SOLUTIONS MANUAL

Second Edition

Data Networks

DIMITRI BERTSEKAS

Massachusetts Institute of Technology

ROBERT GALLAGER

Massachusetts Institute of Technology

PRENTICE HALL. Englewood Cliffs. New Jersey 07632

A Paramount Communications Company
Englewood Cliffs, New Jersey 07632

10 9 8 7 6 5 4 3

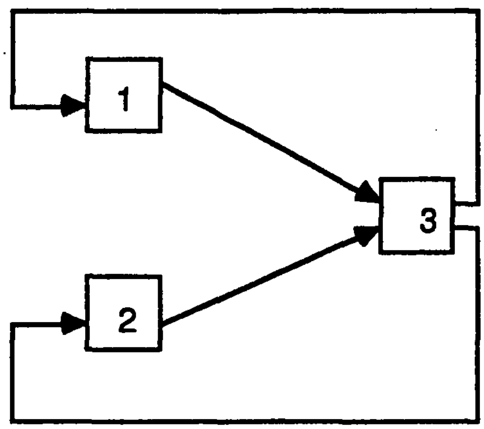
ISBN 0-13-200924-2
Printed in the United States of America

CHAPTER 3 SOLUTIONS

3.1

A customer that carries out the order (eats in the restaurant) stays for 5 mins (25 mins). Therefore the average customer time in the system is $T = 0.5*5 + 0.5*25 = 15$. By Little's Theorem the average number in the system is $N = \lambda*T = 5*15=75$.

3.2



We represent the system as shown in the figure. The number of files in the entire system is exactly one at all times. The average number in node i is $\lambda_i R_i$ and the average number in node 3 is $\lambda_1 P_1 + \lambda_2 P_2$. Therefore the throughput pairs (λ_1, λ_2) must satisfy (in addition to nonnegativity) the constraint

$$\lambda_1(R_1 + P_1) + \lambda_2(R_2 + P_2) = 1.$$

If the system were slightly different and queueing were allowed at node 3, while nodes 1 and 2 could transmit at will, a different analysis would apply. The transmission bottleneck for the files of node 1 implies that

$$\lambda_1 \leq \frac{1}{R_1}$$

Similarly for node 2 we get that

$$\lambda_2 \leq \frac{1}{R_2}$$

Node 3 can work on only one file at a time. If we look at the file receiving service at node 3 as a system and let N be the average number receiving service at node 3, we conclude from Little's theorem that

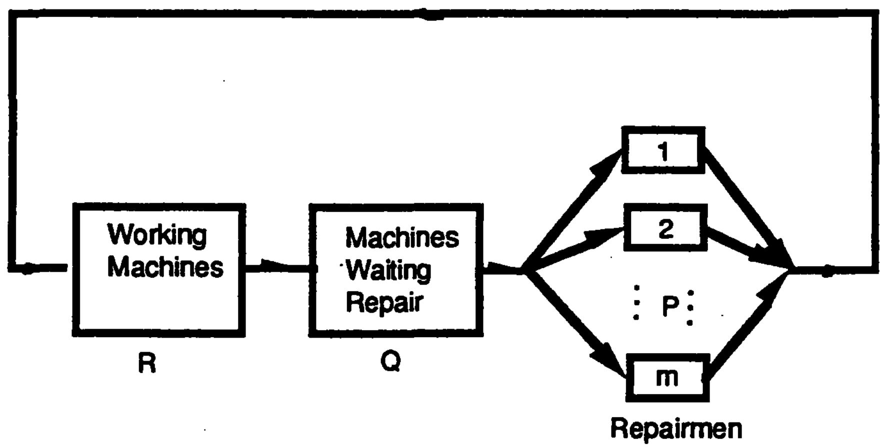
$$\lambda_1 P_1 + \lambda_2 P_2 = N$$

and $N \le 1$

This implies that

$$\lambda_1 P_1 + \lambda_2 P_2 \le 1$$

3.3



We represent the system as shown in the figure. In particular, once a machine breaks down, it goes into repair if a repairperson is available at the time, and otherwise waits in a queue for a repairperson to become free. Note that if m=1 this system is identical to the one of Example 3.7.

Let λ be the throughput of the system and let Q be the average time a broken down machine waits for a repairperson to become free. Applying Little's theorem to the entire system, we obtain

$$\lambda(R+Q+P) = N \qquad (1)$$

from which

$$\lambda(R+P) \le N \qquad (2)$$

Since the number of machines waiting repair can be at most (N-m), the average waiting time λQ is at most the average time to repair (N-m) machines, which is (N-m)P. Thus, from Eq. (1) we obtain

$$\lambda(R+ (N - m)P + P) \ge N \qquad (3)$$

Applying Little's theorem to the repairpersons, we obtain

$$\lambda P \le m \qquad (4)$$

The relations (2)-(4) give the following bounds for the throughput λ

$$\frac{N}{R + (N - m + 1)P} \leq \lambda \leq \min\left\{\frac{m}{P}, \frac{N}{R + P}\right\} \tag{5}$$

Note that these bounds generalize the ones obtained n Example 3.7 (see Eq. (3.9)).
By using the equation $T=N/\lambda$ for the average time between repairs, we obtain from Eq. (5)

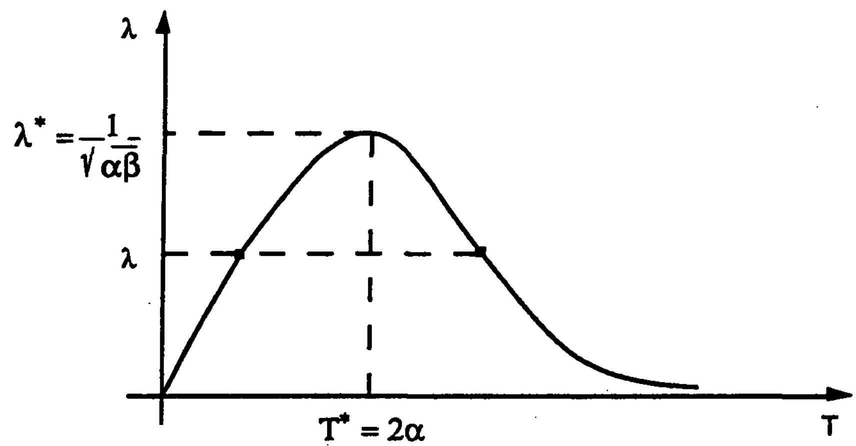
$$\min\{NP/m, R + P\} \leq T \leq R + (N - m + 1)P$$

3.4

If λ is the throughput of the system, Little's theorem gives $N = \lambda T$, so from the relation $T= \alpha + \beta N^2$ we obtain $T = \alpha + \beta\lambda^2 T^2$ or

$$\lambda = \sqrt{\frac{T - \alpha}{\beta T^2}} \tag{1}$$

This relation betweeen λ ands T is plotted below.



The maximum value of λ is attained for the value T^* for which the derivative of $(T - \alpha)/\beta T^2$ is zero (or $1/(\beta T^2) - 2(T - \alpha)/(\beta T^3) = 0$). This yields $T^* = 2\alpha$ and from Eq. (1), the corresponding maximal throughput value

$$\lambda^* = \frac{1}{\sqrt{\alpha\beta}} \tag{2}$$

(b) When $\lambda < \lambda^*$, there are two corresponding values of T: a low value corresponding to an uncongested system where N is relatively low, and a high value corresponding to a congested system where N is relatively high. This assumes that the system reaches a steady-state. However, it can be argued that when the system is congested a small increase in the number of cars in the system due to statistical fluctuations will cause an increase in the time in the system, which will tend to decrease the rate of departure of cars from the system. This will cause a further increase in the number in the system and a further increase in the time in the system, etc. In other words, when we are operating on the right side of the curve of the figure, there is a tendency for *instability* in the system, whereby a steady-state is never reached: the system tends to drift towards a traffic jam where the car depature rate from the system tends towards zero and the time a car spends in the system tends towards infinity. Phenomena of this type are analyzed in the context of the Aloha multiaccess system in Chapter 4.

3.5

The expected time in question equals

$$\begin{aligned}E\{\text{Time}\} = &(5 + E\{\text{stay of 2nd student}\})*P\{\text{1st stays less or equal to 5 minutes}\}\\ &+ (E\{\text{stay of 1st} \mid \text{stay of 1st} \geq 5\} + E\{\text{stay of 2nd}\})*\\ &\quad P\{\text{1st stays more than 5 minutes}\}.\end{aligned}$$

We have E{stay of 2nd student} = 30, and, using the memoryless property of the exponential distribution,

$$E\{\text{stay of 1st} \mid \text{stay of 1st} \geq 5\} = 5 + E\{\text{stay of 1st}\} = 35.$$

Also

$$P\{\text{1st student stays less or equal to 5 minutes}\} = 1 - e^{-5/30}$$
$$P\{\text{1st student stays more than 5 minutes}\} = e^{-5/30}.$$

By substitution we obtain

$$E\{\text{Time}\} = (5 + 30)*(1 - e^{-5/30}) + (35 + 30)*\ e^{-5/30} = 35 + 30*e^{-5/30} = 60.394.$$

3.6

(a) The probability that the person will be the last to leave is 1/4 because the exponential distribution is memoryless, and all customers have identical service time distribution. In particular, at the instant the customer enters service, the remaining service time of each of the other three customers served has the same distribution as the service time of the customer.

(b) The average time in the bank is 1 (the average customer service time) plus the expected time for the first customer to finish service. The latter time is 1/4 since the departure process is statistically identical to that of a single server facility with 4 times larger service rate. More precisely we have

$$\begin{aligned}P\{\text{no customer departs in the next t mins}\} &= P\{\text{1st does not depart in next t mins}\}\\ &\quad * P\{\text{2nd does not depart in next t mins}\}\\ &\quad * P\{\text{3rd does not depart in next t mins}\}\\ &\quad * P\{\text{4th does not depart in next t mins}\}\\ &= (e^{-t})^4 = e^{-4t}.\end{aligned}$$

Therefore

$$P\{\text{the first departure occurs within the next t mins}\} = 1 - e^{-4t},$$

and the expected time to the next depature is 1/4. So the answer is 5/4 minutes.

(c) The answer will not change because the situation at the instant when the customer begins service will be the same under the conditions for (a) and the conditions for (c).

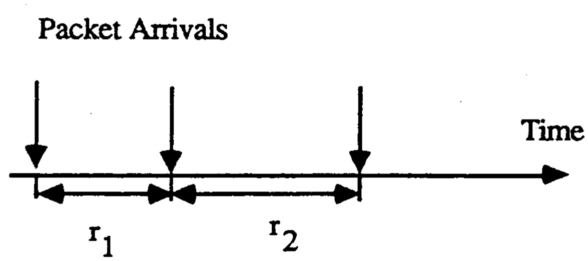
3.7

In the statistical multiplexing case the packets of at most one of the streams will wait upon arrival for a packet of the other stream to finish transmission. Let W be the waiting time , and note that $0 \leq W \leq T/2$. We have that one half of the packets have system time $T/2 + W$ and waiting time in queue W. Therefore

$$\text{Average System Time} = (1/2)T/2 + (1/2)(T/2+W) = (T+W)/2$$
$$\text{Average Waiting Time in Queue} = W/2$$
$$\text{Variance of Waiting Time} = (1/2)(W/2)^2+(1/2)(W/2)^2 = W^2/4.$$

So the average system time is between T/2 and 3T/4 and the variance of waiting time is between 0 and $T^2/16$.

3.8



Fix a packet. Let r_1 and r_2 be the interarrival times between a packet and its immediate predecessor, and successor respectively as shown in the figure above. Let X_1 and X_2 be the lengths of the predecessor packet, and of the packet itself respectively. We have:

$$P\{\text{No collision w/ predecessor or successor}\} = P\{r_1 > X_1, r_2 > X_2\}$$
$$= P\{r_1 > X_1\}P\{r_2 > X_2\}.$$

$$P\{\text{No collision with any other packet}\} = P_1 \, P\{r_2 > X_2\}$$

where

$$P_1 = P\{\text{No collision with all preceding packets}\}.$$

(a) For fixed packet lengths (= 20 msec)

$$P\{r_1 > X_1\} = P\{r_2 > X_2\} = e^{-\lambda*20} = e^{-0.01*20} = e^{-0.2}$$
$$P_1 = P\{r_1 \leq X_1\}.$$

Therefore the two probabilities of collision are both equal to $e^{-0.4} = 0.67$.

(b) For X exponentially distributed packet length with mean $1/\mu$ we have

$$P\{r_1 > X_1\} = P\{r_2 > X_2\} = \int_0^{\infty} P\{r_1 > X \mid X_1 = X\} p\{X_1 = X\} dX$$

$$= \int_0^{\infty} e^{-\lambda X} \mu e^{-\mu X} dX = \frac{\mu}{\lambda + \mu}$$

Substituting $\lambda = 0.01$ and $\mu = 0.05$ we obtain $P\{r_1 > X_1\} = P\{r_2 > X_2\} = 5/6$, and

$P\{\text{No collision w/ predecessor or successor}\} = (5/6)^2 = 0.694.$

Also P_1 is seen to be the steady-state probability of a customer finding an empty system in the $M/M/\infty$ system with arrival and service rate λ and μ respectively. Therefore $P_1 = e^{-\lambda/\mu} = e^{-0.2}$. Therefore

$P\{\text{No collision with any other packet}\} = e^{-0.2} 5/6 = 0.682.$

3.9

(a) For each session the arrival rate is $\lambda = 150/60 = 2.5$ packets/sec. When the line is divided into 10 lines of capacity 5 Kbits/sec, the average packet transmission time is $1/\mu = 0.2$ secs. The corresponding utilization factor is $\rho = \lambda/\mu = 0.5$. We have for each session $N_Q = \rho^2/(1 - \rho) = 0.5$, $N = \rho/(1 - \rho) = 1$, and $T = N/\lambda = 0.4$ secs. For all sessions collectively N_Q and N must be multiplied by 10 to give $N_Q = 5$ and $N = 10$.

When statistical multiplexing is used, all sessions are merged into a single session with 10 times larger λ and μ; $\lambda = 25$, $1/\mu = 0.02$. We obtain $\rho = 0.5$, $N_Q = 0.5$, $N = 1$, and $T = 0.04$ secs. Therefore N_Q, N, and T have been reduced by a factor of 10 over the TDM case.

(b) For the sessions transmitting at 250 packets/min we have $\rho = (250/60)*0.2 = 0.833$ and we have $N_Q = (0.833)^2/(1 - 0.833) = 4.158$, $N = 5$, $T = N/\lambda = 5/(250/60) = 1.197$ secs. For the sessions transmitting at 50 packets/min we have $\rho = (50/60)*0.2 = 0.166$, $N_Q = 0.033$, $N = 0.199$, $T = 0.199/(50/60) = 0.239$.

The corresponding averages over all sessions are $N_Q = 5*(4.158 + 0.033) = 21$, $N = 5*(5 + 0.199) = 26$, $T = N/\lambda = N/(5*\lambda_1 + 5*\lambda_2) = 26/(5*(250/60)+5*(50/60)) = 1.038$ secs.

When statistical multiplexing is used the arrival rate of the combined session is $5*(250 + 50) = 1500$ packets/sec and the same values for N_Q, N, and T as in (a) are obtained.

3.10

(a) Let t_n be the time of the nth arrival, and $\tau_n = t_{n+1} - t_n$. We have for $s \geq 0$

$$P\{\tau_n > s\} = P\{A(t_n + s) - A(t_n) = 0\} = e^{-\lambda s}$$

(by the Poisson distribution of arrivals in an interval). So

$$P\{\tau_n \leq s\} = 1 - e^{-\lambda s}$$

which is (3.11).

To show that $\tau_1, \tau_2, \ldots$ are independent, note that (using the independence of the numbers of arrivals in disjoint intervals)

$$P\{\tau_2 > s \mid \tau_1 = \tau\} = P\{0 \text{ arrivals in } (\tau, \tau+s] \mid \tau_1 = \tau\}$$
$$= P\{0 \text{ arrivals in } (\tau, \tau+s]\} = e^{-\lambda s} = P\{\tau_2 > s\}$$

Therefore τ_2 and τ_1 are independent.

To verify (3.12), we observe that

$$P\{A(t + \delta) - A(t) = 0\} = e^{-\lambda\delta}$$

so (3.12) will be shown if

$$\lim_{\delta\to 0} (e^{-\lambda\delta} - 1 + \lambda\delta)/\delta = 0$$

Indeed, using L'Hospital's rule we have

$$\lim_{\delta\to 0} (e^{-\lambda\delta} - 1 + \lambda\delta)/\delta = \lim_{\delta\to 0} (-\lambda e^{-\lambda\delta} + \lambda) = 0$$

To verify (3.13) we note that

$$P\{A(t + \delta) - A(t) = 1\} = \lambda\delta e^{-\lambda\delta}$$

so (3.13) will be shown if

$$\lim_{\delta\to 0} (\lambda\delta e^{-\lambda\delta} - \lambda\delta)/\delta = 0$$

This is equivalent to

$$\lim_{\delta\to 0} (\lambda e^{-\lambda\delta} - \lambda) = 0$$

which is clearly true.

To verify (3.14) we note that

$$P\{A(t + \delta) - A(t) \geq 2\} = 1 - P\{A(t + \delta) - A(t) = 0\} - P\{A(t + \delta) - A(t) = 1\}$$

$$= 1 - (1 - \lambda\delta + o(\delta)) - (\lambda\delta + o(\delta)) = o(\delta)$$

(b) Let N_1, N_2 be the number of arrivals in two disjoint intervals of lengths τ_1 and τ_2. Then

$$\begin{aligned}
P\{N_1+N_2 = n\} &= \Sigma^n_{k=0} P\{N_1 = k, N_2 = n-k\} = \Sigma^n_{k=0} P\{N_1 = k\}P\{N_2 = n-k\} \\
&= \Sigma^n_{k=0} e^{-\lambda\tau_1}[(\lambda\tau_1)^k/k!]e^{-\lambda\tau_2}[(\lambda\tau_2)^{(n-k)}/(n-k)!] \\
&= e^{-\lambda(\tau_1+\tau_2)}\Sigma^n_{k=0}[(\lambda\tau_1)^k(\lambda\tau_2)^{(n-k)}]/[k!(n-k)!] \\
&= e^{-\lambda(\tau_1+\tau_2)}[(\lambda\tau_1 + \lambda\tau_2)^n/n!]
\end{aligned}$$

(The identity

$$\Sigma^n_{k=0}[a^k b^{(n-k)}]/[k!(n-k)!] = (a + b)^n/n!$$

can be shown by induction.)

