# ON THE RELATION BETWEEN THE DISTRIBUTIONS OF THE QUEUE SIZE AND THE WAITING TIME

BY HIDENORI MORIMURA

## §1. Introduction.

In many articles on queuing theory, two measures of effectiveness, that is, queue size and waiting time are dealt with. However, it seems that their handlings are separated in many cases.

In this paper, we shall remark on the relation between the distribution of queue size and that of waiting time (especially the relation of the expectations). In some text books on operations research (e. g. [8]) by the rough and intuitive argument, the relation: $E(L)=\lambda E(W)$ is described, where $1/\lambda$ is the mean interarrival time, $E(W)$ is the expected waiting time and $E(L)$ is the mean queue size in the equilibrium state. And, the exact proof of this relation were done by calculating the both sides of this equality separately in some special cases. For instance, Morse [8] showed that the relation is valid in the cases $M/M/s$, $M/E_k/1$ and $E_2/M/1$.

We shall consider here four types of queue size in the equilibrium state which are denoted by $L$, $L^*$, $Q$ and $Q^*$ as follows:

$L$: queue size (not include the customer being served) observed at any time,

$L^*$: queue size observed at the epoch just before a customer arrives,

$Q$: queue size observed at the epoch just before the service of a customer begins,

$Q^*$: queue size observed at the epoch just after the service of a customer has finished.

In the above if we try to describe more exactly, (for example, to say about $L$), we must define $L$ as the random variable obeying to the limit distribution of $L(t)$ as $t\to\infty$, where $L(t)$ means the queue size at time $t$.

Throughout this paper we shall assume that $sE(X_i)>E(Y_i)$, which guarantees the existence of the limit distribution of $L(t)$, where $s$, $E(X_i)$ and $E(Y_i)$ mean the number of servers, the expected interarrival time and the expected service time, respectively.

Furthermore, this assumption will guarantee the existence of the equilibrium distribution (i. e., the limit distribution) of the other quantities mentioned in the above. (This facts were shown in [1] and [5].)

In the present paper we shall show that

( i ) the distribution of $L^*$ will be expressed using the distribution of the waiting time $W$ in the equilibrium state;

Received September 9, 1961.

( ii ) the distributions of $L^*$, $Q^*$ and $Q$ are identical in the single server case;

( iii) the same fact will hold in the many server case under the assumption that the distributions of the interarrival time and service time are both absolutely continuous;

( iv) the relation $E(L^*)=E(L)=E(Q)=E(Q^*)=\lambda E(W)$ is valid for the case $M/G/s$ (but the statement is somewhat rough);

( v ) in the case $M/G/1$, $E(L)$ and $\mathrm{Var}(L)$ will be expressed in some exact forms;

( vi) in the case of non-Poissonian arrivals, there are two examples that the inequalities $E(Q)>E(W)/E(X_i)$ and $E(Q)>E(W)/E(Y_i)$ hold respectively;

(vii) but the inequality $E(Q)<E(W)/E(Y_i)$ will hold in a general case;

and remark on some other facts;

## §2. Some properties of the function $G(x)$.

In this section, we shall restrict ourselves in the single server case, and use the method of the imbedded queuing process of general type due to Kawata [3]. Then, first of all, some notations and relations needed to do the following discussions will be introduced here.

As in [3], we shall use the following notations. Let

$$t_0<t_1<t_2<\cdots$$

be a sequence of instants when customers successively arrive at the service station, and set

$$t_j-t_{j-1}=X_j, \qquad j=1, 2, \cdots$$

which are interarrival times. Furthermore let $Y_j$ $(j=1, 2, \cdots)$ be the service time which is required by the $j$-th customer who has arrived at the epoch $t_{j-1}$. Throughout the paper, we assume that each of $\{X_j\}$ and $\{Y_j\}$ is a sequence of independent random variables having identical distributions, and $X_j$ and $Y_j$ are also mutually independent. Set $Z_j=Y_j-X_j$ $(j=1, 2, \cdots)$ and

$$\begin{aligned}
&S_n=\sum_{j=1}^{n} Z_j,\\
&a_n=P(S_1>0, S_2>0, \cdots, S_n>0) \qquad (n\geqq 1),\\
&a_0=1,\\
&b_n=P(S_n>0) \qquad (n\geqq 1).
\end{aligned}$$

We assume also $-\infty<E(Z_j)<0$.

