European Journal of Operational Research 201 (2010) 247–258

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

Queues with slow servers and impatient customers

Nir Perel [a,*], Uri Yechiali [a,b]

[a] Department of Statistics and Operations Research, School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel
[b] Afeka Academic College of Engineering, Afeka, Tel Aviv 69107, Israel

ARTICLE INFO

Article history:
Received 6 August 2008
Accepted 18 February 2009
Available online 27 February 2009

Keywords:
Alternating queue
$M/M/1$
$M/M/c$
$M/M/\infty$
Slow server(s)
Impatient customers
Abandonment

ABSTRACT

We study $M/M/c$ queues ($c = 1$, $1 < c < \infty$ and $c = \infty$) in a 2-phase (fast and slow) Markovian random environment, with impatient customers. The system resides in the fast phase (phase 1) an exponentially distributed random time with parameter η and the arrival and service rates are λ and μ, respectively. The corresponding parameters for the slow phase (phase 0) are γ, λ_0, and μ_0 ($\leqslant \mu$). When in the slow phase, customers become impatient. That is, each customer, upon arrival, activates an individual timer, exponentially distributed with parameter ξ. If the system does not change its environment from 0 to 1 before the customer's timer expires, the customer abandons the queue never to return.

We concentrate on deriving analytic solutions to the queue-length distributions. We derive, for each case of c, the corresponding probability generating function, and calculate the mean queue size. Several extreme cases are investigated and numerical results are presented.

1. Introduction

Models with customers' impatience in queues have been studied by various authors in the past, where the source of impatience was either a long wait already experienced in the queue, or a long wait anticipated by a customer upon arrival. There is an extensive literature on this kind of models and we refer the reader to [3,6,8,10,15,16,18,19] and references there. However, recently, Altman and Yechiali [1,2] analyzed models where customers become impatient *only* when the server(s) is (are) on 'vacation' and unavailable for service. That is, customer's impatience arises only when, upon arrival, no servers are ready to serve. The $M/M/1$, $M/G/1$, $M/M/c$ and $M/M/\infty$ queues were investigated and various performance measures calculated. Yechiali [21] then analyzed $M/M/c$ systems (for $1 \leqslant c \leqslant \infty$) that suffer *disastrous breakdowns*, resulting in the loss of all customers present (e.g. all running and waiting sessions). While a repair process is taking place, the flow of new customers continues but they become *impatient* since no server is available. Recently, Martin and Mitrani [12] studied an $M/M/1$ model, with an intermittently available server (the server goes through breakdowns and repairs). While the server is unavailable, the stream of new arrivals continues while customers may abandon the system. The main difference between [12,21] is that in the former abandonments occur also when the server is active.

In this work, we examine the case where customers' impatience is due to a *slow* service rate. For example, the server might be occupied with other, higher priority, tasks, but is not totally unavailable. In other words, the server keeps on working but with a slower rate than before. In order to analyze the model, we consider an $M/M/c$ queue ($c = 1,\ 1 < c < \infty,\ c = \infty$) operating in a 2-phase random environment. That is, the system oscillates between two phases, denoted by 0 and 1, residing in phase (environment) j, $j = 0, 1$, an exponentially distributed random time with parameters γ and η, respectively. Under environment 1, the Poisson arrival rate is λ and the service time is exponentially distributed with parameter μ. However, when operating under environment 0, the Poisson arrival rate is λ_0, the service rate drops to $\mu_0 \leqslant \mu$, and customers become impatient. That is, each customer, upon arrival, activates an individual timer, exponentially distributed with parameter ξ. If the system does not change its environment from 0 to 1 before the customer's timer expires, the customer abandons the queue never to return.

Queues in random environment have been long studied in the literature. We mention works by Yechiali and Naor [22], Yechiali [20], Neuts [13], O'Cinneide and Purdue [14], Baykal-Gursoy and Xiao [4] and Gupta et al. [11]. We indicate that, if customers are patient and do not leave the system when the server becomes slower, the current model reduces to the original $M/M/1$ queue in 2-phase random environment studied in [22,11]. All the above random-environment models can be formulated as a level (environment) dependent quasi

* Corresponding author.
E-mail addresses: perelnir@post.tau.ac.il (N. Perel), uriy@afeka.ac.il, uriy@post.tau.ac.il (U. Yechiali).

0377-2217/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2009.02.024

birth-and-death processes. Bright and Taylor [7] presented such a formulation, using Neuts Matrix Geometric approach, and proposed algorithms for computing the equilibrium queue distributions. However, their methods are based on a *truncation* of the possible number of levels (=number of customers in the system). In contrast, using classical approach of probability generating functions, we present analytic formulas for various characteristics of the processes considered.

The paper consists of the following models: In Section 2 we consider the $M/M/1$ queue. After deriving the system's steady-state balance equations we obtain and solve a differential equation for $G_0(z)$, the (partial) probability generating function (PGF) of the queue size when the server is slow. $G_0(z)$ is a function of P_{00} and P_{10}, the fraction of time the system is empty in each environment, respectively. The calculation of those probabilities completes the derivation of $G_0(z)$. Following that, the PGF of environment $j = 1$, $G_1(z)$, is derived by using a direct relation between the two PGFs. The mean total number of customers in the system is then calculated. Various extreme cases, resulting when some of the parameters approach 0 or ∞, are examined, and numerical examples are presented. Section 3 deals with the $M/M/c$ model, for a finite c. We derive the PGF $G_0(z)$ by solving a differential equation and calculating the required $2c$ boundary probabilities for the complete representation of $G_0(z)$, and then of $G_1(z)$. Performance measures similar to those of Section 2 are calculated. In Section 4 we study the $M/M/\infty$ case. In order to derive the corresponding PGFs, we utilize the (no abandonment) 2-phase model studied by Baykal-Gursoy and Xiao [4], where an $M/M/\infty$ queueing system subject to partial failures is investigated.

2. The single server case

2.1. The model

Consider an $M/M/1$ type queue operating in a 2-phase random environment, where the underlying process is a 2-state continuous-time Markov chain as described in the introduction. It is assumed that the underlying 2-phase environment Markov process is independent of the arrival, service and impatience processes, and we investigate the system in steady-state.

Let L denote the total number of customers present in the system and let J denote the server's environment (0 or 1). Then the pair (J, L) defines a continuous-time Markov process with transition-rate diagram as shown in Fig. 2.1.

2.2. Balance equations and generating functions

Let $P_{jn} = P(J = j, L = n)$ $(j = 0, 1, n = 0, 1, 2, \ldots)$ denote the steady-state probabilities of the random process (J, L). Then, the set of balance equations is given by

$$\underline{j=0} \quad \begin{cases} n = 0 : (\lambda_0 + \gamma)P_{00} = \eta P_{10} + (\mu_0 + \xi)P_{01}, \\ n \geqslant 1 : (\lambda_0 + \gamma + \mu_0 + n\xi)P_{0n} = \lambda_0 P_{0,n-1} + \eta P_{1n} + (\mu_0 + (n+1)\xi)P_{0,n+1}, \end{cases} \tag{2.1}$$

$$\underline{j=1} \quad \begin{cases} n = 0 : (\lambda + \eta)P_{10} = \gamma P_{00} + \mu P_{11}, \\ n \geqslant 1 : (\lambda + \mu + \eta)P_{1n} = \lambda P_{1,n-1} + \gamma P_{0n} + \mu P_{1,n+1}. \end{cases} \tag{2.2}$$

For $j = 0, 1$ let $P_{j\bullet} = \sum_{n=0}^{\infty} P_{jn} = P(J = j)$. Then, by summing (2.2) over n we get

$$(\lambda + \eta)P_{1\bullet} + \mu(P_{1\bullet} - P_{10}) = \lambda P_{1\bullet} + \mu(P_{1\bullet} - P_{10}) + \gamma P_{0\bullet},$$

which leads to

$$\eta P_{1\bullet} = \gamma P_{0\bullet}.$$

Since $P_{0\bullet} + P_{1\bullet} = 1$, we get

$$P_{0\bullet} = \frac{\eta}{\gamma + \eta}, \quad P_{1\bullet} = \frac{\gamma}{\gamma + \eta}. \tag{2.3}$$

Eq. (2.3) can also be obtained by taking horizontal cuts between the two environments in Fig. 2.1. Clearly, (2.3) can be derived directly by considering the environment fluctuations as an alternating renewal process.

Now, define the (partial) probability generating functions (PGFs)

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

By multiplying each equation for n in (2.1) by z^n, respectively, summing over n and rearranging terms we get

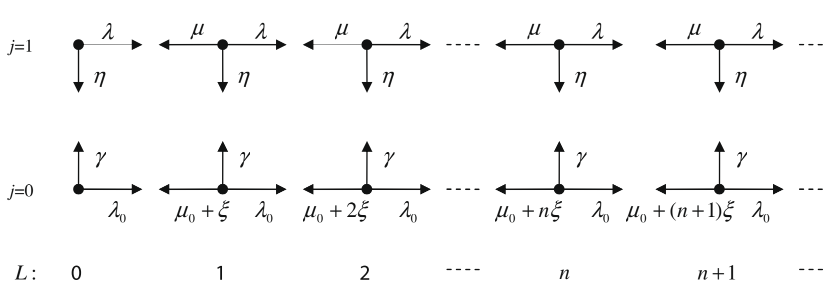

Fig. 2.1. Transition-rate diagram for the $M/M/1$ case.

$$G_0'(z)[\xi(1-z)z] = G_0(z)[(\lambda_0 z - \mu_0)(1-z) + \gamma z] - \eta z G_1(z) + \mu_0(1-z)P_{00}, \tag{2.4}$$

where $G_0'(z) = \frac{d}{dz}G_0(z)$.