(c) The number of arrivals of the combined process in disjoint intervals is clearly independent, so we need to show that the number of arrivals in an interval is Poisson distributed, i.e.

$$\begin{aligned}
P\{A_1(t+\tau) + \ldots + A_k(t+\tau) - A_1(t) - \ldots - A_k(t) = n\} \\
= e^{-(\lambda_1 + \ldots + \lambda_k)\tau}[(\lambda_1 + \ldots + \lambda_k)\tau]^n/n!
\end{aligned}$$

For simplicity let k=2; a similar proof applies for $k > 2$. Then

$$\begin{aligned}
P\{A_1(t+\tau) + A_2(t+\tau) - A_1(t) - A_2(t) = n\} \\
&= \Sigma^n_{m=0} P\{A_1(t+\tau) - A_1(t) = m, A_2(t+\tau) - A_2(t) = n-m\} \\
&= \Sigma^n_{m=0} P\{A_1(t+\tau) - A_1(t) = m\}P\{A_2(t+\tau) - A_2(t) = n-m\}
\end{aligned}$$

and the calculation continues as in part (b). Also

$$\begin{aligned}
P\{1 \text{ arrival from } A_1 \text{ prior to } t \mid 1 \text{ occured}\} \\
&= P\{1 \text{ arrival from } A_1, 0 \text{ from } A_2\}/P\{1 \text{ occured}\} \\
&= (\lambda_1 t e^{-\lambda_1 t} e^{-\lambda_2 t})/(\lambda t e^{-\lambda t}) = \lambda_1/\lambda
\end{aligned}$$

(d) Let t be the time of arrival. We have

$$\begin{aligned}
P\{t < s \mid 1 \text{ arrival occured}\} &= P\{t < s, 1 \text{ arrival occured}\}/P\{1 \text{ arrival occured}\} \\
&= P\{1 \text{ arrival occured in } [t_1, s), 0 \text{ arrivals occured in } [s, t_2]\}/P\{1 \text{ arrival occured}\} \\
&= (\lambda(s - t_1)e^{-\lambda(s - t_1)}e^{-\lambda(s - t_2)})/(\lambda(t_2 - t_1)e^{-\lambda(t_2 - t_1)}) = (s - t_1)/(t_2 - t_1)
\end{aligned}$$

This shows that the arrival time t is uniformly distributed in $[t_1, t_2]$.

3.11

(a) Let us call the two transmission lines 1 and 2, and let $N_1(t)$ and $N_2(t)$ denote the respective numbers of packet arrivals in the interval [0,t]. Let also $N(t) = N_1(t) + N_2(t)$. We calculate the joint probability $P\{N_1(t) = n, N_2(t) = m\}$. To do this we first condition on N(t) to obtain

$$P\{N_1(t) = n, N_2(t) = m\} = \Sigma_{k=0}^{\infty} P\{N_1(t) = n, N_2(t) = m \mid N(t) = k\}P\{N(t) = k\}.$$

Since

$$P\{N_1(t) = n, N_2(t) = m \mid N(t) = k\} = 0 \qquad \text{when} \quad k \neq n+m$$

we obtain

$$P\{N_1(t) = n, N_2(t) = m\} = P\{N_1(t) = n, N_2(t) = m \mid N(t) = n + m\}P\{N(t) = n + m\}$$
$$= P\{N_1(t) = n, N_2(t) = m \mid N(t) = n + m\}e^{-\lambda t}[(\lambda t)^{n+m}/(n + m)!]$$

However, given that n+m arrivals occurred, since each arrival has probability p of being a line 1 arrival and probability 1-p of being a line 2 arrival, it follows that the probability that n of them will be line 1 and m of them will be line 2 arrivals is the binomial probability

$$\binom{n+m}{n}p^n(1-p)^m$$

Thus

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

$$= e^{-\lambda tp}\frac{(\lambda tp)^n}{n\,!} \quad e^{-\lambda t(1-p)}\frac{(\lambda t(1-p))^m}{m!} \qquad (1)$$

Hence

$$P\{N_1(t) = n\} = \sum_{m=0}^{\infty} P\{N_1(t) = n,\ N_2(t) = m\}$$

$$= e^{-\lambda tp}\frac{(\lambda tp)^n}{(n)!}\sum_{m=0}^{\infty} e^{-\lambda t(1-p)}\frac{(\lambda t(1-p))^m}{m!}$$

$$= e^{-\lambda tp}\frac{(\lambda tp)^n}{(n)!}$$

That is, $\{N_1(t), t \geq 0\}$ is a Poisson process having rate λp. Similarly we argue that $\{N_2(t), t \geq 0\}$ is a Poisson process having rate $\lambda(1 - p)$. Finally from Eq. (1) it follows that the two processes are independent since the joint distribution factors into the marginal distributions.

(b) Let A, A_1, and A_2 be as in the hint. Let I be an interarrival interval of A_2 and consider the number of arrivals of A_1 that lie in I. The probability that this number is n is the probability of n successive arrivals of A_1 followed by an arrival of A_2, which is $\rho^n(1 - \rho)$. This is also the probability that a customer finds upon arrival n other customers waiting in an M/M/1 queue. The service time of each of these customers is exponentially distributed with parameter μ, just like the interarrival times of process A. Therefore the waiting time of the customer in the M/M/1 system has the same distribution as the interarrival time of process A_2. Since by part (a), the process A_2 is Poisson with rate $\mu - \lambda$, it follows that the waiting time of the customer in the M/M/1 system is exponentially distributed with parameter $\mu - \lambda$.

3.12

For any scalar s we have using also the independence of τ_1 and τ_2

$$P(\min\{\tau_1,\tau_2\} \geq s) = P(\tau_1 \geq s, \tau_2 \geq s) = P(\tau_1 \geq s)\, P(\tau_2 \geq s)$$

$$= e^{-\lambda_1 s} e^{-\lambda_2 s} = e^{-(\lambda_1 + \lambda_2)s}$$

Therefore the distribution of $\min\{\tau_1,\tau_2\}$ is exponential with mean $1/(\lambda_1 + \lambda_2)$.

By viewing τ_1 and τ_2 as the arrival times of the first arrivals from two independent Poisson processes fwith rates λ_1 and λ_2, we see that the equation $P(\tau_1 < \tau_2) = \lambda_1/(\lambda_1 + \lambda_2)$ follows from Problem 3.10(c).

Consider the M/M/1 queue and the amount of time spent in a state k>0 between transition into the state and transition out of the state. This time is $\min\{\tau_1,\tau_2\}$, where τ_1 is the time between entry to the state k and the next customer arrival and τ_2 is the time between entry to the state k and the next service completion. Because of the memoryless property of the exponential distribution, τ_1 and τ_2 are exponentially distributed with means $1/\lambda$ and $1/\mu$, respectively. It follows using the fact shown above that the time between entry and exit from stae k is exponentially distributed with mean $1/(\lambda+\mu)$. The probability that the transition will be from k to k+1 is $\lambda/(\lambda+\mu)$ and that the transition will be from k to k-1 is $\mu/(\lambda+\mu)$. For state 0 the amount of time spent is exponentially distributed with mean $1/\lambda$ and the probability of a transition to state 1 is 1. Because of this it can be seen that M/M/1 queue can be described as a continuous Markov chain with the given properties.

3.13

(a) Consider a Markov chain with state

n = Number of people waiting + number of empty taxi positions

Then the state goes from n to n+1 each time a person arrives and goes from n to n-1 (if $n \geq 1$) when a taxi arrives. Thus the system behaves like an M/M/1 queue with arrival rate 1 per min and departure rate 2 per min. Therefore the occupancy distribution is

$p_n = (1-\rho)/\rho^n$

where $\rho = 1/2$. State n, for $0 \le n \le 4$ corresponds to 5, 4, 3, 2, 1 taxis waiting while $n > 4$ corresponds to no taxi waiting. Therefore

P{5 taxis waiting} = 1/2
P{4 taxis waiting} = 1/4
P{3 taxis waiting} = 1/8
P{2 taxis waiting} = 1/16
P{1 taxi waiting} = 1/32

and P{no taxi waiting} is obtained by subtracting the sum of the probabilities above from unity. This gives P{no taxi waiting} = 1/32.

(b) See the hint.

(c) This system corresponds to taxis arriving periodically instead of arriving according to a Poisson process. It is the slotted M/D/1 system analyzed in Section 6.3.

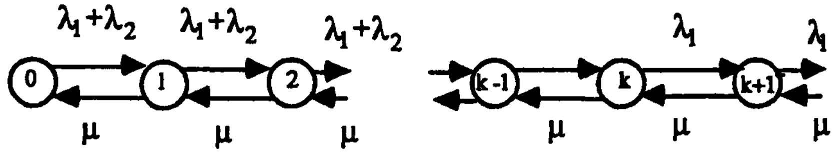
3.14

(a) The average message transmission time is $1/\mu = L/C$ so the service rate is $\mu = C/L$. When the number of packets in the system is larger than K, the arrival rate is λ_1. We must have

$$0 \le \lambda_1 < \mu$$
$$0 \le \lambda_2$$

in order for the arrival rate at node A to be less than the service rate for large state values. For these values, therefore, the average number of packets in the system will stay bounded.

(b) The corresponding Markov chain is as shown in the figure below. The steady-state probabilities satisfy



$p_n = \rho^n p_0$ for $n \le k$
$p_n = \rho_1^{n-k}\rho^k p_0$ for $n > k$

where $\rho = (\lambda_1 + \lambda_2)/\mu$, $\rho_1 = \lambda_1/\mu$. We have

$$\Sigma^{\infty}_{n=0}\, p_n = 1$$

or

$$p_0 \Sigma^k_{n=0} \rho^n + \Sigma^{\infty}_{n=k+1} \rho_1^{n-k} \rho^k = 1$$

from which we obtain after some calculation

$$p_0 = [(1 - \rho)(1 - \rho_1)]/[1 - \rho_1 - \rho^k(\rho-\rho_1)] \qquad \text{for } \rho < 1$$

and

$$p_0 = (1 - \rho_1)/[1 + k(1 - \rho_1)] \qquad \text{for } \rho = 1$$

For packets of source 1 the average time in A is

$$T_1 = (1/\mu)(1 + N)$$

where

$$N = \Sigma^{\infty}_{n=0}\, np_n$$

is the average number in the system upon arrival. The average number in A from source 1 is

$$N_1 = \lambda_1 T_1$$

For packets of source 2 the average time in A is

$$T_2 = (1/\mu)(1+ N')$$

where

$$N' = \frac{\sum_{n=0}^{k-1} np_n}{\sum_{n=0}^{k-1} p_n}$$

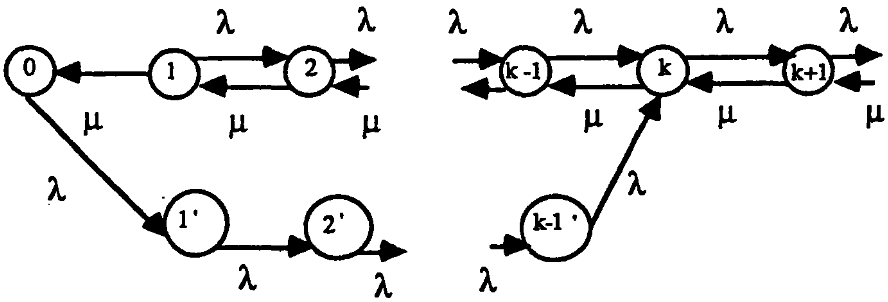
is the average number in the system found by an accepted packet of source 2. To find the average number in the system from source 2 we must find the arrival rate into node A of packets from source 2. This is

$$\lambda'_2 = \lambda_2 P\{\text{arriving packet from source 2 is accepted}\} = \lambda_2 \Sigma^{k-1}_{n=0}\, p_n$$

and the average number from source 2 in A is

$$N_2 = \lambda'_2 T_2$$

3.15



The transition diagram of the corresponding Markov chain is shown in the figure. We have introduced states 1', 2', . . ., (k-1)' corresponding to situations where there are customers in the system waiting for service to begin again after the system has emptied out. Using global balance equations between the set of states {1', 2', . . . ,i') and all other states, for i' = 1', . . . , (k-1)', we obtain $\lambda p_0 = \lambda p_{1'} = \lambda p_{2'} = \ldots = \lambda p_{(k-1)'}$, so

$$p_0 = p_{1'} = p_{2'} = \ldots = p_{(k-1)'}$$

Also by using global balance equations we have

$$\mu p_1 = \lambda p_0$$
$$\mu p_2 = \lambda(p_1 + p_{1'}) = \lambda(p_1 + p_0)$$
$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$
$$\mu p_k = \lambda(p_{k-1} + p_{(k-1)'}) = \lambda(p_{k-1} + p_0)$$
$$\mu p_{i+1} = \lambda p_i \qquad i \geq k.$$

By denoting $\rho = \lambda/\mu$ we obtain

$$p_i = \rho(1 + \rho + \ldots + \rho^{i-1})p_0 \qquad 1 \leq i \leq k$$

$$p_i = \rho^{1+i-k}(1 + \rho + \ldots + \rho^{k-1})p_0 \qquad i > k.$$

Substituting these expressions in the equation $p_{1'} + \ldots + p_{(k-1)'} + p_0 + p_1 + \ldots = 1$ we obtain p_0

$$p_0\left(k + \sum_{i=1}^{k} \frac{\rho(1-\rho^i)}{1-\rho} + \frac{1-\rho^k}{1-\rho}\rho^2(1+\rho + \ldots)\right) = 1$$

$$p_0 = \left(k + \frac{\rho}{1-\rho}\sum_{i=1}^{k}(1-\rho^i) + \frac{\rho^2}{(1-\rho)^2}(1-\rho^k)\right)^{-1}$$

After some calculation this gives $p_0 = (1 - \rho)/k$ (An alternative way to verify this formula is to observe that the fraction of time the server is busy is equal to ρ by Little's theorem). Therefore, the fraction of time the server is idle is $(1 - \rho)$. When this is divided among the k equiprobable states 0, 1', . . ., (k-1)' we obtain $p_0 = (1 - \rho)/k$. The average number in the system is

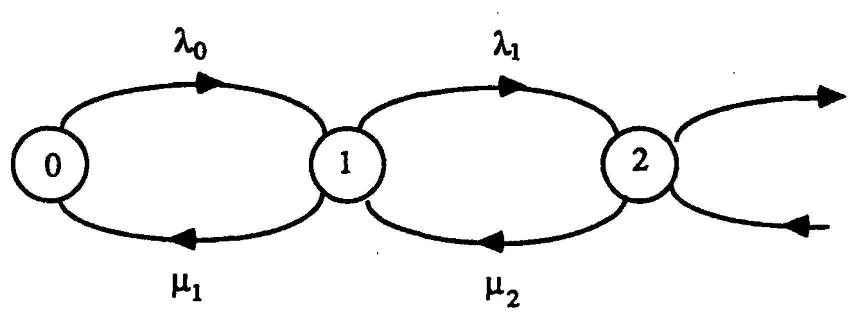
$$N = p_{1'} + 2p_{2'} + \ldots + (k-1)p_{(k-1)'} + \sum_{i=0}^{\infty} ip_i = p_0 \frac{k(k-1)}{2} + \sum_{i=0}^{\infty} ip_i$$

where the probabilities p_i are given in the equations above. After some calculation this yields

$$N = (k-1)/2 + \rho/(1 - \rho).$$

The average time in the system is (by Little's Theorem) $T = N/\lambda$.

3.16



The figure shows the Markov chain corresponding to the given system. The local balance equation for it can be written down as :

$$\rho_0 p_0 = p_1$$

$$\rho_1 p_1 = p_2$$

$$\ldots \qquad \ldots$$

$$\Rightarrow p_{n+1} = \rho_n p_n = \rho_{n-1} \rho_n p_{n-1} = \ldots = (\rho_0 \rho_1 \cdots \rho_n) p_0$$

but,

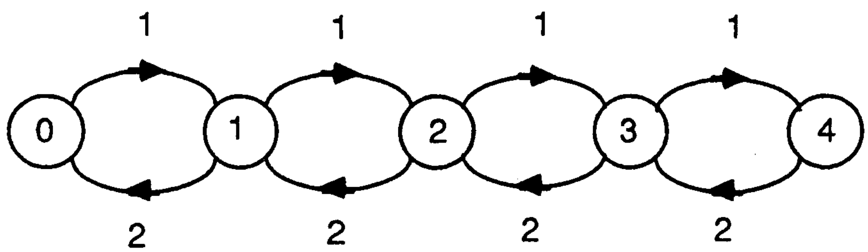
$$\sum_{i=0}^{\infty} p_i = p_0(1 + \rho_0 + \rho_0 \rho_1 + \ldots.) = 1$$

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

3.17

The discrete time version of the M/M/1 system can be characterized by the same Markov chain as the continuous time M/M/1 system and hence will have the same occupancy distribution.

