18

w/ loose pages

The Single Server Queue with Poisson Input and Semi-Markov Service Times I.

by

Marcel F. Neuts (*)

Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 18

July 1964

(*) This research was sponsored in part by the Purdue Research Foundation through an XL-award.

Summary

A general queueing model is obtained, when one assumes that the service times of the successive customers form an m-state semi-Markov process. If the interarrival times are independent, identically distributed negative exponential random variables, it is possible to obtain general equations for the time-dependent characteristics of the queue. The known results in the case of independent, identically distributed service times and in the case of a queue with bulk service are obtained as special cases.

The stationary distributions, associated with the queue, are also obtained.

In this first paper the process of successive busy periods and the virtual waiting time process are discussed.

1. Introduction

Considerable attention has been given to the theory of single server queueing processes under the assumptions that the successive interarrival times and the successive service times are independent sequences of independent, identically distributed random variables. We refer to the treatises by L. Takàcs [14] and T. L. Saaty [10] and their impressive lists of references.

Apart from the general discussion by V. E. Beneš [2] and a relatively small number of technical papers, little work has been done on the explicit solution of queueing models with some form of dependence in the input process or in the process of service times.

In this paper, the strong, but convenient, assumption of independent, identically, negative exponentially distributed interarrival times is retained, but the successive service times form an m-state semi-Markov sequence. This can be interpreted as a queueing model with m types of customers for which the successive customer types form an m-state Markov chain. Semi-Markov processes were defined and extensively studied by P. Levy [3], W. L. Smith [11], R. Pyke [6, 7, 8] and others.

In the particular case, when $m = 1$, the semi-Markov process becomes a renewal process and our results become those of L. Takàcs on the transient and stationary behavior of single server queues with Poisson input and recurrent service times. The queueing process in which customers are served in batches of size m - Takàcs [14] - is also a special case of the present model, as we will indicate later.

In order to make this paper relatively self-contained, we will now review some of the terminology and state some of the properties of semi-Markov processes. For a detailed treatment, we refer to R. Pyke [6, 7].

2. Semi-Markov Processes with a finite Number of States.

We consider a double sequence of random variables $\{(J_n, X_n), n = 0,1,\ldots\}$, defined on a complete probability space, having the following properties.

(1) $X_o = o$ a.s, $P\{J_o = k\} = p_k^o$, for $k = 1, \ldots, m < \infty$.

(2) $P\{J_n = k, X_n \leq x \mid X_1, X_2, \ldots, X_{n-1}, J_o, J_1, \ldots, J_{n-1}\} =$

$$P\{J_n = k, X_n \leq x \mid J_{n-1}\} = Q_{J_{n-1}k}(x),$$

for $n = 1, 2,\ldots$

The functions $Q_{ij}(x)$, $(i, j = 1,\ldots, m)$ are mass functions, which are non-decreasing, right-continuous and satisfy the following relations:

(3) $Q_{ij}(x) = o$ for $x \leq o$, $Q_{ij}(+\infty) = p_{ij}$, $\sum_{j=1}^{m} Q_{ij}(\infty) \leq 1$, $(i,j=1,\ldots,m)$

The latter inequality usually is replaced by equality. If strict inequality holds for some i, we call the semi-Markov process ''improper''.

We will refer to the sequence $\{(J_n, X_n), n = o,1,\ldots\}$ as a ''semi-Markov sequence''.

It follows from (2) that the random variables $X_n(n=1,\ldots)$ are conditionally independent, given the random variables $J_n(n=o,1,\ldots)$. Pyke [6].

We denote the matrix $\{Q_{ij}(x)\}$ by $Q(x)$ and the matrix $\{p_{ij}\}$ by P.

In this paper, we will assume throughout that the successive service times of the customers entering the queue can be described by a sequence of pairs of random variables $\{(J_n, X_n), n = o,1,\ldots\}$ forming an m-state semi-Markov sequence with initial probability vector p^o and a matrix of transition probability distributions $Q(x)$, such that $P = Q(\infty)$ is a stochastic matrix. The random variable $X_n(n=1,2,\ldots)$ may be interpreted as the service time of the nth customer and the random variable J_n as the type of the (n+1)st customer. Formula (2) expresses that X_n and the type of the next customer depend on the entire past history of the process only through the type of the nth customer. We note that the semi-Markov sequence is entirely characterized by the vector p^o and by the matrix of distributions $Q(x)$.

In order to avoid inessential difficulties, we henceforth assume that the semi-Markov sequence is irreducible, which means that the matrix P is irreducible and hence that the marginal process $\{J_n, n=o,1,\dots\}$ is a single class Markov chain with m states.

We will also use the term ''general semi-Markov sequence''. This term will refer to a process, defined as above, except that the initial random variable X_1 is described by a different matrix of transition probability distributions $\tilde{Q}(x)$, where

$$\tilde{Q}_{ij}(x) = P\left\{J_1=j,\ X_1 \le x \mid J_o = i\right\}$$

The term ''semi-Markov Process'' is usually reserved for the continuous parameter process $J_t = J_{N(t)}$, in which

$$N(t) = \sup\left\{n: X_o + \dots + X_n \le t\right\},$$

It should be noted that the process J_t, which we will introduce below is not so defined. In order to avoid confusion, we have coined the term ''semi-Markov sequence.''

3. Formal definition of the Queueing Process

Customers arrive at a single server counter in the instants $\tau_1, \tau_2, \dots$ The interarrival times $\tau_n - \tau_{n-1}$ ($\tau_o = o$, $n = 1,2,\dots$) are identically distributed, independent positive random variables with distribution function

$F(x) = P\left\{\tau_n - \tau_{n-1} \leq x\right\}$, given by:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq o, \quad \lambda > o \\ 0 & x < o \end{cases} \tag{4}$$

that is, the input process is a homogeneous Poisson process with parameter λ. If a customer arrives at the counter at an instant when the server is idle, then his service starts immediately. If upon his arrival the server is busy, he joins the queue. Unless otherwise stated, the order in which the customers are served is immaterial.

Let the type of the first customer (J_o) be described by the initial probability vector p^o. The type of the $(n+1)$st customer is J_n and the service time of the nth customer is denoted by X_n. We assume that the double sequence $\left\{(J_n, X_n);\ n=o,1,\ldots\right\}$ is a proper m-state semi-Markov sequence with matrix of transition probability distributions $Q(x)$.

We propose for this queueing process the symbol $F(x)/p^o, Q(x)/1$.

We denote by J_t the type of the last customer to join the queue, before the instant of time t. In order to make this definition meaningful, we define $J_t = J_o$, the type of the first customer, in the interval from $t = o$ to the time of the first arrival.

In this paper, we will show that the sequence of the successive busy periods, together with the type of the first customer of each busy period, is again a general semi-Markov sequence and we will obtain general equations for its transition probability matrices $\tilde{G}(y)$ and $G(y)$.

We will also define the virtual waiting time $\eta(t)$ as the time a customer would have to wait if he joined the queue at the instant t and service is in order of arrival. The bivariate, continuous time process $\left\{J_t, \eta(t);\ t \geq o\right\}$ will be referred to as the ''virtual waiting time process''.

Other processes related to this queueing model, such as the queue length process and the waiting time of the nth customer, will be discussed in a subsequent paper.

4. A Useful Theorem

We introduce the matrix $\Psi(s)$ with entries $\psi_{ij}(s)$ $(i, j=1,\ldots,m)$, defined as the Laplace transforms:

$$\psi_{ij}(s) = \int_0^{\infty} e^{-sx}\, d\, Q_{ij}(x), \qquad \text{Re } s \geq 0 \tag{5}$$

and we denote the m eigenvalues of $\Psi(s)$ by $\eta_\rho(s)$. $(\rho = 1, \ldots, m)$.

Throughout this paper we will make extensive use of the following theorem.

<u>Theorem 1</u>

The equation

$$\det\left[z\, I - w\, \Psi(s+\lambda-\lambda z)\right] = 0 \tag{6}$$

has exactly m roots in the unit circle $|z| < 1$ if either $\text{Re } s \geq 0$ and $|w| < 1$ or $\text{Re } s > 0$ and $|w| \leq 1$.

