PRIORITY QUEUES

BY
RUPERT G. MILLER, JR.

TECHNICAL REPORT NO. 6

PREPARED UNDER CONTRACT Nonr-225(28)
(NR-047-019)
FOR
OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

SEPTEMBER 10, 1958
PRIORITY QUEUES

BY

RUPERT G. MILLER, JR.

TECHNICAL REPORT NO. 6

PREPARED UNDER CONTRACT Nonr-225(28)
(NR-047-019)
FOR
OFFICE OF NAVAL RESEARCH

Reproduction in Whole or in Part is Permitted for any Purpose of the United States Government

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

SEPTEMBER 10, 1958
§0. Introduction and Summary.

In a priority queue different types of items (individuals or elements) arrive at a service mechanism, and each type item has a relative priority for order of service. To be precise, let there be $K$ classes of items, $1, 2, \ldots, K$. If the service mechanism is to select an item for service, a type $i$ item will be selected in preference to a type $j$ item for $i < j$ even if the type $j$ item arrived before the type $i$ item, and within each class the "first come, first served" policy determines the order of service. When a type $j$ item is in service and a type $i$ item arrives ($i < j$), there are two major disciplines for handling the priority demand. The "head-of-the-line" discipline allows the type $j$ item to complete service but places the type $i$ item ahead of any other lower priority items. The "preemptive" discipline withdraws the type $j$ item from service and replaces it by the type $i$ item. Note that under the preemptive scheme the only time at which a type $j$ item ($1 < j$) can be in service is when there are no items of types $1, \ldots, j-1$ in the queue. When a lower priority item which has been preempted returns to service, the preemptive discipline must distinguish two cases. The "preemptive resume" policy allows the preempted item to resume service at the point at which it was preempted so that its service time upon reentry has been reduced by the amount of time the item has already spent in service. The "preemptive repeat" policy requires the preempted item to commence
service again at the beginning. A priority queue with an indifferent server is of course a special case of the preemptive resume discipline.

In the special case $K = 2$ the type 1 items will be referred to as priority items and the type 2 items as non-priority items.

It will be assumed throughout this paper that the input process for type 1 items, $i = 1, \ldots, K$, is Poisson with arrival rate $\lambda_i$ and the input processes operate independently. The service time distribution for a type 1 item (in isolation) will be denoted by $F_{S_1}$ and unless explicitly stated to the contrary will be assumed to be general subject only to the restriction $F_{S_1}(0+) = 0$. The service mechanism consists of a single channel or server.

A. Cobham in [1], [2] introduced the head-of-the-line priority queue and derived the expected waiting times in equilibrium for a single channel queue with general service distributions and for a multiple channel queue with common exponential service. J. Holley [3] produced a simplification of Cobham's method, and Dressin and Reich ([4],[5]) derived the characteristic functions for the equilibrium waiting time distributions of a single channel queue with common exponential service. H. Kesten and J. Th. Runnenburg ([6],[7]) obtained the Laplace-Stieltjes transforms and first two moments of the steady state waiting time distributions for the case of general service distributions, and P. M. Morse ([8], Ch. 9) derived the generating function of the equilibrium probability distribution on the number of items in the $K = 2$ priority
queue with different exponential service.

The first published results for the preemptive discipline were those of H. White and L. S. Christie [9]. White and Christie derived the generating function for the stationary probabilities of the number of items in a $K = 2$ priority queue with different exponential service. They also obtained the expected steady state waiting times and in the special case of equal effective service rates the Laplace-Stieltjes transforms of the actual waiting time distributions. J. Y. Barry [10] and F. F. Stephan [11] have unpublished results for the preemptive queue. In [12] E. Koenigsberg has generalized this priority model to a continuous number of priority types with application to machine breakdown problems.

In this report the following quantities have either been obtained explicitly or characterized as the unique (subject to regularity conditions) solution to a functional equation:

§1. Generating function for the stationary probabilities on the number of priority and non-priority items ($K = 2$) in the queue for the

1.) preemptive discipline with different exponential service rates (White and Christie [9]).

2.) head-of-the-line discipline with different exponential service rates (Morse [8]).

3.) imbedded Markov chain, head-of-the-line discipline, general service distributions.
§2. Laplace-Stieltjes transforms of the waiting time distributions for
1.) priority and non-priority items \((K = 2)\) in the steady state,
head-of-the-line discipline.
2.) the lowest priority item \((K \geq 2)\) at arbitrary time \(t\) and in
equilibrium, head-of-the-line discipline.
3.) any type item \((K \geq 2)\) at arbitrary time \(t\) and in equilibrium
for the preemptive resume priority queue.

§3. Laplace-Stieltjes transform of the distribution of a busy period
\((K \geq 2)\) which commences with
1.) a type \(j\) item in line initially (for both the head-of-the-line
and preemptive resume disciplines).
2.) a randomly selected item in line initially (for both the head-
of-the-line and preemptive resume disciplines).

§4. Generating function for the probabilities on the number of items
serviced during a busy period where the number is
1.) the number of type \(j\) items and the initial arrival is type \(j\)
(both head-of-the-line and preemptive resume disciplines).
2.) the total number of items, irrespective of class, and the
initial arrival is random (both head-of-the-line and preemptive
resume disciplines).

The generating functions 1.) and 2.) of §1 are given explicitly. The
generating function 3.), §1, is also given explicitly except in that it
involves a transform characterized as the unique solution of a functional
equation. Similarly, except for involving transforms from §3, the transforms of §2 have an explicit expression. The transforms of §3 and §4 are characterized as the unique solutions to functional equations. In most of the cases the first two moments of the distributions are obtained.

In §1 it is pointed out that in the event the service distributions are exponential, the stationary probabilities for the imbedded Markov chain do not agree with those obtained from the infinitesimal generator. Although this may not be intuitively appealing, it is not the first time the discrepancy has occurred. This phenomenon has been noted earlier by L. Takács [13], H. Scarf [14], and the author [15] for different queues.

§1. Stationary Distributions.

Except for only the simplest queues the problem of determining the probability distribution on the number of items in the queue at a general time $t$ remains unsolved, and the solutions for those queues which admit a solution are for the most part non-trivial. As a result considerable attention has been paid to the stationary or steady state distributions. This problem as well has not been solved for general distributions, but both for the case of Poisson arrivals, exponential service and the case of Poisson arrivals, general service methods for obtaining a stationary distribution are known.
If the arrivals are Poisson and the service exponential, the queue process is a generalized birth and death process. If $P$ is a stationary distribution of the queue process, it must be a solution to the steady state equations which, symbolically, can be represented as $PA = 0$ where $A$ is the infinitesimal matrix of the process. For birth and death processes the system of equations $PA = 0$ need not have a unique solution, and each solution, even if it is a probability distribution, need not be a stationary distribution of the process (see[16]). For the two examples to be considered here the solution, subject to the condition that it be a probability distribution, is unique so if the more difficult problem of verifying the existence of a stationary distribution is shelved for the present, a study of the equations $PA = 0$ will yield a characterization of the (assumed existent) stationary distribution.

This is the method employed by White and Christie in [9] to study the preemptive priority queue for the case $K = 2$. Through algebraic manipulation of the steady state equations White and Christie obtained the generating function and thereby the first and second moments for this priority queue. $\mu_1$ and $\mu_2$ will denote the service rate parameters for the priority and non-priority items, respectively. Justification of a non-priority Poisson service process with the parameter $\mu_2$ from assumptions on the non-priority service process in isolation is discussed in detail in [9] with regard to the resume and repeat
disciplines and the indifferent server queue.

Let $\rho_i = \lambda_i / \mu_i$, $i = 1, 2$. If $\rho_1 + \rho_2 > 1$, the queue will become saturated with priority and/or non-priority items so no stationary distribution will exist. Hence, it will be assumed that the equilibrium condition $\rho_1 + \rho_2 < 1$ holds.

Define $P_{nm}$ to be the stationary probability that there are $n$ priority and $m$ non-priority items in the queue and $F(s,t) = \sum_{n,m} P_{nm} s^n t^m$.