Moreover, we shall denote the number of customers in the system at time $t$ by $\eta(t)$. It was already shown (e. g. [3]) that the distribution of $\eta(t_n-0)$ converges to the limit distribution $\{p^{(m)}\}$ $(m=0, 1, 2, \cdots)$ as $n\to\infty$ under the condition $-\infty<E(Z_j)<0$.

Now, for the sake of expression of the limit distribution in the exact form, Kawata [3] introduced the following function:

$$G(x)=\sum_{n=1}^{\infty}F^{(n)}(x), \tag{2.1}$$

where $F^{(n)}(x)$ is defined recurrently as

$$\begin{aligned} F^{(1)}(x) &= P(S_1<x),\\ F^{(2)}(x) &= \int_0^\infty F(x-y)\,dF^{(1)}(y),\\ &\cdots\cdots\cdots\cdots\cdots,\\ F^{(n)}(x) &= \int_0^\infty F(x-y)\,dF^{(n-1)}(y), \end{aligned} \tag{2.2}$$

from which we have

$$P(S_1>0,\ S_2>0,\ \cdots,\ S_{k-1}>0,\ S_k>x)=\int_x^\infty dF^{(k)}(y).$$

Using the function $G(x)$ he gets the expression of the limit distribution as follows:

$$p^{(0)}=e^{-K}=\exp\left(-\sum_{n=1}^{\infty}\frac{b_n}{n}\right), \tag{2.3}$$

$$p^{(m)}=e^{-K}\int_0^\infty\left(\int_x^\infty dG(y)\right)d[F_{m-1}(x)-F_m(x)] \qquad (m\geqq 1) \tag{2.4}$$

where

$$F_m(x)=P\left(\sum_{i=1}^{m}X_i<x\right), \tag{2.5}$$

Furthermore, recently he shows [4] that the all $\nu$-th moments of $G(x)$ exist $(\nu=1,2,\cdots,n)$ if $E\{Z_j^\nu\}<\infty$ $(\nu=1,2,\cdots,n+1)$ under some analytical condition.

First of all, we shall show that $G(x)$ multiplying by the constant $e^{-K}$ means the disiribution function of the waiting time in the steady state for $x>0$.

In fact, for some $\nu$ if the service station which was vacant at the arrival epoch of the $(n-\nu)$th customer has been not vacant continuously from the epoch, then the waiting time of the $n$-th customer $W_n$ will be equal to $Z_{n-\nu+1}+Z_{n-\nu+2}+\cdots+Z_n$. Then, we have[1)]

$$P\{W_n>x\}=\sum_{\nu=1}^{n}p_{n*\nu}{}^{(0)}P(S_1>0,\ S_2>0,\ \cdots,\ S_{\nu-1}>0,\ S_\nu>x) \quad \text{for } x\geqq 0, \tag{2.6}$$

where $p_\nu{}^{(0)}$ means the probability that the system is empty at the time when $\nu$-th customer arrives. Then we can easily see from the results in [3] that

1) In this formula, $P(S_1>0,\ S_2>0,\ \cdots,\ S_{\nu-1}>0,\ S_\nu>x)$ must be interpreted as $P(S_1>x)$ for $\nu=1$. For convenience' sake, the same notation will be used below.

$$\lim_{n\to\infty} P(W_n > x)$$

$$= e^{-K}\sum_{\nu=1}^{\infty} P(S_1 > 0,\ S_2 > 0,\ \cdots,\ S_{\nu-1} > 0,\ S_\nu > x) \tag{2.7}$$

$$= e^{-K}\int_x^\infty dG(y).$$

On the other hand, we have

$$G(\infty) = \int_{-\infty}^\infty dG(x) = \int_{-\infty}^\infty d\left\{\sum_{\nu=1}^{\infty} P(S_1 > 0,\ \cdots,\ S_{\nu-1} > 0,\ S_\nu < x\right\}$$

$$= \sum_{\nu=1}^{\infty}\int_{-\infty}^\infty dP(S_1 > 0,\ \cdots,\ S_{\nu-1} > 0,\ S_\nu < x) \tag{2.8}$$

$$= \sum_{\nu=0}^{\infty} P(S_1 > 0,\ \cdots,\ S_{\nu-1} > 0) = \sum_{k=0}^{\infty} a_k = e^K,$$

since the termwise integrability in above was guaranteed in [3].

