Queueing Systems 26 (1997) 151–167

Formulas and representations for cyclic Markovian networks via Palm calculus

Michael A. Zazanis *

Department of MIE, University of Massachusetts, Amherst, MA 01003, USA
E-mail: zazanis@ecs.umass.edu

Received 6 June 1996; revised 6 February 1997

We present an extension of the arrival theorem for the output process from a node in closed Markovian networks which we use to obtain simple representations and explicit expressions for the throughput, the distribution of the cycle time, and the joint distribution of interoutput times from a node in single class closed networks with exponential servers. Our approach uses tools from Palm calculus to obtain a recursion on the number of customers in the system. The analysis relies on a non-overtake condition and thus many of the results obtained here apply only to cyclic, single server networks. One of the surprising conclusions of our analysis is that the interoutput times that comprise the cycle time of a customer are (finitely) exchangeable, i.e., that their joint distribution is invariant under permutations.

Keywords: closed queueing networks, cycle time distribution, Palm probabilities, stationary point processes

1. Introduction and brief description of the main results

Consider a closed, single class, queueing network in which c customers visit M nodes in a cyclic fashion. The nodes are single server stations with unlimited buffer space and exponential service times with rates μ_i, $i = 1, 2, \dots, M$. In this paper we will present an analysis of such networks using tools of Palm calculus and reversibility arguments to obtain novel representations for the throughput, cycle time distribution, and joint distribution of consecutive interoutput times from a node. These representations provide new insights into the structure of these relatively simple queueing processes and lead to closed form expressions for various performance criteria. The significance of the representations and formulas obtained here is not diminished by the availability of efficient computational algorithms such as Mean Value Analysis [23] and the convolution algorithm [5]. Rather, as it will become apparent from our analysis, our results extend the applicability of these algorithms to new performance criteria.

* Research supported in part by NSF grant EEC-9527519.

It has been known since the mid sixties [12] that the equilibrium distribution of such networks admits a *product form solution* of the form

$$\frac{1}{G(c)}\mu_1^{-n_1}\mu_2^{-n_2}\cdots\mu_M^{-n_M}, \quad n_i \in \mathbf{Z}_+,\ i = 1, \ldots, M,$$

where $n_1 + \cdots + n_M = c$, and $G(c)$ is a normalization constant (the *partition function*) given by

$$G(c) := \sum_{n_1+\cdots+n_M=c} \mu_1^{-n_1}\mu_2^{-n_2}\cdots\mu_M^{-n_M}.$$

The development of efficient algorithms for computing the partition function (the convolution algorithm [5] and Mean Value Analysis [23]) signaled the beginning of the use of these models in a host of applications, most notably in performance evaluation of computer systems, communication networks, and manufacturing systems.

The representations obtained in this paper are based on the arrival theorem together with tools of Palm calculus and involve the sum of processing times as follows. Let Y_i, $i = 1, 2, \ldots, M$, be independent, exponential random variables with rates μ_i. Denote by $Y = \sum_{i=1}^{M} Y_i$. Then the partition function of a cyclic Gordon–Newell network with c customers is given by

$$G(c) = \frac{1}{c!} EY^c$$

and its throughput by

$$c\,\frac{E[Y^{c-1}]}{E[Y^c]}.$$

Similarly, consider the cycle time of an arbitrary customer in such a network with c customers, i.e., the time that elapses in steady state from the moment a tagged customer joins the queue in station 1 until he completes service at station M. It will be shown that its Laplace transform is given by

$$\frac{E[Y^{c-1}\mathrm{e}^{-sY}]}{E[Y^{c-1}]},$$

with corresponding distribution function

$$F_c(x) = \frac{E[Y^{c-1}\mathbf{1}(Y \leqslant x)]}{E[Y^{c-1}]}.$$

Analogous representations in terms of Y, are obtained for the joint distribution of the c consecutive interoutput times that comprise the cycle time of a customer in such a network with population size c; denoting these consecutive interdeparture times by $\tau_0, \tau_1, \ldots, \tau_{c-1}$, the Palm probability of the event $\{\tau_0 > x_0, \tau_1 > x_1, \ldots, \tau_{c-1} > x_{c-1}\}$, is equal to

$$\frac{E(Y - x_0 - x_1 - \cdots - x_{c-1})_+^{c-1}}{EY^{c-1}},$$

where $(y)_+$ is the positive part of a real number y. From the above expression we reach the surprising conclusion that (under the Palm probability measure with respect to the output process from node M) the interoutput times that comprise the cycle time of a customer in a network in equilibrium are exchangeable random variables.

While the representations for the partition function and the throughput can be readily extended from cyclic Gordon–Newell networks to networks with arbitrary markovian routing, the same is not true for the cycle and interoutput time representations. This is a consequence of the fact that the later depend crucially on the "non-overtake" property enjoyed by cyclic single server networks.

The representations presented in this paper and the results regarding the interoutput times are new. On the other hand, the closed form expressions for the partition function and some of the results on cycle times have already been derived by other means. Koenigsberg [16–18] was the first to study extensively cyclic networks and to obtain closed form expressions for the partition function, though not always in minimal form. Harrison [14] was the first to provide a closed form expression for the partition function for single node Gordon–Newell networks with markovian routing. Later Gordon [11] provided a simpler approach based on generating functions as well as extension to multi-server nodes (at the expense of a considerable increase in the complexity of the expressions obtained). Gerasimov [9,10] has obtained similar results and has provided extensions to multiclass product form networks using techniques from complex analysis.

The main results on cycle times were obtained by Daduna and Schassberger [24,25] who in a series of papers showed that, in single node cyclic networks, the joint Laplace transform of the consecutive sojourn times of a tagged customer through the nodes of the network satisfy the same recursion as the partition function of a network with one less customer (see also the analysis of a two-node Markovian network in [6]). This surprising result provided a convolution algorithm similar to "Buzen's table" [5] for the joint Laplace transform of sojourn times. Boxma, Kelly, and Könheim [2] provided an insightful probabilistic explanation for this result using the arrival theorem and a time reversal argument. Finally, Harrison [13,14] combined the explicit expression for the partition function with the above results to provide an explicit expression for the cycle time of a customer in a cyclic Gordon–Newell network. We also mention the integral representation and related asymptotic expansions for the partition function of McKenna and Mitra [19,21,22].

