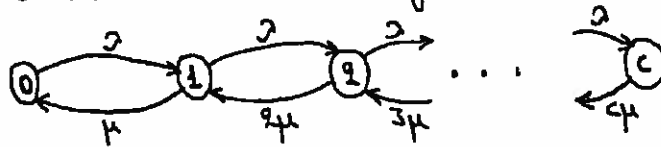


Problem 8.4

let the state be the number of busy servers. The state space  
 $\{0, 1, 2, \dots, c\}$

The state transition rate diagram is



From the diagram, we can write the balance equations

$$\lambda P_0 = \mu P_1$$

$$\lambda P_1 = 2\mu P_2$$

...

$$\lambda P_{c-1} = c\mu P_c$$

where  $P_i = P[\text{the number of busy servers at steady state} = i]$   
 $(i = 0, 1, \dots, c)$

and

$$\lambda = 40 \text{ customers/hour}, \quad \mu = 60 \text{ customers/hour}$$

Solving the balance equations, we have

$$P_i = \frac{\lambda^i}{\mu^i i!} P_0 = \left(\frac{\lambda}{\mu}\right)^i \frac{1}{i!} P_0 \quad (i = 0, 1, 2, \dots, c) \quad (1)$$

Using the fact that  $\sum_{i=0}^c P_i = 1$ , we have that

$$P_0 = \left[ \sum_{i=0}^c \left(\frac{\lambda}{\mu}\right)^i \left(\frac{1}{i!}\right) \right]^{-1} \quad (2)$$

We know that

$P_c = \text{Prob. that all ATM are busy}$

$= \text{Prob. arrivals find all ATM's busy}$

Since, in one hour, there are 40 customers arriving, <sup>on average</sup> the <sup>expected</sup> loss per hour is  $40p_c$  dollars.

We want the loss per hour to be less than 5 dollars. That is

$$40p_c < 5 \Rightarrow p_c < \frac{1}{8}$$

Using (1) and (2) :

$$\frac{\left(\frac{2}{3}\right)^c \frac{1}{c!}}{\sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n \frac{1}{n!}} < \frac{1}{8} = 0.1250$$

When  $c=1$ ,  $p_c=0.4$

$c=2$ ,  $p_c=0.1176 < 0.1250$

Hence,  
 $c=2$  is recommended.

### Problem 8.5

(1)

As suggested in the hint that was provided in the homework, this workstation can be modeled as an  $M/M/1/N$  queue and analyzed according to the results provided in Section 8.6.6 of your textbook by *Cassandras and Lafortune*.

In our problem, we have  $\rho = 12 \text{ part/hr}$  and  $\beta = \frac{1}{15 \text{ min}} = \frac{1}{\frac{1}{4} \text{ hr}} = 4 \text{ part/hr}$ . So,  $\rho = \frac{\lambda}{\mu} = \frac{12}{4} = 3$ .

Now, the average waiting time at the queue equals the average time spent at the station minus the average service time. Following the textbook notation:

$$E[W] = E[R] - \frac{1}{\mu}$$

and from (8.70)

$$E[R] = \frac{N}{\mu(1-\pi_0)} - \frac{1}{\mu}$$

Also, from (8.87)

$$\pi_0 = \left[ \sum_{n=0}^N \frac{N!}{(N-n)!} \rho^n \right]^{-1}$$

Hence, the posed requirement can be expressed by:

$$E[W] = \frac{N}{12 \left( 1 - \frac{1}{\sum_{n=0}^N \frac{N!}{(N-n)!} \rho^n} \right)} - \frac{1}{4} - \frac{1}{12} < \frac{2}{60} \Leftrightarrow$$

$$\Leftrightarrow \frac{N}{1 - \left[ \sum_{n=0}^N \frac{N!}{(N-n)!} \rho^n \right]^{-1}} < \frac{22}{5} = 4.4 \quad (1)$$

The corresponding utilization of the workstation server is:

$$U(N) = 1 - \pi_0(N) \quad (2)$$

②

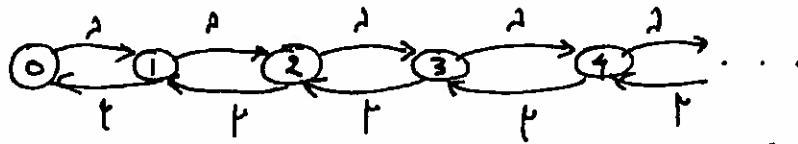
The required  $N$  can be identified from (1) through binary search and the corresponding utilization can be obtained from (2).

