Section 3.7. The notion of reversibility was used in Markov chain analysis by Kolmogorov [Kol36], and was explored in depth in [Kel79] and [Wal88].

Section 3.8. There is an extensive literature on product form solutions of queueing networks following Jackson's original paper [Jac57]. The survey [DiK85] lists 314 references. There are also several books on the subject: [Kel79], [BrB80], [GeP87], [Wal88], and [CoG89]. The heuristic explanation of Jackson's theorem is due to [Wal83].

PROBLEMS

3.1 Customers arrive at a fast-food restaurant at a rate of five per minute and wait to receive their order for an average of 5 minutes. Customers eat in the restaurant with probability 0.5 and carry out their order without eating with probability 0.5. A meal requires an average of 20 minutes. What is the average number of customers in the restaurant? (Answer: 75.)

3.2 Two communication nodes 1 and 2 send files to another node 3. Files from 1 and 2 require on the average R_1 and R_2 time units for transmission, respectively. Node 3 processes a file of node i ($i = 1, 2$) in an average of P_i time units and then requests another file from either node 1 or node 2 (the rule of choice is left unspecified). If λ_i is the throughput of node i in files sent per unit time, what is the region of all feasible throughput pairs (λ_1, λ_2) for this system?

3.3 A machine shop consists of N machines that occasionally fail and get repaired by one of the shop's m repairpersons. A machine will fail after an average of R time units following its previous repair and requires an average of P time units to get repaired. Obtain upper and lower bounds (functions of R, N, P, and m) on the number of machine failures per unit time and on the average time between repairs of the same machine.

3.4 The average time T a car spends in a certain traffic system is related to the average number of cars N in the system by a relation of the form $T = \alpha + \beta N^2$, where $\alpha > 0$, $\beta > 0$ are given scalars.

(a) What is the maximal car arrival rate λ^* that the system can sustain?

(b) When the car arrival rate is less than λ^*, what is the average time a car spends in the system assuming that the system reaches a statistical steady state? Is there a unique answer? Try to argue against the validity of the statistical steady-state assumption.

3.5 An absent-minded professor schedules two student appointments for the same time. The appointment durations are independent and exponentially distributed with mean 30 minutes. The first student arrives on time, but the second student arrives 5 minutes late. What is the expected time between the arrival of the first student and the departure of the second student? (Answer: 60.394 minutes.)

3.6 A person enters a bank and finds all of the four clerks busy serving customers. There are no other customers in the bank, so the person will start service as soon as one of the customers in service leaves. Customers have independent, identical, exponential distribution of service time.

(a) What is the probability that the person will be the last to leave the bank assuming that no other customers arrive?

(b) If the average service time is 1 minute, what is the average time the person will spend in the bank?

(c) Will the answer in part (a) change if there are some additional customers waiting in a common queue and customers begin service in the order of their arrival?

3.7 A communication line is divided in two identical channels each of which will serve a packet traffic stream where all packets have equal transmission time T and equal interarrival time $R > T$. Consider, alternatively, statistical multiplexing of the two traffic streams by combining the two channels into a single channel with transmission time $T/2$ for each packet. Show that the average system time of a packet will be decreased from T to something between $T/2$ and $3T/4$, while the variance of waiting time in queue will be increased from 0 to as much as $T^2/16$.

3.8 Consider a packet stream whereby packets arrive according to a Poisson process with rate 10 packets/sec. If the interarrival time between any two packets is less than the transmission time of the first to arrive, the two packets are said to collide. (This notion will be made more meaningful in Chapter 4 when we discuss multiaccess schemes.) Find the probabilities that a packet does not collide with either its predecessor or its successor, and that a packet does not collide with another packet assuming:

(a) All packets have a transmission time of 20 msec. (Answer: Both probabilities are equal to 0.67.)

(b) Packets have independent, exponentially distributed transmission times with mean 20 msec. (This part requires the $M/M/\infty$ results.) (Answer: The probability of no collision with predecessor or successor is 0.694. The probability of no collision is 0.682.)

3.9 A communication line capable of transmitting at a rate of 50 Kbits/sec will be used to accommodate 10 sessions each generating Poisson traffic at a rate 150 packets/min. Packet lengths are exponentially distributed with mean 1000 bits.

(a) For each session, find the average number of packets in queue, the average number in the system, and the average delay per packet when the line is allocated to the sessions by using:

(1) 10 equal-capacity time-division multiplexed channels. (Answer: $N_Q = 5$, $N = 10$, $T = 0.4$ sec.)

(2) Statistical multiplexing. (Answer: $N_Q = 0.5$, $N = 1$, $T = 0.04$ sec.)

(b) Repeat part (a) for the case where five of the sessions transmit at a rate of 250 packets/min while the other five transmit at a rate of 50 packets/min. (Answer: $N_Q = 21$, $N = 26$, $T = 1.038$ sec.)

3.10 This problem deals with some of the basic properties of the Poisson process.

(a) Derive Eqs. (3.11) to (3.14).

(b) Show that if the arrivals in two disjoint time intervals are independent and Poisson distributed with parameters $\lambda\tau_1$, $\lambda\tau_2$, then the number of arrivals in the union of the intervals is Poisson distributed with parameter $\lambda(\tau_1 + \tau_2)$. (This shows in particular that the Poisson distribution of the number of arrivals in any interval [cf. Eq. (3.10)] is consistent with the independence requirement in the definition of the Poisson process.) *Hint*: Verify the correctness of the following calculation, where N_1 and N_2 are the number of arrivals in the two disjoint intervals:

$$P\{N_1 + N_2 = n\} = \sum_{k=0}^{n} P\{N_1 = k\}P\{N_2 = n - k\}$$
$$= e^{-\lambda(\tau_1+\tau_2)} \sum_{k=0}^{n} \frac{(\lambda\tau_1)^k(\lambda\tau_2)^{n-k}}{k!(n-k)!}$$
$$= e^{-\lambda(\tau_1+\tau_2)} \frac{(\lambda\tau_1 + \lambda\tau_2)^n}{n!}$$

