[1]This research was supported by the Office of Naval Research, Contract No. ONR-N00014-67-A-0321-0002.

SOME RELATED QUEUEING MODELS WITH DEPENDENT SERVICE AND INTER-ARRIVAL TIMES

Colin E. Jeffcoat

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 876
June, 1973

TABLE OF CONTENTS

COLIN E. JEFFCOAT. Some Related Queueing Models with Dependent Service and Inter-arrival times. (Under the direction of WALTER L. SMITH).

A B S T R A C T

In this dissertation we discuss some related queueing models which allow for a type of dependence between the service and inter-arrival times. We assume that there are a finite number of customer types and that the service time of a customer depends on either the preceding inter-arrival time (backward models) or the following arrival time (forward models) through the customer type. Particular cases of the general models are the transition models, for which arrivals are generated by transitions in a Markov chain in continuous time. Possible applications for these models are production lines, doctor's appointments and storage situations.

We analyse the stationary behaviour of the forward and backward models. The case of the forward model has been examined before, and quite general results are possible. We consider how these results relate to the forward transition model. A complete analysis of the stationary behaviour for backward models is achieved only in the particular case of exponential arrivals, where we show how to compute the Laplace transform of the stationary distribution of total waiting times.

We also analyse the busy period distribution for this last model, and show how to compute the busy period transform.

ACKNOWLEDGMENTS

I wish to thank Dr. W. L. Smith, my thesis adviser for his help and encouragement during the preparation of this work. I would like to thank the members of my committee, Dr. W. Hoeffding, Dr. R. J. Cannon, Dr. K. J. C. Smith and Dr. E. J. Wegman, for their cooperation, and Dr. S. Cambanis for his assistance at a crucial stage.

Mrs. Pamela Smith did an excellent job of the typing for which I am grateful. I would like to thank my wife, Mary, for her constant support. In helping me through my stay at the University of North Carolina I would like to thank the Department of Statistics, the taxpayers of North Carolina and most of all my friends.

NOTATION

arg(s)	argument of s
a.s.	almost surely, i.e. with probability 1
def	definition
deg	degree (of a polynomial)
d.f.	distribution function
iid	independent and identically distributed
Im(s)	imaginary part of s
I_n	the $n \times n$ unit matrix
$\mathbb{R}$	the set of real numbers
Re(s)	the real part of s

We abuse the usual set notation by writing $\{s : s$ has property $P\}$ as $\{s$ has property $P\}$.

$F^{o}(s), F^{*}(s)$ denote the Laplace and Laplace-Stieltjes transform of F (see Appendix).

All integrals of the form $\int_0^\infty g(x)\, F(dx)$ are understood to include the lower limit.

$f(x) = o(g(x))$ as $x \to 0$ or ∞ means

$$\frac{f(x)}{g(x)} \to 0 \quad \text{as} \quad x \to 0 \text{ or } \infty$$

$f(x) = O(g(x))$ as $x \to 0$ or ∞ means

$$\frac{f(x)}{g(x)} \text{ is bounded as } x \to 0 \text{ or } \infty .$$

□ denotes the end of a proof or result.

CHAPTER 1

INTRODUCTION

1.1 *Introduction to Queueing Theory*

A queue is created when customers arrive at a service facility and wait for service in the event that it is not immediately available. The probabilistic description of a queueing model is specified by

a) the distribution of arrival times,

b) the number of servers,

c) the queue discipline, i.e. the order of service,

d) the distribution of service times.

Throughout this dissertation we consider only single-server queues in which the customers are served in order of arrival (first come, first served). If two or more customers arrive simultaneously, we assume that there exists a procedure for deciding precedence.

Observation of the queue is assumed to begin at time 0. Let U_o be the time of the first arrival, and U_n the time between the n^{th} and $(n+1)^{th}$ arrivals $(n = 1,2,\ldots)$. V_n will denote the service time of the n^{th} customer $(n = 1,2,\ldots)$, with V_o the sum of the service times of any customers already in the system at time 0.

The mathematically interesting aspects of a queue are the number of customers in the system at one time, the amount of time a customer has to wait for service and the duration of busy periods, i.e. the time intervals during which the server is continuously busy. We define the queueing time of a customer to be the time between his arrival and commencement of service. Total waiting time is defined to be the time from arrival to completion of service.

Analysis of the transient (time-dependent) behaviour of a queue is usually quite complicated. A common procedure is to show that a unique equilibrium distribution exists and assume in applications that the system has reached equilibrium. We use the terms stationary distribution and equilibrium distribution interchangeably to mean a limiting distribution in some sense. For example, if $\{Q_n\}$ is a sequence of random variables with F_n the d.f. of Q_n, then the stationary distribution of $\{Q_n\}$ is $\lim_{n\to\infty} F_n$, provided this limit exists and is a proper d.f.

1.2 *Correlated Queues*

Most investigations of the single-server queue assume that the inter-arrival times $\{U_n\}$ and service times $\{V_n\}$ are independent sequences of iid random variables. We wish to consider queueing models for which this independence assumption does not hold; following Conolly [2] we call these correlated queues.

Several authors have considered various types of dependence between service and inter-arrival times. D. R. Cox and W. L. Smith analysed a model which had $\{U_n\}$ and $\{V_n\}$ related through the queue size at arrival epochs [5; § 2.4]. C. M. Harris discussed in [7] a similar case where service time is dependent on the queue size at the moment when service begins.

M.F. Neuts investigated in [16] a model where arrivals and service times were related through an extraneous phase process: during phase i arrivals are Poisson with parameter λ_i and services initiated during this phase have d.f. H_i.

R. M. Loynes generalized some results of D. V. Lindley concerning equilibrium distributions to the case where $\{(U_n, V_n)\}$ is a strictly stationary metrically transitive sequence [12]. In a subsequent paper [13], he showed how the stationary distribution of queueing time may be found when the queue has a certain structure.

We are interested in two related subclasses of correlated queues, which may be characterized as follows:

(I) queues with backward dependence: U_n and V_{n+1} are dependent $n = 0,1,2,\ldots$, all other pairs U_n, V_m being independent;

(II) queues with forward dependence: U_n and V_n are dependent, $n = 0,1,2,\ldots$, all other pairs U_n, V_m being independent.

With backward dependence, the service time of a customer is related to the arrival interval preceding his arrival. With forward dependence the service time is related to the following arrival interval.

B. W. Conolly in [2] and [3] (with N. Hadidi) analysed a particular backward model with V_{n+1} directly proportional to U_n. The motivation given for studying this type of model was that such queues may be made self-regulating: the server may adjust V_{n+1} according to the value of U_n in order to stabilize the queue, i.e. to reduce waiting times for a fixed traffic intensity.

F. I. John discussed a general method applicable to queues with forward dependence. As possible applications he mentions a production line and appointments at a doctor's office. In both cases the time between arrivals "should depend strongly on the estimated service times" [8].

Conolly claims in [2] that forward models cannot be interpreted usefully as self-regulating systems. However, such an interpretation can certainly be made if service times are ascertainable immediately on arrival (as in storage applications), or if an estimate of service time is available then, as in John's examples.

Backward dependence does seem more natural, however, since it is quite plausible that the service time of a customer might influence the time required for arrival; this is especially so when the queue is interpreted as an infinite store (or dam) as in [20], [14], for then service times correspond to size of deliveries.

1.3 *Definition of the basic models*

We now define the queueing models that we deal with in this dissertation. First we define the general backward and forward models, then the particular cases of the transition models.

The so-called general models are not the most general models as described in the previous section. Instead they are the special cases of (I) and (II) where there are a finite number of customer types, and $\{U_n\}$ and $\{V_n\}$ are related only through the customer types. This does restrict the range of application considerably, but some such simplification appears necessary if usable results are to be obtained (c.f. § 1 of [13]).

We define the *general backward model* as follows: There are κ customer types represented by $1,2,\ldots,\kappa$. The (unconditional) probability that an arriving customer be of type j is π_j, where $\pi_j > 0$ for each j and $\sum_j \pi_j = 1$. Let Z_n be the type of the n^{th} customer. Then we suppose that $\{U_n : n\geq 1\}$ and $\{V_n : n\geq 1\}$ are iid with

$$\Pr\{U_n \leq x \mid \{Z_m\},\{V_m\}\} = A_{Z_{n+1}}(x) \quad (n = 1,2,\ldots) ,$$

$$\Pr\{V_n \leq x \mid \{Z_m\},\{U_m\}\} = B_{Z_n}(x) \quad (n = 1,2,\ldots) .$$

Clearly U_n and V_{n+1} are related through Z_{n+1}, but conditionally independent given Z_{n+1}.

The definition of the *general forward model* is the same except that U_n is related to Z_n, i.e.

$$\Pr\{U_n \leq x \mid \{Z_m\},\{V_m\}\} = A_{Z_n}(x) \quad (n = 1,2,\ldots) .$$ Thus for the general forward model, U_n and V_n are related through Z_n, but conditionally independent given Z_n.

The transition models are special cases of these general models, with inter-arrival distributions belonging to a subclass of those distributions with rational Laplace transforms. In particular, the arrivals correspond to certain transitions in a finite Markov chain in continuous time.

We suppose that $\{Z_t, t\geq 0\}$ is an homogeneous Markov chain with state space $\{0,1,\ldots,\kappa\}$ and infinitesimal transition probabilities

$$\Pr\{Z_{t+h} = j \mid Z_t = i\} = (1-\lambda_i h)\delta_{ij} + \lambda_i h p_{ij} + o(h)$$

for $t,h>0$, where $\lambda_0,\lambda_1,\ldots,\lambda_\kappa$ are positive real numbers and $\mathbb{P} = [p_{ij}]$ is an irreducible Markov transition matrix.

Realizations of (Z_t) have the following form: the process stays in state i for a time period which is exponentially distributed with parameter λ_i, then jumps to state j according to the probabilities $\{p_{ij}\}$, $j \in \{0,1,\ldots,\kappa\}$ (see [9; pp. 226-9]) . Note that a "jump" is not necessarily a change of state, since transitions $i \to i$ may be possible. We call (Z_t) the guiding chain.

We define the *backward transition model* as follows. There are $\kappa+1$ customer types represented by $0,1,\ldots,\kappa$. Arrivals are generated by transitions into state 0 of the guiding chain; in particular, a transition $j \to 0$ is identified with the arrival of a type j customer, with service time d.f. B_j .

The *forward transition model* is defined similarly, except that arrivals correspond to transitions out of state 0: a transition $0 \to j$ is identified with the arrival of a type j customer.

The inter-arrival distributions for the transition models are (possibly infinite) mixtures of Erlang distributions, since each arrival interval is the sum of a variable number of exponential phases. Actually, it is possible to show (as in Lemma 2.12) that each A_j has a rational Laplace transform for these models. As noted by Loynes in [14], it is possible to construct a large class of inter-arrival distributions by varying $\kappa,\{\lambda_j\}$ and $\mathbb{P}$ (see also [4]).

An annoying feature of the transition models is that every arrival interval begins or ends with an exponentially distributed phase with parameter λ_o . The resulting difficulties in interpretation may be minimised by taking λ_o very large; it would be convenient to be able to set $\lambda_o = \infty$. Instantaneous states may lead to theoretical difficulties; however, it is possible to redefine the transition models so as to incorporate the effect of having $\lambda_o = \infty$, provided $p_{oo} = 0$. This is done for the backward model in Section 4.4.

With $\lambda_o = \infty$, we may represent the backward model with exponential arrivals (as treated in Chapters 3 and 4) as a transition model with

$$\mathbb{P} = \begin{bmatrix} 0 & \pi_1 & \pi_2 \cdots \pi_K \\ 1 & 0 & 0 \ldots 0 \\ \cdots & & \\ 1 & 0 & 0 \ldots 0 \end{bmatrix} .$$

More complicated inter-arrival distributions can be constructed by choosing submatrices of $\mathbb{P}$ appropriately. For example, Erlangian and hyperexponential inter-arrival distributions are generated by submatrices of the form

$$\begin{bmatrix} 0 & 1 & 0 \ldots 0 \\ 0 & 0 & 1 \ldots 0 \\ \cdots & & \\ 0 & 0 & 0 \ldots 1 \\ 1 & 0 & 0 \ldots 0 \end{bmatrix} , \qquad \begin{bmatrix} 0 & p_1 & p_2 \cdots p_n & 0 \\ 0 & 0 & 0 \ldots 0 & 1 \\ \cdots & & & \\ 0 & 0 & 0 \ldots 0 & 1 \\ 1 & 0 & 0 \ldots 0 & 0 \end{bmatrix}$$

1.4 *Summary of Known Results*

The general backward model does not seem to have been investigated before. Results concerning the stationary distribution of queueing time are implied for the general forward model by results in [13].

W. L. Smith analysed the transient behaviour of the forward transition model in some unpublished work (c. 1958 - see [21]). He obtained an explicit expression for the Laplace transform of the busy period distribution.

R. M. Loynes considered in [14] what is essentially a more general transition model, including both the forward and backward transition models as special cases. His analysis of the transient behaviour did not lead to any explicit solutions.

1.5 *Outline of the dissertation*

The main objective of this dissertation is to obtain some useful results for the backward models. We have not been able to derive explicit solutions for the busy period transform for the backward transition model as in the case of the forward transition model.

In Chapter 2 we apply the methods of W. L. Smith [20] to the general forward model to obtain results on the stationary distribution of queueing time. Actually, the main results of this chapter are contained in [13]; they are derived independently here and included for completeness and comparison with the results of Chapter 3. We also show how these results apply to the forward transition model.

In Chapter 3 we consider the stationary behaviour of the general backward model. Instead of proceeding as in Chapter 2 we must consider here the total waiting times. To solve the resulting integral equation we are forced to restrict ourselves to the case of exponential arrivals.

In Chapter 4 we consider busy periods for the backward models. In the case of exponential arrivals we generalize a procedure used in [5] to obtain the busy period transform. This procedure may be partially generalized to the backward transition model with $\lambda_o = \infty$.

CHAPTER 2

THE STATIONARY DISTRIBUTION OF QUEUEING TIME FOR THE FORWARD MODELS

2.1 *Introduction*

The stationary distribution of queueing time was investigated for the G/G/1 queue by D. V. Lindley in [11] and W. L. Smith in [20]. In this chapter we apply Smith's methods to the general forward model.

Section 2.2 is essentially a complete derivation of Smith's main results with only a few minor changes.

In Section 2.3 we obtain some more practical results (following §§ 3.4, 3.5, 4.2 and 4.3 in [20]) by making further assumptions on the service time or inter-arrival time distributions. We develop the latter case in more detail and show how it applies to the forward transition model.

2.2 In this section we consider the general forward model with κ customer types. We describe this model as follows: at each arrival epoch the probability that the arriving customer be of type j is π_j, independently of any previous events ($j = 1,2,\ldots,\kappa$). We assume $\pi_j > 0$ for all j. The service time d.f. of a type j customer is B_j, and the time to the next arrival after a type j customer has d.f. A_j.

Let Z_n denote the type of the n^{th} customer, V_n his service time, and U_n the time between the n^{th} and $(n+1)^{th}$ arrivals. Then for $n \geq 1$, we have $\Pr\{Z_n = j\} = \pi_j$, $\Pr\{V_n \leq x \mid Z_n = j\} = B_j(x)$, and $\Pr\{U_n \leq x \mid Z_n = j\} = A_j(x)$. $V_1, V_2, \ldots$ are iid random variables, as are $U_1, U_2, \ldots$; we leave unspecified the distributions of U_o, the time to the first arrival, and V_o, the residual service time at 0. U_n and V_n are conditionally independent given Z_n for $n \geq 1$. Let $X_n = V_n - U_n$ $(n = 0,1,2,\ldots)$. Then $X_1, X_2, \ldots$ are iid random variables, which enables us to apply the methods of Lindley and Smith. We assume that A_j and B_j have finite means

$$a_j = \int_0^\infty x \, A_j(dx), \; b_j = \int_0^\infty x \, B_j(dx) \; .$$

Let $a = \sum_j \pi_j a_j$, $b = \sum_j \pi_j b_j$. Set $H_j(x) = \int_0^\infty B_j(x+y) \, A_j(dy)$, and $H(x) = \sum_{j=1}^{K} \pi_j H_j(x)$. Then $H_j(x) = \Pr\{X_n \leq x \mid Z_n = j\}$ and H is the d.f. of X_n $(n = 1,2,\ldots)$.

Let Q_n be the queueing time of the n^{th} customer, i.e. the time from his arrival to commencement of service $(n = 1,2,\ldots)$. Take $Q_o = 0$. Then we have the familiar relation (see [11])

$$Q_{n+1} = \max (Q_n + X_n, 0) \quad (n = 0,1,2,\ldots) \; . \tag{2.1}$$

Q_{n+1} is determined by $\{(U_k, V_k): 0 \leq k \leq n\}$, so clearly $\{Q_n\}$ is a Markov process. Note that Q_n is independent of Z_n and X_n. For $n = 0,1,2,\ldots$ let $F_n(x) = \Pr\{Q_n \leq x\}$. Then for $x \geq 0$

$$F_{n+1}(x) = \int_0^\infty \Pr\{Q_{n+1} \le x \mid Q_n = u\} \, F_n(du)$$

$$= \int_0^\infty \Pr\{X_n \le x-u\} \, F_n(du)$$

by (2.1), since Q_n is independent of X_n. Thus

$$F_{n+1}(x) = \int_0^\infty H(x-u) \, F_n(du), \quad \text{for} \quad x \ge 0, \; n = 0,1,2,\ldots \tag{2.2}$$

We show that under the usual first moment condition $\{Q_n\}$ has a stationary distribution, i.e. $\lim_{n\to\infty} F_n(x)$ exists and gives a proper distribution function. Assume

$$\sum_{j=1}^{K} \pi_j (b_j - a_j) < 0, \quad \text{i.e.} \quad EX_1 < 0 \, . \tag{A1}$$

LEMMA 2.1: Under assumption (A1), $\{Q_n\}$ has a stationary distribution.

PROOF: See §§ 3,4 of [11]. Lindley's proof applies to our model also [see note on p. 278 of [11]]. □

Let F be the (unique) stationary distribution function, i.e. $F(x) = \lim_{n\to\infty} F_n(x)$. Then F satisfies

$$F(x) = \int_0^\infty H(x-u) \, F(du) \quad (x \ge 0) \tag{2.3}$$

from (2.2) and the dominated convergence theorem. (2.3) corresponds to Equation (4) in [20].

We now follow the Wiener-Hopf procedure as outlined by Smith. First we define

$$F_1(x) = \begin{cases} \int_0^\infty H(x-u)\ F(du) & \text{for } x < 0 , \\ 0 & \text{for } x \geq 0 . \end{cases}$$

Then for any real x we have

$$F(x) + F_1(x) = \int_0^\infty H(x-u)\ F(du). \tag{2.4}$$

If we take Laplace-Stieltjes transforms of both sides of (2.4), then we obtain

$$F^*(s) + F_1{}^*(s) = H^*(s)\ F^*(s)\ . \tag{2.5}$$

We know that both sides of (2.5) are finite at least for s on the imaginary axis; we can extend this range by making the following assumption: (A2) There exists $\mu' > 0$ such that for each $j \in \{1,2,\ldots,\kappa\}$

$$\int_0^\infty e^{\mu' x}\ B_j(dx) < \infty,\ \int_0^\infty e^{\mu' x}\ A_j(dx) < \infty\ .$$

Assumption (A2) implies that $B_j{}^*(s)$ is analytic in $\{Re(s) > -\mu'\}$ and $A_j{}^*(-s)$ is analytic in $\{Re(s) < \mu'\}$ for each j. It follows that $H^*(s) = \sum_{j=1}^{\kappa} \pi_j A_j{}^*(-s) B_j{}^*(s)$ is analytic in $S_\mu \underset{\text{def}}{=} \{s : -\mu \leq Re(s) \leq \mu\}$ for some $\mu > 0$. Fix μ with this property, e.g. $\mu = \mu'/_2$.

