Queueing Syst (2006) 52: 261–279
DOI 10.1007/s11134-006-6134-x

Analysis of customers' impatience in queues with server vacations

Eitan Altman · Uri Yechiali

Received: 29 September 2004 / Revised: 10 November 2005

Abstract Many models for customers impatience in queueing systems have been studied in the past; the source of impatience has always been taken to be either a long wait already experienced at a queue, or a long wait anticipated by a customer upon arrival. In this paper we consider systems with servers vacations where customers' impatience is due to an *absentee* of servers upon arrival. Such a model, representing frequent behavior by waiting customers in service systems, has never been treated before in the literature. We present a comprehensive analysis of the single-server, *M/M/*1 and *M/G/*1 queues, as well as of the multi-server *M/M/c* queue, for both the multiple and the single-vacation cases, and obtain various closed-form results. In particular, we show that the proportion of customer abandonments under the single-vacation regime is smaller than that under the multiple-vacation discipline.

Keywords Queueing · Single and multiple vacations · Impatience · Abandonment · $M/M/1$ · $M/G/1$ · $M/M/c$.

1. Introduction

Customers impatience has been dealt with in the queueing literature mainly in the context of customers abandoning the queue due to either a long wait already experienced, or a long wait anticipated upon arrival. Many authors treated the impatience phenomenon under various assumptions. We mention a few of the works. Palm's pioneering work [1953, 1957] seems to be the first to analyze queueing systems with impatient customers by considering the unlimited buffer *M/M/c* queue and assuming that each individual customer stays in the queue as long as his waiting time does not exceed an exponentially distributed impatience time. Daley (1965) studied the *GI/G/*1 queue "in which customers entering the system, if they are obliged to wait too long, may leave the system before starting or completing their service". He derived an integral equation for the limiting waiting-time distribution function and investigated its solution for the cases of deterministic and of distributed impatience.

Takacs (1974) further studied the *M/G/*1 queue in which customers have a fixed threshold on their sojourn time, and derived the limiting distributions of the actual and of the virtual waiting times. Baccelli et al. (1984) considered the *GI/G/*1 queue where each 'aware' customer, upon arrival, leaves immediately if he knows that his total waiting time is beyond his impatience threshold. They devoted the analysis to characterization of waiting times in the system. Boxma and de Waal (1994) studied the *M/M/c* queue with generally distributed impatience times, developed several approximations for the abandonment probability and tested them via simulation. Altman and Borovkov (1997) considered impatience of customers in a retrial queue, in which a customer leaves the system if its commulative sojourn time exceeds some random threshold; impatience is shown to have an important impact on the system stability. Van Houdt et al. (2003) presented "an algorithmic procedure to calculate the delay distribution of im(patient) customers in a discrete time *D–MAP/PH/*1 queue, where the service time distribution of a customer depends on his waiting time." They consider deterministic impatience time in the waiting room and in the system.

This work was supported by the Euro-Ngi network of excellence.

E. Altman (✉)
INRIA, 2004 route des Lucioles, B.P. 93, 06902 Sophia Antipolis, France
e-mail: altman@sophia.inria.fr

U. Yechiali
Department of Statistics and Operations Research, School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel
e-mail: uriy@post.tau.ac.il

The abandonment phenomenon and its importance with respect to stability of so called “call centers” has been studied extensively recently by various authors. We direct the reader to the survey paper by Gans et al. (2003) and the many references there. For the relevance of impatience in telecommunication systems, see e.g. Bonald and Roberts (2001).

However, there are situations where customers impatience is due to an *absentee* of servers upon arrival. This situation is encountered, in particular, when observing human behaviour in service systems: if an arriving customer sees no server present in the system, he/she may abandon the queue if no server shows up within some time.

As a consequence, we analyze in this work queueing systems with servers vacations, where each arriving customer who finds no servers on duty, activates an *independent* random impatience timer. If a server does not show up by the time the timer expires, the customer *abandons* the queue. (This timer procedure is *different* from the *server's* timer procedure studied by Boxma et al. (2002) and by Yechiali (2004), in which the server, upon returning from a vacation to an *empty* system, activates a timer before taking another vacation.) We study both the multiple-vacation and the single-vacation models of the single-server *M/M/*1 and *M/G/*1 queues, as well as of the multi-server *M/M/c* systems. The analyses of the *M/M/*1 queues, as well as of the *M/M/c* systems, require the solution of a differential equation for the partial generating functions. This is not a common occurrence when employing the partial generating functions method. The analyses of the *M/G/*1 queues are achieved by an interesting use of a related *M/G/*∞ model. The *M/M/c* queues further require the calculation of certain probabilities, which are derived by finding the roots of a 2*c*-degree polynomial being the determinant of a certain matrix whose entries are functions of the system's parameters.

The paper consists of the following models: In Section 2 we consider the *M/M/*1 queue with multiple server's exponentially distributed vacations and with exponentially distributed impatience times. After deriving the balance equations we obtain and solve a differential equation for $G_0(z)$, the (conditional) generating function of the queue size when the server is on vacation. This enables us to calculate the fractions of time the server is vacationing or busy. In addition, we calculate P_{00}, the fraction of time the server is on vacation and the system is empty. Section 3 deals with the multiple-vacation *M/G/*1 queue where the vacation times, the service times and the impatience times are generally distributed. We derive the Laplace-Stieltjes transform (LST) and the mean of the vacation period; the probability generating function (PGF) and mean of the number of customers at the start of a busy period; the LST and the mean duration of a busy period, and calculate P_{00}. Furthermore, we derive the PGF of the number of customers at a service completion instant, and present a decomposition result. Also, we have calculated the mean number of customers in the system at an arbitrary moment. Finally, we derive a closed-form expression for *P*(*served*), the fraction of customers served without abandoning the system. Section 4 treats the multiple-vacation $M/M/c$ queue with exponentially distributed vacation and impatience times. Again, the balance equations lead to a differential equation for the PGF $G_0(z)$. Interestingly enough, its solution is similar to the solution of the $M/M/1$ case. In order to obtain a complete solution for the unknown probabilities of the system state, the roots of a polynomial, obtained via the calculation of a square matrix, are determined. Section 5, 6 and 7 are the counterparts of Section 2, 3 and 4, respectively, where the scenario is that of the server following the single-vacation service procedure.

2. Multiple vacations: *M/M/*1 queue with exponentially distributed vacation and impatience times

2.1. The model

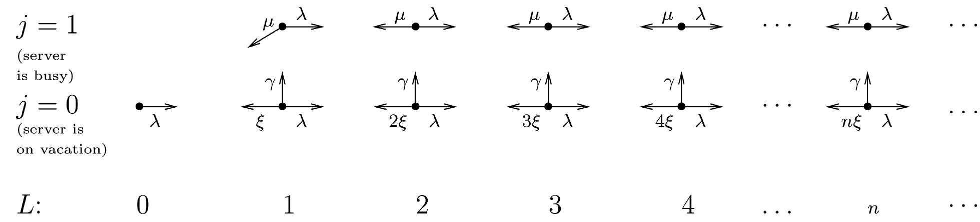
The underlying process is a $M/M/1$ queue with multiple server vacations (Levy and Yechiali, 1975). The Poisson arrival rate is λ, service times B are exponentially distributed with parameter μ, and each vacation duration U is exponentially distributed with parameter γ. Customers are impatient. That is, whenever a customer arrives to the system and realizes that the server is on vacation he activates an ‘impatience timer’ T, exponentially distributed with parameter ξ, which is *independent* of the queue size at that moment. If the server returns from his vacation before the time T expires (and starts rendering service), the customer stays in the system until his service is completed. However, if T expires while the server is still on vacation, our customer *abandons* the queue, never to return.

2.2. Balance equations

Let L denote the total number of customers in the system, and let J denote the number of working servers ($J = 0$ implies that the server is on vacation, while $J = 1$ denotes that the server is active). Then, the pair (J, L) defines a continuous-time Markov process with transition-rate diagram as depicted in Figure 1. Let $P_{jn} = P\{J = j, L = n\}$ ($j = 0, 1; n = 0, 1, 2, \ldots$) denote the (steady state) system-state probabilities. Then, the set of balance equations is given as follows:

$$\underline{j=0} \quad \begin{cases} n = 0 & \lambda P_{00} = \xi P_{01} + \mu P_{11} \\ n \geq 1 & (\lambda + n\xi + \gamma) P_{0n} = \lambda P_{0,n-1} \\ & \quad + (n+1)\xi P_{0,n+1}, \end{cases} \tag{2.1}$$

Fig. 1 Transition-rate diagram



$$\underline{j=1}\quad \begin{cases} n=1 & (\lambda+\mu)P_{11}=\mu P_{12}+\gamma P_{01} \\ n\geq 2 & (\lambda+\mu)P_{1n}=\lambda P_{1,n-1} \\ & +\mu P_{1,n+1}+\gamma P_{0,n}. \end{cases} \tag{2.2}$$

Define the Probability Generating Functions (PGFs)

$$G_0(z)=\sum_{n=0}^{\infty}P_{0n}z^n\ ,\ G_1(z)=\sum_{n=1}^{\infty}P_{1n}z^n.$$

Then, by multiplying each equation for n in (2.2) by z^n, summing over n and rearrange terms, we get

$$G_1(z)[(\lambda z-\mu)(1-z)]=\gamma z G_0(z)-(\mu P_{11}+\gamma P_{00})z \tag{2.3}$$

In a similar manner we obtain from (2.1)

$$\xi(1-z)G_0'(z)=[\lambda(1-z)+\gamma]G_0(z)-(\gamma P_{00}+\mu P_{11}). \tag{2.4}$$

where $G_0'(z)=\frac{d}{dz}G_0(z)$. We derive the solution of the differential Eq. (2.4) in the following section.

2.3. Solution of the differential equation

Set

$$A=\gamma P_{00}+\mu P_{11}. \tag{2.5}$$

Then, for $z\neq 1$,

$$G_0'(z)-\left[\frac{\lambda}{\xi}+\frac{\gamma}{\xi(1-z)}\right]G_0(z)=\frac{-A}{\xi(1-z)} \tag{2.6}$$

Multiplying both sides of (2.6) by $e^{-\frac{\lambda}{\xi}z}(1-z)^{\frac{\gamma}{\xi}}$ we get

$$\frac{d}{dz}\left[(e^{-\frac{\lambda}{\xi}z}\cdot(1-z)^{\gamma/\xi})G_0(z)\right]=\frac{-A}{\xi}e^{-\frac{\lambda}{\xi}z}(1-z)^{\frac{\gamma}{\xi}-1}.$$

Integrating from 0 to z we have

$$\begin{aligned} &e^{-\frac{\lambda}{\xi}z}\cdot(1-z)^{\gamma/\xi}G_0(z)-G_0(0) \\ &\quad=\frac{-A}{\xi}\int_{s=0}^{z}(1-s)^{\frac{\gamma}{\xi}-1}\cdot e^{-\frac{\lambda}{\xi}s}\cdot ds \end{aligned} \tag{2.7}$$

Thus,

$$\begin{aligned} G_0(z)&=G_0(0)\cdot e^{\frac{\lambda}{\xi}z}\cdot(1-z)^{-\frac{\gamma}{\xi}}-\frac{A}{\xi}e^{\frac{\lambda}{\xi}z}\cdot(1-z)^{-\frac{\gamma}{\xi}}\cdot \\ &\quad\int_{s=0}^{z}(1-s)^{\frac{\gamma}{\xi}-1}\cdot e^{-\frac{\lambda}{\xi}s}\cdot ds \end{aligned} \tag{2.8}$$

