ZOR – Methods and Models of Operations Research (1993) 37:303–331

Queueing Systems on a Circle

DIRK P. KROESE

Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

VOLKER SCHMIDT

Department of Stochastics, University of Ulm, Helmholtzstraße 18, 7900 Ulm, Germany

Abstract: Consider a ring on which customers arrive according to a Poisson process. Arriving customers drop somewhere on the circle and wait there for a server who travels on the ring. Whenever this server encounters a customer, he stops and serves the customer according to an arbitrary service time distribution. After the service is completed, the server removes the client from the circle and resumes his journey.

We are interested in the number and the locations of customers that are waiting for service. These locations are modeled as random counting measures on the circle. Two different types of servers are considered: The polling server and the Brownian (or drunken) server. It is shown that under both server motions the system is stable if the traffic intensity is less than 1. Furthermore, several earlier results on the configuration of waiting customers are extended, by combining results from random measure theory, stochastic integration and renewal theory.

CYCLIC SERVER SYSTEM; POLLING SERVER; BROWNIAN SERVER; RANDOM COUNTING MEASURES; REGENERATIVE PROCESSES; STABILITY; STATIONARY POINT PROCESSES; COUPLING; LIMIT DISTRIBUTIONS; LAPLACE FUNCTIONALS; STOCHASTIC DECOMPOSITION; STOCHASTIC INTEGRATION; ITO'S FORMULA.

AMS 1990 subject classifications: Primary 60K25, 90B22, 60G55; Secondary 60G57, 60H05.

1 Introduction

Recent developments in tele-communication and computer architecture have given rise to several exotic queueing systems. An important sub-class of such non-standard queueing models is formed by the *queueing systems on a circle.* Instead of waiting in a queue, arriving customers now choose positions on a circle. There they wait until they are visited by a server who travels on the ring. When the server is only allowed to travel in a fixed direction, the system is usually referred to as a *polling system.* Over the last few years a wide variety of cyclic service models has found application in tele-communication (e.g. Local Area Networks) and reliability (inspection policies). Usually such models are formulated in discrete time, cf. for example [5], [12] and [14]. However, a

0340-9422/93/3/303-331 $2.50 © 1993 Physica-Verlag, Heidelberg

continuous approach, such as in [2], [3], [8], [10] and [13], often reveals much more of the underlying structure of the model.

Here we consider two such queueing systems on the circle: the polling-server system and the Brownian-server system. In the first system, the server moves uni-directionally at constant speed, whereas the Brownian server carries out a Brownian motion with zero drift on the ring. Customers arrive according to a Poisson process, and are dropped on the circle (relatively to the position of the server) according to an arbitrary diffuse distribution (which is not necessarily uniform). We analyze the configuration of waiting customers on the ring. This is done by representing the waiting customers through a random counting measure. It is proved that both systems are stable, provided that the mean service time is smaller than the mean inter-arrival time. Moreover, results on the configuration of waiting customers obtained in [10] are extended to the polling- and Brownian-server systems. The Laplace functional for the random measure of waiting customers is given for the polling system. For the Brownian-server system we derive the mean measure of waiting customers and the expected number of waiting customers.

In the analysis we use results from stochastic integration theory (see the appendix) and renewal theory (cf. [1]). For basic definitions and results on random counting measures (point processes) we refer to the appendix. Note that we will frequently use the notation μf for the integral of a function f with respect to a (random) measure μ.

2 The Model(s)

In this section we describe the polling-server and Brownian-server systems in more detail. Both systems are based on the following model. Let C be a circle with circumference one. Fix the orientation on C, clockwise, say. A server travels on C, in a way which is specified below, and stops to serve a customer whenever he encounters one on his way. Starting with an empty system, these customers arrive according to a Poisson process with intensity a and drop somewhere on the circle, independently of everything else. Specifically: we assume that the distance from the server to a newly arrived customer, measured along the orientation on C, has a fixed distribution π. For convenience we assume that π is a *diffuse* distribution on $[0, 1]$.

During a service the server does not move. The consecutive service times are i.i.d. random variables with distribution function (d.f.) F and ith moment e_i, $i \in \mathbb{N}$ (here we take $\mathbb{N} = \{1, 2, \dots\}$). When a service has been completed, the customer is removed from the circle, after which the server resumes his journey. We assume that the service times are independent of the arrival process and the locations of the clients on C.

The two models are specified by the following rules that govern the servers motion on the circle, when the server is idle:

Polling Server: In this model the server (when not busy) travels at constant speed α^{-1} in the direction of the orientation on C.

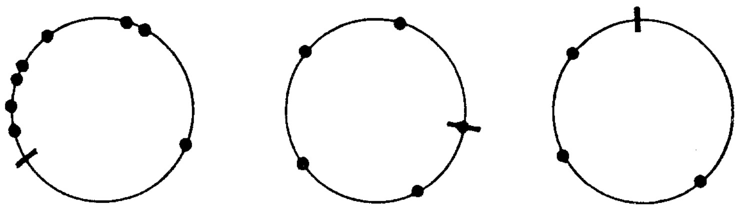
Brownian Server: Here the (idle) server carries out a Brownian motion on the circle, with zero drift and variance parameter σ^2. This server movement is assumed to be completely independent of everything else.

3 Waiting Customers

In fig. 1 a number of possible configurations of waiting customers on the ring is depicted. In both models above we are interested in describing this "distribution" of waiting customers probabilistically.

A first observation to make is that the actual position of the server is not very relevant to the analysis. It is the configuration of customers *relatively* to the position of the server that is important. We therefore analyze the system from the point of view of the server. For every time t we identify the circle C with the interval $[0, 1]$ in the obvious way, such that both 0 and 1 are identified with the position of the server. That is, for every $t \geq 0$ we cut the circle at the current position of the server and stretch it onto the interval $[0, 1]$, hereby preserving the orientation from 0 to 1.

From the point of view of the server, the customer paths form a stochastic flow on $\mathbb{R}_+ \times [0, 1]$, as is shown in fig. 2 (the paths depicted in fig. 2 correspond to the Brownian-server model). Notice that from the perspective of the server, the positions of *arriving* customers form a Poisson random measure on $\mathbb{R}_+ \times [0, 1]$ with mean measure $aLeb(\mathbb{R}_+) \times \pi$ (see appendix). It will be convenient to de-



(a) (b) (c)

Fig. 1. A number of possible customer configurations on the circle. The waiting customers are depicted by dots (●). The position of the server is given by the symbol "|". Note that in (b) a customer is being served.

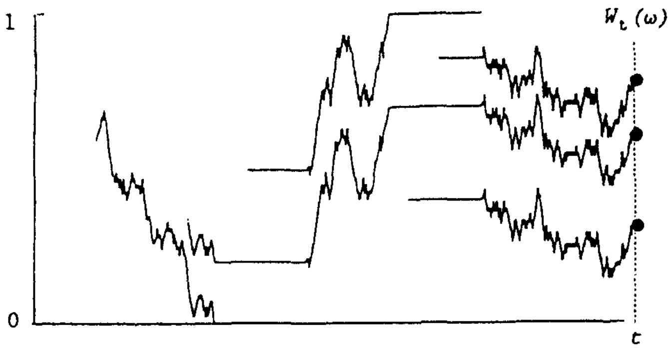


Fig. 2. From the *point of view of the server*, customers arrive according to a Poisson random measure on $\mathbb{R}_+ \times [0, 1]$ with mean measure $aLeb(\mathbb{R}_+) \times \pi$, and move all in the same fashion (when the server is not busy). Whenever a customer "reaches the server" (the corresponding customer path hits 0 or 1), all customers stop moving for a certain service period. After the service has been completed, the customer that has been served is removed and the other customers resume their journeys. The atoms of $W_t(\omega)$ (•) form the relative positions of waiting customers at time t. The paths depicted here correspond to the Brownian-server model.

scribe the positions of waiting customers at time t as atoms of a random counting measure (r.c.m.) W_t on $[0, 1]$. These atoms are formed by the intersections of the customer paths (as seen by the server) and the line $x = t$, see fig. 2. Notice that in fig. 2, the actual configuration of customers on C at time t (corresponding to W_t) is similar to the last configuration of fig. 1, up to a rotation of the picture.

