ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XVIII, 3 (1984), p. 343-350

R. SZEKLI (Wrocław)

LITTLE'S FORMULA FOR WORK-CONSERVING NORMAL $G/G/1$ QUEUES IN SERIES

1. Introduction. In this paper we deal with a single server queue which operates according to a work-conserving normal discipline. In Section 3 we give a definition of this discipline convenient for further considerations. We say that in the queue which operates according to a work-conserving normal discipline the knowledge of the generic sequence in a busy cycle is sufficient to determine the waiting times of customers in this busy cycle. The well-known disciplines as first-in-first-out, last-in-first-out, shortest remaining processing time, round robin, instantaneous feed-back are work-conserving normal ones (for definitions see, e.g., [2], [5]).

The relations between the time average of the number of customers L in the queue, the arrival intensity λ^{-1}, and the average sojourn time v in the queue, of the form $L = \lambda v$, have been investigated for single stable queues with various queue disciplines. This kind of relations is called *Little's formula.*

Our aim is to prove Little's formula for single server queues in series which operate according to work-conserving normal disciplines.

Following Stidham [6] we use the sample path approach to the problem. The methods used are simpler than those given in the papers which use the theory of point processes.

2. Assumptions and notation. We use the following notation:
$\{t_n, n \geqslant 1\}$ — sequence of arrival moments;
$\{t'_n, n \geqslant 1\}$ — sequence of departure moments;
$\{w_n, n \geqslant 1\}$ — actual waiting time process;
S_n — service time of the n-th customer;
$v_n = S_n + w_n$ — sojourn time of the n-th customer;
$T_n = t_{n+1} - t_n$ — time between the n-th and $(n+1)$-st arrival;
$T'_n = t'_{n+1} - t'_n$ — time between the n-th and $(n+1)$-st departure;
$\{L(t), t \geqslant 0\}$ — queue size process;

$L = \lim_{t\to\infty} t^{-1} \int_0^t L(s)ds$ — time average of the number of customers in the queue;

$\lambda^{-1} = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} T_i$ — arrival rate;

$v = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} v_i$ — average sojourn time in the queue.

We assume that both S_n and T_n are non-negative and finite. Then in any realization the values of S_n and T_n for all n, together with the initial conditions, determine completely the development of the queue. The basic assumption is that $\{(T_n, S_n), n \geqslant 1\}$ forms a strictly stationary ergodic sequence.

According to Breiman [1] the stationary process $\{(T_n, S_n), n \geqslant 1\}$ may be extended to a stationary process $\{(T_n, S_n), -\infty < n < \infty\}$ which is again ergodic. Suppose this has been done. Then $\{(T_n, S_n), n \geqslant 1\}$ is called a *generic sequence*.

From now on, all random variables are considered to be functions on a single fixed underlying probability space Ω.

The assumption that the sequence is ergodic is not necessary, but allows us to avoid irrelevant details in the proofs and undue complication in the statements of the theorems. It seems in any case a reasonable assumption in practical situations. When the sequence $\{(T_n, S_n), n \geqslant 1\}$ is not ergodic, the limit

$$v = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} v_i,$$

given by the ergodic theorem, is no longer constant but a random variable $\mathrm{E}(v_1 | I_v)$, where I_v is the σ-field of invariant sets.

We use the well-known property that busy cycles are the same for two queues with a common generic sequence, the first operating according to a first-in-first-out discipline and the second according to a work-conserving discipline.

We assume that at the beginning the queues are empty. A queue which is not empty works.

Denote by $\{\tilde{w}_n, n \geqslant 1\}$ and $\{L^*(t), t \geqslant 0\}$ the stationary waiting time and the stationary queue size process, respectively.

We write WCN queue and FIFO queue for the queue which operates according to work-conserving normal and first-in-first-out disciplines, respectively.

3. Preliminaries. In this section we recall some properties of FIFO queues. Then for an arbitrary WCN queue we construct a FIFO queue which has the same arrival and departure moments. In the next section this queue will be used to show the validity of Little's formula for the WCN queue.

