4-1979

Markovian Multiserver Queueing Systems with Servers in Series

Nancy Jean Boynton
Western Michigan University

MARKOVIAN MULTISERVER QUEUEING SYSTEMS WITH SERVERS IN SERIES

by
Nancy Jean Boynton

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfullment
of the
Degree of Doctor of Philosophy

Western Michigan University
Kalamazoo, Michigan
April, 1979

ACKNOWLEDGEMENTS

I wish to thank the faculty of the Mathematics Department of Western Michigan University for all their time and interest. In particular, I would like to thank Dean A. Bruce Clarke, my dissertation advisor, for his guidance and encouragement. I would also like to thank Professor Ralph L. Disney of Virginia State University and Polytechnic Institute, Professor Gerald L. Seivers, Professor Michael R. Stoline, and Professor Christian R. Hirsch of Western Michigan University for their many helpful suggestions in the preparation of this manuscript. Finally, I would like to thank Margaret Johnson for typing this manuscript.

Nancy Jean Boynton

For My Parents

and

My Grandparents

7913673

BOYNTON, NANCY JEAN
MARKOVIAN MULTISERVER QUEUEING SYSTEMS WITH
SERVERS IN SERIES.

WESTERN MICHIGAN UNIVERSITY, PH.D., 1979

University
Microfilms
International 300 N. ZEEB ROAD, ANN ARBOR, MI 48106

TABLE OF CONTENTS

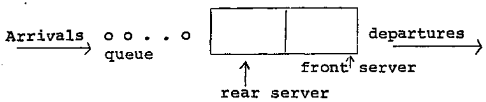
CHAPTER PAGE

0. INTRODUCTION

In this paper queues with the servers in series will be examined. Customers must pass through each server station but are served by only one. Examples of this would be gas stations with several pumps in a line and certain assembly or production lines.

Since each customer must pass through each server station either before or after being served, one has a possibility of blocking. (E.g. A server is empty but no customer can reach it since access is blocked by a prior busy server, or a customer may complete service but may be forced to remain in the server station, blocking access to it, since a server in front is still busy.) Customers are served on a first-come-first-served basis. In this paper arrivals are assumed to form a Poisson stream with mean arrival rate λ . Each server is assumed to have a negative exponential service time distribution with the same service rate μ = 1/mean service time.

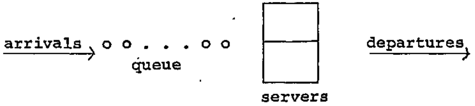
First the queue with 2 servers in series will be examined.



The values of ρ, the ratio of the input rate to the total output rate, $\rho = \lambda/(2\mu)$, for which the system is positive recurrent, recurrent null, and transient are determined. A number of formulas and numerical examples of this system are then examined.

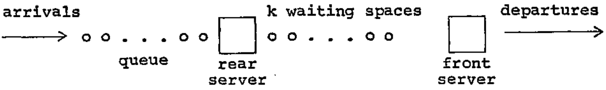
The queue with 2 servers in series is related to a queue with bulk service where 2 customers are served at once as long as there are 2 customers in the system [3, pp. 412-414]. However, in the bulk system a customer arriving to an empty system must wait until another arrives before beginning service. Whereas, in the series system service begins immediately.

There are well-known results for queues where the servers are in parallel instead of in series. Such a system has no blocking. As soon as a customer completes service he leaves and another customer begins service.



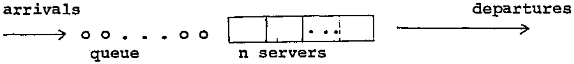
The efficiency of the 2-server parallel queue will be compared to the 2-server series queue. Because of the possibility of blocking, the series queue clearly is less efficient than the parallel queue. However placing the servers in a single line often offers construction advantages.

The 2-server series queueing problem can be generalized to a queueing system with 2 servers in series but with k waiting spaces between the servers.

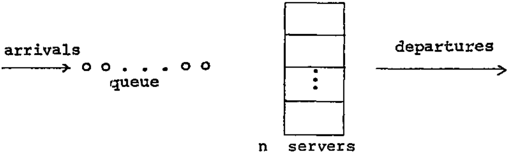


This avoids some of the blocking since if a customer completes service in the rear server and the front server is busy, there may be a vacant storage space into which he can move temporarily thus avoiding the blocking of this server. The values of ρ which make the queue positive recurrent and transient are found. As would be expected this is an improvement over the 2-server series system but is still not as efficient as the 2-server parallel system.

Finally the queueing system with n servers in series will be examined.



The values of ρ which make this system positive recurrent and transient are found. These can be compared to the values of ρ which make the n server parallel queueing system positive recurrent, recurrent null, and transient.



As in the two server case, the n server parallel queue is more efficient than the n server series queue. The larger n becomes, the greater the difference in relative efficiency, since blocking becomes increasingly likely.

I. MARKOV CHAIN PRELIMINARIES

1.1 Markov Chains

The processes considered in this paper are all continuous-parameter, discrete-state, time-homogeneous Markov chains X_t, $0 \le t < \infty$, in which the state space S consists of a denumerable set of states, $S = \{s_0, s_1, s_2, \ldots\}$. Denote the transition function of such a Markov chain by

$$p_{uv}(t) = \Pr[X_{t+s} = v \mid X_s = u]$$

for $u, v \in S$, which are independent of s by the time-homogeneity. These functions will satisfy the Chapman-Kolmogorov equations

$$p_{uw}(t+s) = \sum_{v \in S} p_{uv}(t) p_{vw}(s) \quad \text{for} \quad s > 0, t > 0 .$$

For a detailed discussion of such processes the reader is refered to any standard reference, for example [2] or [6].

1.2 The Rate Matrix

The Markov chains considered here are all standard in the sense of Chung [2, p. 128] and satisfy the following assumptions:

(a) Each $p_{uv}(t)$ is continuous at 0 .

(b) The derivatives $\lambda_{ij} = p'_{s_i s_j}(0)$ exist for each $s_i , s_j \in S$, are finite, and satisfy $\lambda_{ii} \leq 0 , \lambda_{ij} \geq 0$ for $i \neq j$, $\sum_j \lambda_{ij} = 0$.

(c) $p_{s_i s_j}(h) = \delta_{ij} + \lambda_{ij} h + o(h)$, where the functions $o(h)$ satisfy $\frac{o(h)}{h} \to 0$ as $h \to 0^+$, uniformly in i and j .

(d) The time T that the process remains in any given state s_i , given that $X_o = s_i$, has a negative exponential distribution with parameter λ_{ii} , $\Pr[T \leq t] = 1 - e^{\lambda_{ii} t}$

(e) The probability that the process will have two or more one-step transitions in time h is $o(h)$.

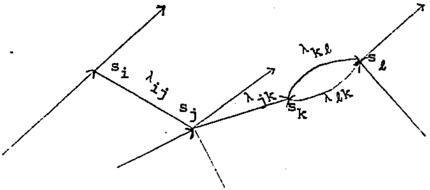
The quantities λ_{ij} , $i \neq j$, are termed the <u>transition rates</u>. λ_{ij} is proportional to the probability that a system in state s_i will be in state s_j h time units later, up to a term of order $o(h)$.

The matrix $\Lambda = (\lambda_{ij})$ is termed the <u>rate matrix</u> of the system.

Under the preceeding assumptions systems of differential equations, <u>the Kolmogorov systems</u>, [2, p. 245] can be constructed for the transition probabilities $p_{s_i s_j}(t)$. For the processes considered in this paper, however, these systems are too complex to be easily analyzed.

1.3 Rate Diagram

The behavior of such a Markov chain and its rates can be described graphically by considering the weighted digraph with vertices at the various states $s_1, s_2, \ldots$, and with a directed edge indicated from state s_i to state s_j weighted with the quantity λ_{ij} provided $i \neq j$ and $\lambda_{ij} > 0$. (The quantities λ_{ii} need not be indicated since by section 1.3(b) $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$.)



This weighted diagraph is termed the rate diagram of the system. The limiting, long-run, and connectivity properties of the Markov chain can be analyzed in terms of this rate diagram.

1.4 Structure and Classification of States

A state s_j is said to be accessible from state s_i if $p_{s_i s_j}(t) > 0$ for some $t > 0$. It can be shown [9, p. 274] that s_j is accessible from state s_i if and only if $s_i = s_j$ or a directed path exists in the diagraph from s_i to s_j. Two states s_i and s_j are said to communicate if each is accessible from the other. Such communication forms an equivalence relation over the state space.

A collection R of states is termed closed if s_j is not accessible from s_i whenever $s_i \in R$, $s_j \notin R$.

A minimal closed set is a nonempty closed set of states which has no closed proper subset. Within such a set every state communicates with every other state.

A Markov chain is termed irreducible if its state space consists of a single minimal closed set.

A state s_i is termed absorbing if

$$\Pr[X_t = s_i | X_0 = s_i] = 1 \quad \text{for} \quad t > 0,$$

which is equivalent to $\lambda_{ij} = 0$ for all $j \neq i$. Once state i is reached, the process stays there with probability 1.

A state s_i is termed <u>recurrent</u> if either it is absorbing or

$$\Pr[\exists t > \mu \ni X_t = s_i \mid X_u \neq s_i \text{ for some } u > 0 \text{ and } X_o = s_i] = 1$$

i.e. if the process leaves state s_i, return is certain.

For a recurrent nonabsorbing state s_i, let the random variable T represent the first return time, i.e. $T = \inf\{t : X_t = s_i$ where $X_o = s_i$ and $X_u \neq s_i$ for some $0 < u < t\}$. The quantity $E[T]$ is termed the <u>mean return time</u>. A state is termed <u>positive recurrent</u> if and only if it is absorbing or $E[T] < \infty$.

The property of positive recurrence is extremely important. When a queueing system is positive recurrent the queue does not grow infinitely long and return to the empty queue is certain. Much of this paper will be devoted to obtaining conditions for positive recurrence in several important processes.

A recurrent nonabsorbing state is termed <u>recurrent null</u> if $E[T] = \infty$, i.e. recurrence is certain (with probability 1) but the mean recurrence time is infinite.

A nonrecurrent state is termed transient, i.e. the state is not absorbing and return is not certain, $Pr[X_t = s_i \text{ for some } t > 0 | X_o = s_i] < 1$. Thus every state is one of these three types.

An important theorem is that within a minimal closed set all states are of the same type: positive recurrent, recurrent null, or transient. Thus one speaks of a positive recurrent, recurrent null or transient Markov chain.

The definitions given here are not in their most general form, but are equivalent to the general definitions for processes satisfying the assumptions of this chapter and of this paper.

1.5 Limiting Distributions

Since, as previously mentioned, study of the transition probabilities $p_{s_i s_j}(t)$ for $0 \le t < \infty$ is difficult, usually (and in this paper, always) one studies the behavior of these quantities for large values of t , i.e. one assumes the system has been operating for a long time.

A key theorem is the following (which is a special case of a more general result [2, pp. 183-186]):

Theorem 1.1 For an irreducible Markov chain (satisfying the conditions of 1.1 and 1.2) the following limits exist for all states i and j

$$p(j) = \lim_{t\to\infty} p_{s_i s_j}(t)$$

with values independent of the initial state s_i. The quantities p(j) are termed the limiting probabilities.

(a) If the process is positive recurrent, then

$$p(j) > 0 \quad \text{for each } j, \quad \text{and}$$
$$\sum_j p(j) = 1$$

In this case the sequence {p(j)} is termed the limiting probability distribution of the process.

(b) If the process is null recurrent or transient, then

$$p(j) \equiv 0 \quad \text{for each } j .$$

The determination of the limiting probability distributions for various positive recurrent processes is a major goal here.

1.6 The Normal Equations

Assume that the Markov chain under consideration is irreducible.

A system of equations that must be satisfied by the limiting probabilities $p(j)$ can be obtained formally from the Chapman-Kolmogorov equations (section 1.1) by differentiating with respect to s, setting $s = 0$, letting $t \to \infty$, and determining that

$$\lim_{t\to\infty} \frac{d}{dt} p_{s_i s_j}(t) = \frac{d}{dt} \lim_{t\to\infty} p_{s_i s_j}(t) = \frac{d}{dt} p(j) = 0 .$$

(Of course the proof involves a substantial argument, see [3, p. 55]). The resulting system of equations are termed the <u>normal equations</u> of the process

$$\text{(1.1)} \qquad \sum_i p(i)\lambda_{ij} = 0 \quad \text{for each } j .$$

In matrix form this system can be written as

$$\text{(1.2)} \qquad \vec{p}\,\Lambda = \vec{0}$$

where $\vec{p} = (p(0), p(1), p(2), \ldots)$.

Theorem 1.2 <u>For an irreducible Markov chain with the assumptions of section</u> 1.2 <u>the limiting probabilities satisfy the normal equations</u> (1.1) <u>or</u> (1.2).

The normal equations can be displayed more intuitively in a form that facilitates setting them up by writing (1.1) in the form

$$-p(j)\lambda_{jj} = \sum_{i \neq j} p(i)\lambda_{ij}$$

or

$$p(j) \sum_{k \neq j} \lambda_{jk} = \sum_{i \neq j} p(i)\lambda_{ij} \tag{1.3}$$

Equation (1.3) can be interpreted as a balance equation, which is easily set up using the rate diagram. On the left one has the limiting probability of a fixed state s_j multiplied by the total rate out from that state s_j to all other states s_k. On the right one has the total rate into the state s_j from all other states s_i, weighted in each case by the limiting probability of s_i. By setting up this balance equation at each vertex of the rate diagram the normal equations can be obtained readily.

1.7 Conditions for Positive Recurrence

Consider the normal equations (1.1) or (1.2) with a general unknown vector $\vec{x}$ replacing $\vec{p}$.

$$(1.4) \quad \text{or} \quad \begin{cases} \sum_i x(i)\lambda_{ij} = 0 \quad \text{for each } j \\ \vec{x}\,\Lambda = \vec{0} \end{cases}$$

The key theorem is the following, see [4].

Theorem 1.3 <u>An irreducible Markov chain (satisfying the assumptions of sections</u> 1.1 <u>and</u> 1.2) <u>is positive recurrent if and only if the normal equations</u> (1.4) <u>have a nontrivial solution which forms an absolutely convergent series,</u> $\Sigma|x(i)| < \infty$. <u>In this case there exists a nonzero constant</u> c <u>such that</u>

$$(1.5) \qquad p(j) = cx(j) \quad \text{for each } j$$

i.e. the limiting probability distribution is proportional to any nontrivial absolutely convergent solution to (1.4). Such a Markov chain is positive recurrent only if any nonnegative solution of the inequalities

$$(1.6) \qquad \sum_i x(i)\lambda_{ij} \le 0 \quad \text{for each } j$$

is convergent, $\sum_{i=o}^{\infty} x(i) < \infty$.

This theorem not only gives a method for determining when an irreducible process is positive recurrent, but also for determining the limiting probabilities $p(j)$. (Choose c so that $c \sum_j x(j) = \sum_j p(j) = 1$.)

1.8 Conditions for Transience or Null Recurrence

A necessary and sufficient condition for an irreducible Markov chain to be transient can be found by considering the following system of equations:

$$\sum_{j=o}^{\infty} \lambda_{ij}\, x(j) = 0 \quad \text{for} \quad i = 1, 2, \ldots \tag{1.7}$$

or

$$\Lambda \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} y \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{1.8}$$

where y can be any number. Note that the equation for $i = 0$, (state s_o), is omitted from the system. (Any other distinguished state could be omitted instead.)

Theorem 1.4 (Foster [4]). An irreducible Markov chain (satisfying the assumptions of sections 1.1 and 1.2) is transient if and only if the system (1.7) or (1.8) has a bounded nonconstant solution.

One also uses the following:

Theorem 1.5 (Foster [4]) An irreducible Markov chain (satisfying the assumptions of sections 1.1 and 1.2) is recurrent if and only if the system of inequalities

$$\sum_{j=0}^{\infty} \lambda_{ij}\, x(j) \le 0 \quad \text{for} \quad i = 1, 2, \ldots \tag{1.9}$$

has a solution satisfying $x(j) \to \infty$ as $j \to \infty$.

II. THE STANDARD 2-SERVER SERIES SYSTEM

2.1 Definition of the Problem

Consider a queueing system with two servers in series. The customers will be served on a first-come-first-served basis. Each customer is served by one of the two servers, and the queue discipline requires that any customer entering the servers move to the furthest forward accessible server.

The various states of the system can conveniently be represented by symbols of the form $(x\ y)_n$ where n represents the number of customers who have not yet completed service, the <u>queue length</u>, x and y are server sybmols representing the status of the two servers: y indicates the status of the front server and x indicates the status of the rear server. y will have the symbolic values 'e' or 's' depending on whether the front server is <u>empty</u> or <u>busy</u>. x will be 'e','s' or 'b' depending on whether the rear server is <u>empty</u>, <u>busy</u> or <u>blocked</u> (i.e. contains a customer who has completed service but is blocked from exit by a busy front server).

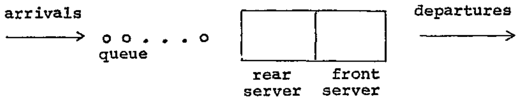


Diagram of the System: Figure 2.1

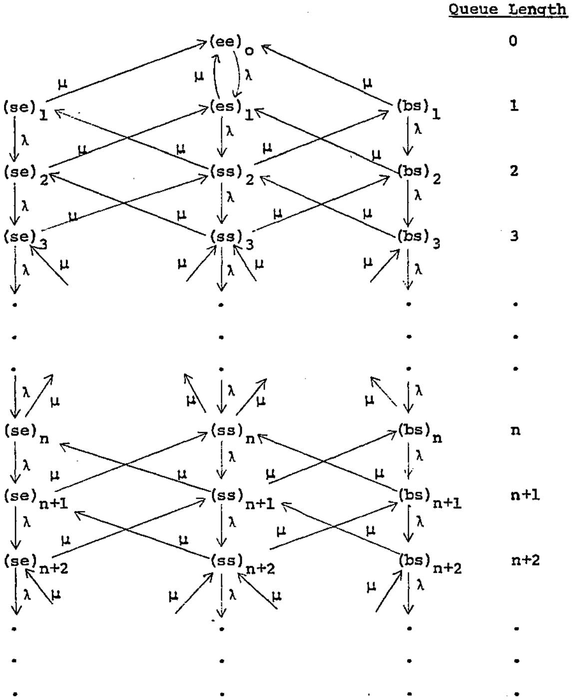
The states of the system are $(ee)_o$, $(es)_1$, $(bs)_1$, $(se)_1$, $(ss)_2$, $(bs)_2$, $(se)_2$, ..., $(ss)_n$, $(bs)_n$, $(se)_n$, ... Note that the queue length subscript n represents the number of unserved customers waiting in line plus the number of busy servers.

