Vacation and Polling Models with Retrials

Onno Boxma and Jacques Resing

EURANDOM and Department of Mathematics and Computer Science
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
{o.j.boxma,j.a.c.resing}@tue.nl

Abstract. We study a vacation-type queueing model, and a single-server multi-queue polling model, with the special feature of retrials. Just before the server arrives at a station there is some deterministic glue period. Customers (both new arrivals and retrials) arriving at the station during this glue period will be served during the visit of the server. Customers arriving in any other period leave immediately and will retry after an exponentially distributed time. Our main focus is on queue length analysis, both at embedded time points (beginnings of glue periods, visit periods and switch- or vacation periods) and at arbitrary time points.

Keywords: Vacation queue, polling model, retrials.

1 Introduction

Queueing systems with retrials are characterized by the fact that arriving customers, who find the server busy, do not wait in an ordinary queue. Instead of that they go into an orbit, retrying to obtain service after a random amount of time. These systems have received considerable attention in the literature, see e.g. the book by Falin and Templeton [7], and the more recent book by Artalejo and Gomez-Corral [3].

Polling systems are queueing models in which a single server, alternatingly, visits a number of queues in some prescribed order. Polling systems, too, have been extensively studied in the literature. For example, various different service disciplines (rules which describe the server's behaviour while visiting a queue) and both models with and without switchover times have been considered. We refer to Takagi [20,21] and Vishnevskii and Semenova [22] for some literature reviews and to Boon, van der Mei and Winands [5], Levy and Sidi [12] and Takagi [18] for overviews of the applicability of polling systems.

In this paper, motivated by questions regarding the performance modelling of optical networks, we consider vacation and polling systems with retrials. Despite the enormous amount of literature on both types of models, there are hardly any papers having both the features of retrials of customers and of a single server polling a number of queues. In fact, the authors are only aware of a sequence of papers by Langaris [9,10,11] on this topic. In all these papers the author

A. Horváth and K. Wolter (Eds.): EPEW 2014, LNCS 8721, pp. 45–58, 2014.

determines the mean number of retrial customers in the different stations. In [9] the author studies a model in which the server, upon polling a station, stays there for an exponential period of time and if a customer asks for service before this time expires, the customer is served and a new exponential stay period at the station begins. In [10] the author studies a model with two types of customers: primary customers and secondary customers. Primary customers are all customers present in the station at the instant the server polls the station. Secondary customers are customers who arrive during the sojourn time of the server in the station. The server, upon polling a station, first serves all the primary customers present and after that stays an exponential period of time to wait for and serve secondary customers. Finally, in [11] the author considers a model with Markovian routing and stations that could be either of the type of [9] or of the type of [10].

In this paper we consider a polling station with retrials and so-called glue periods. Just before the server arrives at a station there is some deterministic glue period. Customers (both new arrivals and retrials) arriving at the station during this glue period "stick" and will be served during the visit of the server. Customers arriving in any other period leave immediately and will retry after an exponentially distributed time.

The study of queueing systems with retrials and glue periods is motivated by questions regarding the performance modelling and analysis of optical networks. Performance analysis of optical networks is a challenging topic (see e.g. Maier [13] and Rogiest [17]). In a telecommunication network, packets must be routed from source to destination, passing through a series of links and nodes. In copper-based transmission links, packets from different sources are time-multiplexed. This is often modeled by a single server polling system. Optical fibre offers some big advantages for communication w.r.t. copper cables: huge bandwidth, ultra-low losses, and an extra dimension – the wavelength of light. However, in an optical routing node, opposite to electronics, it is difficult to store photons, and hence buffering in optical routers can only be very limited. Buffering in these networks is typically realized by sending optical packets into fibre delay loops, i.e., letting them circulate in a fibre loop and extracting them after a certain number of circulations. This feature can be modelled by retrial queues. Recent experiments with 'slow light', where light is slowed down by significantly increasing the refractive index in waveguides, have up to now shown very modest buffering times [8]. It should be noted that with the very high speeds achievable in fibre, packet durations are very short, so that small buffering times may already allow sufficient storage of small packets. We represent the effect of slowing down light by introducing a glue period at a queue just before the server arrives.

The paper is organized as follows. In Section 2 we consider the case of a single queue with vacations and retrials; arrivals and retrials only "stick" during a glue period. We study this case separately because (i) it is of interest in its own right, (ii) it allows us to explain the analytic approach as well as the probabilistic meaning of the main components in considerable detail, (iii) it makes the analysis of the multi-queue case more accessible, and (iv) results for the one-queue case may serve as a first-order approximation for the behaviour of a particular queue in the N-queue case, switchover periods now also representing glue and visit periods at other queues. In Section 3 the two-queue case is analyzed. We do not have the space to treat the N-queue case in this paper, but the analysis in Sections 2 and 3 lays the groundwork for analyzing the N-queue case. Section 4 presents some conclusions and suggestions for future research.

2 Queue Length Analysis for the Single-Queue Case

2.1 Model Description

In this section we consider a single queue Q in isolation. Customers arrive at Q according to a Poisson process with rate λ. The service times of successive customers are independent, identically distributed (i.i.d.) random variables (r.v.), with distribution $B(\cdot)$ and Laplace-Stieltjes transform (LST) $\tilde{B}(\cdot)$. A generic service time is denoted by B. After a visit period of the server at Q it takes a vacation. Successive vacation lengths are i.i.d. r.v., with S a generic vacation length, with distribution $S(\cdot)$ and LST $\tilde{S}(\cdot)$. We make all the usual independence assumptions about interarrival times, service times and vacation lengths at the queues. After the server's vacation, a *glue* period of deterministic (i.e., constant) length begins. Its significance stems from the following assumption. Customers who arrive at Q do not receive service immediately. When customers arrive at Q during a glue period G, they stick, joining the queue of Q. When they arrive in any other period, they immediately leave and retry after a retrial interval which is independent of everything else, and which is exponentially distributed with rate ν. The glue period is immediately followed by a visit period of the server at Q.