For FIFO queues we have the following lemma:

LEMMA 1 (Loynes [3]). *Let the random variables* w_n $(n \geqslant 1)$ *be related by the transform*

$$w_{n+1} = \max(0, w_n + S_n - T_n), \tag{1}$$

where $\{(T_n, S_n), n \geqslant 1\}$ *is a strictly stationary ergodic sequence for which* $\mathrm{E}S_1 < \mathrm{E}T_1$ *and* $w_1 = 0$. *Then there exists just one almost everywhere (a.e.) finite strictly stationary ergodic sequence of random variables* $\{\tilde{w}_n, -\infty < n < \infty\}$ *such that*

$$\tilde{w}_{n+1} = \max(0, \tilde{w}_n + S_n - T_n).$$

Furthermore, $\tilde{w}_n$ *is the minimal sequence satisfying* (1) *for all* n, w_n *is such for* $n \geqslant 1$ *in the sense that if* $\{x_n\}$ *is another such sequence, then* $x_n \geqslant \tilde{w}_n \geqslant w_n$.

We write

$$a_1 = \inf\{i \geqslant 1 : \tilde{w}_i = 0\} \quad \text{and} \quad a_{n+1} = \inf\{i > a_n : \tilde{w}_i = 0\}.$$

The sequence $\{a_n, n \geqslant 1\}$ denotes consecutive indices of customers arriving at the empty queue. Note that under the assumptions of Lemma 1 the random variables a_n are finite a.e. From Lemma 1 it also follows that $\tilde{w}_i = w_i$ a.e. for $i \geqslant a_1$ (see [3]).

Let

$$C_n = t_{a_{n+1}} - t_{a_n}, \quad k_n = a_{n+1} - a_n,$$

$$N_n = \begin{cases} 0 & \text{for } \tilde{w}_n \neq 0, \\ 1 & \text{for } \tilde{w}_n = 0. \end{cases}$$

The random variable C_n is called the *n-th busy cycle.*

Suppose that the random variables a_n are finite a.e. for all n. Since both the work-conserving and FIFO queues have the same busy cycles, provided they have the same generic sequence, we can define a work-conserving normal discipline.

Definition. A work-conserving queue discipline is called *normal* if there exists a sequence of real Borel functions $\{\varphi_n, n \geqslant 1\}$,

$$\varphi_n : (\boldsymbol{R}_+ \times \boldsymbol{R}_+)^n \to \boldsymbol{R}_+^n,$$

for which

$$(w_{a_n}, \ldots, w_{a_{n+1}-1}) = \varphi_{k_n}\big((T_{a_n}, S_{a_n}), \ldots, (T_{a_{n+1}-1}, S_{a_{n+1}-1})\big) \tag{2}$$

a.e. for all n.

The class of WCN queue disciplines was introduced by Rolski [4]. We recall a property of a WCN queue proved by him.

LEMMA 2 (Rolski [4]). *Let the random variables* w_n *be related by* (2) *for* $n \geq 1$, *where* $\{(T_n, S_n), -\infty < n < \infty\}$ *is a strictly stationary ergodic sequence for which* $\mathrm{E}S_1 < \mathrm{E}T_1$. *Then there exists a strictly stationary ergodic sequence of random variables* $\{\tilde{w}_n, -\infty < n < \infty\}$ *satisfying* (2) *for all* n *and such that the sequence* $\{(T_n, S_n, \tilde{w}_n), -\infty < n < \infty\}$ *is strictly stationary ergodic.*

Rolski [4] proved also that there exists a strictly stationary WCN queue size process $\{L^*(t), t \in \boldsymbol{R}\}$. Consequently, note that for

$$L_1 = \inf\{t > 0 : L^*(t) = 0\}, \qquad a_1 = \inf\Big\{i \geq 1 : \sum_{k=1}^{i-1} T_k > L_1\Big\},$$

$$L_{n+1} = \inf\{t > t_{a_n} : L^*(t) = 0\}$$

we have

$$a_{n+1} = \inf\Big\{i > a_n : \sum_{k=1}^{i-1} T_k > L_{n+1}\Big\}.$$

Consider an arbitrary WCN queue. By the following lemma we can construct an FIFO queue which has the same consecutive arriving and departure moments.

It will be convenient to use the symbols w_i^{F} and S_i^{F} for the waiting time and the service time, respectively, in this FIFO queue.