2. Implications of the arrival theorem for the departure process from a node

In this section we study the implications of the arrival theorem for the structure of the departure process from a node in a closed, cyclic network with single, exponential servers. We will show that, in a network with c customers, the Palm distribution of the $c-1$ departure times preceding and following the departure of a "tagged" customer at time $t=0$ is the same as the stationary distribution of the $c-1$ departures before time 0 and after 0 in a stationary network with $c-1$ customers. Based on this fact

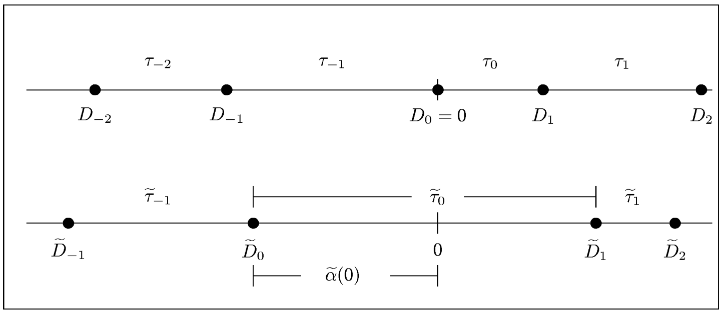

Figure 1.

(theorem 2) and the Palm inversion formula, a recursion formula leading to a simple representation for the cycle time distribution and to explicit formulas will be obtained.

Consider a cyclic closed queueing network with M single server exponential nodes (with service rates $\mu_1, \ldots, \mu_M$) and c customers and denote the corresponding continuous time Markov process by $\mathbf{X}(t) := (X_1(t), \ldots, X_M(t))$, where $X_i(t)$ designates the number of customers in node i at time t. In particular, the paths of $\mathbf{X}$ are assumed *right-continuous* P-a.s. Let $\mathbf{Z}_+$ designate the nonnegative integers. The state space of this process is the simplex

$$\mathcal{S}(M,c) := \left\{ \mathbf{n} = (n_1, n_2, \ldots, n_M)\colon\ n_i \in \mathbf{Z}_+,\ \sum_{i=1}^{M} n_i = c \right\}.$$

Let $\{D_n\}_{n\in\mathbf{Z}}$ denote the *output process* from a given node of the network (say, node M), i.e., the point process of successive departure epochs (service completions) from that node. We will designate the interoutput times by $\tau_n := D_{n+1} - D_n$. When necessary, the number of customers in the network will appear as a second index in these processes.

Consider now the family of Markov processes $\{\mathbf{X}_c\}_{c=1,2,\ldots}$ defined on the probability space $(\Omega, \mathcal{F}, P)$ and assume that all these processes are *stationary* and *independent* under P. We will designate by P_c^0 the Palm transformation of the probability measure P with respect to the point process $\{D_{n,c}\}$. We follow the standard convention in numbering the points of $\{D_{n,c}\}$, i.e., $D_{0,c} \leqslant 0 < D_{1,c}$ P-a.s. (see figure 1). (As it will become apparent, the joint statistics of the family of processes $\mathbf{X}_c$ will not play any role in our analysis. We define these processes on the same probability space only because it will be convenient to consider the Palm transformations of P with respect to the family of point processes $\{(D_{n,c})_{n\in\mathbf{Z}}\}_{c=1,2,\ldots}$.)

To simplify the notation we will drop the subscript c that refers to the number of customers in the network for the rest of this section. All quantities will refer to the network with population c, unless marked by a tilde, in which case they refer to the same network with population $c-1$.

The arrival theorem ([23,26]; see also [2,3,27]) states that, in steady state, a customer completing service at a node sees the network in equilibrium with one less customer. This can be made precise as follows: Let $\mathbf{e}_i$ designate the unit vector in the ith direction, $i = 1, \ldots, M$, and note that $\mathbf{X}(0) = \mathbf{n} + \mathbf{e}_1$ P^0-a.s. for some $\mathbf{n} \in \mathcal{S}(M, c-1)$ by the right continuity of the paths of $\mathbf{X}$ since, in a cyclic network, right after a departure from node M one customer will always be in node 1. The arrival theorem as it applies to our system can then be summarized by:

Theorem 1. In cyclic closed networks with exponential servers

$$P^0\{\mathbf{X}(0) = \mathbf{n} + \mathbf{e}_1\} = P\{\widetilde{\mathbf{X}}(0) = \mathbf{n}\}, \quad \forall\, \mathbf{n} \in \mathcal{S}(M, c-1). \tag{1}$$

The arrival theorem has immediate implications for the structure of the departure process from a node (say node M). Suppose that in a *cyclic single node* network with c customers a tagged customer leaves node M at time $t = 0$. We will show that the joint statistics of the $c-1$ departure epochs preceding 0 and the $c-1$ departures following 0 are identical to the $c-1$ departures preceding and following zero in the same network *in equilibrium* with $c-1$ customers. (For instance, in figure 1 the upper sample path depicts the Palm version of the departure process from node M in a network with $c = 3$ customers while the sample path below shows the stationary version of the departure process from the same node in a network with 2 customers. We will show that the Palm distribution of $(D_{-2}, D_{-1}, D_1, D_2)$ and the stationary distribution of $(\widetilde{D}_{-1}, \widetilde{D}_0, \widetilde{D}_1, \widetilde{D}_2)$ are the same.) The above statement is formalized in the following

Theorem 2. The distribution of $(D_{-c+1}, \ldots, D_{-1}, D_1, \ldots, D_{c-1})$ under P^0 is the same as the distribution of $(\widetilde{D}_{-c+2}, \ldots, \widetilde{D}_0, \widetilde{D}_1, \ldots, \widetilde{D}_{c-1})$ under P. Suppose that, A_n, B_n, $1 \leqslant n \leqslant c$ are Borel subsets of $\mathbf{R}$. Then

$$\begin{aligned} &P^0\big(D_{-c+1} \in A_{c-1}, \ldots, D_{-1} \in A_1, D_1 \in B_1, \ldots, D_{c-1} \in B_{c-1}\big) \\ &\quad = P\big(\widetilde{D}_{-c+2} \in A_{c-1}, \ldots, \widetilde{D}_0 \in A_1, \widetilde{D}_1 \in B_1, \ldots, \widetilde{D}_{c-1} \in B_{c-1}\big). \end{aligned} \tag{2}$$

Remark. One of the implications of the above statement is that the stationary forward recurrence time of the departure process from node M in the network with population $c-1$ has the same distribution as the typical interdeparture interval in the network with population c.