If, in addition, the m eigenvalues $\eta_\rho(s+\lambda-\lambda z)$, $\rho = 1,\ldots, m$ of the matrix $\Psi(s+\lambda-\lambda z)$ are distinct for all values of s and z, such that Re $s > 0$ and $|z| \leq 1$ or if some collection of eigenvalues are identical for all such values of s and z, while the remaining eigenvalues are distinct, then they can be defined so as to be analytic functions of $s+\lambda-\lambda z$ in this region.

In this case, each of the equations

$$z - w\, \eta_\rho(s+\lambda-\lambda z) = 0 \tag{7}$$

has a unique root $\gamma_\rho(s,w)$ in the open disk $|z| < 1$ for either $\text{Re } s \geq 0$ and $|w| < 1$ or $\text{Re } s > 0$ and $|w| \leq 1$. These roots $\gamma_\rho(s,w)$ are then regular functions of s and w for $\text{Re } s > o$ and $|w| \leq 1$ and continuous on the boundary. Furthermore:

(8) $$\gamma_\rho(s,w) = w \sum_{j=1}^{\infty} \frac{(-\lambda w)^{j-1}}{j!} \left[\frac{d^{j-1}[\eta_\rho(\lambda+s)]^j}{ds^{j-1}} \right],$$

For $s+\lambda-\lambda z$ real, the matrix $\Psi(s+\lambda-\lambda z)$ is a nonnegative matrix. Let us denote by $\eta_m(s+\lambda-\lambda z)$ the root, which is equal to the Perron eigenvalue for real values of $s+\lambda-\lambda z$. For real $s+\lambda-\lambda z$, $\eta_m(s+\lambda-\lambda z)$ is strictly positive. Let us denote by $w = \gamma_m(o,1)$ the smallest positive real root of the equation

(9) $$\omega = \eta_m(\lambda - \lambda\omega),$$

It then follows that $\omega < 1$ if and only if $\lambda\eta_m'(o+) < -1$ and $\omega = 1$ if and only if $\lambda\eta_m'(o+) \geq -1$.

<u>Proof</u>

We first assume that in the region $\text{Re } s > o$, $|z| \leq 1$ the m eigenvalues $\eta_\rho(s+\lambda-\lambda z)$ of the matrix $\Psi(s+\lambda-\lambda z)$ are distinct and defined analytically.

For $\text{Re } s > o$ and $|z| \leq 1$ we have:

(10) $$|\psi_{ij}(s+\lambda-\lambda z)| \leq p_{ij}, \qquad i,j=1,\ldots, m$$

with equality holding only if $p_{ij} = o$. This implies that the spectral radius of the matrix $\Psi(s+\lambda-\lambda z)$ is strictly less than that of the matrix P, which is a stochastic matrix and hence has its spectral radius equal to one - Wielandt [15], Rosenblatt [9]. It follows that:

(11) $$|\eta_\rho(s+\lambda-\lambda z)| < 1, \qquad \rho = 1,\ldots, m \qquad \text{Re } s > o, \quad |z| \leq 1$$

Equation (6) may now be written as:

$$(12) \qquad \prod_{\rho=1}^{m} \left[z - w\eta_\rho(s+\lambda-\lambda z)\right] = o$$

Rouché's theorem may now be applied to each of the factors, since $|w| \leq 1$, to yield that each one of them has one and only one root in the open unit disk.

In the region $\text{Re } s \geq o$, $|z| \leq 1$ the weaker inequality holds in (11), but $|w| < 1$ and Rouché's theorem again yield the desired result.

The analyticity of the root $\gamma_\rho(s,w)$, $\rho = 1,\ldots, m$ follows from the analyticity of the eigenvalue $\eta_\rho(s+\lambda-\lambda z)$ as does the Lagrange expansion, given in formula (8) Takacs [14].

If we set $s = o$ and x real $(o \leq x)$ the matrix $\Psi(\lambda-\lambda x)$ is nonnegative and irreducible and has monotone increasing entries on $o \leq x < \infty$. The largest eigenvalue (the Perron root) is simple and is a monotone increasing convex function of x. - Miller [5], Bellman [2]. Since $\eta_m(o+) = 1$ the equation (9) has a unique positive root $\omega < 1$ in $[o,1]$ if and only if $\lambda\eta_m'(o+) < -1$. The unique zero will be equal to one if and only if $\lambda\eta_m'(o+) \geq -1$. In addition the equation $z = \eta_m(\lambda-\lambda z)$ has only one root in the open disk if $\lambda\eta_m'(o+) < -1$ and has no root in the open disk in the other case. It follows that $\omega = \gamma_m(o+,1)$.

If the matrix $\Psi(s+\lambda-\lambda z)$ has multiple eigenvalues at all points of the region $\text{Re } s \geq o$, $|z| \leq 1$, while the remaining eigenvalues are distinct, then the above argument remains valid. We will assume later that in this case the matrix $\Psi(s+\lambda-\lambda z)$ is diagonalizable.

If the matrix $\Psi(s+\lambda-\lambda z)$ has multiple eigenvalues at some points in the region $\text{Re } s \geq o$, $|z| \leq 1$, then it is no longer possible to define the eigenvalues analytically, since branch-points are present. We will exclude this case later, but the result on the number of roots in the unit disk remains valid as is seen by a perturbation argument and an application of Hurwitz's theorem. One may also apply the argument of R.M. Loynes [4], who uses an extended version of Rouché's theorem.

Remark

If we write $\gamma_m(s)$ for $\gamma_m(s,1)$, then it follows from

$$\gamma_m(s) = \eta_m[s+\lambda-\lambda\gamma_m(s)]$$

that if $\omega = 1$, we have

$$(13) \qquad \gamma_m'(o+) = \eta_m'(o+)\left[1+\lambda\eta_m'(o+)\right]^{-1} \qquad \text{for } \lambda\eta_m'(o+) > -1$$

$$= \infty \qquad \text{for } \lambda\eta_m'(o+) = -1.$$

The following theorem throws light on the meaning of the number $\eta_m'(o+)$.

Theorem 2

Let $\bar{\pi} = (\pi_1, \pi_2, \ldots, \pi_m)$ be the stationary transition probability distribution for the Markov chain with matrix P and let ξ_{ij} be defined as:

$$(14) \qquad \xi_{ij} = \int_o^\infty x \, d\, Q_{ij}(x), \qquad i,j = 1,\ldots, m$$

and $\zeta_i = \sum_{j=1}^{m} \xi_{ij}$. We assume $\zeta_i < \infty$ for $i = 1,\ldots, m$ then

$$(15) \qquad \eta_m'(o+) = -\sum_{j=1}^{m} \pi_j \zeta_j$$

Moreover, let S_n denote the sum of the service times of the first n customers, then:

$$(16) \qquad \frac{S_n}{n} \xrightarrow[n]{} \sum_{j=1}^{m} \pi_j \zeta_j \quad \text{a.s.}$$

Proof

The distribution $\bar{\pi}$ is the unique probability vector for which

$$\bar{\pi} P = \bar{\pi},$$

Let $\bar{\alpha}(s)$ denote the left eigenvector of $\Psi(s)$, corresponding to $\eta_m(s)$, normalized in such a manner that $\bar{\alpha}(o+) = \bar{\pi}$ then

$$\bar{\alpha}(s)\ \Psi(s) = \eta_m(s)\ \bar{\alpha}(s) \tag{17}$$

Differentiating in (17) and setting $s = o+$, we obtain:

$$\sum_{j=1}^{m} \alpha_j'(o+)\ p_{ji} - \sum_{j=1}^{m} \pi_j \xi_{ji} = \alpha_i'(o+) + \eta_m'(o+)\ \pi_i,$$

It follows from the finiteness of the ξ_{ij} that the numbers $\alpha_j'(o+)$ are also finite. Summing on i in the above equations yields

$$\eta_m'(o+) = - \sum_{j=1}^{m} \pi_j \zeta_j$$

The statement (16) is one of the many versions of the strong law of large numbers for positive recurrent semi-Markov process and a detailed proof is completely analogous to those of Pyke [8].