3.18



$$p_1 = \frac{1}{2} p_0$$

$$p_n = \frac{1}{2} p_{n-1} \quad \text{for } 1 \le n \le 4$$

$$\sum_{i=0}^{4} p_i = 1$$

Solving the above equations we get,

$$p_n = \frac{2^{4-n}}{31} \quad \text{for } 0 \le n \le 4$$

$$N = \sum_{n=0}^{4} n\, p_n = \frac{26}{31}$$

P(a customer arrives but has to leave) = 1/31

Hence the arrival rate of passengers who join the queue =

$$(1-p_4)\, \lambda = \frac{30}{31} \text{ per minute } = \lambda_a \text{ (say)}$$

$$T = N/\lambda_a = \frac{26/31}{30/31} = \frac{13}{15} \text{ minutes}$$

3.19

We have here an M/M/m/m system where m is the number of circuits provided by the company. Therefore we must find the smallest m for which $p_m < 0.01$ where p_m is given by the Erlang B formula

$$p_m = \frac{(\lambda/\mu)^m/m!}{\sum_{n=0}^{m} (\lambda/\mu)^n/n!}.$$

We have $\lambda = 30$ and $\mu = 1/3$, so $\lambda/\mu = 30 \cdot 3 = 90$. By substitution in the equation above we can calculate the required value of m.

3.20

We view this as an M/M/m problem. We have

λ=0.5, E(X) = $1/\mu$ = 3, m=? so that W<0.5

We know that the utilization factor has to be less than 1 or m has to be greater than or equal to 2. By the M/M/n results we have

$$W = \frac{\frac{\lambda}{m\mu} P_Q}{\lambda\left(1 - \frac{\lambda}{m\mu}\right)} = \frac{P_Q}{m\mu - \lambda}$$

$$\text{where } P_Q = \frac{p_0 \left(\frac{\lambda}{\mu}\right)^m}{m!\left(1 - \frac{\lambda}{m\mu}\right)}$$

$$\text{and } \quad p_0 = \left[\sum_{n=0}^{m-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^m}{m!\,(1 - \lambda/\mu)}\right]^{-1}$$

As can be seen from the expressions above m should be at most 5 because at m=5 , W is less than 0.5 because P_Q is less than 1.

The following C program calculates the optimum m.

```
double P0(lambda,mu,m){
        mrho = lambda/mu;
        rho = mrho/m;
        for(n=0; n<m; n++)
                temp1 = pow(mrho,n)/ fact(n);
```

```
        temp2 = pow(mrho,m)/(fact(m)*(1-rho));
        return(1/( temp1 + temp2 )); /* this returns p0 */
}

int fact(n){
        if (n==0) return (1);
        else
                return(n* fact (n-1));
}

double W(lambda,mu,m){
        PQ = P0(lambda,mu,m) * pow(mrho,m) /
                                                (fact(m) * (1-rho));
return(PQ/(m *mu - lambda));
}       /* this returns W for a given m */

main() {
        lambda = 0.5; mu = 0.333; previous_W = 100.0;
        for(m=2; m<=5; m++)
                if ((temp = W(lambda,mu,m)) < Previous_W)
                        previous_W = temp;
                else
                        { print(m-1);
                        break;
                }
}
```

3.21

We have $p_n = \rho^n p_0$ where $\rho = \lambda/\mu$. Using the relation

$$\sum_{n=0}^{m} p_n = 1$$

we obtain

$$p_0 = \frac{1}{\sum_{n=0}^{m} \rho^n} = \frac{1-\rho}{1-\rho^{m+1}}$$

Thus

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

3.22

(a) When all the courts are busy, the expected time between two departures is 40/5 = 8 minutes. If a pair sees k pairs waiting in the queue, there must be exactly k+1 departures from the system before they get a court. Since all the courts would be busy during this whole time, the average waing time required before k+1 departures is 8(k+1) minutes.

(b) Let X be the expected waiting time given that the courts are found busy. We have

$$\lambda = 1/10, \qquad \mu = 1/40, \qquad \rho = \lambda/(5\mu) = 0.8$$

and by the M/M/m results

$$W = \frac{\rho P_Q}{\lambda(1 - \rho)}$$

Since $W = XP_Q$, we obtain $X = W/P_Q = \rho/[\lambda(1 - \rho)] = 40$ min.

3.23

Let

$$p_m = P\{\text{the 1st m servers are busy}\}$$

as given by the Erlang B formula. Denote

r_m = Arrival rate to servers (m+1) and above
λ_m = Arrival rate to server m.

We have

$$r_m = p_m\lambda$$
$$\lambda_m = r_{m-1} - r_m = (p_{m-1} - p_m)\lambda.$$

The fraction of time server m is busy is

$$b_m = \lambda_m/\mu.$$

3.24

We will show that the system is described by a Markov chain that is identical to the M/M/1 chain. For small δ we have

$$P\{\text{k arrivals and j departures}\} = 0(\delta) \qquad \text{if } k + j \geq 2$$

$$P\{\text{0 arrivals and 1 departure} \mid \text{starting state} = i \geq i\}$$
$$= P\{\text{0 arrivals} \mid \text{starting state } i \geq 1\} \bullet P\{\text{1 departure} \mid \text{starting state } i \geq 1\}$$

We have

$$P\{\text{ 0 arrivals} \mid \text{starting state } i \geq 1\} = P\{0 \text{ arrivals}\} = 1 - \lambda\delta + 0(\delta).$$

The probability P{1 departure | starting state i > 1} is obtained from the binomial distribution or sum of i Bernoulli trials, each with a probability of success equal to $(\mu/i)\,\delta + 0(\delta)$. We need the probability of one success, which is

$$\binom{i}{1}(1-(\mu/i)\,\delta + 0(\delta))^{i-1}\,((\mu/i)\,\delta + 0(\delta))$$

Therefore

$$P\{\text{ 0 arrivals and 1 departure} \mid \text{starting state} = i \geq 1\}$$
$$= \binom{i}{1}(1-(\mu/i)\,\delta + 0(\delta))^{i-1}\,((\mu/i)\,\delta + 0(\delta)) \bullet (1-\lambda\delta + 0(\delta)) = \mu\delta + 0(\delta)$$

Similarly

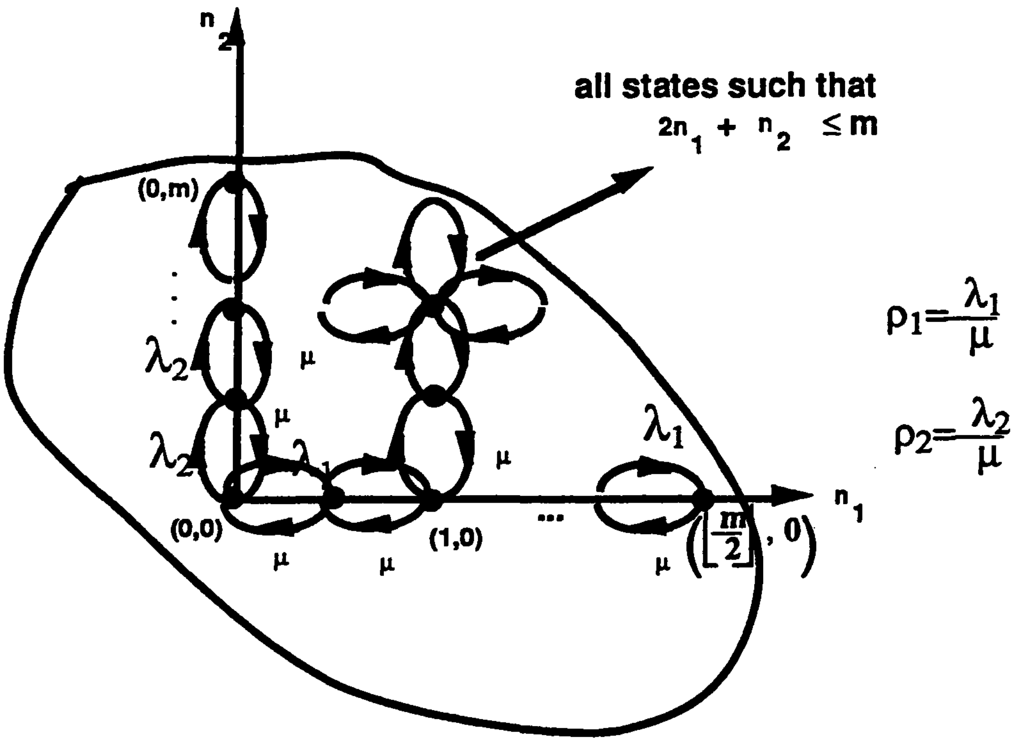
$$P\{1 \text{ arrival and 0 departure} \mid \text{starting state} = i\}$$
$$= P\{1 \text{ arrival}\} \bullet P\{0 \text{ departure} \mid \text{starting state} = i\}$$

$$= (\lambda\delta + 0(\delta)) \bullet \left[\binom{i}{0}(1-(\mu/i)\,\delta + 0(\delta))^{i}\right] = \lambda\delta + 0(\delta)$$

Thus the transition rates are the same as for the M/M/1 system.

3.25

Let n_1 be the number of radio-to-radio calls and n_2 be the number of radio-to-nonradio calls which have not finished yet. Then we have the following Markov chain:



The occupancy distribution $p(n_1,n_2)$ is of the form

$$p(n_1,n_2) = \rho_1^{n_1}(1-\rho_1)\rho_2^{n_2}(1-\rho_2)/G, \text{ for } 2n_1+n_2 \le m$$

and 0 otherwise (it is easy to check that this formula satisfies the global balance equations). To calculate G note that

$$\sum_{\{(n_1,n_2)|2n_1+n_2\le m\}} \sum p(n_1,n_2) = 1 \Rightarrow G = \sum_{\{(n_1,n_2)|2n_2+n_2\le m\}} \rho_1^{n_1}(1-\rho_1)\rho_2^{n_2}(1-\rho_2) =$$

$$\sum_{n_1=0}^{\left\lfloor \frac{m}{2}\right\rfloor}\sum_{n_2=0}^{m-2n_1}\rho_1^{n_1}(1-\rho_1)\rho_1^{n_2}(1-\rho_2) = \sum_{n_1=0}^{\left\lfloor \frac{m}{2}\right\rfloor}\rho_1^{n}(1-\rho_1)(1-\rho_2)\frac{1-\rho_2^{m-2n_1+1}}{1-\rho_2}$$

$$= (1-\rho_1)\sum_{n_1=0}^{\left\lfloor \frac{m}{2}\right\rfloor}\rho_1^{n_1} - \sum_{n_1=0}^{\left\lfloor \frac{m}{2}\right\rfloor}(1-\rho_1)\rho_2^{m+1}\left(\frac{\rho_1}{\rho_2^2}\right)^{n_1}$$

$$= 1-\rho_1^{\left\lfloor \frac{m}{2} \right\rfloor +1} - (1-\rho_1)\rho_2^{m+1} \frac{1-\left(\frac{\rho_1}{\rho_2^2}\right)^{\left\lfloor \frac{m}{2} \right\rfloor +1}}{1-\frac{\rho_1}{\rho_2^2}}$$

$$= 1-\rho_1^{\left\lfloor \frac{m}{2} \right\rfloor +1} - (1-\rho_1)\rho_2^{m+1-2\left\lfloor \frac{m}{2} \right\rfloor} \frac{\rho_2^{2\left\lfloor \frac{m}{2} \right\rfloor +2} - \rho_1^{\left\lfloor \frac{m}{2} \right\rfloor +1}}{\rho_2^2 - \rho_1}$$

$$\Rightarrow G = \begin{cases} 1-\rho_1^{\frac{m}{2}+1} - (1-\rho_1)\rho_2 \dfrac{\rho_2^{m+2} - \rho_1^{m/2+1}}{\rho_2^2-\rho_1} & \text{if m even} \\[2ex] 1-\rho_1^{\frac{m+1}{2}} - (1-\rho_1)\rho_2^2 \dfrac{\rho_2^{m+1} - \rho_1^{\frac{m+1}{2}}}{\rho_2^2-\rho_1} & \text{if m odd} \end{cases}$$

Let

p_1 = blocking probability of radio-to-radio calls
p_2 = blocking probability of radio-to-nonradio calls

Then

$$p_2 = \sum_{2n_1+n_2=m} p(n_1,n_2)$$

$$p_1 = \sum_{m-1\leq 2n_1+n_2\leq m} p(n_1,n_2) = p_2 + \sum_{2n_1+n_2=m-1} p(n_1,n_2).$$

But

$$p_2 = \sum_{n_1=0}^{\left\lfloor \frac{m}{2} \right\rfloor} p(n_1, m-2n_1) = \sum_{n_1=0}^{\left\lfloor \frac{m}{2} \right\rfloor} \rho_1^{n}(1-\rho_1)\rho_2^{m-2n}(1-\rho_2)/G =$$

$$= \frac{(1-\rho_1)(1-\rho_2)\rho_2^m}{G} \frac{1-\left(\rho_1/\rho_2^2\right)^{\left\lfloor \frac{m}{2} \right\rfloor +1}}{1-\frac{\rho_1}{\rho_2^2}}$$

and

$$p_1 = p_2 + \sum_{n_1=0}^{\lfloor \frac{m-1}{2} \rfloor} p(n_1, m-1-2n_1) = p_2 + \frac{(1-\rho_1)(1-\rho_2)\rho_2^{m-1}}{G} \cdot \frac{1-\left(\frac{\rho_1}{\rho_2^2}\right)^{\lfloor \frac{m+1}{2} \rfloor}}{1-\frac{\rho_1}{\rho_2^2}}$$

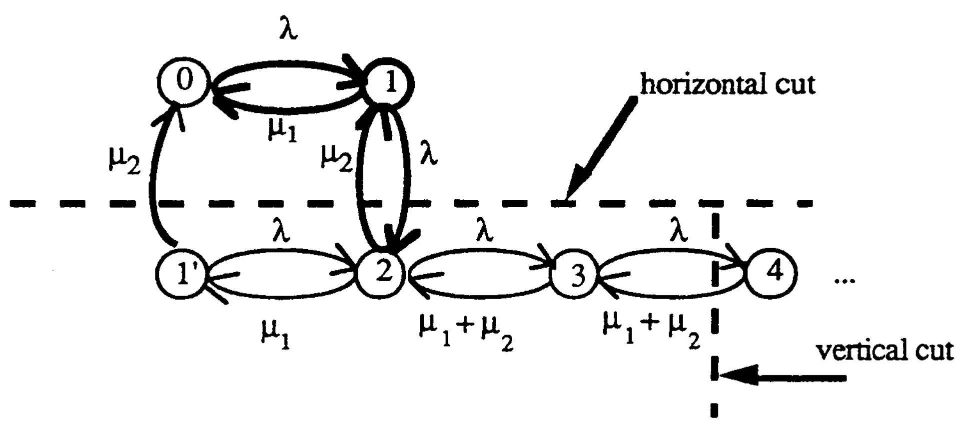
3.26

Define the state to be the number of operational machines. This gives a Markov chain, which is the same as in an M/M/1/m queue with arrival rate λ and service rate μ. The required probability is simply p_0 for such a queue.

3.27

Assume $\mu_1 > \mu_2$.

We have the following Markov chain:



Let state 1 represent 1 customer in the system being served by server 1
Let state 1' represent 1 customer in the system being served by server 2

i) Flow across a vertical cut

$$p_i = \frac{\lambda}{\mu_1 + \mu_2} p_{i-1} \qquad \text{for } i \geq 2$$

Therefore

$$\bar{p}_i = \left(\frac{\lambda}{\mu_1 + \mu_2}\right)^{i-2} p_2 \qquad \text{for } i \geq 2$$

ii) Flow in and out of state 1'

$$(\lambda + \mu_2)\, p_{1'} = p_2\, \mu_1$$

Therefore

$$p_{1'} = p_2 \; \frac{\mu_1}{\lambda + \mu_2}$$

iii) Flow across horizontal cut

$$p_1\, \lambda = (p_{1'} + P_2)\, \mu_2$$

Therefore

$$p_1 = \frac{\mu_2}{\lambda}\left(p_2 + p_2 \; \frac{\mu_1}{\lambda + \mu_2}\right) = p_2\, \frac{\mu_2}{\lambda}\left(1 + \frac{\mu_1}{\lambda + \mu_2}\right)$$

iv) Flow in and out of state 0

$$p_0\, \lambda = \pi_1\, \mu_1 + p_{1'}\, \mu_2$$

Therefore

$$p_0 = \frac{1}{\lambda}\, p_2 \left(\frac{\mu_1 \mu_2}{\lambda}\left(1 + \frac{\mu_1}{\lambda + \mu_2}\right) + \frac{\mu_1 \mu_2}{\lambda + \mu_2}\right)$$

We have

$$\sum_i p_i = 1$$

from which

$$p_2 = \left(\frac{1}{1 - \dfrac{\lambda}{\mu_1 + \mu_2}} + \frac{\left(1 + (\mu_2/\lambda)\right)\mu_1}{\lambda + \mu_2} + \frac{\left(1 + (\mu_1/\lambda)\right)\mu_2}{\lambda}\left(1 + \frac{\mu_1}{\lambda + \mu_2}\right)\right)^{-1}$$

3.28

We have

$$E\{f^2\} = E\left\{\left(\sum_{i=1}^{n}\gamma_i\right)^2\right\} = E\left\{E\left\{\left(\sum_{i=1}^{n}\gamma_i\right)^2 \mid n\right\}\right\} = E\{ns_\gamma^2 + n(n-1)\Gamma^2\}$$

$$= E\{n\}(s_\gamma^2 - \Gamma^2) + E\{n^2\}\Gamma^2$$

Since

$$E\{n\} = \lambda/\mu, \qquad E\{n\ \} = \sigma_n^2 + (\lambda/\mu)\ = \lambda/\mu + (\lambda/\mu)$$

we obtain

$$\sigma_f^2 = E\{f^2\} - F^2 = E\{f^2\} - (\lambda/\mu)^2\Gamma^2 = (\lambda/\mu)(s_\gamma^2 - \Gamma^2) + [(\lambda/\mu) + (\lambda/\mu)^2]\Gamma^2 - (\lambda/\mu)^2\Gamma^2$$

$$= (\lambda/\mu)s_\gamma^2$$

so finally

$$\sigma_f = (\lambda/\mu)^{1/2}s_\gamma$$

3.29

For each value of x, the average customer waiting time for each of the two types of customers (x items or less, vs more than x) is obtained by the P-K formula for an M/G/1 system with arrival rate and service time parameters depending on x. By computing the overall customer waiting time for each x in the feasible range [1,40], we can numerically compute the optimal value of x.