Proof. Condition the left hand side of (2) on the state of the network right after a customer has jumped from node M to node 1:

$$P^0\big(D_{-c+1} \in A_{c-1}, \ldots, D_{-1} \in A_1, D_1 \in B_1, \ldots, D_{c-1} \in B_{c-1} \mid \mathbf{X}(0) = \mathbf{n}+\mathbf{e}_1\big). \tag{3}$$

By the Markov property, past and future departures are conditionally independent given the present state and the above display becomes

$$\begin{aligned} &P^0\big(D_{-c+1} \in A_{c-1}, \ldots, D_{-1} \in A_1 \mid \mathbf{X}(0) = \mathbf{n} + \mathbf{e}_1\big) \\ &\quad \times P^0\big(D_1 \in B_1, \ldots, D_{c-1} \in B_{c-1} \mid \mathbf{X}(0) = \mathbf{n} + \mathbf{e}_1\big). \end{aligned} \tag{4}$$

Let us consider first the second term in (4). We will use the arrival theorem *and the non-overtake property of cyclic single server networks* to show that

$$
\begin{aligned}
&P^0\big(D_1 \in B_1, \ldots, D_{c-1} \in B_{c-1} \mid \mathbf{X}(0) = \mathbf{n} + \mathbf{e}_1\big) \\
&\quad = P\big(\widetilde{D}_1 \in B_1, \ldots, \widetilde{D}_{c-1} \in B_{c-1} \mid \widetilde{\mathbf{X}}(0) = \mathbf{n}\big). \qquad (5)
\end{aligned}
$$

The role of the arrival theorem in the above equality is obvious. To see the necessity for the non-overtake condition note that after a departure of the "tagged" customer from node M at time 0 (which is necessarily an arrival to node 1) the next departure from node M cannot be the tagged customer, who in fact will not influence in any way whatsoever the next $c-1$ departures from node M since overtaking is not possible.

We now turn our attention to the first term in (4). Let $-A_n = \{-x\colon\ x \in A_n\}$. We will also need to consider the *time reversed* network; *all quantities referring to time reversed processes will be designated by primes.* In particular $\{D'_n\}$ (respectively $\{\widetilde{D}'_n\}$) is the point process of customers jumping from node 1 to node M in the reversed network with c (respectively $c-1$) customers. Also, both the forward and reversed time process are assumed to have *right-continuous* sample paths. Then

$$
\begin{aligned}
&P^0\big(D_{-c+1} \in A_{c-1}, \ldots, D_{-1} \in A_1 \mid \mathbf{X}(0) = \mathbf{n} + \mathbf{e}_1\big) \\
&\quad = P^0\big(D'_1 \in -A_1, \ldots, D'_{c-1} \in -A_{c-1} \mid \mathbf{X}'(0) = \mathbf{n} + \mathbf{e}_M\big) \\
&\quad = P\big(\widetilde{D}'_1 \in -A_1, \ldots, \widetilde{D}'_{c-1} \in -A_{c-1} \mid \widetilde{\mathbf{X}}'(0) = \mathbf{n}\big) \\
&\quad = P\big(\widetilde{D}_1 \in A_1, \ldots, \widetilde{D}_{c-1} \in A_{c-1} \mid \widetilde{\mathbf{X}}(0) = \mathbf{n}\big).
\end{aligned}
$$

The first step above is obtained by considering the time reversed process: The statistics of the $c-1$ epochs of customer jumps from M to 1 before 0, given that at time 0 a jump occurs from M to 1 and the position of the rest of the customers is described by $\mathbf{n}$ are the same as those of the $c-1$ epochs of customer jumps *from 1 to M after time 0* in the time reversed network given that, at time 0, a customer jumps from 1 to M and the rest of the customers are distributed in the network according to $\mathbf{n}$. The second step follows directly by (5) applied to the reversed network, and the last one by reversing time once more in the network in equilibrium with one less customer.

Thus (4) is equal to

$$
\begin{aligned}
&P\big(\widetilde{D}_{-c+2} \in A_{c-1}, \ldots, \widetilde{D}_0 \in A_1 \mid \widetilde{\mathbf{X}}(0) = \mathbf{n}\big) \\
&\quad \times P\big(\widetilde{D}_1 \in B_1, \ldots, \widetilde{D}_{c-1} \in B_{c-1} \mid \widetilde{\mathbf{X}}(0) = \mathbf{n}\big),
\end{aligned}
$$

which, by virtue of the Markov property, is in turn equal to

$$
P\big(\widetilde{D}_{-c+2} \in A_{c-1}, \ldots, \widetilde{D}_0 \in A_1, \widetilde{D}_1 \in B_1, \ldots, \widetilde{D}_{c-1} \in B_{c-1}) \mid \widetilde{\mathbf{X}}(0) = \mathbf{n}\big). \qquad (6)
$$

Then (2) follows directly from (1) and the equality of the conditional probabilities in (3) and (6). □

Theorem 3. In the above network with population c let $T_0 := D_c - D_0$ denote the cycle time of the customer who leaves node M at time D_0. Similarly, $\widetilde{T}_0 := \widetilde{D}_{c-1} - \widetilde{D}_0$ is the cycle time of a typical customer in the network with population $c-1$. Then, for any measurable, nonnegative function $g : \mathbf{R}_+ \to \mathbf{R}_+$,

$$E^0 g(T_0) = \frac{\widetilde{E}^0[\widetilde{T}_0 g(\widetilde{T}_0)]}{\widetilde{E}^0 \widetilde{T}_0}. \tag{7}$$

Remark. The above theorem shows that the statistics of the cycle time in a network with c customers can be obtained from those of a network with $c-1$ customers via the change of measure given by $\widetilde{T}_0/(\widetilde{E}^0\widetilde{T}_0)$. The representations of the next section are based on a repeated application of this idea.

