The Annals of Mathematical Statistics
1970, Vol. 41, No. 1, 182–187

ON THE WAITING TIME IN THE QUEUING SYSTEM GI/G/1

BY R. SCHASSBERGER

University of Calgary

1. Introduction. This paper deals with the waiting time $X(t)$ in the queuing model $GI/G/1$. $X(t)$ is defined as the time needed to complete the serving of all those units which are present in the system at time t. In order to obtain information about the distribution of $X(t)$, we use the auxiliary variable $Y(t)$, defined as the time between t and the first arrival after t. The vector $(X(t), Y(t))$ forms a Markov process. We consider the distribution function $L_{y_0}^{x_0}(t; y, x) = P_r\{Y(t) \leqq y, X(t) \leqq x \mid Y(0) = y_0, X(0) = x_0\}$ and obtain in the case $x_0 = 0$, $y_0 \geqq 0$ a closed expression for

$$\hat{L}_{y_0}^{x_0}(\theta; s, w) = \int_{t=0}^{\infty} \int_{y=0}^{\infty} \int_{x=0-}^{\infty} e^{-\theta t} e^{-sy} e^{-wx} L_{y_0}^{x_0}(t; dy, dx)\, dt.$$

This contains as a special case an expression for $\hat{L}_{y_0}^{x_0}(\theta; 0, w)$, which is the Laplace-Stieltjes transform with respect to x and the Laplace transform with respect to t of the distribution function of $X(t)$, thus determining this distribution function completely. The results are valid for arbitrary service-time distribution function $B(t)$ concentrated on $[0, \infty)$ and for any interarrival-time distribution function $A(t)$ with $A(0) = 0$. Our analysis is based on the method of stages, described in [4]. This method exploits the fact that every distribution function $F(t)$ concentrated on $[0, \infty)$ can be approximated weakly as u tends to infinity by distribution functions

$$F_u(t) \doteq F(0) + \sum_{k=1}^{\infty} \{F(k/u) - F((k-1)/u)\} E_u^{k*}(t),$$

where $E_u^{k*}(t)$ is the k-fold convolution of the distribution function $1 - e^{-ut}$ (see [4]).

The results seem to be new. Keilson and Kooharian ([3]) investigated the system and derived expressions for Laplace transfcrms of the regeneration and server occupation time distributions. They were led to Wiener-Hopf type equations, and this aspect of the problem appears in our analysis as well, although we do not use the corresponding techniques. Takács ([5]) has derived the limiting distribution of $X(t)$ as t tends to infinity, using the same auxiliary random variable $Y(t)$ that we use.

The method of stages, as applied in this paper, consists essentially of associating to the Markov process $(X(t), Y(t))$ a family of discrete-state Markov processes, whose members approximate it. The theory of this procedure is intended to be presented in a wider context at a future time and we feel therefore justified in omitting the proof for its validity in this paper. Unfortunately, our method does not cover the case $x_0 > 0$, corresponding to a non-empty queue at $t = 0$. Although we are able to derive the desired expressions in this case for the approximating system, their unwieldy appearance makes the limiting procedure seem intractable.

Received March 10, 1969.

2. A factorization problem. Let $A(t)$ and $B(t)$ be distribution functions concentrated on $[0, \infty)$, and let $a(s)$, $b(s)$ be their Laplace-Stieltjes transforms. For real $\theta > 0$, the function $a(s)b(\theta-s)$ has to be considered in the analysis of the following sections. This function is regular in the strip $0 \leqq Re(s) \leqq \theta$. (In this paper, a function is called regular in a closed region, if it is regular in the usual sense and continuous on the boundary.) If $C(t)$ is the distribution function determined by setting $C(t) = 1-B(-t)$ for the continuity points $t \geqq 0$ of $B(t)$, then

$$(b(\theta))^{-1}e^{\theta t}\,dC(t) \doteq dD(t) \tag{2.1}$$

determines a distribution function concentrated on $(-\infty, 0]$. For the convolution $F(t) \doteq A(t)*D(t)$ we obtain

