

THE STATIONARY G/G/s QUEUE

PIERRE LE GALL
France Telecom, CNET
4 Parc de la Bérengère
F-92210 Saint-Cloud, France

(Received November, 1997; Revised February, 1998)

The distribution of the queueing delay in the stationary G/G/s queue is given with an application to the GI/G/s queue and to the M/G/s queue.

Key words: G/G/s Queue, GI/G/s Queue, First Come-First Served, Factorization, Singular Points.

AMS subject classifications: 60K25, 90B22.

1. Introduction

In this paper we evaluate the distribution of the queueing delay in the stationary G/G/s queue. The particular case of the GI/G/s queue was extensively studied earlier by Pollaczek [3]: it proves difficult to write down the equations for the GI/G/s queue, whose partial solution can only be derived after long and complex calculations involving multiple contour integrals in a multi-dimensional complex plane.

In Section 2, we state the underlying assumptions and introduce notation before evaluating all singularities of the Laplace-Stieltjes transform of this distribution for the GI/G/s queue for the limited case of the stationary regime. The method is then extended to the case of the G/G/s queue (Section 3).

In Section 4, we derive the constraints to be satisfied in order that this Laplace-Stieltjes transform be holomorphic. To do so we propose a *factorization method*, which is more general than the Wiener-Hopf type decomposition.

In Section 5, we derive an expression for the distribution of the queueing delay in the stationary G/G/s queue together with its asymptotic behavior for long delays. In Section 6, we apply our results to the case of the GI/G/s queue. Finally, in Section 7, we consider the M/G/s queue.

The method presented requires relatively simple calculations making it possible to consider the evaluation of local queueing delays in multiserver queueing networks.

2. Notation, Assumptions and Preliminary Results

2.1 The Stationary G/G/s Queue

We consider a queue handled by a multiserver of s *identical servers*.

a) The Arrival Process: We assume a metrically transitive, strictly stationary process of successive, non-negative interarrival times. Let $\mathbf{N}(t)$ denote the random number of arrivals in the interval $(0, t]$. We write $d\mathbf{N}(t)=\mathbf{1}$ or $\mathbf{0}$ dependent on whether or not there is an arrival in the infinitesimal time interval $(t, t + dt)$. We exclude the possibility of simultaneous arrivals. We can then write:

$$E[d\mathbf{N}(t_0) \cdot d\mathbf{N}(t_0 + t)] = E[d\mathbf{N}(t_0)] \cdot \rho(t)dt, \quad (1)$$

where $\rho(t)$ is the arrival rate at time $(t + t_0)$ if an arbitrary arrival occurred at time t_0 . We let, for $R_\rho(z) < 0$:

$$\int_0^\infty e^{zt} \cdot \rho(t)dt = \alpha_1(z) = \sum_{x=1}^\infty \varphi_{0,x}(z), \quad (2)$$

where $\varphi_{0,x}(z)$ corresponds to the x^{th} arrival following the epoch t_0 . However, the stationary assumption and Abelian theorem gives: $\lim_{z \rightarrow 0^+} z \cdot \alpha_1(z) = \Lambda$, where Λ is the mean arrival rate. In a more general way, we may write, for $j = 1, 2, \dots$:

$$\begin{aligned} E[d\mathbf{N}(t_0) \cdot d\mathbf{N}(t_0 + t_1) \dots d\mathbf{N}(t_0 + t_1 + \dots + t_j)] \\ = [E d\mathbf{N}(t_0)] \cdot f_j(t_1 \dots t_j) \cdot dt_1 \dots dt_j, \end{aligned} \quad (3)$$

and for $R_\rho(z_j) < 0$, $j = 1, 2, \dots$:

$$\int_0^\infty e^{z_1 t_1} \cdot dt_1 \dots \int_0^\infty e^{z_j t_j} \cdot dt_j \cdot f_j(t_1 \dots t_j) = \alpha_j(z_1 \dots z_j). \quad (4)$$

In the case of a *renewal process*, the successive arrival intervals \mathbf{Y}_n are *mutually independent* and identically distributed and we let: $\varphi_0(z) = E e^{-zY}$ for $R(z) < 0$. Expression (2) becomes:

$$\alpha_1(z) = \frac{\varphi_0(z)}{1 - \varphi_0(z)}, \quad (5)$$

and (4) becomes:

$$\alpha_j(z_1 \dots z_j) = \alpha_1(z_1) \dots \alpha_1(z_j). \quad (6)$$

b) The Service Times: The successive service times \mathbf{T}_n are *mutually independent* and independent of the arrival process. They are identically distributed in accordance with a distribution function $F_1(t)$ and we let $\varphi_1(z) = E e^{-zT}$, for $R_\rho(z) < 0$. We exclude the possibility of bulk service; consequently:

$$F_1(0) = F_1(+0) = 0. \quad (7)$$

c) The Service Discipline: The service discipline is “*first come-first served*”, and the s servers are *indistinguishable*.

d) Traffic Intensity: The traffic intensity (per server) is supposed to be less than one:

$$\eta = \frac{[EdN(t)] \cdot [E(T_n)]}{s} < 1. \quad (8)$$

Under this condition, Loynes [1] demonstrated the existence of the stationary regime.

e) Queuing Delay: Let τ_n denote the queuing delay of the n^{th} customer, and τ of an arbitrary customer. *Note:* Since the term “waiting time” means “sojourn time” in Little’s formula, for clarity we prefer to use the term “queuing delay” for the queuing process only.

f) Contour Integrals: In this paper, we use (Cauchy) contour integrals along the imaginary axis in the complex plane. If the contour (followed by the bottom to the top) is to the right of the imaginary axis (the contour is closed at infinity to the right), we write \int_{+0} . If the contour is to the left of the imaginary axis, we write \int_{-0} . Unless it is necessary to specify whether the contour is to the right or to the left of the imaginary axis, we write \int_0 .

