A queueing system with random processing batches and vacations

George K. Mytalas and Michael A. Zazanis*
Department of Statistics
Athens University of Economics and Business
76 Patission str., Athens 104 34, Greece.

Abstract

In this paper we examine a queueing model with Poisson arrivals, random processing batches, and vacations is examined. is interrupted service process along the lines of the Queue-Cart model of Coffman and Gilbert [7] and we extend their results to the case where the cart's capacity varies stochastically and customers arrive in batches. We also examine processing batch distributions with unbounded support and provide a solution via Wiener-Hopf techniques. In all cases care is taken in the analysis in order to obtain the steady state distribution without the assumption that the service and cart delivery time distributions are light-tailed. If fact, our results are obtained under the natural conditions of finite first moments, together with the stability condition which gurantees the existence of a stationary version of the process. We finally provide further results regarding waiting time distributions using the distributional version of Little's law.

KEYWORDS: QUEUEING, MANUFACTURING, BULK SERVICE QUEUES, MATERIALS HANDLING.

1 Model description

We analyze a model which consists of an M/G/1 queue with vacations in the service mechanism. Customers arrive according to a Poisson process with rate $\lambda > 0$ to the system and have i.i.d. service requirements which we will denote by $\{\sigma_n\}_{n=1}^{\infty}$. These are assumed to be independent of the arrival process and their common distribution will be denoted by $B(x) := P(\sigma \leq x)$ with finite mean $E\sigma$. The capacity of the queue is assumed to be infinite. At specific time epochs the server initiates a "vacation period" during which it is unavailable to serve customers, while arriving customers accumulate in the waiting area. Successive vacation times form a sequence of i.i.d. random variables, independent of the arrival process and service requirements, denoted by $\{G_n\}_{n=1}^{\infty}$, with common distribution, $G(x)$, and finite mean, EG. The server's operation alternates thus between *service phases* and *vacation phases.*

At the beginning of each service phase a processing batch the size of the processing batch is set. Processing batches are assumed to be i.i.d. random variables, independent of the arrival process, service

*Corresponding author. `email:zazanis@aueb.gr`

requirements, and vacation lengths, and will be denoted by $\{\Theta_m\}$. Their common distribution will be denoted by $\theta_k = P(\Theta = k)$, $k = 1, 2, \ldots$, and will be assumed to have finite mean $E\Theta = \sum_{k=1}^{\infty} k\theta_k < \infty$. During the service phase customers are served in a FIFO fashion until either the number of customers served in the phase becomes equal to the processing batch size, Θ_m or the queue empties, whichever happens first. In both cases the server initiates a new vacation phase. We shall call this the *partial batch* policy, since it is possible that fewer customers than required by the processing batch are served in the phase. When the server returns from the vacation an new processing batch is set and a new service phase begins. If, upon returning from a vacation, the server finds the queue empty then we will assume that he immediately takes a new vacation (Thus we allow service phases to have zero duration.) Variations in the behavior of the server when, upon returning from a vacation, finds the queue empty are possible. For instance we may suppose that in such cases the server waits for a fixed period of time before and only if this elapses without arrivals he leaves again or that he waits until a fixed number of customers arrives etc. Such variations do not burden the analysis but, since they have been studied extensively in the vacations literature, they will not be considered here.

We will also consider the "*complete batch*" policy according to which, when a service phase is initiated and a production batch is set the server remains available, waiting for a customer to arrive if necessary, and works until the production batch is complete. After this, the vacation phase begins during which the server is unavailable. At the end of the vacation a new cycle begins with a new production batch determined at random from the given distribution.

The system described above is a type of an M/G/1 queue in a random environment. Under the assumption that the production batch has fixed size, say $\Theta_m = N$ with probability 1, this system has been studied in Coffman and Gilbert [7]. There, the fixed production batch is interpreted as the capacity of a cart placed next to the processing station. Finished parts are placed in the cart and when the cart is full the server takes the cart to its destination. Thus, server vacations in that model correspond to the time it takes the server to deliver the cart. If we suppose that the same cart is used to store the output of two or more stations served by the same server then the need for a cart with stochastic capacity arrises naturally.

The model we propose has also applications to queueing systems with unreliable servers. Suppose that the server is subject to failures. These failures are assumed to manifest themselves at the initialization phase of service and to be independent of the service requirement of the customer. Under these conditions the random processing batch model proposed constitutes an accurate model. Vacation periods correspond then to down time for the system while the server is being repaired. In this context the complete batch policy described above is more appropriate. The partial batch policy may be appropriate if we assume that idle periods are used for preventive maintenance. In this case a model with vacations whose duration distribution depends on whether the preceding processing batch has been completed or not may be more appropriate. The analysis of such models will be sketched in section 4.4.

Throughout the paper the analysis is carried out by distinguishing two cases, according to whether the production batch distribution has bounded or unbounded support. In the first case, where the support of the production batch size distribution is bounded above by a constant N (this could be the cart's capacity in the first model mentioned above) the analysis is based on an argument using Rouché's theorem which is typical of the analysis of queues with batch service (see [4]). In this respect attention has been paid in order to establish our results under the natural conditions for the existence of a stationary version of the process i.e. *the finiteness of first moments* plus the stability condition of the system.

The second case, where the distribution of the production batch size has unbounded support, is in general harder and can only be dealt with by Wiener-Hopf techniques. We indicate how to carry out this procedure and we also provide explicit solutions for the case of production batches whose distribution is either geometric or a combination of geometric factors. In a variation of the above model, we will also consider briefly the case where customers do not arrive singly but in i.i.d. batches.

In all cases the analysis of the system proceeds by first analyzing an embedded Markov chain by means of generating functions and then using standard results from semi-regenerative processes in order to obtain the stationary distribution of the number of customers in the system.

2 The stability condition

Here we sketch the argument that gives the stability condition for this system both for the "partial batch" and for the "complete batch" policy. While the argument is expressed in a heuristic fashion it can easily be turned into a rigorous proof.

Let us consider first the partial batch policy. The queueing process can be thought of as consisting of cycles, each cycle comprising a vacation phase and the following service phase. (The service phase may have zero duration with positive probability corresponding to the case where a returning server finds the queue empty.) Let A be a random variable distributed according to the total number of customers who *arrive* during the a typical cycle and L a random variable distributed according to the number of customers *served* during a typical cycle. If a stationary version of the system exists then the expected number of customers who arrive during a cycle must be equal to the expected number of customers served during that cycle, i.e. $EA = EL$. The number of arrivals during the cycle can be distinguished into arrivals during the vacation phase and arrivals during the service phase. The expected number of arrivals during the vacation phase is equal to λEG. During the service phase these customers are to be served, together with all the customers who arrive during *their* service time (this of course holding only on the average). However, each one of the customers who arrives during the vacation phase generates a busy period which contains on the average $\frac{1}{1-\lambda E\sigma}$ customers, including himself. Thus we have $EA = \frac{\lambda EG}{1-\lambda E\sigma}$. In order for the balance to be maintained, a total of $EL = EA$ customers must be served during the service phase. However, the number served during the service phase cannot exceed the "service capacity" of the cycle i.e. the expected production batch size which is $E\Theta$. In fact, since under the partial batch policy the cart leaves with a partial batch with positive probability, $E\Theta > EL$. Thus the necessary and sufficient condition for stability is

$$E\Theta > \frac{\lambda EG}{1 - \lambda E\sigma}. \tag{1}$$

We now turn to the complete batch policy. Here the number served in each cycle is equal to the production batch size and hence $EL = E\Theta$. For the same reasons as above, $EL = EA$. However, the expected number of customers arriving within a cycle will be greater than $\frac{\lambda EG}{1-\lambda E\sigma}$ since, with positive probability the queue empties before the production batch is complete and the server has to wait for new arrivals. The above heuristic derivation echoes that of Coffman and Gilbert [7] and can be justified by the detailed analysis in the following section. We included it here because we believe that it adds insight into the operation of the system.

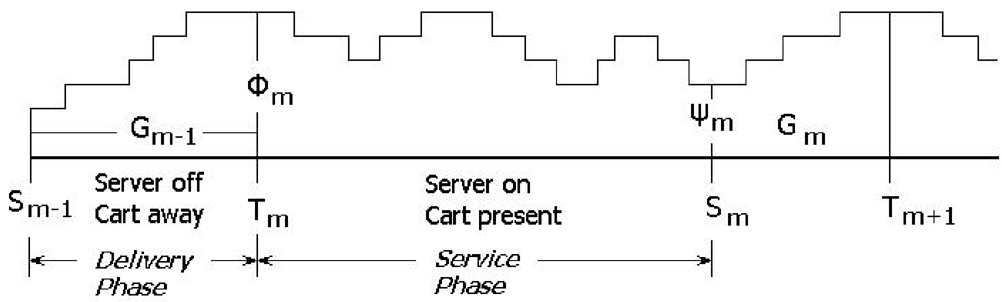

Figure 1: Sample path of the queue.

3 Analysis of the embedded Markov chain of the system under a *partial batch* policy

3.1 Notation

We consider the embedded point process of the epochs when the server returns to the queue at the end of a vacation phase, i.e. at the end of a vacation. We denote these points by $\{T_m; m \in \mathbf{Z}\}$. Let us also denote by $\{S_m; m \in \mathbf{Z}\}$ the corresponding epochs when the server leaves the queue to deliver the cart, i.e. the beginnings of vacations. We shall think of the sample path of the process as consisting of cycles. Each cycle comprises a *service phase* where the server is present and serving customers, and a *vacation phase* during which the server is away, delivering the cart to its destination. The number of customers in the system at time t is denoted by X_t and the process $\{X_t;\, t \in \mathbf{R}\}$ is assumed to have right–continuous sample paths. The mth cycle starts at T_m with the end of a vacation. We denote by Φ_m the number of customers in the system at epoch T_m, (i.e. $\Phi_m = X_{T_m}$). This means that at the start of the mth cycle, i.e. at the moment when the server returns with the cart to the queue, he finds Φ_m customers waiting for service. Clearly, (T_m, Φ_m), $m \in \mathbf{Z}$, is a Markov–renewal process and $\{X_t;\, t \in \mathbf{R}\}$ is a *semi–regenerative process* with respect to it. Also denote by Ψ_m the number of customers left behind in the queue at epoch S_m when the server leaves the system to deliver the cart, i.e. $\Psi_m = X_{S_m}$. Finally we will denote by L_m the number of services in the mth cycle which is equal to the contents of the cart when it leaves. Clearly we have $L_m \leq \Theta_m$, and $\Psi_m = 0$ if $L_m < \Theta_m$ since we assume that a partial batch policy is used. Also recall that, according to this policy, if $\Phi_m = 0$ then the server does not stay in the queue at all but immediately takes another vacation. Hence, in that case $S_m = T_m$, and $\Psi_m = \Phi_m = 0 = L_m$. Figure 1 illustrates these definitions.