Thus, we have for $x > 0$,

$$P(W \leqq x) \equiv \lim_{n\to\infty} P(W_n \leqq x) = e^{-K}\left(e^K - \int_x^\infty dG(y)\right) = e^{-K}G(x). \tag{2.9}$$

Further, since

$$G(0) = \int_{-\infty}^0 dG(x) = \sum_{k=1}^{\infty} P(S_1 > 0,\ \cdots,\ S_{k-1} > 0,\ S_k < 0) \tag{2.10}$$

equal to unity, the probability that a customer will find an empty system when he has just arrived corresponds to $e^{-K}G(0)$. For $x < 0$, the interpretation of the physical meaning of $G(x)$ is troublesome.

Since the steady state distribution of the number of customers in the system at the time just before the arrival is given by (2.4), we can calculate it using (2.9) if the steady state distribution of the waiting time was known, which will be calculated easer than the queue size distribution in many cases.

For example, we have easily the following results based on the waiting time distribution obtained by Pollaczek [9] as examples of Poisson arrival case with single server. In these examples since the distributions of interarrival time are always negative exponential with parameter $\lambda$,

$$d\{F_{m-1}(x) - F_m(x)\} = \begin{cases} \left\{\dfrac{\lambda^{m-1}x^{m-2}}{(m-2)!} - \dfrac{\lambda^m x^{m-1}}{(m-1)!}\right\} e^{-\lambda x}\,dx & (x > 0), \\ 0 & (x \leqq 0). \end{cases} \tag{2.11}$$

EXAMPLE 1 (Negative exponential service).

$$P(Y_j \leqq x) = 1 - e^{-x}, \tag{2.12}$$

$$P(W > x) = \lambda e^{(1-\lambda)x} \qquad (x > 0). \tag{2.13}$$

Thus,

$$\int_0^\infty \lambda e^{-(1-\lambda)x} \frac{\lambda^m x^{m-1}}{(m-1)!} e^{-\lambda x} dx = \lambda^{m+1}, \tag{2.14}$$

hence

$$p^{(m)} = (1-\lambda)\lambda^m, \tag{2.15}$$

which is the well known result.

EXAMPLE 2 (Hyper-exponential service).

$$P(Y_j < x) = a_1(1-e^{-b_1 x}) + a_2(1-e^{-b_2 x}), \qquad a_1 + a_2 = 1;\ b_1, b_2 > 0, \tag{2.16}$$

$$P(W > x) = \lambda\left(\frac{\beta_2 - 1 + \lambda}{\beta_2 - \beta_1} e^{-\beta_1 x} + \frac{\beta_1 - 1 + \lambda}{\beta_1 - \beta_2} e^{-\beta_2 x}\right), \tag{2.17}$$

where $\beta_1$ and $\beta_2$ are the roots of the following quadratic equation:

$$x^2 + (b_1 - b_2 - \lambda)x + b_1 b_2 (1-\lambda) = 0. \tag{2.18}$$

Thus, since

$$\int_0^\infty P(W > x) \frac{\lambda^m x^{m-1}}{(m-1)!} e^{-\lambda x} dx = \lambda^{m+1}\left\{\frac{\beta_2 - 1 + \lambda}{\beta_2 - \beta_1} \frac{1}{(\beta_1 + \lambda)^m} + \frac{\beta_1 - 1 + \lambda}{\beta_1 - \beta_2} \frac{1}{(\beta_2 + \lambda)^m}\right\},$$

we have

$$p^{(m)} = \frac{\lambda^m}{\beta_2 - \beta_1}\left\{\frac{\beta_1(\beta_2 - 1 + \lambda)}{(\beta_1 + \lambda)^m} - \frac{\beta_2(\beta_1 - 1 + \lambda)}{(\beta_2 + \lambda)^m}\right\}. \tag{2.19}$$

## § 3. Relations between the various types of queue size.

In the case of many servers, the above discussion can not be applied. But, Kiefer and Wolfowitz [5] showed the similar and general relation as follows:

$$P(Q \geqq n) = \int_0^\infty P(X_1 + \cdots + X_n < x)\, dG^*(x) \tag{3.1}$$

where $Q$ denotes the queue size just before the service of a customer begins, and $G^*(x)$ is the distribution of the waiting time in the steady state.

Comparing the above relation to ours in the single server case, we can easily see that (2.4) and (3.1) are essentially similar except a slight difference on the observation epochs of queue size. In fact, if we denote by $L^*$ the queue size in the equilibrium state at the epoch just before a customer arrives, we can see that

$$P(L^* = m) = \begin{cases} p^{(m+1)} & \text{for } m \geqq 1, \\ p^{(0)} + p^{(1)} & \text{for } m = 0 \end{cases} \tag{3.2}$$

and

$$\begin{aligned} P(L^* \geqq n) &= \sum_{m=n+1}^{\infty} p^{(m)} \\ &= e^{-K}\int_0^\infty F_n(x)\, dG(x) = P(Q \geqq n) \qquad \text{for } n \geqq 1, \end{aligned} \tag{3.3}$$

from the relations (3.2), (2.4), (2.5), (2.9) and (3.1). Of course, the equality (3.3) will be meaningful only in the case of single server. But we shall note below that this relation will hold in the case of Poisson arrival, many servers and absolutely continuous service time distribution.