Here is a program to find x to minimize the average waiting time:

```
Lambda=1; Past_T= 1000000; T=0; x=-1;
while (x<=40) do
if (T> Past_T) do
        begin
                Past_T = T;
                x = x+1;
                lambda1 = lambda * x/40;
                E_service_time_1 = (1+x)/2;
                E_service_time_2 =  (41+x)/2;
                E_service_time_square1 = 0;
                E_service_time_square2 = 0;
                for i=1 to x do
                        E_service_time_square1 =
                        E_service_time_square1+(i*i);
```

```
                for i=x+1 to 40 do
                        E_service_time_square2 =
                                E_service_time_square2+(i*i);
                E_service_time_square1 =
                        E_service_time_square1/x;
                E_service_time_square2 =
                                E_service_time_square2/(40-x);
                T1 = E_service_time_1 +
                (lambda*E_service_time_square1/(2.0*(1-
                lambda1*E_service_time_1)));
                T2 = E_service_time_2 +
                        (lambda*E_service_time_square2/(2.0*(1-
                                        lambda2*E_service_time_2)));
                T = (T1*x + T2*(40-x))/40;
end;
        print(x);
```

3.30

From Little's Theorem (Example 1) we have that P{the system is busy} = $\lambda E\{X\}$. Therefore P{the system is empty} = $1 - \lambda E\{X\}$.

The length of an idle period is the interarrival time between two typical customer arrivals. Therefore it has an exponential distribution with parameter λ, and its average length is $1/\lambda$.

Let B be the average length of a busy period and let I be the average length of an idle period. By expressing the proportion of time the system is busy as $B/(I + B)$ and also as $\lambda E\{X\}$ we obtain

$$B = E\{X\}/(1 - \lambda E\{X\}).$$

From this the expression $1/(1 - \lambda E\{X\})$ for the average number of customers served in a busy period is evident.

3.31

The problem with the argument given is that more customers arrive while long-service customers are served, so the average service time of a customer found in service by another customer upon arrival is more than $E\{X\}$.

3.32

Following the hint we write for the ith packet

$$U_i = R_i + \sum_{j=1}^{N_i} X_{i-j}$$

where

U_i : Unfinished work at the time of arrival of the ith customer
R_i : Residual service time of the ith customer
N_i : Number found in queue by the ith customer
X_j : Service time of the jth customer

Hence

$$E\{U_i\} = E\{R_i\} + E\left\{\sum_{j=1}^{N_i} E\{X_{i-j} \mid N_i\}\right\}$$

Since X_{i-j} and N_i are independent

$$E\{U_i\} = E\{R_i\} + E\{X\}E\{N_i\}$$

and by taking limit as $i \to \infty$ we obtain $U = R + (1/\mu)N_Q = R + (\lambda/\mu)W = R + \rho W$, so

$$W = (U - R)/\rho.$$

Now the result follows by noting that both U and R are independent of the order of customer service (the unfinished work is independent of the order of customer service, and the steady state mean residual time is also independent of the customer service since the graphical argument of Fig. 3.16 does not depend on the order of customer service).

3.33

Consider the limited service gated system with zero packet length and reservation interval equal to a slot. We have

T_{TDM} = Waiting time in the gated system

For $E\{X^2\} = 0$, $E\{V\} = 1$, $\sigma_V^2 = 0$, $\rho = 0$ we have from the gated system formula (3.77)

Waiting time in the gated system = $(m + 2 - 2\lambda)/(2(1 - \lambda)) = m/(2(1 - \lambda)) + 1$

which is the formula for T_{TDM} given by Eq. (3.59).

3.34

(a) The system utilization is ρ, so the fraction of time the system transmits data is ρ. Therefore the portion of time occupied by reservation intervals is $1 - \rho$.

(b) If

p: Fraction of time a reservation interval is followed by an empty data interval

and M(t) is the number of reservation intervals up to time t, then the number of packets transmitted up to time t is $\approx (1 - p)M(t)$. The time used for reservation intervals is $\approx M(t)E\{V\}$, and for data intervals $\approx (1 - p)M(t)E\{X\}$. Since the ratio of these times must be $(1 - \rho)/\rho$ we obtain

$$(1 - \rho)/\rho = (M(t)E\{V\})/((1 - p)M(t)E\{X\}) = E\{V\}/((1 - p)E\{X\})$$

or

$$1 - p = (\rho E\{V\})/((1 - \rho)E\{X\})$$

which using $\lambda = \rho/E\{X\}$, yields $p = (1 - \rho - \lambda E\{V\})/(1 - \rho)$

3.35

Consider a gated all-at-once version of the limited service reservation system. Here there are m users, each with independent Poisson arrival rate λ/μ. Each user has a separate queue, and is allowed to make a reservation for at most one packet in each reservation interval. This packet is then transmitted in the subsequent data interval. The difference with the limited service system of Section 3.5.2 is that here users share reservation and data intervals.

Consider the ith packet arrival into the system and suppose that the user associated with packet i is user j. We have as in Section 3.5.2

$$E\{W_i\} = E\{R_i\} + E\{N_i\}/\mu + (1 + E\{Q_i\} - E\{m_i\})E\{V\}$$

where W_i, R_i, N_i, μ, $E\{V\}$ are as in Section 3.5.2, Q_i is the number of packets in the queue of user j found by packet i upon arrival, and m_i is the number (0 or 1) of packets of user j that will start transmission between the time of arrival of packet i and the end of the frame in which packet i arrives. We have as in Section 3.5.2

$$R = \lim_{i\to\infty} E\{R_i\} + E\{N_i\}/\mu + (1 + E\{Q_i\} - E\{m_i\})E\{V\}$$

$$N = \lim_{i\to\infty} E\{N_i\} = \lambda W$$

$$Q = \lim_{i\to\infty} E\{Q_i\} = \lambda W/m$$

so there remains to calculate $\lim_{i\to\infty} E\{m_i\}$.

There are two possibilities regarding the time of arrival of packet i.

a) Packet i arrives during a reservation interval. This event, call it A, has steady state probability $(1-\rho)$

$$P\{A\} = 1-\rho.$$

Since the ratio of average data interval length to average reservation interval length is $\rho/(1-\rho)$ we see that the average steady state length of a data interval is $\rho E\{V\}/(1-\rho)$. Therefore the average steady state number of packets per user in a data interval is $\rho E\{V\}/((1-\rho)mE\{X\}) = \lambda E\{V\}/((1-\rho)m)$. This also equals the steady state value of $E\{m_i| A\}$ in view of the system symmetry with respect to users

$$\lim_{i\to\infty} E\{m_i| A\} = \frac{\lambda E\{V\}}{(1-\rho)m}.$$

b) Packet i arrives during a data interval. This event, call it B, has steady state probability ρ

$$P\{B\} = \rho.$$

Denote

$$\alpha = \lim_{i\to\infty} E\{m_i \mid B\},$$

$$\alpha_k = \lim_{i\to\infty} E\{m_i \mid B, \text{the data interval of arrival of packet i contains k packets}\}.$$

Assuming $k > 0$ packets are contained in the data interval of arrival, there is equal probability $1/k$ of arrival during the transmission of any one of these packets. Therefore

$$\alpha_k = \sum_{n=1}^{k} \frac{1}{k}\frac{k-n}{m} = \frac{k(k-1)}{2km} = \frac{k-1}{m}.$$

Let P(k) be the unconditional steady-state probability that a nonempty data interval contains k packets, and $E\{k\}$ and $E\{k^2\}$ be the corresponding first two moments. Then we have using Bayes' rule

$\lim_{i\to\infty} P\{\text{The data interval of arrival of packet i contains k packets}\} = kP(k)/E\{k\}$.

Combining the preceding equations we have

$$\alpha = \sum_{k=1}^{m} \frac{kP(k)}{E\{k\}}\alpha_k = \sum_{k=1}^{m} \frac{P(k)k(k-1)}{2E\{k\}m} = \frac{E\{k^2\}}{2mE\{k\}} - \frac{1}{2m}.$$

We have already shown as part of the analysis of case a) above that

$$E\{k\} = \lambda E\{V\}/(1 - \rho)$$

so there remains to estimate $E\{k^2\}$. We have

$$E\{k^2\} = \sum_{k=1}^{m} k^2 P(k)$$

If we maximize the quantity above over the distribution $P(k)$, $k = 0,1,\ldots, m$ subject to the constraints $\sum_{k=1}^{m} kP(k) = E\{k\}$, $\sum_{k=0}^{m} P(k) = 1$, $P(k) \geq 0$ (a simple linear programming problem) we find that the maximum is obtained for $P(m) = E\{k\}/m$, $P(0) = 1 - E\{k\}/m$, and $P(k) = 0$, $k = 1,2,\ldots,m-1$. Therefore

$$E\{k^2\} \leq mE\{k\}.$$

Similarly if we minimize $E\{k^2\}$ subject to the same constraints we find that the minimum is obtained for $P(k'-1) = k' - E\{k\}$, $P(k) = 1 - (k' - E\{k\})$ and $P(k') = 0$ for $k \neq k' - 1, k'$ where k' is the integer for which $k' - 1 \leq E\{k\} < k'$. Therefore

$$E\{k^2\} \geq (k' - 1)^2 (k' - E\{k\}) + (k')^2 [1 - (k' - E\{k\})]$$

After some calculation this relation can also be written

$$E\{k^2\} \geq E\{k\} + (k' - 1)(2E\{k\} - k') \quad \text{for } E\{k\} \in (k' - 1, k'), \quad k' = 1, 2, \ldots, m$$

Note that the lower bound above is a piecewise linear function of $E\{k\}$, and equals $(E\{k\})^2$ at the breakpoints $k' = 1,2,\ldots,m$. Summarizing the bounds we have

$$\frac{E\{k\} + (k' - 1)(2E\{k\} - k')}{2mE\{k\}} - \frac{1}{2m} \leq \alpha \leq \frac{1}{2} - \frac{1}{2m},$$

where k' is the positive integer for which

$$k' - 1 \leq E\{k\} < k'.$$

Note that as $E\{k\}$ approaches its maximum value m (i.e., the system is heavily loaded), the upper and lower bounds coincide. By combining the results for cases a) and b) above we have

$$\lim_{i\to\infty} E\{m_i\} = P\{A\} \lim_{i\to\infty} E\{m_i \mid A\} + P\{B\} \lim_{i\to\infty} E\{m_i \mid B\}$$

$$= (1-\rho)\frac{\lambda E\{V\}}{(1-\rho)m} + \rho\alpha$$

or finally

$$\lim_{i\to\infty} E\{m_i\} = \frac{\lambda E\{V\}}{m} + \rho\alpha$$

where α satisfies the upper and lower bounds given earlier. By taking limit as $i \to \infty$ in the equation

$$E\{W_i\} = E\{R_i\} + E\{N_i\}/\mu + (1 + E\{Q_i\} - E\{m_i\})E\{V\}$$

and using the expressions derived we finally obtain

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

where α satisfies

$$\frac{E\{k\} + (k'-1)(2E\{k\} - k')}{2mE\{k\}} - \frac{1}{2m} \le \alpha \le \frac{1}{2} - \frac{1}{2m},$$

$E\{k\}$ is the average number of packets per data interval

$$E\{k\} = \lambda E\{V\}/(1-\rho)$$

and k' is the integer for which $k' - 1 \le E\{k\} < k'$. Note that the formula for the waiting time W becomes exact in the limit both as $\rho \to 0$ (light load), and as $\rho \to 1 - \lambda E\{V\}/m$ (heavy load) in which case $E\{k\} \to m$ and $\alpha \to 1/2 - 1/2m$. When $m = 1$ the formula for W is also exact and coincides with the one derived for the corresponding single user one-at-a-time limited service system.

3.36

For each session, the arrival rates, average transmission times and utilization factors for the short packets (class 1), and the long packets (class 2) are

$\lambda_1 = 0.25$ packets/sec, $1/\mu_1 = 0.02$ secs, $\rho_1 = 0.005$
$\lambda_2 = 2.25$ packets/sec, $1/\mu_2 = 0.3$ secs, $\rho_2 = 0.675.$

The corresponding second moments of transmission time are

$$E\{X_1^2\} = 0.0004 \qquad E\{X_2^2\} = 0.09.$$

The total arrival rate for each session is $\lambda = 2.5$ packets/sec. The overall 1st and 2nd moments of the transmission time, and overall utilization factors are given by

$$1/\mu = 0.1*(1/\mu_1) + 0.9*(1/\mu_2) = 0.272$$
$$E\{X^2\} = 0.1*E\{X_1^2\} + 0.9*E\{X_2^2\} = 0.081$$
$$\rho = \lambda/\mu = 2.5*0.272 = 0.68.$$

We obtain the average time in queue W via the P - K formula $W = (\lambda E\{X^2\})/(2*(1 - \rho)) = 0.3164$. The average time in the system is $T = 1/\mu + W = 0.588$. The average number in queue and in the system are $N_Q = \lambda W = 0.791$, and $N = \lambda T = 1.47$.

The quantities above correspond to each session in the case where the sessions are time - division multiplexed on the line. In the statistical multiplexing case W, T , N_Q and N are decreased by a factor of 10 (for each session).

In the nonpreemptive priority case we obtain using the corresponding formulas:

$$W_1 = (\lambda_1 E\{X_1^2\} + \lambda_2 E\{X_2^2\})/(2*(1 - \rho_1)) = 0.108$$
$$W_2 = (\lambda_1 E\{X_1^2\} + \lambda_2 E\{X_2^2\})/(2*(1 - \rho_1)*(1 - \rho_1 - \rho_2)) = 0.38$$
$$T_1 = 1/\mu_1 + W_1 = 0.128$$
$$T_2 = 1/\mu_2 + W_2 = 1.055$$
$$N_{Q1} = \lambda_1 * W_1 = 0.027 \qquad N_{Q2} = \lambda_2 * W_2 = 0.855$$
$$N_1 = \lambda_1 * T_1 = 0.032 \qquad N_2 = \lambda_2 * T_2 = 2.273.$$

3.37

(a)

$\lambda = 1/60$ per second
$E(X) = 16.5$ seconds
$E(X^2) = 346.5$ seconds
$$T = E(X) + \lambda E(X^2)/2(1-\lambda E(X))$$
$$= 16.5 + (346.5/60)/2(1- 16.5/60) = 20.48 \text{ seconds}$$

(b) Non-Preemptive Priority

In the following calculation, subscript 1 will imply the quantities for the priority 1 customers and 2 for priority 2 customers. Unsubscripted quantities will refer to the overall system.

$$\lambda = \frac{1}{60} \ , \quad \lambda_1 = \frac{1}{300} \ , \quad \lambda_2 = \frac{1}{75}$$

$$E(X) = 16.5 \ , \ E(X_1) = 4.5 \ , E(X_2) = 19.5$$

$$E(X^2) = 346.5$$

$$R = \frac{1}{2} \lambda E(X^2) = 2.8875$$

$$\rho_1 = \lambda_1 E(X_1) = 0.015$$

$$\rho_2 = \lambda_2 E(X_2) = 0.26$$

$$W_1 = \frac{R}{1-\rho_1} = 2.931$$

$$W_2 = \frac{R}{1-\rho_2} = 4.043$$

$$T_1 = 7.4315, \ T_2 = 23.543$$

$$T = \frac{\lambda_1 T_1 + \lambda_2 T_2}{\lambda} = 20.217$$

(c) Preemptive Queueing

The arrival rates and service rates for the two priorities are the same for preemptive system as the non-preemptive system solved above.

$$E(X_1^2) = 22.5, \ E(X_2^2) = 427.5$$

$$R_1 = \frac{1}{2} \lambda_1 E(X_1^2) = 0.0075$$

$$R_2 = R_1 + \frac{1}{2} \lambda_2 E(X_2^2) = 2.8575$$

$$T_1 = \frac{E(X_1)(1-\rho_1) + R_1}{1-\rho_1}$$

$$T_2 = \frac{E(X_2)(1-\rho_1-\rho_2) + R_2}{(1-\rho_1)(1-\rho_1-\rho_2)}$$

$$T = (\lambda_1 T_1 + \lambda_2 T_2)/\lambda = 19.94$$

3.38

(a) The same derivation as in Section 3.5.3 applies for W_k, i.e.

$$W_k = R/(1 - \rho_1 - \ldots - \rho_{k-1})(1 - \rho_1 - \ldots - \rho_k)$$

where $\rho_i = \lambda_i/(m\mu)$, and R is the mean residual service time. Think of the system as being comprised of a serving section and a waiting section. The residual service time is just the time until any of the customers in the waiting section can enter the serving section. Thus, the residual service time of a customer is zero if the customer enters service immediately because there is a free server at the time of arrival, and is otherwise equal to the time between the customer's arrival, and the first subsequent service completion. Using the memoryless property of the exponential distribution it is seen that

$$R = P_Q E\{\text{Residual service time} \mid \text{queueing occurs}\} = P_Q/(m\mu).$$

(b) The waiting time of classes 1, . . ., k is not influenced by the presence of classes (k+1), . . . , n. All priority classes have the same service time distribution, thus, interchanging the order of service does not change the average waiting time. We have

$$W_{(k)} = \text{Average waiting time for the M/M/m system with rate } \lambda_1 + \ldots + \lambda_k.$$

By Little's theorem we have

Average number in queue of class k = Average number in queue of classes 1 to k
- Average number in queue of classes 1 to k-1

and the desired result follows.