(c) Show that if k independent Poisson processes $A_1, \dots, A_k$ are combined into a single process $A = A_1 + A_2 + \cdots + A_k$, then A is Poisson with rate λ equal to the sum of the rates $\lambda_1, \dots \lambda_k$ of $A_1, \dots, A_k$. Show also that the probability that the first arrival of the combined process comes from A_1 is λ_1/λ independently of the time of arrival. *Hint*: For $k = 2$ write

$$P\{A_1(t+\tau) + A_2(t+\tau) - A_1(t) - A_2(t) = n\}$$
$$= \sum_{m=0}^{n} P\{A_1(t+\tau) - A_1(t) = m\}P\{A_2(t+\tau) - A_2(t) = n - m\}$$

and continue as in the hint for part (b). Also write for any t

$$P\{1 \text{ arrival from } A_1 \text{ prior to } t \mid 1 \text{ occurred}\}$$
$$= \frac{P\{1 \text{ arrival from } A_1 \text{ prior to } t, 0 \text{ from } A_2\}}{P\{1 \text{ occurred}\}}$$
$$= \frac{\lambda_1 t e^{-\lambda_1 t} e^{-\lambda_2 t}}{\lambda t e^{-\lambda t}} = \frac{\lambda_1}{\lambda}$$

(d) Suppose we know that in an interval $[t_1, t_2]$ only one arrival of a Poisson process has occurred. Show that, conditional on this knowledge, the time of this arrival is uniformly distributed in $[t_1, t_2]$. *Hint*: Verify that if t is the time of arrival, we have for all $s \in [t_1, t_2]$,

$$P\{t < s \mid 1 \text{ arrival occurred in } [t_1, t_2]\}$$
$$= \frac{P\{1 \text{ arrival occurred in } [t_1, s), 0 \text{ arrivals occurred in } [s, t_2]\}}{P\{1 \text{ arrival occurred}\}}$$
$$= \frac{s - t_1}{t_2 - t_1}$$

3.11 Packets arrive at a transmission facility according to a Poisson process with rate λ. Each packet is independently routed with probability p to one of two transmission lines and with probability $(1 - p)$ to the other.

(a) Show that the arrival processes at the two transmission lines are Poisson with rates λp and $\lambda(1 - p)$, respectively. Furthermore, the two processes are independent. *Hint*: Let $N_1(t)$ and $N_2(t)$ be the number of arrivals in $[0, t]$ in lines 1 and 2, respectively. Verify the correctness of the following calculation:

$$
\begin{aligned}
P\{N_1(t) = n, N_2(t) = m\} & \\
&= P\{N_1(t) = n, N_2(t) = m \mid N(t) = n + m\} \frac{e^{-\lambda t}(\lambda t)^{n+m}}{(n+m)!} \\
&= \binom{n+m}{n} p^n (1-p)^m \frac{e^{-\lambda t}(\lambda t)^{n+m}}{(n+m)!} \\
&= \frac{e^{-\lambda t p}(\lambda t p)^n}{n!} \frac{e^{-\lambda t(1-p)}(\lambda t(1-p))^m}{m!}
\end{aligned}
$$

$$
P\{N_1(t) = n\} = \sum_{m=0}^{\infty} P\{N_1(t) = n, N_2(t) = m\} = \frac{e^{-\lambda t p}(\lambda t p)^n}{n!}
$$

(b) Use the result of part (a) to show that the probability distribution of the customer delay in a (first-come first-serve) $M/M/1$ queue with arrival rate λ and service rate μ is exponential, that is, in steady-state we have

$$
P\{T_i \geq \tau\} = e^{-(\mu-\lambda)\tau}
$$

where T_i is the delay of the i^{th} customer. *Hint*: Consider a Poisson process A with arrival rate μ, which is split into two processes, A_1 and A_2, by randomization according to a probability $\rho = \lambda/\mu$; that is, each arrival of A is an arrival of A_1 with probability ρ and an arrival of A_2 with probability $(1-\rho)$, independently of other arrivals. Show that the interarrival times of A_2 have the same distribution as T_i.

3.12 Let τ_1 and τ_2 be two exponentially distributed, independent random variables with means $1/\lambda_1$ and $1/\lambda_2$. Show that the random variable $\min\{\tau_1, \tau_2\}$ is exponentially distributed with mean $1/(\lambda_1 + \lambda_2)$ and that $P\{\tau_1 < \tau_2\} = \lambda_1/(\lambda_1 + \lambda_2)$. Use these facts to show that the $M/M/1$ queue can be described by a continuous-time Markov chain with transition rates $q_{n(n+1)} = \lambda$, $q_{(n+1)n} = \mu$, $n = 0, 1, \ldots$. (See Appendix A for material on continuous-time Markov chains.)

3.13 Persons arrive at a taxi stand with room for W taxis according to a Poisson process with rate λ. A person boards a taxi upon arrival if one is available and otherwise waits in a line. Taxis arrive at the stand according to a Poisson process with rate μ. An arriving taxi that finds the stand full departs immediately; otherwise, it picks up a customer if at least one is waiting, or else joins the queue of waiting taxis.

(a) Use an $M/M/1$ queue formulation to obtain the steady-state distribution of the person's queue. What is the steady-state probability distribution of the taxi queue size when $W = 5$ and λ and μ are equal to 1 and 2 per minute, respectively? (Answer: Let p_i = Probability of i taxis waiting. Then $p_0 = 1/32$, $p_1 = 1/32$, $p_2 = 1/16$, $p_3 = 1/8$, $p_4 = 1/4$, $p_5 = 1/2$.)

(b) In the leaky bucket flow control scheme to be discussed in Chapter 6, packets arrive at a network entry point and must wait in a queue to obtain a permit before entering the network. Assume that permits are generated by a Poisson process with given rate and can be stored up to a given maximum number; permits generated while the maximum number of permits is available are discarded. Assume also that packets arrive according to a Poisson process with given rate. Show how to obtain the occupancy distribution of the queue of packets waiting for permits. *Hint*: This is the same system as the one of part (a).

(c) Consider the flow control system of part (b) with the difference that permits are not generated according to a Poisson process but are instead generated periodically at a given rate. (This is a more realistic assumption.) Formulate the problem of finding the type of occupancy distribution of the packet queue as an $M/D/1$ problem.

3.14 A communication node A receives Poisson packet traffic from two other nodes, 1 and 2, at rates λ_1 and λ_2, respectively, and transmits it, on a first-come first-serve basis, using a link with capacity C bits/sec. The two input streams are assumed independent and their packet lengths are identically and exponentially distributed with mean L bits. A packet from node 1 is always accepted by A. A packet from node 2 is accepted only if the number of packets in A (in queue or under transmission) is less than a given number $K > 0$; otherwise, it is assumed lost.

(a) What is the range of values of λ_1 and λ_2 for which the expected number of packets in A will stay bounded as time increases?

(b) For λ_1 and λ_2 in the range of part (a) find the steady-state probability of having n packets in $A\,(0 \leq n < \infty)$. Find the average time needed by a packet from source 1 to clear A once it enters A, and the average number of packets in A from source 1. Repeat for packets from source 2.