Remark

The quantity $\sum_{j=1}^{m} \pi_j \zeta_j$ may be interpreted as the long range average service time of a customer. We will find results on the stationary behavior of the queue, that are completely analogous to those for the queue with renewal service times. The relative magnitude of $\frac{1}{\lambda}$ and $\sum_{j=1}^{m} \pi_j \zeta_j$ will completely characterize whether the queue becomes stationary, becomes transient or is in unstable equilibrium.

We note that the condition $\zeta_i < \infty$ $(i = 1,\ldots, m)$ is a necessary and sufficient condition for the positive recurrence of the semi-Markov process of service times - Pyke [6].

5. The Process of Successive Busy Periods

Let $Y_1, Y_2,\ldots, Y_n,\ldots$ denote the lengths of the successive intervals of time, during which the server is busy. Let I_o denote the type of the first customer to join the queue $(I_o = J_o)$ and let I_n be the type of the first customer to be served during the $(n + 1)$st busy period. If we define $Y_o = o$ a.s., then it follows from the assumptions imposed on the queue, that the bivariate sequence of random variables $\{(I_n, Y_n), n = o,1,\ldots\}$ is a general semi-Markov sequence. If the virtual waiting time at $t = o$ is zero, then the sequence is an ordinary semi-Markov sequence and the process of busy periods is completely characterized by the vector p^o of initial probabilities and by the matrix $G(x) = \{G_{ij}(x)\}$ $(i,j,=1,\ldots, m)$ of transition probability distributions.

If the virtual waiting time is different from zero, the transition probability matrix, corresponding to Y_1, will be different from $G(x)$. In this case we denote this initial matrix by $\tilde{G}(x)$ and then the semi-Markov sequence of successive busy periods is fully characterized by p^o, $\tilde{G}(x)$ and $G(x)$.

We will now express the matrices $\tilde{G}(x)$ and $G(x)$ in terms of the input rate λ and the matrix $Q(x)$. We will also obtain the joint distributions of the lengths of the busy periods and the number of customers served during each busy period. It is clear that these distributions will not depend on the order in which the customers are served so long as service starts with the first customer to arrive after an idle period.

We assume throughout this discussion that the matrix $\Psi(s+\lambda-\lambda z)$ has m distinct eigenvalues $\eta_\rho(s+\lambda-\lambda z)$, which are not identically zero, throughout the domain $\text{Re } s > o$, $|z| \leq 1$ ($\rho = 1,\ldots, m$). This and other non-degeneracy assumptions will be discussed below.

We introduce the following notations:

$G_{ij}(n,k;x)$ denotes the probability that a busy period, other than the first, consist of at least n services, that the total service time of the first n customers is at most x, that at the end of the nth service k customers are waiting and that the next customer to arrive after the end of the nth service is of type j, given that the first customer of the busy period was of type i.

We introduce the Laplace transform:

$$\Gamma_{ij}(n,k;s) = \int_o^\infty e^{-sx}\, d\, G_{ij}(n,k;x), \quad n \geq 1, \text{ Re } s \geq o \tag{18}$$

and the generating functions:

$$C_{ij}(n,z,s) = \sum_{k=o}^{\infty} \Gamma_{ij}(n,k;s)z^k, \quad |z| \leq 1. \tag{19}$$

$$D_{ij}(w,z,s) = \sum_{n=1}^{\infty} C_{ij}(n,s,z)w^n \quad |w| \leq 1 \tag{20}$$

We set $\Gamma_{ij}(n,o,s) = \Gamma_{ij}(n,s)$ and define the generating function $E_{ij}(w,s)$ as:

$$E_{ij}(w,s) = \sum_{n=1}^{\infty} \Gamma_{ij}(n,s)w^n, \quad |w| \leq 1 \tag{21}$$

Clearly $E_{ij}(1,s)$ is the Laplace-Stieltjes transform of the distribution $G_{ij}(x)$.

Finally, let the vectors $\bar{\alpha}_\rho(s+\lambda-\lambda z) = [\alpha_{1\rho}, \alpha_{2\rho}, \ldots, \alpha_{m\rho}]$ denote the right eigenvectors of the matrix $\Psi(s+\lambda-\lambda z)$ corresponding to the m eigenvalues $\eta_\rho(s+\lambda-\lambda z)$ ($\rho=1,\ldots, m$). We then have the following theorem:

Theorem 3

Provided the $m \times m$ matrix T with columns $\bar{\alpha}_\rho[s+\lambda-\lambda\gamma_\rho(s,w)]$ $(\rho=1,\ldots, m)$ is non-singular for all s and w in the region $\mathrm{Re}\ s > o$, $|w| \leq 1$ the matrix $E(w,s)$ with entries $E_{ij}(w,s)$ is uniquely determined by the fact that its m eigenvalues are the roots $\gamma_\rho(s,w)$ of equation (6) and the corresponding right eigenvectors are the m vectors $\bar{\alpha}_\rho[s+\lambda-\lambda\gamma_\rho(s,w)]$. $(\rho=1,\ldots,m)$

Proof

The probabilities $G_{ij}(n,k;x)$ satisfy the following recurrence relations:

(22)
$$G_{ij}(1,k;x) = \int_o^x e^{-\lambda y} \frac{(\lambda y)^k}{k!} dQ_{ij}(y),$$

and

$$G_{ij}(n,k;x) = \sum_{\nu=1}^{m} \sum_{r=1}^{k+1} \int_o^x G_{i\nu}(n-1,r;x-y)e^{-\lambda y} \frac{(\lambda y)^{k-r+1}}{(k-r+1)!} dQ_{\nu j}(y),$$

for $n > 1$.

Taking Laplace-Stieltjes transforms in formulae (22) we obtain:

(23)
$$\Gamma_{ij}(1,k,s) = \int_o^\infty e^{-(\lambda+s)y} \frac{(\lambda y)^k}{k!} d\, Q_{ij}(y)$$

and:

$$\Gamma_{ij}(n,k,s) = \sum_{r=1}^{k+1} \sum_{\nu=1}^{m} \Gamma_{i\nu}(n-1,r,s) \int_o^\infty e^{-(\lambda+s)y} \frac{(\lambda y)^{k-r+1}}{(k-r+1)!} d\, Q_{\nu j}(y)$$

for $n > 1$ and hence:

(24)
$$C_{ij}(1,z,s) = \psi_{ij}(s+\lambda-\lambda z)$$

and

$$z\, C_{ij}(n,z,s) = \sum_{\nu=1}^{m} C_{i\nu}(n-1,z,s)\ \psi_{\nu j}(s+\lambda-\lambda z) - \sum_{\nu=1}^{m} \Gamma_{i\nu}(n-1,s)\ \psi_{\nu j}(s+\lambda-\lambda z)$$

for $n > 1$.

Finally, we obtain:

(25) $$z\, D_{ij}(w,z,s) = w \sum_{\nu=1}^{m} D_{i\nu}(w,z,s)\ \psi_{\nu j}(s+\lambda-\lambda z)$$

$$+wz\ \psi_{ij}(s+\lambda-\lambda z) - w \sum_{\nu=1}^{m} E_{i\nu}(w,s)\ \psi_{\nu j}(s+\lambda-\lambda z)$$

or in matrix notation:

(26) $$D(w,z,s)\ \left[z\ I-w\ \psi(s+\lambda-\lambda z)\right] = w\left[z\ I-E(w,s)\right]\ \Psi(s+\lambda-\lambda z)$$

The inverse of the matrix $z\ I-w\ \Psi(s+\lambda-\lambda z)$ exists for all z, w and s in the region $|z| \leq 1$, $|w| \leq 1$ and $\text{Re } s > o$, except for the roots:

$$z = \gamma_\rho(s,w) \qquad \rho = 1,\dots, m$$

Under the non-degeneracy assumptions the matrix $\Psi(s+\lambda-\lambda z)$ may be written as:

(27) $$\Psi(s+\lambda-\lambda z) = R(s+\lambda-\lambda z)\ H(s+\lambda-\lambda z)\ R^{-1}(s+\lambda-\lambda z),$$

in which $H_{ij}(s+\lambda-\lambda z) = \delta_{ij}\ \eta_i(s+\lambda-\lambda z)$ and in which the columns of $R(s+\lambda-\lambda z)$ are the right eigenvectors of the matrix $\Psi(s+\lambda-\lambda z)$. The components of the eigenvectors may be defined to be analytic functions of $s+\lambda-\lambda z$ in the region $\text{Re } s \geq o$, $|z| \leq 1$.

