Single-Server Queues with Spatially Distributed Arrivals

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, 89069 Ulm
Germany

Abstract

Consider a queueing system where customers arrive at a circle according to a homogeneous Poisson process. After choosing their positions on the circle, according to a uniform distribution, they wait for a single server who travels on the circle. The server's movement is modelled by a Brownian motion with drift. Whenever the server encounters a customer, he stops and serves this customer. The service times are independent, but arbitrarily distributed. The model generalizes the continuous cyclic polling system (the diffusion coefficient of the Brownian motion is zero in this case) and can be interpreted as a continuous version of a Markov polling system. Using Tweedie's lemma for positive recurrence of Markov chains with general state space, we show that the system is stable if and only if the traffic intensity is less than one. Moreover, we derive a stochastic decomposition result which leads to equilibrium equations for the stationary configuration of customers on the circle. Steady-state performance characteristics are determined, in particular the expected number of customers in the system as seen by a travelling server and at an arbitrary point in time.

Keywords: Single-server queue, spatially distributed arrival points, travelling server, Brownian motion, embedded Markov chain, stability, Tweedie's lemma, regenerative processes, stochastic decomposition, equilibrium equations, mean queue length.

1 Introduction

Service systems with spatially distributed arrivals have been investigated in queueing theory for a long time. In the earlier papers it is assumed that customers arrive at a finite number N of spatially distributed service stations (see e.g. Boxma and Groenendijk [11], Cooper and Murray [15], Eisenberg [17], Fuhrmann[20], Koheim and Meister [28], Kühn [32], Takagi [42]). In most of these papers, a server visits consecutive stations according to a (deterministic) *polling table*. This is a function $f : \{1,\ldots,m\} \to \{1,\ldots,N\}$, where m denotes the length of the polling period and $f(k)$ the number of the station that is served in the k th step. An important special case is the (deterministic) cyclic polling system, where $m = N$ and $f(k) = (k \bmod N)+1$. Later on, also models were investigated in which the server visits the stations in a random way. For example in Altman and Levy [3], Keilson and Servi [25], and Tedijanto [43] a cyclic Bernoulli polling model is considered. Although the server uses here a deterministic routing scheme to travel from station to station, the server only actually 'visits' a station (and performs service if necessary) with a certain probability which can depend on that station. A more general model of Markov polling is considered in Borovkov and Schassberger [10] and Boxma and Westrate [12] where it is assumed that the server chooses the next station according to a Markov transition matrix $(p_{ij}\ ;\ i,j \in \{1,\ldots,N\})$. For the special case $p_{ij} = p_j$, see Kleinrock and Levy [26].

When the number N of stations becomes large, it may be difficult to determine performance characteristics of such systems, in particular when dealing with a random routing scheme. Thus, the question arises whether it is easier to investigate an approximative continuous model (describing the case where $N \gg 1$). In the present paper we give a partial solution of this problem, assuming that arriving customers are distributed over a circle C according to a uniform distribution on C and that the server's movement on the circle is governed by a Brownian motion with drift. This queueing system can be interpreted as a continuous version of a special case of the Markov polling models investigated in Borovkov and Schassberger [10], Boxma and Westrate [12] and Kleinrock and Levy [26]. Namely, the Markov polling model in which the arrival rates, the distributions of service times and the walking times are the same for each station and where the server visits the stations according to a random walk, with $p_{ij} > 0$ only if $j = (i \bmod N) + 1$ or $i = (j \bmod N)+1$. In order to describe general Markov polling by an approximative continuous model, it should be assumed that the server's movement on the circle is governed by a more general diffusion process. This will be the subject of further research.

In Section 2 we present the model of the Brownian server, where the state of the system is described by a random point configuration on the unit interval. Next, considering the system only at those points in time when a customer departs from the system, we arrive at a discrete-time Markov chain (R_n) with non-countable state space. Using Tweedie's lemma for positive recurrence of Markov chains with general state space, we show in Section 4 that (R_n) is ergodic if the traffic intensity is less than one. This gives the existence of a stationary regenerative process which describes the steady-state behavior of the system in continuous time, see Section 5. In Section 6 we derive a stochastic decomposition result by introducing three stochastic clock processes. This leads to equilibrium equations for the stationary configuration of customers in the system. In a sense, these equilibrium equations correspond to the rate conservation law for stationary semi-martingales (see e.g. Bardhan and Sigman [7], Mazumdar et al. [35], Miyazawa [36]). With the help of the equilibrium equations, we can determine not only the stationary expected number of customers in the system given that the server is travelling, but also (if the second moment of service times is finite) the stationary expected number of customers being in the system at an 'arbitrary point in time', see Sections 7 and 8. It turns out that these two performance characteristics depend in a quite natural way on the parameters that govern the server's movement (see Theorems 4 and 6). Moreover, formula (8.6) for the latter expectation is in accordance with the decomposition property derived in Fuhrmann and Cooper [21] for the M/GI/1 queue with generalized vacations.

2 The Model

We consider a queueing system where customers arrive at a circle according to a homogeneous Poisson process with intensity a. The arrival epoch of the nth customer is denoted by $T_n,\ n \in I\!N = \{1, 2, \ldots\}$. The arriving customers drop on a circle C with circumference one, according to a uniform distribution independently of each other and of any other random mechanism. Here they wait for a single server who travels on the circle. The server carries out a Brownian motion with drift on C. We assume that the Brownian motion has drift parameter $\alpha^{-1} \geq 0$ and variance parameter $\sigma^2 \geq 0$, and that it is stochastically independent of everything else. Whenever the server meets a customer he stops and serves this customer. The consecutive service times are i.i.d. non-negative random variables with distribution function F and first moment $e_1 < \infty$. The service times are also assumed to be independent of the arrival process and of the movement of the server (when travelling). After service completion, the customer is removed from the circle and the server resumes his walk. Notice that the movement of the server neither depends on the number of customers being actually present in the system nor on their positions.

The *cyclic polling system* considered in Kroese and Schmidt [29], where the server moves uni-directionally with constant speed $\alpha^{-1} > 0$ (when not serving) can be seen as a special case ($\sigma^2 = 0$) of the *Brownian server system,* described above. In other words, the Brownian server could also be interpreted as a polling one whose movement is disturbed from the outside. Furthermore, for $\sigma^2 > 0$ and $\alpha^{-1} = 0$ the *drunken server model* of Kroese and Schmidt [30] appears, i.e. the server's movement is governed by a Brownian motion with zero drift.

We analyze this non-standard class of queueing systems by describing the positions of waiting customers at time t, relative to the actual position of the server, by the atoms of a random counting measure W_t on C, where we assume that at time $t = 0$ we start with the empty state, i.e. $W_0(C) = 0$. We thus 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 way that both 0 and 1 are identified with the actual 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 $[0, 1]$. Specifically, if for a realization $\omega \in \Omega$, the server is in s at time t, and if n denotes the number of customers on the circle at that time, at locations $\xi_1, \ldots, \xi_n$ (for $n \geq 1$), then $W_t(\omega, .)$ is the counting measure on $[0, 1]$ with atoms at $v_i, i = 1, \ldots, n$, where

$$v_i = \begin{cases} \xi_i - s & \text{if } \xi_i \geq s \\ 1 - (s - \xi_i) & \text{if } \xi_i < s. \end{cases}$$

In other words,

$$W_t(\omega, B) = \begin{cases} \sum_{i=1}^{n} \delta_{v_i}(B) & \text{if } n \geq 0 \\ 0 & \text{if } n = 0, \end{cases} \tag{2.1}$$

for every Borel set $B \in \mathcal{B}([0, 1])$, where

$$\delta_v(B) = \begin{cases} 1 & \text{if } v \in B \\ 0 & \text{if } v \notin B, \end{cases}$$

see Kroese and Schmidt [29], [30] for further details. From the view point of the server (which we will take from now on), the location of the server is fixed at 0 (or 1) while the circle 'rotates' when the server is not busy. The customer paths form a stochastic flow on $\mathbb{R}_+ \times [0, 1]$. Moreover, from the server's perspective, the positions of arriving customers form a Poisson random measure on $\mathbb{R}_+ \times [0,1]$ with intensity measure $(a\nu) \times \pi$, where ν denotes the Lebesgue measure on $\mathcal{B}(\mathbb{R}_+)$, and π the uniform distribution on C.

We adopt the following notational conventions throughout this paper: (X_t) indicates the continuous-time stochastic process $(X_t\,; t \in \mathbb{R}_+)$, and (X_n) denotes the random sequence $(X_n; n \in \mathbb{N})$.

3 Formulation of the Stability Problem

In Section 5 we investigate the question under what condition there exists a time-stationary measure-valued process $(\tilde{W}_t)$ such that the finite-dimensional distributions of the processes $(W_{t+h}; t \geq 0)$ converge in variation to the corresponding finite-dimensional distributions of $(\tilde{W}_t)$ as h tends to infinity. We show that this question can be answered positively provided that

$$ae_1 < 1\,. \tag{3.1}$$

With the help of a certain majorization technique, we considered related problems in Kroese and Schmidt [29], [30] where we showed for the special cases $\alpha^{-1} > 0$, $\sigma^2 = 0$ and $\alpha^{-1} = 0$, $\sigma^2 > 0$ that, under (3.1), the measure-valued process (W_t) is regenerative with regeneration periods that have absolutely continuous distribution and finite expectation. Now, we present a new and, as we believe, more elegant proof of this fact for the general case $(\alpha^{-1}, \sigma^2) \in \mathbb{R}_+^2 \setminus \{(0,0)\}$, by applying Tweedie's generalization of Foster's criterion for positive recurrence of a certain embedded measure-valued Markov chain (see Tweedie [45]).