Working like this we find that the required number of fixtures is  $N=2$  and the resulting utilization is

$$u(2) \approx 0.47.$$

b)

For parts (i) and (ii), let  $\{X_t, t \geq 0\}$  denote the stochastic process <sup>①</sup> that tracks the number of taxis waiting at the station. Then, this process is a CTMC with the following structure:



where the transition rates  $\lambda$  and  $\mu$  are defined as follows:

- \*  $\lambda$  = taxi arrival rate =  $1/\text{min}$
- \*  $\mu$  = customer " " =  $2/\text{min}$

In fact, the above CTMC is a birth-death process with identical structure to the birth-death process that models the dynamics of an M/M/1 queue with arrival rate  $\lambda$  and processing rate  $\mu$ .

Hence, working as in that case, we can infer the following:

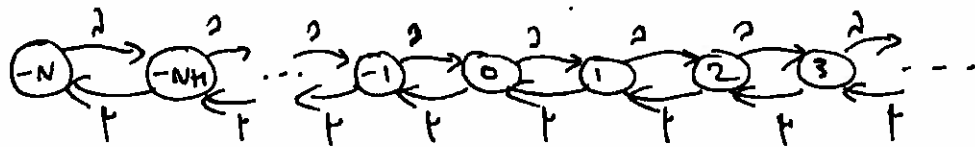
The above CTMC is ergodic (i.e., it reaches steady-state) iff  $\rho = \lambda/\mu < 1$ . Then, the steady-state distribution is given by

$\pi_i = (1-\rho)\rho^i$ ,  $i=0,1,2,\dots$ . Furthermore, the average # of taxis waiting is given by  $\sum_{i=0}^{\infty} i\pi_i = \dots = \frac{\rho}{1-\rho}$ . Finally, an arriving customer will get a taxi as long as the considered CTMC is in a state  $X_t \geq 1$ . But the total probability for this class of states is  $1 - \pi_0 = 1 - (1-\rho) = \rho$ .

For the provided values for  $\lambda$  and  $\mu$ , we can see that ~~the considered~~  $\rho = \frac{1}{2} < 1$  and the considered system is ergodic. ~~the~~ The average # of taxis waiting is equal to  $\frac{1/2}{1-1/2} = 1$ . Also, an arriving customer will be served with prob.  $1 - 1/2 = 0.5$ .

(2)

To answer part (iii), we need also to leave the evolution of the customer queue in our definition of the system state. Recognizing, however, that only one of the two queues (i.e., taxis or customers) can be non-empty at any time point, the necessary extension of the CTMC structure can be performed as follows:



In the above state space, a negative state models accumulation of customers and a positive state models accumulation of taxis.

Furthermore, under an appropriate relabeling of the states, the process retains the structure of the CTMC that models an M/M/L queue with respective arrival and service rates  $\lambda$  and  $\mu$ . Hence, under the assumption that  $\rho = \lambda/\mu < 1$ , the steady-state distribution of this process is characterized as follows:

$$\pi_i = (1-\rho) \rho^{N+i}, \quad i = -N, -N+1, -N+2, \dots$$

But then,

$$\begin{aligned} \text{Avg \# of taxis waiting} &= \sum_{i=1}^{\infty} i (1-\rho) \rho^{N+i} = \rho^N (1-\rho) \sum_{i=1}^{\infty} i \rho^i = \\ &= \rho^N (1-\rho) \frac{\rho}{(1-\rho)^2} = \frac{\rho^{N+1}}{1-\rho} \end{aligned}$$

$$\begin{aligned} \text{Also, Prob an arriving customer will get a taxi} &= 1 - \pi_N = \\ &= 1 - (1-\rho) = \rho \end{aligned}$$

(3)

The average waiting time for a customer who joined the queue can be expressed as follows:

$$W = \sum_{i=-N}^{\infty} a_i \tau_i$$

where

\*  $a_i$  = prob. that the ~~new~~ <sup>joining</sup> customer found the system in state  $i$ ,

$i = -N, -N+1, -N+2, \dots$

\*  $\tau_i$  = expected waiting time given that the <sup>joining</sup> customer found the system in state  $i$ .

