5 $G/M/1$ queue

In this chapter we study the $G/M/1$ queue, which forms the dual of the $M/G/1$ queue. In this system customers arrive one by one with interarrival times identically and independently distributed according to an arbitrary distribution function $F_A(\cdot)$ with density $f_A(\cdot)$. The mean interarrival time is equal to $1/\lambda$. The service times are exponentially distributed with mean $1/\mu$. For stability we again require that the occupation rate $\rho = \lambda/\mu$ is less than one.

The state of the $G/M/1$ queue can be described by the pair (n, x) where n denotes the number of customers in the system and x the elapsed time since the last arrival. So we need a complicated two-dimensional state description. However, like for the $M/G/1$ queue, the state description is much easier at special points in time. If we look at the system on arrival instants, then the state description can be simplified to n only, because $x = 0$ at an arrival. Denote by L_k^a the number of customers in the system just before the kth arriving customer. In the next section we will determine the limiting distribution

$$a_n = \lim_{k\to\infty} P(L_k^a = n).$$

From this distribution we will be able to calculate the distribution of the sojourn time.

5.1 Arrival distribution

In this section we will determine the distribution of the number of customers found in the system just before an arriving customer when the system is in equilibrium.

We first derive a relation between the random variables L_{k+1}^a and L_k^a. Defining the random variable D_{k+1} as the number of customers served between the arrival of the kth and $k+1$th customer, it follows that

$$L_{k+1}^a = L_k^a + 1 - D_{k+1}.$$

From this equation it is immediately clear that the sequence $\{L_k^a\}_{k=0}^{\infty}$ forms a Markov chain. This Markov chain is called the $G/M/1$ *imbedded Markov chain*.

We must now calculate the associated transition probabilities

$$p_{i,j} = P(L_{k+1}^a = j | L_k^a = i).$$

Clearly $p_{i,j} = 0$ for all $j > i+1$ and $p_{i,j}$ for $j \le i+1$ is equal to the probability that exactly $i+1-j$ customers are served during the interarrival time of the $k+1$th customer. Hence the matrix P of transition probabilities takes the form

$$P = \begin{pmatrix} p_{0,0} & \beta_0 & 0 & \cdots & \\ p_{1,0} & \beta_1 & \beta_0 & 0 & \cdots \\ p_{2,0} & \beta_2 & \beta_1 & \beta_0 & 0 \\ p_{3,0} & \beta_3 & \beta_2 & \beta_1 & \beta_0 \\ \vdots & & & & \ddots \end{pmatrix},$$

where β_i denotes the probability of serving i customers during an interarrival time given that the server remains busy during this interval (thus there are more than i customers present). To calculate β_i we note that given the duration of the interarrival time, t say, the number of customers served during this interval is Poisson distributed with parameter μt. Hence, we have

$$\beta_i = \int_{t=0}^{\infty} \frac{(\mu t)^i}{i!} e^{-\mu t} f_A(t) dt. \tag{1}$$

Since the transition probabilities from state j should add up to one, it follows that

$$p_{i,0} = 1 - \sum_{j=0}^{i} \beta_j = \sum_{j=i+1}^{\infty} \beta_j .$$

The transition probability diagram is shown in figure 1.

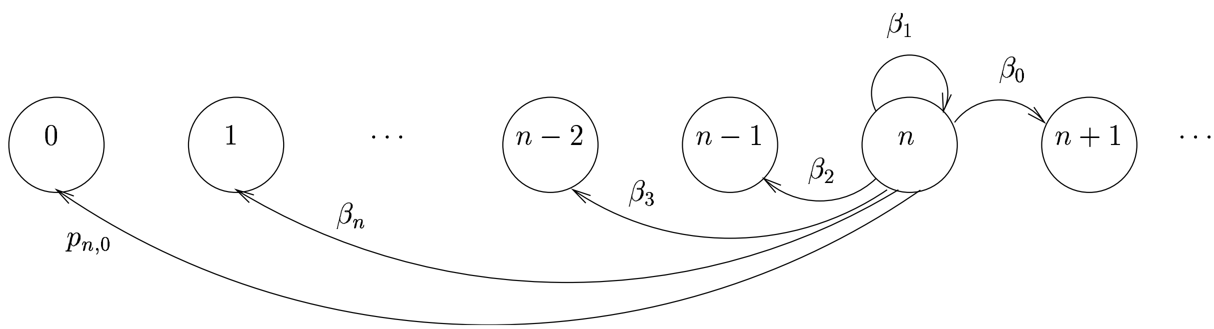

Figure 1: Transition probability diagram for the $G/M/1$ imbedded Markov chain

This completes the specification of the imbedded Markov chain. We now wish to determine its limiting distribution $\{a_n\}_{n=0}^{\infty}$. The limiting probabilities a_n satisfy the equilibrium equations