LEMMA 3. *Let* $\{(T_n, S_n), -\infty < n < \infty\}$ *be a strictly stationary ergodic sequence for which* $\mathrm{E}S_1 < \mathrm{E}T_1$. *Then there exists a strictly stationary ergodic sequence* $\{S_n^{\mathrm{F}}, -\infty < n < \infty\}$ *such that consecutive departure moments from the FIFO queue with the generic sequence* $\{(T_n, S_n^{\mathrm{F}}), -\infty < n < \infty\}$ *are equal to those from the WCN queue with the generic sequence* $\{(T_n, S_n), -\infty < n < \infty\}$. *Furthermore,*

$$\lim_{n\to\infty} n^{-1} w_n^{\mathrm{F}} = 0 \ \textit{a.e.} \quad \textit{and} \quad w_n^{\mathrm{F}} = \tilde{w}_n^{\mathrm{F}} \quad \textit{for } n \geq a_1 \ \textit{a.e.}$$

Proof. For a fixed $\omega \in \Omega$ such that the sequence $\{w_m(\omega)\}$ satisfies (2) for all m and $n \in \{a_i, \ldots, a_{i+1}-1\}$, where $i \geq 1$, there exist indices $i_1, i_2 \in \{a_i, \ldots, a_{i+1}-1\}$ such that

$$T_n' = (t_{i_1} + v_{i_1}) - (t_{i_2} + v_{i_2}) = \begin{cases} -\sum\limits_{j=i_1+1}^{i_2} T_j + v_{i_1} - v_{i_2} & \text{for } i_2 > i_1, \\ \sum\limits_{j=i_2+1}^{i_1} T_j + v_{i_1} - v_{i_2} & \text{for } i_1 \geq i_2. \end{cases}$$

From the equalities

$$v_n = w_n + S_n,$$

$$(w_{a_i}, \ldots, w_{a_{i+1}-1}) = \varphi_{k_i}\big((T_{a_i}, S_{a_i}), \ldots, (T_{a_{i+1}-1}, S_{a_{i+1}-1})\big)$$

we infer that there exists a measurable mapping f_{k_i} for which

(3) $$(T'_{a_i}, \ldots, T'_{a_{i+1}-1}) = f_{k_i}\big((T_{a_i}, S_{a_i}), \ldots, (T_{a_{i+1}-1}, S_{a_{i+1}-1})\big).$$

From Lemma 2 it follows that there exists a strictly stationary ergodic sequence $\{\tilde{T}'_n, -\infty < n < \infty\}$ satisfying (3) a.e. for all i. Note that $\tilde{T}'_n = T'_n$ for $n \geqslant a_1$.

For all i put

(4) $$(S^F_{a_i}, \ldots, S^F_{a_{i+1}-1}) = (v^F_{a_i}, \tilde{T}'_{a_i+1}, \ldots, \tilde{T}'_{a_{i+1}-1}),$$

where

$$v^F_{a_i} = w_k + S_k + \sum_{j=a_i+1}^{k} T_j,$$

and k denotes the index of the customer which first completed the service in the i-th busy cycle.

From Lemma 2 it follows that there exists a strictly stationary ergodic sequence $\{S^F_i, -\infty < i < \infty\}$ satisfying (4) for all i. Note that the sequence $\{(T_i, S^F_i), -\infty < i < \infty\}$ is strictly stationary ergodic.

Let us put $v^F_n = t'_n - t_n$. For $n \in \{a_i, \ldots, a_{i+1}-1\}$ there exists an index $i_1 \in \{a_i, \ldots, a_{i+1}-1\}$ such that

$$v^F_n = \begin{cases} v_{i_1} + \sum_{j=n}^{i_1} T_j & \text{for } n \leqslant i_1, \\ v_{i_1} - \sum_{j=i_1+1}^{n} T_j & \text{for } n > i_1. \end{cases}$$

The same argument as before shows that there exists a strictly stationary ergodic sequence $\{\tilde{v}^F_n, -\infty < n < \infty\}$ such that $v^F_n = \tilde{v}^F_n$ for $n \geqslant a_1$.

Set $\tilde{w}^F_n = \tilde{v}^F_n - S^F_n$. Then the sequence $\{\tilde{w}^F_n, -\infty < n < \infty\}$ is strictly stationary ergodic.

Note that $\mathrm{E}S^F_1 \leqslant \mathrm{E}T_1$. For a contradiction, assume that $\mathrm{E}S^F_1 > \mathrm{E}T_1$. Then either $\mathrm{E}S^F_1 < \infty$, which yields

$$\lim_{n\to\infty} n^{-1} w^F_n = \max(0, \mathrm{E}S^F_1 - \mathrm{E}T_1) > 0 \text{ a.e.},$$

or $\mathrm{E}S^F_1 = \infty$, which yields

$$\lim_{n\to\infty} n^{-1} w^F_n = \infty \text{ a.e.}$$

Both cases contradict the fact that $w^F_n = 0$ infinitely often a.e. (a_n are finite a.e. for all n).