In the next section it is proved that the polling-server system and the Brownian-server system are both stable, in the sense that the measure-valued process (W_t) is regenerative, provided that the traffic intensity ae_1 is smaller than 1. Hence W_t converges in distribution to a limiting random measure W on $[0, 1]$. The probabilistic description of the waiting customers will be given in terms of the law W. W can be interpreted as the random measure of customers that are waiting to be served, in the "stationary situation". Analysis of the process (W_t) is facilitated by the introduction of a new "clock" which stops whenever the server is busy. For any realization $\omega \in \Omega$, we introduce the clock process (S_t) by

$$S_t(\omega) = \int_0^t dx \, I\{W_x(\{0\}, \omega) = 0, W_x(\{1\}, \omega) = 0\} , \qquad t \geq 0 , \tag{3.1}$$

That is, we only run the clock when the server is not busy. Let (ν_t) denote the right-continuous functional inverse of (S_t), and define

$$Q_t = W_{\nu_t} . \tag{3.2}$$

Q_t is again a r.c.m. on [0, 1]. In the next section it will be shown that (Q_t) is a regenerative process as well. Hence there exists a limiting random counting measure Q on [0, 1]. Q can be interpreted as the random measure of waiting customers in the stationary situation *given* that the server is not busy (see Section 5). The relationships between the various measures Q_t, W_t, Q and W will be specified in Section 5.

4 Stability

In this section we prove that the Brownian-server and the polling-server systems are stable, provided that the traffic intensity ae_1 is smaller than 1. Specifically, the main result of this section is the following theorem.

Theorem 4.1: If $ae_1 < 1$, then the processes (W_t) and (Q_t) for the polling-server and Brownian-server systems are regenerative with regeneration periods that have absolutely continuous distributions and finite expectations.

A short proof of Theorem 4.1 for the polling-server case with π uniform, can be found in Section 3 of [10]. This proof can be easily extended to the polling-server case where π is an arbitrary diffuse distribution on [0, 1]. However, for the Brownian-server case the proof of Theorem 4.1 is much more difficult and forms a substantial part of this paper. In the proof we will make use of a stable auxiliary queueing system, which processes all customers slower than the original system. The actual proof is deferred to the end of the section.

Notice that stability does *not* depend on the "speed" of the server. For slow servers the circle "fills up" with a lot of customers, and hence the server does not loose much time traveling to the next customer.

Assume that at time zero we start with the empty state. Clearly, each point in time at which the system returns into this state after a busy period is a regeneration epoch of both the processes (W_t) and (Q_t). We show that the distance between two consecutive regeneration epochs of this type have absolutely continuous distributions and finite expectations.

Notice that a regeneration period of (W_t) and (Q_t) can be represented as the sum of two independent random variables: the residual time to the next arrival (the distribution of which is exponential) plus the length of a busy period. We will refer to the regeneration periods as the *busy cycles* of the queue. Thus, in order to prove that a busy cycle has an absolute continuous distribution, it suffices to show that the length of the corresponding busy period is a proper random variable, e.g. by showing that the expectation of the busy cycle is finite. This can be shown by introducing the following auxiliary queueing system.

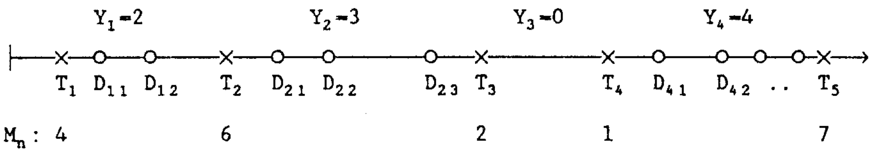


Fig. 3. An illustration of the definitions of the auxiliary batch queue. T_5 is a regeneration epoch.

Consider the following $M/G/\infty$-type batch queue. The batches of customers arrive according to a Poisson process with intensity a, at times $T_1, T_2, \ldots,$ with batch sizes $M_1, M_2, \ldots$ We assume that the M_n's form an i.i.d. sequence of random variables with finite expectation, independent of the arrival times. Let $(X_{nk}, n, k \in \mathbb{N})$ be a further sequence of i.i.d. positive random variables, independent of the arrival times and the batch sizes. Within each interval $[T_n, T_{n+1})$ we construct a sequence of "departure times" in the following way. Let $S_{nk} = T_n + X_{n1} + \cdots + X_{nk}$, $k = 1, 2, \ldots$ and let Y_n be the number of S_{nk}'s in the interval $[T_n, T_{n+1})$, $n = 1, 2, \ldots$ Denote these S_{nk}'s by D_{nk}, $k = 1, \ldots, Y_n$. These D_{nk}'s will be the departure times of customers. Note that then the Y_n's form an i.i.d. sequence of random variables with $0 < EY_n < \infty$, independent of the batch sizes. An illustration of these definitions is given in fig. 3.

Service is performed in the following way: At time of arrival every batch is assigned to a free server. For every $n = 1, 2, \ldots,$ the kth customer of the batch that arrived at T_n is completely served (and therefore departs) at the kth time D_{ij} after T_n, $k = 1, \ldots, M_n$. For example, the six customers that arrive at time T_2 in fig. 3, depart at times $D_{21}, D_{22}, D_{23}, D_{41}, D_{42}$ and D_{43}. The two customers that arrive at time T_3 depart at D_{41} and D_{42}.

The question whether this queueing system is stable can be answered affirmatively. Furthermore, for proving that the busy cycles of the processes (W_t) and (Q_t) of the original polling- or Brownian-server system have finite expectations if $ae_1 < 1$, we will make use of the fact that the number of batches and the number of customers served during a busy period of the auxiliary queue have finite expectation. For showing this finiteness property, we extend the sequence $(T_n, M_n, n \in \mathbb{N})$ to the sequence $(T_n, M_n, n \in \mathbb{Z})$, where the T_n's form a stationary (Poisson) point process $\cdots < T_{-1} < T_0 < 0 \leq T_1 < T_2 < \cdots$ on $\mathbb{R}$ with intensity a, and the M_n's form an independent i.i.d. sequence of random variables, each distributed as M_1. Analogously, we extend the sequence $(X_{nk}, n, k \in \mathbb{N})$ to the i.i.d. sequence $(X_{nk}, n \in \mathbb{Z}, k \in \mathbb{N})$. For $n \leq 0$ we define the random variables S_{nk}, D_{nk} and Y_n in the same way as above. Note, however, that although the Y_n's are still independent, Y_0 has a different distribution than the other Y_n's. But this is of no importance for the proof. For every $n \in \mathbb{Z}$ define

$$T_n^* = \min\left\{T_m > T_n \colon M_n \leq \sum_{j=n}^{m-1} Y_j\right\},$$

and let, for $t \in \mathbb{R}$

$$Z_t = \sum_{k=-\infty}^{\infty} I_{(-\infty,t)}(T_k) I_{(t,\infty)}(T_k^*) . \tag{4.1}$$

T_n^* can be interpreted as the arrival time of the first customer after the time that the batch that arrived at time T_n has been completely served. Z_{T_n}, consequently, can be interpreted as the total number of batches in the system, seen by the batch that arrives at time T_n (itself not included). In the above interpretation, the auxiliary queuing system is in a "stationary situation".

Consider the partial point process of (T_n) consisting of those arrival times T_n such that $Z_{T_n} = 0$. Such arrival times are usually called *empty points* (see e.g. p. 71 in [6]). Using similar arguments as in [9], where the $G/GI/\infty$ queue with independent service times has been considered, we get the following result.

Lemma 4.1: Almost every realization of the arrival process $(T_n,\ n \le 0)$ *on the negative half-line has infinitely many empty points.*

Proof: Note that the event $H_n = \{Z_{T_n} = 0\}$ can be written in the form

$$H_n = \{M_{n-1} \le Y_{n-1}, M_{n-2} \le Y_{n-1} + Y_{n-2}, M_{n-3} \le Y_{n-1} + Y_{n-2} + Y_{n-3}, \ldots\} .$$

Let R be the number of arrival times T_n in $(-\infty, 0)$ for which $Z_{T_n} = 0$, and let $E = (R = \infty\}$. Observe that E only depends on the tail of the i.i.d. sequence $(M_{-1}, Y_{-1}), (M_{-2}, Y_{-2}), \ldots$. Thus by Kolmogorov's zero-one law we get that the probability of event E is either 0 or 1. On the other hand, for the conditional probability $P_Y(E)$ of the event E given (Y_n), we have

$$P_Y(E) = P_Y\left(\bigcap_{n=1}^{\infty} \bigcup_{m \ge n} H_{-m}\right) = \lim_{n\to\infty} P_Y\left(\bigcup_{m\ge n} H_{-m}\right) \ge \lim_{n\to\infty} P_Y(H_{-n}) = P_Y(H_0) . \tag{4.2}$$

Define $B_n = \{M_{-n} \le Y_{-1} + \cdots + Y_{-n}\}$, $n = 1, 2, \ldots$, then

$$P_Y(H_0) = P_Y\left(\bigcap_{n=1}^{\infty} B_n\right) = \prod_{n=1}^{\infty} P_Y(B_n) . \tag{4.3}$$

The last term in (4.3) is an infinite product of positive factors. It is well-known that such a product is positive if and only if

$$\sum_{n=1}^{\infty} \{1 - P_Y(B_n)\} < \infty \ . \tag{4.4}$$

Hence by (4.2)–(4.4) we have that $P_Y(E) > 0$ if and only if

$$\sum_{n=1}^{\infty} P_Y\left(M_{-n} \geq \sum_{j=1}^{n} Y_{-j}\right) < \infty \ . \tag{4.5}$$