If we set

$$R_{ij}(s+\lambda-\lambda z) = \alpha_{ij}(s+\lambda-\lambda z)$$

$$(R^{-1})_{ij}(s+\lambda-\lambda z) = \beta_{ij}(s+\lambda-\lambda z)$$

then (26) and (27) lead to:

$$(28) \quad D_{ij}(w,z,s) = w \sum_{\nu=1}^{m} \left[z\delta_{i\nu} - E_{i\nu}(w,s)\right] \sum_{\rho=1}^{m} \frac{\alpha_{\nu\rho}(s+\lambda-\lambda z)\ \eta_\rho(s+\lambda-\lambda z)\ \beta_{\rho j}(s+\lambda-\lambda z)}{z - w\ \eta_\rho(s+\lambda-\lambda z)}$$

for all $z \neq \gamma_\rho(s,w)$ $(\rho = 1,\ldots, m)$

The functions $D_{ij}(w,z,s)$ must be analytic for all $|w| \leq 1$, Re $s > o$, $|z| \leq 1$. This implies that the zeros of the denominators must also be zeros of the numerators and hence for all $\rho = 1,\ldots, m$ we must have:

$$(29) \quad \sum_{\nu=1}^{m} \left[\delta_{i\nu}\gamma_\rho(s,w) - E_{i\nu}(w,s)\right] \alpha_{\nu\rho}\left[s+\lambda-\lambda\ \gamma_\rho(s,w)\right]$$

$$\cdot\ \eta_\rho\left[s+\lambda-\lambda\ \gamma_\rho(s,w)\right]\ \beta_{\rho j}\left[s+\lambda-\lambda\ \gamma_\rho(s,w)\right] = o$$

for all $i, j = 1,\ldots, m$.

By assumption $\eta_\rho(s+\lambda-\lambda\ \gamma_\rho)$ does not vanish identically and $\beta_{\rho j}\left[s+\lambda-\lambda\gamma_\rho(s,w)\right]$ is different from zero for at least one j, so we obtain:

$$(30) \quad \sum_{\nu=1}^{m} E_{i\nu}(w,s)\ \alpha_{\nu\rho}\left[s+\lambda-\lambda\gamma_\rho(s,w)\right] = \gamma_\rho(s,w)\ \alpha_{i\rho}\left[s+\lambda-\lambda\ \gamma_\rho(s,w)\right]$$

for all $i, \rho = 1,\ldots, m$.

Since we have assumed that the matrix $T(w,s)$ is nonsingular the system (30) can be solved uniquely for the m^2 functions $E_{ij}(w,s)$, which are then analytic in the region Re $s \geq o$, $|w| \leq 1$.

Equation (30) leads to:

$$E(w,s) = T(w,s)\ \Delta(w,s)\ T^{-1}(w,s) \tag{31}$$

in which $\Delta_{ij}(w,s) = \delta_{ij}\ \gamma_i(s,w)$.

This completes the proof of the theorem.

Remarks

a) The assumption that the matrix $T(w,s)$ is nonsingular is not essential for this theorem, since the entries of the matrix $E(w,s)$, given by (31) will only have removable singularities at points where $\det T(w,s)$ vanishes. This condition seems to be necessary however to guarantee the uniqueness and the analyticity of the solutions for the analogous equations for the matrix $\tilde{G}(x)$.

The case of multiple eigenvalues for the matrix $\Psi(s+\lambda-\lambda z)$ seems to be intractable in general. In some cases (Q_{ij} negative exponential, $m = 2$) it was possible to verify that this cannot occur except in degenerate cases. As an illustration of such a degeneracy, we mention the cases in which $Q_{ij}(x) = p_j Q_j(x)$ or $Q_{ij}(x) = p_j Q_i(x)$. In these cases, all but one of the eigenvalues are identically equal to zero, but it is easy to see that these queueing models may be solved directly by considering only one customer type, whose service time distribution is a mixture of m distributions.

b) This argument generalizes one of the many derivations of Takàcs' functional equation for the case $m = 1$. Takacs [12]. The more direct combinatorial arguments, which are available in the case $m = 1$, appear not to extend easily to our model [13].

Theorem 4

The semi-Markov sequence of successive service times is irreducible. It is proper if and only if $\lambda\eta_m'(0+) \geq -1$. It is positive recurrent if and only if $\lambda\eta_m'(0+) > -1$. In this case the means

$$M_{ij} = \int_0^{\infty} x \; d \; G_{ij}(x)$$

are finite.

Proof:

In order to show irreducibility it is sufficient to prove that there is positive probability of a transition from state i into state j in the semi-Markov sequence of successive busy periods, for any two states i and j. Since the Markov chain P is irreducible, there is an integer ν for which $p_{ij}^{(\nu)} > o$.

Assume first that the service time of the customer of type i is different from zero with positive probability. Then there is positive probability that during the service time of the customer of type i exactly $\nu-1$ customers arrive such that during their accumulated service time no further customers arrive and such that the next customer thereafter is of type j.

If the service time of the customer of type i is zero almost surely, then there is either a sequence of busy periods of zero length, which leads from i to j in ν steps or from i a type of customer with non-zero service time is reached. The argument given above then shows that there is positive probability of reaching j in the remaining steps.

If we consider equation (30) with $\rho = m$, $w = 1$ and let s tend to $o+$, we obtain by continuity:

$$(33) \qquad \sum_{\nu=1}^{m} G_{i\nu}(\infty) \; \alpha_{\nu m}(\lambda-\lambda\omega) = \omega\alpha_{im}(\lambda-\lambda\omega), \qquad (i = 1,\dots, m)$$

The vector $\bar{\alpha}_m(\lambda-\lambda\omega)$ is the eigenvector corresponding to the largest eigenvalue of the matrix $\Psi(\lambda-\lambda\omega)$, which is a nonnegative matrix. The components of $\bar{\alpha}_m(\lambda-\lambda\omega)$ can therefore be chosen to be strictly positive. Equation (33) expresses that $\bar{\alpha}_m(\lambda-\lambda\omega)$ is also the eigenvector of the matrix $G(\infty)$ corresponding to the eigenvalue ω. Since $G(\infty)$ is irreducible and since ω is a positive number it must be the Perron eigenvalue of $G(\infty)$, since the Perron eigenvalue of an irreducible nonnegative matrix is the only eigenvalue that can have a strictly positive eigenvector. Therefore $G(\infty)$ can be a stochastic matrix only if $\omega = 1$. Conversely if $\omega = 1$ we have $\bar{\alpha}_m(\lambda-\lambda\omega) = \bar{\alpha}_m(o+)$, where $\bar{\alpha}_m(o+)$ is a right eigenvector corresponding to the eigenvalue one of the matrix P. Hence all components of $\bar{\alpha}_m(o+)$ are equal and it follows that

$$\sum_{\nu=1}^{m} G_{i\nu}(\infty) = 1, \qquad (i = 1,\dots, m)$$

Finally we know that $\omega = 1$ if and only if $\lambda\eta_m'(o+) \geq -1$.

Let us normalize the vector $\bar{\alpha}_m(o+)$ by setting $\alpha_{im}(o+) = 1$ for all $i = 1,\dots, m$. It is easy to verify that in this case the finiteness of the ζ_j $(j = 1,\dots, m)$ implies that $\alpha_{im}'(o+)$ is finite for all i.