Then,

$$\begin{aligned} G_0(1)&=e^{\frac{\lambda}{\xi}}\left[G_0(0)-\frac{A}{\xi}\int_{s=0}^{1}(1-s)^{\frac{\gamma}{\xi}-1}e^{-\frac{\lambda}{\xi}s}ds\right]\cdot \\ &\quad\lim_{z\to 1}\left[(1-z)^{-\gamma/\xi}\right]. \end{aligned}$$

Since $G_0(1)=\sum_{n=0}^{\infty}P_{0n}\overset{def}{=}P_{0\bullet}>0$ and $\lim_{z\to 1}(1-z)^{-\gamma/\xi}=\infty$, we must have that

$$G_0(0)=\frac{A}{\xi}\int_{s=0}^{1}(1-s)^{\frac{\gamma}{\xi}-1}\cdot e^{-\frac{\lambda}{\xi}s}ds \tag{2.9}$$

Define

$$Z(\lambda,\gamma):=-\lambda^{-\gamma}e^{-\lambda}\left(-\Gamma(\gamma,-\lambda)+\Gamma(\gamma)\right),$$

where $\Gamma(z)$ is the Γ-function that has the representation $\Gamma(z):=\int_{t=0}^{\infty}\exp(-t)t^{z-1}dt$, and $\Gamma(a,z):=\int_{t=z}^{\infty}\exp(-t)\,t^{a-1}dt$. Further define

$$K:=\int_{s=0}^{1}(1-s)^{\frac{\gamma}{\xi}-1}e^{-\frac{\lambda}{\xi}s}ds.$$

Some computations give

$$K=Z\left(\frac{\lambda}{\xi},\frac{\gamma}{\xi}\right). \tag{2.10}$$

We then write, using (2.5),

$$G_0(0)=P_{00}=\frac{\gamma P_{00}+\mu P_{11}}{\xi}\cdot K=\frac{K\mu}{\xi-K\gamma}P_{11}. \tag{2.11}$$

(It is easy to check that $\xi-\gamma K>0$. Indeed, $K=\int_0^1(1-s)^{\gamma/\xi-1}e^{-\frac{\lambda}{\xi}s}\cdot ds<\int_0^1(1-s)^{\gamma/\xi-1}\ \cdot ds=\frac{\xi}{\gamma}$). Now, substituting in (2.8) the value of A from equation (2.9) we obtain

$$G_0(z) = G_0(0) \cdot e^{\frac{\lambda}{\xi}z}\left[1 - \frac{\int_0^z (1-s)^{\frac{\gamma}{\xi}-1} e^{-\frac{\lambda}{\xi}s} \cdot ds}{\int_0^1 (1-s)^{\frac{\gamma}{\xi}-,1} \cdot e^{-\frac{\lambda}{\xi}s} \cdot ds}\right] \Big/ (1-z)^{\gamma/\xi}. \tag{2.12}$$

Using L'Hopital rule, the probability that the server is on vacation, $P_{0\bullet} = \sum_{n=0}^{\infty} P_{0n}$, is derived:

$$G_0(1) = P_{0\bullet} = G_0(0) \frac{1}{\frac{\gamma}{\xi}\int_0^1 (1-s)^{\frac{\gamma}{\xi}-1} \cdot e^{-\frac{\lambda}{\xi}s} ds} = \frac{\xi}{\gamma K} G_0(0) = \frac{\xi}{\gamma K} P_{00}. \tag{2.13}$$

Clearly, the probability that the server is working is $P_{1\bullet} = \sum_{n=1}^{\infty} P_{1n} = 1 - P_{0\bullet}$. By substituting $G_0(0)$ from (2.11) we get the relation

$$G_0(1) = P_{0\bullet} = \frac{\frac{A}{\xi}K}{\frac{\gamma}{\xi}K} = \frac{A}{\gamma} = \frac{\gamma P_{00} + \mu P_{11}}{\gamma}.$$

Clearly, the above expression,

$$\gamma P_{0\bullet} = \gamma P_{00} + \mu P_{11}. \tag{2.14}$$

can be obtained by considering a horizontal 'cut' between the two levels of the transition rate diagram (Figure 1).

Equation (2.12) expresses $G_0(z)$ in terms of $G_0(0) = P_{00}$, the proportion of time the server is on vacation and there are no customers in the system. Also, $G_1(z)$ is a function of $G_0(z)$ and P_{00}. Thus, once P_{00} is calculated, $G_0(z)$ and $G_1(z)$ are completely determined, P_{11} is obtained from (2.11), and $P_{0\bullet}$ is given by (2.13). Finally, the proportion of customers *abandoning* the system is $P(T < U)P_{0\bullet} = \frac{\xi}{\gamma+\xi} P_{0\bullet}$. Nevertheless, the necessary and sufficient condition for stability is $\lambda < \mu$, for otherwise the busy period has an infinite expectation.

We derive the value of P_{00} in the next section.

2.4. Derivation of $P_{0\bullet}$, $P_{1\bullet}$, P_{00}, $E[L_0]$ and of $E[L_1]$.

From (2.3),

$$G_1(z) = \frac{[\gamma G_0(z) - (\mu P_{11} + \gamma P_{00})]z}{(\lambda z - \mu)(1-z)}.$$

Applying L'Hopital rule, we get

$$G_1(1) = \frac{[\gamma G_0(1) - (\mu P_{11} + \gamma P_{00})] + \gamma G_0'(1)}{\mu - \lambda},$$

where $G_0'(1) \equiv E[L_0] = \sum_{n=1}^{\infty} n P_{0n}$.

Since $G_j(1) = P_{j\bullet}$ $(j = 1, 2)$, by applying equation (2.14) to the numerator of $G_1(1)$ above, we have

$$E[L_0] = \frac{\mu - \lambda}{\gamma} P_{1\bullet}. \tag{2.15}$$

On the other hand, from equation (2.4),

$$E[L_0] = \lim_{z \to 1} G_0'(z) = \frac{-\lambda G_0(1) + \gamma G_0'(1)}{-\xi} = \frac{-\lambda P_{0\bullet} + \gamma E[L_0]}{-\xi},$$

implying that

$$E[L_0] = \frac{\lambda P_{0\bullet}}{\gamma + \xi}. \tag{2.16}$$

Equating the two expressions (2.15) and (2.16) for $E[L_0]$, and using $1 = P_{0\bullet} + P_{1\bullet}$, we get

$$P_{1\bullet} = \frac{\lambda\gamma}{\mu\gamma + \xi(\mu - \lambda)}, \quad P_{0\bullet} = \frac{(\gamma + \xi)(\mu - \lambda)}{\mu\gamma + \xi(\mu - \lambda)}, \tag{2.17}$$

implying that

$$E[L_0] = \frac{\lambda(\mu - \lambda)}{\mu\gamma + \xi(\mu - \lambda)}.$$

Now, using equation (2.13), we finally obtain

$$P_{00} = \frac{\gamma K}{\xi} P_{0\bullet} = \frac{\gamma K}{\xi} \cdot \frac{(\gamma + \xi)(\mu - \lambda)}{\mu\gamma + \xi(\mu - \lambda)}, \tag{2.18}$$

where K is given by equation (2.10).

$P_{1\bullet}$ [$P_{0\bullet}$] is a decreasing [increasing] convex [concave] function of ξ, having its limits at $\frac{\lambda}{\mu}$ [at $1 - \lambda/\mu$] when $\xi \to 0$, and at 0 [at 1, respectively] when $\xi \to \infty$. $E[L_0]$ behaves similar to $P_{1\bullet}$. Indeed, $E[L_0] \to \frac{\lambda}{\gamma}(1 - \frac{\lambda}{\mu})$ when $\xi \to 0$, and $E[L_0] \to 0$ when $\xi \to \infty$ (that is, every arrival who finds the server on vacation leaves immediately). As for P_{00}, its behavior as a function of ξ is given in Figure 2 for parameter values $\lambda = 1, \mu = 2$, and $\gamma = 1$. In general, it is an increasing concave function of ξ with asymptote at 1 (since, as $\xi \to \infty$, the probability that at the end of a vacation the system is non empty, converges to zero). When $\xi \to 0$ then P_{00} approaches 0.25 (see also section 3).

Derivation of $E[L_1]$

From (2.3) we have

$$G_1(z) = \frac{\gamma G_0(z)z - Az}{(\lambda z - \mu)(1-z)},$$

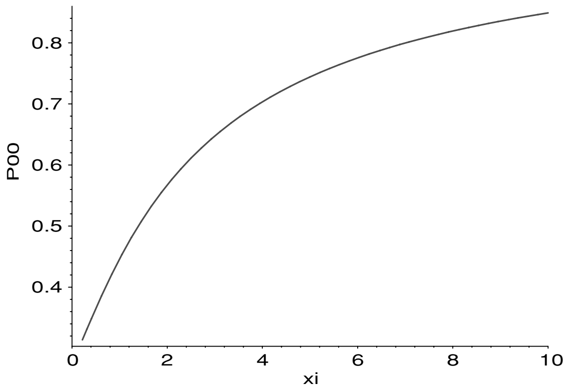


Fig. 2 P_{00} as a function of ξ

By using L'Hopital rule and $\gamma P_{0\bullet} = A$, we derive

$$E[L_1] = \lim_{z\to 1} G_1'(z) = \frac{\gamma E[L_0(L_0-1)] + 2\gamma E[L_0]}{2(\mu-\lambda)^2},$$

where $E[L_0(L_0-1)]$ is obtained by differentiating twice $G_0(z)$ at $z=1$.

2.5. Sojourn times

Let S be the total sojourn time of a customer in the system, measured from the moment of arrival until departure, either after completion of service or as a result of abandonment. By Little's law,

$$E[S] = \frac{E[L]}{\lambda} = \frac{E[L_0]+E[L_1]}{\lambda} \tag{2.19}$$

However, a more important measure of performance is S_{served}, defined as the total sojourn time of a customer who *completes* his service.

Let S_{jn} denote the conditional sojourn time of a customer who do *not* abandon the system, given that the state upon his arrival is (j, n). Clearly, $E[S_{1n}] = \frac{n+1}{\mu}, \quad n = 1, 2, 3, \ldots$ When $J = 0$, for $n \geq 1$,

$$\begin{aligned} E[S_{0n}] = {} & \frac{\gamma}{\gamma+(n+1)\xi+\lambda}\left(\frac{1}{\gamma+(n+1)\xi+\lambda} + E[S_{1n}]\right) \\ & + \frac{\lambda}{\gamma+(n+1)\xi+\lambda} \cdot \\ & \times \left(\frac{1}{\gamma+(n+1)\xi+\lambda} + E[S_{0n}]\right) \\ & + \frac{(n+1)\xi}{\gamma+(n+1)\xi+\lambda} \cdot \frac{n}{n+1} \cdot \\ & \times \left(\frac{1}{\gamma+(n+1)\xi+\lambda} + E[S_{0,n-1}]\right) \end{aligned}$$

The second term above follows since a new arrival does not change the sojourn time of a customer present in the system, while the third term takes into account the probability $n/(n+1)$ that, when there is an abandonment among $(n+1)$ waiting customers, our customer will not be the one to leave.