Following the approach of Coffman and Gilbert [7] we let d_k^m be the epoch of the kth service completion during the mth cycle. We will agree to set $d_0^m = T_m$. Clearly, in the mth cycle we have $T_m = d_0^m < d_1^m < d_2^m < \cdots < d_{L_m}^m$. Let $X_{d_k^m}$ be the number of customers left behind at the kth epoch of the mth cycle and in particular note that $X_{d_0^m} = \Phi_m$. We will assume that the system is stationary and we will analyze its behavior over "a typical cycle". Therefore, without risk of confusion, we will drop the subscript m referring to a particular cycle in what follows. Suppose that the system has been operating in stationarity and that time $t = 0$ coincides with $d_0^m = T_m$ (in other words consider the Palm version of the process with respect to the point process $\{T_m\}$). Note that, under the partial batch policy,

$$\{L \geq k\} = \{X_{d_0} > 0, X_{d_1} > 0, \ldots, X_{d_{k-1}} > 0\} \cap \{\Theta \geq k\} \tag{2}$$

and

$$\{L = k\} = \{X_{d_0} > 0, X_{d_1} > 0, \ldots, X_{d_{k-1}} > 0\} \cap (\{\Theta = k\} \cup \{X_{d_k} = 0\}), \qquad k = 1, 2, \ldots.$$

whereas $\{L = 0\} = \{X_{d_0} = 0\}$. We define the generating functions

$$Q_k(z) = E[z^{X_{d_k}}; L \geq k] \tag{3}$$

and set

$$\begin{aligned} F_k &= Q_k(0) = P\left(X_{d_k} = 0; L \geq k\right) \;=\; P\left(X_{d_k} = 0; L = k\right) \;=\; P\left(X_{d_k} = 0; L = k; \Theta \geq k\right) \\ &= P\big(X_{d_0} > 0, X_{d_1} > 0, \ldots, X_{d_{k-1}} > 0, X_{d_k} = 0; \Theta \geq k\big). \end{aligned} \tag{4}$$

Note that F_k is the probability that the typical service phase consists of precisely k services and that the next vacation phase starts with an empty queue. In section 9 their role in determining the statistics on the cart contents is examined in detail. We also point out that, in view of (2) and (3)

$$Q_k(z) = E[z^{X_{d_k}}; L \geq k \mid \Theta \geq k] = E[z^{X_{d_k}}; L \geq k \mid \Theta = n] \quad \text{for } n = k, k+1, k+1, \ldots. \tag{5}$$

Furthermore, with B denoting the service time distribution and B^* the corresponding Laplace transform,

$$U(z) := B^*(\lambda(1-z)),$$

is the p.g.f. (probability generating function) of the number of arrivals during a service time. Similarly, with G and G^* denoting the distribution and Laplace transform respectively of the vacation period for the server,

$$D(z) := G^*(\lambda(1-z))$$

is the p.g.f. of the number of arrivals during a server vacation time. We also define for convenience the quantities

$$\alpha(z) := U(z)z^{-1}, \qquad y(z) := \alpha^{-1}(z). \tag{6}$$

3.2 Random production batch size with finite support

The "dynamics" of the process during a service period (i.e. during intervals of the form $(S_m, T_{m+1}), m \in$ **Z**) are described by the following basic recursive relationship which involves the generating functions defined in (3) and (6)

$$Q_{k+1}(z) = (Q_k(z) - F_k)\,\alpha(z), \qquad k = 0, 1, \ldots, N-1. \tag{7}$$

This recursion expresses the fact that the number of customers left behind at the end of the $(k+1)$th service completion are equal to the number left behind at the kth service completion minus one plus the number that arrived during this service time, *provided that the queue has not emptied* and that the production batch size is at least equal to $k+1$ or greater. From it we readily obtain

$$Q_n(z) = \alpha(z)^n Q_0(z) - \sum_{k=0}^{n-1} F_k \alpha(z)^{n-k}, \tag{8}$$

$n = 1, 2, \ldots, N$. By definition

$$Q_0(z) = E\left[z^{X_{d_0}}; L \geq 0\right] = E\left[z^{X_{d_0}}\right] = E[z^{\Phi}] \tag{9}$$

is the p.g.f. of the number of customers in the queue at an epoch when the service phase begins. (Of course $P(L \geq 0) = 1$). Also, the p.g.f. of the number of customers left behind in the queue after the server leaves in order to deliver the cart is given by

$$\Pi(z) := E[z^{\Psi}] = \sum_{n=1}^{N} \theta_n \left(Q_n(z) + \sum_{k=0}^{n-1} F_k \right). \tag{10}$$

Indeed, conditioning on the production batch size to be equal to n, for the typical cycle in stationarity, F_k, $k = 0, 1, \ldots, n-1$, is the probability that the server leaves behind an empty queue and the cart contains k customers, i.e. a partial production batch, while $F_n = Q_n(0)$ is the probability that the server leaves behind an empty queue and the cart leaves with a complete batch of n customers. Thus $E[z^{\Psi} \mid \Theta = n] = Q_n(z) + \sum_{k=0}^{n-1} F_k$ and (10) follows by taking expectation over Θ. Taking into account (8) we obtain

$$\Pi(z) = \sum_{n=1}^{N} \theta_n \left(\alpha(z)^n Q_0(z) + \sum_{k=0}^{n-1} F_k \left(1 - \alpha(z)^{n-k}\right) \right). \tag{11}$$

On the other hand the number of customers in the system at the beginning of the typical service phase is equal to the number left behind at the end of the previous service phase plus the number of customers who arrived during the intervening vacation phase. The p.g.f. of these latter is $D(z)$ and thus we have, under stationarity,

$$\Pi(z)D(z) = Q_0(z). \tag{12}$$

We also note that $P(\Psi = 0 \mid \Theta = n) = \sum_{k=0}^{n} F_k$ and thus (11) is established with

$$Q_0(z) = \sum_{n=1}^{N} \theta_n \left(\alpha(z)^n Q_0(z) + \sum_{k=0}^{n-1} F_k \left(1 - \alpha(z)^{n-k}\right) \right) D(z). \tag{13}$$

Before proceeding we point out that in the sequel we will occasionally be *dropping the dependence of some generating functions on* z for notational convenience. Thus we will be writing y instead of $y(z)$, D instead of $D(z)$, and so forth. From (8), (13), and (6), we conclude that

$$Q_0 \left(y^N - D \sum_{n=1}^{N} \theta_n y^{N-n} \right) = D \sum_{n=1}^{N} \theta_n \sum_{k=0}^{n-1} F_k \left(y^N - y^{N-n+k} \right) \tag{14}$$

or equivalently

$$\Pi(z) = \frac{\sum_{k=0}^{N-1} F_k \sum_{n=k+1}^{N} \theta_n \left(y^N - y^{N-n+k} \right)}{y^N - D \sum_{n=1}^{N} \theta_n y^{N-n}}. \tag{15}$$

The above can also be written as

$$\Pi(z) = \frac{\sum_{k=0}^{N-1} F_k \sum_{n=k+1}^{N} \theta_n \left(z^N - z^{N-n+k} U^{n-k} \right)}{z^N - D \sum_{n=1}^{N} \theta_n z^{N-n} U^n}. \tag{16}$$

The N constants, $F_0, F_1, \ldots, F_{N-1}$, can be obtained from Rouché's theorem as follows. It is shown in the Appendix (cf. [7]) that the equation

$$z^N - D(z) \sum_{n=1}^{N} \theta_n z^{N-n} U(z)^n = 0 \tag{17}$$

has N complex roots, $z_0, z_1, \ldots, z_N$, where $z_0 = 1$ and the remaining $N-1$ roots are within the unit circle, i.e. $|z_i| < 1$ for $i = 1, 2, \ldots, N-1$, provided that the stability condition holds. We thus know that equation (17) has precisely N zeros that satisfy $|z| \leq 1$. One of them is $z = 1$ which obviously satisfies $z^N - D(z) \sum_{n=1}^{N} \theta_n z^{N-n} U(z)^n = 0$ and is a single root. Thus there remain $N-1$ roots of the denominator in the unit disk which we shall call z_i, $i = 1, 2, \ldots, N-1$. Since $Q_0(z)$ does not have any singularities within the unit disk these must also be zeros of the numerator of (16). Hence the N unknown constants, $F_0, F_1, \ldots, F_{N-1}$ must satisfy the $N-1$ equations

$$\sum_{k=0}^{N-1} F_k \sum_{n=k+1}^{N} \theta_n \left(z_i^N - z_i^{N-n+k} U(z_i)^{n-k} \right) = 0, \qquad i = 1, 2, \ldots, N-1.$$

Let

$$y_i := \frac{z_i}{U(z_i)}, \qquad i = 1, 2, \ldots, N-1. \tag{18}$$

Considering Q_0 as a function of y, the y_i 's must also be zeros of the numerator of (15), or equivalently, taking into account (18), together with the fact that the z_i's satisfy (17), and $U(z_i) \neq 0$ we have

$$\sum_{k=0}^{N-1} F_k \sum_{n=k+1}^{N} \theta_n \left(y_i^N - y_i^{N-n+k} \right) = 0, \qquad i = 1, 2, \ldots, N-1.$$

The polynomial in y

$$P(y) := \sum_{k=0}^{N-1} F_k \sum_{n=k+1}^{N} \theta_n \left(y^N - y^{N-n+k} \right)$$

has degree N and its N roots are $1, y_1, y_2, \ldots, y_{N-1}$. Thus

$$P(y) = C(y-1) \prod_{i=1}^{N-1} (y - y_i) \, . \tag{19}$$

The constant C can be determined by noting that

$$C \prod_{i=1}^{N-1} (1 - y_i) = \lim_{y \to 1} \frac{\sum_{k=0}^{N-1} F_k \sum_{n=k+1}^{N} \theta_n \left(y^N - y^{N-n+k} \right)}{y - 1} \tag{20}$$

$$= \sum_{k=0}^{N-1} F_k \sum_{n=k+1}^{N} \theta_n \, (n - k) \, , \tag{21}$$

where in the last equation we have used de l' Hospital's rule. The quantity on the right hand side of (20) is obtained by determining the value of $Q_0(z)$ when $z = 1$ as follows. Letting $z \to 1$ (or equivalently $y \to 1$) and applying de l'Hospital's rule in (16), we obtain

$$\sum_{k=0}^{N-1} F_k \sum_{n=k+1}^{N} \theta_n \, (n - k) = E\Theta - \frac{\lambda E G}{1 - \rho}. \tag{22}$$

From (19), (21), and (22) we obtain the value of the constant in (19)

$$C = \frac{E\Theta - \frac{\lambda EG}{1-\lambda E\sigma}}{\prod_{i=1}^{N-1}(1-y_i)}.$$

We have thus established the following

Theorem 1. *The number of customers left behind at the end of a typical service phase in steady state is given by*

$$\Pi(z) = \frac{E\Theta - \frac{\lambda EG}{1-\lambda E\sigma}}{y^N - D\sum_{n=1}^{N}\theta_n y^{N-n}}(y-1)\prod_{i=1}^{N-1}\frac{y-y_i}{1-y_i} \tag{23}$$

where y is given by (6) and the y_i's by (18).