2.2 Preliminary Results (GI/G/1 Queue)

It will be useful to refer to Pollaczek [2] for the queue GI/G/1. For $\text{Re}(q) \geq 0$, we have:

$$\left\{ \begin{array}{l} Ee^{-q\tau} = \text{Exp}\left\{\frac{-1}{2\pi i} \int_{+0} \left\{\frac{1}{q+\zeta_1} - \frac{1}{\zeta_1}\right\} \cdot \log N_0(\zeta_1) \cdot d\zeta_1\right\}, \\ \text{with: } N_0(\zeta_1) = 1 - \varphi_0(-\zeta_1) \cdot \varphi_1(\zeta_1). \end{array} \right. \quad (9)$$

Note that:

$$\text{Exp}\left\{\frac{-1}{2\pi i} \int_{+0} \left[\frac{1}{q+\zeta_1} - \frac{1}{\zeta_1}\right] \cdot \log\left(\frac{1}{1-\varphi_0(-\zeta_1)}\right) \cdot d\zeta_1\right\} = 1, \quad (9a)$$

since we have $|\varphi_0(-\zeta_1)| < 1$ for $\text{Re}(\zeta_1) > 0$, and, consequently, there are no singular points in this region. We then multiply (9a) by (9), which leads to the substitution:

$$N_0(\zeta_1) \rightarrow N_1(\zeta_1) = \frac{N_0(\zeta_1)}{1-\varphi_0(-\zeta_1)} = 1 - \alpha_1(-\zeta_1) \cdot [\varphi_1(\zeta_1) - 1], \quad (10)$$

where $\alpha(-\zeta_1)$ is defined by (5). Finally, the functional $Ee^{-q\tau}$, defined by (9), may be written, for $\text{Re}(q) \geq 0$, as

$$\left\{ \begin{array}{l} \Phi^+(q) = Ee^{-q\tau} = \text{Exp}\left[\frac{-1}{2\pi i} \int_{+0} \left\{\frac{1}{q+\zeta_1} - \frac{1}{\zeta_1}\right\} \cdot \log N_1(\zeta_1) \cdot d\zeta_1\right], \\ \text{with: } N_1(\zeta_1) = 1 - [\zeta_1 \cdot \alpha_1(-\zeta_1)] \cdot \left(\frac{\varphi_1(\zeta_1) - 1}{\zeta_1}\right) = 1 - \alpha_1(-\zeta_1) \cdot [\varphi_1(\zeta_1) - 1]. \end{array} \right. \quad (11)$$

To find the singular points of this expression, we cross the poles $\zeta_1 = 0$ and $\zeta_1 = -q$ in the integrand, and have $\Phi^-(q)$ denote new Expression (11), for $\text{Re}(\zeta_1) < 0$ and $\text{Re}(\zeta_1) < -q$. We get the following Wiener-Hopf type of decomposi-

tion for $\text{Re}(q) < 0$, where $\Phi^-(q)$ is holomorphic for $\text{Re}(q) < 0$, and

$$\Phi^+(q) = (1 - \eta) \cdot \frac{\Phi^-(q)}{N_1(-q)}. \quad (12)$$

Here, η denotes the traffic intensity and the probability of the delay. The roots of the denominator define the singular points of $Ee^{-q\tau}$. Conversely, from (12) we may deduce Expression (11). We intend to extend this procedure to the case $s > 1$ by initially defining the singular points, and then defining the kind of decomposition.

3. The Singular Points [for $\text{Re}(q) < 0$]

From Expressions (10) and (12), the singular points of $Ee^{-q\tau}$ in the case of the stationary GI/G/1 queue are also the singular points of

$$\frac{1}{N_1(-q)} = \frac{1 - \varphi_0(q)}{1 - \varphi_0(q) \cdot \varphi_1(-q)} = [1 - \varphi_0(q)] \cdot \sum_{n=0}^{\infty} [\varphi_0(q)]^n \cdot [\varphi_1(-q)]^n. \quad (13)$$

The latter may be rewritten as:

$$\frac{1}{N_1(-q)} = 1 - \frac{1}{2\pi i} \cdot \int_{+0} \frac{dz_1}{z_1} \cdot \frac{q}{q + z_1} \cdot \frac{1 - \varphi_0(q)}{1 - \varphi_0(q) \cdot \varphi_1(z_1)}. \quad (14)$$

In the stationary regime, the distribution of the queueing delay is independent of the initial conditions. For instance, in the case of a large number of arrivals prior to time zero, the busy period is then very long. The terms $\sum_{\lambda=1}^n (\mathbf{T}_\lambda - \mathbf{Y}_\lambda)$ and $[\varphi_0(q)]^n \cdot [\varphi_1(-q)]^n$ serve to evaluate the queueing delay of the n^{th} customer following the earlier arrivals.