LEMMA 2.2: <u>Under assumption (A2), $F_1{}^*(s)$ converges for $Re(s) \leq \mu$.</u>

PROOF: Let $A(x) = \sum_{j=1}^{K} \pi_j A_j(x)$, i.e. A is the unconditional d.f. of U_n, for $n \geq 1$. Assumption (A2) implies that $1 - A(x) = o(e^{-\mu' x})$ as $x \to +\infty$ (Theorem 2.2b in [23]). Now for $x > 0$

$$(0\leq)\ F_1(-x) = \int_0^\infty H(-x-u)\ F(du)$$

$$\leq H(-x)$$

$$= \Pr\{V_1 - U_1 \leq -x\}$$

$$\leq \Pr\{U_1 \geq x\}$$

$$= 1 - A(x-0)\ .$$

Therefore $F_1(-x) = o(e^{-\mu' x})$ as $x \to +\infty$, so by Theorem 2.1 in [23], $F_1^*(s)$ converges for $\mathrm{Re}(s) < \mu'$. □

Thus, under assumption (A2) we see that (2.5) is valid at least for all values of s in the strip $\{0 \leq \mathrm{Re}(s) \leq \mu\}$.

Our next step is to factorize $1 - H^*(s)$. We need a further assumption: (A3) There exists $\delta > 0$ such that $H^*(s) = O(|s|^{-\delta})$ as $|s| \to \infty$ in S_μ.

REMARK: Assumptions (A2) and (A3) are satisfied in most practical applications of queueing theory, e.g. if $A_1,\ldots,A_K$, $B_1,\ldots B_K$ are exponential, hyperexponential or Erlangian distribution functions. (A3) certainly implies that H cannot be a lattice d.f., for if H puts all probability at points $a + md$, $m = 0, \pm 1, \pm 2, \ldots$, then $\left|H^*\left(\frac{2n\pi i}{d}\right)\right| = 1$ for $n = 1,2,\ldots,$ which contradicts (A3). Throughout the rest of this section we will assume (A1), (A2) and (A3) implicitly.

LEMMA 2.3: We can choose $\mu > 0$ so that $1 - H^*(s)$ has only one zero in S_μ.

PROOF: $1 - H^*(s)$ has a single zero at $s = 0$, since $H^*(s) = 1 + (a-b)s + o(|s|)$ as $|s| \to 0$. There are no other zeros on the imaginary axis, since H is a lattice d.f. if $|H^*(i\theta)| = 1$ for some $\theta \in \mathbb{R} - \{0\}$, and this is ruled out by assumption (A3).

It also follows from (A3) that there exists $\Delta > 0$ such that $1 - H^*(s)$ has no zeros in $S_\mu \cap \{s : |\mathrm{Im}(s)| > \Delta\}$. Inside the closed, bounded rectangle $S_\mu \cap \{s : |\mathrm{Im}(s)| \le \Delta\}$, $1 - H^*(s)$ is analytic and therefore can have only a finite number of zeros. Since all but one of these zeros must lie off the imaginary axis, we can choose μ so that the zero at $s = 0$ is the only one contained in S_μ. □

In order to factorize $1 - H^*(s)$ we consider the function $g(s)$ defined as follows: fix λ real so that $\lambda > \mu$ and set

$$g(s) = \frac{s-\lambda}{s}\{1 - H^*(s)\} \quad \text{for} \quad s \in S_\mu - \{0\},$$

$$g(0) = \lambda(a-b).$$

Clearly $g(s)$ is analytic in $S_\mu - \{0\}$. It is easy to show that $g(s)$ has a power series expansion around $s = 0$, so that $g(s)$ is analytic everywhere in S_μ.

LEMMA 2.4: $g(s)$ is bounded in S_μ, and also bounded away from zero.

PROOF: For $|s|$ large, $\left|\frac{s-\lambda}{s}\right| \sim 1$ and $|1 - H^*(s)| \le 1 + M|s|^{-\delta}$

by (A3). Thus $g(s)$ is bounded in some set $S_\mu \cap \{s : |s| > \Delta\}$. Clearly $g(s)$ is bounded in the closed, bounded set $S_\mu \cap \{s : |s| \le \Delta\}$. Thus $g(s)$ is bounded in S_μ. Since $g(s)$ has no zeros in S_μ and $|g(s)| \sim 1$ for $|s|$ large, it follows by the continuity of g that $g(s)$ is bounded away from zero in S_μ. □

LEMMA 2.5: <u>It is possible to define $\arg\{g(s)\}$ so that</u>

$$\arg\{g(s)\} < \pi \quad \underline{\text{for } s \in S_\mu} .$$

PROOF. Consider first points s on the imaginary axis: suppose $s = i\tau, \tau \in \mathbb{R} - \{0\}$. In the proof of Lemma 2.3 we showed that $|H^*(i\tau)| \neq 1$, which implies $|H^*(i\tau)| < 1$. Therefore $1 - H^*(i\tau) \in \{z : |z-1| < 1\}$ and $\mathrm{Re}\{1 - H^*(i\tau)\} > 0$. Since

$$\mathrm{Re}\left\{\frac{i\tau-\lambda}{i\tau}\right\} = 1$$

we can define $\arg\{1 - H^*(i\tau)\}$ and $\arg\left\{\frac{i\tau-\lambda}{i\tau}\right\}$ continuously so that

$$|\arg\{1 - H^*(i\tau)\}| < \pi/2 , \quad \left|\arg\frac{i\tau-\lambda}{i\tau}\right| < \pi/2 .$$

Since $g(0) = \lambda(a-b) > 0$, it follows that we can define $\arg\{g(s)\}$ continuously on the imaginary axis so that $|\arg\{g(s)\}| < \pi$.

For $|s|$ large, we have from the proof of Lemma 2.4 that $g(s) = 1 + O(|s|^{-\delta}) \to 1$ as $|s| \to \infty$ in S_μ. Thus $\arg\{g(s)\} \to 0$, so that there exists $\Delta > 0$ such that $|\arg\{g(s)\}| < \pi$ for all $s \in S_\mu \cap \{z : |z| > \Delta\}$.

The set $S_\mu \cap \{z : |z| \le \Delta\}$ is compact, and $|\arg\{g(s)\}| < \pi$ for all s in the closed subset $\{z : \mathrm{Re}(z) = 0, |z| \le \Delta\}$. Therefore, by Lemma A3 , since $\arg\{g(s)\}$ is a continuous function of s, there exists $\varepsilon > 0$ such that $|\arg\{g(s)\}| < \pi$ for all $s \in S_\varepsilon \cap \{z : |z| \le \Delta\}$. Since ε is independent of μ, we may replace μ by $\min(\mu,\varepsilon)$. Then $|\arg\{g(s)\}| < \pi$ for all $s \in S_\mu$. □

From Lemmas 2.4 and 2.5 it follows that $G(s) \underset{\mathrm{def}}{=} \log g(s)$ is a well-defined, single-valued, analytic, bounded function of s in S_μ.

LEMMA 2.6: There exists $M \ge 0$ such that for all $s \in S_\mu$

$$|G(s)| \le \frac{M}{1 + |s|^\delta} .$$

PROOF: As $|s| \to \infty$ with $s \in S_\mu$ we have

$$\begin{aligned} |G(s)| &\le |\log|g(s)|| + |\arg\{g(s)\}| \\ &= |\log|1 + O(|s|^{-\delta})|| + |\tan^{-1} O(|s|^{-\delta})| \\ &= O(|s|^{-\delta}) . \end{aligned}$$

Thus there exist $\Delta > 0$, $M_1 \ge 0$ such that $|G(s)| \le M_1 |s|^{-\delta}$ for $|s| > \Delta$. Since $G(s)$ is bounded in S_μ, there exists $M_2 \ge 0$ such that $|G(s)| \le M_2$. If we choose $M = \max[M_1(1+\Delta^{-\delta}), M_2(1+\Delta^{\delta})]$, then the result follows. □

Consider the rectangular contour Λ_T obtained by truncating S_μ at $\mathrm{Im}(s) = \pm T$ $(T>0)$, i.e. Λ_T is the rectangle with vertices

$\pm\,\mu \pm iT$. Suppose z is any point inside S_μ. Then z will be inside Λ_T for T sufficiently large; since $G(s)$ is analytic on and inside Λ_T we have by Cauchy's integral formula

$$G(z) = \frac{1}{2\pi i} \int_{\Lambda_T} \frac{G(s)}{s-z}\, ds \; .$$

Now from the proof of Lemma 2.6

$$\left| \int_{-\mu+iT}^{\mu+iT} \frac{G(s)}{s-z}\, ds \right| \le 2\pi \sup \left\{ \frac{G(s)}{s-z} : \operatorname{Im}(s) = T, s \in S_\mu \right\}$$

$$= O(T^{-1-\delta}) \quad \text{as} \quad T \to \infty \; .$$

Thus

$$\int_{-\mu+iT}^{\mu+iT} \frac{G(s)}{s-z}\, ds \to 0 \quad \text{as} \quad T \to \infty \; .$$

Similarly

$$\int_{-\mu-iT}^{\mu-iT} \frac{G(s)}{s-z}\, ds \to 0 \quad \text{as} \quad T \to \infty \; .$$

Define

$$I_1(z) = \int_{\mu-i\infty}^{\mu+i\infty} \frac{G(s)}{s-z}\, ds, \quad I_2(z) = \int_{-\mu-i\infty}^{-\mu+i\infty} \frac{G(s)}{s-z}\, ds \quad (|\operatorname{Re}(z)| < \mu) \; .$$

Then $G(z) = \frac{1}{2\pi i} \{I_1(z) - I_2(z)\}$.

LEMMA 2.7: <u>$I_1(z)$ (resp. $I_2(z)$) is bounded and analytic for $\operatorname{Re}(z) \le \mu - \varepsilon$ (resp. $\operatorname{Re}(z) \ge -\mu + \varepsilon$) for any $\varepsilon > 0$.</u>

PROOF: Suppose $z = x + iy$ where $x \le \mu - \varepsilon$. Then

$$|I_1(z)| = \left| \int_{-\infty}^{\infty} \frac{G(\mu+it)}{\mu-x+i(t-y)} \, dt \right|$$

$$\le \int_{-\infty}^{\infty} \frac{M}{1+|\mu+it|^{\delta}} \cdot \frac{1}{[(\mu-x)^2+(t-y)^2]^{1/2}} \, dt \quad \text{by Lemma 2.6}$$

$$\le \int_{-\infty}^{\infty} \frac{M}{1+|t|^{\delta}} \cdot \frac{1}{[\varepsilon^2+(t-y)^2]^{1/2}} \, dt \, .$$

Choose $p,q > 1$ so that $p\delta > 1$ and $p^{-1} + q^{-1} = 1$. Then, since $\frac{M}{1+|t|^{\delta}} \in L_p$ and $[\varepsilon^2+(t-y)^2]^{-1/2} \in L_q$, we have by Hölder's inequality

$$|I_1(z)| \le \left\{ \int_{-\infty}^{\infty} \frac{M^p}{(1+|t|^{\delta})^p} \, dt \right\}^{1/p} \cdot \left\{ \int_{-\infty}^{\infty} \frac{dt}{[\varepsilon^2+(t-y)^2]^{q/2}} \right\}^{1/q}$$

$$= \left\{ \int_{-\infty}^{\infty} \frac{M^p}{(1+|t|^{\delta})^p} \, dt \right\}^{1/p} \left\{ \int_{-\infty}^{\infty} \frac{du}{(\varepsilon^2+u^2)^{q/2}} \right\}^{1/q} .$$

Since this last bound is finite and independent of z it follows that $I_1(z)$ is bounded for $\mathrm{Re}(z) \le \mu - \varepsilon$.

To show that $I_1(z)$ is analytic in $\{\mathrm{Re}(z) \le \mu - \varepsilon\}$ consider $z = x + iy$, where $x \le \mu - \varepsilon$, and $h \ne 0$. We have

$$\frac{1}{h}\{I_1(z+h) - I_1(z)\} = \int_{\mu-i\infty}^{\mu+i\infty} \frac{G(s)}{(s-z)(s-z-h)}\,ds$$

$$= \int_{-\infty}^{\infty} \frac{G(\mu+it)}{\{\mu-x+i(t-y)\}\{\mu-x-h+i(t-y)\}}\,dt\,.$$

If $h \le \varepsilon/2$, then the integrand is bounded by

$$\frac{M}{(1+|t|^{\delta})\{[\varepsilon^2+(t-y)^2][\varepsilon^2/4+(t-y)^2]\}^{1/2}} \in L_1 \ .$$

Therefore by the dominated convergence theorem

$$\lim_{h\to 0}\frac{1}{h}\{I_1(z+h) - I_1(z)\} = \int_{\mu-i\infty}^{\mu+i\infty} \frac{G(s)}{(s-z)^2}\,ds \ (< \infty)\ .$$

Thus $I_1(z)$ is analytic in $\{\mathrm{Re}(z) \le \mu - \varepsilon\}$. The proofs of the corresponding results for $I_2(z)$ are similar. □

We may now prove the main result of this section.

THEOREM 2.1: Under assumptions (A1), (A2), (A3), there exists $\mu > 0$ such that for fixed $\lambda > \mu$

$$1 - H^*(s) = \frac{s\Phi^+(s)}{(s-\lambda)\Phi^-(s)} \quad \text{for } s \text{ in } \{|\mathrm{Re}(s)| < \mu\}, \tag{2.6}$$

where $\Phi^+(s)$ (resp. $\Phi^-(s)$) is analytic, bounded and zero-free in $\{\mathrm{Re}(s) \ge -\mu + \varepsilon\}$ for any $\varepsilon > 0$ (resp. $\{\mathrm{Re}(s) \le \mu - \varepsilon\}$). The Laplace-Stieltjes transform of the stationary queueing time distribution is

$$F^*(s) = \frac{\Phi^+(0)}{\Phi^+(s)} \quad \underline{\text{for} \quad \mathrm{Re}(s) \geq 0} \; . \tag{2.7}$$

PROOF: We know that for $s \in \{|\mathrm{Re}(s)| < \mu\}$

$$1 - H^*(s) = \frac{s}{s-\lambda} \circ e^{G(s)}$$

and we have shown above that $G(s) = \frac{1}{2\pi i} \{I_1(s) - I_2(s)\}$. Let

$$\Phi^+(s) = \exp\left\{-\frac{1}{2\pi i} I_2(s)\right\}, \; \Phi^-(s) = \exp\left\{-\frac{1}{2\pi i} I_1(s)\right\} .$$

Then Equation (2.6) follows. By Lemma 2.7 $\Phi^+(s)$ is analytic and bounded in $\{\mathrm{Re}(s) \geq -\mu + \varepsilon\}$. Clearly $\Phi^+(s)$ has no zeros in $\{\mathrm{Re}(s) \geq -\mu + \varepsilon\}$. The properties of $\Phi^-(s)$ follow similarly.

From (2.5) and (2.6) we have

$$s\Phi^+(s)F^*(s) = -(s-\lambda)\Phi^-(s)F_1^{\;*}(s) \quad \text{for} \quad 0 \leq \mathrm{Re}(s) < \mu \; . \tag{2.8}$$

Now $\Phi^+(s)$ and $F^*(s)$ are analytic and bounded for $\mathrm{Re}(s) > 0$, so $s\Phi^+(s)F^*(s)$ is analytic and $0(|s|)$ as $|s| \to \infty$ in $\{\mathrm{Re}(s) > 0\}$. $F_1^{\;*}(s)$ is analytic and bounded for $\mathrm{Re}(s) \leq \mu$ by Lemma 2.2. Also $\Phi^-(s)$ is analytic and bounded for $\mathrm{Re}(s) \leq {}^{\mu}/_{2}$. Therefore $(s-\lambda)\Phi^-(s)F_1^{\;*}(s)$ is analytic and $0(|s|)$ as $|s| \to \infty$ in $\{\mathrm{Re}(s) \leq {}^{\mu}/_{2}\}$.

Thus the two sides of (2.8) represent a function which is analytic in the finite complex plane and $0(|s|)$ as $|s| \to \infty$. It follows from a variant of Liouville's Theorem [22; 2.52] that this function must be linear, i.e. there exist complex numbers α, β such that

$$s\Phi^{+}(s)F^{*}(s) = \alpha + \beta s \quad \text{for} \quad \mathrm{Re}(s) \geq 0 \ .$$

Putting $s = 0$ shows that $\alpha = 0$, and by continuity $\beta = \Phi^{+}(0)$. Therefore (2.8) follows. □

COROLLARY: The cumulant-generating function of queueing time is

$$\log F^{*}(s) = \frac{s}{2\pi i} \int_{-\mu-i\infty}^{-\mu+i\infty} \frac{1}{z(z-s)} \log\left\{\frac{z-\lambda}{z}\,(1-H^{*}(s))\right\} dz \ .$$

PROOF:

$$\log F^{*}(s) = \frac{1}{2\pi i}\{I_2(s) - I_2(0)\}$$

$$= \frac{1}{2\pi i} \int_{-\mu-i\infty}^{-\mu+i\infty} \frac{sG(z)}{z(z-s)}\, dz \ . \qquad \square$$

2.3 We may obtain more practical results by making further assumptions on either the service time or the inter-arrival time distributions. We prove one result with conditions on the service times on the basis of some results in [20]. Then we develop in more detail the case of conditions on the inter-arrival distributions; we show that the required conditions are always satisfied for the forward transition model. Throughout this section we retain assumptions (A1), (A2), (A3).

Let T be the class of distribution functions with rational Laplace-Stieltjes transform. We show that the stationary distribution of queueing time is in T if the service time distribution functions are in T and no customer has zero service time a.s. In addition, we are led to a formula for $F^*(s)$ which may be inverted easily in any particular example.

THEOREM 2.2: If

(i) $B_1,\ldots,B_K \in T$,

(ii) $B_j(0) < 1$ for each j ,

then $F \in T$; in particular we have

$$F^*(s) = \prod_{j=1}^{M} \left\{ \frac{1 - {}^s/c_j}{1 - {}^s/d_j} \right\} \quad \text{for } s \text{ in } \{Re(s) \geq 0\} , \tag{2.9}$$

where $c_1,\ldots,c_M$ are the poles of $H^*(s)$ in $\{Re(s) < 0\}$ and $d_1,\ldots,d_M$ are the zeros of $1 - H^*(s)$ in $\{Re(s) < 0\}$.

PROOF: Condition (i) implies that $H^*(s)$ is meromorphic in $\{Re(s) < 0\}$, i.e. has no singularities there other than a finite number of poles $c_1,\ldots,c_M$. We have $B_j^*(s) = \frac{P_j(s)}{Q_j(s)}$, where $P_j(s)$ and $Q_j(s)$ are polynomials in s such that $\deg\{P_j\} \leq \deg\{Q_j\}$. If $\deg\{P_j\} < \deg\{Q_j\}$, then $B_j^*(s) \to 0$ as $|s| \to \infty$ in any direction. If $\deg\{P_j\} = \deg\{Q_j\}$, then the ratio of the leading coefficient of P_j to that of Q_j is $B_j(0)$, since $B_j^*(s) \to B_j(0)$ as $s \to +\infty$ along the real axis. In either case, condition (ii) implies that $|B_j^*(s)| < 1$ for $|s|$ sufficiently large in $\{Re(s) < 0\}$. Since $|A_j^*(-s)| \leq 1$ for $Re(s) < 0$, it follows that $|H^*(s)| < 1$ for $|s|$ sufficiently large in $\{Re(s) < 0\}$. The theorem then follows by Theorems 1 and 3 in [20]. □

COROLLARY: The stationary probability of an arriving customer finding the server free is $\prod_{j=1}^{M} d_j/c_j$.

PROOF: $F(0) = \lim_{s\to\infty} F^*(s)$ (s real). □

APPLICATIONS

Suppose we know A_j, B_j $(j = 1,2,\ldots,\kappa)$. Then we can find the poles of $H^*(s)$ by inspection. To apply formula (2.9) we need only solve

$$\sum_{j=1}^{\kappa} \pi_j A_j^*(-s) B_j^*(s) = 1$$

in the left half-plane for $d_1,\ldots,d_M$. We can then find $F(x)$ by expanding

$$\prod_{j=1}^{M} \left\{\frac{1 - {}^s/c_j}{1 - {}^s/d_j}\right\}$$

in partial fractions and inverting termwise.