Proof. In view of theorem 2,

$$E^0 g\big(D_{c-k} - D_{-k}\big) = Eg\big(\widetilde{D}_{c-k} - \widetilde{D}_{-k+1}\big) \quad \text{for } k = 1, 2, \ldots, c-1, \tag{8}$$

while from the Palm inversion formula we obtain

$$\begin{aligned} Eg\big(\widetilde{D}_{c-k} - \widetilde{D}_{1-k}\big) &= \lambda_{c-1}\widetilde{E}^0 \int_{\widetilde{D}_0}^{\widetilde{D}_1} g\big(\widetilde{D}_{c-k} - \widetilde{D}_{1-k}\big)\mathrm{d}u \\ &= \lambda_{c-1}\widetilde{E}^0\big[\big(\widetilde{D}_1 - \widetilde{D}_0\big) g\big(\widetilde{D}_{c-k} - \widetilde{D}_{1-k}\big)\big], \end{aligned} \tag{9}$$

where $\lambda_{c-1} := 1/(\widetilde{E}^0[\widetilde{D}_1 - \widetilde{D}_0])$ is the throughput of the network with $c-1$ customers. From (8) and (9), expressing all the quantities involved in terms of the interoutput times $\widetilde{\tau}_n = \widetilde{D}_{n+1} - \widetilde{D}_n$,

$$E^0 g\big(D_{c-k} - D_{-k}\big) = \lambda_{c-1}\widetilde{E}^0\big[\widetilde{\tau}_0 g(\widetilde{\tau}_{1-k} + \cdots + \widetilde{\tau}_{c-k-1})\big].$$

Now use the invariance of P^0 under shifts along the points of $\{D_n\}$ in the left hand side of the above equation and the invariance of $\widetilde{P}^0$ under shifts along $\{\widetilde{D}_n\}$ in the right hand side to rewrite the above as

$$E^0 g\big(D_c - D_0\big) = \lambda_{c-1}\widetilde{E}^0\big[\widetilde{\tau}_k \ g(\widetilde{\tau}_0 + \cdots + \widetilde{\tau}_{c-2})\big] \quad \text{for } k = 1, 2, \ldots, c-1,$$

or equivalently

$$E^0 g(T_0) = \lambda_{c-1}\widetilde{E}^0\big[\widetilde{\tau}_k \ g\big(\widetilde{T}_0\big)\big] \quad \text{for } k = 1, 2, \ldots, c-1,$$

since $\widetilde{T}_0 := \widetilde{D}_{c-1} - \widetilde{D}_0 = \widetilde{\tau}_0 + \cdots + \widetilde{\tau}_{c-2}$. Adding term by term the above $c-1$ equations we obtain

$$E^0 g(T_0) = \frac{\lambda_{c-1}}{c-1}\widetilde{E}^0\big[\widetilde{T}_0 g\big(\widetilde{T}_0\big)\big]. \tag{10}$$

Letting $g(x) \equiv 1$ in the above equation we obtain $\lambda_{c-1}\widetilde{E}^0[\widetilde{T}_0] = c-1$ (Little's law). A direct substitution back in (10) completes the proof. $\square$

We also point out that, with $g(x) = 1/x$ the above theorem gives the following "harmonic mean" formula:

$$E^0\left[\frac{1}{T_0}\right] = \frac{1}{\widetilde{E}^0\widetilde{T}_0},$$

reminiscent of the corresponding harmonic mean formula of Palm theory.

3. Representations for the cycle time, interoutput times, and throughput

In this section we will apply recursively the "change of measure" result of theorem 3 to express the cycle time in the network with c customers in terms of the cycle time in the network with a single customer. Throughout the section we will use a subscript c to designate quantities referring to the network with c customers. In particular, the nth departure epoch from node M will be denoted by $D_{n,c}$ and the cycle time of the customer who leaves node M by $T_{0,c} = D_{c,c} - D_{0,c}$.

Theorem 4. Let Y_i, $i = 1, \dots, M$, be independent, exponential random variables with rates $\mu_1, \dots, \mu_M$, defined on the probability space $(\Omega, \mathcal{F}, P)$ and $Y := \sum_{i=1}^M Y_i$. Then the cycle time distribution is given by

$$P_c^0(T_{0,c} \leqslant x) = \frac{E[Y^{c-1}\mathbf{1}(Y \leqslant x)]}{E[Y^{c-1}]}, \tag{11}$$

with the corresponding Laplace transform

$$E_c^0 \mathrm{e}^{-sT_{0,c}} = \frac{E[Y^{c-1}\mathrm{e}^{-sY}]}{EY^{c-1}}. \tag{12}$$

In particular, the moments of the cycle time distribution are given by the expression

$$E_c^0 T_{0,c}^k = \frac{EY^{c-1+k}}{EY^{c-1}}, \quad k = 1, 2, \dots, \tag{13}$$

and the throughput by

$$\lambda_c = c\frac{EY^{c-1}}{EY^c}. \tag{14}$$

Proof. Start with (7) of theorem 3 and apply the relationship recursively (with $g(y) = y^n\mathbf{1}\ (y \leqslant x)$, $n = 0, 1, 2, \dots$) to obtain

$$\begin{aligned} E_c^0\mathbf{1}(T_{0,c} \leqslant x) &= \frac{E_{c-1}^0[T_{0,c-1}\mathbf{1}(T_{0,c-1} \leqslant x)]}{E_{c-1}^0[T_{0,c-1}]} = \cdots \\ &= \frac{E_1^0[(T_{0,1})^{c-1}\mathbf{1}(T_{0,1} \leqslant x)]}{E_{c-1}^0[T_{0,c-1}]E_{c-2}^0[T_{0,c-2}]\cdots E_1^0[T_1]}. \end{aligned} \tag{15}$$

Letting $x \to \infty$ in the above equation we obtain (by monotone convergence and the fact that $T_{0,c} < \infty$, P_c^0-a.s.)

$$E_1^0\big[(T_{0,1})^{c-1}\big] = E_{c-1}^0[T_{0,c-1}]E_{c-2}^0[T_{0,c-2}]\cdots E_1^0[T_1] \tag{16}$$

and substituting in (15),

$$P_c^0(T_{0,c} \leqslant x) = \frac{E_1^0[T_{0,1}^{c-1}\mathbf{1}(T_{0,c} \leqslant x)]}{E_1^0(T_{0,1})^{c-1}}. \tag{17}$$

To conclude the proof of (11) we need only recall that $T_{0,1}$, the cycle time in a cyclic network with a single customer, is simply the sum of the processing times at the M stations, which have been assumed to be independent exponential random variables.