From PASTA we have that:

$$a_i = \begin{cases} 0, & i = -N \\ \frac{\pi_i}{1 - \pi_{-N}}, & i = -N+1, -N+2, \dots \end{cases}$$

Also, from the memoryless property of the exp. distribution that characterizes the taxi inter-arrival times, we get that

$$\tau_i = \begin{cases} (-i+1)/\lambda, & i = -N+1, -N+2, \dots, 0 \\ 0, & \text{o.w.} \end{cases}$$

Hence,

$$W = \frac{1}{1 - \pi_{-N}} \left[ \pi_{-N+1} \frac{N}{\lambda} + \pi_{-N+2} \frac{N-1}{\lambda} + \dots + \pi_{-1} \frac{2}{\lambda} + \pi_0 \frac{1}{\lambda} \right] =$$

$$= \frac{1-p}{p\lambda} \left[ Np + (N-1)p^2 + \dots + 2p^{N-1} + p^N \right] =$$

$$= \frac{1-p}{\lambda} \sum_{i=0}^{N-1} (N-i)p^i = \frac{1-p}{\lambda} \left[ N \sum_{i=0}^{N-1} p^i - \sum_{i=0}^{N-1} ip^i \right]$$

(4)

Since we need  $p < 1$  for stability (ergodicity), we proceed as follows:

$$\sum_{i=0}^{N-1} p^i = \frac{1-p^N}{1-p}$$

$$\begin{aligned} \sum_{i=0}^{N-1} i p^i &= \sum_{i=1}^{N-1} i p^i = p \sum_{i=1}^{N-1} i p^{i-1} = p \frac{d}{dp} \left( \sum_{i=1}^{N-1} p^i \right) = p \frac{d}{dp} \left( \frac{1-p^N}{1-p} - 1 \right) \\ &= p \frac{d}{dp} \left( \frac{1-p^N}{1-p} \right) = \frac{p}{(1-p)^2} \left[ (1-p)(-N p^{N-1}) - (1-p^N)(-1) \right] \\ &= \frac{p}{(1-p)^2} \left[ (N-1)p^N - N p^{N-1} + 1 \right] \end{aligned}$$

Hence,

$$\begin{aligned} W &= \frac{1}{2} \left\{ N(1-p^N) - \frac{p}{1-p} \left[ (N-1)p^N - N p^{N-1} + 1 \right] \right\} = \\ &= \frac{1}{2(1-p)} \left[ N - N p^N - N p + N p^{N+1} - (N-1)p^{N+1} + N p^N - p \right] = \\ &= \frac{1}{2(1-p)} \left[ N - (N+1)p + p^{N+1} \right] \end{aligned}$$

The above expression is pretty compact and therefore useful for computation. On the other hand, it can also be proven further as follows:

$$\begin{aligned} W &= \frac{1}{2(1-p)} \left\{ N(1-p) - p(1-p^N) \right\} = \frac{1}{2} \left[ N - p \sum_{i=0}^{N-1} p^i \right] = \\ &= \frac{1}{2} \sum_{i=1}^N (1-p^i) \end{aligned}$$

What is an intuitive interpretation of the last expression?

(5)

$$\begin{aligned}
 W = \frac{1}{\lambda} \sum_{i=1}^N (1-p^i) &= \frac{1}{\lambda} \times \text{Prob}(\text{joining customer waits for his own cab}) + \\
 &+ \frac{1}{\lambda} \times \text{Prob}(\text{joining "also" " " ~~also~~ the cab of another customer}) + \\
 &+ \frac{1}{\lambda} \times \text{Prob}(\text{joining customer also waits for the cab of a third customer}) + \dots + \\
 &+ \frac{1}{\lambda} \times \text{Prob}(\text{joining customer also waits for the cab of the (N-1)st customer in the queue})
 \end{aligned}$$