Note that condition (3.1) is in accordance with the corresponding stability condition for the 'usual' M/G/1 queue. Nevertheless, the classical results for the M/G/1 queue can not be used directly because in our model, besides service times, walking times appear. Furthermore, note that in the stability condition (3.1) the parameters α^{-1} and σ^2 do not appear. At the first sight, this might be surprising. However, an explanation for this is that a slowly walking server causes the circle to 'fill up' with a lot of customers. The fact that walking times become negligible when the number of customers in the system is large, is essentially the reason why the stability of the queue is not influenced by the walk times and, in particular, not by α^{-1} and σ^2. The necessity of condition (3.1) is quite clear. This can be seen by comparing our queue with an ordinary M/G/1 queue (with the same arrival rate and service time distribution). Such an M/G/1 queue processes the customers faster than our queue, and is stable only if $ae_1 < 1$.

For queueing systems where arriving customers are continuously distributed in a non-countable space C, a formal proof of stability is often omitted in the literature (see e.g. Bertsimas and van Ryzin [9], Coffman and Gilbert [13], Fuhrmann and Cooper [22], where this problem has been left open). However, for a related class of queues stability conditions have been known for a long time (see Eisenberg [17], Kühn [32], Takagi [42]). In these queues customers arrive at a finite number of points in C, so-called stations. Only recently, stability conditions for such systems have been systematically analyzed and generalized (see Altman et al. [1], [2], Borovkov and Schassberger [10], Fricker and Jaibi [19], Georgiadis and Szpankowski [23], Resing [38], Spieksma and Tweedie [41]). For the continuous cyclic polling system, we gave a proof of stability in Kroese and Schmidt [29], see also Coffman and Stolyar [14]. There a stable auxiliary $M/GI/\infty$ queue was constructed which works slower than the original polling system and for which each empty point (i.e. an arrival epoch at which the system is empty) is simultaneously an empty point of polling system. Unfortunately, this construction does not work for the Brownian server. Next, in Kroese and Schmidt [30] a different, rather cumbersome majorization has been used to show that also in the drunken-server case (W_t) is regenerative if and only if (3.1) holds.

The present approach to prove stability of the Brownian server model has the advantage that it works equally elegant for both the cyclic polling and the drunken (zero drift) server. The idea is to consider the embedded Markov chain (R_n) with $R_0 = W_0$ and, for $n \geq 1, R_n = W_{U_n}$ where U_n denotes the epoch of the n th service completion. The state space $(E, \mathcal{E})$ of this Markov chain is the set E of all finite counting measures on $[0,1]$ endowed with the σ-algebra $\mathcal{E}$ of subsets of E generated by the family of sets $\{\varphi \in E : \varphi([a,b)) = n\}$ with $0 \leq a \leq b \leq 1, n \in I\!N_0 = \{0,1,2,\ldots\}$. It is well-known that E is a metric space which is separable and complete, and that $\mathcal{E}$ is the σ-algebra of its Borel sets (see e.g. Section 1.15 of Matthes et al. [34]).

Next, Tweedie's lemma (see [45]) is used to show that, for a certain integer $\zeta > 0$, the Markov chain (X_n) with $X_n = R_{\zeta n}$ is ergodic and, consequently, that the empty state is positive recurrent provided that (3.1) holds. Because empty points are regeneration epochs, we are able to construct a time-stationary regenerative process $(\tilde{W}_t)$ such that the finite-dimensional distributions of the processes $(W_{t+h}; t \geq 0)$ converge in variation to the corresponding finite-dimensional distributions of $(\tilde{W}_t)$ as $h \to \infty$.

Independently of the present paper, in Altman and Levy [4] stability of another continuous polling system has been investigated where the server is assumed to move according to certain 'gated-greedy' and 'gated-scan' policy, respectively.

4 The Embedded Markov Chain (X_n)

Let $\boldsymbol{P}_x$ be a probability measure under which the embedded Markov chain (X_n) (as defined in the previous section) starts at state $x \in E$ at time $n = 0$. The expectation with respect to $\boldsymbol{P}_x$ is denoted by $\boldsymbol{E}_x$. Let $\tau_A = \min\{k \geq 1 : X_k \in A\}$ denote the hitting time of a set $A \in \mathcal{E}$. We show that if (3.1) holds, the following conditions (4.1)–(4.4) are satisfied: There exist a set $A \in \mathcal{E}$, a probability measure η on $\mathcal{E}_A = \{D \cap A : D \in \mathcal{E}\}$, a real number $\beta \in (0,1)$ and an integer $m_0 \in I\!N$ such that

$$\boldsymbol{P}_x(\tau_A < \infty) = 1 \quad \text{for every} \quad x \in E, \tag{4.1}$$

$$\sup_{x \in A} \boldsymbol{E}_x \tau_A < \infty, \tag{4.2}$$

and

$$\inf_{x \in A} \boldsymbol{P}_x(X_{m_0} \in B) \geq \beta \eta(B) \quad \text{for every} \quad B \in \mathcal{E}_A. \tag{4.3}$$

Furthermore, there exist integers $n_1, n_2, \ldots$ and an integer k such that

$$\int_A \boldsymbol{P}_x(\tau_A = n_i)\, \eta(dx) > 0, \quad \gcd\{m_0 + n_1, m_0 + n_2, \ldots, m_0 + n_k\} = 1 \tag{4.4}$$

where $\gcd(D)$ denotes the greatest common divisor of the set D.

Let

$$P^n(x, D) = \boldsymbol{P}_x(X_n \in D) \tag{4.5}$$

denote the n-step transition kernels of (X_n), and denote $P^1(x, D)$ by $P(x, D)$. It is well-known (see e.g. Athreya and Ney [6], Nummelin [37]) that from (4.1)–(4.4) it follows that the integral equation

$$\mu(D) = \int_E P(x, D)\, \mu(dx), \quad D \in \mathcal{E}, \tag{4.6}$$

has a unique solution μ in the set of all probability measures on $\mathcal{E}$ and that

$$\sup_{D \in \mathcal{E}} |P^n(x, D) - \mu(D)| \longrightarrow 0 \quad \text{if} \quad n \to \infty \tag{4.7}$$

for every $x \in E$. Moreover, μ is a stationary initial distribution of the Markov chain (X_n), and no further stationary initial distribution exists, see also Lindvall [33], Sigman and Wolff [40].

Now, assuming that (3.1) holds we show that the conditions (4.1)–(4.4) are fulfilled. In connection with this we use the following generalization of Foster's criterion for positive recurrence (cf. Section 6 of Tweedie [45]).

Lemma 1 Let $A \in \mathcal{E}, \epsilon > 0$ and $g : E \to I\!R_+$ be a $\mathcal{E} - \mathcal{B}(I\!R_+)$ measurable function such that

$$\int_E g(y)P(x,dy) \le g(x) - \epsilon \quad \text{for every} \quad x \in A^c. \tag{4.8}$$

Then

$$\boldsymbol{E}_x \tau_A \le \frac{g(x)}{\epsilon} \quad \text{for every} \quad x \in A^c \tag{4.9}$$

and

$$\boldsymbol{E}_x \tau_A \le 1 + \epsilon^{-1} \int_{A^c} g(y)P(x,dy) \quad \text{for every} \quad x \in A. \tag{4.10}$$

Thus, (4.1) and (4.2) are proved if we find a set $A \in \mathcal{E}$, a constant $\epsilon > 0$, and a function $g : E \to I\!R_+$ such that (4.8) and

$$\sup_{x \in A} \int_{A^c} g(y)P(x,dy) < \infty \tag{4.11}$$

hold. For every $j \in I\!N_0$, let $A_j = \{\varphi \in E : \varphi([0,1)) = j\}$ denote the set of counting measures on $[0,1)$ having exactly j atoms.