$$\int_{-\infty}^{+\infty} e^{-st}\,dF(t) = (b(\theta))^{-1}\,a(s)b(\theta-s) \tag{2.2}$$

in the mentioned strip. We need in Sections 3 through 5 a representation

$$1-a(s)b(\theta-s) = (1-f^{+}(\theta;s))(1-f^{-}(\theta;s)), \tag{2.3}$$

where $1-f^{+}(\theta;s)$ is regular and non-zero in $Re(s) \geqq 0$ and $1-f^{-}(\theta;s)$ is regular and non-zero in $Re(s) \leqq \theta$. Writing $1-a(s)b(\theta-s)$ as $1-b(\theta)a(s)b(\theta-s)/b(\theta)$ reveals the possibility of applying known results from the fluctuation theory of random walks. A brief description of the required results follows.

Let $X_1, X_2, \cdots$ be independent random variables with a common distribution $F(t)$ and define $S_0 = 0$, $S_1 = X_1$, $S_2 = X_1+X_2, \cdots$. If N^{+} and N^{-} are the epochs of first entry of $\{S_n\}, n > 0$, into $(0, \infty)$ and $(-\infty, 0]$ respectively, then for the characteristic function $\phi(\zeta)$ of $F(t)$ the relation $1-z\phi(\zeta) = (1-E[z^{N^+}\exp(i\zeta S_{N^+})])$ $(1-E[z^{N^-}\exp(i\zeta S_{N^-})])$ holds for $|z| \leqq 1$. E stands for expectation. For $|z| < 1$

$$\log[1-E[z^{N^+}\exp(i\zeta S_{N^+})]]^{-1} = \sum_{n=1}^{\infty} n^{-1}z^n\int_{0+}^{\infty} e^{i\zeta t}\,dF^{n*}(t),$$
$$\log[1-E[z^{N^-}\exp(i\zeta S_{N^-})]]^{-1} = \sum_{n=1}^{\infty} n^{-1}z^n\int_{-\infty}^{0+} e^{i\zeta t}\,dF^{n*}(t),$$

(Feller [2] page 569).

Observing that for $Re(s) \leqq \theta$ and $F(t)$ as in (2.2) $\int_{-\infty}^{0+} e^{-st}\,dF^{n*}(t)$ converges by influence of the factor $e^{\theta t}$ in (2.1), and that $\int_{0+}^{\infty} e^{-st}\,dF^{n*}(t)$ converges for $Re(s) \geqq 0$, we conclude that, for $|z| < 1$, the functions $1-E[z^{N^+}e^{-sS_{N^+}}]$ and $1-E[z^{N^-}e^{-sS_{N^-}}]$ are regular and non-zero in s for $Re(s) \geqq 0$ and $Re(s) \leqq \theta$, respectively, and that the relation $1-zE[e^{-sX}] = (1-E[z^{N^+}e^{-sS_{N^+}}])(1-E[z^{N^-}e^{-sS_{N^-}}])$ holds for $|z| \leqq 1$ and $0 \leqq Re(s) \leqq \theta$. Here $E[e^{-sX}] = (b(\theta))^{-1}a(s)b(\theta-s)$. Putting $z = b(\theta)$ we obtain

$$1-a(s)b(\theta-s) = \exp\Big(-\sum_{n=1}^{\infty} n^{-1}(b(\theta))^n\int_{0+}^{\infty} e^{-st}\,dF^{n*}(t)\Big)$$
$$\cdot\exp\Big(-\sum_{n=1}^{\infty} n^{-1}(b(\theta))^n\int_{-\infty}^{0+} e^{-st}\,dF^{n*}(t)\Big).$$

The first factor on the right-hand side will be denoted in the following sections by $1-f^{+}(\theta;s)$, the other one by $1-f^{-}(\theta;s)$.