The service discipline at Q is gated: During the visit period at Q, the server serves all "glued" customers in that queue, i.e., all customers waiting at the end of the glue period (but none of those in orbit, and neither any new arrivals).

We are interested in the steady-state behaviour of this vacation model with retrials. We hence make the assumption that $\rho := \lambda \mathbb{E}B < 1$; it may be verified that this is indeed the condition for this steady-state behaviour to exist.

Some more notation:
G_n denotes the nth glue period of Q.
V_n denotes the nth visit period of Q (immediately following the nth glue period).
S_n denotes the nth vacation of the server (immediately following the nth visit period).
X_n denotes the number of customers in the system (hence in orbit) at the start of G_n.
Y_n denotes the number of customers in the system at the start of V_n. Notice that here we should distinguish between those who are queueing and those who are in orbit: We write $Y_n = Y_n^{(q)} + Y_n^{(o)}$, where q denotes queueing and o denotes in orbit.
Finally,
Z_n denotes the number of customers in the system (hence in orbit) at the start of S_n.

2.2 Queue Length Analysis at Embedded Time Points

In this subsection we study the steady-state distributions of the numbers of customers at the beginning of (i) glue periods, (ii) visit periods, and (iii) vacation periods. Denote by X a r.v. with as distribution the limiting distribution of X_n. Y and Z are similarly defined, and $Y = Y^{(q)} + Y^{(o)}$, the steady-state numbers of customers in queue and in orbit at the beginning of a visit period (which coincides with the end of a glue period). In the sequel we shall introduce several generating functions, throughout assuming that their parameter $|z| \leq 1$. For conciseness of notation, let $\beta(z) := \tilde{B}(\lambda(1-z))$ and $\sigma(z) := \tilde{S}(\lambda(1-z))$. Then it is easily seen that

$$\mathbb{E}[z^X] = \sigma(z)\mathbb{E}[z^Z], \tag{2.1}$$

since X equals Z plus the new arrivals during the vacation;

$$\mathbb{E}[z^Z] = \mathbb{E}[\beta(z)^{Y^{(q)}} z^{Y^{(o)}}], \tag{2.2}$$

since Z equals $Y^{(o)}$ plus the new arrivals during the $Y^{(q)}$ services; and

$$\mathbb{E}[z_q^{Y^{(q)}} z_o^{Y^{(o)}}] = \mathrm{e}^{-\lambda(1-z_q)G}\mathbb{E}[\{(1-\mathrm{e}^{-\nu G})z_q + \mathrm{e}^{-\nu G}z_o\}^X]. \tag{2.3}$$

The last equation follows since $Y^{(q)}$ is the sum of new arrivals during G and retrials who return during G; each of the X customers which were in orbit at the beginning of the glue period have a probability $1-\mathrm{e}^{-\nu G}$ of returning before the end of that glue period.

Combining Equations (2.1)-(2.3), and introducing

$$f(z) := (1-\mathrm{e}^{-\nu G})\beta(z) + \mathrm{e}^{-\nu G}z, \tag{2.4}$$

we obtain the following functional equation for $\mathbb{E}[z^X]$:

$$\mathbb{E}[z^X] = \sigma(z)\mathrm{e}^{-\lambda(1-\beta(z))G}\mathbb{E}[f(z)^X].$$

Introducing $K(z) := \sigma(z)\mathrm{e}^{-\lambda(1-\beta(z))G}$ and $X(z) := \mathbb{E}[z^X]$, we have:

$$X(z) = K(z)X(f(z)). \tag{2.5}$$

This is a functional equation that naturally occurs in the study of queueing models which have a branching-type structure; see, e.g., [6] and [16]. Typically, one may view customers who newly arrive into the system during a service as children of the served customer ("branching"), and customers who newly arrive into the system during a vacation or glue period as immigrants. Such a functional equation may be solved by iteration, giving rise to an infinite product – where the jth term in the product typically corresponds to customers who descend from an ancestor of j generations before. In this particular case we have after n iterations:

$$X(z) = \prod_{j=0}^{n} K(f^{(j)}(z))X(f^{(n+1)}(z)), \tag{2.6}$$

where $f^{(0)}(z) := z$ and $f^{(j)}(z) := f(f^{(j-1)}(z))$, $j = 1, 2, \ldots$. Below we show that this product converges for $n \to \infty$ iff $\rho < 1$, thus proving the following theorem:

Theorem 1. *If $\rho < 1$ then the generating function $X(z) = \mathbb{E}[z^X]$ is given by*

$$X(z) = \prod_{j=0}^{\infty} K(f^{(j)}(z)). \tag{2.7}$$

Proof. Equation (2.5) is an equation for a branching process with immigration, where the number of immigrants has generating function $K(z)$ and the number of children in the branching process has generating function $f(z)$. Clearly, $K'(1) = \lambda\mathbb{E}S + \lambda\rho G < \infty$ and $f'(1) = \mathrm{e}^{-\nu G} + \left(1 - \mathrm{e}^{-\nu G}\right)\rho < 1$, if $\rho < 1$. The result of the theorem now follows directly from the theory of branching processes with immigration (see e.g., Theorem 1 on page 263 in Athreya and Ney [4]).