But (4.5) holds for almost every realization of (Y_n). This follows from the fact that the M_{-n}'s are identically distributed with finite expectation and independent of the (Y_n), and that by the strong law of large numbers, with probability 1, and for arbitrarily small $\varepsilon > 0$, we have

$$\sum_{j=1}^{n} Y_{-j} \geq n(EY_1 - \varepsilon) > 0 \ ,$$

for sufficiently large n, because $0 < EY_1 < \infty$. Thus the validity of (4.5) is a consequence of the obvious inequalities

$$\sum_{n=1}^{\infty} P_Y(M_{-n} \geq n(EY_{-1} - \varepsilon)) \leq \frac{EM_{-1}}{EY_{-1} - \varepsilon} < \infty \ .$$

Summarizing the above considerations, we showed that $P_Y(E) > 0$ with probability one. Consequently, $P(E) = 1$, which had to be shown. ■

Let $(T_n^e, n \in \mathbb{Z})$ denote the (partial) point process of empty points of the arrival process (T_n), i.e. $\cdots < T_{-1}^e < T_0^e < 0 \leq T_1^e < \cdots$ is the subsequence of those arrival times T_n, for which $Z_{T_n} = 0$. Note that (T_n^e) is a stationary point process on $\mathbb{R}$. From Lemma 4.1 it follows easily that, with probability one, (T_n^e) has infinitely many points both on the negative and positive half-lines. Thus the intensity λ_e of (T_n^e) is strictly positive. Moreover, the intensity is finite, since $\lambda_e \leq a < \infty$. Consequently, from the general theory of stationary point processes it follows that the Palm distribution P^0 of (T_n^e) is well-defined (see e.g. Sect. 12.3 in [4], or Sect. 1.2 in [6]). Namely,

$$P^0(T_1^e - T_0^e < x_1, T_2^e - T_1^e < x_2, \ldots, T_n^e - T_{n-1}^e < x_n)$$

$$= \frac{1}{\lambda_e} \sum_{k=1}^{\infty} P(T_k^e < 1, T_{k+1}^e - T_k^e < x_1, T_{k+2}^e - T_{k+1}^e < x_2, \ldots, T_{k+n}^e - T_{k+n-1}^e < x_n) \ , \tag{4.6}$$

for $x_1, \ldots, x_n > 0$, $n = 1, 2, \ldots$ Furthermore, for every $n \in \mathbb{Z}$ the expectation of $T^e_{n+1} - T^e_n$ taken with respect to the Palm distribution is finite, that is

$$E_{P^0}(T^e_{n+1} - T^e_n) = \lambda_e^{-1} < \infty \ . \tag{4.7}$$

Finally, we have

$$P(T^e_1 \leq v) = \lambda_e \int_0^v dx\, P^0(T^e_1 > x) \ , \qquad v \geq 0 \ , \tag{4.8}$$

which shows that $T^e_1 < \infty$ P-almost surely.

Lemma 4.2: For every $n \geq 1$ we have

$$E(T^e_{n+1} - T^e_n) = E_{P^0}(T^e_{n+1} - T^e_n) = \lambda_e^{-1} \ . \tag{4.9}$$

Proof: Note that the random variables $(T^e_{n+1} - T^e_n)$ are independent, and for $n \neq 0$ identically distributed. Moreover, the pair (T^e_0, T^e_1) is independent of the sequence $(T^e_{n+1} - T^e_n, n \neq 0)$. Thus from the definition (4.6) of the Palm distribution we get

$$P^0(T^e_{n+1} - T^e_n < x) = \frac{1}{\lambda_e} \sum_{k=1}^{\infty} P(T^e_k < 1,\, T^e_{n+k+1} - T^e_{n+k} < x)$$

$$= \frac{1}{\lambda_e} P(T^e_{n+1} - T^e_n < x) \sum_{k=1}^{\infty} P(T^e_k < 1) = P(T^e_{n+1} - T^e_n < x) \ . \qquad \blacksquare$$

Now we are in a position to show that the auxiliary $M/G/\infty$-type batch queue which starts at time zero in the empty state is stable, and that the number of customers served during a busy period of this queue has finite expectation. For doing this, we introduce similarly to (4.1) the process $(Z^0_t, t \geq 0)$ by

$$Z^0_t = \sum_{k=-\infty}^{\infty} I_{(0,t)}(T_k) I_{(t,\infty)}(T^*_k) \ , \tag{4.10}$$

where the T^*_n's are defined as before, $n = 1, 2, \ldots$ Consequently, $Z^0_{T_n}$ can be interpreted as the total number of batches in the system, seen by the batch that arrives at time T_n (itself not included). This all for the auxiliary queuing system starting empty at time zero. The following two results hold for the auxiliary queue.

Lemma 4.3: The expectation of the length of the busy cycle of the auxiliary $M/G/\infty$-type batch queue is finite.

Proof: Note that

$$Z_t^0(\omega) = Z_t(\omega) \qquad \text{for all } t \geq T_1^e(\omega) \ , \tag{4.11}$$

where $T_1^e < \infty$ P-a.s., by (4.8). The expectation of the busy cycle of the auxiliary queue is equal to the expected distance between two consecutive arrival times T_n, with $n \geq 1$, such that $Z_{T_n}^0 = 0$, which, in view of the coupling equation (4.11), is equal to $E(T_2^e - T_1^e)$, which is finite by (4.7) and (4.9). This completes the proof. ∎

Lemma 4.4: The mean number EJ^0 of batches served during a busy period of the auxiliary $M/G/\infty$-type batch queue is finite.

Proof: Let N_k denote the number of arrivals in the interval $[T_1^e, T_k^e)$. By using for example the individual ergodic theorem, we have, with probability one

$$EJ^0 = \lim_{k\to\infty} \frac{N_k}{k} = \lim_{k\to\infty} \frac{N_k}{T_{k+1}^e - T_1^e} \lim_{k\to\infty} \frac{T_{k+1}^e - T_1^e}{k}$$

$$= aE(T_2^e - T_1^e) = a/\lambda_e < \infty \ . \qquad ∎$$

Returning now to the original Brownian- and polling-server systems we give the proof of Theorem 4.1.

Proof of Theorem 4.1: Consider as regenerations epochs the departure times of those customers that leave the system empty. It suffices to prove that the regenerations periods of (W_t) and (Q_t) have finite expectations (see the discussion at the beginning of this section). Proceeding similarly as in [10], we first consider the process (Q_t). Clearly, for showing that the regeneration period of (Q_t) has finite expectation, it suffices to prove that the process $(|Q_t|)$ has this property, where $|Q_t| = Q_t([0, 1])$ is the total number of customers in the system at time t (measured with respect to the new clock introduced with the definition of Q_t).

Next we majorize $(|Q_t|)$ by the process $(Z_t^0, t \geq 0)$ defined in (4.10) where the characteristics of the corresponding auxiliary $M/G/\infty$-type batch queue are chosen in the following specific way: We subdivide the customers arriving in the original system into disjoint classes. The group of *first-generation customers* is formed by those customers that arrive when the server is travelling. The group of *second-generation customers* is formed by those customers that arrive when a first-generation customer is being served. Third-, fourth-, etc. generation customers are defined in a similar way. Notice that any nth-generation customer is a "descendant" from only one first-generation customer. Let $T_1, T_2, \ldots$ denote the arrival times of the first-generation customers in the new time scale. They form a stationary Poisson process on the positive half line with intensity a. Furthermore, for $n \geq 1$, let M_n be the total number of descendants produced by the nth first-generation customer (including this initiating customer). It is easy to show (cf. [10]) that the M_n's form an i.i.d. sequence of random variables that are independent of the arrival times of first-generation customers, and that $EM_n < \infty$ if $ae_1 < 1$. To complete the specification of the auxiliary queue, consider first the Brownian-server case.