$$\begin{aligned}
 \text{Prob}(\text{joining customer waits for his own cab}) &= \sum_{i=-NH}^0 a_i = 1 - \sum_{i=1}^{\infty} a_i = \\
 &= 1 - \sum_{i=1}^{\infty} \frac{\pi_i}{1 - \pi_{-N}} = 1 - \frac{1}{p} \sum_{i=1}^{\infty} (1-p) p^{Ni} = 1 - (1-p) p^N \sum_{i=0}^{\infty} p^i = \\
 &= 1 - p^N
 \end{aligned}$$

$$\begin{aligned}
 \text{Similarly, Prob}(\text{joining customer also waits for the cab of the (N-1)st customer}) &= \\
 &= \sum_{i=-NH}^{-NH} a_i = 1 - \sum_{i=-NH+1}^{\infty} a_i = 1 - \frac{1}{p} \sum_{i=-NH+1}^{\infty} \pi_i = 1 - \frac{1}{p} \sum_{i=-NH+1}^{\infty} (1-p) p^{Ni} = \\
 &= 1 - \frac{(1-p)}{p} p^2 \sum_{i=0}^{\infty} p^i = 1 - p
 \end{aligned}$$

The probabilities for the other terms in the above sum can be computed in the same manner.

(6)

An alternative computation of  $W$  can be developed through Little's law. Let  $Q$  denote the average # of customers waiting for a taxi. Then, from Little's law,

$$W = \frac{Q}{p\mu} = \frac{Q}{\lambda}$$

where we have taken into consideration that from the stream of the arriving customers, only a percentage  $p$  of them join the queue.

From its definition,  $Q$  itself can be computed as follows:

$$\begin{aligned} Q &= \sum_{i=1}^N (-i) \pi_i = - \sum_{i=1}^N i (1-p) p^{N+i} = (1-p) \sum_{i=1}^N i p^{N-i} = \\ &= (1-p) p^N \sum_{i=1}^N i p^{-i} = (1-p) p^{N-1} \sum_{i=1}^N i p^{-(i-1)} = (1-p) p^{N-1} \sum_{i=1}^N i u^{(i-1)} \end{aligned}$$

where  $u = 1/p$

$$\text{But } \sum_{i=1}^N i u^{i-1} = \frac{d}{du} \sum_{i=0}^N u^i = \frac{d}{du} \frac{1-u^{N+1}}{1-u} = \frac{-(N+1)u^N(1-u) + (1-u^{N+1})}{(1-u)^2}$$

$$= \frac{-(N+1)(u^N - u^{N+1}) + 1 - u^{N+1}}{(1-u)^2} = \frac{Nu^{N+1} - (N+1)u^N + 1}{(1-u)^2}$$

$$= \frac{N/p^{N+1} - (N+1)/p^N + 1}{(1-1/p)^2} = \frac{N - (N+1)p + p^{N+1}}{p^{N+1} \frac{(1-p)^2}{p^2}} = \frac{N - (N+1)p + p^{N+1}}{p^{N-1} (1-p)^2}$$

Substituting back to the expression for  $Q$  we get:

$$Q = (1-p) p^{N-1} \frac{N - (N+1)p + p^{N+1}}{p^{N-1} (1-p)^2} = \frac{N - (N+1)p + p^{N+1}}{1-p}$$

And finally,

$$W = \frac{1}{\lambda(1-p)} [N - (N+1)p + p^{N+1}]$$

7

For part (iv), notice that when  $\lambda$  increases to 2 taxis per minute,  $\rho = 1$ , and therefore, the considered birth-death process becomes unstable. In this regime, the # of taxis waiting will grow to infinity and therefore arriving customers will always get a taxi immediately.

(1)

## Part C

- Consider the  $M/M/1$  queueing system operating in steady state and let:
- $W$  denote the random time spent in system by an arriving customer,
  - $N$  denote the random number of customers in the system that are encountered by the aforementioned customer upon his arrival;
  - $F_W(t)$  denote the cdf for  $W$ .