With the same initial conditions as before, the only change for the GI/G/s queue is to note that the sequence t_n of successive terminations of service times is now defined by s strings j of successive service times $\mathbf{T}_\lambda(j)$, $j = 1, \dots, s$ in parallel and by the expression

$$t_n = \mathbf{Min} \left[\sum_{\lambda_1=0}^{n_1} \mathbf{T}_{\lambda_1}(1), \dots, \sum_{\lambda_s=0}^{n_s} \mathbf{T}_{\lambda_s}(s) \right], \quad n_1 + \dots + n_s = n, \quad (15)$$

where all possible groups $(n_1 \dots n_s)$ are considered. To evaluate this sequence we will use the following expression given by Pollaczek [3]:

$$\left\{ \begin{array}{l} \exp[-q \cdot \min^+(a_1 \dots a_s)] - 1 = \frac{1}{(2\pi i)^s} \int_{+0} \frac{dz_1}{z_1} \dots \int_{+0} \frac{dz_s}{z_s} \cdot \frac{q}{q + \sum_{\nu=1}^s z_\nu} \exp\left(\sum_{\nu=1}^s a_\nu \cdot z_\nu\right), \\ \text{with: } \text{Re}\left(q + \sum_{\nu=1}^s z_\nu\right) > 0, \end{array} \right. \quad (16)$$

where: $\min^+(a_1 \dots a_s) = \mathbf{Max}[0, \min(a_1 \dots a_s)]$. Consequently, for the stationary GI/G/s queue, (14) becomes:

$$\frac{1}{N_s(-q)} = 1 - \frac{1}{(2\pi i)^s} \cdot \int_{+0} \frac{dz_1}{z_1} \cdots \int_{+0} \frac{dz_s}{z_s} \cdot \frac{q}{q + \sum_{\nu=1}^s z_\nu} \prod_{j=1}^s \frac{1 - \varphi_0(q)}{1 - \varphi_0(q) \cdot \varphi_1(z_j)}, \quad (17)$$

which may be written, due to (5), as

$$\frac{1}{N_s(-q)} = 1 - \frac{1}{(2\pi i)^s} \cdot \int_{+0} \frac{dz_1}{z_1} \cdots \int_{+0} \frac{dz_s}{z_s} \cdot \frac{q}{q + \sum_{\nu=1}^s z_\nu} \prod_{j=1}^s \frac{1}{1 - \alpha_1(q) \cdot [\varphi_1(z_j) - 1]}. \quad (18)$$

Note that the transformation from (14) to (18) depends on the arrival process through $[\alpha_1(q)]^{n_1} \dots [\alpha_1(q)]^{n_s}$. For stationary G/G/s queue (with the same initial conditions) we have to make the following substitution, due to (6):

$$[\alpha_1(q)]^{n_1} \dots [\alpha_1(q)]^{n_s} \rightarrow \alpha_n(q \dots q), \quad n_1 + \dots + n_s = \mathbf{n}.$$

More simply, it may be noted that (11) depends on $\{1/N_1(\zeta_1)\}$ through $-\log N_1(\zeta_1) = \log[1/N_1(\zeta_1)]$. We will see that this is also true for the stationary G/G/s queue [see (33) below]. Consequently, it is sufficient to consider the following substitution:

$$\prod_{j=1}^s [1 - \alpha_1(q) \cdot (\varphi_1(z_j) - 1)] = 1 + \sum_{\lambda=0}^{s-1} \binom{s}{\lambda} \cdot (-1)^{s-\lambda} \cdot [\alpha_1(q)]^{s-\lambda} \prod_{j=\lambda+1}^s [\varphi_1(z_j) - 1] \quad (19)$$

$$\rightarrow \mathbf{R}_s(z_1 \dots z_s; q) = 1 + \sum_{\lambda=0}^{s-1} \binom{s}{\lambda} \cdot (-1)^{s-\lambda} \cdot \alpha_{s-\lambda}(q \dots q) \cdot \prod_{j=\lambda+1}^s [\varphi_1(z_j) - 1].$$

This substitution is very simple. It is due also to a typical property of the arrival point processes (with non-simultaneous arrivals); the logarithm of the characteristic functional with respect to $N(t)$ allows to replace all the higher moments (3) by a single integration of $E[N(t)]$. Finally, we can state:

Theorem 1: (Singular points) *For the stationary G/G/s queue, the singular points of $Ee^{-q\tau}$, for $Re(q) < 0$, are those of the function:*

$$G_s(q) = 1 - \frac{1}{(2\pi i)^s} \cdot \int_{+0} \frac{dz_1}{z_1} \cdots \int_{+0} \frac{dz_s}{z_s} \cdot \frac{q}{q + \sum_{\nu=1}^s z_\nu} \cdot \frac{1}{\mathbf{R}_s(z_1 \dots z_s, q)}, \quad (20)$$

with:

$$\mathbf{R}_s(z_1 \dots z_s; q) = 1 + \sum_{\lambda=0}^{s-1} \binom{s}{\lambda} \cdot (-1)^{s-\lambda} \cdot \alpha_{s-\lambda}(q \dots q) \cdot \prod_{j=\lambda+1}^s [\varphi_1(z_j) - 1],$$

and:

$$Re(q + \sum_{\nu=1}^s z_\nu) > 0,$$

where $\alpha_{s-\lambda}(q \dots q)$ is defined by (4) for $Re(q) < 0$:

$$\int_0^\infty e^{qt\lambda+1} \cdot dt_{\lambda+1} \dots \int_0^\infty e^{qt_s} \cdot dt_s \cdot f_{s-\lambda}(t_{\lambda+1} \dots t_s).$$