Assume that arrivals form a Poisson process with intensity λ (λ = mean number of arrivals per unit time). Assume that each server has a mean service rate of μ, where service times are negative exponentially distributed with probability density function $\mu e^{-\mu t}$. Therefore the mean service time is $1/\mu$. Assume also that all interarrival and service times are mutually independent.

With these assumptions the process forms a continuous-time, discrete-state, time-homogeneous Markov chain of the type considered in Chapter 1. The state space is denumerable and the states are represented by the state symbols $(xy)_n$.

2.2 Rate Diagram

The one-step transitions for this Markov chain are generated by the arrivals and the service completions. Each arrival has the effect of increasing the queue length by 1 (and changing a server symbol from an e to an s if a server is accessible). Each service completion has the effect of reducing the queue length by 1 (and changing one or both of the server symbols). Thus, for example, a completion in the front server would change state $(bs)_3$ to $(ss)_2$. The probability of an arrival in time h is $\lambda h + o(h)$, so (see section 1.2(d)) the <u>transition rate</u> λ_{ij} for a transition generated by an arrival will be λ. The probability of a completion from one busy server in time h is $\int_0^h \mu e^{-\mu t}\, dt = \mu h + o(h)$, so the <u>transition rate</u> λ_{ij} for a transition generated by a service completion will be μ. Thus the rate diagram can be constructed as follows.



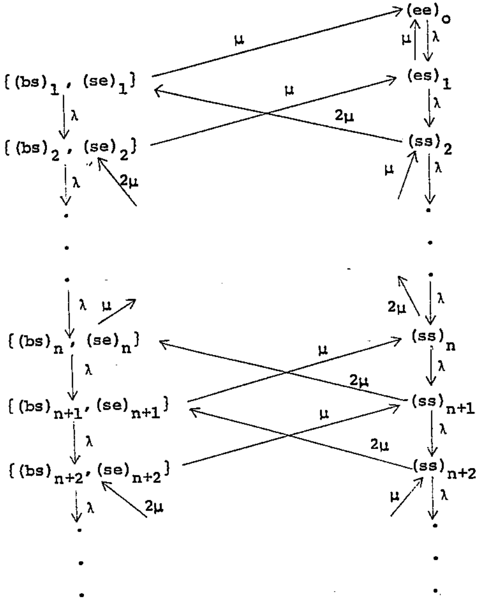
<u>Rate Diagram:</u> Figure 2.2

2.3 Lumping

Note that arrivals are represented in the rate diagram (Fig. 2.2) by vertical arrows. Departures from either the right or the left state in the diagram with the same queue length lead to the same middle state at the same rate μ . This means that the left and right states at each queue length can be combined into a single state $\{(bs)_k, (se)_k\}$ and the Markov property retained. It is unnecessary to know which of these two states holds in order to predict the transition probability to the next state. We say that the states $(bs)_k$ and $(se)_k$ can be lumped into a single state, and that the process is lumpable, [7, p. 124]. When possible such lumping is highly desirable, since it simplifies the discussion by reducing the size of the state space. After finding the limiting probabilities for the state $(ss)_n$ and the lumped states $\{(bs)_n, (se)_n\}$, the limiting probabilities for the states $(bs)_n$ and $(se)_n$ can be determined. The limiting probabilities of the lumped states is equal to the sum of the limiting probabilities of the individual states. Service completions in the state $(ss)_{n+1}$ lead to $(bs)_n$ or $(se)_n$ at equal rates. Arrivals to $(bs)_{n-1}$ lead to $(bs)_n$ at the same rate arrivals to

$(se)_{n-1}$ lead to $(se)_n$. Since these are the only transitions leading to $(bs)_n$ or $(se)_n$ and the limiting probabilities are independent of the state the process starts in, the limiting probabilities of the states $(bs)_n$ and $(se)_n$ will each be one-half of the limiting probabilities of the lumped superstate $\{(bs)_n, (se)_n\}$.

The rate diagram for the lumped process is as follows:



<u>Rate Diagram for Lumped Process</u>: Figure 2.3

Denote the long-run probabilities (see section 1.5) of this irreducible Markov chain by

$$r(n) = p\{(bs)_n, (se)_n\} = \lim_{t\to\infty} P_{\alpha,\{(bs)_n,(se)_n\}}(t) , \quad \text{for } n \geq 1$$

(2.1)

$$p(n) = p\{(ss)_n\} = \lim_{t\to\infty} P_{\alpha,\{(ss)_n\}}(t) , \quad \text{for } n \geq 2$$

$$p(1) = p\{(es)_1\}$$

$$p(0) = p\{(ee)_o\}$$

where α represents an arbitrary initial state.

2.4 Normal Equations

The normal equations can be derived from Figure 2.3 by balancing input and output at each vertex. For example, at vertex $\{(bs)_{n+1}, (se)_{n+1}\}$ one has

$$p\{(bs)_{n+1},(se)_{n+1}\}(\lambda+\mu) = p\{(bs)_n, (se)_n\}\lambda + p\{(ss)_{n+2}\}2\mu$$

or

$$(\lambda + \mu)r(n+1) = \lambda r(n) + 2\mu p(n+2) \quad .$$

Balancing the input and output at vertex $\{(ss)_{n+1}\}$ one has

$$p\{(ss)_{n+1}\}(\lambda+2\mu) = p\{(ss)_n\}\lambda + p\{(bs)_{n+2}, (se)_{n+2}\}\mu$$

or

$$(\lambda + 2\mu)p(n+1) = \lambda p(n) + \mu r(n+2) .$$

If this balancing is also done for the vertices $\{(ee)_o\}$, $\{(bs)_1, (se)_1\}$, and $\{(es)_1\}$, one obtains the following system of normal equations:

$$-(\lambda+\mu)r(n+1) + \lambda r(n) + 2\mu p(n+2) = 0 , \quad n \geq 1$$
$$\mu r(n+2) - (\lambda+2\mu)p(n+1) + \lambda p(n) = 0 , \quad n \geq 1$$

plus 3 special equations for the first 3 states:

$$-(\lambda+\mu)r(1) + 2\mu p(2) = 0$$

$$\mu r(2) - (\lambda+\mu)p(1) + \lambda p(0) = 0$$

$$\mu r(1) - \lambda p(0) + \mu p(1) = 0$$

This system can be simplified by introducing the parameter

$$\rho = \frac{\lambda}{2\mu} .$$

This is the ratio of the arrival rate to the total service rate of the two servers, and represents the traffic intensity of the corresponding 2-server parallel queueing system.

The final normal equations are

$$(2.2) \qquad -(2\rho+1)r(n+1) + 2\rho r(n) + 2p(n+2) = 0\,, \quad n \ge 1$$

$$(2.3) \qquad r(n+2) - 2(\rho+1)p(n+1) + 2\rho p(n) = 0\,, \quad n \ge 1$$

$$(2.4) \qquad -(2\rho+1)r(1) + 2p(2) = 0$$

$$(2.5) \qquad r(2) - (2\rho+1)p(1) + 2\rho p(0) = 0$$

$$(2.6) \qquad -\rho p(0) = \rho(1) <-> -2\rho p(0) + \rho(1)$$

The system is dependent, so that any solutions of all but one of these equations will satisfy the remaining one. Thus one of the equations can be disregarded. Let us disregard (2.5) and study the system (2.2)-(2.4) and (2.6).

By Theorem 1.3 the Markov chain will be positive recurrent iff the system (2.2)-(2.4) and (2.6) with $r(n)$ and $p(n)$, replaced by the arbitrary variables $x_1(n)$ and $x_2(x)$ respectively has a nontrivial absolutely convergent solution. If this is the case, any such solution to this homogeneous system can be normalized to form a probability distribution satisfying

$$(2.7) \qquad \sum_{n=1}^{\infty} r(n) + \sum_{n=0}^{\infty} p(n) = 1\,.$$

Note that $p(0)$ now appears only in equation (2.6). (2.6) can be used to determine $p(0)$ in terms of the other quantities, (2.8) $$p(0) = \frac{1}{2\rho}[r(1) + p(1)] \quad .$$

Eliminating $p(0)$, (2.7) becomes

(2.9) $$\frac{1}{2\rho}[r(1) + p(1)] + \sum_{n=1}^{\infty} [r(n) + p(n)] = 1 \; .$$

Thus the problem reduces to finding necessary and sufficient conditions that the system (2.2) - (2.4) has a nontrivial absolutely convergent solution.

2.5 Difference Equation Approach

The general system (2.2) and (2.3) (omitting the special equation (2.4)) constitutes a vector difference equation. Using the shift operator E, $E[x(n)] = x(n+1)$, this system can be written in the form

(2.10) $$F(E)\vec{x}(n) = \vec{0}, \quad n \geq 1$$

where

(2.11) $$F(E) = \begin{bmatrix} -(2\rho+1)E + 2\rho & 2E^2 \\ E^2 & -2(\rho+1)E + 2\rho \end{bmatrix}$$

and

(2.12) $$\vec{x}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} \quad .$$

The fundamental solutions of such a linear homogeneous difference equation are known [8, p. 601] to have the form

$$\vec{c}\zeta^n$$

where ζ is an appropriate, possibly complex nonzero scalar, and $\vec{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$ is an appropriate constant vector.

If this is substituted into (2.10) one obtains

$$\begin{aligned} \vec{0} &= F(E)\vec{c}\zeta^n && \text{for } n \geq 1 \\ &= F(E)\zeta^n\vec{c} \\ &= F(\zeta)\zeta^n\vec{c} && \text{since } E\zeta^n = \zeta\cdot\zeta^n \\ &= F(\zeta)\vec{c}\zeta^n \end{aligned}$$

Thus $F(\zeta)\vec{c} = \vec{0}$, and one concludes $\vec{c}\zeta^n$ is a solution iff $F(\zeta)$ is a vector in its null space. Thus ζ must be a zero of the <u>fundamental polynomial</u>

$$f(z) = \det F(z)$$

$$= \begin{vmatrix} -(2\rho+1)z + 2\rho & 2z^2 \\ z^2 & -2(\rho+1)z+2\rho \end{vmatrix}$$

and $\vec{c}$ must satisfy

(2.14) $$[-(2\rho+1)\zeta+2\rho]c_1+2\zeta^2c_2 = 0$$

or, equivalently

$$(2.15) \qquad \zeta^2 c_1 + [-2(\rho+1)\zeta + 2\rho] c_2 = 0$$

(Certain modifications are necessary if $f(z)$ has multiple zeros. This will be shown not to be the case.)

The general solution to the system (2.10) has the form

$$(2.16) \qquad \vec{x}(n) = \sum_j \vec{c}_j \zeta_j^{\,n}$$

where the ζ_j's are the zeros of $f(z)$, and the constant vectors $\vec{c}_j$ satisfy (2.14) and (2.15) for $\zeta = \zeta_j$.

In order that the terms of (2.16) form an absolutely convergent series, only the zeros ζ_j with $|\zeta_j| < 1$ can be included. Denote these zeros by $\zeta_1, \ldots, \zeta_m$. Thus, by Theorem 1.3 the system will be positive recurrent iff constant vectors $\vec{c}_1, \ldots, \vec{c}_m$ can be obtained, satisfying the above conditions, and such that the terms of (2.16) also satisfy the remaining equation (2.4).

Denote the components of the vectors $\vec{c}_j$ by $\begin{bmatrix} c_{1j} \\ c_{2j} \end{bmatrix}$.

One can set $b_j = c_{1j}$ and by (2.14) and (2.15), (2.17) $$c_{2j} = b_j \delta_j$$

where $$\delta_j = \frac{(2\rho+1)\zeta_j - 2\rho}{2\zeta_j^2} = \frac{\zeta_j^2}{2(\rho+1)\zeta_j - 2\rho} .$$

Thus

$$c_j = b_j \begin{bmatrix} 1 \\ \delta_j \end{bmatrix}$$

Substituting (2.16) into (2.4) yields, using (2.7),

(2.18) $$2\rho \sum_{j=1}^{m} b_j = 0 .$$

This system will be positive recurrent iff one can obtain a nontrivial set of constants $b_1, \ldots, b_m$ satisfying (2.18). Clearly this will be possible if and only if $m \geq 2$.

2.6 The Fundamental Polynomial

The preceeding discussion shows that it is necessary to locate the zeros of the fundamental polynomial

$$f(z) = \begin{vmatrix} -(2\rho+1)z + 2\rho & 2z^2 \\ z^2 & -2(\rho+1)z + 2\rho \end{vmatrix}$$

(2.19)

$$= 2\{[(2\rho+1)z - 2\rho][(\rho+1)z - \rho] - z^4\}$$

in relation to the unit circle.

The zeros of this polynomial can be located by elementary methods. The method used here, however, has the advantage of permitting generalization to more complicated systems.

Lemma 2.1: $f(z)$ _has one zero at_ $z = 1$.

Proof: Direct substitution.

Theorem 2.1 (Rouché's Theorem [1, p. 88]): _If_ u _and_ v _are analytic inside and on a closed contour_ C , _and if_ $|v| < |u|$ _on_ C , _then_ u _and_ $u \pm v$ _will have the same number of zeros inside_ C .

Lemma 2.2: _If_ $\rho < 2/3$, $f(z)$ _has exactly two zeros inside the open unit circle_. _If_ $\rho \geq 2/3$, $f(z)$ _has at least two zeros inside the closed unit circle_ $|z| \leq 1$ _and at most one zero inside the open unit circle_ $|z| < 1$.

Proof: On the circle $|z| = 1 + \varepsilon$, $\varepsilon \neq 0$ positive or negative, one has

$$|[(2\rho+1)z-2\rho][(\rho+1)z-\rho]| - |z^4|$$

$$= |(2\rho+1)z-2\rho||(\rho+1)z-\rho| - |z^4|$$

$$\geq [(2\rho+1)(1+\varepsilon) - 2\rho][(\rho+1)(1+\varepsilon) - \rho] - (1+\varepsilon)^4$$

$$= 1 + (3\rho+2)\varepsilon + (2\rho+1)(\rho+1)\varepsilon^2 - (1 + 4\varepsilon + 6\varepsilon^2 + 4\varepsilon^3 + \varepsilon^4)$$

$$= (3\rho - 2)\varepsilon + 0(\varepsilon)$$

> 0 for sufficiently small ε provided $\rho < 2/3$ and $\varepsilon < 0$, or $\rho > 2/3$ and $\varepsilon > 0$.

Thus, if $\rho < 2/3$ we have $|[(2\rho+1)z-2\rho][(\rho+1)z-\rho]| - |z^4| > 0$ on every circle inside the unit circle and sufficiently close to it. By Rouché's Theorem (Theorem 2.1) $f(z)$ will have the same number of zeros as $[(2\rho+1)z-2\rho][(\rho+1)z-\rho]$, namely 2, inside each such circle and hence inside the unit circle $|z| < 1$.

In case $\rho > 2/3$, $\varepsilon > 0$, the same argument shows that $f(z)$ will have exactly two zeros inside the closed unit circle $|z| \leq 1$. One of these is at $z = 1$ so at most one zero can be inside the open unit circle.

To consider the case $\rho = 2/3$, note that the zeros are analytic and hence continuous functions of ρ. Choose $\rho > 2/3$. As $\rho \to 2/3^+$, no zero can leave the closed region $|z| \geq 1$. Thus when $\rho = 2/3$ $f(z)$ will have at least three zeros in $|z| \geq 1$ and hence at most one zero inside $|z| < 1$.

Q.E.D.

Actually one can say much more.

Lemma 2.3: <u>The four zeros</u> $\zeta_1, \zeta_2, \zeta_3$, <u>and</u> ζ_4 <u>of</u> $f(z)$ <u>are all real and satisfy</u>:

$$\zeta_4 = 1$$

$$\zeta_3 < -1$$

$$0 < \zeta_2 < \frac{\rho}{\rho+1}$$

$$\frac{2\rho}{2\rho+1} < \zeta_1 < 1 \quad \text{if} \quad \rho < 2/3$$

$$\zeta_1 = 1 \quad \text{if} \quad \rho = 2/3$$

$$\zeta_1 > 1 \quad \text{if} \quad \rho > 2/3$$

Proof: By direct calculation we see that

$$f(-\infty) = -\infty, \quad f(-1) > 0, \quad f(0) > 0$$

$$f\left(\frac{\rho}{\rho+1}\right) < 0, \quad f\left(\frac{2\rho}{2\rho+1}\right) < 0, \quad f(1) = 0,$$

$$\text{and} \quad f'(1) = 3\rho - 2 .$$

The results follow.

2.7 Conditions for Positive Recurrence

By Theorem 1.3 the system will be positive recurrent if and only if the normal equations have a nontrivial absolutely convergent solution. The discussion in section 2.5 shows that this is true if and only if (2.18) has a nontrivial solution which is true if and only if the number of zeros of $f(z)$ inside the open unit circle is at least 2. Combining this with Lemma 2.2 one obtains.

Theorem 2.2. The 2-server series queueing system is positive recurrent if and only if $\rho < 2/3$.

2.8 Conditions for the System to be Transient

Foster's Theorem (Theorem 1.4) asserts that a necessary and sufficient condition for an irreducible Markov chain to be transient is the existence of a nonconstant bounded solution to a certain system of equations (1.7). This system can be written in the form

$$(2.20) \quad \Big(\sum_{j\neq i} \lambda_{ij}\Big)y(i) = \sum_{j\neq i} \lambda_{ij}\, y(j)\,, \quad \text{for} \quad i = 1\,,2\,,\ldots$$

In terms of the rate diagram one assigns an arbitrary variable $y(i)$ to each vertex (or state) i. At each vertex i (except one, the 0-state), the left side of the equation is $y(i)$ multiplied by the total rate out from that vertex. The right side represents the sum of the variables $x(j)$ over all vertices that can be reached in one step from $x(i)$ each multiplied by the transition rate for this one-step transition.

In the Markov chain of this chapter, assign the arbitrary variables to the states as follows:

> To the lumped state $\{(bs)_n, (se)_n\}$ assign the variable $y_1(n)$, $n \geq 1$. To the simple state $(ss)_n$ assign the variable $y_2(n)$, $n \geq 2$. Also assign $y_2(1)$ to $(es)_1$ and $y_2(0)$ to $(ee)_o$.

