

$$r_i = r_i(0), \quad r_i^+ = r_i(0+), \quad r_i^- = r_i(0-) \qquad (i = 1, \cdots, s),$$

$$p_n = P(l = n), \quad p_n^0 = P_0(l^- = n), \quad p_n^* = P_1(l^+ = n) \qquad (n = 0, 1, \cdots),$$

$$G_e(x) = \int_0^x (1 - G(u))du/ES, \quad \tilde{G}_e(\theta) = \int \exp(-\theta x) G_e(dx),$$

$$\phi_j(\theta) = E(\exp(-\theta r) \mid r > 0, l = j) \qquad (j = 1, 2, \cdots),$$

$$\phi_j^0(\theta) = E_0(\exp(-\theta r) \mid r^- > 0, l^- = j) \qquad (j = 1, 2, \cdots),$$

$$\phi_j^*(\theta) = E_1(\exp(-\theta r) \mid r^- > 0, l^+ = j) \qquad (j = 1, 2, \cdots),$$

where $r_1, \cdots, r_s$ are identically distributed and so their suffixes and those of the $r_i^{\pm}$'s are omitted when this causes no confusion. Note that G_e is known as the stationary residual distribution of G. $\phi_j(\theta)$ is the Laplace transform of the residual service time of a busy server at an arbitrary time t under the condition that $l(t) = j$. $\phi_j^0(\theta)$ and $\phi_j^*(\theta)$ have similar meanings at the arrival and departure epochs of customers, respectively.

Now we consider $M/GI/s$ and $M/GI/s/k$ simultaneously. In this case, the notation $M/GI/s/k$ with $k = +\infty$ for $M/GI/s$ is convenient. Note that $p_n = p_n^0 = p_n^*$ for any $n = 0, 1, \cdots, s + k$ in $M/GI/s/k$. Our main concern is the distribution of q, and so we wish to obtain equations involving the p_n's. The following lemma is used for this; it is an immediate consequence of Corollary 3.1 of Miyazawa (1983).

Lemma 2.1. Let h be a non-negative differentiable function on R^{s+1} (($s+1$)-dimensional real space) and define

$$X(t) = h(l(t), r_1(t), \cdots, r_s(t)) \qquad (t \in R).$$

Then, if $E_0X(0-)$ and $E_1X(0-)$ are finite, we have

$$\text{(2.1)} \qquad EX'(0) = \lambda(E_0X(0-) - E_0X(0+)) + \lambda(E_1X(0-) - E_1X(0+)),$$

where $X'(0)$ denotes the derivative of $X(t)$ at $t = 0$, which exists a.s. P.

This lemma shows the equality of the intensities of state changes of the process $\{X(t)\}$, and so it is called the *intensity conservation law*. Now let, for $j = 1, 2, \cdots, s + k$,

$$X_j(t) = I_{\{l(t)=j\}} \sum_{i=1}^{s} \exp(-\theta r_i(t)) I_{\{r_i(t)>0\}},$$

where I_A is an indicator function of a set A. We note the following equations. For $j = 1, 2, \cdots$,

$$EX_j(0) = \min(s, j)\theta\phi_j(\theta)p_j,$$

$$E_0X_j(0-) = \min(s, j)\phi_j^0(\theta)p_j^0,$$

$$E_1X_j(0-) = ((\min(s, j) - 1)\phi_j^*(\theta) + 1)p_{j-1}^*,$$

$$E_0X_j(0+) = (\tilde{G}(\theta)I_{\{j<s+1\}} + \min(s, j-1)\phi_{j-1}^0(\theta))p_{j-1}^0,$$

$$E_1X_j(0+) = (\tilde{G}(\theta)I_{\{j>s-1\}} + \min(s-1, j)\phi_j^*(\theta))p_j^*.$$

By substituting these values for the corresponding terms of (2.1) we have the next result, since $p_n^0 = p_n$ and $\phi_n^0(\theta) = \phi_n(\theta)$ for any $n > 0$ in M/GI-type queues and $E_0X_{s+k}(0-) = E_1X_{s+k}(0+)$.

Theorem 2.1. In $M/GI/s/k$ ($k < \infty$ or $k = +\infty$), for $j = 1, 2, \cdots, s + k - 1$,

$$\begin{aligned} &\theta \min(s, j)\phi_j(\theta)p_j \\ \text{(2.2)} \quad &= \lambda[\{(\min(s, j) - 1)\phi_{j-1}^*(\theta) + 1 - (\tilde{G}(\theta)I_{\{j<s+1\}} + \min(s, j-1)\phi_{j-1}(\theta))\}p_{j-1} \\ &\qquad - \{\tilde{G}(\theta)I_{\{j>s-1\}} + \min(s-1, j)\phi_j^*(\theta) - \min(s, j)\phi_j(\theta)\}p_j], \end{aligned}$$

and, if $k < +\infty$, then, for $j = s + k$,

$$\text{(2.3)} \qquad \theta s\phi_j(\theta)p_j = \lambda[1 + (s-1)\phi_{j-1}^*(\theta) - \min(s, j-1)\phi_{j-1}(\theta) - \tilde{G}(\theta)I_{\{k=0\}}]p_{j-1}.$$