3.39

Let k be such that

$$\rho_1 + \ldots + \rho_{k-1} \le 1 < \rho_1 + \ldots + \rho_k.$$

Then the queue of packets of priority k will grow infinitely, the arrival rate of each priority up to and including k-1 will be accomodated, the departure rate of priorities above k will be zero while the departure rate of priority k will be

$$\tilde{\lambda}_k = \frac{(1 - \rho_1 - \cdots - \rho_{k-1})}{\bar{X}_k}$$

In effect we have a priority system with k priorities and arrival rates

$$\tilde{\lambda}_i = \lambda_i \qquad \text{for } i < k$$

$$\tilde{\lambda}_k = \frac{(1 - \rho_1 - \cdots - \rho_{k-1})}{\bar{X}_k}$$

For priorities $i < k$ the arrival process is Poisson so the same calculation for the waiting time as before gives

$$W_i = \frac{\sum_{i=1}^{k} \tilde{\lambda}_i \overline{X_i^2}}{2(1 - \rho_1 - \cdots - \rho_{i-1})(1 - \rho_1 - \cdots - \rho_i)}, \qquad i < k$$

For priority k and above we have infinite average waiting time in queue.

3.40

(a) The algebraic verification using Eq. (3.79) listed below

$$W_k = R/(1 - \rho_1 - \ldots - \rho_{k-1})(1 - \rho_1 - \ldots - \rho_k)$$

is straightforward. In particular by induction we show that

$$\rho_1 W_1 + \cdots + \rho_k W_k = \frac{R(\rho_1 + \quad + \rho_k)}{1 - \rho_1 - \cdots - \rho_k}$$

The induction step is carried out by verifying the identity

$$\rho_1 W_1 + \cdots + \rho_k W_k + \rho_{k+1} W_{k+1} = \frac{R(\rho_1 + \cdots + \rho_k)}{1 - \rho_1 - \cdots - \rho_k} + \frac{\rho_{k+1} R}{(1 - \rho_1 - \cdots \rho_k)(1 - \rho_1 - \cdots - \rho_{k+1})}$$

The alternate argument suggested in the hint is straightforward.

(b) Cost

$$C = \sum_{k=1}^{n} c_k N_Q^k = \sum_{k=1}^{n} c_k \lambda_k W_k = \sum_{k=1}^{n} \left(\frac{c_k}{\bar{X}_k}\right) \rho_k W_k$$

We know that $W_1 \le W_2 \le \ldots\ldots \le W_n$. Now exchange the priority of two neighboring classes i and j=i+1 and compare C with the new cost

$$C' = \sum_{k=1}^{n} \left(\frac{c_k}{\bar{X}_k}\right) \rho_k W'_k$$

In C' all the terms except $k = i$ and j will be the same as in C because the interchange does not affect the waiting time for other priority class customers. Therefore

$$C'-C = \frac{c_j}{\overline{X}_j}\rho_j W'_j + \frac{c_i}{\overline{X}_i}\rho_i W'_i - \frac{c_i}{\overline{X}_i}\rho_i W_i - \frac{c_j}{\overline{X}_j}\rho_j W_j.$$

We know from part (a) that

$$\sum_{k=1}^{n} \rho_k W_k = \text{constant.}$$

Since W_k is unchanged for all k except $k = i$ and j $(= i+1)$ we have

$$\rho_i W_i + \rho_j W_j = \rho_i W'_i + \rho_j W'_j.$$

Denote

$$B = \rho_i W'_i - \rho_i W_i = \rho_j W_j - \rho_j W'_j$$

Clearly we have $B \geq 0$ since the average waiting time of customer class i will be increased if class i is given lower priority. Now let us assume that

$$\frac{c_i}{\overline{X}_i} \leq \frac{c_j}{\overline{X}_j}$$

Then

$$C'-C = \frac{c_i}{\overline{X}_i}(\rho_i W'_i - \rho_i W_i) - \frac{c_j}{\overline{X}_j}(\rho_j W_j - \rho_j W'_j) = B\left(\frac{c_i}{\overline{X}_i} - \frac{c_j}{\overline{X}_j}\right)$$

Therefore, only if $\frac{c_i}{\overline{X}_i} < \frac{c_{i+1}}{\overline{X}_{i+1}}$ can we reduce the cost by exchanging the priority order of i and i+1. Thus, if (1,2,3,...,n) is an optimal order we must have

$$\frac{c_1}{\overline{X}_1} \geq \frac{c_2}{\overline{X}_2} \geq \frac{c_3}{\overline{X}_3} \geq \ldots \geq \frac{c_n}{\overline{X}_n}$$

3.41

Let D(t) and $T_i(t)$ be as in the solution of Problem 3.31. The inequality in the hint is evident from Figure 3.30, and therefore it will suffice to show that

$$\lim_{t\to\infty} \frac{1}{t}\sum_{i\in D(t)} T_i = \lim_{t\to\infty} \frac{1}{t}\sum_{i=1}^{\alpha(t)} T_i \qquad (1)$$

We first show that

$$T_k/t_k \to 0 \qquad \text{as } k \to \infty \qquad (2)$$

where t_k is the arrival time of the kth customer. We have by assumption

$$\lim_{k\to\infty} (k/t_k) = \lambda,$$

and by taking the limit as $k\to\infty$ in the relation

$$(k+1)/t_{k+1} - k/t_k = 1/t_{k+1} - ((t_{k+1} - t_k)/t_{k+1})(k/t_k)$$

we obtain

$$t_k/t_{k+1} \to 1 \qquad \text{as } k \to \infty \qquad (3)$$

We also have

$$\frac{\sum_{i=1}^{k} T_i}{t_k} = \frac{k}{t_k}\frac{\sum_{i=1}^{k} T_i}{k} \to \lambda T \qquad \text{as } k \to \infty \qquad (4)$$

so

$$\frac{\sum_{i=1}^{k+1} T_i}{t_{k+1}} - \frac{\sum_{i=1}^{k} T_i}{t_k} \to 0$$

or

$$\frac{T_{k+1}}{t_{k+1}} + \frac{\sum_{i=1}^{k} T_i}{t_k}\left(\frac{t_k}{t_{k+1}} - 1\right) \to 0$$

which proves (2).

Let $\varepsilon > 0$ be given. Then, from (2), there exists k such that $T_i < t_i\,\varepsilon$ for all $i > k$. Choose t large enough so that $\alpha(t) > k$. Then

$$\sum_{i=1}^{\beta(t)} M_i \leq \int_0^t r(\tau)d\tau \leq \sum_{i=1}^{\alpha(t)} M_i$$

or

$$\frac{\beta(t)}{t}\frac{\sum_{i=1}^{\beta(t)} M_i}{\beta(t)} \leq \frac{1}{t}\int_0^t r(\tau)d\tau \leq \frac{\alpha(t)}{t}\frac{\sum_{i=1}^{\alpha(t)} M_i}{\alpha(t)}$$

Under the assumptions

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

$$M = \lim_{k\to\infty}\frac{1}{k}\sum_{i=1}^{k} M_i$$

we have

$$R = \lambda M$$

where

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

is the time average rate at which the system earns.

(b) Take $r_i(t) = 1$ for all t for which customer i resides in the system, and $r_i(t) = 0$ for all other t.

(c) If X_i and W_i are the service and queueing times of the ith customer we have

$$M_i = X_i W_i + (X_i^2)/2$$

where the two terms on the right above correspond to payment while in queue and service respectively. Applying part (a) while taking expected values and taking the limit as $i \to \infty$, and using the fact that X_i and W_i are independent we obtain

$$U = \lambda(E\{X\}W + E\{(X_i^2)/2\})$$

where U, the rate at which the system earns, is the average unfinished work in the system. Since the arrival process is Poisson, U is also the average unfinished work seen by an arriving customer. Therefore U = W, and the P - K formula follows.

3.43

We have similar to Section 3.5

$$W = R + \rho W + \bar{W}_B \qquad (1)$$

where the mean residual service time is

$$R = \frac{\lambda \bar{n} \overline{X^2}}{2}$$

We derive the average waiting time of a customer for other customers that arrived in the same batch

$$\bar{W}_B = \sum_j r_j E\{W_B \mid \text{batch has size } j\}$$

where

P_j = Probability a batch has size j
r_j = Proportion of customers arriving in a batch of size j

We have

$$r_j = \frac{jP_j}{\sum_{n=1}^{\infty} nP_n} = \frac{jP_j}{\bar{n}}$$

Also since the customer is equally likely to be in any position within the batch

$$E\{W_B \mid \text{batch is of size } j\} = \sum_{k=1}^{j} (k-1)\bar{X}\frac{1}{j} = \frac{j-1}{2}\bar{X}$$

Thus from (2)

$$E\{W_B\} = \sum_j \frac{jP_j(j-1)\bar{X}}{2\bar{n}} = \frac{\bar{X}(\overline{n^2} - \bar{n})}{2\bar{n}}$$

Substituting in (1) we obtain

$$W = \frac{R}{1-\rho} + \frac{\bar{W}_B}{1-\rho}$$

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

3.44

(a) Let p_0 be the steady state probability that an arriving packet finds the system empty. Then, in effect, a packet has service time X with probability 1 - p_0, and service time X + Δ with probability p_0. The fraction of time the system is busy is obtained by applying Little's Theorem to the service portion of the system

$$\lambda[E\{X\}(1 - p_0) + (E\{X\} + E\{\Delta\})p_0] = \lambda(E\{X\} + E\{\Delta\}p_0)$$

This is also equal to P{system busy} $\doteq$ 1 - p_0, so by solving for p_0, we obtain

$$p_0 = (1 - \lambda E\{X\})/(1 + \lambda E\{\Delta\}) = (1 - \rho)/(1 + \lambda E\{\Delta\})$$

where $\rho = \lambda E\{X\}$.

(b) Let

E{I} = average length of an idle period
E{B} = average length of a busy period

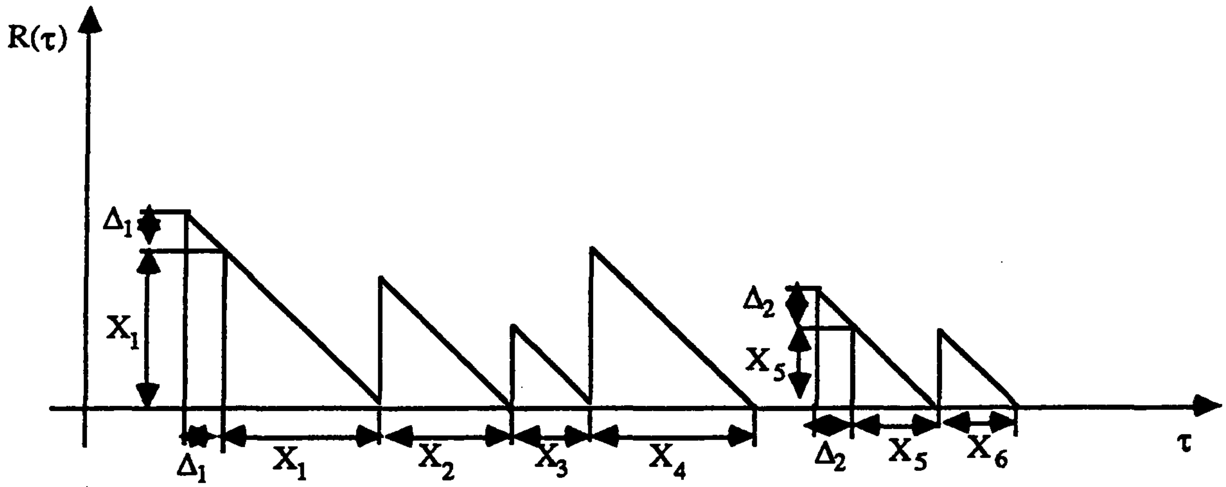
Since the arrival process is Poisson we have

E{I} = 1/λ = average time betwen last departure in a busy period and the next arrival in the system

$$p_0 = \frac{E\{I\}}{E\{I\} + E\{B\}} = \frac{1/\lambda}{1/\lambda + E\{B\}} = \frac{1 - \lambda E\{X\}}{1 + \lambda E\{\Delta\}}$$

$$E\{B\} = \frac{E\{X\} + E\{\Delta\}}{1 - E\{X\}\lambda} = \frac{E\{X\} + E\{\Delta\}}{1 - \rho}$$

(c) We calculate the mean residual time using a graphical argument



From the figure we have

$$\int_0^t R(\tau)d\tau = \sum_{i=1}^{M(t)} \frac{1}{2}X_i^2 + \sum_{i=1}^{N(t)} X_{j(i)}\Delta_i + \frac{1}{2}\Delta_i^2$$

where $X_{j(i)}$ is the service time of the first packet of the ith busy period, and

M(t) = # of arrivals up to t
N(t) = # of busy periods up to t

Taking the limit as $t \to \infty$, and using the fact

$$\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1 - p_0}{E\{B\}} = \frac{\lambda(1-\rho)}{1+\lambda E\{\Delta\}}$$

we obtain

$$R = \lim_{t\to\infty} \frac{\int_0^t R(\tau)d\tau}{t} = \lim_{t\to\infty}\left[\frac{M(t)}{t}\frac{\sum_{i=1}^{M(t)} \frac{1}{2}X_i^2}{M(t)} + \frac{N(t)}{t}\frac{\sum_{i=1}^{N(t)} X_{j(i)}\Delta_i + \frac{1}{2}\Delta_i^2}{N(t)}\right]$$

We have, as in Section 3.5, $W = R + \rho W$ or

$$W = R/(1 - \rho)$$

Substituting the expression for R obtained previously we obtain

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

3.45

(a) It is easy to see that

Pr (system busy serving customers) = ρ

Pr (system idle) = 1-ρ = P(0 in system) + P(1 in idle system) + ...
+ P(k-1 in idle system)

It can be seen that

P(0 in system) = P(1 in idle system) = ... = P(k-1 in idle system) = (1-ρ)/k

implying that

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

(b) A busy period in this problem is the length of time the system is nonempty. The end of each busy period is a renewal point for the system. Between two renewal points, the system is either waiting (with 0 or more customers) or serving.

Let $\overline{W}$ be the expected length of a waiting period.
Since arrivals are poisson, we have

$$\overline{W} = \frac{k}{\lambda} .$$

Let $\overline{S}$ be the expected length of a serving period.
Then the probability that the system is serving = $\rho = \frac{\overline{S}}{\overline{S} + \overline{W}}$
implying that

$$\frac{1}{\rho} - 1 = \frac{\overline{W}}{\overline{S}}$$

or

$$\overline{S} = \frac{\rho \overline{W}}{1-\rho} = \frac{\rho \frac{k}{\lambda}}{1-\rho}$$

Let $\bar{I}$ be the expected length of time the system is empty.

The expected length of a busy period $= \bar{S} + \bar{W} - \bar{I}$

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

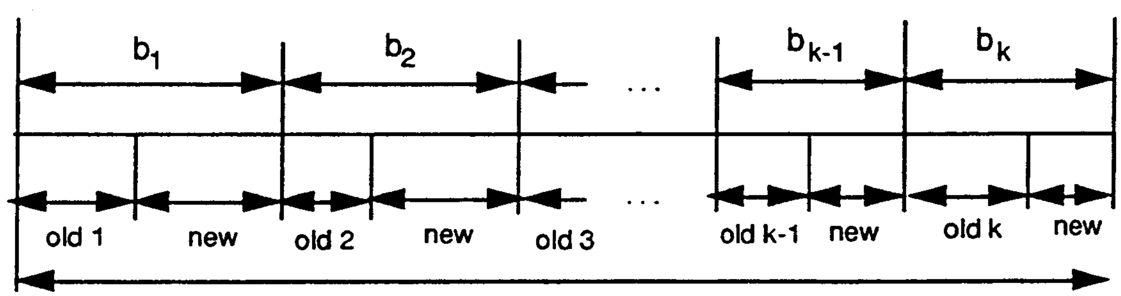
$$= \frac{\rho\, k + (1-\rho)\,(k-1)}{\lambda(1-\rho)} = \frac{k + \rho - 1}{\lambda\,(1-\rho)}$$

$\frac{\rho\,(k/\lambda)}{1-\rho}$ is k times the average length of an M/G/1 busy period and $\frac{k-1}{\lambda}$

is the average time from the first arrival until the system starts serving.

(c) We will call the k packets that are in the system at the beginning of the busy period "old" packets. Those that come during the busy period (and are therefore served during this period) are called "new" packets.