Thus we have variables $y_1(n)$, $n \geq 1$ and $y_2(n)$, $n \geq 0$, to appear in the Foster equations.

By refering to the rate diagram Figure 2.3 the equations (2.20) can be set up using the previous discussion. For example, at the state $\{(bs)_{n+1}, (se)_{n+1}\}$ the equation becomes

$$(\lambda+\mu)y_1(n+1) = \lambda y_1(n+2)+\mu y_2(n) .$$

If this is done at every state, except the 0-state $(ee)_o$, and again setting $\rho = \frac{\lambda}{2\mu}$, one obtains

$$(2.21) \qquad -(2\rho+1)y_1(n+1)+2\rho y_1(n+2)+y_2(n) = 0$$

$$(2.22) \qquad 2y_1(n) - 2(\rho+1)y_2(n+1)+2\rho y_2(n+2) = 0$$

for $n \geq 1$, plus the two special equations:

$$(2.23) \qquad -(2\rho+1)y_1(1)+2\rho y_1(2)+y_2(0) = 0$$

$$(2.24) \qquad -(2\rho+1)y_2(1)+2\rho y_2(2)+y_2(0) = 0$$

Actually (2.23) coincides with (2.21) for $n = 0$, so that any solution of the system (2.21) and (2.22) for all n will satisfy (2.23).

The system (2.21) and (2.22) can be written in the form

$$(2.25) \qquad G(E)\vec{y}(n) = \vec{0} \qquad \text{for} \quad n \geq 1$$

where

$$(2.26) \qquad G(E) = \begin{bmatrix} -(2\rho+1)E+2\rho E^2 & 1 \\ 2 & -2(\rho+1)E+2\rho E^2 \end{bmatrix}$$

and

$$(2.27) \qquad \vec{y}(n) = \begin{bmatrix} y_1(n) \\ y_2(n) \end{bmatrix}, \quad \text{for} \quad n \geq 1$$

A similar argument to that in section 2.5 shows that the bounded solutions of the system (2.25) will have the form

$$\vec{y}(n) = \sum_{j=1}^{m} \vec{d}_j \, \zeta_j^n \tag{2.28}$$

where $\zeta_1, \ldots, \zeta_m$ are the zeros of the polynomial

$$g(z) = \det G(z)$$

$$= \begin{vmatrix} -(2\rho+1)z+2\rho z^2 & 1 \\ 2 & -2(\rho+1)z+2\rho z^2 \end{vmatrix} \tag{2.29}$$

inside the open unit circle $|z| < 1$. The constant vectors $\vec{d}_j = \begin{bmatrix} d_{ij} \\ d_{2j} \end{bmatrix}$ are in the null space of $G(z)$.

Taking d_{2j} arbitrary one can solve for d_{ij} in terms of d_{2j} as follows:

$$0 = [-(2\rho+1)\zeta_j+2\rho\zeta_j^2]d_{1j}+d_{2j}=2d_{1j}+2[-(\rho+1)\zeta_j+\rho\zeta_i^2]d_{2j} \tag{2.30}$$

$$d_{1j} = d_{2j}[(\rho+1)\zeta_i-\rho\zeta_j^2] = \frac{d_{2j}}{[(2\rho+1)\zeta_j-2\rho\zeta_j^2]}$$

Substitution into (2.24) yields

$$\sum_{j=1}^{m} d_{2j}(2\rho\zeta_j-1)(\zeta_j-1) = 0 \tag{2.31}$$

The polynomial $g(z)$ is easily seen to be

$$g(z) = z^4 f(1/z) \tag{2.32}$$

where $f(z)$ is the polynomial (2.19) of section 2.6. Thus the number of zeros of $g(z)$ inside the unit circle is the same as the number of zeros of $f(z)$ outside the unit circle.

Theorem 2.3 <u>The 2-server series queueing system is transient if</u> $\rho > 2/3$.

Proof: By Lemma 2.2, if $\rho > 2/3$ $f(z)$ has two zeros inside the closed unit circle. Thus $g(z) = z^4 f(1/z)$ has two zeros inside the open unit circle. Thus $m \geq 2$ in (2.31) and (2.31) has a nontrivial solution for the d_{2j}'s with $|\zeta_j| < 1$. Note that the coefficient $(2\rho\zeta_j - 1)(\zeta_j - 1)$ will not be zero for these terms since $\zeta = 1$ is not inside the circle and $\zeta = 1/(2\rho)$ is not a zero, as is easily seen by direct substitution. The theorem follows from Foster's Theorem, Theorem 1.4.

Q.E.D.

2.9 Conditions for the System to be Recurrent Null

Only the case $\rho = 2/3$ needs to be considered. To do this we use Foster's Criteria, Theorem 1.5. The inequalities

$$\sum_{j=o}^{\infty} \lambda_{ij} y(j) \leq 0 \quad \text{for} \quad i = 1, 2, \ldots$$

are identical to the system (2.21) - (2.24) with the = signs replaced by $\leq$.

Theorem 2.4 _If_ $\rho = 2/3$, _the 2-server series queueing system is recurrent null._

Proof: The inequalities for Foster's Criteria in the case $\rho = 2/3$ become:

$$-7/3\ y_1(n+1) + 4/3\ y_1(n+2) + y_2(n) \leq 0 \ , \quad n \geq 0$$

$$2\ y_1(n) - 10/3\ y_2(n+1) + 4/3\ y_2(n+2) \leq 0 \ , \ n \geq 1$$

and

$$-7/3\ y_2(1) + 4/3\ y_2(2) + y_2(0) \leq 0 \ .$$

By direct substitution one sees that

$$y_1(n) = n - 1 \ , \ y_2(n) = n - 4/3$$

satisfy these inequalities, (actually this substitution makes all the left sides equal to zero). As $n \to \infty$, $y_1(n) \to \infty$ and $y_2(n) \to \infty$. Thus by Foster's Criteria the system is recurrent.

Since by Theorem 2.2 the system is known to not be positive recurrent in the case $\rho = 2/3$, it follows that it must be recurrent null .

Q.E.D.

2.10 Summary of Results

The preceeding results can be summarized by the following:

Theorem 2.5: The 2-server series queueing system is

(i) positive recurrent iff $\rho < 2/3$;

(ii) transient iff $\rho > 2/3$;

(iii) recurrent null iff $\rho = 2/3$.

This can be compared to the well-known [10, p. 116] results for the 2-server parallel queueing system. This system is

(i) positive recurrent iff $\rho < 1$;

(ii) transient iff $\rho > 1$;

(iii) recurrent null iff $\rho = 1$.

Thus confining the customers to a single path has resulted in a system 2/3 as efficient (in terms of the saturation point for recurrence) as a parallel system.

III. FORMULAS AND NUMERICAL EXAMPLES

3.1 General Discussion

In this chapter attention is confined to the positive recurrent case in the two server series queueing system, $\rho < 2/3$. In this case formulas for the limiting probabilities are obtained. Also formulas are obtained for certain other quantities of interest, such as the mean queue length, and the mean waiting time.

Recall (2.19) that the fundamental polynomial is

$$f(z) = 2\{[(2\rho + 1)z - 2\rho][(\rho + 1)z - \rho] - z^4\} .$$

Let ζ_1 and ζ_2 be the two zeros of $f(z)$ inside the unit circle. By Lemma 2.4

$$0 < \zeta_2 < \zeta_1 < 1 .$$

Equation (2.18) is satisfied if $b_1 + b_2 = 0$, so take $b_1 = b$ and $b_2 = -b$. Hence for $n \geq 1$

$$r(n) = b(\zeta_1^n - \zeta_2^n) \tag{3.1}$$

$$p(n) = b(\delta_1\zeta_1^n - \delta_2\zeta_2^n) \tag{3.2}$$

where $\delta_j = \dfrac{(2\rho + 1)\zeta_j - 2\rho}{2\zeta_j^2} = \dfrac{\zeta_j^2}{2(\rho + 1)\zeta_j - 2\rho}$,

$r(n)$ represents the long run probability of a blocked state with queue length n, and $p(n)$ represents the long run probability of an unblocked state with queue length n.

3.2 Explicit Formulas for Limiting Probabilities

The limiting probabilities must satisfy (2.9)

$\frac{1}{2\rho}[r(1) + p(1)] + \sum_{n=1}^{\infty} [r(n) + p(n)] = 1$. Using (3.1)

and (3.2) implies

$$b\left\{\frac{1}{2\rho}\left[\zeta_1(1+\delta_1) - \zeta_2(1+\delta_2)\right] + \sum_{n=1}^{\infty} [(1+\delta_1)\zeta_1^n - (1+\delta_2)\zeta_2^n]\right\} = 1.$$

$$b\left\{\frac{1}{2\rho}\left[\zeta_1(1+\delta_1) - \zeta_2(1+\delta_2)\right] + \frac{(1+\delta_1)\zeta_1}{1-\zeta_1} - \frac{(1+\delta_2)\zeta_2}{1-\zeta_2}\right\} = 1 .$$

$$(3.3)\quad b = \frac{1}{\left\{(1+\delta_1)\zeta_1\left[\frac{1}{2\rho} + \frac{1}{1-\zeta_1}\right] - (1+\delta_2)\zeta_2\left[\frac{1}{2\rho} + \frac{1}{1-\zeta_2}\right]\right\}} .$$

Hence for $n \geq 1$, $r(n)$ and $p(n)$ are given by (3.1), (3.2) and (3.3). The limiting probability that the queue length is n, for $n \geq 1$, thus is

$$(3.4)\qquad r(n) + p(n) = b[(1+\delta_1)\zeta_1^n - (1+\delta_2)\zeta_2^n]$$

From (2.8) one finds that the probability, that the system is idle, is

$$p(o) = \frac{b}{2\rho}[(1+\delta_1)\zeta_1 - (1+\delta_2)\zeta_2] \quad (3.5)$$

3.3 Other Useful Measurements

3.3.1 Mean Queue Length

The mean queue length is

$$(3.6) \quad \begin{cases} m = \sum_{n=1}^{\infty} n[r(n) + p(n)] \\ \quad = b\left\{ \frac{(1+\delta_1)\zeta_1}{(1-\zeta_1)^2} - \frac{(1+\delta_2)\zeta_2}{(1-\zeta_2)^2} \right\} . \end{cases}$$

where b is given by (3.3).

3.3.2 Mean Waiting Times

Let T_1 be the random variable representing the waiting time until service begins, for an arrival at a random time in the stationary case. Define $m_1 = E[T_1]$ as the <u>mean time until service begins</u>. Let N be a random variable representing the queue length in the stationary case.

The time until service begins for an arrival at an instant when the system is in an unblocked state clearly is

$\sum_{i=1}^{[\frac{N}{2}]} X_i$ where the X_i's are independent, identically distributed with X_i having the distribution of $\max(S_1, S_2)$, where S_1 and S_2 are independent negative exponential random variables with parameter μ. X_i represents the time for a system with 2 busy servers to clear out and admit 2 new customers. X_i has distribution function $(1 - e^{-\mu x})^2$, and $E[X_i] = \frac{3}{2\mu}$.

The time until service begins for an arrival when the system is in a blocked state is $S + \sum_{i=1}^{[\frac{N-1}{2}]} X_i$, with X_i the same as above, and S again having a negative exponential distribution with parameter μ, since one server has to finish before the system goes into an unblocked state. Thus

$$m_1 = E[E[T_1|N]]$$

$$(3.7)\begin{cases} = \sum_{n=1}^{\infty} \{[\frac{n}{2}]\frac{3}{2\mu} p(n) + (\frac{1}{\mu} + [\frac{n-1}{2}]\frac{3}{2\mu}) r(n)\} \\ = \frac{b}{2\mu} \sum_{k=1}^{\infty} \Big\{3k[(1+\delta_1)(1+1/\zeta_1)\zeta_1^{2k} - (1+\delta_2)(1+1/\zeta_2)\zeta_2^{2k}] \\ \quad - [\zeta_1^{2k-1}(3\delta_1 + \zeta_1 + 1) - \zeta_2^{2k-1}(3\delta_2 + \zeta_2 + 1)]\Big\} \\ = \frac{b}{2\mu}\left\{ \frac{\zeta_1[2 + 3\delta_1\zeta_1 + \zeta_1^2]}{(1-\zeta_1^2)(1-\zeta_1)} - \frac{\zeta_2[2 + 3\delta_2\zeta_2 + \zeta_2^2]}{(1-\zeta_2^2)(1-\zeta_2)} \right\} \end{cases}$$

where b is given by (3.3).

Another measure of waiting time is $m_2 = E[T_2]$ where T_2 is the time until service is <u>completed</u>. Thus $T_2 = T_1 + S$ where S has a negative exponential distribution with parameter μ. Hence the <u>mean waiting time until service is completed</u> is

$$(3.8) \qquad \begin{cases} m_2 = E[T_2] \\ \quad = E[T_1] + E[S] \\ \quad = 1/\mu + m_1 \end{cases} .$$

Still another measure of waiting time is $m_3 = E[T_3]$, where T_3 is the time until exit from the system. Thus

$$T_3 = T_2 + B$$

where B is a random variable representing the amount of time the customer remains in the system after having completed service (exit blocked). In the case of an arrival to an unblocked system

$$B = \begin{cases} 0 & \text{if } N \text{ is even} \\ X - S_2 & \text{if } N \text{ is odd} \end{cases}$$

and in the case of an arrival to a blocked system

$$B = \begin{cases} X - S_2 & \text{if } N \text{ is even} \\ 0 & \text{if } N \text{ is odd} \end{cases}$$

where as before $X = \max(S_1, S_2)$ and S_1 and S_2 are independent, negative exponential random variables. S_1 represents service time in the front server, and S_2 represents service time in the rear server. The <u>mean time until exit from the system is</u>

$$(3.9) \quad \begin{cases} m_3 = E[T_3] \\ \quad = E[T_2] + E[B] \end{cases}$$

where

$$E[B] = E[E[B|N]]$$

$$(3.1) \quad \begin{cases} = \sum_{k=1}^{\infty} \left\{ (\frac{3}{2\mu} - \frac{1}{\mu}) p(2k-1) + (\frac{3}{2\mu} - \frac{1}{\mu}) r(2k) \right\} \\ = \frac{b}{2\mu} \sum_{k=1}^{\infty} \left\{ (1 + \delta_1/\zeta_1)\zeta_1^{2k} - (1 + \delta_2/\zeta_2)\zeta_2^{2k} \right\} \\ = \frac{b}{2\mu} \left\{ \frac{(\zeta_1 + \delta_1)\zeta_1}{1 - \zeta_1^2} - \frac{(\zeta_2 + \delta_2)\zeta_2}{1 - \zeta_2^2} \right\} . \end{cases}$$

Thus m_3 is given explicitly by (3.3), (3.9) and (3.10).

3.3.3 Other Measurements

One measure of interest is the probability that one of the servers is blocked, p_b . This is of interest as a measure of the inefficiency of the system. When one of the servers is blocked the system cannot operate at full capacity.

$$p_b = \sum_{1}^{\infty} r(n)$$

$$= \sum_{1}^{\infty} b(\zeta_1^n - \zeta_2^n)$$

$$= b\left\{\frac{\zeta_1}{1-\zeta_1} - \frac{\zeta_2}{1-\zeta_2}\right\} .$$

A measure of customer frustration is the probability that the system is blocked given that a customer is waiting, p_w . Customers become frustrated when they must wait even though a server is not busy.

$$p_w = \frac{\sum_{2}^{\infty} r(n)}{\sum_{2}^{\infty} r(n) + \sum_{3}^{\infty} p(n)}$$

$$= \frac{\frac{\zeta_1^2}{1-\zeta_1} - \frac{\zeta_2^2}{1-\zeta_2}}{\frac{\zeta_1^2}{1-\zeta_1}[1 + \delta_1\zeta_1] - \frac{\zeta_2^2}{1-\zeta_2}[1 + \delta_2\zeta_2]} .$$

Another measure of interest is the ratio of probability of the front server being busy to the probability of the rear server being busy, r . This measure would be important if each server had a total service capacity and one server could be given more of this capacity.

$$r = \frac{\sum_{1}^{\infty} p(n) + \sum_{1}^{\infty} p((bs)_n)}{\sum_{2}^{\infty} p(n) + \sum_{1}^{\infty} p((se)_n)}$$

$$= \frac{\sum_{1}^{\infty} p(n) + \frac{1}{2} \sum_{1}^{\infty} r(n)}{\sum_{2}^{\infty} p(n) + \frac{1}{2} \sum_{1}^{\infty} r(n)}$$

$$= 1 + \frac{p(1)}{\sum_{2}^{\infty} p(n) + \frac{1}{2} \sum_{1}^{\infty} r(n)}$$

$$= 1 + \frac{\delta_1\zeta_1 - \delta_2\zeta_2}{\frac{\zeta_1(\delta_1\zeta_1 + 1/2)}{1 - \zeta_1} - \frac{\zeta_2(\delta_2\zeta_2 + 1/2)}{1 - \zeta_2}}$$

3.4 . Numerical Results and Comparisions

In this section, for given values of ρ , certain probabilities and expected values are found. When both the series and parallel systems are positive recurrent, the two systems can be compared by observing the probabilities of certain queue lengths.

Table 3.1: Probabilities of Queue Length n for 2-Server Systems.

$\rho = 1/3$

n	Series System	Parallel System
0	.4460	.5000
1	.2974	.3333
2	.1303	.1111
3	.0621	.0370
4	.0310	.0123
5	.0159	.0041
6	.0082	.0014
7	.0043	.0005
8	.0022	.0002
9	.0012	.0001
10	.0006	.0000

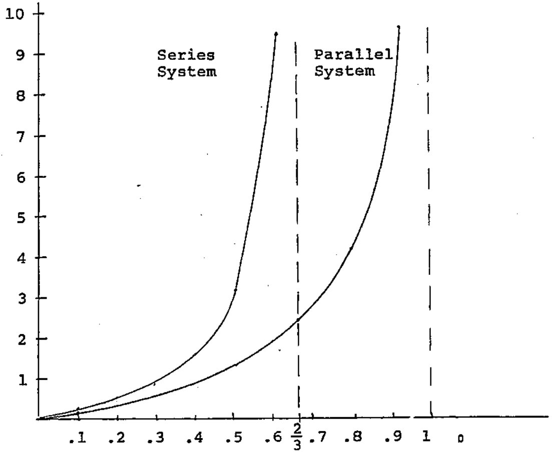
$\rho = 1/2$

n	Series System	Parallel System
0	.2139	.3333
1	.2139	.3333
2	.1444	.1667
3	.1042	.0833
4	.0776	.0417
5	.0586	.0208
6	.0445	.0104
7	.0339	.0052
8	.0258	.0026
9	.0197	.0013
10	.0150	.0006

The probabilities for short queue lengths are higher in the parallel case. For longer queue lengths the probabilities are higher in the series case.

