A Product-Form Solution for a Two-Class $Geo^{Geo}/D/1$ Queue with Random Routing and Randomly Alternating Service

Arnaud Devos(✉), Michiel De Muynck, Herwig Bruneel, and Joris Walraevens

Department of Telecommunications and Information Processing (EA07), Ghent University - UGent, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium
{Arnaud.Devos,MichielR.DeMuynck,Herwig.Bruneel,Joris.Walraevens}@ugent.be

Abstract. We analyze a discrete-time queueing system, consisting of two queues and a single server. The server randomly distributes its time between the two queues. Service times of any customer of either queue are deterministically equal to 1 time slot. In general, the joint analysis of such a two-queue system turns out to be very hard. In this paper, we assume that the total number of arrivals into the system constitutes a series of i.i.d. random variables with common geometric distribution. Each arriving customer is routed probabilistically to a queue. By means of a state-of-the-art approach, we obtain a closed-form expression of the steady-state joint PGF of the number of customers present ("system contents") in both queues, at the beginning of a random slot. We find that the joint PGF is of product form, which proves that the system contents in both queues are independent. We provide an additional intuitive stochastic explanation for this remarkable result. We discuss several model extensions using the stochastic analysis.

Keywords: Two-class queueing model · Non-work-conserving · Product-form solution

1 Introduction

Two-class queueing systems have received great attention in the queueing literature over the last decades. On the one hand, this is partially because these models can be used for many practical applications. On the other hand, the analysis of such a model usually provides formidable mathematical difficulties which are of interest in their own right. The latter is also the scope and focus of this work. In the most common two-class queueing models, the scheduling disciplines (e.g. global first-come-first-served [7], fixed priority and round robin) depend on the number of customers present in the system, typically to keep the system work-conserving. However, some multi-class queueing models with

E. Hyytiä and V. Kavitha (Eds.): VALUETOOLS 2022, LNICST 482, pp. 81–95, 2023.
https://doi.org/10.1007/978-3-031-31234-2_6

a non-work-conserving scheduler have been studied in the past, as well. See for example [4,8,11,16] and references therein. Besides possible practical motivations, a non-work-conserving scheduler usually has the (mathematical) benefit that the per-class queues can be analyzed relatively easily when viewed in isolation [8,16]. Unfortunately, a joint analysis of both per-class queues is still considered to be notoriously difficult, or even impossible, if one is interested in a closed-form solution [8,16]. Notice that a joint analysis is for instance necessary if one is interested in joint performance measures such as the correlation coefficient between the two queue contents, or the variance of the total queue content. In an earlier paper [9], we studied a discrete-time two-class model with the following non-work conserving scheduling discipline: at each service opportunity, the server is willing to serve queue 1 (resp. queue 2) with probability α (resp. $1-\alpha$). If the server chooses an empty queue, no-one gets service in that time slot. For this model, we were also confronted with the notorious problem of solving a functional equation [6], which we were unable to solve exactly. However, for the same service model, but for Bernoulli distributed numbers of class-1 and class-2 arrivals per time slot, we found a closed-form analysis in [10]. In [9], we assume that the numbers of per-slot class-1 and class-2 arrivals have a *generic* distribution. The surprising result of [10] has made us question if there exist other arrival processes such that a closed-form analysis is feasible. The answer is yes: we have in particular found a product-form solution for the case that the total numbers of arrivals to the queueing system during consecutive slots are independent geometrically distributed random variables and customers are probabilistically routed to one of the two queues (independently from customer to customer). This analysis is the subject of this paper.

The outline of the paper is as follows. In Sect. 2, the mathematical model is presented. Some preliminaries are given in Sect. 3. In Sect. 4, we analyze the joint system contents by using tools from complex analysis. Alternatively and complementarily, an intuitive stochastic analysis is presented in Sect. 5. Finally, we discuss several extensions for this model in Sect. 6.

2 Mathematical Model

We review the model of [9] in this section. We consider a discrete-time single-server system with two infinite waiting rooms. Time is assumed to be slotted, i.e. the time axis is divided into fixed-length intervals referred to as (time) slots. New customers may enter the system at an given (continuous) point on the time axis, but services can only start and end at slot boundaries. There are two types of customers arriving to the system, namely customers of class 1 (resp. class 2) joining queue 1 (resp. queue 2). We assume that the *total* numbers of newly generated arrivals during consecutive slots are independent and geometrically distributed with mean λ_T. More specifically,

$$\begin{aligned} a_T(n) &\triangleq \Pr[n \text{ customer arrivals in one slot}], \\ &= \frac{1}{1+\lambda_T}\left(\frac{\lambda_T}{1+\lambda_T}\right)^n, \quad n = 0, 1, \ldots . \end{aligned} \tag{1}$$

One easily verifies that this arrival process corresponds to a batch arrival process with geometrically distributed inter-arrival times and the batch sizes are geometrically distributed (see for example [2, Chapter 1]).