The random variables X_{nk} appearing in the general definition of the auxiliary $M/G/\infty$-type batch queue are determined in the following way by the Brownian motion of the server on the circle. Namely, from the independence properties of the Brownian motion with zero drift it follows that, in the original Brownian-server system, we can start the Brownian motion anew at each arrival time T_n, $n \in \mathbb{N}$ of a first-generation customer without changing the random mechanism of the system. By $(B_t^{(n)}, t \geq 0)$ we denote the Brownian motion which starts at T_n. Furthermore, we can assume that, for each $n \in \mathbb{N}$, the processes $(B_t^{(n)}, t \geq 0)$ are i.i.d. copies of a Brownian motion with zero drift and diffusion parameter $\sigma^2 > 0$ and, moreover, that they are independent of the T_n's and the M_n's. Then, let X_{nk} be the random amount of time which the Brownian motion $(B_t^{(n)}, t \geq 0)$ needs for achieving, for the kth time, a (positive or negative) increase of size 1, i.e.

$$X_{nk} = \min\left\{t > 0: \left|B^{(n)}\left(\sum_{j=1}^{k-1} X_{nj} + t\right) - B^{(n)}\left(\sum_{j=1}^{k-1} X_{nj}\right)\right| = 1\right\}.$$

Clearly, the X_{nk}'s given in this way form an i.i.d. sequence of proper positive random variables which are independent of the T_n's and M_n's. Hence the auxiliary queue fulfills all the assumption used in the proofs of Lemmas 4.1–4.4. Furthermore, from the construction of this special auxiliary $M/G/\infty$-type batch queue it follows that, for every $n \geq 1$, the last descendant of the nth first-generation customer leaves this system not eariler than he does in the original Brownian-server system (considering here the new time scale introduced with (Q_t)). This is because of two reasons: First, it is ruled out that more than one descendant of one and the same first-generation customer is served simultaneously, whereas in the Brownian-server case the server can walk towards more than one descendant of one and the same first-generation customer. Second, the walking distance to the next descendant is always less than the distance 1 which has been considered in the definition of the "service times" (X_{nk}). Consequently, the process $(Z_t^0, t \geq 0)$ of our specific auxiliary queue majorizes the process $(|Q_t|, t \geq 0)$, in the sense that

$$\min\{t > T_1 : |Q_t| = 0\} \leq \min\{t > T_1 : Z_t^0 = 0\} , \tag{4.12}$$

with probability one. Thus, from Lemma 4.3 it follows that the process (Q_t) has a regeneration period with finite expected length.

Similarly, for the polling-server system, the specification of the special auxiliary $M/G/\infty$-type batch queue such that (4.12) holds, is completed by taking $X_{nk} = 1$, for all $k, n \geq 1$.

Finally, for showing that also the process (W_t) is regenerative it suffices to notice that the difference between the length of the (first) regeneration period of (W_t) and that of (Q_t) is equal to the sum of the service times of all customers served during this regeneration period. Thus it remains to be shown that the expectation of this sum is finite. Let J denote the number of first-generation customers served during the (first) regeneration period, and let U_{jk} denote the service time of the $(k-1)$th descendant of the jth first-generation customer, $j = 1, \ldots, J$, $k = 1, \ldots, M_j$. The expected sum of the service times can now be written as

$$E \sum_{j=1}^{J} \sum_{k=1}^{M_j} U_{jk} = EJE \sum_{k=1}^{M_1} U_{1k} = EJEM_1EU_{11}$$

by Wald's lemma. From (4.13) and Lemma 4.4 we get $EJ \leq EJ^0 < \infty$, and by our assumptions we have $EM_1 < \infty$ and $EU_{11} < \infty$, which completes the proof. ■

Remark 4.1: In [13] the following modification of the polling-server has been considered. The server is assumed to scan an interval of fixed length at constant speed. The scan follows a fixed path, where the interval is divided into a finite number of subintervals which not necessarily must be disjoint. Stability of this system (in particular, Theorems 1 and 4 in [13]) immediately follows from the results we obtained in the present section. Namely, it suffices to specify the auxiliary $M/G/\infty$-type batch queue in the same way as we did for the polling-server and the Brownian-server system, with the only difference that now the X_{nk}'s are taken equal to the (fixed) total length of a scanning path, for all $k, n \in \mathbb{N}$.

5 The Measures Q and W

As a result of Theorem 4.1 and the Key Renewal Theorem (KRT) (cf. [1]), there exist limiting r.c.m.'s W and Q on [0, 1] such that

$$W_t \xrightarrow{\mathscr{D}} W \qquad \text{and} \qquad Q_t \xrightarrow{\mathscr{D}} Q \ ,$$

as $t \to \infty$, where $\xrightarrow{\mathscr{D}}$ denotes convergence in distribution. W represents the positions of waiting customers (relatively to the position of the server) in the *stationary* situation. Q can be interpreted as the random measure of waiting customers in the stationary situation *given* that the server is not busy (see (5.4)). It will be convenient to introduce, similarly to (3.1), two new "clocks" (S_t^0) and (S_t^1), which are defined by

$$S_t^i = \int_0^t dx\, I\{W_x\{i\} = 1\} \ , \qquad t \geq 0 \ , \quad i = 0, 1 \ . \tag{5.1}$$

In other words, we run (S_t^0) if there is a customer being served at 0, and we run (S_t^1) only when the server is busy with a customer at 1. Let (ν_t^i) denote the right-continuous functional inverses of (S_t^i), $i = 0, 1$, and define

$$Q_t^i = W_{\nu_t^i} \ , \qquad i = 0, 1 \ . \tag{5.2}$$

Then Q_t^0 and Q_t^1 are again r.c.m.'s on $[0, 1]$ and the processes (Q_t^0) and (Q_t^1) are regenerative with absolutely continuous regeneration periods of finite expected length, because (W_t) has this property. The corresponding limiting random measures to which these processes converge in distribution are denoted by Q^0 and Q^1, respectively. The next theorem gives another interpretation of these measures.

Theorem 5.1: The limiting probability of being idle is given by

$$\lim_{t \to \infty} P(W_t\{0\} = 0, W_t\{1\} = 0) = 1 - ae_1 \ . \tag{5.3}$$

Moreover, for any $f \in \mathscr{B}[0, 1]$ (see appendix), we have

$$\lim_{t \to \infty} E(e^{-W_t f} \mid W_t\{0\} = 0, W_t\{1\} = 0) = Ee^{-Qf} \ . \tag{5.4}$$

and, for $i = 0, 1$,

$$\lim_{t \to \infty} E(e^{-W_t f} \mid W_t\{i\} = 1) = Ee^{-Q^i f} \ . \tag{5.5}$$

Proof: As regeneration epochs for (Q_t) we take those times t for which $|Q_{t-}| = 1$ and $|Q_t| = 0$ (and similarly for (W_t)). The regeneration epochs of (W_t) are just the departure times of customers that leave the system empty. Let Y and Y' denote the length of the first regeneration period of (Q_t) and (W_t), respectively. Notice that since the system starts empty, we can take $t = 0$ as the first regeneration epoch for both processes. Since Y' is also the length of the first regeneration period of $(e^{-W_t f} I_{\{W_t\{0\}=0, W_t\{1\}=0\}})$, we obtain from the KRT that

$$\lim_{t\to\infty} E(e^{-W_t f} I_{\{W_t\{0\}=0, W_t\{1\}=0\}}) = \frac{1}{EY'} E \int_0^{Y'} ds\, e^{-W_s f} I_{\{W_s\{0\}=0, W_s\{1\}=0\}}$$

$$= \frac{1}{EY'} E \int_0^{Y} ds\, e^{-Q_s f} = \frac{EY}{EY'} E e^{-Qf} \,. \tag{5.6}$$

Similarly, we obtain

$$\lim_{t\to\infty} EI_{\{W_t\{0\}=0, W_t\{1\}=0\}} = \frac{EY}{EY'}\,, \tag{5.7}$$

so that (5.4) follows from (5.6) and (5.7), and (5.5) follows by analogy. The limiting probability of being idle follows directly from the KRT, as in Corollary 3.1 of [10]. ■

Remark 5.1: By the KRT and Theorem 5.1 it easy to see that the distribution of W is a mixture of the distributions of Q, Q^0 and Q^1, the mixing factors being $1 - ae_1$, $P(W\{0\} = 1)$ and $P(W\{1\} = 1)$, respectively.

Using Theorem 5.1 we can interpret Q as the r.c.m. of waiting customers in the stationary situation, *given* that the server is *not busy*. Moreover, a customer who (in the Brownian-server case) is under service could have been approached by the server from one of two directions. Remembering that we fixed an orientation on C, we can interpret Q^0 as the r.c.m. of waiting customers in the stationary situation, *given* that the server is busy *and* that the server travelled in the *direction* of the *orientation* on the circle before reaching the customer. A similar interpretation holds for Q^1 for the opposite direction.

Before we come to the main result of this section, that relates the various random measures in yet another way to one another, we need to introduce some more notation. Let $f \in C_+[0, 1]$ (the set of positive, continuous functions on $[0, 1]$), and consider the stochastic process $(W_t f)$, starting at 0. A typical realization of $(W_t f)$, for the Brownian-server case, is given in fig. 4.