$$\begin{aligned} a_0 &= a_0 p_{0,0} + a_1 p_{1,0} + a_2 p_{2,0} + \cdots \\ &= \sum_{i=0}^{\infty} a_i p_{i,0} \qquad (2) \\ a_n &= a_{n-1}\beta_0 + a_n \beta_1 + a_{n+1}\beta_2 + \cdots \\ &= \sum_{i=0}^{\infty} a_{n-1+i}\beta_i, \qquad n = 1, 2, \ldots \qquad (3) \end{aligned}$$

To find the solution of the equilibrium equations it appears that the generating function approach does not work here (verify). Instead we adopt the direct approach by trying to find solutions of the form

$$a_n = \sigma^n, \quad n = 0, 1, 2, \ldots \tag{4}$$

Substitution of this form into equation (3) and dividing by the common power σ^{n-1} yields

$$\sigma = \sum_{i=0}^{\infty} \sigma^i \beta_i \, .$$

Of course we know that β_i is given by (1). Hence we have

$$\begin{aligned}\sigma &= \sum_{i=0}^{\infty} \sigma^i \int_{t=0}^{\infty} \frac{(\mu t)^i}{i!} e^{-\mu t} f_A(t)dt \\ &= \int_{t=0}^{\infty} e^{-(\mu-\mu\sigma)t} f_A(t)dt.\end{aligned}$$

The last integral can be recognised as the Laplace-Stieltjes transform of the interarrival time. Thus we arrive at the following equation

$$\sigma = \widetilde{A}(\mu - \mu\sigma). \tag{5}$$

We immediately see that $\sigma = 1$ is a root of equation (5), since $\widetilde{A}(0) = 1$. But this root is not useful, because we must be able to normalize the solution of the equilibrium equations. It can be shown that as long as $\rho < 1$ equation (5) has a unique root σ in the range $0 < \sigma < 1$, and this is the root which we seek. Note that the remaining equilibrium equation (2) is also satisfied by (4) since the equilibrium equations are dependent. We finally have to normalize solution (4) yielding

$$a_n = (1 - \sigma)\sigma^n\,, \quad n = 0, 1, 2, \ldots \tag{6}$$

Thus we can conclude that the queue length distribution found just before an arriving customer is *geometric* with parameter σ, where σ is the unique root of equation (5) in the interval $(0, 1)$.

Example 5.1 *(M/M/1)*
For exponentially distributed interarrival times we have

$$\widetilde{A}(s) = \frac{\lambda}{\lambda + s}\,.$$

Hence equation (5) reduces to

$$\sigma = \frac{\lambda}{\lambda + \mu - \mu\sigma}\,,$$

so

$$\sigma(\lambda + \mu - \mu\sigma) - \lambda = (\sigma - 1)(\lambda - \mu\sigma) = 0.$$

Thus the desired root is $\sigma = \rho$ and the arrival distribution is given by

$$a_n = (1 - \rho)\rho^n\,, \quad n = 0, 1, 2, \ldots$$

Note that this distribution is exactly the same as the equilibrium distribution of the $M/M/1$. This is of course no surprise, because here we have Poisson arrivals.