Table 3.2: Mean Queue Length

ρ	Series System	Parallel System
.05	.1015	.1003
.10	.2109	.2020
.15	.3348	.3069
.20	.4807	.4167
.25	.6587	.5333
.30	.8834	.6593
.35	1.1786	.7977
.40	1.5856	.9524
.45	2.1831	1.1285
.50	3.1444	1.3333
.55	4.9394	1.5771
.60	9.4474	1.8750
.65	41.0970	2.2511
2/3	∞	2.4000



Mean Queue Length: Figure 3.1

For small values of ρ the mean queue length of the series queue is about the same as the parallel queue. As ρ approaches 2/3 the mean queue length of the series queue becomes much greater than that of the parallel queue.

Table 3.3: <u>Values of</u> ζ_1 <u>and</u> ζ_2

ρ	ζ_1	ζ_2
.05	.0923	.0475
.10	.1751	.0903
.15	.2532	.1286
.20	.3286	.1632
.25	.4024	.1945
.30	.4754	.2230
.35	.5477	.2490
.40	.6196	.2729
.45	.6913	.2950
.50	.7627	.3154
.55	.8340	.3345
.60	.9052	.3523
.65	.9763	.3690
2/3	1.0000	.3744

As ρ approaches 2/3, ζ_1 approaches 1 and ζ_2 approaches $\frac{-3 + \sqrt{17}}{3} = .3744$.

Theorem 3.1: *As* $\rho \to 0$, ζ_1 *is asymptotic to* 2ρ *and* ζ_2 *is asymptotic to* ρ, *that is* $\zeta_1 = 2\rho + o(\rho)$ *and* $\zeta_2 = \rho + o(\rho)$ *as* $\rho \to 0$.

Proof: The zeros ζ_1 and ζ_2, being analytic functions of ρ, can be represented in the form $\zeta_j = a_o + a_1\rho + a_2\rho^2 + o(\rho^2)$. ζ_1 and ζ_2 are the zeros of $f(z)$ inside the unit circle. $f(z) = (z-1)g(z)$ where $g(z) = z^3 + z^2 - \rho(2\rho + 3)z + 2\rho^2$ and so ζ_1 and ζ_2 are zeros of $g(z)$. Lemma 2.4 states that ζ_1 and ζ_2 are positive, hence substituting $a_o + a_1 + a_2\rho^2 + o(\rho^2)$ for z in $g(z)$ shows that $a_o = 0$. Hence

$$(a_1\rho + a_2\rho^2 + o(\rho^2))^3 + (a_1\rho + a_2\rho^2 + o(\rho^2))^2$$

$$-\rho(2\rho + 3)(a_1\rho + a_2\rho^2 + o(\rho^2)) + 2\rho^2 = 0$$

$$\rho^2[a_1^2 - 3a_1 + 2] + o(\rho^2) = 0$$

$$a_1^2 - 3a_1 + 2 + \frac{o(\rho^2)}{\rho^2} = 0$$

Let $\rho \to 0$

$$a_1^2 - 3a_1 + 2 = 0$$

Therefore $a_1 = 1$ or 2. Thus for small ρ

$$\zeta_1 = 2\rho + o(\rho)$$

$$\zeta_2 = \rho + o(\rho)$$

Q.E.D.

The conclusion of Theorem 3.1 is observed in Table 3.3 for $\rho = .05$, i.e. $\zeta_1 = .0923 \approx 2\rho$, and $\zeta_2 = .0475 \approx \rho$.

Table 3.4 presents the mean waiting times for the 2-server parallel and series queueing systems. Columns four and five refer to the parallel system, in which w_1 is the mean time until service begins and w_2 is the mean time until service is finished and the customer leaves the system. The values in the table below are for $\mu = 1$. For other μ simply divide the values by μ.

Table 3.4: Mean Waiting Times

($\mu = 1$)

ρ	m_1	m_2	m_3	w_1	w_2
.05	.0150	1.0150	1.0571	.0025	1.0025
.10	.0544	1.0544	1.1285	.0101	1.0101
.15	.1160	1.1160	1.2161	.0230	1.0230
.20	.2018	1.2018	1.3239	.0417	1.0417
.25	.3173	1.3173	1.4587	.0667	1.0667
.30	.4723	1.4723	1.6308	.0989	1.0989
.35	.6838	1.6838	1.8579	.1396	1.1396
.40	.9820	1.9820	2.1703	.1905	1.1905
.45	1.4257	2.4257	2.6272	.2539	1.2539
.50	2.1444	3.1444	3.3583	.3333	1.3333
.55	3.4904	4.4904	4.7159	.4337	1.4337
.60	6.8728	7.8728	8.1092	.5625	1.5625
.65	30.6131	31.6131	31.8597	.7316	1.7316
2/3	∞	∞	∞	.8000	1.8000

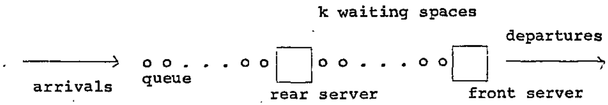
Just as with mean queue lengths one sees that for small values of ρ the series and parallel queues have about the same waiting times, however for large values of ρ the mean waiting times for the series queue are much larger than those for the parallel queue.

IV. THE 2-SERVER k WAITING SPACES SERIES SYSTEM

4.1 Definition of the Problem

One possible extension of the simple two-server series queueing problem is a system with two servers in series but with k waiting spaces between the two servers. Customers are served on a first-come-first-served basis. Each customer is served by one of the two servers. Customers entering the servers move to the furthest forward accessible server. A customer having completed service at the rear server moves into the furthest forward available waiting space to wait until he can leave the system.

The states of the system can be represented by symbols of the form $(xy)_{n,i}$ where n represents the number of customers who have not yet completed service (the queue length), x and y are server symbols representing the status of the rear and front servers respectively, and i indicates the number of occupied waiting spaces. y can be 's' or 'e' depending on whether the front server is busy or empty. x can be 's', 'e', or 'b' depending on whether the rear server is busy, empty, or blocked. (Blocking of the rear server is possible only if all waiting spaces are full.)



<u>Diagram of System:</u> Figure 4.1

The states of the systen are

$$(ee)_{o,o}$$

$$(es)_{1,o},\ (es)_{1,1},\ \ldots,\ (es)_{1,k},\ (bs)_{1,k},\ (s,e)_{1,o}$$

$$(ss)_{2,o},\ (ss)_{2,1},\ \ldots,\ (ss)_{2,k},\ (bs)_{2,k},\ (s,e)_{2,o}$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$(ss)_{n,o},\ (ss)_{n,1},\ \ldots,\ (ss)_{n,k},\ (bs)_{n,k},\ (s,e)_{n,o}$$

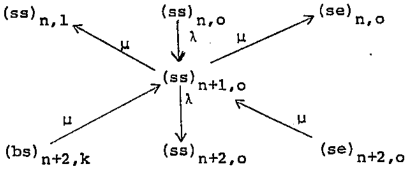
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

Assume that arrivals form a Poisson process with intensity λ, and assume that each server has a mean service rate of μ, where service times are exponentially distributed with probability density function $\mu e^{-\mu t}$. Assume also that all interarrival and service times are mutually independent.

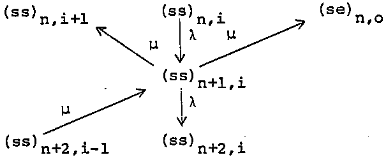
With these assumptions the process forms a continuous-time, discrete-state, time-homogeneous Markov chain of the type considered in Chapter 1.

4.2 Rate Diagram

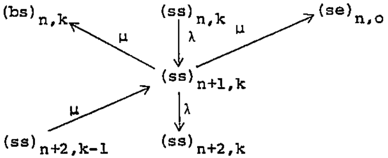
The one-step transitions for this Markov chain are generated by arrivals and service completions. Each arrival increases the queue length by one (and changes a server symbol from e to s if a server is accessible). Each service completion reduces the queue length by one and may change the number in the waiting space as well as one or both server symbols. For example a completion of the rear server would change the state $(ss)_{5,4}$ to $(ss)_{4,5}$ provided $k \geq 5$. The probability of an arrival in time h is $\lambda h + o(h)$, so the transition rate for a transition generated by an arrival will be λ. The probability of a completion from one busy server in time h is $\mu h + o(h)$ so the transition rate for a transition generated by a service completion will be μ. Hence the rate diagram will have vertices and weighted edges made up by fitting together subgraphs of the following types.



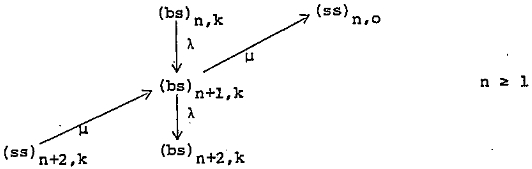
$n \geq 2$

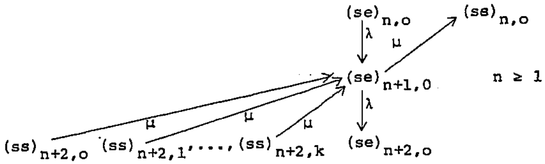


$o < i < k$, $n \geq 2$



$n \geq 2$



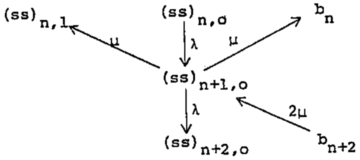


<u>Some Vertices of the Rate Diagram</u>: Figure 4.2

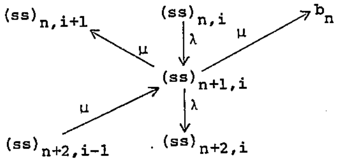
In addition there will be special subgraphs around the vertices $(ss)_{2,i}$, $(es)_{1,i}$, $(bs)_{1,i}$ for $o \le i \le k$, as well as $(se)_{1,o}$, and $(ee)_{o,o}$.

4.3 Lumping

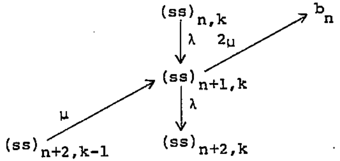
Just as the states $(se)_n$ and $(bs)_n$ were lumped in Chapter 2, the states $(se)_{n,o}$ and $(bs)_{n,k}$ can be lumped here. Departures from either of these two states lead to the same state $(ss)_{n-1,o}$ at the same rate μ . Thus we can lump these two states into a single state $b_n = \{(se)_{n,o}, (bs)_{n,k}\}$ and the Markov property is retained. Some of the vertices of the rate diagram, for the lumped process, are as follows:



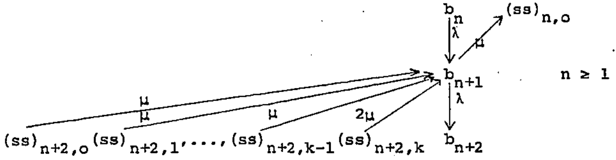
$n \geq 2$



$o < i < k$, $n \geq 2$



$n \geq 2$



Some Vertices of the Rate Diagram for the Lumped Process: Figure 4.3

Again the vertices $(xy)_{n,i}$ with $n = 0, 1, 2$ require special treatment.

Denote the long-run probabilities of this irreducible Markov chain by

$$p_i(n) = P\{(ss)_{n,i}\}$$

$$= \lim_{t\to\infty} p_{\alpha,(ss)_{n,i}}(t) \quad \text{for} \quad n \geq 2 \quad 0 \leq i \leq k$$

$$p_{k+1}(n) = P\{b_n\} = \text{probability of queue length } n \text{ with one server blocked}$$

$$= \lim_{t\to\infty} p_{\alpha,b_n}(t) \qquad \text{for} \quad n \geq 1$$

$$p_o(0) = P\{(ee)_{o,o}\}$$

$$p_i(1) = P\{(es)_{1,i}\} \;, \quad 0 \le i \le k$$

$$p_{k+1}(1) = P\{b_1\}$$

where α represents an arbitrary initial state.

4.4 Normal Equations

The normal equations are obtained from figure 4.3, balancing input and output rates at each state, as in Chapter 2. The result for general n, $n+1 \ge 2$ is:

$$-(\lambda + 2\mu)p_i(n+1) + \lambda p_i(n) + \mu p_{i-1}(n+2) = 0$$

$$1 \le i \le k$$

$$-(\lambda + \mu)p_{k+1}(n+1) + \lambda p_{k+1}(n) + \sum_{i=o}^{k} \mu p_i(n+2)$$

$$+ \mu p_k(n+2) = 0$$

$$-(\lambda + 2\mu)p_o(n+1) + \lambda p_o(n) + \mu p_{k+1}(n+2) = 0 \,.$$

In addition, the normal equations for queue length 0 and 1 are:

$$-(\lambda+\mu)p_o(1) + \lambda p_o(0) + \mu p_{k+1}(2) = 0$$

$$-(\lambda+\mu)p_i(1) + \mu p_{i-1}(2) = 0 \quad 1 \le i \le k$$

$$-\lambda p_o(0) + \sum_{i=o}^{k+1} \mu p_i(1) = 0$$

This system can be rewritten, setting $\rho = \frac{\lambda}{2\mu}$,

(4.1) $$-2(\rho+1)p_i(n+1) + 2\rho p_i(n) + p_{i-1}(n+2) = 0$$

$$1 \le i \le k, \quad n \ge 1$$

(4.2) $$-(2\rho+1)p_{k+1}(n+1) + 2\rho\, p_{k+1}(n) + \sum_{i=o}^{k} p_i(n+2)$$

$$+ \rho_k(n+2) = 0, \quad n \ge 1$$

(4.3) $$-2(\rho+1)p_o(n+1) + 2\rho\, p_o(n) + p_{k+1}(n+2) = 0, n \ge 1 .$$

Also the normal equations for queue length 0 and 1 become:

(4.4) $$-(2\rho+1)p_o(1) + 2\rho\, p_o(0) + p_{k+1}(2) = 0$$

(4.5) $$-(2\rho+1)p_i(1) + p_{i-1}(2) = 0 \quad 1 \le i \le k$$

(4.6) $$-(2\rho+1)p_{k+1}(1) + \sum_{i=o}^{k} p_i(2) + p_k(2) = 0$$

$$(4.7) \qquad -2\rho\, p_o(0) + \sum_{i=o}^{k+1} p_i(1) = 0 .$$

This system is dependent and thus one equation say (4.4), can be disregarded. Note that $p_o(0)$ appears only in (4.7). Thus (4.7) is automatically satisfied if it is used to determine $p_o(0)$ in terms of the other quantities,

$$p_o(0) = \frac{1}{2\rho} \sum_{i=o}^{k+1} p_i(1)$$

By Theorem 1.3 the Markov chain will be positive recurrent if the normal equations, with the probabilities $p_i(n)$ replaced by arbitrary variables $x_i(n)$, have a nontrivial absolutely convergent solution. If this is the case any such solution, normalized to form a probability distribution, will give the limiting probabilities:

$$p_i(n) = cx_i(n) , \quad \text{and} \quad \sum_n \sum_i p_i(n) + p_o(0) = 1 .$$

Substituting $\frac{1}{2\rho} \sum p_i(1)$ for $p_o(0)$ one obtains

$$(4.8) \qquad \frac{1}{2\rho} \sum_{i=o}^{k+1} p_i(1) + \sum_{n=1}^{\infty} \sum_{i=o}^{k+1} p_i(n) = 1$$

as the normalizing equation. Hence the problem reduces to finding sufficient conditions for the system (4.1) - (4.3) and (4.5) - (4.6) to have a nontrivial, absolutely convergent solution.

4.5 Difference Equations

The general system (4.1) - (4.3) (omitting the special equations (4.5) - (4.6)) consitutes a vector difference equation. Using the shift operator E this system can be written in the form

$$(4.9) \qquad F_k(E)\vec{x}(n) = \vec{0}$$

$n \geq 1$, where

$$F_k(E) = \begin{bmatrix} R_E & 0 & 0 & \cdots & 0 & 0 & E^2 \\ E^2 & R_E & 0 & \cdots & 0 & 0 & 0 \\ 0 & E^2 & R_E & \cdots & 0 & 0 & 0 \\ 0 & 0 & E^2 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & E^2 & R_E & 0 \\ E^2 & E^2 & E^2 & \cdots & E^2 & 2E^2 & S_E \end{bmatrix}_{(k+2)\times(k+2)}$$

where $R_E = [-2(\rho+1)E + 2\rho]$ and $S_E = [-(2\rho+1)E + 2\rho]$.

and

$$\vec{x}(n) = \begin{bmatrix} x_o(n) \\ x_1(n) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_k(n) \\ x_{k+1}(n) \end{bmatrix}$$

The fundamental solutions of such a linear homogeneous difference equation have the form $\vec{c}\,\zeta^n$ where ζ is an appropriate nonzero scalar and

$$\vec{c} = \begin{bmatrix} c_o \\ c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_k \\ c_{k+1} \end{bmatrix}$$

is an appropriate constant vector.

Substituting this into (4.9) one obtains

$$\begin{aligned}\vec{0} &= F_k(E)\vec{c}\ \zeta^n \\ &= F_k(E)\,\zeta^n\,\vec{c} \\ &= F_k(\zeta)\,\zeta^n\,\vec{c} \\ &= F_k(\zeta)\vec{c}\ \zeta^n\end{aligned}$$

Thus $F_k(\zeta)\ \vec{c} = \vec{0}$ and one concludes that $\vec{c}\ \zeta^n$ is a solution iff $F_k(\zeta)$ is a singular matrix and $\vec{c}$ is a vector in its null space. Hence ζ must be a zero of the <u>fundamental</u> <u>polynomial</u> $f_k(z) = \det F_k(z)$, and the corresponding $\vec{c}$ must satisfy

$$F_k(\zeta)\ \vec{c} = \vec{0}\ . \tag{4.10}$$

4.6 The Zeros of $f_k(z)$

Further analysis depends on knowledge of the location of the zeros of $f_k(z)$. The following is an explicit representation of $f_k(z)$.