The equations (2.2) and (2.3) are called the *basic equations* in this paper. Define

$$\psi(\theta, x) = \sum_{j=s}^{s+k} \phi_j(\theta)p_j x^{(j-s)}.$$

That is, $\psi(\theta, x) = E(\exp(-\theta r)x^{l-s}; l > s-1)$. By multiplying X^{j-s} by (2.2), (2.3) and summing them up for $j > s - 1$, we have

$$(2.4)\qquad \begin{aligned}(\theta s-\lambda(1-x))\psi(\theta,x)&=\lambda(1-\tilde{G}(\theta))p_{s-1}+\lambda(x-\tilde{G}(\theta))\psi(0,x)\\&\quad+\lambda(L(\theta)+M(\theta,x)+N(\theta,x)),\end{aligned}$$

where we put $p_{s+k}=0$ for $k=+\infty$, and

$$L(\theta)=(s-1)(\phi^*_{s-1}(\theta)-\phi_{s-1}(\theta))p_{s-1}$$

$$M(\theta,x)=(s-1)(1-x)\sum_{j=s}^{s+k-1}(\phi_j(\theta)-\phi^*_j(\theta))x^{j-s}p_j$$

$$N(\theta,x)=(\tilde{G}(\theta)-\phi_{s+k}(\theta)+x(\phi_{s+k}(\theta)-1))x^k p_{s+k}.$$

Define $\theta(x)=\lambda(1-x)/s$ and let $\theta=\theta(x)$ in (2.4). Then we have the following expression for $\psi(0,x)$, which is equivalent to the generating function of q.

Corollary 2.1. In $M/GI/s/k$, we have, for any x $(0<x<1)$,

$$(2.5)\qquad \psi(0,x)=\frac{(1-\tilde{G}(\theta(x)))p_{s-1}+L(\theta(x))+M(\theta(x),x)+N(\theta(x),x)}{(\tilde{G}(\theta(x))-x)}.$$

The expression (2.5) involves many unknown quantities, but it has very interesting properties. For example, the denominator of the right-hand side of (2.5) vanishes at $x=\eta$, where η $(\neq 1)$ is a positive solution of the equation

$$(2.6)\qquad \frac{\lambda}{\lambda-s\theta(\eta)}\tilde{G}(\theta(\eta))=1.$$

On the other hand, η is known as the decreasing rate of p_n, that is, $p_n\sim K\eta^{-n}$ as n tends to ∞ for some constant K, in $PH/PH/s$ queues (cf. Theorem 5.1 of Takahashi (1981)).

Finally, we present some lemmas for the asymptotic properties of p_j's in light and heavy traffic.

Lemma 2.2 (Burman and Smith (1983)). In $M/PH/s$,

$$(2.7)\qquad \lim_{\lambda\downarrow 0}p_j/\lambda p_{j-1}=ES/j\qquad (j=1,2,\cdots,s),$$

$$(2.8)\qquad \lim_{\lambda\downarrow 0}p_{s+1}/\lambda p_s=\gamma,$$

where

$$\gamma=\int_0^{+\infty}(1-G_e(u))^s du.$$

Let $W(z)$ and $Q(z)$ be the Laplace transforms of W and q, respectively, where W denotes the waiting time in the steady state. Then, the next equation for $M/GI/s$ is well known (cf. Haji and Newell (1971)):

$$Q(z) = W(\lambda(1 - \exp(-z))). \tag{2.9}$$

From this and the heavy-traffic result of Köllerström (1974), we can get the next lemma.

Lemma 2.3. For $M/GI/s$, if the service-time distribution is fixed and if $E(S^3)$ is finite, we have, for $z > 0$,

$$\lim_{\rho \uparrow 1} \psi(0, \exp(-(1-\rho)z)) = \lim_{\rho \uparrow 1} Q((1-\rho)z) = \frac{1}{1+(1+\delta^2)z/2}, \tag{2.10}$$

where $\delta^2 = \mathrm{Var}(S)/(E^2 S)$.

It is desirable that approximation formulas satisfy these asymptotic properties, so the above lemmas are helpful in reaching good approximations.

3. Approximation formulas for $M/GI/s$

In this section, we are concerned only with $M/GI/s$. First we examine why some assumptions are needed in the approximations for the queue-length distribution, i.e., for $\psi(0, x)$. The following assumptions are typical in this paper. For each $i \geqq 1$,

$$\phi_i(\theta) = \phi_i^*(\theta) \quad \text{for any } \theta > 0, \tag{A.i}$$

$$\frac{d}{d\theta}(\phi_i(\theta) - \phi_i^*(\theta))\Big|_{\theta=0} = 0, \tag{B.i}$$

where the derivative in (B.i) is the right derivative. (B.i) implies the coincidence of the expectations if they exist, while (A.i) means that of the distributions. The meaning of these assumptions is considered. Next we derive various approximation formulas for $\psi(0, x)$ from (2.5) with the basic equations (Theorem 2.1) and additional assumptions. Note that $N(\theta, x) = 0$ for $M/GI/s$, and so we can determine $\psi(0, x)$ if we give p_{s-1}, $L(\theta)$ and $M(\theta, x)$.