Thus,

$$\begin{aligned} (\gamma+(n+1)\xi)E[S_{0n}] = {} & \frac{\gamma+n\xi+\lambda}{\gamma+(n+1)\xi+\lambda} + \frac{\gamma(n+1)}{\mu} \\ & + n\xi E[S_{0,n-1}] \end{aligned} \tag{2.20}$$

We also have

$$\begin{aligned} E[S_{00}] = {} & \frac{\gamma}{\gamma+\xi+\lambda}\left(\frac{1}{\gamma+\xi+\lambda} + \frac{1}{\mu}\right) \\ & + \frac{\lambda}{\gamma+\xi+\lambda}\left(\frac{1}{\gamma+\xi+\lambda} + E[S_{00}]\right), \end{aligned}$$

implying that

$$E[S_{00}] = \frac{1}{\gamma+\xi}\left(\frac{\gamma+\lambda}{\gamma+\xi+\lambda} + \frac{\gamma}{\mu}\right)$$

Iterating (2.20) we obtain, for $n \geq 0$,

$$\begin{aligned} E[S_{0n}] = {} & \frac{1}{\gamma+(n+1)\xi}\left[\sum_{k=1}^{n}\left(\frac{\gamma+(k-1)\xi+\lambda}{\gamma+k\xi+\lambda} + \frac{k\gamma}{\mu}\right)\right. \\ & \prod_{j=k}^{n}\left(\frac{j\xi}{\gamma+j\xi}\right) \\ & \left. + \left(\frac{\gamma+n\xi+\lambda}{\gamma+(n+1)\xi+\lambda} + \frac{(n+1)\gamma}{\mu}\right)\right] \end{aligned} \tag{2.21}$$

Finally, using the expression for $E[S_{1n}]$, we write

$$E[S_{served}] = \sum_{n=1}^{\infty} P_{1n}E[S_{1n}] + \sum_{n=0}^{\infty} P_{0n}E[S_{0n}] \tag{2.22}$$

That is,

$$E[S_{served}] = \frac{E[L_1]+P_{1\bullet}}{\mu} + \sum_{n=0}^{\infty} P_{0n}E[S_{0n}] \tag{2.23}$$

One might be interested in the (conditional) expectation of $W(0,n)$, the total sojourn time of a customer in the system, *regardless* if served or not, given that upon arrival he observes the state $(0, n)$. Then, for $n \geq 1$,

$$\begin{aligned} E[W(0,n)] = {} & \frac{1}{\gamma+(n+1)\xi+\lambda} \\ & + \frac{\gamma}{\gamma+(n+1)\xi+\lambda}E[S_{1n}] \end{aligned}$$

$$+\frac{\lambda}{\gamma+(n+1)\xi+\lambda}E[W(0,n)]+\frac{(n+1)\xi}{\gamma+(n+1)\xi+\lambda}$$
$$\times\left(\frac{1}{n+1}\cdot 0+\frac{n}{n+1}E[W(0,n-1)]\right),$$

and, for $n=0$,

$$E[W(0,0)]=\frac{1}{\gamma+\xi+\lambda}+\frac{\gamma}{\gamma+\xi+\lambda}\frac{1}{\mu}+\frac{\xi}{\gamma+\xi+\lambda}\cdot 0$$
$$+\frac{\lambda}{\gamma+\xi+\lambda}E[W(0,0)].$$

After calculations, the above yields

$$E[W(0,n)]=\sum_{k=0}^{n}\frac{n!}{k!}\frac{\xi^{n-k}}{\mu}\cdot\frac{\mu+(k+1)\gamma}{\prod_{j=k+1}^{n+1}(\gamma+j\xi)}.$$

3. Multiple vacations: *M/G/*1 queue with generally distributed vacation and impatience times

In this section we consider the case of generally distributed service times, i.e., the underlying process is the M/G/1 queue with multiple server vacations (c.f. Levy and Yechiali 1975). The arrival process is Poisson with rate λ. Service times are i.i.d random variables, all copy of B, having first moment $E[B]$, second moment $E[B^2]$ and Laplace Stieltjes transform (LST) $B^*(s)=E[e^{-sB}]$. At the end of a busy period the server takes a vacation U, having finite moments $E[U]$ and $E[U^2]$, and LST $U^*(s)$. If the system is empty at the end of a vacation, the server takes another vacation. If there are $n\geq 1$ customers at the end of a vacation, the server starts immediately a busy period. When the server is on vacation and is not available for service, arriving customers are *impatient*. An arrival who finds that the server is away on vacation, activates an 'impatience timer,' T. If the server does not return by the time T, the customer abandons the system. Each customer activates its own timer and the T_i's are i.i.d. random variables, independent of the number of customers waiting.

Let the starting time of a vacation be $t=0$. Then, a *key observation* is that, within U, the evolution of the system is the same as that of a *M/G/*∞ queue with service times all distributed as T. For time $t\leq U$, it is well known (Takacs, 1962) that the number of customers in the system has a Poisson distribution with parameter

$$\Lambda(t)=\lambda\int_0^t[1-P(T\leq y)]dy\,,\quad t\leq U. \tag{3.1}$$

We will use this observation extensively in the sequel.

3.1. Duration of a vacation period, τ

Consider the time $t=0$ when the server first leaves for a vacation of duration U_1. If at time $t=U_1$ the queue is empty, the server takes another vacation U_2, and so on. This sequence of events terminates at the first time when the server returns and finds a non-empty system. We call this entire length of time, τ, a Vacation Period.

Using the *M/G/*∞ analogy, the probability of an empty system at time U is $e^{-\Lambda(U)}$. Thus,

$$\tau=\sum_{i=1}^{k}U_i+U_{k+1}\ \text{ with probability }\ \left(\prod_{i=1}^{k}e^{-\Lambda(U_i)}\right)$$
$$\left(1-e^{-\Lambda(U_{k+1})}\right).$$

Therefore, the LST, $\tilde{\tau}(s)$, of the Vacation Period is given by

$$\tau^*(s)$$
$$=\sum_{k=0}^{\infty}E\left[e^{-s(\sum_{i=1}^{k}U_i)}e^{-sU_{k+1}}\left(e^{-\sum_{i=1}^{k}\Lambda(U_i)}\right)\left(1-e^{-\Lambda(U_{k+1})}\right)\right]$$
$$=\sum_{k=0}^{\infty}\left(E\left[e^{-(sU+\Lambda(U))}\right]\right)^k\left(E\left[e^{-sU}\right]-E\left[e^{-(sU+\Lambda(U))}\right]\right)$$
$$=\frac{U^*(s)-E\left[e^{-(sU+\Lambda(U))}\right]}{1-E\left[e^{-(sU+\Lambda(U))}\right]} \tag{3.2}$$

It follows that the mean length of a Vacation Period is

$$E[\tau]=\frac{E[U]}{1-E\left[e^{-\Lambda(U)}\right]} \tag{3.3}$$

3.2. Number of customers at a start of a busy period

A busy period starts with $N(\tau)\geq 1$ customers. We now derive the Probability Generating Function (PGF) of $N(\tau)$. It should be pointed out that $N(\tau)$ is *not* distributed as a Poisson variable with parameter $\Lambda(\cdot)$. This follows since the last vacation U in τ (in which there is at least one arrival) is not a regular one. Indeed,

$$U^*(s)\big|_{N(U)\geq 1}=E\left[e^{-sU}\middle|N(U)\geq 1\right]$$
$$=\frac{E\left[e^{-sU}\cdot I\{N(U)\geq 1\}\right]}{E[P(N(U)\geq 1)]}$$
$$=\frac{U^*(s)-E\left[e^{-(sU+\Lambda(U))}\right]}{1-E\left[e^{-\Lambda(U)}\right]}$$

This results in

$$E\left[U\big|_{N(U)\geq 1}\right] = \frac{E[U] - E\left[Ue^{-\Lambda(U)}\right]}{\left(1 - E\left[e^{-\Lambda(U)}\right]\right)}.$$

We write

$$N(\tau) = \begin{cases} N(U_1) & \text{if } N(U_1) \geq 1 \\ N'(\tau') & \text{if } N(U_1) = 0 \end{cases} \tag{3.4}$$

where $N'(\tau')$ and τ' are i.i.d. replicas of $N(\tau)$ and τ, respectively. Then, the PGF of $N(\tau)$ is given by

$$\begin{aligned} G_{N(\tau)}(z) &= E[z^{N(\tau)}] \\ &= E\left\{E[z^{N(U)}|N(U) \geq 1]P(N(U) \geq 1)\right\} \\ &\quad + E\left\{E[z^{N(\tau)}|N(U) = 0]P(N(U) = 0)\right\} \\ &= E\left[\sum_{n=1}^{\infty} z^n e^{-\Lambda(U)} \frac{(\Lambda(U))^n}{n!}\right] \\ &\quad + E[z^{N(\tau)}] \cdot E[e^{-\Lambda(U)}] \end{aligned}$$

Thus,

$$\begin{aligned} G_{N(\tau)}(z) &= \frac{E[e^{-(1-z)\Lambda(U)}] - E[e^{-\Lambda(U)}]}{1 - E[e^{-\Lambda(U)}]} \\ &= \frac{\sum_{n=1}^{\infty} \frac{1}{n!} E[e^{-\Lambda(U)}(\Lambda(U))^n] z^n}{1 - E[e^{-\Lambda(U)}]}. \end{aligned} \tag{3.5}$$

It readily follows that

$$P(N(\tau) = n) = \frac{\frac{1}{n!} E[e^{-\Lambda(U)}(\Lambda(U))^n]}{1 - E[e^{-\Lambda(U)}]} \quad (n = 1, 2, 3, \ldots) \tag{3.6}$$

and

$$E[N(\tau)] = \frac{E[\Lambda(U)]}{1 - E[e^{-\Lambda(U)}]} \tag{3.7}$$

3.3. The busy period and $P_{(busy)}$

Let Γ denote the duration of a busy period.

A busy period starts with $N(\tau) \geq 1$ customers, and hence is equal to the sum of $N(\tau)$ i.i.d regular *M/G/*1 periods θ_1, $\theta_2, \ldots, \theta_{N(\tau)}$, all distributed like θ, where $\theta^*(s) = B^*[s + \Lambda(1 - \theta^*(s))]$ (Kleinrock, 1975, p. 212). Thus, the LST of Γ is given by

$$\begin{aligned} \Gamma^*(s) &= E[e^{-s\Gamma}] = E[e^{-s(\sum_{i=1}^{N(\tau)} \theta_i)}] \\ &= [\theta^*(s)]^{N(\tau)} = G_{N(\tau)}(\theta^*(s)). \end{aligned}$$

Using (3.5) we get

$$\Gamma^*(s) = \frac{E[e^{-(1-\theta^*(s)\Lambda(U)}] - E[e^{-\Lambda(U)}]}{1 - E[e^{-\Lambda(U)}]}. \tag{3.8}$$

In particular, with $\rho = \lambda E[B]$,

$$E[\Gamma] = E[N(\tau)]E[\theta] = \frac{E[\Lambda(U)]}{1 - E[e^{-\Lambda(U)}]} \cdot \frac{E[B]}{1 - \rho}. \tag{3.9}$$

Now, the proportion of time the server is busy, $P_{(busy)}$, is give by

$$P_{(busy)} = \frac{E[\Gamma]}{E[\Gamma] + E[\tau]} = \frac{E[\Lambda(U)]E[B]}{E[\Lambda(U)]E[B] + (1 - \rho)E[U]} \tag{3.10}$$

Indeed, for the *M/M/*1 case, $P_{(busy)} = P_{1\bullet}$ as given by (2.17). Furthermore, $P_{(busy)} < \rho$. To see this, it is enough to show that $E[\Lambda(U)] < \lambda E[U]$. In fact,

$$\lambda E\left[\int_0^U [1 - P(T \leq y)] \mathrm{d}y\right] \leq \lambda E\left[\int_0^U 1 \cdot \mathrm{d}y\right] = \lambda E[U]. \tag{3.11}$$

3.4. Calculation of P_{00}

We now calculate P_{00}, defined in section 1.