Similarly, using (2.2) we obtain

$$G_1(z)[(\lambda z - \mu)(1-z) + \eta z] = \gamma z G_0(z) - \mu(1-z)P_{10}. \tag{2.5}$$

Define

$$\alpha(z) = (\lambda_0 z - \mu_0)(1-z) + \gamma z, \quad \beta(z) = (\lambda z - \mu)(1-z) + \eta z.$$

Then, (2.5) can be written as

$$G_1(z) = \frac{\gamma z G_0(z) - \mu(1-z)P_{10}}{\beta(z)}. \tag{2.6}$$

Substituting (2.6) in (2.4) leads to the following differential equation:

$$G_0'(z) - \frac{\alpha(z)\beta(z) - \eta\gamma z^2}{\xi z(1-z)\beta(z)} G_0(z) = \frac{\eta\mu z P_{10} + \mu_0 \beta(z) P_{00}}{\xi\beta(z)z}. \tag{2.7}$$

Before solving the differential equation (2.7), let us examine $\beta(z)$. The roots z_1, z_2 of the quadratic polynomial $\beta(z) = -\lambda(z-z_1)(z-z_2)$ are

$$z_{1,2} = \frac{\lambda + \mu + \eta \mp \sqrt{(\lambda+\mu+\eta)^2 - 4\lambda\mu}}{2\lambda},$$

where $z_1 \cdot z_2 = \frac{\mu}{\lambda}$ and $z_1 + z_2 = \frac{\lambda+\mu+\eta}{\lambda}$. Furthermore, z_1 as a function of the parameter η represents the Laplace Stieltjes transform of the busy period in a regular $M/M/1$ queue with arrival rate λ and service rate μ. In addition, $\beta(0) = -\mu < 0, \beta(1) = \eta > 0$, and $\beta(+\infty) < 0$. Therefore $z_2 > 1 > z_1 > 0$.

2.3. Solution of the differential equation (2.7)

Define the coefficient of $G_0(z)$ in (2.7) as $f(z) = -\frac{\alpha(z)\beta(z) - \eta\gamma z^2}{\xi(1-z)z\beta(z)}$. $f(z)$ can be represented as $f(z) = -\frac{\lambda_0}{\xi} + \frac{\mu_0}{z\xi} + \frac{\gamma}{\xi}\left(\frac{M}{z-z_1} + \frac{N}{z_2-z}\right)$, where

$$M = \frac{\frac{\mu}{\lambda} - z_1}{z_2 - z_1} = \frac{z_1 z_2 - z_1}{z_2 - z_1} = \frac{z_1(z_2-1)}{z_2 - z_1} > 0, \quad N = \frac{\frac{\mu}{\lambda} - z_2}{z_2 - z_1} = \frac{z_1 z_2 - z_2}{z_2 - z_1} = \frac{z_2(z_1-1)}{z_2 - z_1} < 0.$$

In order to solve (2.7), we multiply its both sides by $e^{\int f(z)dz}$ (see [5, p. 30]). Therefore,

$$\int f(z)dz = -\frac{\lambda_0 z}{\xi} + \frac{\mu_0}{\xi}\ln(z) + \frac{\gamma M}{\xi}\ln|z-z_1| - \frac{\gamma N}{\xi}\ln(z_2 - z),$$

and

$$e^{\int f(z)dz} = e^{-\frac{\lambda_0 z}{\xi}} z^{\frac{\mu_0}{\xi}} |z-z_1|^{\frac{\gamma M}{\xi}} (z_2 - z)^{-\frac{\gamma N}{\xi}}. \tag{2.8}$$

Multiplying both sides of (2.7) by (2.8) gives the following:

$$\frac{d}{dz}\left[\left(e^{-\frac{\lambda_0 z}{\xi}} z^{\frac{\mu_0}{\xi}} |z-z_1|^{\frac{\gamma M}{\xi}} (z_2-z)^{-\frac{\gamma N}{\xi}}\right) G_0(z)\right] = \left(\frac{\eta\mu P_{10} z + \mu_0\beta(z)P_{00}}{\xi\beta(z)z}\right) e^{-\frac{\lambda_0 z}{\xi}} z^{\frac{\mu_0}{\xi}} |z-z_1|^{\frac{\gamma M}{\xi}} (z_2-z)^{-\frac{\gamma N}{\xi}}. \tag{2.9}$$

It seems convenient to split Eq. (2.9) into two intervals, as follows. Define

$$k_1(z) = e^{-\frac{\lambda_0 z}{\xi}} z^{\frac{\mu_0}{\xi}} (z_1 - z)^{\frac{\gamma M}{\xi}} (z_2 - z)^{-\frac{\gamma u N}{\xi}}, \quad z \leqslant z_1,$$
$$k_2(z) = e^{-\frac{\lambda_0 z}{\xi}} z^{\frac{\mu_0}{\xi}} (z - z_1)^{\frac{\gamma M}{\xi}} (z_2 - z)^{-\frac{\gamma N}{\xi}}, \quad z \geqslant z_1,$$

then Eq. (2.9) can be written as a set of two equations,

$$\begin{cases} \frac{d}{dz}[k_1(z)G_0(z)] = \left(\frac{\eta\mu P_{10} z + \mu_0\beta(z)P_{00}}{\xi\beta(z)z}\right) k_1(z), & z \leqslant z_1, \\ \frac{d}{dz}[k_2(z)G_0(z)] = \left(\frac{\eta\mu P_{10} z + \mu_0\beta(z)P_{00}}{\xi\beta(z)z}\right) k_2(z), & z \geqslant z_1. \end{cases} \tag{2.10}$$

Integrating the upper part of (2.10) from 0 to z $(z \leqslant z_1)$ and the lower part from z_1 to z and rearranging terms gives

$$G_0(z) = \begin{cases} \dfrac{\frac{\eta\mu}{\xi} P_{10} \int_0^z \frac{k_1(x)}{\beta(x)} dx + \frac{\mu_0}{\xi} P_{00} \int_0^z \frac{k_1(x)}{x} dx}{k_1(z)}, & z \leqslant z_1, \\ \dfrac{\frac{\eta\mu}{\xi} P_{10} \int_{z_1}^z \frac{k_2(x)}{\beta(x)} dx + \frac{\mu_0}{\xi} P_{00} \int_{z_1}^z \frac{k_2(x)}{x} dx}{k_2(z)}, & z \geqslant z_1. \end{cases} \tag{2.11}$$

Eq. (2.11) expresses $G_0(z)$ in terms of $G_0(0) = P_{00}$ (the proportion of time the server is in environment 0 and there are no customers in the system) and in terms of $G_1(0) = P_{10}$ (the proportion of time the server is in environment 1 and the system is empty). Thus, once P_{00} and P_{10} are calculated, $G_0(z)$ is completely determined, and $G_1(z)$ is obtained by using Eq. (2.6).

2.4. Derivation of $P_{00}, P_{10}, E[L_0]$ and $E[L_1]$

We write

$$\frac{\eta}{\eta+\gamma} = P_{0\bullet} = G_0(1) = \frac{\frac{\eta\mu}{\xi} P_{10} \int_{z_1}^{1} \frac{k_2(x)}{\beta(x)} dx + \frac{\mu_0}{\xi} P_{00} \int_{z_1}^{1} \frac{k_2(x)}{x} dx}{k_2(1)}.$$

This gives

$$\frac{\eta\xi k_2(1)}{\eta+\gamma} = \eta\mu P_{10} \int_{z_1}^{1} \frac{k_2(x)}{\beta(x)} dx + \mu_0 P_{00} \int_{z_1}^{1} \frac{k_2(x)}{x} dx. \tag{2.12}$$

Next, by setting $z = z_1$ in the upper part of Eq. (2.11), we get

$$\frac{\eta\mu}{\xi} P_{10} \int_{0}^{z_1} \frac{k_1(x)}{\beta(x)} dx + \frac{\mu_0}{\xi} P_{00} \int_{0}^{z_1} \frac{k_1(x)}{x} dx = 0. \tag{2.13}$$

Since both numerator and denominator vanish at $z = z_1$.

Define

$$S = \int_0^{z_1} \frac{k_1(x)}{\beta(x)} dx, \quad T = \int_0^{z_1} \frac{k_1(x)}{x} dx, \quad U = \int_{z_1}^{1} \frac{k_2(x)}{\beta(x)} dx, \quad V = \int_{z_1}^{1} \frac{k_2(x)}{x} dx,$$

then, by the definitions of $k_1(z)$, $k_2(z)$ and $\beta(z)$, it follows that $T, U, V > 0$ and $S < 0$.

