

Non-coincidence probabilities and the time-dependent behavior of tandem queues with deterministic input

W. Böhm

Department of Mathematical Statistics, University of Economics, A-1090 Vienna, Augasse 2-6, Austria

Received 6 December 1996; received in revised form 1 March 2000; accepted 1 March 2000

Abstract

In this paper we derive a formula for zero-avoiding transition probabilities of an r -node tandem queue with exponential servers and deterministic input. In particular, we show that these transition probabilities may be interpreted as non-coincidence probabilities of a set of dissimilar Poisson processes restricted by a time-dependent boundary. © 2000 Elsevier Science B.V. All rights reserved.

MSC: Primary 60K25; secondary 60C05

Keywords: Tandem queues; Non-coincidence probabilities; Poisson process

1. Introduction

The queueing system we consider in this paper has the following structure: there are $r \geq 1$ independent service stations arranged in a line, such that the output of one station is the input of the next station, each server has waiting room with infinite capacity. Customers arrive from outside the system at the first server at predetermined time instants τ_1, τ_2, \dots , the inter-arrival times $\tau_{n+1} - \tau_n$, need not be equal. The service times at the servers are independent exponential random variables with means $1/\lambda_i$, $i = 1, \dots, r$. Once a job finishes service at the last server, it leaves the system.

Let $Q_i(t)$ denote the number of customers waiting at server i at time t and assume that there is a positive number of customers waiting at this server initially, i.e. $Q_i(0) = m_i > 0$. Furthermore, let

$$T_i = \inf\{t: Q_i(t) \leq 0\} \tag{1}$$

and define

$$T = \min_i T_i. \tag{2}$$

Thus T is the time from system startup until a server becomes idle for the first time and in this way generalizes the notion of a busy period familiar from single-server systems.

It will be convenient to introduce the vector notation

$$\mathbf{Q}(t) = (Q_1(t), Q_2(t), \dots, Q_r(t))$$

$$\mathbf{m} = (m_1, m_2, \dots, m_r), \quad m_i \in \mathbb{N},$$

$$\mathbf{k} = (k_1, k_2, \dots, k_r), \quad k_i \in \mathbb{N}.$$

Our primary concern will be to derive a formula for

$$P(\mathbf{Q}(t) = \mathbf{k}, T > t | \mathbf{Q}(0) = \mathbf{m}), \quad (3)$$

the probability that the vector process $\mathbf{Q}(t)$ moves from state \mathbf{m} to state \mathbf{k} in $(0, t)$ in such a way that none of the servers becomes idle in the intervening time.

A determinant formula for (3) in the special case of a system with one service station only has been published by Stadje (1995) and by Böhm and Mohanty (1997), who deal also with the case of a non-stationary server.

However, if there is more than one server, such queueing systems seem to be notoriously difficult to analyze. This impression is strongly supported, if we look at results already available for tandem queues with a Poisson input stream. For this case Massey (1984) has given formulas for the functions (3) which turn out to be linear combinations of lattice Bessel functions. In a subsequent paper, Baccelli and Massey (1990) have analyzed the general transient behavior of Markovian tandem queues by operator analytic methods. But unfortunately, numerical evaluation of lattice Bessel functions turns out to be of exponential complexity with respect to the number of service stations. Böhm et al. (1994) found a determinant formula for the functions (3) using a combinatorial approach based on a multivariate generalization of the classical ballot theorem. As a result, the numerical computation of zero-avoiding transition probabilities for Markovian tandem systems was now possible in polynomial time. It is interesting to note that Massey's lattice Bessel functions are essentially infinite sums of those special determinants which are discussed in Böhm et al. (1994).

Is there any special reason for the prevalence of determinants in Markovian tandem queues and single-server systems with deterministic input? The affirmative answer is: yes. The transient behavior of such queueing systems during time periods in which the servers are continuously busy may be interpreted as a non-coincidence problem. This has been demonstrated by Böhm and Mohanty (1997). Therefore, an application of the classical Karlin–McGregor Theorem on non-coincidence probabilities leads quite naturally to those ubiquitous determinants.