Corollary 1: For the stationary GI/G/s queue, (6) gives $\alpha_{s-\lambda}(q..q) = [\alpha_1(q)]^{s-\lambda}$ and (18) gives:

$$G_s(q) = 1 - \frac{1}{(2\pi i)^s} \cdot \int_{+0} \frac{dz_1}{z_1} \cdots \int_{+0} \frac{dz_s}{z_s} \cdot \frac{q}{q + \sum_{\nu=1}^s z_\nu} \cdot \prod_{j=1}^s \frac{1}{1 - \alpha_1(q) \cdot [\varphi_1(z_j) - 1]}. \quad (20a)$$

4. The Factorization

We want to get a holomorphic function for $\operatorname{Re}(q) \geq 0$, with singular points ($\operatorname{Re}(q) < 0$) as defined by (20). We will be obliged to introduce some auxiliary complex variables z_i ($i = 1..s$). Finally, we want to define a function $U_s(z_1..z_s; q)$, holomorphic for $\operatorname{Re}(z_i) \geq 0$ (i.e., $\geq -\delta$), $i = 1..s$, and $\operatorname{Re}(q) \geq 0$, such as the singular points of $U_s(0..0; q)$ are defined by (20).

Consider the following integral, for $\operatorname{Re}(q + \sum_{\nu=1}^s z_\nu) > 0$:

$$I(q) = \frac{1}{(2\pi i)^s} \cdot \int_{+0} \frac{dz_1}{z_1} \cdots \int_{+0} \frac{dz_s}{z_s} \cdot \frac{q}{q + \sum_{\nu=1}^s z_\nu} \cdot U_s(z_1..z_s; q). \quad (21)$$

With our assumption for U_s , the integrand is holomorphic for $\operatorname{R}(z_j) > 0$, $j = 1..s$. Therefore, we have: $I(q) \equiv 0$. But, if we cross over poles $z_j = 0$ ($j = 1..s$) from the left hand side of the right-hand side, the residue of the integrand is: $(-1)^s \cdot U_s(0..0; q)$. We therefore deduce that

$$U_s(0..0; q) = \frac{(-1)^s}{(2\pi i)^s} \cdot \int_{-0} \frac{dz_1}{z_1} \cdots \int_{-0} \frac{dz_s}{z_s} \cdot \frac{q}{q + \sum_{\nu=1}^s z_\nu} U_s(z_1..z_s; q). \quad (22)$$

In this integral, q is not a variable; rather it is only a parameter. Introduce the following factorization:

Factorization: We set

$$U_s(z_1..z_s; q) = \frac{(-1)^s \cdot \alpha_s(q..q) \cdot \prod_{j=1}^s [\varphi_1(z_j) - 1]}{R_s(z_1..z_s, q)} \prod_{i=1}^s M_i(z_1..z_s; q), \quad (23)$$

where R_s is defined by (20) and U_s is holomorphic for $\operatorname{Re}(z_i) \geq 0$ ($i = 1..s$) and $\operatorname{Re}(q) \geq 0$, with the following conditions for M_i :

- M_i is holomorphic for $\operatorname{Re}(z_i) < 0$, $i = 1..s$;
- $M_i(z_1..z_{i-1}, -q - \sum_{\nu=1}^{i-1} z_\nu, z_{i+1}..z_s; q) \equiv 1$.

Now, we want to evaluate the integral (22) in the region $\operatorname{Re}(z_i) < 0$, $i = 1..s$. In this region, the equation $R_s(z_1..z_s; q) = 0$ has no root because of the inequality $|\varphi_1(z_j)| < 1$; thus the product

$$(-1)^{s-\lambda} \prod_{j=\lambda+1}^s [\varphi_1(z_j) - 1],$$

in (20), always has a positive real part. Consequently, it is the same for R_s .

When we integrate in complex plane z_s , M_s has no singularities for $\operatorname{R}(z_s) < 0$ due

to *condition* (a). In this plane, we find only the pole: $z_s = -q - \sum_{\nu=1}^{s-1} z_\nu$. To evaluate the residue, we have to apply *condition* (b): as a result, \mathbf{M}_s disappears. This will be the same for the integration in successive planes z_{s-1}, \dots, z_1 . Finally, (22) becomes:

$$\mathbf{U}_s(0 \dots 0; q) = \frac{1}{(2\pi i)^s} \int_{+0} \frac{dz_1}{z_1} \dots \int_{+0} \frac{dz_s}{z_s} \cdot \frac{q}{q + \sum_{\nu=1}^s z_\nu} \cdot \frac{\alpha_s(q \dots q) \cdot \prod_{j=1}^s [\varphi_1(z_j) - 1]}{\mathbf{R}_s(z_1 \dots z_s, q)},$$

where $z_j = 0$ ($j = 1 \dots s$) is not a pole. To simplify the integrand, firstly we consider the GI/G/s queue. We write, due to the symmetry with respect to variables z_j :

$$\begin{aligned} \alpha_s(q \dots q) \frac{\prod_{j=1}^s [\varphi_1(z_j) - 1]}{\mathbf{R}_s(z_1 \dots z_s, q)} &= \prod_{j=1}^s \left(\frac{\alpha_1(a) \cdot [\varphi_1(z_j) - 1]}{1 - \alpha_1(q) \cdot [\varphi_1(z_j) - 1]} \right) \\ &= (-1)^s \cdot \prod_{j=1}^s \left(1 - \frac{1}{1 - \alpha_1(q) \cdot [\varphi_1(z_j) - 1]} \right) \\ &= (-1)^s \cdot \left[1 + \sum_{\lambda=0}^{s-1} \binom{s}{\lambda} \cdot (-1)^{s-\lambda} \frac{1}{\prod_{j=\lambda+1}^s [1 - \alpha_1(q) \cdot [\varphi_1(z_j) - 1]]} \right]. \end{aligned}$$