From (2.13) we get

$$P_{10} = -\frac{\mu_0 P_{00} T}{\eta\mu S}. \tag{2.14}$$

Substituting (2.14) in (2.12) yields

$$P_{00} = \frac{\eta\xi k_2(1) S}{\mu_0(\eta+\gamma)(SV - TU)}. \tag{2.15}$$

Finally, from (2.14) we get

$$P_{10} = -\frac{\xi k_2(1) T}{\mu(\eta+\gamma)(SV - TU)}. \tag{2.16}$$

Notice that $S < 0$ and $SV - TU < 0$, so P_{00} and P_{10} are positive.

One can show formally that the system is ergodic. Intuitively, we indicate that the system is always stable since, with any set of parameters $\lambda > 0, \mu > 0, \lambda_0 \geqslant 0, \mu_0 \geqslant 0, \gamma > 0, \eta > 0, \xi > 0$, the abandonment process, whose overall rate increases with L, prevents explosion. Alternatively, the system is stable if and only if P_{00} and P_{10} are positive, which always holds for the above set of parameters.

We now calculate mean queue sizes. Employing vertical cuts in Fig. 2.1 gives

$$\lambda_0 P_{0n} + \lambda P_{1n} = \mu P_{1,n+1} + (\mu_0 + (n+1)\xi) P_{0,n+1}, \quad n \geqslant 0. \tag{2.17}$$

Summing (2.17) over n yields

$$\lambda_0 P_{0\bullet} + \lambda P_{1\bullet} = \mu(P_{1\bullet} - P_{10}) + \mu_0(P_{0\bullet} - P_{00}) + \xi \sum_{n=0}^{\infty} (n+1) P_{0,n+1}. \tag{2.18}$$

Define $G'_j(z)|_{z=1} = E[L_j] = \sum_{n=0}^{\infty} n P_{jn}, j = 0, 1$. Then, Eq. (2.18) is written as

$$\lambda_0 P_{0\bullet} + \lambda P_{1\bullet} = \mu(P_{1\bullet} - P_{10}) + \mu_0(P_{0\bullet} - P_{00}) + \xi E[L_0]. \tag{2.19}$$

Eq. (2.19) simplify testifies that the mean arrival rate (left hand side) equals the sum of the effective service rate and the abandonment rate. Thus

$$E[L_0] = \frac{\lambda_0 P_{0\bullet} + \lambda P_{1\bullet} - \mu(P_{1\bullet} - P_{10}) - \mu_0(P_{0\bullet} - P_{00})}{\xi}. \tag{2.20}$$

By defining $\hat{\lambda} = \lambda_0 P_{0\bullet} + \lambda P_{1\bullet}$, $\hat{\mu} = \mu_0 P_{0\bullet} + \mu P_{1\bullet}$, Eq. (2.20) can be written as

$$E[L_0] = \frac{\hat{\lambda} - \hat{\mu} + \mu P_{10} + \mu_0 P_{00}}{\xi}. \tag{2.21}$$

Differentiation of $G_1(z)$ in (2.6), setting $z = 1$ and using (2.21) give

$$E[L_1] = \frac{\gamma(\hat{\lambda} - \hat{\mu}) + \xi(\lambda - \mu) P_{1\bullet} + \gamma\mu_0 P_{00} + \mu P_{10}(\gamma + \xi)}{\xi\eta}. \tag{2.22}$$

The mean number of customers in the system, $E[L]$, is given by $E[L] = E[L_0] + E[L_1]$.

2.4.1. Numerical examples

Example 1. Consider the following set of parameter values:

$$\lambda = 4 < \mu = 7, \quad \lambda_0 = 2 < \mu_0 = 5, \quad \gamma = 2, \quad \eta = 2, \quad \xi = 1.$$

With the aid of "Maple", we get

$$T = 0.0019837, \quad S = -0.0008534, \quad U = 0.00473321, \quad V = 0.00687647, \quad k_2(1) = 0.05479,$$

resulting in

$$P_{00} = 0.3064, \quad P_{10} = 0.2544, \quad P_{0\bullet} = P_{1\bullet} = 0.5, \quad E[L_0] = 0.3131, \quad E[L_1] = 0.4536, \quad E[L] = 0.7667.$$

Example 2. Next consider the values $\lambda = 7 > \mu = 4, \lambda_0 = 5 > \mu_0 = 2, \gamma = 5, \eta = 2$ and a small abandonment rate $\xi = 0.1$. The calculations lead to

$$T = 5.724 \times 10^{-21}, \quad S = -4.881 \times 10^{-22}, \quad U = 5.035 \times 10^{-27}, \quad V = 8.735 \times 10^{-27}, \quad k_2(1) = 5.854 \times 10^{-38},$$
$$P_{00} = 1.234 \times 10^{-14}, \quad P_{10} = 3.617 \times 10^{-14}, \quad P_{0\bullet} = \frac{2}{7}, \quad P_{1\bullet} = \frac{5}{7}, \quad E[L_0] = 30, \quad E[L_1] = 76.0714, \quad E[L] = 106.0714.$$

That is, even with $\lambda > \mu$, $\lambda_0 > \mu_0$, and with small ξ, the system does not explode.

2.5. Extreme cases

We denote by $G_j^{(i)}(z), P_{jn}^{(i)}, P_{j\bullet}^{(i)}, E[L_j^{(i)}], E[L^{(i)}], k_j^{(i)}(z)$ for $j = 0, 1$ the PGFs, steady-state probabilities, expected values and the functions k_j corresponding to the following extreme cases, where $i = 1, 2, \ldots, 7$.

Since it is not a straightforward procedure to derive the results for the following extreme cases directly from the general case, we treat each case by itself.

1. $\mu_0 \to 0$.

Assume that $\mu_0 \to 0$. That is, when $j = 0$, no service is rendered. The system in this case alternates between a regular $M/M/1$ queue and an $M/M/\infty$-type queue in which the service rate is replaced by the abandonment rate.

In this case, $\beta(z), z_1, z_2, M$ and N are as given in Sections 2.2 and 2.3, and

$$k_1^{(1)}(z) = e^{-\frac{\lambda_0 z}{\xi}}(z_1 - z)^{\frac{\gamma M}{\xi}}(z_2 - z)^{-\frac{\gamma N}{\xi}}, \quad k_2^{(1)}(z) = e^{-\frac{\lambda_0 z}{\xi}}(z - z_1)^{\frac{\gamma M}{\xi}}(z_2 - z)^{-\frac{\gamma N}{\xi}}.$$

Solving a differential equation similarly as in Section 2.3 gives

$$G_0^{(1)}(z) = \begin{cases} \dfrac{\eta\mu P_{10}^{(1)} \int_0^z \frac{k_1^{(1)}(x)}{\beta(x)} dx + \xi P_{00}^{(1)} z_1^{\frac{\gamma M}{\xi}} z_2^{-\frac{\gamma N}{\xi}}}{\xi k_1^{(1)}(z)}, & z \leqslant z_1, \\[2ex] \dfrac{\eta\mu P_{10}^{(1)} \int_{z_1}^z \frac{k_2^{(1)}(x)}{\beta(x)} dx}{\xi k_2^{(1)}(z)}, & z \geqslant z_1. \end{cases}$$

To obtain the probabilities $P_{00}^{(1)}$ and $P_{10}^{(1)}$ we repeat the process from Section 2.4 and derive two equations connecting $P_{00}^{(1)}$ and $P_{10}^{(1)}$,

$$P_{10}^{(1)} = \xi k_2^{(1)}(1) \cdot \left[\mu(\gamma + \eta) \int_{z_1}^1 \frac{k_2^{(1)}(x)}{\beta(x)} dx\right]^{-1},$$
$$P_{00}^{(1)} = -\eta k_2^{(1)}(1) \int_0^{z_1} \frac{k_1^{(1)}(x)}{\beta(x)} dx \cdot \left[(\gamma + \eta) z_1^{\frac{\gamma M}{\xi}} z_2^{-\frac{\gamma N}{\xi}} \int_{z_1}^1 \frac{k_2^{(1)}(x)}{\beta(x)} dx\right]^{-1}.$$

Now, knowing $P_{00}^{(1)}$ and $P_{10}^{(1)}$, any probability $P_{0n}^{(1)}$ and $P_{1n}^{(1)}$, for $n \geqslant 1$, can be calculated progressively by using the balance equations, or by differentiating $G_0^{(1)}(z)$ and $G_1^{(1)}(z)$, respectively.

For the set of parameters $\lambda = 4, \mu = 7, \lambda_0 = 2, \gamma = 2, \eta = 2$ and $\xi = 1$ we get

$$\int_0^{z_1} \frac{k_1^{(1)}(x)}{\beta(x)} dx = -0.072553, \quad \int_{z_1}^1 \frac{k_2^{(1)}(x)}{\beta(x)} dx = 0.010429, \quad P_{00}^{(1)} = 0.12983, \quad P_{10}^{(1)} = 0.18762.$$

2. $\gamma \to 0, \eta > 0$.

With $\gamma \to 0$ we have $P_{1\bullet}^{(2)} = 0$, and therefore $P_{1n}^{(2)} = 0$ for all n.