3.15 Consider a system that is identical to $M/M/1$ except that when the system empties out, service does not begin again until k customers are present in the system (k is given). Once service begins it proceeds normally until the system becomes empty again. Find the steady-state probabilities of the number in the system, the average number in the system, and the average delay per customer. [Answer: The average number in the system is $N = \rho/(1-\rho) + (k-1)/2$.]

3.16 *M/M/1-Like System with State-Dependent Arrival and Service Rate.* Consider a system which is the same as $M/M/1$ except that the rate λ_n and service rate μ_n when there are n customers in the system depend on n. Show that

$$p_{n+1} = (\rho_0 \cdots \rho_n)p_0$$

where $\rho_k = \lambda_k/\mu_{k+1}$ and

$$p_0 = \left[1 + \sum_{k=0}^{\infty}(\rho_0 \cdots \rho_k)\right]^{-1}$$

3.17 *Discrete-Time Version of the M/M/1 System.* Consider a queueing system where interarrival and service times are integer valued, so customer arrivals and departures occur at integer times. Let λ be the probability that an arrival occurs at any time k, and assume that at most one arrival can occur. Also let μ be the probability that a customer who was in service at time k will complete service at time $k+1$. Find the occupancy distribution p_n in terms of λ and μ.

3.18 Empty taxis pass by a street corner at a Poisson rate of 2 per minute and pick up a passenger if one is waiting there. Passengers arrive at the street corner at a Poisson rate of 1 per minute and wait for a taxi only if there are fewer than four persons waiting; otherwise, they leave and never return. Find the average waiting time of a passenger who joins the queue. (Answer: 13/15 min.)

3.19 A telephone company establishes a direct connection between two cities expecting Poisson traffic with rate 30 calls/min. The durations of calls are independent and exponentially distributed with mean 3 min. Interarrival times are independent of call durations. How many circuits should the company provide to ensure that an attempted call is blocked (because all circuits are busy) with probability less than 0.01? It is assumed that blocked calls are lost (*i.e.*, a blocked call is not attempted again).

3.20 A mail-order company receives calls at a Poisson rate of one per 2 min and the duration of the calls is exponentially distributed with mean 3 min. A caller who finds all telephone operators busy patiently waits until one becomes available. Write a computer program to determine how many operators the company should use so that the average waiting time of a customer is half a minute or less?

3.21 Consider the $M/M/1/m$ system which is the same as $M/M/1$ except that there can be no more than m customers in the system and customers arriving when the system is full are lost. Show that the steady-state occupancy probabilities are given by

$$p_n = \frac{\rho^n(1-\rho)}{1-\rho^{m+1}}, \qquad 0 \le n \le m$$

3.22 An athletic facility has five tennis courts. Players arrive at the courts at a Poisson rate of one pair per 10 min and use a court for an exponentially distributed time with mean 40 min.

(a) Suppose that a pair of players arrives and finds all courts busy and k other pairs waiting in queue. How long will they have to wait to get a court on the average?

(b) What is the average waiting time in queue for players who find all courts busy on arrival?

3.23 Consider an $M/M/\infty$ queue with servers numbered $1, 2, \ldots$ There is an additional restriction that upon arrival a customer will choose the lowest-numbered server that is idle at the time. Find the fraction of time that each server is busy. Will the answer change if the number of servers is finite? *Hint*: Argue that in steady-state the probability that all of the first m servers are busy is given by the Erlang B formula of the $M/M/m/m$ system. Find the total arrival rate to servers $(m+1)$ and higher, and from this, the arrival rate to each server.

3.24 $M/M/1$ *Shared Service System.* Consider a system which is the same as $M/M/1$ except that whenever there are n customers in the system they are all served simultaneously at an equal rate $1/n$ per unit time. Argue that the steady-state occupancy distribution is the same as for the $M/M/1$ system. *Note*: It can be shown that the steady-state occupancy distribution is the same as for $M/M/1$ even if the service time distribution is not exponential (*i.e.*, for an $M/G/1$ type of system) ([Ros83], p. 171).

3.25 *Blocking Probability for Single-Cell Radio Systems ([BaA81] and [BaA82]).* A cellular radiotelephone system serves a given geographical area with m radiotelephone channels connected to a single switching center. There are two types of calls: radio-to-radio calls, which occur with a Poisson rate λ_1 and require two radiochannels per call, and radio-to-nonradio calls, which occur with a Poisson rate λ_2 and require one radiochannel per call. The duration of all calls is exponentially distributed with mean $1/\mu$. Calls that cannot be accommodated by the system are blocked. Give formulas for the blocking probability of the two types of calls.

3.26 A facility of m identical machines is sharing a single repairperson. The time to repair a failed machine is exponentially distributed with mean $1/\lambda$. A machine, once operational, fails after a time that is exponentially distributed with mean $1/\mu$. All failure and repair times are independent. What is the steady-state proportion of time where there is no operational machine?

3.27 $M/M/2$ *System with Heterogeneous Servers.* Derive the stationary distribution of an $M/M/2$ system where the two servers have different service rates. A customer that arrives when the system is empty is routed to the faster server.

3.28 In Example 3.11, verify the formula $\sigma_f = (\lambda/\mu)^{1/2} s_\gamma$. *Hint*: Write

$$E\{f^2\} = E\left\{\left(\sum_{i=1}^{n} \gamma_i\right)^2\right\} = E\left\{E\left\{\left(\sum_{i=1}^{n} \gamma_i\right)^2 \mid n\right\}\right\},$$

and use the fact that n is Poisson distributed.

3.29 Customers arrive at a grocery store's checkout counter according to a Poisson process with rate 1 per minute. Each customer carries a number of items that is uniformly distibuted between 1 and 40. The store has two checkout counters, each capable of processing items at a rate of 15 per minute. To reduce the customer waiting time in queue, the store manager considers dedicating one of the two counters to customers with x items or less and dedicating the other counter to customers with more than x items. Write a small computer program to find the value of x that minimizes the average customer waiting time.

3.30 In the $M/G/1$ system, show that

$$P\{\text{the system is empty}\} = 1 - \lambda\overline{X}$$

$$\text{Average length of time between busy periods} = \frac{1}{\lambda}$$

$$\text{Average length of busy period} = \frac{\overline{X}}{1 - \lambda\overline{X}}$$

$$\text{Average number of customers served in a busy period} = \frac{1}{1 - \lambda\overline{X}}$$

3.31 Consider the following argument in the $M/G/1$ system: When a customer arrives, the probability that another customer is being served is $\lambda\overline{X}$. Since the served customer has mean service time $\overline{X}$, the average time to complete the service is $\overline{X}/2$. Therefore, the mean residual service time is $\lambda\overline{X}^2/2$. What is wrong with this argument?