Theorem 4.1 $f_k(z) = (-1)^{k+1}\{[-(2\rho+1)z+2\rho][2(\rho+1)z-2\rho]^{k+1} \cdot$

$$+ \sum_{i=1}^{k} z^{2(i+1)}[2(\rho+1)z-2\rho]^{k+1-i}+2z^{2(k+2)}\}$$

Proof:

$f_k(z) =$

$$\begin{vmatrix} R_z & 0 & 0 & \cdots & 0 & 0 & z^2 \\ z^2 & R_z & 0 & \cdots & 0 & 0 & 0 \\ 0 & z^2 & R_z & \cdots & 0 & 0 & 0 \\ 0 & 0 & z^2 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & z^2 & R_z & 0 \\ z^2 & z^2 & z^2 & \cdots & z^2 & 2z^2 & S_z \end{vmatrix}_{(k+2)x(k+2)}$$

where $R_z = [-2(\rho+1)z+2\rho]$ and $S_z = [-(2\rho+1)z + 2\rho]$.

From this one obtains

$$f_k(z) = [-(2\rho+1)z+2\rho][-2(\rho+1)z+2\rho]^{k+1}+(-1)^{k+1}z^2H_{k+1}$$

by expanding the last column, where

$H_\ell =$

$$\begin{vmatrix} z^2 & R_z & 0 & \cdots & 0 & 0 \\ 0 & z^2 & R_z & \cdots & 0 & 0 \\ 0 & 0 & z^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & z^2 & R_z \\ z^2 & z^2 & z^2 & \cdots & z^2 & 2z^2 \end{vmatrix}$$

$2 \le \ell \le k+1 \qquad\qquad \ell \times \ell$

where $R_z = [-2(\rho+1)z + 2\rho]$. The quantity H_ℓ is seen to satisfy the recursion relationship by expanding by the first column:

$$H_\ell = (-1)^{\ell-1}\,[-2(\rho+1)z+2\rho]^{\ell-1} + z^2 H_{\ell-1}$$

where $H_1 = 2z^2$

Hence

$$f_k(z) = [-(2\rho+1)z+2\rho][-2(\rho+1)z+2\rho]^{k+1}$$

$$+ (-1)^{k+1}z^2\{(-1)^k z^2[-2(\rho+1)z+2\rho]^k$$

$$+ z^2\{(-1)^{k-1}z^2[-2(\rho+1)z+2\rho]^{k-1}$$

$$+ \dots + z^2\{(-1)z^2[-2(\rho+1)z+2\rho] + 2z^4\}\cdots\}\} .$$

$$= (-1)^{k+1}\{[-(2\rho+1)z+2\rho][2(\rho+1)z-2\rho]^{k+1}$$

$$+ \sum_{i=1}^{k} z^{2(i+1)}[2(\rho+1)z-2\rho]^{k+1-i} + 2z^{2(k+2)}\} .$$

Q.E.D.

Certain modifications in the subsequent arguments are necessary if $f_k(z)$ has multiple zeros inside the unit circle. For simplicity only the case where $f_k(z)$ has distinct zeros inside $|z| < 1$ will be considered in detail. Hence the null space of $F_k(\zeta)$ (see [2, p. 253]), where ζ is a zero of $f_k(z)$, is 1-dimensional and the corresponding $\vec{c}$ is nontrivial and unique up to a multiplicative constant. The general solution to the system (4.1) - (4.7) has the form

$$\vec{p}(x) = \sum_j \vec{c}_j \, \zeta_j^{\,n} \tag{4.11}$$

where the ζ_j's are the distinct zeros of $f_k(z)$ and the constant vectors $\vec{c}_j$ satisfy (4.10) for $\zeta = \zeta_j$.

In order that the terms of (4.11) form an absolutely convergent series, only the zeros ζ_j with $|\zeta_j| < 1$ can be included. Denote these zeros by $\zeta_1, \ldots, \zeta_m$. Thus by Theorem 1.3, the system will be positive recurrent iff constant vectors $\vec{c}_1, \ldots, \vec{c}_m$ can be obtained, satisfying the above conditions and also the remaining equations (4.5) and (4.6).

In order for $\vec{c}$ to satisfy (4.10)

$$(4.12) \qquad [-2(\rho + 1)\zeta + 2\rho]c_o + c_{k+1}\zeta^2 = 0$$

$$(4.13) \qquad [-2(\rho + 1)\zeta + 2\rho]c_i + c_{i-1}\zeta^2 = 0 \quad 1 \le i \le k$$

$$(4.14) \qquad [-(2\rho + 1)\zeta + 2\rho]c_{k+1} + \sum_{i=o}^{k} c_i\zeta^2 + c_k\zeta^2 = 0$$

Using equation (4.13), for each j the components c_{ij} can be represented in terms of c_{oj}

$$c_{ij} = \left[\frac{\zeta_j^2}{2(\rho+1)\zeta_j - 2\rho}\right] c_{i-1j} = \left[\frac{\zeta_j^2}{2(\rho+1)\zeta_j - 2\rho}\right]^i c_{oj}$$

for $0 \le i \le k$

From equation (4.12), c_{oj} can be represented in terms of

$$c_{k+1j} = a_j \ , \quad c_{oj} = \left[\frac{\zeta_j^2}{[2(\rho+1)\zeta_j - 2\rho]}\right] a_j \quad \text{and thus}$$

$$c_{ij} = \delta_j^{i+1} \ a_j \quad \text{for} \quad 0 \le i \le k$$

where $\delta_j = \dfrac{\zeta_j^2}{2(\rho+1)\zeta_j - 2\rho}$. Equation (4.14) must also be shown to hold for these c_{ij} . Substituting $\left[\dfrac{\zeta_j^2}{2(\rho+1)\zeta_i - 2\rho}\right]^{i+1} a_j$ for c_{ij} into the left side of (4.14) results in

$$[-(2\rho+1)\zeta_j + 2\rho]a_j + \zeta_j^2 \left[\frac{\zeta_j^2}{2(\rho+1)\zeta_j - 2\rho}\right]^{k+1} a_j$$

$$+ \sum_{i=0}^{k} \zeta_j^2 \left[\frac{\zeta_j^2}{2(\rho+1)\zeta_j - 2\rho}\right]^{(i+1)} a_j \ .$$

To show this is zero, divide through by $\dfrac{a_j}{[2(\rho+1)\zeta_j - 2\rho]^{k+1}}$

yielding

$$[-(2\rho+1)\zeta_j+2\rho][2(\rho+1)\zeta_j-2\rho]^{k+1} + \zeta_j^{2(k+2)}$$

$$+ \sum_{i=o}^{k} \zeta_j^{2(i+2)}[2(\rho+1)\zeta_j-2\rho]^{(k+1)-(i+1)}$$

$$= [-(2\rho+1)\zeta_j+2\rho][2(\rho+1)\zeta_j-2\rho]^{k+1} + 2\zeta_j^{2(k+2)}$$

$$+ \sum_{i=o}^{k-1} \zeta_j^{2(i+2)}[2(\rho+1)\zeta_j-2\rho]^{(k+1)-(i+1)}$$

$$= [-(2\rho+1)\zeta_j+2\rho][2(\rho+1)\zeta_j-2\rho]^{k+1} + 2\zeta_i^{2(k+2)}$$

$$+ \sum_{i=1}^{k} \zeta_j^{2(i+1)} [2(\rho+1)\zeta_j-2\rho]^{k+1-i}$$

$$= f(\zeta_i)$$

$$= 0$$

since ζ_j is a zero of $f_k(z)$. Thus equation (4.14) holds with

$$c_{ij} = \left[\frac{\zeta_j^2}{2(\rho+1)\zeta_j-2\rho}\right]^{i+1} a_j = \delta_j^{i+1} a_j \quad 0 \le j \le k, \quad \text{and}$$

$$c_{k+1j} = a_j$$

It was shown above that only zeros inside the unit circle can be used.

Theorem 4.2 <u>A nontrivial solution to the system</u> (4.1)-(4.7) <u>exists if</u> $f_k(z)$ <u>has at least</u> $k + 2$ <u>zeros inside the unit circle</u>.

Proof: There are $k + 1$ initial conditions on the a_j's that must be satisfied. There are the same number of a_j's as zeros of $f_k(z)$ inside the open unit circle. Let B be the coefficient matrix of this system. If there are $m \geq k + 2$ zeros then B is a $(k + 1)x(m)$ matrix and hence there is a nontrivial solution to the system $BA = 0$ where $A^t = (a_1, a_2, \ldots, a_m)$.

4.7 The Fundamental Polynomial

From the preceeding one sees that it is necessary to obtain further information about the zeros of the fundamental polynomial.

$f_k(z) =$

$$\begin{vmatrix} R_z & 0 & 0 & \cdots & 0 & 0 & z^2 \\ z^2 & R_z & 0 & \cdots & 0 & 0 & 0 \\ 0 & z^2 & R_z & \cdots & 0 & 0 & 0 \\ 0 & 0 & z^2 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & z^2 & R_z & 0 \\ z^2 & z^2 & z^2 & \cdots & z^2 & 2z^2 & S_z \end{vmatrix}_{(k+2)\times(k+2)}$$

where $R_z = [-2(\rho + 1)z + 2\rho]$ and $S_z = [-(2\rho+1)z + 2\rho]$.

Lemma 4.1: $f_k(z)$ _has at least one zero at_ $z = 1$.

Proof: Direct substitution.

Definition: $$\rho_k = \frac{3\cdot 2^k-1}{4\cdot 2^k-1} .$$

Lemma 4.2: _If_ $\rho < \rho_k$, $f_k(z)$ _has exactly_ $k + 2$ _zeros inside the open unit circle._ _If_ $\rho \geq \rho_k$, $f_k(z)$ _has at most_ $k + 1$ _zeros inside the open unit circle._

Proof: A recurrence relation for $f_k(z)$ can be obtained by noting that $f_k(z) = (-1)^{k+1}h(z)$ where

$$h(z) = \begin{vmatrix} z^2 & R_z & 0 & \cdots & 0 & 0 \\ 0 & z^2 & R_z & \cdots & 0 & 0 \\ 0 & 0 & z^2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & z^2 & R_z \\ S_z & z^2 & z^2 & \cdots & z^2 & 2z^2 \end{vmatrix}_{(k+2)\times(k+2)}$$

where $R_z = [-2(\rho+1)z+2\rho]$ and $S_z = [-(2\rho+1)z+2\rho]$.

The zeros of $f_k(z)$ are the same as the zeros of $h(z)$. Now expanding by the last column results in

$$h(z) = 2z^{2(k+2)} - [-2(\rho+1)z+2\rho]w_k \tag{4.15}$$

where

$$w_k = \begin{vmatrix} z^2 & R_z & 0 & \cdots & 0 & 0 \\ 0 & z^2 & R_z & \cdots & 0 & 0 \\ 0 & 0 & z^2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & z^2 & R_z \\ S_z & z^2 & z^2 & \cdots & z^2 & z^2 \end{vmatrix}_{(k+1)\times(k+1)}$$

where $R_z = [-2(\rho+1)z + 2\rho]$ and $S_z = [-(2\rho+1)z + 2\rho]$.

If one continues to expand by the last column, the following recurrence relation is obtained.

$$w_\ell = z^{2(\ell+1)} - [-2(\rho+1)+2\rho]w_{\ell-1} \quad \text{for} \quad 1 \le \ell \le k$$

where $w_o = -(2\rho+1)z+2\rho$.

On the circle $|z| = 1 + \varepsilon$, $\varepsilon \neq 0$, the inequality $|a-b| \geq |a| - |b|$ yields $|w_o| \geq (2\rho+1)(1+\varepsilon) - 2\rho$ $= 1 + (2\rho+1)\varepsilon$

Similarly $|w_1| \geq -(1+\varepsilon)^4 + [2(\rho+1)(1+\varepsilon)-2\rho]\,|w_o|$

$$\geq 1 + 6\rho\varepsilon + O(\varepsilon^2)$$

By continuing this process one obtains a sequence $d_o, d_1, \ldots, d_k$ of constants such that

$$|w_\ell| \geq 1 + d_\ell\varepsilon + O(\varepsilon^2)$$

for $1 \leq \ell \leq k$ on the circle $|z| = 1 + \varepsilon$. From above one can take $d_o = 2\rho + 1$ and $d_1 = 6\rho$. In general

$$|w_{\ell+1}| \geq -(1+\varepsilon)^{2(\ell+2)} + [2(\rho+1)(1+\varepsilon)-2\rho]\,|w_\ell|$$

$$\geq -(1+\varepsilon)^{2(\ell+2)} + [2(\rho+1)(1+\varepsilon) - 2\rho][1+d_\ell\varepsilon+O(\varepsilon^2)$$

$$= 1 + \varepsilon(2\rho+2d_\ell-2\ell-2) + O(\varepsilon^2)$$

Hence

$$|w_{\ell+1}| \geq 1 + d_{\ell+1} + O(\varepsilon^2)$$

if $$d_{\ell+1} = 2\rho + 2d_\ell - 2\ell - 2\ .$$

This difference equation for the d_ℓ's has the solution

(4.16) $$d_\ell = (4\rho - 3)2^\ell - 2\rho + 4 + 2\ell$$

which can be checked inductively.

Thus, on $|z| = 1 + \varepsilon$,

$$|w_k| \geq 1 + \{(4\rho-3)2^k + (-2\rho+4+2k)\}\varepsilon + O(\varepsilon^2)$$

and so

$$\begin{aligned}|\{-2(\rho+1)+2\rho\}w_k| &\geq (2+2(\rho+1)\varepsilon)(1+d_k\varepsilon+O(\varepsilon^2))\\ &= 2 + 2(\rho+1)\varepsilon + 2d_k\varepsilon + O(\varepsilon^2)\\ &= 2+2(\rho+1)\varepsilon+2^{k+1}(4\rho-3)\varepsilon-4\rho\varepsilon+8\varepsilon+4k\varepsilon+O(\varepsilon^2)\\ &= 2+[2^{k+1}(4\rho-3)-2\rho+10+4k]\,\varepsilon+O(\varepsilon^2) .\end{aligned}$$

Hence the magnitudes, of the two terms making up $h(z)$, see (4.15), can be compared on $|z| = 1 + \varepsilon$.

(4.17) $$|\{-2(\rho+1)z+2\rho\}w_k| - |2z^{2(k+2)}|$$

$$\geq 2+2[(4\rho-3)2^k-\rho+5+2k]\varepsilon-2(1+(2k+4)\varepsilon) + O(\varepsilon^2)$$

$$= 2[(4\cdot 2^k-1)\rho-3\cdot 2^k+1]\varepsilon + O(\varepsilon^2)$$

<u>Case</u> (i) $\rho < \rho_k$

Whenever $\rho < \rho_k = \dfrac{3\cdot 2^k-1}{4\cdot 2^k-1}$, and $\varepsilon < 0$ and

sufficiently small, the right side of (4.17) is positive. By Rouché's Theorem $h(z)$, given by (4.11), must have the same number of zeros inside the circle $|z| = 1 + \varepsilon$ as $[-2(\rho+1)z+2\rho]w_k$. By letting $\varepsilon \to 0^-$ one concludes that the same is true inside the open unit circle. Thus we need only determine the number of zeros of $[-2(\rho+1)z+2\rho]w_k$ inside the open unit circle. Now

$$\begin{aligned}
|[-2(\rho+1)z+2\rho]w_{k-1}| &\geq [2(\rho+1)(1+\varepsilon)-2\rho]\,|w_{k-1}| \\
&\geq [2+2(\rho+1)\varepsilon][1+d_{k-1}\varepsilon+O(\varepsilon^2)] \\
&= [2+2(\rho+1)\varepsilon]\{1+[(4\rho-3)2^{k-1}-2\rho+4 \\
&\quad +2(k-1)]\varepsilon+O(\varepsilon^2)\} \\
&= 2 + O(\varepsilon) \\
&> 1 + O(\varepsilon) \\
&= |z^{2(k+1)}| .
\end{aligned}$$

Hence $|[-2(\rho+1)z+2\rho]w_{k-1}| - |z^{2(k+1)}| > 0$, and it follows from Rouché's Theorem that $w_k = -z^{2(k+1)} + [-2(\rho+1)z+2\rho]w_{k-1}$ has the same number of zeros inside the circle $|z| = 1 + \varepsilon$ as $[-2(\rho+1)z+2\rho]w_{k-1}$ has. Since the first factor has one zero, $2\rho/(2\rho+2)$ inside the unit circle, on letting $\varepsilon \to 0$ one sees that w_k has one more zero inside the unit circle than w_{k-1}.

$$\begin{aligned}|[-2(\rho+1)z+2\rho]w_{\ell-1}| &\geq [2(\rho+1)(1+\varepsilon)-2\rho]|w_{\ell-1}| \\ &\geq [2+2(\rho+1)\varepsilon][1+[(4\rho-3)2^{\ell-1}-2\rho+4 \\ &\qquad +2(\ell-1]\varepsilon+O(\varepsilon^2)] \\ &= 2+O(\varepsilon) \\ &> 1+O(\varepsilon) \\ &= |z^{2(\ell+1)}| \quad .\end{aligned}$$

Therefore $w_\ell = [-2(\rho+1)z+2\rho]w_{\ell-1}-z^{2(\ell+1)}$ $1 \leq \ell \leq k$ has the same number of zeros inside the circle $|z| = 1+\varepsilon$, $\varepsilon < 0$, as $[-2(\rho+1)z+2\rho]w_{\ell-1}$. Letting $\varepsilon \to 0^-$ it can be seen that w_ℓ has one more zero inside the unit circle than $w_{\ell-1}$ for $1 \leq \ell \leq k$. Since $w_0 = -(2\rho+1)z+2\rho$ has one zero inside the unit circle, one concludes that $h(z)$, and hence $f_k(z)$, has exactly $k+2$ zeros inside the unit circle.

<u>Case</u> (ii) $\rho > \rho_k$

If $\rho > \rho_k = \dfrac{3\cdot 2^k-1}{4\cdot 2^k-1}$, $\varepsilon > 0$ and sufficiently small,

then $|[-2(\rho+1)z+2\rho]w_k| - |2z^{2(k+2)}| > 0$ from (4.17). By Rouché's Theorem $h(z)$ has the same number of zeros inside the circle $|z| = 1+\varepsilon$, $\varepsilon > 0$, as $[-2(\rho+1)z+2\rho]w_k$. By letting $\varepsilon \to 0^+$ one concludes that the same is true inside the closed unit circle. As above we see that w_k has $k+1$ zeros inside the closed unit circle and so

$h(z)$ has $k+2$ zeros inside the closed unit circle. Since one zero is $z = 1$ there are at most $k+1$ zeros of $h(z)$ and hence of $f_k(z)$ inside the open unit circle.