3.1. *Some fundamental properties of the assumptions.* It is known that $p_1, \cdots, p_{s-1}$ of $M/M/s$ give good approximations for $M/GI/s$ (cf. Tijms et al. (1981)). Lemma 2.2 confirms it, at least for light traffic. We show how weak assumptions have helped in obtaining these results. Let $p_j(\exp)$ ($j = 0, 1, \cdots, s-1$) be those probabilities of $M/M/s$. That is, $p_j(\exp) = (\lambda ES)^j p_0(\exp)/j!$ for $j = 1, \cdots, s-1$ and

$$p_0(\exp) = \left[\sum_{j=0}^{s-1} (\lambda ES)^j/j! + (\lambda ES)^s/((1-\rho)s!)^{-1}\right].$$

Proposition 3.1. In $M/GI/s$, (B.i) holds for $i = 1, \cdots, s-1$ if and only if

$$p_j = p_j(\exp) \qquad (j = 0, 1, \cdots, s-1). \tag{3.1}$$

Proof. Let, for $j = 1, 2, \cdots, s$,

$$h_j(\theta) = (\phi_j^*(\theta) - \phi_j(\theta))/\theta,$$

and let $h_0(\theta) = 0$. Then, from Theorem 2.1, we have, for $j = 1, 2, \cdots, s-1$,

$$(j\phi_j(\theta) + \lambda h_j(\theta))p_j = \lambda((1 - \tilde{G}(\theta))/\theta + h_{j-1}(\theta))p_{j-1}. \tag{3.2}$$

Assume (B.i) for $i = 1, \cdots, s$. Then, by letting $\theta \downarrow 0$ in (3.2), we have that $p_j = (\lambda ES)^j p_0/j!$ for $j = 1, 2, \cdots, s-1$. Hence, from the well-known relation $s(1-\rho)P(l > s-1) = \Sigma_{j=0}^{s-1}(s\rho - j)p_j$, we obtain that $P(l > s-1) = (\lambda ES)^s p_0/((1-\rho)s!)$. Now p_0 is determined by $\Sigma_{j=0}^{+\infty} p_j = 1$, and we get (3.1). Conversely, assume (3.1). Let θ tend to 0 in (3.2) for $j = 1$; we then have (B.i) since $h_0(\theta) = 0$. In a similar way, we obtain (B.i) for $i = 2, 3, \cdots, s-1$ inductively.

The assumption that ϕ_j (or ϕ_j^*) $= \tilde{G}_e$ for some set of j's has often been used in the literature. For example, in Case A of Tijms et al. (1981), to obtain the $M/M/s$ approximation for $p_1, \cdots, p_{s-1}$ it is assumed that the residual service times at departure epochs are i.i.d. having the distribution G_e when fewer than s customers are in the system. The stronger assumption was used in Nozaki and Ross (1978). Proposition 3.1 shows that those assumptions can be weakened considerably: no independence assumptions are needed. Concerning G_e, the next result explains the appearance of this distribution.

Proposition 3.2. In $M/GI/s$, (A.i) holds for $i = 1, \cdots, s-1$ if and only if $\phi_j(\theta) = \tilde{G}_e(\theta)$ for $j = 1, \cdots, s-1$ and $p_j = p_j(\exp)$ for $j = 0, 1, \cdots, s-1$.

Proof. Suppose (A.i) for $i = 1, \cdots, s$. Then, from Proposition 3.1, we have that $p_j = p_j(\exp)$ for $j = 0, 1, \cdots, s-1$. Thus, the proposition follows from (3.2) in the proof of Proposition 3.1. The converse is also implied by (3.2).

Remark. For $j \geqq s$, $\phi_j(\theta) = \phi_j^*(\theta)$ does not imply $\phi_j(\theta) = \tilde{G}_e(\theta)$ in general. For example, if, for some j_0 and any $j > j_0$, $\phi_j(\theta) = \phi_j^*(\theta)$, then, from (2.2),

$$\lim_{j\to\infty} \phi_j(\theta) = \frac{\lambda(\eta - \tilde{G}(\theta))}{\lambda(\eta - 1) + s\theta},$$

where η is given by (2.6). Note that this is exact in $M/PH/s$, as shown by Takahashi (1981). Those facts show that the assumption that $\phi_j(\theta) = \tilde{G}_e(\theta)$ does not hold for large j.

The following results show what we can say about the asymptotic behaviour of the probabilities in light traffic by the basic equations.

Proposition 3.3.

(i) (B.$s-1$) and (2.2) for $j = s-1$ imply that

(3.3) $$\lim_{\lambda \downarrow 0} p_s / \lambda p_{s-1} = ES/s.$$

(ii) (2.2) for $j = s$ implies that

(3.4) $$\lim_{\lambda \downarrow 0} [p_{s+1}/\rho p_s + (s\phi_s'(0) - (s-1)\phi_s^{*\prime}(0))/ES] = 0,$$

where the existence of $\phi_j'(0)$ and $\phi_j^{*\prime}(0)$ are assumed for $j = s-1, s, s+1$ and for sufficiently small $\lambda > 0$ in all cases.