To obtain the corresponding result with conditions on the inter-arrival distributions we need to express $F^*(s)$ in terms of $\Phi^-(s)$ and consider the behaviour of $H^*(s)$ and $\Phi^-(s)$ in the right half-plane (see § 4.3 in [20]).

To prove an analogue of Smith's Theorem 1 we need a family of simple closed contours in the right half-plane on which the inequality $|H^*(s)| < 1$ is valid. We saw in the proof of Lemma 2.3 that $|H^*(i\theta)| < 1$ for $\theta \in \mathbb{R} - \{0\}$. In our next result we show that $|H^*(s)| < 1$ on a small semicircular arc to the left of $s = 0$.

LEMMA 2.9: There exists $\varepsilon > 0$ such that $|H^*(s)| < 1$ for all s in $\{|s| = \varepsilon, \mathrm{Re}(s) < 0\}$.

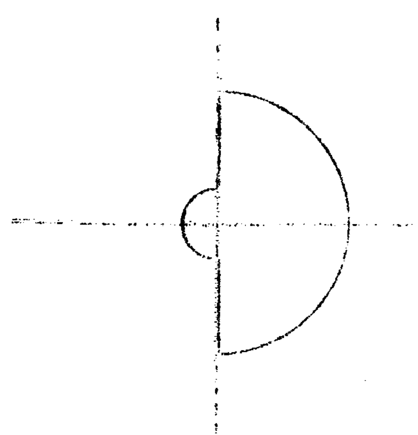
PROOF: $H^*(0) = 1$ and $H^{*\prime}(0) = a - b > 0$, so there exists ε in $(0,\mu)$ such that

$$0 < H^*(x) < 1 \text{ for all } x \text{ (real) in } (-\varepsilon,0) .$$

Suppose $\mathrm{Re}(s) = \sigma \in [-\varepsilon,0)$. Then

$$\begin{aligned} |H^*(s)| &\le \sum_{j=1}^{K} \pi_j |A_j^*(-s)||B_j^*(s)| \\ &\le \sum_{j=1}^{K} \pi_j A_j^*(-\sigma)B_j^*(\sigma) \\ &= H^*(\sigma) \\ &< 1 . \end{aligned}$$ □

Fix ε in $(0,\mu)$ with this property. For any $R > \varepsilon$, we define C_R to be the simple closed contour formed by joining the semi-circular arc of Lemma 2.9 with the arc $\{|z| = R, \mathrm{Re}(z) > 0\}$ along the imaginary axis, i.e. $C_R = \{|z|=\varepsilon,\mathrm{Re}(z)<0\} \cup \{\varepsilon<|z|<R,\mathrm{Re}(z)=0\} \cup \{|z|=R,\mathrm{Re}(z)>0\}$. C_R may be represented by the following diagram:



We have the following analogue of Smith's Lemma 2:

LEMMA 2.10: <u>If, for some $R > \varepsilon$,</u>

<u>(i) $H^*(s)$ is meromorphic inside C_R</u> ,

<u>(ii) $|H^*(s)| < 1$ on the arc $\{|s|=R, Re(s)>0\}$</u> ,

<u>then $1 - H^*(s)$ has the same number of zeros inside C_R as $H^*(s)$ has poles</u> .

PROOF: The proof is very similar to that of Smith's Lemma 2. $H^*(s)$ has a finite number of poles inside C_R and no other singularities in or on C_R . Let these poles be $s_1,\ldots,s_N$. Define

$$P(s) = \prod_{j=1}^{N} (s-s_j), \; R(s) = P(s)\, H^*(s) \; .$$

We suppose here that $H^*(s)$ is expressed in lowest terms as $R(s)/P(s)$, i.e. $P(s)$ and $R(s)$ have no common factors. Since the only singularities of $H^*(s)$ are poles, this means that

$$P(s) = 0 \Longrightarrow R(s) \neq 0 \; .$$

$P(s)$ and $R(s)$ are analytic inside and on C_R . By Lemma 2.8 and the remarks preceding it, condition (ii) implies that $|H^*(s)| < 1$ for all $s \in C_R$. Therefore, since $P(s) \neq 0$ on C_R, we have

$$|R(s)| < |P(s)| \quad \text{for all} \quad s \in C_R \; .$$

By Rouché's Theorem, $P(s)\{1-H^*(s)\} = P(s) - R(s)$ has exactly N zeros inside C_R . No zero of $P(s)\{1-H^*(s)\}$ can be a zero of $P(s)$ as well, since $P(s) = 0 \Longrightarrow R(s) \neq 0$. Thus $1 - H^*(s)$ has

N zeros inside C_R .

N.B. Note that we do not require $s_1,\ldots,s_N$ to be distinct. $\square$

In Theorem 2.1 we showed that Equation (2.6), viz.

$$1 - H^*(s) = \frac{s\Phi^+(s)}{(s-\lambda)\Phi^-(s)} ,$$

is valid for all s such that $-\mu < \mathrm{Re}(s) < \mu$. In the following lemma we extend the domain of validity of this equation.

LEMMA 2.11: <u>If $H^*(s)$ is meromorphic in $\{\mathrm{Re}(s) > 0\}$, then</u>

$$(s-\lambda)\Phi^-(s)[1-H^*(s)] = s\Phi^+(s) \tag{2.10}$$

<u>throughout $\{\mathrm{Re}(s) > -\mu\}$ except at isolated points in $\{\mathrm{Re}(s) > \mu\}$</u> .

PROOF: From (2.6) we know that (2.10) is valid for s in $\{-\mu < \mathrm{Re}(s) < \mu\}$. Now $s\Phi^+(s)$ is analytic in $\{\mathrm{Re}(s) > -\mu\}$ and

$h(s) \underset{\text{def}}{=} (s-\lambda)\Phi^-(s)[1-H^*(s)]$ is analytic in $\{-\mu < \mathrm{Re}(s) < \mu\}$.
Therefore $h(s)$ has an analytic continuation throughout $\{\mathrm{Re}(s) \geq \mu\}$. If $H^*(s)$ is meromorphic in $\{\mathrm{Re}(s) > 0\}$, then $1 - H^*(s)$ may be continued analytically throughout $\{\mathrm{Re}(s) \geq \mu\}$ except at a finite number of poles $s_1,\ldots,s_N$. It follows that $(s-\lambda)\Phi^-(s)$ has an analytic continuation defined by

$$(s-\lambda)\Phi^-(s) = \frac{s\Phi^+(s)}{1-H^*(s)}$$

throughout $\{\mathrm{Re}(s) \geq \mu\}$ except at $s_1,\ldots,s_N$ or at the zeros of $1-H^*(s)$, which are isolated. It follows that (2.10) is valid throughout $\{Re(s) > -\mu\}$ except at these points. □

COROLLARY: The zeros of $1 - H^*(s)$ in $\{Re(s) > 0\}$ are the poles of $(s-\lambda)\Phi^-(s)$ in $\{Re(s) > 0\}$; the poles of $H^*(s)$ $(s_1,\ldots,s_N)$ are the zeros of $(s-\lambda)\Phi^-(s)$ in $\{Re(s) > 0\}$. □

We may now prove the analogue of Theorem 2.2 with conditions on the inter-arrival distributions.

THEOREM 2.3: <u>If</u>

<u>(i) $A_1,\ldots,A_K \in T$</u>

<u>(ii) $A_j(0) < 1$ for each j,</u>

<u>and $s_1,\ldots,s_N$ are the poles of $H^*(s)$ in $\{Re(s) > 0\}$, then $1 - H^*(s)$ has $N-1$ zeros in $\{Re(s) > 0\}$, $r_1,\ldots,r_{N-1}$ say, and</u>

$$F^*(s) = \frac{(a-b)s \prod_{j=1}^{N-1} (s/r_j - 1)}{\{1-H^*(s)\} \prod_{j=1}^{N} (s/s_j - 1)} \quad \text{for all } s \text{ in } \{Re(s) \geq 0\} \tag{2.11}$$

<u>(the right-hand side of (2.11) being defined by continuity when terms in the denominator vanish).</u>

PROOF: Condition (i) implies that $H^*(s)$ is meromorphic in $\{Re(s) > 0\}$. Since $\varepsilon < \mu$ and $H^*(s)$ is analytic in S_μ, we see that $H^*(s)$ has a finite number of poles inside any C_R . We may choose R sufficiently large that C_R contains all the poles of $H^*(s)$ in $\{Re(s) > 0\}$.

As in the proof of Theorem 2.2, condition (ii) implies that $|H^*(s)| < 1$ for all s on the arc $\{|s| = R, \mathrm{Re}(s) > 0\}$ for R sufficiently large. Since the only zero of $1 - H^*(s)$ inside S_μ is a single zero at $s = 0$, it follows by Lemma 2.10 that $1 - H^*(s)$ has $N-1$ zeros in $\{\mathrm{Re}(s) > 0\}$. Denote these by $r_1,\ldots,r_{N-1}$. From (2.10) it follows that

$$\Phi^+(0) = \lim_{s\to 0+} (s-\lambda)\Phi^-(s)\frac{1-H^*(s)}{s}$$

$$= \lambda(a-b)\Phi^-(0) ,$$

and therefore for all but isolated points in $\{\mathrm{Re}(s) > 0\}$

$$F^*(s) = \frac{\lambda(a-b)\Phi^-(0)s}{\{1-H^*(s)\}(s-\lambda)\Phi^-(s)} . \tag{2.12}$$

From the corollary to Lemma 2.11, the function $\psi(s)$ defined for $s \notin \{s_j\} \cup \{r_j\}$ by

$$\psi(s) = (s-\lambda)\Phi^-(s)\frac{\prod_1^{N-1}(s-r_j)}{\prod_1^N(s-s_j)}$$

and by continuity elsewhere, is zero-free and entire in the finite plane. We show that $\psi(s)$ is bounded and therefore constant.

In $\{\mathrm{Re}(s) \le 0\}$, $(s-\lambda)\dfrac{\prod_1^{N-1}(s-r_j)}{\prod_1^N(s-s_j)}$ is bounded since $\mathrm{Re}(s_j) > 0$, $\mathrm{Re}(r_j) > 0$ for each j. Also $\Phi^-(s)$ is bounded in $\{\mathrm{Re}(s) \le \mu/2\}$, so $\psi(s)$ is bounded in the left half-plane. Choose R so that

$\lambda, s_1, \ldots, s_N, r_1, \ldots, r_{N-1}$ all lie inside C_R. Now

$(s-\lambda) \dfrac{\prod_1^{N-1} (s-r_j)}{\prod_1^{N} (s-s_j)}$ is bounded in $\{|s| \geq R, \mathrm{Re}(s) > 0\}$, and $\Phi^-(s)$ is bounded there also since by (2.10)

$$\Phi^-(s) = \frac{s\Phi^+(s)}{(s-\lambda)\{1-H^*(s)\}}$$

and $\Phi^+(s), \frac{s}{s-\lambda}$ and $\{1-H^*(s)\}^{-1}$ are all bounded there. Therefore $\psi(s)$ is bounded in $\{|s| \geq R, \mathrm{Re}(s) > 0\}$, i.e. outside C_R in the right half-plane. $\psi(s)$ is bounded inside C_R since analytic there and bounded on C_R. Thus $\psi(s)$ is bounded everywhere, and therefore is constant by Liouville's Theorem. Suppose $\psi(s) \equiv \rho \neq 0$. Then we have

$$(s-\lambda)\Phi^-(s) = \frac{\rho \prod_1^N (s-s_j)}{\prod_1^{N-1} (s-r_j)} \quad \text{for } s \notin \{r_j\},\ \mathrm{Re}(s) \geq 0 .$$

(2.11) now follows from equation (2.12). $\square$

COROLLARY: The stationary probability of an arriving customer finding the server free is

$$\frac{(a-b)\prod_1^N s_j}{\prod_1^{N-1} r_j} .$$

<u>Application to the forward transition model:</u>

Consider the forward transition model with customer types $0,1,\ldots,\kappa$, transition matrix $[p_{jk}]$ and parameters $\lambda_0, \lambda_1, \ldots, \lambda_\kappa$.

We recall that for this model a type j arrival corresponds to a transition $0 \to j$ in the guiding chain (Z_t); the time spent by (Z_t) in state j between transitions is exponentially distributed with parameter λ_j (assume $0 < \lambda_j < \infty$), $j = 0,1,\ldots,\kappa$. We shall demonstrate that the conditions of Theorem 2.3 are always satisfied for this model.

Condition (ii) is clearly satisfied since each type j arrival interval begins with an exponentially distributed period in state j of the guiding chain. [We may allow λ_o to be $+\infty$ here if $p_{oo} = 0$.] In the following result we show that condition (i) is always satisfied.

LEMMA 2.12: <u>The inter-arrival distributions for the forward transition model have rational Laplace transforms.</u>

PROOF: Let A_j be the d.f. of a type j arrival interval as usual ($j = 0,1,\ldots,\kappa$). The length of a type j arrival interval is the time when the guiding chain first leaves state 0 given $Z(0) = j$. Clearly $A_o^*(s) = \frac{\lambda_o}{\lambda_o+s}$. For $j = 1,\ldots,\kappa$ we have the relation

$$A_j^*(s) = \frac{\lambda_j}{\lambda_j+s} \sum_{k=0}^{\kappa} p_{jk}A_k^*(s) ,$$

since a period in state j is followed by a transition to some state k. We may write this as

$$\frac{\lambda_j+s}{\lambda_j} A_j^*(s) - \sum_{k=1}^{\kappa} p_{jk}A_k^*(s) = p_{jo} \frac{\lambda_o}{\lambda_o+s} , \quad (j = 1,2,\ldots,\kappa). \qquad (2.13)$$

Thus we have κ linear equations in $A_1^*(s),\ldots,A_\kappa^*(s)$. The determinant of this system of equations is a polynomial in s which is zero only at isolated points. By Cramer's rule, $A_j^*(s)$ is seen to be a ratio of polynomials in s. □

NOTE: The proof of Lemma 2.12 goes through when $\lambda_o = +\infty$.

We conclude this chapter with a numerical example.

Example: Consider the forward transition model with $\kappa=2, \lambda_o=4, \lambda_1=3, \lambda_2=3$ and

$$\mathbb{P} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}.$$

We suppose that service times are constant, with

$$b_o = \frac{1}{2}, \quad b_1 = \frac{1}{3}, \quad b_2 = \frac{1}{4} .$$

The $B_j^*(s) = e^{-sb_j}$, so that Theorem 2.2. does not apply. From (2.13) we have

$$A_o^*(s) = \frac{4}{4+s}, \quad A_1^*(s) = \frac{3.4}{(3+s)(4+s)}, \quad A_2^*(s) = \frac{4(6+s)}{(2+s)(3+s)(4+s)} .$$

Hence $a_o = \frac{1}{4}$, $a_1 = \frac{7}{12}$, $a_2 = \frac{11}{12}$ and

$$a = \Sigma p_{oj} a_j = \frac{7}{12}, \quad b = \Sigma p_{oj} b_j = \frac{17}{48} .$$

Now $b-a = -\frac{11}{48} < 0$, so (A1) is satisfied. (A2) and (A3) may be verified with $\mu' = \frac{3}{2}$, $\delta = 1$. $H^*(s) = \Sigma p_{oj} A_j^*(-s) B_j^*(s)$ has 3 poles in $\{\mathrm{Re}(s) > 0\}$ at $s = 2,3,4$. $1-H^*(s)$ has 2 zeros in $\{\mathrm{Re}(s) > 0\}$ which are computed by Newton's method to be 2.76, 4.60. We may now evaluate $F^*(s)$ by (2.11).

CHAPTER 3

STATIONARY BEHAVIOUR OF THE BACKWARD MODELS

3.1 *Introduction*

The queueing time process $\{Q_n\}$ is not Markovian for the backward model, so we cannot expect to obtain a simple integral equation like (2.3) for the stationary queueing time distribution. However, the total waiting time process is Markovian; in Section 3.2 we show that this process has a stationary distribution G under the usual first moment condition, and derive an integral equation for G.

In Section 3.3 we solve this equation in the case of exponential arrivals. As a by-product we see how to calculate some stationary probabilities that will be useful in dealing with busy periods in Chapter 4.

3.2 In this section we consider the general backward model, so-called because at each arrival epoch the customer type (and therefore service time) is related to the preceding arrival interval. Specifically, we assume that there are κ customer types: at each arrival epoch the type of the *next* customer is selected from $\{1,2,\ldots,\kappa\}$ according to the probabilities $\pi_1,\pi_2,\ldots,\pi_\kappa$, independently of any previous events. The service time of a type j customer is sampled from a distribution with d.f. B_j, and the length of the arrival interval preceding a type j customer has d.f. A_j $(j = 1,2,\ldots,\kappa)$.

Let Z_n be the type of the n^{th} customer, V_n his service time and U_{n-1} the time between the $(n-1)^{th}$ and n^{th} arrivals $(n = 1,2,\ldots)$ (U_o being the time from 0 to the first arrival). Let V_o be the residual service time at time 0. We leave unspecified the distribution of V_o. We may summarize our assumptions as follows: for $n = 1,2,\ldots$

$$\Pr\{Z_n = j\} = \pi_j \ ,$$

$$\Pr\{V_n \le x \mid Z_n = j\} = B_j(x) \ ,$$

$$\Pr\{U_{n-1} \le x \mid Z_n = j\} = A_j(x) \ .$$

Thus $V_1, V_2, \ldots,$ are iid random variables, and $U_o, U_1, \ldots$ are iid also. V_n and U_{n-1} are dependent through Z_n, but conditionally independent given Z_n. We suppose that the mean service and inter-arrival times are finite, i.e.

$$a_j = \int_0^\infty u\, A_j(du) < \infty \ , \quad b_j = \int_0^\infty u\, B_j(du) < \infty \ .$$

Let $a = \sum_{j=1}^{\kappa} \pi_j a_j$, $b = \sum_{j=1}^{\kappa} \pi_j b_j$. Then $E\{U_n\} = a$ for $n=0,1,2,\ldots,$ $E\{V_n\} = b$ for $n = 1,2,\ldots$. Let

$$X_n = V_n - U_{n-1} \quad (n=1,2,\ldots)$$

Then $X_1, X_2, \ldots$ are iid random variables, with $E\{X_n\} = b - a$.

The queueing time process $\{Q_n\}$ (as defined in Chapter 2) is not Markovian for the backward model. This is evident from a consideration of the relation (c.f. (2.1))

$$Q_{n+1} = \max (Q_n + V_n - U_n, 0) \; . \tag{3.1}$$

Since $Q_n = \max (Q_{n-1} + V_{n-1} - U_{n-1}, 0)$, we see that knowledge of both Q_{n-1} and Q_n yields information about U_{n-1} (and therefore V_n) that is not contained in knowledge of Q_n alone. As a result the conditional distribution of Q_{n+1} given Q_{n-1}, Q_n will be different from the distribution of Q_{n+1} conditional on Q_n alone.

However, a related process is Markovian. Let W_n be the total waiting time of the n^{th} customer, i.e. the time from his arrival to completion of service. Then, as remarked in [2 ; p. 1013], $\{W_n\}$ is a Markov process. This is clear from the recursive relation

$$W_{n+1} = \max (W_n - U_n + V_{n+1}, V_{n+1}) \quad (n = 0,1,2,\ldots) \tag{3.2}$$

which follows from (3.1) since $W_n = Q_n + V_n$. (Take $W_o = V_o$). Clearly W_{n+1} is determined by $\{V_o, U_o, V_1, U_1, \ldots, U_n, V_{n+1}\}$, and the Markovian property follows. We note that W_{n+1} is independent of U_{n+1} and Z_{n+1}.

Let G_n be the d.f. of W_n $(n = 0,1,2,\ldots)$. In the following result, we derive a recursive relation for $\{G_n\}$.