In both models upward jumps of $(W_t f)$ occur via a Poisson process (A_t) with rate a on $\mathbb{R}_+$. The size of a jump is independent of everything else and is

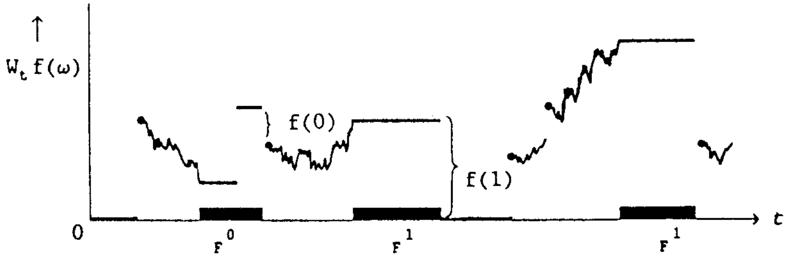


Fig. 4. A realization of $(W_t f)$. For every ω, $F^i(\omega)$ denotes the set of times where W_t has an atom at i, $i = 0, 1$. Upward jumps correspond to arrivals of customers and have size $f(\xi_i)$, where (ξ_i) is an i.i.d. sequence of r.v.'s on $[0, 1]$, all having distribution π. Downward jumps correspond to departures of customers and have size $f(0)$ or $f(1)$.

distributed as $f(\xi)$, where ξ has distribution π on $[0, 1]$. Downward jumps have sizes $f(0)$ or $f(1)$, depending on whether W_t had an atom at 0 or 1 just before the jump. Let (D_t^i) denote the departure counting process that counts downward jumps of size $f(i)$, $i = 0, 1$. Let $(\tilde{A}_t)$ be the compound Poisson process that jumps at arrival times (T_i), with jump sizes $(f(\xi_i))$, the ξ_i's having distribution π, being i.i.d. and independent of everything else. The continuous part of $(W_t f)$ is denoted by (C_t). Let $F^i = F^i(\omega)$ denote the set of times that W_t has an atom at i, i.e. $F^i(\omega) = \{t \geq 0\colon W_t(\omega, \{i\}) = 1\}$, $i = 0, 1$. And let $Z(\omega) = \mathbb{R}_+ \backslash \{F^0(\omega) \cup F^1(\omega)\}$. For all $t \geq 0$ we have

$$W_t f = C_t - D_t^0 f(0) - D_t^1 f(1) + \tilde{A}_t \ , \tag{5.8}$$

The next theorem, the main theorem of this section, shows the relationship between the various random measures. It could be regarded as a kind of stochastic decomposition result for random measures (see also [7] and [8] for stochastic decomposition results on cyclic server systems). By specifying process (C_t) for the two queueing systems at hand, we will be able to derive several important performance measures for the different models. In connection with this, we use the fact that (in both models) (C_t) is a semi-martingale w.r.t. the filtration generated by (W_t), and that for $f \in V$, for some set $V \subseteq C_+[0, 1]$, the limits

$$L_1 = -\lim_{t\to\infty} \frac{1}{t} E \oint_0^t e^{-W_s f}\, dC_s \qquad \text{and} \qquad L_2 = \lim_{t\to\infty} \frac{1}{t} E \frac{1}{2} \int_0^t e^{-W_s f} d[C, C]_s$$

exist (here $\oint$ denotes the Itô-integral sign, and $[C, C]$ the quadratic variation process of (C_t)). The proof that (C_t) is a semi-martingale is given in Sections 6 and 7. Specifically: for the polling system (C_t) is of bounded variation for every continuously differentiable f on $[0, 1]$ (see (6.1)), and for the Brownian-server system we have that for every twice-continuously differentiable function f on

[0, 1], C_t can be written in the form $\oint_0^t F_s dB_s + H_t$, where (F_s) is a left-continuous process, (B_t) a Brownian motion on $\{t \geq 0\colon W_t\{0\} = W_t\{1\} = 0,\ W_t[0, 1] > 0\}$ and constant elsewhere, and H_t a process of bounded variation, all processes being adapted to the to the filtration generated by (W_t), see (7.1). In Sections 6 and 7 it is also shown that L_1 and L_2 are well-defined. Notice that since (C_t) is of bounded variation for the polling-server model, the Itô-integral in L_1 becomes in this case an ordinary (stochastic) integral, and, moreover, $L_2 = 0$.

Theorem 5.2: Let $p_i = P(W\{i\} = 1)$, $i = 0, 1$, *then the following relationship holds for all* $f \in V$:

$$0 = L_1 + L_2 - \beta E e^{-Wf} + \frac{\beta L_F(\beta)}{1 - L_F(\beta)}(p_0(e^{f(0)} - 1)Ee^{-Q^0 f}$$

$$+ p_1(e^{f(1)} - 1)Ee^{-Q^1 f}) \ , \tag{5.9}$$

where $\beta = a\int_0^1 \pi(dx)(1 - e^{-f(x)})$, *and* L_F *denotes the Laplace-Stieltjes transform of* F.

Proof: Stochastic intensities, compensators etc. are always w.r.t. the filtration generated by (W_t). The theorem is proved analogously to Theorem 3.3 of [10]. We therefore only sketch the main ideas. Because (C_t) is a semi-martingale, it follows from (5.8) that $(W_t f)$ is also a semi-martingale. By Itô's formula (see appendix) and (5.8) we have

$$e^{-W_t f} = e^{-W_0 f} - \oint_0^t e^{-W_s f}\, dC_s + \frac{1}{2}\int_0^t e^{-W_s f} d[C, C]_s$$

$$+ \sum_{0 < s \leq t} [e^{-W_s f} - e^{-W_{s-} f}] \ , \tag{5.10}$$

where this last sum can be written as

$$\sum_{i=1}^{\infty} (e^{-f(\xi_i)} - 1)e^{-W_{T_i-} f} I_{(0,t]}(T_i) + (e^{f(0)} - 1)\int_0^t e^{-W_{s-} f}\, dD_s^0$$

$$+ (e^{f(1)} - 1)\int_0^t e^{-W_{s-} f}\, dD_s^1 \ . \tag{5.11}$$

Notice that ξ_i is independent of T_i and W_{T_i-}, and has distribution π on [0, 1]. The next step is to take expectations on both sides of (5.10), with (5.11) substituted into (5.10). The expectation of the infinite sum in (5.11) is equal to $-\beta a^{-1} E \int_0^t e^{-W_{s-}f}\, dA_s$, so that we get

$$Ee^{-W_t f} = Ee^{-W_0 f} - E\oint_0^t e^{-W_s f}\, dC_s + \frac{1}{2} E\int_0^t e^{-W_s f} d[C, C]_s - \frac{\beta}{a} E\int_0^t e^{-W_{s-} f}\, dA_s$$

$$+ (e^{f(0)} - 1) E\int_0^t e^{-W_{s-} f}\, dD_s^0 + (e^{f(1)} - 1) E\int_0^t e^{-W_{s-} f}\, dD_s^1 \,. \qquad (5.12)$$

We convert, by (A.8), the expectations of the stochastic integrals in the last three terms of (5.12) into expectations of (stochastic) integrals w.r.t. the compensators of the counting processes (A_t), (D_t^0) and (D_t^1), respectively. The compensator of (D_t^i), $i = 1, 2$, is very similar to the compensator of a renewal process, and is given (in an adjusted form) by Lemma 3.1 of [10]. The compensator of (A_t) is just (at). Finally, after the conversion, we divide the left- and right-hand side of (5.12) by t and let $t \to \infty$. The proof is then completed through application of the KRT and Theorem 3.1, analogously to the proof Theorem 5.3 of [10]. ■

Remark 5.2: Notice that in the polling-server case, by (5.3), we have $p_0 = ae_1$ and $p_1 = 0$ in (5.9). For the Brownian-server model we limit ourselves to distributions π that are symmetric around 1/2, in which case, by symmetry, $p_0 = p_1 = ae_1/2$.

Next, we specify the continuous process (C_t) for the two different server types. Using Theorem 5.1, this will give us important information about the queueing systems.

6 Polling-Server

In this model the server (when not busy) travels at constant speed α^{-1} in the direction of the orientation on C. For the case that π is the uniform distribution on [0, 1], a complete solution was found in [10]. The more general polling model that we consider here (π is an arbitrary diffuse distribution on [0, 1]) can be analyzed in exactly the same way as the former model. We therefore only present the main results.

Since the server only travels in the direction of the orientation on C, we immediately have that $P(W\{1\} = 1) = 0$. Therefore Q^1 plays no role here. Clearly, (C_t) (defined in Section 5) is a process of bounded variation, hence its quadratic variation vanishes. Moreover, we have almost surely, for $f \in C_+^1[0, 1]$ (the set of positive continuously differentiable functions on $[0, 1]$)

$$\frac{d}{dt} C_t = -\alpha^{-1} I_Z(t) W_t f' \ . \tag{6.1}$$

The proof of (6.1) goes as follows. Recall that $Z(\omega)$ is the set of times where the server is not busy. First observe that $(C_t(\omega))$ is constant on $\bar{Z}(\omega)$. Next, suppose that $|W_t(\omega)| = n > 0$ on $Z(\omega)$. Let $x_1, \ldots, x_n$ be the atoms of $W_t(\omega)$, then

$$W_t(\omega) f = \sum_{i=1}^{n} f(x_i)$$

and, for sufficiently small $h > 0$,

$$\begin{aligned} W_{t+h}(\omega) f &= \sum_{i=1}^{n} f(x_i - \alpha^{-1} h) = \sum_{i=1}^{n} f(x_i) - h\alpha^{-1} \sum_{i=1}^{n} f'(x_i) + o(h) \\ &= W_t(\omega) f - h\alpha^{-1} W_t(\omega) f' + o(h) \ . \end{aligned}$$

Therefore

$$\lim_{h \downarrow 0} \frac{W_{t+h} f - W_t f}{h} = -\alpha^{-1} W_t f' \ , \qquad \text{on } Z \ .$$

This is valid if $|W_t(\omega)| > 0$, but trivially also when $|W_t(\omega)| = 0$, so that (6.1) follows. The following theorem is a specification of Theorem 5.2.