Let D denote the sum of time intervals, within τ, where the system is empty. That is,

$$D = \int_0^{\tau} I\{N(t) = 0\} dt. \tag{3.12}$$

Due to the regenerative property of the system we can write

$$D = \int_0^{U_1} I\{N(t) = 0\} dt + D' \cdot I\{N(U_1) = 0\}$$

where D' has the same distribution of D. Since $E[I\{N(t) = 0\}] = e^{-\Lambda(t)}$ we have

$$E[D] = \frac{E[\int_0^U e^{-\Lambda(t)} dt]}{1 - E[e^{-\Lambda(U)}]}. \tag{3.13}$$

Now, since P_{00} is the fraction of time in which both the system is empty and the server is on vacation, we have

$$P_{00} = \frac{E[D]}{E[\Gamma] + E[\tau]}. \tag{3.14}$$

Using (3.13), (3.10) and (3.3) we finally get

$$\begin{aligned} P_{00} &= \frac{E[\int_0^U e^{-\Lambda(t)}dt]}{E[\Lambda(U)]E[\theta] + E[U]} \\ &= (1-\rho)\frac{E[\int_0^U e^{-\Lambda(t)}dt]}{E[\Lambda(U)]E[B] + (1-\rho)E[U]} \end{aligned} \tag{3.15}$$

For the case where the impatience variable T is exponentially distributed with parameter ξ,

$$P_{00} = \frac{(1-\rho)E\left[\int_0^U e^{-\frac{\lambda}{\xi}(1-e^{-\xi t})}dt\right]}{\frac{\rho}{\xi}[1-U^*(\xi)] + (1-\rho)E[U]}. \tag{3.16}$$

Note that when $\xi \to 0$ we get:

$$P_{00} = \frac{1-\rho}{E[U]} \cdot \frac{1-U^*(\lambda)}{\lambda}.$$

If U is exponentially distributed with parameter γ then this simplifies to

$$P_{00} = \frac{1-\rho}{E[U](\lambda+\gamma)} = (1-\rho)\frac{\gamma}{\lambda+\gamma}$$

In particular, if we substitute the parameters values used in Figure 2 we obtain 0.25, which indeed coincides with what we see in the figure.

3.4.1. *Exponentially distributed vacation and impatience times*

Supposing that $U \sim Exp(\gamma)$, equation (3.16) yields

$$P_{00} = \frac{(1-\rho)\int_{u=0}^{\infty} \gamma e^{-\gamma u}\left(\int_{t=0}^{u} e^{-\frac{\lambda}{\xi}(1-e^{-\xi t})}dt\right)du}{\frac{\rho}{\xi}\left[1-\frac{\gamma}{\gamma+\xi}\right] + (1-\rho)\frac{1}{\gamma}}.$$

By changing order of integration and applying change of variable: $s = 1 - e^{-\xi t}$ in the numerator above, we get

$$\begin{aligned} \int_{s=0}^{1} e^{-\frac{\lambda}{\xi}s}(1-s)^{\gamma/\xi} \cdot \frac{ds}{\xi(1-s)} &= \int_{s=0}^{1} \frac{1}{\xi}(1-s)^{\frac{\gamma}{\xi}-1}e^{-\frac{\lambda}{\xi}s}ds \\ &= \frac{K}{\xi}, \end{aligned}$$

where the last equality comes from equation (2.10). Thus,

$$P_{00} = \frac{(1-\rho)\frac{K}{\xi}}{\frac{\rho}{\gamma+\xi} + \frac{1-\rho}{\gamma}} = \frac{\gamma K}{\xi} \cdot \frac{(\gamma+\xi)(\mu-\lambda)}{\gamma\lambda + (\gamma+\xi)(\mu-\lambda)},$$

which is the expression for P_{00} given by eq (2.19).

3.5. Number of customers at a service completion instant

Let X_n denote the number of customers present *after* the completion of the n-th service. We have

$$X_{n+1} =_d \begin{cases} X_n - 1 + A(B) & \text{if } X_n \geq 1 \\ N(\tau) - 1 + A(B) & \text{if } X_n = 0 \end{cases} \tag{3.17}$$

where $A(t)$ is the number of Poisson arrivals in $(0, t]$. The symbol $=_d$ means "equal in distribution." Therefore, in steady state, the PGF of $X = \lim_{n\to\infty} X_n$ is given by

$$\begin{aligned} \hat{X}(z) = E[z^X] &= E[z^X | X > 0]z^{-1}E[z^{A(B)}](1-P_0) \\ &\quad + E[z^{N(\tau}]z^{-1}E[z^{A(B)}]P_0 \\ &= z^{-1}B^*[\lambda(1-z)]\left[(\hat{X}(z) - P_0)\right] \\ &\quad + E[z^{N(\tau)}]P_0\Big] \end{aligned}$$

where $P_0 = (X = 0)$. Thus,

$$\hat{X}(z) = P_0\frac{E[z^{N(\tau)}] - 1}{z - B^*[\lambda(1-z)]}B^*[\lambda(1-z)] \tag{3.18}$$

where $E[z^{N(\tau)}] = G_{N(\tau}(z)$ is given by (3.5).

To calculate P_0 we substitute $z = 1$ in (3.18) and apply L'Hopital rule to get

$$P_0 = P(X=0) = \frac{1-\rho}{E[N(\tau)]} = (1-\rho)\frac{1-E[e^{-\Lambda(U)}]}{E[\Lambda(U)]} \tag{3.19}$$

where $\rho = \lambda E[B]$ and $E[N(\tau)]$ is given by (3.7).

Note that $P_0 \neq P_{00}$ since P_{00} is the *fraction of time* the system is *empty*, while P_0 is the *relative frequency of occurrences*, *among service completion instants*, when the system becomes empty.

Finally,

$$\begin{aligned} \hat{X}(z) &= (1-\rho)\frac{(G_{N(\tau)}(z) - 1)B^*[\lambda(1-z)]}{E[N(\tau)](z - B^*[\lambda(1-z)])} \\ &= (1-\rho)\frac{\left(E[e^{(1-z)\lambda(U)}] - 1\right)B^*[\lambda(1-z)]}{E[\Lambda(U)](z - B^*[\lambda(1-z)])} \end{aligned} \tag{3.20}$$

When $T = \infty$, the process transforms into the classical *M/G/*1 queue with *multiple server vacations*. Then, $\Lambda(U) = \lambda U$ and $(G_{N(\tau)}(z) - 1)/E[N(\tau)] = (U^*[\lambda(1-z)] - 1)/(\lambda E[U])$.

Equation (3.20) then leads to the known expression (see Levy and Yechiali (1975) and Boxma, et al. (2002))

$$\hat{X}(z) = (1-\rho)\frac{U^*[\lambda(1-z)] - 1}{\lambda E(U)(z - B^*[\lambda(1-z)])}B^*[\lambda(1-z)]). \tag{3.21}$$

3.6. A decomposition representation

Equation (3.20) can be written in a decomposition form, namely,

$$\hat{X}(z) = \hat{L}_{M/G/1}(z) \cdot \frac{G_{N(\tau)}(z) - 1}{(z-1)E[N(\tau)]} \tag{3.22}$$

where $\hat{L}_{M/G/}1(z)$ is the PGF of the system's state (occupancy) at an arbitrary moment in the corresponding regular *M/G/*1 queue, given by

$$\hat{L}_{M/G/1}(z) = (1-\rho)\frac{(z-1)B^*[\lambda(1-z)]}{z - B^*[\lambda(1-z)]}$$

That is, X is the sum of two independent random variables, $L_{M/G/1}$ and Y, where the PGF of Y is given by

$$\hat{Y}(z) = \frac{G_{N(\tau)}(z) - 1}{(z-1)E[N(\tau)]}$$

Note that, because of customer abandonments due to impatience, one cannot use PASTA and Burke's theorem to deduce that the distribution of X, the number of customers at service completions, is the same as the distribution of L, the number of customers at an arbitrary instant.

Now, from (3.22) and using (3.5), it follows that

$$\begin{aligned} E[X] &= E[L_{M/G/1}] + E[Y] \\ &= \left[\frac{\lambda^2 E[B^2]}{2(1-\rho)} + \rho\right] + \frac{E[\Lambda(U)^2]}{2E[\Lambda(U)]} \end{aligned} \tag{3.23}$$

When $T \to \infty$, $\frac{E[\Lambda(U)^2]}{2E[\Lambda(U)]} = \frac{\lambda E[U^2]}{2E[U]} = \Lambda E[R_U]$, where R_U is the residual life time of U.

3.7. Mean number of customers at an arbitrary moment

We now calculate $E[L]$, the mean number of customer in the system at an arbitrary moment. We write

$$\begin{aligned} E[L] &= E[L|\textit{Vacation Period}](1 - P_{(busy)}) \\ &\quad + E[L|\textit{Busy}]P_{(busy)}. \end{aligned} \tag{3.24}$$

Consider a vacation period τ. Let $N(t)$ be the number of customers in the system at time $t \in [0, \tau]$. Let $\Delta = \int_0^\tau N(t)dt$. Then,

$$\Delta = \int_0^{U_1} N(t)dt + \Delta' \cdot I\{N(U_1) = 0\}$$

where Δ' has the same distribution as Δ, and $N(t)$ has a Poisson distribution with parameter $\Lambda(t)$. Taking expectations we get

$$E[\Delta] = \frac{E\left[\int_0^U N(t)dt\right]}{1 - E\left[e^{-\Lambda(U)}\right]}.$$

Then, using (3.3),

$$E[L|\textit{Vacatoin Period}] = \frac{E[\Delta]}{E[\tau]} = \frac{E\left[\int_0^U N(t)dt\right]}{E[U]}.$$

Now, $E\left[\int_0^U N(t)dt\right] = E\left[\Lambda(t)\right]$, implying that

$$E[L|\textit{Vacation Period}] = \frac{E\left[\int_0^U \Lambda(t)dt\right]}{E[U]}. \tag{3.25}$$

In particular, when $U \sim Exp(\gamma)$ and $T \sim Exp(\xi)$, then $E[L|\textit{Vacation Period}] = \frac{\lambda}{\gamma+\xi}$. Now consider a busy period (which starts with $N(\tau) \geq 1$ customers). Then,

$$\begin{aligned} &E[L|\textit{Busy}] \\ &\quad = E_{N(\tau)}\left[\sum_{n=1}^{N(\tau)}\{E[L_{M/G/1}|\textit{Busy}] + (N(\tau) - n)\}\right] \\ &\quad = E[N(\tau)]E[L_{M/G/1}|\textit{Busy}] \\ &\qquad + \frac{1}{2}E_{N(\tau)}[N(\tau)(N(\tau) - 1)]. \end{aligned} \tag{3.26}$$

Clearly, $E[L_{M/G/1}|\textit{Busy}] = E[L_{M/G/1}]/\rho$. Collecting terms, we obtain

$$E[L] = E\left[\int_0^U \Delta(t)dt\right] \frac{1-\rho}{(1-\rho)E[U] + E[\Lambda(U)]E[B]}$$

$$+ \left\{E[N(\tau)]E[L_{M/G/1}]/\rho + \frac{1}{2}E[N(\tau)(N(\tau)-1)]\right\}$$

$$\frac{E[\Lambda(U)]E[B]}{(1-\rho)E[U] + E[\Lambda(U)]E[B]} \quad (3.27)$$

where $E[N(\tau)]$ is given by (3.7), $E[L_{M/G/1}]$ is given by the first term in the right hand side of (3.23) and

$$E[N(\tau)(N(\tau)-1)] = \frac{E[\Lambda(U)^2]}{1 - E\left[e^{-\lambda(U)}\right]}, \quad (3.28)$$

which is derived by differentiating (3.5) twice at $z = 1$. In particular, when $T \sim Exp(\xi)$, and $U \sim Exp(\gamma)$, then

- (i) $E[\Lambda(U)^2] = \frac{\lambda^2}{\xi^2}\left[1 - \frac{2\gamma}{\xi+\gamma} + \frac{\gamma}{2\xi+\gamma}\right]$,
- (ii) $E[e^{-\Lambda(U)}] = \frac{\gamma}{\xi}K$,
- (iii) $E\left[\int_0^U \Delta(t)dt\right] = \frac{\lambda}{\xi}\left[E[U] - \frac{1}{\xi} + \frac{\gamma}{\xi(\xi+\gamma)}\right] = \frac{\lambda}{\gamma(\xi+\gamma)}$.