Proof. From (2.2), for $j = s$ and $s+1$, we have

(3.5) $$\frac{p_j}{\lambda p_{j-1}} = \frac{(1 - K_j(\theta)) + (s-1)(\phi_{j-1}^*(\theta) - \phi_{j-1}(\theta))}{(\theta s - \lambda)\phi_j(\theta) + \lambda \tilde{G}(\theta) + \lambda(s-1)(\phi_j^*(\theta) - \phi_j(\theta))},$$

where $K_j(\theta) = \tilde{G}(\theta)$ for $j = s$ and $\phi_s(\theta)$ for $j = s+1$. Divide the numerator and denominator of the right-hand side of (3.5) by θ and let it tend to 0. Then the right-hand side of (3.5) is expressed by $\tilde{G}'(0)$, $\phi_j'(0)$, and $\phi_j^{*\prime}(0)$ $(j = s-1, s, s+1)$. Hence, by letting λ tend to 0, we have (3.3) and (3.4).

Part (i) of this proposition shows compatibility with the exact result in Lemma 2.2; Part (ii) is used in the following subsections. Finally, we give the heavy-traffic result.

Proposition 3.4. For any $j \geqq s$, (B.i) for $i = 1, \cdots, s$ and $i \geqq j$ imply that (2.5) satisfies (2.10).

Proof. From (2.5) and Proposition 3.1, this proposition follows if $L(\theta(x))$ and $M(\theta(x), x)$ are $o(x)$. (B.$s-1$) implies that $L(\theta(x)) = o(x)$, and (B.i) for $i \geqq j$ implies $M(\theta(x), x) = o(x)$. Thus we obtain the proposition.

3.2. *Approximations of Hokstad–Tijms and Tijms et al.*

From now on, we indicate characteristics by (app), e.g. $\psi(0, x)$ (app), when we use the additional assumptions to obtain them. In this section, we rederive some known approximations in a unified way. In our all approximations, we assume (B.i) for $i = 1, 2, \cdots, s-1$. From Proposition 3.2 and the remark following it, we can assume (B.i) for $i \geqq s-1$, which implies $L(\theta) + M(\theta, x) = 0$. Thus, from (2.5), we can get a very simple approximation.

Theorem 3.1. In $M/GI/s$, the assumptions $L(\theta(x)) + M(\theta(x), x) = 0$ and (B.i) for $i = 1, \cdots, s-1$ imply the Hokstad–Tijms approximation, i.e.,

(3.6) $$\psi(0, x)(\text{app}) = \frac{1 - \tilde{G}(\theta(x))}{\tilde{G}(\theta(x)) - x} p_{s-1}(\exp) \qquad (|x| < 1).$$

This approximation was first obtained by Hokstad (1978) and later, in a different way, by Tijms et al. (1981). We call it the Hokstad–Tijms approximation. By Proposition 3.4, (3.6) is compatible with the heavy-traffic approximation. For light traffic, it is easily seen that (2.9) holds, but we have that

$$\lim_{\lambda \downarrow 0} p_{s+1}(\text{app})/\rho p_s(\text{app}) = E(S^2)/2(ES)^2, \tag{3.7}$$

so it does not satisfy (2.10). This is one reason why the approximation does not hold good for light traffic. Boxma et al. (1980) also checked the same property as (3.7). From (3.6), we can easily get an approximation for the mean queue length and so on. The formula for the mean has been rederived by many authors. For example, Nozaki and Ross (1978) had given definite conditions for obtaining it; the assumptions of Theorem 3.1 above are much weaker than theirs.

Tijms et al. (1981) pointed out that the mean queue length given by (3.6) is an over (or under)estimate according to whether δ is larger or smaller than 1. They proposed another approximation.

We now consider the conditions for obtaining Case A of Tijms et al. (1981), which we call the Tijms et al. approximation. To give their approximation in our formulation, we introduce some notation. Define $V = s \min(U_1, \cdots, U_s)$, where $U_1, \cdots, U_s$ are i.i.d. having the distribution G_e. We denote the Laplace transform of V by $V(\theta)$. Then, from (3.25) of Van Hoorn (1983), the Tijms et al. approximation is given by (3.1) and

$$\psi(0, x)(\text{app}) = \frac{\rho(1-x)V(\theta(x))}{\tilde{G}(\theta(x)) - x} p_{s-1}(\exp). \tag{3.8}$$

We can see that (3.8) satisfies the asymptotic properties of Lemmas 2.2 and 2.3. In fact, numerical tests show that it much improves the Hokstad–Tijms approximation, particularly in light traffic. Tijms et al. (1981) used the following assumptions.

(T.1) For $i = 1, \cdots, s-1$, at the epoch when $l^+ = i$ and a customer has just departed, the residual service times at busy servers are independent and have the same distribution G_e.

(T.2) For $i \geqq s$, at the epoch when $l^+ = i$ and a customer has just departed, the residual time to the next departure epoch has the common distribution $G(sx)$.