If we differentiate with respect to s in equation (30) and take $\rho = m$, $w = 1$ and $s = o+$, we obtain:

$$(34) \qquad \sum_{\nu=1}^{m} M_{i\nu} = -\left[1+\lambda\eta_m'(o+)\right]^{-1}\left[\alpha_{im}'(o+) - \sum_{\nu=1}^{m} G_{i\nu}(\infty)\, \alpha_{\nu m}'(o+) + \eta_m'(o+)\right]$$

The result follows if we show that the denominator and all the numerators in (34) cannot vanish simultaneously. The denominator vanishes if and only if $\eta_m'(o+) = -\lambda^{-1} < o$. Let us denote the vector $\bar{\alpha}_m'(o+)$ by $\bar{\delta}$, then if all the numerators in (34) vanish simultaneously we have the equation

$$(35) \qquad \left[I - G(\infty)\right] \bar{\delta} = -\eta_m'(o+)\, \bar{1}.$$

where $\bar{1}$ is the columnvector with all its components equal to one. Let $\bar{\pi}$ denote the stationary probability distribution vector corresponding to the stochastic matrix $G(\infty)$. We note that $\bar{\pi}$ is the same vector as for the matrix P. Upon taking the scalar product of the rowvector $\bar{\pi}$ and the columnvector in (35) we obtain:

$$o = \bar{\pi}\left[I - G(\infty)\right] \bar{\delta} = -\eta_m'(o+) \; \bar{\pi} \, \bar{1} = -\eta_m'(o+) = \frac{1}{\lambda} \neq o$$

which is impossible.

So if $\lambda\eta_m'(o+) > -1$ all M_{ij} are finite $(i,j = 1,\dots, m)$ and it follows by a theorem of Pyke [6] that this is equivalent to positive recurrence of the semi-Markov process of busy periods. When $\lambda\eta_m'(o+) = -1$ at least one of the M_{ij} must be infinite and hence the semi-Markov process is null-recurrent.

We now proceed with the derivation of the equations for the matrix $\tilde{G}(x)$. This matrix will depend on the random variable $\eta(o)$, the virtual waiting time at $t = o$. If $\eta(o) = o$ a.s. then the matrix $\tilde{G}(x)$ may be set equal to the identity matrix. We recall that J_o is the type of the first customer to join the queue and that $J_t = J_o$ for values of t up to the first transition. In many applications $\eta(o)$ and J_o will be dependent on each other, so we introduce the conditional distribution

$$(36) \qquad \tilde{W}_i(x) = P\left\{\eta(o) \leq x \mid J_o = i\right\}, \quad (i=1,\dots, m)$$

There are several ways to define the initial busy period but we will agree to say that a transition occurs from state i into state i in the case where no customers arrive during the interval $(o,\eta(o))$. The initial service, before the arrival of the first customer is counted as one service in the definition of $\tilde{G}_{ij}(n,k;x)$. The definitions of $\tilde{G}_{ij}(n,k;x)$ and of its transforms are then the same as for $G_{ij}(n,k;x)$. We only write down these recursion relations which are formally different from those for $G_{ij}(n,k;x)$. We have

(37) $$\hat{G}'_{ij}(1,k,x) = \delta_{ij} \int_0^\infty e^{-\lambda y} \frac{(\lambda y)^k}{k!} d\,\tilde{W}_i(y), \quad k = 0,1,\ldots$$

If we introduce the Laplace transform

(38) $$\int_0^\infty e^{-sx}\, d\,\tilde{W}_i(x) = \tilde{\Omega}_i(s)$$

then

(39) $$\tilde{C}_{ij}(1,z,s) = \delta_{ij}\,\tilde{\Omega}_i(s+\lambda-\lambda z),$$

$$\tilde{E}_{ij}(w,s) = \sum_{n=1}^{\infty} \tilde{\Gamma}_{ij}(n,s)\, w^n,$$

and

(40) $$z\,\tilde{D}_{ij}(w,z,s) = w \sum_{\nu=1}^{m} \tilde{D}_{i\nu}(w,z,s)\;\psi_{\nu j}(s+\lambda-\lambda z) +$$

$$\delta_{ij}\; wz\,\tilde{\Omega}_i(s+\lambda-\lambda z) - w \sum_{\nu=1}^{m} \tilde{E}_{i\nu}(w,s)\;\psi_{\nu j}(s+\lambda-\lambda z),$$

for $i,\ j = 1,\ldots,\ m$.

In matrix form we get:

(41) $$\tilde{D}(w,z,s)\left[z\,I - w\,\Psi(s+\lambda-\lambda z)\right] = w\left[z\,\Delta_{\tilde{\Omega}}(s+\lambda-\lambda z) - \tilde{E}(w,s)\;\Psi(s+\lambda-\lambda z)\right]$$

in which

$$(42) \qquad \Delta_{\widetilde{\Omega}}(s+\lambda-\lambda z) = \operatorname{diag}\left\{\widetilde{\Omega}_1(s+\lambda-\lambda z),\ldots,\widetilde{\Omega}_m(s+\lambda-\lambda z)\right\},$$

Reasoning as before, and under the same nondegeneracy assumptions, we get:

$$(43) \qquad \sum_{\nu=1}^{m} \widetilde{E}_{i\nu}(w,s)\,\alpha_{\nu\rho}(s+\lambda-\lambda\gamma_\rho) = w\,\widetilde{\Omega}_i\left[s+\lambda-\lambda\gamma_\rho(s,w)\right]\,\alpha_{i\rho}\left[s+\lambda-\lambda\gamma_\rho(s,w)\right]$$

for $i, \rho = 1,\ldots, m$.

Note that formula (43) yields $\widetilde{E}_{ij}(w,s) = \delta_{ij}\,w$ in case $\eta(o) = o$ a.s. As before we find that $\widetilde{G}(\infty)$ is a stochastic matrix when $\omega = 1$.

We summarize these findings in the following theorem.

Theorem 5

Provided the $m \times m$ matrix $T(w,s)$ is nonsingular for all s and w with $\operatorname{Re} s > o$ and $|w| \leq 1$, the matrix $\widetilde{E}(w,s)$ is uniquely determined by the system of equations (43). If $\omega = 1$, the matrix $\widetilde{G}(\infty) = \widetilde{E}(1, o+)$ is stochastic.

6. The Virtual Waiting time process

Consider the bivariate continuous parameter process $\left\{ J_t,\ \eta(t): t \geq o \right\}$, in which J_t is the type of the last customer to join the queue before the time instant t and in which $\eta(t)$ is the virtual waiting time at time t. Clearly the process $\left\{ J_t,\ \eta(t): t \geq o \right\}$ is a Markov process and its path functions can be described as follows: In the instants $\tau_n (n = 1,2,\ldots)$ the value of $\eta(t)$ has a jump of magnitude X_n, where X_n is the service time of the nth customer. Also in these instants the process J_t, which takes values in the set $\left\{ 1,2,\ldots, m \right\}$ changes value. At all other points the value of $\eta(t)$ is either zero or decreases linearly with slope -1 until it jumps or reaches zero. If at the instant t, $\eta(t)$ is zero, it remains zero until the next customer arrives in the queue.

We introduce the following notations:

$$(44) \quad W_{ij}(t,x) = P\left\{ \eta(t) \leq x,\ J_t = j \mid J_o = i \right\}, \qquad (i,j = 1,\dots, m)$$

$$(45) \quad W_{ij}(o,x) = \delta_{ij}\, \widetilde{W}_i(x), \qquad (i,j = 1,\dots, m)$$

$$(46) \quad \Omega_{ij}(t,s) = \int_o^\infty e^{-sx}\, d_x W_{ij}(t,x), \quad \text{Re } s \geq o \quad i,j = 1,\dots, m$$

$$(47) \quad \widetilde{\Omega}_i(s) = \int_o^\infty e^{-sx}\, d\, \widetilde{W}_i(x), \quad \text{Re } s \geq o \quad i = 1,\dots, m$$

$$(48) \quad \Omega^*_{ij}(s,\zeta) = \int_o^\infty \Omega_{ij}(t,\zeta)\, e^{-st}\, dt, \quad \text{Re } s > o, \quad \text{Re } \zeta \geq o$$

$$(49) \quad W^*_{ij}(s) = \int_o^\infty W_{ij}(t,o)\, e^{-st}\, dt, \quad \text{Re } s > o, \quad i,j = 1,\dots, m$$

as well as the matrices

$$\Omega^*(s,\zeta) = \left(\Omega^*_{ij}(s,\zeta)\right) \qquad W^*(s) = \left(W^*_{ij}(s)\right)$$

We have the following theorem.