For clarity we consider the special case of the *cyclic polling server* first, i.e. the case where $\sigma^2 = 0$. Let

$$\zeta = \min\{j \in I\!N : j > \frac{a\alpha}{1 - ae_1}\} \tag{4.12}$$

We put

$$A = \bigcup_{j=0}^{\zeta-1} A_j \quad \text{and} \quad g(x) = j \quad \text{for} \quad x \in A_j. \tag{4.13}$$

Then, for every $j \ge \zeta$ and for every $x \in A_j$ we have

$$\int_E g(y)P(x,dy) = \sum_{k=j-\zeta}^{\infty} \int_{A_k} k\,P(x,dy) = \sum_{k=0}^{\infty} \int_{A_{j-\zeta+k}} (j-\zeta+k)P(x,dy)$$

$$= (j-\zeta) + \sum_{k=0}^{\infty} k\,P(x, A_{j-\zeta+k}),$$

where the sum $\sum_{k=0}^{\infty} kP(x, A_{j-\zeta+k})$ is equal to the expected number of customers which arrive in the time interval $(U_n, U_{n+\zeta})$ under the condition that $W_{U_n} = x$. The length of this time interval is the sum of ζ service times, which are independent of the Poisson input, and of a random variable which is bounded by the time which the server would need to go around the whole circle when there is no customer. This time is equal to α. Thus, for every $x \in A_j$ with $j \geq \zeta$ we have

$$\sum_{k=0}^{\infty} k\, P(x, A_{j-\zeta+k}) \leq a(\zeta e_1 + \alpha) \tag{4.14}$$

and, consequently,

$$\int_E g(y)P(x, dy) \leq j + \{a(\zeta e_1 + \alpha) - \zeta\} = g(x) - \epsilon\ ,$$

where $\epsilon = \zeta(1 - ae_1) - a\alpha > 0$, i.e. (4.8) holds. Furthermore, for every $x \in A_j$ with $j < \zeta$ we have

$$\int_{A^c} g(y)P(x, dy) = \sum_{k=\zeta}^{\infty} \int_{A_k} k\, P(x, dy) \leq \sum_{k=0}^{\infty} kP(x, A_k),$$

where the last term is the expected number of customers in the system at time $U_{n+\zeta}$ under the condition that $W_{U_n} = x$. Because $x \in A_j$ with $j < \zeta$, this expected value is not greater than the sum of the following expectations: The expected number of all customers, which arrive during the ζ service times between U_n and $U_{n+\zeta}$, and the expected number of those customers, which arrive during the corresponding ζ walk times. Thus, because the sum of these walk times is not greater than $\zeta\alpha$, we have

$$\int_{A^c} g(y)P(x, dy) \leq a(\zeta e_1 + \zeta\alpha) \tag{4.15}$$

for every $x \in \bigcup_{j=0}^{\zeta-1} A_j$ and, consequently, (4.11) holds.

Next, we show that (4.3) is fulfilled with $m_0 = 1$ and η the probability measure on $\mathcal{E}_A$ defined by

$$\eta(B) = \begin{cases} 1 & \text{if } \mathbf{0} \in B \\ 0 & \text{if } \mathbf{0} \notin B, \end{cases} \tag{4.16}$$

where $\mathbf{0}$ denotes the zero-measure on [0,1], with $\mathbf{0}([0,1]) = 0$. We therefore have to show that

$$\inf_{x\in A} \boldsymbol{P}_x(X_1 = \mathbf{0}) > 0. \tag{4.17}$$

The assumption $x \in \bigcup_{j=0}^{\zeta-1} A_j$ means that, at the considered departure epoch, less than ζ customers are left in the system. Thus, for realizing the event $\{X_1 = \mathbf{0}\}$, it suffices that during ζ service times and during ζ walk times no customer arrives. Consequently,

$$\inf_{x\in A} \boldsymbol{P}_x(X_1 = \mathbf{0}) \geq \int_0^\infty \mathrm{e}^{-a(t+\zeta\alpha)}\, dF^\zeta(t) > 0, \tag{4.18}$$

where F^ζ denotes the ζ-fold convolution of the distribution function F of service times.

Now, it remains to show that (4.4) holds with $m_0 = 1$. However, this easily follows from the fact that $\boldsymbol{P}(\tau_A = n | X_0 = \mathbf{0}) > 0$ for every integer $n \in I\!N$.

For the general *Brownian server* case, we will use similar arguments to show that the conditions (4.1)–(4.4) are fulfilled provided that (3.1) holds. Let (B_t) be a Brownian motion with drift $-\alpha^{-1}$ and variance parameter $\sigma^2 \geq 0$. Assume that $B_0 = 0$. Let V denote the time at which (B_t) first attains one of the values 1 or -1, i.e.

$$\boldsymbol{P}(V \leq t) = \boldsymbol{P}(\max_{0<u\leq t} |B_u| \geq 1).$$

By $\gamma = \boldsymbol{E}\, V$ we denote the expectation of the non-negative random variable V. It is well-known (see e.g. Section 15.3 in Karlin and Taylor [24]) that $\gamma < \infty$. Now, by replacing the definition (4.12) of ζ by

$$\zeta = \min\{j \in I\!N : j > \frac{a\gamma}{1 - ae_1}\} \tag{4.12'}$$

we can proceed completely analogously in showing that, for the Brownian server, the conditions (4.1)–(4.4) are fulfilled when (3.1) holds. In particular, this is due to the fact that, for $x \in A_j$ with $j \geq \zeta$, the sum $\sum_{k=0}^\infty kP(x, A_{j-\zeta+k})$ is not greater than the expected number of customers which arrive during ζ service times and during ζ switching times, where the sum of these switching times is stochastically bounded by V. Thus,

$$\sum_{k=0}^\infty k\, P(x, A_{j-\zeta+k}) \leq a(\zeta e_1 + \gamma),$$

and (4.8) follows in the same way as in the case of the cyclic polling server. Moreover, the inequalities (4.15) and (4.18) remain true in a slightly modified form. Namely, putting again $A = \bigcup_{j=0}^{\zeta-1} A_j$ and $g(x) = j$ for $x \in A_j$, we have

$$\int_{A^c} g(y) P(x, dy) \leq a(\zeta e_1 + \zeta\gamma)$$

and

$$\inf_{x\in A} \boldsymbol{P}_x(X_1 = \boldsymbol{0}) \geq \int_0^\infty \int_0^\infty \mathrm{e}^{-a(t+s)} dF^\zeta(t) dG^\zeta(s) > 0,$$

where G^ζ denotes the ζ-fold convolution of the distribution function G of V. This gives (4.11) and, consequently, (4.1)–(4.4).

Remark Unfortunately, for a further kind of server, the so-called *greedy server*, who always walks (at constant speed α^{-1}) towards the *nearest* customer, the question remains open whether analogous arguments can be used. In Coffman and Gilbert [13] the conjecture is formulated that (3.1) is sufficient for stability of the greedy server, but no proof is given. The main difficulty with the greedy-server model is that the walking discipline is dynamic, in the sense that it depends on the actual state of the system. Any newly arriving customer could change the direction of the server's movement. Clearly, it would be possible to provide an analogous proof as above if also for the greedy server one could show that the mean number of customers which arrive during the walk times of the server within the interval $(U_n, U_{n+\zeta})$ is bounded uniformly with respect to ζ under the condition that at time U_n we start from a state $x \in A_j$ with $j \geq \zeta$, i.e. $W_{U_n}([0,1]) \geq \zeta$. On the other hand, for a certain *gated-greedy server* it has been recently shown in Altman and Levy [4] that (3.1) is sufficient for stability. In this case the server operates in certain cycles, where in each cycle the server serves a customer if and only if he is present in the system at the beginning of the cycle. Within the cycles, the greedy-server approach is considered, i.e. within the cycles the server always picks for service the closest customer for it. Note however that from the results of Altman and Levy [4] it can not be concluded that (3.1) is sufficient for stability of the usual greedy server. A light-traffic approximation for the greedy server has been derived in Kroese and Schmidt [31].

5 Limiting Behavior in Continuous Time

By the above consideration it follows in particular that we have for the general Brownian server model

$$\sup_{x\in A} \boldsymbol{E}_x \tau_A < \infty \quad \text{with} \quad A = \bigcup_{j=0}^{\zeta-1} A_j, \tag{5.1}$$

for a certain natural number ζ, provided that (3.1) holds. Now we use this result to show that, in continuous time, the expected amount of time which the system needs to return to the empty state is finite.

We assume again that at time $t = 0$ we start with the empty state. By $S_1, S_2, \ldots$ we denote the subsequence of those departure epochs at which the Markov chain (X_n) takes values in the set $A = \bigcup_{j=0}^{\zeta-1} A_j$. Then, with the notation $S_0 = 0$, we have for every $n \in I\!N$,

$$\boldsymbol{E}\,(S_n - S_{n-1}) = \int_A \boldsymbol{E}_x(S_n - S_{n-1})\, \boldsymbol{P}\,(W_{S_{n-1}} \in dx).$$

Observe that the random variable $S_n - S_{n-1}$ is the sum of $\zeta\tau_A$ service times and of $\zeta\tau_A$ walk times, where the latter ones can be bounded from the above by i.i.d. copies of the random time V at which the Brownian motion (B_t) first attains one of the values 1 or -1. Thus, by Wald's lemma, we get

$$\boldsymbol{E}_x(S_n - S_{n-1}) \le \zeta(e_1 + \gamma)\ \boldsymbol{E}_x\tau_A.$$

Because of (5.1) this gives

$$\boldsymbol{E}\,(S_n - S_{n-1}) \le c < \infty \tag{5.2}$$

uniformly in n. Denote by τ_0 the smallest (random) natural number such that the interval $(S_{\tau_0-1}, S_{\tau_0}]$ contains a departure epoch immediately after which the system is empty. Then, for the time T which the system needs to return to the empty state, the inequality $T \le \sum_{n=1}^{\tau_0}(S_n - S_{n-1})$ holds. Thus, we get the bound

$$\boldsymbol{E}\,T \le c\ \boldsymbol{E}\,\tau_0$$

where again Wald's lemma and (5.2) have been used. Furthermore, we have

$$\boldsymbol{E}\,\tau_0 = \sum_{k=1}^{\infty} \boldsymbol{P}\,(\tau_0 \ge k) \le \sum_{k=1}^{\infty} \left(1 - \int_0^{\infty}\int_0^{\infty} \mathrm{e}^{-a(t+s)} dF^{\zeta}(t) dG^{\zeta}(s)\right)^k < \infty.$$

This gives that the expected return time $\boldsymbol{E}\,T$ is finite. Thus, using the fact that each departure epoch immediately after which the system is empty, is a regeneration epoch of the queueing process (W_t), we arrive at the following result.