The expression for the Laplace transform (12) is obtained by writing

$$E_c^0\mathrm{e}^{-sT_{0,c}} = 1 - s\int_0^\infty P_c^0(T_{0,c} > x)\mathrm{e}^{-sx}\,\mathrm{d}x$$

and using (11) and Fubini's theorem. The expression for the moments can also be obtained readily from (11) again using Fubini's theorem. Finally, (14) follows from (13) with $k = 1$ and Little's law. □

Remark. The random variable Y was introduced to illustrate the simplicity of the final representations. (11) is of course equivalent to (17).

3.1. Joint distribution of the interoutput times

The above approach can be used to obtain the joint distribution of the interoutput times that comprise the cycle time of a customer. When the population size is c, the cycle time of the tagged customer that departs from node M at time 0 is $T_{0,c} := D_{c,c} - D_{0,c} = \tau_{0,c} + \cdots + \tau_{c-1,c}$. The following theorem gives an expression for the joint distribution of the c consecutive interoutput times $\tau_{n,c}$, $n = 0, 1, \ldots, c-1$.

Theorem 5. The joint distribution of c consecutive interoutput times in a cyclic network with c customers can be expressed as

$$P_c^0(\tau_{0,c} > x_0, \ldots, \tau_{c-1,c} > x_{c-1}) = \frac{E(Y - x_0 - x_1 - \cdots - x_{c-1})_+^{c-1}}{EY^{c-1}}, \tag{18}$$

where $x_0, \ldots, x_{c-1} \in \mathbf{R}_+$, and Y is defined as in theorem 4. In particular, we note that the joint distribution any c consecutive interoutput times is *invariant under permutations.*

Proof. By the invariance of P_c^0 under shifts along the points of $\{D_{n,c}\}$ the left hand side of (18) can be written as $P_c^0(\tau_{-1,c} > x_0, \tau_{0,c} > x_1, \ldots, \tau_{c-2,c} > x_{c-1})$ which in turn, from the arrival theorem, is equal to

$$P\big(\alpha_{c-1}(0) > x_0, \tau_{0,c-1} - \alpha_{c-1}(0) > x_1, \tau_{1,c-1} > x_2, \ldots, \tau_{c-2,c-1} > x_{c-1}\big),$$

where $\{\alpha_{c-1}(t);\ t\in\mathbf{R}\}$ is the age process associated with the point process $\{D_{n,c-1}\}$ (see figure 1). Using now the Palm inversion formula and denoting by $(x)_+$ the positive part of a real number x, the above display becomes

$$\begin{aligned}&\lambda_{c-1}E^0_{c-1}\int_0^{\tau_{0,c-1}}\mathbf{1}(u>x_0,\tau_{0,c-1}-u>x_1,\tau_{1,c-1}>x_2,\dots,\tau_{c-2,c-1}>x_{c-1})\,\mathrm{d}u\\ &\quad=\lambda_{c-1}E^0_{c-1}\big[(\tau_{0,c-1}-x_0-x_1)_+\ \mathbf{1}(\tau_{1,c-1}>x_2,\dots,\tau_{c-2,c-1}>x_{c-1})\big].\end{aligned}$$

Use once more the invariance of P^0_{c-1} under shifts along the points of $\{D_{n,c-1}\}$ to write the last expectation in the above display as

$$E^0_{c-1}\big[(\tau_{-1,c-1}-x_0-x_1)_+\ \mathbf{1}(\tau_{0,c-1}>x_2,\dots,\tau_{c-3,c-1}>x_{c-1})\big]. \tag{19}$$

We now repeat the same cycle of arguments. We use again the arrival theorem, to express (19) as a *stationary* expectation in a network with $c-2$ customers and then apply the Palm inversion formula:

$$\begin{aligned}&E\big[\big(\alpha_{c-2}(0)-x_0-x_1\big)_+\\ &\qquad\times\mathbf{1}\big(\tau_{0,c-2}-\alpha_{c-2}(0)>x_2,\tau_{1,c-2}>x_3,\dots,\tau_{c-3,c-2}>x_{c-1}\big)\big]\\ &\quad=\lambda_{c-2}E^0_{c-2}\int_0^{\tau_{0,c-2}}(u-x_0-x_1)_+\ \mathbf{1}(\tau_{0,c-2}-u>x_2)\\ &\qquad\times\mathbf{1}(\tau_{1,c-2}>x_3,\dots,\tau_{c-3,c-2}>x_{c-1})\,\mathrm{d}u\\ &\quad=\frac{1}{2}\lambda_{c-2}E^0_{c-2}\big[(\tau_{0,c-2}-x_0-x_1-x_2)^2_+\ \mathbf{1}(\tau_{1,c-2}>x_3,\dots,\tau_{c-3,c-2}>x_{c-1})\big].\end{aligned}$$

Repeated application of the above arguments leads to the following expression for the joint distribution of the interoutput times:

$$P^0_c(\tau_{0,c}>x_0,\dots,\tau_{c-1,c}>x_{c-1})=\frac{\lambda_{c-1}\cdots\lambda_1}{(c-1)!}E^0_1\big[(\tau_{0,1}-x_0-x_1-\cdots-x_{c-1})^{c-1}_+\big].$$

Now let $x_0,\dots,x_{c-1}\downarrow 0$ in the above equation and conclude (by monotone convergence and the fact that the interoutput time distribution does not have atoms at zero)

$$1=\frac{\lambda_{c-1}\lambda_{c-2}\cdots\lambda_1}{(c-1)!}E^0_1\big[(\tau_{0,1})^{c-1}\big]. \tag{20}$$

(18) follows then from the above two equations. □

Corollary 6. The interoutput time distribution is given by

$$P^0_c(\tau_{0,c}>x)=\frac{E(Y-x)^{c-1}_+}{EY^{c-1}}, \tag{21}$$

with moments

$$E^0_c(\tau_{0,c})^k=\frac{1}{\binom{k+c-1}{c-1}}\frac{EY^{c+k-1}}{EY^{c-1}}. \tag{22}$$