Example 5.2 *($E_2/M/1$)*
Suppose that the interarrival times are Erlang-2 distributed with mean 2/3, so

$$\widetilde{A}(s) = \left(\frac{3}{3+s}\right)^2 .$$

Further assume that $\mu = 4$ (so $\rho = 3/2 \cdot 1/4 = 3/8 < 1$). Then equation (5) reduces to

$$\sigma = \left(\frac{3}{7-4\sigma}\right)^2 .$$

Thus

$$\sigma(7-4\sigma)^2 - 9 = (\sigma - 1)(4\sigma - 9)(4\sigma - 1) = 0.$$

Hence the desired root is $\sigma = 1/4$ and

$$a_n = \frac{3}{4}\left(\frac{1}{4}\right)^n , \quad n = 0, 1, 2, \ldots$$

Example 5.3
Suppose that the interarrival time consist of two exponential phases, the first phase with parameter μ and the second one with parameter 2μ (so it is slightly more complicated than Erlang-2 where both phases have the same parameter), where μ is also the parameter of the exponential service time. The Laplace-Stieltjes transform of the interarrival time is given by

$$\widetilde{A}(s) = \frac{2\mu^2}{(\mu + s)(2\mu + s)} .$$

For this transform equation (5) reduces to

$$\sigma = \frac{2\mu^2}{(2\mu - \mu\sigma)(3\mu - \mu\sigma)} = \frac{2}{(2-\sigma)(3-\sigma)} .$$

This leads directly to

$$\sigma^3 - 5\sigma^2 + 6\sigma - 2 = (\sigma - 1)(\sigma - 2 - \sqrt{2})(\sigma - 2 + \sqrt{2}) = 0.$$

Clearly only the root $\sigma = 2 - \sqrt{2}$ is acceptable. Therefore we have

$$a_n = (\sqrt{2} - 1)(2 - \sqrt{2})^n , \quad n = 0, 1, 2, \ldots$$

5.2 Distribution of the sojourn time

Since the arrival distribution is geometric, it is easy to determine the distribution of the sojourn time. In fact, the analysis is similar to the one for for the $M/M/1$ queue. With probability a_n an arriving customer finds n customers in the system. Then his sojourn time is the sum of $n+1$ exponentially distributed service times, each with mean $1/\mu$. Hence,

$$\begin{aligned}\widetilde{S}(s) &= E(e^{-sS}) \\ &= \sum_{n=0}^{\infty} a_n \left(\frac{\mu}{\mu+s}\right)^{n+1} \\ &= \sum_{n=0}^{\infty} (1-\sigma)\sigma^n \left(\frac{\mu}{\mu+s}\right)^{n+1} \\ &= \frac{\mu(1-\sigma)}{\mu+s} \sum_{n=0}^{\infty} \left(\frac{\mu\sigma}{\mu+s}\right)^{n} \\ &= \frac{\mu(1-\sigma)}{\mu(1-\sigma)+s}\,.\end{aligned}$$

From this we can conclude that the sojourn time S is exponentially distributed with parameter $\mu(1-\sigma)$, i.e.,

$$P(S \leq t) = 1 - e^{-\mu(1-\sigma)t}, \quad t \geq 0.$$

Clearly the sojourn time distribution for the $G/M/1$ is of the *same form* as for the $M/M/1$, the only difference being that ρ is replaced by σ.

Along the same lines it can be shown that the distributon of the waiting time W is given by

$$P(W \leq t) = 1 - \sigma e^{-\mu(1-\sigma)t}, \quad t \geq 0.$$

Note that the probability that a customer does not have to wait is given by $1-\sigma$ (and *not* by $1-\rho$).

5.3 Mean sojourn time

It is tempting to determine the mean sojourn time directly by the mean value approach. For an arriving customer we have

$$E(S) = E(L^a)\frac{1}{\mu} + \frac{1}{\mu}\,, \tag{7}$$

where the random variable L^a denotes the number of customers in the system found on arrival. According to Little's law it holds that

$$E(L) = \lambda E(S). \tag{8}$$

Unfortunately, we do not have Poisson arrivals, so

$$E(L^a) \neq E(L).$$

Hence the mean value approach does not work here, since we end up with only two equations for three unknowns. Additional information is needed in the form of (6), yielding

$$E(L^a) = \sum_{n=0}^{\infty} n a_n = \sum_{n=0}^{\infty} n(1-\sigma)\sigma^n = \frac{\sigma}{1-\sigma}\,.$$

Then it follows from (7) and (8) that

$$E(S) = \frac{\sigma}{(1-\sigma)\mu} + \frac{1}{\mu} = \frac{1}{(1-\sigma)\mu}\,, \qquad E(L) = \frac{\lambda}{(1-\sigma)\mu} = \frac{\rho}{(1-\sigma)}\,.$$