The Karlin–McGregor Theorem, one of the most remarkable results in the theory of stochastic processes, states that under certain continuity conditions the probability of non-coincidence of a set of $r > 1$ independent and identical Markov processes is given by a determinant formed by the transition functions of these processes. Here non-coincidence means, that no two of the processes occupy the same state at the same time. The result is valid also for non-stationary processes having the strong Markov property, the state space being an arbitrary metric space. However, it is required that the processes have identical transition functions. For more details see the papers by Karlin and McGregor (1959) and Karlin (1988).

In this paper we will prove two interesting facts: (i) the functions (3) still can be interpreted as non-coincidence probabilities associated with a set of *dissimilar* Poisson processes, dissimilarity meaning that the transition rates need not be identical. In addition to non-coincidence, the processes are also restricted by a time-dependent boundary induced by the arrival instants τ_1, τ_2, \dots . We will show (ii), that even for dissimilar Poisson processes restricted by a moving boundary a determinant formula analogous to the Karlin–McGregor theorem holds. Remarkably, this formula consists of a *single* determinant. So the situation is considerably simpler than in the case of a Markovian tandem system where we have an infinite sum of determinants. Moreover, we will show that our determinant is of particularly simple structure, if the inter-arrival times are equal.

2. An auxiliary result — the dummy path lemma

We will see shortly that the transition functions (3) are intimately connected with the probability distribution of Poisson processes whose sample paths do not touch a certain time dependent boundary. Such a boundary, say $C(t)$, can always be represented by a discrete set of points, if $C(t)$ is a continuous and non-decreasing function. Thus we are left with the problem to determine the probability that the sample paths of a Poisson process do not pass through points of a given set. A solution to this problem is provided by the more general *Dummy Path Lemma*, which we state here for completeness:

Lemma 2.1 (Dummy Path Lemma, Böhm and Mohanty, 1997). *Let X_t be a strong Markov process with discrete state-space $\mathcal{S} \subset \mathbb{Z}$ and transition function*

$$P(X_t = k | X_s = m) = p(s, t; m, k), \quad s \leq t.$$

Furthermore, let

$$\mathcal{C}_M = \{(u_1, a_1), (u_2, a_2), \dots, (u_M, a_M)\}, \quad u_i \in \mathbb{R}^+, a_i \in \mathcal{S}$$

denote a set of points ordered such that $u_1 \leq u_2 \leq \dots \leq u_M$. The first coordinate of these points represents time, the second a possible state of the process X_t .

Then the probability that X_t moves from state m to state k in $(0, t)$ without touching any of the points in \mathcal{C}_M is given by the determinant

$$P(X_t = k, X_s \notin \mathcal{C}_M, 0 \leq s \leq t | X_0 = m) = \det \|d_{ij}\|_{(M+1) \times (M+1)}, \tag{4}$$

where

$$\begin{aligned} d_{00} &= p(0, t; m, k), \\ d_{0j} &= p(u_j, t; a_j, k), \quad j = 1, 2, \dots, M, \\ d_{i0} &= p(0, u_i; m, a_i), \quad i = 1, 2, \dots, M, \\ d_{ij} &= p(u_j, u_i; a_j, a_i), \quad i \geq j, \\ &0 \quad \text{otherwise} \end{aligned} \tag{5}$$

and $X_t \notin \mathcal{C}_M$ is a short-hand notation for $(t, X_t) \notin \mathcal{C}_M$.

The name *Dummy Path Lemma* may be justified by the fact that the points of the set \mathcal{C}_M may be regarded as degenerate sample paths of M probabilistically identical copies of X_t , degenerate in the sense that the starting and end points of these paths coincide. The determinant (5) then follows from Karlin’s Theorem (1988). Observe that by an interchange of rows, (5) becomes a Hessenberg form and therefore a recursive evaluation of this determinant is straightforward.