Theorem 1 The measure-valued process (W_t) of the configuration of customers on the circle is regenerative with regeneration periods that have absolutely continuous distributions and finite expectations.

Corollary 1 There exists a time-stationary regenerative process $(\tilde{W}_t)$ which takes its values in the set of all finite counting measures on the interval [0,1] such that the finite-dimensional distributions of the processes $(W_{t+h}; t \ge 0)$ converge in variation to the corresponding finite-dimensional distributions of $(\tilde{W}_t)$ as h tends to infinity.

The *proof* of this corollary follows from Theorem 6.3.1 in Berbee [8], where it is shown that it suffices to prove that a regeneration period has finite expectation and that its distribution is *spread out*, i.e. that for a certain natural number n the n fold convolution of this distribution is nonsingular with respect to the Lebesgue measure on the real line (see also Corollary VI.1.4 in Asmussen [5] for the convergence of the one-dimensional distributions). For a proof that the limit process $(\tilde{W}_t)$ is again regenerative, see Thorisson [44].

6 Equilibrium Equations: Stochastic Decomposition

Important information about the (general) Brownian server system can be obtained via equilibrium equations. Under the assumptions that the drift α^{-1} or the variance parameter σ^2 are equal to zero, characteristics of the stationary configuration of waiting customers have already been investigated in Kroese and Schmidt [29], [30]. It turns out that some results obtained in these two previous papers can be extended to the Brownian-server case with general (non-zero) drift.

In connection with this, besides the process (W_t), we will consider three other measure-valued processes (Q_t), (Q_t^0) and (Q_t^1). They are defined by means of random clocks. First, for any realization $\omega \in \Omega$, let $\bar{B} = \bar{B}_\omega$ denote the set of times where the server is not busy. And for $i \in \{0,1\}$, let $B^i = B^i_\omega$ denote the set of times $\{t \geq 0 : W_t(\omega, \{i\}) = 1\}$ where the server is busy with a client at i. Now define the *clock process* (S_t) by

$$S_t(\omega) = \int_0^t \mathbf{1}_{\bar{B}}(x)dx, \tag{6.1}$$

where $\mathbf{1}_{\bar{B}}$ denotes the indicator function of $\bar{B}$. Thus, the clock (S_t) only runs when the server is not busy. Furthermore, let (ν_t) denote the right-continuous functional inverse of (S_t), i.e.

$$\nu_t = \inf\{u \geq 0 : S_u > t\}. \tag{6.2}$$

Then, the process (Q_t) with $Q_t = W_{\nu_t}$ will describe the evolution of the configuration of waiting customers given that the server is walking. In a similar way we define the processes (Q_t^i) with $Q_t^i = W_{\nu_t^i}$, where

$$\nu_t^i = \inf\{u \geq 0 : S_u^i > t\} \tag{6.3}$$

and

$$S_t^i(\omega) = \int_0^t \mathbf{1}_{B^i}(x)dx, \tag{6.4}$$

for $i = 0, 1$. The process (Q^0_t) describes the evolution of the configuration of waiting customers *given* that the server is *busy* serving a customer and *given* that the server reached this customer walking *in* the direction of the orientation of the circle. Analogously, the process (Q^1_t) concerns the case that the customer who is under service was reached from the opposite direction.

By using similar arguments as in Section 3 one can show that the process (Q_t) is regenerative with regeneration periods which have absolutely continuous distribution and finite expectation, provided that (3.1) holds.

Assumption For the rest of this paper we assume that the distribution of the service times is spread out.

Then, if (3.1) is satisfied, proceeding as in Sections 3 to 5 we get that the processes (Q^0_t) and (Q^1_t) are regenerative with regeneration periods having spread-out distribution and finite expectation. Thus, from Theorem 6.3.1 in Berbee [8], it follows that there exist time-stationary limit processes $(\tilde{Q}_t), (\tilde{Q}^0_t)$ and $(\tilde{Q}^1_t)$ of $(Q_t), (Q^0_t)$ and (Q^1_t), respectively. For the rest of the paper we use the notation $W = \tilde{W}_0, Q = \tilde{Q}_0, Q^0 = \tilde{Q}^0_0$ and $Q^1 = \tilde{Q}^1_0$. Let $C_+[0,1]$ and $C^2_+[0,1]$ denote the family of non-negative continuous and twice-continuously differentiable functions on $[0,1]$, respectively. And, for every $f \in C_+[0,1]$ and for every random measure M on $[0,1]$, let Mf denote the integral $\int_0^1 f(x)M(dx)$.

The following two theorems are generalizations of results obtained in Kroese and Schmidt [29] and [30]. Theorem 2 gives an intuitive interpretation of Q, Q^0 and Q^1, whereas Theorem 3 is a kind of *stochastic decomposition* result, linking the Laplace functional of Q to those of Q^0 and Q^1. The proof is completely analogous to the proofs of similar theorems in Kroese and Schmidt [29], [30] and we therefore omit them.

Theorem 2 The stationary probability that the server is not busy serving is given by

$$(6.5) \qquad \lim_{t\to\infty} \boldsymbol{P}(W_t(\{0,1\}) = 0) = \boldsymbol{P}(W(\{0,1\}) = 0) = 1 - ae_1.$$

Moreover, for any $f \in C_+[0,1]$,

$$(6.6) \qquad \lim_{t\to\infty} \boldsymbol{E}(\mathrm{e}^{-W_t f} | W_t(\{0,1\}) = 0) = \lim_{t\to\infty} \boldsymbol{E}\,\mathrm{e}^{-Q_t f} = \boldsymbol{E}\,\mathrm{e}^{-Qf}$$

and, for $i = 0, 1$,

$$(6.7) \qquad \lim_{t\to\infty} \boldsymbol{E}(\mathrm{e}^{-W_t f} | W_t(\{i\}) = 1) = \lim_{t\to\infty} \boldsymbol{E}\,\mathrm{e}^{-Q^i_t f} = \boldsymbol{E}\,\mathrm{e}^{-Q^i f}.$$

Theorem 3 Let $p_i = \boldsymbol{P}(W(\{i\}) = 1)$ for $i \in \{0,1\}$. Then, for any $f \in C_+^2[0,1]$ it holds

$$
\begin{aligned}
0 &= (1-ae_1)\sigma^2(-\frac{1}{2}\boldsymbol{E}\mathrm{e}^{-Qf}Qf'' + \frac{1}{2}\boldsymbol{E}\mathrm{e}^{-Qf}(Qf')^2) \\
&+ (1-ae_1)\boldsymbol{E}\mathrm{e}^{-Qf}(\alpha^{-1}Qf' - \beta) \\
&+ p_0\,\boldsymbol{E}\mathrm{e}^{-Q^0 f}((\mathrm{e}^{f(0)}-1)\frac{\beta L_F(\beta)}{1-L_F(\beta)} - \beta) \\
&+ p_1\,\boldsymbol{E}\mathrm{e}^{-Q^1 f}((\mathrm{e}^{f(1)}-1)\frac{\beta L_F(\beta)}{1-L_F(\beta)} - \beta),
\end{aligned}
\tag{6.8}
$$

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

In the general case it seems to be difficult to determine the steady-state probabilities p_0 and p_1. Note however that this problem does not appear in the special cases $\sigma^2 = 0$ and $\alpha^{-1} = 0$, respectively.

Corollary 2 (Kroese and Schmidt [29]) For the cyclic polling server ($\sigma^2 = 0$) we have $p_0 = ae_1$ and $p_1 = 0$. Consequently,

$$
\begin{aligned}
0 &= (1-ae_1)\boldsymbol{E}\mathrm{e}^{-Qf}(\alpha^{-1}Qf' - \beta) \\
&+ ae_1\boldsymbol{E}\mathrm{e}^{-Q^0 f}((\mathrm{e}^{f(0)}-1)\frac{\beta L_F(\beta)}{1-L_F(\beta)} - \beta).
\end{aligned}
\tag{6.8$'$}
$$

Corollary 3 (Kroese and Schmidt [30]) For the drunken server ($\alpha^{-1} = 0$) we have $p_0 = p_1 = \frac{ae_1}{2}$. Thus,

$$
\begin{aligned}
0 &= (1-ae_1)\sigma^2(-\frac{1}{2}\boldsymbol{E}\mathrm{e}^{-Qf}Qf'' + \frac{1}{2}\boldsymbol{E}\mathrm{e}^{-Qf}(Qf')^2) \\
&- (1-ae_1)\beta\boldsymbol{E}\mathrm{e}^{-Qf} \\
&+ \frac{ae_1}{2}[\boldsymbol{E}\mathrm{e}^{-Q^0 f}((\mathrm{e}^{f(0)}-1)\frac{\beta L_F(\beta)}{1-L_F(\beta)} - \beta) \\
&+ \boldsymbol{E}\mathrm{e}^{-Q^1 f}((\mathrm{e}^{f(1)}-1)\frac{\beta L_F(\beta)}{1-L_F(\beta)} - \beta)].
\end{aligned}
\tag{6.8$''$}
$$

Remark Note that the technique used in Kroese and Schmidt [29], [30] for proving the *equilibrium equations* (6.8′) and (6.8″) is similar to the method used in Bardhan and Sigman [7] for proving a rate conservation law for stationary semi-martingales (see also Mazumdar et al. [35] and Miyazawa

[36]). In both cases, Itô's formula of stochastic integration is used. However, observe that two different kinds of stochastic decomposition have been considered. We have split the time axis with respect to three *non-atomic* random measures (by means of the clocks (S_t), (S_t^0), (S_t^1)), whereas in Bardhan and Sigman [7] *only purely atomic* embedded random measures (induced by jump processes) have been considered.