By applying the assumption (T.1) to our basic equations (2.2), we can get (3.1) and (B.i) for $i = 1, \cdots, s-1$ as in Propositions 3.1 and 3.2. Hence (T.1) implies $L(\theta) = 0$. Thus, comparing (3.8) with (2.5), we see that

$$M(\theta(x), x)(\text{app}) = \rho(1-x)(V(\theta(x)) - \tilde{G}_e(\theta(x)))p_{s-1}(\exp) \tag{3.9}$$

implies (3.8). It seems to be difficult to get (3.9) from (T.2) by using our basic equations, but, we can give one interpretation of it. Suppose (A.i) for $i \geqq s+1$, and

$$(3.10) \quad (s-1)(\phi_s(\theta)(\text{app})-\phi_s^*(\theta)(\text{app}))p_s(\text{app})=(V(\theta)-\tilde{G}_e(\theta))\rho p_{s-1}(\text{exp}).$$

Then (3.9) is obtained. Hence we get the following theorem.

Theorem 3.2. In $M/GI/s$, the assumptions (B.i) for $i=1,\cdots,s-2$, (A.$s-1$), (A.j) for $j \geqq s+1$ and (3.10) imply the Tijms et al. approximation (3.8).

Remark. (A.$s-1$) is equivalent to $L(\theta)=0$ for any $\theta>0$, and the assumptions (A.j) for $j \geqq s+1$ can be weakened such that

$$(3.11) \quad \sum_{j=s+1}^{+\infty}(\phi_j(\theta(x))-\phi_j^*(\theta(x)))x^{j-s}p_j=0.$$

From Theorems 3.1 and 3.2, we can say that the Tijms et al. approximation attempts to weaken the assumption (A.s) in $L(\theta(x))+M(\theta(x),x)=0$. In the next section, we try similar techniques and try to weaken it for both (A.$s-1$) and (A.s).

3.3. *New approximation formulas.* We derive new approximation formulas from (2.5). Define

$$D(\theta)=s\phi_s(\theta)-(s-1)\phi_s^*(\theta).$$

Then we have the next theorem.

Theorem 3.3. In $M/GI/s$, under the assumptions (B.i) for $i=1,\cdots,s-1$ and (A.j) for $j \geqq s+1$, if $D(\theta)(\text{app})$ is given, then we have, for any x $(x<1)$,

$$(3.12) \quad \psi(0,x)(\text{app})=\frac{\tilde{G}(\theta(x))-xD(\theta(x))(\text{app})}{(\tilde{G}(\theta(x))-x)}p_s(\text{app}),$$

where the existence of $D'(0)(\text{app})$ is assumed and $p_s(\text{app})$ is given by

$$(3.13) \quad p_s(\text{app})=\frac{\rho}{1-\rho-\lambda D'(0)(\text{app})/s}p_{s-1}(\text{exp}).$$

Proof. In this proof, we omit (app) and (exp) for simplicity. Firstly we show (3.13). From (2.2) of Theorem 2.1 (or (3.5)), we have

$$(3.14) \quad \begin{aligned}&\lambda(1-G(\theta))p_{s-1}+\lambda L(\theta)\\&=[(\theta s-\lambda)\phi_s(\theta)+\lambda\tilde{G}(\theta)+\lambda(s-1)(\phi_s^*(\theta)-\phi_s(\theta))]p_s.\end{aligned}$$

Hence we get

$$\theta s \phi_s(\theta) p_s = \lambda[(D(\theta) - \tilde{G}(\theta))p_s + (1 - \tilde{G}(\theta))p_{s-1} + L(\theta)]. \tag{3.15}$$

Dividing both sides of (3.15) by θ and letting θ tend to 0, we obtain (3.13). On the other hand, from (3.13) and (3.15), we have

$$\begin{aligned}
&(1-x)(s-1)(\phi_s(\theta(x)) - \phi_s^*(\theta(x)))p_s \\
&\quad = (1-x)(D(\theta(x)) - \phi_s(\theta(x)))p_s \\
&\quad = [(1-x)D(\theta(x)) - \theta(x)s\phi_s(\theta(x))/\lambda]p_s \\
&\quad = [\tilde{G}(\theta(x)) - xD(\theta(x))]p_s - (1 - \tilde{G}(\theta(x)))p_{s-1} - L(\theta(x)).
\end{aligned}$$

Substituting this result into (2.5), we get (3.12) by the assumption (A.j) for $j \geqq s+1$.

Remark. As seen from the proof of Theorem 3.3, (3.13) is exact if the true p_s, $D'(0)$ and p_{s-1} are used instead of p_s(app), $D'(0)$(app) and p_{s-1}(exp), respectively. From this, (2.8) and (3.4), we can get the following second-order limiting result,

$$\lim_{\lambda \downarrow 0} \frac{\rho p_{s-1} - p_s}{\lambda p_s} = \gamma - ES/s. \tag{3.16}$$

Since the assumption (A.$s-1$) and (A.s) is removed to get (3.12), we can expect that it gives a good approximation if $D(\theta)$(app) is properly chosen. Now let us consider $D(\theta)$. From (ii) of Proposition 3.3, (2.8) holds if and only if

$$\lim_{\lambda \downarrow 0} D'(0)(\text{app}) = -s\gamma. \tag{3.17}$$

This information can be used to determine $D(\theta)$(app), but it is not enough. Further information can be obtained from (3.14). That is, by the relation

$$\lambda D(\theta)p_s = (\theta s \phi_s(\theta) + \lambda \tilde{G}(\theta))p_s - \lambda((1 - \tilde{G}(\theta))p_{s-1} + L(\theta)), \tag{3.18}$$

$D(\theta)$ can be determined from $\phi_s(\theta)$, $L(\theta)$ and p_s, where p_s can be replaced by $D'(0)$. We consider three cases.