As we saw above, the explicit determination of the N constants, $F_0, \ldots, F_{N-1}$, is not necessary for the determination of $\Pi(z)$. Nonetheless, these constants are useful in order to obtain, among other things, statistics for the cart contents when it is delivered. Their computation is given in the appendix. The detailed analysis of the statistics of the cart's contents is undertaken in section 9. Here we confine ourselves to the observation that the probability that a production batch is delivered incomplete is equal to $p_e := \sum_{n=1}^{N}\theta_n\sum_{k=0}^{n-1}F_k$. Changing the order of summation and using the above equations we have

$$p_e = \sum_{k=0}^{N-1}F_k\sum_{n=k+1}^{N}\theta_n = \frac{E\Theta - \frac{\lambda EG}{1-\lambda E\sigma}}{\prod_{i=1}^{N-1}(y_i-1)}\sum_{i=1}^{N-1}y_i.$$

The expected number of customers in the cart when it is delivered can be computed by first conditioning on the size of the production batch:

$$E[L \mid \Theta = n] = \sum_{k=0}^{n-1}kF_k + n\left(1 - \sum_{k=0}^{n-1}F_k\right) = n - \sum_{k=0}^{n-1}F_k(n-k).$$

Taking expectation over the size of the production batch, we then have

$$\begin{aligned} EL &= E\Theta - \sum_{n=1}^{N}\theta_n\sum_{k=0}^{n-1}F_k(n-k) = E\Theta - \sum_{k=0}^{N-1}F_k\sum_{n=k+1}^{N}\theta_n(n-k) \\ &= \frac{\lambda EG}{1-\rho}. \end{aligned} \tag{24}$$

where in the last equation we have made use of (22). Note that the expected contents of the cart i.e. the expected "actual production batch size" is of course less than $E\Theta$ (because of the occurrence of incomplete production batches when the queue empties) and does not depend on the production batch size distribution $\{\theta_n\}$, provided that the stability condition (1) holds.

4 Production batch size with unbounded support

The analysis of the previous section depended on the assumption that the production batch size had a distribution with finite support. As it will readily become clear, no conceptual difficulties are involved in dropping this assumption. However, from a computational point of view, new difficulties arise as the argument based on Rouché's theorem can no longer be used.

Suppose that the cart capacity is, from transfer to transfer, a random variable with distribution

$$P(\Theta = n) = \theta_n, \qquad n = 1, 2, \ldots$$

and corresponding generating function

$$\Theta(z) := \sum_{n=1}^{\infty} \theta_n z^n.$$

(We no longer assume the existence of an integer N such that $\theta_n = 0$ for $n > N$.) We again apply the analysis of the previous section with the same notation as before. Once more the epochs when the server returns after delivering the cart back to queue for the mth time is denoted by T_m while the epoch right after T_m when the server takes the cart (together with any customers that it contains) to be delivered and starts a vacation is denoted by S_m. Here a typical cycle starts, say at T_m, the server serves L_m customers (where $L_m \leq \Theta_m$ and Θ_m is the size of the cart during the mth cycle) and then departs to deliver the cart at time S_m. Let, as in the previous section, $Q_0(z) = E\left[z^{\Phi_m}\right]$, $Q_n(z) = E\left[z^{X_{d_n}}; L \geq n\right]$, $F_n = Q_n(0)$, and $\Pi(z) = E\left[z^{\Psi_m}\right]$ be the p.g.f. of the number of customers left behind when the cart leaves the queue. The fundamental relationship becomes

$$\Pi(z) = \sum_{n=1}^{\infty} \left(Q_n(z) + \sum_{k=0}^{n-1} F_k \right) \theta_n. \tag{25}$$

The basic recursion (7) still holds and thus we have (8) for $n = 1, 2, \ldots$ .Thus

$$\sum_{n=1}^{\infty} \left(\alpha(z)^n Q_0(z) + \sum_{k=0}^{n-1} F_k \left(1 - \alpha(z)^{n-k}\right) \right) \theta_n = \Pi(z).$$

Also, (12) still holds as before and using Fubini's theorem to change the order of summation we can rewrite the above expression as

$$\Pi(z) D(z) \Theta(\alpha(z)) + \sum_{k=0}^{\infty} F_k \sum_{n=1}^{\infty} \theta_{n+k} \left(1 - \alpha(z)^n\right) = \Pi(z)$$

or

$$\Pi(z) = \frac{\sum_{k=0}^{\infty} F_k \sum_{n=1}^{\infty} \theta_{n+k} \left(1 - \alpha(z)^n\right)}{1 - D(z)\Theta(\alpha(z))}. \tag{26}$$

The above equation is the counterpart of equation (16) of the previous section. Note however that the numerator depends on a whole sequence of unknown constants F_k, $k = 0, 1, 2, \ldots$. Clearly the techniques of the previous section cannot be applied here. In fact in this general case a solution can be obtained, at least in principle, using the Wiener-Hopf decomposition technique as described in the sequel.

4.1 Wiener-Hopf decomposition

From equation (6) we have

$$z = yU(z). \tag{27}$$

Using Lagrange's series expansion (e.g. see Copson [6]) if D is an analytic function in a domain containing the origin then $D(z(y))$ is an analytic function of y with series expansion around the origin given by

$$D(z(y)) = \sum_{n=0}^{\infty} y^n \kappa_n$$

where

$$\kappa_0 = D(0) = G^*(\lambda) \quad \text{and} \quad \kappa_n = \frac{1}{n!}\,\frac{d^{n-1}}{dt^{n-1}}\left(D'(t)U(t)\right)\Big|_{t=0}, \qquad n = 1, 2, \ldots. \tag{28}$$

In particular, when $D(z) = z$ the above expression gives

$$z(y) = \sum_{n=1}^{\infty} \frac{y^n}{n!}\,\frac{d^{n-1}}{dt^{n-1}} U(t)\Big|_{t=0}.$$

With the change of variables from z to y and setting $\widetilde{\Pi}(y) := \Pi(z(y))$ equation (26) becomes

$$\widetilde{\Pi}(y) = \frac{\sum_{k=0}^{\infty} F_k \sum_{n=1}^{\infty} \theta_{n+k}\left(1 - y^{-n}\right)}{1 - D(z(y))\Theta(y^{-1})}. \tag{29}$$

Note from (28) that $\kappa_n \geq 0$ and also from (27) that when $y = 1$ then $z = 1$. Thus, $D(z(1)) = \sum_{n=0}^{\infty} \kappa_n = 1$ and hence κ_n, $n = 0, 1, 2, \ldots$, is a probability distribution on the natural numbers. Setting

$$K(y) := \sum_{n=0}^{\infty} y^n \kappa_n,$$

we see that $K(y)$ is the p.g.f. of this distribution. Then, we can use the standard Wiener-Hopf decomposition argument as follows. Let us denote by κ^{*n} the n–fold convolution of the sequence $\{\kappa_m; m = 0, 1, 2, \ldots\}$ with itself, i.e. $\kappa^{*1}_m = \kappa_m$ for all $m = 0, 1, 2, \ldots$, and $\kappa^{*n}_m = \sum_{l=0}^{m} \kappa^{*(n-1)}_{m-l}\kappa_l$, $m = 0, 1, 2, \ldots$, and similarly θ^{*n} will denote the n–fold convolution of $\{\theta_m\}$ with itself. Then we can write

$$\begin{aligned}
\frac{1}{1 - K(y)\Theta(y^{-1})} &= \exp\left(\sum_{n=1}^{\infty} \frac{1}{n} K^n(y)\Theta^n(y^{-1})\right) \\
&= \exp\left(\sum_{n=1}^{\infty} \frac{1}{n} \sum_{m=0}^{\infty} \kappa^{*n}_m y^m \sum_{l=0}^{\infty} \theta^{*n}_l y^{-l}\right) \\
&= \exp\left(\sum_{r=-\infty}^{\infty} y^r \sum_{n=1}^{\infty} \frac{1}{n} \sum_{\{m-l=r\}} \kappa^{*n}_m \theta^{*n}_l\right) \\
&= \exp\left(\sum_{r=1}^{\infty} y^r \sum_{n=1}^{\infty} \frac{1}{n} \sum_{l=1}^{\infty} \kappa^{*n}_{l+r}\theta^{*n}_l\right) \exp\left(\sum_{r=0}^{\infty} y^{-r} \sum_{n=1}^{\infty} \frac{1}{n} \sum_{m=0}^{\infty} \kappa^{*n}_m \theta^{*n}_{m+r}\right)
\end{aligned}$$

and thus

$$\frac{1}{1 - K(y)\Theta(y^{-1})} = \frac{J^-(y^{-1})}{J^+(y)} \tag{30}$$

where, of course,

$$J^+(\zeta) := \exp\left(-\sum_{r=1}^{\infty}\zeta^r\sum_{n=1}^{\infty}\frac{1}{n}\sum_{l=1}^{\infty}\kappa_{l+r}^{*n}\theta_l^{*n}\right), \qquad J^-(\zeta) := \exp\left(\sum_{r=0}^{\infty}\zeta^r\sum_{n=1}^{\infty}\frac{1}{n}\sum_{m=0}^{\infty}\kappa_m^{*n}\theta_{m+r}^{*n}\right),$$

are two functions that are analytic at least within the unit disk, $|\zeta| \le 1$. Then from (29) and (30) we have

$$\widetilde{\Pi}(y)J^+(y) = J^-(y^{-1})\sum_{k=0}^{\infty}F_k\sum_{n=1}^{\infty}\theta_{n+k}\left(1-y^{-n}\right). \tag{31}$$

Since the left hand side is obviously bounded for $|y| \le 1$ and the right hand side is bounded for $\left|y^{-1}\right| \le 1$ or $|y| \ge 1$ it follows from Liouville's theorem that both sides of (31) are equal to a constant, say Λ. Thus

$$P(y) = \frac{\Lambda}{J^+(y)}$$

and

$$\Pi(z) = \frac{\Lambda}{J^+(z/U(z))} = \Lambda\exp\left(\sum_{r=1}^{\infty}z^rU^{-r}(z)\sum_{n=1}^{\infty}\frac{1}{n}\sum_{l=1}^{\infty}\kappa_{l+r}^{*n}\theta_l^{*n}\right).$$

Setting $z = 1$ in the above expression we readily determine the value of B from the requirement that $\Pi(1) = 1$. Thus we have

$$\Pi(z) = \exp\left(\sum_{r=1}^{\infty}\left(z^rU^{-r}(z)-1\right)\sum_{n=1}^{\infty}\frac{1}{n}\sum_{l=1}^{\infty}\kappa_{l+r}^{*n}\theta_l^{*n}\right). \tag{32}$$

The above analysis parallels the analysis of M/G/1 queues with bulk service when the batch size has not bounded support. We refer the reader to Prabhu [22, p. 164]. (See also Kemperman [20] and Keilson [15], [16].)