Case (iii) $\rho = \rho_k$

Note that the zeros of $h(z)$ are analytic and hence continuous functions of ρ. Choose $\rho > \rho_k$. As $\rho \to \rho_k^+$ no zero can leave the closed region $|z| \geq 1$. Note $h(z)$ is of degree $2(k+2)$ and $h(z)$ has at least $k+3$ zeros in $|z| \geq 1$ since there are at most $k+1$ zeros in $|z| < 1$. Thus when $\rho = \rho_k$ $h(z)$ and hence $f_k(z)$ has at least $k+3$ zeros in $|z| \geq 1$ and hence at most $k+1$ zeros in $|z| < 1$.

Q.E.D.

Lemma 4.3. There are no zeros of $f_k(z)$ on the unit circle except $z = 1$.

Proof: $f_k(z) = (-1)^{k+1}\{-[(2\rho+1)z-2\rho][2(\rho+1)z-2\rho]^{k+1}$

$$+ 2z^{2(k+2)} + \sum_{i=1}^{k} z^{2(i+1)}[2(\rho+1)z-2\rho]^{k+1-i}\} .$$

Suppose $f_k(z) = 0$, $|z| = 1$, $z \neq 1$ then

$$|2(\rho+1)z-2\rho| = 2(\rho+1)\left|z - \frac{\rho}{\rho+1}\right|$$

$$> 2(\rho+1)\left(1 - \frac{\rho}{\rho+1}\right)$$

$$= 2$$

$$|(2\rho+1)z-2\rho| = (2\rho+1)\left|z - \frac{2\rho}{2\rho+1}\right|$$

$$> (2\rho+1)\left(1 - \frac{2\rho}{2\rho+1}\right)$$

$$= 1 .$$

Since $f_k(z) = 0$,

$$[(2\rho+1)z-2\rho] = \frac{2z^{2(k+2)}}{[2(\rho+1)z-2\rho]^{k+1}} + \sum_{i=1}^{k} \frac{z^{2(i+1)}}{[2(\rho+1)z-2\rho]^{i}}$$

Now

$$1 < |[(2\rho+1)z-2\rho|$$

$$= \left| \frac{2z^{2(k+2)}}{[2(\rho+1)z-2\rho]^{k+1}} + \sum_{i=1}^{k} \frac{z^{2(i+1)}}{[2(\rho+1)z-2\rho]^{i}} \right|$$

$$< \left| \frac{2z^{2(k+2)}}{[2(\rho+1)z-2\rho]^{k+1}} \right| + \sum_{i=1}^{k} \left| \frac{z^{2(i+1)}}{[2(\rho+1)z-2\rho]^{i}} \right|$$

$$< \frac{2}{2^{k+1}} + \sum_{i=1}^{k} \frac{1}{2^i}$$

$$= \frac{1}{2^k} + \frac{\frac{1}{2} - (\frac{1}{2})^{k+1}}{1 - \frac{1}{2}}$$

$$= \frac{1}{2^k} + 1 - \frac{1}{2^k}$$

$$= 1 .$$

This is a contradiction and hence the only zero on the unit circle is $z = 1$.

Q.E.D.

Note that further calculation shows that the zero $z = 1$ is simple whenever $\rho \neq \rho_k$ and double when $\rho = \rho_k$. Using this fact the last sentence in Lemma 4.2 can be strengthened by replacing "at most" by "exactly".

4.8 Conditions for Positive Recurrence

By Theorem 1.3 the system will be positive recurrent if (4.1)-(4.3) and (4.5)-(4.6) have a nontrivial absolutely convergent solution. From Theorem 4.2 this is true if there are at least $k + 2$ zeros of $f_k(z)$ inside the open unit circle. Combining this with Lemma 4.2 one obtains the following.

Theorem 4.3: The 2-server, k storage space, series queueing system is positive recurrent if

$$\rho < \rho_k \ .$$

4.9 Conditions for the System to be Transient

Foster's Theorem (Theorem 1.4) asserts that a necessary and sufficient condition for an irreducible Markov chain to be transient is the existence of a non-constant bounded solution to a certain system of equations. As in section 2.8 one can set up the equations from the rate diagram. For this Markov chain assign the arbitrary variables to the states as follows:

To the simple state $(ss)_{n,i}$ assign the variable $y_i(n)$, $n \geq 2$, $0 \leq i \leq k$. To the lumped state b_n assign the variable $y_{k+1}(n)$, $n \geq 1$. Also assign the variable $y_i(1)$ to the state $(es)_{1,i}$ and $y_o(0)$ to the state $(ee)_{o,o}$.

Foster's equations then become the following with $\rho = \lambda/2\mu$.

For $n \geq 1$

$$(4.18) \qquad y_o(n) - (2\rho+1)y_{k+1}(n+1) + 2\rho y_{k+1}(n+2) = 0$$

$$(4.19) \qquad y_{k+1}(n) + y_{i+1}(n) - 2(\rho+1)y_i(n+1) + 2\rho y_i(n+2) = 0$$

$$0 \leq i \leq k$$

and

(4.20) $y_0(0) - (2\rho+1)y_i(1) + 2\rho y_i(2) = 0 \quad 0 \le i \le k+1$

Equation (4.20) with $i = k+1$ is equation (4.18) with $n = 0$, so any solution of the system (4.18) - (4.19) for all n will satisfy (4.20) for $i = k+1$. Thus we need only determine whether there is a nonconstant bounded solution of (4.18) - (4.19) which satisfies (4.20) for $0 \le i \le k$.

The general system (4.18) - (4.19) can be written in the form

(4.21) $$G_k(E)\vec{y}(n) = \vec{0}$$

where

$$G_k(E) = \begin{bmatrix} T_E & 1 & \cdots & 0 & 1 \\ 0 & T_E & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & \cdots & T_E & 2 \\ 1 & 0 & \cdots & 0 & U_E \end{bmatrix}_{(k+2)\times(k+2)}$$

where $T_E = [-2(\rho+1)E + 2\rho E^2]$ and $U_E = [-(2\rho+1)E + 2\rho E^2]$.

and

$$\vec{y}(n) = \begin{bmatrix} y_o(n) \\ y_1(n) \\ \cdot \\ \cdot \\ \cdot \\ y_k(n) \\ y_{k+1}(n) \end{bmatrix}$$

The fundamental solutions of such a linear homogeneous difference equation have the form $\vec{d}\zeta^n$ where ζ is an appropriate nonzero scalar and

$$\vec{d} = \begin{bmatrix} d_o \\ d_1 \\ \cdot \\ \cdot \\ \cdot \\ d_k \\ d_{k+1} \end{bmatrix}$$

is an appropriate constant vector.

Substituting this into (4.21) one obtains

$$\vec{0} = G_k(E)\vec{d}\zeta^n$$
$$= G_k(E)\,\zeta^n\vec{d}$$
$$= G_k(\zeta)\,\zeta^n\vec{d}$$
$$= G_k(\zeta)\vec{d}\zeta^n \quad .$$

This $G_k(\zeta)\vec{d} = \vec{0}$ and one concludes that $\vec{d}\zeta^n$ is a solution if and only if $G_k(\zeta)$ is a singular matrix and $\vec{d}$ is a vector in its null space. Hence ζ must be a zero of the fundamental polynomial $g(z) = \det G(z)$, and the corresponding $\vec{d}$ must satisfy

(4.22) $$G_k(\zeta)\vec{d} = \vec{0} \; .$$

Bounded solutions of this system will have the form $\vec{y}(n) = \sum_j \vec{d}_j\, \zeta_j^{\,n}$ where the ζ_i are the zeros of the polynomial $g_k(z) = \det G_k(z)$ inside the closed unit circle $|z| \le 1$.

The following is an explicit representation of $g_k(z)$.

Theorem 4.4: $g_k(z) = z^{2(k+2)} f_k(1/z) =$

$$= (-1)^{k+1}\{[-(2\rho+1)z+2\rho z^2][2(\rho+1)z-2\rho z^2]^{k+1} + 2$$
$$+ \sum_{i=1}^{k} [2(\rho+1)z-2\rho z^2]^{i}\}$$

Proof:

$g_k(z) =$

$$\begin{vmatrix} T_z & 1 & \cdots & 0 & 1 \\ 0 & T_z & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & \cdots & T_z & 2 \\ 1 & 0 & \cdots & 0 & U_z \end{vmatrix}_{(k+2)x(k+2)}$$

Where $T_z = [-2(\rho+1)z+2\rho z^2]$ and $U_z = [-(2\rho+1)z+2\rho z^2]$.

$$= (z)^{2(k+2)} \begin{vmatrix} T & 1/z^2 & \cdots & 0 & 1/z^2 \\ 0 & T & \cdots & 0 & 1/z^2 \\ 0 & 0 & \cdots & 0 & 1/z^2 \\ 0 & 0 & \cdots & 0 & 1/z^2 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & 1/z^2 & 1/z^2 \\ 0 & 0 & \cdots & T & 2/z^2 \\ 1 & 0 & \cdots & 0 & U \end{vmatrix}_{(k+2)\times(k+2)}$$

where $T = [-2(\rho+1)/z+2\rho]$ and $U = [-(2\rho+1)/z+2\rho]$.

Since the determinant of the transpose of a matrix equals the determinant of the matrix

$$g_k(z) = z^{2(k+2)} f_k(1/z)$$

$$= (-1)^{k+1}\{[-(2\rho+1)z+2\rho z^2][2(\rho+1)-2\rho z^2]^{k+1} + 2$$

$$+ \sum_{i=1}^{k} [2(\rho+1)z-2\rho z^2]^{i}\} .$$

Q.E.D.

In order for $\vec{d}$ to satisfy (4.22)

$$(4.23) \qquad [-(2\rho+1)\zeta+2\rho\zeta^2]\,d_{k+1} + d_o = 0$$

$$(4.24) \qquad d_{k+1} + [-2(\rho+1)\rho+2\rho\zeta^2]\,d_i + d_{i+1} = 0 \quad 0 \le i \le k .$$

Using equation (4.24) for each j the components d_{ij} can be represented in terms of $d_{k+1j} = e_j$

$$d_{ij} = [2(\rho+1)\zeta_j-2\rho\zeta_j{}^2]\,d_{i-1j} - d_{k+1j}$$

$$= [2(\rho+1)\zeta_j-2\rho\zeta_j^2]^i\, d_{oj} - \sum_{\ell=o}^{i-1} [2(\rho+1)\zeta_j-2\rho\zeta_j^2]^{\ell} e_j$$

for $1 \le i \le k$

$$d_{oj} = [(2\rho+1)\zeta_j-2\rho\zeta_j^2]\,e_j$$

$$d_{ij} = e_j\{[2(\rho+1)\zeta_j-2\rho\zeta_j^2]^i[(2\rho+1)\zeta_j-2\rho\zeta_j^2]$$

$$- \sum_{\ell=o}^{i-1} [2(\rho+1)\zeta_j-2\rho\zeta_j{}^2]^{\ell}\}$$

for $1 \le i \le k$

Equation (4.24) for $i = k$ must be shown to hold for these d_{ij} . Substituting

$$e_j\{[2(\rho+1)\zeta_j-2\rho\zeta_j^2]^i[(2\rho+1)\zeta_j-2\rho\zeta_j^2] - \sum_{\ell=o}^{i-1}[2(\rho+1)\zeta_j-2\rho\zeta_j^2]^{\ell}\}$$

for d_{ij} into the left side of (4.23) with $i = k$ results in

$$2e_j+e_j[-2(\rho+1)\zeta_j+2\rho\zeta_j^2]\{[2(\rho+1)\zeta_j-2\rho\zeta_j^2]^k[(2\rho+1)\zeta_j-2\rho\zeta_j^2]$$

$$-\sum_{\ell=o}^{k-1}[2(\rho+1)\zeta_j-2\rho\zeta_j^2]^{\ell}\} \quad .$$

To show this is zero, multiply by $\frac{(-1)^{k+1}}{e_j}$ yielding

$$(-1)^{k+1}\{2+[2(\rho+1)\zeta_j-2\rho\zeta_j^2]^{k+1}[-(2\rho+1)\zeta_j+2\rho\zeta_j^2]$$

$$+\sum_{\ell=o}^{k-1}[2(\rho+1)\zeta_j-2\rho\zeta_j^2]^{\ell+1}\}$$

$$= (-1)^{k+1}\{2+[2(\rho+1)\zeta_j-2\rho\zeta_j^2]^{k+1}[-(2\rho+1)\zeta_j+2\rho\zeta_j^2]$$

$$+\sum_{\ell=1}^{k}[2(\rho+1)\zeta_j-2\rho\zeta_j^2]^{\ell}\}$$

$$= g_k(\zeta_j)$$

$$= 0$$

since ζ_j is a zero of $g_k(z)$.

Solutions of (4.18) - (4.19) are of the form

$y_i(n) = \sum_j d_{ij} \zeta_j^n$. If the ζ_j's are such that

$|\zeta_j| < 1$ and d_{ij} are not all zero then there is a bounded nonconstant solution to the general system. There are $k + 1$ initial conditions. If there are $k + 2$ zeros of $g_k(z)$ inside the open unit circle and hence $k + 2$ e_j's then there is a nontrivial solution for the e_j's .

The polynomial $g_k(z)$ has been shown to be $g_k(z) = z^{2(k+2)} f_k(1/z)$. Thus the number of zeros of $g_k(z)$ inside the open unit circle is the same as the number of zeros of $f_k(z)$ outside the unit circle, for a given ρ . By Lemma 4.2 if $\rho < \rho_k$, $f_k(z)$ has $k + 2$ zeros inside the closed unit circle and hence, having degree $2(k+2)$, $k + 2$ zeros outside the closed unit circle. Thus there is a nontrivial solution for the e_j's and so from Foster's Theorem (Theorem 1.4) we obtain the following.

Theorem 4.5: <u>The</u> 2-<u>server</u>, k <u>storage space series queueing system is transient if</u> $\rho > \rho_k$.

The only remaining case is $\rho = \rho_k$. To consider this case we use Foster's Criteria, Theorem 1.5. The inequalities $\sum_{j=o}^{\infty} \lambda_{\ell j} y_{\ell}(j) \le 0$ for $\ell = 1, 2, \ldots$ are identical to the system (4.18)-(4.20) with the equal signs replaced by $\le$.

Lemma 4.4: _If_ $\rho = \rho_k$, _the_ 2-_server_, k _storage space series queueing system is recurrent_.

Proof: The inequalities for Foster's Criteria become:

(4.25) $y_o(n) - (2\rho_k+1)y_{k+1}(n+1) + 2\rho_k y_{k+1}(n+2) \le 0$

for $n \ge 0$

(4.26) $y_{k+1}(n) + y_{i+1}(n) - 2(\rho_k+1)y_i(n+1) + 2\rho_k y_i(n+2) \le 0$

for $0 \le i \le k$ and $n \ge 1$

Also

(4.27) $y_o(0) - (2\rho_k+1)y_i(1) + 2\rho_k y_i(2) \le 0$

for $0 \le i \le k$

where $\rho_k = \dfrac{3 \cdot 2^k - 1}{4 \cdot 2^k - 1}$. It can be seen by direct subsitution that

$$y_i(n) = n + (1-\rho_k)2^{i-k} \quad 0 \le i \le k+1, \quad n \ge 1$$

$$y_0(0) = 1 = 2\rho_k + (1-\rho_k)2^{-k}$$

satisfy the inequalities (4.25)-(4.27). As $n \to \infty$ $y_i(n) \to \infty$ for $0 \le i \le k+1$. Thus by Foster's Criteria the system is recurrent.

4.10 Summary of Results

The preceeding results can be summarized by the following:

Theorem 4.6: *The 2-server, k storage series queueing system is*

i) *positive recurrent if* $\rho < \rho_k$,

ii) *transient iff* $\rho < \rho_k$,

where $\rho_k = \dfrac{3 \cdot 2^k - 1}{4 \cdot 2^k - 1}$.

This can be compared to the well-known [10, p. 116] results for the 2-server parallel queueing system. *This system is*

i) *positive recurrent iff* $\rho < 1$;

ii) *transient iff* $\rho > 1$;

iii) *recurrent null iff* $\rho = 1$.

Note that $2/3 \le \rho_k < 3/4$. $\rho_o = 2/3$ and $\rho_k \to 3/4$ as $k \to \infty$, thus the saturation value of ρ cannot exceed 3/4 with intermediate storage, a comparatively small increase over the value of 2/3 with no storage.

V. THE n-SERVER SERIES QUEUEING PROBLEM

5.1 Definition of the Problem

A second possible extension of the simple 2-server series queueing problem is the simple n-server series queueing problem. The customers will be served on a first-come-first-served basis. Each customer is served by one of the n servers. There are no storage spaces between servers. The queue discipline requires that any customer entering the servers move as far forward as possible.

The various essential[1] states of the system can be represented by symbols of the form $(i)_{m,b}$ where m represents the number of customers, who have not yet completed service, (the queue length), i is the number of accessible servers, and b is the number of blocked servers.[2] A server may be blocked by a customer, who has completed service but cannot leave the system because his

1. A state that communicates with every state that it leads to is termed essential, otherwise it is termed inessential.
2. Note that if $i = 0$, state $(0)_{m,b}$ exists for $0 \leq b \leq n-1$, $m \geq 0$; if $i > 0$, state $(i)_{m,b}$ exists only for $0 \leq b \leq n - i - 1$ and $m = n - i - b$.

way is blocked. A server may also be blocked if he has no customer but customers cannot get to him because the way is blocked by customers at other servers. All empty servers are either to the right or left of the busy servers. We will see that any other arrangement is inessential.