Case 1. We assume that (B.$s-1$) holds, $\phi_s(\theta)(\text{app}) = \tilde{G}_e(\theta)$ and $\phi_s(\lambda)(\text{app}) = \phi_s^*(\lambda)(\text{app})$. From these, we have

$$p_s(\text{app}) = \frac{(1 - \tilde{G}(\lambda))}{(s-1)\tilde{G}_e(\lambda) + \tilde{G}(\lambda)} p_{s-1}(\text{exp}). \tag{3.19}$$

Case 2. This case is an improvement of Case 1 for the light traffic. We replace the third assumption of Case 1 by $-D'(0) = s\gamma$.

Case 3. We assume that

$$(3.20)\qquad D(\theta)(\mathrm{app}) = \begin{cases} (1-\rho)V(\theta)+\rho\tilde{G}_e(\theta) & \text{for } \delta_3 \leqq 1 \\ V(\theta) & \text{for } \delta_3 > 1, \end{cases}$$

where $\delta_3 = E((S-ES)^3)/E^3S$.

In Case 1, the third assumption may be artificial; it is selected from among the assumptions that $\phi_s(\theta)(\mathrm{app}) = \phi_s^*(\theta)(\mathrm{app})$ for some θ by numerical tests. This case does not need any information on $V(\theta)$. In Case 2, the value $V'(0)$ is needed, and, in Case 3, the full information on $V(\theta)$ is necessary. The assumption of Case 3 is obtained as follows. From (3.17), $D(\theta)$ might be approximated by $V(\theta)$ for small ρ and by $\tilde{G}_e(\theta)$ for large ρ. Hence, it might seem reasonable to take the linear interpolation of those two distributions for $D(\theta)(\mathrm{app})$. However, this does not hold good for large $E(S^3)$ since Eq(app) contains $D''(0)$ and so $E(S^3)$ in this case (cf. Appendix), which contradicts the fact that Eq(app) is finite if and only if ES^2 is finite (cf. Miyazawa (1979)). Thus we give $D(\theta)(\mathrm{app})$ from (3.20). In Cases 2 and 3, the properties of Lemmas 2.2 and 2.3 are satisfied. In the appendix, we give the formulas for the mean queue length of these three cases. The numerical tests shows that their values compete in quality with that of Tjims et al.'s approximation. In particular, the behaviour in Case 1 is very similar to that of Tijms et al.'s approximation except in the low-traffic case. Some of those numerical values are given in Tables 1 and 2.

TABLE 1
The mean queue length of $M/E_3/5$

ρ	Hokstad	Tijms	Case 1	Case 2	Case 3	Exact
0.1	$0.13000E-04$	$0.16368E-04$	$0.13617E-04$	$0.16417E-04$	$0.16208E-04$	$0.16161E-04$
0.3	$0.57541E-01$	$0.69136E-02$	$0.63262E-02$	$0.69650E-02$	$0.66925E-02$	$0.66699E-02$
0.6	$0.23651E+00$	$0.26335E+00$	$0.25901E+00$	$0.26587E+00$	$0.24806E+00$	$0.25196E+00$
0.9	$0.45750E+00$	$0.47067E+01$	$0.47179E+01$	$0.47259E+01$	$0.45435E+01$	$0.46341E+01$

TABLE 2
The mean queue length of $M/H_2/5$ with $q = 0.81$ and $\delta^2 = 2.25$
$(G(x) = q(1-\exp(-f1x)) + (1-q)(1-\exp(-f2x)): f1 = 0.1.62017\ f2 = 0.37983)$

ρ	Hokstad	Tijms	Case 1	Case 2	Case 3	Exact
0.1	$0.31687E-04$	$0.22457E-04$	$0.26934E-04$	$0.22547E-04$	$0.21811E-04$	$0.22244E-04$
0.3	$0.14025E-01$	$0.10848E-01$	$0.11486E-01$	$0.10940E-01$	$0.10194E-01$	$0.10742E-01$
0.6	$0.57562E+00$	$0.50109E+00$	$0.50654E+00$	$0.50527E+00$	$0.47129E+00$	$0.50462E+00$
0.9	$0.11151E+02$	$0.10790E+02$	$0.10798E+02$	$0.10820E+02$	$0.10580E+02$	$0.10896E+02$

4. Approximation formulas for $M/GI/s/k$

Consider the case of $M/GI/s/k$. This case was studied in Hokstad (1978) and Van Hoorn (1983), but explicit expressions such as (3.6) and so on have not been obtained. In this section, we give explicit approximation formulas for it by using (2.5) in a similar way as in the case of $M/GI/s$. In this case, we need additional assumptions to determine $N(\theta(x), x)$ of (2.5). The next assumption is used for this.

$$\text{(C)} \qquad \phi_{s+k}(\theta) = \tilde{G}_e(\theta) \quad \text{for any } \theta > 0.$$

Here, we only consider the most simple case, i.e., we assume the assumptions of Theorem 3.1 for L and M in (2.5). In this section we always assume $p \neq 1$.