<u>Theorem 6</u>

The Laplace-Stieltjes transforms of the virtual waiting-time distributions are given by the following formulae:

If $\Omega(t,s) = \left(\Omega_{ij}(t,s)\right)$ then

$$(50) \quad \Omega(t,s) = \left[\Delta_{\widetilde{\Omega}}(s) - \int_o^t sW(u,o)e^{-u\left[(s-\lambda)I+\lambda\Psi(s)\right]}du\right]e^{t\left[(s-\lambda)I+\lambda\Psi(s)\right]}$$

in which the matrix $W(u,o)$ remains to be determined and the matrix $\Delta_{\tilde{\Omega}}$ is defined in formula (42).

Provided the matrix of coefficients does not vanish anywhere in $\operatorname{Re} s > o$, the Laplace transforms of the entries of the matrix $W(u,o)$ are given by the unique solutions to the system of linear equations

$$(51) \qquad \sum_{\nu=1}^{m} W^*_{i\nu}(s)\, \alpha_{\nu\rho}(\zeta_\rho(s)) = \frac{\tilde{\Omega}_i(\zeta_\rho(s))}{\zeta_\rho(s)}\, \alpha_{i\rho}(\zeta_\rho(s))$$

$(i, \rho = 1, \dots, m)$ in which the numbers

$$(52) \qquad \zeta_\rho(s) = s + \lambda\left[1 - \gamma_\rho(s,1)\right] \quad (\rho = 1, \dots, m)$$

are the roots with smallest absolute values of the equations

$$(53) \qquad \zeta = s + \lambda - \lambda\, \eta_\rho(\zeta)$$

Proof:

Our proof parallels Takàcs' proof for the case $m = 1$. Let $\delta_{\Delta t}$ denote the number of customers, joining the queue during the time interval $(t, t+\Delta t|$. Then by assumption we have

$$(54) \qquad P\left\{\delta_{\Delta t} = k\right\} = e^{-\lambda\Delta t}\, \frac{(\lambda\Delta t)^k}{k!}, \quad (k = o, 1, \dots)$$

Using the theorem of total probability, we obtain:

$$(55) \quad \Omega_{ij}(t+\Delta t,s) = \int_o^\infty e^{-sx} d_x\, P\left\{\eta(t+\Delta t \le x,\ J_{t+\Delta t} = j \,|\, J_o = i\right\}$$

$$= \sum_{k=o}^{\infty} P\left\{\delta_{\Delta t} = k\right\} \int_o^\infty e^{-sx} d_x\, P\left\{\eta(t+\Delta t) \le x,\ J_{t+\Delta t} = j \,|\, J_o = i,\ \delta_{\Delta t} = k\right\}$$

$$= (1-\lambda\Delta t) \int_o^\infty e^{-sx} d_x\, P\left\{\eta(t+\Delta t) \le x,\ J_{t+\Delta t} = j \,|\, J_o = i,\ \delta_{\Delta t} = o\right\}$$

$$+ \lambda\Delta t \int_o^\infty e^{-sx} d_x\, P\left\{\eta(t+\Delta t) \le x,\ J_{t+\Delta t} = j \,|\, J_o = i,\ \delta_{\Delta t} = 1\right\} + o(\Delta t),$$

We evaluate the two integrals separately:

$$(56) \quad \int_o^\infty e^{-sx} d_x\, P\left\{\eta(t+\Delta t) \le x,\ J_{t+\Delta t} = j \,|\, J_o = i,\ \delta_{\Delta t} = o\right\} =$$

$$W_{ij}(t,\Delta t) + e^{s\Delta t} \int_{\Delta t}^\infty e^{-sx} d_x\, W_{ij}(t,x)\ ,$$

Since $W_{ij}(t,x)$ is right-continuous in x, we have:

$$W_{ij}(t,\Delta t) = W_{ij}(t,o) + O(\Delta t)$$

and

$$o \le \int_o^{\Delta t} x\, d_x\, W_{ij}(t,x) \le \Delta t\left[W_{ij}(t,\Delta t) - W_{ij}(t,o)\right] = o(\Delta t)$$

It follows that:

$$(57) \quad \int_0^\infty e^{-sx} d_x\, P\left\{\eta(t+\Delta t) \leq x,\ J_{t+\Delta t} = j \,|\, J_o = o,\ \delta_{\Delta t} = o\right\}$$

$$= (1+s\Delta t)\ \Omega_{ij}(t,s) - s\Delta t\ W_{ij}(t,o) + o(\Delta t)$$

The second integral can be written as:

$$(58) \quad \int_0^\infty e^{-sx} d_x\, P\left\{\eta(t+\Delta t) \leq x,\ J_{t+\Delta t} = j \,|\, J_o = i,\ \delta_{\Delta t} = 1\right\} =$$

$$\sum_{\nu=1}^{m} \int_0^\infty d_y\, W_{i\nu}(t,y) \int_0^\infty e^{-sx} d_x\, P\left\{\eta(t+\Delta t) \leq x,\ J_{t+\Delta t} = j \,|\, J_o = i,\right.$$

$$\left.\delta_{\Delta t} = 1,\ J_t = \nu,\ \eta(t) = y\right\}$$

$$= \sum_{\nu=1}^{m} \int_{\Delta t}^\infty e^{-s(y-\Delta t)} d_y\, W_{i\nu}(t,y) \int_0^\infty e^{-sz} d\, Q_{\nu j}(z) +$$

$$\sum_{\nu=1}^{m} e^{-s\epsilon_t \Delta t} \int_0^{\Delta t} d_y\, W_{i\nu}(t,y) \int_0^\infty e^{-sz} d\, Q_{\nu j}(z)$$

with $o \leq \epsilon_t \leq 1$.

$$= \sum_{\nu=1}^{m} (1+s\Delta t) \left[\Omega_{i\nu}(t,s) - \int_0^{\Delta t} (1-sy) d_y\, W_{i\nu}(t,y)\right] \psi_{\nu j}(s) +$$

$$+ \sum_{\nu=1}^{m} (1-s\epsilon_t \Delta t)\ W_{i\nu}(t,\Delta t)\ \psi_{\nu j}(s),$$

Summing up in formula (55), we obtain:

$$(59) \quad \Omega_{ij}(t+\Delta t,s) = \Omega_{ij}(t,s) + s\Delta t\left[\Omega_{ij}(t,s) - W_{ij}(t,o)\right]$$

$$- \lambda\Delta t\ \Omega_{ij}(t,s) + \lambda\Delta t \sum_{\nu=1}^{m} \Omega_{i\nu}(t,s)\ \psi_{\nu j}(s) + o(\Delta t)$$

As $\Delta t \to o$, we obtain:

$$(60) \quad \frac{\partial}{\partial t}\Omega_{ij}(t,s) = (s-\lambda)\ \Omega_{ij}(t,s) + \lambda \sum_{\nu=1}^{m} \Omega_{i\nu}(t,s)\ \psi_{\nu j}(s) - s\ W_{ij}(t,o),$$

$$(i,j = 1,\ldots, m)$$

Equation (60) can be written as a matrix differential equation

$$(61) \quad \frac{\partial}{\partial t}\Omega(t,s) = \Omega(t,s)\left[(s-\lambda)I + \lambda\ \Psi(s)\right] - s\ \bar{W}(t,o)$$

with the initial conditions:

$$\Omega(o,s) = \Delta_{\widetilde{\Omega}}(s)$$

Its solution is given in formula (50).