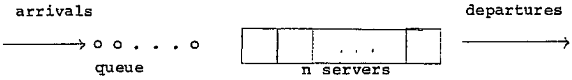


Diagram of the System: Figure 5.1

The states of the system are

$(n)_{o,o}$

$(n-1)_{1,o}$ $(n-2)_{1,1}, \ldots, (0)_{1,n-1}$

$(n-2)_{2,o}$ $(n-3)_{2,1}, \ldots, (0)_{2,n-1}$

$\vdots$

$(2)_{n-2,o},$ $(1)_{n-2,1}, \ldots, (0)_{n-2,n-1}$

$(1)_{n-1,o},$ $(0)_{n-1,1}, \ldots, (0)_{n-1,n-1}$

$(0)_{n,o},$ $(0)_{n,1}, \ldots,$ $(0)_{n,n-1}$

$(0)_{n+1,o},$ $(0)_{n+1,1}, \ldots, (0)_{n+1,n-1}$

$\vdots$

$(0)_{m,o}$ $(0)_{m,1}, \ldots,$ $(0)_{m,n-1}$

$\vdots$

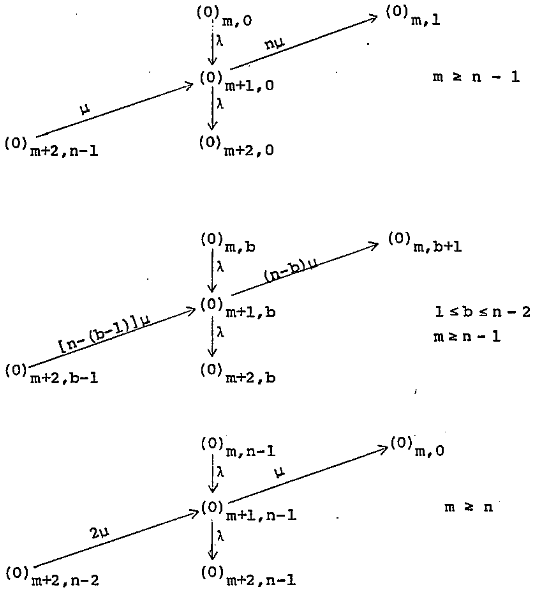
A state with empty servers between busy or blocked servers is inessential. This is seen by noting that whenever the system starts in such a state it will leave this class of states once each customer, who was originally in the servers has left the system. No such inessential state is accessible from an essential state of the form $(i)_{m,b}$. Hence, once the system has left the inessential states it never returns. Since the long-run probabilities are considered here it is not necessary to consider the inessential states.

Assume arrivals form a Poisson process with intensity λ; each server has a mean service rate of μ where service times are exponentially distributed with probability density function $\mu e^{-\mu t}$; all interarrival and service times are mutually independent.

With these assumptions the process forms a continuous-time, discrete-state, time-homogeneous Markov chain of the type considered in Chapter 1. The state space is denumerable and the states are represented by the state symbols $(i)_{m,b}$.

5.2 Rate Diagram

The one-step transitions for this Markov chain are generated by arrivals and service completions. Each arrival increases the queue length by 1. Each service completion reduces the queue length by 1. For example if $n = 7$ a service completion from one server might change the state $(0)_{10,5}$ to $(0)_{9,6}$. The probability of an arrival in time h is $\lambda h + o(h)$ so the transition rate for a transition generated by an arrival will be λ. The probability of a completion from one busy server in time h is $\mu h + o(h)$ so the transition rate for a transition generated by a service completion will be μ. Thus the rate diagram will have vertices and weighted edges made up by fitting together subgraphs of the following types:



<u>Some Vertices of the Rate Diagram</u>: Figure 5.2

In addition there will be special subgraphs around the vertices $(i)_{m,b}$ where $m < n$.

Denote the long-run probabilities of this irreducible Markov chain by $p_b^i(m) = P[(i)_{m,b}] = \lim_{t\to\infty} p_{\alpha,(i)_{m,b}}(t)$

where α represents an arbitrary state, m is the queue length, i is the number of accessible servers, and b is the number of blocked states.

5.3 Normal Equations

The normal equations, with $\rho = \frac{\lambda}{n\mu}$, are the following:

(5.1)
$$-[n\rho+n-b-(n-m-b-1)^+]\,p_b^{(n-m-b-1)^+}(m+1)$$
$$+\ n\rho p_b^{(n-m-b)^+}(m)$$
$$+\ [n-b+1-(n-m-b-1)^+]\,p_{b-1}^{(n-m-b-1)^+}(m+2) = 0$$

for $m \geq 0$, $1 \leq b \leq n-1$

(5.2)
$$-[n\rho+n-(n-m-1)^+]\,p_o^{(n-m-1)^+}(m+1)+n\rho p_o^{(n-m)^+}(m)$$
$$+\ p_{n-1}^o(m+2) = 0 \quad \text{for} \quad m \geq 0$$

(5.3)
$$-n\rho p_o^n + \sum_{b=o}^{n-1} p_b^{n-1-b}(1) = 0$$

where $x^+ = \max(x, 0)$ and $p_b^{\,i}(0) = 0$ except when $i = n$ and $b = 0$. For $m \geq n$ these equations become

$$(5.4) \quad -[n\rho+n-b]\, p_b^o(m+1) + n\rho p_b^o(m) + [n-b+1]\, p_{b-1}^o(m+2) = 0$$

$$1 \leq b \leq n - 1$$

$$(5.5) \qquad -[n\rho+n]\, p_o^{\,o}(m+1) + n\rho p_o^{\,o}(m) + p_{n-1}^o(m+2) = 0$$

This system is dependent, so any solution of all but one of these equations is a solution of all. Hence one of the equations can be disregarded. (5.2) with $m = 0$ will be disregarded. Further (5.3) is automatically satisfied if it is used to determined $p_o^n(0)$ in terms of the other quantities,

$p_o^n(0) = \frac{1}{n\rho} \sum_{b=0}^{n-1} p_b^{n-1-b}(1)$. Hence only the system (5.1)-(5.2) need be studied.

By Theorem 1.3 the Markov chain is positive recurrent if the normal equations with the probabilities $p_b^i(m)$ replaced by arbitrary variables $x_b^i\ (m)$, have a non-trivial absolutely convergent solution. If this is the case any such solution to this homogeneous system can be normalized to form a probability distribution satisfying

(5.6) $$p_o^n(0) + \sum_{m=1}^{\infty} \sum_{b=0}^{n-1} p_b^{(n-b-m)^+}(m) = 1$$

Thus the problem reduces to finding sufficient conditions that the system has a nontrivial absolutely convergent solution.

5.4 Difference Equation Approach

The general system consitutes a vector difference equation. Using the shift operator E this system can be written in the form

(5.7) $$F^{(n)}(E)\vec{p}(m) = \vec{0} \qquad m \geq n$$

where

$F^{(n)}(E) =$

$$
\begin{bmatrix}
T_E(n) & 0 & \cdots & 0 & E^2 \\
nE^2 & T_E(n-1) & \cdots & 0 & 0 \\
0 & (n-1)E^2 & \cdots & 0 & 0 \\
\cdot & \cdot & \cdots & \cdot & \cdot \\
\cdot & \cdot & \cdots & \cdot & \cdot \\
\cdot & \cdot & \cdots & \cdot & \cdot \\
\cdot & \cdot & \cdots & \cdot & \cdot \\
0 & 0 & \cdots & T_E(2) & 0 \\
0 & 0 & \cdots & 2E^2 & T_E(1)
\end{bmatrix}_{n \times n}
$$

where $T_E(i) = [-(n\rho+i)E+n\rho]$.

$$
\vec{p}(m) = \begin{bmatrix}
p_0^o(m) \\
p_1^o(m) \\
\cdot \\
\cdot \\
\cdot \\
\cdot \\
p_{n-1}^o(m)
\end{bmatrix}
$$

The fundamental solutions of such a linear homogeneous difference equation are known to have the form $\vec{c}\zeta^m$ where ζ is an appropriate nonzero scalar and

$$\vec{c} = \begin{bmatrix} c_n \\ c_{n-1} \\ \cdot \\ \cdot \\ \cdot \\ c_1 \end{bmatrix}$$

is an appropriate constant vector.

If this is substituted into (5.7) as in Chapter 4 one obtains

$$\vec{0} = F^{(n)}(\zeta)\vec{c}\ \zeta^n \ . \tag{5.8}$$

Thus $F^{(n)}(\zeta)\vec{c} = \vec{0}$ and one concludes that $\vec{c}\zeta^m$ is a solution if and only if $F^{(n)}(\zeta)$ is a singular matrix and $\vec{c}$ is a vector in its null space. Thus ζ must be a zero of the fundamental polynomial

$$f^{(n)}(z) = \det F^{(n)}(z) \tag{5.9}$$

and $\vec{c}$ must satisfy

$$f^{(n)}(\zeta)\vec{c} = \vec{0} \ . \tag{5.10}$$

5.5 The Zeros of $f^{(n)}(z)$

Further analysis depends on knowledge of the location of the zeros of $f^{(n)}(z)$. The following is an explicit representation of $f^{(n)}(z)$.

Theorem 5.1: $f^{(n)}(z) = (-1)^{n-1}\{n!z^{2n}-[(n\rho+1)z-n\rho][(n\rho+2)z-n\rho]\ldots$

$\ldots[(n\rho+n)z-n\rho]\}$.

Proof:

$f^{(n)}(z) =$

$$\begin{vmatrix} T_z(n) & 0 & \cdots & 0 & z^2 \\ nz^2 & T_z(n-1) & \cdots & 0 & 0 \\ 0 & (n-1)z^2 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & T_z(2) & 0 \\ 0 & 0 & \cdots & 2z^2 & T_z(1) \end{vmatrix}$$

$n \times n$

where $T_z(i) = [-(n\rho+i)z+n\rho]$.

From this one obtains

$$f^{(n)}(z) = (-1)^{n-1} n! z^{2n} + [-(n\rho+1)z+n\rho][-(n\rho+2)z+n\rho] \cdots$$

$$\cdots [-(n\rho+n)z+n\rho]$$

by expanding by the last column.

Hence

$$f^{(n)}(z) = (-1)^{n-1}\{n! z^{2n} - [(n\rho+1)z-n\rho][(n\rho+2)z-n\rho] \cdots$$

$$\cdots [(n\rho+n)z-n\rho]\}$$

Q.E.D.

Certain modifications in the subsequent arguments are necessary if $f^{(n)}(z)$ has multiple zeros inside the unit circle. For simplicity only the case where $f^{(n)}(z)$ has distinct zeros inside $|z| < 1$ will be considered in detail. Hence the null space of $F^{(n)}(\zeta)$ (see [2, p. 253]), where ζ is a zero of $f^{(n)}(z)$, is 1-dimensional and the corresponding $\vec{c}$ is nontrivial and unique up to a multiplicative constant. The general solution to the system (5.4) - (5.5) has the form

(5.11) $$\vec{p}(m) = \sum_j \vec{c}_j \zeta_j^m$$

where the ζ_j's are the distinct zeros of $f^{(n)}(z)$, and the constant vectors $\vec{c}_j$ satisfy (5.10) for $\zeta = \zeta_j$.

In order that the terms of (5.11) form an absolutely convergent series, only the zeros ζ_j with $|\zeta_j| < 1$ can be included. Denote these zeros by $\zeta_1, \ldots, \zeta_\ell$. Thus by Theorem 1.3, the system will be positive recurrent if constant vectors $\vec{c}_1, \ldots, \vec{c}_\ell$ can be obtained, satisfying the previous conditions and also the remaining equations.

In order for $\vec{c}$ to satisfy (5.10)

$$(5.12) \qquad [-(n\rho+n)\zeta+n\rho]c_n + \zeta^2 c_1 = 0$$

$$(5.13) \qquad [-(n\rho+i)\zeta+n\rho]c_i + (i+1)\zeta^2 c_{i+1} = 0 \qquad 1 \le i \le n-1.$$

Using equation (5.13) for each j the components c_{ij} can be represented in terms of $c_{1j} = a_j$

$$c_{ij} = \left[\frac{(n\rho+i-1)\zeta_j - n\rho}{i\zeta_j^2}\right] c_{i-1\,j}$$

$$= \left[\frac{(n\rho+i-1)\zeta_j - n\rho}{i\zeta_j^2}\right] \left[\frac{(n\rho+i-2)\zeta_j - n\rho}{(i-1)\zeta_j^2}\right]$$

$$\cdots \left[\frac{(n\rho+1)\zeta_j - n\rho}{2\zeta_j^2}\right] a_j$$

for $2 \le i \le n$

$$c_{ij} = \delta_{ij}\delta_{i-1j} \cdots \delta_{2j}a_j$$

where $\delta_{ij} = \dfrac{(n\rho+i-1)\zeta_j - n\rho}{i\zeta_j^2}$ for $2 \le i \le n$.

Equation (5.12) must also be shown to hold for these c_{ij} .

Substituting

$$\left[\frac{(n\rho+n-1)\zeta_j - n\rho}{n\zeta_j^2}\right] \cdots \left[\frac{(n\rho+1)\zeta_j - n\rho}{2\zeta_j^2}\right] a_j \quad \text{for } c_{nj},$$

and a_j for c_{1j}

into the left side of (5.12) results in

$$[-(n\rho+n)\zeta_j+n\rho]\left[\frac{(n\rho+n-1)\zeta_j-n\rho}{n\zeta_j^2}\right]\cdots\left[\frac{(n\rho+1)\zeta_j-n\rho}{2\zeta_j^2}\right]a_j+\zeta_j^2\,a_j$$

To show this is zero multiply through by

$$\frac{(-1)^{n-1}n!\,\zeta_j^{2(n-1)}}{a_j},$$ which yields

$$[-(n\rho+n)\zeta_j+n\rho][(n\rho+n-1)\zeta_j-n\rho]\cdots[(n\rho+1)\zeta_j-n\rho]\ +n!\,\zeta_j^{2n}$$

$$=(-1)^{n-1}\{n!\,\zeta_1^{2n}-[(n\rho+n)\zeta_j-n\rho]\cdots[(n\rho+1)\zeta_j-n\rho]\}$$

$$=f^{(n)}(\zeta_j)$$

$$=0$$

Since ζ_j is a zero of $f^{(n)}(z)$.

Thus equation (5.12) holds with

$$c_{ij}=\left[\frac{(n\rho+i-1)\zeta_j-n\rho}{i\zeta_j^2}\right]\cdots\left[\frac{(n\rho+1)\zeta_j-n\rho}{2\zeta_j^2}\right]a_j$$

$$=\delta_{ij}\cdots\delta_{2j}\,a_j \qquad 2\le i\le n$$

$$c_{ij}=a_j$$

Let $\zeta_1, \ldots, \zeta_q$ represent the zeros of $f^{(n)}(z)$ inside the unit circle. On substitution of the above representations for c_{ij} into the initial conditions (5.2)-(5.3) yield a single homogeneous linear system for $\vec{a} = (a_1, \ldots, a_q)^t$ of the form

$$B\vec{a} = \vec{0}, \tag{5.14}$$

where B is $k \times q$ for some fixed $k \geq q$.

5.6 The Fundamental Polynomial

Before proceeding to obtain conditions under which the system (5.2)-(5.3) has a nontrivial solution it is necessary to obtain further information about the zeros of the fundamental polynomial.

$$f^{(n)}(z) = \begin{vmatrix} T_z(n) & 0 & \cdots & 0 & z^2 \\ nz^2 & T_z(n-1) & \cdots & 0 & 0 \\ 0 & (n-1)z^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & T_z(2) & 0 \\ 0 & 0 & \cdots & 2z^2 & T_z(1) \end{vmatrix}_{n \times n}$$

where $T_z(i) = [-(n\rho+i)z + n\rho]$.

Note that $f^{(2)} = f_o = f$.

Lemma 5.1: $f^{(n)}$ _has at least one zero at_ $z = 1$.

Proof: Direct substitution

Definition: $$\rho^{(n)} = \frac{1}{\sum_{i=1}^{n} \frac{1}{i}} .$$

Lemma 5.2 _If_ $\rho < \rho^{(n)}$, $f^{(n)}(z)$ _has exactly_ n _zeros inside the open unit circle_. _If_ $\rho > \rho^{(n)}$, $f^{(n)}(z)$ _has at most_ $n - 1$ _zeros inside the open unit circle_.

Proof: Recall from Theorem 5.1.

$$f^{(n)}(z) = (-1)^{n-1}\{n!z^{2n}-[(n\rho+1)z-n\rho][(n\rho+2)z-n\rho] \cdots [(n\rho+n)z-n\rho]\} .$$

On the circle $|z| = 1 + \varepsilon$, $\varepsilon \neq 0$ the inequality $|a - b| \geq |a| - |b|$ yields

$$(5.15) \quad |[(n\rho+1)z-n\rho]\cdots[(n\rho+n)z-n\rho]| - |n!\, z^{2n}|$$

$$= |(n\rho+1)z-n\rho|\cdots|(n\rho+n)z-n\rho| - n!\,|z|^{2n}$$

$$\geq [(n\rho+1)(1+\varepsilon)-n\rho]\cdots[(n\rho+n)(1+\varepsilon)-n\rho]$$

$$- n![1+2n\varepsilon+O(\varepsilon^2)]$$

$$= [1+(n\rho+1)\varepsilon]\cdots[n+(n\rho+n)\varepsilon]-n![1+2n\varepsilon+O(\varepsilon^2)]$$

$$= n! + \sum_{i=1}^{n} \frac{n!(n\rho+i)\varepsilon}{i} - n!-2n\varepsilon n!+O(\varepsilon^2)$$

$$= \varepsilon n! \left[\sum_{i=1}^{n} \frac{n\rho+1}{i} - 2n\right] + O(\varepsilon^2)$$

$$= \varepsilon n! \left[n\rho \sum_{i=1}^{n} \frac{1}{i} + n - 2n\right] + O(\varepsilon^2)$$

$$= n\varepsilon n! \left[\rho \sum_{i=1}^{n} \frac{1}{i} - 1\right] + O(\varepsilon^2)$$

<u>Case</u> (i) $\rho < \rho^{(n)}$

Whenever $\rho>\rho^{(n)} = \dfrac{1}{1/\sum_{i=1}^{n}\frac{1}{i}}$, $\varepsilon < 0$ and sufficiently small, (5.15) is positive. By Rouché's Theorem $f^{(n)}(z)$, given by (5.9), must have the same number of zeros inside the circle $|z| = 1 + \varepsilon$ as $[(n\rho+1)z-n\rho]\cdots[(n\rho+n)z-n\rho]$ has. Letting $\varepsilon \to 0^-$ one concludes that the same is true inside the open unit circle. Therefore $f^{(n)}(z)$ has n zeros inside the open unit circle when $\rho < \rho^{(n)}$.

<u>Case</u> (ii) $\rho > \rho^{(n)}$

If $\rho > \rho^{(n)}\ 1/\dfrac{1}{\sum_{i=1}^{n}\frac{1}{i}}$, $\varepsilon > 0$ and sufficiently small, then (5.15) is positive. By Rouché's Theorem $f^{(n)}(z)$ has the same number of zeros inside the circle $|z| = 1 + \varepsilon$, $\varepsilon > 0$ as $[(n\rho+1)z-n\rho]\cdots[(n\rho+n)z-n\rho]$ has. By letting $\varepsilon \to 0^+$ one concludes that the same is true inside the closed unit circle. One of these is at $z = 1$ so at most $n - 1$ zeros can be inside the open unit circle.