3.32 *$M/G/1$ System with Arbitrary Order of Service.* Consider the $M/G/1$ system with the difference that customers are not served in the order they arrive. Instead, upon completion of a customer's service, one of the waiting customers in queue is chosen according to some rule, and is served next. Show that the P-K formula for the average waiting time in queue W remains valid provided that the relative order of arrival of the customer chosen is independent of the service times of the customers waiting in queue. *Hint*: Argue that the independence hypothesis above implies that at any time t, the number $N_Q(t)$ of customers waiting in queue is independent of the service times of these customers. Show that this in turn implies that $U = R + \rho W$, where R is the mean residual time and U is the average steady-state unfinished work in the system (total remaining service time of the customers in the system). Argue that U and R are independent of the order of customer service.

3.33 Show that Eq. (3.59) for the average delay of time-division multiplexing on a slot basis can be obtained as a special case of the results for the limited service reservation system. *Hint*: Consider the gated system with zero packet length.

3.34 Consider the limited service reservation system. Show that for both the gated and the partially gated versions:

(a) The steady-state probability of arrival of a packet during a reservation interval is $1 - \rho$.

(b) The steady-state probability of a reservation interval being followed by an empty data interval is $(1 - \rho - \lambda\overline{V})/(1 - \rho)$. *Hint:* If p is the required probability, argue that the ratio of the times used for data intervals and for reservation intervals is $(1 - p)\overline{X}/\overline{V}$.

3.35 *Limited Service Reservation System with Shared Reservation and Data Intervals.* Consider the gated version of the limited service reservation system with the difference that the m users share reservation and data intervals, (*i.e.*, all users make reservations in the same interval and transmit at most one packet each in the subsequent data interval). Show that

$$W = \frac{\lambda\overline{X^2}}{2(1-\rho-\lambda\overline{V}/m)} + \frac{(1-\rho)\overline{V^2}}{2(1-\rho-\lambda\overline{V}/m)\overline{V}} + \frac{(1-\rho\alpha-\lambda\overline{V}/m)\overline{V}}{1-\rho-\lambda\overline{V}/m}$$

where $\overline{V}$ and $\overline{V^2}$ are the first two moments of the reservation interval, and α satisfies

$$\frac{\overline{K}+(\hat{K}-1)(2\overline{K}-\hat{K})}{2m\overline{K}} - \frac{1}{2m} \le \alpha \le \frac{1}{2} - \frac{1}{2m}$$

where

$$\overline{K} = \frac{\lambda\overline{V}}{1-\rho}$$

is the average number of packets per data interval, and $\hat{K}$ is the smallest integer which is larger than $\overline{K}$. Verify that the formula for W becomes exact as $\rho \to 0$ (light load) and as $\rho \to 1-\lambda\overline{V}/m$ (heavy load). *Hint:* Verify that

$$W = R + \lambda W + \left(1 + \frac{\lambda W}{m} - S\right)\overline{V}$$

where $S = \lim_{i\to\infty} E\{S_i\}$ and S_i is the number (0 or 1) of packets of the owner of packet i that will start transmission between the time of arrival of packet i and the end of the cycle in which packet i arrives. Try to obtain bounds for S by considering separately the cases where packet i arrives in a reservation and in a data interval.

3.36 Repeat part (a) of Problem 3.9 for the case where packet lengths are not exponentially distributed, but 10% of the packets are 100 bits long and the rest are 1500 bits long. Repeat the problem for the case where the short packets are given nonpreemptive priority over the long packets. (Answer: $N_Q = 0.791$, $N = 1.47$, $T = 0.588$ sec.)

3.37 Persons arrive at a Xerox machine according to a Poisson process with rate one per minute. The number of copies to be made by each person is uniformly distributed between 1 and 10. Each copy requires 3 sec. Find the average waiting time in queue when:

(a) Each person uses the machine on a first-come first-serve basis. (Answer: $W = 3.98$.)

(b) Persons with no more than 2 copies to make are given nonpreemptive priority over other persons.

3.38 *Priority Systems with Multiple Servers.* Consider the priority systems of Section 3.5.3 assuming that there are m servers and that all priority classes have exponentially distributed service times with common mean $1/\mu$.

(a) Consider the nonpreemptive system. Show that Eq. (3.79) yields the average queueing times with the mean residual time R given by

$$R = \frac{P_Q}{m\mu}$$

where P_Q is the steady-state probability of queueing given by the Erlang C formula of Eq. (3.36). [Here $\rho_i = \lambda_i/(m\mu)$ and $\rho = \sum_{i=1}^{n} \rho_i$.]

(b) Consider the preemptive resume system. Argue that $W_{(k)}$, defined as the average time in queue averaged over the first k priority classes, is the same as for an $M/M/m$ system with arrival rate $\lambda_1 + \cdots + \lambda_k$ and mean service time $1/\mu$. Use Little's Theorem to show that the average time in queue of a k^{th} priority class customer can be obtained recursively from

$$W_1 = W_{(1)}$$

$$W_k = \frac{1}{\lambda_k}\left[W_{(k)}\sum_{i=1}^{k}\lambda_i - W_{(k-1)}\sum_{i=1}^{k-1}\lambda_i\right], \qquad k = 2, 3, \ldots, n$$

3.39 Consider the nonpreemptive priority queueing system of Section 3.5.3 for the case where the available capacity is sufficient to handle the highest-priority traffic but cannot handle the traffic of all priorities, that is,

$$\rho_1 < 1 < \rho_1 + \rho_2 + \cdots + \rho_n$$

Find the average delay per customer of each priority class. *Hint*: Determine the departure rate of the highest-priority class that will experience infinite average delay and the mean residual service time.

3.40 *Optimization of Class Ordering in a Nonpreemptive System.* Consider an n-class, nonpreemptive priority system:

(a) Show that the sum $\sum_{k=1}^{n} \rho_k W_k$ is independent of the priority order of classes, and in fact

$$\sum_{k=1}^{n} \rho_k W_k = \frac{R\rho}{1-\rho}$$

where $\rho = \rho_1 + \rho_2 + \cdots + \rho_n$. (This is known as the $M/G/1$ conservation law [Kle64].) *Hint:* Use Eq. (3.79). Alternatively, argue that $U = R + \sum_{k=1}^{n} \rho_k W_k$, where U is the average steady-state unfinished work in the system (total remaining service time of customers in the system), and U and R are independent of the priority order of the classes.