LEMMA 3.1: <u>For $x \geq 0$, $n = 0,1,2,\ldots$ we have</u>

$$G_{n+1}(x) = \sum_{j=1}^{K} \pi_j \int_0^x \int_0^\infty G_n(x-u+z) \, A_j(dz) \, B_j(du) \; . \tag{3.3}$$

PROOF:

$$\Pr\{W_{n+1} \le x\} = \sum_{j=1}^{\kappa} \pi_j \Pr\{W_{n+1} \le x \mid Z_{n+1} = j\}$$

and

$$\Pr\{W_{n+1} \le x \mid Z_{n+1} = j\} = \int_0^x \Pr\{W_{n+1} \le x \mid Z_{n+1} = j,\ V_{n+1} = u\}\ B_j(du).$$

If $V_{n+1} \le x$, then from relation (3.2) we have

$$W_{n+1} \le x \iff W_n - U_n + V_{n+1} \le x\ .$$

Hence it follows that

$$\Pr\{W_{n+1} \le x \mid Z_{n+1} = j,\ V_{n+1} = u\} = \Pr\{W_n - U_n \le x - u \mid Z_{n+1} = j\}$$

$$= \int_0^\infty G_n(x - u + z)\ A_j(dz)$$

since W_n and U_n are independent. Therefore (3.3) follows. □

We may show that $\{W_n\}$ has a stationary distribution if $EX_1 < 0$.

LEMMA 3.2: <u>If $EX_1 = b - a < 0$, then $G(x) \underset{\text{def}}{=} \lim_{n\to\infty} G_n(x)$ exists for all x and is a proper d.f.</u>

PROOF: Applying relation (3.2) repeatedly, we have

$$W_n = \max\left(V_o + X_1 + X_2 + \dots + X_n,\ V_1 + X_2 + \dots + X_n, \dots, V_{n-1} + X_n, V_n\right),$$

$$\Pr\{W_n \le x\} = \Pr\{V_n \le x, V_{n-1} + X_n \le x, \dots, V_1 + X_2 + \dots + X_n \le x, V_o + X_1 + \dots + X_n \le x\}\ .$$

Since $V_1,V_2,\ldots$ are iid and $U_o,U_1,\ldots$ are iid with U_{k-1},V_k dependent, we may replace V_j by V_{n+1-j} $(j = 1,2,\ldots,n)$ and U_{j-1} by U_{n-j} $(j = 1,2,\ldots,n)$ without changing the above probability. This means replacing X_j by X_{n+1-j} for $j = 1,2,\ldots,n$. Let $S_j = X_1 + \ldots + X_j$ $(j = 1,2,\ldots)$; $S_o = 0$. Then we have

$$G_n(x) = \Pr\{V_1 \le x, S_1+V_2 \le x, S_2+V_3 \le x, \ldots, S_{n-1}+V_n \le x, S_n+V_o \le x\}.$$

Let $H_n(x) = \Pr\{V_1 \le x, S_1+V_2 \le x, S_2+V_3 \le x, \ldots, S_{n-1}+V_n \le x\}$. Then $0 \le G_n(x) \le H_n(x) \le 1$ and $H_{n+1}(x) \le H_n(x)$ $(n = 1,2,\ldots)$. It follows that $\lim_{n\to\infty} H_n(x)$ exists: define $G(x) = \lim_{n\to\infty} H_n(x)$.

Now $0 \le H_n(x) - G_n(x) \le \Pr\{S_n + V_o > x\}$ and since S_n is independent of V_o we have

$$\Pr\{S_n + V_o > x\} = \int_0^\infty \Pr\{S_n > x-y\}\, G_o(dy) \ .$$

By the strong law of large numbers, $\Pr\{S_n > x-y\} \to 0$ as $n \to \infty$ for all x,y , so by the dominated convergence theorem, $\Pr\{S_n + V_o > x\} \to 0$ as $n \to \infty$. Therefore, $\lim_{n\to\infty} G_n(x)$ exists and equals $G(x)$.

We have $G(x) = 0$ for all $x < 0$, and $G(x)$ is clearly a non-decreasing function. Therefore, to show that $G(x)$ is a proper d.f. we need only show that $\lim_{x\to\infty} G(x) = 1$. This follows by the same methods as used in §4, (ii) of [11], since $S_n + V_{n+1} \to -\infty$ a.s. by the strong law of large numbers and

$$G(x) = \Pr\{\sup_{n\ge 0}(S_n + V_{n+1}) \le x\} \ . \qquad \square$$

We assume throughout the rest of this chapter that the stationarity condition, $b - a < 0$, is satisfied. It follows from (3.3) that G satisfies the integral equation

$$G(x) = \sum_{j=1}^{\kappa} \pi_j \int_0^x \int_0^\infty G(x-u+z)\ A_j(dz)\ B_j(du) \quad (x \geq 0). \tag{3.4}$$

This equation can be written in Wiener-Hopf form as

$$G(x) = \int_0^\infty H(x-u)\ G(du)\ ,$$

but the kernel function H does not bear such a simple relation to $\{A_j, B_j\}$ as does the kernel for the forward model. Therefore, the methods of Chapter 2 are not appropriate.

Let $K_j(x)$ be defined for each j as

$$\begin{aligned} K_j(x) &= \int_0^\infty G(x+z)\ A_j(dz) \quad (x \geq 0)\ , \\ &= 0 \qquad\qquad\qquad\qquad (x < 0)\ . \end{aligned}$$

Then (3.4) may be written as

$$G(x) = \sum_{j=1}^{\kappa} \pi_j \int_0^x K_j(x-u)\ B_j(du) \quad (x \geq 0)\ .$$

If we take Laplace transforms, then by Lemma A2 we have

$$G^o(s) = \sum_{j=1}^{\kappa} \pi_j\ K_j{}^o(s)\ B_j{}^*(s)\ , \tag{3.5}$$

wherever both sides are defined.

To proceed further with the solution of (3.5) we need to express $K_j{}^o(s)$ in terms of $G^o(s)$ and $A_j{}^*(s)$. We have been able to do this only in the case of exponential arrivals, i.e. when $A_j(x) = 1-e^{-\lambda_j x}$ $(j = 1,2,\ldots,\kappa)$. We examine this case in the following section.

3.3 In this section we continue the investigation of the stationary distribution of total waiting time for the particular case of the backward model with exponential arrivals. In addition to the assumptions of Section 3.2 we suppose that for each j $A_j(x)$ is of the form

$$A_j(x) = 1 - e^{-\lambda_j x} \quad (x \geq 0) \; .$$

Clearly $\lambda_j = {}^1/_{a_j}$, since the mean of an exponential distribution with parameter λ is ${}^1/_\lambda$.

This restriction leads us to a simple expression for $K_j^{\,o}(s)$ in terms of $G^o(s)$.

LEMMA 3.3: <u>For all s in $\{\mathrm{Re}(s) \geq 0\}$ such that $s \neq \lambda_j$</u>

$$K_j^{\,o}(s) = \frac{\lambda_j}{\lambda_j - s} \{G^o(s) - G^o(\lambda_j)\} \quad (j = 1,\ldots,\kappa) \; . \tag{3.6}$$

PROOF: We have

$$K_j^{\,o}(s) = \lambda_j \int_0^\infty \int_0^\infty e^{-sy-\lambda_j z} \, G(y+z) \, dz \, dy \; .$$

Put $y+z = u$, $z = t$. Then we have

$$\begin{aligned} K_j^{\,o}(s) &= \lambda_j \int_0^\infty \int_0^u e^{-s(u-t)-\lambda_j t} \, G(u) \, dt \, du \\ &= \lambda_j \int_0^\infty e^{-su} \, G(u) \left\{ \int_0^u e^{-(\lambda_j - s)t} \, dt \right\} du \\ &= \frac{\lambda_j}{\lambda_j - s} \int_0^\infty e^{-su} \left\{ 1 - e^{-(\lambda_j - s)u} \right\} G(u) \, du \\ &= \frac{\lambda_j}{\lambda_j - s} \{G^o(s) - G^o(\lambda_j)\} \; . \qquad \square \end{aligned}$$

From (3.5) and (3.6) we have

$$G^o(s) = \sum_{j=1}^{\kappa} \frac{\pi_j \lambda_j}{\lambda_j - s} \{G^o(s) - G^o(\lambda_j)\} B_j^*(s) \tag{3.7}$$

for all $s \notin \{\lambda_1, \lambda_2, \ldots, \lambda_\kappa\}$ in $\{\mathrm{Re}(s) \geq 0\}$.

Since $G^*(s) = s\, G^o(s)$ we have the following result:

THEOREM 3.1: <u>If $\sum_{j=1}^{\kappa} \pi_j \left(b_j - {}^1/_{\lambda_j} \right) < 0$, then the stationary distribution of total waiting time for the backward model with exponential arrivals has Laplace transform</u>

$$G^*(s) = \frac{s \sum_{j=1}^{\kappa} \frac{\pi_j}{\lambda_j - s} G^*(\lambda_j)\, B_j^*(s)}{\sum_{j=1}^{\kappa} \frac{\pi_j \lambda_j}{\lambda_j - s}\, B_j^*(s) - 1} \quad . \tag{3.8}$$

To make use of (3.8) to calculate $G^*(s)$, we need some way of determining the quantities $G^*(\lambda_j)$. We proceed as follows:

Suppose $\lambda_1, \ldots, \lambda_\kappa$ are distinct. (If not, then we could amalgamate the appropriate customer types as is done in Chapter 4 without affecting $G^*(s)$.) Consider the function $\phi(s)$ defined by

$$\phi(s) = \sum_{j=1}^{\kappa} \frac{\pi_j \lambda_j B_j^*(s)}{\lambda_j - s} \qquad (\mathrm{Re}(s) \geq 0,\ s \notin \{\lambda_1, \ldots, \lambda_\kappa\}) \ .$$

In Lemma 4.13 we show that the equation $\phi(s) = 1$ has exactly $\kappa-1$ (real) positive roots $\mu_2, \ldots, \mu_\kappa$. Since $G^*(s)$ is analytic throughout $\{\mathrm{Re}(s) > 0\}$, it follows that

$$\sum_{j=1}^{\kappa} \frac{\pi_j}{\lambda_j - s} \, G^*(\lambda_j) B_j^*(s) = 0 \quad \text{for} \quad s = \mu_2, \mu_3, \ldots, \mu_\kappa \, .$$

Let $\tau_j = \pi_j \, G^*(\lambda_j)$. Then we have $\kappa - 1$ linear equations in $\tau_1, \tau_2, \ldots, \tau_\kappa$. We may derive one more equation by considering (3.8) as $s \to 0+$. Since $G^*(0) = 1$, we have

$$1 = \lim_{s \to 0+} \frac{\sum_j \frac{\tau_j}{\lambda_j - s} B_j^*(s)}{\frac{1}{s}\left\{\sum_j \frac{\pi_j \lambda_j}{\lambda_j - s} B_j^*(s) - 1\right\}}$$

$$= \frac{\sum_j {}^{\tau_j}/_{\lambda_j}}{\sum_j \pi_j \lim_{s \to 0+} \left\{(1 - {}^{s}/_{\lambda_j})^{-1} B_j^*(s) - 1\right\}}$$

$$= \frac{\sum {}^{\tau_j}/\lambda_j}{\sum \pi_j ({}^{1}/\lambda_j - b_j)}$$

$$= \frac{\sum {}^{\tau_j}/_{\lambda_j}}{a - b} \, .$$

Thus we have κ linear equations in $\tau_1, \tau_2, \ldots, \tau_\kappa$, viz.

$$\sum_{j=1}^{\kappa} {}^{\tau_j}/_{\lambda_j} = a - b \, , \tag{3.9}$$

$$\sum_{j=1}^{\kappa} \frac{\tau_j}{\lambda_j - \mu_k} B_j^*(\mu_k) = 0 \quad (k = 2,3,\ldots,\kappa) \, . \tag{3.10}$$

To summarize our procedure, we solve $\phi(s) = 1$ for $\mu_2, \ldots, \mu_\kappa$ as in Section 4.3, then calculate $\tau_1, \ldots, \tau_\kappa$ from (3.9), (3.10). Hence we can evaluate $G^*(s)$ as

$$G^*(s) = \frac{s \sum_j \frac{\tau_j B_j^*(s)}{\lambda_j - s}}{\sum_j \frac{\pi_j \lambda_j}{\lambda_j - s} B_j^*(s) - 1} \qquad \text{for} \quad s \notin \{\lambda_j\}, \mathrm{Re}(s) \geq 0 \ .$$

We may interpret τ_j in terms of stationary probabilities.

LEMMA 3.4: <u>τ_j is the stationary probability that an arriving customer be of type j and have zero queueing time.</u>

PROOF: $G^*(\lambda_j) = \lambda_j G^o(\lambda_j)$

$$= \int_0^\infty \lambda_j \, e^{-\lambda_j x} \, G(x) \, dx$$

$$= \int_0^\infty G(x) \, A_j(dx) \ .$$

A customer has zero queueing time iff his total waiting time is equal to his service time. From (3.2) this occurs iff the total waiting time of the previous customer W is no greater than the intervening arrival time U. Since W_n is independent of U_n for any n, it follows that the stationary probability of a type j customer having zero queueing time is

$$\Pr\{W - U \leq 0\} = \int_0^\infty G(x) \, A_j(dx) = G^*(\lambda_j) \ .$$

Since π_j is the probability that an arriving customer be of type j, the result follows. □

COROLLARY: The stationary probability that the server be free is $\sum_{j=1}^{K} \tau_j$. □

Let $\gamma_j = \frac{\tau_j}{\Sigma\tau_j}$. Then γ_j is clearly the stationary probability that a customer with zero queueing time be of type j. As the arrival of a customer with zero queueing time is equivalent to the beginning of a busy period (see Chapter 4), γ_j may be regarded as the stationary probability that the initial customer of a busy period be of type j.

We may show that the stationary probability that a busy period start with a type j customer is in general different from the probability π_j that *any* arriving customer be of type j.

LEMMA 3.5: It is not possible that $\gamma_j = \pi_j$ for two distinct values of j .

PROOF: Since $\tau_j = \pi_j G^*(\lambda_j)$ and $\gamma_j = \frac{\tau_j}{\Sigma\tau_j}$, $\gamma_j = \pi_j \Rightarrow G^*(\lambda_j) = \sum_j \tau_j$. Therefore, if $\gamma_1 = \pi_1$ and $\gamma_2 = \pi_2$ where $\lambda_1 < \lambda_2$, we have $G^*(\lambda_1) = G^*(\lambda_2)$. But

$$G^*(\lambda_1) - G^*(\lambda_2) = \int_0^\infty \left\{e^{-\lambda_1 x} - e^{-\lambda_2 x}\right\} G(dx)$$

and $\exp\{-\lambda_1 x\} \geq \exp\{-\lambda_2 x\}$ with equality only at $x = 0$. Therefore $G^*(\lambda_1) = G^*(\lambda_2)$ iff G puts all probability at 0, i.e. $G(0) = 1$. This is impossible with exponential arrivals, so the result follows. □

NOTE: For the general *forward* model the stationary probability that a customer with zero queueing time be of type j is π_j, i.e. the same as the unconditional probability that a customer be of type j. That is because Q_n is independent of Z_n for this model. It can be shown that the stationary probability that the server be free is simply $1 - {}^{b}/_{a}$. and the mean length of a general busy period is $\frac{b}{1-b/a}$ for the general forward model (c.f. Section 4.3) .

From (3.9) we have the following expression for the stationary probability that the server be free:

$$\sum_j \tau_j = \frac{a-b}{\sum_j \gamma_j/\lambda_j} \quad . \tag{3.11}$$

CHAPTER 4

BUSY PERIODS FOR THE BACKWARD MODELS

4.1 A busy period begins when a customer arrives to find the server free, and ends when the server becomes free again. In this chapter we work in terms of busy periods generated by a fixed initial service time; in particular, we denote by $Y(w)$ the length of a busy period generated by an initial service time w. If $\phi(w,s)$ is the Laplace transform $E\{e^{-sY(w)}\}$ and Y_j denotes the length of a busy period begun by a type j customer, then

$$E\left\{e^{-sY_j}\right\} = \int_0^\infty \phi(w,s)\; B_j(dw) \; .$$

In Section 4.2 we consider busy periods for the backward model with exponential arrivals, generalizing a procedure applied by Cox and Smith in [5 ; §5.6] to the random arrivals case. We show that the busy period transform $\phi(w,s)$ is a mixture of exponential terms, i.e.

$$\phi(w,s) = \sum_{j=1}^{K} H_j(s)\; e^{-w\mu_j(s)} \; ,$$

and derive equations enabling us to solve for $H_j(s)$, $\mu_j(s)$ for all j. In Section 4.3 we show how to compute the means $m_j = EY_j$, $m(w) = EY(w)$. In Section 4.4 we extend some of the results in Section 4.2 to the more general case of the backward transition model with $\lambda_o = \infty$.

Similar results could be obtained for $\phi(w,s)$ for the corresponding forward models, but the simple expression for

$$E\left\{e^{-sY_j}\right\}$$

above is no longer valid, since the initial service time and the time to the second arrival are now dependent. Therefore, the procedures of this chapter are of little value for the forward models.

4.2 *The backward model with exponential arrivals*

In this section we consider the general backward model with κ customer types, general service times and exponential inter-arrival times. At each arrival epoch the type of the next customer is selected from $\{1,2,\ldots,\kappa\}$ according to the (non-zero) probabilities $\pi_1,\pi_2,\ldots,\pi_\kappa$ and independently of any previous events. If the next customer is chosen to be of type j, then the time to his arrival is sampled from an exponential distribution with parameter λ_j, and his service time has d.f. B_j. If $\lambda_j = \lambda_k$ for some $j \neq k$, then we could amalgamate the two customer types without affecting the busy period distribution by taking

$$\frac{\pi_j B_j + \pi_k B_k}{\pi_j + \pi_k}$$

as the new service time d.f. Therefore, we may assume that $\lambda_1,\lambda_2,\ldots,\lambda_\kappa$ are distinct; suppose $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_\kappa$. We assume also that the mean service times

$$b_j \underset{\text{def}}{=} \int_0^\infty u\, B_j(du)$$

are finite.

To consider the busy periods for this model, we suppose that at time 0 a customer with service time w has arrived to find the server free. Let $W_t(w)$ be the virtual waiting time at time $t \geq 0$ under this initial assumption. Define $Y(w) = \inf \{t : W_t(w) = 0\}$, i.e. $Y(w)$ is the length of the busy period induced by the arrival of a customer with service time w. $Y(w)$ is measurable since it is a hitting time for the process $(W_t, t \geq 0)$, [15; p. 72], and $Y(w)$ is finite a.s. provided the stationarity condition holds, i.e.

$$\sum_{j=1}^{K} \pi_j (b_j - 1/\lambda_j) < 0 ,$$

since then W_t reaches 0 eventually with probability 1 (see Chapter 3).

At any instant $t > 0$, there will be a customer in the process of arriving, of the type selected at the previous arrival epoch. We define Z_t to be this customer type and make $(Z_t, t \geq 0)$ right-continuous, so that if t is an arrival epoch, then Z_t is the type of the next customer to arrive. Then (Z_t) is a stationary Markov process, with $\Pr\{Z_t = j\} = \pi_j$. We will refer to Z_t as the arrival state at time t.