The model now transforms into a state independent $M/M/1$-type queue augmented with state-dependent abandonment rates.

In this case, since $\gamma \to 0$, $k^{(2)}(z) = e^{-\frac{\lambda_0 z}{\xi}} z^{\frac{\mu_0}{\xi}}$ and therefore, utilizing (2.11), we can write $G_0^{(2)}(z)$ as

$$G_0^{(2)}(z) = \frac{\mu_0}{\xi} P_{00}^{(2)} e^{\frac{\lambda_0 z}{\xi}} z^{-\frac{\mu_0}{\xi}} \int_0^z e^{-\frac{\lambda_0 x}{\xi}} x^{\frac{\mu_0}{\xi} - 1} dx.$$

We write

$$\int_0^z e^{-\frac{\lambda_0 x}{\xi}} x^{\frac{\mu_0}{\xi} - 1} dx = \left(\frac{\lambda_0}{\xi}\right)^{-\frac{\mu_0}{\xi}} \left(Gamma\left(\frac{\mu_0}{\xi}\right) - Gamma\left(\frac{\mu_0}{\xi}, \frac{\lambda_0 z}{\xi}\right)\right),$$

where $Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$, and $Gamma(\alpha, z) = \int_z^\infty t^{\alpha-1} e^{-t} dt$ is the incomplete $Gamma$ function (see [9, p. 47]).

The above gives

$$\int_0^z e^{-\frac{\lambda_0 x}{\xi}} x^{\frac{\mu_0}{\xi}-1}\,dx = \left(\frac{\lambda_0}{\xi}\right)^{-\frac{\mu_0}{\xi}} \int_0^{\frac{\lambda_0 z}{\xi}} e^{-t} t^{\frac{\mu_0}{\xi}-1}\,dt,$$

and therefore

$$G_0^{(2)}(z) = e^{\frac{\lambda_0 z}{\xi}} \mu_0 (\lambda_0 z)^{-\frac{\mu_0}{\xi}} \xi^{\frac{\mu_0}{\xi}-1} P_{00}^{(2)} \int_0^{\frac{\lambda_0 z}{\xi}} e^{-t} t^{\frac{\mu_0}{\xi}-1}\,dt. \tag{2.23}$$

By setting $z = 1$ in (2.23) (using $G_0^{(2)}(1) = 1$) we get $P_{00}^{(2)}$. That is,

$$P_{00}^{(2)} = e^{-\frac{\lambda_0}{\xi}} \lambda_0^{\frac{\mu_0}{\xi}} \xi^{1-\frac{\mu_0}{\xi}} \left[\mu_0 \int_0^{\frac{\lambda_0}{\xi}} e^{-t} t^{\frac{\mu_0}{\xi}-1}\,dt\right]^{-1},$$

and therefore

$$G_0^{(2)}(z) = e^{-\frac{\lambda_0}{\xi}(1-z)} z^{-\frac{\mu_0}{\xi}} \left[\int_0^{\frac{\lambda_0}{\xi}} e^{-t} t^{\frac{\mu_0}{\xi}-1}\,dt\right]^{-1} \int_0^{\frac{\lambda_0 z}{\xi}} e^{-t} t^{\frac{\mu_0}{\xi}-1}\,dt. \tag{2.24}$$

Once $P_{00}^{(2)}$ is known, we can derive $P_{0n}^{(2)}$ for all n by using the formula

$$P_{0n}^{(2)} = P_{00}^{(2)} (\lambda_0)^n \prod_{j=1}^{n} \frac{1}{\mu_0 + j\xi}, \quad n = 1, 2, \ldots,$$

where

$$\prod_{j=1}^{n} \frac{1}{\mu_0 + j\xi} = \xi \cdot Gamma\left(1 + \frac{\mu_0}{\xi}\right) \left(\xi^{n+1} \cdot Gamma\left(1 + n + \frac{\mu_0}{\xi}\right)\right)^{-1}.$$

We thus have

$$P_{0n}^{(2)} = P_{00}^{(2)} (\lambda_0)^n \xi \cdot Gamma\left(1 + \frac{\mu_0}{\xi}\right) \left(\xi^{n+1} \cdot Gamma\left(1 + n + \frac{\mu_0}{\xi}\right)\right)^{-1}, \quad n = 1, 2, \ldots$$

3. $\gamma > 0, \eta \to 0$.

In this case we have $P_{0\bullet}^{(3)} = 0$, and therefore $P_{0n}^{(3)} = 0$ for all $n \geqslant 0$, resulting in $G_0^{(3)}(z) = 0$. Our system then becomes a regular $M(\lambda)/M(\mu)/1$ system.

4. $\xi \to \infty$.

This system can be described as an $M/M/1$-type queue, where the system suffers disastrous breakdowns, occurring when the server is at its functioning phase. A failure of the system rejects all customers present, and a repair process starts immediately. There are no arrivals during the repair process.

A transition-rate diagram for this case is depicted in Fig. 2.2.

With $\xi \to \infty$, the differential equation (2.7) gives $G_0^{(4)\prime}(z) = 0$. Also, $G_0^{(4)}(z) = P_{00}^{(4)}$ for all $0 \leqslant z \leqslant 1$. Substituting $z = 1$ gives

$$G_0^{(4)}(1) = P_{0\bullet}^{(4)} = P_{00}^{(4)} = \frac{\eta}{\gamma + \eta}.$$

To calculate $P_{10}^{(4)}$ we use Eq. (2.6) and get

$$G_1^{(4)}(z) = \frac{\gamma P_{00}^{(4)} z - \mu(1-z) P_{10}^{(4)}}{\beta(z)}. \tag{2.25}$$

By setting $z = z_1$ in (2.25) we get

$$\gamma P_{00}^{(4)} z_1 = \mu(1 - z_1) P_{10}^{(4)},$$

and therefore

$$P_{10}^{(4)} = \frac{\gamma \eta z_1}{\mu(\gamma + \eta)(1 - z_1)}.$$

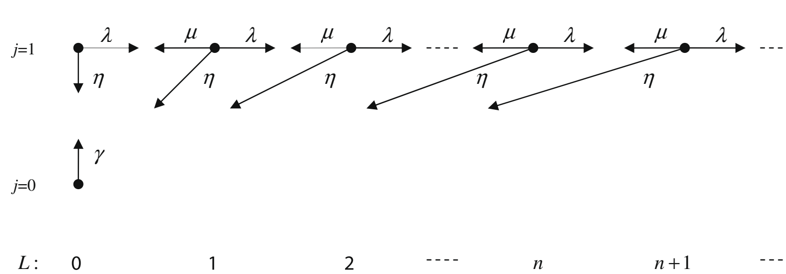

Fig. 2.2. Transition-rate diagram when $\xi \to \infty$.

With known $P_{00}^{(4)}$ and $P_{10}^{(4)}$, $G_1^{(4)}(z)$ is obtained in a closed form as follows:

$$G_1^{(4)}(z) = \frac{\frac{z\eta\gamma}{\gamma+\eta} - \mu(1-z)\left(\frac{\gamma\eta z_1}{\mu(\gamma+\eta)(1-z_1)}\right)}{\beta(z)} = \frac{\gamma\eta}{\lambda(\gamma+\eta)(z_2-z)(1-z_1)},$$

since $\beta(z) = -\lambda(z-z_1)(z-z_2)$.

$G_1^{(4)}(z)$ can also be represented as a power series in the following way:

$$G_1^{(4)}(z) = \frac{\gamma\eta}{\lambda(\gamma+\eta)(1-z_1)} \cdot \frac{1}{z_2\left(1-\frac{z}{z_2}\right)} = \sum_{n=0}^{\infty} \frac{\gamma\eta}{\lambda(\gamma+\eta)(1-z_1)z_2^{n+1}} \cdot z^n.$$

Since $G_1^{(4)}(z) = \sum_{n=0}^{\infty} P_{1n}^{(4)} z^n$, it follows that $P_{1n}^{(4)} = \frac{\gamma\eta}{\lambda(\gamma+\eta)(1-z_1)z_2^{n+1}}$ for all $n \geqslant 0$.

As an illustration we take the set $\lambda = 2, \mu = 2, \eta = 1$ and $\gamma = 3$. Then, $z_1 = 0.5$, $z_2 = 2, P_{00}^{(4)} = 0.25$ and $P_{10}^{(4)} = 0.375$.

Comparing to the general model, we use Eqs. (2.15) and (2.16) with $\lambda_0 = 3$, $\mu_0 = 0.5$ and $\xi = 10,000$, and get

$$T = 20000.69, \quad S = -3332.6613, \quad U = 3332.892, \quad V = 0.6928888, \quad k_2(1) = 0.99963, \quad P_{00} = 0.249873, \quad P_{10} = 0.374898.$$

5. $\xi \to 0$.

Assume $\xi \to 0$. Then the system reduces to the model described in Yechiali and Naor [22]. The condition for stability is $\hat{\mu} - \hat{\lambda} > 0$ ($\hat{\mu}$ and $\hat{\lambda}$ are as defined before).