7 The Expected Number of Customers Seen by the Walking Server

In Section 8 we derive a formula for $\boldsymbol{E}\, W([0,1])$. In Kroese and Schmidt [29], [30] we obtained such formulas for the special cases $\sigma^2 = 0$ and $\alpha^{-1} = 0$, respectively. In both cases we started by analyzing random measure Q, the conditional steady-state configuration of customers given that the server is walking. Using the equilibrium equations $(6.8')$ and $(6.8'')$, respectively, in combination with some further specific features of the cyclic polling server and the zero-drift Brownian server, we derived in both cases the mean measure m_Q of Q, that is the measure on $\mathcal{B}[0,1]$ defined by $m_Q(A) = \boldsymbol{E}\, Q(A)$ for every $A \in \mathcal{B}[0,1]$. Unfortunately, for the general Brownian server case, the derivation of m_Q via the equilibrium equations does not seem to work, mainly because then the weighting probabilities p_0 and p_1 are typically unknown.

Thus, we present a new approach for determining m_Q in the general case. Let (B_t) be a Brownian motion with drift $-\alpha^{-1}$ and variance parameter σ^2, where $(\alpha^{-1}, \sigma^2) \in \mathbb{R}_+^2 \setminus \{(0,0)\}$; $B_0 = 0$. Consider the following *particle system* on $\mathbb{R}_+ \times [0,1]$ which is defined recursively and is governed by (B_t). The particle system starts off with a 'parent particle' at position $x \in (0,1)$ at time $t = 0$. The parent 'walks' in the strip $\mathbb{R}_+ \times [0,1]$ until it hits the level 0 or 1. Then it dies, creating simultaneously a non-negative (random) number of new particles. These particles also move around and they die when they hit 0 or 1, possibly creating second-generation offspring, and so on. The movement of the particles is governed by (B_t) in the following way: Suppose that at time s there are n particles 'alive' at positions $y_1, \ldots, y_n$. Let V denote the first time that one of the processes $(y_i + B_t - B_s;\ t \geq s)$, $i = 1, \ldots, n$ hits the level 0 or 1. Then, the positions of the particles at time $t \in [s, V)$ are given by $y_1 + B_t - B_s, \ldots, y_n + B_t - B_s$. At time V the particle that hits the boundary of the strip $\mathbb{R}_+ \times [0,1]$ dies, and, simultaneously, J new particles are created at positions (if $J(\omega) > 0$)

$$(t, \xi_1), \ldots, (t, \xi_J).$$

Here, J has a compound Poisson distribution with generating function given by

$$\boldsymbol{E}\, z^J = \int_0^\infty \mathrm{e}^{-a(1-z)u} dF(u), \quad 0 \le z \le 1, \tag{7.1}$$

where a is the arrival intensity and F the distribution function of service times. The ξ_i 's are i.i.d. random variables, uniformly distributed on $[0,1]$ and independent of the other stochastic components of the model.

Let L_t^x denote the counting measure describing the positions at time t of those particles that are alive at time t and created (possibly via some intermediate generations) by a single parent particle at time 0 and position x. For any $t \ge 0$ and $f \in C_+[0,1]$, define

$$H_f(t) = \int_0^1 G_f(t,x)\, dx, \tag{7.2}$$

where $G_f(t,x) = \boldsymbol{E}\, L_t^x f$.

From the independence properties of the Poisson arrival process and from the definition (6.1) of the clock process (S_t) it follows that the particle system described above can be used to express the mean measure m_Q in terms of H_f.

Lemma 2 For every $f \in C_+[0,1]$ it holds that

$$\boldsymbol{E}\, Qf = a \int_0^\infty H_f(t)\, dt. \tag{7.3}$$

Proof For every $t \ge 0$ we have

$$\begin{aligned} \boldsymbol{E} Q_t f &= \boldsymbol{E} \sum_{i=1}^\infty \mathbf{1}_{\{T_i \ge t\}} \int_0^1 G_f(t - T_i, x)\, dx \\ &= \boldsymbol{E} \sum_{i=1}^\infty \mathbf{1}_{\{T_i \ge t\}} H_f(t - T_i) = \boldsymbol{E} \int_0^t H_f(t-s)\, A(ds) \\ &= a \int_0^t H_f(t-s)\, ds, \end{aligned}$$

where $T_1, T_2, \ldots$ are the arrival epochs of the homogeneous Poisson arrival measure A with intensity a. Here, the last equality follows from the fact that the mean measure of the Poisson random measure A is given by $a\, ds$. Letting t tend to infinity completes the proof. □

Next we determine the mean measure m_Q for the special case that $e_1 = 0$, i.e. service times vanish. This turns out to be a crucial step in determining

m_Q for general $e_1 \geq 0$. In connection with this, we use a well-known result from the theory of diffusion processes which gives the expectation

$$e_f^x = \boldsymbol{E} \int_0^{V_x} f(B_t + x)\, dt$$

for any $x \in (0,1)$ and $f \in C_+[0,1]$, where V_x denotes the exit time from $[0,1]$ of the stochastic process $(B_t + x, t \geq 0)$.

Lemma 3 Let (B_t) be a Brownian motion with parameter vector $(\mu, \sigma_2) \in I\!R^2 \setminus \{(0,0)\}$. Then, for every $x \in (0,1)$ and $f \in C_+[0,1]$, it holds

$$(7.4) \qquad e_f^x = \begin{cases} \frac{1}{\mu}\{\frac{e^{\theta x}-1}{e^{\theta}-1}\int_x^1 (1-e^{\theta(1-y)})f(y)dy \\ +\frac{e^{\theta}-e^{\theta x}}{e^{\theta}-1}\int_0^x (e^{-\theta y}-1)f(y)dy\} & \text{if } \mu \neq 0, \sigma^2 > 0 \\ \frac{1}{\mu}\int_x^1 f(y)dy & \text{if } \mu > 0, \sigma^2 = 0 \\ \frac{1}{|\mu|}\int_0^x f(y)dy & \text{if } \mu < 0, \sigma^2 = 0 \\ \frac{2}{\sigma^2}\{x\int_x^1 (1-y)f(y)dy \\ +(1-x)\int_0^x yf(y)dy\} & \text{if } \mu = 0, \sigma^2 > 0 \end{cases}$$

where $\theta = -2\mu/\sigma^2$.

For a *proof* of this lemma and for further details we refer to Sections 15.3 and 15.4 of Karlin and Taylor [24].

For convenience, we will use the notation $q_f = \boldsymbol{E}\, Qf$ when $e_1 = 0$.

Lemma 4 Assume that $e_1 = 0$. Then,

$$(7.5) \qquad q_f = \begin{cases} a\alpha \int_0^1 (\frac{1-e^{-2\alpha^{-1}y/\sigma^2}}{1-e^{-2\alpha^{-1}/\sigma^2}} - y)f(y)dy & \text{if } \alpha^{-1} > 0, \sigma^2 > 0 \\ a\alpha \int_0^1 (1-y)f(y)dy & \text{if } \alpha^{-1} > 0, \sigma^2 = 0 \\ a\sigma^{-2} \int_0^1 y(1-y)f(y)dy & \text{if } \alpha^{-1} = 0, \sigma^2 > 0. \end{cases}$$

Proof Observe that, for $e_1 = 0$, the counting measure L_t^x has the following simple form

$$(7.6) \qquad L_t^x([c,d]) = \mathbf{1}_{\{V_x > t,\, c \leq B_t + x \leq d\}}$$

where $[c,d]$ is an arbitrary interval in $[0,1]$. Thus, from (7.3) and (7.6) we get

$$\begin{aligned} \boldsymbol{E}Qf &= a\int_0^\infty \int_0^1 \boldsymbol{E}L_t^x f\, dx\, dt \\ &= a\int_0^\infty \int_0^1 \boldsymbol{E}\mathbf{1}_{\{V_x>t\}} f(B_t+x)\, dxdt \end{aligned}$$

i.e.

$$\boldsymbol{E}\,Qf = a\int_0^1 \boldsymbol{E}\int_0^{V_x} f(B_t+x)\,dt\,dx, \tag{7.7}$$

for every $f\in C_+[0,1]$. Consequently, substituting (7.4) into the last integral and integrating with respect to x gives (7.5). □