Proof. The expression for the tail of the distribution follows immediately from (18) by setting $x_0 = x$, $x_1 = \cdots = x_{c-1} = 0$. To obtain the expression for the moments note that

$$\begin{aligned}
E_c^0(\tau_{0,c})^k &= k\int_0^\infty x^{k-1}P_c^0(\tau_{0,c} > x)\,\mathrm{d}x = k\int_0^\infty \frac{E(Y-x)_+^{c-1}}{EY^{c-1}}\,\mathrm{d}x \\
&= \frac{k}{EY^{c-1}}E\int_0^Y x^{k-1}(Y-x)^{c-1}\,\mathrm{d}x \\
&= \frac{k}{EY^{c-1}}E\left[Y^{c+k-1}\int_0^1 u^{k-1}(1-u)^{c-1}\,\mathrm{d}x\right] \\
&= \frac{E[Y^{c+k-1}]}{EY^{c-1}}\frac{k(k-1)!(c-1)!}{(k+c-1)!},
\end{aligned}$$

where in the above computation we have used Fubini's theorem and a well known integration formula involving the Beta function. (22) follows immediately from the above. □

Corollary 7. The covariance of two interoutput times, $\tau_{i,c}$, $\tau_{i+j,c}$, such that $|j| < c$, is given by

$$\mathrm{Cov}(\tau_{i,c}, \tau_{i+j,c}) = E_c^0[\tau_{0,c}\tau_{j,c}] - \left(E_c^0[\tau_{0,c}]\right)^2 = \frac{1}{\lambda_c}\left(\frac{1}{\lambda_{c+1}} - \frac{1}{\lambda_c}\right). \tag{23}$$

The corresponding correlation coefficient is given by $\mathrm{Corr}(\tau_{0,c}, \tau_{j,c}) = (\lambda_c - \lambda_{c+1})/(2\lambda_c - \lambda_{c+1})$. We note in particular that the covariance is negative (since $\lambda_{c+1} > \lambda_c$, and that it remains constant for $|j| < c$ (which can of course be inferred directly from the invariance of the joint distribution (18) under permutations).

Proof. A straightforward application of Fubini's theorem shows that

$$\begin{aligned}
E_c^0[\tau_{0,c}\tau_{j,c}] &= \int_0^\infty\int_0^\infty P_c^0(\tau_{0,c} > x, \tau_{j,c} > y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \frac{1}{EY^{c-1}}E\int_0^\infty\int_0^\infty (Y-x-y)_+^{c-1}\,\mathrm{d}x\,\mathrm{d}y \\
&= \frac{1}{EY^{c-1}}E\int_0^Y\int_0^{Y-y} (Y-x-y)^{c-1}\,\mathrm{d}x\,\mathrm{d}y \\
&= \frac{E[Y^{c+1}]}{c(c+1)E[Y^{c-1}]} = \frac{E[Y^{c+1}]}{(c+1)E[Y^c]}\,\frac{E[Y^c]}{cE[Y^{c-1}]} = \frac{1}{\lambda_{c+1}\lambda_c}.
\end{aligned}$$

Taking into account the obvious relationship $E_c^0\tau_{0,c} = 1/\lambda_c$ (which also follows from (19) with $k = 1$ and (14) we obtain (23). □

4. Closed form expressions

In this section we obtain closed form expressions for the cycle time distribution, the interoutput distribution, the throughput, and the partition function, in terms of the service rates at the nodes. To simplify the analysis we will assume throughout this section that $\mu_i \neq \mu_j$ for $i \neq j$. This assumption is of course not essential and the same techniques can be applied when the rates are not different, at the expense of the lack of simplicity of the final formulas.

4.1. *The cycle time*

Let $\Phi_c(s) := E_c^0 \mathrm{e}^{-sT_{0,c}}$ be the Laplace transform of the cycle time in the network with c customers. From (10) we have

$$\begin{aligned}\Phi_c(s) &= \frac{\lambda_{c-1}}{c-1} E_{c-1}^0 \left[T_{0,c-1} \mathrm{e}^{-sT_{0,c-1}}\right] = -\frac{\lambda_{c-1}}{c-1} E_{c-1}^0 \left[\frac{\mathrm{d}}{\mathrm{d}s} \mathrm{e}^{-sT_{0,c-1}}\right] \\ &= -\frac{\lambda_{c-1}}{c-1} \frac{\mathrm{d}}{\mathrm{d}s} \Phi_{c-1}(s),\end{aligned}$$

where in the last equation above the interchange between expectation and differentiation can be justified by a simple dominated convergence argument. From the above recursion we obtain

$$\Phi_c(s) = (-1)^{c-1} \frac{\prod_{n=1}^{c-1} \lambda_n}{(c-1)!} \frac{\mathrm{d}^{c-1}}{\mathrm{d}s^{c-1}} \Phi_1(s), \tag{24}$$

where $\Phi_1(s)$ is the Laplace transform of the cycle time with a single customer in the network:

$$\Phi_{\mathbf{1}}(s) = \prod_{i=1}^{M} \frac{\mu_i}{\mu_i + s} = \sum_{i=1}^{M} \alpha_i \frac{\mu_i}{\mu_i + s}, \tag{25}$$

with

$$\alpha_i := \prod_{\substack{j=1 \\ j\neq i}}^{M} \frac{\mu_j}{\mu_j - \mu_i}. \tag{26}$$

The second equation in (25) is obtained by a straightforward partial fractions expansion, valid provided that $\mu_i \neq \mu_j$ when $i \neq j$.

Remark. Note that (25) is clearly *symmetric in the μ_i's* (i.e., invariant under permutations) and hence without loss of generality we can assume that $\mu_1 < \mu_2 < \cdots < \mu_M$. In that case $\alpha_i = (-1)^{i-1}|\alpha_i|$, $i = 1, \ldots, M$, i.e., the signs alternate. Define now the quantities

$$\beta_i(c) := \frac{\alpha_i \mu_i^{-c}}{\sum_{i=1}^{M} \alpha_i \mu_i^{-c}}. \tag{27}$$

These "signed weights" will play an important role in a number of formulas that follow.