That $P_{nm}$ is uniquely defined (if at all) will become evident from the equations.

The steady state equations $PA = 0$ become

$$0 = -(\lambda_1 + \lambda_2 + \mu_1)P_{nm} + \lambda_1 P_{n-1,m} + \lambda_2 P_{n,m-1} + \mu_1 P_{n+1,m} \quad (n,m > 0)$$

$$0 = -(\lambda_1 + \lambda_2)P_{00} + \mu_1 P_{10} + \mu_2 P_{01}$$

$$0 = -(\lambda_1 + \lambda_2 + \mu_1)P_{n0} + \lambda_1 P_{n-1,0} + \mu_1 P_{n+1,0} \quad (n > 0)$$

$$0 = -(\lambda_1 + \lambda_2 + \mu_2)P_{0m} + \lambda_2 P_{0,m-1} + \mu_2 P_{0,m+1} + \mu_1 P_{1m} \quad (m > 0)$$

Equations (1.1) imply that the $P_n(t) = \sum_m P_{nm} t^m$ satisfy

$$0 = -[\lambda_1 + \lambda_2(1-t) + \mu_1]P_n(t) + \lambda_1 P_{n-1}(t) + \mu_1 P_{n+1}(t) \quad (n > 0)$$
(1.3) \[ 0 = -[\lambda_1 + \lambda_2(1-t) + \mu_2(1-t^{-1})]P_0(t) + \mu_1P_1(t) + \mu_2(1-t^{-1})P_{00} \]

Equation (1.2) is a homogeneous, linear, second order difference equation so the solution is \( P_n(t) = \alpha^n(t)P_0(t) \) where

(1.4) \[ \alpha(t) = \frac{\lambda_1 + \lambda_2(1-t) + \mu_1}{2\mu_1} \sqrt{\frac{(\lambda_1 + \lambda_2(1-t) + \mu_1)^2 - 4\lambda_1\mu_1}{2\mu_1}} \]

is the root of the quadratic equation in \( \alpha \) resulting from (1.2) which lies in the interval \((0,1)\). From equation (1.3) it is evident that

(1.5) \[ P_0(t) = P_{00} \mu_2(t^{-1}-1)[\mu_1\alpha(t) - \lambda_1 - \lambda_2(1-t) - \mu_2(1-t^{-1})]^{-1} \]

Since \( P(s,t) = \sum_{n=0}^{\infty} P_n(t)s^n \), \( P(s,t) = P_0(t)[1-\alpha(t)s]^{-1} \). \( P_{00} \) is evaluated from the restriction \( P(1,1) = 1 \); \( P_{00} = 1 - \rho_1 - \rho_2 \). Therefore,

(1.6) \[ P(s,t) = \frac{(1-\rho_1-\rho_2)\mu_2(t^{-1}-1)}{[\mu_1\alpha(t) - \lambda_1 - \lambda_2(1-t) - \mu_2(1-t^{-1})][1-\alpha(t)s]} \]

The moments of the number of priority items in the queue are the same as those for priority items in isolation; e.g.,
\[
E(n) = \frac{\rho_1}{1-\rho_1} \quad ; \quad E(n^2) = \frac{2\rho_1^2}{(1-\rho_1)^2} + \frac{\rho_1}{(1-\rho_1)} .
\]

The moments of the number of non-priority items in the queue can be evaluated from (1.6). In particular, White and Christie have calculated

\[
E(m) = \left[ \frac{\rho_2}{1-\rho_1-\rho_2} \right] [1 + \frac{\mu_2}{\mu_1} \frac{\rho_1}{1-\rho_1} ]
\]

\[
E(m^2) = \frac{2\rho_1(\lambda_2/\mu_1)^2}{(1-\rho_1)^3(1-\rho_1-\rho_2)}
\]

\[
(1.8) \quad \rho_2(1-\rho_1)^3 + \rho_1^2(\lambda_2/\mu_1)^2 + \rho_1(1-\rho_1)(1-\rho_1+\rho_2)(\lambda_2/\mu_1)
\]

\[
+ \frac{\rho_2^2}{(1-\rho_1-\rho_2)^2} \left[ 1 + \frac{\mu_2}{\mu_1} \frac{\rho_1}{1-\rho_1} \right]^2 .
\]

In addition, White and Christie in [9] have given recursive methods for explicitly calculating the \( P_{nm} \).

For a head-of-the-line priority queue with exponential service times \((\mu_1, \mu_2)\) P. M. Morse ([8], Ch. 9) has derived the generating function and first moments of the stationary probabilities through the same technique. Since a non-priority item can be in service while there are priority items in the queue, it is necessary to add an
additional subscript to describe the state of the process. To be exact, let $P_{1nm}$ be the stationary probability that a priority item is in service and there are $n$ priority ($n > 0$), $m$ non-priority items in the queue, and let $P_{2nm}$ be the stationary probability that a non-priority item is in service with $n$ priority, $m$ non-priority ($m > 0$) items in the queue. $P_{00}$ is the probability the queue is empty. As for the preemptive queue, the steady state equations will show that $P_{1nm}$ is uniquely defined. Also, define

$$P_{1n}(t) = \sum_{m=0}^{\infty} P_{1nm} t^m \quad (n > 0), \quad P_{2n}(t) = \sum_{m=1}^{\infty} P_{2nm} t^m \quad (n > 0),$$

$$P_1(s, t) = \sum_{n=1}^{\infty} P_{1n}(t) s^n, \quad \text{and} \quad P_2(s, t) = \sum_{n=0}^{\infty} P_{2n}(t) s^n. \quad \text{Then, } P(s, t),$$

the generating function, is equal to $P_{00} + P_1(s, t) + P_2(s, t)$. For the queue to reach a steady state the equilibrium condition $\rho_1 + \rho_2 < 1$ must be fulfilled.

The steady state equations when written out explicitly become

$$(1.9) \quad 0 = -(\lambda_1 + \lambda_2 + \mu_1) P_{1nm} + \lambda_1 P_{1, n-1, m} + \lambda_2 P_{1, n, m-1} + \mu_1 P_{1, n+1, m} + \mu_2 P_{2, n, m+1} \quad (n > 1, m > 0)$$

$$(1.10) \quad 0 = -(\lambda_1 + \lambda_2 + \mu_2) P_{2nm} + \lambda_1 P_{2, n-1, m} + \lambda_2 P_{2, n, m-1} \quad (n > 0, m > 1)$$
(1.11) \[ 0 = - (\lambda_1 + \lambda_2 + \mu_1)P_{110} + \lambda_2 P_{111, -1, 0} + \mu_2 P_{210} + \mu_3 P_{211} \quad (m > 0) \]

(1.12) \[ 0 = - (\lambda_1 + \lambda_2 + \mu_2)P_{210} + \lambda_2 P_{211} \quad (n > 0) \]

(1.13) \[ 0 = - (\lambda_1 + \lambda_2 + \mu_1)P_{110} + \lambda_2 P_{111, -1, 0} + \mu_2 P_{210} + \mu_3 P_{211} \quad (n > 1) \]

(1.14) \[ 0 = - (\lambda_1 + \lambda_2 + \mu_2)P_{210} + \lambda_2 P_{211} \quad (m > 1) \]

(1.15) \[ 0 = - (\lambda_1 + \lambda_2 + \mu_1)P_{110} + \lambda_2 P_{111, -1, 0} + \mu_2 P_{210} + \mu_3 P_{211} \]

(1.16) \[ 0 = - (\lambda_1 + \lambda_2 + \mu_2)P_{210} + \lambda_2 P_{211} \quad (m > 1) \]

(1.17) \[ 0 = - (\lambda_1 + \lambda_2)P_{110} + \mu_1 P_{111} + \mu_2 P_{211} \]

Equations (1.9), (1.11), (1.13), (1.15) when multiplied by the appropriate powers of \( s \) and \( t \) and summed yield the expression