First of all, we shall show that the distributions of $L^*$, $Q^*$ and $Q$ are all identical in the many server case under the condition that the distributions of the interarrival time and the service time are both absolutely continuous. And, we must quote the result due to Finch [1] to do this. He showed in the $s$ server case, if the distributions of the interarrival time and service time are both absolutely continuous and $E(Y_i) < sE(X_i)$, then the following limit exists:

$$\alpha_k \delta t = \lim_{t\to\infty} P\{\eta(t+\delta t) = \eta(t)+1 \mid \eta(t) = k\} \qquad (k = 0, 1, 2, \cdots) \tag{3.4}$$

and

$$p^{(k)} = q^{(k)} = E(X_i)\cdot\alpha_k\cdot\lim_{t\to\infty} P\{\eta(t) = k\}. \tag{3.5}$$

In the following, we shall denote as

$$\lim_{n\to\infty} P\{\eta(t_n + W_n + Y_n + 0) = k\} = q^{(k)},$$

the existence of which was known (for instance, [1]).

By the definition of $Q$, we can see that the event $Q = m\ (m \geqq 1)$ will happen if and only if the system is busy when the costomer has arrived. Then we have

$$P(Q = m) = q^{(m+s)} \qquad (m \geqq 1) \tag{3.6}$$

since the epochs when the service of a customer begins and that of the previous customer finishes consist with each other in this case. And, the event $Q = 0$ will happen in two ways, one of which is the case to exist free servers in the system when a customer arrives and the other is the case of no queue behind him when the service of the customer begins. Then, we have

$$P(Q = 0) = p^{(0)} + p^{(1)} + \cdots + p^{(s-1)} + q^{(s)} \tag{3.7}$$

$$= q^{(0)} + q^{(1)} + \cdots + q^{(s)} \tag{3.8}$$

by (3.5). Obviously, $P(Q^* = 0)$ and $P(Q^* = m)$ $(m \geqq 1)$ are equal to the right hand of (3.7) and (3.6), respectively. Thus, we can say from (3.5) that the distributions of $L^*$, $Q^*$ and $Q$ are all identical in our case.

If we restrict ourselves to the single server case, then we can see that the distributions of $Q$, $Q^*$ and $L^*$ are all identical without the condition that the distributions of the interarrival time and the service time are both absolutely continuous. In fact, in this case (3.7) will be reduced as

$$P(Q = 0) = p^{(0)} + q^{(1)}. \tag{3.9}$$

But the event $\eta(t_n + W_n + Y_n + 0) = 0$ is equivalent to the event $\eta(t_{n+1} - 0) = 0$, then we can easily see that $p^{(0)} = q^{(0)}$. Thus we may assert that the distribution of $Q$ and $Q^*$ are mutually identical in the single server case. Furthermore, we know (for instance, [1]) that the limit distribution of $\eta(t_n-0)$ will equal to that of $\eta(t_n+W_n+Y_n+0)$. Thus we have an assertion that the distributions of $Q$, $Q^*$ and $L^*$ are all identical in the single server case which is an alternative proof of (3.3).

By the way, in the many server case with the condition that the distribution of service time is absolutely continuous, if the arrival is Poissonian, it is easily seen that $\alpha_k=\lambda$ and $E(X_i)=1/\lambda$, then the three limit distributions in (3.5)are identical. Thus, we can say that the distributions of $L^*$, $L$, $Q$ and $Q^*$ consist with each other in the case of many servers in which the distribution of the service time is absolutely continuous, and arrival is Poissonian.