Substituting the above into (3.27), together with $E[\Lambda(U)] = \lambda/(\xi+\gamma)$ (which we shall establish in (3.31)), yields an explicit solution for $E[L]$ when T and U are distributed exponentially. Further note that it can readily be shown that in the exponential case, and with $E[B] = 1/\mu$, the first term in (3.27) coincides with $E[L_0]$ given in Section 2.4.

3.8. Proportion of customers served

An important performance measure is the proportion of customers served, denoted by $P(served)$. We can write,

$$P(served) = \frac{\text{Expected number of customers served during a cycle}}{\text{Expected number of arrivals during a cycle}} \quad (3.29)$$

Using (3.10) and then (3.3) and (3.7), we get

$$P(served) = \frac{E[\Gamma]/E[B]}{\lambda\left[E[\Gamma] + E[\tau]\right]} = \frac{P_{(busy)}}{\rho} = \frac{1}{\rho + \lambda(1-\rho)\frac{E[U]}{E[\Lambda(U)]}} \quad (3.30)$$

Clearly, $P(served) \to 1$ when $\rho \to 1$. For exponentially distributed vacations and impatience times, where $U \sim Exp(\gamma)$ and $T \sim Exp(\xi)$, we have

$$\lambda(U) = \lambda \int_{y=0}^U e^{-\xi y}dy = \frac{\lambda}{\xi}\left(1 - e^{-\xi U}\right)$$

leading to

$$E[\Lambda(U)] = \int_{u=0}^\infty \frac{\lambda}{\xi}\left(1 - e^{-\xi u}\right)\gamma e^{-\gamma u}du = \frac{\lambda}{\xi+\gamma}. \quad (3.31)$$

It follows that

$$P(served) = \frac{1}{\rho + (1-\rho)\frac{\xi+\gamma}{\gamma}} = \frac{\gamma}{\gamma + (1-\rho)\xi}. \quad (3.32)$$

4. Multiple vacations: *M/M/c* queue with exponentially distributed vacation and impatience times

4.1. The model

We consider a *M/M/c* type queue with $c \geq 1$ server and multiple vacations (see Levy and Yechiali (1976) and recent studies using matrix geometric methods by Zhang and Tian (2003)). Service time, B, of each individual customer is exponentially distributed with mean $1/\mu$. The arrival process is Poisson with rate λ, and we assume that $\lambda < c\mu$. Servers never stay idle in the service station: when a server finishes a service and finds no waiting customers in the queue, he immediately leaves for a vacation. The duration of a vacation by a server is exponentially distributed with mean $1/\gamma$. If there are $1 \leq j \leq c-1$ customers attended by j servers and one of the $c-j$ vacationing servers returns from a vacation to find an empty queue, he immediately leaves for another vacation.

Customers are impatient: an arriving customer who finds that all servers are on vacation activates an independent timer, with exponentially distributed duration, having mean $1/\xi$. If no server becomes available by the moment the timer expires, the customer *abandons* the queue (and the system).

4.2. Balance equations

Let L denote the total number of customers in the system, and let J denote the number of working servers. Then the pair (J, L) defines a continuous-time Markov process with stationary probabilities $P_{jn} = P\{J = j, L = n\}$ $(0 \leq j \leq c, n \geq j)$. The set of (steady state) balance equations for this process is given as follows:

$$\underline{j=0}\quad\begin{cases} n=0 & \lambda P_{00}=\mu P_{11}+\xi P_{01}\\ n\geq 1 & (\lambda+n\xi+c\gamma)P_{0n}\\ & \quad =\lambda P_{0,n-1}+(n+1)\xi P_{0,n+1}\end{cases}\tag{4.1}$$

$$\underline{1\leq j\leq c-1}\quad\begin{cases} n=j & (\lambda+j\mu)P_{jj}\\ & \quad =j\mu P_{j,j+1}+(c-j+1)\gamma P_{j-1,n}\\ & \quad\ +(j+1)\mu P_{j+1,j+1}\\ n>j & [\lambda+j\mu+(c-j)\gamma]P_{jn}\\ & =\lambda P_{j,n-1}+j\mu P_{j,n+1}\\ & +(c-j+1)\gamma P_{j-1,n}\end{cases}\tag{4.2}$$

$$\underline{j=c}\quad\begin{cases} n=c & (\lambda+c\mu)P_{cc}=c\mu P_{c,c+1}\\ & \qquad +\gamma P_{c-1,c}\\ n=c+k & (\lambda+c\mu)P_{c,c+k}=\lambda P_{c,c+k-1}\\ & \qquad +c\mu P_{c,c+k+1}\\ & \qquad +\gamma P_{c-1,c+k}\\ (k=1,2,3,\ldots) & \end{cases}\tag{4.3}$$

4.3. Generating functions

For every j ($j=0,1,\ldots,c$) we define a (partial) Generating Function (GF) and its derivative:

$$G_j(z)=\sum_{n=j}^{\infty}P_{jn}z^n\quad G'_j(z)=\frac{d}{dz}G_j(z).\tag{4.4}$$

By multiplying by z^n each equation for j and n in (4.1), (4.2), and (4.3), and summing over n we obtain the following:

From (4.1), for $\underline{j=0}$, we derive a differential equation for $G_0(z)$:

$$\begin{aligned}\xi(1-z)G'_0(z)&=[\lambda(1-z)+c\gamma]G_0(z)\\ &\quad-(c\gamma P_{00}+\mu P_{11}).\end{aligned}\tag{4.5}$$

From (4.2), for $\underline{1\leq j\leq c-1}$, we obtain a set of linear equations in $G_j(z)$:

$$\begin{aligned}&[(\lambda z-j\mu)(1-z)+(c-j)\gamma z]G_j(z)\\ &\quad-(c-j+1)\gamma zG_{j-1}(z)\\ &=\left[\left((c-j)\gamma z^j-j\mu z^{j-1}\right)P_{jj}\right.\\ &\quad\left.-(c-j+1)\gamma z^{j-1}P_{j-1,j-1}\right]z\\ &\quad+(j+1)\mu P_{j+1,j+1}z^{j+1}\\ &=z^j\left[((c-j)\gamma z-j\mu)P_{jj}\right.\\ &\quad\left.-(c-j+1)\gamma P_{j-1,j-1}+(j+1)\mu zP_{j+1,j+1}\right]\end{aligned}\tag{4.6}$$

Finally, from (4.3), for $\underline{j=c}$, we get

$$\begin{aligned}&[\lambda z(1-z)-c\mu(1-z)]G_c(z)-\gamma zG_{c-1}(z)\\ &\quad=z^c[-c\mu P_{cc}-\gamma P_{c-1,c-1}]\end{aligned}\tag{4.7}$$

Define, for every j, the marginal probability $P_{j\bullet}=\sum_{n=j}^{\infty}P_{jn}=Prob(J=j)$. Then, by substituting $z=1$ in each of the equations (4.5), (4.6) and (4.7), and repeated substitution, we get, for $0\leq j\leq c-1$,

$$(c-j)\gamma[P_{j\bullet}-P_{jj}]=(j+1)\mu P_{j+1,j+1}.\tag{4.8}$$

4.4. Solution of the differential equation for $G_0(z)$

Interesting enough, the differential equation (4.5) is *similar* to equation (2.4), where the only difference is that the term $c\gamma$ replaces the term γ. Therefore, the solution is given by (see (2.8))

$$\begin{aligned}G_0(z)&=P_{00}e^{\frac{\lambda}{\xi}z}\cdot(1-z)^{-\frac{c\gamma}{\xi}}-\frac{A_c}{\xi}e^{\frac{\lambda}{\xi}z}\\ &\quad\cdot(1-z)^{-\frac{c\gamma}{\xi}}\int_{s=0}^{z}(1-s)^{\frac{c\gamma}{\xi}-1}\cdot e^{-\frac{\lambda}{\xi}s}\cdot ds\end{aligned}\tag{4.9}$$

where, similarly to the single server case, $A_c=c\gamma P_{00}+\mu P_{11}$, and

$$\begin{aligned}P_{00}=G_0(0)&=\frac{A_c}{\xi}\cdot\int_{s=0}^{1}(1-s)^{\frac{c\gamma}{\xi}-1}\cdot e^{-\frac{\lambda}{\xi}s}\cdot ds\\ &\stackrel{\text{def}}{=}\frac{A_c}{\xi}\cdot K_c.\end{aligned}\tag{4.10}$$

(K_c can be computed using (2.10) with $c\gamma$ replacing γ. Thus, $\xi P_{00}=A_cK_c=(c\gamma P_{00}+\mu P_{11})K_c$, implying that $A_c=\xi P_{00}/K_c$ and

$$(\xi-c\gamma K_c)P_{00}=\mu K_cP_{11}.\tag{4.11}$$

In order to obtain $G_0(z)$ completely we need to calculate P_{00}. This is accomplished in section 4.7, together with the calculation of *all* P_{jj} and $P_{j\bullet}$.

4.5. Evaluating the mean busy period duration, $E[\Gamma_c]$

In the *M/M/c* with vacations model, when all servers are on vacation ($J = 0$), the equivalent of a single vacation U of the *M/G/*1 queue is the variable V, distributed exponentially with parameter $c\gamma$. Thus, the duration of a Vacation Period is

$$\tau_c = \sum_{i=1}^{k} V_i + V_{k+1} \text{ with probability } \left(\prod_{i=1}^{k} e^{-\Lambda(V_i)}\right) \left(1 - e^{-\Lambda(V_{k+1})}\right).$$

Therefore, similarly to equation (3.2), the LST of τ_c is given by

$$\tau_c^*(s) = \frac{V^*(s) - E[e^{-(sV+\Lambda(V))}]}{1 - E[e^{-(sV+\Lambda(V))}]} \tag{4.12}$$

where $V^*(s) = \frac{c\gamma}{c\gamma+s}$. Also, similarly to (3.3),

$$E[\tau_c] = \frac{E[V]}{1 - E[e^{-\Lambda(V)}]} = \left[c\gamma(1 - E[e^{-\Lambda(V)}])\right]^{-1} \tag{4.13}$$

The PGF, probability mass function and mean of $N(\tau_c)$ are given, respectively, by equations (3.5), (3.6) and (3.7), where $V \sim Exp(c\gamma)$ replaces U.

The equivalent of D (Section 3.4) is a random variable D_c, with mean (similarly to (3.13)),

$$E[D_c] = \frac{E\left[\int_0^V e^{-\Lambda(t)} dt\right]}{1 - E[e^{-\Lambda(V)}]} \tag{4.14}$$

The *busy period*, Γ_c, is defined as the length of time from the moment, starting at τ_c, when one of the c vacationing servers returns from a vacation to find $N(\tau_c)$ waiting customers, until the first moment thereafter in which there are no customers in the system anymore. Thus, as in (3.14),

$$P_{00} = \frac{E[D_c]}{E[\Gamma_c] + E[\tau_c]} \tag{4.15}$$

Once P_{00} is calculated, $E[\Gamma_c]$ is directly determined from (4.15).

4.6. Solution of the set of generating functions

The set of equations (4.6) and (4.7) is a set of c linear equations with $(c+1)$ variables $G_j(z)$, $j = 0, 1, 2, \ldots, c$, where $G_0(z)$ is given by (4.9) and the c unknowns are the PGFs $G_j(z)$, $j = 1, 2, \ldots, c$.