Now let $N(t)$ be a Poisson process with rate $\lambda > 0$ and let $P_{\mathcal{C}}(N(t) = b | N(0) = a)$ denote the probability that $N(t)$ moves from state a to state b in $(0, t)$ such that the sample paths of $N(t)$ do not pass through any of the points in \mathcal{C} . We may think of the set \mathcal{C} as a time dependent or moving boundary imposed on $N(t)$. A Laplace-expansion of (5) yields the formula

$$P_{\mathcal{C}}(N(t) = b | N(0) = a) = \frac{e^{-\lambda t} (\lambda t)^{b-a}}{(b-a)!} R(a, b), \tag{6}$$

where $R(a, b)$ can be evaluated recursively to

$$R(a, b) = R_{n+1} \tag{7}$$

with

$$R_{\ell} = - \sum_{i=0}^{\ell-1} \binom{c_{\ell}}{c_i} \left(1 - \frac{u_i}{u_{\ell}}\right)^{c_{\ell}-c_i} \left(\frac{u_i}{u_{\ell}}\right)^{c_i} R_i, \quad \ell = 1, 2, \dots, n+1, \tag{8}$$

subject to the boundary conditions

$$R_0 = -1, \tag{9}$$

$$u_0 = 0, \quad c_0 = a \quad \text{and} \quad u_{n+1} = t, \quad c_{n+1} = b.$$

3. The main results

In this section we will show (i) that the zero-avoiding transition probabilities (3) are just non-coincidence probabilities of a set of independent and dissimilar Poisson processes, restricted by a time-dependent boundary, and (ii) that there is still a determinant formula similar to the Karlin–McGregor formula. Let us deal with (i) first.

Assume that customers arrive at time instants τ_1, τ_2, \dots and assume also that $\tau_n < t \leq \tau_{n+1}$. Consider the service process $N_r(t)$ at the last node. During a time period in which server r is continuously busy, this process is Poisson with rate λ_r . It starts at height zero, and the corresponding sample paths are right continuous step functions with lefthand limits. The paths terminate at time t at height equal to the number of completed services in $(0, t)$. The service process $N_{r-1}(t)$ at node $r - 1$ is Poisson with rate λ_{r-1} . This process forms the input stream at node r . Since we require that the server at node r is continuously busy, it is necessary that the sample paths of $N_{r-1}(t)$ and $N_r(t)$ do not touch in $(0, t)$, i.e. they do not have a coincidence in the time interval $(0, t)$. Furthermore, since there are $m_r > 0$ customers already waiting at node r at time zero, we assume that $N_{r-1}(t)$ starts at height m_r .

Similarly, the service process at node $r - 2$ must start at height $m_{r-1} + m_r$ and there must not be any coincidence of the paths of $N_{r-2}(t)$ and $N_{r-1}(t)$. In general, the processes $N_i(t)$ start at heights

$$a_i = \sum_{\ell=i+1}^r m_\ell, \quad i = 1, 2, \dots, r \tag{10}$$

and we agree that an empty sum has to be interpreted as zero. Let $\mathbf{a} = (a_1, \dots, a_r)$.

Customers arrive at node 1 at predetermined time instants $\tau_1, \tau_2, \dots, \tau_n$. These time instants give rise to a step function $C(t)$, say, which is left continuous with right-hand limits and has jumps of size $+1$ at times $u_i = \tau_i, i = 1, \dots, n$. This function starts at height $\sum_{\ell=1}^r m_\ell$, because there are m_1 customers already waiting at node 1 at time zero. Since we require that the server at node 1 is permanently busy in $(0, t)$, the sample paths of $N_1(t)$ must not touch or cross this step function. However, because $N_1(t)$ is non-decreasing and has jumps of unit size only, it is sufficient to consider the right-hand limits of $C(t)$ only, because any excursion of $N_1(t)$ above $C(t)$ which terminates below $C(t)$ necessarily passes through any of the points

$$(u_i, c_i) = (\tau_i, c + i - 1), \quad i = 1, 2, \dots, n \tag{11}$$

with

$$c = \sum_{\ell=1}^r m_\ell. \tag{12}$$

Thus regarding $N_1(t)$, we encounter exactly the same situation as in Section 2: $N_1(t)$ is restricted by a time-dependent boundary $C(t)$ which is represented by the set of points

$$\mathcal{C} = \{(\tau_i, c + i - 1), 1 \leq i \leq n\}, \tag{13}$$

where c is defined above.

Let us next determine the terminal positions of the processes $N_i(t)$. Here we have to take care of the fact that there are $k_i > 0$ customers waiting at nodes $i = 1, 2, \dots, r$.