Then,

$$\begin{aligned}
 F_W(t) &\equiv P(W \leq t) = \\
 &= \sum_{n=0}^{\infty} P(W \leq t \mid N=n) P(N=n) = \\
 &= \sum_{n=0}^{\infty} F_{\text{Erlang}(n+1, \mu)}(t) \cdot a_n
 \end{aligned}$$

where the last inequality results by

- setting  $a_n \equiv P(N=n)$ , and
  - taking into consideration that in an  $M/M/1$  queue
    - \* customers are processed in a FCFS basis, and
    - \* their processing times are iid r.v.'s distributed according to an exponential distribution with rate  $\mu$ .
- Furthermore, from the Markovian nature of the arrival process and PASTA, we have:

$$a_n = \pi_n = (1-\rho)\rho^n \quad \text{where} \quad \rho = \lambda/\mu < 1 \quad (\text{from stability})$$

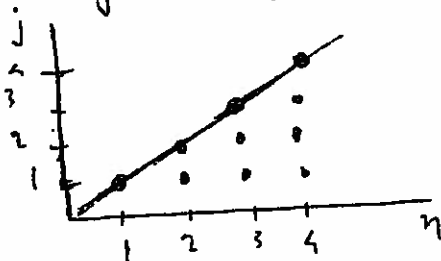
Then,

(2)

$$\begin{aligned}
F_n(t) &= (1-p) \sum_{n=0}^{\infty} p^n \left[ 1 - e^{-\lambda t} \sum_{j=0}^n \frac{(\lambda t)^j}{j!} \right] = \\
&= (1-p) \sum_{n=0}^{\infty} p^n - (1-p) e^{-\lambda t} \sum_{n=0}^{\infty} p^n \sum_{j=0}^n \frac{(\lambda t)^j}{j!} = * \\
&= 1 - (1-p) e^{-\lambda t} \sum_{j=0}^{\infty} \sum_{n=0}^{\infty} p^{j+n} \frac{(\lambda t)^j}{j!} = \\
&= 1 - (1-p) e^{-\lambda t} \sum_{j=0}^{\infty} \frac{(\lambda t)^j}{j!} \sum_{n=0}^{\infty} p^n = \\
&= 1 - (1-p) e^{-\lambda t} \frac{1}{1-p} \sum_{j=0}^{\infty} \frac{(\lambda t)^j}{j!} = \\
&= 1 - e^{-\lambda t} e^{\lambda t} = 1 - e^{-\lambda t} e^{\lambda t} = \\
&= 1 - e^{-(\lambda-2)t}
\end{aligned}$$

But this is the cdf of an exponential distribution with rate  $\lambda-2$ .

To understand better the step denoted by (\*) in the above derivation, consider the following figure:



The double sums are computed for the pairs indicated by the dotted matrix above, and in the first expression the summation is computed on a column by column basis, while in the second expression on a row by row basis.

An insightful interpretation of this result is as follows:

From the dynamics of the  $M/M/1$  queue we have that

$$\begin{aligned} & \text{Prob (an arriving customer leaves at the } n\text{-th processing completed} \\ & \quad \text{after his arrival)} = \\ & = \text{Prob (arriving customer encountered } n-1 \text{ customers when he} \\ & \quad \text{joined the queue)} = \\ & = \pi_{n-1} = \rho^{n-1}(1-\rho), \quad n=1, 2, \dots \end{aligned}$$

The next-to-last step in the above derivation results from PASTA.

Furthermore, the above result implies a geometric distribution with "success" prob. equal to  $1-\rho$ ; each "trial" in the interpretation of this distribution corresponds to a job completion, and the "successful trial" is the job completion that corresponds to the "tagged" customer.

While the customer is in the system, the server must be continuously ~~be~~ busy, and therefore, during this time job completions take place with rate  $\mu$ . The above discussion further implies that any of these completions will be the completion of our tagged customer with prob  $1-\rho$ , and therefore, by the properties of the Poisson process, the completion of our tagged customer occurs with rate  $\mu(1-\rho) = \mu - \lambda$ . Since this rate is constant, the <sup>continuous</sup> distribution that characterizes the "time-to-completion" (i.e., otherwise, the sojourn time) for our customer is exponential with the same rate.

## Part d

①

For parts (1-3) of this problem, you can see the pages 7 and 8 from the excerpt by Chen and Yao on "Birth-Death processes and Jackson queueing networks" that has been ~~posted at the library~~ <sup>provided in the next pages</sup> academic reserves. Part (4) follows directly from the result of part (3).