Set

$$a_j(z) = (\lambda z - j\mu)(1 - z) + (c - j)\gamma z \quad (1 \le j \le c) \tag{4.16}$$

and

$$b_j(z) = z^j \left[((c - j)\gamma z - j\mu) P_{jj} - (c - j + 1)\gamma P_{j-1,j-1} + (j+1)\mu z P_{j+1,j+1}\right] \quad (1 \le j \le c - 1) \tag{4.17}$$

$$b_c(z) = z^c \left[-c\mu P_{cc} - \gamma P_{c-1,c-1}\right].$$

Then, equations (4.6) and (4.7) can be written as

$$a_j(z)G_j(z) - (c - j + 1)\gamma z G_{j-1}(z) = b_j(z) \quad (1 \le j \le c) \tag{4.18}$$

where the terms $b_j(z)$ are functions of the probabilities P_{jj} for $j = 1, 2, \ldots, c$.

Define a $c \times c$ square matrix $Q(z)$ as

$$Q(z) = \begin{bmatrix} a_1(z) & 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ -(c-1)\gamma z & a_2(z) & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & -(c-2)\gamma z & a_3(z) & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & -(c-3)\gamma z & a_4(z) & & & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & -2\gamma z & a_{c-1}(z) & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\gamma z & a_c(z) \end{bmatrix}$$

Set the vector $\underline{d}^T(z) = (d_1(z), b_2(z), b_3(z), \ldots, b_c(z))$, where $d_1(z) = b_1(z) + c\gamma z G_0(z)$. Also, set $\underline{g}^T(z) = (G_1(z), G_2(z), \ldots, G_c(z))$. Then, the set (4.18) can be written in a matrix form as

$$Q(z)\underline{g}(z) = \underline{d}(z). \tag{4.19}$$

To obtain $G_j(z)$ we use Cramer's rule and write

$$|Q(z)|G_j(z) = |Q_j(z)|, \quad 0 \le j \le c \tag{4.20}$$

where $|Q|$ denotes the determinant of the matrix Q, and $Q_j(z)$ is a matrix obtained from $Q(z)$ by replacing the jth column with the rhs vector of 4.19, $\underline{d}(z)$. It thus follows that the functions $G_j(z)$ are expressed in terms of the $c+1$ probabilities P_{jj}, $0 \le j \le c$. Once P_{00} is calculated, P_{11} is extracted from (4.11), but we still have to find $c - 1$ additional equations for the unknowns P_{jj}. The following theorem answers this requirement.

Theorem. The $2c$-degree polynomial $|Q(z)|$ has exactly $(c-1)$ distinct roots in the interval $(0, 1)$.

Proof: $|Q(z)| = \prod_{j=1}^{c} a_j(z)$. For $j = c$, $a_c(z) = (1-z)(\lambda z - c\mu)$, implying that the two roots of $a_c(z)$ are $z_c^{(1)} = 1$ and $z_c^{(2)} = c\mu/\lambda > 1$. For $1 \le j \le c-1$, $a_j(z) = -\lambda z^2 + (\lambda + j\mu + (c-j)\gamma)z - j\mu$, implying that $a_j(0) = -j\mu < 0$, $a_j(1) = (c-j)\gamma > 0$, while $a_j(\infty) = -\infty$. Thus, the two roots of $a_j(z)$ are $0 < z_j^{(1)} < 1$ and $1 < z_j^{(2)} < \infty$. Clearly, all $z_j^{(1)}$ roots for $1 \le j \le c-1$ are distinct, which completes the proof. □

Now, since $G_j(z)$ is positive for $0 < z < 1$, then $|Q_j(z)|$ vanishes whenever $|Q(z)|$ does. Thus, each equation $|Q_j(z_j^{(1)})| = 0$, $j = 1, 2, \ldots, c-1$, yields an independent equation in the probabilities P_{jj}, $0 \le j \le c$. Once P_{00} and P_{11} are known, the values of the required P_{jj} for $2 \le j \le c$ are directly obtained.

Given P_{jj} for $0 \le j \le c$, the PGFs $G_j(z)$ are completely determined by equation (4.20).

4.7. Simultaneous calculation of P_{jj}, $P_{j\bullet}$ and $G'_j(1) \equiv E[L_j]$ for $0 \le j \le c$

The calculation of P_{00} for the multiple-vacations *M/M/c* queue is more elaborate than the calculation in the multiple-vacation *M/M/*1 case. Indeed, P_{00} is derived *simultaneously* with all unknowns P_{jj} and $P_{j\bullet}$ as follows.

Equation 4.11 gives a relation between P_{00} and P_{11}, namely,

$$(\xi - c\gamma K_c)P_{00} = \mu K_c P_{11}.$$

Similarly to the derivation of equation (2.13), by using equations (4.9) and (4.10) we get

$$P_{0\bullet} = \frac{\xi}{c\gamma K_c} P_{00}. \tag{4.21}$$

This implies, similarly to (2.14), that

$$c\gamma P_{0\bullet} = c\gamma P_{00} + \mu P_{11} = A_c. \tag{4.22}$$

Also, from (4.5), and similarly to (2.16), we obtain

$$E[L_0] = \frac{\lambda P_{0\bullet}}{c\gamma + \xi}. \tag{4.23}$$

where

$$E[L_j] \equiv G'_j(1) = \sum_{n=j}^{\infty} n P_{jn} \qquad j = 0, 1, 2, \ldots, c. \tag{4.24}$$

From (4.18), after taking derivatives at $z = 1$, we have, for $1 \le j \le c$

$$a'_j(1)P_{j\bullet} + a_j(1)G'_j(1) - [(c-j+1)\gamma](P_{j\bullet} + G'_{j-1}(1)) = b'_j(1) \tag{4.25}$$

where, from (4.16),

$$a'_j(1) = -\lambda + j\mu + (c-j)\gamma \qquad 1 \le j \le c, \tag{4.26}$$

and, from 4.17, for $j = 1, 2, \ldots, c-1$,

$$b'_j(1) = j\left[((c-j)\gamma - j\mu)P_{jj} - (c-j+1)\gamma P_{j-1,j-1} + (j+1)\mu P_{j+1,j+1}\right] + (c-j)\gamma P_{jj} + (j+1)\mu P_{j+1,j+1} \tag{4.27}$$

while, for $j = c$,

$$b'_c(1) = c[-c\mu P_{cc} - \gamma P_{c-1,c-1}]. \tag{4.28}$$

Thus, there are $3(c+1)$ unknowns involved: P_{jj}, $P_{j\bullet}$ and $E[L_j] = G'_j(1)$, for $j = 0, 1, 2, \ldots, c$. Their values are obtained by solving a set of 3(c+1) linear equations as follows:

Relations (4.11), (4.21) and (4.23) give 3 independent equations involving P_{00}, P_{11}, $P_{0\bullet}$ and $E[L_0]$. The set (4.25) yields another c independent equations relating the unknowns $P_{j\bullet}$ (for $1 \le j \le c$) and $E[L_j]$ (for $0 \le j \le c$). There are additional $c-1$ independent equations resulting from the Theorem in section 4.6. The set (4.8) adds c equations relating to the P_{jj}'s and the $P'_{j\bullet}$'s. The last equation is $\sum_{j=0}^{c} P_{j\bullet} = 1$.

All in all we have $(3 + c + (c-1) + c + 1) =)3c + 3$ independent equations to solve for the $3(c+1)$ unknown variables.

5. Single vacation: *M/M/*1 queue with exponentially distributed vacation and impatience times

5.1. The model

We consider now the *M/M/*1 queue where the server takes only a single vacation (see [Levy and Yechiali, 1975)]) at the end of a busy period. If the server returns from a vacation to an empty system he waits dormant to the first arrival thereafter, who opens a busy period. Otherwise, if the queue is positive at the vacation's termination, the server starts a busy period with no delay. Customers are, as before, impatient when they find, upon arrival, that the server is vacationing. Each customer activates his independent, exponentially distributed 'impatience time', T.

5.2. Balance equations and PGFs

The system state is, as before, (J, L). Figure 1 can represent the transition rate diagram of the single vacation case if we add the state (1, 0) to it. The balance equations are now

$$\underline{j=0}\begin{cases} n=0 & (\lambda+\gamma)P_{00} = \xi P_{01} + \mu P_{11} \\ n\geq 1 & (\lambda+n\xi+\gamma)P_{0n} \\ & \quad = \lambda P_{0,n-1} + (n+1)\xi P_{0,n+1} \end{cases} \tag{5.1}$$

$$\underline{j=1}\begin{cases} n=0 & \lambda P_{10} = \gamma P_{00} \\ n\geq 1 & (\lambda+\mu)P_{1n} \\ & \quad = \lambda P_{1,n-1} + \mu P_{1,n+1} + \gamma P_{0n} \end{cases} \tag{5.2}$$

Summing (5.2) over n we obtain, similarly to (2.14), that in this case $\gamma P_{0\bullet} = \mu P_{11}$.

Define the PGFs: $G_j(z) = \sum_{n=0}^{\infty} P_{jn} z^n$, $j = 1, 2$.

Then, similarly to section 2.2, we obtain from (5.1)

$$(\lambda+\gamma)G_0(z) + \xi z G_0'(z) = \lambda z G_0(z) + \xi G_0'(z) + \mu P_{11} \tag{5.3}$$

and from (5.2)

$$\lambda G_1(z) + \mu(G_1(z) - P_{10}) = \lambda z G_1(z) + \frac{\mu}{z}(G_1(z) - P_{11}z - P_{10}) + \gamma G_0(z). \tag{5.4}$$

Equation (5.3) can be written as

$$\xi(1-z)G_0'(z) = [\lambda(1-z)+\gamma]G_0(z) - \mu P_{11} \tag{5.5}$$

implying ($z = 1$) that, indeed,

$$\gamma P_{0\bullet} = \mu P_{11}. \tag{5.6}$$

The differential Eq. (5.5) is very similar to equation (2.4) and therefore its solution is given by equation (2.12) with $G_0(0) = P_{00} = \frac{\mu P_{11}}{\xi} K$ (see (2.11)).

Using the last relation, as well as $\lambda P_{10} = \gamma P_{00}$ from (5.2), equation (5.4) is written as

$$G_1(z)[(\lambda z-\mu)(1-z)] = \gamma z G_0(z) - \left(\mu\frac{\gamma}{\lambda}(1-z) + \frac{\xi}{K}z\right)P_{00}. \tag{5.7}$$

Also, from (5.7), when $z = 1$ we get, as in (2.13),

$$P_{0\bullet} = \frac{\xi}{\gamma K} P_{00}. \tag{5.8}$$

Again, once P_{00} is calculated, both $G_0(z)$ and $G_1(z)$ are completely determined, as well as $P_{0\bullet}$, which is given by (2.13).

We'll calculate P_{00} in the next section.