Theorem 6.1: For all $f \in C_+^1[0, 1]$ we have,

$$0 = (1 - ae_1) E e^{-Qf} \alpha^{-1} Q f' - \beta E e^{-Wf} + ae_1 \frac{\beta L_F(\beta)}{1 - L_F(\beta)} (e^{f(0)} - 1) E e^{-Q^0 f} \ , \tag{6.2}$$

where $\beta = a \int_0^1 \pi(dx)(1 - e^{-f(x)})$, and L_F denotes the Laplace-Stieltjes transform of F.

Proof: By (6.1) we have that (C_t) is a (continuous) process of bounded variation (hence a semi-martingale), and consequently that $L_2 = 0$ in Theorem 5.2. Moreover, by Remark 5.2 we have $p_0 = ae_1$ and $p_1 = 0$ in (5.9). It remains therefore to be shown that, for all $f \in C_+^1[0, 1]$

$$L_1 = \lim_{t\to\infty} \frac{1}{t} E \oint_0^t e^{-W_s f}\, dC_s = -(1 - ae_1)\alpha^{-1} Ee^{-Qf} Qf' \ , \tag{6.3}$$

since then (6.2) is immediate from Theorem 5.2. Let Y and Y' be the lengths of the first regeneration periods of (Q_t) and (W_t), respectively, as in the proof of Theorem 5.1. Then (6.3) follows by the following set of equalities:

$$\begin{aligned} L_1 &= -\alpha^{-1} \lim_{t\to\infty} \frac{1}{t} E \int_0^t ds\, e^{-W_s f} I_Z(s) W_s f' \\ &= -\alpha^{-1} \lim_{t\to\infty} E(e^{-W_t f} I_Z(t) W_t f') = -\alpha^{-1} \frac{1}{EY'} E \int_0^{Y'} ds\, e^{-W_s f} I_Z(s) W_s f' \\ &= -\alpha^{-1} \frac{1}{EY'} E \int_0^{Y} ds\, e^{-Q_s f} Q_s f' = -\alpha^{-1}(1 - ae_1) Ee^{-Qf} Qf' \ . \end{aligned}$$

The first equality follows from (6.1), the second one from the time-averaging property of the regenerative process $(e^{-W_t f} I_Z(t) W_t f')$ (cf. Theorem V.3.1. of [1]). The third equality is a result of the KRT applied to $(e^{-W_t f} I_Z(t) W_t f')$. The fourth equality follows from the definition of Q_t as a time-change of W_t. And the last equality is a result of the KRT, the Continuous Mapping Theorem (CMT) (cf. [4], p. 618) and Theorem 5.1, taking into consideration that Y is also the first regeneration cycle of $(e^{-Q_t f} Q_t f')$. ∎

Since Theorem 6.1 shows that the distribution of Q^0 (and hence also the distribution of W) is completely determined by that of Q, it suffices to concentrate on Q. The next theorem determines the law of Q in terms of a computable expression for its Laplace functional.

Theorem 6.2: Let G be the function defined by

$$G(z) = Ee^{-a(1-z)X} \ , \qquad z \in [0, 1] \ , \tag{6.3}$$

X being a random variable with d.f. F. Then the Laplace functional of Q is given by

$$Ee^{-Qf} = \exp - a\alpha \int_0^\infty ds \{1 - K(s)\} \ , \tag{6.4}$$

where K is the unique solution to

$$K(t) = \begin{cases} \int_0^t \pi(dx) G(K(t-x)) + \int_t^1 \pi(dx) e^{-f(x-t)} , & \text{for } 0 \le t \le 1 \\ \int_0^1 \pi(dx) G(K(t-x)) , & \text{for } t \ge 1 , \end{cases} \tag{6.5}$$

Theorems 6.1 and 6.2 (in combination with Remark 5.1) completely describe the probability law of the r.c.m. of waiting customers. Moment measures of Q and W can be derived at least numerically, but some performance measures can be given explicitly. The next corollary can be derived from Theorem 6.1 alone, analogously to Theorem 5.1 in [10]. A similar approach is used in Corollary 7.1 for the Brownian-server case.

Corollary 6.1: The mean measure of Q is given by

$$EQ(dx) = \frac{a\alpha}{1 - ae_1} \pi[x, 1] \, dx \ . \tag{6.6}$$

The expected number of customers on the ring is

$$E|W| = ae_1 + \frac{a^2 e_2/2 + \alpha a \int_0^1 \pi(dx) x}{1 - ae_1} \ . \tag{6.7}$$

Corollary 6.2: When π is the uniform distribution on [0, 1], *we can even find explicit second-order results, namely* (*see Corollary* 6.1 *of* [8])

$$\begin{aligned} \operatorname{var} Qf = {} & \frac{\alpha a}{1 - ae_1} \int_0^1 dx (1-x) f^2(x) + \frac{\alpha a^3 e_2}{1 - ae_1} \int_0^1 dy \left(\int_0^y dx \, f(x) \right)^2 \\ & + \frac{\alpha a^4 e_1 e_2}{(1 - ae_1)^2} \left(\int_0^1 dx (1-x) f(x) \right)^2 . \end{aligned} \tag{6.8}$$

Moreover, when π is the uniform distribution on [0, 1] *and when the service is constant, an explicit expression for the Laplace functional of Q is given by*

$$Ee^{-Qf} = e^{-c_f} \left(\frac{1 - ae_1}{1 - ae_1 \int_0^1 dy \, e^{-h(y)}} \right)^{\alpha/e_1} , \tag{6.9}$$

where $c_f = \alpha a \int_0^1 dx(1-x)(1-e^{-f(x)})$, $h(y) = ae_1 \int_0^y dx(1-e^{-f(x)})$, *for every positive measurable function* f *on* $[0, 1]$. ■

7 Brownian-Server

In this model the (idle) server carries out a Brownian motion on the circle which is completely independent of everything else, with zero drift and variance parameter σ^2. We take π here symmetrical on $[0, 1]$ w.r.t. $1/2$, so that, in view of (5.3), by symmetry we have $P(W\{0\} = 1) = P(W\{1\} = 1) = ae_1/2$. Unlike the polling model, no complete solution (in terms of the Laplace functional of W) is known for the Brownian-server model. One could still, as in the polling case (cf. [10]), analyze the system as a stochastic flow, in which particles are born, move and die in some random way, but the dependencies between the particles complicate the analysis severely. However, important information about the behavior of the system can still be derived. The following theorem is again a direct consequence of Theorem 5.2.

Theorem 7.1: For all $f \in C_+^2[0, 1]$ *we have,*

$$0 = (1 - ae_1)\sigma^2\left(-\frac{1}{2}Ee^{-Qf}Qf'' + \frac{1}{2}Ee^{-Qf}(Qf')^2\right) - \beta Ee^{-Wf}$$

$$+ \frac{1}{2}ae_1\frac{\beta L_F(\beta)}{1 - L_F(\beta)}((e^{f(0)} - 1)Ee^{-Q^0 f} + (e^{f(1)} - 1)Ee^{-Q^1 f}) ,$$

where, $\beta = a\int_0^1 \pi(dx)(1 - e^{-f(x)})$, *and* L_F *is the Laplace-Stieltjes transform of* F.

Proof: Notice first that (C_t) is no longer of bounded variation. Therefore, specifying the behaviour (C_t) by pathwise arguments, as in (6.2), is no longer valid. However, using Itô's formula we can show that (C_t) satisfies a simple stochastic differential equation.