3. The approximating system. We consider the queuing system $GI/G/1$ under the restriction, that the distribution functions are of the types

$$A(t) = \sum_{l=1}^{\infty} a_l E_{\lambda}^{l*}(t) \tag{3.1}$$

for the interarrival times and

$$B(t) = \sum_{k=0}^{\infty} b_k E_{\mu}^{k*}(t) \tag{3.2}$$

for the service times, where $\{a_l\}$, $\{b_k\}$ are probability distributions and $E_{\lambda}^{l*}(t)$ is the l-fold convolution of the distribution function $1-e^{-\lambda t}$. If the service for a unit is performed with probability b_k in k consecutive independent stages with common exponential distribution of mean μ^{-1}, then the service time distribution function is of type (3.2). If a unit starts its arrival procedure immediately after the arrival of the preceding unit, and if this procedure lasts with probability a_l for l consecutive independent time phases with common exponential distribution of mean λ^{-1}, then the interarrival time distribution function is of type (3.1). We may take the point of view, that a unit at its time of arrival creates k stages and l phases, with probabilities b_k and a_l, respectively.

Let $k(t) \geqq 0$ be the number of stages waiting at time $t \geqq 0$, including the one under performance. Let $l(t) \geqq 1$ be the number of phases the unit to arrive next has yet to go through, including the one it is in at time t. Then $\{z(t) = (k(t), l(t)), t \geqq 0\}$ is a Markov process with state space $E = \{(k, l) \mid k = 0, 1, 2, \cdots; l = 1, 2, 3, \cdots\}$ and stationary transition probabilities $p_{vl}^{uk}(t) = \Pr\{z(t) = (k, l) \mid z(0) = (u, v)\}$, $(u, v), (k, l) \in E$. These probabilities satisfy the system

$$\begin{aligned}
(d/dt)p_{vl}^{u0}(t) &= -\lambda p_{vl}^{u0}(t) + \mu p_{vl}^{u1}(t) + \lambda p_{v,l+1}^{u,0}(t) + \lambda p_{v1}^{u0}(t) a_l b_0 \\
(d/dt)p_{vl}^{uk}(t) &= -(\lambda+\mu)p_{vl}^{uk}(t) + \mu p_{v,l}^{u,k+1}(t) + \lambda p_{v,l+1}^{u,k}(t) \\
&\qquad + \lambda \textstyle\sum_{i=0}^{k} p_{v,1}^{u,i}(t) a_l b_{k-i}, \qquad k \geqq 0, \\
p_{vl}^{uk}(0) = \delta_{vl}^{uk} &= 0 \quad \text{for } (u, v) \neq (k, l), \\
&= 1 \quad \text{for } (u, v) = (k, l).
\end{aligned} \tag{3.3}$$

This system can be derived in the usual manner. If we write it as $(d/dt)P(t) = P(t)Q$, where $P(t) = \{p_{vl}^{uk}(t)\}$ is the transition matrix, then the matrix Q thus defined is conservative and bounded, whence it follows that $P(t)$ is standard (terminology of Chung [1]).

Taking Laplace transforms

$$\hat{P}_{vl}^{uk}(\theta) = \int_0^{\infty} e^{-\theta t} p_{vl}^{uk}(t)\, dt, \qquad \theta > 0, \tag{3.4}$$

we derive from (3.3) for

$$\hat{L}_v^u(\theta; s, w) \doteq \sum_{k=0}^{\infty} \sum_{l=1}^{\infty} \hat{P}_{vl}^{uk}(\theta)(1+s/\lambda)^{-l}(1+w/\mu)^{-k} \tag{3.5}$$

the relation

$$\begin{aligned}
(\theta - w - s)\hat{L}_v^u(\theta; s, w) &= (1+s/\lambda)^{-v}(1+w/\mu)^{-u} - w\textstyle\sum_{l=1}^{\infty}(1+s/\lambda)^{-l}\hat{p}_{vl}^{u0}(\theta) \\
&\quad - \lambda(1-a(s)b(w))\textstyle\sum_{k=0}^{\infty}(1+w/\mu)^{-k}\hat{p}_{v1}^{uk}(\theta),
\end{aligned} \tag{3.6}$$

$\theta > 0$, $\operatorname{Re}(s) \geqq 0$, $\operatorname{Re}(w) \geqq 0$, where $a(s)$, $b(w)$ are the Laplace-Stieltjes transforms of $A(t)$, $B(t)$, respectively. $\hat{L}_v^u(\theta; s, w)$ is obtained in terms of known expressions in