When $|s| < \min\{\mu_1, \ldots, \mu_M\}$, a series expansion of (25) gives

$$\Phi_{\mathbf{1}}(s) = \sum_{k=0}^{\infty} (-1)^k s^k \sum_{i=1}^{M} \alpha_i \mu_i^{-k}$$

and hence

$$\begin{aligned}\frac{\mathrm{d}^{c-1}}{\mathrm{d}s^{c-1}} \Phi_1(s) &= \sum_{k=c-1}^{\infty} (-1)^k k(k-1)\cdots(k-c+2) s^{k-c+1} \sum_{i=1}^{M} \alpha_i \mu_i^{-k} \\ &= (-1)^{c-1} \sum_{i=1}^{M} \alpha_i \mu_i^{-c+1} \sum_{k=0}^{\infty} (-1)^k (k+c-1)\cdots(k+1) \left(\frac{s}{\mu_i}\right)^k \\ &= (-1)^{c-1}(c-1)! \sum_{i=1}^{M} \alpha_i \mu_i^{-c+1} \left(\frac{\mu_i}{\mu_i + s}\right)^c. \qquad (28)\end{aligned}$$

The interchange in the order of summations above is justified by the absolute and uniform convergence of the series. Also, (28) holds for all $s > -\min\{\mu_1, \ldots, \mu_M\}$ by analytic continuation. Thus, from (24)

$$\Phi_c(s) = \lambda_1 \lambda_2 \cdots \lambda_{c-1} \sum_{i=1}^{M} \alpha_i \mu_i^{-c+1} \left(\frac{\mu_i}{\mu_i + s}\right)^c. \qquad (29)$$

Set $s = 0$ in the equation above to obtain

$$\lambda_1 \lambda_2 \cdots \lambda_{c-1} = \frac{1}{\sum_{i=1}^{M} \alpha_i \mu_i^{-c+1}} \qquad (30)$$

and hence

$$\Phi_c(s) = \frac{\sum_{i=1}^{M} \alpha_i \mu_i^{-(c-1)} (\mu_i/(\mu_i + s))^c}{\sum_{i=1}^{M} \alpha_i \mu_i^{-(c-1)}} = \sum_{i=1}^{M} \beta_i(c-1) \left(\frac{\mu_i}{\mu_i + s}\right)^c, \qquad (31)$$

where in the second expression we have used (27) to express the cycle time as a weighted sum or Erlang random variables. The corresponding density is

$$\sum_{i=1}^{M} \beta_i(c-1) \mu_i \frac{(\mu_i x)^{c-1}}{(c-1)!} \mathrm{e}^{-\mu_i x}.$$

4.2. Expressions for the throughput and partition function

We will need again Y, the cycle time of a customer alone in the network or, equivalently, the sum of the processing times: $Y = \sum_{i=1}^{M} Y_i$ where the Y_i's are independent, exponential random variables with rates μ_i. From (25) we see that the distribution of Y can be written as a mixture of exponential distributions with alternating signs:

$$P(Y \leqslant x) = 1 - \sum_{i=1}^{M} \alpha_i \mathrm{e}^{-\mu_i x}. \tag{32}$$

Hence, the moments of Y are given by

$$E\big[Y^c\big] = c! \sum_{i=1}^{M} \alpha_i \mu_i^{-c}, \quad c = 1, 2, \ldots. \tag{33}$$

From the above expression and (14) we obtain the following formula for the throughput

$$\lambda_c = \frac{\sum_{i=1}^{M} \alpha_i \mu_i^{-(c-1)}}{\sum_{i=1}^{M} \alpha_i \mu_i^{-c}} = \sum_{i=1}^{M} \beta_i(c) \mu_i. \tag{34}$$

The second equation above uses (27) to express the system throughput as a weighted average of the rates of the individual stations.

Now we return to (20) which we can rewrite as $\lambda_1 \lambda_2 \cdots \lambda_c = c! E Y^c$ (in view of the fact that the interoutput time in the network with a single customer, $\tau_{0,1}$ is the sum of the processing times in the M stations). Since the throughput can be expressed in terms of the partition function as $\lambda_c = G(c-1)/G(c)$, the above product telescopes to give the representation

$$G(c) = \frac{1}{c!} E\big[Y^c\big]. \tag{35}$$

The above, together with (33) leads to the following explicit expression for the partition function

$$G(c) = \sum_{i=1}^{M} \alpha_i \mu_i^{-c}. \tag{36}$$

(36) was derived first in [13] using direct computational arguments and later in [11] by a generating function approach.

We point out that the expressions in this section do not depend on the non-overtake condition and hence apply to single server Gordon–Newell networks with arbitrary Markovian routing. Let $R = [r_{ij}]$ be the (irreducible) routing matrix of the network and $\boldsymbol{\nu} := (\nu_1, \ldots, \nu_M)$ the stationary distribution of R. The partition function

$$G(c) = \sum_{n_1 + \cdots + n_M = c} \left(\frac{\nu_1}{\mu_1}\right)^{n_1} \cdots \left(\frac{\nu_M}{\mu_M}\right)^{n_M}$$

can still be expressed as $(1/c!)EY^c$ with $Y = \sum_{i=1}^{M} Y_i$, Y_i being again independent exponential random variables with rates μ_i/ν_i. (34) and (36) are also valid, with μ_i replaced by μ_i/ν_i and

$$\alpha_i = \prod_{j \neq i} \frac{\mu_j/\nu_j}{\mu_j/\nu_j - \mu_i/\nu_i}.$$

Algorithms and techniques for computing the partition function in closed queueing networks is of course a vast subject. In addition to the references in §1 we also point out the asymptotic techniques proposed in [21,22].

4.3. *Moments of the cycle time distribution*

From (13), (34), and (36) we readily obtain the following expressions for the moments of the cycle time:

$$E_c^0 T_{0,c}^k = c(c+1)\cdots(c+k-1)\frac{\sum_{i=1}^{M} \alpha_i \mu_i^{-(c+k-1)}}{\sum_{i=1}^{M} \alpha_i \mu_i^{-(c-1)}} \tag{37}$$

$$= c(c+1)\cdots(c+k-1)\frac{G(c+k-1)}{G(c-1)}, \quad k = 1, 2, \dots\ . \tag{38}$$

From (14), (34), and (37) we also obtain the following version of Little's law for the moments of the cycle time in a closed network:

$$\lambda_c \lambda_{c+1} \cdots \lambda_{c+k-1} E_c^0 T_{0,c}^k = c(c+1)\cdots(c+k-1). \tag{39}$$

(The above was obtained first in [20].) The above expression implies that the kth moment of the cycle time in a cyclic network with c customers is equal to the product of the mean cycle times in networks with $c, c+1, \dots, c+k-1$ customers

$$E_c^0 (T_{0,c})^k = E_c^0[T_{0,c}]\ E_{c+1}^0[T_{0,c+1}] \cdots E_{c+k-1}^0[T_{0,c+k-1}].$$