Obviously (C_t) is constant on $\overline{Z}$ (Z as is Sections 5 and 6). Next, suppose that $|W_t| = n > 0$, with atoms at $X_1(t), \ldots, X_n(t)$, for some $t \in Z$. Let $f \in C_+^2[0, 1]$. We can write $W_t f$ as $f(X_1(t)) + \cdots + f(X_n(t))$. By Itô's formula (in stochastic differential form) we have for every $i = 1, \ldots, n$

$$df(X_i(t)) = f'(X_i(t))\,dX_i(t) + \tfrac{1}{2}f''(X_i(t))d[X_i, X_i]_t$$

$$= \sigma f'(X_i(t))\,dB_t + \tfrac{1}{2}\sigma^2 f''(X_i(t))\,dt ,$$

where B is a standard Brownian motion. Consequently,

$$dC_t = (\sigma W_t f' \, dB_t + \tfrac{1}{2}\sigma^2 W_t f'' \, dt) I_Z(t) \ , \tag{7.1}$$

which is also true for t for which $|W_t| = 0$. This shows that (C_t) is a (continuous) semi-martingale (see also Remark 5.3). As a consequence of (7.1) we have, by (A.6), the following quadratic variation of (C_t)

$$[C, C]_t = \sigma^2 \int_0^t du \, I_Z(u)(W_u f')^2 \ . \tag{7.2}$$

Since by symmetry $P(W\{0\} = 1) = P(W\{1\} = 1) = ae_1/2$, Theorem 7.1 follows from Theorem 5.2, if we can prove that

$$\begin{aligned} L_1 + L_2 &= \lim_{t\to\infty} \frac{1}{t}\left(-E \oint_0^t e^{-W_s f} \, dC_s + E \frac{1}{2} \int_0^t e^{-W_s f} d[C, C]_s \right) \\ &= (1 - ae_1)\sigma^2 \left(-\frac{1}{2} Ee^{-Qf} Qf'' + \frac{1}{2} Ee^{-Qf}(Qf')^2 \right) . \end{aligned} \tag{7.3}$$

By (A.7) we have $E \oint_0^t e^{-W_s f} W_t f' \, dB_t = 0$, so that by (7.1) and (7.2) $L_1 + L_2$ is equal to

$$\lim_{t\to\infty} \frac{1}{t} \frac{\sigma^2}{2} \left(-E \int_0^t ds \, e^{-W_s f} W_s f'' I_Z(s) + E \int_0^t ds \, e^{-W_s f}(W_s f')^2 I_Z(s) \right) . \tag{7.4}$$

Let Y and Y' be defined as in the proof of Theorem 5.1. The proof of (7.3) now goes analogously to the proof of Theorem 6.1. Consider the regenerative processes $(X_t') = (e^{-W_t f} I_Z(t)\{-W_t f'' + (W_t f')^2\})$ and $(X_t) = (e^{-Q_t f}\{-Q_t f'' + (Q_t f')^2\})$. By the KRT and the CMT we have that (X_t) converges in distribution to a random variable $U = e^{-Qf}\{-Qf'' + (Qf')^2\}$. By the time-average properties of regenerative processes, the definition of Q_t and Theorem 5.1, (7.4) therefore becomes

$$\frac{\sigma^2}{2} \frac{1}{EY'} E \int_0^{Y'} ds \, X_s' = \frac{\sigma^2}{2} \frac{EY}{EY'} \frac{1}{EY} E \int_0^{Y} ds \, X_s = \frac{\sigma^2}{2}(1 - ae_1)EU \ , \tag{7.5}$$

which proves (7.3). ∎

Corollary 7.1: The mean measure of Q is given by

$$EQ(dx) = \frac{a\sigma^{-2}}{(1-ae_1)}\left(2\int_0^x \pi(dy)(y-x) + x\right)dx \ , \tag{7.6}$$

and the expected number of waiting customers is

$$E|W| = \frac{a\sigma^{-2}}{(1-ae_1)}\int_0^1 \pi(dy)y(1-y) + ae_1 + \frac{a^2e_2}{2(1-ae_1)} \ . \tag{7.7}$$

Proof: Since Theorem 7.1 also holds for functions pf where p is an arbitrary small positive number we have in particular for all $f \in C_+^2[0, 1]$

$$0 = -\frac{1}{2}\sigma^2 EQf''(1-ae_1) - a\int_0^1 \pi(dx)f(x) + \frac{1}{2}a(f(0)+f(1)) \ ,$$

which can be rewritten as

$$\frac{(1-ae_1)}{a\sigma^{-2}}EQf'' = 2\int_0^1 (dx - \pi(dx))f(x) + \int_0^1 dx\, x(1-x)f''(x) \ . \tag{7.8}$$

By partial integration (recall here that π is symmetric w.r.t. symmetry point 1/2) we can write the first integral on the right-hand side of (7.8) as

$$2\int_0^1 dx\left(\frac{1}{2}x^2 - \int_0^x dy\, \pi[0,y]\right)f''(x) \ , \tag{7.9}$$

so that, after interchanging the integration-order, we find from (7.8) and (7.9) that, for arbitrary $f \in C_+^2[0, 1]$,

$$\int_0^1 EQ(dx)f''(x) = \int_0^1 dx\left\{\frac{a\sigma^{-2}}{(1-ae_1)}\left(-2x\pi[0,x] + 2\int_0^x \pi(dy)y + x\right)\right\}f''(x) \ ,$$

which proves (7.6). Moreover, let $f(x) \equiv p$, $p > 0$ in Theorem 7.1, then for all $p > 0$,

$$0 = -a(1-e^{-p})Ee^{-p|W|} + \frac{1}{2}ae_1\frac{a(1-e^{-p})L_F(a(1-e^{-p}))}{1-L_F(a(1-e^{-p}))}(e^p - 1)$$
$$\times\ (Ee^{-p|Q^0|} + Ee^{-p|Q^1|}) \ . \tag{7.10}$$

Since by symmetry $Ee^{-p|Q^0|} = Ee^{-p|Q^1|}$, we have by Remark 5.1,

$$Ee^{-p|W|} = (1 - ae_1)Ee^{-p|Q|} + ae_1 Ee^{-p|Q^0|} \ . \tag{7.11}$$

Substituting (7.11) into (7.10) yields, after dividing right- and left-hand side of (7.10) by $-a(1 - e^{-p})$, that

$$0 = (1 - ae_1)Ee^{-p|Q|} + e_1\left(\frac{a(1 - e^p)L_F(a(1 - e^{-p}))}{1 - L_F(a(1 - e^{-p}))} + a\right)Ee^{-p|Q^0|} \ .$$

Since this is true for all $p > 0$, we must have in particular that

$$0 = -(1 - ae_1)E|Q| + e_1(e_1^{-1} - a)E|Q^0| - \frac{1}{2}ae_1^{-1}e_2 + (a - e_1^{-1})e_1 \ . \tag{7.12}$$

Moreover, by (7.11)

$$E|W| = (1 - ae_1)E|Q| + ae_1 E|Q^0| \ , \tag{7.13}$$

and by (7.6)

$$E|Q| = \frac{a\sigma^{-2}}{(1 - ae_1)}\int_0^1 \pi(dy)y(1 - y) \ , \tag{7.14}$$

so that (7.7) follows from (7.12)–(7.14). ■

8 Conclusion/Remarks

Random counting measures provide a convenient way to describe the positions of waiting customers for these zero-buffer cyclic server models. A complete solution for the polling-server model was found, in terms of Theorems 6.1 and 6.2. Numerical results will be given in a separate paper. When π is absolutely continuous w.r.t. the Lebesgue measure we can rewrite (6.5) in terms of a (non-linear) difference-differential equation, which can be solved efficiently by Runge-Kutta methods. However, it is not unlikely that an analytic solution of (6.5) exists for every F (remember that in the constant service time case (6.9) gives such a solution). In fact, by taking $e_1 = 1$ and expanding K and (6.4) in terms of $a < 1$, one can, by subsequently solving (6.5), find as many terms of the expansion of K as one wishes. This is, unfortunately, a very laborious work, and no clear structure in the solution procedure has been found yet.

For the Brownian-server model we derived the mean measure of Q, which gives a good indication how the customers are "on the average" distributed over the circle. Moreover, the exact expectation for the number of waiting customers was derived, which is one of the main performance measures. If π is not symmetric, we are left with one unknown parameter $P(W\{0\} = 1)$ in Theorem 7.1, but apart from that, the analysis stays the same. The unknown constant could for example be estimated by simulation, which then gives an estimation for EQ.

Other performance measures (such as the waiting times of the customers) can be tackled by the same approach. For some results on waiting times in the constant service time, with π uniform, we refer to [2]. Formulas for the mean waiting time in scanning-server systems have been derived in [13].

Remark 8.1: When the service times are zero, the measures W and Q can be taken as one and the same. Moreover, it is easy to see that W must be a *Poisson Random Measure* on [0, 1] in the polling-server case. In particular, W is completely specified by its *mean measure*, which is given in (6.6) with $e_1 = 0$. In the Brownian-server case, W is in general not Poisson when $e_1 = 0$. However, the mean measure of W can still be given (in (7.6), with $e_1 = 0$). For example, suppose that π is the uniform distribution on [0, 1] and $e_1 = 0$. Then W has mean measure

$$EW(dx) = a\alpha(1 - x)\,dx\ , \qquad \text{for } x \in [0, 1]\ ,$$

in the polling-server case, and

$$EW(dx) = a\sigma^{-2}x(1 - x)\,dx\ , \qquad \text{for } x \in [0, 1]\ ,$$

in the Brownian-server case. Hence the mean measure has a *linear* density in the polling-server case and a *quadratic* density in the Brownian-server case. The models in which the service times are negligible, are also referred to as *snow-plow* models, cf. [3].