Section 4. If $L_v^u(t; y, x)$ denotes the probability, that at time t the waiting time for the next arrival is at most y and the time needed to finish the service for all present units is at most x, then

$$\hat{L}_v^u(\theta; s, w) = \int_{t=0}^{\infty}\int_{y=0}^{\infty}\int_{x=0-}^{\infty} \exp(-\theta t - sy - wx) L_v^u(t; dy, dx)\, dt.$$

Hence $\hat{L}_v^u(\theta; s, w)$ completely determines the desired $L_v^u(t; y, x)$. This inversion problem, however, is not a topic of this paper.

4. Calculation of $\hat{L}_v^u(\theta; s, w)$. The functions

$$\hat{U}_v^u(\theta; s) = \sum_{l=1}^{\infty} (1+s/\lambda)^{-l} \hat{p}_{vl}^{u0}(\theta) \tag{4.1}$$

and

$$\hat{V}_v^u(\theta; w) = \sum_{k=0}^{\infty} (1+w/\mu)^{-k} \hat{p}_{v1}^{uk}(\theta) \tag{4.2}$$

have to be determined in order to obtain $\hat{L}_v^u(\theta; s, w)$ from (3.6). $\hat{U}_v^u(\theta; s)$ is of independent interest, because

$$\hat{U}_v^u(\theta; s) = \int_{t=0}^{\infty}\int_{y=0}^{\infty} e^{-\theta t - sy} L_v^u(t; dy, 0)\, dt. \tag{4.3}$$

Putting now $w = \theta - s$, we derive from (3.6), using the regularity properties of $\hat{L}_v^u(\theta; s, w)$, the relation

$$0 = (1+s/\lambda)^{-v}(1+(\theta-s)/\mu)^{-u} - (\theta-s)\hat{U}_v^u(\theta; s) - \lambda(1-a(s)b(\theta-s))\hat{V}_v^u(\theta; \theta-s)$$

for $0 \leqq \operatorname{Re}(s) \leqq \theta$. The results of Section 2 allow this to be rewritten as

$$\begin{aligned}(1+s/\lambda)^{-v}[1-f^+(\theta; s)]^{-1} - (1+(\theta-s)/\mu)^u(\theta-s)[1-f^+(\theta; s)]^{-1}\hat{U}_v^u(\theta; s)\\ = \lambda(1-f^-(\theta; s))(1+(\theta-s)/\mu)^u \hat{V}_v^u(\theta; \theta-s)\end{aligned} \tag{4.4}$$

for $0 \leqq \operatorname{Re}(s) \leqq \theta$. The left-hand side of (4.4) is regular in $\operatorname{Re}(s) \geqq 0$, whereas the right-hand side is regular in $\operatorname{Re}(s) \leqq \theta$. Since there is a common strip to both sides, they are but different representations of the same function $\hat{P}_v^u(\theta; s)$, regular for all finite s. Moreover, examination of both sides yields that $\lim_{|s|\to\infty}(\hat{P}_v^u(\theta; s)/s^{u+1}) = 0$. Thus, by the theorem of Liouville-Hadamard, $\hat{P}_v^u(\theta; s)$ is a polynomial of degree less than or equal to u. The knowledge of its coefficients would imply the knowledge of $\hat{U}_v^u(\theta; s)$ and $\hat{V}_v^u(\theta; \theta-s)$ and hence of $\hat{V}_v^u(\theta; w)$. The left-hand side of (4.4) yields the relation

$$\begin{aligned}(1+(\theta-s)/\mu)^u(\theta-s)[1-f^+(\theta; s)]^{-1}\hat{U}_v^u(\theta; s) = (1+s/\lambda)^{-v}[1-f^+(\theta; s)]^{-1}\\ -\hat{P}_v^u(\theta; s)\end{aligned} \tag{4.5}$$

for $\operatorname{Re}(s) \geqq 0$. The regularity of $\hat{U}_v^u(\theta; s)$ in $\operatorname{Re}(s) \geqq 0$ requires the right-hand side of this relation to have a zero at $s = \theta$ and another one of degree u at $s = \theta + \mu$. These conditions determine the polynomial $\hat{P}_v^u(\theta; s)$ completely.