While the above expression gives the p.g.f. in explicit form, in practice even computation of the first moment would be very arduous. The situation however becomes much simpler if we assume that the production batch size is geometric or a combination of geometric factors. These cases will be examined in the following subsections.

Finally we compute the expected "actual production batch size" i.e. the expected contents of the cart each time it is delivered. The argument is the same as in the finite support case and thus $EL = E\Theta - \sum_{k=0}^{\infty}F_k\sum_{n=k+1}^{\infty}\theta_n(n-k)$. This expectation is can be explicitly computed from (26) since $\Pi(1) = 1$ by an application of de l'Hospital's rule. Again, $EL = \frac{\lambda EG}{1-\rho}$ regardless of the production batch distribution, provided that the stability condition holds.

4.2 Geometric production batch size

As we saw in the previous subsection, the determination of $\Pi(z)$ for general production batch distribution is computationally difficult. However, when the production batch size is geometrically distributed, one can obtain an explicit, computationally tractable solution. One can start in this case with the factorization problem (30) which has a simple solution. Alternatively one could determine the unknown constants F_k, $k = 0, 1, 2, \dots$ in (26) directly as follows. Suppose that

$$\theta_n = (1-\gamma)\gamma^{n-1}, \qquad n = 1, 2, \dots \tag{33}$$

and thus $\Theta(z) = \frac{(1-\gamma)z}{1-\gamma z}$. Then

$$\Pi(z) = \frac{(1-\gamma)\sum_{k=0}^{\infty} F_k \sum_{n=1}^{\infty} \gamma^{n+k-1}\left(1-\alpha(z)^n\right)}{1 - D(z)\frac{(1-\gamma)\alpha(z)}{1-\gamma\alpha(z)}} = \frac{(z-U(z))F(\gamma)}{z-\gamma U(z)-(1-\gamma)D(z)U(z)} \tag{34}$$

where, in the above equation we have used the generating function

$$F(z) = \sum_{k=0}^{\infty} F_k z^k. \tag{35}$$

To determine the unknown quantity $F(\gamma)$ in (34) we use de l'Hospital's rule and the fact that $\Pi(1) = 1$ to obtain

$$F(\gamma) = 1 - (1-\gamma)\frac{\lambda EG}{1-\rho}. \tag{36}$$

Once $\Pi(z)$ has been determined, it is straightforward to evaluate the steady state distribution for the number of customers in the system as we will see in the sequel. We point out that the above model corresponds to the situation where, after each service completion the server "flips a coin" and with probability γ he decides to serve another customer, if one is available or take a vacation if a customer is not available. With probability $1-\gamma$ the server takes a vacation regardless of whether there are customers waiting in line or not. At the end of each vacation the server returns to the queue and, if empty, he immediately takes another vacation whereas if not then the "coin-flipping procedure" begins again. This is the Bernoulli vacation model (see Keilson and Servi [17] and Doshi [8]).

4.3 Linear combination of geometric factors

More generally, we may assume that the production batch size is a linear combination of geometric factors, i.e.

$$\theta_n = \sum_{s=1}^{S} c_s(1-\gamma_s)\gamma_s^{n-1}, \qquad n = 1, 2, \dots,$$

where $0 < \gamma_s < 1$, the γ_s's are assumed to be different from each other, and the c_s's are such that $c_s \neq 0$, $\sum_{s=1}^{S} c_s = 1$, and $\theta_n \geq 0$, $\forall n \in \mathbf{N}$. Then

$$\Theta(z) = \sum_{s=1}^{S} c_s \frac{(1-\gamma_s)z}{1-\gamma_s z}$$

and thus, using again (35),

$$\begin{aligned}\Pi(z) &= \frac{\sum_{k=0}^{\infty} F_k \sum_{n=1}^{\infty} \sum_{s=1}^{S} c_s(1-\gamma_s)\gamma_s^{n+k-1}\left(1-\alpha(z)^n\right)}{1 - D(z)\sum_{s=1}^{S} c_s \frac{(1-\gamma_s)\alpha(z)}{1-\gamma_s\alpha(z)}} \\ &= \frac{(z-U(z))\sum_{s=1}^{S} F(\gamma_s)\frac{c_s}{z-\gamma_s U(z)}}{1 - D(z)\sum_{s=1}^{S} c_s \frac{(1-\gamma_s)U(z)}{z-\gamma_s U(z)}} \\ &= \frac{(z-U(z))\sum\limits_{s=1}^{S} F(\gamma_s)c_s \prod_{r\neq s}(z-\gamma_r U(z))}{\prod\limits_{s=1}^{S}(z-\gamma_s U(z)) \; - \; D(z)U(z)\sum\limits_{s=1}^{S} c_s(1-\gamma_s)\prod_{r\neq s}(z-\gamma_r U(z))}. \end{aligned} \tag{37}$$

The S unknown constants, $F(\gamma_s)$, $s = 1, 2, \ldots, S$, can be determined from a standard argument using Rouché's theorem as follows. If we set $f(z) := \prod_{s=1}^{S}(z-\gamma_s U(z))$ and $g(z) := -\,D(z)U(z)\sum_{s=1}^{S} c_s(1-\gamma_s)\prod_{r\neq s}(z-\gamma_r U(z)) = -D(z)U(z)\Theta(z)f(z)$ then, it is easy to see that the function f has precisely S roots within the disc $|z| < 1$. Indeed, the equation $z = \gamma_s U(z)$ has a unique, real solution $r_s \in (\gamma_s, 1)$ when $\gamma_s \in (0, 1)$. On the circle $|z| = 1 - \varepsilon$ (where ε is chosen so small that the contour contains $r_1, \ldots, r_S$) $|g(z)| \leq |D(z)|\,|U(z)|\,|\Theta(z)|\,|f(z)| \leq (1-\varepsilon)^3\,|f(z)| < |f(z)|$, thus Rouché's theorem applies. Hence the denominator has precisely S roots within the circle $|z| = 1 - \varepsilon$, say $z_1, z_2, \ldots, z_S$. These must also be roots of the numerator. The equation $z = U(z)$ has precisely two roots, 1, and a real root greater than 1, when $U'(1) = \rho < 1$. Thus the factor $(z - U(z))$ cannot vanish for inside the circle $|z| = 1 - \varepsilon$. Furthermore,

$$\prod_{s=1}^{S}(z_t - \gamma_s U(z_t)) \neq 0 \quad \text{for } t = 1, 2, \ldots, S. \tag{38}$$

Indeed, if $\prod_{s=1}^{S}(z_{t_1} - \gamma_s U(z_{t_1})) = 0$ then $z_{t_1} - \gamma_{s_1} U(z_{t_1}) = 0$ for some s_1. Since z_{t_1} is a root of the denominator, $D(z_{t_1})U(z_{t_1})\sum_{s=1}^{S} c_s(1-\gamma_s)\prod_{r\neq s}(z_{t_1} - \gamma_r U(z)) = 0$ and hence, $c_{s_1}(1-\gamma_{s_1})\prod_{r\neq s_1}(z_{t_1} - \gamma_r U(z_{t_1})) = 0$. This implies in turn that $z_{t_1} - \gamma_{s_2} U(z_{t_1}) = 0$ for some $s_2 \neq s_1$. But then $z_{t_1} - \gamma_{s_1} U(z_{t_1}) = 0 = z_{t_1} - \gamma_{s_2} U(z_{t_1})$ which implies $\gamma_{s_1} = \gamma_{s_2}$ which is impossible. Thus, dividing the numerator with the left hand side of (38) we have

$$\sum_{s=1}^{S} F(\gamma_s)\frac{c_s}{z_t - \gamma_s U(z_t)} = 0, \qquad t = 1, 2, \ldots, S.$$

One of the above equations is in fact redundant and has to be replaced by the condition obtained by the requirement that $\Pi(1) = 1$ which, applying de l'Hospital's rule gives

$$\sum_{s=1}^{S} F(\gamma_s)\frac{c_s}{1-\gamma_s} = E\Theta - \frac{\lambda EG}{1-\rho}.$$

4.4 Vacation length depending on whether the production batch is complete

Here we examine a variation of the above model according to which the distribution of the vacation length depends on whether the server completed the production batch that preceded it or whether it was incomplete. In the context of the server failure model we suppose that, if $L_m = \Theta_m$ then a failure has occured and therefore the subsequent vacation period has distribution G (corresponding to full repair) whereas if $L_m < \Theta_m$ this means that the subsequent vacation period will have distribution $\widetilde{G}$. One easily sees that (25) still holds while now

$$Q_0(z) = \sum_{n=1}^{\infty} \theta_n \left(\widetilde{D}(z) \sum_{k=1}^{n-1} F_k + D(z) Q_n(z) \right). \tag{39}$$

which we can also write as

$$Q_0(z) = \left(\widetilde{D}(z) - D(z)\right) \sum_{n=1}^{\infty} \theta_n \sum_{k=1}^{n-1} F_k + D(z)\Pi(z)$$

Thus we have

$$\sum_{n=1}^{\infty} \left(\alpha(z)^n Q_0(z) + \sum_{k=0}^{n-1} F_k \left(1 - \alpha(z)^{n-k}\right) \right) \theta_n = \Pi(z).$$

$$\left(\left(\widetilde{D}(z) - D(z)\right) \sum_{n=1}^{\infty} \theta_n \sum_{k=1}^{n-1} F_k + D(z)\Pi(z) \right) \Theta(\alpha(z)) + \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} F_k \left(1 - \alpha(z)^{n-k}\right) \theta_n = \Pi(z)$$

or

$$\Pi(z) = \frac{\Theta(\alpha(z)) \left(\widetilde{D}(z) - D(z)\right) \sum_{n=1}^{\infty} \theta_n \sum_{k=1}^{n-1} F_k + \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} F_k \left(1 - \alpha(z)^{n-k}\right) \theta_n}{1 - \Theta(\alpha(z)) D(z)}.$$

In the case of geometric production batches (i.e. constant probability of failure) where θ_n is given by (33) we obtain

$$\Pi(z) = \frac{((1-\gamma)U(z)(\widetilde{D}(z) - D(z)) + z - U(z))F(\gamma)}{z - \gamma U(z) - (1-\gamma)D(z)U(z)}.$$

The unknown $F(\gamma)$ is again determined by de l'Hospital's rule and is seen to be equal to $F(\gamma) = \frac{1-\rho-(1-\gamma)\lambda EG}{1-\rho+(1-\gamma)\lambda\left(E\widetilde{G}-EG\right)}$. Of course, in a reliability context, preventive maintenance would be useless in this case and hence $\widetilde{D}(z) = 1$ and $E\widetilde{G} = 0$. The case of combination of geometric batches as well as the general approach via the Wiener-Hopf decomposition can be treated by adopting the analysis of sections 4.3 and 4.1 *mutatis mutandis*.