For the G/G/s queue, we may apply the same reasoning which led to substitution (19) and used the following substitution, corresponding to formula (6): $[\alpha_1(q)]^{s-\lambda} \rightarrow \alpha_{s-\lambda}(q \dots q)$. Under the integrand we may write:

$$\begin{aligned} &\frac{\alpha_s(q \dots q) \cdot \prod_{j=1}^s [\varphi_1(z_j) - 1]}{\mathbf{R}_s(z_1 \dots z_s; q)} \\ &= (-1)^s \cdot \left[1 + \sum_{\lambda=0}^{s-1} \binom{s}{\lambda} \cdot (-1)^{s-\lambda} \cdot \frac{1}{\mathbf{R}_{s-\lambda}(z_{\lambda+1} \dots z_s; q)} \right]. \end{aligned}$$

Between brackets, term 1 and terms $\lambda (> 0)$, where at least one z_j is missing, do not contribute to the integral [for $\text{Re}(z_j) > 0$]. The expression of $\mathbf{U}_s(0 \dots 0; q)$ becomes with the only term $\lambda = 0$:

$$\mathbf{U}_s(0 \dots 0; q) = \frac{1}{(2\pi i)^s} \cdot \int_{+0} \frac{dz_1}{z_1} \dots \int_{+0} \frac{dz_s}{z_s} \cdot \frac{q}{q + \sum_{\nu=1}^s z_\nu} \cdot \frac{1}{\mathbf{R}_s(z_1 \dots z_s; q)}. \tag{24}$$

Let

$$\mathbf{V}_s(z_1 \dots z_s; q) = 1 - U_s(z_1 \dots z_s; q). \tag{25}$$

We can finally write, for $\text{Re}(q) \geq 0$:

$$V_s(0 \dots 0; q) = G_s(q), \tag{26}$$

where $G_s(q)$ is Expression (20) giving the singular points of $Ee^{-q\tau}$. To summarize:

The holomorphic function $V_s(z_1 \dots z_s; q)$, defined by (25) and by factorization, in (23), with conditions (a) and (b), leads to (26) giving the singular points of $Ee^{-q\tau}$ for the stationary G/G/s queue.

Note: 1) For $s = 1$, the factorization, in (23), is of the Wiener-Hopf type:

$$-U_1(z_1; q) = -\frac{R_1(z_1; q) - 1}{R_1(z_1; q)} M_1(z_1; q) \rightarrow \frac{M_1(z_1; q)}{R_1(z_1; q)}.$$

2) For $s > 1$, it is not sufficient to define $G_s(q)$ and the singular points; to define $Ee^{-q\tau}$, it is also necessary to use the factorization, in (23), different from the Wiener-Hopf type.

5. The Queuing Delay (Stationary G/G/s Queue)

5.1 The Distribution

We introduce the following holomorphic function for $Re(z_j) \geq 0$ ($j = 1 \dots s$) and $R(q) \geq 0$:

$$\left\{ \begin{array}{l} V_s(z_1 \dots z_s; q) = 1 - U_s(z_1 \dots z_s; q) \\ = 1 - \text{Exp} \left[\frac{1}{(2\pi i)^s} \int_{-0} \left[\frac{1}{q + \zeta_1} + \frac{1}{z_1 - \zeta_1} \right] \right. \\ \left. \cdot d\zeta_1 \dots \int_{-0} \left[\frac{1}{q + \sum_{\nu=1}^{s-1} z_\nu + \zeta_s} + \frac{1}{z_s - \zeta_s} \right] \cdot d\zeta_s \cdot \log N_s(\zeta_1 \dots \zeta_s), \right. \end{array} \right. \quad (27)$$

with $Re(q + \sum_{\nu=1}^{i-1} z_\nu + \zeta_i) > 0$, $i = 1 \dots s$, and

$$\frac{1}{N_s(z_1 \dots z_s)} = \frac{A_s(z_1 \dots z_s)}{B_s(z_1 \dots z_s)}, \quad (28)$$

where: $A_s(z_1 \dots z_s) = (-1)^s \cdot \left\{ \prod_{j=1}^s [\varphi_1(z_j) - 1] \right\} \cdot \alpha_s(-z_1, -z_1 - z_2, \dots, -\sum_{\nu=1}^s z_\nu)$,

$$B_s(z_1 \dots z_s) = 1 + \sum_{\lambda=0}^{s-1} \binom{s}{\lambda} \cdot (-1)^{s-\lambda}$$

$$\cdot \left\{ \prod_{j=\lambda+1}^s [\varphi_1(z_j) - 1] \right\} \cdot \alpha_{s-\lambda} \left(-\sum_{\nu=1}^{\lambda+1} z_\nu, \dots, -\sum_{\nu=1}^s z_\nu \right),$$

and $\alpha_j(z_1 \dots z_j)$ is defined by (4).