In particular, the coefficient of variation of the cycle time with c customers in the system can be written as

$$C_v(c) = \sqrt{\left(1 + \frac{1}{c}\right)\frac{\lambda_c}{\lambda_{c+1}} - 1}. \tag{40}$$

This last result has implications for the design and operation of manufacturing systems that rely on controlling the level of work-in-process inventory (see [4] and the references therein). Given that λ_c is increasing in c we have the (asymptotically tight) bound:

$$C_v(c) \leqslant \sqrt{1/c}.$$

4.4. *Interoutput times*

Finally we will provide explicit expressions for the joint distribution of the c interoutput times that comprise the cycle time of a customer. From (18) and (32)

$$P_c^0(\tau_{0,c} > x_0, \ldots, \tau_{c-1,c} > x_{c-1}) = \frac{1}{EY^{c-1}} \sum_{i=1}^{M} \alpha_i \int_{x_0+x_1+\cdots+x_{c-1}}^{\infty} y^{c-1} \mu_i \mathrm{e}^{-\mu_i y}\,\mathrm{d}y$$
$$= \frac{(c-1)!}{EY^{c-1}} \sum_{i=1}^{M} \alpha_i \mu_i^{-c+1} \mathrm{e}^{-\mu_i(x_0+x_1+\cdots+x_{c-1})}$$
$$= \sum_{i=1}^{M} \beta_i(c-1) \mathrm{e}^{-\mu_i(x_0+x_1+\cdots+x_{c-1})},$$

where in the last equation we have used (32) and (27).

Acknowledgement

The author would like to thank Takis Konstantopoulos for many helpful comments.

References

[1] F. Baccelli and P. Brémaud, *Elements of Queueing Theory: Palm-Martingale Calculus and Stochastic Recurrences* (Springer, Berlin, 1994).

[2] O.J. Boxma, F.P. Kelly and A.G. Könheim, The product form for sojourn time distributions in cyclic exponential queues, J. Assoc. Comput. Mach. 31(1) (1984) 128–133.

[3] P. Brémaud, Characteristics of queueing systems observed at events and the connection between stochastic intensity and Palm probabilities, Queueing Systems 5 (1989) 99–112.

[4] J.A. Buzacott, J.G. Shanthikumar and D. Yao, Jackson network models of manufacturing systems, in: *Stochastic Modeling and Analysis of Manufacturing Systems*, ed. D.D. Yao (Springer, New York, 1994).

[5] J.P. Buzen, Computational algorithms for closed queueing networks with exponential servers, Commun. Assoc. Comput. Mach. 16 (1973) 527–531.

[6] W.M. Chow, The cycle time distribution of exponential cyclic queues, J. Assoc. Comput. Mach. 27 (1981) 281–286.

[7] H. Daduna, Passage times for overtake-free paths in Gordon–Newell networks, Adv. Appl. Probab. 14 (1982) 672–686.

[8] P. Franken, D. Köning, U. Arndt and V. Schmidt, *Queues and Point Processes* (Wiley, Chichester, 1982).

[9] A.I. Gerasimov, An analytical method for obtaining performance measures of queueing networks with different classes of customers, Cybernetics 3 (1991) 98–102 (in Russian).

[10] A.I. Gerasimov, On normalizing constants in multi-class queueing networks, Oper. Res. 43 (1995) 704–711.

[11] J.J. Gordon, The evaluation of normalizing constants in closed queueing networks, Oper. Res. 38 (1990) 863–869.

[12] W.J. Gordon and G.F. Newell, Closed queueing networks with exponential servers, Oper. Res. 15 (1967) 252–267.
[13] P.G. Harrison, A note on cycle times in tree-like queueing networks, Adv. Appl. Probab. 16 (1984) 216–219.
[14] P.G. Harrison, On normalizing constants in queueing networks, Oper. Res. 33 (1985) 464–468.
[15] F.P. Kelly and P.K. Pollet, Sojourn times in closed queueing networks, Adv. Appl. Probab. 15 (1983) 638–656.
[16] E. Koenigsberg, Cyclic queues, Oper. Res. Quarterly 9 (1958) 22–35.
[17] E. Koenigsberg, Production lines and internal storage: a review, Mgmt. Sci. 5 (1959) 410–433.
[18] E. Koenigsberg, Twenty-five years of cyclic queues and queueing networks: a review, J. Oper. Res. Soc. 33 (1982) 605–619.
[19] J. McKenna, Asymptotic expansions of the sojourn time distribution functions of jobs in closed, product-form queueing networks, J. Assoc. Comp. Mach. 34 (1987) 985–1003.
[20] J. McKenna, A generalization of Little's law to moments of queue lengths and waiting times in closed, product-form queueing networks, J. Appl. Probab. 26 (1989) 121–133.
[21] J. McKenna and D. Mitra, Integral representation and asymptotic expansions for closed queueing networks: normal usage, Bell System Tech. J. 61 (1982) 346–360.
[22] J. McKenna and D. Mitra, Asymptotic expansions and integral representation of queue lengths in closed Markovian networks, J. Assoc. Comput. Mach. 31 (1982) 346–360.
[23] M. Reiser and S.S. Lavenberg, Mean-value analysis of closed multichain queueing networks, J. Assoc. Comput. Mach. 27 (1988) 313–322.
[24] R. Schassberger and H. Daduna, The time for a round trip in a cycle of exponential queues, J. Assoc. Comput. Mach. 30 (1980) 146–150.
[25] R. Schassberger and H. Daduna, Sojourn times in queueing networks with multiserver nodes, J. Appl. Probab. 24(2) (1987) 511–521.
[26] K.C. Sevcik and I. Mitrani, The distribution of queueing network states at input and output instants, J. Assoc. Comput. Mach. 28 (1981) 358–371.
[27] J. Walrand, *An Introduction to Queueing Networks* (Prentice-Hall, Englewood Cliffs, NJ, 1988).
[28] M.A. Zazanis, Push and pull systems with external demands, in: *Proceedings of 32nd Allerton Conference on Communication, Control, and Computing* (1994) pp. 691–698.
[29] M.A. Zazanis, Closed form expressions for the partition function and cycle time distribution in Gordon–Newel networks, Preprint, MIE Dept., University of Massachusetts, Amherst, MA (1995).