Having obtained an expression for $\mathbb{E}[z^X]$ in Theorem 1, expressions for $\mathbb{E}[z^Z]$ and $\mathbb{E}[z_q^{Y^{(q)}} z_o^{Y^{(o)}}]$ immediately follow from (2.2) and (2.3). Moments of X may be obtained from Theorem 1, but it is also straightforward to obtain $\mathbb{E}X$ from Equations (2.1)-(2.3):

$$\mathbb{E}X = \lambda\mathbb{E}S + \mathbb{E}Z, \tag{2.8}$$

$$\mathbb{E}Z = \rho\mathbb{E}Y^{(q)} + \mathbb{E}Y^{(o)}, \tag{2.9}$$

$$\mathbb{E}Y^{(q)} = \lambda G + (1 - \mathrm{e}^{-\nu G})\mathbb{E}X, \tag{2.10}$$

$$\mathbb{E}Y^{(o)} = \mathrm{e}^{-\nu G}\mathbb{E}X, \tag{2.11}$$

yielding

$$\mathbb{E}X = \frac{\lambda\mathbb{E}S + \lambda\rho G}{(1-\rho)(1 - \mathrm{e}^{-\nu G})}. \tag{2.12}$$

Hence

$$\mathbb{E}Y^{(q)} = \lambda G + (1 - \mathrm{e}^{-\nu G})\frac{\lambda\mathbb{E}S + \lambda\rho G}{(1-\rho)(1 - \mathrm{e}^{-\nu G})} = \frac{\lambda\mathbb{E}S + \lambda G}{1-\rho}, \tag{2.13}$$

$$\mathbb{E}Y^{(o)} = \mathrm{e}^{-\nu G}\frac{\lambda\mathbb{E}S + \lambda\rho G}{(1-\rho)(1 - \mathrm{e}^{-\nu G})}, \tag{2.14}$$

$$\mathbb{E}Z = \frac{\lambda\rho G + \lambda\mathbb{E}S[\rho(1 - \mathrm{e}^{-\nu G}) + \mathrm{e}^{-\nu G}]}{(1-\rho)(1 - \mathrm{e}^{-\nu G})}. \tag{2.15}$$

Notice that the denominators of the above expressions equal $1 - f'(1)$. Also notice that it makes sense that the denominators contain both the factor $1-\rho$ and the probability $1 - \mathrm{e}^{-\nu G}$ that a retrial returns during a glue period.

In a similar way as the first moments of X, $Y^{(q)}$, $Y^{(o)}$ and Z have been obtained, we can also obtain their second moment. Here we only mention $\mathbb{E}X^2$:

$$\begin{aligned}\mathbb{E}X^2 &= \frac{K''(1)}{(1-\rho)(1-\mathrm{e}^{-\nu G})(1+\rho(1-\mathrm{e}^{-\nu G})+\mathrm{e}^{-\nu G})} \\ &+ \frac{K'(1)[1-(\rho(1-\mathrm{e}^{-\nu G})+\mathrm{e}^{-\nu G})^2 + 2K'(1)(\rho(1-\mathrm{e}^{-\nu G})+\mathrm{e}^{-\nu G}) + (1-\mathrm{e}^{-\nu G})\lambda^2\mathbb{E}B^2]}{(1-\rho)^2(1-\mathrm{e}^{-\nu G})^2(1+\rho(1-\mathrm{e}^{-\nu G})+\mathrm{e}^{-\nu G})},\end{aligned} \tag{2.16}$$

where $K'(1) = \lambda\mathbb{E}S + \lambda\rho G$ and $K''(1) = \lambda^2\mathbb{E}S^2 + 2\rho\lambda^2 G\mathbb{E}S + \lambda^3 G\mathbb{E}B^2 + (\lambda G\rho)^2$.

Remark 1. Special cases of the above analysis are, e.g.:
(i) Vacations of length zero. Simply take $\sigma(z) \equiv 1$ and $\mathbb{E}S = 0$ in the above formulas.
(ii) $\nu = \infty$. Retrials now always return during a glue period. We then have $f(z) = \beta(z)$, which leads to minor simplifications.

Remark 2. It seems difficult to handle the case of non-constant glue periods, as it seems to lead to a process with complicated dependencies. If G takes a few distinct values $G_1, \dots, G_N$ with different probabilities, then one might still be able to obtain a kind of multinomial generalization of the infinite product featuring in Theorem 1. One would then have several functions $f_i(z) := (1-\mathrm{e}^{-\nu G_i})\beta(z) + \mathrm{e}^{-\nu G_i} z$, and all possible combinations of iterations $f_i(f_h(f_k(\dots(z))))$ arising in functions $K_i(z) := \sigma(z)\mathrm{e}^{-\lambda(1-\beta(z))G_i}$, $i = 1, 2, \dots, N$. By way of approximation, one might stop the iterations after a certain number of terms, the number depending on the speed of convergence (hence on $1-\rho$ and on $1-\mathrm{e}^{-\nu G_i}$).

2.3 Queue Length Analysis at Arbitrary Time Points

Having found the generating functions of the number of customers at the beginning of (i) glue periods ($\mathbb{E}[z^X]$), (ii) visit periods ($\mathbb{E}[z_q^{Y^{(q)}} z_o^{Y^{(o)}}]$), and (iii) vacation periods ($\mathbb{E}[z^Z]$), we can also obtain the generating function of the number of customers at arbitrary time points.