Now, we decompose U_s for $Re(z_i) < 0$, $i = 1 \dots s$. First, consider variable z_s and complex plane ζ_s in (27). We go to the region $Re(\zeta_s) > Re(z_s)$ and, consequently, we cross pole $\zeta_s = z_s$ and set

$$U_s(z_1 \dots z_s; q) = H_s(z_1 \dots z_s; q) \cdot M_s(z_1 \dots z_s; q), \quad (29)$$

where the integrand of H_s is the residue of the integrand of $U_s(z_1 \dots z_s; q)$. We have:

$$\begin{aligned}
 H_s(z_1 \dots z_s; q) &= \text{Exp} \left[\frac{-1}{(2\pi i)^{s-1}} \cdot \int_{-0} [\frac{1}{q + \zeta_1} + \frac{1}{z_1 - \zeta_1}] \cdot d\zeta_1 \dots \right. \\
 &\dots \int_{-0} \left[\frac{1}{q + \sum_{\nu=1}^{s-2} z_\nu + \zeta_{s-1}} + \frac{1}{z_{s-1} - \zeta_{s-1}} \right] \cdot d\zeta_{s-1} \cdot \log \mathbf{N}_s(\zeta_1 \dots \zeta_{s-1}, z_s) \Big].
 \end{aligned} \tag{30}$$

$M_s(z_1 \dots z_s; q)$ is still defined by (27) but with $R(\zeta_s) > R(z_s)$. M_s is therefore holomorphic for $\text{Re}(z_s) < 0$. Proceeding in the same way for variable z_{s-1} and plane ζ_{s-1} in (30), we set as above

$$H_s(z_1 \dots z_s; q) = H_{s-1}(z_1 \dots z_s; q) \cdot M_{s-1}(z_1 \dots z_s; q), \tag{31}$$

where the integrand of H_{s-1} is the residue of that of H_s . We therefore have:

$$\left\{ \begin{aligned}
 H_{s-1}(z_1 \dots z_s; q) &= \text{Exp} \left[\frac{-1}{(2\pi i)^{s-1}} \cdot \int_{-0} \left[\frac{1}{q + \zeta_1} + \frac{1}{z_1 - \zeta_1} \right] \cdot d\zeta_1 \dots \right. \\
 &\dots \int_{-0} \left[\frac{1}{q + \sum_{\nu=1}^{s-3} z_\nu + \zeta_{s-2}} + \frac{1}{z_{s-2} - \zeta_{s-2}} \right] \cdot d\zeta_{s-2} \cdot \log \mathbf{N}_s(\zeta_1 \dots \zeta_{s-2}, z_{s-1}, z_s) \Big].
 \end{aligned} \right. \tag{32}$$

M_{s-1} is defined by (30), with $R(\zeta_{s-1}) > R(z_{s-1})$. M_{s-1} is holomorphic for $\text{Re}(z_{s-1}) < 0$. Continuing in this way, we derive the decomposition, in (23) with $j = 1 \dots s$:

$$\left\{ \begin{aligned}
 M_j(z_1 \dots z_s; q) &= \text{Exp} \left[\frac{-1}{(2\pi i)^j} \int_{-0} \left[\frac{1}{q + \zeta_1} + \frac{1}{z_1 - \zeta_1} \right] \cdot d\zeta_1 \dots \right. \\
 &\dots \int_{-0} \left[\frac{1}{q + \sum_{\nu=1}^{j-1} z_\nu + \zeta_j} + \frac{1}{z_j - \zeta_j} \right] \cdot d\zeta_j \cdot \log \mathbf{N}_s(\zeta_1 \dots \zeta_j, z_{j+1} \dots z_s) \Big].
 \end{aligned} \right. \tag{33}$$

M_j satisfies conditions (a) and (b); finally \mathbf{U}_s satisfies the factorization, in (23), and consequently, \mathbf{V}_s satisfies (25). Moreover, from (25), we have $V_s(0 \dots 0; 0) \equiv 1$. We will consider specially the function $(Ee^{-q\tau})$ relating to the queueing delay τ of *delayed* customers, defined by $\mathbf{V}_s(0 \dots 0; q) = (1 - P) + P \cdot Ee^{-q\tau}$, where P is the *probability of delay* corresponding to $|q|$ increasing indefinitely. Consequently, we may drop the terms $[1/(z_j - \zeta_j)]$ in integrand (27). Since we know that the solution is unique, we can state:

Theorem 2: (G/G/s Queue) *For the stationary G/G/s queue, the queueing delay τ of an arbitrary delayed customer is given, for $\text{Re}(q) > 0$, by:*

$$Ee^{-q\tau} = 1 - \text{Exp} \left\{ \frac{-1}{(2\pi i)^s} \cdot \int_{-0} \frac{d\zeta_1}{q + \zeta_1} \dots \int_{-0} \frac{d\zeta_s}{q + \zeta_s} \right\} \cdot \log \mathbf{N}_s(\zeta_1 \dots \zeta_s), \tag{34}$$

where $\mathbf{N}_s(\zeta_1 \dots \zeta_s)$ is defined by the long Expression (28).