## §4. Relations between the expected waiting time and mean queue size.

Based upon (2.4) or (3.1), we can calculate the expectation of queue size observed at some epochs. For instance, in the single server case, from (2.4), we have

$$\begin{aligned} E(L^*) &= \sum_{m=1}^{\infty} m p^{(m+1)} \\ &= \sum_{m=1}^{\infty} e^{-K} \int_0^{\infty} \left( \int_x^{\infty} dG(y) \right) \cdot dF_m(x) \qquad (4.1) \\ &= \sum_{m=1}^{\infty} e^{-K} \int_0^{\infty} F_m(x)\, dG(x), \qquad (4.2) \end{aligned}$$

then, we have

$$E(L^*) = e^{-K} \int_0^{\infty} H(x)\, dG(x), \tag{4.3}$$

where

$$H(x) \equiv \sum_{m=1}^{\infty} F_m(x)$$

is the expectation of the renewal number in $(0, x)$, which converges uniformly in any finite interval of $x$. Again, in the many server case, from (3.1), same argument will imply

$$E(Q) = \int_0^{\infty} H(x)\, dG^*(x). \tag{4.4}$$

Thus, we have the following

THEOREM 1. *In a queuing system with $s$ servers ($s \geqq 1$), if the arrivals are in Poisson fashion and if $E(Y_i) < sE(X_i)$, then we have*

$$E(Q) = \lambda E(W)$$

*for any distribution of the service time.*

*Proof.* From the above discussion, we can easily see that the assertion

$E(Q)=\lambda E(W)$ is equivalent to

$$\int_0^\infty H(x)\,dG^*(x)=\lambda\int_0^\infty x\,dG^*(x). \tag{4.5}$$

If the distribution of $X_i$ is the negative exponential with parameter $\lambda$:

$$F(x)=\begin{cases}1-e^{-\lambda x} & \text{for} \quad x>0,\\ 0 & \text{for} \quad x\leqq 0,\end{cases} \tag{4.6}$$

it is well known that

$$H(x)=\begin{cases}\sum_{m=1}^{\infty}F_m(x)=\lambda x & \text{for} \quad x>0,\\ 0 & \text{for} \quad x\leqq 0,\end{cases} \tag{4.7}$$

thus we have (4.5) directly.

COROLLARY 1. *In a queuing system with s servers* ($s\geqq 1$), *if the arrivals are in Poisson fashion and the distribution of service time is absolutely continuous, then*

$$E(Q)=E(L)=E(L^*)=E(Q^*)=\lambda E(W).$$

*Proof.* The above remark in § 3 will imply that $E(Q)=E(L^*)=E(Q^*)=E(L)$, then the assertion is obvious.

This theorem unify and generalize the previous results on the relation in the case of Poisson arrivals.

By the way, Morse [8] ($E_2/M/1$) and Kawamura [2] ($E_l/M/s$, $E_l/E_k/1$) showed that the relation: $E(L)=\lambda E(W)$ is valid. However, in our case on $E(L^*)$ this relation is not true. To show the fact, we shall show the following

EXAMPLE 3. In the queuing system with single server we shall assume that the distribution of interarrival time is Erlangian and $E(Y_i)<E(X_i)$. Thus, let the density function of $X_i$ be

$$f(x)=\begin{cases}\dfrac{(\lambda l)^l}{\Gamma(l)}x^{l-1}e^{-\lambda l x} & \text{for} \quad x>0,\\ 0 & \text{for} \quad x\leqq 0\end{cases} \tag{4.8}$$

and let $\{\xi_i\}$ be a sequence of mutually independent and identically distributed random variables obeying to the negative exponential distribution with parameter $l\lambda$.

Putting $U_n=\sum_{i=1}^n\xi_i$, we have

$$H(x)=\sum_{n=1}^{\infty}P(X_1+\cdots+X_n<x)=\sum_{n=1}^{\infty}P(U_{nl}<x), \tag{4.9}$$

since we can rewrite as $X_1=\sum_{i=1}^l\xi_i$ and so on. Noting

$$P(U_{nl} < x) \leqq P(U_m < x) \tag{4.10}$$

for all $m \leqq nl$, (4.9) will be evaluated as

$$H(x) \leqq \frac{1}{l} \sum_{m=1}^{\infty} P(U_m < x) = \frac{1}{l} \cdot l\lambda \cdot x = \lambda x. \tag{4.11}$$

Since the equality in (4.10) and (4.11) holds if and only if $l = 1$, then we can see that

$$E(L^*) = e^{-K} \int_0^\infty H(x)\, dG(x) < e^{-K} \int_0^\infty \lambda x\, dG(x) = \lambda E(W), \tag{4.12}$$

if $l > 1$. In the other words, in the case of Erlangian input except the Poisson input case, the relation $E(L^*) = \lambda E(W)$ will not hold.