For node 1, this means, that $N_1(t)$ must terminate k_1 units below $C(t)$. But since by assumption $\tau_n < t \leq \tau_{n+1}$, $C(t)$ has already reached height $c + n$, thus $N_1(t)$ terminates at height $c + n - k_1$. Similarly, there are k_2 customers at node 2 at time t . Therefore, $N_1(t) - N_2(t) = k_2$, or $N_2(t)$ terminates at height $c + n - k_1 - k_2$. In general, the processes $N_i(t)$ terminate at time t at heights

$$b_i = \sum_{\ell=1}^r m_\ell + n - \sum_{\ell=1}^i k_\ell = c + n - \sum_{\ell=1}^i k_\ell, \quad i = 1, 2, \dots, r. \tag{14}$$

Define $\mathbf{b} = (b_1, b_2, \dots, b_r)$.

Let us summarize the basic facts: the event $\{T > t\}$ requires

1. The process $N_1(t)$ must not touch or cross the boundary $C(t)$, as explained above. If this happens, then the queue at the first station has become empty at least once, i.e. $T \leq t$.
2. There must not be any coincidence between the processes $N_{i-1}(t)$ and $N_i(t), 2 \leq i \leq r$, because a coincidence implies that the queue at station i has become empty at least once and therefore also $T \leq t$.

But observe that

$$\begin{aligned}
 P(N_1(s) \notin \mathcal{C}, N_i(s) < N_{i-1}(s), i = 2, \dots, r, 0 \leq s \leq t) \\
 = P(N_i(s) \notin \mathcal{C}, N_i(s) < N_{i-1}(s), i = 2, \dots, r, 0 \leq s \leq t).
 \end{aligned}$$

Thus, $P(T > t)$ is a non-coincidence probability of the processes $N_i(t)$, $i = 1, \dots, r$, whose sample paths are restricted to stay ultimately below $\mathcal{C}(t)$.

We may summarize these findings in

Theorem 3.1. *Let $N(t) = (N_1(t), \dots, N_r(t))$ and $\tau_n < t \leq \tau_{n+1}$. Then*

$$\begin{aligned}
 P(Q(t) = \mathbf{k}, T > t | Q(0) = \mathbf{m}) \\
 = P(N(t) = \mathbf{b}, \text{non-coincidence} \cap \mathcal{M} | N(0) = \mathbf{a}),
 \end{aligned} \tag{15}$$

where \mathcal{M} is the event

$$\mathcal{M} = \{\text{none of the processes } N_i(t) \text{ touches the boundary } \mathcal{C}\},$$

and \mathcal{C}, \mathbf{a} and \mathbf{b} are given by (13), (10) and (14).

What remains to be done, is to show that (15) is in fact a determinant:

Theorem 3.2.

$$\begin{aligned}
 P(N(t) = \mathbf{b}, \text{no coincidence} \cap \mathcal{M} | N(0) = \mathbf{a}) \\
 = \det \left\| \frac{R(a_j, b_i)}{(b_i - a_j)!} \right\|_{r \times r} e^{-\sum_{v=1}^r \lambda_v t} \prod_{v=1}^r (\lambda_v t)^{b_v - a_v}
 \end{aligned} \tag{16}$$

where $R(a_j, b_i)$ is defined by (7).

Proof. Let us denote by $P_{\lambda_1, \dots, \lambda_r}(A)$ the probability of the event A when the processes $N_i(t)$, $1 \leq i \leq r$, are independent Poisson with rates λ_i . Given $N_i(0) = a_i$ and $N_i(t) = b_i$, the $b_i - a_i$ points where $N_i(t)$ has jumps are uniformly distributed in $(0, t)$. Hence, it follows that

$$\begin{aligned}
 P_{\lambda_1, \dots, \lambda_r}(\text{no coincidence} \cap \mathcal{M} | N(0) = \mathbf{a}, N(t) = \mathbf{b}) \\
 = P_{1, \dots, 1}(\text{no coincidence} \cap \mathcal{M} | N(0) = \mathbf{a}, N(t) = \mathbf{b}).
 \end{aligned} \tag{17}$$