(b) Suppose that there is a cost c_k per unit time for each class k customer that waits in queue. Show that cost is minimized when classes are ordered so that

$$\frac{\overline{X_1}}{c_1} \le \frac{\overline{X_2}}{c_2} \le \cdots \le \frac{\overline{X_n}}{c_n}$$

Hint: Express the cost as $\sum_{k=1}^{n}(c_k/\overline{X}_k)(\rho_k W_k)$ and use part (a). Also use the fact that interchanging the order of any two adjacent classes leaves the waiting time of all other classes unchanged.

3.41 *Little's Theorem for Arbitrary Order of Service; Analytical Proof [Sti74].* Consider the analysis of Little's Theorem in Section 3.2 and the notation introduced there. We allow the possibility that the initial number in the system is positive [*i.e.*, $N(0) > 0$]. Assume that the time-average arrival and departure rates exist and are equal:

$$\lambda = \lim_{t\to\infty}\frac{\alpha(t)}{t} = \lim_{t\to\infty}\frac{\beta(t)}{t}$$

and that the following limit defining the time-average system time exists:

$$T = \lim_{k\to\infty}\frac{1}{N(0)+\alpha(t)}\left(\sum_{i\in D(t)} T_i + \sum_{i\in \overline{D}(t)} (t - t_i)\right)$$

where $D(t)$ is the set of customers departed by time t and $\overline{D}(t)$ is the set of customers that are in the system at time t. (For all customers that are initially in the system, the time T_i is counted starting at time 0.) Show that regardless of the order in which customers are served, Little's Theorem ($N = \lambda T$) holds with

$$N = \lim_{t \to \infty} \frac{1}{t} \int_0^t N(\tau)\, d\tau$$

Show also that

$$T = \lim_{k \to \infty} \frac{1}{k} \sum_{i=1}^{k} T_i$$

Hint: Take $t \to \infty$ below:

$$\frac{1}{t} \sum_{i \in D(t)} T_i \leq \frac{1}{t} \int_0^t N(\tau)\, d\tau \leq \frac{1}{t} \sum_{i \in D(t) \cup \overline{D}(t)} T_i$$

3.42 *A Generalization of Little's Theorem.* Consider an arrival–departure system with arrival rate λ, where entering customers are forced to pay money to the system according to some rule.

(a) Argue that the following identity holds:

Average rate at which the system earns $= \lambda \cdot$ (Average amount a customer pays)

(b) Show that Little's Theorem is a special case.

(c) Consider the $M/G/1$ system and the following cost rule: Each customer pays at a rate of y per unit time when its remaining service time is y, whether in queue or in service. Show that the formula in (a) can be written as

$$W = \lambda \left(\overline{X} W + \frac{\overline{X^2}}{2} \right)$$

which is the Pollaczek–Khinchin formula.

3.43 *$M/G/1$ Queue with Random-Sized Batch Arrivals.* Consider the $M/G/1$ system with the difference that customers are arriving in batches according to a Poisson process with rate λ. Each batch has n customers, where n has a given distribution and is independent of customer service times. Adapt the proof of Section 3.5 to show that the waiting time in queue is given by

$$W = \frac{\lambda \overline{n}\, \overline{X^2}}{2(1-\rho)} + \frac{\overline{X}(\overline{n^2} - \overline{n})}{2\overline{n}(1-\rho)}$$

Hint: Use the equation $W = R + \rho W + W_B$, where W_B is the average waiting time of a customer for other customers who arrived in the same batch.

3.44 *$M/G/1$ Queue with Overhead for Each Busy Period.* Consider the $M/G/1$ queue with the difference that the service of the first customer in each busy period requires an increment Δ over the ordinary service time of the customer. We assume that Δ has a given distribution and is independent of all other random variables in the model. Let $\rho = \lambda \overline{X}$ be the utilization factor. Show that

(a) $p_0 = P\{\text{the system is empty}\} = (1-\rho)/(1+\lambda\overline{\Delta})$.

(b) Average length of busy period $= (\overline{X} + \overline{\Delta})/(1-\rho)$.

(c) The average waiting time in queue is

$$W = \frac{\lambda\overline{X^2}}{2(1-\rho)} + \frac{\lambda[\overline{(X+\Delta)^2} - \overline{X^2}]}{2(1+\lambda\overline{\Delta})}$$

(d) Parts (a), (b), and (c) also hold in the case where Δ may depend on the interarrival and service time of the first customer in the corresponding busy period.

3.45 Consider a system that is identical to $M/G/1$ except that when the system empties out, service does not begin again until k customers are present in the system (k is given). Once service begins, it proceeds normally until the system becomes empty again. Show that:

(a) In steady-state:

$$P\{\text{system empty}\} = \frac{1-\rho}{k}$$

$$P\{\text{system nonempty and waiting}\} = \frac{(k-1)(1-\rho)}{k}$$

$$P\{\text{system nonempty and serving}\} = \rho$$

(b) The average length of a busy period is

$$\frac{\rho + k - 1}{\lambda(1-\rho)}$$

Verify that this average length is equal to the average time between arrival and start of service of the first customer in a busy period, plus k times the average length of a busy period for the corresponding $M/G/1$ system ($k = 1$).

(c) Suppose that we divide a busy period into a busy/waiting portion and a busy/serving portion. Show that the average number in the system during a busy/waiting portion is $k/2$ and the average number in the system during a busy/serving portion is

$$\frac{N_{M/G/1}}{\rho} + \frac{k-1}{2}$$

where $N_{M/G/1}$ is the average number in the system for the corresponding $M/G/1$ system ($k = 1$). *Hint*: Relate a busy/serving portion of a busy period with k independent busy periods of the corresponding $M/G/1$ system where $k = 1$.

(d) The average number in the system is

$$N_{M/G/1} + \frac{k-1}{2}$$

3.46 *Single-Vacation $M/G/1$ System.* Consider the $M/G/1$ system with the difference that each busy period is followed by a single vacation interval. Once this vacation is over, an arriving customer to an empty system starts service immediately. Assume that vacation intervals are independent, identically distributed, and independent of the customer interarrival and service times. Prove that the average waiting time in queue is

$$W = \frac{\lambda\overline{X^2}}{2(1-\rho)} + \frac{\overline{V^2}}{2I}$$

where I is the average length of an idle period, and show how to calculate I.