## § 5. Mean and variance of the queue size in the Poisson arrival case.

In the last section, we showed that the expected queue size at service beginning epochs equals $\lambda$ times of mean waiting time in the case of Poisson arrivals. Further, in this case, when the service time distribution is absolutely continuous, it is also noted that the expected queue size at any time not necessarily particular epochs is also $\lambda$ times of mean waiting time. In the single server case with Poisson arrival, the condition on the service time will be not necessary as was shown by Khinchin (see [10]). And in this case mean waiting time may be found by the so-called Pollaczek-Khinchin-Kendall's formula:

$$E(W) = \frac{\lambda E(Y_\iota^2)}{2\{1 - \lambda E(X_\iota)\}}.$$

In this connection, we shall find some concrete forms for the $E(L)$ and $\mathrm{Var}(L)$ in Poisson arrival case.

THEOREM 2. *In the case of Poisson arrivals with single server, we have in the equilibrium state*

$$E(L) = \lambda E(W) = \frac{\lambda^2 b_2}{2(1 - \lambda b_1)}, \tag{5.1}$$

$$\mathrm{Var}(L) = \lambda^2 \mathrm{Var}(W) + \lambda E(W) = \frac{\lambda^2 b_2}{2(1 - \lambda b_1)} + \frac{\lambda^3 b_3}{3(1 - \lambda b_1)} + \frac{\lambda^4 b_2^2}{4(1 - \lambda b_1)^2}, \tag{5.2}$$

*where $b_i$ $(i = 1, 2, 3)$ are the i-th moments of service time which are assumed to exist.*

*Proof.* (5.1) is evident from the Theorem 1 and the Pollaczek-Khinchin-Kendall's formula. The first assertion of (5.2) will be proved in the following theorem in the more general case. The second assertion of (5.2) is also evident from the result by Pollaczek [9] such as

$$\mathrm{Var}(W) = \frac{\lambda b_3}{3(1 - \lambda b_1)} + \frac{\lambda^2 b_2^2}{4(1 - \lambda b_1)^2}.$$

THEOREM 3. *In the case of Poisson arrivals with many servers,*

$$\mathrm{Var}(Q) = \lambda^2 \mathrm{Var}(W) + \lambda E(W). \tag{5.3}$$

*Proof.* Since $Q$ denote the queue size at the service beginning epochs, we have

$$\begin{aligned} E(Q^2) &= \sum_{n=1}^{\infty} n^2 p_n = \sum_{n=1}^{\infty} n^2\{P(Q \geqq n) - P(Q \geqq n+1)\} \\ &= \sum_{n=1}^{\infty} \{[(n-1)^2 + 2n - 1]P(Q \geqq n) - n^2 P(Q \geqq n+1)\} \qquad (5.4) \\ &= 2\sum_{n=1}^{\infty} nP(Q \geqq n) - \sum_{n=1}^{\infty} P(Q \geqq n) \\ &= 2\int_0^{\infty} \sum_{n=1}^{\infty} nP\{X_1 + \cdots + X_n < x\}\, dG^*(x) - E(Q). \qquad (5.5) \end{aligned}$$

By the way, since

$$P\{X_1 + \cdots + X_n < x\} = \int_0^x \lambda^n e^{-\lambda t} \frac{t^{n-1}}{(n-1)!}\, dt,$$

we have

$$\begin{aligned} \sum_{n=1}^{\infty} nP\{X_1 + \cdots + X_n < x\} &= \int_0^x e^{-\lambda t} \left\{ \frac{d}{dt} \sum_{n=1}^{\infty} \frac{\lambda^n x^n}{(n-1)!} \right\} dt \\ &= \int_0^x (\lambda^2 t + \lambda)\, dt = \frac{\lambda^2 x^2}{2} + \lambda x. \end{aligned} \tag{5.6}$$

Hence, inserting (5.6) into (5.5), we have

$$\begin{aligned} E(Q^2) &= 2\int_0^{\infty} \left\{ \frac{\lambda^2 x^2}{2} + \lambda x \right\} dG^*(x) - E(Q) \\ &= \lambda^2 E(W^2) + 2\lambda E(W) - E(Q). \end{aligned}$$

Thus, Theorem 1 implies that

$$E(Q^2) = \lambda^2 E(W^2) + \lambda E(W). \tag{5.7}$$

Hence,

$$\mathrm{Var}(Q) = E(Q^2) - \{E(Q)\}^2 = \lambda^2 \mathrm{Var}(W) + \lambda E(W),$$

which is (5.3).