5 Time-stationary distribution of the number of customers in the queue and sojourn times

As we saw in the previous section (T_m, Φ_m), $m \in \mathbf{Z}$, is a Markov–renewal process and that the process $\{X_t;\, t \in \mathbf{R}\}$ is semi–regenerative with respect to this Markov–renewal process. Furthermore, it is possible to see that, under the stability condition (1), the Markov chain $\{\Phi_m; m \in \mathbf{Z}\}$ is positive recurrent.

Consider the basic epochs $\{T_m\}$ when the server leaves the queue in order to deliver the cart and a vacation period begins. Under the stability condition, it is clear that there exists a steady–state regime since this is a semi–regenerative system. (Alternatively, we could identify ordinary regeneration cycles corresponding to the epochs when the server leaves the queue *empty* to deliver the cart.) It is also possible to show that these regenerative cycles have finite mean and thus there exists a steady state random variable, say X_∞, such that $X_t \xrightarrow{d} X_\infty$ where $\xrightarrow{d}$ denotes convergence in distribution. We shall establish the following

Theorem 2. *The stationary number of customers in the system when the server uses a partial batch policy has p.g.f. given by*

$$Ez^{X_\infty} = \Pi(z) G_I^*(\lambda(1-z)) \frac{(1-\rho)B^*(\lambda(1-z))}{1-\rho B_I^*(\lambda(1-z))}, \tag{40}$$

where $\Pi(z)$ *is the p.g.f. of the number of customers present in the system at the beginning of a typical vacation. Depending on whether the production batch size distribution has bounded or unbounded support,* $\Pi(z)$ *is given by (23) or by (32).*

Proof: We will establish the theorem assuming that the production batch size distribution does not necessarily have bounded support. We begin with a version of the process which satisfies the following conditions: *(i)* The time origin coincides with the beginning of a "typical" cycle, i.e. $\cdots < T_{-2} < T_{-1} < T_0 = 0 < T_1 < T_2 < \cdots$ and *(ii)* $\Phi_0 = X_{T_0} = X_0$ is distributed according to the (jump) stationary distribution of the Markov Chain $\{\Phi_m; m \in \mathbf{Z}\}$. If we denote by λ^* the rate of the process $\{T_m\}$ we then have the following formula connecting the distribution of X_∞ to that of $\{X_t; t \in [T_0, T_1)\}$: For any bounded function $f : \mathbf{N} \to \mathbf{R}$,

$$Ef(X_\infty) = \lambda^* E \int_{T_0}^{T_1} f(X_s) ds.$$

In particular, if we take $f(x) = z^x$ (where $0 \le z \le 1$) we have the following expression for the p.g.f. of the time stationary distribution of the number of customers in the queue:

$$Ez^{X\infty} = \lambda^* E \int_{T_0}^{T_1} z^{X_s} ds. \tag{41}$$

The formulae above can be thought of as consequence of the semi-regenerative nature of the system (see [**?**]). Alternatively, if one is willing to use the language of stationary and ergodic processes these are special cases of the Palm inversion formula (see Baccelli and Brémaud [1]). The integral on the right hand side of (41) can be split into two parts,

$$I_1 := \int_{T_0}^{S_0} z^{X_s} ds; \qquad \text{and} \qquad I_2 := \int_{S_0}^{T_1} z^{X_s} ds.$$

The first term is analyzed by conditioning on the size of the production batch. On the event $\{\Theta = n\}$ it splits into a sum of n terms as follows

$$I_1 = \sum_{i=0}^{n-1} \mathbf{1}(L > i, \Theta = n) \int_{d_i}^{d_{i+1}} z^{X_s} ds.$$

Since $X_s = X_{d_i} + A(d_i, s]$ where $A(d_i, s]$ is the number of Poisson arrivals in the interval $(d_i, s]$, we can write

$$EI_1 = \sum_{n=1}^{\infty} \sum_{i=0}^{n-1} E\left[\mathbf{1}(L > i, \Theta = n) \int_{d_i}^{d_{i+1}} z^{X_{d_i} + A(d_i,s]} ds\right].$$

Note that, because of the independent increments property of the Poisson arrival process,

$$\begin{aligned}
E\left[\mathbf{1}(L > i, \Theta = n) \int_{d_i}^{d_{i+1}} z^{X_{d_i} + A(d_i,s]} ds\right] &= E\left[\mathbf{1}(L > i, \Theta = n) z^{X_{d_i}} \int_{d_i}^{d_{i+1}} e^{-\lambda(s-d_i)(1-z)} ds\right] \\
&= E\left[\mathbf{1}(L > i, \Theta = n) z^{X_{d_i}} \frac{1 - e^{-\lambda(d_{i+1}-d_i)(1-z)}}{\lambda(1-z)}\right] \\
&= \frac{1 - B^*(\lambda(1-z))}{\lambda(1-z)} E\left[\mathbf{1}(L > i, \Theta = n) z^{X_{d_i}}\right]
\end{aligned}$$

where, in the above derivation we have used the fact that $E[e^{-s(d_{i+1}-d_i)} \mid L > i, \Theta = n] = B^*(s)$ and $d_{i+1} - d_i$ is independent of X_{d_i} on $\{L > i\}$. Also, taking into account (2), (3), (4), (5), and the fact that $i < n$ we have that

$$\begin{aligned}
E\left[\mathbf{1}(L > i, \Theta = n) z^{X_{d_i}}\right] &= E\left[\mathbf{1}(L \geq i, \Theta = n) z^{X_{d_i}}\right] - E\left[\mathbf{1}(L = i, \Theta = n) z^{X_{d_i}}\right] \\
&= E\left[\mathbf{1}(L \geq i) z^{X_{d_i}}\right] P(\Theta = n) - F_i P(\Theta = n) \\
&= \left(Q_i(z) - F_i\right) \theta_n.
\end{aligned}$$

Hence

$$EI_1 = \frac{1 - B^*(\lambda(1-z))}{\lambda(1-z)} \sum_{n=1}^{\infty} \theta_n \sum_{i=0}^{n-1} \left(Q_i(z) - F_i\right),$$

and, using (7) (which as we saw holds regardless of whether the production batch size has bounded support or not) we obtain

$$\sum_{i=0}^{n-1} \left(Q_i(z) - F_i\right) = \sum_{i=1}^{n} Q_i(z) y.$$

Elementary manipulations yield

$$\sum_{i=1}^{n} Q_i(z) = Q_n(z) \frac{y^n - 1}{y - 1} + \sum_{j=1}^{n-1} F_j \frac{y^j - 1}{y - 1}$$

and thus the first term can be written as

$$EI_1 = \frac{y - z}{\lambda(1-z)(y-1)} \sum_{n=1}^{\infty} \theta_n \left(Q_n(z) \left(y^n - 1\right) + \sum_{j=1}^{n-1} F_j \left(y^j - 1\right) \right).$$

Using (14) in the above expression, we can rewrite EI_1 after some algebraic manipulations as

$$EI_1 = (D-1) \frac{y - z}{\lambda(1-z)(y-1)} \sum_{n=1}^{\infty} \theta_n \left(Q_n(z) + \sum_{j=0}^{n-1} F_j \right) = \frac{1-D}{\lambda(1-z)} \frac{y-z}{1-y} \Pi(z)$$

where, in the second equation we have used (25).

On the other hand, the expectation of I_2, is given by

$$EI_2 = E[z^{\Psi} \int_0^G z^{A(0,s]} ds] = \Pi(z) \frac{1 - D(z)}{\lambda(1 - z)}. \tag{42}$$

Thus adding the two equations above term by term we have

$$EI_1 + EI_2 = \Pi(z) \frac{1 - D(z)}{\lambda(1 - z)} \left(1 + \frac{y - z}{1 - y}\right).$$

From the above, after some elementary manipulations we obtain

$$Ez^{X_\infty} = \lambda^*(EI_1 + EI_2) = \lambda^* \Pi(z) G_I^*(\lambda(1 - z)) EG \frac{B^*(\lambda(1 - z))}{1 - \rho B_I^*(\lambda(1 - z))},$$

where the rate λ^* can be computed from the normalization requirement by setting $z = 1$ in the above relationship. Indeed,

$$\lambda^* = \frac{1 - \rho}{EG} \tag{43}$$

and this completes the proof of the theorem. ∎

Remark: The representation of the p.g.f. of the number of customers in stationarity can be interpreted as a *decomposition into three parts* of the type one should expect in view of the well known properties of M/G/1 queues with vacations (see [13] and also [8], [12], and [19]). The term $\frac{(1-\rho)B^*(\lambda(1-z))}{1-\rho B_I^*(\lambda(1-z))}$ is of course the p.g.f. the number of customers in a steady state M/G/1 queue without vacations, the term $G_I^*(\lambda(1 - z))$ is the p.g.f. of the number of Poisson arrivals during the forward recurrence time of a typical vacation, and finally $\Pi(z)$ is the p.g.f. of the number of customers present in the system at the beginning of a typical vacation. Of course, this decomposition holds because of the partial batch policy used.