The right-hand side of (17) can be evaluated using the Karlin–McGregor Theorem. In particular, we find that

$$\begin{aligned}
 P_{1, \dots, 1}(\text{no coincidence} \cap \mathcal{M} | N(0) = \mathbf{a}, N(t) = \mathbf{b}) \\
 = \frac{P_{1, \dots, 1}(N(t) = \mathbf{b}, \text{no coincidence} \cap \mathcal{M} | N(0) = \mathbf{a})}{P_{1, \dots, 1}(N(t) = \mathbf{b} | N(0) = \mathbf{a})} \\
 = \frac{\prod_{v=1}^r (b_v - a_v)!}{e^{-rt} \prod_{v=1}^r t^{b_v - a_v}} \det \left\| \frac{e^{-t} t^{b_i - a_j}}{(b_i - a_j)!} R(a_j, b_i) \right\| \\
 = \det \left\| \frac{R(a_j, b_i)}{(b_i - a_j)!} \right\| \prod_{v=1}^r (b_v - a_v)!.
 \end{aligned} \tag{18}$$

If we multiply (18) by $P(N(t) = \mathbf{b} | N(0) = \mathbf{a})$ we get (16). \square

4. The special case of equal inter-arrival times

Some simplifications are possible, if we assume that inter-arrival times $\tau_{i+1} - \tau_i$ are equal to some constant $\alpha > 0$. The boundary that the processes $N_1(t), \dots, N_r(t)$ must not touch or cross becomes now the straight line $\mathcal{C}(t) = c - 1 + t/\alpha$, where $c = \sum_{\ell=1}^r m_\ell$, the sum of all customers in the system at time zero.

As before, let $P_{\mathcal{C}}(N(t) = b | N(0) = a)$ denote the probability that the Poisson process $N(t)$, which starts in state a and terminates at time t in state b , does not touch or cross the boundary $\mathcal{C}(t)$. Then again

$$P_{\mathcal{C}}(N(t) = b | N(0) = a) = \frac{e^{-\lambda t} (\lambda t)^{b-a}}{(b-a)!} R(a, b),$$

as in the case of variable inter-arrival times. But now we have an explicit expression for functions $R(a, b)$. This formula can be found by means of Takács' Ballot Theorem. Of course, the recurrence (8) will yield the same result, but it requires some messy calculations (see also Zacks, 1991).

Takács' Ballot Theorem (Takács, 1967, Theorem 1 of Section 13, p. 37) states that

$$P(X(s) < s, 0 < s < t | X(t) = k, X(0) = 0) = 1 - \frac{k}{t} \tag{19}$$

for any process $X(t)$ with non-decreasing sample paths and independent increments. For real $\alpha > 0$, $\alpha N(t)$ is a process satisfying exactly these requirements. Thus, (19) applies and we have

$$P(N(s) < s/\alpha, 0 < s < t | N(t) = k, N(0) = 0) = 1 - \frac{k\alpha}{t}. \tag{20}$$

Conditioning on the point where paths touch or cross the boundary for the *last* time, we get

$$\begin{aligned} P_{\mathcal{C}}(N(t) = b | N(0) = a) &= P(N(t) = b | N(0) = a) \\ &\quad - \sum_{i=1}^{b-c+1} P(N(\alpha i) = c + i - 1 | N(0) = a) \\ &\quad \times P(N(t - \alpha i) = b - c - i + 1, N(s) < s/\alpha, 0 < s < t - \alpha i | N(0) = 0). \end{aligned}$$

Using (20) and dividing by $P(N(t) = b | N(0) = a)$ we obtain

$$\begin{aligned} R(a, b) &= 1 - \left(1 - \frac{\alpha(b-c+1)}{t} \right) \\ &\quad \times \sum_{i=1}^{b-c+1} \binom{b-a}{c+i-a-1} \left(\frac{\alpha i}{t} \right)^{c+i-a-1} \left(1 - \frac{\alpha i}{t} \right)^{b-c-i}. \end{aligned} \tag{21}$$

If we reverse the order of summation in (21) and observe that

$$\sum_{i=0}^{b-c} = \sum_{i=0}^{b-a} - \sum_{i=b-c+1}^{b-a}, \tag{22}$$

then the first summation on the right-hand side of (22) simplifies to

$$\frac{t}{t - \alpha(b - c + 1)}$$

according to Abel’s identity (Riordan, 1968, p. 23). Thus, we get the alternative representation

$$R(a, b) = \left(1 - \frac{\alpha(b - c + 1)}{t}\right) \sum_{j=0}^{c-a-2} \left(-\frac{\alpha}{t}(c - a - j - 1)\right)^j \times \left(1 + \frac{\alpha}{t}(c - a - j - 1)\right)^{b-a-j-1}, \tag{23}$$

which is much more convenient from a computational point of view for large values of b .