An arriving customer is routed to queue j with probability $\frac{\lambda_j}{\lambda_T}$, $j = 1, 2$. In particular, note that

$$\lambda_1 + \lambda_2 = \lambda_T. \tag{2}$$

A simple calculation shows that

$$\begin{aligned} a(n_1, n_2) &\triangleq \Pr[n_1 \text{ class-1 arrivals and } n_2 \text{ class-2 arrivals in one slot}] \\ &= \frac{1}{1+\lambda_T}\left(\frac{\lambda_T}{1+\lambda_T}\right)^{n_1+n_2}\binom{n_1+n_2}{n_1}\left(\frac{\lambda_1}{\lambda_T}\right)^{n_1}\left(\frac{\lambda_2}{\lambda_T}\right)^{n_2}. \end{aligned} \tag{3}$$

The joint probability generating function is then given by

$$\begin{aligned} A(z_1, z_2) &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} a(n_1, n_2) z_1^{n_1} z_2^{n_2} \\ &= \frac{1}{1+\lambda_1+\lambda_2-\lambda_1 z_1-\lambda_2 z_2}. \end{aligned} \tag{4}$$

The service times of customers equal one slot. At the beginning of every time slot, the single server randomly selects either queue 1 or queue 2 to serve. This selection occurs independently of the system state and is modeled by a single parameter α $(0 < \alpha < 1)$, that is defined as

$$\alpha = \Pr[\text{server is available to class-1 customers during a slot}].$$

This directly means that the server is available to class-2 customers during a slot with probability $1 - \alpha$. Moreover, it is assumed that the state of the server (available to either class-1 or class-2 customers) during a certain slot is independent of the state of the server during previous slots, and also of the other random variables present in the model. Notice that when the server is available to an empty queue, no service occurs in that slot, even when the other queue is non-empty. Hence, the system is not work-conserving.

Finally, since we are interested in a stochastic equilibrium, it is required that (see Equation (5) in [9])

$$\lambda_1 < \alpha \quad \text{and} \quad \lambda_2 < 1 - \alpha. \tag{5}$$

3 Preliminaries

In this section, we write down the fundamental functional equation for the joint PGF of the steady-state system contents. Next, we provide the marginal PGFs of the system contents, as these are readily obtained.

3.1 The Functional Equation

In order to analyze the above queueing model, let $u_{1,k}$ and $u_{2,k}$ denote the numbers of class-1 and class-2 customers in the system at the beginning of slot k. It is easy to see that $(u_{1,k}, u_{2,k})$ is a discrete-time Markov chain with state space

$$\mathcal{S} = \{0, 1, \ldots\}^2.$$

We assume the stationary distribution of this stochastic process exists, see also (5). Therefore, we define the limiting joint probability mass function (pmf)

$$p(n_1, n_2) \triangleq \lim_{k\to\infty} \Pr[u_{1,k} = n_1, u_{2,k} = n_2], \quad (n_1, n_2) \in \mathcal{S}. \tag{6}$$

The balance equations read as follows:

$$\begin{aligned} p(n_1, n_2) = \alpha &\sum_{j=0}^{n_2} a(n_1, n_2 - j)p(0, j) \\ &+ \alpha \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} a(n_1 - i, n_2 - j)p(i + 1, j) \\ &+ (1 - \alpha) \sum_{i=0}^{n_1} a(n_1 - i, n_2)p(i, 0) \\ &+ (1 - \alpha) \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} a(n_1 - i, n_2 - j)p(i, j + 1). \end{aligned} \tag{7}$$

Now let us define the joint PGF

$$U(z_1, z_2) \triangleq \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} p(n_1, n_2) z_1^{n_1} z_2^{n_2}. \tag{8}$$

It follows from Eq. (7) that

$$K(z_1, z_2)U(z_1, z_2) = A(z_1, z_2)[(1 - \alpha)(z_2 - 1)z_1 U(z_1, 0) + \alpha(z_1 - 1)z_2 U(0, z_2)], \tag{9}$$

with

$$K(z_1, z_2) = z_1 z_2 - [(1 - \alpha)z_1 + \alpha z_2]A(z_1, z_2). \tag{10}$$

Equation (9) is in accordance with the functional equation obtained in Section 2.1 in [9]. Recall that in this work we assume that $A(z_1, z_2)$ is given by (4), in contrast to [9] (where the exact form of $A(z_1, z_2)$ is not specified).