(1.18) \[ \left[ \lambda_1(1-s) + \lambda_2(1-t) + \mu_1(1-s^{-1}) \right]P_1(s,t) \]

\[ = \mu_2 t^{-1}P_2(s,t) - \mu_1 P_{11}(t) - \mu_2 t^{-1}P_{20}(t) + \lambda_1 sP_{00} . \]
Equations (1.10), (1.12), (1.14), (1.16), and (1.17) imply

\begin{equation}
\begin{align*}
\lambda_1(1-s) + \lambda_2(1-t) + \mu_2 P_2(s,t) \\
= \mu_1 P_{11}(t) + \mu_2 t^{-1} P_{20}(t) - [\lambda_1 + \lambda_2(1-t)] P_{00} .
\end{align*}
\end{equation}

The required relationship between \( P_{11}(t) \) and \( P_{20}(t) \) is obtained from (1.14), (1.16) and is given by

\begin{equation}
\begin{align*}
\mu_1 P_{11}(t) = [\lambda_1 + \lambda_2(1-t) + \mu_2](1-t^{-1})] P_{20}(t) + [\lambda_1 + \lambda_2(1-t)] P_{00} .
\end{align*}
\end{equation}

Combination of (1.18), (1.19), and (1.20) yields the following expression for \( P(s,t) \):

\begin{equation}
\begin{align*}
P(s,t) = \frac{\mu_1(1-s^{-1})}{\lambda_1(1-s) + \lambda_2(1-t) + \mu_1(1-s^{-1})} P_{00} \\
+ \frac{[\lambda_1 + \lambda_2(1-t) + \mu_2][\mu_1(1-s^{-1}) - \mu_2(1-t^{-1})]}{[\lambda_1(1-s) + \lambda_2(1-t) + \mu_2][\lambda_1(1-s) + \lambda_2(1-t) + \mu_1(1-s^{-1})]} P_{20}(t) .
\end{align*}
\end{equation}

From (1.21) it becomes evident that in order to determine \( P(s,t) \) it is first necessary to determine \( P_{20}(t) \). This can be accomplished by considering several additional relationships obtainable from the steady state equations. From (1.10), (1.12)
(1.22) \[ p_{2n}(t) = \left( \frac{\lambda_1}{\lambda_1 + \lambda_2(1-t) + \mu_2} \right) p_{2n-1}(t) \] \quad (n \geq 1),

from (1.9), (1.13)

(1.23) \[ -\mu_1 p_{1,n+1}(t) + [\lambda_1 + \lambda_2(1-t) + \mu_1] p_{1n}(t) - \lambda_1 p_{1,n-1}(t) = \mu_2 t^{-1} p_{2n}(t) \] \quad (n > 1),

and from (1.11), (1.15)

(1.24) \[ [\lambda_1 + \lambda_2(1-t) + \mu_1] p_{11}(t) - \mu_1 p_{12}(t) - \lambda_1 p_{00} = \mu_2 t^{-1} p_{21}(t). \]

Equation (1.22) implies immediately

(1.25) \[ p_{2n}(t) = \left( \frac{\lambda_1}{\lambda_1 + \lambda_2(1-t) + \mu_2} \right)^n p_{20}(t). \]

Equation (1.23) is a non-homogeneous, linear, second order difference equation. The solution to the corresponding homogeneous equation is \( p_{1n}(t) = \alpha^n(t)A(t) \) where \( \alpha(t) \) is defined by (1.4) and \( A(t) \) is a function still to be determined; a particular solution for the non-homogeneous case is
(1.26) \[ P_{01}(t) = \frac{\lambda_1 \mu_2 t^{-1}}{\left(\mu_1 - \mu_2\right)\left[\lambda_1 + \frac{\lambda_2}{\mu_2}(1-t) + \mu_2\right] - \lambda_1 \mu_2} \left[\frac{\lambda_1}{\lambda_1 + \frac{\lambda_2}{\mu_2}(1-t) + \mu_2}\right] P_{20}(t). \]

Finally, with the results of (1.25) and (1.26) \( A(t) \) and \( P_{20}(t) \) can be determined from (1.20), (1.24), and the relation (1.2) which defines \( \alpha(t) \).

(1.27) \[ A(t) = P_{00} + \frac{\mu_2 t^{-1}[\lambda_1 + \lambda_2(1-t) + \mu_2]}{\lambda_1 \mu_2 - (\mu_1 - \mu_2)[\lambda_1 + \lambda_2(1-t) + \mu_2]} P_{20}(t). \]

(1.28) \[ P_{20}(t) = \frac{P_{00} t[\lambda_1 + \lambda_2(1-t) - \mu_1 \alpha(t)]((\mu_1 - \mu_2)[\lambda_1 + \lambda_2(1-t) + \mu_2] - \lambda_1 \mu_1)}{[\lambda_1 + \lambda_2(1-t) + \mu_2][\lambda_1 \mu_1 t - (\mu_1 - \mu_2) t[\lambda_1 + \lambda_2(1-t) + \mu_2] + \mu_2[\mu_1(1-\alpha(t)) - \mu_2]]}. \]

(1.21), (1.26) and the restraint \( P(1,1) = 1 \) determine \( P(s,t) \) uniquely; \( P_{00} = 1 - \rho_1 - \rho_2 \).

P. M. Morse has calculated the mean numbers of priority and non-priority items in the queue.

(1.29) \[ E(n) = \frac{\lambda_1}{\mu_1 - \lambda_1} \left[ 1 + \rho_2 \frac{\mu_1}{\mu_2} \right] \]

(1.30) \[ E(m) = \rho_2 + \frac{\lambda_2}{\mu_1 - \lambda_1} \left[ \frac{\rho_1 + \rho_2 \mu_1/\mu_2}{1 - \rho_1 - \rho_2} \right]. \]
The previous technique employed by Morse is not applicable for a head-of-the-line priority queue with Poisson arrivals but non-exponential service. The queue process as characterized by the number of items in the queue no longer has the Markov property, and infinitesimal changes cease to have significance. The Markov property can be restored, however, by reducing the continuous time parameter process to discrete time. This technique was introduced by D. G. Kendall ([17], [18]) and has been utilized by other workers (cf. [15], [19]) to obtain a stationary distribution for queues in which either the arrival or service time is non-exponential. If the service times are non-exponential in a head-of-the-line priority queue, a discrete time Markov process is generated if the queue process is observed only at those points in time which are the termination points of a service period - priority or non-priority. The state of the queue is \((n,m)\) where \(n\) is the number of priority and \(m\) the number of non-priority items in the queue (at the end of the service period). It is readily apparent that the discrete time process has the Markov property since both priority and non-priority arrivals are Poisson.

The behavior of the discrete time Markov chain "imbedded" in the continuous time process is determined by the transition probability matrix which is expressible in terms of

\[
P_{ij} = \text{probability that } i \text{ priority and } j \text{ non-priority items arrive during a priority service period}
\]
and \( q_{ij} \) is probability that \( i \) priority and \( j \) non-priority items arrive during a non-priority service period.

Let \( F_{S_1} \) be the distribution function of the service time for a type 1 item, \( i = 1, 2 \), and let \( \tilde{S}_1 \) be the Laplace-Stieltjes transform of \( F_{S_1} \). Clearly,

\[
p_{ij} = \int_0^\infty e^{-(\lambda_1 + \lambda_2)t} \frac{\lambda_1^i}{i!} \frac{\lambda_2^j}{j!} \, dF_{S_1}(t)
\]

and

\[
P(s, t) = \sum_{i,j} p_{ij} s^i t^j = \tilde{S}_1(\lambda_1(1-s) + \lambda_2(1-t))
\]

so

\[
q_{ij} = \int_0^\infty e^{-(\lambda_1 + \lambda_2)t} \frac{\lambda_1^i}{i!} \frac{\lambda_2^j}{j!} \, dF_{S_2}(t)
\]

and

\[
Q(s, t) = \sum_{i,j} q_{ij} s^i t^j = \tilde{S}_2(\lambda_1(1-s) + \lambda_2(1-t))
\]
Let \( P((n,m) \rightarrow (n',m')) \) be the probability the queue moves from state \( (n,m) \) to state \( (n',m') \) in one transition. In terms of the \( p_{ij}, q_{ij} \), the \( P((n,m) \rightarrow (n',m')) \) are

1. \( P((n,m) \rightarrow (n',m')) = 0 \) for \( n' < n - 1, \ n > 1, \ all \ m, m' \).