Consider realizations of the process $(W_t(w), t \geq 0)$. (In the rest of this paragraph we suppress w). (W_t) has discrete increments at each arrival epoch equal to the service time of the arriving customer. Elsewhere (W_t) decreases continuously at a unit rate until it reaches zero, where it remains until the next arrival epoch. (W_t) is not a Markov process since $\{W_s : 0 \leq s \leq t\}$ contains information about the time of the last arrival before t that is not given by knowledge of W_t alone, and this information affects the distribution of the time to the next arrival after t, except when the (overall) inter-arrival distribution is exponential. [The unconditional inter-arrival distribution has d.f.

$$\sum_{j=1}^{\kappa} \pi_j \left(1 - e^{-\lambda_j x}\right).\Big]$$

Suppose (W_s, Z_s) is given for $0 \leq s \leq t$. The distribution of inter-arrival time conditional on the arrival type is exponential, so the time to the next arrival after t is now independent of the time since the last arrival before t, as Z_t is given. Therefore, the future behaviour of (W_u, Z_u) for $u > t$ depends on $\{(W_s, Z_s), 0 \leq s \leq t\}$ only through (W_t, Z_t). This shows that $((W_t, Z_t), t \geq 0)$ is a Markov process. It is clear from the discussion at the beginning of this paragraph that the behaviour of (W_u, Z_u) for $u > t$ given (W_t, Z_t) is independent of t, i.e. $((W_t, Z_t))$ is homogeneous [1].

Define $\phi(w,s) = E\{e^{-sY(w)}\}$ for $\mathrm{Re}(s) \geq 0, w \geq 0$; i.e. $\phi(w,s)$ is the Laplace-Stieltjes transform of the d.f. of $Y(w)$. In the classical random arrivals case, which is what our present model reduces to when $\kappa = 1$, it is possible to show that

$$\phi(w_1 + w_2, s) = \phi(w_1, s)\, \phi(w_2, s) \quad \text{for any } w_1, w_2 > 0,$$

and consequently that $\phi(w,s)$ is necessarily of the form $\phi(w,s) = e^{-w\theta(s)}$ [5 ; p. 145]. In generalizing this procedure to our present model we must split $\phi(w,s)$ into components corresponding to the conditional transforms of the busy period distribution given the initial and final arrival states.

Let

$$\phi_{ij}(w,s) = E\left\{e^{-sY(w)} \middle| Z_o = i, Z_{Y(w)} = j\right\} \cdot \Pr\left\{Z_o = i, Z_{Y(w)} = j\right\} .$$

Then

$$\phi(w,s) = \sum_{i=1}^{\kappa} \sum_{j=1}^{\kappa} \phi_{ij}(w,s) .$$

In the following lemma we derive a relation satisfied by the components $\phi_{ij}(w,s)$ and corresponding to the relation

$$\phi(w_1 + w_2, s) = \phi(w_1, s)\ \phi(w_2, s)$$

for the random arrivals case.

LEMMA 4.1 <u>For any $w_1, w_2 > 0$ and $i,j \in \{1,2,\ldots,\kappa\}$, $\mathrm{Re}(s) \geq 0$</u> ,

$$\phi_{ij}(w_1 + w_2, s) = \sum_{\ell=1}^{\kappa} \frac{1}{\pi_\ell}\, \phi_{i\ell}(w_1, s)\ \phi_{\ell j}(w_2, s) . \tag{4.1}$$

PROOF: We may split the busy period generated by an initial service time $w_1 + w_2$ into two parts, the first part being the time required to expend amount w_1 of the virtual waiting time, the second being the time required to expend the remaining amount w_2. Let

$$Y_1 = \inf\{t : W_t(w_1 + w_2) = w_2\},\ Y_2 = \inf\{t : W_{t+Y_1}(w_1 + w_2) = 0\} .$$

Then

$$Y(w_1 + w_2) = Y_1 + Y_2 .$$

Let E_{ij} be the event

$$\left[Z_o = i, Z_{Y_1+Y_2} = j\right] ,$$

and $E_{ij\ell}$ the event

$$\left[Z_o = i, Z_{Y_1} = \ell, Z_{Y_1+Y_2} = j\right] .$$

Then

$$E_{ij} = \sum_{\ell=1}^{K} E_{ij\ell}$$

and so we have

$$\phi_{ij}(w_1 + w_2, s) = E\left\{e^{-s(Y_1 + Y_2)} | E_{ij}\right\} \cdot \Pr(E_{ij})$$

$$= \int_{E_{ij}} e^{-s(Y_1 + Y_2)} dP$$

$$= \sum_{\ell=1}^{K} \int_{E_{ij\ell}} e^{-s(Y_1 + Y_2)} dP$$

$$= \sum_{\ell=1}^{K} E\left\{e^{-s(Y_1 + Y_2)} | E_{ij\ell}\right\} \cdot \Pr(E_{ij\ell})$$

(where P is the probability measure of the underlying probability space).

If $E_{ij\ell}$ has occurred, then $W_{Y_1} = w_2$ and $Z_{Y_1} = \ell$. Because of the homogeneity of the (W_t, Z_t) process, the conditional distribution of Y_2 given $Y_1, E_{ij\ell}$ is the same as the distribution of the length of the busy period induced by an initial service time w_2 conditional on the initial arrival state being ℓ and the final arrival state j, i.e.

$$\Pr\left\{Y_2 \le z \mid Y_1 = y, E_{ij\ell}\right\} = \Pr\left\{Y(w_2) \le z \mid Z_o = \ell, Z_{Y(w_2)} = j\right\} .$$

(We note that the conditional probability

$$\Pr\left\{Y_2 \le z \mid Y_1 = y, E_{ij\ell}\right\}$$

can be defined for almost all y as

$$\lim_{\delta \to 0} \Pr\left\{Y_2 \le z \mid y \le Y_1 < y + \delta, E_{ij\ell}\right\}$$

[17; p. 260].) Since

$$\Pr\left\{Y_2 \le z \mid Y_1 = y, E_{ij\ell}\right\}$$

is therefore independent of y, it follows that Y_1 and Y_2 are conditionally independent given $E_{ij\ell}$, for if

$$G_{ij\ell}(y) = \Pr\{Y_1 \le y \mid E_{ij\ell}\} ,$$

then we have

$$\begin{aligned}
\Pr\{Y_1 \le y_1, Y_2 \le y_2 \mid E_{ij\ell}\} &= \int_0^{y_1} \Pr\{Y_2 \le y_2 \mid E_{ij\ell}\}\, G_{ij\ell}(dy) \\
&= \Pr\{Y_2 \le y_2 \mid E_{ij\ell}\} \int_0^{y_1} G_{ij\ell}(dy) \\
&= \Pr\{Y_2 \le y_2 \mid E_{ij\ell}\} \Pr\{Y_1 \le y_1 \mid E_{ij\ell}\} .
\end{aligned}$$

Because of this conditional independence of Y_1, Y_2 we have

$$E\left\{e^{-s(Y_1 + Y_2)} \middle| E_{ij\ell}\right\} = E\left\{e^{-sY_1} \middle| E_{ij\ell}\right\} \cdot E\left\{e^{-sY_2} \middle| E_{ij\ell}\right\} .$$

We defined Y_1 as $\inf \{t : W_t(w_1 + w_2) = w_2\}$; for every realization of the process $(W_t(w_1 + w_2))$ there is a realization of $(W_t(w_1))$ such that $W_t(w_1 + w_2) = W_t(w_1) + w_2$ for all t such that $0 \le t \le \inf \{u : W_u(w_1) = 0\}$. Thus the conditional distribution of Y_1 given $E_{ij\ell}$ is the same as the conditional distribution of $Y(w_1)$ given $Z_o = i, Z_{Y(w_1)} = \ell$. Hence

$$E\left\{e^{-sY_1} \middle| E_{ij\ell}\right\} = \frac{\phi_{i\ell}(w_1, s)}{\Pr\left\{Z_o = i, Z_{Y(w_1)} = \ell\right\}} .$$

From our discussion of the conditional independence of Y_1 and Y_2 given $E_{ij\ell}$, it follows that the conditional distribution of Y_2 given $E_{ij\ell}$ is the same as the conditional distribution of $Y(w_2)$ given $Z_o = \ell, Z_{Y(w_2)} = j$. Hence

$$E\left\{e^{-sY_2} \middle| E_{ij\ell}\right\} = \frac{\phi_{\ell j}(w_2, s)}{\Pr\left\{Z_o = \ell, Z_{Y(w_2)} = j\right\}} .$$

Now, since (Z_t) is stationary, we have

$$\begin{aligned}\Pr(E_{ij\ell}) &= \Pr\left\{Z_0 = i, Z_{Y_1} = \ell, Z_{Y_1+Y_2} = j\right\}\\ &= \Pr\left\{Z_0 = i, Z_{Y_1} = \ell\right\}\cdot \Pr\left\{Z_{Y_1+Y_2} = j \mid Z_0 = i, Z_{Y_1} = \ell\right\}\\ &= \Pr\left\{Z_0 = i, Z_{Y_1} = \ell\right\}\cdot \Pr\left\{Z_{Y_2} = j \mid Z_0 = \ell\right\}\\ &= \Pr\left\{Z_0 = i, Z_{Y(w_1)} = \ell\right\}\cdot \Pr\left\{Z_{Y(w_2)} = j \mid Z_0 = \ell\right\}.\end{aligned}$$

Therefore, we have shown that

$$\phi_{ij}(w_1 + w_2, s) = \sum_{\ell=1}^{\kappa} \phi_{i\ell}(w_1, s)\ \phi_{\ell j}(w_2, s)\cdot \frac{\Pr\left\{Z_{Y(w_2)} = j \mid Z_0 = \ell\right\}}{\Pr\left\{Z_0 = \ell, Z_{Y(w_2)} = j\right\}}$$

$$= \sum_{\ell=1}^{\kappa} \phi_{i\ell}(w_1, s)\ \phi_{\ell j}(w_2, s)\circ \frac{1}{\Pr\{Z_0 = \ell\}}$$

which gives us Equation (4.1), since $\Pr\{Z_0 = \ell\} = \pi_\ell$. □

Let $\Phi(w,s)$ be the $\kappa \times \kappa$ matrix with (i,j)-element $\phi_{ij}(w,s)$. Then (4.1) is equivalent to the matrix relation

$$\Phi(w_1 + w_2, s) = \Phi(w_1, s)\ \Lambda^{-1}\ \Phi(w_2, s)\ , \tag{4.2}$$

where $\Lambda = \text{diag}\ (\pi_1, \pi_2, \ldots, \pi_\kappa)$. From (4.2) we may derive a partial differential (matrix) equation for $\Phi(w,s)$. Let

$$\phi_{.j}(w,s) = \sum_i \phi_{ij}(w,s),\quad \omega_{ij}(s) = \int_0^\infty \phi_{.j}(u,s)\ B_i(du)\ .$$

Then

$$\omega_{ij}(s) = E\left\{e^{-sY_i} \mid Z_{Y_i} = j\right\} \cdot \Pr\left\{Z_{Y_i} = j\right\}, \qquad \text{i.e. } \omega_{ij}(s)$$

is a multiple of the Laplace transform of the length of the busy period initiated by a type i customer, conditional on the busy period ending in arrival state j. Let $\Psi(s)$ be the $\kappa \times \kappa$ matrix with (i,j)-element $\psi_{ij}(s) = (\lambda_i + s)\,\delta_{ij} - \lambda_i\omega_{ij}(s)$, δ_{ij} being the Kronecker delta: $\delta_{ij} = 1$ if $i = j$, $= 0$ otherwise.

THEOREM 4.1 $\Phi(w,s)$ satisfies the matrix differential equation

$$\frac{\partial}{\partial w}\Phi(w,s) = -\Phi(w,s)\,\Psi(s) \quad \underline{\text{for}} \quad w > 0,\ \mathrm{Re}(s) \geq 0 . \tag{4.3}$$

PROOF: Suppose $w,h > 0$. Then from (4.2) we have

$$h^{-1}\{\Phi(w + h,s) - \Phi(w,s)\} = \Phi(w,s)\,h^{-1}\{\Lambda^{-1}\Phi(h,s) - I_\kappa\} .$$

We consider the behaviour of $\Phi(h,s)$ as $h \to 0+$. Let $N(t)$ be the number of arrivals in $(0,t]$. Since the inter-arrival distributions are exponential, we have $\Pr\{N(h) = 0 \mid Z_0 = i\} = 1 - \lambda_i h + o(h)$ as $h \to 0+$, $\Pr\{N(h) = 1 \mid Z_0 = i\} = \lambda_i h + o(h)$ and $\Pr\{N(h) > 1 \mid Z_0 = i\} = o(h)$. If $N(h) = 0$, then $Y(h) = h$ and $Z_{Y(h)} = Z_0$; if $N(h) = 1$, then at time h the virtual waiting time is equal to the service time of the newly-arrived customer, who is necessarily of type $Z_0 = i$. Therefore, we have

$$\phi_{ij}(h,s) = \sum_{m=0}^{1} E\left\{e^{-sY(h)} \mid Z_o=i, Z_{Y(h)}=j, N(h)=m\right\}. \Pr\left\{Z_o=i, Z_{Y(h)}=j, N(h)=m\right\}$$

$$+ E\left\{e^{-sY(h)} \mid Z_o=i, Z_{Y(h)}=j, N(h)>1\right\} \Pr\left\{Z_o=i, Z_{Y(h)}=j, N(h)>1\right\}$$

$$= \pi_i(1 - \lambda_i h)e^{-sh}\delta_{ij} + \pi_i\lambda_i h e^{-sh}\int_0^\infty \phi_{.j}(u,s)B_i(du) + o(h)$$

$$= \pi_i\{1 - (\lambda_i+s)h\}\delta_{ij} + \pi_i\lambda_i h\, \omega_{ij}(s) + o(h)$$

Hence

$$\lim_{h\to 0+} h^{-1}\left\{\pi_i^{-1}\, \phi_{ij}(h,s) - \delta_{ij}\right\} = -(\lambda_i+s)\delta_{ij} + \lambda_i\omega_{ij}(s) = -\psi_{ij}(s)$$

so

$$\lim_{h\to 0+} h^{-1}\{\Lambda^{-1}\Phi(h,s) - I_\kappa\} = -\Psi(s) .$$

It follows that $\Phi(w,s)$ is differentiable with respect to w and satisfies (4.3). □

In solving Equation (4.3) we take as boundary condition the equation $\Phi(0,s) = \Lambda$, since

$$\phi_{ij}(0,s) = \Pr\{Z_o = i, Z_{Y(0)} = j\} = \pi_i\delta_{ij} .$$

For fixed s, the j^{th} column of Φ' satisfies the differential equation

$$\frac{d\underset{\sim}{x}}{dw} = -\Psi'\underset{\sim}{x} ,$$

with boundary condition $\underset{\sim}{x}(0) = \pi_j\underset{\sim}{e}_j$, where $\underset{\sim}{e}_j$ is the j^{th} column of I_κ. The unique solution of this equation satisfying the boundary condition is $\underset{\sim}{x} = e^{-w\Psi'}(\pi_j \underset{\sim}{e}_j)$ [10; p. 190]. It follows that the unique solution of (4.3) satisfying $\Phi(0,s) = \Lambda$ is

$$\Phi(w,s) = \Lambda\left\{e^{-w\Psi'}\right\}' = \Lambda e^{-w\Psi(s)} .$$

LEMMA 4.2 <u>If A is an $n \times n$ matrix with n distinct eigenvalues $\mu_1,\mu_2,\ldots,\mu_n$, then</u>

$$e^{tA} = \sum_{j=1}^{n} e^{t\mu_j}\,\ell_j(A) ,$$

<u>where ℓ_j is a fundamental polynomial of Lagrange type with degree $n - 1$ such that $\ell_j(\mu_i) = \delta_{ij}$</u> .

PROOF: A is simple [10; p. 61], so the lemma follows from results in [10; p. 169]. □

Suppose $\Psi(s)$ has κ distinct eigenvalues $\mu_1(s),\ldots,\mu_\kappa(s)$ at some point s in $\{\mathrm{Re}(s) \geq 0\}$. Then by Lemma 4.2

$$e^{-w\Psi(s)} = \sum_{j=1}^{\kappa} e^{-w\mu_j(s)}\,\ell_j(\Psi(s),s) ,$$

where $\ell_j(u,s)$ is a polynomial in u of degree $\kappa - 1$ such that $\ell_j(\mu_i(s),s) = \delta_{ij}$. Specifically, we have

$$\ell_j(u,s) = \prod_{k\neq j}\left\{\frac{u - \mu_k(s)}{\mu_j(s) - \mu_k(s)}\right\} .$$

It follows then that $\phi(w,s)$, which is the sum of the elements of the matrix $\Phi(w,s)$, can be written as

$$\phi(w,s) = \sum_{j=1}^{\kappa} e^{-w\mu_j(s)} H_j(s) ,$$

where $H_j(s)$ is the sum of the elements of the matrix $\Lambda\ell_j(\Psi(s),s)$. Let $S = \{s : \mathrm{Re}(s) \geq 0$ and $\Psi(s)$ has κ distinct eigenvalues$\}$. Then we have proved

THEOREM 4.2 <u>For all $s \in S$ and all $w \geq 0$</u>

$$\phi(w,s) = \sum_{j=1}^{\kappa} H_j(s)\, e^{-w\mu_j(s)} . \tag{4.4}$$

Let $S^+ = \{s : \mathrm{Re}(s) > 0\} \cap S$. We shall prove that Equation (4.4) is valid throughout S^+ except at isolated points. First we note some properties of $\mu_j(s), H_j(s)$.

LEMMA 4.3: (i) <u>We may number the eigenvalues of $\Psi(s)$ as s varies so that $\mu_j(s)$ is a continuous function of s in $\{\mathrm{Re}(s) \geq 0\}$.</u>

(ii) <u>$\mu_j(s)$ is analytic throughout S^+.</u>

(iii) <u>$H_j(s)$ is analytic throughout S^+.</u>

PROOF: Since each $\omega_{ij}(s)$ is a multiple of a Laplace transform, the elements of $\Psi(s)$ are continuous functions of s in $\{\mathrm{Re}(s) \geq 0\}$. The eigenvalues of $\Psi(s)$ are continuous functions of the elements of $\Psi(s)$ [10; p. 225], so (i) is clear.

Since $\omega_{ij}(s)$ is analytic throughout $\{\mathrm{Re}(s) > 0\}$, each $\psi_{ij}(s)$ is analytic there. It follows from the theory of analytic functions that each $\mu_j(s)$ is analytic at every point s in $\{\mathrm{Re}(s) > 0\}$ such that $\Psi(s)$ has non-repeated eigenvalues [10; p. 237]. This proves (ii); (iii) follows directly from (ii) and the definition of $H_j(s)$. We note that $H_j(s)$ may not be defined at points where $\Psi(s)$ has repeated eigenvalues. □

We show now that points where $\Psi(s)$ has repeated eigenvalues correspond to the zeros of an analytic function, and are therefore isolated. It is clear that $\Psi(s)$ has repeated eigenvalues iff the discriminant of the equation $|\mu I_K - \Psi(s)| = 0$ is zero, i.e. iff $0 = D(s) \underset{\text{def}}{=} \prod_{i<j} \{\mu_i(s) - \mu_j(s)\}^2$.

LEMMA 4.4: <u>$D(s)$ is analytic throughout $\{Re(s) > 0\}$.</u>

PROOF: From (ii) of Lemma 4.3, $D(s)$ is analytic throughout S^+. Since repeated eigenvalues may fail to be analytic, we need some results from [6; pp. 233-7] to the effect that symmetric polynomials in multiple roots are analytic, in order to show that $D(s)$ is analytic everywhere in $\{Re(s) > 0\}$.

We summarize the results we need from [6]: Let $F(x,y)$ be an analytic function of x and y in a neighbourhood of a point (α,β) such that the equation $F(\alpha,y) = 0$ has $y = \beta$ as a root of multiplicity $n > 1$. Then there exist positive real numbers ρ and r such that for all x with $|x - \alpha| \leq r$ there are exactly n roots $y_1(x),\ldots,y_n(x)$ of $F(x,y) = 0$ in $\{y : |y - \beta| \leq \rho\}$. Every expression of the form

$$\sum_{j=1}^{n} \{y_j(x)\}^k$$

(with $k = 1,2,\ldots$) is analytic in x throughout $\{x : |x - \alpha| \leq r\}$.