3.47 *The $M/G/\infty$ System.* Consider a queueing system with Poisson arrivals at rate λ. There are an infinite number of servers, so that each arrival starts service at an idle server immediately on arrival. Each server has a general service time distribution and $F_X(x) = P\{X \le x\}$

denotes the probability that a service starting at any given time τ is completed by time $\tau + x$ [$F_X(x) = 0$ for $x \leq 0$]. The servers have independent and identical service time distributions.

(a) For x and δ ($0 < \delta < x$) very small, find the probability that there was an arrival in the interval $[\tau - x,\ \tau - x + \delta]$ *and* that this arrival is still being served at time τ.

(b) Show that the mean service time for any arrival is given by

$$\overline{X} = \int_0^\infty [1 - F_X(x)]\, dx$$

Hint: Use a graphical argument.

(c) Use parts (a) and (b) to verify that the number in the system is Poisson distributed with mean $\lambda\overline{X}$.

3.48 *An Improved Bound for the G/G/1 Queue.*

(a) Let r be a nonnegative random variable and let x be a nonnegative scalar. Show that

$$\frac{\overline{\left(\max\{0, r - x\}\right)^2}}{\left(\overline{\max\{0, r - x\}}\right)^2} \geq \frac{\overline{r^2}}{(\overline{r})^2}$$

where overbar denotes expected value. *Hint*: Prove that the left-hand expression is monotonically nondecreasing as a function of x.

(b) Using the notation of Section 3.5.4, show that

$$\sigma_I^2 \geq (1 - \rho)^2 \sigma_a^2$$

and that

$$W \leq \frac{\lambda(\sigma_a^2 + \sigma_b^2)}{2(1 - \rho)} - \frac{\lambda(1 - \rho)\sigma_a^2}{2}$$

Hint: Use part (a) with r being the customer interarrival time and x equal to the time in the system [cf. Eq. (3.93)].

3.49 *Last-Come First-Serve M/G/1 System.* Consider an $M/G/1$ system with the difference that upon arrival at the queue, a customer goes immediately into service, replacing the customer who is in service at the time (if any) on a preemptive-resume basis. When a customer completes service, the customer most recently preempted resumes service. Show that:

(a) The expected length of a busy period, denoted $E\{B\}$, is the same as in the ordinary $M/G/1$ queue.

(b) Show that the expected time in the system of a customer is equal to $E\{B\}$. *Hint*: Argue that a customer who starts a busy period stays in the system for the entire duration of the busy period.

(c) Let C be the average time in the system of a customer requiring one unit of service time. Argue that the average time in the system of a customer requiring X units of service time is XC. *Hint*: Argue that a customer requiring two units of service time is "equivalent" to two customers with one unit service time each, and with the second customer arriving at the time that the first departs.

(d) Show that

$$C = \frac{E\{B\}}{E\{X\}} = \frac{1}{1 - \rho}$$

3.50 *Truncation of Queues.* This problem illustrates one way to use simple queues to obtain results about more complicated queues.

(a) Consider a continuous-time Markov chain with state space S, stationary distribution $\{p_j\}$, and transition rates q_{ij}. Suppose that we have a truncated version of this chain, that is, a new chain with space $\overline{S}$, which is a subset of S and has the same transition rates q_{ij} between states i and j of $\overline{S}$. Assume that for all $j \in \overline{S}$, we have

$$p_j \sum_{i \notin \overline{S}} q_{ji} = \sum_{i \notin \overline{S}} p_i q_{ij}$$

Show that if the truncated chain is irreducible, then its stationary distribution $\{\overline{p}_j\}$ satisfies $\overline{p}_j = p_j / \sum_{i \in \overline{S}} p_i$ for all $j \in \overline{S}$. (Note that $\overline{p}_j$ is the conditional probability for the state of the original chain to be j conditioned on the fact that it lies within $\overline{S}$.)

(b) Show that the condition of part (a) on the stationary distribution $\{p_j\}$ and the transition rates $\{q_{ij}\}$ is satisfied if the original chain is time reversible, and that in this case, the truncated chain is also time reversible.

(c) Consider two queues with independent Poisson arrivals and independent exponentially distributed service times. The arrival and service rates are denoted λ_i, μ_i, for $i = 1, 2$, respectively. The two queues share a waiting room with finite capacity B (including customers in service). Arriving customers that find the waiting room full are lost. Use part (b) to show that the system is reversible and that for $m + n \leq B$, the steady-state probabilities are

$$P\{m \text{ in queue } 1,\ n \text{ in queue } 2\} = \frac{\rho_1^m \rho_2^n}{G}$$

where $\rho_i = \lambda_i / \mu_i$, $i = 1, 2$, and G is a normalizing constant.

3.51 *Decomposition/Aggregation of Reversible Chains.* Consider a time reversible continuous-time Markov chain in equilibrium, with state space S, transition rates q_{ij}, and stationary probabilities p_j. Let $S = \cup_{k=1}^{K} S_k$ be a partition of S in mutually disjoint sets, and denote for all k and $j \in S_k$:

u_k = Probability of the state being in S_k (*i.e.*, $u_k = \sum_{j \in S_k} p_j$)

π_j = Probability of the state being equal to j conditioned on the fact that the state belongs to S_k (*i.e.*, $\pi_j = P\{X_n = j \mid X_n \in S_k\} = p_j / u_k$)

Assume that all states in S_k communicate with all other states in S_k.

(a) Show that $\{\pi_j \mid j \in S_k\}$ is the stationary distribution of the truncated chain with state space S_k (cf. Problem 3.50).

(b) Show that $\{u_k \mid k = 1, \ldots, K\}$ is the stationary distribution of the so-called *aggregate chain*, which is the Markov chain with states $k = 1, \ldots, K$ and transition rates

$$\tilde{q}_{km} = \sum_{j \in S_k,\ i \in S_m} \pi_j q_{ji}, \qquad k, m = 1, \ldots, K$$

Show also that the aggregate chain is reversible. (Note that the aggregate chain corresponds to a fictitious process; the actual process, corresponding to transitions between sets of states, need not be Markov.)

(c) Outline a divide-and-conquer solution method that first solves for the distributions of the truncated chains and then solves for the distribution of the aggregate chain. Apply this method to Examples 3.12 and 3.13.

(d) Suppose that the truncated chains are reversible but the original chain is not. Show that the results of parts (a) and (b) hold except that the aggregate chain need not be reversible.

3.52 *An Extension of Burke's Theorem.* Consider an $M/M/1$ system in steady state where customers are served in the order that they arrive. Show that given that a customer departs at time t, the arrival time of that customer is independent of the departure process prior to t. *Hint*: Consider a customer arriving at time t_1 and departing at time t_2. In reversed system terms, the arrival process is independent Poisson, so the arrival process to the left of t_2 is independent of the times spent in the system of customers that arrived at or to the right of t_2.