2. \( P((n,m) \rightarrow (n',m')) = 0 \) for \( m' < m, \ n > 1, \ all \ n' \).

3. \( P((n,m) \rightarrow (n-1+i, m+j)) = p_{ij} \) for \( i, j \geq 0, \ n \geq 1, \ all \ m \).

4. \( P((0,m) \rightarrow (n,m')) = 0 \) for \( m' < m-1, \ all \ n \).

5. \( P((0,m) \rightarrow (i,m-1+j)) = q_{ij} \) for \( i, j \geq 0, \ m > 0 \).

6. \( P((0,0) \rightarrow (i,j)) = \tau_1 p_{ij} + \tau_2 q_{ij} \) for \( i,j \geq 0 \) where

\[
\tau_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2}, \quad \tau_2 = \frac{\lambda_2}{\lambda_1 + \lambda_2}.
\]

The transition probabilities under (6) have their special form because if the state is \((0,0)\) the queue is next observed at the end of the service period for the first arrival so the probability of the new state \((n',m')\) depends on which type of item was first to arrive.

Let \( E(S_i) \) be the expected service time for a type \( i \) item, \( i = 1, 2 \). An argument can be constructed to prove that the state \((0,0)\)
is ergodic if the equilibrium condition \( \lambda_1 E(S_1) + \lambda_2 E(S_2) < 1 \) is satisfied. Since the proof is closely analogous to those in [15] and [20] for different queues, it will be omitted for the sake of brevity. Whereas the existence of the stationary distribution is still an open question for the two earlier problems, the equilibrium condition guarantees the existence of the stationary distribution for the imbedded Markov chain.

The stationary probability of there being \( n \) priority, \( m \) non-priority items in the queue will be denoted by \( \pi_{nm} \). By definition, the stationary distribution \( \pi = \{\pi_{nm}\} \) must satisfy the system of equations

\[
(1.32) \quad \pi_{n'm'} = \sum_{n,m} \pi_{nm} P((n,m) \to (n',m')), \quad \text{all} \quad n',m'.
\]

This system of equations can be utilized to yield an expression for

\[
\pi(s,t) = \sum_{n,m} \pi_{nm} s^n t^m. \quad \text{From (1.32)}
\]

\[
(1.33) \quad \pi(s,t) = \sum_{n',m'} \sum_{n,m} \pi_{nm} P((n,m) \to (n',m')) s^{n'} t^{m'}
\]

which, when simplified, gives the following expression for \( \pi(s,t) \):
\[ \pi(s,t) = [\pi_{00}(t_1 P(s,t) + t_2 Q(s,t) - t^{-1} Q(s,t)) \\
+ \pi_0(t)(t^{-1} Q(s,t) - s^{-1} P(s,t))][1 - s^{-1} P(s,t)]^{-1} \]

(1.34)

where \( \pi_0(t) = \sum_m \pi_{0m} t^m \) is analogous to \( p_{20}(t) \) of (1.21).

To determine \( \pi(s,t) \) it is necessary and sufficient to determine \( \pi_0(t) \). This can be accomplished by imbedding a second Markov chain within the original imbedded Markov chain. The second Markov chain is defined by taking cognizance of the state of the process only at those time points which are the termination points of a service period leaving 0 priority items in the queue. The state of the queue is \( (m) \), the number of non-priority items in the queue. It is clear that the Markov chain so defined is imbedded within the original chain since a trial for the second chain occurs at the end of a service period only if there are 0 priority items left in line whereas, previously, the termination of any service period constituted a trial.

Let \( P(m \rightarrow m') \) denote the transition probability of moving from state \( m \) to state \( m' \) in succeeding trials. For \( m > 0, \ j \geq 0 \)

\[ P(m \rightarrow m-1 + j) = P(m \rightarrow m-1 + j, \ \text{no priority arrivals in interim}) \]

\[ + P(m \rightarrow m-1 + j, \ \text{priority arrivals in interim}) \]
where

\begin{equation}
(1.35) \quad P(m \rightarrow m-1 + j, \text{ no priority arrivals in interim}) = q_{0j}
\end{equation}

\begin{equation}
(1.36) \quad P(m \rightarrow m - 1 + j, \text{ priority arrivals in interim})
= \sum_{j=0}^{\infty} \sum_{n=1}^{\infty} \left[ \int_{0}^{\infty} e^{-(\lambda_1 + \lambda_2)u} \frac{(\lambda_1 u)^n}{n!} \frac{(\lambda_2 u)^j}{j!} \, dF_{S_2}(u) \right] \cdot \left[ \int_{0}^{\infty} e^{-\lambda_2 v} \frac{(\lambda_2 v)^{j-k}}{(j-k)!} \, dF_B^{(n)}(v) \right].
\end{equation}

\(F_B\) is the distribution of the busy period for priority items in isolation (see [21]) and \(F_B^{(n)}\) is its \(n\)-fold convolution. If \(P(t)\) is defined by \(P(t) = \sum_{j=0}^{\infty} P(m \rightarrow m - 1 + j) \, t^j\) for \(m > 0\), it is readily verified that

\begin{equation}
(1.37) \quad P(t) = \tilde{S}_2 (\lambda_1 (1 - \tilde{B} (\lambda_2 (1 - t)))) + \lambda_2 (1 - t))
\end{equation}

where \(\tilde{B}\) is the Laplace-Stieltjes transform of \(F_B\).

For \(m = 0, j \geq 0\)
\[ P(0 \to j) = P(0 \to j, \text{ first arrival is priority}) + P(0 \to j, \text{ first arrival is non-priority}) \]

where

\[ (1.38) \quad P(0 \to j, \text{ first arrival is priority}) = \tau_1 \int_0^\infty e^{-\lambda_2 u} \frac{(\lambda_2 u)^j}{j!} dF_B(u) \]

\[ (1.39) \quad P(0 \to j, \text{ first arrival is non-priority}) = \tau_2 \sum_{j=0}^\infty \sum_{n=0}^\infty \left[ \int_0^\infty e^{-(\lambda_1 + \lambda_2) u} \frac{(\lambda_1 u)^n}{n!} \frac{(\lambda_2 u)^j}{j!} dF_S(u) \right] \]

\[ \cdot \left[ \int_0^\infty e^{-\lambda_2 v} \frac{(\lambda_2 v)^{j-k}}{(j-k)!} dF_B^{(n)}(v) \right] \]

where \( F(0) \) is the distribution which concentrates its total mass at 0.

If \( Q(t) \) is defined by \( Q(t) = \sum_{j=0}^\infty P(0 \to j) t^j \), then

\[ (1.40) \quad Q(t) = \tau_1 B(\lambda_2 (1-t)) + \tau_2 B(\lambda_1 (1-B(\lambda_2 (1-t)))) + \lambda_2 (1-t)) \]
Let \( \pi_m^o \) be the stationary probability of there being \( m \) non-priority items in the queue (for the second imbedded Markov chain). Algebraic manipulation of the system of equations

\[
(1.41) \quad \pi_{m'}^o = \sum_m \pi_m^o \ P\{m \to m'\}, \quad \text{all } m'
\]

yields the following expression for \( \pi^o(t) = \sum_{m=0}^\infty \pi_m^o \ t^m \):

\[
(1.42) \quad \pi^o(t) = \pi_0^o \ [Q(t) - t^{-1} P(t)] [1 - t^{-1} P(t)]^{-1}.
\]

\( \pi_0^o \) determines the normalization for the distribution \( \pi^o \). If the second Markov chain is to be viewed as imbedded within the first, the proper normalization is \( \pi_0^o = \pi_{00} \) which implies \( \pi_0^o = \pi_{0m} \) for all \( m \). Hence,

\[
(1.43) \quad \pi^o(t) = \pi_{00} [Q(t) - t^{-1} P(t)] [1 - t^{-1} P(t)]^{-1}.
\]

In conjunction with (1.34), (1.43) yields
\[ (1.44) \quad \pi(s,t) = \pi_{00} [1 - s^{-1} P(s,t)]^{-1} \left\{ \tau_1 P(s,t) + \tau_2 Q(s,t) - t^{-1} Q(s,t) ight. \\
\left. + (t^{-1} Q(s,t) - s^{-1} P(s,t)) [1 - t^{-1} P(t)]^{-1} Q(t) - t^{-1} P(t) \right\} \]

\( \pi_{00} \) is determined by the restraint \( \pi(1,1) = 1; \pi_{00} = 1 - \lambda_1 E(S_1) - \lambda_2 E(S_2) \).