We consider the following service discipline: the server transmits old packets only when it doesn't have new packets available at that moment (so it's not FCFS). Since this discipline doesn't depend on the length of the packets it will give the same average number of packets in the system. Thus a busy period looks as illustrated below:



Busy period

In a subperiod b_i of the busy period, the old packet i and the new packets combine to give the same distribution as a normal M/G/1 busy period except that there are an extra k-i old packets in the system. It is easy to see that the distribution of the length of $b_1, b_2, \ldots b_k$ is the same since each of them is exactly like an M/G/1 busy period.

$$\Rightarrow E(N \mid \text{serving}) = E(N \mid b_1)\, P(b_1 \mid \text{serving}) + \ldots + E(N \mid b_k)\, P(b_k \mid \text{serving})$$

$$P(b_i \mid \text{serving}) = \frac{1}{k}$$

$$E(N \mid b_i) = E(N_{M/G/1} \mid \text{busy}) + k\text{-}i$$

implying that

$$E(N \mid \text{serving}) = \frac{1}{k}\left(k\ E(N_{M/G/1} \mid \text{busy}) + \sum_{i=1}^{k}(k\text{-}i)\right)$$

$$= \frac{k\text{-}1}{2} + E(N_{M/G/1} \mid \text{busy})$$

We have

$$E(N_{M/G/1}) = E(N_{M/G/1} \mid \text{busy})\ \rho$$

from which

$$E(N_{M/G/1} \mid \text{busy}) = \frac{E(N_{M/G/1})}{\rho}$$

or

$$E(N \mid \text{serving}) = \frac{E(N_{M/G/1})}{\rho} + \frac{k\text{-}1}{2}$$

Also

$$E(N \mid \text{busy waiting}) = E(N \mid \text{waiting with } 1)\ P(\text{waiting with } 1 \mid \text{busy waiting}) + \ldots$$
$$+ E(N \mid \text{waiting with } k\text{-}1)\ P(\text{waiting with } k\text{-}1 \mid \text{busy waiting})$$

$$P(\text{waiting with } i \mid \text{busy waiting})$$
$$= P(\text{waiting with } j \mid \text{busy waiting}) = \frac{1}{k\text{-}1} \quad \text{for all } 0 < i, j < k$$

from which

$$E(N \mid \text{busy waiting}) = \frac{1}{k\text{-}1}\sum_{i=1}^{k\text{-}1} i \quad = \frac{k}{2}$$

(d) $E(N) = E(N \mid \text{busy waiting})\ P(\text{busy waiting}) + E(N \mid \text{busy serving})\ P(\text{busy serving})$

$$= \frac{k}{2}\ \frac{(k\text{-}1)\ (1\text{-}\rho)}{k} + \left(\frac{E(N_{M/G/1})}{\rho} + \frac{k\text{-}1}{2}\right)\rho$$

$$= E(N_{M/G/1}) + \frac{k\text{-}1}{2}$$

3.46

We have

$$W = R/(1 - \rho)$$

where

$$R = \lim_{t\to\infty}\left\{\frac{1}{t}\sum_{i=1}^{M(t)}\frac{1}{2}X_i^2 + \frac{1}{t}\sum_{i=1}^{L(t)}\frac{V_i^2}{2}\right\}$$

where L(t) is the number of vacations (or busy periods) up to time t. The average length of an idle period is

$$I = \int_0^\infty p(v)\left[\int_0^V v\lambda e^{-\lambda\tau}d\tau + \int_V^\infty \tau\lambda e^{-\lambda\tau}d\tau\right]dv$$

and it can be seen that the steady-state time average number of vacations per unit time

$$\lim_{t\to\infty}\frac{L(t)}{t} = \frac{1-\rho}{I}$$

We have

$$\lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{L(t)}\frac{V_i^2}{2} = \lim_{t\to\infty}\frac{L(t)}{t}\frac{\sum_{i=1}^{L(t)}\frac{V_i^2}{2}}{L(t)} = \lim_{t\to\infty}\frac{L(t)}{t}\frac{\overline{V^2}}{2I} = \frac{\overline{V^2}(1-\rho)}{2I}$$

Therefore

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

and

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

3.47

(a) Since arrival times and service times are independent, the probability that there was an arrival in a small interval δ at time τ - x and that this arrival is still being served at time τ is $\lambda\delta[1 - F_X(x)]$.

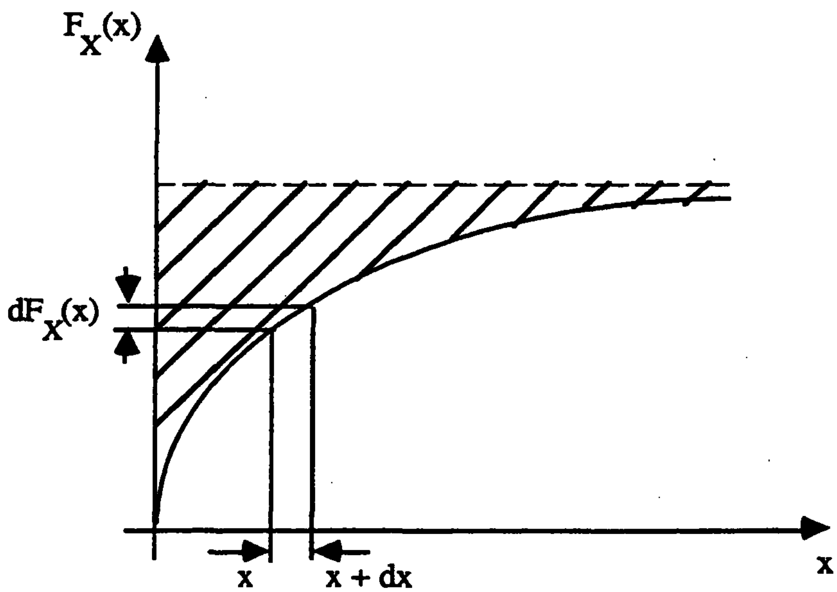
(b) We have

$$\overline{X} = \int_0^\infty x dF_X(x)$$

and by calculating the shaded area of the figure below in two different ways we obtain

$$\int_0^\infty x dF_X(x) = \int_0^\infty [1 - F_X(x)]dx$$

This proves the desired expression.



(c) Let $p_n(x)$ be the steady state probability that the number of arrivals that occurred prior to time $\tau - x$ and are still present at time τ is exactly n.

For $n \geq 1$ we have

$$p_n(x - \delta) = \{1 - \lambda[1 - F_X(x)]\delta\}p_n(x) + \lambda[1 - F_X(x)]\delta p_{n-1}(x)$$

and for n =0 we have

$$p_0(x - \delta) = \{1 - \lambda[1 - F_X(x)]\delta\}p_0(x).$$

Thus $p_n(x)$, n = 0, 1, 2, . . . are the solution of the differential equations

$$dp_n/dx = a(x)p_n(x) - a(x)p_{n-1}(x) \qquad \text{for } n \geq 1$$

$$dp_0/dx = a(x)p_0(x) \qquad \text{for } n = 0$$

where

$$a(x) = \lambda[1 - F_X(x)].$$

Using the known conditions

$$p_n(\infty) = 0 \qquad \text{for } n \geq 1$$
$$p_0(\infty) = 1$$

it can be verified by induction starting with n = 0 that the solution is

$$p_n(x) = [e^{-\int_x^\infty a(y)dy}]\frac{[\int_x^\infty a(y)dy]^n}{n!}, \qquad x \geq 0,\ n = 0, 1, 2, \ldots$$

Since

$$\int_0^\infty a(y)dy = \lambda\int_0^\infty [1 - F_X(x)]dy = \lambda E\{X\}$$

we obtain

$$p_n(0) = e^{-\lambda E\{X\}}\frac{[\lambda E\{X\}]^n}{n!}, \qquad n = 0, 1, 2, \ldots$$

Thus the number of arrivals that are still in the system have a steady state Poisson distribution with mean $\lambda E\{X\}$.

3.48

(a) Denote

$$f(x) = E_r[(\max\{0,r-x\})^2]$$

and

$$g(x) = (E_r[\max\{0,r-x\}])^2,$$

where $E_r[\cdot]$ denotes expected value with respect to r (x is considered constant). We will prove that f(x)/g(x) is montonically nondecreasing for x nonnegative and thus attain its minimum value for x=0. We have

$$\frac{\partial f(x)}{\partial x} = E_r\left[\frac{\partial}{\partial x}(\max\{0,r-x\})^2\right] = 2E_r\left[\max\{0,r-x\}\cdot \frac{\partial}{\partial x}(\max\{0,r-x\})\right],$$

where

$$\frac{\partial(\max\{0,r-x\})}{\partial x} = -\, u(r-x)$$

where $u(\cdot)$ is the step function. Thus

$$\frac{\partial f(x)}{\partial x} = -2E_r\,[(\max\{0,r-x\})\cdot u(r-x)] = -2E_r[\max\{0,r-x\}]$$

Assume for simplicity that r has a probability density function (the solution is similar in the more general case). Then

$$\frac{\partial g(x)}{\partial x} = 2E_r[\max\{0,r-x\}]\, E_r\left[\frac{\partial}{\partial x}\max\{0,r-x\}\right] = -2E_r[\max\{0,r-x\}]\cdot \int_{r>x} p(r)dr$$

Thus

$$\frac{\partial f(x)}{\partial x}g(x) - f(x)\frac{\partial g(x)}{\partial x} = 2E_r[\max\{0,r-x\}]E_r\left[(\max\{0,r-x\})^2\right]\int_{r>x} p(r)dr$$

$$- 2E_r[\max\{0,r-x\}]E_r([\max\{0,r-x\}])^2$$

For $\frac{f(x)}{g(x)}$ monotonically nondecreasing we must have $\frac{\partial f(x)}{\partial x}g(x) - f(x)\frac{\partial g(x)}{\partial x} \geq 0$ or equivalently

$$E_r[(\max\{0,r-x\})^2]\int_{r>x} p(r)dr - (E_r[\max\{0,r-x\}])^2$$

$$= \int_{r>x}(r-x)^2 p(r)dr\int_{r>x} p(r)dr - \left(\int_{r>x}(r-x)p(r)dr\right)^2 \geq 0$$

which is true by Schwartz's inequality. Thus the ratio

$$\frac{f(x)}{g(x)} = \frac{E_r[(\max\{0,r-x\})^2]}{(E_r[\max\{0,r-x\}])^2}$$

is monotonically nondecreasing and attains its minimum value at x=0. On the other hand, we have

$$\frac{f(0)}{g(0)} = \frac{E(r^2)}{[E(r)]^2},$$

since $r \geq 0$, and the result follows.

(b) We know (cf. Eq. (3.93)) that

$$I_k = -\min\{0, W_k + X_k - \tau_k\} = \max\{0, \tau_k - W_k - X_k\}$$

$$= \max\{0, \tau_k - S_k\},$$

where S_k is the time in the system. Since the relation in part (a) holds for any nonnegative scalar x, we can take expected values with respect to x as well, and use the fact that for any function of x, g(x), $E(g(x)^2) \geq E^2(g(x))$, and find that

$$E_{r,x}[(\max\{0,r-x\})^2] \geq \frac{\overline{r^2}}{(\overline{r}^2)} (E_{r,x}[\max\{0,r-x\}])^2, \tag{1}$$

where x is considered to be a random variable this time. By letting $r = \tau_k$, $x = S_k$, and $k \to \infty$, we find from (1) that

$$\overline{I^2} \geq \frac{\overline{\tau^2}}{(\bar{\tau})^2} (\bar{I})^2$$

or

$$\frac{\overline{I^2} - (\bar{I})^2}{(\bar{I})^2} \geq \frac{\overline{\tau^2} - (\bar{\tau})^2}{(\bar{\tau})^2}$$

or

$$\sigma_I^2 \geq \frac{(\bar{I})^2}{(\bar{\tau})^2} \sigma_a^2 \quad \text{(since } \sigma_a^2 \text{ is defined as the variance of interarrival times)}$$

Since $\bar{I} = \frac{1-\rho}{\lambda}$ and $\bar{\tau} = \frac{1}{\lambda}$ we get

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

By using Eq. (3.97), we then obtain

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

3.49

(a) Since the arrivals are Poisson with rate λ, the mean time until the next arrival starting from any given time (such as the time when the system becomes empty) is $1/\lambda$. The time average fraction of busy time is $\lambda E[X]$. This can be seen by Little's theorem applied to the service facility (the time average number of customers in the server is just the time average fraction of busy time), or it can be seen by letting $\sum_{i=1}^{n} X_i$ represent the time the server is busy with the first n customers, dividing by the arrival time of the n^{th} customer, and going to the limit.

Let E[B] be the mean duration of a busy period and $E[I] = 1/\lambda$ be the mean duration of an idle period. The time average fraction of busy time must be E[B]/(E[B]+E[I]). Thus

$$\lambda E[X] = E[B]/(E[B]+1/\lambda)\,; \qquad E[B] = \frac{E[X]}{1 - \lambda E[X]}$$

This is the same as for the FCFS M/G/1 system (Problem 3.30).

(b) If a second customer arrives while the first customer in a busy period is being served, that customer (and all subsequent customers that arrive while the second customer is in the system) are served before the first customer resumes service. The same thing happens for any subsequent customer that arrives while the first customer is actually in service. Thus when the first customer leaves, the system is empty. One can view the queue here as a stack, and the first customer is at the bottom of the stack. It follows that E[B] is the expected system time given a customer arriving to an empty system.

The customers already in the system when a given customer arrives receive no service until the given customer departs. Thus the system time of the given customer depends only on its own service time and the new customers that arrive while the given customer is in the system. Because of the memoryless property of the Poisson arrivals and the independence of service times, the system time of the given customer is independent of the number of customers (and their remaining service times) in the system when the given customer arrives. Since the expected system time of a given customer is independent of the number of customers it sees upon arrival in the system, the expected time is equal to the expected system time when the given customer sees an empty system; this is E[B] as shown above.

(c) Given that a customer requires 2 units of service time, look first at the expected system time until 1 unit of service is completed. This is the same as the expected system time of a customer requiring one unit of service (i.e., it is one unit of time plus the service time of all customers who arrive during that unit and during the service of other such customers). When one unit of service is completed for the given customer, the given customer is in service with one unit of service still required, which is the same as if a new customer arrived requiring one unit of service. Thus the given customer requiring 2 units of service has an expected system time of 2C. Extending the argument to a customer requiring n units of service, the expected system time is nC. Doing the argument backwards for a customer requiring 1/n of service, the expected system time is C/n. We thus conclude that E[system time | X=x] = Cx.

(d) We have

$$E[B] = \int_0^{\infty} Cx\, dF(x) = CE[X]; \qquad C = \frac{1}{1 - \lambda E[X]}$$

3.50

(a) Since $\{p_j\}$ is the stationary distribution, we have for all $j \in S$

$$p_j\left(\sum_{i \in \bar{S}} q_{ji} + \sum_{i \notin \bar{S}} q_{ji}\right) = \sum_{i \in \bar{S}} p_i q_{ij} + \sum_{i \notin \bar{S}} p_i q_{ij}$$

Using the given relation, we obtain for all $j \in \bar{S}$

$$p_j \sum_{i \in \bar{S}} q_{ji} = \sum_{i \in \bar{S}} p_i q_{ij}$$

Dividing by $\sum_{i \in \bar{S}} p_i$, it follows that

$$\bar{p}_j \sum_{i \in \bar{S}} q_{ji} = \sum_{i \in \bar{S}} \bar{p}_i q_{ij}$$

for all $j \in \bar{S}$, showing that $\{\bar{p}_j\}$ is the stationary distribution of the truncated chain.

(b) If the original chain is time reversible, we have $p_j q_{ji} = p_i q_{ij}$ for all i and j, so the condition of part (a) holds. Therefore, we have $\bar{p}_j q_{ji} = \bar{p}_i q_{ij}$ for all states i and j of the truncated chain.

(c) The finite capacity system is a truncation of the two independent M/M/1 queues system, which is time reversible. Therefore, by part (b), the truncated chain is also time reversible. The formula for the steady state probabilities is a special case of Eq. (3.39) of Section 3.4.

3.51

(a) Since the detailed balance equations hold, we have

$$p_j q_{ji} = p_i q_{ij}$$

Thus for $i, j \in S_k$, we have

$$\frac{p_j}{u_k} q_{ji} = \frac{p_i}{u_k} q_{ij} \Leftrightarrow \pi_j q_{ji} = \pi_i q_{ij}$$

and it follows that the π_i, $i \in S_k$ satisfy the detailed balance equations. Also

$$\sum_{i\in S_k}\pi_i = \frac{\sum_{i\in S_k} p_i}{u_k} = \frac{u_k}{u_k} = 1$$

Therefore, $\{\pi_i \mid i \in S_k\}$ as defined above, is the stationary distribution of the Markov chain with state space S_k.

(b) Obviously

$$\sum_{k=1}^{K} u_k = \sum_{k=1}^{K}\sum_{j\in S_k} p_j = 1. \tag{1}$$

Also we have

$$u_k\tilde{q}_{km} = \sum_{\substack{j\in S_k\\ i\in S_m}} \pi_j q_{ji} u_k,$$

which in view of the fact $\pi_j u_k = p_j$ for $j \in S_k$, implies that

$$u_k\tilde{q}_{km} = \sum_{\substack{j\in S_k\\ i\in S_m}} p_j q_{ji} \tag{2}$$

and

$$u_m\tilde{q}_{mk} = \sum_{\substack{j\in S_m\\ i\in S_k}} \pi_j q_{ji} u_m = \sum_{\substack{j\in S_m\\ i\in S_k}} q_{ji} p_j = \sum_{\substack{i\in S_m\\ j\in S_k}} q_{ij}\, p_j.$$

(3)

Since the detailed balance equations $p_j q_{ji} = q_{ij} p_i$ hold, we have

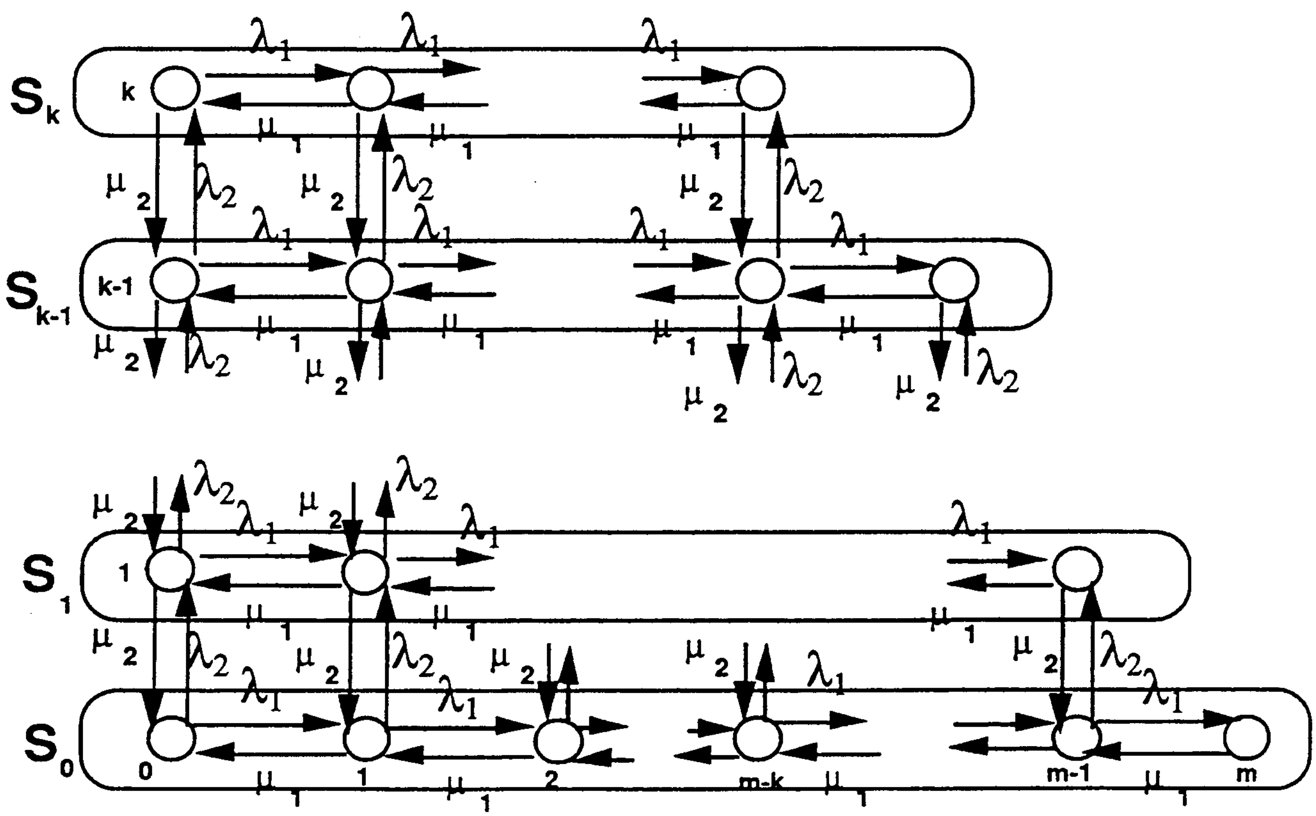
$$\sum_{\substack{j\in S_k\\ i\in S_m}} p_j q_{ji} = \sum_{\substack{i\in S_m\\ j\in S_k}} q_{ij} p_i \tag{4}$$

Equations (2)-(4) give

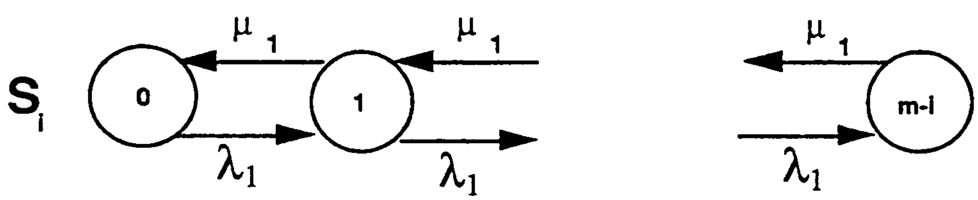
$$u_k\tilde{q}_{km} = u_m\tilde{q}_{mk}. \tag{5}$$

Equations (1) and (5) imply that $\{u_k \mid k=1,\ldots,K\}$ is the stationary distribution of the Markov chain with states $1,\ldots,k$, and transition rates $\tilde{q}_{km}$.