Theorem 4.1. In $M/GI/s/k$, the assumptions of Theorem 3.1 and (C) imply

$$p_0(\text{app}) = \frac{1}{\sum_{j=0}^{s-1} (\lambda ES)^j/j! + (1-\rho\eta^{-k})(\lambda ES)^s/(1-\rho)s!}, \tag{4.1}$$

$$p_j(\text{app}) = p_0(\text{app})(\lambda ES)^j/j! \qquad (j = 1, \cdots, s-1), \tag{4.2}$$

$$p_{s+k}(\text{app}) = \rho\eta^{-k} p_{s-1}(\text{app}), \tag{4.3}$$

$$\psi(0, x)(\text{app}) = \frac{1-\tilde{G}(\theta(x))}{\tilde{G}(\theta(x)) - x}(p_{s-1}(\text{app}) - x^k p_{s+k}(\text{app})/\rho) + x^k p_{s+k}(\text{app}). \tag{4.4}$$

This theorem can be obtained by using (C) and the fact that $\psi(0, x)$ exists for all x and so the numerator of (2.5) equals 0 for $x = \eta$. Note that this approximation is exact for $k = 0$ and it coincides with the Hokstad–Tijms approximation for $k = +\infty$. Of course, it is exact for all k in the case of the exponential service time. Except those extremal cases, $\psi(0, x)(\text{app})$ does not have good properties as a generating function since it is not in general a polynomial of order k. Nevertheless, moments obtained from it give good approximations in numerical tests. In particular, the mean queue length given below much improves the approximation by Nozaki and Ross (1978), which is also true in the theoretical sense since our assumptions weaken theirs (cf. Table 3). The following results are easily obtained from Theorem 4.1.

Corollary 4.1. In $M/GI/s/k$, under the same assumptions of Theorem 4.1, we have

$$\text{Eq(app)} = \frac{\rho^2 m_2}{2(1-\rho)^2}\{p_{s-1}(\text{app}) - p_{s+k}(\text{app})/\rho\} - k\rho p_{s+k}(\text{app})/(1-\rho). \tag{4.5}$$

In Tables 3 and 4, we give some numerical values, in which (app) means our approximation. We note that if we use better assumptions, such as those of Tijms et al.'s Case A and so on, then we can expect better approximations.

TABLE 3
The mean queue length in $M/E_3/3/k$

	$k=2$	$k=5$	$k=10$	$k=20$	$k=30$	
	0.01880	0.02196	0.02202	0.02202		(exact)
$\rho=0.3$	0.01777	0.01998	0.02001	0.02001		(app)
	0.01452	0.01971	0.02001	0.02001		(Nozaki and Ross)
	0.17720	0.32808	0.36752	0.36943	0.36943	(exact)
$\rho=0.6$	0.18478	0.32072	0.35322	0.35474	0.35474	(app)
	0.08459	0.25134	0.34046	0.35458	0.35458	(Nozaki and Ross)
	0.4833	1.429	2.744	4.19538	4.71849	(exact)
$\rho=0.9$	0.52701	1.48026	2.76944	4.18212	4.68857	(app)
	−0.76830	−0.31598	0.87958	2.79606	3.89426	(Nozaki and Ross)

TABLE 4
The loss probability of customers in $M/E_3/3/k$

	$k=2$	$k=5$	$k=10$	$k=20$	$k=30$	
$\rho=0.3$	$2.84E-3$	$2.48E-5$	$6.65E-9$	$4.23E-16$		(exact)
	$1.78E-3$	$1.24E-5$	$3.11E-9$	$1.98E-16$		(app)
$\rho=0.6$	$4.20E-2$	$4.60E-3$	$1.16E-4$	$7.34E-8$	$4.66E-11$	(exact)
	$3.42E-2$	$3.59E-3$	$9.00E-5$	$5.71E-8$	$3.62E-11$	(app)
$\rho=0.9$	$1.4\ E-1$	$6.02E-2$	$3.85E-3$	$2.12E-3$	$7.83E-4$	(exact)
	$1.29E-1$	$5.62E-2$	$3.68E-3$	$2.01E-3$	$7.47E-4$	(app)

Appendix

We give approximation formulas of Cases 1, 2 and 3 for the mean queue length of $M/GI/s$.

Corollary A.1. Under the same assumptions as in Theorem 3.3, if $D^{(2)}(0)$ exists, we have

$$\text{(A.1)}\qquad \mathrm{Eq(app)} = \left[\frac{1}{2s^2(1-\rho)^2}\lambda d_1/s + \frac{1}{s(1-\rho)}(d_1 + \lambda d_2/2s)\right] p_s(\mathrm{app}),$$

where $d_i = -D^{(i)}(0)$ for $i = 1,2,3$.