Theorem 2. *If $\rho < 1$, we have the following results:*

a) The joint generating function, $R^{va}(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point in a vacation period is given by

$$R^{va}(z_q, z_o) = \mathbb{E}[z_o^Z] \cdot \frac{1 - \tilde{S}(\lambda(1-z_o))}{\lambda(1-z_o)\mathbb{E}S}. \tag{2.17}$$

b) The joint generating function, $R^{gl}(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point in a glue period is given by

$$R^{gl}(z_q, z_o) = \int_{t=0}^{G} \mathrm{e}^{-\lambda(1-z_q)t} \mathbb{E}[\{(1-\mathrm{e}^{-\nu t})z_q + \mathrm{e}^{-\nu t} z_o\}^X] \frac{dt}{G}. \tag{2.18}$$

c) The joint generating function, $R^{vi}(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point in a visit period is given by

$$R^{vi}(z_q, z_o) = \frac{z_q \left[\mathbb{E}[z_q^{Y^{(q)}} z_o^{Y^{(o)}}] - \mathbb{E}[\tilde{B}(\lambda(1-z_o))^{Y^{(q)}} z_o^{Y^{(o)}}]\right]}{\mathbb{E}[Y^{(q)}]\left(z_q - \tilde{B}(\lambda(1-z_o))\right)} \cdot \frac{1 - \tilde{B}(\lambda(1-z_o))}{\lambda(1-z_o)\mathbb{E}B}. \tag{2.19}$$

d) The joint generating function, $R(z_q, z_o)$, of the number of customers in the queue and in the orbit at an arbitrary time point is given by

$$R(z_q, z_o) = \rho R^{vi}(z_q, z_o) + (1-\rho)\tfrac{G}{G+\mathbb{E}S} R^{gl}(z_q, z_o) + (1-\rho)\tfrac{\mathbb{E}S}{G+\mathbb{E}S} R^{va}(z_q, z_o). \tag{2.20}$$

Proof.

a) Follows from the fact that during vacation periods the number of customers in the queue is 0 and the fact that the number of customers at an arbitrary time point in the orbit is the sum of two independent terms: The number of customers at the beginning of the vacation period and the number that arrived during the past part of the vacation period. The generating function of the latter is given by

$$\frac{1-\tilde{S}(\lambda(1-z_o))}{\lambda(1-z_o)\mathbb{E}S}.$$

b) Follows from the fact that if the past part of the glue period is equal to t, the generating function of the number of new arrivals in the queue during this period is equal to $\mathrm{e}^{-\lambda(1-z_q)t}$ and each customer present in the orbit at the beginning of the glue period is, independent of the others, still in orbit with probability $\mathrm{e}^{-\nu t}$ and has moved to the queue with probability $1-\mathrm{e}^{-\nu t}$.

c) During an arbitrary point in time in a visit period the number of customers in the system consists of two parts:
   - the number of customers in the system at the beginning of the service time of the customer currently in service, leading to the term

$$\frac{z_q\left(\mathbb{E}[z_q^{Y^{(q)}}z_o^{Y^{(o)}}]-\mathbb{E}[\tilde{B}(\lambda(1-z_o))^{Y^{(q)}}z_o^{Y^{(o)}}]\right)}{\mathbb{E}[Y^{(q)}]\left(z_q-\tilde{B}(\lambda(1-z_o))\right)};$$

   (see Takagi [19], formula (5.14) on page 206, for a similar formula in the ordinary $M/G/1$ vacation queue with gated service but without retrials).
   - the number of customers that arrived during the past part of the service of the customer currently in service, leading to the term

$$\frac{1-\tilde{B}(\lambda(1-z_o))}{\lambda(1-z_o)\mathbb{E}B}.$$

d) Follows from the fact that the fraction of time the server is visiting Q is equal to ρ, and if the server is not visiting Q, with probability $\mathbb{E}S/(G+\mathbb{E}S)$ the server is on vacation and with probability $G/(G+\mathbb{E}S)$ the system is in a glue phase.

From Theorem 2, we now can obtain the steady-state mean number of customers in the system at arbitrary time points in vacation periods ($\mathbb{E}[R_{va}]$), in glue periods ($\mathbb{E}[R_{gl}]$), in visit periods ($\mathbb{E}[R_{vi}]$) and in arbitrary periods ($\mathbb{E}[R]$). These are given by

$$\begin{aligned}
\mathbb{E}[R_{va}] &= \mathbb{E}[Z]+\lambda\tfrac{\mathbb{E}[S^2]}{2\mathbb{E}[S]},\\
\mathbb{E}[R_{gl}] &= \mathbb{E}[X]+\lambda\tfrac{G}{2},\\
\mathbb{E}[R_{vi}] &= 1+\lambda\tfrac{\mathbb{E}[B^2]}{2\mathbb{E}[B]}+\tfrac{\mathbb{E}[Y^{(q)}Y^{(o)}]}{\mathbb{E}[Y^{(q)}]}+\tfrac{(1+\rho)\mathbb{E}[Y^{(q)}(Y^{(q)}-1)]}{2\mathbb{E}[Y^{(q)}]},\\
\mathbb{E}[R] &= \rho\mathbb{E}[R_{vi}]+(1-\rho)\tfrac{G}{G+\mathbb{E}S}\mathbb{E}[R_{gl}]+(1-\rho)\tfrac{\mathbb{E}S}{G+\mathbb{E}S}\mathbb{E}[R_{va}].
\end{aligned}$$

Remark that the quantities $\mathbb{E}[Y^{(q)}Y^{(o)}]$ and $\mathbb{E}[Y^{(q)}(Y^{(q)}-1)]$ can be obtained using (2.3):

$$\mathbb{E}[Y^{(q)}Y^{(o)}] = \lambda G \mathrm{e}^{-\nu G}\mathbb{E}[X] + \left(1-\mathrm{e}^{-\nu G}\right)\mathrm{e}^{-\nu G}\mathbb{E}[X(X-1)],$$
$$\mathbb{E}[Y^{(q)}(Y^{(q)}-1)] = (\lambda G)^2 + \left(1-\mathrm{e}^{-\nu G}\right)^2\mathbb{E}[X(X-1)] + 2\lambda G\left(1-\mathrm{e}^{-\nu G}\right)\mathbb{E}[X].$$

Finally, the mean sojourn time of an arbitrary customer now immediately follows from Little's formula. The results of this section can, e.g., be used to determine the value of G which minimizes the mean sojourn time of an arbitrary customer.