3.2 The Marginal PGFs

In this section, we define the pmf $p_i(n)$, $i = 1, 2$, of the number of class-i customers in the system at the beginning of a random slot in steady state, i.e.,

$$p_i(n) = \lim_{k\to\infty} \Pr[u_{i,k} = n] \qquad n = 0, 1, \ldots. \tag{11}$$

The marginal PGF $U_1(z)$ of the number of class-1 customers in the system at the beginning of a random slot in steady state can be easily deducted from (9) by choosing $\{z_1 = z, z_2 = 1\}$, which results in

$$U_1(z) \triangleq \sum_{n=0}^{\infty} p_1(n) z^n \tag{12}$$

$$= U(z, 1)$$

$$= \frac{\alpha - \lambda_1}{\alpha - \lambda_1 z}. \tag{13}$$

The reader will recognize the expression above as the PGF of a geometrically distributed random variable with parameter $\frac{\lambda_1}{\alpha}$. Hence the pmf $p_1(n)$ is given by

$$p_1(n) = \left(1 - \frac{\lambda_1}{\alpha}\right)\left(\frac{\lambda_1}{\alpha}\right)^n, \quad n = 0, 1, \dots . \tag{14}$$

It easily follows that the radius of convergence of (12), say τ_1, is equal to

$$\tau_1 = \frac{\alpha}{\lambda_1}. \tag{15}$$

For reasons of symmetry, we have that

$$U_2(z) \triangleq \sum_{n=0}^{\infty} p_2(n) z^n \tag{16}$$

$$= U(1, z)$$

$$= \frac{1 - \alpha - \lambda_2}{1 - \alpha - \lambda_2 z} \tag{17}$$

and

$$p_2(n) = \left(1 - \frac{\lambda_2}{1-\alpha}\right)\left(\frac{\lambda_2}{1-\alpha}\right)^n, \quad n = 0, 1, \dots . \tag{18}$$

The radius of convergence of (16), say τ_2, is given by

$$\tau_2 = \frac{1 - \alpha}{\lambda_2}. \tag{19}$$

4 Complex-Analytic Analysis

Two-class queueing models often give rise to the problem of solving a functional equation similar to (9). In the pioneering paper [12], a solution technique for such equations is demonstrated. It was shown that the solution can be found as the solution of a boundary-value problem for analytic functions. This technique is nowadays considered as the state-of-the art technique for dealing with such equations, and this technique goes nowadays by the name of the *boundary-value method*. A more detailed contribution of the analysis of such functional equations was later provided in [5,13]. The method that we employ in this work covers the same principle ideas as the one of the boundary-value method, which we now describe below.

In order to obtain the joint PGF $U(z_1, z_2)$ for arbitrary values (z_1, z_2) from (9), the partial PGFs $U(z_1, 0)$ and $U(0, z_2)$ have to be obtained, which is the non-straightforward objective of the analysis. It is worth mentioning that substitution of $\{z_1 = z, z_2 = 0\}$ or $\{z_1 = 0, z_2 = z\}$ into (9) always leads to the tautology "$0 = 0$". The crucial part of the analysis is studying the function K, which is referred to as the *kernel* of the functional equation. This is because whenever a zero $(\hat{z}_1, \hat{z}_2)$ of K lies inside the region of convergence of the PGF $U(z_1, z_2)$, this relates $U(\hat{z}_1, 0)$ with $U(0, \hat{z}_2)$. More precisely, for any $(\hat{z}_1, \hat{z}_2)$ which lies inside the region of convergence of $U(\cdot, \cdot)$ such that $K(\hat{z}_1, \hat{z}_2) = 0$ and $A(\hat{z}_1, \hat{z}_2) \neq 0$, we have that

$$(1 - \alpha)(\hat{z}_2 - 1)\hat{z}_1 U(\hat{z}_1, 0) + \alpha(\hat{z}_1 - 1)\hat{z}_2 U(0, \hat{z}_2) = 0,$$

which gives us an equation in terms of $U(\hat{z}_1, 0)$ and $U(0, \hat{z}_2)$.

Therefore, we will first discuss the regions of convergence of the PGFs of interest. Next, we investigate the kernel K. We emphasize that the function K has the same structure as the one in [15]. Therefore the analysis of the kernel K is comparable to the one in [15]. The difference between [15] and our work is that the RHS of the functional equation does not have the same structure. Hence, at some point in the analysis, a different approach is used (in comparison with [15]).