Proof. Firstly let us define

$$Y(\theta) = \frac{\lambda(1 - D(\theta)(\mathrm{app}))}{\lambda(\tilde{G}(\theta) - 1) + s\theta}.$$

Then, from (3.12), we have

$$\text{(A.2)}\qquad \psi(0, x) = [1 + xY(\theta(x))]p_s(\mathrm{app}).$$

Recall the relation $\theta(x)=\lambda(1-x)/s$. Hence, from (A.2), we get

$$\text{Eq(app)}=[Y(0)+\theta'(1)Y'(0)]p_s(\text{app}). \tag{A.3}$$

From the definition of Y,

$$Y(\theta)(\lambda(\tilde{G}(\theta)-1)+s\theta)=\lambda(1-D(\theta)(\text{app})).$$

By twice differentiating both sides of this equation and letting θ tend to 0, we have

$$Y(0)=\frac{\lambda d_1}{s(1-\rho)},$$

$$Y'(0)=-\frac{\lambda(Y(0)E(S^2)+d_2)}{2s(1-\rho)}.$$

Thus we obtain (A.1).

Next, we give the formula for Cases 1 and 2. Of course, this is a special case of (A.1), but it is useful for numerical calculations to give it separately, since $D(\theta)(\text{app})$ is complicated in those cases. The only difference between Case 1 and Case 2 is $p_s(\text{app})$, so we have left $p_s(\text{app})$ as a parameter.

Corollary A.2. In Cases 1 and 2, we have

$$\begin{aligned}\text{Eq(app)}&=\frac{\lambda^2E(S^2)}{2s^2(1-\rho)^2}p_{s-1}(\exp)\\&\quad+(\rho p_{s-1}(\exp)-(1+\rho(\delta^2-1)/2)p_s(\text{app})/(1-\rho).\end{aligned} \tag{A.4}$$

We remark that the second term on the right-hand side of (A.4) amends the property pointed out for the Hokstad–Tijms approximation by Tijms et al. (1981). Of course, all our approximations agree with the exact values if G is an exponential distribution.

Acknowledgement

The author expresses his hearty thanks to Professor Y. Takahashi, who calculated exact values for the hyperexponential cases, and also to the referee, whose comments were helpful in revising the first version of this paper.

References

Boxma, O. J., Cohen, J. W. and Huffels, N. (1980) Approximations of the mean waiting time in an $M/G/s$ queueing system. *Operat. Res.* **27**, 1115–1127.

Burman, D. Y. and Smith, D. R. (1983) A light-traffic theorem for multi-server queues. *Math. Operat. Res.* **8**, 15–25.

Groenevelt, H., van Hoorn, M. H. and Tijms, H. C. (1984) Tables for $M/G/c$ queueing systems with phase-type service. *European. J. Operat. Res.* **16**, 257–269.

HAJI, R. AND NEWELL, G. F. (1971) A relation between stationary queue and waiting time distributions. *J. Appl. Prob.* **8**, 617–620.

HILLIER, F. S. AND YU, O. S. (1981) *Queueing Tables and Graphs.* North-Holland, New York.

HOKSTAD, P. (1978) Approximations for the $M/G/m$ queue. *Operat. Res.* **26**, 510–523.

KIMURA, T. (1983) Diffusion approximation for an $M/G/m$ queue. *Operat. Res.* **31**, 304–321.

KÖLLERSTRÖM, J. (1974) Heavy traffic theory for queues with several servers. I. *J. Appl. Prob.* **11**, 544–552.

MIYAZAWA, M. (1979) A formal approach to queueing processes in the steady state and their applications. *J. Appl. Prob.* **16**, 332–346.

MIYAZAWA, M. (1983) The derivation of invariance relations in complex queueing systems with stationary inputs. *Adv. Appl. Prob.* **15**, 874–885.

NOZAKI, S. A. AND ROSS, S. M. (1978) Approximations in finite-capacity multi-server queues with Poisson arrivals. *J. Appl. Prob.* **15**, 826–834.

STOYAN, D. (1976) Approximations for $M/G/s$ queues. *Math. Operationsforsch. Statist. Ser. Optimization* **7**, 587–594.

TAKAHASHI, Y. (1977) An approximation formula for the mean waiting time of an $M/G/c$ queue. *J. Operat. Res. Soc. Japan* **20**, 150–163.

TAKAHASHI, Y. (1981) Asymptotic exponentiality of the tail of the waiting time distribution in a $PH/PH/c$ queue. *Adv. Appl. Prob.* **13**, 619–630.

TIJMS, H. C., VAN HOORN, W. H. AND FEDERGRUEN, A. (1981) Approximations for the steady-state probabilities in the $M/G/c$ queue. *Adv. Appl. Prob.* **13**, 186–206.

VAN HOORN, M. H. (1983) Algorithms and Approximations for Queueing Systems. Doctoral Thesis, Vrije Universiteit te Amsterdam.