5.3. Calculation of P_{00}

From equation (5.7),

$$G_1(z) = \frac{\gamma z G_0(z) - \left(\mu\frac{\gamma}{\lambda}(1-z) + \frac{\xi}{K}z\right)P_{00}}{(\lambda z-\mu)(1-z)} \tag{5.9}$$

Applying L'Hopital rule, we get

$$G_1(1) = \frac{\gamma G_0(1) + \gamma G_0'(1) + \left(\frac{\mu\gamma}{\lambda} - \frac{\xi}{K}\right)P_{00}}{\mu-\lambda}.$$

That is,

$$G_0'(1) = E[L_0] = \frac{\mu-\lambda}{\gamma}P_{1\bullet} - P_{0\bullet} - \left(\frac{\mu}{\lambda} - \frac{\xi}{\gamma K}\right)P_{00}. \tag{5.10}$$

On the other hand, from (5.5)

$$E[L_0] = \lim_{z\to 1} G_0'(z) = \frac{-\lambda G_0(1) + \gamma G_0'(1)}{-\xi},$$

implying that

$$E[L_0] = \frac{\lambda}{\gamma+\xi}P_{0\bullet}. \tag{5.11}$$

Equating the two expressions (5.10) and (5.11) for $E[L_0]$, using equation (5.8) and $1 = P_{0\bullet} + P_{1\bullet}$, we obtain the needed expression for P_{00} (by which $P_{0\bullet}$ and $E[L_0]$ are obtained):

$$P_{00}\left[\frac{\lambda\xi}{(\xi+\gamma)K} + \frac{(\mu-\lambda)\xi}{\gamma K} + \frac{\gamma\mu}{\lambda}\right] = \mu - \lambda. \tag{5.12}$$

Note the similarities and differences between (5.12) and 2.18. Clearly, P_{00}(single vacation) $<$ P_{00}(multiple vacation).

5.4. Sojourn times

Let S and S_{jn} be defined as is section 2.5. Then, $E[S]$ is given by expression (2.19), with $E[L_1]$ derived from (5.9). $E[S_{1n}] = (n+1)/\mu$, but this time for $n = 0, 1, 2, \ldots$ rather than for $n \geq 1$. $E[S_{0n}]$ is given by (2.21). Finally, $E[S_{served}]$ is given by (2.22) and (2.23), but with the first sum in (2.22) starting from $n = 0$.

6. Single vacation: *M/G/*1 queue with generally distributed vacation and impatience times

The setting is as in Section 3, but under the single-vacation policy.

6.1. The busy period

A busy period starts either with $N(U) = n \geq 1$ customers (with probability $e^{-\Lambda(U)}(\Lambda(U))^n/n!$ or with a single customer (with probability $e^{-\Lambda(U)}$). Thus, the LST of the busy period is

$$\begin{aligned}\Gamma^*(s) &= E\left[\sum_{n=1}^{\infty}(\theta^*(s))^n e^{-\Lambda(U)}(\Lambda(U))^n/n!\right] \\ &\quad + E\left[\theta^*(s)e^{-\Lambda(U)}\right] \\ &= E\left[e^{-\Lambda(U)}\left(e^{\Lambda(U)\theta^*(s)}-1\right)\right] + \theta^*(s)E\left[e^{-\Lambda(U)}\right] \\ &= E\left[e^{-\Lambda(U)(1-\theta^*(s))}\right] - (1-\theta^*(s))E\left[e^{-\Lambda(U)}\right] \qquad (6.1)\end{aligned}$$

From (6.1) it follows that

$$E[\Gamma] = E[\theta]\left(E[\Lambda(U)] + E\left[e^{-\Lambda(U)}\right]\right) \qquad (6.2)$$

where $E[\theta] = E[B]/(1-\rho)$.

6.2. Calculation of P_{00}, P_{10} and $P_{1\bullet}$

A cycle C consists of a busy period Γ, a single vacation U, and, with probability $E\left[e^{-\Lambda(U)}\right]$, an exponential inter-arrival time with mean $1/\lambda$. Thus,

$$E[C] = E[\Gamma] + E[U] + E\left[e^{-\Lambda(U)}\right](1/\lambda). \qquad (6.3)$$

In the single vacation case the sum D of time intervals, within a vacation U, in which the system is empty is given by

$$D = \int_0^U I\{N(t) = 0\}dt \qquad (6.4)$$

and

$$E[D] = E\left[\int_0^U e^{-\Lambda(t)}dt\right]. \qquad (6.5)$$

The fraction of time in which there are no customers in the system and the server is on vacation is, using (6.5), (6.3) and (6.2),

$$\begin{aligned}P_{00} &= \frac{E[D]}{E[C]} \\ &= (1-\rho)\frac{E\left[\int_0^U e^{-\Lambda(t)}dt\right]}{E[B]\left(E[\Lambda(U)] + E\left[e^{-\Lambda(U)}\right]\right) + (1-\rho)\left(E[U] + E\left[e^{-\Lambda(U)}\right]/\lambda\right)} \qquad (6.6)\end{aligned}$$

while the fraction of time in which the server is ready to serve but there are no customers present in the system is

$$P_{10} = \frac{E\left[e^{-\Lambda(U)}\right]/\lambda}{E[C]} \qquad (6.7)$$

Finally,

$$P_{1\bullet} = \frac{E[\Gamma]}{E[C]}.$$

Remark (Exponentially distributed service, vacation and impatience times).

In the case where $B \sim Exp(\mu)$, $U \sim Exp(\gamma)$ and $T \sim Exp(\xi)$, similarly to the derivation in Section 3.4.1,

$$E\left[\int_{t=0}^U e^{-\Lambda(t)}dt\right] = \frac{K}{\xi}, \quad E\left[e^{-\Lambda(U)}\right] = \gamma\frac{K}{\xi} \qquad (6.8)$$

Using (3.31), it follows that eq.(6.6) reduces to (5.12).

6.3. Number of customers at a service completion instant

Define X_n and P_0 as in section 3.5. Then, for a given U,

$$X_{n+1} =_d \times \begin{cases} X_n - 1 + A(B) & \text{if } X_n \geq 1 \\ N(U)\big|_{N(U)\geq 1} - 1 + A(B) & \text{if } X_n = 0,\ N(U) \geq 1 \\ A(B)\big|_{N(U)=0} & \text{if } X_n = 0,\ N(U) = 0 \end{cases} \qquad (6.9)$$

Thus, in steady state, the PGF of X is given by

$$\begin{aligned}\hat{X}(z) &= (1-P_0)E[z^X|X>0]z^{-1}B^*[\lambda(1-z)] \\ &\quad + P_0 \cdot E[E[z^{N(U)}|N(U)>0]z^{-1}B^*[\lambda(1-z)] \\ &\quad P(N(U)>0) + e^{-\Lambda(U)}B^*[\lambda(1-z)]] \\ &= (\hat{X}(z) - P_0)z^{-1}B^*[\lambda(1-z)] \\ &\quad + P_0 \cdot E\{[\hat{N}(U) - e^{-\Lambda(U)}]z^{-1}B^*[\lambda(1-z)] \\ &\quad + e^{-\Lambda(U)}B^*[\lambda(1-z)]\}\end{aligned}$$

That is,

$$\hat{X}(z) = P_0 \frac{U^*[\lambda(1-z)] - (1-z)E[e^{-\Lambda(U)}] - 1}{z - B^*[\lambda(1-z)]} \times B^*[\lambda(1-z)] \tag{6.10}$$

P_0 is given by

$$P_0 = \frac{1-\rho}{\lambda E[U] + E\left[e^{-\lambda(U)}\right]}. \tag{6.11}$$

Finally,

$$\hat{X}(z) = (1-\rho) \frac{U^*[\lambda(1-z)] - (1-z)E\left[e^{-\Lambda(U)}\right] - 1}{\left(\lambda E[U] + E\left[e^{-\Lambda(U)}\right]\right)(z - B^*[\lambda(1-z)])} \times B^*[\lambda(1-z)]. \tag{6.12}$$

6.4. Proportion of customers served

The proportion of customers served without abandoning the system is defined as in equation (3.29). Then

$$P(served) = \frac{E[\Gamma]/E[B]}{\lambda E[C]} = \frac{P_{(busy)}}{\rho}$$

Using (6.2) and (6.3) we have

$$P(served) = \frac{E[\Lambda(U)] + E\left[e^{-\Lambda(U)}\right]}{(1-\rho)\left[\frac{\rho}{1-\rho}\left(E[\Lambda(U)] + E\left[e^{-\Lambda(U)}\right]\right) + \lambda E[U] + E\left[e^{-\Lambda(U)}\right]\right]}$$
$$= \frac{E[\Lambda(U)] + E\left[e^{-\Lambda(U)}\right]}{\rho E[\Lambda(U)] + E\left[e^{-\Lambda(U)}\right] + \lambda(1-\rho)E[U]} \tag{6.13}$$

Again, when $\rho \to 1$, then $P(served) \to 1$. For $U \sim Exp(\gamma)$ and $T \sim Exp(\xi)$, we have from (3.31) that $E[\Lambda(U)] = \frac{\lambda}{\xi+\gamma}$, and from (6.8) that $E\left[e^{-\Lambda(U)}\right] = \frac{\gamma K}{\xi}$. Thus,

$$P(served) = \frac{\frac{\lambda}{\xi+\gamma} + \frac{\gamma K}{\xi}}{\frac{\rho\lambda}{\xi+\gamma} + \frac{\gamma K}{\xi} + \frac{\lambda(1-\rho)}{\gamma}} \tag{6.14}$$

Comparing (6.14) with (3.32), it follows that

$$P(served|single\ vacation) > P(served|multiple\ vacations).$$

7. Single vacation: *M/M/c* queue with exponentially distributed vacation and impatience times

7.1. The Model

In this section we consider an *M/M/c*-type queue, similar to the one described in section 4.1, but with each server taking only a *single* vacation at a time.

7.2. Balance equations

As in section 4.2, the pair (J, L) defines a continuous-time Markov process where, now, for every j, the state space consists of all $n \geq 0$ for every j, rather than only of $n \geq j$. Specifically, let $P_{jn} = P\{J = j, L = n\}$ $(0 \leq j \leq c; n \geq 0)$. Then, the set of balance equations is given as follows:

$$\underline{j=0}\ \begin{cases} n = 0 & (\lambda + c\gamma)P_{00} = \mu P_{11} + \xi P_{01} \\ n \geq 1 & (\lambda + n\xi + c\gamma)P_{0n} = \lambda P_{0,n-1} \\ & \quad + (n+1)\xi P_{0,n+1} \end{cases} \tag{7.1}$$

$\underline{1 \leq j \leq c-1}$

$$\times \begin{cases} n = 0 & [\lambda + (c-j)\gamma]P_{j0} \\ & = (c-j+1)\gamma P_{j-1,0} \\ & + \mu P_{j+1,1} \\ 1 \leq n \leq j-1 & [\lambda + n\mu + (c-j)\gamma]P_{jn} \\ & = \lambda P_{j,n-1} \\ & + (c-j+1)\gamma P_{j-1,n} \\ & + (n+1)\mu P_{j+1,n+1} \\ n = j & [\lambda + j\mu + (c-j)\gamma]P_{jn} \\ & = \lambda P_{j,n-1} \\ & + (c-j+1)\gamma P_{j-1,n} \\ & + j\mu P_{j+1,n+1} \\ & + (j+1)\mu P_{j+1,n+1} \\ n > j & [\lambda + j\mu + (c-j)\gamma]P_{jn} \\ & = \lambda P_{j,n-1} \\ & + (c-j+1)\gamma P_{j-1,n} \\ & + j\mu P_{j,n+1} \end{cases} \tag{7.2}$$

$\underline{j = c}$

$$\times \begin{cases} n = 0 & \lambda P_{c0} = \gamma P_{c-1,0} \\ 1 \le n \le c-1 & (\lambda + n\mu) P_{cn} \\ & = \lambda P_{c,n-1} + \gamma P_{c-1,n} \\ n \ge c & (\lambda + c\mu) P_{cn} \\ & = \lambda P_{c,n-1} + \gamma P_{c-1,n} + c\mu P_{c,n+1} \end{cases} \tag{7.3}$$