3 Queue Length Analysis for The Two-Queue Case

3.1 Model Description

In this section we consider a one-server polling model with two queues, Q_1 and Q_2. Customers arrive at Q_i according to a Poisson process with rate λ_i; they are called type-i customers, $i = 1, 2$. The service times at Q_i are i.i.d. r.v., with B_i a generic r.v., with distribution $B_i(\cdot)$ and LST $\tilde{B}_i(\cdot)$, $i = 1, 2$. After a visit of the server at Q_i, it switches to the other queue. Successive switchover times from Q_i to the other queue are i.i.d. r.v., with S_i a generic r.v., with distribution $S_i(\cdot)$ and LST $\tilde{S}_i(\cdot)$, $i = 1, 2$. We make all the usual independence assumptions about interarrival times, service times and switchover times at the queues. After a switch of the server to Q_i, there first is a deterministic (i.e., constant) glue period G_i, $i = 1, 2$, before the visit of the server at Q_i begins. As in the one-queue case, the significance of the glue period stems from the following assumption. Customers who arrive at Q_i do not receive service immediately. When customers arrive at Q_i during a glue period G_i, they stick, joining the queue of Q_i. When they arrive in any other period, they immediately leave and retry after a retrial interval which is independent of everything else, and which is exponentially distributed with rate ν_i, $i = 1, 2$.

The service discipline at both queues is gated: During the visit period at Q_i, the server serves all "glued" customers in that queue, i.e., all type-i customers waiting at the end of the glue period – but none of those in orbit, and neither any new arrivals.

We are interested in the steady-state behaviour of this polling model with retrials. We hence assume that the stability condition $\sum_{i=1}^{2}\rho_i < 1$ holds, where $\rho_i := \lambda_i \mathbb{E}B_i$.

Some more notation:

G_{ni} denotes the nth glue period of Q_i.
V_{ni} denotes the nth visit period of Q_i.
S_{ni} denotes the nth switch period out of Q_i, $i = 1, 2$.
$(X_{n1}^{(i)}, X_{n2}^{(i)})$ denotes the vector of numbers of customers of type 1 and of type 2 in the system (hence in orbit) at the start of G_{ni}, $i = 1, 2$.
$(Y_{n1}^{(i)}, Y_{n2}^{(i)})$ denotes the vector of numbers of customers of type 1 and of type 2 in the system at the start of V_{ni}, $i = 1, 2$. We distinguish between those who are

queueing in Q_i and those who are in orbit for Q_i: We write $Y_{n1}^{(1)} = Y_{n1}^{(1q)} + Y_{n1}^{(1o)}$ and $Y_{n2}^{(2)} = Y_{n2}^{(2q)} + Y_{n2}^{(2o)}$, where q denotes queueing and o denotes in orbit. Finally,
$(Z_{n1}^{(i)}, Z_{n2}^{(i)})$ denotes the vector of numbers of customers of type 1 and of type 2 in the system (hence in orbit) at the start of S_{ni}, $i = 1, 2$.

3.2 Queue Length Analysis

In this section we study the steady-state joint distribution of the numbers of customers in the system at beginnings of glue periods. This will also immediately yield the steady-state joint distributions of the numbers of customers in the system at the beginnings of visit periods and of switch periods. We follow a similar generating function approach as in the one-queue case, throughout making the following assumption regarding the parameters of the generating functions: $|z_i| \leq 1$, $|z_{iq}| \leq 1$, $|z_{io}| \leq 1$. Observe that the generating function of the vector of numbers of arrivals at Q_1 and Q_2 during a type-i service time B_i is $\beta_i(z_1, z_2) := \tilde{B}_i(\lambda_1(1 - z_1) + \lambda_2(1 - z_2))$. Similarly, the generating function of the vector of numbers of arrivals at Q_1 and Q_2 during a type-i switchover time S_i is $\sigma_i(z_1, z_2) := \tilde{S}_i(\lambda_1(1 - z_1) + \lambda_2(1 - z_2))$. We can successively express (in terms of generating functions) $(X_{n1}^{(2)}, X_{n2}^{(2)})$ into $(Z_{n1}^{(1)}, Z_{n2}^{(1)})$, $(Z_{n1}^{(1)}, Z_{n2}^{(1)})$ into $(Y_{n1}^{(1q)}, Y_{n1}^{(1o)}, Y_{n2}^{(1)})$, and $(Y_{n1}^{(1q)}, Y_{n1}^{(1o)}, Y_{n2}^{(1)})$ into $(X_{n1}^{(1)}, X_{n2}^{(1)})$; etc. Denote by $(X_1^{(i)}, X_2^{(i)})$ the vector with as distribution the limiting distribution of $(X_{n1}^{(i)}, X_{n2}^{(i)})$, $i = 1, 2$, and similarly introduce $(Z_1^{(i)}, Z_2^{(i)})$ and $(Y_1^{(i)}, Y_2^{(i)})$, with $Y_1^{(1)} = Y_1^{(1q)} + Y_1^{(o)}$ and with $Y_2^{(2)} = Y_2^{(2q)} + Y_2^{(o)}$, for $i = 1, 2$. We have:

$$\mathbb{E}[z_1^{X_1^{(2)}} z_2^{X_2^{(2)}}] = \sigma_1(z_1, z_2)\mathbb{E}[z_1^{Z_1^{(1)}} z_2^{Z_2^{(1)}}]. \tag{3.1}$$

$$\mathbb{E}[z_1^{Z_1^{(1)}} z_2^{Z_2^{(1)}} | Y_1^{(1q)} = h_{1q}, Y_1^{(1o)} = h_{1o}, Y_2^{(1)} = h_2] = z_1^{h_{1o}} z_2^{h_2} \beta_1^{h_{1q}}(z_1, z_2), \tag{3.2}$$

yielding

$$\mathbb{E}[z_1^{Z_1^{(1)}} z_2^{Z_2^{(1)}}] = \mathbb{E}[\beta_1(z_1, z_2)^{Y_1^{(1q)}} z_1^{Y_1^{(1o)}} z_2^{Y_2^{(1)}}]. \tag{3.3}$$