A similar approach can be used to study the "greedy-server" model, where the server always travels to the nearest customer (at constant speed). The stability of this model seems to be difficult to prove, and one usually (carefully) ignores this issue. In particular, it seems to us that the method of Section 4 (proving stability through an auxiliary $M/G/\infty$-type batch queue), does not work in this case. A variant on the greedy-server system is the "semi-greedy-server" model where the greedy server can only decide to change direction when he is busy. For this model stability is perhaps easier to prove. However, when one assumes that the greedy-server system is stable when the traffic intensity is less than 1, one can, by the methods developed here and in [10], find lower bounds for the expected number of waiting customers and their mean measure. Finally, the case where π is an atomic distributions on [0, 1] could shed some new light on the cyclic server model.

Appendix

Throughout the appendix and throughout the paper $(\Omega, \mathscr{H}, P)$ denotes the probability space in the background. For any topological space E, $\mathscr{B}(E)$ denotes either the Borel σ-algebra on E or the set of non-negative measurable functions on E. The indicator function corresponding to a set A is written as I_A. The Lebesgue measure of a Borel set A of $\mathbb{R}^m$ $(m \in \mathbb{N})$ is denoted by $Leb(A)$. We will frequently write μf for the integral of a function f with respect to a (random) measure μ. First, we give some basic definitions and results on random measures. A reference is for example [4].

Let $(E, \mathscr{E})$ be a measurable space, for definiteness we assume that E is Polish and that $\mathscr{E}$ is the Borel σ-algebra on E (or the set of non-negative $\mathscr{E}$-measurable numerical functions). A mapping M from $\Omega \times \mathscr{E}$ into $\overline{\mathbb{R}}_+$ is called a *random measure* on $(E, \mathscr{E})$ if

a) $B \to M(\omega, B)$ is a measure on $(E, \mathscr{E})$ for every $\omega \in \Omega$, and
b) $\omega \to M(\omega, B)$ is a random variable for every $B \in \mathscr{E}$.

According to Fubini's theorem

$$Mf(\omega) = \int_E M(\omega, dx) f(x) \ , \qquad \omega \in \Omega \ , \tag{A.1}$$

defines a positive random variable Mf for each positive $\mathscr{E}$-measurable function f, and

$$\mu(A) = EM(A) = \int P(d\omega) M(\omega, A) \ , \qquad A \in \mathscr{E} \tag{A.2}$$

defines a measure μ on $(E, \mathscr{E})$, which is called the *mean measure* of M.

M is called a *random counting measure* if for almost every ω, there exists a countable set $D(\omega)$ such that

$$M(\omega, A) = \sum_{x \in D(\omega)} \delta_x(A) \ , \tag{A.3}$$

where δ_x denotes the Dirac measure at $x \in E$. When the sets $D(\omega)$ are locally finite, M is called a *point process*. A random measure M is said to be a *Poisson random measure* (on $(E, \mathscr{E})$) with *mean measure* (or *intensity measure*) μ if

(a) $M(A)$ has the Poisson distribution with mean $\mu(A)$ for all $A \in \mathscr{E}$, and
(b) $M(A_1), \ldots, M(A_n)$ are independent whenever $A_1, \ldots, A_n \in \mathscr{E}$ are disjoint, this being true for every $n \geq 2$.

Theorem A.1: The probability law of random measure M on $(E, \mathscr{E})$ is completely specified by its Laplace functional L defined by

$$Lf = Ee^{-Mf} \ , \qquad f \in \mathscr{E} \ .$$

Moreover, the Laplace functional of a Poisson random measure on $(E, \mathscr{E})$ with mean measure μ is given by

$$Lf = \exp - \int_E \mu(dx)(1 - e^{-f(x)}) \ , \qquad \text{for all } f \in \mathscr{E} \ . \tag{A.4}$$

We further restrict ourselves to the case where $E = \mathbb{R}_+$. Let $\mathscr{F} = (\mathscr{F}_t)_{t \geq 0}$ be an augmented and right-continuous filtration (cf. [11]). Adaptedness, martingales, compensators etc. are always with respect to this filtration. Let $\mathbb{D}$ denote the collection of all real valued adapted processes on $\mathbb{R}_+$ whose every path $t \to X_t(\omega)$ is right-continuous and left-limited. Let $\mathbf{L}$ denote the collection of all adapted real valued processes on $\mathbb{R}_+$ whose every path is left-continuous and right-limited.

We give some basic results in stochastic integration. The definitions and proofs can be found for example in [11].

Theorem A.2: (Itô's formula) *Let $X \in \mathbb{D}$ be a semi-martingale and let $f \in$ be a twice continuously differentiable function, then the process $f(X)$ is a semi-martingale as well, and*

$$f(X_t) = f(X_0) + \oint_0^t f'(X_{s-})\, dX_s + \frac{1}{2}\int_0^t f''(X_{s-}) d[X, X]_s$$

$$+ \sum_{0 < s \leq t} \left\{ f(X_s) - f(X_{s-}) - f'(X_{s-})\Delta X_s - \frac{1}{2} f''(X_{s-})(\Delta X_s)^2 \right\} \ , \tag{A.5}$$

where $\oint$ denotes the Itô-integral sign, $[X, X]$ the quadratic variation process of X, $\Delta X_s = X_s - X_{s-}$ for $s > 0$ and $\Delta X_0 = 0$.

Let B denote the standard Brownian motion, and let $F \in \mathbf{L}$. It is well-known that, since B is a local L^2-martingale, the stochastic process $\oint F\,dB$ is a local L^2-martingale as well, with quadratic variation

$$[\oint F\,dB, \oint F\,dB]_t = \int_0^t F_s^2\,ds \ . \tag{A.6}$$

In particular we have

$$E \oint_0^t F_s\,dB_s = 0 \ , \qquad \text{for all } t \geq 0 \ . \tag{A.7}$$

The next theorem is one of the main theorems of stochastic integration w.r.t. point processes, where now a point process N is considered as a counting process instead of a random counting measure.

Theorem A.3: Let $N \in \mathbb{D}$ be a point process with compensator A. Then for all $F \in \mathbf{L}$,

$$E \int F_u\,dN_u = E \int F_u\,dA_u \ . \tag{A.8}$$

In many cases of practical interest A is given by

$$A_t = \int_0^t \lambda_s\,ds \ , \qquad t \geq 0 \ ,$$

where (λ_s) is called the *stochastic intensity* of N. When N is a renewal (counting) process, then A is given by Theorem 13.2 III of [4].

References

[1] Asmussen S (1987) *Applied Probability and Queues.* Wiley & Sons, New York

[2] Coffman Jr. EG, Gilbert EN (1986) A Continuous Polling System with Constant Service Times. *IEEE Trans. Inf. Th.* it-32 (4):584–591

[3] Coffman Jr EG, Gilbert EN (1987) Polling and Greedy Servers on a Line. *Queueing Systems* 2:115–145

[4] Daley DJ, Vere-Jones D (1988) *An Introduction to the Theory of Point Processes.* Springer-Verlag, New York
[5] Eisenberg M (1972) Queues with periodic service and change-over times. *Oper. Res.* 20:440–451
[6] Franken P, König D, Arndt U, Schmidt V (1981) *Queues and Point Processes.* Akademie-Verlag, Berlin
[7] Fuhrmann S, Cooper RB (1985a) Stochastic Decomposition in the $M/G/1$ Queue with Generalized Vacations. *Oper. Res.* 33:1117–1119
[8] Fuhrmann S, Cooper RB (1985b) Application of Decomposition Principle in $M/G/1$ Vacation Model to Two Continuum Cyclic Queueing Models—Especially Token-Ring LANs. *AT&T Technical Journal* 64 (5):1091–1098
[9] Gutjahr R (1977) Eine Bemerkung zur Existenz unendlich vieler "Leerpunkte" in Bedienungssystemen mit unendlich vielen Bedienungsgeräten. *Math. Operationsforsch. Statist. Ser. Optimization* 8:245–251
[10] Kroese DP, Schmidt V (1992) A Continuous Polling System with General Service Times. *Ann. of Appl. Probab* 2(4):906–927
[11] Liptser RS, Shiryayev AN (1978) *Statistics of Random Processes. II: Applications.* Springer Verlag, New York
[12] Mack C, Murphy T, Webb NL (1957) The efficiency of N machines uni-directionally patrolled by one operative when walking time and repair times are constant. *J. Roy. Statist. Soc.* (*B*) 19:166–172
[13] Stolyar A, Coffman Jr. EG (1991) Polling on a Line: General Service Times. Preprint
[14] Takagi H (1986) *Analysis of Polling Systems.* The MIT Press, Cambridge Massachusetts

Received: July 1992