4. The single server queue. We consider now a single server WCN queue. In the following lemma we prove that the average sojourn time in the WCN queue equals that in the FIFO queue constructed in Section 3.

LEMMA 4. *Assume that the WCN queue has the generic sequence* $\{(T_n, S_n), -\infty < n < \infty\}$ *strictly stationary ergodic and such that* $ES_1 < ET_1$. *Then*

$$\lim_{n\to\infty} n^{-1}\sum_{i=1}^{n} v_i^{\mathrm{F}} = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n} v_i \ \textit{a.e.}$$

Proof. Write $D_i = \{a_i, \ldots, a_{i+1}-1\}$. First we prove the equality (for a fixed $\omega \in \Omega$ as in Lemma 3)

$$\sum_{j\in D_i} v_j^{\mathrm{F}} = \sum_{j\in D_i} v_j.$$

Assume that the customer which departed as the n-th in the i-th busy cycle is the $\pi_n(i)$-th arrived. Note that

$$\pi(i) = \big(\pi_1(i), \ldots, \pi_{k_i}(i)\big)$$

is a permutation of the set D_i.

Moreover, we have

$$v_j^{\mathrm{F}} = \begin{cases} v_{\pi_j(i)} - T_{\pi_j(i),j} & \text{for } j > \pi_j(i), \\ v_{\pi_j(i)} + T_{j,\pi_j(i)} & \text{for } \pi_j(i) > j, \end{cases} \tag{5}$$

where $j \in D_i$, $T_{i,j} = \sum_{m=i+1}^{j} T_m$.

We can write $\pi(i)$ as a finite composition of transpositions:

$$\pi(i) = p_1 \circ \ldots \circ p_m.$$

In the next part of the proof we use induction. Assume that $\pi(i) = p_1$. Then there exist indices $i_1, i_2 \in D_i$ such that $i_1 \leqslant i_2$ and $\pi_{i_1}(i) = i_2$, $\pi_{i_2}(i) = i_1$. From (5) we have

$$v_{i_1}^{\mathrm{F}} = v_{i_2} + T_{i_1,i_2}, \quad v_{i_2}^{\mathrm{F}} = v_{i_1} - T_{i_1,i_2}, \quad v_j^{\mathrm{F}} = v_j,\ j \notin \{i_1, i_2\}.$$

Hence

$$\sum_{j\in D_i} v_j^{\mathrm{F}} = \sum_{j\in D_i} v_j.$$

Assume that $\pi(i) = p_1 \circ \ldots \circ p_m \circ p_{m+1} = \hat{\pi}(i) \circ p_{m+1}$. Denote by $\hat{v}_i$ the quantities obtained from (5) after changing $\pi(i)$ into $\hat{\pi}(i)$. By the inductive assumption we have

$$\sum_{j \in D_i} \hat{v}_j = \sum_{j \in D_i} v_j.$$

There exist indices $i_1, i_2 \in D_i$ such that $i_1 < i_2$ and

$$v^{\mathrm{F}}_{i_1} = \hat{v}_{i_2} + T_{i_1, i_2}, \quad v^{\mathrm{F}}_{i_2} = \hat{v}_{i_1} - T_{i_1, i_2}, \quad v^{\mathrm{F}}_j = \hat{v}_j, \; j \notin \{i_1, i_2\}.$$

Hence

$$\sum_{j \in D_i} v^{\mathrm{F}}_j = \sum_{j \in D_i} \hat{v}_j \quad \text{and} \quad \sum_{j \in D_i} v^{\mathrm{F}}_j = \sum_{j \in D_i} v_j.$$

Using the same argument for all i we have

$$a_n^{-1} \sum_{i=a_1}^{a_n} v^{\mathrm{F}}_i = a_n^{-1} \sum_{i=a_1}^{a_n} v_i \quad \text{for all } n.$$

Since the equalities $v_n = \tilde{v}_n$ and $v^{\mathrm{F}}_n = \tilde{v}^{\mathrm{F}}_n$ hold a.e. for $n \geqslant a_1$ and since the sequences $\{\tilde{v}_n, -\infty < n < \infty\}$, $\{\tilde{v}^{\mathrm{F}}_n, -\infty < n < \infty\}$ are strictly stationary ergodic, both limits in Lemma 4 exist a.e. The limits are equal because they coincide on the subsequence $\{a_n, n \geqslant 1\}$.