Furthermore,

$$\begin{aligned} &\mathbb{E}[z_{1q}^{Y_1^{(1q)}} z_{1o}^{Y_1^{(1o)}} z_2^{Y_2^{(1)}} | X_1^{(1)} = i_1, X_2^{(1)} = i_2] \\ &= z_2^{i_2} \mathrm{e}^{-\lambda_2(1-z_2)G_1} \mathrm{e}^{-\lambda_1(1-z_{1q})G_1} [(1 - \mathrm{e}^{-\nu_1 G_1}) z_{1q} + \mathrm{e}^{-\nu_1 G_1} z_{1o}]^{i_1}, \end{aligned} \tag{3.4}$$

yielding

$$\begin{aligned} \mathbb{E}[z_{1q}^{Y_1^{(1q)}} z_{1o}^{Y_1^{(1o)}} z_2^{Y_2^{(1)}}] &= \mathrm{e}^{-\lambda_2(1-z_2)G_1} \mathrm{e}^{-\lambda_1(1-z_{1q})G_1} \\ &\times \mathbb{E}[[(1 - \mathrm{e}^{-\nu_1 G_1}) z_{1q} + \mathrm{e}^{-\nu_1 G_1} z_{1o}]^{X_1^{(1)}} z_2^{X_2^{(1)}}]. \end{aligned} \tag{3.5}$$

It follows from (3.1), (3.3) and (3.5), with

$$f_1(z_1, z_2) := (1 - \mathrm{e}^{-\nu_1 G_1})\beta_1(z_1, z_2) + \mathrm{e}^{-\nu_1 G_1} z_1, \tag{3.6}$$

that

$$\mathbb{E}[z_1^{X_1^{(2)}} z_2^{X_2^{(2)}}] = \sigma_1(z_1, z_2)\mathrm{e}^{-\lambda_1(1-\beta_1(z_1,z_2))G_1 - \lambda_2(1-z_2)G_1} \times\mathbb{E}[f_1(z_1, z_2)^{X_1^{(1)}} z_2^{X_2^{(1)}}]. \tag{3.7}$$

Similarly we have, with

$$f_2(z_1, z_2) := (1 - \mathrm{e}^{-\nu_2 G_2})\beta_2(z_1, z_2) + \mathrm{e}^{-\nu_2 G_2} z_2, \tag{3.8}$$

that

$$\mathbb{E}[z_1^{X_1^{(1)}} z_2^{X_2^{(1)}}] = \sigma_2(z_1, z_2)\mathrm{e}^{-\lambda_1(1-z_1)G_2 - \lambda_2(1-\beta_2(z_1,z_2))G_2} \times\mathbb{E}[z_1^{X_1^{(2)}} f_2(z_1, z_2)^{X_2^{(2)}}]. \tag{3.9}$$

It follows from (3.7) and (3.9) that

$$\begin{aligned}\mathbb{E}[z_1^{X_1^{(1)}} z_2^{X_2^{(1)}}] = {} & \sigma_1(z_1, f_2(z_1, z_2))\sigma_2(z_1, z_2)\mathrm{e}^{-\lambda_1(1-z_1)G_2 - \lambda_2(1-\beta_2(z_1,z_2))G_2} \\ & \times\mathrm{e}^{-\lambda_1(1-\beta_1(z_1,f_2(z_1,z_2)))G_1 - \lambda_2(1-f_2(z_1,z_2))G_1} \\ & \times\mathbb{E}[f_1(z_1, f_2(z_1, z_2))^{X_1^{(1)}} f_2(z_1, z_2)^{X_2^{(1)}}].\end{aligned} \tag{3.10}$$

We can rewrite this, with

$$h_1(z_1, z_2) := f_1(z_1, f_2(z_1, z_2)), \quad h_2(z_1, z_2) := f_2(z_1, z_2), \tag{3.11}$$

and

$$X(z_1, z_2) := \mathbb{E}[z_1^{X_1^{(1)}} z_2^{X_2^{(1)}}], \tag{3.12}$$

and with an obvious definition of $K(\cdot, \cdot)$, as

$$X(z_1, z_2) = K(z_1, z_2)X(h_1(z_1, z_2), h_2(z_1, z_2)). \tag{3.13}$$

Define

$$h_i^{(0)}(z_1, z_2) := z_i, \quad h_i^{(n)}(z_1, z_2) := h_i(h_1^{(n-1)}(z_1, z_2), h_2^{(n-1)}(z_1, z_2)), \quad i = 1, 2. \tag{3.14}$$

Theorem 3. *If $\rho_1 + \rho_2 < 1$, then the generating function $X(z_1, z_2)$ is given by*

$$X(z_1, z_2) = \prod_{m=0}^{\infty} K(h_1^{(m)}(z_1, z_2), h_2^{(m)}(z_1, z_2)). \tag{3.15}$$

Proof. Equation (3.15) follows from (3.13) by iteration. We still need to prove that the infinite product converges if $\rho_1 + \rho_2 < 1$. Equation (3.13) is an equation for a multi-type branching process with immigration, where the number of immigrants of different types has generating function $K(z_1, z_2)$ and the number of children of different types of a type 1 individual in the branching process has generating function $h_1(z_1, z_2)$ and the number of children of different types of a type 2 individual in the branching process has generating function $h_2(z_1, z_2)$. An important role in the analysis of such a process is played by the mean matrix M of the branching process,