Note that (7.5) implies that $\boldsymbol{E}\,Q$ (for the case $e_1=0$) has a density v with respect to the Lebesgue measure on [0,1]. This density has an intuitive interpretation, yielding another way to prove (7.7) without using the general representation formula (7.3) in Lemma 2. For this, one should consider a time-reversed model, looking backward from the actual position of the server. Notice that for every $t\ge 0$ the process $(B_u^*; 0\le u\le t)$, given by $B_u^* = B_{t-u}-B_t$, is again a Brownian motion on $[0,t]$, starting at 0, but now with drift α^{-1} and (the variance parameter is again σ^2). Let V_x^* denote the exit time from $[-x,1-x]$ of the Brownian motion (B_t^*). We now calculate an upper and lower bound for $\boldsymbol{E}\,Q_t([x-h,x+h])$, for all $x\in(0,1)$ and for large t and small $h>0$. Let $V_{x,h}^*$ and $V_{x,-h}^*$ denote the exit time from $[-x-h,1-x+h]$ and $[-x+h,1-x-h]$, respectively, of the Brownian motion (B_t^*). Clearly, in the time interval $(\{t-V_{x,-h}^*\}^+,t)$ the server did not visit that part of the circle which, seen from the actual position of the server at time t, is described by the interval $[x-h,x+h]$ (here $\{u\}^+=\max\{0,u\}$). Thus,

$$2ah\,\boldsymbol{E}\,(V_{x,-h}^*\wedge t)\le \boldsymbol{E}\,Q_t([x-h,x+h]),$$

where $s\wedge t=\min\{s,t\}$. Analogously,

$$\boldsymbol{E}\,Q_t([x-h,x+h])\le 2ah\,\boldsymbol{E}\,(V_{x,h}^*\wedge t).$$

Moreover, both $\boldsymbol{E}\,(V_{x,-h}^*\wedge t)$ and $\boldsymbol{E}\,(V_{x,h}^*\wedge t)$ converge to $\boldsymbol{E}\,(V_x^*\wedge t)$ uniformly in x when h goes to zero. Consequently,

$$\boldsymbol{E}\,Q_t(dx) = a\,\boldsymbol{E}\,(V_x^*\wedge t)\,dx.$$

Letting $t\to\infty$ in the last expression gives

$$\boldsymbol{E}\,Q(dx) = a\,\boldsymbol{E}\,V_x^*\,dx,$$

so that $v(x) = a\,\boldsymbol{E}\,V_x^*$, and (7.7) follows by the standard theory of Brownian motion.

Theorem 4 In the general case, i.e. if $e_1 \geq 0$, we have

$$\boldsymbol{E}\,Qf = \frac{1}{1 - ae_1}\,q_f, \tag{7.8}$$

for every $f \in C_+[0,1]$, where q_f is given by (7.5).

Proof We use the general representation formula (7.3), together with (7.7). In connection with this, we first analyze the expectation $G_f(t,x) = \boldsymbol{E}\,L_t^x f$ appearing in (7.2). Clearly, we have

$$G_f(t,x) = \boldsymbol{E}\,L_t^x f \mathbf{1}_{\{V_x > t\}} + \boldsymbol{E}\,L_t^x f \mathbf{1}_{\{V_x \leq t\}}. \tag{7.9}$$

Moreover, we get

$$\boldsymbol{E}\,L_t^x f \mathbf{1}_{\{V_x > t\}} = \boldsymbol{E}\,f(B_t + x)\mathbf{1}_{\{V_x > t\}}$$

for the first summand in (7.9), and, from (7.1),

$$\boldsymbol{E}\,L_t^x f \mathbf{1}_{\{V_x \leq t\}} = \boldsymbol{E}\,\mathbf{1}_{\{V_x \leq t\}} ae_1 H_f(t - V_x)$$

for the second summand. Because of (7.2) and (7.3), this gives

$$\begin{aligned}
\boldsymbol{E}Qf &= a\int_0^\infty H_f(t)\,dt \\
&= a\int_0^\infty \int_0^1 \boldsymbol{E} f(B_t + x)\mathbf{1}_{\{V_x > t\}}\,dx\,dt \\
&+ a^2 e_1 \int_0^\infty \int_0^1 \boldsymbol{E}\mathbf{1}_{\{V_x \leq t\}} H_f(t - V_x)\,dx\,dt \\
&= a\int_0^1 \boldsymbol{E} \int_0^{V_x} f(B_t + x)\,dt\,dx + a^2 e_1 \int_0^1 \boldsymbol{E} \int_{V_x}^\infty H_f(t - V_x)\,dt\,dx.
\end{aligned}$$

where

$$\boldsymbol{E} \int_{V_x}^\infty H_f(t - V_x)dt = \boldsymbol{E} \int_0^\infty H_f(t)\,dt.$$

Thus, from (7.3) and (7.7) we get

$$\boldsymbol{E}\,Qf = q_f + ae_1\,\boldsymbol{E}\,Qf.$$

This completes the proof. □

8 The Expected Number of Customers at an Arbitrary Point in Time

We now use the equilibrium equation (6.8) and Theorem 4 to derive a formula for the (unconditional) expectation $\boldsymbol{E}\, W([0,1])$ of the number of customers on the circle at 'an arbitrary point in time'. Note that, in general, this expectation may be infinite, even when the system is stable, i.e. when condition (3.1) is fulfilled. In this section we will, for any (random) measure M on $\mathcal{B}[0,1]$, abbreviate $M([0,1])$ to $|M|$, and, consequently, $\boldsymbol{E}\, M([0,1])$ to $\boldsymbol{E}\,|M|$.

Theorem 5 If (3.1) holds and if the second moment $e_2 = \int t^2 dF(t)$ of service time distribution is finite, then $\boldsymbol{E}\,|W| < \infty$.

Remark Note that one possibility to prove Theorem 5 is given by the elegant concepts of 'ancestral line' and 'offspring' considered in Fuhrmann and Cooper [21]. Then, the finiteness of the expectation $\boldsymbol{E}\,|W|$ follows from the decomposition formula (4) in Fuhrmann and Cooper [21], because Theorem 4 of the present paper yields that $\boldsymbol{E}\,|Q| < \infty$ and because the stationary mean queue length in the 'usual' M/G/1 queue is finite provided that $e_2 < \infty$. We remark however that the decomposition formula mentioned has been derived in Fuhrmann and Cooper [21] in a somewhat informal way because the authors of that paper simply assume the existence of a steady state giving no arguments which would show under what conditions this assumption is satisfied. Moreover, the notions 'random departing customer' and 'random point in time' considered in Fuhrmann and Cooper [21] remain vague without using a general point-process setting which includes the notion of Palm distribution, see e.g. Franken et al. [18], König and Schmidt [27]. In the present paper we decided to give a separate proof of Theorem 5 (and of Theorem 6) because our approach seems to work also for certain non-Poisson arrival processes, e.g. in the case when the arrival epochs of customers form a Markov modulated Poisson process. Furthermore, observe that the expected total workload at an arbitrary point in time can be derived fairly easily, once we have a formula for $\boldsymbol{E}\,|Q|$. See Kroese and Schmidt [31], where this has been discussed in detail.

Proof of Theorem 5 Consider again the embedded Markov chain (R_n) introduced in Section 3, where R_n denotes the configuration of customers on the circle at (= immediately after) the n th service completion. From the results obtained in Section 4 it follows that also the Markov chain (R_n) is ergodic in the sense that the corresponding n step transition probabilities converge in variation to the limit distribution μ given by (4.6) (that is, to the same limit distribution as (X_n)).

Next, let us denote by $([U_n, W_{U_n}])$ the time-stationary point process of departure epochs marked by the customer configurations at departure epochs, where the distribution of $([U_n, W_{U_n}])$ is given by the inversion formula for Palm distributions of marked point processes (see e.g. formula (6.3.2) in Berbee [8] or, in a more general context, Theorem 1.2.9 in Franken et al. [18]). Note that, in Section 3, we used the same notation U_n for the non-stationary case. But, it seems that there will be no confusion.

The limit distribution μ given by (4.6) of the embedded Markov chain (R_n) can be seen as the Palm mark distribution of the stationary marked point process $([U_n, W_{U_n}])$. Thus, from the PASTA property of stationary queueing systems with Poisson arrival process (see e.g. Theorems 4.1.1 and 4.3.1 in Franken et al. [18]) it follows that the distribution of W is equal to μ. Consequently, we have

$$\boldsymbol{E}\,|W| = \int |\varphi|\,\mu(d\varphi). \tag{8.1}$$

Thus, it suffices to show that the right-hand side of (8.1) is finite. For proving this we can use a general criterion for the existence of moments of stationary Markov chains with general state space (see Theorem 1 of Tweedie [46]). Namely, it suffices that for some set $A \in \mathcal{E}$ with $0 < \mu(A)$ and $\int_A |\varphi|\,\mu(d\varphi) < \infty$, some $\epsilon \in (0,1)$ and some measurable function $g : E \to I\!R_+$ with $g(\varphi) \geq \epsilon\,|\varphi|$, $\varphi \in A^c$ the following holds:

$$\int_{A^c} g(y)\,P(\varphi, dy) \leq g(\varphi) - \epsilon|\varphi| \quad \text{for every } \varphi \in A^c \tag{8.2}$$

and

$$\sup_{\varphi \in A} \int_{A^c} g(y) P(\varphi, dy) < \infty, \tag{8.3}$$

where $P(\varphi, dy)$ denotes the transition probabilities of the Markov chain (X_n) with $X_n = R_{\zeta n}$ considered in Section 4. Note that μ is the stationary initial distribution both of (R_n) and (X_n), where ζ is given by $(4.12')$.