Furthermore, the discussions in § 3 and Theorem 3 will imply the following

COROLLARY 2. *In the many server case with Poisson arrival, if the distribution of the service time is absolutely continuous, then we have*

$$\mathrm{Var}(L) = \lambda^2 \mathrm{Var}(W) + \lambda E(W). \tag{5.8}$$

*In* (5.8), *we can replace* $L^*$ *or* $Q^*$ *for* $L$.

COROLLARY 3. *In the single server case with Poisson arrival,* (5.8) *is valid without the additional condition in Corollary* 2.

## § 6. Some further remarks.

REMARK 1. Kiefer and Wolfowitz [5] noted the following relation without any proof:

$$\frac{E(W)}{E(X_i)}-1<E(Q)\leqq\frac{E(W)}{E(X_i)}. \tag{6.1}$$

We shall remark here something on the second inequality. In § 4, we noted two examples each of which is the case of the equality and of the strict inequality. But we shall give an example which will give the converse inequality.[2)]

In a single server queuing system, let the distribution of the interarrival time $X_i$ be

$$\begin{cases} P(X_i=h)=\dfrac{1}{2}, \\ P(X_i=29h)=\dfrac{1}{2} \end{cases} \tag{6.2}$$

where $h$ is a fixed positive number. Furthermore let the service time be identically $3h$. Then, obviously, we have

$$E(X_i)=15h>3h=E(Y_i) \tag{6.3}$$

which satisfies the condition of ergodicity.

In this case, since $E(Q)=E(L^*)$ we shall consider the same inequality on $E(L^*)$. Let $L_n^*$ be the queue size at the epoch just before the $n$-th customer arrives, while $W_n$ be his waiting time.

Then we shall have

$$W_n\leqq Y_{n-k-1}+Y_{n-k}+\cdots+Y_{n-1}=3h(k+1)$$

if $L_n^*=k$, so that

$$E(W)\leqq 3h[E(L^*)+1]. \tag{6.4}$$

By the way, it is evident that

$$P(L^*=1)<\frac{1}{4}$$

and

$$E(L^*)>\frac{1}{4}. \tag{6.5}$$

Thus, we have

$$E(L^*)>\frac{1}{4}=\frac{3h}{E(X_i)-3h}, \tag{6.6}$$

from (6.3). (6.6) implies that

2) The construction of this example is due to Dr. H. Hatori. The original one by the author was more complicated.

$$E(L^*) > \frac{3h[E(L^*)+1]}{E(X_i)}. \tag{6.7}$$

Combining (6.7) and (6.4) we have

$$E(Q) = E(L^*) > \frac{E(W)}{E(X_i)}$$

which is a contrary relation to (6.1).

REMARK 2.[3] When we replace $E(Y_i)$ for $E(X_i)$, the resulting inequality $E(Q) < E(W)/E(Y_i)$ will hold in the single server case.

Analogously to the case of the above example, we shall consider the same inequality on $E(L^*)$ instead of $E(Q)$ in this case too.

The event $L_n^* = k$ is relevant to the sample values of the random variables $X_1, X_2, \cdots, X_{n-1}$ and $Y_1, Y_2, \cdots, Y_{n-k-1}$, that is, it is independent of the random variables $Y_{n-k}, \cdots, Y_{n-1}$ $(k = 1, 2, \cdots, n-2)$. Thus, noting $W_n \geqq Y_{n-k} + \cdots + Y_{n-1}$, we can see that

$$E(W_n) \gneqq E(Y_{n-k} + \cdots + Y_{n-1}) = E(L_n^*)E(Y_i). \tag{6.8}$$

Letting $n \to \infty$, we can say

$$E(W) \geqq E(L^*) \cdot E(Y_i). \tag{6.9}$$

In (6.8) the equality sign will be deleted except a trivial case when $D/D/1$ with $E(X_i) \geqq E(Y_i)$, in the other words, $W_n = L_n^* = 0$ for all $n$.

In fact, if the probability of the event $W_n > Y_{n-k} + \cdots + Y_{n-1}$ is positive, then the inequality in (6.8) will hold strictly. But if it is always zero it means that the service of $(n-k)$-th customer always begins at the same time when the $n$-th customer arrives. This is impossible except the trivial case mentioned above.

REMARK 3. In the example 3, we discussed under the constraint of input distribution as Erlangian. But, we may regard the fact as a converse theorem of Theorem 1 in the practical sense, because we know that the relation $E(L) = \lambda E(W)$ is valid only in Erlangian input case [2], [8], and in the case of general input, it may be known from the results by Kendall [7] or Wishart [11] that the relation $E(L^*) = \lambda E(W)$ does not hold.