In addition to (4.5) we have from (3.6)

$$\lambda(1-f^-(\theta; s))(1+(\theta-s)/\mu)^u \hat{V}_v^u(\theta; \theta-s) = \hat{P}_v^u(\theta; s) \tag{4.6}$$

for $\operatorname{Re}(s) \leqq \theta$, or, after replacing $\theta - s$ by w,

$$\hat{V}_v^u(\theta; w) = \lambda^{-1}[1-f^-(\theta; \theta-w)]^{-1}(1+w/\mu)^{-u}\hat{P}_v^u(\theta; \theta-w) \tag{4.7}$$

for $\mathrm{Re}(w) \geqq 0$. Substitution of (4.5) and (4.7) into (3.6) results in a relation between $\hat{L}_v^u(\theta; s, w)$ and known expressions. The actual calculation of $\hat{P}_v^u(\theta; s)$ is unpleasant and does not seem, in view of the generalization in mind, to lead to simple expressions. Fortunately, however, the case $u = 0$, corresponding to a system starting without a unit in service, leads further. For $u = 0$, the polynomial $\hat{P}_v^0(\theta; s)$ degenerates to the constant $(1+\theta/\lambda)^{-v}[1-f^+(\theta;\theta)]^{-1}$, which can be readily seen from (4.5). We obtain

$$(4.8)\quad (\theta-s)\hat{U}_v^0(\theta;s) = (1+s/\lambda)^{-v}-(1+\theta/\lambda)^{-v}[1-f^+(\theta;\theta)]^{-1}[1-f^+(\theta;s)] \quad \text{and}$$

$$(4.9)\quad \hat{V}_v^0(\theta;w) = \lambda^{-1}[1-f^-(\theta;\theta-w)]^{-1}(1+\theta/\lambda)^{-v}[1-f^+(\theta;\theta)]^{-1}$$

With these relations, (3.6) becomes

$$(4.10)\quad \begin{aligned}(\theta-s)(\theta-s-w)\hat{L}_v^0(\theta;s,w) &= \frac{\theta-s-w}{(1+s/\lambda)^v}+\frac{1}{(1+\theta/\lambda)^v}\cdot\frac{1}{1-f^+(\theta;\theta)}\\ &\cdot\left[w\frac{1-a(s)b(\theta-s)}{1-f^-(\theta;s)}-(\theta-s)\frac{1-a(s)b(w)}{1-f^-(\theta;\theta-w)}\right].\end{aligned}$$

Finally, for this section, we establish the transforms for the distribution of principal interest, that is the distribution of the virtual waiting time. These transforms are obtained by putting $s = 0$. Thus

$$(4.11)\quad \begin{aligned}\theta(\theta-w)\hat{L}_v^0(\theta;0,w) &= \theta-w+\frac{1}{(1+\theta/\lambda)^v}\cdot\frac{1}{1-f^+(\theta;\theta)}\\ &\cdot\left[w\frac{1-b(\theta)}{1-f^-(\theta;0)}-\theta\frac{1-b(w)}{1-f^-(\theta;\theta-w)}\right],\end{aligned}$$

where $\hat{L}_v^0(\theta;0,w) = \int_{t=0}^{\infty}\int_{x=0-}^{\infty} e^{-\theta t-wx} L_v^0(t;\infty,dx)\,dt$, and

$$(4.12)\quad \hat{U}_v^0(\theta;0) = \frac{1}{\theta}-\frac{1}{\theta}\frac{1}{(1+\theta/\lambda)^v}\frac{1-f^+(\theta;0)}{1-f^+(\theta;\theta)}$$

where $\hat{U}_v^0(\theta;0) = \int_{t=0}^{\infty} e^{-\theta t} L_v^0(t;\infty,0)\,dt$.