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}, \tag{3.16}$$

where m_{ij} represents the mean number of children of type j of a type i individual. In our case, the elements of the matrix M are given by

$$m_{11} = \mathrm{e}^{-\nu_1 G_1} + \left(1 - \mathrm{e}^{-\nu_1 G_1}\right)\rho_1 + \left(1 - \mathrm{e}^{-\nu_1 G_1}\right)\left(1 - \mathrm{e}^{-\nu_2 G_2}\right)\rho_1\rho_2, \tag{3.17}$$

$$m_{12} = \left(1 - \mathrm{e}^{-\nu_1 G_1}\right)\lambda_2 \mathbb{E}B_1 \left(\mathrm{e}^{-\nu_2 G_2} + \left(1 - \mathrm{e}^{-\nu_2 G_2}\right)\rho_2\right), \tag{3.18}$$

$$m_{21} = \left(1 - \mathrm{e}^{-\nu_2 G_2}\right)\lambda_1 \mathbb{E}B_2, \tag{3.19}$$

$$m_{22} = \mathrm{e}^{-\nu_2 G_2} + \left(1 - \mathrm{e}^{-\nu_2 G_2}\right)\rho_2. \tag{3.20}$$

For example, formula (3.18) can be explained as follows. A type 1 customer present in the system at the beginning of a glue period of Q_1 is served with probability $1 - \mathrm{e}^{-\nu_1 G_1}$. If he is served, on average $\lambda_2 \mathbb{E}B_1$ type 2 customers will arrive during his service time. During the visit period of Q_2 each of these customers is not served with probability $\mathrm{e}^{-\nu_2 G_2}$ or served with probability $1 - \mathrm{e}^{-\nu_2 G_2}$, in which case on average ρ_2 type 2 customers will arrive during this service time.

The theory of multitype branching processes with immigration (see Quine [15] and Resing [16]) now states that if (i) the expected total number of immigrants in a generation is finite and (ii) the maximal eigenvalue λ_{max} of the mean matrix M satisfies $\lambda_{max} < 1$, then the generating function of the steady state distribution of the process is given by (3.15). To complete the proof of Theorem 3, we shall now verify (i) and (ii).

Ad (i): The expected total number of immigrants in a generation is

$$\begin{aligned}
&(\lambda_1 + \lambda_2)\mathbb{E}S_2 + \lambda_1 G_2 + \lambda_2 G_2 (\lambda_1 + \lambda_2)\mathbb{E}B_2 \\
&+\lambda_1 \mathbb{E}S_1 + \lambda_2 \mathbb{E}S_1 \left(\mathrm{e}^{-\nu_2 G_2} + \left(1 - \mathrm{e}^{-\nu_2 G_2}\right)(\lambda_1 + \lambda_2)\mathbb{E}B_2\right) \\
&+\lambda_1 G_1 \left(\lambda_1 \mathbb{E}B_1 + \lambda_2 \mathbb{E}B_1 \left(\mathrm{e}^{-\nu_2 G_2} + \left(1 - \mathrm{e}^{-\nu_2 G_2}\right)(\lambda_1 + \lambda_2)\mathbb{E}B_2\right)\right) \\
&+\lambda_2 G_1 \left(\mathrm{e}^{-\nu_2 G_2} + \left(1 - \mathrm{e}^{-\nu_2 G_2}\right)(\lambda_1 + \lambda_2)\mathbb{E}B_2\right),
\end{aligned} \tag{3.21}$$

and hence indeed finite. Here, the term $(\lambda_1 + \lambda_2)\mathbb{E}S_2$ corresponds to the customers arriving during the switch period out of Q_2. The term $\lambda_1 G_2$ corresponds to the type 1 customers arriving during the glue period of Q_2. The term $\lambda_2 G_2 (\lambda_1 + \lambda_2)\mathbb{E}B_2$ corresponds to the type 2 customers arriving during the glue period of Q_2. These customers are served during the visit period of Q_2 and during their service time other customers will arrive. The term $\lambda_1 \mathbb{E}S_1$ corresponds to the type 1 customers arriving during the switch period out of Q_1. The term $\lambda_2 \mathbb{E}S_1 \left(\mathrm{e}^{-\nu_2 G_2} + \left(1 - \mathrm{e}^{-\nu_2 G_2}\right)(\lambda_1 + \lambda_2)\mathbb{E}B_2\right)$ corresponds to the type

2 customers arriving during the switch period out of Q_1. These customers are served during the visit period of Q_2 with probability $1 - \mathrm{e}^{-\nu_2 G_2}$ (in which case again other customers will arrive during their service time) and with probability $\mathrm{e}^{-\nu_2 G_2}$ they are not served during this visit period. The last two terms in (3.21) correspond to the type 1 and type 2 customers arriving during the glue period of Q_1.

Ad (ii): Define the matrix

$$H = \begin{pmatrix} \mathrm{e}^{-\nu_1 G_1} + \left(1 - \mathrm{e}^{-\nu_1 G_1}\right)\rho_1 & \left(1 - \mathrm{e}^{-\nu_1 G_1}\right)\lambda_2 \mathbb{E}B_1 \\ \left(1 - \mathrm{e}^{-\nu_2 G_2}\right)\lambda_1 \mathbb{E}B_2 & \mathrm{e}^{-\nu_2 G_2} + \left(1 - \mathrm{e}^{-\nu_2 G_2}\right)\rho_2 \end{pmatrix}, \tag{3.22}$$

where the elements h_{ij} of the matrix H represent the mean number of type j customers that replace a type i customer during a visit period of Q_i (either new arrivals if the customer is served, or the customer itself if it is not served). We have that

$$H \begin{pmatrix} \mathbb{E}B_1 \\ \mathbb{E}B_2 \end{pmatrix} = \begin{pmatrix} \left[\mathrm{e}^{-\nu_1 G_1} + \left(1 - \mathrm{e}^{-\nu_1 G_1}\right)(\rho_1 + \rho_2)\right]\mathbb{E}B_1 \\ \left[\mathrm{e}^{-\nu_2 G_2} + \left(1 - \mathrm{e}^{-\nu_2 G_2}\right)(\rho_1 + \rho_2)\right]\mathbb{E}B_2 \end{pmatrix} < \begin{pmatrix} \mathbb{E}B_1 \\ \mathbb{E}B_2 \end{pmatrix} \tag{3.23}$$

if and only if $\rho_1 + \rho_2 < 1$. Using this result and following the same line of proof as in Section 5 of Resing [16], we can show that the stability condition $\rho_1 + \rho_2 < 1$ implies that also the maximal eigenvalue λ_{max} of the mean matrix M satisfies $\lambda_{max} < 1$. This concludes the proof.