In order to determine the unknown matrix $W(t,o)$, we write ζ in place of s in equation (60) and take the Laplace transform of both sides with respect to the variable t. We so obtain:

$$(62) \quad (s-\zeta+\lambda)\ \Omega^{*}_{ij}(s,\zeta) - \lambda \sum_{\nu=1}^{m} \psi_{\nu j}(\zeta)\ \Omega^{*}_{i\nu}(s,\zeta) =$$

$$\delta_{ij}\widetilde{\Omega}_{i}(\zeta) - \zeta \int_{o}^{\infty} W_{ij}(t,o)\ e^{-st}dt, \quad (i,j = 1,\ldots,m;\mathrm{Re}\ s>o)$$

In matrix form, we obtain:

$$(63) \qquad \Omega^*(s,\zeta) = \left[\Delta_{\underset{\sim}{\Omega}}(\zeta) - \zeta\, W^*(s)\right]\left[(s-\zeta+\lambda)\; I - \lambda\; \Psi(\zeta)\right]^{-1}$$

As before, we assume that the matrix $\Psi(\zeta)$ has distinct eigenvalues for $\mathrm{Re}\;\zeta \geq o$ and can be written as:

$$(64) \qquad \Psi(\zeta) = R(\zeta)\; H(\zeta)\; R^{-1}(\zeta)$$

where R and H are defined in formula (27).

Formulae (63) and (64) lead to:

$$(65) \qquad \Omega^*_{ij}(s,\zeta) = \sum_{\nu=1}^{m}\left[\delta_{i\nu}\Omega_i(\zeta) - \zeta\; W^*_{i\nu}(s)\right] \cdot \sum_{\rho=1}^{m} \frac{\alpha_{\nu\rho}(\zeta)\beta_{\rho j}(\zeta)}{s+\lambda-\zeta-\lambda\eta_\rho(\zeta)}$$

The right hand side of (65) must be an analytic function in s and ζ for $\mathrm{Re}\; s > o$, $\mathrm{Re}\;\zeta \geq o$. By Rouché's theorem we see that the denominators $s+\lambda-\zeta-\lambda\eta_\rho(\zeta)$, $\rho = 1,\ldots, m$ have one and only one zero in ζ in this domain. It is directly verified that these roots are given by:

$$(66) \qquad \zeta_\rho(s) = s+\lambda-\lambda\gamma_\rho(s,1), \quad (\rho = 1,\ldots, m)$$

where $\gamma_\rho(s,1)$ is the root of the equation

$$z = \eta_\rho(s+\lambda-\lambda z)$$

which has smallest absolute value.

Arguing as before, we see that the matrix $W^*(s)$ must satisfy the system of m^2 equations in m^2 unknowns:

$$(67) \qquad \sum_{\nu=1}^{m} W^*_{i\nu}(s)\,\alpha_{\nu\rho}(\zeta_\rho(s)) = \frac{\tilde{\Omega}_i(\zeta_\rho(s))}{\zeta_\rho(s)}\,\alpha_{i\rho}\,(\zeta_\rho(s)),$$

$$(\rho, i = 1, \ldots, m).$$

<u>Remark</u>

The distribution functions $W_{ij}(t,x)$ satisfy the following system of integro-differential equations:

$$(68) \qquad \frac{\partial W_{ij}(t,x)}{\partial t} = \frac{\partial W_{ij}(t,x)}{\partial x} - \lambda\left[W_{ij}(t,x) - \sum_{\nu=1}^{m}\int_{o}^{x} Q_{\nu j}(x-y)\,d_y W_{i\nu}(t,y)\right]$$

$(i,j = 1,\ldots, m)$ for almost all $t \geq o$ and $x \geq o$.

This can be argued as follows: By the theorem on total probability, we have:

$$(69) \qquad W_{ij}(t+\Delta t,x) = (1-\lambda\Delta t)\,W_{ij}(t,x+\Delta t) + \lambda\Delta t \sum_{\nu=1}^{m}\int_{o}^{x} Q_{\nu j}(x-y)\,d_y\,W_{i\nu}(t,y) + o(\Delta t)$$

for the event $\left\{\eta(t,\Delta t) \leq x,\ J_{t+\Delta t} = j\right\}$ may happen in the following, mutually exclusive ways.

1. In the time interval $(t,t+\Delta t]$ no customer arrives at the counter and $\eta(t) \leq x + \Delta t$, $J_t = j$. The probability of this event is

$$(1 - \lambda\Delta t)\,W_{ij}(t,x + \Delta t) + o(\Delta t)$$

2. In the time-interval $\left[t, t+\Delta t\right]$ one customer arrives at the counter. His type must be j and type of the previous customer is ν $(\nu = 1, \ldots, m)$. His service time is less than $x - \eta(t) + \epsilon_t \Delta t$, where $o \leq \epsilon_t \leq 1$. The probability of this occurence is:

$$\lambda \Delta t \int_o^x \sum_{\nu=1}^{m} Q_{\nu j}(x-y)\, d_y\, W_{i\nu}(y) + o(\Delta t)$$

3. In the time-interval $(t, t+\Delta t]$, more than one customer arrivesin the queue. The probability of this event is $o(\Delta t)$.

Now for each fixed t and almost all $x \geq o$, we have

$$W_{ij}(t, x + \Delta t) = W_{ij}(t,x) + \frac{\partial W_{ij}(t,x)}{\partial x} \Delta t + o(\Delta t)$$

Substituting in (69) and taking

$$\lim_{\Delta t \to o} \frac{W_{ij}(t+\Delta t, x) - W_{ij}(t,x)}{\Delta t}$$

we obtain formula (68).

It is possible to obtain the equations (67) for the matrix $W^*(s)$ directly from the process of successive busy periods.

We consider the sequence $\left\{(I_n, Z_n);\ n = o,1,\ldots\right\}$ in which the random variables I_n are defined as $Z_o = o$ a.s. and $Z_n = Y_n + U_n$ $(n = 1, \ldots)$ in which the U_n are the lengths of the idle periods, immediately following the busy periods. The idle periods have lengths, which are independent of the busy periods and are negative exponentially distributed with parameter λ. The sequence $\left\{(I_n, Z_n);\ n = o,1,\ldots\right\}$ is clearly a semi-Markov sequence with initial probability vector p^o and with transition probability matrices, whose entrywise Laplace transforms are given by $\lambda(\lambda+s)^{-1}\tilde{E}(1,s)$ and $\lambda(\lambda+s)^{-1} E(1,s)$.

The matrices $\tilde{E}$ and E were defined in formulae (39) and (21) respectively.

Let us denote by Z_t the continuous time semi-Markov process associated with the sequence $\{(I_n, Z_n)\ n = o,1,\ldots\}$ and let $M^o_{ij}(t)$ be the expected number of visits to the state j before time t, given that the initial state is i.

We set:

$$m^o_{ij}(s) = \int_o^{\infty} e^{-st}\, d\, M^o_{ij}(t), \qquad (i,j = 1,\ldots, m) \tag{70}$$

and denote the matrix $\{ m^o_{ij}(s) \}$ by $m^o(s)$. It then follows from a slight extension of a formula of Pyke [7] that

$$m^o(s) = \frac{\lambda}{\lambda+s} \tilde{E}(1,s) \left[I - \frac{\lambda}{\lambda+s} E(1,s)\right]^{-1} = \lambda\, \tilde{E}(1,s)\left[(\lambda+s)\ I - \lambda\ E(1,s)\right]^{-1} \tag{71}$$

On the other hand

$$M^o_{ij}(t) = \lambda \int_o^t W_{ij}(\omega,o)\ du, \tag{72}$$

for, if we introduce a new time variable, involved only when the server is idle and $I_t = j$, then in this process the successive visits to j form a Poisson process with parameter λ.