<u>Case</u> (iii) $\rho = \rho^{(n)}$

Note that the zeros of $f^{(n)}(z)$ are analytic and hence continuous functions of ρ. Choose $\rho > \rho^{(n)}$.

As $\rho \to \rho^{(n)^+}$ no zero can leave the closed region $|z| \geq 1$. Note $f^{(n)}(z)$ is of degree $2n$ and $f^{(n)}(z)$ has at least $n + 1$ zeros in $|z| \geq 1$ since there are at most $n - 1$ zeros in $|z| < 1$. Thus when $\rho = \rho^{(n)}$ $f^{(n)}(z)$ has at least $n + 1$ zeros in $|z| \geq 1$ and hence at most $n - 1$ zeros in $|z| < 1$.

Q.E.D.

Theorem 5.2: <u>There are no zeros of</u> $f^{(n)}(z)$ <u>on the unit circle except at</u> $z = 1$.

Proof: $f^{(n)}(z) = (-1)^{n-1}\{n!z^{2n} - [(n\rho+1)z-n\rho]\cdots[(n\rho+n)z-n\rho]\}$.

Suppose $f(z) = 0$, $|z| = 1$, $z \neq 1$. Then

$$|(n\rho+i)z-n\rho| = (n\rho+i)\left|z - \frac{n\rho}{n\rho+i}\right|$$

$$> (n\rho+i)\left[1 - \frac{n\rho}{n\rho+i}\right]$$

$$= i$$

Since $f^{(n)}(z) = 0$

$$|n!z^{2n}| = |[(n\rho+1)z-n\rho]\cdots[(n\rho+n)z-n\rho]|$$

Now

$$n! < |[(n\rho+1)z-n\rho]\cdots[(n\rho+n)z-n\rho]|$$

$$= |n!z^{2n}|$$

$$= n!$$

This is a contradiction and hence the only zero on the unit circle is at $z = 1$.

Q.E.D.

Further calculation shows that the zero $z = 1$ is simple whenever $\rho \neq \rho^{(n)}$ and double when $\rho = \rho^{(n)}$. Using this fact the last sentence in the statement of Lemma 5.2 can be strengthened by replacing "at most" by "exactly".

5.7 Conditions for Positive Recurrence

Theorem 5.3: A nontrivial solution to the system (5.2)-(5.3) exists if $f^{(n)}(z)$ has at least n zeros inside the unit circle. Thus the system is positive recurrent if $\rho < \rho^{(n)}$.

Proof: It was shown previously that only zeros inside the unit circle can be used. There are the same number of a_j's as zeros of $f^{(n)}(z)$ inside the open unit circle. Let B be the coefficient matrix of the system of initial conditions and $\vec{a} = (a_1, a_2, \ldots, a_q)^t$ where q is the number of zeros inside the unit circle. This system has the form $B\vec{a} = \vec{0}$.

Consider the case $\rho < \rho^{(n)}$. It was proved in Lemma 5.2 that in this case $q = n$. $B\vec{a} = \vec{0}$ has a nontrivial solution if and only if the rank of B is less than n . This is true if and only if $\det B' = 0$

for every $n \times n$ submatrix B' of B. Choose one of these submatrices B'. Let $\varphi(\rho) = \det B'$. The system is positive recurrent for sufficiently small ρ since the system clearly is more efficient than the system with one server and a service rate of μ. Hence for all sufficiently small ρ, $\varphi(\rho) = 0$. $\varphi(\rho)$ is an analytic function of ρ, thus $\varphi(\rho) \equiv 0$ for all $\rho < \rho^{(n)}$. This is true for each B'. Therefore the rank of B is less than n for all $\rho < \rho^{(n)}$, and the system has a nontrivial solution if $\rho < \rho^{(n)}$. Thus the system is positive recurrent if $\rho < \rho^{(n)}$.

5.8 Conditions for the System to be Transient

Foster's Theorem (Theorem 1.4) asserts that a necessary and sufficient condition for an irreducible Markov chain to be transient is the existence of a nonconstant bounded solution to a certain system of equations. As in section 2.8 one can set up the equations from the rate diagram. For this Markov chain assign the variable $y_b^i(m)$ to the state $(i)_{m,b}$ where m is the queue length, b the number of blocked servers and i the number of accessible servers. Note that if $i > 0$ then $m = n - i - b$.

Foster's equations then become the following with

$\rho = \frac{\lambda}{n\mu}$.

(5.16) $$-[n\rho+n-b-(n-m-b-1)^{+}]\,y_b^{(n-m-n-1)^{+}}(m+1)$$

$$+\ n\rho y_b^{(n-m-b-2)^{+}}(m+2)$$

$$+\ [n-b-(n-m-b-1)^{+}]\,y_{b+1}^{(n-m-b-1)^{+}}(m) = 0$$

for $m \geq 0 \quad 0 \leq b \leq n-2$

(5.17) $$-[n\rho+1]\,y_{n-1}^{o}(m+1)+n\rho y_{n-1}^{o}(m+2)+y_o^{(n-m)^{+}}(m) = 0$$

for $m \geq 0$

For $m \geq n$ these equations become

(5.18) $$-[n\rho+n-b]\,y_b^{o}(m+1) + n\rho y_b^{o}(m+2)+(n-b)y_{b+1}^{o}(m) = 0$$

for $0 \leq b \leq n-2$ and

(5.19) $$-[n\rho+1]\,y_{n-1}^{o}(m+1)+n\rho y_{n-1}^{o}(m+2) + y_o^{o}(m) = 0 .$$

This system can be written in the form

(5.20) $$G^{(n)}(E)\vec{y}(n) = \vec{0}$$

where

$G^{(n)}(E) =$

$$\begin{bmatrix} U_E(n) & n & \cdots & 0 & 0 \\ 0 & U_E(n-1) & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & U_E(2) & 2 \\ 1 & 0 & \cdots & 0 & U_E(1) \end{bmatrix}_{n \times n}$$

where $U_E(i) = [-(n\rho+i)E + n\rho E^2]$.

and

$$\vec{y}(m) = \begin{bmatrix} y_0^o(m) \\ y_1^o(m) \\ \vdots \\ y_{n-1}^o(m) \end{bmatrix}$$

A argument similar to that in section 4.9 shows the fundamental solutions of this system have the form

(5.21) $$\vec{y}(m) = \vec{d}\zeta^m$$

and the general bounded solution will be a linear combination of such terms,

(5.22) $$\vec{y}(m) = \sum_{j=1}^{q} \vec{d}_j \, \zeta_j^m$$

where $\zeta_1, \ldots, \zeta_q$ are the zeros of the polynomial $g^{(n)}(z) = \det g^{(n)}(z)$ inside the closed unit circle. The constant vectors

$$\vec{d}_j = \begin{bmatrix} d_{nj} \\ d_{n-1\,j} \\ \cdot \\ \cdot \\ \cdot \\ d_{1j} \end{bmatrix}$$

are in the null space of $G^{(n)}(\zeta_j)$.

The following is an explicit representation of $g^{(n)}(z)$.

Theorem 5.4: $g^{(n)}(z) = z^{2n} f^{(n)}(1/z)$

$$= (-1)^{n-1} \{n! - [(n\rho+1)z - n\rho z^2] \cdots [(n\rho+n)z - n\rho z^2]\} .$$

Proof:

$$g^{(n)}(z) =$$

$$\begin{vmatrix} U_z(n) & n & \cdots & 0 & 0 \\ 0 & U_z(n-1) & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & U_z(2) & 2 \\ 1 & 0 & \cdots & 0 & U_z(1) \end{vmatrix}_{n \times n}$$

where $U_z(i) = [-(n\rho+i)z+n\rho z^2]$.

$$= z^{2n}$$

$$\begin{vmatrix} U(n) & n/z^2 & \cdots & 0 & 0 \\ 0 & U(n-1) & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & U(2) & 2/z^2 \\ 1/z^2 & 0 & \cdots & 0 & U(1) \end{vmatrix}_{n \times n}$$

Where $U(i) = [-(n\rho + i)/z + n\rho]$.

Since the determinant of the transpose of a matrix equals the determinant of that matrix, one sees by comparision with Theorem 5.1

$$g^{(n)}(z) = z^{2n} f^{(n)}(1/z)$$

$$= (-1)^{n-1}\{n! - [(n\rho+1)z - n\rho z^2]\cdots[(n\rho+n)z - n\rho z^2]\}$$

Q.E.D

In order for $\vec{y}(n)$ to satisfy (5.20)

(5.23) $$[-(n\rho+1)\zeta + n\rho\zeta^2]d_1 = d_n = 0$$

(5.24) $$[-(n\rho+b)\zeta + n\rho\zeta^2]d_b + bd_{b-1} = 0 \qquad 2 \le b \le n$$

For the general solution (5.22) one must use equation (5.24) for each j. The components d_{ij} of $d_j = (d_{nj}, \ldots, d_{1j})^t$ can be represented in terms of $d_{nj} = e_j$

$$d_{bj} = \frac{[(n\rho+b+1)\zeta_j - n\rho\zeta_j^2]\, d_{b+1,j}}{b+1}$$

$$= \frac{[(n\rho+b+1)\zeta_j - n\rho\zeta_j^2]\cdots[(n\rho+n)\zeta_j - n\rho\zeta_j^2]\, e_j}{(b+1)\cdots n}$$

for $1 \le b \le n-1$

Equation (5.23) can be shown to hold for these d_{bj} .

Substitute

$$\frac{[(n\rho+2)\zeta_j-n\rho\zeta_j^2]\cdots[(n\rho+n)\zeta_j-n\rho\zeta_j^2]e_j}{n!}$$

for d_{bj} into the left side of (5.23). This results in

$$\frac{-[(n\rho+1)\zeta_j-n\rho\zeta_j^2]\cdots[(n\rho+n)\zeta_j-n\rho\zeta_j^2]e_j + e_j}{n!}$$

To show this is zero, multiply by $\frac{(-1)^{n-1}n!}{e_j}$ yielding

$$(-1)^n\{n!-[(n\rho+1)\zeta_j-n\rho\zeta_j^2]\cdots[(n\rho+n)\zeta_j-n\rho\zeta_j^2]\}$$

$$= g^{(n)}(\zeta_j)$$

$$= 0$$

since ζ_j is a zero of $g^{(n)}(z)$.

The polynomial $g^{(n)}(z)$ has been shown to be $g^{(n)}(z) = z^{2n}f^{(n)}(1/z)$. Thus, for a given ρ, the number of zeros of $g^{(n)}(z)$ inside the open unit circle is the same as the number of zeros of $f^{(n)}(z)$ outside the unit circle. By Lemma 5.2 if $\rho > \rho^{(n)}$ $f^{(n)}(z)$ has n zeros inside the closed unit circle and hence, having degree $2n$, n zeros outside the closed unit circle.

Theorem 5.5: *The n-server series queueing system is transient if* $\rho > \rho^{(n)}$.

Proof: If $\rho > \rho^{(n)}$, then there are n zeros of $g^{(n)}(z)$ inside the unit circle and thus ne_j's. The system of initial equations for the e_j has the form $\tilde{B}\vec{e} = \vec{0}$ where $\tilde{B}$ is the coefficient matrix of the kxn system and $\vec{e} = (e_1, e_2, \ldots, e_n)^t$. The system is transient if this system of equations has a nontrivial solution. $\tilde{B}\vec{e} = \vec{0}$ has a nontrivial solution if and only if the rank of $\tilde{B}$ is less than n . This is true if and only if det $\tilde{B}' = 0$ for every $n \times n$ submatrix $\tilde{B}'$ of $\tilde{B}$. If ρ is sufficiently large the system is known to be transient. Therefore $\varphi(\rho) = \det B' = 0$ for all $n \times n$ submatrices B' of B, for sufficiently large ρ .
Hence $\varphi(\rho) \equiv 0$ for all $\rho > \rho^{(n)}$ since $\varphi(\rho)$ is an analytic function of ρ . Thus the system $\tilde{B}\vec{e} = \vec{0}$ has a nontrivial solution for $\rho > \rho^{(n)}$ and so $\rho > \rho^{(n)}$ implies the system is transient.

The only remaining case is $\rho = \rho^{(n)}$. This case can be shown to be recurrent as in to section 4.9.

5.9 Summary of Results

The preceeding results can be summarized by the following with $\rho = \frac{\lambda}{n\mu}$:

Theorem 5.6: <u>The</u> n-<u>server series queueing system is</u>

i) <u>positive recurrent if</u> $\rho > \rho^{(n)}$,

ii) <u>transient</u> <u>iff</u> $\rho > \rho^{(n)}$,

iii) <u>recurrent</u> <u>if</u> $\rho = \rho^{(n)}$,

<u>where</u> $\rho^{(n)} = \dfrac{1}{\sum_{i=1}^{n} \frac{1}{i}}$.

This can be compared to the well-known [10, p. 116] results for tne n-server parallel queueing system. <u>This system is</u>

i) <u>positive recurrent</u> <u>iff</u> $\rho < 1$;

ii) <u>transient</u> <u>iff</u> $\rho > 1$;

iii) <u>recurrent null</u> <u>iff</u> $\rho = 1$.

Note that for $n = 1$ the system is the same as the parallel system with $n = 1$ and $\rho^{(1)} = 1$. As $n \to \infty$ $\rho^{(n)} \to 0$.

The n-server series queueing system is positive recurrent if $\rho < \rho^{(n)} = 1/\sum_{i=1}^{n} \sim \dfrac{1}{\log n + \frac{1}{2}}$.

Thus the system is positive recurrent if $\frac{\lambda}{n\mu} < \rho^{(n)} \sim \dfrac{1}{\log n+\frac{1}{2}}$

and hence if

$$\frac{\lambda}{\mu} < n_{\rho}^{(n)} \sim \frac{n}{\log n + \frac{1}{2}} \to \infty . \qquad (5.25)$$

A regular s server parallel system is positive recurrent if $\frac{\lambda}{s\mu} < 1$ and hence if

$$\frac{\lambda}{\mu} < s \qquad (5.26)$$

Comparing (5.25) and (5.26) we see that for large n, the series system has approximately the same saturation point as the regular s-server parallel system with

$$s = \frac{n}{\log n + \frac{1}{2}} .$$

VI SUMMARY

6.1 The 2-Server Series Queueing System

In this paper it has been determined (Theorem 2.5) that the 2-server series queueing system is

<u>positive recurrent</u> <u>iff</u> $\rho < 2/3$;
<u>recurrent null</u> <u>iff</u> $\rho = 2/3$;
<u>transient</u> <u>iff</u> $\rho > 2/3$,

where $\rho = \lambda/(2\mu)$, λ = mean input rate, μ = mean service rate of <u>each server</u>.

If the servers are located in parallel instead of in series the system is

<u>positive recurrent</u> <u>iff</u> $\rho < 1$;
<u>recurrent null</u> <u>iff</u> $\rho = 1$;
<u>transient</u> <u>iff</u> $\rho > 1$.

Thus blocking causes a drop of 1/3 in the efficiency of the system. For small values of ρ the two systems are similar. However, as $\rho \to 2/3$ the series case becomes much less efficient.

6.2 The 2-Server, k Storage Space Series Queueing System

In an attempt to relieve some of the blocking of the rear server, k waiting spaces are placed between the servers. This system was determined to be (Theorem 4.6)

<u>positive recurrent</u>	<u>if</u>	$\rho < \rho_k$;
<u>transient</u>	<u>iff</u>	$\rho > \rho_k$,

where $\rho_k = \dfrac{3 \cdot 2^k - 1}{4 \cdot 2^k - 1}$.

$\rho_k \to 3/4$ as $k \to \infty$, hence this system is still less than 3/4 as efficient as the parallel system.

6.3 The n Server Series Queueing System

Another generalization of the 2-server series queueing system is the n server series queueing system. It was determined (Theorem 5.6) that this system is

<u>positive recurrent</u>	<u>if</u>	$\rho < \rho^{(n)}$;
<u>transient</u>	<u>iff</u>	$\rho > \rho^{(n)}$,

where $\rho^{(n)} = \dfrac{1}{\sum_{i=1}^{n} \frac{1}{i}}$, and $\rho = \lambda/(n\mu)$.

These results can be compared to the results for the n server parallel queueing system. This system is

positive recurrent	iff	$\rho < 1$;
recurrent null	iff	$\rho = 1$;
transient	iff	$\rho > 1$.

The n server series queueing system is comparable in efficiency with a parallel system having $[\frac{n}{\log n + 1/2}]$ servers, in a precisely defined sense (see section 5.9).

6.4 Future Research

Many other questions can be asked about series queueing systems. It would be of interest to know whether the system is positive recurrent or recurrent null, in the case $\rho = \rho_k$ in the 2-server, k waiting spaces series queueing problem and in the case $\rho = \rho^{(n)}$ in the n server series queueing problem. A system with more than two servers and intermediate storage could be investigated. The queueing systems examined in this paper could be generalized to Possion arrivals and general service times, or a general distribution for interarrival times and negative exponential service times, etc. The output process of these systems could be investigated. Also it would be interesting to know how a finite waiting room effects the system.

REFERENCES

1. Ash, Robert B., Complex Variables. New York: Academic Press, 1971. Pp. viii + 255.

2. Chung, Kai Lai, Markov Chains. New York: Springer-Verlag, 1967. Pp. xi + 301.

3. Cohen, J. W., The Single Server Queue. Amsterdam: North-Holland Publishing Company, 1969. Pp. xiv + 657.

4. Foster, F. G., "On the Stochastic Matrices Associated with Certain Queueing Processes." Annals of Mathematical Statistics, 24(1953), 355-360.

5. Jordan, Charles, Calculus of Finite Differences. New York: Chelsea Publishing Company, 1965. pp. xxi + 652.

6. Karlin, Samuel, A First Course in Stochastic Processes. New York: Academic Press, 1966. Pp. xi + 502.

7. Kemeny, John G. and Snell, J. Laurie, Finite Markov Chains. Princeton: D. Van Nostrand Company, Inc. 1960. Pp. viii + 210.

8. Miller, Kenneth S., Linear Difference Equations. New York: W. A. Benjamin, Inc., 1968. Pp. x + 105.

9. Roberts, Fred S., Discrete Mathematical Models. Englewood Cliffs: Prentice-Hall, Inc., 1976. Pp. xvi + 559.

10. Saaty, Thomas L., Elements of Queueing Theory with Applications. New York: McGraw-Hill Book Company, Inc., 1961. Pp. xv + 423.