4 Conclusions and Suggestions for Future Research

In this paper we have studied vacation queues and two-queue polling models with the gated service discipline and with retrials. Motivated by optical communications, we have introduced a glue period just before a server visit; during such a glue period, new customers and retrials "stick" instead of immediately going into orbit. For both the vacation queue and the two-queue polling model, we have derived steady-state queue length distributions at an arbitrary epoch and at various specific epochs. This was accomplished by establishing a relation to branching processes. We have thus laid the groundwork for the performance analysis of an N-queue polling model with retrials.

In future studies, we shall not only turn to that N-queue model; we also would like to consider other service disciplines. Furthermore, the following model variants seem to fall within our framework: (i) customers may *not* retry with a certain probability; (ii) the arrival rates may be different for visit, vacation and glue periods; (iii) one might allow that new arrivals during a glue period are already served during that glue period.

We would also like to explore the possibility to study the heavy traffic behavior of these models via the relation to branching processes, cf. [14].

Finally, we would like to point out an important advantage of optical fibre: the wavelength of light. A fibre-based network node may thus route incoming packets not only by switching in the time-domain, but also by wavelength division multiplexing. In queueing terms, this gives rise to *multiserver* polling models, each server representing a wavelength. We refer to [1] for the stability analysis of multiserver polling models, and to [2] for a mean field approximation of large passive optical networks. It would be very interesting to study multiserver polling models with the additional features of retrials and glue periods.

Acknowledgment. The authors gratefully acknowledge fruitful discussions with Kevin ten Braak and Tuan Phung-Duc about retrial queues and with Ton Koonen about optical networks. The research is supported by the IAP program BESTCOM, funded by the Belgian government.

References

1. Antunes, N., Fricker, C., Roberts, J.: Stability of multi-server polling system with server limits. Queueing Systems 68, 229–235 (2011)
2. Antunes, N., Fricker, C., Robert, P., Roberts, J.: Traffic capacity of large WDM Passive Optical Networks. In: Proceedings 22nd International Teletraffic Congress, ITC 22, Amsterdam (September 2010)
3. Artalejo, J.R., Gomez-Corral, A.: Retrial Queueing Systems: A Computational Approach. Springer, Berlin (2008)
4. Athreya, K.B., Ney, P.E.: Branching Processes. Springer, Berlin (1972)
5. Boon, M.A.A., van der Mei, R.D., Winands, E.M.M.: Applications of polling systems. SORMS 16, 67–82 (2011)
6. Boxma, O.J., Cohen, J.W.: The $M/G/1$ queue with permanent customers. IEEE J. Sel. Areas in Commun. 9, 179–184 (1991)
7. Falin, G.I., Templeton, J.G.C.: Retrial Queues. Chapman and Hall, London (1997)
8. Koonen, A.M.J.: Personal communication (2014)
9. Langaris, C.: A polling model with retrial of customers. Journal of the Operations Research Society of Japan 40, 489–507 (1997)
10. Langaris, C.: Gated polling models with customers in orbit. Mathematical and Computer Modelling 30, 171–187 (1999)
11. Langaris, C.: Markovian polling system with mixed service disciplines and retrial customers. Top 7, 305–322 (1999)
12. Levy, H., Sidi, M.: Polling models: Applications, modeling and optimization. IEEE Trans. Commun. 38, 1750–1760 (1990)
13. Maier, M.: Optical Switching Networks. Cambridge University Press, Cambridge (2008)
14. Olsen, T.L., van der Mei, R.D.: Periodic polling systems in heavy traffic: Distribution of the delay. Journal of Applied Probability 40, 305–326 (2003)
15. Quine, M.P.: The multitype Galton-Watson process with immigration. Journal of Applied Probability 7, 411–422 (1970)
16. Resing, J.A.C.: Polling systems and multitype branching processes. Queueing Systems 13, 409–426 (1993)
17. Rogiest, W.: Stochastic Modeling of Optical Buffers. Ph.D. Thesis, Ghent University, Ghent, Belgium (2008)

18. Takagi, H.: Application of polling models to computer networks. Comput. Netw. ISDN Syst. 22, 193–211 (1991)
19. Takagi, H.: Queueing Analysis: A Foundation of Performance Evaluation. Vacation and Priority Systems, vol. 1. Elsevier Science Publishers, Amsterdam (1991)
20. Takagi, H.: Queueing analysis of polling models: progress in 1990-1994. In: Dshalalow, J.H. (ed.) Frontiers in Queueing: Models, Methods and Problems, pp. 119–146. CRC Press, Boca Raton (1997)
21. Takagi, H.: Analysis and application of polling models. In: Reiser, M., Haring, G., Lindemann, C. (eds.) Dagstuhl Seminar 1997. LNCS, vol. 1769, pp. 424–442. Springer, Heidelberg (2000)
22. Vishnevskii, V.M., Semenova, O.V.: Mathematical methods to study the polling systems. Autom. Remote Control 67, 173–220 (2006)