(c) We will deal only with Example 3.13. (Example 3.12 is a special case with k=m).



For i =0,1,...,k we define S_i as the set of states (i,0), (i,1),..., (i,m-i), (see Figure 1). Then the truncated chain with state space S_i is



We denote by $\pi_j^{(i)}$ the stationary probability of state j of the truncated chain S_i and we let

$$\rho_1 = \frac{\lambda_1}{\mu_1}.$$

Then

$$\pi_{j+1}^{(i)} = \rho_1 \pi_j^{(i)}$$

Thus

$$\pi_0^{(i)} \sum_{j=0}^{m-i} \rho_1^j = 1.$$

or

$$\pi_0^{(i)} = \frac{1-\rho_1}{1-\rho_1^{m-i+1}}$$

Therefore,

$$\pi_j^{(i)} = \frac{1-\rho_1}{1-\rho_1^{m-i+1}} \rho_1^j \qquad i=0,1,2,\dots,k, \quad j=0,1,2,\dots,m-i$$

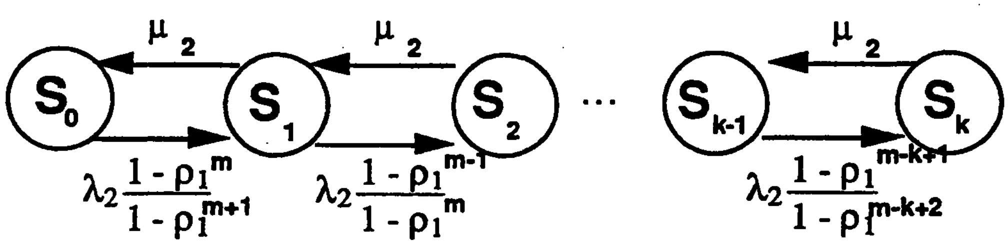
The transition probabilities of the aggregate chain are

$$\tilde{q}_{l,\,l+1} = \sum_{j\in S_l, i\in S_{l+1}} \pi_j^{(l)} q_{ji} = \sum_{j=0}^{m-l-1} \pi_j^{(l)} \lambda_2 = \lambda_2(1-\pi_{m-l}^{(l)})$$

$$= \lambda_2\left(1-\frac{1-\rho_1}{1-\rho_1^{m-l+1}}\rho_1^{m-l}\right) = \lambda_2 \frac{1-\rho_1^{m-l}}{1-\rho_1^{m-l+1}}$$

$$\tilde{q}_{l+1,\,l} = \sum_{\substack{j\in S_{l+1}\\ i\in S_l}} \pi_j^{(l+1)} q_{ji} = \mu_2 \sum_{j\in S_{l+1}} \pi_j^{(l+1)} = \mu_2.$$

The aggregate chain is given by the following figure.



Thus we have

$$u_{l+1} = \rho_2 \cdot \frac{1-\rho_1^{m-l}}{1-\rho_1^{m-l+1}} u_l$$

from which

$$u_l = \rho_2^{l_2} \Pi_{j=0}^{l-1} \left(\frac{1-\rho_1^{m-j}}{1-\rho_1^{m-j+1}} \right) = \rho_2 \frac{1-\rho_1^{m-l+1}}{1-\rho_1^{m+1}} u_0$$

Furthermore, we have

$$\sum_{l=0}^{k} u_l = 1,$$

or

$$u_0 \sum_{l=0}^{k} \rho_2^l \frac{1-\rho_1^{m-l+1}}{1-\rho_1^{m+1}} = 1$$

from which

$$u_0 = \frac{1-\rho_1^{m+1}}{\frac{1-\rho_2^{k+1}}{1-\rho_2} - \rho_1^{m+1} \cdot \frac{1-\left(\frac{\rho_2}{\rho_1}\right)^{k+1}}{1-\frac{\rho_2}{\rho_1}}}$$

and

$$u_l = u_0 \, \rho_2^l \frac{1-\rho_1^{m-l+1}}{1-\rho_1^{m+1}}$$

Thus

$$p(n_1 n_2) = \pi_{n_1}^{(n_2)} u_{n_2} = \frac{u_0 \rho_2^{n_2} (1-\rho_1) \rho_1^{n_1}}{1-\rho_1^{m+1}}$$

from which we obtain the product form

$$p(n_1 n_2) = \left(\frac{1-\rho_2^{k+1}}{1-\rho_2} - \rho_1^{m+1} \frac{1-\left(\frac{\rho_2}{\rho_1}\right)^{k+1}}{1-\frac{\rho_2}{\rho_1}} \right)^{-1} (1-\rho_1) \rho_1^{n_1} \rho_2^{n_2}$$

(d) We are given that the detailed balance equations hold for the truncated chains. Thus for $i, j \in S_k$, we have $p_i q_{ij} = p_j q_{ji}$. Furthermore,

$$\sum_{i\in S_k}\pi_i = \sum_{i\in S_k}\frac{p_i}{u_k} = \frac{u_k}{u_k} = 1$$

Thus $\{\pi_j \mid j\in S_k\}$ is the distribution for the truncated chain S_k and the result of part (a) holds.

To prove the result of part (b), we have to prove that the global balance equations hold for the aggregate chain, i.e.,

$$\sum_{m=1}^{k}\tilde{q}_{km}u_k = \sum_{m=1}^{k} u_m\tilde{q}_{mk},$$

or equivalently

$$\sum_{m=1}^{k}\sum_{j\in S_k, i\in S_m} \pi_j u_k q_{ji} = \sum_{m=1}^{k}\sum_{j\in S_k, i\in S_m} q_{ij}\pi_i u_m$$

For $j \in S_k$, we have $\pi_j u_k = p_j$, and for $i \in S_m$, we have $\pi_i u_m = p_i$, so we must show

$$\sum_{m=1}^{k}\sum_{j\in S_k, i\in S_m} p_j q_{ji} = \sum_{m=1}^{k}\sum_{j\in S_k, i\in S_m} q_{ij} p_i$$

or

$$\sum_{j\in S_k}\sum_{\text{all } i} p_j q_{ji} = \sum_{j\in S_k}\sum_{\text{all } i} p_i q_{ij} \tag{6}$$

Since $\{p_i\}$ is the distribution of the original chain, the global balance equations

$$\sum_{\text{all } i} p_j q_{ji} = \sum_{\text{all } i} p_i q_{ij}$$

By summing over all $j\in S_k$, we see that Eq. (6) holds. Since

$$\sum_{k=1} u_k = 1$$

and we just proved that the u_k's satisfy the global balance equations for the aggregate chain, $\{u_k \mid k = 1,\ldots,K\}$ is the distribution of the aggregate chain. This proves the result of part (b).

3.52

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. In particular, $t_2 - t_1$ is independent of the (reversed system) arrival process to the left of t_2. In forward system terms, this means that $t_2 - t_1$ is independent of the departure process to the left of t_2.

3.53

(a) If customers are served in the order they arrive then given that a customer departs at time t from queue 1, the arrival time of that customer at queue 1 (and therefore the time spent at queue 1), is independent of the departure process from queue 1 prior to t. Since the departures from queue 1 are arrivals at queue 2, we conclude that the time spent by a customer at queue 1 is independent of the arrival times at queue 2 prior to the customer's arrival at queue2. These arrival times, together with the corresponding independent (by Kleinrock's approximation) service times determine the time the customer spends at queue 2 and the departure process from queue 2 before the customer's departure from queue 2.

(b) Suppose packets with mean service time $1/\mu$ arrive at the tandem queues with some rate λ which is very small ($\lambda << \mu$). Then, the apriori probability that a packet will wait in queue 2 is very small.

Assume now that we know that the waiting time of a packet in queue 1 was nonzero. This information changes the aposteriori probability that the packet will wait in queue 2 to at least 1/2 (because its service time in queue 1 will be less than the service time at queue 2 of the packet in front of it with probability 1/2). Thus the knowledge that the waiting time in queue 1 is nonzero, provides a lot of information about the waiting time in queue 2.

3.54

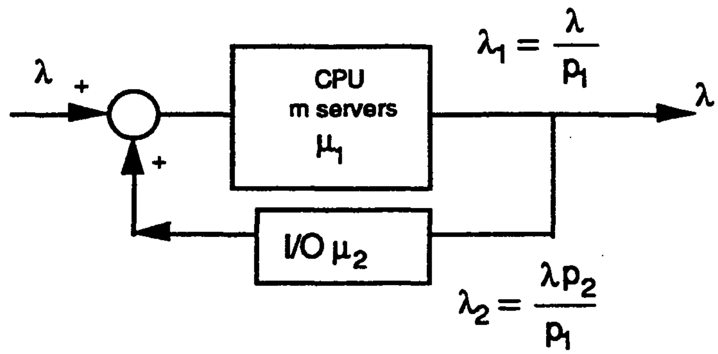
It can be verified by checking the detailed balance equations that the M/M/1/m queue is reversible. Hence the arrival and departure process are the same. The arrival process is Poisson but with probability p_m an arrival does not enter the system. Also since an external arrival process is independent of the state of the system, the arrival process to the sytem is still Poisson but with rate $\lambda(1 - p_m)$. By reversibility we conclude that the departure process is also Poisson with rate $\lambda(1 - p_m)$.

3.55

Let

$$\rho_1 = \frac{\lambda_1}{\mu_1}$$

$$\rho_2 = \frac{\lambda_2}{\mu_2}$$



Using Jackson's Theorem and Eqs. (3.34)-(3.35) we find that

$$P(n_1,n_2) = \begin{cases} p_0 \dfrac{(m\rho_1)^{n_1}}{n_1!} \cdot \rho_2^{n_2}(1-\rho_2), & n_1 \le m \\ p_0 \dfrac{m^m \rho_1^{n_1}}{m!} \rho_2^{n_2}(1-\rho_2), & n_1 > m \end{cases}$$

where

$$p_0 = \left[\sum_{n=0}^{m-1} \frac{(m\rho_1)^n}{n!} + \frac{(m\rho_1)^m}{m!(1-\rho_1)} \right]^{-1}$$

3.56

(a) We have

$$P(X_n=i) = (1-\rho)\rho^i; \quad i \ge 0; \ \rho = \lambda/\mu$$

$$\begin{aligned} P(X_n=i, D_n=j) = P(D_n=j \mid X_n=i)\,P(X_n=i) &= \mu\Delta(1-\rho)\rho^i; & i\ge 1, j=1 \\ &= 0; & i=0, j=1 \\ &= (1-\mu\Delta)(1-\rho)\rho^i; & i\ge 1, j=0 \\ &= 1-\rho; & i=0, j=0 \end{aligned}$$

(b) $P(D_n=1) = \sum_{i=1} \mu\Delta((1-\rho)\rho^i = \mu\Delta\rho = \lambda\Delta$

(c) $$\begin{aligned} P(X_n=i \mid D_n=1) = \frac{P(X_n=i, D_n=1)}{P(D_n=1)} &= \{\mu\Delta(1-\rho)\rho^i\}/\lambda\Delta = (1-\rho)\rho^{i-1}; \ i\ge 1 \\ &= 0; \quad i=0 \end{aligned}$$

(d) $P(X_{n+1}=i | D_n=1) = P(X_n=i+1 | D_n=1) = (1-\rho)\rho^i$; $i \geq 0$

In the first equality above, we use the fact that, given a departure between $n\Delta$ and $(n+1)\Delta$, the state at $(n+1)\Delta$ is one less than the state at $n\Delta$; in the second equality, we use part d). Since $P(X_{n+1}=i) = (1-\rho)\rho^i$, we see that X_{n+1} is statistically independent of the event $D_n=1$. It is thus also independent of the complementary event $D_n=0$, and thus is independent of the random variable D_n.

(e)
$$
\begin{aligned}
P(X_{n+1}=i, D_{n+1}=j | D_n) &= P(D_{n+1}=j | X_{n+1}=i, D_n)P(X_{n+1}=i | D_n) \\
&= P(D_{n+1}=j | X_{n+1}=i)P(X_{n+1}=i)
\end{aligned}
$$

The first part of the above equality follows because X_{n+1} is the state of the Markov process at time $(n+1)\Delta$, so that, conditional on that state, D_{n+1} is independent of everything in the past. The second part of the equality follows from the independence established in e). This establishes that X_{n+1}, D_{n+1} are independent of D_n; thus their joint distribution is given by b).

(f) We assume the inductive result for k-1 and prove it for k; note that part f establishes the result for k=1. Using the hint,

$$
\begin{aligned}
P(X_{n+k}=i | D_{n+k-1}=1, D_{n+k-2},..., D_n) &= P(X_{n+k-1}=i+1 | D_{n+k-1}=1, D_{n+k-2},..., D_n) \\
&= \frac{P(X_{n+k-1}=i+1, D_{n+k-1}=1 | D_{n+k-2},..., D_n)}{P(D_{n+k-1}=1 | D_{n+k-2},..., D_n)} \\
&= \frac{P(X_{n+k-1}=i+1, D_{n+k-1}=1)}{P(D_{n+k-1}=1)} \\
&= P(X_{n+k-1}=i+1 | D_{n+k-1}=1) = P(X_{n+k}=i | D_{n+k-1}=1)
\end{aligned}
$$

The third equality above used the inductive hypothesis for the independence of the pair (X_{n+k-1}, D_{n+k-1}) from $D_{n+k-2},...D_n$ in the numerator and the corresponding independence of D_{n+k-1} in the denominator. From part e), with n+k-1 replacing n, $P(X_{n+k}=i | D_{n+k-1}) = P(X_{n+k}=i)$, so

$$P(X_{n+k}=i | D_{n+k-1}=1, D_{n+k-2},..., D_n) = P(X_{n+k})$$

Using the argument in e), this shows that conditional on $D_{n+k-2},..., D_n$, the variable X_{n+k} is independent of the event $D_{n+k-1}=1$ and thus also independent of $D_{n+k-1}=0$. Thus X_{n+k} is independent of $D_{n+k-1},..., D_n$. Finally,

$$P(X_{n+k}=i, D_{n+k}=j | D_{n+k-1},..., D_n) = P(D_{n+k}=j | X_{n+k}=i)P(X_{n+k}=i | D_{n+k-1},..., D_n)$$

$$= P(D_{n+k}=j | X_{n+k}=i)P(X_{n+k}=i)$$

which shows the desired independence for k.

(g) This shows that the departure process is Bernoulli and that the state is independent of past departures; i.e., we have proved the first two parts of Burke's theorem without using reversibility. What is curious here is that the state independence is critical in establishing the Bernoulli property.