4.1 Regions of Convergence

The boundary function $U(\cdot, 0)$ is defined by

$$U(z, 0) = \sum_{n=0}^{\infty} p(n, 0) z^n, \tag{20}$$

the power series of the horizontal boundary probabilities. Similarly, the boundary function $U(0, \cdot)$ is defined by

$$U(0, z) = \sum_{n=0}^{\infty} p(0, n) z^n, \tag{21}$$

the power series of the vertical boundary probabilities. We now investigate for which values of z these two infinite series converge. To accomplish this, we observe that for every $n = 0, 1, \ldots$

$$\begin{aligned} p(n, 0) &\leq \sum_{j=0}^{\infty} p(n, j) \\ &= p_1(n). \end{aligned}$$

Hence, the radius of convergence of (12) is a lower bound for the radius of convergence of (20). Analogously, the radius of convergence of (16) is a lower bound for the radius of convergence of (21). Recall that we have determined the radius of convergence τ_1 (resp. τ_2) of (12) (resp. (16)) in the Sect. 3.2.

Next, we investigate the joint PGF $U(z_1, z_2)$. For any z_2 with modulus smaller than or equal to 1, we have

$$
\begin{aligned}
|U(z_1, z_2)| &\le \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p(i,j)|z_1|^i|z_2|^j \\
&\le \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} p(i,j)|z_1|^i \\
&= \sum_{i=0}^{\infty} p_1(i)|z_1|^i \\
&= U_1(|z_1|).
\end{aligned}
$$

From the fact that the radius of convergence of $U_1(z)$ is given by τ_1, we further obtain that

$$|U(z_1, z_2)| < \infty, \quad \text{if } |z_1| < \tau_1, |z_2| \le 1.$$

For reasons of symmetry, we can similarly prove that

$$|U(z_1, z_2)| < \infty, \quad \text{if } |z_1| \le 1, |z_2| < \tau_2.$$

In summary, we have shown that:

Lemma 1

(i) $U(z, 0)$ *converges absolutely for* $|z| < \tau_1$,
(ii) $U(0, z)$ *converges absolutely for* $|z| < \tau_2$,
(iii) $U(z_1, z_2)$ *converges absolutely for* $|z_1| < \tau_1$, $|z_2| \le 1$,
(iv) $U(z_1, z_2)$ *converges absolutely for* $|z_1| \le 1$, $|z_2| < \tau_2$.

4.2 Analysis of the Kernel $K(z_1, z_2)$

In this subsection, we investigate the zeros of the kernel K. We have that

$$
\begin{aligned}
&K(z_1, z_2) = 0 \\
\Leftrightarrow &\frac{-z_1^2 z_2\lambda_1 - z_1 z_2^2\lambda_2 + z_1 z_2\lambda_1 + z_1 z_2\lambda_2 + z_1\alpha - z_2\alpha + z_1 z_2 - z_1}{-\lambda_1 z_1 - \lambda_2 z_2 + \lambda_1 + \lambda_2 + 1} = 0 \\
\Leftrightarrow &- z_1^2 z_2\lambda_1 - z_1 z_2^2\lambda_2 + z_1 z_2\lambda_1 + z_1 z_2\lambda_2 + z_1\alpha - z_2\alpha + z_1 z_2 - z_1 = 0 \\
\Leftrightarrow &z_1 z_2\left(-z_1\lambda_1 - z_2\lambda_2 + \lambda_1 + \lambda_2 + \frac{\alpha}{z_2} - \frac{\alpha}{z_1} + 1 - \frac{1}{z_2}\right) = 0.
\end{aligned}
$$

The left-hand side in the equation above has the same form as the kernel that is studied in [15]. We now proceed as per [15] to analyze the solutions of this equation. Letting

$$H_1(z_1) = \alpha + \lambda_1 - \lambda_1 z_1 - \frac{\alpha}{z_1}, \tag{22}$$

$$H_2(z_2) = 1 - \alpha + \lambda_2 - \lambda_2 z_2 - \frac{1-\alpha}{z_2}, \tag{23}$$

we see that

$$H_1(z_1) + H(z_2) = 0 \Rightarrow K(z_1, z_2) = 0.$$

For complex values of z_1, we will only observe that for $z_1 = \sqrt{\frac{\alpha}{\lambda_1}} e^{\pm \mathrm{i}\theta}$, $H_1(z_1)$ is real and equal to

$$\alpha + \lambda_1 - \sqrt{\lambda_1 \alpha} 2 \cos(\theta).$$

Moreover, we have that

$$\alpha + \lambda_1 - \sqrt{\lambda_1 \alpha} 2 \cos(\theta) > 0, \tag{24}$$

since

$$\begin{aligned} \alpha + \lambda_1 - \sqrt{\lambda_1 \alpha} 2 \cos(\theta) &\geq \alpha + \lambda_1 - \sqrt{\lambda_1 \alpha} 2 \\ &= (\sqrt{\alpha} - \sqrt{\lambda_1})^2 \\ &> 0. \end{aligned}$$

We are ready to prove the following lemma.