The first moments of \( n \) and \( m \) (and also higher moments) can be calculated from (1.44). Let \( \rho_i = \lambda_i E(S_i), i = 1,2 \), and let \( E(S_i^2), i = 1,2 \), be the second moment of the service time distribution for type \( i \) items. Then,

\[ (1.45) \quad E(n) = \tau_1 (\rho_1 + \rho_2) + \frac{\lambda_1^2 \left[ \tau_1 E(S_1^2) + \tau_2 E(S_2^2) \right]}{2(1 - \rho_1)} \]

\[ (1.46) \quad E(m) = \tau_2 (\rho_1 + \rho_2) + \frac{\lambda_2^2 \left[ \tau_1 E(S_1^2) + \tau_2 E(S_2^2) \right]}{2\mu_1} \left[ \frac{\lambda_2 (\mu_1 + \mu_2) + \rho_1 (1 - \rho_1 - \rho_2)}{(1 - \rho_1 - \rho_2)(1 - \rho_1)} \right] \]

From (1.44)-(1.46) it becomes apparent that (1.44) does not agree with (1.21) when the service times are exponentially distributed. At first glance this could appear surprising since it might be felt that \( \pi_{nm} \) would equal \( P_{1nm} + P_{2nm} (n,m > 0) \) with a corresponding relationship
for the cases $n=0$, $m=0$. However, as more is learned about queueing problems, it becomes increasingly apparent that the stationary distribution for the imbedded Markov chain and the equilibrium distribution for $t \to \infty$ concur for only the simpler queues. Such discrepancies have already been observed in [13], [14], and [15]. For the more complex queues those points in time which are termination points for a service period seem to have a special character for which the behavior of the queue size is different from general time $t$.

The above work on the imbedded Markov chain was for a priority queue with head-of-the-line discipline. It might be hoped to duplicate these procedures for the preemptive priority queue. However, there does not exist a natural imbedding procedure for the preemptive queue. The discrete time process generated by considering the service terminal points and defining the state of the system to be the number of priority, non-priority items in the queue is non-Markov. To restore the Markov property it would be necessary to incorporate into the definition of the state the amount of time the leading non-priority item has already spent in service for the preemptive resume discipline and the maximum period of time spent in service before preemption for the preemptive repeat discipline. The only method of avoiding incorporation of an additional time quantity into the definition of the state would be to observe the process only at the termination
of service of a non-priority item. This, however, is a one-dimensional queue and has no significance for the priority queue.

For those one-dimensional queues in which the arrival distribution is general and the service exponential the natural imbedding considers those times at which a new arrival enters the queue. This particular imbedding pattern leads to a non-Markov process in the priority case. For the process to be Markov it is necessary to specify with the arrival of a new item (priority or non-priority) the number of priority, non-priority items in the queue and the time of the last arrival of the other type item. The addition of an extra time quantity into the state definition prevents any simple analysis.

Extension of the results of this section to the case of three or more types should be feasible except for the prohibitive amount of algebra involved.
§2. Waiting Time Distributions

The waiting time of an item is defined to be the length of time the item must wait in the queue before it is taken into service. The time an item spends in service is not included in the waiting time. For a priority queue with head-of-the-line discipline the time in service of an item is just the length of its service period, but for the preemptive discipline the term "time in service" will mean the total time from the moment the item first enters service to the moment it completes service including those periods of time in which it is waiting for reentry into service after having been preempted.

The equilibrium condition $1-\rho_1-\ldots-\rho_K > 0$ will be assumed throughout this section so that it is meaningful to discuss stationary distributions. With slight modification the discussion for general time $t$ also applies to the transient case.

The steady state waiting time distributions for a head-of-the-line priority queue, $K=2$, are easily obtained and will be discussed first. The method introduced by D. G. Kendall for the simple queue with a single class can be applied in the priority queue to derive the waiting time distribution for a priority item. Suppose an item has just completed service. Since the queue is assumed to be operating in a state of equilibrium, with probability $\tau_1$ the item was a priority item and with probability $\tau_2$ the item was non-priority.
If the item was a priority item, the number of priority items remaining in the queue must be the number which arrived during its waiting time and service period. If the item was non-priority, the number of priority items in the queue is just the number which arrived during its service period. But the probability that there are \( n \) priority items remaining in the queue is 

\[
\sum_{m=0}^{\infty} \pi_{nm} = \tau_1 \int_0^\infty \frac{-\lambda_1 t (\lambda_1 t)^n}{n!} dF_{W_1} * F_{S_1}(t) + \tau_2 \int_0^\infty \frac{-\lambda_1 t (\lambda_1 t)^n}{n!} dF_{S_2}(t)
\]

where \( F_{W_1} \) is the waiting time distribution for a priority item and 

\( F_{W_1} * F_{S_1} \) denotes the convolution of \( F_{W_1} \) and \( F_{S_1} \). From (2.1)

\[
\tilde{W}_1(s) = \frac{\pi(-\lambda_1^{-s}, 1) - \tau_2 \tilde{S}_2(s)}{\tau_1 \tilde{S}_1(s)}
\]

where \( \tilde{W}_1 \) is the Laplace-Stieltjes transform of \( F_{W_1} \). When the actual value of \( \pi(-\lambda_1^{-s}, 1) \) as given by (1.44) is substituted into (2.2),
(2.3) \[ \tilde{W}_1(s) = \frac{-(1-p_1-p_2)s-\lambda_2[1-S_2(s)]}{\lambda_1-s-\lambda_1\tilde{S}_1(s)}. \]

The moments of \( \tilde{W}_1 \) can be readily calculated from (2.3). In particular,

(2.4) \[ E(\tilde{W}_1) = \frac{\lambda_1E(S_1^2) + \lambda_2E(S_2^2)}{2(1-p_1)}. \]

(2.5) \[ E(\tilde{W}_1^2) = \frac{\lambda_1E(S_1^2)}{3(1-p_1)} + \frac{\lambda_2E(S_2^2)[\lambda_1E(S_1^2) + \lambda_2E(S_2^2)]}{2(1-p_1)^2}. \]

The Laplace-Stieltjes transform of the non-priority waiting time distribution can be obtained by an appropriate modification of the results of L. Takács [21]. For the simple queue with a single type, Poisson arrivals (\( \lambda \)), and general service distribution, F. S. Takács established that the Laplace-Stieltjes transform \( \tilde{W}(s) \) of the steady state waiting time distribution was given by

(2.6) \[ \tilde{W}(s) = \frac{1-\lambda E(s)}{1-\lambda E(s) - \frac{1-S(s)}{s}}. \]
where $\tilde{S}$ is the Laplace-Stieltjes transform of $F_S$ and $E(S)$ its first moment.

The waiting time of a non-priority item is the sum of two waiting times, $W_2^*$ and $W_2^{**}$. $W_2^*$ is the time required to service all priority and non-priority items already in the queue at the arrival of the non-priority item, and $W_2^{**}$ is the time consumed in servicing all subsequent priority arrivals which precede the entrance into service of the non-priority item. As far as the waiting time of the non-priority item is concerned, the following queue discipline could be in effect at its arrival. Service all priority and non-priority items in the queue ahead of the non-priority item at its arrival. Any priority arrivals occurring during this time interval are refused service until the items initially in the queue have been serviced — even if this means servicing a non-priority item in preference to a priority item. After the initial group has been serviced, commence service on the by-passed priority items and continue service until the queue has been emptied of priority items. At this moment the non-priority item whose waiting time is in question may enter service. $W_2^{**}$ is defined to be the service time for the left-over priority items.