An alternative derivation of part (4) that reveals the connection of this result to some other results presented in class, and also to part (c) of this homework, is as follows:

Let  $X(\bar{t}_a) = \#$  of customers that are encountered by an arrival at steady state

$S =$  the sojourn time of this arrival

Then,

$$\begin{aligned} P(X(\bar{t}_a) = n | S = t) &= \lim_{\Delta t \rightarrow 0} \frac{P[X(\bar{t}_a) = n \wedge S \in [t, t + \Delta t)]}{P[S \in [t, t + \Delta t)]} = \\ &= \lim_{\Delta t \rightarrow 0} \frac{P[X(\bar{t}_a) = n] \cdot P[S \in [t, t + \Delta t) | X(\bar{t}_a) = n] / \Delta t}{P[S \in [t, t + \Delta t)] / \Delta t} = \\ &= \frac{\text{pmf}[\text{Geometric}(n; 2/\mu)] \text{pdf}[\text{Erlang}(t; n+1, \mu)]}{\text{pdf}[\text{Exponential}(t; \mu-2)]} = \\ &= \frac{\left[\left(\frac{2}{\mu}\right)^n \left(1 - \frac{2}{\mu}\right)\right] \cdot \left[\mu^{n+1} t^n \cdot e^{-\mu t} / n!\right]}{(\mu-2) e^{-(\mu-2)t}} = \frac{(2t)^n e^{-2t}}{n!} = \\ &= \text{pmf}[\text{Poisson}(n; 2t)]. \end{aligned}$$

**Proof.** Pick any pair of nodes  $(i, j)$ . If  $q(i, j) = 0$ , then  $q(j, i) = 0$ , following (1.9), and the detailed balance equation (1.8) is trivially satisfied. If  $q(i, j) > 0$ , then there is an edge linking the pair. Since the graph is a tree, the only probability flow between  $i$  and  $j$  is through this edge. In other words, the full balance equation for state  $i$  (or for state  $j$ ) reduces to the detailed balance equation between  $i$  and  $j$ . That is, (1.8) is satisfied.  $\square$

**Example 1.7** A special case of the above is the birth-death process. The associated graph is a line, or a single-branch tree. Here,  $q(i, j) = \lambda(i)$  or  $\mu(i)$ , respectively for  $j = i + 1$  and  $j = i - 1$ , the birth and the death rates; and the familiar relation [cf. (1.1)]

$$\lambda(i)\pi(i) = \mu(i+1)\pi(i+1)$$

is nothing but the detailed balance equation.

Below are some quick implications of Theorem 1.4, with the proofs left as exercise problems.

**Corollary 1.8** Suppose  $\{X(t)\}$  is a reversible Markov chain on state space  $S$ . Suppose we truncate the state space to  $S' \subset S$  (by letting some transition rates be zero, for instance). Then, the modified Markov chain is still reversible, with invariant distribution

$$\pi'(i) = \pi(i) / \sum_{j \in S'} \pi(j), \quad \forall i \in S'.$$

**Corollary 1.9** Suppose a stationary Markov chain  $\{X(t)\}$  has a finite state space and is irreducible. Then,  $\{X(t)\}$  is reversible if and only if its rate matrix  $Q$  can be expressed as  $Q = AD$  with  $A$  being a symmetric matrix and  $D$  a diagonal matrix.

**Example 1.10** Consider the  $M/M(n)/1$  queue, with Poisson arrivals at a constant rate  $\lambda$  and state-dependent service rates  $\mu(n)$ . Let  $\{X(t)\}$  be the queue process:  $X(t)$  denotes the total number of jobs in the system at time  $t$ . This is a special case of the birth-death process in Example 1.7. Hence, in equilibrium  $\{X(t)\}$  is reversible.

Now consider the time-reversal  $\{\bar{X}(t)\}$ , which represents another birth-death queue, whose departure process is the arrival process of the original queue, i.e., a Poisson process with constant rate  $\lambda$ . But because of reversibility, the original queue and its time-reversal have the same probabilistic behavior; and in particular, the departure processes in the two queues are exactly the same. Hence, we can conclude that the departure process from a stationary  $M/M(n)/1$  queue is Poisson with constant rate  $\lambda$ , exactly the same as the arrival process.

Furthermore, in the original queue the number of future arrivals after  $t$  is independent of the state—the number of jobs in the system—at  $t$ .