5. The unrestricted system. The unrestricted system $GI/G/1$ allows any distribution function $A(t)$ concentrated on $(0, \infty)$ and any $B(t)$ concentrated on $[0, \infty)$. For

$$A_\lambda(t) \doteq \textstyle\sum_{l=1}^{\infty}(A(l/\lambda)-A((l-1)/\lambda))\,E_\lambda^{l*}(t) \quad \text{and}$$

$$B_\mu(t) \doteq B(0)+\textstyle\sum_{k=1}^{\infty}(B(k/\mu)-B((k-1)/\mu))\,E_\mu^{k*}(t)$$

we have, by the theorem mentioned in the introduction, $\lim_{\lambda\to\infty} A_\lambda(t) = \frac{1}{2}(A(t-)+A(t+))$ and $\lim_{\mu\to\infty} B_\mu(t) = \frac{1}{2}(B(t-)+B(t+))$. Clearly $A_\lambda(t)$ and $B_\mu(t)$ are distribution functions of the types (3.1) and (3.2), respectively. This implies weak convergence and hence $\lim_{\lambda\to\infty} a_\lambda(s) = a(s)$, $\lim_{\mu\to\infty} b_\mu(s) = b(s)$. It is not hard to see that the factors of $1-a_\lambda(s)b_\mu(\theta-s)$ (as defined in (2.3)) tend to the corresponding factors of $1-a(s)b(\theta-s)$. We intend now to carry the results (4.10), (4.11), and (4.12), obtained for the approximating $A_\lambda(t)$ and $B_\mu(t)$, over for $A(t)$, $B(t)$ and interpret them as results for the unrestricted system. This requires taking care of the initial condition v, which has in this form no meaning in the unrestricted system. v is the number of initially outstanding arrival phases. Each phase has an expected length λ^{-1} and variance λ^{-2}. As λ tends to infinity, we have to increase v in order to maintain a positive expected length of delay before the first arrival. Keeping v/λ constant thus means keeping the expected value of the time before the first arrival constant. The variance of this random variable however tends to 0 and the results obtained by letting $\lambda \to \infty$, $v/\lambda = y_0$, are to be interpreted therefore as results for $GI/G/1$ starting at $t = 0$ under the virtual waiting time 0 and a delay y_0 of the first arrival. Transition to the limits in (4.10), (4.11), and (4.12) now yields

$$(\theta - s)(\theta - s - w)\hat{L}^0_{y_0}(\theta; s, w) = (\theta - s - w)\, e^{-sy_0} + e^{-\theta y_0} \frac{1}{1 - f^+(\theta;\theta)} \cdot \left[w \frac{1 - a(s)b(\theta - s)}{1 - f^-(\theta; s)} - (\theta - s) \frac{1 - a(s)b(w)}{1 - f^-(\theta; \theta - w)} \right]$$

$$\theta(\theta - w)\hat{L}^0_{y_0}(\theta; 0, w) = (\theta - w) + e^{-\theta y_0} \frac{1}{1 - f^+(\theta;\theta)} \left[w \frac{1 - b(\theta)}{1 - f^-(\theta; 0)} - \theta \frac{1 - b(w)}{1 - f^-(\theta; \theta - w)} \right]$$

and $\theta \hat{U}^0_{y_0}(\theta; 0) = 1 - e^{-\theta y_0}[1 - f^+(\theta; 0)]/[1 - f^+(\theta;\theta)]$.

As has been mentioned in the introduction, a rigorous analysis justifying this limiting procedure is omitted.

REFERENCES

[1] Chung, K. L. (1960). *Markov Chains with Stationary Transition Probabilities*. Springer' Berlin.

[2] Feller, W. (1966). *An Introduction to Probability Theory and its Applications*, **2**. Wiley, New York.

[3] Keilson, J. and Kooharian, A. (1962). On the general time dependent queue with a single server. *Ann. Math. Statist.* **33** 767–791.

[4] Schassberger, R. (1968). A note on the method of stages. Research Paper No. 53, Department of Mathematics, Univ. of Calgary.

[5] Takács, L. (1963). The limiting distribution of the virtual waiting time and the queue size for a single-server queue with recurrent input and general servicc times. *Sankhyā, Ser. A* **25** 91–100.