For a certain natural number $\zeta_0 \geq \zeta$, which will be specified later, we put

$$A = \bigcup_{j=0}^{\zeta_0 - 1} A_j \quad \text{and} \quad g(\varphi) = j^2 \text{ for } \varphi \in A_j, \tag{8.4}$$

where $A_j = \{\varphi \in E : |\varphi| = j\}$. Then, similarly as in Section 4, for every $j \geq \zeta_0$ and for every $\varphi \in A_j$, we have

$$\int_E g(y) P(\varphi, dy) = \sum_{k=0}^{\infty} \int_{A_{j-\zeta+k}} (j - \zeta + k)^2 P(\varphi, dy)$$

$$\begin{aligned} &= (j-\zeta)^2 + 2(j-\zeta)\sum_{k=0}^{\infty} kP(\varphi, A_{j-\zeta+k}) + \sum_{k=0}^{\infty} k^2 P(\varphi, A_{j-\zeta+k}) \\ &\le j^2 - 2j(\zeta - ae_1\zeta - a\gamma) + c, \end{aligned}$$

where

$$c = \zeta^2 - (2\zeta - 1)a(\zeta e_1 + \alpha) + a^2(\zeta e_2 + \zeta^2 e_1^2 + 2\alpha\zeta e_1 + \alpha^2)$$

is a finite constant. Consequently, for $\epsilon = \zeta - ae_1\zeta - a\gamma > 0$ we have

$$\int_E g(y)P(\varphi, dy) \le j^2 - \epsilon j \tag{8.5}$$

for every $j \ge \zeta_0$ and $\varphi \in A_j$, where $\zeta_0 = \min\{j \in I\!N : j \ge \zeta,\ j > c\,\epsilon^{-1}\}$. Thus, the conditions (8.2) and (8.3) are satisfied. This finishes the proof of Theorem 5. □

Corollary 4 The expectations $\boldsymbol{E}\,|Q^0|$ and $\boldsymbol{E}\,|Q^1|$ are finite provided that the conditions of Theorem 5 are fulfilled.

Remark Note that the notion PASTA (Poisson Arrivals See Time Averages) was coined in Wolff [48] whereas the priciple itself, applied in (8.1), to prove that certain embedded and non-embedded stationary queueing characteristics coincide, without calculating them explicitely, was introduced considerably earlier (see Schassberger [39], Wolff [47]). A general point-process approach to this question is given in Franken et al. [18], see also Chapters 9 and 10 of König and Schmidt [27].

Theorem 6 Under the conditions of Theorem 5, the expected number of customers on the ring in the 'stationary situation' is

$$\boldsymbol{E}\,|W| = \frac{1}{1-ae_1} q_1 + ae_1 + \frac{a^2 e_2}{2(1-ae_1)}, \tag{8.6}$$

where q_1 is given by (7.5) with $f(x) \equiv 1$, i.e.

$$q_1 = \begin{cases} a\alpha\left(\dfrac{1}{1-\mathrm{e}^{-2\alpha^{-1}/\sigma^2}} - \dfrac{\alpha\sigma^2+1}{2}\right) & \text{if } \alpha^{-1} > 0, \sigma^2 > 0 \\ \dfrac{a\alpha}{2} & \text{if } \alpha^{-1} > 0, \sigma^2 = 0 \\ \dfrac{a}{6\sigma^2} & \text{if } \alpha^{-1} = 0, \sigma^2 > 0. \end{cases}$$

Remarks 1^o Note that from the individual ergodic theorem (see e.g. Theorem 1.3.12 of Franken et al. [18]) it follows that

$$\lim_{s\to\infty} \frac{1}{s}\int_0^s |W_t|dt = \boldsymbol{E}\,|W| \tag{8.7}$$

with probability one. Thus, $\boldsymbol{E}\,|W|$ is an important performance characteristic.

2^o It is well-known that, for the stationary mean queue length $\boldsymbol{E}\,W_{ssq}$ in the 'usual' single-server queue M/G/1 with arrival intensity a and first two moments e_1, e_2 of service times, it holds

$$\boldsymbol{E}\,W_{ssq} = ae_1 + \frac{a^2 e_2}{2(1-ae_1)}.$$

Thus, because of (7.8), the formula (4,23) can be written in the form

$$\boldsymbol{E}\,|W| = \boldsymbol{E}\,|Q| + \boldsymbol{E}\,W_{ssq} \tag{8.8}$$

which is in accordance with the decomposition property derived in Fuhrmann and Cooper [21] for the M/G/1 queue with generalized server vacations. In particular, it follows from (8.8) that the difference between the mean queue lengths $\boldsymbol{E}\,|W|$ and $\boldsymbol{E}\,|Q|$, at an arbitrary point in time and given that the server is walking, respectively, does *not* depend on the parameters $(\alpha^{-1}, \sigma^2) \in I\!R_+^2 \setminus \{(0,0)\}$.

3^o From (7.8) and (8.6) we see that $\boldsymbol{E}\,|Q|$ and $\boldsymbol{E}\,|W|$, depend on the parameter vector $(\alpha^{-1}, \sigma^2) \in I\!R_+^2 \setminus \{(0,0)\}$ in the same way as the expectation $\boldsymbol{E}\int_0^1 V_x dx$ of the time that a (α^{-1}, σ^2)-Brownian motion, starting from a uniform distribution on (0,1), hits 0 or 1. In particular, for fixed $\sigma^2 \geq 0$, the expectations $\boldsymbol{E}\,|Q|$ and $\boldsymbol{E}\,|W|$ are, as a function of α^{-1}, increasing on a certain interval $(0,\delta)$ and decreasing on (δ,∞), where $\delta = 0$ if $\sigma^2 = 0$, and $\delta > 0$ if $\sigma^2 > 0$. Moreover, for fixed $\alpha^{-1} \geq 0$, $\boldsymbol{E}\,|Q|$ and $\boldsymbol{E}\,|W|$ behave analogously as a function of σ^2. If α^{-1} and σ^2, respectively, tend to infinity, then $\boldsymbol{E}\,|W|$ converges to the stationary mean queue length in the 'usual' M/G/1-queue.

Proof of Theorem 6 For $p \in (0,1)$ we consider the function $f \in C_+^2[0,1]$ with $f(x) \equiv p$. Then, (6.8) takes the form

$$\begin{aligned} 0 &= -\beta(1-ae_1)\boldsymbol{E}\mathrm{e}^{-p|Q|} \\ &+ p_0\boldsymbol{E}\mathrm{e}^{-p|Q^0|}((\mathrm{e}^p - 1)\frac{\beta L_F(\beta)}{1-L_F(\beta)} - \beta) \\ &+ p_1\boldsymbol{E}\mathrm{e}^{-p|Q^1|}((\mathrm{e}^p - 1)\frac{\beta L_F(\beta)}{1-L_F(\beta)} - \beta), \end{aligned}$$

where $\beta = a(1 - \mathrm{e}^{-p})$. This is equivalent to

$$
\begin{aligned}
0 &= (1 - ae_1)\boldsymbol{E}\mathrm{e}^{-p|Q|}(1 - L_F(\beta) \\
&+ p_0\boldsymbol{E}\mathrm{e}^{-p|Q^0|}[(1 - L_F(\beta)) - (\mathrm{e}^p - 1)L_F(\beta)] \qquad (8.9)\\
&+ p_1\boldsymbol{E}\mathrm{e}^{-p|Q^1|}[(1 - L_F(\beta)) - (\mathrm{e}^p - 1)L_F(\beta)]
\end{aligned}
$$

Because $\boldsymbol{E}\,|Q| < \infty$ (see Theorem 4) and because from Theorem 5 it follows that the expectations $\boldsymbol{E}\,|Q^0|$ and $\boldsymbol{E}\,|Q^1|$ are finite as well, we get from the quadratic terms of a Taylor series expansion of (8.9) with respect to p that

$$
\begin{aligned}
0 &= -(1 - ae_1)ae_1\boldsymbol{E}|Q| - ae_1(1 - ae_1) - \frac{a^2e_2}{2} \\
&+ (1 - ae_1)(p_0\boldsymbol{E}|Q^0| + p_1\boldsymbol{E}|Q^1|),
\end{aligned}
$$

taking into account that $p_0 + p_1 = ae_1$. Thus,

$$
p_0\,\boldsymbol{E}\,|Q^0| + p_1\,\boldsymbol{E}\,|Q^1| = ae_1\,\boldsymbol{E}\,|Q| + ae_1 + \frac{a^2e_2}{2(1 - ae_1)}. \qquad (8.10)
$$