Lemma 2. *For values $z_1 = \sqrt{\frac{\alpha}{\lambda_1}} e^{\pm \mathrm{i}\theta}$, there is a unique $z_2 =: y(z_1) \in]0,1[$ such that*

$$H_1(z_1) + H_2(y(z_1)) = 0.$$

Proof. For $x \in [0, 1]$, $H_2(x)$ increases monotonically from $-\infty$ at $x = 0$ to 0 at $x = 1$. Moreover, we have that

$$H_1\left(\sqrt{\frac{\alpha}{\lambda_1}} e^{\pm \mathrm{i}\theta}\right) > 0,$$

cf. (24). In particular, $H_2(x)$ is continuous in $]0,1[$. By virtue of the intermediate value theorem, there is a unique value z_2 in the interval $]0,1[$ such that

$$H_2(z_2) = -H_1\left(\sqrt{\frac{\alpha}{\lambda_1}} e^{\pm \mathrm{i}\theta}\right).$$

4.3 Analytic Continuation of $U(z,0)$ and $U(0,z)$

We can now proceed to determine the functions $U(z, 0)$ and $U(0, z)$ and hence solve the functional equation (9). To accomplish this, it is crucial to note that

$$y(z) = y(\bar{z}), \quad |z| = \sqrt{\frac{\alpha}{\lambda_1}},$$

with y defined in the previous subsection. The reason why the equality above holds, is simply because

$$\begin{aligned} H_1\left(\sqrt{\frac{\alpha}{\lambda_1}}e^{\mathrm{i}\theta}\right) &= \alpha + \lambda_1 - \sqrt{\lambda_1\alpha}2\cos(\theta) \\ &= \alpha + \lambda_1 - \sqrt{\lambda_1\alpha}2\cos(-\theta) \\ &= H_1\left(\sqrt{\frac{\alpha}{\lambda_1}}e^{-\mathrm{i}\theta}\right). \end{aligned}$$

Further, since $\frac{\alpha}{\lambda_1} > 1$ it obviously holds that

$$\sqrt{\frac{\alpha}{\lambda_1}} = \sqrt{\tau_1} < \tau_1$$

Due to the inequality above and the fact that $y(z) \in]0,1[$, we have that $U(z, y(z))$ remains finite by virtue of Lemma 1 on page 7. Hence, substituting $\{z_1 = z, z_2 = y(z)\}$, $|z|^2 = \frac{\alpha}{\lambda_1}$ into the functional equation (9) yields

$$(1-\alpha)(y(z)-1)zU(z,0) + \alpha(z-1)y(z)U(0,y(z)) = 0,$$

and substitution of $(\bar{z}, y(z))$ for values of z such that $|z|^2 = \frac{\alpha}{\lambda_1}$ yields

$$(1-\alpha)(y(z)-1)\bar{z}U(\bar{z},0) + \alpha(\bar{z}-1)y(z)U(0,y(z)) = 0.$$

Eliminating $U(0, y(z))$ gives us

$$(\bar{z}-1)zU(z,0) = (z-1)\bar{z}U(\bar{z},0), \quad |z| = \sqrt{\frac{\alpha}{\lambda_1}}. \tag{25}$$

Finally, when multiplying both sides of the relation above by z and using the relations $z\bar{z} = \frac{\alpha}{\lambda_1} \Leftrightarrow |z| = |\bar{z}| = \sqrt{\frac{\alpha}{\lambda_1}}$ we find that

$$z\left(\frac{\alpha}{\lambda_1} - z\right)U(z,0) = \frac{\alpha}{\lambda_1}(z-1)U\left(\frac{\alpha}{\lambda_1}z^{-1}, 0\right), \quad |z| = \sqrt{\frac{\alpha}{\lambda_1}}. \tag{26}$$

This functional equation for $U(\cdot, 0)$ can for instance be solved as follows. Let us substitute the series expression (20) into (26), we obtain that

$$z\left(\frac{\alpha}{\lambda_1} - z\right)\sum_{n=0}^{\infty} p(n,0)z^n = \frac{\alpha}{\lambda_1}(z-1)\sum_{n=0}^{\infty} p(n,0)\left(\frac{\alpha}{\lambda_1}\right)^n z^{-n},$$

which can be rewritten as

$$\begin{aligned} &\sum_{n=1}^{\infty}\frac{\alpha}{\lambda_1}p(n-1,0)z^n - \sum_{n=2}^{\infty}p(n-2,0)z^n - \sum_{n=0}^{\infty}\left(\frac{\alpha}{\lambda_1}\right)^{n+2}p(n+1,0)z^{-n} \\ &- \frac{\alpha}{\lambda_1}p(0,0)z + \sum_{n=0}^{\infty}\left(\frac{\alpha}{\lambda_1}\right)^{n+1}p(n,0)z^{-n} = 0. \end{aligned}$$