This is simply because the Poisson arrival process has a constant rate. By reversibility, this immediately translates into the fact that the number of past departures up to time  $t$  in the time-reversed queue, and hence also in the original queue, is independent of the number of jobs in the system at  $t$ . This is a counterintuitive result, since one would tend to think that the number of jobs in the system at time  $t$ , e.g., whether there is zero or at least one job to keep the server busy, would depend on the number of past departures up to  $t$ .

### 1.3 Stochastic Orders

Suppose  $X$  is a nonnegative, integer-valued random variable. Let  $p(n) = P(X = n) > 0$  for all  $n \in \mathcal{N} = \{0, 1, \dots, N\}$ , where  $N$  is a given integer. We allow  $N$  to be infinite.

**Definition 1.11** The *equilibrium rate* of  $X$  is a real-valued, nonnegative function,  $r : \mathcal{N} \mapsto \mathbb{R}_+ = [0, \infty)$  defined as

$$r(0) = 0, \quad r(n) = p(n-1)/p(n), \quad n = 1, \dots, N.$$

It follows that the equilibrium rate of  $X$  and its probability mass function (pmf) uniquely define each other. In particular,

$$p(n) = p(0)/[r(1) \cdots r(n)], \quad n = 1, \dots, N,$$

and

$$p(0) = \left[ 1 + \sum_{n=1}^N 1/(r(1) \cdots r(n)) \right]^{-1}.$$

(When  $N = \infty$ , convergence of the summation is required.)

**Example 1.12** For a birth-death queue in equilibrium, with unit arrival rate and state-dependent service rate  $\mu(n)$ , we have

$$\mu(n) = P[Y = n-1]/P[Y = n] = r_Y(n),$$

where  $Y$  denotes the number of jobs in the system, and  $r_Y(n)$  its equilibrium rate. That is, the equilibrium rate is simply the service rate.

It turns out that the equilibrium rates are useful in comparing random variables under the *likelihood ratio ordering*.

**Definition 1.13**  $X$  and  $Y$  are two discrete random variables. Suppose their pmfs have a common support set  $\mathcal{N}$ . Let  $r_X$  and  $r_Y$  denote their

(1)

Part e

Consider the  $i$ -th arrival. Due to the preemptive nature of the applied LIFO policy, this job goes immediately in service and it will finish in  $Z_i$  time units iff  $Z_i \leq Y_{i+1}$ .

If, on the other hand,  $Z_i > Y_{i+1}$ , then this job will be preempted by the next arrival, and it will finish in  $Z_i + Z_{i+1}$  time units iff  $Z_i + Z_{i+1} \leq Y_{i+1} + Y_{i+2}$ .

More generally, let  $\bar{j} = \min \{ j=0,1,2,\dots \text{ s.t. } \sum_{k=i}^{i+j} Z_k \leq \sum_{k=i+1}^{i+j+1} Y_k \}$

Then, reasoning as above, we can see that the sojourn time for job  $i$  is

$$S_i = \sum_{k=i}^{i+\bar{j}} Z_k$$

3)

- (1) Consider the  $n$ -th arrival at this queue. Assuming that the first arrival took place at time  $t = 0$ , the  $n$ -th arrival takes place at time  $t = (n-1)\tau$ . Since service times are exponentially distributed with rate  $\mu$ , the  $i$ -th arrival, for  $i = 1, \dots, n-1$ , will be still in service upon the arrival of the  $n$ -th job with prob.  $p_i = e^{-\mu(n-i)\tau}$ . Hence, setting  $A_n = \#$  of job encountered by the  $n$ -th arrival, we have that

$$\begin{aligned} E[A_n] &= E\left[\sum_{i=1}^{n-1} I_{\{i\text{-th job still in service upon } n\text{-th arrival}\}}\right] = \\ &= \sum_{i=1}^{n-1} E[I_{\{i\text{-th job still in service}\}}] = \sum_{i=1}^{n-1} p_i = \sum_{i=1}^{n-1} e^{-\mu(n-i)\tau} = \\ &= \sum_{j=1}^{n-1} e^{-\mu j\tau}, \text{ where in the last step we have set } j = n-i. \end{aligned}$$