5.2 The Asymptotic Distribution

The asymptotic distribution corresponds to the (real) singularity closest to the origin. We wish to evaluate it. To more readily evaluate this singular point, we transform (20) by noting a certain symmetry with respect to variables z_j close to this real point. We successively set:

$$z_j = \frac{1}{s} \cdot (-q + \zeta_j), \quad j = 1 \dots s; \quad (35)$$

and

$$\varphi_1^*(z) = \varphi_1\left(\frac{z}{s}\right). \quad (36)$$

We deduce, for $\text{Re}(-q + \zeta_j) > 0$ ($j = 1 \dots s$):

$$G_s(q) = 1 - \frac{1}{(2\pi i)^s} \cdot \int_{+0} \frac{d\zeta_1}{-q + \zeta_1} \dots \int_{+0} \frac{d\zeta_s}{-q + \zeta_s} \cdot \frac{s \cdot q}{\sum_{\nu=1}^s \zeta_\nu} \cdot \frac{1}{R'_s(\zeta_1 \dots \zeta_s, q)}, \quad (37)$$

with

$$R'_s(\zeta_1 \dots \zeta_s; q) = 1 + \sum_{\lambda=0}^{s-1} \binom{s}{\lambda} \cdot (-1)^{s-\lambda} \cdot \alpha_{s-\lambda}(q \dots q) \cdot \prod_{j=\lambda+1}^s [\varphi_1^*(-q + \zeta_j) - 1].$$

We go now to the region $\text{Re}(-q + \zeta_j) < 0$ ($j = 1 \dots s$) and, consequently, cross poles $\zeta_1 = \dots = \zeta_s = 0$ and $\zeta_1 = \dots = \zeta_s = q$. Expression (37) becomes, for $\text{Re}(-q + \zeta_j) < 0$ ($j = 1 \dots s$):

$$G_s(q) = \left(\frac{-1}{q}\right)^{s-1} \cdot \frac{1}{R'_s(0 \dots 0, q)} - \frac{s}{(2\pi i)^s} \int_{-0} \frac{d\zeta_1}{-q + \zeta_1} \dots \int_{-0} \frac{d\zeta_s}{-q + \zeta_s} \cdot \frac{q}{\sum_{\nu=1}^s \zeta_\nu} \cdot \frac{1}{R'_s(\zeta_1 \dots \zeta_s, q)}. \quad (38)$$

For $\text{Re}(-q + \zeta_j) < 0$, we have $|\varphi_1^*(-q + \zeta_j)| < 1$. It follows that the real part of $-\varphi_1^*(-q + \zeta_j) - 1$ is positive, and it is the same for $R'_s(\zeta_1 \dots \zeta_s; q)$, $R_e(q) > 0$. It does not appear that any singularity exists in the integrand of (38) for $R_e(\zeta_j) < 0$ ($j = 1 \dots s$). The integral is equal to zero, and we deduce for our transformation that around the wanted singular point,

$$G_s(q) = \left(\frac{-1}{q}\right)^{s-1} \cdot \frac{1}{R'_s(0 \dots 0, q)}. \quad (39)$$

Rule: Finally, the *singular point* (for the asymptotic expression of the queueing delay distribution) is the real root q_0 (*closest to the origin*) of the following equation, for $\text{Re}(q) < 0$:

$$R'_s(0 \dots 0; q) = 1 + \sum_{\lambda=0}^{s-1} \binom{s}{\lambda} \cdot \alpha_{s-\lambda}(q \dots q) \cdot [1 - \varphi_1^*(-q)]^{s-\lambda} = 0, \quad (40)$$

where $\varphi_1^*(-q)$ is defined by (36).

6. The Stationary GI/G/s Queue

To get $Ee^{-q\tau}$, we apply Theorem 2 where (28) becomes, due to (6):

$$\frac{1}{N_s(z_1 \dots z_s)} = \frac{A_s(z_1 \dots z_s)}{B_s(z_1 \dots z_s)}, \tag{41}$$

with

$$A_s(z_1 \dots z_s) = (-1)^s \cdot \prod_{j=1}^s \{[\varphi_1(z_j) - 1] \cdot \alpha_1(-\sum_{\nu=1}^j z_\nu)\},$$

$$B_s(z_1 \dots z_s) = 1 + \sum_{\lambda=0}^{s-1} \binom{s}{\lambda} \cdot (-1)^{s-\lambda} \cdot \prod_{j=\lambda+1}^s \{[\varphi_1(z_j) - 1] \cdot \alpha_1(-\sum_{\nu=1}^j z_\nu)\}.$$

To get the asymptotic distribution, equation (40) becomes:

$$R'_s(0 \dots 0; q) = [1 - \alpha_1(q) \cdot [\varphi_1^*(-q) - 1]]^s = \frac{1 - \varphi_0(q) \cdot \varphi_1^*(-q)}{1 - \varphi_0(q)}. \tag{42}$$

This leads to the equation:

$$1 - \varphi_0(q) \cdot \varphi_1^*(-q) = 0. \tag{43}$$

The asymptotic expression of the queueing delay distribution corresponds to the real root q_0 (closest to the origin) of (43). We may conclude:

Rule for the asymptotic distribution: The arbitrary *delayed* customer of the stationary GI/G/s queue has the following queueing delay *asymptotic* distribution:

$$F(t, s) = F(st, 1), \tag{44}$$

where $F(t, 1)$ is the queueing delay distribution in the GI/G/1 queue corresponding to the couple $\varphi_0(z)$ and $\varphi_1(\frac{z}{s})$. But, to evaluate the probability of delay, we have to use the factorization, in (23) for $s > 1$ rather than the Wiener-Hopf decomposition (corresponding to $s = 1$).