Corollary 1. *In particular, when the production batch size is geometric, i.e.* $\theta_n = (1 - \gamma)\gamma^{n-1}$, $n = 1, 2, 3, \ldots,$ *the p.g.f. of the number of customers in the system in steady state is given by*

$$Ez^{X_\infty} = \frac{1}{\lambda EG} \frac{(1 - \rho)U(z)F(\gamma)(1 - D(z))}{z - \gamma U(z) - (1 - \gamma)D(z)U(z)} \tag{44}$$

where $F(\gamma)$ *is given by (36).*

Proof: Use (34) for $\Pi(z)$ in theorem 2. ∎

5.1 Sojourn time distribution

The sojourn time is obtained easily from the above formula via the distributional version of Little's law (see [2], [18], [19], and [26].) Indeed, setting $s = \lambda(1 - z)$ in (40) we obtain

$$T(s) = \Pi(1 - s/\lambda) G_I^*(s) \frac{(1 - \rho)B^*(s)}{1 - \rho B_I^*(s)}, \tag{45}$$

where, $\Pi(1-s/\lambda)$ can be computed from (23). After the necessary simplifications, taking into account that $\alpha(1-s/\lambda) = B^*(s)\frac{\lambda}{\lambda+s}$, we have

$$\Pi(1-s/\lambda) = \frac{\sum_{k=0}^{N-1} F_k \sum_{n=k+1}^{N-1} \theta_n \left(1 - B^*(s)^{n-k}\left(\frac{\lambda}{\lambda-s}\right)^{n-k}\right)}{1 - G^*(s)\sum_{n=1}^{N} \theta_n B^*(s)^n \left(\frac{\lambda}{\lambda-s}\right)^n}.$$

It should be pointed out that (45) gives the total time from the moment a customer enters the queue to the moment he enters the cart. The additional delay due to the time the customer has to wait until the cart is delivered is not included. In fact it is not possible to do this using the distributional version of Little's law, since the total sojourn time of a customer in this case depends on future arrivals as well.

In the case of the geometric batch transfer size, setting $z = 1 - s/\lambda$ in (44) and carrying out the necessary simplifications we obtain

$$T(s) = \frac{(1-\rho)G_I^*(s)B^*(s)F(\gamma)}{1 - \rho B_I^*(s) - (1-\gamma)\rho_G B^*(s)G_I^*(s)}$$

6 The "complete batch" policy

So far we have carried out the analysis assuming a partial batch policy. Alternative strategies can also be analyzed, as in [7]. In this section we sketch the analysis for what we will call the "complete batch" policy. According to this policy, each time the server returns with the cart to the system a random variable representing the production batch size is realized. The server keeps serving customers until this production batch size is completed (waiting for new arrivals if the queue empties) and as soon as the batch is completed he departs to deliver the cart thus initiating a vacation period. Upon returning to the system a new production batch is set and the whole process repeats itself. The starting point in our analysis is to realize that, with the given policy, each service phase consists of a complete batch so that $L_m = \Theta_m$. If we define

$$R_k(z) := E\left[z^{X_{d_k}} | \Theta \geq k\right] = E\left[z^{X_{d_k}} | \Theta = k\right],$$

the system dynamics in this case are described by

$$R_k(z) = \left(\frac{R_{k-1} - H_{k-1}}{z} + H_{k-1}\right) U(z), \tag{46}$$

where

$$H_k := R_k(0).$$

This in turn, upon iteration, gives

$$\alpha^n R_0(z) = R_n(z) + (1-z)\left(H_0\alpha^n + H_1\alpha^{n-1} + \cdots + H_{n-1}\alpha\right). \tag{47}$$

Since we still have $\Pi(z) = \sum_{n=1}^{\infty} \theta_n R_n(z)$ and

$$R_0(z) = \Pi(z)D(z), \tag{48}$$

from the above recursion we obtain

$$\Pi(z) = (z-1)\frac{\sum_{n=1}^{\infty}\theta_n\left(\sum_{j=0}^{n-1}H_j\alpha^{n-j}\right)}{1-D(z)\Theta(\alpha)}. \tag{49}$$

When the batch size distribution has finite support, say the set $\{1,2,\ldots,N\}$, then the denominator is the same as in the corresponding expression for the partial batch policy in section 3.2. Thus the N unknown constants, $H_0, H_1, \ldots, H_{N-1}$, on which $\Pi(z)$ depends in this case is obtained by Rouché's theorem, as before. When the batch size distribution has infinite support, in general one has to resort to Wiener-Hopf factorization techniques in order to determine $\Pi(z)$. Of course one can analyze easily the case where the batch size distribution is a combination of geometric factors as in section 4.3. Here we will restrict ourselves to the analysis of the case of geometric batches, i.e. $\theta_n = (1-\gamma)\gamma^n$, $n = 1, 2, \ldots$.. Then, arguing as in section 4.2 we see that

$$\Pi(z) = \frac{(z-1)U(z)(1-\gamma)H(\gamma)}{z-\gamma U(z)-(1-\gamma)D(z)U(z)}$$

where

$$H(\gamma) = \frac{1-\rho}{1-\gamma} - \lambda EG$$

as can be seen from an argument using the fact that $\Pi(1) = 1$ and de l'Hospital's rule.

Finally we determine the stationary distribution of the number of customers in the system (excluding the cart) under the complete batch policy. We indicate the differences in this case, illustrating the case of geometric production batches. With the notation of the section 5 we have

$$\begin{aligned}
EI_1 &= \sum_{n=1}^{\infty}\theta_n\sum_{i=0}^{n-1}E\left[\int_{d_i}^{d_{i+1}} z^{X_s}ds\big|\,\Theta=n\right] \\
&= \sum_{n=1}^{\infty}\theta_n\sum_{i=0}^{n-1}\left(\lambda^{-1}\left(1+z\tfrac{1-U(z)}{1-z}\right)P(X_{d_i}=0|\,\Theta=n)+E\left[z^{X_{d_i}}\mathbf{1}(X_{d_i}>0)\big|\,\Theta=n\right]\lambda^{-1}\tfrac{1-U(z)}{1-z}\right) \\
&= \sum_{n=1}^{\infty}\theta_n\sum_{i=0}^{n-1}\left(\lambda^{-1}\left(1+z\frac{1-U(z)}{1-z}-\frac{1-U(z)}{1-z}\right)H_i+R_i(z)\lambda^{-1}\frac{1-U(z)}{1-z}\right) \\
&= \sum_{n=1}^{\infty}\theta_n\sum_{i=0}^{n-1}\left(\lambda^{-1}U(z)H_i+R_i(z)\lambda^{-1}\frac{1-U(z)}{1-z}\right)
\end{aligned}$$

Recall that, by definition $R_i(z) := E[z^{X_{d_i}}|\Theta=i]$ and $H_i = P(X_{d_i}=0 \mid \Theta=i)$. Rewrite (46) as

$$yR_i = R_{i-1} + (z-1)H_{i-1}$$

and obtain

$$\sum_{i=0}^{n-1}R_i(z) = \frac{y}{1-y}\left(R_n(z)-R_0(z)\right) - \frac{z-1}{1-y}\sum_{i=0}^{n-1}H_i.$$

Thus we have

$$\begin{aligned}
EI_1 &= \lambda^{-1}\left(U(z)+\frac{1-U(z)}{1-y}\right)\sum_{n=1}^{\infty}\theta_n\sum_{i=0}^{n-1}H_i + \lambda^{-1}\frac{1-U(z)}{1-z}\sum_{n=1}^{\infty}\theta_n\frac{y}{1-y}\left(R_n(z)-R_0(z)\right)\\
&= \lambda^{-1}U\frac{1-z}{U-z}C' + \lambda^{-1}\frac{1-U(z)}{1-z}\frac{z}{U-z}\sum_{n=1}^{\infty}\theta_n\left(R_n(z)-R_0(z)\right)\\
&= \lambda^{-1}U\frac{1-z}{U-z}C' + \lambda^{-1}\frac{1-U}{1-z}\frac{z}{U-z}\left(1-D\right)\Pi(z)
\end{aligned}$$

where we have set $C' := \sum_{n=1}^{\infty}\theta_n\sum_{i=0}^{n-1}H_i$ and taken into account that $\Pi(z)=\sum_{n=1}^{\infty}\theta_n R_n(z)$ and $R_0(z)=\Pi(z)D(z)$. On the other hand (42) still holds and thus

$$\lambda^*(EI_1+EI_2) = \frac{\lambda^*}{\lambda}U\frac{1-z}{U-z}C' + \frac{\lambda^*}{\lambda}\Pi(z)\frac{1-D}{U-z}. \tag{50}$$

The value of C' can be determined from (49) using the observation that $\Pi(1)=1$ and de l' Hospital's rule:

$$C'=(1-\rho)E\Theta-\lambda EG.$$

Since, as in section 5 $Ez^{X_\infty}=\lambda^*(EI_1+EI_2)$, we can determine λ^* by setting $z=1$ in (50) and using once more de l' Hospital's rule. Thus we obtain

$$\frac{\lambda}{\lambda^*}=E\Theta. \tag{51}$$

Putting things together we obtain

$$\begin{aligned}
Ez^{X_\infty} &= (1-\rho)U\frac{1-z}{U-z}\left(1-\frac{\lambda EG}{1-\rho}\right)+\frac{1}{E\Theta}\Pi(z)\frac{1-D}{U-z}\\
&= (1-\rho)U\frac{1-z}{U-z}\left(1-\frac{\lambda EG}{E\Theta(1-\rho)}\right)+\frac{\lambda EG}{E\Theta(1-\rho)}\Pi(z)\frac{1-D}{\lambda EG(1-z)}(1-\rho)\frac{1-z}{U-z}
\end{aligned}$$

or

$$Ez^{X_\infty} = (1-p)(1-\rho)U\frac{1-z}{U-z}+p\Pi(z)G_I^*(\lambda(1-z))(1-\rho)\frac{1-z}{U-z} \tag{52}$$

where $p=\frac{\lambda EG}{E\Theta(1-\rho)}$, $\Pi(z)$ as given in (49) is the p.g.f. of the number of customers left behind at the end of the typical service phase, $G_I^*(\lambda(1-z))$ the p.g.f. of the number of Poisson arrivals during the residual service time of a vacation period and finally $(1-\rho)U\frac{1-z}{U-z}$ is the p.g.f. of the stationary number of customers in the corresponding M/G/1 system without vacations (in that case the size of the production batch becomes irrelevant). Note that the second term on the right hand side of (52) includes the term $(1-\rho)\frac{1-z}{U-z}$ which is the generating function of the number of Poisson arrivals during the *waiting time* in the corresponding M/G/1 system without vacations.