Applying these results to our situation, let

$$J_k = \{j : \mu_j(s) = \mu_k(s)\}.$$

If $\mu_k(s)$ is a repeated eigenvalue of $\Psi(s)$, then $\{\mu_j(s) : j \in J_k\}$ are multiple roots of the equation $|yI_\kappa - \Psi(s)| = 0$. Since

$$F(s,y) \underset{\text{def}}{=} |yI_\kappa - \Psi(s)|$$

is analytic in y and s in a neighbourhood of $(s,\mu_k(s))$ for any s in $\{\text{Re}(s) > 0\}$, it follows that

$$\sum_{j\in J_k} \{\mu_j(s)\}^\ell$$

is analytic in some neighbourhood of s for any $\ell = 1,2,\ldots$.
It follows that

$$\sum_{j=1}^{\kappa} \{\mu_j(s)\}^\ell$$

is analytic throughout $\{\text{Re}(s) > 0\}$ for any $\ell = 1,2,\ldots$. As noted in [6 ; p. 235], this is sufficient to prove that all symmetric polynomials in $\mu_1(s),\ldots,\mu_\kappa(s)$ are analytic in $\{\text{Re}(s) > 0\}$; in particular, $D(s)$ is analytic throughout $\{\text{Re}(s) > 0\}$. $\square$

It will follow that $D(s)$ has isolated zeros if $D(s)$ is not identically zero. In the following lemma we show that $\Psi(s)$ has κ distinct eigenvalues for all s with $\text{Re}(s)$ sufficiently large, so that $D(s)$ cannot be identically zero.

LEMMA 4.5: <u>If $B_j(0) = 0$ for each j, then $\Psi(s)$ has κ distinct eigenvalues when $\mathrm{Re}(s)$ is sufficiently large.</u>

PROOF: We restrict our proof to the case where s is real; the proof in the general case is similar. We show first that $\omega_{ij}(s) = o(1)$ as $s \to \infty$. We have

$$|\omega_{ij}(s)| = \left|\int_0^\infty \phi_{.j}(u,s)\ B_i(du)\right|$$

$$\leq \int_0^\infty |\phi_{.j}(u,s)|\ B_i(du)$$

$$\leq \int_0^\infty e^{-su}\ B_i(du) \qquad \text{since } Y(u) \geq u$$

$$= B_i^*(s) \quad .$$

Since

$$\lim_{s\to\infty} B_i^*(s) = B_i(0)$$

it follows from the assumption $B_i(0) = 0$ for each i that

$$\lim_{s\to\infty} \omega_{ij}(s) = 0 \ ,$$

i.e. $\omega_{ij}(s) = o(1)$.

By Geršgorin's Theorem [10; p. 226], each eigenvalue of $\Psi(s)$ lies in at least one of the discs

$$D_j(s) = \left\{z : |z - (\lambda_j + s) + \lambda_j\omega_{jj}(s)| \leq \sum_{k\neq j} \lambda_j |\omega_{jk}(s)|\right\}$$

(and in exactly one of these discs if $D_1,\ldots,D_\kappa$ are disjoint).

Now $\omega_{ij}(s) = o(1)$ implies $\lambda_j + s - \lambda_j\omega_{jj}(s) = \lambda_j + s + o(1)$ and

$$\sum_{k\neq j} \lambda_j |\omega_{jk}(s)| = o(1)$$

as $s \to \infty$, so that $D_1(s),\ldots,D_K(s)$ will be disjoint for s sufficiently large. It follows that $\mu_1(s),\ldots,\mu_K(s)$ can be numbered so that $\mu_j(s) = \lambda_j + s + o(1)$ as $s \to \infty$, so they must be distinct for s sufficiently large. □

Thus we have proved

THEOREM 4.3: <u>If $B_i(0) = 0$ for each i, then equation (4.4) is valid throughout $\{Re(s) > 0\}$ except possibly at isolated points.</u>

REMARK: $H_j(s)$ and $\mu_j(s)$ have been defined only in terms of the matrix $\Psi(s)$, which is usually unknown. In order to solve for $H_j(s),\mu_j(s)$ we need to derive equations satisfied by $H_j(s),\mu_j(s)$. We shall obtain such equations by substituting

$$\phi(w,s) = \sum_{j=1}^{K} H_j(s)\, e^{-w\mu_j(s)}$$

into an integral equation for $\phi(w,s)$.

First we prove some properties of $H_j(s),\mu_j(s)$:

LEMMA 4.6 (i) $\sum_{j=1}^{K} H_j(s) = 1$ <u>for all $s \in S$</u>

(ii) <u>If $B_i(0) = 0$ for each i, then</u> $\lim_{Re(s)\to\infty} H_j(s) = \pi_j$ <u>for each j.</u>

(iii) $\underline{\mathrm{Re}\{\mu_j(s)\} \geq 0 \text{ for } s \text{ in } \{\mathrm{Re}(s) \geq 0\};\ \mathrm{Re}\{\mu_j(s)\} > 0}$
$\underline{\text{for } s \text{ in } \{\mathrm{Re}(s) > 0\}.}$

PROOF: (i) is clear since $\phi(0,s) = 1$ for all s. To prove (ii) we restrict our proof to the case s real. From the proof of Lemma 4.5 we know that $\mu_\ell(s) = \lambda_\ell + s + o(1)$ as $s \to \infty$. Therefore, $\mu_j(s) - \mu_k(s) = \lambda_j - \lambda_k + o(1)$; since $\omega_{ij}(s) = o(1)$ it follows that the (m,n) - element of the matrix $\Psi(s) - \mu_k(s)I_\kappa$ is $\psi_{mn}(s) - \mu_k(s)\,\delta_{mn} = (\lambda_m - \lambda_k)\,\delta_{mn} + o(1)$ as $s \to \infty$. For s sufficiently large, $s \in S$ so that $H_j(s)$ is the sum of the elements of the matrix

$$\Lambda \prod_{k \neq j} \left\{ \frac{\Psi(s) - \mu_k(s) I_\kappa}{\mu_j(s) - \mu_k(s)} \right\}.$$

Since all off-diagonal elements of $\Psi(s) - \mu_k(s)I_\kappa$ are $o(1)$, it follows that the (m,n) - element of $\prod_{k\neq j} \{\Psi(s) - \mu_k(s)I_\kappa\}$ is

$$\prod_{k\neq j} (\lambda_m - \lambda_k)\,\delta_{mn} + o(1) = \delta_{jm}\delta_{mn} + o(1),$$

Hence $H_j(s) = \pi_j + o(1)$.

Suppose $\mu(s)$ is an eigenvalue of $\Psi(s)$. By Geršgorin's Theorem there exists $j \in \{1,2,\ldots,\kappa\}$ such that

$$|\mu(s) - (\lambda_j + s) + \lambda_j \omega_{jj}(s)| \leq \sum_{k\neq j} \lambda_j |\omega_{jk}(s)|$$

(see proof of Lemma 4.5). Hence

$$\operatorname{Re}\{\mu(s)\} \geq \operatorname{Re}\{\lambda_j + s - \lambda_j\omega_{jj}(s)\} - \sum_{k\neq j} \lambda_j |\omega_{jk}(s)| .$$

Now

$$|\omega_{jk}(s)| \leq \int_0^\infty |\phi_{.k}(u,s)| B_j(du) \leq \int_0^\infty \Pr\{Z_{Y(u)} = k\} B_j(du) = q_{jk},$$

say. q_{jk} is the probability that a busy period initiated by a type j customer ends during arrival state k, so $q_{jk} \geq 0$ and $\sum_k q_{jk} = 1$. Since $\operatorname{Re}\{\omega_{jj}(s)\} \leq |\omega_{jj}(s)| \leq q_{jj}$, we have

$$\operatorname{Re}\{\mu(s)\} \geq \lambda_j + \operatorname{Re}(s) - \lambda_j q_{jj} - \lambda_j \sum_{k\neq j} q_{jk} = \operatorname{Re}(s).$$

Thus $\operatorname{Re}(s) \geq 0 \Rightarrow \operatorname{Re}\{\mu(s)\} \geq 0$, $\operatorname{Re}(s) > 0 \Rightarrow \operatorname{Re}\{\mu(s)\} > 0$. □

REMARKS: 1) For each j, the zeros of $H_j(s)$ must be isolated in S^+, since Lemma 4.6 (ii) implies that $H_j(s)$ is not identically zero, and we already know that $H_j(s)$ is analytic throughout S^+ (Lemma 4.3).

NOTE: This presumes that $B_i(0) = 0$ for all i; we assume henceforth that this is true.

2) Part (ii) of Lemma 4.6 also implies that

$$\lim_{\operatorname{Re}(s)\to\infty} e^{sw}\phi(w,s) = \sum_{j=1}^{\kappa} \pi_j e^{-\lambda_j w} ,$$

a result that we can obtain from another direction as follows: Let $\xi(w) = Y(w) - w$, i.e. $\xi(w)$ is the excess length over the initial service time of a busy period induced by an initial service time w. Then $\Pr\{\xi(w) = 0\} = \Pr\{\text{no arrivals in } (0,w]\} = \sum_{j=1}^{\kappa} \pi_j e^{-\lambda_j w}$. But $E\{e^{-s\xi(w)}\} = e^{sw}\phi(w,s)$, so that

$$\Pr\{\xi(w) = 0\} = \lim_{Re(s)\to\infty} e^{sw}\,\phi(w,s) .$$

We now derive an integral equation for $\phi(w,s)$:

LEMMA 4.7: <u>$\phi(w,s)$ satisfies the integral equation</u>

$$\phi(w,s) = \sum_{j=1}^{K} \pi_j \left\{ e^{-(s+\lambda_j)w} + \int_0^w \lambda_j\, e^{-(\lambda_j+s)z} \int_0^\infty \phi(w-z+v,s)B_j(dv)dz \right\} \quad (4.5)$$

<u>for</u> $w \geq 0$, $Re(s) \geq 0$.

PROOF: The derivation of this equation follows a standard pattern for integral equations in queueing theory, so we merely sketch the proof. We have

$$\phi(w,s) = \sum_{j=1}^{K} \pi_j\, E\{e^{-sY(w)} | Z_0 = j\} .$$

Given $Z_0 = j$, either no arrival occurs in $(0,w]$ with probability $e^{-\lambda_j w}$, in which case $Y(w) = w$, or the first arrival in $(0,w]$ takes place at time z with probability $\lambda_j e^{-\lambda_j z} dz$, increasing the virtual waiting time at z to $w - z + V$ where V has d.f. B_j, so that in this case

$$E\left\{e^{-sY(w)}\right\} = \int_0^\infty e^{-sz}\, \phi(w-z+v,s)B_j(dv) . \qquad \square$$

To derive equations for $H_j(s), \mu_j(s)$ we now substitute the expression for $\phi(w,s)$ from equation (4.4) into equation (4.5). Then for each $s \in S$ we have

$$\sum_{i=1}^{K} H_i(s)\, e^{-w\mu_i(s)} = \sum_{j=1}^{K} \pi_j \Big\{ e^{-(\lambda_j+s)w}$$

$$+ \int_0^w \lambda_j\, e^{-(\lambda_j+s)z} \sum_{i=1}^{K} H_i(s) \int_0^\infty e^{-(w-z+v)\mu_i(s)} B_j(dv)\,dz \Big\}.$$

Now
$$\int_0^w \lambda_j\, e^{-(\lambda_j+s)z} \sum_{i=1}^{K} H_i(s) \int_0^\infty e^{-(w-z+v)\mu_i(s)} B_j(dv)\,dz$$

$$= \lambda_j \sum_{i=1}^{K} H_i(s)\, e^{-w\mu_i(s)} \int_0^w e^{-(\lambda_j+s-\mu_i(s))z} \int_0^\infty e^{-v\mu_i(s)} B_j(dv)\,dz$$

$$= \lambda_j \sum_{i=1}^{K} H_i(s)\, e^{-w\mu_i(s)}\, B_j^*\{\mu_i(s)\} \int_0^w e^{-(\lambda_j+s-\mu_i(s))z}\, dz$$

$$= \lambda_j \sum_{i=1}^{K} H_i(s)\, e^{-w\,\mu_i(s)}\, B_j^*\{\mu_i(s)\}\, \frac{1 - e^{-(\lambda_j+s-\mu_i(s))w}}{\lambda_j + s - \mu_i(s)} \quad .$$

Therefore for each $s \in S$ we have

$$\sum_{i=1}^{K} H_i(s)\, e^{-w\mu_i(s)}$$

$$= \sum_{j=1}^{K} \pi_j \left\{ e^{-(s+\lambda_j)w} + \lambda_j \sum_{i=1}^{K} H_i(s)\, B_j^*\{\mu_i(s)\}\, \frac{e^{-w\mu_i(s)} - e^{-(\lambda_j+s)w}}{\lambda_j + s - \mu_i(s)} \right\} \qquad (4.6)$$

By Lemma A4 , we may equate the coefficients of $e^{-w\mu_j(s)}$ and $e^{-(s+\lambda_j)w}$ to zero, since for $s \in S$ $\mu_1(s),\ldots,\mu_K(s)$ are distinct, $\lambda_1,\ldots,\lambda_K$ are distinct and $\mu_i(s)$ can equal $\lambda_j + s$ for some j only at isolated points in S^+ .

Equating the coefficient of $e^{-(s+\lambda_j)w}$ to zero in equation (4.6), we obtain for $s \in S$, $j = 1,2,\ldots,\kappa$

$$\lambda_j \sum_{i=1}^{\kappa} \frac{H_i(s)\; B_j^*\{\mu_i(s)\}}{\lambda_j + s - \mu_i(s)} = 1 \; . \tag{4.7}$$

The coefficient of $e^{-w\mu_i(s)}$ in equation (4.6) is

$$H_i(s)\left\{\sum_{j=1}^{\kappa} \frac{\pi_j\lambda_j\; B_j^*\{\mu_i(s)\}}{\lambda_j + s - \mu_i(s)} - 1\right\} .$$

Since each $\mu_i(s)$ is continuous throughout $\{\mathrm{Re}(s) \geq 0\}$ and $H_i(s)$ has isolated zeros in S^+, for this coefficient to be zero we must have for every $s \in S^+$

$$\sum_{j=1}^{\kappa} \frac{\pi_j\lambda_j\; B_j^*\{\mu_i(s)\}}{\lambda_j + s - \mu_i(s)} = 1 \quad \text{for } i = 1,2,\ldots,\kappa \; . \tag{4.8}$$

We may extend the domain of validity of equation (4.8) to the entire right half-plane $\{\mathrm{Re}(s) \geq 0\}$ in two steps: every s in $\{\mathrm{Re}(s) > 0\} - S^+$ can be surrounded by a deleted neighbourhood in S^+, so that (4.8) holds for all s in $\{\mathrm{Re}(s) > 0\}$ by the continuity of $\mu_i(s)$; since $\mu_i(s)$ is continuous to the right at the imaginary axis, it follows that (4.8) holds for s on the imaginary axis as well. To summarize, we have shown that

$$\lambda_j \sum_{i=1}^{\kappa} \frac{H_i(s)\; B_j^*\{\mu_i(s)\}}{\lambda_j + s - \mu_i(s)} = 1 \quad \text{for } s \in S,\; j = 1,2,\ldots,\kappa \; ; \tag{4.7}$$

$$\sum_{j=1}^{\kappa} \frac{\pi_j\lambda_j\; B_j^*\{\mu_i(s)\}}{\lambda_j + s - \mu_i(s)} = 1 \quad \text{for } \mathrm{Re}(s) \geq 0,\; i = 1,2,\ldots,\kappa \; . \tag{4.8}$$

Define

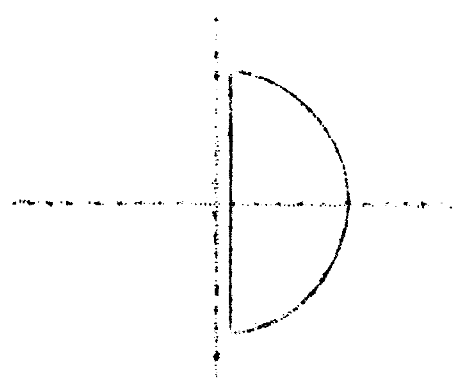
$$Q(z,s) = \sum_{j=1}^{\kappa} \frac{\pi_j \lambda_j \; B_j^*(z)}{\lambda_j + s - z} \quad \text{for} \quad Re(z) \geq 0, \; Re(s) \geq 0.$$

LEMMA 4.8: For any fixed positive s, the equation $Q(z,s) = 1$ has exactly κ roots in $\{Re(z) > 0\}$.

PROOF: Once again we apply Rouché's Theorem, our proof following the pattern of the proof of Lemma 2.10. Consider the simple closed contour $\Gamma(\varepsilon,R)$ defined for $0 < \varepsilon < R$ by

$$\Gamma(\varepsilon,R) = \{z : Re(z) = \varepsilon, |z| \leq R\} \cup \{z : |z| = R, Re(z) > \varepsilon\} .$$

$\Gamma(\varepsilon,R)$ is almost a semicircle in the right half-plane, as in the following diagram:



Since $Q(z,s)$ is analytic in $\{\mathrm{Re}(s) > 0\}$ except for a finite number of singularities at $z = \lambda_j + s$, $j = 1,2,\ldots,\kappa$, $Q(z,s)$ is analytic on $\Gamma(\varepsilon,R)$ for all ε sufficiently small and all R sufficiently large. We show next that we may choose ε,R so that $|Q(z,s)| < 1$ for all $z \in \Gamma(\varepsilon,R)$.

For $|z| = R, \mathrm{Re}(z) > 0$ we have $B_j^*(z) \le 1, |\lambda_j+s-z| \ge R-|\lambda_j+s|$,

$$\text{so that} \quad |Q(z,s)| \le \sum_{j=1}^{\kappa} \frac{\pi_j\lambda_j}{|R-|\lambda_j+s||} \to 0 \text{ as } R \to \infty .$$

Hence for R sufficiently large, $|Q(z,s)| < 1$ on this arc. For z on the imaginary axis, $z = i\theta$, say,

$$|Q(z,s)| \le \sum_{j=1}^{\kappa} \frac{\pi_j\lambda_j|B_j^*(i\theta)|}{|\lambda_j + s - i\theta|}$$

$$\le \sum_{j=1}^{\kappa} \frac{\pi_j\lambda_j}{\{(\lambda_j+\sigma)^2 + (\tau-\theta)^2\}^{1/2}} \quad \text{where } s=\sigma+i\tau$$

$$< \sum_{j=1}^{\kappa} \pi_j = 1 \text{ since } \sigma > 0 .$$

It follows by Lemma A3 that for ε sufficiently small $|Q(z,s)| < 1$ for all z in $\{\mathrm{Re}(z) = \varepsilon, |z| \le R\}$. Thus we may choose ε,R so that $|Q(z,s)| < 1$ for all $z \in \Gamma(\varepsilon,R)$.

Let $P(z) = \prod_{j=1}^{\kappa} (\lambda_j + s - z)$, $M(z) = P(z)Q(z,s)$. Then

$$M(z) = \sum_{j=1}^{\kappa} \pi_j \lambda_j B_j^*(z) \prod_{\ell \neq j} (\lambda_\ell + s - z) ,$$

so $P(z)$ and $M(z)$ are analytic inside and on $\Gamma(\varepsilon,R)$, and $|M(z)| = |P(z)||Q(z,s)| < |P(z)|$ for all $z \in \Gamma(\varepsilon,R)$. Therefore, by Rouché's Theorem, $P(z) - M(z) = P(z)\{1 - Q(z,s)\}$ has the same number of zeros inside $\Gamma(\varepsilon,R)$ as does $P(z)$. It follows that $P(z) - M(z)$ has exactly κ zeros in $\{Re(z) > 0\}$.