From (71) and (72) we obtain:

$$W^*(s) = \tilde{E}(1,s)\left[(\lambda+s)\ I - \lambda\ E(1,s)\right]^{-1}, \tag{73}$$

In particular if $P\left\{\eta(o) = o \,|\, I_o = i\right\} = 1$ for $i = 1,\ldots, m$, then we have

$$\tilde{E}(1,s) = I \qquad \Delta_{\tilde{\Omega}}(s) = I,$$

Formulae (73) and (31) then yield:

$$W^{*}(s) = T(1,s)\left[(\lambda+s)\ I - \lambda\ddot{\Delta}\ (1,s)\right]^{-1} T^{-1}(1,s) \tag{74}$$

in which the matrix $\Delta(1,s)$ is defined as in (31). But we have

$$\zeta_\rho(s) = s + \lambda - \lambda\gamma_\rho(s,1)$$

which yields formula (51) after substitution in (74) in the case where $\eta(o) = o$ a.s.

If we use the formula for the matrix $E(1,s)$ we likewise obtain the formula (51) for the general case.

If we know (51) and $E(1,s)$ we may by the same token derive the equations for $E(1,s)$.

Equation (51) may be written as.

$$W^{*}(s)\ T(1,s) = B(s)$$

in which

$$B_{i\rho}(s) = \frac{\tilde{\Omega}_i(\zeta_\rho(s))}{\zeta_\rho(s)}\ \alpha_{i\rho}(\zeta_\rho(s))$$

This leads to:

$$\tilde{E}(1,s) = B(s)\left[(\lambda+s)\ I - \lambda\Delta(1,s)\right]^{-1} T^{-1}(1,s)$$

which is clearly the matrix version of the system of equations (43) with $w = 1$.

7. The Limiting Behavior of the Virtual Waiting Time

We first study the limiting behavior of the probabilities $W_{ij}(t,o)$ as t tends to infinity.

Theorem 7

The limits $\lim_{t \to \infty} W_{ij}(t,o)$ always exist and are independent of i. If we set

$$\lim_{t \to \infty} W_{ij}(t,o) = P_j^*$$

then $P_j^* = o$, $j = 1,\dots, m$ if $\lambda\eta_m'(o+) \leq -1$. The limits P_j^* are the solutions of the system of linear equations

$$(75) \qquad \sum_{\nu=1}^{m} P_\nu^* \alpha_{\nu\rho}\left[\zeta_\rho(o+)\right] = o \qquad \text{for} \quad \rho \neq m$$

$$\sum_{\nu=1}^{m} P_\nu^* = 1 + \lambda\eta_m'(o+) \qquad \text{for} \quad \rho = m.$$

The latter equation gives the asymptotic value of the probability that the queue is empty at a given instant of time t for $t \to \infty$.

Proof:

We consider the time intervals between epochs at which the server becomes idle. If we associate with each such interval the type of the first customer to be served during the busy period, which it contains, then we again obtain a semi-Markov process. It is characterized by two matrices $\tilde{\mathcal{H}}$ and $\mathcal{H}$ and by the vector $\bar{p}^o$ of initial probabilities. Clearly

$$\tilde{\mathcal{H}}_{ij}(x) = \tilde{G}_{ij}(x)$$

and

recurrent processes - Takacs [14] we obtain from (76) and (77) that

$$\lim_{t \to \infty} W_{i\nu}(t,o) = \frac{1}{\lambda M^*_{\nu\nu}}$$

which is equal to zero in the null-recurrent case and strictly positive in the positive recurrent case.

Knowing that the limits exist, we may obtain the equations (75) by applying a standard Abelian theorem.

Since the limit exists, we have:

$$\lim_{t \to \infty} \frac{1}{t} \int_o^t W_{ij}(u,o)\, du = P^*_j$$

and hence

$$\lim_{s \to o} s \int_o^\infty e^{-st} W_{ij}(t,o)\, dt = P^*_j$$

Using (49) and multiplying by s in both sides of equation (67), we obtain as $\to$ o+, that:

$$\sum_{\nu=1}^{m} P^*_\nu \alpha_{\nu\rho} \left[\lambda - \lambda\gamma_\rho(o+,1)\right] = 0, \quad \text{for} \quad \rho \neq m.$$

and for $\rho = m$:

$$\sum_{\nu=1}^{m} P^*_\nu \alpha_{\nu m}(o+) = \alpha_{im}(o+) \lim_{s \to o+} \frac{s}{s+\lambda-\lambda\gamma_m(s,1)} = \alpha_{im}(o+) \left[1+\lambda\eta'_m(o+)\right]$$

However, because of the interpretation of the $\alpha_{\nu m}(o+)$, we have

$$\mathcal{H}_{ij}(x) = F * G_{ij}(x)$$

where F is the negative exponential distribution with parameter λ.

Let $N_{i\nu}(t)$ $(i, \nu = 1, \ldots, m)$ denote the expected number of times emptiness is reached in the interval $(o, t]$, whereby the first customer of the next busy period is of type ν. We then have:

$$(76) \qquad W_{i\nu}(t, o) = \delta_{i\nu} W_{i\nu}(o, o)\, e^{-\lambda t} + \int_o^t e^{-(t-u)\lambda} dN_{i\nu}(u)$$

for $i, \nu = 1, \ldots, m$.

The semi-Markov process $[\bar{p}^o, \tilde{\mathcal{H}}, \mathcal{H}]$ clearly has the same ergodic properties as the process of busy periods. If $\lambda \eta_m'(o+) < -1$, then all states are transient and $N_{i\nu}(u) = o$ for all i and ν and $u \geq o$.

It follows that if $\lambda \eta_m'(o+) < -1$, we have $\lim_{t \to \infty} W_{i\nu}(t, o) = \lim_{t \to \infty} e^{-\lambda t} W_{i\nu}(o, o) = o$.

If $\lambda \eta_m'(o+) \geq -1$, then the semi-Markov process is recurrent. The epochs at which the state ν is entered, form a renewal process. Applying Blackwell's theorem, we obtain that for all $h > o$.

$$(77) \qquad \lim_{t \to \infty} \frac{N_{i\nu}(t+h) - N_{i\nu}(t)}{h} = \lim_{t \to \infty} \frac{N_{i\nu}(t)}{t} = \frac{1}{M^*_{\nu\nu}}$$

in which $M^*_{\nu\nu}$ is the expected length of time between successive visits to state ν in the semi-Markov process $[\bar{p}^o, \tilde{\mathcal{H}}, \mathcal{H}]$. It is known that $M^*_{\nu\nu}$ $(\nu = 1, \ldots, m)$ is finite if and only if the semi-Markov process is positive recurrent and infinite otherwise. - Pyke [6].

We note that the distribution of the recurrence time of the state ν $(\nu = 1, \ldots, m)$ is non-lattice, so, applying the fundamental theorem on recurrent processes - Takacs [14] we obtain from (76) and (77) that

$$\lim_{t \to \infty} W_{i\nu}(t,o) = \frac{1}{\lambda M^*_{\nu\nu}}$$

which is equal to zero in the null-recurrent case and strictly positive in the positive recurrent case.

Knowing that the limits exist, we may obtain the equations (75) by applying a standard Abelian theorem.

Since the limit exists, we have:

$$\lim_{t \to \infty} \frac{1}{t} \int_o^t W_{ij}(u,o)\, du = P^*_j$$

and hence

$$\lim_{s \to o} s \int_o^\infty e^{-st}\, W_{ij}(t,o)\, dt = P^*_j$$

Using (49) and multiplying by s in both sides of equation (67), we obtain as $s \to o+$, that:

$$\sum_{\nu=1}^{m} P^*_\nu\, \alpha_{\nu\rho}\left[\lambda - \lambda\gamma_\rho(o+,1)\right] = 0, \quad \text{for } \rho \neq m.$$

and for $\rho = m$:

$$\sum_{\nu=1}^{m} P^*_\nu\, \alpha_{\nu m}(o+) = \alpha_{im}(o+) \lim_{s \to o+} \frac{s}{s+\lambda-\lambda\gamma_m(s,1)} = \alpha_{im}(o+)\left[1+\lambda\Pi'_m(o+)\right]$$

However, because of the interpretation of the $\alpha_{\nu m}(o+)$, we have