7 Bulk arrivals

There are no significant changes in the above analysis if we assume that customers arrive not singly but in batches. Of course arrival epochs are still Poisson (λ) and the arriving batches are an i.i.d. sequence of random variables $\{\beta_n\}$, independent of the Poisson arrival process, with common distribution $P(\beta = k) = b_k$, $k = 1, 2, 3, \ldots$. The corresponding p.g.f. will be denoted by $b(z) := \sum_{k=1}^{\infty} b_k z^k$ and the mean batch size by $m_b = \sum_{k=1}^{\infty} k b_k$. In order not to obscure the main features of the problem we will introduce here the simplifying assumption that the production batch size sequence $\{\Theta_m\}$ is deterministic and equal to the cart capacity N. It is easy to see that, in this case, the stability condition becomes

$$N > \frac{\lambda EG}{1 - \lambda E\beta E\sigma}.$$

We can analyze this system in precisely the same way as the single customer arrival case. Indeed, equations (8) and (13) hold unchanged, if we substitute for $U(z)$ and $D(z)$ the p.g.f.'s

$$U_b(z) := B^*(\lambda(1 - b(z))), \qquad D_b(z) := G^*(\lambda(1 - b(z))).$$

Then $\Pi_b(z)$, the p.g.f. of the number of customers left behind in the queue at a typical vacation start is given by the relationship

$$\Pi_b(z) = \frac{\sum_{k=0}^{N-1} F_{b,k} z^k \left(z^{N-k} - U_b(z)^{N-k}\right)}{z^N - D_b(z) U_b(z)^N}. \tag{53}$$

where, again, the N constants $F_{b,k}$, $k = 0, 1, 2, \ldots, N-1$ are obtained by Rouché's theorem.

Let us again denote by $\{T_n\}$ the basic epochs when the server leaves the queue in order to deliver the cart and a vacation period begins. $\{X_t; t \in \mathbf{R}\}$ denotes the number of customers in the system process, and λ^* the rate of the point process $\{T_n\}$. As usual, X_∞ denotes a random variable with the steady state distribution of the process $\{X_t\}$. An analysis entirely analogous to that of section 5 gives the following expression for the p.g.f. of the stationary number of customers in the system

$$Ez^{X_\infty} = \Pi_b(z) G_I^*(\lambda(1 - b(z))) \frac{(1 - \rho m_b) B^*(\lambda(1 - z))}{1 - \rho B_I^*(\lambda(1 - z))},$$

which assumes again the form of a three way decomposition. The term $\frac{(1-\rho m_b)B^*(\lambda(1-z))}{1-\rho B_I^*(\lambda(1-z))}$ is the p.g.f. of the time-stationary number of customers in an M/G/1 queue with bulk arrivals and without vacations, the term $G_I^*(\lambda(1 - b(z)))$ is the p.g.f. of the total number of arrivals during the forward recurrence time of a typical vacation; and finally $\Pi_b(z)$ is the p.g.f. of the number of customers present in the system at the beginning of a typical vacation.

8 The contents of the cart when it is delivered

When the partial batch policy is used the contents of the cart when it is delivered or "actual production batch size" is a random variable stochastically smaller than the production batch size. Its distribution in stationarity can be determined as follows. Let $\Upsilon(z, w) := \sum_{k=0}^{\infty} Q_k(z) w^k$ and $F(w) := \sum_{k=0}^{\infty} F_k w^k = \Upsilon(0, w)$. Then, from (7) it follows that

$$\Upsilon(z, w) - Q_0(z) = \sum_{k=0}^{\infty} Q_{k+1}(z) w^{k+1} = w\alpha \left(\sum_{k=0}^{\infty} Q_k(z) w^k - \sum_{k=0}^{\infty} F_k w^k \right)$$

or, recalling definition (6), after some elementary manipulations,

$$\Upsilon(z,w) = \frac{zQ_0(z) - F(w)wU(z)}{z - wU(z)}. \tag{54}$$

The above expression involves the unknown function $F(w)$ which can be determined as follows. Suppose that $|w| < 1$. The equation $z - wU(z) = 0$ has for each fixed value of w in the unit disk a unique solution $\zeta(w)$. (This can be seen by an application of Rouche's theorem, see Takács [24]). In fact $\zeta(w) = \sum_{n=1}^{\infty} \frac{1}{n!} w^n \ (d/dt)^{n-1} U^n(t)\big|_{t=0}$ according to the Bürman-Lagrange inversion formula. The numerator must also vanish when $z = \zeta(w)$ and thus $\zeta(w)Q_0(\zeta(w)) = F(w)wU(\zeta(w))$ or $F(w) = Q_0(\zeta(w))$. Taking into account (12) as well we have

$$F(w) = \Pi(\zeta(w))D(\zeta(w)). \tag{55}$$

As we saw in section 4.1 $D(\zeta(w)) = \sum_{n=0}^{\infty} \kappa_n w^n$ with $\kappa_0 = D(0)$ and $\kappa_n = \frac{1}{n!} \ (d/dt)^{n-1} D'(t) U^n(t)\big|_{t=0}$, $n = 1, 2, \ldots$. Using (32) and the fact that $\zeta(w)/U(\zeta(w)) = w$ we thus obtain the generating function for the sequence $\{F_n\}$ as follows

$$F(w) = \sum_{n=0}^{\infty} \kappa_n w^n \exp\left(\sum_{r=1}^{\infty} (w^r - 1) \sum_{n=1}^{\infty} \frac{1}{n} \sum_{l=1}^{\infty} \kappa_{l+r}^{*n} \theta_l^{*n} \right) \tag{56}$$

Once the sequence $\{F_n\}$ has been determined, the number of customers in the actual production batch size is obtained by first conditioning on the production batch size as follows. We have

$$E\left[w^L \,|\Theta = n\right] = \sum_{k=0}^{n-1} w^k F_k + w^n \left(1 - \sum_{k=0}^{n-1} F_k\right) = w^n + \sum_{k=0}^{n-1} F_k \left(w^k - w^n\right)$$

and thus

$$Ew^L = \Theta(w) + \sum_{k=0}^{\infty} w^k F_k \sum_{n=k+1}^{\infty} \theta_n - \sum_{k=0}^{\infty} F_k \sum_{n=k+1}^{\infty} w^n \theta_n. \tag{57}$$

Things of course become simpler when the production batch size is geometric, as in section 4.2. Then, setting $K(w) := D(\zeta(w)) = \sum_{n=0}^{\infty} \kappa_n w^n$ (55) becomes

$$F(w) = K(w) \frac{(\zeta(w) - U(\zeta(w)))F(\gamma)}{\zeta(w) - \gamma U(\zeta(w)) - (1-\gamma)K(w)U(\zeta(w))} = \frac{F(\gamma)K(w)(1-w)}{-w + \gamma + (1-\gamma)K(w)}$$

with $F(\gamma)$ given by (36). The mean of the probability distribution $\{\kappa_n\}$ is given by $K'(0) = D'(0)\zeta'(0)$ and of course $\zeta'(0) = U(0) = 1$, thus $K'(0) = \lambda EG$. Hence, if we define the distribution function

$$K_I(w) := \frac{1}{K'(0)} \frac{1 - K(w)}{1 - w}$$

we have

$$F(w) = \frac{F(\gamma)K(w)}{1 - (1-\gamma)\lambda EG K_I(w)}. \tag{58}$$

When the production batch size is geometric (57) simplifies into the following expression

$$Ew^L = \frac{1 - \gamma + (1-w)F(\gamma w)}{1 - \gamma w}$$

which, together with (58) gives the generating function of the number of items delivered.

9 The contents of the cart in steady state

This section refers to the queue and cart model and we are interested in the cart as long as it is ”next to the server”, receiving customers, so we will suppose that the number in the cart process becomes equal to zero as soon as the server takes the cart to deliver it (see figure 2). *In order not to obscure the main aspects of the problem here and in the sequel we assume that the production batch size is deterministic and equal to the cart's capacity,* N. Random production batches can be analyzed in a similar fashion as in the previous sections. In the first subsection we examine the marginal distribution of the cart contents under the partial batch policy while in the second the joint distribution of the number of customers in the queue and the cart.

9.1 The marginal distribution under the partial batch policy

If we denote by $Y \equiv Y_\infty$ a random variable with the steady state distribution of the number of customers in the cart and by $C(w) := Ew^{Y_\infty}$ the corresponding p.g.f., then the cycle formula gives

$$\begin{aligned}
C(w) &= Ew^{Y_\infty} = \frac{E\int_{T_0}^{T_1} w^{Y_s} ds}{E(T_1 - T_0)} = \frac{EG + E\sigma E \sum_{k=1}^{L} w^{k-1}}{EG + ELE\sigma} \\
&= \frac{EG + E\sigma E\left[\frac{1-w^L}{1-w}\right]}{EG + ELE\sigma} \\
&= \frac{1 - w^N - \sum_{k=0}^{N-1} F_k \left(w^k - w^N\right)}{1-w} \, \frac{E\sigma}{EG + ELE\sigma} \; + \; \frac{EG}{EG + ELE\sigma} \\
&= \rho \frac{1}{EL} \sum_{k=0}^{N-1} w^k \left(1 - \sum_{i=0}^{k} F_i\right) \; + \; (1-\rho),
\end{aligned}$$

where, in the above string of equalities we have taken into account the fact that $Q_N(1) = 1 - \sum_{k=0}^{N-1} F_k$ and $EL = \frac{\lambda EG}{1-\lambda E\sigma}$. (This last equation is (24).) Thus the steady–state number of customers in the cart is

$$\begin{aligned}
P(Y=0) &= \frac{\rho}{EL}(1 - F_0) + (1-\rho), \\
P(Y=k) &= \frac{\rho}{EL}\left(1 - \sum_{i=0}^{k} F_i\right), \qquad k = 1, 2, \ldots, N-1.
\end{aligned}$$

The expected number of customers in the steady state is then equal to

$$EY = \sum_{k=1}^{N-1} k \frac{\rho}{EL}\left(1 - \sum_{i=0}^{k} F_i\right) = \rho \frac{E[L(L-1)]}{2EL}.$$

9.2 Joint distribution of the number of customers in the queue and the cart

Arguing as above, we can obtain with a little more effort the joint distribution of the number of customers in the queue and the cart, $V(z,w) := Ez^{X_\infty} w^{Y_\infty}$ by using the same method as in the analysis of §5.

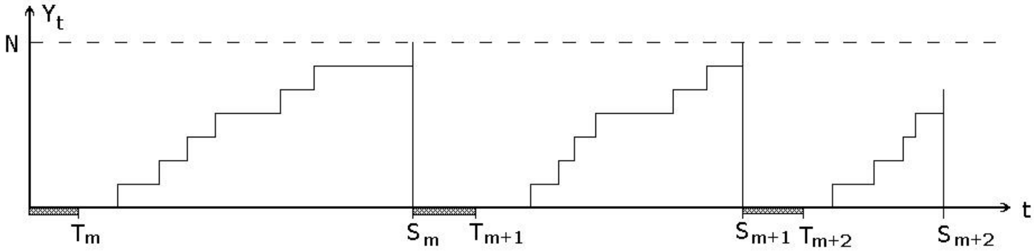

Figure 2: Sample path of cart contents.