Note: 1) For the GI/D/s queue, the cyclic nature of service time terminations leads to a factorization, similar to the Wiener-Hopf decomposition, but for the couple $[\varphi_0(z)]^s$ and $\varphi_1(z)$, instead.

2) For the GI/M/s queue, there is only one singular point. Consequently, (44) is valid for any value of $t (> 0)$.

7. The Stationary M/G/s Queue

In (41), the function $\alpha_1(z_i)$ becomes, in case of Poisson arrivals with Λ for the *arrival rate*:

$$\alpha_1(z_i) = -\frac{\Lambda}{z_i}. \tag{45}$$

Expression (41) becomes:

$$\frac{1}{N_s(z_1 \dots z_s)} = \frac{A_s(z_1 \dots z_s)}{B_s(z_1 \dots z_s)}, \quad (46)$$

with

$$A_s(z_1 \dots z_s) = (-1)^s \cdot \prod_{j=1}^s \left\{ \frac{\Lambda \cdot [\varphi_1(z_j) - 1]}{z_1 + \dots + z_j} \right\},$$

$$B_s(z_1 \dots z_s) = 1 + \sum_{\lambda=0}^{s-1} \binom{s}{\lambda} \cdot (-1)^{s-\lambda} \cdot \prod_{j=\lambda+1}^s \left\{ \frac{\Lambda \cdot [\varphi_1(z_j) - 1]}{z_1 + \dots + z_j} \right\}.$$

Consider the term for $j = \lambda + 1$ by writing it as

$$\Lambda \cdot \left\{ \frac{\varphi_1(z_{\lambda+1}) - 1}{z_{\lambda+1}} \right\} \cdot \frac{z_{\lambda+1}}{z_1 + \dots + z_{\lambda+1}}.$$

Due to the symmetry in (34) and (46), we do not change anything for the value of integrals in (27) by writing successively, for z_j with $j \leq \lambda + 1$:

$$\Lambda \cdot \left\{ \frac{\varphi_1(z_{\lambda+1}) - 1}{z_{\lambda+1}} \cdot \frac{z_j}{z_1 + \dots + z_{\lambda+1}} \right\} \rightarrow \frac{1}{\lambda+1} \Lambda \cdot \left\{ \frac{\varphi_1(z_{\lambda+1}) - 1}{z_{\lambda+1}} \right\} \cdot \frac{z_1 + \dots + z_{\lambda+1}}{z_1 + \dots + z_{\lambda+1}}.$$

If we apply this transformation to the successive terms in (46), Theorem 2 allows us to state:

Theorem 3: (M/G/s Queue) *For the stationary M/G/s queue, the function $Ee^{-q\tau}$ relating to the queueing delay τ of an arbitrary delayed customer is given by (34) with*

$$\frac{1}{N_s(z_1 \dots z_s)} = \frac{A_s(z_1 \dots z_s)}{B_s(z_1 \dots z_s)}, \quad (47)$$

with

$$A_s(z_1 \dots z_s) = \frac{(-1)^s}{s!} \cdot \prod_{j=1}^s \left\{ \frac{\Lambda \cdot [\varphi_1(z_j) - 1]}{z_j} \right\},$$

$$B_s(z_1 \dots z_s) = 1 + \sum_{\lambda=0}^{s-1} \frac{(-1)^{s-\lambda}}{(s-\lambda)!} \cdot \prod_{j=\lambda+1}^s \left\{ \Lambda \cdot \frac{\varphi_1(z_j) - 1}{z_j} \right\}.$$

8. Conclusion

We may note a discrepancy in the case of the GI/M/s queue. We recall that a key point in this paper relating to (15) is the one that defines the *sequence t_n of successive terminations of service times* during the congestion period; and it is not a Markov process. The traditional Markovian assumption was sufficient to evaluate the events for a single arrival, and for the asymptotic expression of the queueing delay distribution of a *delayed* customer. However, to evaluate the probability of the delay and occupancy probabilities it is necessary to properly evaluate the correlations between two successive arrivals by taking into account the starting epoch and age (in number of service times) of a busy (quite congested) period. It is also imperative to introduce the *sequence t_n* defined in (15), and modify the basic Poisson process to evaluate the service time terminations during congestion. Surprisingly, for a very long time, the discrepancy between the formulas presented in Pollaczek [3] and Takács formula (for the GI/M/s queue) has never been stressed.

References

- [1] Loynes, R.M., The stability of a queue with nonindependent interarrival and service times, *Proc. Cambridge Philos. Soc.* **58** (1962), 494-520.
- [2] Pollaczek, F., Problèmes stochastiques posés par le phénomène de formation d'une queue d'attente à un guichet et par des phénomènes apparentés, *Mémorial des Sciences Mathématiques*, Gauthier-Villars, Paris **CXXXVI** (1957). (= GI/G/1 queue; in French).
- [3] Pollaczek, F., Théorie analytique des problèmes stochastiques relatifs à un groupe de lignes téléphoniques avec dispositif d'attente, *Mémorial des Sciences Mathématiques*, Gauthier-Villars, Paris **CL** (1961). (=GI/G/s queue; in French).