This equation is valid for z-values such that $|z| = \sqrt{\frac{\alpha}{\lambda_1}}$. Hence, by multiplying by appropriate powers of z and integrating over the positively oriented complex circle centered at 0 with radius $\sqrt{\tau_1}$, we may identify the coefficients of equal (positive or negative) powers of z of the above equation. We obtain that

$$p(n,0) = \frac{\lambda_1}{\alpha} p(n-1,0), \qquad n = 1, 2, \ldots.$$

Solving this elementary difference equation yields

$$p(n,0) = \left(\frac{\lambda_1}{\alpha}\right)^n p(0,0).$$

Finally, using the condition $\sum_{n=0}^{\infty} p(n,0) = p_2(0) = 1 - \frac{\lambda_2}{1-\alpha}$ gives us

$$p(0,0) = \frac{(\alpha - \lambda_1)(1 - \alpha - \lambda_2)}{\alpha(1-\alpha)}.$$

Hence, the probabilities $p(n,0)$ ($n = 0, 1, \ldots$) are completely determined. A simple calculation then shows that $U(z,0)$ is given by the following rational function

$$U(z,0) = \frac{\left(1 - \frac{\lambda_2}{1-\alpha}\right)(\alpha - \lambda_1)}{\alpha - \lambda_1 z}. \tag{27}$$

For reasons of symmetry, it can be shown that $U(0,z)$ is given by

$$U(0,z) = \frac{\left(1 - \frac{\lambda_1}{\alpha}\right)(1 - \alpha - \lambda_2)}{1 - \alpha - \lambda_2 z}. \tag{28}$$

The reader might have noticed that

$$U(z_1, 0) = U_1(z_1)U_2(0) \text{ and that } U(0, z_2) = U_1(0)U_2(z_2).$$

This suggests that the steady-state system contents are statistically independent.

4.4 The Joint PGF

If we substitute the expressions (27) and (28) for $U(z,0)$ and $U(0,z)$, respectively, into the functional equation for $U(z_1, z_2)$, we get

$$U(z_1, z_2) = \frac{(\alpha - \lambda_1)(1 - \alpha - \lambda_2)}{(\alpha - \lambda_1 z_1)(1 - \alpha - \lambda_2 z_2)}. \tag{29}$$

We can conclude that $U(z_1, z_2) = U(z_1, 1)U(1, z_2)$, i.e., the steady-state system contents are indeed statistically independent. We emphasize that this is a striking result. This result is primarily obtained through a complex-analytic analysis. In Sect. 5, we provide a more intuitive and stochastic analysis of the system contents to obtain the same result.

4.5 The Steady-State Joint Probabilities

The joint probability mass function $p(n_1, n_2)$ of the numbers of class-1 and class-2 customers is thus easily obtained because the joint pmf can be factorized as

$$p(n_1, n_2) = p_1(n_1)p_2(n_2).$$

We obtain that, using (14) and (18),

$$p(n_1, n_2) = \frac{(\alpha - \lambda_1)(1 - \alpha - \lambda_2)}{\alpha(1 - \alpha)} \left(\frac{\lambda_1}{\alpha}\right)^{n_1} \left(\frac{\lambda_2}{1 - \alpha}\right)^{n_2}. \tag{30}$$

5 Intuitive Stochastic Analysis

We now present an intuitive explanation of the mutual independence of the two system contents in the model under study. We have assumed that the total numbers of arrivals during consecutive slots constitute a sequence of i.i.d. geometric random variables. Observing the system contents from slot to slot thus implies that the system contents can increase to any level, since there is a non-zero probability of having $n \in \mathbb{N}$ arrivals during a slot. However, note that a geometric random variable can be viewed as the outcome of a sequence of independent and identical Bernoulli experiments. In this case, the two possible outcomes of the Bernoulli experiments are "a customer arrives" and "no customer arrives". Note that "no customer arrives" implies the end of arrivals in a slot. Because we assume that an arriving customer is randomly routed to a queue with fixed probabilities, independently from customer to customer, we thus have that the arrivals constitute a sequence of independent events. These events are the following

(i) A class-1 customer arrives;
(ii) A class-2 customer arrives;
(iii) End of arrivals in a slot.

The probability of having event (i), event (ii) or event (iii) (at a given point in time) is independent of the events that have occurred before. These three probabilities can be computed, but the exact expressions are not needed for the remainder of this section. Note that we have introduced an explicit sequence of arrivals in a slot, but that the total numbers of arrivals per class are given by the distribution (3). As a consequence, and because of the randomly alternating service discipline studied in this paper, changes of the *system contents* can be represented by a sequence of independent events, which are the following

1. A class-1 customer arrives;
2. A class-2 customer arrives;
3. A new slot begins and a class-1 customer is served (if any);
4. A new slot begins and a class-2 customer is served (if any).