3.53 Consider the model of two queues in tandem of Section 3.7 and assume that customers are served at each queue in the order they arrive.

(a) Show that the times (including service) spent by a customer in queue 1 and in queue 2 are mutually independent, and independent of the departure process from queue 2 prior to the customer's departure from the system. *Hint:* By Burke's Theorem, the time spent by a customer in queue 1 is independent of the sequence of arrival times at queue 2 prior to the customer's arrival at queue 2. These arrival times (together with the corresponding independent service times) determine the time the customer spends at queue 2 as well as the departure process from queue 2 prior to the customer's departure from the system.

(b) Argue by example that the times a customer spends waiting *before entering service* at the two queues are *not* independent.

3.54 Use reversibility to characterize the departure process of the $M/M/1/m$ queue.

3.55 Consider the feedback model of a CPU and I/O device of Example 3.19 with the difference that the CPU consists of m identical parallel processors. The service time of a job at each parallel processor is exponentially distributed with mean $1/\mu_1$. Derive the stationary distribution of the system.

3.56 Consider the discrete-time approximation to the $M/M/1$ queue of Fig. 3.6. Let X_n be the state of the system at time $n\Delta$ and let D_n be a random variable taking on the value 1 if a departure occurs between $n\Delta$ and $(n+1)\Delta$, and the value 0 if no departure occurs. Assume that the system is in steady-state at time $n\Delta$. Answer the following without using reversibility.

(a) Find $P\{X_n = i, D_n = j\}$ for $i \geq 0$, $j = 0, 1$.

(b) Find $P\{D_n = 1\}$.

(c) Find $P\{X_n = i, D_n = 1\}$ for $i \geq 0$.

(d) Find $P\{X_{n+1} = i, D_n = 1\}$ and show that X_{n+1} is statistically independent of D_n. *Hint*: Use part (c); also show that $P\{X_{n+1} = i\} = P\{X_{n+1} = i \mid D_n = 1\}$ for all $i \geq 0$ is sufficient to show independence.

(e) Find $P\{X_{n+k} = i, D_{n+1} = j \mid D_n\}$ and show that the pair of variables (X_{n+1}, D_{n+1}) is statistically independent of D_n.

(f) For each $k > 1$, find $P\{X_{n+k} = i, D_{n+k} = j \mid D_{n+k-1}, D_{n+k-2}, \ldots, D_n\}$ and show that the pair (X_{n+k}, D_{n+k}) is statistically independent of $(D_{n+k-1}, D_{n+k-2}, \ldots, D_n)$. *Hint*: Use induction on k.

(g) Deduce a discrete-time analog to Burke's Theorem.

3.57 Consider the network in Fig. 3.39. There are four sessions: ACE, ADE, BCEF, and BDEF sending Poisson traffic at rates 100, 200, 500, and 600 packets/min, respectively. Packet lengths are exponentially distributed with mean 1000 bits. All transmission lines have capacity 50 kbits/sec, and there is a propagation delay of 2 msec on each line. Using the Kleinrock independence approximation, find the average number of packets in the system, the average delay per packet (regardless of session), and the average delay per packet of each session.

3.58 *Jackson Networks with a Limit on the Total Number of Customers.* Consider an open Jackson network as described in the beginning of Section 3.8, with the difference that all customers who arrive when there are a total of M customers in the network are blocked from entering and are lost for the system. Derive the stationary distribution. *Hint*: Convert the system into a closed network with M customers by introducing an additional queue $K+1$ with service rate equal to $\sum_{j=1}^{K} r_j$. A customer exiting queue $i \in \{1, \ldots, K\}$ enters queue $K+1$ with probability $1 - \sum_j P_{ij}$, and a customer exiting queue $K+1$ enters queue $i \in \{1, \ldots, K\}$ with probability $r_i / \sum_{j=1}^{K} r_j$.

3.59 Justify the Arrival Theorem for closed networks by inserting a very fast $M/M/1$ queue between every pair of queues. Argue that conditioning on a customer moving from one queue to another is essentially equivalent to conditioning on a single customer being in the fast $M/M/1$ queue that lies between the two queues.

3.60 Consider a closed Jackson network where the service time at each queue is independent of the number of customers at the queue. Suppose that for a given number of customers, the utilization factor of one of the queues, say queue 1, is strictly larger than the utilization factors of the other queues. Show that as the number of customers increases, the proportion of time that a customer spends in queue 1 approaches unity.

3.61 Consider a model of a computer CPU connected to m I/O devices as shown in Fig. 3.40. Jobs enter the system according to a Poisson process with rate λ, use the CPU and with

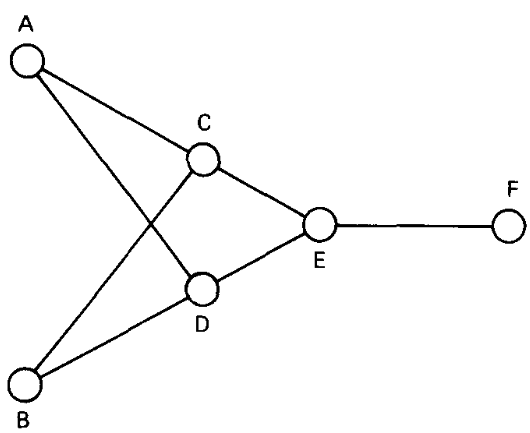

Figure 3.39 Network of transmission lines for Problem 3.57.

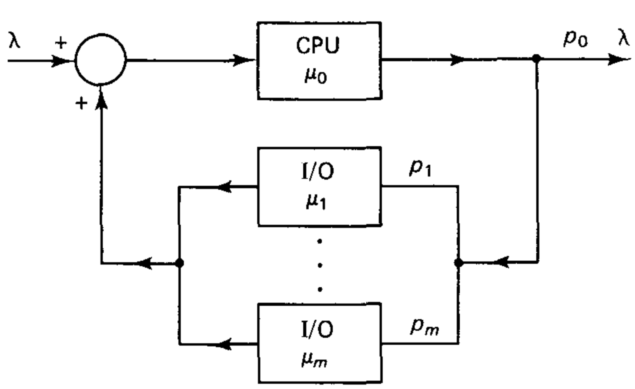

Figure 3.40 Model of a computer CPU connected to m I/O devices for Problem 3.61.

probability p_i, $i = 1, \ldots, m$, are routed to the i^{th} I/O device, while with probability p_0 they exit the system. The service time of a job at the CPU (or the i^{th} I/O device) is exponentially distributed with mean $1/\mu_0$ (or $1/\mu_i$, respectively). We assume that all job service times at all queues are independent (including the times of successive visits to the CPU and I/O devices of the same job). Find the occupancy distribution of the system and construct an "equivalent" system with $m+1$ queues in tandem that has the same occupancy distribution.