3.57

The session numbers and their rates are shown below:

Session	Session number p	Session rate x_p
ACE	1	100/60 = 5/3
ADE	2	200/60 = 10/3
BCEF	3	500/60 = 25/3
BDEF	4	600/60 = 30/3

The link numbers and the total link rates calculated as the sum of the rates of the sessions crossing the links are shown below:

Link	Total link rate
AC	$x_1 = 5/3$
CE	$x_1 + x_3 = 30/3$
AD	$x_2 = 10/3$
BD	$x_4 = 10$
DE	$x_2 + x_4 = 40/3$
BC	$x_3 = 25/3$
EF	$x_3 + x_4 = 55/3$

For each link (i,j) the service rate is

$$\mu_{ij} = 50000/1000 = 50 \text{ packets/sec},$$

and the propagation delay is $D_{ij} = 2 \times 10^{-3}$ secs. The total arrival rate to the system is

$$\gamma = \Sigma_i x_i = 5/3 + 10/3 + 25/3 + 30/3 = 70/3$$

The average number on each link (i, j) (based on the Kleinrock approximation formula) is:

$$N_{ij} = \frac{\lambda_{ij}}{\mu_{ij} - \lambda_{ij}} + \lambda_{ij} D_{ij}$$

From this we obtain:

Link	Average Number of Packets on the Link
AC	(5/3)/(150/3 - 5/3) + (5/3)(2/1000) = 5/145 + 1/300
CE	1/4 + 1/50
AD	1/14 + 1/150
BD	1/4 + 1/50

DE	4/11 + 2/75
BC	1/5 + 1/60
EF	11/19 + 11/300

The average total number in the system is $N = \Sigma_{(i,j)} N_{ij} \cong 1.84$ packet. The average delay over all sessions is $T = N/\gamma = 1.84 \times (3/70) = 0.0789$ secs. The average delay of the packets of an individual session are obtained from the formula

$$T_p = \sum_{(i,j) \text{ on } p} \left[\frac{\lambda_{ij}}{\mu_{ij}(\mu_{ij} - \lambda_{ij})} + \frac{1}{\mu_{ij}} + D_{ij} \right]$$

For the given sessions we obtain applying this formula

Session p	Average Delay T_p
1	0.050
2	0.053
3	0.087
4	0.090

3.58

We convert the system into a closed network with M customers as indicated in the hint. The (k+1)st queue corresponds to the "outside world". It is easy to see that the queues of the open systems are equivalent to the first k queues of the closed system. For example, when there is at least one customer in the (k+1)st queue (equivalently, there are less than M customers in the open system) the arrival rate at queue i is

$$\sum_{m=1}^{k} r_m \frac{r_i}{\sum_{j=1}^{k} r_j} = r_i .$$

Furthermore, when the (k+1)st queue is empty no external arrivals can occur at any queue i, i = 1,2,...,k. If we denote with $p(n_1,...,n_k)$ the steady state distribution for the open system, we get

$$p(n_1,n_2,\ldots,n_k) = \begin{cases} 0 & \text{if } \sum_{i=1}^{k} n_i > M \\ \dfrac{\rho_1^{n_1} \rho_2^{n_2} \cdots \rho_k^{n_k} \rho_{k+1}^{\left(M - \sum_{i=1}^{k} n_i\right)}}{G(M)} & \text{otherwise} \end{cases}$$

where

$$\rho_i = \frac{r_i}{\mu}, \; i = 1,2,\ldots,k,$$

$$\rho_{k+1} = \frac{\sum_{i=1}^{k} r_i(1-\sum_{j=1}^{k} p_{ij})}{\sum_{i=1}^{k} r_i}$$

and G(M) is the normalizing factor.

3.59

If we insert a very fast M/M/1 queue ($\mu \to \infty$) between a pair of queues, then the probability distribution for the packets in the rest of the queues is not affected. If we condition on a single customer being in the fast queue, since this customer will remain in this queue for $1/\mu$ ($\to 0$) time on the average, it is equivalent to conditioning on a customer moving from one queue to the other in the original system.

If $P(n_1,\ldots,n_k)$ is the stationary distribution of the original system of k queues and $P'(n_1,\ldots,n_k,n_{k+1})$ is the corresponding probability distribution after the insertion of the fast queue k+1, then

$$P(n_1,\ldots,n_k \mid \text{arrival}) = P'(n_1,\ldots,n_k, n_{k+1} = 1 \mid n_{k+1} = 1),$$

which by independence of $n_1, \ldots, n_k, n_{k+1},$ is equal to $P(n_1,\ldots,n_k)$.

3.60

Let U_j = utility function of j^{th} queue.

We have to prove that

$$\lim_{M\to\infty}(U_j(M)) = \lim_{M\to\infty} \frac{\lambda_j(M)}{\mu_j} = 1$$

But from problem 3.65 we have

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

Thus it is enough to prove that

$$\lim_{M\to\infty}\frac{G(M)}{G(M-1)} = \rho_j$$

where $\rho_j = \max\{\rho_1,\ldots,\rho_k\}$. We have

$$G(M) = \sum_{n_1+\ldots+n_k=M} \rho_1^{n_1}\ldots\rho_j^{n_j}\ldots\rho_k^{n_k} = \sum_{\substack{n_1+\ldots+n_k=M\\ n_j=0}} \rho_1^{n_1}\ldots\rho_j^{n_j}\ldots\rho_k^{n_k} + \sum_{\substack{n_1+\ldots+n_k=M\\ n_j\neq 0}} \rho_1^{n_1}\ldots\rho_j^{n_j}\ldots\rho_k^{n_k}$$

$$= A(M) + B(M) \qquad (1)$$

Since $\rho_j = \max\{\rho_1,\ldots,\rho_k\}$ we have that

$$\lim_{M\to\infty}\frac{A(M)}{B(M)} = 0.$$

Thus, Eq. (1) implies that

$$\lim_{M\to\infty}\frac{\displaystyle\sum_{\substack{n_1+\ldots+n_k=M\\ n_j\neq 0}} \rho_1^{n_1}\ldots\rho_j^{n_j}\ldots\rho_k^{u_k}}{G(M)} = 1$$

or, denoting $n'_j = n_j - 1$,

$$\lim_{M\to\infty}\frac{\rho_j \displaystyle\sum_{n_1+\ldots+n'_j+\ldots+n_k=M-1} \rho_1^{n_1}\ldots\rho_j^{n'_j}\ldots\rho_k^{n_k}}{G(M)} = 1$$

or

$$\lim_{M\to\infty}\frac{\rho_j\, G(M-1)}{G(M)} = 1$$

3.61

We have $\sum_{i=0}^{m} p_i = 1$.

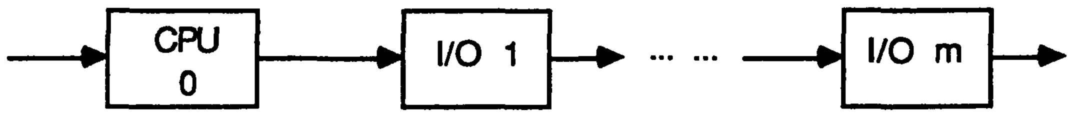
The arrival rate at the CPU is λ/p_0 and the arrival rate at the i^{th} I/O port is $\lambda p_i/p_0$. By Jackson's Theorem, we have

$$P(n_0, n_1, \dots, n_m) = \prod_{i=0}^{m} \rho_i^{n_i}(1 - \rho_i)$$

$$\text{where } \rho = \frac{\lambda}{\mu_0 \, p_0}$$

$$\text{and } \rho_i = \frac{\lambda \, p_i}{\mu_i \, p_0} \qquad \text{for } i > 0$$

The equivalent tandem system is as follows:



The arrival rate is λ. The service rate for queue 0 is $\mu_0 p_0$ and for queue i (i > 0) is $\mu_i p_0/p_i$.

3.62

Let λ_0 be the arrival rate at the CPU and let λ_i be the arrival rate at I/O unit i. We have

$$\lambda_i = p_i \lambda_0, \qquad i = 1,\dots,m.$$

Let

$$\bar{\lambda}_0 = 1, \quad \bar{\lambda}_i = p_i, \quad i = 1,\dots,m,$$

and

$$\rho_i = \frac{\bar{\lambda}_i}{\mu_i}, \quad i = 0,1,\dots,m,$$

By Jackson's Theorem, the occupancy distribution is

$$P(n_0, n_1, \dots, n_m) = \frac{\rho_0^{n_0} \rho_1^{n_1} \cdots \rho_m^{n_m}}{G(M)},$$

where G(M) is the normalization constant corresponding to M customers,

$$G(M) = \sum_{n_0+n_1+\dots+n_m=M} \rho_0^{n_0} \rho_1^{n_1} \cdots \rho_m^{n_m}$$

Let

$$U_0 = \frac{\lambda_0}{\mu_0}$$

be the utilization factor of the CPU. We have

$$U_0 = P(n_0 \geq 1) = \sum_{\substack{n_0+n_1+\ldots+n_m=M \\ n_0 \geq 1}} P(n_0, n_1, \ldots, n_m)$$

$$= \sum_{n'_0+n_1+\ldots+n_m=M-1} \frac{\rho_0^{n'_0+1} \rho_1^{n_1} \ldots \rho_m^{n_m}}{G(M)}$$

$$= \rho_0 \frac{G(M-1)}{G(M)} = \frac{1}{\mu_0} \frac{G(M-1)}{G(M)},$$

where we used the change of variables $n'_0 = n_0 - 1$. Thus the arrival rate at the CPU is

$$\lambda_0 = \frac{G(M-1)}{G(M)}$$

and the arrival rate at the I/O unit i is

$$\lambda_i = \frac{p_i G(M-1)}{G(M)}, \quad i=1,\ldots,m.$$

3.63

(a) We have $\lambda = N/T$ and

$$T = T_1 + T_2 + T_3$$

where

T_1 = Average time at first transmission line
T_2 = Average time at second transmission line
$T_3 = \overline{Z}$

We have

$$\overline{X} \leq T_1 \leq N\overline{X} \tag{1}$$

$$\overline{Y} \leq T_2 \leq N\overline{Y}$$

so

$$\frac{N}{N(\bar{X}+\bar{Y})+\bar{Z}} \le \lambda \le \frac{N}{\bar{X}+\bar{Y}+\bar{Z}} .$$

Also

$$\lambda \le \frac{K}{\bar{X}}, \qquad \lambda \le \frac{1}{\bar{Y}}$$

so finally

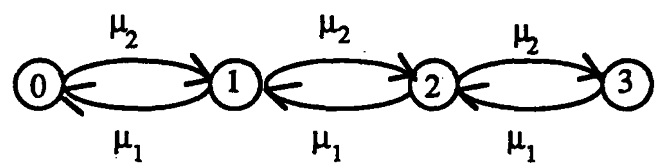
$$\frac{N}{N(\bar{X}+\bar{Y})+\bar{Z}} \le \lambda \le \min\left\{\frac{K}{\bar{X}}, \frac{1}{\bar{Y}}, \frac{N}{\bar{X}+\bar{Y}+\bar{Z}}\right\}$$

(b) The same line of argument applies except that in place of (1) we have

$$\bar{X} \le T_1 \le (N - K + 1)\bar{X}$$

3.64

(a) The state is determined by the number of customers at node 1 (one could use node 2 just as easily). When there are customers at node 1 (which is the case for states 1, 2, and 3), the departure rate from node 1 is μ_1; each such departure causes the state to decrease as shown below. When there are customers in node 2 (which is the case for states 0, 1, and 2), the departure rate from node 2 is μ_2; each such departure causes the state to increase.



(b) Letting p_i be the steady state probability of state i, we have $p_i = p_{i-1}\,\rho$, where $\rho = \mu_2/\mu_1$. Thus $p_i = p_0\rho^i$. Solving for p_0,

$$p_0 = [1 + \rho + \rho^2 + \rho^3]^{-1}, \quad p_i = p_0\,\rho^i; \quad i=1,2,3.$$

(c) Customers leave node 1 at rate μ_1 for all states other than state 0. Thus the time average rate r at which customers leave node 1 is $\mu_1(1-P_0)$, which is

$$r = \frac{\rho+\rho^2+\rho^3}{1+\rho+\rho^2+\rho^3}\,\mu_1$$

(d) Since there are three customers in the system, each customer cycles at one third the rate at which departures occur from node 1. Thus a customer cycles at rate r/3.

(e) The Markov process is a birth-death process and thus reversible. What appears as a departure from node i in the forward process appears as an arrival to node i in the backward process. If we order the customers 1, 2, and 3 in the order in which they depart a node, and note that this order never changes (because of the FCFS service at each node), then we see that in the backward process, the customers keep their identity, but the order is reversed with backward departures from node i in the order 3, 2, 1, 3, 2, 1, ...

3.65

Since $\mu_j(m) = \mu_j$ for all m, and the probability distribution for state $n = (n_1,...,n_k)$ is

$$P(n) = \frac{\rho_1^{n_1}\cdots\rho_k^{n_k}}{G(M)},$$

where

$$\rho_j = \frac{\overline{\lambda}_j}{\mu_j}$$

The utilization factor $U_j(M)$ for queue j is

$$U_j(M) = \sum_{\substack{n_1+\dots+n_k=M \\ \text{s.t. } n_j\neq 0}} P(n) = \frac{\sum_{\substack{n_1+\dots+n_k=M \\ n_j\neq 0}} \rho_1^{n_1}\dots\rho_k^{n_k}}{G(M)},$$

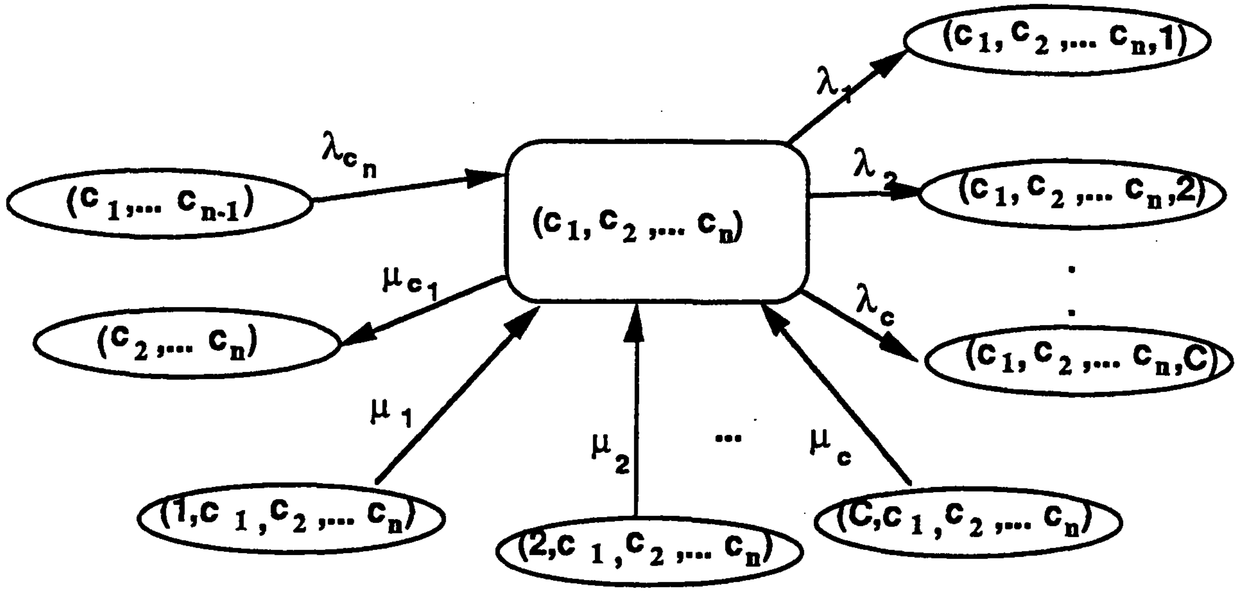
Denoting $n'_j = n_j-1$, we get

$$U_j(M) = \frac{\sum_{n_1+\dots+n'_j+\dots+n_k=M-1} \rho_1^{n_1}\dots\rho_j^{n'_j}\dots\rho_k^{n_k}\cdot\rho_j}{G(M)} = \frac{\rho_j G(M-1)}{G(M)}.$$

3.66

Let c_i indicate the class of the i^{th} customer in queue.

We consider a state $(c_1,c_2,...,c_n)$ such that $\mu_{c_1} \neq \mu_{c_n}$.



If the steady-state distribution has a product form then the global balance equations for this state give

$$p(c_1)\cdots p(c_n)(\mu_{c_1} + \lambda_1 + \cdots + \lambda c) = p(c_1)\cdots p(c_{n-1})\lambda c_n$$
$$+ (\mu_1 p(1) + \mu_2 p(2) + \cdots + \mu_c p(c))\, p(c_1)\cdot\ldots\cdot p(c_n)$$

or

$$p(c_n)(\mu c_1 + \lambda_1 + \cdots + \lambda_c) = \lambda c_n + (\mu_1 p(1) + \mu_2 p(2) + \cdots + \mu_c p(c))\cdot p(c_n)$$

Denote

$$M = \mu_1 p(1) + \mu_2 p(2) + \cdots + \mu_c p(c)$$

$$\lambda = \lambda_1 + \cdots + \lambda_c.$$

Then

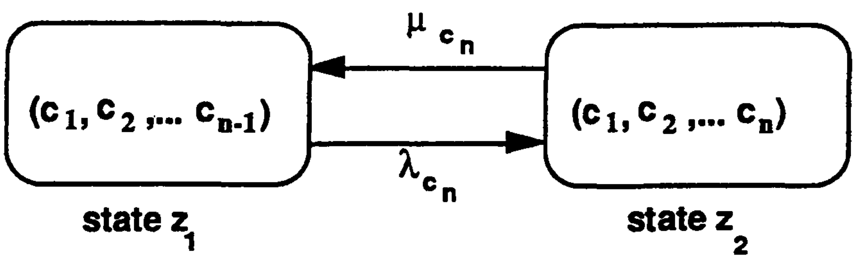
$$p(c_n)(\mu_{c1} + \lambda) = \lambda_{cn} + M\cdot p(c_n)$$

or

$$p(c_n) = \frac{\lambda_{c_n}}{\mu_{c_1} + \lambda - M}$$

This is a contradiction because even if $\lambda_{c_1} = 0$, $p(c_n)$ still depends on μ_{c_1}. The contradiction gives that $\mu_{c_1} = \mu =$ constant for every class. Thus we can model this system by a Markov chain only if $\mu_1 = \mu_2 = \cdots = \mu_c$.

(b) We will prove that the detailed balance equations hold. Based on the following figure



the detailed balance equations are

$$\lambda_{c_n} \cdot p(z_1) = \mu_{c_n} \cdot p(z_2)$$

or

$$\lambda_{c_n} \rho_{c_1} \cdots \rho_{c_{n-1}} = \mu_{c_n} \rho_{c_1} \cdots \rho_{c_n},$$

which obviously holds.