In this case no differential equations are involved and one has to solve two algebraic equations, (2.4) and (2.5), connecting $G_0^{(5)}(z)$ and $G_1^{(5)}(z)$. The solution gives

$$G_0^{(5)}(z) = \frac{(1-z)[\eta\mu P_{10}^{(5)} z + \mu_0\beta(z)P_{00}^{(5)}]}{\gamma\eta z^2 - \alpha(z)\beta(z)}, \quad G_1^{(5)}(z) = \frac{(1-z)[\gamma\mu_0 P_{00}^{(5)} z + \mu\alpha(z)P_{10}^{(5)}]}{\gamma\eta z^2 - \alpha(z)\beta(z)}.$$

Defining

$$h(z) = -\lambda\lambda_0 z^3 + (\lambda\lambda_0 + \mu\lambda_0 + \lambda\mu_0 + \eta\lambda_0 + \gamma\lambda)z^2 - (\mu\lambda_0 + \lambda\mu_0 + \mu\mu_0 + \eta\mu_0 + \gamma\mu)z + \mu\mu_0,$$

we get

$$G_0^{(5)}(z) = \frac{\eta\mu P_{10}^{(5)} z + \mu_0\beta(z)P_{00}^{(5)}}{h(z)}, \quad G_1^{(5)}(z) = \frac{\gamma\mu_0 P_{00}^{(5)} z + \mu\alpha(z)P_{10}^{(5)}}{h(z)}.$$

By using the single root of $h(z)$ in (0,1), denoted as z_0, $P_{00}^{(5)}$ and $P_{10}^{(5)}$ are obtained:

$$P_{00}^{(5)} = \frac{\eta(\hat{\mu}-\hat{\lambda})z_0}{\mu_0(1-z_0)(\mu-\lambda z_0)}, \quad P_{10}^{(5)} = \frac{\gamma(\hat{\mu}-\hat{\lambda})z_0}{\mu(1-z_0)(\mu_0-\lambda_0 z_0)}.$$

6. Both γ and η tend to 0.

Assume $\gamma \to 0, \eta \to 0$, while $\frac{\eta}{\gamma} \to r$ for some constant $r > 0$. Then,

$$P_{0\bullet}^{(6)} = \frac{\eta}{\gamma+\eta} \to \frac{r}{1+r}, \quad P_{1\bullet}^{(6)} = \frac{\gamma}{\gamma+\eta} \to \frac{1}{1+r}.$$

The overall generating function, $G^{(6)}(z) = G_0^{(6)}(z) + G_1^{(6)}(z)$, can be expressed as a probabilistic mixture between cases 2 and 3 as follows:

$$G^{(6)}(z) = P_{0\bullet}^{(6)} \cdot G_0^{(2)}(z) + P_{1\bullet}^{(6)} \cdot G_1^{(3)}(z)$$

implying that

$$G^{(6)}(z) = \frac{r}{1+r} \cdot e^{-\frac{\lambda_0}{\xi}(1-z)} z^{-\frac{\mu_0}{\xi}} \left[\int_0^{\frac{\lambda_0}{\xi}} e^{-t} t^{\frac{\mu_0}{\xi}-1}\,dt\right]^{-1} \int_0^{\frac{\lambda_0 z}{\xi}} e^{-t} t^{\frac{\mu_0}{\xi}-1}\,dt + \frac{1}{1+r} \cdot \frac{\mu-\lambda}{\mu-\lambda z}.$$

Also, in this case we have $P_{00}^{(6)} = \frac{r}{1+r} \cdot P_{00}^{(2)}$, and $P_{10}^{(6)} = \frac{1}{1+r} \cdot P_{10}^{(3)}$.

Define $G^{(6)}(0) = P_{\bullet 0}^{(6)}$ as the proportion of time the system is empty. Then, $P_{\bullet 0}^{(6)} = P_{00}^{(6)} + P_{10}^{(6)} = \frac{r}{1+r} P_{00}^{(2)} + \frac{1}{1+r} P_{10}^{(3)}$. That is,

$$P_{\bullet 0}^{(6)} = \frac{r}{1+r} e^{-\frac{\lambda_0}{\xi}} \lambda_0^{\frac{\mu_0}{\xi}} \xi^{1-\frac{\mu_0}{\xi}} \left[\mu_0 \int_0^{\frac{\lambda_0}{\xi}} e^{-t} t^{\frac{\mu_0}{\xi}-1}\,dt\right]^{-1} + \frac{1}{1+r}\left(1-\frac{\lambda}{\mu}\right).$$

As an illustration, take $\lambda = 4, \mu = 8, \lambda_0 = 3, \mu_0 = 3$ and $\xi = 2$. $P_{00}^{(2)} = 0.507$, and $P_{10}^{(3)} = 0.5$. Now, for $r = 2$ $P_{\bullet 0}^{(6)} = 2/3 \cdot 0.506859 + 1/3 \cdot 0.5 = 0.50457$.

7. Both γ and η tend to ∞.

Assume $\gamma \to \infty, \eta \to \infty$, such that $\eta/\gamma \to r$ for some constant $r > 0$. That is, the system oscillates rapidly between phases 0 and 1. It follows that

$$P_{0n}^{(7)} = rP_{1n}^{(7)} \quad \forall n \geqslant 0.$$

That is, the fraction of time where there are n customers in the system is divided between the phases by the ratio r, implying that

$$G_0^{(7)}(z) = rG_1^{(7)}(z) \tag{2.26}$$

and

$$P_{0\bullet}^{(7)} = rP_{1\bullet}^{(7)}.$$

Alternatively,

$$P_{0\bullet}^{(7)} = \frac{\eta}{\gamma+\eta} \to \frac{r}{1+r}, \quad P_{1\bullet}^{(7)} = \frac{\gamma}{\gamma+\eta} \to \frac{1}{1+r}.$$

Now, in order to calculate $G_0^{(7)}(z)$ and $G_1^{(7)}(z)$ we utilize Eq. (2.7), which transforms into

$$G_0^{(7)'}(z) - \left(\frac{\lambda_0+\lambda/r}{\xi} - \frac{\mu_0+\mu/r}{z\xi}\right) G_0^{(7)}(z) = \frac{\mu P_{10}^{(7)} + \mu_0 P_{00}^{(7)}}{z\xi}. \tag{2.27}$$

The solution of Eq. (2.27) is

$$G_0^{(7)}(z) = \frac{\left(\mu P_{10}^{(7)} + \mu_0 P_{00}^{(7)}\right)}{\xi} e^{\frac{(\lambda_0+\lambda/r)z}{\xi}} z^{-\frac{\mu_0+\mu/r}{\xi}} \int_0^z e^{-\frac{(\lambda_0+\lambda/r)u}{\xi}} u^{\frac{\mu_0+\mu/r}{\xi}-1} du. \tag{2.28}$$

Note that, as opposed to Eq. (2.23) (which describes $G_0^{(2)}(z)$ when $\gamma \to 0$ and $\eta > 0$), in (2.28) the parameters λ and μ play a role.

Finally, for deriving $P_{00}^{(7)}$ and $P_{10}^{(7)}$, we use the following two equations:

The first equation results from substituting $z = 1$ in (2.28) and using $G_0^{(7)}(1) = r/1 + r$.

The second equation relating $P_{00}^{(7)}$ and $P_{10}^{(7)}$ is $P_{00}^{(7)} = rP_{10}^{(7)}$.

3. The *c*-server case

3.1. The model

Consider now the multi-server case with $1 \leqslant c < \infty$ servers. As before, the system alternates between phases 0 and 1, as described in Section 2.1. When the system operates in phase 0, each and every server becomes slower with the same rate μ_0. A transition-rate diagram is depicted in Fig. 3.1.

3.2. Balance equations and generating functions

The set of balance equations is given by

$$\begin{aligned} &\underline{j=0} \\ &\quad n=0 : (\lambda_0+\gamma)P_{00} = \eta P_{10} + (\mu_0+\xi)P_{01}, \\ &\quad 1 \leqslant n \leqslant c-1 : (\lambda_0+\gamma+n(\mu_0+\xi))P_{0n} = \lambda_0 P_{0,n-1} + \eta P_{1n} + (n+1)(\mu_0+\xi)P_{0,n+1}, \\ &\quad n \geqslant c : (\lambda_0+\gamma+c\mu_0+n\xi)P_{0n} = \lambda_0 P_{0,n-1} + \eta P_{1n} + (c\mu_0+(n+1)\xi)P_{0,n+1}, \end{aligned} \tag{3.1}$$

$$\begin{aligned} &\underline{j=1} \\ &\quad n=0 : (\lambda+\eta)P_{10} = \mu P_{11} + \gamma P_{00}, \\ &\quad 1 \leqslant n \leqslant c-1 : (\lambda+n\mu+\eta)P_{1n} = \lambda P_{1,n-1} + \gamma P_{0n} + (n+1)\mu P_{1,n+1}, \\ &\quad n \geqslant c : (\lambda+c\mu+\eta)P_{1n} = \lambda P_{1,n-1} + \gamma P_{0n} + c\mu P_{1,n+1}. \end{aligned} \tag{3.2}$$

Clearly, the proportions of time the system stays in each of the two levels remain unaffected and are given by Eq. (2.3).