3.62 Consider a closed version of the queueing system of Problem 3.61, shown in Fig. 3.41. There are M jobs in the system at all times. A job uses the CPU and with probability p_i, $i = 1, \ldots, m$, is routed to the i^{th} I/O device. The service time of a job at the CPU (or the i^{th} I/O device) is exponentially distributed with mean $1/\mu_0$ (or $1/\mu_i$, respectively). We assume that all job service times at all queues are independent (including the times of successive visits to the CPU and I/O devices of the same job). Find the arrival rate of jobs at the CPU and the occupancy distribution of the system.

3.63 *Bounds on the Throughput of a Closed Queueing Network.* Packets enter the network of transmission lines shown in Fig. 3.42 at point A and exit at point B. A packet is first transmitted on one of the lines $L_1, \ldots, L_K$, where it requires on the average a transmission time $\overline{X}$, and is then transmitted in line L_{K+1}, where it requires on the average a transmission time $\overline{Y}$. To effect flow control, a maximum of $N \geq K$ packets are admitted into the system.

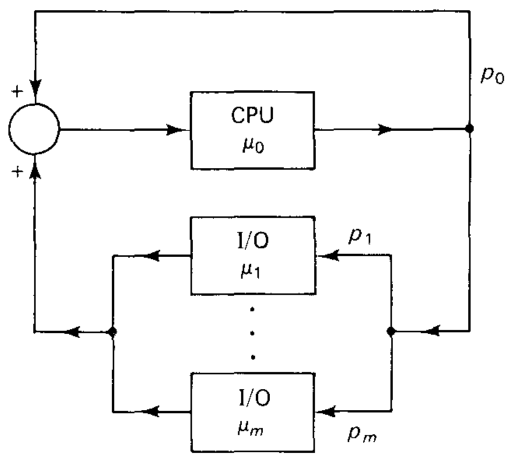

Figure 3.41 Closed queueing system for Problem 3.62.

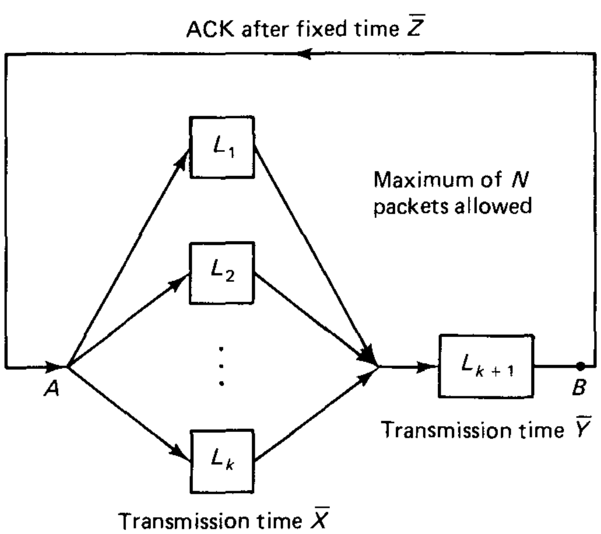

Figure 3.42 Closed queueing network for Problem 3.63.

Each time a packet exits the system at point B, an acknowledgment is sent back and reaches point A after a fixed time $\overline{Z}$. At that time, a new packet is allowed to enter the system. Use Little's Theorem to find upper and lower bounds for the system throughput under two circumstances:

(a) The method of routing a packet to one of the lines $L_1, \ldots, L_K$ is unspecified.

(b) The routing method is such that whenever one of the lines $L_1, \ldots, L_K$ is idle, there is no packet waiting at any of the other lines.

3.64 Consider the closed queueing network in Fig. 3.43. There are three customers who are doomed forever to cycle between queue 1 and queue 2. The service times at the queues are independent and exponentially distributed with mean μ_1 and μ_2. Assume that $\mu_2 < \mu_1$.

(a) The system can be represented by a four-state Markov chain. Find the transition rates of the chain.

(b) Find the steady-state probabilities of the states.

(c) Find the customer arrival rate at queue 1.

(d) Find the rate at which a customer cycles through the system.

(e) Show that the Markov chain is reversible. What does a departure from queue 1 in the forward process correspond to in the reversed process? Can the transitions of a single customer in the forward proces be associated with transitions of a single customer in the reverse process?

3.65 Consider the closed queueing network of Section 3.8.2 and assume that the service rate $\mu_j(m)$ at the j^{th} queue is independent of the number of customers m in the queue $[\mu_j(m) = \mu_j$ for all $m]$. Show that the utilization factor $U_j(M) = \lambda_j(M)/\mu_j$ of the j^{th} queue is given by

$$U_j(M) = \rho_j \frac{G(M-1)}{G(M)}$$

where $\rho_j = \overline{\lambda}_j/\mu_j$ (compare with Examples 3.21 and 3.22).

3.66 *$M/M/1$ System with Multiple Classes of Customers.* Consider an $M/M/1$-like system with first-come first-serve service and multiple classes of customers denoted $c = 1, 2, \ldots, C$. Let λ_i and μ_i be the arrival and service rate of class i.

(a) Model this system by a Markov chain and show that unless $\mu_1 = \mu_2 = \cdots = \mu_C$, its steady-state distribution does not have a product form. *Hint:* Consider a state $z = (c_1, c_2, \ldots, c_n)$ such that $\mu_{c_1} \neq \mu_{c_n}$. Write the global balance equations for state z.

(b) Suppose instead that the service discipline is last-come first-serve (as defined in Problem 3.49). Model the system by a Markov chain and show that the steady-state distribution has the product form

$$P(z) = P(c_1, c_2, \ldots, c_n) = \frac{\rho_{c_1}\rho_{c_2}\cdots\rho_{c_n}}{G}$$

where $\rho_c = \lambda_c/\mu_c$ and G is a normalizing constant.

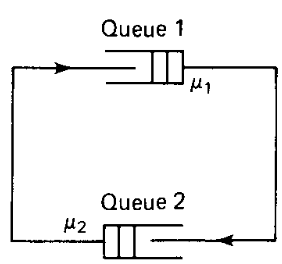

Figure 3.43 Closed queueing network for Problem 3.64.