THEOREM 1. *Assume that the WCN queue has the generic sequence* $\{(T_n, S_n), -\infty < n < \infty\}$ *strictly stationary ergodic and such that* $\mathrm{E}S_1 < \mathrm{E}T_1$. *Then* $L = \lambda v$ *a.e.*

Proof. By Lemma 3 we have

$$\lim_{n \to \infty} n^{-1} v^{\mathrm{F}}_n = 0$$

and the averages λ and v^{F} do exist. Stidham [6] proved that under these assumptions the average L^{F} exists and $L^{\mathrm{F}} = \lambda v^{\mathrm{F}}$ a.e. Now the equation $L = \lambda v$ a.e. follows from Lemma 4 and the fact that $L^{\mathrm{F}}(t) = L(t)$ for $t \geqslant 0$.

5. Single server queues in series. Let N be a fixed positive integer. We say that N queues are *in series* if the customer passes in turn through these queues spending a waiting time and the service time in any particular one, proceeding to the next succeeding queue immediately when the service time is completed.

Consider N of WCN queues in series. It is convenient to use the superscript i for the random variables describing the i-th queue, e.g., w^i_n, a^i_n, $L^i(t)$.

THEOREM 2. *Let the sequence* $\{(T_n^1, S_n^1, \ldots, S_n^N), -\infty < n < \infty\}$ *be strictly stationary ergodic and such that* $\mathrm{E} S_1^i < \mathrm{E} T_1^i$ *for* $i = 1, \ldots, N$. *Then* $L^i = \lambda v^i$ *a.e. for* $i = 1, \ldots, N$.

Proof. From Theorem 1 we have $L^1 = \lambda v^1$ a.e. By Lemma 3 there exists a strictly stationary ergodic sequence $\{\tilde{T}_n^{1'}, -\infty < n < \infty\}$ such that $\tilde{T}_n^{1'} = T_n^{1'}$ for $n \geqslant a_1^1$. It is clear that $T_n^{1'} = T_n^2$. Hence $T_n^2 = \tilde{T}_n^{1'}$ a.e. for $n \geqslant a_1^1$.

The arrival intensity for the second queue is λ^{-1} because

$$\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} T_i^2 = \lim_{n\to\infty} n^{-1} t_n^1 + \lim_{n\to\infty} n^{-1} v_n^{1\mathrm{F}},$$

and the second limit on the right-hand side equals zero by Lemma 3.

Let $\tilde{T}_n^2 = \tilde{T}_n^{1'}$ for all n. Then for $n \geqslant a_1^1$ we have $\tilde{T}_n^2 = T_n^2$.

The sequence $\{(\tilde{T}_n^2, S_n^2), -\infty < n < \infty\}$ is strictly stationary ergodic. Consider the second queue with this sequence as the generic sequence. Hence the equality $L^2 = \lambda v^2$ a.e. follows from Theorem 1 and the fact that $T_n^2 = \tilde{T}_n^2$ for $n \geqslant a_1^1$ a.e.

The argument can clearly be carried from queue to queue. Thus the proof is complete.

Acknowledgement. This paper is based on a master degree work written at the University of Wrocław, May 1980. The author is grateful to Dr. T. Rolski for his help in the presentation of this work.

References

[1] L. Breiman, *Probability*, Addison-Wesley, New York 1968.

[2] R. Conway, W. Maxwell and L. Miller, *The theory of scheduling*, Addison-Wesley, New York 1967.

[3] R. M. Loynes, *The stability of queues with non-independent inter-arrival and service times*, Proc. Camb. Phil. Soc. 58 (1962), p. 467-520.

[4] T. Rolski, *Stationary random processes associated with point processes*, Lecture Notes in Statistics, Springer Verlag, New York 1981.

[5] L. Miller and L. Schrage, *The M/G/1 queue with SRPT discipline*, Operations Res. 14 (1966), p. 670-684.

[6] S. Stidham, Jr., *A last word on* $L = \lambda w$, ibidem 22 (1974), p. 417-422.

MATHEMATICAL INSTITUTE
UNIVERSITY OF WROCŁAW
50-384 WROCŁAW

Received on 18. 6. 1982