If $n$ priority items arrive during the $W_2^*$ units of time, the distribution of $W_2^{**}$ is the same as the distribution of $B_n$ where $B_n$ is the length of a busy period (see[21] and §3) for a one-dimensional queue with only priority items in which there are $n$
priority items initially. Thus,

\[ P\{W_2 \leq x\} = \int_0^x \left[ \sum_{n=0}^{\infty} \frac{e^{-\lambda_1 y} (\lambda_1 y)^n}{n!} P\{B_n \leq x - y\}\right] dP\{W_2^\# \leq y\} \]

The Laplace-Stieltjes transform equivalent of (2.7) is

\[ \tilde{W}_2(s) = \tilde{W}_2^\# (s + \lambda_1 (1 - \tilde{F}_1(s))) \]

where \( \tilde{F}_1(s) \) is the transform of the distribution \( F_{B_1} \).

The distribution of \( W_2^\# \) is obtainable from the result of L. Takács (2.6) with the identifications \( \lambda = \Lambda_2 \) and \( \tilde{S}(s) = \tilde{S}_2^\#(s) \) where

\[ \Lambda_2 = \lambda_1 + \lambda_2 \quad , \quad \tilde{S}_2^\#(s) = \tau_1 \tilde{S}_1(s) + \tau_2 \tilde{S}_2(s) \]

To a non-priority item arriving at the queue the distinction between previously arrived priority and non-priority items is immaterial. All are serviced ahead of the non-priority item and could just as well be viewed as having the average service time distribution given in (2.9).
Hence,

\[
\tilde{W}_2(s) = \frac{1 - \Lambda_2 E(S_2^*)}{1 - \Lambda_2 \frac{1 - \tilde{S}_2^*(s + \lambda_1 (1 - \tilde{B}_1(s)))}{s + \lambda_1 (1 - \tilde{B}_1(s))}}
\]

\[
= \frac{1 - \rho_1 - \rho_2}{1 + \sum_{i=1}^{2} \lambda_i [1 - \tilde{S}_1(s + \lambda_1 (1 - \tilde{B}_1(s)))]} \frac{1}{s + \lambda_1 (1 - \tilde{B}_1(s))}
\]

\[\tilde{B}_1(s)\] has been characterized by L. Takács [21] as the unique analytic solution to the functional equation

\[
f(s) = \tilde{S}_1(s + \lambda_1 (1 - f(s))) \quad \text{Re}\{s\} > 0
\]

which satisfies the restriction \(\lim_{s \to \infty} f(s) = 0\) for real \(s\).

The moments of \(W_2\) can be determined from (2.10) and (2.11). The first two are given by

\[
E(W_2) = \frac{\lambda_1 E(S_1^2) + \lambda_2 E(S_2^2)}{2(1 - \rho_1)(1 - \rho_1 - \rho_2)}
\]
\begin{equation}
E(W_2^2) = \frac{\lambda_1 E(S^3_1) + \lambda_2 E(S^3_2)}{3(1-\rho_1)(1-\rho_1-\rho_2)} + \frac{[\lambda_1 E(S^2_1) + \lambda_2 E(S^2_2)]^2}{2(1-\rho_1)^2(1-\rho_1-\rho_2)^2} + \frac{\lambda_1 E(S^2_1) [\lambda_1 E(S^2_1) + \lambda_2 E(S^2_2)]}{2(1-\rho_1)^3(1-\rho_1-\rho_2)}
\end{equation}

H. Kesten and J. Th. Runnenburg ([6], [7]) have obtained an alternative characterization of \(\tilde{W}_1\) and \(\tilde{W}_2\). The first two moments as computed by their method agree with (2.4)-(2.5) and (2.12)-(2.13). In addition, Kesten and Runnenburg have derived a characterization for the transform of the steady state waiting time distribution of any type \(j\) item for a priority queue with general \(K\).

The method employed to characterize \(\tilde{W}_2(s)\) above can be extended to characterize the waiting time transform of the lowest priority, type \(K\) item for arbitrary time \(t\) and in equilibrium. Let \(W_K(t)\) denote the waiting time for a lowest priority item, type \(K\), if it were to arrive at time \(t\). \(W_K(t)\) is the sum of two components, \(W^*_K(t)\) and \(W^{**}_K(t)\), which are defined analogously to \(W^*_2\) and \(W^{**}_2\). The same argument verifies that
\begin{equation}
(2.14) \quad P\{ \hat{W}_K(t) \leq x \} = \int_0^\infty \left[ \sum_{n_1, \ldots, n_{K-1}} e^{-\sum_{i=1}^{K-1} \lambda_i y_i} \frac{\lambda_1^{n_1}}{n_1!} \ldots \frac{\lambda_{K-1}^{n_{K-1}}}{n_{K-1}!} \cdot P\{ B_{K-1} \leq x-y \} \right] \, dP\{ \hat{W}_K^*(t) \leq y \}
\end{equation}

where $B_{K-1,n_1,\ldots,n_{K-1}}$ is the length of a busy period for a priority queue with just $K-1$ types $1, \ldots, K-1$ which commences with $n_i$ type $i$ items, $i=1, \ldots, K-1$, in line initially (see §3). In terms of Laplace-Stieltjes transforms (2.14) becomes

\begin{equation}
(2.15) \quad \hat{W}_K(s;t) = \hat{W}_K(s + \lambda_1 (1 - \hat{B}_{K-1,1}(s)) + \ldots + \lambda_{K-1} (1 - \hat{B}_{K-1,K-1}(s));t)
= \hat{W}_K(s + \hat{\Lambda}_{K-1}(1 - \hat{B}_{K-1}^*(s));t)
\end{equation}

where $\hat{B}_{K-1,i}(s)$ is the transform of the busy period distribution for the $K-1$ dimensional priority queue which commences with a single type $i$ item in line (see §3) and

\begin{equation}
\hat{B}_{K-1}^*(s) = \sum_{i=1}^{K-1} \tau_i \hat{B}_{K-1,i}(s), \quad \hat{\Lambda}_{K-1} = \lambda_1 + \ldots + \lambda_{K-1} \quad \text{and} \quad \tau_i = \frac{\lambda_i}{\hat{\Lambda}_{K-1}}, \ i=1, \ldots, K-1.
\end{equation}
\( \tilde{\beta}_{K-1,i}(s) \) and \( \tilde{\beta}^*_K(s) \) will be characterized in the next section.

\( \tilde{W}_K(s;t) \) can be obtained from the results of L. Takács for a simple queue. For a simple, single class, queue with Poisson arrivals \( \lambda \) and general service distribution \( F_S \) the transform \( \tilde{\tilde{W}}(s;t) \) of the waiting time distribution at time \( t \) is given by

\[
(2.16) \quad \tilde{\tilde{W}}(s;t) = e^{[s-\lambda(1-\tilde{S}(s))]} \int_{1-s}^{t} e^{-u[s-\lambda(1-\tilde{S}(s))]} F_W(0^+;u) du
\]

where \( F_W(0^+;u) \) is the probability that the queue is empty at time \( u \).