Again, the probability of having event 1, event 2, event 3 or event 4 (at a given point in time) is independent of the events that have occurred before. These four probabilities can be computed, but the exact expressions are not needed for the remainder of this section. The evolution of the system contents can for instance be described by means of a sequence of numbers, taken from the set $\{1, 2, 3, 4\}$, where the numbers corresponds to the events as described above. For example, a slot with two class-1 arrivals, one class-2 arrival, followed by a slot with a class-1 service could be "represented" by the sequence 1123. Important to note is that we now assume that services in a slot happen before arrivals. However, it is not difficult to see that this again boils down to the model we studied earlier.

Let v_1 (resp. v_2) denote the number of class-1 customers (resp. class-2 customers) in the system (right after the *occurrence* of an event $\in \{1, 2, 3, 4\}$), when a stochastic equilibrium has been reached. In such an equilibrium, the system contents are represented by an **infinite** sequence, say ω, of numbers of the set $\{1, 2, 3, 4\}$. It is easy to see that the random variable v_1 is completely determined by the events 1 and 3, while v_2 is completely determined by the events 2 and 4: for given ω, only the numbers 1 and 3 (resp. 2 and 4) are sufficient and necessary to determine v_1 (resp. v_2). Now it can be seen that information of v_2 (and hence information about the occurrences of the events 2 and 4 in ω) gives us no information about v_1. Even knowing the exact order and amount of 2*s* and 4*s* in the infinite sequence ω provides us no information about the (number and order of) 1*s* and 3*s* in ω, and v_1 is solely determined by the latter. Similarly, information about v_1 (and hence information about the occurrences the events 1 and 3 in ω) gives us no information about v_2. From this reasoning, we may conclude that v_1 and v_2 are independent random variables. Finally, because the beginning of slots are just moments after specific events, the stationary system contents of class 1 and class 2 at the beginning of slots, u_1 and u_2, are also independent because of the BASTA property (Bernoulli-Arrivals-See-Time-Averages [14]).

6 Direct Extensions of the Queueing Model

Inspired by the intuitive explanation for the product-form result, we studied several extensions to our queueing model such that the system contents remain independent.

6.1 Multiple Arrivals

Consider the queueing model as described in Sect. 3.2, but with the following (more general) form for $A(z_1, z_2)$:

$$A(z_1, z_2) = \frac{1}{1 + \mu_1 + \mu_2 - \mu_1 L_1(z_1) - \mu_2 L_2(z_2)}, \tag{31}$$

with $L_1(z)$ and $L_2(z)$ being PGFs, but not specified. Note that the case where

$$L_1(z) = L_2(z) = z$$

corresponds to the arrival PGF (4). We can apply the same reasoning as per Sect. 5. The only difference is that the events 1 and 2 should be replaced by

1. *A number* L_1 of class-1 customer arrive;
2. *A number* L_2 of class-2 customer arrive.

The number of arriving customers L_1 (resp. L_2) is distributed according to the PGF $L_1(z)$ (resp. $L_2(z)$). Because the events are still independent, we can again conclude that the steady-state system contents at the beginning of a random slot are independent, which implies that $U(z_1, z_2) = U(z_1, 1)U(1, z_2)$. Hence,

$$U(z_1, z_2) = \frac{(\alpha - \lambda_1)(1 - \alpha - \lambda_2)(z_1 - 1)(z_2 - 1)A_1(z_1)A_2(z_2)}{(z_1 - A_1(z_1)[\alpha + (1 - \alpha)z_1])(z_2 - A_2(z_2)[1 - \alpha + \alpha z_2])}, \tag{32}$$

with

$$A_i(z) = \frac{1}{1 + \mu_i - \mu_i L_i(z)}, \quad i = 1, 2,$$

the PGF of the number of class-i arrivals in a slot. It can be checked that this solution indeed satisfies the functional equation (9).

6.2 Multiple Servers

Consider the queueing model as described in Sect. 3.2. Instead of serving one class-i customer during a slot (if any), we now assume that the number of available servers for class-i customers (if class i is chosen to be served) is distributed according to a distribution with PGF $S_i(z)$, $i = 1, 2$. Note that the case where

$$S_1(z) = S_2(z) = z$$

corresponds to the original model. We can apply the same reasoning as per Sect. 5. The only difference is that the events 3 and 4 should be replaced by

3. A new slot begins and a *number* of class-1 customers are served (if any);
4. A new slot begins and a *number* of class-2 customers are served (if any).