Using the same procedure as in Section 2.1, we get an algebraic equation for $G_1(z)$ and a differential equation for $G_0(z)$, given by

$$G_1(z) = \frac{\gamma z G_0(z) - \mu(1-z)A_1(z)}{\beta_c(z)},$$

and

$$G_0'(z) - \frac{\alpha_c(z)\beta_c(z) - \eta\gamma z^2}{\xi z(1-z)\beta_c(z)} G_0(z) = \frac{\eta\mu z A_1(z) + \mu_0 \beta_c(z) A_0(z)}{\xi z \beta_c(z)}, \tag{3.3}$$

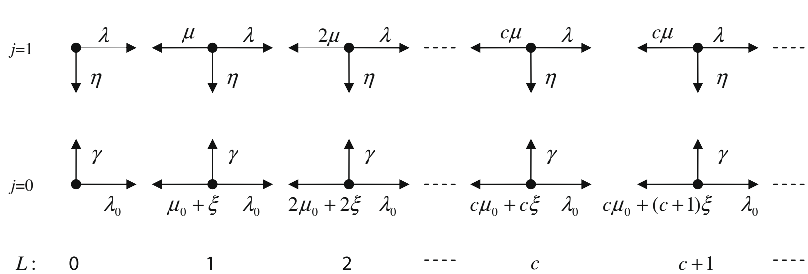

Fig. 3.1. Transition-rate diagram with c servers.

where $A_j(z) = \sum_{n=0}^{c-1}(c-n)P_{jn}z^n$ for $j = 0, 1$; $\alpha_c(z) = (\lambda_0 z - c\mu_0)(1-z) + \gamma z$ and $\beta_c(z) = (\lambda z - c\mu)(1-z) + \eta z$. The roots $z_{1,c}$ and $z_{2,c}$ of $\beta_c(z)$ are $\frac{\lambda+c\mu+\eta\mp\sqrt{(\lambda+c\mu+\eta)^2-4\lambda c\mu}}{2\lambda}$, where $z_{2,c} > 1 > z_{1,c} > 0$.

3.3. Solution of the differential equation (3.3)

The similarity of Eq. (3.1) to Eq. (2.7) implies that the former's solution is given by Eq. (2.11) with the modifications: $A_0(z)$ replacesP_{00}; $A_1(z)$ replacesP_{10}; $z_{j,c}$ replaces z_j for $j = 1, 2$, and $\beta_c(z)$ replaces $\beta(z)$. We thus get

$$G_0(z) = \begin{cases} \dfrac{\frac{\eta\mu}{\xi}\sum_{n=0}^{c-1}(c-n)P_{1n}\int_0^z \frac{k_{1,c}(x)x^n}{\beta_c(x)}dx + \frac{\mu_0}{\xi}\sum_{n=0}^{c-1}(c-n)P_{0n}\int_0^z \frac{k_{1,c}(x)x^n}{x}dx}{k_{1,c}(z)}, & z \leqslant z_{1,c}, \\ \dfrac{\frac{\eta\mu}{\xi}\sum_{n=0}^{c-1}(c-n)P_{1n}\int_{z_{1,c}}^z \frac{k_{2,c}(x)x^n}{\beta_c(x)}dx + \frac{\mu_0}{\xi}\sum_{n=0}^{c-1}(c-n)P_{0n}\int_{z_{1,c}}^z \frac{k_{2,c}(x)x^n}{x}dx}{k_{2,c}(z)}, & z \geqslant z_{1,c}, \end{cases} \tag{3.4}$$

where

$$k_{1,c}(z) = e^{-\frac{\lambda_0 z}{\xi}} z^{\frac{c\mu_0}{\xi}} (z_{1,c} - z)^{\frac{\gamma M_c}{\xi}} (z_{2,c} - z)^{-\frac{\gamma N_c}{\xi}},$$
$$k_{2,c}(z) = e^{-\frac{\lambda_0 z}{\xi}} z^{\frac{c\mu_0}{\xi}} (z - z_{1,c})^{\frac{\gamma M_c}{\xi}} (z_{2,c} - z)^{-\frac{\gamma N_c}{\xi}},$$
$$M_c = \frac{z_{1,c}(z_{2,c} - 1)}{z_{2,c} - z_{1,c}}, \quad N_c = \frac{z_{2,c}(z_{1,c} - 1)}{z_{2,c} - z_{1,c}}.$$

Eq. (3.4) expresses $G_0(z)$ in terms of $2c$ boundary probabilities $P_{00}, P_{01}, \ldots P_{0,c-1}, P_{10}, P_{11}, \ldots P_{1,c-1}$. Those $2c$ probabilities are required in order to completely determine $G_0(z)$.We derive these probabilities in the next section.

3.4. Derivation of the probabilities $P_{00}, P_{01}, \ldots, P_{0,c-1}, P_{10}, P_{11}, \ldots, P_{1,c-1}$

At $z = 1$, $G_0(1) = \frac{\eta}{\eta+\gamma}$, thus from (the lower part of) Eq. (3.4) we get

$$\frac{\eta\xi k_{2,c}(1)}{\eta+\gamma} = \eta\mu\sum_{n=0}^{c-1}(c-n)P_{1n}\int_{z_{1,c}}^1 \frac{k_{2,c}(x)x^n}{\beta_c(x)}dx + \mu_0\sum_{n=0}^{c-1}(c-n)P_{0n}\int_{z_{1,c}}^1 \frac{k_{2,c}(x)x^n}{x}dx.$$

For $z = z_{1,c}$, $k_{1,c}(z_{1,c}) = 0$, implying that the numerator in the upper part of (3.4) vanishes. That is,

$$\frac{\eta\mu}{\xi}\sum_{n=0}^{c-1}(c-n)P_{1n}\int_0^{z_{1,c}} \frac{k_{1,c}(x)x^n}{\beta_c(x)}dx + \frac{\mu_0}{\xi}\sum_{n=0}^{c-1}(c-n)P_{0n}\int_0^{z_{1,c}} \frac{k_{1,c}(x)x^n}{x}dx = 0.$$

The above gives two equations in the boundary probabilities. The remaining $2c-2$ equations connecting $P_{00}, P_{01}, \ldots P_{0,c-1}, P_{10}, P_{11}, \ldots P_{1,c-1}$ are taken from the balance equations (3.1) and (3.2) for $j = 0, 1$ and $n = 0, 1, \ldots, c-2$ (each phase contributes $c-1$ equations).

In order to derive the mean queue sizes, $E[L_0]$ and $E[L_1]$, one can use a direct approach. By taking vertical cuts in Fig. 3.1 we get

$$\lambda P_{1n} + \lambda_0 P_{0n} = (n+1)\mu P_{1,n+1} + (n+1)(\mu_0 + \xi)P_{0,n+1}, \quad 0 \leqslant n \leqslant c-1,$$
$$\lambda P_{1n} + \lambda_0 P_{0n} = c\mu P_{1,n+1} + (c\mu_0 + (n+1)\xi)P_{0,n+1}, \quad c \leqslant n.$$

Summing over n gives

$$\lambda P_{1\bullet} + \lambda_0 P_{0\bullet} = \mu(cP_{1\bullet} - \sum_{n=0}^{c-1}(c-n)P_{1n}) + \mu_0(cP_{0\bullet} - \sum_{n=0}^{c-1}(c-n)P_{0n}) + \xi E[L_0] = \mu(cP_{1\bullet} - A_1(1)) + \mu_0(cP_{0\bullet} - A_0(1)) + \xi E[L_0].$$

That is,

$$E[L_0] = \frac{\hat{\lambda} - c\hat{\mu} + \mu A_1(1) + \mu_0 A_0(1)}{\xi}.$$

The derivation of $E[L_1]$ is done similarly as in Eq. (2.22)

$$E[L_1] = G_1'(z)|_{z=1} = \frac{\gamma(\hat{\lambda} - c\hat{\mu}) + \xi(\lambda - c\mu)P_{1\bullet} + \gamma\mu_0 A_0(1) + \mu A_1(1)(\gamma + \xi)}{\eta\xi}.$$

Finally, the mean total number of customers in the system is $E[L] = E[L_0] + E[L_1]$.