Each zero of $P(z) - M(z)$ must be a root of $Q(z,s) = 1$, for if $P(z_o) = P(z_o) - M(z_o) = 0$, then $z_o = \lambda_k + s$ for some $k \in \{1,2,\dots,\kappa\}$ and

$$0 = M(z_o) = \pi_k \lambda_k B_k^*(\lambda_k + s) \prod_{\ell \neq k} (\lambda_\ell - \lambda_k) .$$

But this is impossible, since $\pi_k, \lambda_k > 0$, $\lambda_1,\dots,\lambda_\kappa$ are distinct and $B_k^*(u)$ cannot be zero for u positive (Lemma A1). Thus we have shown that the equation $Q(z,s) = 1$ has exactly κ roots in $\{Re(z) > 0\}$. □

LEMMA 4.9: For each fixed positive s, the κ roots of the equation $Q(z,s) = 1$ in $\{Re(z) > 0\}$ lie on the positive real axis: one root in the interval $(s, s + \lambda_1)$ and one in each of the intervals $(s + \lambda_i, s + \lambda_{i+1})$ for $i = 1,2,\dots,\kappa-1$.

PROOF: $Q(z,s)$ is real and continuous for z on the positive real axis except at the points $z = \lambda_j + s$. Since $Q(\lambda_i + s + 0,s) = -\infty$

and $Q(\lambda_{i+1} + s - 0, s) = +\infty$, it follows that there is at least one value of z in the interval $(s + \lambda_i, s + \lambda_{i+1})$ such that $Q(z,s) = 1$, for $i = 1,2,\ldots,\kappa-1$. Since

$$Q(s,s) = \sum_{j=1}^{\kappa} \pi_j \, B_j^{\,*}(s) < 1$$

and $Q(s + \lambda_1 - 0) = +\infty$, there is at least one root of $Q(z,s) = 1$ in the interval $(s, s + \lambda_1)$. By Lemma 4.8 there is exactly one root in each of these intervals. □

<u>COROLLARY 1</u>: $\Psi(s)$ has κ distinct eigenvalues for all positive s.

PROOF: The eigenvalues of $\Psi(s)$ are κ roots of the equation $Q(z,s) = 1$ in the region $\{\mathrm{Re}(z) > 0\}$. By Lemma 4.8 there are exactly κ roots in this region, so the eigenvalues must be the κ distinct positive roots of Lemma 4.9.

<u>COROLLARY 2</u>: $\phi(w,s) = \sum_{j=1}^{\kappa} H_j(s)\, e^{-w\mu_j(s)}$ for all positive s.

We can now, in principle, calculate $\phi(w,s)$ for any s in $\{\mathrm{Re}(s) \geq 0\}$: for positive values of s we solve (4.8) for $\mu_1(s),\ldots,\mu_\kappa(s)$ and then (4.7) for $H_1(s),\ldots,H_\kappa(s)$; for s in $\{\mathrm{Re}(s) \geq 0, \mathrm{Im}(s) \neq 0\}$ we need to check whether the κ solutions of $Q(z,s) = 1$ are distinct: if they are, then we conclude that $s \in S$ so that we can solve (4.7) for $H_1(s),\ldots,H_\kappa(s)$ as before; if not then we must evaluate $\phi(w,s)$ as $\lim_{n\to\infty} \phi(w,s_n)$ where $\{s_n\}$ is a sequence in $\{\mathrm{Re}(s) > 0\}$ such that $s_n \to s$.

4.3 *The Mean Length of a Busy Period*

Let $m(w) = E\{Y(w)\}$, i.e. $m(w)$ is the mean length of a busy period induced by an initial service time w. We shall refer to a busy period initiated by a type j customer as a type j busy period. The mean length of a type j busy period is

$$m_j \underset{\text{def}}{=} E\{Y_j\} = \int_0^\infty m(w)\ B_j(dw) \ .$$

In this section we show how $m(w)$ (and consequently $m_1,\ldots,m_K$) may be computed for the model of Section 4.2. In Chapter 3 we showed how to calculate the probabilities $\gamma_1,\ldots,\gamma_K$ where γ_j is the stationary probability that a busy period be of type j. If $\gamma_1,\ldots,\gamma_K$ are known, then we can evaluate such quantities as the overall mean length of a busy period,

$$m = \sum_{j=1}^{K} \gamma_j m_j \ ,$$

the mean length of an idle period, and the stationary probability that the server is free. As usual, we let $b = \sum \pi_j b_j$, $a = \sum \pi_j \lambda_j^{-1}$ be the overall mean service and inter-arrival times.

Procedure A:

We know that

$$m(w) = \lim_{s \to 0+} \frac{1 - \phi(w,s)}{s} \ ,$$

so a possible procedure for computing $m(w)$ is to evaluate

$$\frac{1 - \phi(w,s_n)}{s_n}$$

for a sequence $\{s_n\}$ of positive real numbers such that $s_n \to 0+$, and estimate $m(w)$ as the apparent limit.

Procedure A requires the solution of equations (4.7) and (4.8) at each point $s_n, n = 1,2,\ldots,N$, where N is the number of values required to approximate $m(w)$. A less time-consuming and more exact procedure is available if $\mu_j(s)$ and $H_j(s)$ are differentiable to the right at $s = 0$ for each j, for then we have

$$m(w) = -\sum_{j=1}^{\kappa} \{H_j'(0) - w\mu_j'(0)H_j(0)\} e^{-w\mu_j(0)} . \tag{4.9}$$

For $H_1(s),\ldots,H_\kappa(s)$ to be defined at $s = 0$, we need $\Psi(0)$ to have κ distinct eigenvalues. We prove this fact in the following lemma.

LEMMA 4.10: <u>$\Psi(0)$ has κ distinct eigenvalues.</u>

PROOF: The (i,j) - element of $\Psi(0)$ is $\psi_{ij}(0) = \lambda_i(\delta_{ij} - q_{ij})$, where q_{ij} $(= \omega_{ij}(0))$ is the probability that a type i busy period ends during arrival state j (see the proof of Lemma 4.6). Since $\sum_j q_{ij} = 1$ for each i, the row sums of $\Psi(0)$ are all zero. Therefore $\Psi(0)$ is singular and has 0 as one eigenvalue: write $\mu_1(0) = 0$. From Lemma 4.9 we may number the eigenvalues of $\Psi(s)$ for s positive so that for $i = 2,3,\ldots,\kappa$

$$s + \lambda_{i-1} < \mu_i(s) < s + \lambda_i .$$

Therefore each $\mu_i(0) = \lim_{s\to 0+} \mu_i(s)$ is positive, and

$$\lambda_{i-1} \le \mu_i(0) \le \lambda_i \quad (i = 2,3,\ldots,\kappa) .$$

But we know that each $\mu_i(0)$ is a solution of the equation

$$\sum_{j=1}^{\kappa} \frac{\pi_j \lambda_j \; B_j^*(z)}{\lambda_j - z} = 1,$$

so it is impossible that $\mu_i(0) = \lambda_j$ for any j . Thus

$$\lambda_{i-1} < \mu_i(0) < \lambda_i \quad (i = 2,3,\ldots,\kappa)$$

so that $\mu_1(0),\ldots,\mu_\kappa(0)$ are distinct .

<u>COROLLARY:</u> $H_1(0) = 1$, $H_j(0) = 0$ for $j = 2,3,\ldots,\kappa$.

PROOF: This follows from Lemma A4 , since

$$\sum_{j=1}^{\kappa} H_j(0) \; e^{-w\mu_j(0)} = \phi(w,0) \equiv 1 \quad \text{for all } w ,$$

and we can write this as

$$\{H_1(0) - 1\} \; e^{-w\mu_1(0)} + \sum_{j=2}^{\kappa} H_j(0) \; e^{-w\mu_j(0)} \equiv 0 . \qquad \square$$

For $H_1(s),\ldots,H_\kappa(s)$ to be differentiable to the right at $s = 0$, it is sufficient that $\mu_1(s),\ldots,\mu_\kappa(s)$ be differentiable to the right at $s = 0$, this being clear from the definition of $H_j(s)$. We shall prove that each $\mu_j(s)$ is differentiable to the right at $s = 0$, but first we need the following result:

LEMMA 4.11: <u>The mean length of a type j busy period is finite, i.e. $m_j < \infty$ for each j.</u>

PROOF: Let U_n be the time between the n^{th} and $(n+1)^{th}$ arrivals, and V_n the service time of the n^{th} customer to arrive, assuming that a type j busy period began at time 0. Then $U_o, U_1, U_2, \ldots$ are iid random variables with $E\{U_j\} = a$, and $V_1, V_2, \ldots$ are iid random variables with $E\{V_j\} = b$. Let $X_n = V_n - U_{n-1}$ $(n = 1,2,\ldots)$. Then $X_1, X_2, \ldots$ are iid random variables with $E\{X_j\} = b - a < 0$.

Let M be the number of customers served during the busy period. Let $N = \inf\{k : V_o + X_1 + \ldots + X_k < 0\}$. ($N$ is finite a.s. since $EX_j < 0$). If the busy period is still in progress when the n^{th} customer arrives, then his total waiting time is $V_o + X_1 + \ldots + X_n$ (see the proof of Lemma 3.2). Thus $0 < M \leq N$.

We will show that $EN < \infty$, so that $EM < \infty$. Let $S_k = V_o + X_1 + \ldots + X_k$ $(k = 1,2,\ldots)$. Now $0 > S_N = (S_{N-1} + V_N) - U_{N-1}$ and $S_{N-1} \geq 0$, so $S_{N-1} + V_N \geq 0$. We may show that S_N is distributed as a mixture of exponential random variables, since it is the overshoot below zero of a mixture of exponential random variables. Specifically, we recall that q_{jk} is the probability that a type j busy period end during arrival state k. Suppose the busy period ends during arrival state k. Then U_{N-1} has an exponential distribution with parameter λ_k, and

$$\Pr\{S_N < -x \mid Z_N = k\} = \Pr\{U_{N-1} > x + S_{N-1} + V_N \mid U_{N-1} > S_{N-1} + V_N, Z_N = k\}$$

$$= \Pr\{U_{N-1} > x \mid Z_N = k\} = e^{-\lambda_k x} \quad \text{for } x > 0$$

by the lack of memory of the exponential distribution. Thus

$$\Pr\{S_N < -x\} = \sum_{k=1}^{K} q_{jk}\, e^{-\lambda_k x} \quad \text{for any } x > 0 .$$

But this implies that

$$(0 \geq)\ E\{S_N\} = -\sum_{k=1}^{K} q_{jk}\, \lambda_k^{-1} > -\infty .$$

Next we use a variant of Wald's equation to show that $E\{N\} < \infty$. We apply Lemma A5 with $Z_o = V_o$, $Z_k = X_k$ for $k = 1,2,\ldots,$ and stopping time N. We have

$$E\{S_N\} = E\{V_o\} + E\{X_1\}\, E\{N\}$$
$$= b_j + (b-a) \circ E\{N\} .$$

It follows that $E\{N\} < \infty$, and consequently $E\{M\} < \infty$.

Again we apply Lemma A5 , with $X_j = U_{j-1}$ and F_k being the σ-field generated by $V_o, U_o, V_1, U_1, \ldots, U_{k-1}, V_k$, i.e. for $k \geq 1$ $Z_k = (U_{k-1}, V_k)$. $U_o, U_1, U_2, \ldots$ are iid and U_{k-1} is dependent on Z_k. Since we could define M as $\inf\{k : (V_o - U_o) + (V_1 - U_1) + \ldots + (V_{k-1} - U_{k-1}) < 0\}$, it follows that $[M \leq k] \in F_k$. Therefore

$$E\left\{\sum_{j=0}^{M-1} U_j\right\} = a \cdot E\{M\} < \infty .$$

But $0 \leq Y_j < U_o + U_1 + \ldots + U_{M-1}$, so that $m_j = EY_j < \infty$.

COROLLARY: For all j, k $\omega_{jk}(s)$ is differentiable to the right at $s = 0$.

PROOF: Since $\omega_{jk}(s)$ is a multiple of a Laplace transform of a non-negative random variable,

$$\omega_{jk}'(0) \underset{\text{def}}{=} \lim_{s \to 0+} \frac{\omega_{jk}(s) - \omega_{jk}(0)}{s}$$

is well-defined, either finite or $+\infty$. Now

$$\sum_{k=1}^{K} \omega_{jk}(s)$$

is the Laplace transform of the length of a type j busy period, so

$$-\sum_{k=1}^{K} \omega_{jk}'(0) = EY_j = m_j .$$

It follows from the Lemma that $\omega_{jk}'(0)$ is finite for all j,k. □

We can now prove

LEMMA 4.12: Each $\mu_j(s)$ is differentiable to the right at $s = 0$.

PROOF: Suppose first that $\mu(s)$ is an eigenvalue of $\Psi(s)$ such that $\mu(0) = 0$. Then $\mu(s)$ is defined implicitly by the equation

$$\sum_{j=0}^{K} C_j(s)\{\mu(s)\}^j = 0 \quad (\mathrm{Re}(s) \geq 0) ,$$

where $C_j(s)$ is a polynomial in the elements of $\Psi(s)$. Now $C_o(0) = 0$ and $C_1(0) \neq 0$, since 0 is a non-repeated eigenvalue of $\Psi(0)$.

Recalling that $\psi_{ij}(s) = (\lambda_i + s)\delta_{ij} - \lambda_i\omega_{ij}(s)$, we see by the above corollary that each $C_j(s)$ is differentiable to the right at $s = 0$. Now

$$\sum_{j=0}^{\kappa} \{(C_j(s) - C_j(0))(\mu(s))^j + C_j(0)(\mu(s))^j\} = 0$$

for all s in $\{\mathrm{Re}(s) \geq 0\}$, so that for all $s > 0$

$$\sum_{j=0}^{\kappa} \left\{\frac{C_j(s) - C_j(0)}{s} (\mu(s))^j + C_j(0)\frac{(\mu(s))^j}{s}\right\} = 0 \ .$$

Since $C_j(s)$ is differentiable to the right at $s = 0$ and $\mu(s)$ is continuous to the right at $s = 0$, it follows that

$$\lim_{s\to 0+} \left\{\sum_{j=0}^{\kappa} C_j(0) \frac{(\mu(s))^j}{s}\right\} = \lim_{s\to 0+}\left\{\frac{\mu(s)}{s} \sum_{j=1}^{\kappa} C_j(0)(\mu(s))^{j-1}\right\}$$

exists. But

$$\lim_{s\to 0+} \sum_{j=1}^{\kappa} C_j(0)(\mu(s))^{j-1} = C_1(0) \neq 0 \ , \quad \text{so} \quad \lim_{s\to 0+} \frac{\mu(s)}{s}$$

exists, i.e. $\mu(s)$ is differentiable to the right at $s = 0$.

Suppose now that $\mu(0) = c \neq 0$. Let $\Psi_c(s) = \Psi(s) - cI_\kappa$. Then $\Psi_c(s)$ has an eigenvalue $\eta(s)$ such that $\eta(s) = \mu(s) - c$. Since $\eta(0) = 0$ and $\Psi_c(0)$ has κ distinct eigenvalues, we may proceed as above to show that $\eta(s)$ (and consequently $\mu(s)$) is differentiable to the right at $s = 0$. □

Thus, we have shown by Lemmas 4.10 and 4.12 that (4.9) holds, and by the Corollary to Lemma 4.10, we have

$$m(w) = w\mu_1'(0) - \sum_{j=1}^{\kappa} H_j'(0)\, e^{-w\mu_j(0)} . \qquad (4.10)$$

We need to consider how to compute the quantities $\mu_j(0), \mu_1'(0), H_j'(0)$. We know already that $\mu_1(0) = 0$; the non-zero eigenvalues of $\Psi(0)$ are positive roots of the equation $Q(z,0) = 1$. Write $Q(z)$ for $Q(z,0)$. If this equation had more than $\kappa-1$ positive roots, then we would need to consider the limiting behaviour of the roots of the equation $Q(z,s) = 1$ as $s \to 0+$ in order to decide which positive roots of $Q(z) = 1$ were the eigenvalues of $\Psi(0)$. Our next result shows that this contingency does not arise.

LEMMA 4.13: The equation $Q(z) = 1$ has exactly $\kappa-1$ positive roots.

PROOF: From the proof of Lemma 4.10 we know that the equation $Q(z) = 1$ has at least one root in each of the intervals $(\lambda_i, \lambda_{i+1})$, $i = 1,2,\ldots,\kappa-1$. We may show that once $Q(z)$ reaches 1 in any of these intervals it is strictly increasing throughout the remainder of the interval.

Let

$$g_j(z) = \frac{\lambda_j}{\lambda_j - z} B_j^*(z), \quad j = 1,2,\ldots,\kappa .$$

Then $g_j(z)$ is the (bilateral) Laplace transform of a probability distribution, with region of convergence $\{z : 0 < \mathrm{Re}(z) < \lambda_j\}$, although $g_j(z)$ is actually defined for all z in $\{\mathrm{Re}(z) > 0\}$ such that $z \neq \lambda_j$. In particular, $g_j(z) = E\{e^{-z(V-U)}\}$, where V and U are independent random variables such that V has d.f. B_j and U is exponentially distributed with parameter λ_j. On the non-negative real axis $g_j(x)$ has the following properties:

(P1) $g_j(0) = 1$, $g_j(x) > 0$ for all $x \in (0,\lambda_j)$, $g_j(\lambda_j - 0) = +\infty$;

(P2) $g_j'(0) = 1/\lambda_j - b_j$;

(P3) $g_j''(x) > 0$ for all $x \in (0,\lambda_j)$;

(P4) $g_j(x) < 0$ and $g_j'(x) > 0$ for all $x > \lambda_j$.

(P1),(P2) and (P4) are clear since

$$g_j'(x) = \frac{\lambda_j}{\lambda_j - x} B_j^{*\prime}(x) + \frac{\lambda_j}{(\lambda_j - x)^2} B_j^*(x) .$$

(P3) follows from the equation

$$g_j''(x) = \int_{-\infty}^{\infty} u^2 e^{-ux} G_j(du)$$

where G_j is the d.f. of $V - U$ [23; p. 240], since this integral is finite and positive for all $x \in (0,\lambda_j)$ ($g_j(z)$ is analytic for $0 < \mathrm{Re}(z) < \lambda_j$) .

We now consider the behaviour of

$$Q(x) = \sum_j \pi_j g_j(x)$$

for $x \geq 0$.

$$Q(0) = 1, Q'(0) = \sum_j \pi_j(1/\lambda_j - b_j) = a - b > 0 ,$$

and $Q(x)$ is convex for all $x \in (0,\lambda_1)$ by (P3) (Recall that $\lambda_1 < \lambda_2 < \ldots < \lambda_\kappa$) . Therefore $Q(x) > 1$ for all $x \in (0,\lambda_1)$. From (P4) it is clear that $Q(x) < 0$ for all $x > \lambda_\kappa$, so that the equation $Q(z) = 1$ has no roots in the intervals $(0,\lambda_1)$ or (λ_κ,∞). For the interval $(\lambda_i,\lambda_{i+1})$, $i \in \{1,2,\ldots,\kappa-1\}$, we define

$$f_1(x) = \pi_1 g_1(x) + \ldots + \pi_i g_i(x) \qquad x \neq \lambda_1,\ldots,\lambda_i, x \geq 0,$$

$$f_2(x) = \pi_{i+1} g_{i+1}(x) + \ldots + \pi_\kappa g_\kappa(x) \quad x \neq \lambda_{i+1},\ldots,\lambda_\kappa, x \geq 0.$$

By (P4) $f_1(x)$ is negative and strictly increasing for $x > \lambda_i$. Now $0 < f_2(x) < \infty$ for $0 < x < \lambda_{i+1}$, and $f_2(\lambda_{i+1} - 0) = +\infty$. Therefore we can define $x_i = \inf \{x : x \geq \lambda_i$ and $f_2(x) \geq 1\}$. The equation $Q(z) = 1$ can have no roots in (λ_i, x_i) since $Q(x) = f_1(x) + f_2(x) < 1$ for all x (if any) in this interval. By (P3), $f_2(x)$ is convex in $(0,\lambda_{i+1})$. Therefore, since $f_2(0) = 1$ and $f_2(x_i) \geq 1$, $f_2(x)$ is strictly increasing in the interval (x_i, λ_{i+1}). It follows that $Q(x)$ is strictly increasing for $x \in (x_i, \lambda_{i+1})$. Since $Q(\lambda_{i+1} - 0) = +\infty$, $Q(\lambda_i + 0) = -\infty$ it follows that $Q(z) = 1$ has exactly one root in $(\lambda_i, \lambda_{i+1})$ for each $i \in \{1,2,\ldots,\kappa-1\}$. Thus there are exactly $\kappa-1$ positive roots. $\square$

Thus $\mu_2(0),\ldots,\mu_\kappa(0)$ are unambiguously determined if $\kappa-1$ positive roots are found for the equation $Q(z) = 1$. A simple iterative technique, such as Newton's method, is appropriate here.