\( F_W(0^+;u) \) is determined by the relation

\[
(2.17) \quad \int_0^\infty e^{-su} F_W(0^+;u) du = \frac{1}{s+\lambda(1-\tilde{B}(s))},
\]

\( \tilde{B}(s) \) being the transform of the busy period distribution for the simple queue (see [21]). For the priority queue \( \tilde{\tilde{W}}_K(s;t) = \tilde{\tilde{W}}(s;t) \) with the identifications

\[
\lambda = \lambda_K = \lambda_1 + \ldots + \lambda_K
\]

\[
(2.18) \quad \tilde{S}(s) = \tilde{S}_K^*(s) = \tau_1 \tilde{S}_1(s) + \ldots + \tau_K \tilde{S}_K(s)
\]

\[
\tilde{B}(s) = \tilde{B}_K^*(s) = \tau_1 \tilde{B}_1(s) + \ldots + \tau_K \tilde{B}_K(s)
\]
In the limit as $t \to \infty$

\[
(2.19) \quad \tilde{W}_K(s) = \lim_{t \to \infty} \tilde{W}_K(s;t) = \tilde{W}_K(s + \wedge_{K-1} (1 - B_{K-1}(s)))
\]

where

\[
\tilde{W}_K^m(\alpha) = \frac{1 - \wedge_{K} E(S^*_K)}{1 - \wedge_{K} \frac{1 - \tilde{S}_K^m(\alpha)}{\alpha}}
\]

\[
(2.20) \quad \frac{1 - \rho_1 - \cdots - \rho_K}{\sum_{i=1}^{K} \lambda_i [1 - \tilde{S}_i(\alpha)]} = \frac{1 - \rho_1 - \cdots - \rho_K}{1 - \frac{\sum_{i=1}^{K} \lambda_i [1 - \tilde{S}_i(\alpha)]}{\alpha}}
\]

The moments of the steady state waiting time can be computed from (2.19)-(2.20) and the results of the next section. For example

\[
(2.21) \quad E(W_K) = \frac{\sum_{i=1}^{K} \lambda_i E(S^*_i)}{2(1 - \sum_{i=1}^{K-1} \rho_i)(1 - \sum_{i=1}^{K} \rho_i)}
\]
(2.22) \[ E(w^2_K) = \frac{\sum_{i=1}^{K} \lambda_i E(S^3_i)}{3(1-\sum_{i=1}^{K-1} \rho_i)^2(1-\sum_{i=1}^{K} \rho_i)} \]
\[ \left( \sum_{i=1}^{K} \lambda_i E(S^3_i) \right)^2 \]
\[ \frac{2(1-\sum_{i=1}^{K-1} \rho_i)^2(1-\sum_{i=1}^{K} \rho_i)^2}{2(1-\sum_{i=1}^{K-1} \rho_i)^3(1-\sum_{i=1}^{K} \rho_i)} \]

This same technique can be applied to the preemptive "resume" priority queue to characterize the waiting time distribution for any type item at general time \( t \) and in equilibrium. Let there be \( K \) priority classes in the queue and \( W_j(t) \) be the waiting time for an item in the \( j^{th} \) class if it were to arrive at time \( t \). The distribution of \( W_1(t) \) is the same as the waiting time distribution for the type one items in isolation since priority items preempt any lower class items in service. The waiting time \( W_j(t), j > 1, \) consists of two components, \( W^*_j(t) \) and \( W^*_j^*(t) \). \( W^*_j(t) \) is the time required to service all items of priority \( \leq j \) which are in the queue at time \( t \), and its Laplace-Stieltjes transform is given by (2.16) and (2.17) with
\[ \lambda = \lambda_j, \quad \tilde{S}(s) = \tilde{S}^*_j(s), \text{ and } \tilde{B}(s) = \tilde{B}^*_j(s) = \tau_1 \tilde{B}_{j1}(s) + \ldots + \tau_j \tilde{B}_{jj}(s). \]
\( \tilde{B}_{ji}(s) \) and \( \tilde{B}^*_j(s) \) for the preemptive resume discipline will be characterized in \( \S 3 \). The distribution of \( W^*_j(t) \) is unaffected by the presence of lower priority items because of the preemptive discipline,
and since an item "resumes" service after preemption, the priority
discipline among the items of types 1,...,j could just as well be
abandoned as far as the distribution of $W_j^*(t)$ is concerned. $W_j^{**}(t)$
is the time required to service all arrivals of priority $< j$ which
arrive after $t$ but before the type $j$ item can enter service, and
it is given by a convolution of busy periods $B_{j-1,i}$, $i = 1,...,j-1$,
where the degree of the convolution is determined by the number of
arrivals in the time interval $(0,W_j^*(t))$. Hence,

$$
P\{W_j(t) \leq x\} = \sum_{n_1,\ldots,n_{j-1}} e^{-\sum_{i=1}^{j-1} \frac{\lambda_i y^{n_i}}{n_i!}} \cdot \frac{\lambda_j y^{n_{j-1}}}{n_{j-1}!} \cdot \prod_{i=1}^{j-1} \left[ P\{B_{i-1,n_1,\ldots,n_i} \leq x-y\} \right] \cdot P\{W_j^*(t) \leq y\}
$$

and

$$
\tilde{W}_j(s;t) = \tilde{W}_j^*(s+\sum_{i=1}^{j-1}(1-\tilde{B}_i(s));t).
$$

For the stationary case

$$
\lim_{t \to \infty} \tilde{W}_j(s;t) = \tilde{W}_j^*(s+\sum_{i=1}^{j-1}(1-\tilde{B}_i(s)))
$$
where

\begin{equation}
\hat{W}_j(\alpha) = \frac{1 - \Lambda_j E(S_j^*)}{1 - \frac{1 - \hat{S}_j(\alpha)}{\alpha}}.
\end{equation}

The first two moments of \( W_j \) are given by (2.21) and (2.22) with \( K \) replaced by \( j \).

The quantity \( T_j \), the "time in service" of a type \( j \) priority item, is \( S_j \) only for \( j=1 \) under the preemptive resume discipline. However, the same type of reasoning as used earlier in this section will verify that

\begin{equation}
P\{T_j \leq x\} = \sum_{n_1, \ldots, n_{j-1}} e^{-\Lambda_j y} \frac{(\lambda_{j-1} y)^{n_{j-1}}}{n_{j-1}!} \cdots \frac{\lambda_1 y^{n_1}}{n_1!} 
\end{equation}

\[P\{B_{j-1:n_1 \ldots n_{j-1}} \leq x-y\}\] dF_{S_j}(y)

so

\begin{equation}
\hat{T}_j(s) = \hat{S}_j(s + \Lambda_{j-1} (1 - \hat{B}_{j-1}(s)))
\end{equation}
The first two moments of $T_j$ are

\begin{equation}
E(T_j) = \frac{E(S_j)}{1 - \sum_{i=1}^{j-1} \rho_i}
\end{equation}

\begin{equation}
E(T_j^2) = \frac{E(S_j^2)}{\left[1 - \sum_{i=1}^{j-1} \rho_i\right]^2} + \frac{E(S_j)\left[\lambda_1E(S_1^2) + \cdots + \lambda_{j-1}E(S_{j-1}^2)\right]}{\left[1 - \sum_{i=1}^{j-1} \rho_i\right]^3}
\end{equation}

The preemptive priority queue with indifferent server is a special case of the preemptive resume priority queue so the above results apply as well to the indifferent server queue. The waiting time questions for the preemptive repeat priority queue are for the most part still unsolved.
§3. Busy Period Distributions

A queue is said to be "busy" or "empty" depending upon whether or not there is an item in service. The length of a busy period is the length of time between the arrival of an item at the empty queue and the first subsequent moment at which the queue is again empty. The distribution of the length of a busy period is of fundamental importance in describing the probabilistic behavior of a priority queue and, as was seen in §2, is in fact needed in the determination of the waiting time distributions.

The technique which will be used to characterize the busy period is a modification of that introduced by L. Takács [21] to solve the busy period problem for the simple queue with a single priority class, Poisson arrivals (\( \lambda \)), and a general service distribution \( F_S \). Takács established that \( \tilde{S}(s) \), the Laplace-Stieltjes transform of the busy period distribution, satisfies the functional equation

\[
(3.1) \quad f(s) = \tilde{S}(s + \lambda (1 - f(s)))
\]

and is in fact the unique solution to (3.1) which satisfies in addition
(i) \( f(s) \) analytic for \( \text{Re}\{s\} > 0 \)

\[ (3.2) \]

(ii) \( \lim_{s \to \infty} f(s) = 0 \) on \( s \) real

The similarity of the results for the priority queue to those for the simple queue will be apparent.