The latter situation is particularly interesting, if we study the system under heavy traffic conditions. More precisely, assume that

$$1/\alpha > \lambda_1 > \lambda_2 > \dots > \lambda_r.$$

Set $t = n\alpha + \delta$, $0 < \delta < \alpha$ and put

$$b_i = a_i + \lambda_i n\alpha + \beta_i, \quad i = 1, \dots, r,$$

β_i being the real constants independent of n and chosen such that b_i is an integer. Note that this choice implies for the number of customers waiting at the various stations at time t :

$$k_1 = c + n(1 - \lambda_1\alpha) - a_1 - \beta_1,$$

$$k_i = n\alpha(\lambda_{i-1} - \lambda_i) + a_{i-1} - a_i + \beta_{i-1} - \beta_i, \quad i = 2, \dots, r.$$

Let $n \rightarrow \infty$ in (23). This yields

$$\begin{aligned} \lim_{n \rightarrow \infty} R(a_j, b_i) &= (1 - \lambda_i\alpha) e^{\alpha\lambda_i(c-a_j-1)} \sum_{v=0}^{c-a_j-2} \frac{(-\lambda_i)^v}{v!} (c - a_j - v - 1)^v e^{-\alpha\lambda_i v} \\ &= w_i(a_j) \quad (\text{say}). \end{aligned} \tag{24}$$

This limiting relation enables us to find a useful approximation for the conditional probability $P(T > t | \mathbf{Q}(t) = \mathbf{k}, \mathbf{Q}(0) = \mathbf{m})$. By Stirling’s formula

$$\frac{(b_i - a_i)!}{(b_i - a_j)!} \sim \lambda_i^{a_j - a_i} (n\alpha)^{a_j - a_i}. \tag{25}$$

Furthermore, by (15), (16), (24) and (25)

$$\begin{aligned} P(T > t | \mathbf{Q}(t) = \mathbf{k}, \mathbf{Q}(0) = \mathbf{m}) &= \det \left\| R(a_j, b_i) \frac{(b_i - a_i)!}{(b_i - a_j)!} \right\| \\ &\sim \det \left\| w_i(a_j) \lambda_i^{a_j - a_i} (n\alpha)^{a_j - a_i} \right\| \\ &= \det \left\| w_i(a_j) \lambda_i^{a_j - a_i} \right\|. \end{aligned} \tag{26}$$

Acknowledgements

The author greatly appreciates the constructive and helpful comments of the anonymous referees.

References

- Baccelli, F., Massey, W.A., 1990. A transient analysis of the two-node series Jackson network. Technical Report, INRIA-Sophia, France.
- Böhm, W., Jain, J.L., Mohanty, S.G., 1994. On zero avoiding transition probabilities of an r -node tandem queue. A combinatorial approach. *J. Appl. Probab.* 30, 737–741.
- Böhm, W., Mohanty, S.G., 1997. On the Karlin–McGregor theorem and applications. *Ann. Appl. Probab.* 7 (2), 314–325.
- Karlin, S., McGregor, J.G., 1959. Coincidence probabilities. *Pacific J. Math.* 9, 1141–1164.
- Karlin, S., 1988. Coincident probabilities and applications in combinatorics. *J. Appl. Probab. A* 25, 185–200.
- Massey, W.A., 1984. An operator analytic approach to the Jackson network. *J. Appl. Probab.* 21, 379–393.
- Riordan, J., 1968. *Combinatorial Identities*. Wiley, New York.
- Stadje, W., 1995. The busy period of some queueing systems. *Stochastic Process. Appl.* 55, 159–167.
- Takács, L., 1967. *Combinatorial Methods in the Theory of Stochastic Processes*. Wiley, New York.
- Zacks, S., 1991. Distributions of stopping times for Poisson processes with linear boundaries. *Commun. Statist.-Stochastic Models* 7 (2), 233–242.