Now, because

$$
\boldsymbol{E}\,|W| = (1 - ae_1)\,\boldsymbol{E}\,|Q| + p_0\,\boldsymbol{E}\,|Q^0| + p_1\,\boldsymbol{E}\,|Q^1|,
$$

(8.6) follows from (7.8) and (8.10). This completes the proof. □

Remark 1^o Note that the expression

$$
(ae_1)^{-1}(p_0\,\boldsymbol{E}\,|Q^0| + p_1\,\boldsymbol{E}\,|Q^1|)
$$

can be interpreted as the conditional steady-state expectation of the number of customers on the circle given that the server is busy serving a customer (and, consequently, *not* walking). Let us denote this conditional expectation by $\boldsymbol{E}\,|Q^{(s)}|$. Then, (8.10) says that the relationship between $\boldsymbol{E}\,|Q^{(s)}|$ and the corresponding conditional expectation $\boldsymbol{E}\,|Q|$ given that the server is idle (i.e. walking) has the form

$$
\boldsymbol{E}\,|Q^{(s)}| = \boldsymbol{E}\,|Q| + 1 + \frac{ae_2}{2e_1(1 - ae_1)}. \qquad (8.11)
$$

Thus, similar to the behavior of $\boldsymbol{E}\,|W| - \boldsymbol{E}\,|Q|$, the difference between $\boldsymbol{E}\,|Q^{(s)}|$ and $\boldsymbol{E}\,|Q|$ does *not* depend on the parameters $(\alpha^{-1}, \sigma^2) \in I\!R^2_+ \setminus \{(0,0)\}$.

2^o In Kroese and Schmidt [29], instead of the approach stated in Section 7 of the present paper, the equilibrium equation $(6.8')$ has been used to determine the mean measure m_Q of Q, where the main idea is to consider an appropriately chosen test function f. One might think that, by considering an appropriate f in $(6.8')$, it should be also possible to express the mean measure of Q^o by that of Q. However, this seems to be a complicated problem. Note that, in case that this problem would be solved, one could determine the mean measure of W, too (and not only the expectation $\boldsymbol{E}\,|W|$).

Acknowledgement

The authors thank R. Kiesel for his helpful comments.

References

[1] Altman, E., Foss, S. (1993): Polling on a graph with general arrival and service time distribution. Preprint. INRIA Centre Sophia Antipolis.

[2] Altman, E., Konstantopoulos, P., Liu, Z. (1992): Stability, monotonicity and invariant quantities in general polling systems, QUESTA 11, 35-57.

[3] Altman, E., Yechiali, U. (1993): Cyclic Bernoulli polling. Zeitschrift f. Operations Res. 38, 55-76.

[4] Altman, E., Levy, H. (1993): Queueing in space. Preprint. INRIA Centre Sophia Antipolis/Tel-Aviv University.

[5] Asmussen, S. (1987): Applied Probability and Queues. J. Wiley & Sons, Chichester.

[6] Athreya, K.B., Ney, P. (1978): A new approach to the limit theory of recurrent Markov chains. Trans. Amer. Math. Soc. 245, 493-501.

[7] Bardhan, I., Sigman, K. (1993): Rate conservation law for stationary semimartingales. Probability in Engineering and Inform. Sci. 7, 1-17.

[8] Berbee, H.C.P. (1979): Random Walks with Stationary Increments and Renewal Theory. Mathematical Centre Tracts 112. Mathematisch Centrum, Amsterdam.

[9] Bertsimas, D.J., van Ryzin, G. (1991): A stochastic and dynamic vehicle routing problem in the Euclidean plane. Operations Res. 39, 601-615.

[10] Borovkov, A., Schassberger, R. (1993): Ergodicity of a polling network. Stoch. Proc. Appl. 48

[11] Boxma, O.J., Groenendijk, W.P. (1987): Pseudo-conservation laws in cyclic service systems. J. Appl. Prob. 24, 949-964.

[12] Boxma, O.J., Westrate, J.A. (1989): Waiting times in polling systems with Markovian server routing. In: Siege, G., Lie, J.S. (eds.), Messung, Modellierung und Bewertung von Rechnersystemen und Netzen. Springer, Berlin, 89-104.

[13] Coffman, Jr. E.G., Gilbert, E.N. (1987): Polling and greedy servers on a line. QUESTA 2, 115-145.

[14] Coffman, Jr., E.G., Stolyar, A. (1993): Continuous polling on graphs. Probab. Engin. Inf. Sci. 7, 209-226.

[15] Cooper, R.B., Murray, G. (1969): Queues served in cyclic order. The Bell System Techn. J. 48, 675-689.

[16] Daley, D.J., Vere-Jones, D. (1988): An Introduction to the Theory of Point Processes. Springer, New York.

[17] Eisenberg, M. (1972): Queues with periodic service and changeover times. Operations Res. 20, 440-451.

[18] Franken, P., König, D., Arndt, U., Schmidt, V. (1982): Queues and Point Processes. J. Wiley & Sons, Chichester.

[19] Fricker, C., Jaibi, M.R. (1992): Monotonicity and stability of periodic polling models. Preprint. Tilburg University.

[20] Fuhrmann, S.W. (1985): Symmetric queues served in cyclic order. Operations Res. Letters 4, 139-144.

[21] Fuhrmann, S.W., Cooper, R.B. (1985a): Stochastic decomposition in the M/G/1 queue with generalized vacations. Operations Res. 33, 1117-1129.

[22] Fuhrmann, S.W., Cooper, R.B. (1985b): Application of decomposition principle in M/G/1 vacation model to two continuum cyclic queueing models - especially token-ring LANs. AT & T Techn. J. 64, 1091-1098.

[23] Georgiadis, L., Szpankowski, W. (1992): Stability of token passing rings. QUESTA 11, 7-33.

[24] Karlin, S., Taylor, H.M. (1981): A Second Course in Stochastic Processes. Academic Press, New York.

[25] Keilson, J., Servi, L.D. (1986): Oscillating random walk models for GI/G/1 vacation systems with Bernoulli schedules. J. Appl. Prob. 23, 790-802.

[26] Kleinrock, L., Levy, H. (1988): The analysis of random polling systems. Operations Res. 36, 716-732.

[27] König, D., Schmidt, V. (1992): Random Point Processes. Teubner, Stuttgart (Germ.)

[28] Koheim, A.G., Meister, B. (1974): Waiting lines and times in a system with polling. J. ACM 21, 470-490.

[29] Kroese, D.P., Schmidt, V. (1992): A continuous polling system with general service times. Ann. Appl. Prob. 2, 906-927.

[30] Kroese, D.P., Schmidt, V. (1993): Queueing systems on a circle. Zeitschrift f. Operations Res. 37, 303-331.

[31] Kroese, D.P., Schmidt, V. (1993): Light-traffic analysis for queues with spatially distributed arrivals. Preprint. University of Twente / University of Ulm.

[32] Kühn, P.J. (1979): Multiqueue systems with nonexhaustive cyclic service. The Bell System Techn. J. 58, 671-698.

[33] Lindvall, T. (1992): Lectures on the Coupling Method. J. Wiley & Sons, New York.

[34] Matthes, K., Kerstan, J., Mecke, J. (1978): Infinitely Divisible Point Processes. J. Wiley & Sons, Chichester.

[35] Mazumdar, R., Badrinath, V., Guillemin, F., Rosenberg, C. (1993): On pathwise rate conservation for a class of semi-martingales. Stoch. Proc. Appl. 47, 119-130.

[36] Miyazawa, M. (1993): Rate conservation laws: A survey. QUESTA 14

[37] Nummelin, E. (1978): A splitting technique for Harris recurrent Markov chains. Zeitschrift f. Wahrscheinlichkeitstheorie verw. Geb. 43, 309-318.

[38] Resing, J.A.C. (1993): Polling systems and multitype branching processes. QUESTA 13, 409-426.

[39] Schassberger, R. (1973): Queues. Springer, Wien (Germ.)

[40] Sigman, K., Wolff, R.W. (1993): A review of regenerative processes. SIAM Review 35, 269-2

[41] Spieksma, F.M., Tweedie, R.L. (1992): Strengthening ergodicity to geometric ergodicity for Markov chains. Preprint. University of Leiden/ Colorado State University.

[42] Takagi, H. (1986): Analysis of Polling Systems. MIT Press, Cambridge (MA).

[43] Tedijanto (1990): Exact results for the cyclic service queue with a Bernoulli schedule. Performance Evaluation 11, 107-115.

[44] Thorisson, H. (1992): Construction of a stationary regenerative process. Stoch. Proc. Appl. 42, 237-253.

[45] Tweedie, R.L. (1976): Criteria for classifying general Markov chains. Adv. Appl. Prob. 8, 737-771.

[46] Tweedie, R.L. (1983): The existence of moments for stationary Markov chains. J. Appl. Prob. 20, 191-196.

[47] Wolff, R.W. (1970): Work conserving priorities. J. Appl. Prob. 7, 327-337.

[48] Wolff, R.W. (1982): Poisson arrivals see time averages. Operations Res. 30, 223-231.

[49] Yechiali, U. (1991): Optimal dynamic control of polling systems. In: Cohen, J.W., Pack, C.D. (Eds.) Queueing Performance and Control in ATM, North Holland, 205-217.