The maximum number of served customers of class 1 (resp. class 2) is distributed according to the PGF $S_1(z)$ (resp. $S_2(z)$). Because the events are still independent, we can again conclude that the system contents are independent. We emphasize that the analysis of the marginal system contents for this model is more difficult as compared to the analysis of Sect. 3.2, since a queue in isolation is now a multi-server queue with a variable number of (available) servers. However, several special cases have been studied, such as the cases where the number of available servers never exceeds a certain maximum [1, Section 5], or is geometrically distributed [3].

6.3 Multiple Queues

Consider the queueing model as described in Sect. 3.2, but with an arbitrary number N different customer classes, each with their own queue. The single server is allocated to queue j with probability α_j, $j = 1, \ldots, N$ and $\sum_{j=1}^{N} \alpha_j = 1$. As in Sect. 3.2, we assume that there is a single arrival stream of customers to the system, described by means of a sequence of i.i.d. geometric random variables during the consecutive slots (1). An arriving customer is of class j with probability $\frac{\lambda_j}{\lambda_T}$, $j = 1, 2, \ldots, N$. Each queue is stable as long as $\lambda_j < \alpha_j$.

It is not difficult to see that we may extend the reasoning as in Sect. 5 to multiple queues. We may thus conclude again that the N steady-state system contents are independent at the beginning of a random slot. To the best of our knowledge, this result is one of the few explicit results that exist for the steady-state joint analysis of more than two queues.

Obviously, combining some/all the extensions leads to a product-form solution as well.

References

1. Bruneel, H.: A general model for the behaviour of infinite buffers with periodic service opportunities. Eur. J. Oper. Res. **16**(1), 98–106 (1984)
2. Bruneel, H., Kim, B.G.: Discrete-Time Models for Communication Systems Including ATM. Kluwer Academic Publisher, Boston (1993)
3. Bruneel, H., Wittevrongel, S., Claeys, D., Walraevens, J.: Discrete-time queues with variable service capacity: a basic model and its analysis. Ann. Oper. Res. **239**(2), 359–380 (2016). https://doi.org/10.1007/s10479-013-1428-y
4. Al Hanbali, A., de Haan, R., Boucherie, R.J., van Ommeren, J.: Time-limited polling systems with batch arrivals and phase-type service times. Ann. Oper. Res. **198**(1), 57–82 (2012). https://doi.org/10.1007/s10479-011-0846-y
5. Cohen, J.W., Boxma, O.J.: Boundary Value Problems in Queueing System Analysis. North-Holland, Amsterdam (1983)
6. Cohen, J.W.: Boundary value problems in queueing theory. Queueing Syst. **3**(2), 97–128 (1988). https://doi.org/10.1007/BF01189045
7. De Clercq, S., Laevens, K., Steyaert, B., Bruneel, H.: A multi-class discrete-time queueing system under the FCFS service discipline. Ann. Oper. Res. **202**(1), 59–73 (2013). https://doi.org/10.1007/s10479-011-1051-8
8. de Haan, R., Boucherie, R.J., van Ommeren, J.: A polling model with an autonomous server. Queueing Syst. **62**(3), 279–308 (2009). https://doi.org/10.1007/s11134-009-9131-z
9. Devos, A., Walraevens, J., Fiems, D., Bruneel, H.: Approximations for the performance evaluation of a discrete-time two-class queue with an alternating service discipline. Ann. Oper. Res. **310**(2), 477–503 (2022). https://doi.org/10.1007/s10479-020-03776-5
10. Devos, A., Walraevens, J., Fiems, D., Bruneel, H.: Analysis of a discrete-time two-class randomly alternating service model with Bernoulli arrivals. Queueing Syst. **96**, 133–152 (2020). https://doi.org/10.1007/s11134-020-09663-x
11. Dvir, N., Hassin, R., Yechiali, U.: Strategic behaviour in a tandem queue with alternating server. Queueing Syst. **96**, 205–244 (2020). https://doi.org/10.1007/s11134-020-09665-9

12. Fayolle, G., Iasnogorodski, R.: Two coupled processors: the reduction to a Riemann-Hilbert problem. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **47**(3), 325–351 (1979). https://doi.org/10.1007/BF00535168
13. Fayolle, G., Malyshev, V.A., Iasnogorodski, R.: Random Walks in the Quarter-Plane. Springer, Heidelberg (1999). https://doi.org/10.1007/978-3-642-60001-2
14. Halfin, S.: Batch delays versus customer delays. Bell Syst. Tech. J. **62**(7), 2011–2015 (1983)
15. Konheim, A.G., Meilijson, I., Melkman, A.: Processor-sharing of two parallel lines. J. Appl. Probab. **18**(4), 952–956 (1981)
16. Saxena, M., Boxma, O., Kapodistria, S., Queija, R.: Two queues with random time-limited polling. Probab. Math. Stat. **37**(2), 257–289 (2017)