Numerical Example. We consider the case $c = 2$. The probabilities $P_{00}, P_{01}, P_{10}, P_{11}$ are needed in order to completely determine $G_0(z)$. From Section 3.4, the set of four equations connecting these four probabilities is

$$\frac{\xi\eta k_{2,2}(1)}{\eta+\gamma} = \eta\mu 2P_{10}\int_{z_{1,2}}^1 \frac{k_{2,2}(x)}{\beta_2(x)}dx + \eta\mu P_{11}\int_{z_{1,2}}^1 \frac{k_{2,2}(x)x}{\beta_2(x)}dx + \mu_0 2P_{00}\int_{z_{1,2}}^1 \frac{k_{2,2}(x)}{x}dx + \mu_0 P_{01}\int_{z_{1,2}}^1 \frac{k_{2,2}(x)x}{x}dx,$$
$$\frac{\eta\mu}{\xi}2P_{10}\int_0^{z_{1,2}} \frac{k_{1,2}(x)}{\beta_2(x)}dx + \frac{\eta\mu}{\xi}P_{11}\int_0^{z_{1,2}} \frac{k_{1,2}(x)x}{\beta_2(x)}dx = -\frac{\mu_0}{\xi}2P_{00}\int_0^{z_{1,2}} \frac{k_{1,2}(x)}{x}dx - \frac{\mu_0}{\xi}P_{01}\int_0^{z_{1,2}} \frac{k_{1,2}(x)x}{x}dx,$$
$$(\lambda_0 + \gamma)P_{00} = \eta P_{10} + (\mu_0 + \xi)P_{01}, \quad (\lambda + \eta)P_{10} = \mu P_{11} + \gamma P_{00}.$$

In Table 3.2 we compare numerically the values of the boundary probabilities when $c = 1$ and when $c = 2$, where $\lambda = 3, \mu = 3, \lambda_0 = 2, \mu_0 = 1, \gamma = 2, \eta = 1$ and $\xi = 1$.

Table 3.2
Boundary probabilities and mean queue sizes.

	P_{00}	P_{01}	P_{10}	P_{11}	$E[L_0]$	$E[L_1]$	$E[L]$
$c=1$	0.0689	0.0749	0.1258	0.1218	0.7796	1.9366	2.7162
$c=2$	0.1147	0.1157	0.2276	0.2270	0.3917	0.8300	1.2217

Remark. Extreme cases, similar to those studied in Section 2.5, can be analyzed for the many-server case. We skip those derivations.

4. Infinite number of servers

4.1. The model

We now consider a service system with an infinite number of servers. That is, we deal with customers' abandonments occurring in an $M/M/\infty$ queueing system operating in a 2-phase random environment. As in Sections 2 and 3, when the system switches to phase $j=0$, the service rate of all servers decreases to μ_0. The Transition-rate diagram is given in Fig. 4.1. It follows that, from analytic point of view, one can consider a combined slow service rate and abandonment rate, $\mu_0+\xi$, as a global departure rate for each individual customer when the system is in phase $j=0$.

A related model has been fully investigated by Baykal-Gursoy and Xiao [4], and further studied by Paz and Yechiali [17]. In [4], a service system with an infinite number of servers subject to random interruptions of exponentially distributed durations is considered. During interruptions, all servers work at lower efficiency, compared to their normal functioning rate. However, the Poisson arrival rate in both environments *remains the same*, where in our work the Poisson arrival rate *changes* with the change of environment. Nevertheless, a solution for our model can be obtained from [4] with only a few modifications. For the sake of completeness in the next section we present the balance equations and exhibit a closed-form formula for the partial generating functions, $G_0(z)$ and $G_1(z)$. We also calculate, by a slightly different approach than the one described in [4], the values of $E[L_0]$ and $E[L_1]$, which gives us a formula for $E[L]$, the mean total number of customers in the system.

4.2. Balance equations and generating functions

As in Sections 2 and 3, P_{jn} denotes the steady-state probability of the system being in state (j,n) if the system is in environment j, $j=0,1$, and n customers are present.

The steady-state balance equations are

$$\underline{j=0}\quad \begin{cases} n=0: (\lambda_0+\gamma)P_{00}=\eta P_{10}+(\mu_0+\xi)P_{01}, \\ n\geqslant 0: ((\lambda_0+\gamma+n(\mu_0+\xi))P_{0n}=\lambda_0 P_{0,n-1}+\eta P_{1n}+(n+1)(\mu_0+\xi)P_{0,n+1}, \end{cases} \tag{4.1}$$

$$\underline{j=1}\quad \begin{cases} n=0: (\lambda+\eta)P_{10}=\gamma P_{00}+\mu P_{11}, \\ n\geqslant 0: (\lambda+\eta+n\mu)P_{1n}=\lambda P_{1,n-1}+\gamma P_{0n}+(n+1)\mu P_{1,n+1}. \end{cases} \tag{4.2}$$

Multiplying both sides of (4.1) and (4.2) by z^n and summing over all n yield the differential equations

$$G_0'(z)-\left[\frac{\gamma}{(\mu_0+\xi)(1-z)}+\frac{\lambda_0}{\mu_0+\xi}\right]G_0(z)=-\frac{\eta}{(\mu_0+\xi)(1-z)}G_1(z), \tag{4.3}$$

$$G_1'(z)-\left[\frac{\eta}{\mu(1-z)}+\frac{\lambda}{\mu}\right]G_1(z)=-\frac{\gamma}{\mu(1-z)}G_0(z). \tag{4.4}$$

Before presenting the solution of Eqs. (4.3) and (4.4), we notice that $P_{0\bullet}$ and $P_{1\bullet}$ are the same as in Sections 2 and 3, and are given in (2.3). Furthermore, (4.3) and (4.4) yield, respectively

$$G_0'(z)=\frac{\lambda_0}{\mu_0+\xi}G_0(z)+\frac{\gamma G_0(z)-\eta G_1(z)}{(\mu_0+\xi)(1-z)}, \tag{4.5}$$

and

$$G_1'(z)=\frac{\lambda}{\mu}G_1(z)+\frac{\eta G_1(z)-\gamma G_0(z)}{\mu(1-z)}. \tag{4.6}$$

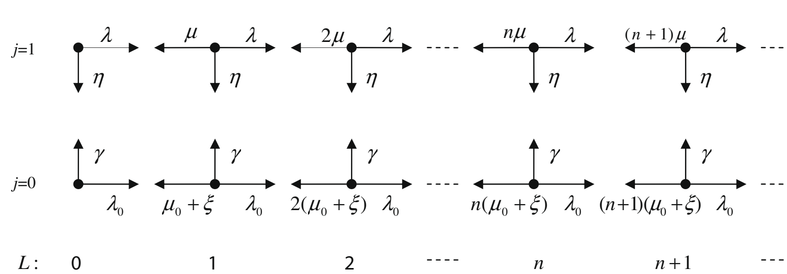

Fig. 4.1. Transition-rate diagram for the $M/M/\infty$ case.

Substituting $z = 1$ in results (4.5) and (4.6), using $G_0(1) = \frac{\eta}{\gamma+\eta}$, $G_1(1) = \frac{\gamma}{\gamma+\eta}$, and applying L'Hopital rule, gives two equations with the two variables, $G_0'(1)$ and $G_1'(1)$, resulting in

$$G_0'(1) = E[L_0] = \frac{\eta(\lambda\gamma + \lambda_0(\eta + \mu))}{(\gamma + \eta)(\gamma\mu + (\eta + \mu)(\mu_0 + \xi))},$$
$$G_1'(1) = E[L_1] = \frac{\gamma(\lambda_0\eta + \lambda(\gamma + \mu_0 + \xi))}{(\gamma + \eta)(\gamma\mu + (\eta + \mu)(\mu_0 + \xi))}.$$

Hence,

$$E[L] = E[L_0] + E[L_1] = \frac{\lambda\gamma^2 + \lambda_0\eta(\eta + \mu) + \gamma(\eta(\lambda + \lambda_0) + \lambda(\mu_0 + \xi))}{(\gamma + \eta)(\gamma\mu + (\eta + \mu)(\mu_0 + \xi))}. \tag{4.7}$$

Setting $\lambda_0 = \lambda$ reduces (4.7) to Eq. (3.6) in [4].

Note that, in this case ($c = \infty$), it is possible to derive the mean total number of customers in the system without actually calculating the PGFs.

To get $G_0(z)$ and $G_1(z)$ we repeat the process described in [4], with a few modifications. We write

$$G_0(z) = \frac{\eta}{\gamma + \eta} \cdot e^{-\frac{\lambda(1-z)}{\mu}} \cdot M(a + 1, b + 1, 2\tilde{\rho}(1 - z)), \tag{4.8}$$
$$G_1(z) = \frac{\eta(\mu_0 + \xi) + \gamma\mu}{(\gamma + \eta)\mu} \cdot e^{-\frac{\lambda(1-z)}{\mu}} \cdot \left[M(a, b, 2\tilde{\rho}(1 - z)) - \frac{a}{b}M(a + 1, b + 1, 2\tilde{\rho}(1 - z))\right], \tag{4.9}$$

where

$$a = \frac{\eta}{\mu}, \quad b = \frac{\eta}{\mu} + \frac{\gamma}{\mu_0 + \xi}, \quad \tilde{\rho} = \frac{1}{2}\left(\frac{\lambda}{\mu} - \frac{\lambda_0}{\mu_0 + \xi}\right),$$

and $M(a, b, z)$ is the *Kummer* function with the following power series representation:

$$M(a, b, z) = \sum_{n=0}^{\infty} \frac{a_{(n)}}{b_{(n)}} \frac{z^n}{n!},$$

where

$$a_{(n)} = a(a + 1)(a + 2) \cdots (a + n - 1) \quad \text{and} \quad a_{(0)} = 1,$$
$$b_{(n)} = b(b + 1)(b + 2) \cdots (b + n - 1) \quad \text{and} \quad b_{(0)} = 1.$$

Now, with known $G_0(z)$ and $G_1(z)$, P_{00} and P_{10} can be calculated directly by substituting $z = 0$ in the generating functions. The rest of the probabilities may be calculated progressively from the balance equations or by repeated differentiation of the PGFs.