With the notation of §5 we have

$$V(z,w) = \lambda^* \left(E \int_{T_0}^{S_0} z^{X_t} w^{Y_t} dt + E \int_{S_0}^{T_1} z^{X_t} w^{Y_t} dt \right). \tag{59}$$

Taking into account that at the beginning of a cycle, when the server returns with the cart to the queue and starts serving, $Y(T_0) = 0$ (i.e. the cart is empty) we have

$$E \int_{T_0}^{S_1} z^{X_t} w^{Y_t} dt = E \sum_{k=0}^{N-1} \mathbf{1}(L > k) w^k \int_{d_k}^{d_{k+1}} z^{X_t} dt = \frac{1 - U(z)}{\lambda(1-z)} \sum_{k=0}^{N-1} w^k Q_k(z).$$

The integral over the vacation phase, where $Y(t) = 0$, is

$$E \int_{S_0}^{T_1} z^{X_t} w^{Y_t} dt = \Pi(z) \frac{1 - G^*(\lambda(1-z))}{\lambda(1-z)},$$

where, as in §5, $\Pi(z) = Q_N(z) + \sum_{j=0}^{N-1} F_j$. Using also the recursion $\alpha(Q_k - F_k) = Q_{k+1}$, we have

$$Q_k = F_k + yF_{k+1} + y^2 F_{k+2} + \cdots + y^{N-k-1} F_{N-1} + y^{N-k} Q_N,$$

and hence

$$\begin{aligned} \sum_{k=0}^{N-1} w^k Q_k &= \sum_{k=0}^{N-1} w^k \left\{ \left(\sum_{l=k}^{N-1} F_l y^{l-k} \right) + y^{N-k} Q_N \right\} \\ &= \Pi(z) \frac{1 - (\alpha w)^N}{1 - \alpha w} + \sum_{l=0}^{N-1} \frac{(\alpha w)^N - (\alpha w)^{l+1}}{1 - \alpha w} F_l, \end{aligned}$$

where we have used the fact that $y = \alpha^{-1}$. Upon substitution in (59), taking into account (43), we obtain after some simplifications the following expression for the joint p.g.f. of the queue and cart contents:

$$\begin{aligned} V(z,w) &= (1-\rho)\Pi(z) G_I^*(\lambda(1-z)) \\ &\quad + (1-\rho) \frac{E\sigma}{EG} \frac{B_I^*(\lambda(1-z))}{1 - \alpha w} \left(\Pi(z) \left(1 - (\alpha w)^N\right) + (\alpha w)^N \sum_{l=0}^{N-1} F_l \; - \; \alpha w \sum_{l=0}^{N-1} F_l (\alpha w)^l \right). \end{aligned}$$

10 Appendix

Here we show that equation (17) has N roots within the unit disk. Variations of this equation abound in the bulk service literature. (See for instance Chaudhry and Templeton [4] and also Coffman and Gilbert [7].) However in these treatments it is (either explicitly or tacitly) assumed that the service and vacation distributions are light-tailed, i.e. that the corresponding moment generating functions exist in an open interval containing the origin. The argument becomes more involved if we assume only the natural conditions for the existence of a stationary version of the process i.e. the finiteness of *first* moments plus the stability condition. Here we shall take this, more general, approach. We begin with the following theorem established in Boudreau, Griffin, and Kac [3].

Theorem 3. *Suppose that* $\varphi(z) := \sum_{n=0}^{\infty} f_n z^n$ *is the p.g.f. of* f_n, $n = 0, 1, 2, \ldots$, *a non-degenerate probability distribution on the non-negative integers with finite mean* $\mu := \sum_{n=0}^{\infty} n f_n$ *and* N *is a natural number. If the condition*

$$N > \mu \tag{60}$$

holds, then the equation

$$z^N - \varphi(z) = 0 \tag{61}$$

has N *roots within the unit disk* $\{z \in \mathbf{C} : |z| \leq 1\}$. $z = 1$ *is a single root of (61) while the remaining* $N - 1$ *roots have modulus strictly smaller than 1.*

In our case, $\varphi(z) = D(z) \sum_{n=1}^{N} \theta_n z^{N-n} U(z)^n$ and $\mu = \varphi'(1) = EG - (1-\rho)E\Theta + N$, thus (60) is equivalent to the stability condition for the system (1).

∎

10.1 Determination of the constants

We give in the sequel an explicit procedure for the computation of the N constants, $F_0, \ldots, F_{N-1}$ in the case of the partial batch policy with finite cart capacity. These constants can be obtained from the identity (19) as follows. If we denote by $S_k := S_k(y_1, y_2, \ldots, y_{N-1})$, $k = 1, 2, \ldots, N-1$ the symmetric polynomials in $N - 1$ variables,

$$\begin{aligned}
S_1 &= \sum_i y_i, \\
S_2 &= \sum_{i<j} y_i y_j, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
S_k &= \sum_{i_1 < i_2 < \cdots < i_k} y_{i_1} y_{i_2} \cdots y_{i_k}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
S_{N-1} &= y_1 y_2 \cdots y_{N-1}.
\end{aligned}$$

Then

$$(y-1)\prod_{i=1}^{N-1}(y-y_i) = y^N - y^{N-1}(1+S_1) + y^{N-2}(S_1+S_2) - y^{N-3}(S_2+S_3) + \cdots$$
$$+(-1)^{N-1}y(S_{N-2}+S_{N-1}) + (-1)^N S_{N-1}$$

On the other hand,

$$\begin{aligned} P(y) &= y^N\left(\sum_{k=0}^{N-1}F_k\sum_{j=k+1}^{N}\theta_j\right) - y^{N-1}\left(\sum_{k=0}^{N-1}F_k\theta_{k+1}\right) - \cdots - y^{N-i}\left(\sum_{k=0}^{N-i}F_k\theta_{k+i}\right) - \cdots \\ &\quad -y^2\left(F_0\theta_{N-2}+F_1\theta_{N-1}+F_2\theta_{N-2}\right) - y\left(F_0\theta_{N-1}+F_1\theta_N\right) - F_0\theta_N. \end{aligned}$$

We thus can obtain the constants F_k from the triangular linear system

$$\begin{aligned} \theta_N F_0 &= C(-1)^N S_{N-1} \\ F_0\theta_{N-1}+F_1\theta_N &= C(-1)^N\left(S_{N-1}+S_{N-2}\right) \\ F_0\theta_{N-2}+F_1\theta_{N-1}+F_2\theta_{N-2} &= C(-1)^N\left(S_{N-2}+S_{N-3}\right) \\ &\vdots \\ F_0\theta_{N-i}+F_1\theta_{N-i-1}+\cdots+F_i\theta_{N-i} &= C(-1)^N\left(S_{N-i}+S_{N-i-1}\right) \\ &\vdots \\ F_0\theta_1+F_1\theta_2+\cdots+F_{N-1}\theta_N &= C(-1)^N\left(S_2+S_1\right) \end{aligned}$$

(One additional equation, namely $F_0\left(\theta_1+\cdots+\theta_N\right)+\cdots+F_k\left(\theta_{k+1}+\cdots+\theta_N\right)+\cdots+F_{N-1}\theta_N = C(-1)^{N-1}S_1$, which is obtained by equating the coefficients of y^N is redundant and has been omitted.)

References

[1] F. Baccelli and P. Brémaud (1994). *Elements of Queueing.* Springer–Verlag, New York.

[2] D. Bertsimas and D. Nakazato (1995). The distributional Little's Law and its applications. *Operations Research* **43**, 298–310.

[3] P.E. Boudreau, J.S. Griffin, Jr., and M. Kac (1962). An elementary queueing problem. *American Mathematical Monthly* **69**, 8, 713-724.

[4] M.L. Chaudhry and J.G.C. Templeton (1983). *A First Course in Bulk Queues.* John Wiley & Sons, New York.

[5] J.W. Cohen (1969). *The Single Server Queue.* North-Holland.

[6] E.T. Copson (1935). *An Introduction to the Theory of Functions of a Complex Variable.* Oxford University Press.

[7] E.G. Coffman and E.N. Gilbert (1992). Service with a queue and a cart. *Management Science* **38**, 867–883.

[8] B.T. Doshi (1986). Queueing system with vacations**.** *Queueing Systems* **1**, 29-66.

[9] A. Federgruen and L. Green (1986). Queueing systems with service interruptions, *Operations Research,* **34,** 5, 752-768.

[10] F.G. Foster and K.M. Nyunt (1961). Queues with batch departures I. *The Annals of Mathematical Statistics*, **32**, 4, 1324-1332.

[11] F.G. Foster and A.G.A.D. Perrera (1964). Queues with batch departures II. *The Annals of Mathematical Statistics*, **35,** 3, 1147-1156.

[12] S.W. Fuhrmann (1984). A note on the M/G/1 queue with server vacations. *Operations Research* **31**, 1368-1373.

[13] S.W. Fuhrmann and R.B. Cooper (1985). Stochastic decompositions in the M/G/1 queue with generalized vacations, *Operations Research,* **33,** 1117-1129.

[14] R. Haji and G.F. Newell (1971). A relation between stationary queue and waiting–time distributions. *Journal of Applied Probability* **8**, 617–620.

[15] J. Keilson (1962). The general bulk queue as a hilbert problem. *Journal of the Royal Statistical Society, Series B,* **24,** 2, 344-358.

[16] J. Keilson (1964). Corrigenda to the above paper. *Journal of the Royal Statistical Society, Series B,* **26,** 3, 487.

[17] J. Keilson and L.D. Servi (1986). Oscillating random walk models for GI/G/1 vacation systems with Bernoulli schedules, *Journal of Applied Probability,* **23**, 790–802.

[18] J. Keilson and L.D. Servi (1988). A distributional form of Little's Law. *Operations Research Letters* **7**, 223–227.

[19] J. Keilson and L.D. Servi (1990). The distributional form of Little's Law and the Fuhrmann–Cooper decomposition. *Operations Research Letters* **9**, 223-227.

[20] J.H.B. Kemperman (1961). *The Passage Problem for a Stationary Markov Chain.* The University of Chicago Press.

[21] T.J. Ott (1987). The single-server queue with independent GI/G and M/G input streams, *Advances in Applied Probability,* **19**, 266-286.

[22] N.U. Prabhu (1965). *Queues and Inventories: A Study of Their Basic Stochastic Processes.* John Wiley.

[23] B. Sengupta (1990). A queue with service interruptions in an alternating random environment. *Operations Research* **38**, 2, 308–318.

[24] L. Takács (1962). *Introduction to the theory of queues,* Oxford University Press.

[25] T. Takine and B. Sengupta (1997). A single server queue with service interruptions, *Queueing Systems*, **26**, 285-300.

[26] Zazanis, M.A. (1998). Functional Versions of Little's Law, *Stochastic Processes and their Applications,* **74,** 2, 195-201.