We may obtain $\mu_1'(0)$ by differentiating equation (4.8) at $s = 0$, setting $i = 1$. We obtain

$$\mu_1'(0) = \frac{a}{a - b} .$$

We may obtain $H_1'(0),\ldots,H_\kappa'(0)$ as follows. If we differentiate equation (4.7) with respect to s and set $s = 0$, then we obtain these κ linear equations:

$$\sum_{k=1}^{\kappa} \frac{\lambda_j H_k{}'(0)\; B_j{}^*\{\mu_k(0)\}}{\lambda_j - \mu_k(0)} = \frac{1 + \mu_1{}'(0)\{\lambda_j b_j - 1\}}{\lambda_j}\ , \quad j = 1,\ldots,\kappa\ . \tag{4.11}$$

To summarize, we have the following procedure for computing $m(w)$:

Procedure B:

1) Solve the equation $Q(z) = 1$ for $\mu_2(0),\ldots,\mu_\kappa(0)$.

2) Compute $\mu_1{}'(0) = \dfrac{a}{a-b}$.

3) Solve Equation (4.11) for $H_1{}'(0),\ldots,H_\kappa{}'(0)$.

Then

$$m(w) = w\mu_1{}'(0) - \sum_{j=1}^{\kappa} H_j{}'(0)\, e^{-w\mu_j(0)} .$$

Hence we can calculate for each i

$$m_i = \int_0^\infty m(w)\; B_i(dw)$$

$$= \mu_1{}'(0) \cdot b_i - \sum_{j=1}^{\kappa} H_j{}'(0)\; B_i{}^*\{\mu_j(0)\}\ . \tag{4.12}$$

If we know the stationary probabilities $\gamma_1,\ldots,\gamma_\kappa$, where γ_j is the probability that a busy period be of type j, then we can calculate the overall mean length of a busy period $m = \sum_j \gamma_j m_j$.

Between busy periods are time intervals during which the server is free (or the queue empty) - we call these idle periods. The length of an idle period is the time to the next arrival after the end of a busy period. If the busy period ended during arrival state j, then the length of the idle period has an exponential distribution with parameter λ_j. The stationary probability that a busy period end during arrival state j is γ_j, since the next busy period must necessarily be of type j. Therefore the mean length of an idle period is

$$\sum_{j=1}^{\kappa} \gamma_j \circ \lambda_j^{-1} = a^* \text{ , say.}$$

By an argument similar to that in [5 ; p. 37], we could derive the intuitively obvious result, that the stationary probability that the server be free is $\frac{a^*}{m + a^*}$.

REMARK: When $\kappa = 1$, the present model reduces to the case of random arrivals (or the M/G/1 queue). A well-known result for the case of random arrivals is that the stationary probability of the server being free is $1 - \rho$, where $\rho = b/a$ [5 ; §2.7]. This result is not true for the present model when $\kappa > 1$: by (3.11) we have

$$\Pr\{\text{server is free}\} = \sum_{j=1}^{\kappa} \tau_j = \frac{a - b}{a^*} .$$

Thus $\Pr\{\text{server is free}\} = 1 - b/a$ iff $a = a^*$, and by Lemma 3.5 this is not true in general.

EXAMPLE: We solve for $m(w), m_1, \ldots, m_\kappa$ in the case $\kappa = 3$, with service time d.f.s $B_j(x) = 1 - e^{-\sigma_j x}$. We take

$$\pi_1 = \frac{1}{4}, \ \pi_2 = \frac{1}{2}, \ \pi_3 = \frac{1}{4}$$

$$\lambda_1 = 1, \ \lambda_2 = 2, \ \lambda_3 = 3$$

$$\sigma_1 = 2, \ \sigma_2 = 6, \ \sigma_3 = 1.$$

Carrying out Procedure B, we find that

$$\mu_2(0) = 1.34(3), \ \mu_3(0) = 2.89(4).$$

The mean overall service and inter-arrival times are

$$b = 11/24, \ a = 7/12$$

so $\mu_1'(0) = 14/3$. Solving (4.11) we find that

$$H_1'(0) = -0.73, \ H_2'(0) = 0.22, \ H_3'(0) = 0.51.$$

Thus $m(w) = 14w/3 + 0.73 - 0.22\, e^{-1.34w} - 0.51\, e^{-2.89w}$ and

$m_1 = 2.73$, $m_2 = 0.98$, $m_3 = 5.17$.

Solving Equations (3.9), (3.10) we obtain

$$\tau_1 = 0.069, \ \tau_2 = 0.090, \ \tau_3 = 0.034.$$

The stationary probability that the server be free is

$$\sum_j \tau_j = 0.19.$$

Therefore the stationary probabilities γ_j are

$$\gamma_1 = 0.36, \ \gamma_2 = 0.47, \ \gamma_3 = 0.18$$

(c.f. $\pi_1 = 0.25$, $\pi_2 = 0.5$ $\pi_3 = 0.25$), and the mean length of a general busy period is

$$\begin{aligned} m &= (0.36)(2.73) + (0.47)(0.98) + (0.18)(5.17) \\ &= 2.37. \end{aligned}$$

4.4 *The backward transition model with* $\lambda_o = \infty$.

The method of Section 4.2 may be applied to the (more general) backward transition model with $\lambda_o = \infty$. We obtain analogues of Theorems 4.1 and 4.2. However the integral equation corresponding to (4.5) involves components of the busy period transform $\phi(w,s)$ rather than $\phi(w,s)$ directly. As a result we are led to much more complicated equations.

We recall that arrivals for this model correspond to transitions into state 0 of an homogeneous Markov chain (Z_t) with state space $\{0,1,\ldots,\kappa\}$. The time spent in state i between transitions has an exponential distribution with parameter λ_i; at a transition the next state is chosen from $\{0,1,\ldots,\kappa\}$ according to the probabilities $\{p_{ij}\}$. We suppose $p_{oo} = 0$ (0 is an instantaneous state since $\lambda_o = \infty$).

We may describe this model in terms of infinitesimal probabilities as follows. Given $Z_t = i \neq 0,\ \tau > 0$:

$$\Pr\{\text{type } i \text{ arrival in } (t,t+\tau)\} = \lambda_i p_{io}\tau + o(\tau) ,$$

$$\Pr\{\text{transition to state } j \neq 0 \text{ in } (t,t+\tau)\} = \lambda_i (p_{ij} + p_{io}p_{oj})\tau + o(\tau)$$

$$\Pr\{\text{no transition in } (t,t+\tau)\} = 1 - \lambda_i\tau + o(\tau) .$$

At each arrival epoch the next state of (Z_t) is selected from $\{1,2,\ldots,\kappa\}$ according to the probabilities $\{p_{oj} : j = 1,2,\ldots,\kappa\}$. Thus we see that this model reduces to the backward model with κ

customer types and exponential arrivals if $p_{io} = 1$ for each i.

As in Section 4.2 we suppose that a customer with service time w has arrived at time 0 to find the server free. Let (W_t), $Y(w)$ and $\phi(w,s)$ be defined as before. (Z_t) is now the guiding chain described above; we take Z_t to be right-continuous, and define $\phi_{ij}(w,s)$ as before. Once again $((W_t, Z_t), t \geq 0)$ is an homogeneous Markov process, so we may prove analogues of (4.1) and (4.2), with π_j replaced by p_{oj}.

Let $\omega_{ij}(s)$ be as before, and $\Psi(s)$ the $\kappa \times \kappa$ matrix with (i,j)-element

$$\psi_{ij}(s) = (\lambda_i + s)\delta_{ij} - \lambda_i p_{io}\, \omega_{ij}(s) - \lambda_i p_{ij} .$$

We may show as in the proof of Theorem 4.1 that

$$\phi_{ij}(h,s) = p_{oi}\{(1 - (\lambda_i + s)h)\delta_{ij} + \lambda_i p_{ij} h + \lambda_i p_{io}\, \omega_{ij}(s) h\} + o(h)$$

as $h \to 0+$. Therefore we have the following analogue of Theorem 4.1:

PROPOSITION: $\Phi(w,s)$ satisfies the matrix differential equation

$$\frac{\partial}{\partial w} \Phi(w,s) = - \Phi(w,s)\, \Psi(s)$$

for all $w > 0$, s in $\{\mathrm{Re}(s) \geq 0\}$. $\square$

Since $\phi(w,s) = \sum_i \sum_j \phi_{ij}(w,s)$, we have the following analogue of Theorem 4.2:

PROPOSITION: For all $w \geq 0$ and all s in $\{\mathrm{Re}(s) \geq 0\}$ such that $\Psi(s)$ has κ distinct eigenvalues, $\phi(w,s)$ is of the form

$$\phi(w,s) = \sum_{j=1}^{\kappa} H_j(s)\, e^{-w\cdot\mu_j(s)} \qquad \square$$

Beyond this point the analogies with Section 4.2 begin to break down. For example, the proof of Lemma 4.5 does not work for the present model. More importantly, when we attempt to solve for $H_j(s), \mu_j(s)$ we find that the integral equation corresponding to (4.5) is

$$\phi(w,s) = \sum_{k=1}^{\kappa} p_{ok}\{e^{-(s+\lambda_k)w} + \lambda_k p_{ko} \int_0^w e^{-(\lambda_k+s)z} \int_0^\infty \phi(w-z+v,s) B_k(dv)\,dz$$

$$+ \sum_{j=1}^{\kappa} \lambda_k p_{kj} \int_0^w e^{-(\lambda_k+s)z}\, \phi_{j.}(w-z,s)\,dz\} .$$

Since this involves the components $\phi_{j.}(w,s)$ we cannot obtain equations for $H_j(s), \mu_j(s)$ as before. Now $\phi_{j.}(w,s)$ is the j^{th} row-sum of the matrix $\Phi(w,s)$. Therefore, as in the derivation of Theorem 4.2, we see that, if $\Psi(s)$ has κ distinct eigenvalues $\mu_1(s),\ldots,\mu_\kappa(s)$,

$$\phi_{j.}(w,s) = \sum_{k=1}^{\kappa} H_{jk}(s)\, e^{-w\,\mu_k(s)} ,$$

where $H_{jk}(s)$ is the j^{th} row-sum of the matrix $\ell_k(\Psi(s),s)$. We may obtain equations for $\mu_j(s), H_{jk}(s)$ by substituting this expression for $\phi_{j.}(w,s)$ into the integral equation

$$\phi_{j.}(w,s) = p_{oj}\{e^{-(s+\lambda_j)w} + \lambda_j p_{jo}\int_0^w e^{-(\lambda_j+s)z}\int_0^\infty \sum_i \phi_{i.}(w-z+v,s)B_j(dv)dz$$

$$+ \sum_{k=1}^{\kappa} \lambda_j p_{jk}\int_0^w e^{-(\lambda_j+s)z}\phi_{k.}(w-z,s)dz\} .$$

However, the equations obtained are much more complicated than (4.7),(4.8), and do not seem amenable to further treatment.

APPENDIX

If $f(x)$ is a real or complex function of a real variable, then we define the Laplace transform of f to be

$$f^{o}(s) = \int_{-\infty}^{\infty} e^{-sx} f(x)\,dx$$

wherever this integral converges for s complex. If F is a non-decreasing function of bounded variation on $\mathbb{R}$, then we define the Laplace-Stieltjes transform of F to be

$$F^{*}(s) = \int_{-\infty}^{\infty} e^{-sx} F(dx) .$$

If $F(x) = 0$ for all $x < 0$, then

$$F^{*}(s) = \int_{0-}^{\infty} e^{-sx} F(dx)$$

$$= sF^{o}(s) \quad \text{(from integration by parts).}$$

If F is the d.f. of a non-negative random variable, then $F^{*}(s)$ converges for all s with $\mathrm{Re}(s) \geq 0$. Also $F^{*}(s)$ is analytic in $\{\mathrm{Re}(s) > 0\}$ and continuous to the right at the imaginary axis. We have

LEMMA A1: If F is the d.f. of a non-negative random variable, then $F^{*}(s)$ cannot vanish on the positive real axis.

PROOF: Suppose $F^*(\sigma_o) = 0$ for some $\sigma_o > 0$. This implies $F^*(s) = 0$ for all s with $Re(s) = \sigma > \sigma_o$, since

$$\begin{aligned} |F^*(s)| &\le \int_0^\infty |e^{-sx}|\ F(dx) \\ &\le \int_0^\infty e^{-\sigma_o x}\ F(dx) \\ &\le F^*(\sigma_o)\ . \end{aligned}$$

But this is impossible since $F^*(s)$ has isolated zeros throughout its region of analyticity ($F^*(0) = 1$ implies $F^*(s) \not\equiv 0$). □

LEMMA A2: If $H(x) = K(x) = 0$ for all $x < 0$, where H is a non-decreasing function of bounded variation, and

$$G(x) = \begin{cases} \int_0^x K(x-u)H(du) & (x \ge 0) \\ 0 & (x<0) \end{cases},$$

then

$$G^o(s) = K^o(s)\ H^*(s)\ .$$

If K is also a non-decreasing function of bounded variation, then $G^*(s) = K^*(s)\ H^*(s)$.

PROOF: If we put $x - u = v$, $u = z$ then we have

$$G^o(s) = \int_0^\infty e^{-sx} \int_0^x K(x-u)H(du)\ dx$$

$$= \int_0^\infty \int_0^\infty e^{-s(v+z)}\ K(v)\ dv\ H(du)$$

$$= K^o(s)H^*(s)\ .$$

The second part follows since G is of bounded variation if H and K are [23; § 11], and $G^*(s) = sG^o(s)$. □

LEMMA A3: <u>Suppose f is a continuous complex-valued function on a compact set E of complex numbers such that $|f(z)| < c$ for all $z \in F$, a closed subset of E. For $z \in E$, let $\rho(z,F) = \inf\{|z-y| : y \in F\}$. Then there exists $\delta > 0$ such that</u>

$$\rho(z,F) < \delta \Rightarrow |f(z)| < c\ .$$

PROOF: $|f|$ is continuous on F and therefore attains its supremum in F. Thus there exists $\varepsilon > 0$ such that $|f(x)| \le c - \varepsilon$ for all $z \in F$. By Theorem F in [19], f is uniformly continuous on E, so there exists $\delta > 0$ such that for any $x,y \in E$

$$|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon\ .$$

Suppose $\rho(z,F) < \delta$. Then $|z-y| < \delta$ for some $y \in F$. Hence

$$\begin{aligned}|f(z)| &\le |f(z) - f(y)| + |f(y)| \\ &< \varepsilon + c - \varepsilon \\ &= c\ .\end{aligned}$$ □

LEMMA A4: If $y_1, y_2, \ldots, y_n$ are distinct complex numbers with

$$\sum_{i=1}^{n} c_i e^{ty_i} = 0 \quad \text{for } t = 1,2,\ldots,n ,$$

then $c_i = 0$ for all i.

PROOF: If $c_j \neq 0$ for some j, then

$$\begin{vmatrix} e^{y_1} & e^{y_2} & \ldots & e^{y_n} \\ e^{2y_1} & e^{2y_2} & \ldots & e^{2y_n} \\ \cdot\ \cdot\ \cdot & & & \\ e^{ny_1} & e^{ny_2} & \ldots & e^{ny_n} \end{vmatrix} = 0$$

Therefore some non-zero linear combination of the rows of this determinant is zero, i.e. there exist $d_1, \ldots, d_n$ such that $\sum_{t=1}^{n} d_t e^{ty_i} = 0$ for each $i = 1,2,\ldots, n$. Hence e^{y_i} is a solution of the equation

$$\sum_{t=1}^{n} d_t x^t = 0$$

for each $i = 1,2,\ldots,n$. But this equation has at most n distinct solutions, one of them being $x = 0$, so $y_1, \ldots, y_n$ cannot be distinct. □

LEMMA A5: Suppose $Z_0, Z_1, Z_2, \ldots$ are independent random vectors with $Z_1, Z_2, \ldots$ identically distributed. Let F_k be the σ-field generated by $Z_0, Z_1, \ldots, Z_k$ $(k \geq 1)$, and suppose that N is a stopping time with respect to $\{F_n\}$, i.e. N is a random variable taking positive integral values such that $[N \leq k] \in F_k$ for each

<u>$k \geq 1$. Suppose $\{X_n\}$ is a sequence of random variables such that</u>

<u>$X_k = f(Z_k)$ for some function f. Let $S_n = X_1 + \ldots + X_n$ $(n \geq 1)$.</u>

<u>Then</u>

$$E\{S_N\} = E\{X_1\} \cdot E\{N\}$$

whenever the right-hand side is well-defined (possibly infinite) .

PROOF: This result may be proved using the method of [18], since $z_1,\ldots,z_n$ of remark (e) in [18] may be random vectors and the presence of Z_o does not affect the argument. $\square$

BIBLIOGRAPHY

[1] K. L. Chung (1968). *A Course in Probability Theory.* Harcourt, Brace and World.

[2] B. W. Conolly (1968). The waiting time process for a certain correlated queue. *Operations Research,* 16, 1006-1015.

[3] B. W. Conolly and N. Hadidi (1969). A correlated queue. *Journal of Applied Probability,* 6, 122-136.

[4] D. R. Cox (1955). A use of complex probabilities in the theory of stochastic processes. *Proceedings of the Cambridge Philosophical Society,* 51, 313-319.

[5] D. R. Cox and W. L. Smith (1961). *Queues.* Methuen.

[6] E. J. B. Goursat (1916). *Mathematical Analysis,* 2, Part I (translated by E. R. Hedrick and O. Dunkel). Ginn and Co.

[7] C. M. Harris (1967). Queues with state-dependent stochastic service rates. *Operations Research,* 15, 117-130.

[8] F. I. John (1963). Single server queues with dependent service and inter-arrival times. *Journal of the Society for Industrial and Applied Mathematics,* 11, 526-533.

[9] S. Karlin (1966). *A First Course in Stochastic Processes.* Academic Press.

[10] P. Lancaster (1969). *Theory of Matrices.* Academic Press.

[11] D. V. Lindley (1952). The theory of queues with a single server. *Proceedings of the Cambridge Philosophical Society,* 48, 277-289.

[12] R. M. Loynes (1962). The stability of a queue with non-independent inter-arrival and service times. Ibid, 58, 497-520.

[13] R. M. Loynes (1962). Stationary waiting-time distributions for single-server queues. *Annals of Mathematical Statistics,* 33, 1323-1339.

[14] R. M. Loynes (1962). A continuous-time treatment of certain queues and infinite dams. *Journal of the Australian Mathematical Society*, 2, 484-497.

[15] P.-A. Meyer (1966). *Probability and Potentials*. Blaisdell.

[16] M. F. Neuts (1971). A queue subject to extraneous phase changes. *Advances in Applied Probability*, 3, 78-119.

[17] A. Rényi (1970). *Probability Theory*. North-Holland.

[18] H. Robbins and E. Samuel (1966). An extension of a lemma of Wald. *Journal of Applied Probability*, 3, 272-273.

[19] G. F. Simmons (1963). *Introduction to Topology and Modern Analysis*. McGraw-Hill.

[20] W. L. Smith (1953). On the distribution of queueing times. *Proceedings of the Cambridge Philosophical Society*, 49, 449-461.

[21] W. L. Smith (1972). Lecture notes for Statistics 280 (University of North Carolina - unpublished).

[22] E. C. Titchmarsh (1932). *The Theory of Functions*. Oxford.

[23] D. V. Widder (1963). *The Laplace Transform*. Princeton.