Numerical Example. With the same values used in Section 3.4, we get

$$P_{00} = G_0(0) = 0.12262, \quad P_{10} = G_1(0) = 0.24524, \quad E[L_0] = \frac{1}{3}, \quad E[L_1] = \frac{2}{3}, \quad E[L] = 1.$$

As expected, P_{00} and P_{10} in the $c = \infty$ case are greater then in the $c = 2$ case, and $E[L]$ for $c = \infty$ is smaller than when $c = 2$.

We note that those neat numerical results for $E[L_0]$ and $E[L_1]$ are a consequence of the following:

Theorem. (Proved in [17]) *In an $M/M/\infty$ queue in a m-phase random environment, with arrival and service rates λ_j and μ_j in phase j, $1 \leqslant j \leqslant m$, the steady-state probabilities satisfy $P_{jn} = P_{j\bullet} \cdot P_{\bullet n}$ (where $P_{j\bullet} = \sum_{n=0}^{\infty} P_{jn}$ and $P_{\bullet n} = \sum_{j=1}^{m} P_{jn}$), if and only if λ_j/μ_j=constant for all j. Furthermore, if λ_j/μ_j=constant, $E[L] = \lambda_j/\mu_j$.*

Indeed, for this example, $\frac{\lambda}{\mu} = \frac{\lambda_0}{\mu_0+\xi} = 1 = E[L]$.

4.3. Extreme cases

Clearly, when considering the cases when $\mu_0 \to 0$, or when $\xi \to 0$, we get the same representations for $G_0(z)$ and $G_1(z)$ as in (4.8) and (4.9), respectively, with only a simple modification. It is also easy to verify, again by utilizing (4.8) and (4.9), that for the cases $\gamma \to 0, \eta > 0$ or $\eta \to 0, \gamma > 0$, we get the generating function of an $M(\lambda_0)/M(\mu_0 + \xi)/\infty$ system or $M(\lambda)/M(\mu)/\infty$ system, respectively.

A more interesting case is when $\xi \to \infty$. A transition rate diagram for this would look similar to Fig. 2.2 with service rates $n\mu$ for all $n \geqslant 0$.

Clearly, $G_0(z) = P_{00} = P_{0\bullet} = \frac{\eta}{\gamma+\eta}$. Substituting this result in Eq. (4.6) leads to a differential equation for $G_1(z)$ as follows:

$$G_1'(z) - \left[\frac{\lambda}{\mu} + \frac{\eta}{\mu(1 - z)}\right]G_1(z) = -\frac{\gamma\eta}{\mu(\gamma + \eta)(1 - z)}. \tag{4.10}$$

The solution of (4.10) is

$$G_1(z) = \frac{\gamma\eta}{\mu(\gamma + \eta)} e^{\frac{\lambda z}{\mu}} (1 - z)^{-\frac{\eta}{\mu}} \int_z^1 e^{-\frac{\lambda u}{\mu}} (1 - u)^{\frac{\eta}{\mu}-1} du.$$

Another interesting extreme case arises when both γ and η tend to ∞, with a constant ratio $\eta/\gamma = r > 0$. Similarly to case 7 in Section 2.5, we get that $G_0(z) = rG_1(z)$. To derive $G_0(z)$ we use Eq. (4.8) and let γ and η tend to ∞, with ratio $r > 0$. We thus get,

$$\frac{a}{b} \rightarrow \frac{(\mu_0+\xi)r}{\mu+(\mu_0+\xi)r},$$

implying that, for all n,

$$\frac{a_{(n)}}{b_{(n)}} \underset{\eta/\gamma\rightarrow r}{\rightarrow} \left(\frac{(\mu_0+\xi)r}{\mu+(\mu_0+\xi)r}\right)^n,$$

and therefore

$$\lim_{\substack{\gamma,\eta\rightarrow\infty \\ \eta/\gamma=r}} M(a+1,b+1,2\tilde{\rho}(1-z)) = \sum_{n=0}^{\infty}\left(\frac{(\mu_0+\xi)r2\tilde{\rho}(1-z)}{\mu+(\mu_0+\xi)r}\right)^n\frac{1}{n!} = \exp\left[\frac{(1-z)r(\lambda(\mu_0+\xi)-\lambda_0\mu)}{\mu^2+\mu(\mu_0+\xi)r}\right].$$

Finally,

$$G_0(z) = \frac{r}{r+1}\cdot\exp\left[-(1-z)\left(\frac{\lambda+\lambda_0 r}{\mu+(\mu_0+\xi)r}\right)\right],$$

and

$$G_1(z) = 1/rG_0(z).$$

We also have

$$E[L] = G'(z)|_{z=1} = (G'_0(z)+G'_1(z))|_{z=1} = \frac{\hat{\lambda}}{\hat{\mu}} \rightarrow \frac{\lambda+\lambda_0 r}{\mu+(\mu_0+\xi)r},$$

which can be obtained directly from (4.7), when $\gamma,\eta\rightarrow\infty$ while $\eta/\gamma\rightarrow r$.

5. Conclusions

We have introduced and analyzed customers' impatience that arises as a result of a slowdown in the servers' service rate. We studied three Markovian models: the single server case, the multiple server case and the infinite-server case. For each model we derived explicit expressions for the PGF of the number of customers in the system, both when the servers are slow and when the system functions normally. We also calculated the mean total number of customers in the system. In the $M/M/1$ and $M/M/c$ ($c<\infty$) queues we solved a differential equation in order to derive the PGFs. When analyzing the $M/M/\infty$ queue, we made use of a related model studied in [4]. Several extreme cases were examined and numerical results were presented.

References

[1] E. Altman, U. Yechiali, Analysis of customers' impatience in queues with server vacations, Queueing Systems 52 (2006) 261–279.
[2] E. Altman, U. Yechiali, Infinite-server queues with system's additional tasks and impatient customers, Probability in the Engineering and Informational Sciences 22 (2008) 477–493.
[3] F. Baccelli, P. Boyer, G. Hebuterne, Single-server queues with impatient customers, Advances in Applied Probability 16 (1984) 887–905.
[4] M. Baykal-Gursoy, W. Xiao, Stochastic decomposition in $M/M/\infty$ queues with Markov modulated service rates, Queueing Systems 48 (2004) 75–88.
[5] W. Boyce, R. DiPrima, Elementary Differential Equations and Boundary Value Problems, seventh ed., John Wiley and Sons, 2001.
[6] O.J. Boxma, P.R. de Waal, Multiserver queues with impatient customers, ITC 14 (1994) 743–756.
[7] L. Bright, P.G. Taylor, Calculating the equilibrium distribution in level dependent quasi-birth-and-death processes, Stochastic Models 11 (3) (1995) 497–525.
[8] D.J. Daley, General customer impatience in the queue $GI/G/1$, Journal of Applied Probability 2 (1965) 186–205.
[9] W. Feller, An Introduction to Probability Theory and its Applications, second ed., Wiley, New York, 1957.
[10] N. Gans, G. Koole, A. Mandelbaum, Telephone call centers: Tutorial, review and research prospects, Manufacturing and Service Operations Management 5 (2003) 79–141.
[11] V. Gupta, M. Harchol-Balter, A. Scheller-Wolf, U. Yechiali, Fundamental characteristics of queues with fluctuating load, in: Proceedings of the ACM SIGMETRICS, vol. 34, 2006, pp. 203–215.
[12] S.P. Martin, I. Mitrani, Job transfers between unreliable servers, Presented in 2nd Madrid Conference on Queueing Theory, July 3–7, 2006.
[13] M.F. Neuts, Matrix-Geometric Solutions in Stochastic Models – An Algorithmic Approach, Johns Hopkins, Baltimore, 1981.
[14] C.A. O'Cinneide, P. Purdue, The $M/M/\infty$ queue in a random environment, Journal of Applied Probability 23 (1986) 175–184.
[15] C. Palm, Methods of judging the annoyance caused by congestion, Tele 4 (1953) 189–208.
[16] C. Palm, Research on Telephone Traffic Carried by Full Availability Groups, Tele 1 (1957) 107. English translation of results first published in 1946 in Swedish in the same journal, which was then entitled Tekniska Meddelanden fran Kungle. Telegrfstyrelsen.
[17] N. Paz, U. Yechiali, A note on the $M/M/\infty$ queue in random environment, Technical report, Department of Statistics and Operations Research, Tel-Aviv University, October 2007.
[18] L. Takacs, A single server queue with limited virtual waiting time, Journal of Applied Probability 11 (1974) 612–617.
[19] B. Van Houdt, R.B. Lenin, C. Blondia, 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.
[20] U. Yechiali, A queueing-type birth-and-death process defined on a continuous-time Markov chain, Operations Research 21 (1973) 604–609.
[21] U. Yechiali, Queues with system disasters and impatient customers when system is down, Queueing Systems 56 (2007) 195–202.
[22] U. Yechiali, P. Naor, Queueing problems with heterogeneous arrivals and service, Operations Research 19 (1971) 722–734.