7.3. Generating functions

Similarly to section 4.3 (multiple vacation case) we define, for every $0 \le j \le c$, the (partial) Generating Function $G_j(z)$ and the marginal probability $P_{j\bullet} = G_j(1)$, where

$$G_j(z) = \sum_{n=0}^{\infty} P_{jn} z^n. \tag{7.4}$$

By multiplying by z^n and summing over n we obtain:
From (7.1), for $\underline{j = 0}$:

$$\xi(1-z)G_0'(z) = [\lambda(1-z) + c\gamma]G_0(z) - \mu P_{11}. \tag{7.5}$$

From (7.2), for $\underline{1 \le j \le c-1}$

$$\begin{aligned} &[(\lambda z - j\mu)(1-z) + (c-j)\gamma z]G_j(z) \\ &- (c-j+1)\gamma z G_{j-1}(z) \\ &= -j\mu(1-z)\sum_{n=0}^{j} P_{jn}z^n - (c-j+1)\gamma z P_{j-1,0} \\ &- \mu z \sum_{n=0}^{j} n P_{jn} z^n + \mu \sum_{n=1}^{j+1} n P_{j+1,n} z^n. \end{aligned} \tag{7.6}$$

From (7.3), for $\underline{j = c}$

$$\begin{aligned} &[(\lambda z - c\mu)(1-z)]G_c(z) - \gamma z G_{c-1}(z) = \\ &- \mu z \sum_{n=1}^{c-1} n P_{cn} z^n - c\mu(1-z)\sum_{n=0}^{c-1} P_{cn} z^n - c\mu P_{cc} z^c. \end{aligned} \tag{7.7}$$

The set (7.6) and (7.7) can be written in a matrix form, similarly to (4.19), as

$$Q(z)\underline{g}(z) = \underline{d}(z)$$

where $Q(z)$ is *exactly* the *same* matrix defined in section 4.6 with the *same* $a_j(z)$ for $1 \le j \le c$, and $\underline{d}^T = (d_1(z), b_2(z), b_3(z), ..., b_c(z))$, where, as before, $d_1(z) = b_1(z) + c\gamma z G_0(z)$, *but*,

$$\begin{aligned} b_j(z) = &-j\mu(1-z)\sum_{n=0}^{j} P_{jn} z^n - (c+j+1)\gamma z P_{j-1,0} \\ &- \mu z \sum_{n=0}^{j} n P_{jn} z^n + \mu \sum_{n=1}^{j+1} n P_{j+1,n} z^n \\ &(1 \le j \le c-1). \end{aligned} \tag{7.8}$$

and

$$\begin{aligned} b_c(z) = &-\mu z \sum_{n=1}^{c-1} n P_{cn} z^n - c\mu(1-z)\sum_{n=0}^{c-1} P_{cn} z^n \\ &- c\mu P_{cc} z^c. \end{aligned} \tag{7.9}$$

7.4. Solution of the differential equation and calculation of the unknown probabilities

The differential equation (7.5) is similar to the differential equation (4.5), implying that the former's solution is given by equation (4.9) with the only modification that the term μP_{11} replaces the term $A_c = c\gamma P_{00} + \mu P_{11}$. That is,

$$\begin{aligned} G_0(z) = e^{\frac{\lambda}{\xi}z}(1-z)^{-\frac{c\gamma}{\xi}} \Bigg(&P_{00} - \frac{\mu P_{11}}{\xi} \\ &\times \int_{s=0}^{z} (1-s)^{\frac{c\gamma}{\xi}-1} e^{\frac{-\lambda}{\xi}s} ds \Bigg). \end{aligned} \tag{7.10}$$

Furthermore, similarly to (4.10)

$$P_{00} = \frac{\mu P_{11}}{\xi} \int_0^1 (1-s)^{\frac{c\gamma}{\xi}-1} e^{-\frac{\lambda}{\xi}s} ds = \frac{\mu P_{11}}{\xi} K_c. \tag{7.11}$$

From equations (7.10) and (7.11), similarly to the derivation of (2.13) and (4.21),

$$P_{0\bullet} = \frac{\xi}{c\gamma K_c} P_{00} = \frac{\mu}{c\gamma} P_{11} \tag{7.12}$$

which can also be obtained directly from (7.5) by setting $z = 1$.

By setting $z = 1$ in equations (7.3) and (7.7) we get, respectively, for $\underline{1 \le j \le c-1}$

$$\begin{aligned} &(c-j)\gamma P_{j\bullet} - (c-j+1)\gamma P_{j-1,\bullet} \\ &\quad = (c-j+1)\gamma P_{j-1,0} - \mu \sum_{n=1}^{j} n P_{jn} \\ &\qquad + \mu \sum_{n=1}^{j+1} n P_{j+1,n} \end{aligned} \tag{7.13}$$

and for $\underline{j = c}$

$$P_{c\bullet} - \gamma P_{c-1,\bullet} = -\mu \sum_{n=1}^{c-1} n P_{cn} - c\mu P_{cc}. \qquad (7.14)$$

Equations (7.12), (7.13) and (7.14) comprise a set of $(c+1)$ equations involving the $(c+1)$ marginal probabilities $P_{j\bullet}$ $(0 \leq j \leq c)$ and the 'boundary' probabilities $\{P_{jn}\}$ for $0 \leq j \leq c$, $0 \leq n \leq j$, excluding P_{00}.

From (7.5), letting $z \to 1$ and using L'Hopital's rule, we get

$$E[L_0] = \frac{-\lambda G_0(1) + c\gamma G_0'(1)}{-\xi} = \frac{\lambda P_{0\bullet}}{c\gamma + \xi}. \qquad (7.15)$$

Since equation (4.18) holds here too (although with different $b_j(z)$, see (7.8) and (7.9)), equations (4.25) and (4.26) hold here also *but*, instead of (4.27) and (4.28) we have, respectively,

$$b_j'(1) = j\mu \sum_{n=0}^{j} P_{jn} - (c+j+1)\gamma P_{j-1,0} - \mu \sum_{n=1}^{j} n P_{jn} - \mu \sum_{n=1}^{j} n^2 P_{jn} + \mu \sum_{n=1}^{j+1} n^2 P_{j+1,n} \qquad (7.16)$$

and

$$b_c'(1) = -\mu \sum_{n=1}^{c-1} n P_{cn} - \mu \sum_{n=1}^{c} n^2 P_{cn} + c\mu \sum_{n=0}^{c-1} P_{cn}. \qquad (7.17)$$

The unknown probabilities are $\{P_{jn}\}$ for $0 \leq j \leq c$, $0 \leq n \leq c$; $\{P_{j\bullet}\}$ for $0 \leq j \leq c$; and $E[L_j]$ for $0 \leq j \leq c$. Altogether there are $((c+1)+1)(c+1)/2 + 2(c+1) = (c^2 + 7c + 6)/2$ unknowns. They are solved by the combination of equations (7.11), (7.12), (7.15); (4.25) with the updated $b_j'(1)$; $c-1$ equations are derived from the roots of $Q(z)$; equations (7.13) and (7.14); $(c+1)c/2$ balance equations for the 'boundary' probabilities $\{P_{jn}\}$, where $1 \leq j \leq c, 0 \leq n \leq j-1$; and the 'total probability' equation $\sum_{j=1}^{c} P_{j\bullet} = 1$. That is, there are $1+1+1+c+(c-1)+c+(c+1)c/2+1 = 3c+3+(c+1)c/2 = (c^2+7c+6)/2$ equations.

Compared with the *M/M/c* multiple vacation case, the solution for the single vacation model requires additional $(c+1)c/2$ equations for the unknown probabilities $\{P_{jn}\}$ where $1 \leq j \leq c, 0 \leq n \leq j-1$.

8. Conclusion

We have introduced and analyzed in this paper a new type of impatience behavior in which customers become impatient (and may leave the system) when the server goes on vacation. This is in contrast with previously studied impatience behavior which did not consider server vacations and where customers may become impatient when the number of customers or the amount of workload queued in front of them is large.

We analyzed both the single and the multiple vacations cases. We studied both Markovian models (the *M/M/*1 and the *M/M/c* queues with exponentially distributed vacation and impatience times) as well as the *M/G/*1 case with generally distributed impatience and vacation times. For the *M/M/*1 case, we derived explicit expressions for the PGF of the number of customers (conditioned on the server state) in the system. For the *M/G/*1 model, we obtained the PGF of the number of customers at various embedded instants (end of service, start of a busy period); we calculated the mean number of customers in the system at an arbitrary moment, and we derived other performance measures, including the proportion of customers being served. For the *M/M/c* queue we derived explicit expressions for the PGF of the number of customers (conditioned on the server state) in terms of several constants, which are derived by finding the roots of a $2c$-degree polynomial being the determinant of a certain matrix whose entries are functions of the system's parameters.

In many queueing problems, working with the PGF of number of customers in the system (or with the LST of the workload) allows one to transform difference equations (resp. differential equations) that represent the balance equations for the steady state probabilities, into algebraic equations. Interestingly, in the Markovian models we introduced in this paper, the difference equations describing the balance equations does not transform into algebraic equations for the PGF: the PGF is characterized by a solution of a differential equation which we solved explicitly for each model.

References

Altman, E. and Borovkov, A. A., "On the stability of retrial queues," Queueing Systems, 26 (1997) 343–363.

Baccelli, F., Boyer, P. and Hebuterne, G., "Single-Server Queues with Impatient Customers," Advances in Applied Probability, 16 (1984) 887–905.

Bonald, T. and Roberts, J., "Performance modeling of elastic traffic in overload," *ACM Sigmetrics*, Cambridge, MA, USA, (2001) 342–343.

Boxma, O.J. and de Waal, P.R. "Multiserver Queues with Impatient Customers," ITC, 14 (1994) 743–756.

Boxma, O.J., Schlegel, S. and Yechiali, U., "A Note on the *M/G/*1 Queue with a Waiting Server, Timer and Vacations," American Mathematical Society Translations, Series 2, 207 (2002) 25–35.

Daley, D.J., "General Customer Impatience in the Queue $GI/G/1$," J. Applied Probability, 2 (1965) 186–205.

Gans, N., Koole, G. and Mandelbaum A., "Telephone call centers: Tutorial, review, and research prospects," Manufacturing and Service Operations Management, 5 (2003) 79–141.

Kleinrock, L. (1975), *Queueing Systems Volume I: Theory*, J. Wiley & Sons, New York.

Levy, Y. and Yechiali, U., “Utilization for the Idle Time in an *M/G/*1 Queuing System,” Management Science, 22 (1975) 202–211.
Levy, Y. and Yechiali, U., “An *M/M/s* Queue with Servers’ Vacations,” Canadian J. of Operational Research and Information Processing, 14 (1976) 153–163.
Palm, C., “Methods of Judging the Annoyance Caused by Congestion,” Tele, 4 (1953) 189–208.
Palm, C., “Research on Telephone Traffic Carried by Full Availability Groups,” Tele, Vol. 1, 107 (1957). (English translation of results first published in 1946 in Swedish in the same journal, which was then entitled Tekniska Meddelanden fran Kungl. Telegrfstyrelsen).
Takacs, L., “A Single-Server Queue with Limited Virtual Waiting Time,” J. Applied Probability, 11 (1974) 612–617.
Takacs, L. “Introduction to the Theory of Queues,” Oxford University Press, New York (1962).
Van Houdt, B., Lenin, R.B. and Blondia, C., “Delay Distribution of (Im)Patient customers in a Discrete Time D–$MAP/PH/1$ Queue with Age-Dependent Service Times,” Queueing Systems, 45 (2003) 59–73.
Yechiali, U., “On the $M^X/G/1$ Queue with a Waiting Server and Vacations,” Sankhya, 66 (2004) 1–17.
Zhang, Z.G. and Tian, N., “Analysis of Queueing Systems with Synchronous Vacations of Partial Servers,” Perfor. Eval., 52(4) (2003) 269–282.
Zhang, Z.G. and Tian, N., “Analysis of Queueing Systems with Synchronous Single Vacation for Some Servers,” Queueing Systems, 45(2) (2003) 161–175.