In steady state, we have:

$$\begin{aligned} E[A] &= E\left[\lim_{n \rightarrow \infty} A_n\right] = E\left[\lim_{n \rightarrow \infty} \sum_{i=1}^{n-1} I_{\{i\text{-th job still in service}\}}\right] = \\ &= \lim_{n \rightarrow \infty} E\left[\sum_{i=1}^{n-1} I_{\{i\text{-th job still in service}\}}\right] = \lim_{n \rightarrow \infty} \sum_{j=1}^{n-1} e^{-\mu j\tau} = \\ &= \sum_{j=1}^{\infty} (e^{-\mu\tau})^j, \text{ where the interchange of the lim and the expectation} \end{aligned}$$

operator is justified by monotone convergence. Furthermore, since  $\mu\tau > 0$ , we have:

$$E[A] = \sum_{j=0}^{\infty} (e^{-\mu\tau})^j - 1 = \frac{1}{1 - e^{-\mu\tau}} - 1 = \frac{1}{e^{\mu\tau} - 1}.$$

(ii) The difference between parts (i) and (ii) is that in the second case, the system is observed at some arbitrary time point  $t$ . Suppose that  $t \in [(n-1)\tau, n\tau)$ , i.e.,  $t$  is the time point of the  $n$ -th arrival or a point between this arrival and the next one, for some arbitrary  $n$ . Hence,  $t = (n-1)\tau + T$ , where  $T$  is a r.v. uniformly distributed in  $[0, \tau)$ . Let  $X(t)$  denote the # of jobs in service at time  $t$ . Then, working as in part (i) above, we can see that:

$$\begin{aligned} E[X] &= E\left[\lim_{t \rightarrow \infty} X(t)\right] = E_T\left[E\left[\lim_{t \rightarrow \infty} X(t) | T\right]\right] = \\ &= E_T\left[E\left[\lim_{j \rightarrow \infty} \sum_{i=0}^{n-1} I_{\{\text{arrival } i \text{ still in service at time } t\}} | T\right]\right] = \\ &= E_T\left[\sum_{j=0}^{\infty} e^{-\mu(\tau+T)}\right], \text{ where again we have interchanged the} \end{aligned}$$

internal expectation and the limit thanks to the monotone convergence theorem. Proceeding with the above calculation, we have:

$$\begin{aligned} E[X] &= E_T\left[\sum_{j=0}^{\infty} e^{-\mu T} (e^{-\mu\tau})^j\right] = E_T\left[e^{-\mu T} \frac{1}{1-e^{-\mu\tau}}\right] = \\ &= \frac{1}{1-e^{-\mu\tau}} \int_0^{\tau} e^{-\mu y} \frac{1}{\tau} dy = \frac{1}{1-e^{-\mu\tau}} \frac{1}{\tau} \left(-\frac{1}{\mu}\right) [e^{-\mu\tau} - 1] = \\ &= \frac{1}{\tau\mu} = \lambda/\mu, \text{ where we set } \lambda = 1/\tau, \text{ i.e., the arrival rate.} \end{aligned}$$

Notice that the obtained result admits the standard interpretation of the traffic intensity  $\rho = \lambda/\mu$ , i.e., the expected number of busy servers (which in this case coincides with the expected number of jobs in the system) is equal to the expected workload that is received each time unit.

(iii) For this part, notice that  $(e^{\mu\tau} - 1) - \mu\tau = e^{\mu\tau} - (1 + \mu\tau) =$  (3)  
 $= (1 + \mu\tau + \sum_{i=2}^{\infty} \frac{(\mu\tau)^i}{i!}) - (1 + \mu\tau) = \sum_{i=2}^{\infty} \frac{(\mu\tau)^i}{i!} > 0.$

Hence,  $E[A] = \frac{1}{e^{\mu\tau} - 1} < \frac{1}{\mu\tau} = E[X].$

This last result further implies that the two distributions that define the above means are different. Hence, the considered system does not present any PASTA-type of effect!

(iv) To answer this part, first notice that by the problem statement, arrivals occur one at a time. Also, the exponential nature of the service times, combined with the independence of the servers, imply that the departure process is (non-homogeneous) Poisson, and therefore, the probability of having more than one departure at the same time is zero. But then, from the relevant theorem discussed in class, we can infer that the sought expectation is equal to  $E[A]$ .