REMARK 4. A sufficient condition for the second inequality of (6.1) is that for all $x$

$$H(x) \leqq \frac{x}{E(X_i)}. \tag{6.10}$$

It is evident that the inequality will deduced from (6.10). Under the condition (6.10), if the distribution of service time is strictly increasing, a converse assertion of Theorem 1 will be proved as follows. In fact, since the distribu-

3) This fact has been suggested by Dr. H. Hatori in his communication to the author.

tion of $Y_i$ is strictly increasing function, we can easily see that the function $F^{(n)}(x)$ are strictly increasing for all $n$, so that the function $G(x)$ is also strictly increasing.

Now, we assume (4.5) to hold, i. e.,

$$\int_0^\infty [\lambda x - H(x)]\, dG(x) = 0 \qquad \text{for all} \quad x > 0. \tag{6.11}$$

Thus, from the relations (6.10) and (6.11), we can say

$$\lambda x = H(x) \qquad \text{for all} \quad x > 0.$$

Furthermore, $H(x)$ must be the solution of the integral equation (renewal equation):

$$H(x) = F(x) + \int_0^x H(x-y)\, dF(y). \tag{6.12}$$

Substituting $\lambda x$ for $H(x)$ in (6.12), we have

$$\lambda x = F(x) + \lambda \int_0^x F(y)\, dy.$$

Since $F(x)$ is the continuous distribution function, it is integrable in any finite interval $(0, x)$. Then we can put as

$$K(x) = \int_0^x F(y)\, dy, \qquad x > 0. \tag{6.13}$$

Using this notation, we can rewrite the equation (4.15) as follows:

$$\frac{dK(x)}{dx} + \lambda K(x) = \lambda x. \tag{6.14}$$

(6.14) is a simple linear differential equation, and we have directly

$$K(x) = x - \frac{1}{\lambda}(1 - e^{-\lambda x}), \qquad x > 0, \tag{6.15}$$

noting $K(0) = 0$. Hence we have

$$F(x) = 1 - e^{-\lambda x}, \qquad x > 0, \tag{6.16}$$

which asserts that if the relation holds, then the input is Poissonian. Though in the case which was shown in Remark 1, (6.10) does not hold for $x = 2h$, at least. It is also valid that (6.10) holds in many cases. It seems that the question "Under what condition does (6.10) hold?" is relevant to study an extension of Wald's equation.

In conclusion, the author expresses his sincerest thanks to Prof. T. Kawata, Prof. K. Kunisawa, Prof. T. Homma, Prof. T. Kawamura and Dr. H. Hatori who have given valuable advices to him.

## REFERENCES

[1] FINCH, P. D., On the distribution of queue size in queuing problem. Acta Math. Sci. Acad. Hung. **10** (1959) 327–336.

[2] KAWAMURA, T., Equilibrium behaviors of certain queuing systems and numerical tables. To appear in J. General Culture, Keio Univ. 8, **9** (1961).

[3] KAWATA, T., On the imbedded queuing process of general type. Bull. Inst. Intern. Statist. 38 (Actes de la 32 session) 3e Livraison (1961), 445–455.

[4] KAWATA, T., General queuing theory with single server, not yet published.

[5] KIEFER, J., AND J. WOLFOWITZ, On the theory of queues with many servers. Trans. Amer. Math. Soc. 78 (1955), 1–19.

[6] KIEFER, J., AND J. WOLFOWITZ, On the characteristics of the general queuing process with applications to random walk. Ann. Math. Statist. 27 (1956), 147–161.

[7] KENDALL, D. G., Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. Ann. Math. Statist. 24 (1953), 338–354.

[8] MORSE, P. M., Queues, inventories and maintenance. J. Wiley, (1958).

[9] POLLACZEK, F., Problèmes stochastiques posés par le phénomène de formation d'une queue d'attente à un guichet et par de phénomènes apparentés. Mémorial des Sci. Math. 136 (1957).

[10] TAKÁCS, L., Investigation of waiting time problems by reduction to Markov processes. Acta Math. Acad. Sci. Hung. **6** (1955) 101–129.

[11] WISHART, D. M. G., A queuing system with $\chi^2$ service-time distribution. Ann. Math. Statist. 27 (1956), 768–779.

DEPARTMENT OF MATHEMATICS,
TOKYO INSTITUTE OF TECHNOLOGY.

Addendum in the proof. Recently, J. D. C. Little gave an excellent proof for the relation $E(L)=\lambda E(W)$ under a quite loose condition (Oper. Res. **9** (1961), 383–387). A part of our Corollary 1 in §4 is included in his result.