Consider a priority queue with \( K \) priority classes and head-of-the-line discipline. As defined in \( \mathcal{S}2 \), let \( B_{Ki} \) be the length of a busy period which commences with the arrival of a type \( i \) item, \( i = 1, \ldots, K \). \( F_{B_{Ki}} \) will denote the distribution function of \( B_{Ki} \) and \( \tilde{F}_{Ki}(s) \) the corresponding Laplace-Stieltjes transform. Let \( B^*_K \) denote the average busy period in which the priority class of the initial arrival is not specified. \( F_{B^*_K} = \tau_1F_{B_{K1}} + \cdots + \tau_KF_{B_{KK}} \), and

\( \tilde{B}^*_K(s) = \tau_1\tilde{F}_{K1}(s) + \cdots + \tau_K\tilde{F}_{KK}(s) \).

The equilibrium condition \( 1-\rho_1 - \cdots - \rho_K > 0 \) will be assumed so that \( F_{B_{Ki}} \) is a bona fide distribution. With modification the discussion applies as well to the transient case. In particular, Theorem 3.1 can be extended to characterize the value \( F_{B^*_K}(+\infty) \).

\( \tilde{B}^*_K(s) \) is the easiest to characterize and will be treated first. Arrivals at the queue constitute a Poisson process with parameter \( \lambda_K \).
and given that an arrival has occurred the probability it belongs to priority class \( j \) is \( r_j \). At the end of the service period of the initial arrival there will be \( n_1, \ldots, n_K \) items of types \( 1, \ldots, K \), respectively, in the queue. The busy period will be prolonged by the amount \( B_{K:n_1, \ldots, n_K} \) which denotes a busy period commencing with \( n_1, \ldots, n_K \) items of types \( 1, \ldots, K \), respectively, in line initially. However, the distribution of \( B_{K:n_1, \ldots, n_K} \) is just a convolution of the distributions \( F_{B_{KL}}^{(n_1)}, \ldots, F_{B_{KK}}^{(n_K)} \) where \( F_{B_{Ki}}^{(n_i)} \) denotes an \( n_i \)-fold convolution of \( F_{B_{Ki}} \). Hence,

\[
P\{ B^*_K \leq x \} = \int_0^x \left[ \sum_{n=0}^{\infty} \frac{e^{-\lambda y} \prod_{l=1}^{n} \frac{(\lambda y)^{n_l}}{n_l!}}{\prod_{l=1}^{n} \sigma(n_l)} \right] \prod_{l=1}^{n} \tau_l^{n_l} \prod_{l=1}^{n} (\tau_l)^{n_l}
\]

(3.3)

\[
P\{ B_{K:n_1, \ldots, n_K} \leq x-y \} \int dF_{B^*_K}(y)
\]

so
\[ \tilde{B}_K^*(s) = \int_0^\infty \left[ \sum_{n=0}^{\infty} e^{-(s + \Lambda K)y} \frac{(\Lambda_K y)^n}{n!} \sum_{n_1, \ldots, n_K} \frac{n!}{n_1! \cdots n_K!} (\tau_1 \tilde{B}_K^*(s))^{n_1} \cdots \right] dF^*_K(y) \]

(3.4)

\[ = \int_0^\infty \left[ \sum_{n=0}^{\infty} e^{-(s + \Lambda K)y} \frac{(\Lambda_K y)^n}{n!} (\tilde{B}_K^*(s))^n \right] dF^*_K(y) \]

\[ = \tilde{S}_K^*(s + \Lambda_K(1 - \tilde{B}_K^*(s))) \]

where \( F_{K}^* = \tau_1 F_{S_1} + \cdots + \tau_K F_{S_K} \) and \( \tilde{S}_K^* \) is its transform.

Moreover, \( \tilde{B}_K^*(s) \) is the unique solution to (3.4) which satisfies the regularity conditions (3.2). The proof of this assertion can be obtained from the proof for the simple queue [21] with the appropriate identifications, but a simpler, alternative proof is presented below. This simpler proof also applies to the simple queue.

**Theorem 3.1:** \( \tilde{B}_K^*(s) \) is the unique solution to the functional equation

(3.5) \[ f(s) = \tilde{S}_K^*(s + \Lambda_K(1 - f(s))) \]

which satisfies, in addition, (i) and (ii) of (3.2).
Proof: It is sufficient to show that (3.5) and (ii) determine $f(s)$ uniquely for real $s > 0$ since by (i) this determines $f(s)$ in the whole half-plane. Suppose there exist two functions $f_1(s)$ and $f_2(s)$ which satisfy (3.5), (i), (ii) but $f_1(s) \neq f_2(s)$ for real $s > 0$.

Let $s_0 > 0$ be a point for which $f_1(s_0) > f_2(s_0) \geq 0$. Since

$$\lim_{s \to \infty} f_1(s) = 0$$

and $f_1$ is continuous, there must exist an $s_1 > s_0$ such that $f_1(s_1) = f_2(s_0) = c$. But this implies that there exist two different values, namely $s_0$ and $s_1$, which satisfy

$$(3.6) \quad c = s_K^*(s + \Lambda_K(1 - c))$$

which is impossible since the right hand side of (3.6) is a strictly decreasing function of $s$. q.e.d.

The moments of $B^*_K$ can be easily computed from (3.4). The first two are given by

$$E(B^*_K) = \frac{E(S^*_K)}{1 - \Lambda_K E(S^*_K)} = \frac{\tau_1 E(S_1) + \ldots + \tau_K E(S_K)}{1 - \rho_1 - \ldots - \rho_K}$$

$$(3.7) \quad E(B^*_K^2) = \frac{E(S^*_K^2)}{(1 - \Lambda_K E(S^*_K))^3} = \frac{\tau_1 E(S_1^2) + \ldots + \tau_K E(S_K^2)}{(1 - \rho_1 - \ldots - \rho_K)^3}$$
The characterization of $\tilde{B}_{Ki}$ will be constructed recursively. Assume that $F_{B_{K-1,i}}$, $i = 1, \ldots, K-1$, and the corresponding $F_{B_{K-1,i}}$, have been determined. $\tilde{B}_{KI}$, $i = 1, \ldots, K$, can be characterized as the solution of a functional equation involving $\tilde{B}_{K-1,i}$, $i = 1, \ldots, K-1$, in the following manner. As far as the distribution of the busy period is concerned, the priority discipline can be disregarded and the items from different priority classes can be served in any order. If the busy period commences with the arrival of a type 1 item, the queue discipline could just as well stipulate that after this item has been serviced no other type 1 items will be serviced until the queue no longer contains any other type items. Let $H_{Ki}$ be the distribution of the time required to empty the queue of items other than type 1 (including the service time of the initial 1 item). If in the time required to clear the queue of non-type 1 items n type 1 items arrive, the busy period is prolonged by an n-fold convolution of busy periods $B_{Ki}$. Hence, $F_{B_{Ki}}$ must satisfy

\begin{equation}
F_{B_{Ki}}(x) = \int_0^x \left[ \sum_{n=0}^{\infty} e^{-\lambda_1 y} \frac{(\lambda_1 y)^n}{n!} F_{B_{Ki}}(x-y) \right] dH_{Ki}(y)
\end{equation}

or

\begin{equation}
\tilde{B}_{Ki}(s) = H_{Ki}(s + \lambda_1(1 - \tilde{B}_{Ki}(s)))
\end{equation}
But $H_{KI}$ is given by

$$H_{KI}(y) = \int \left[ \frac{(\lambda_1 y)^{n_1}}{n_1!} \cdots \frac{(\lambda_{i-1} y)^{n_{i-1}}}{n_{i-1}!} \frac{(\lambda_{i+1} y)^{n_{i+1}}}{n_{i+1}!} \cdots \frac{(\lambda_K y)^{n_K}}{n_K!} \right] \cdot \frac{F_{i_{K-1,1}}^{(n_1)}}{n_1!} \cdots \frac{F_{i_{K-1,i-1}}^{(n_{i-1})}}{n_{i-1}!} \frac{F_{i_{K-1,i+1}}^{(n_{i+1})}}{n_{i+1}!} \cdots \frac{F_{i_{K-1,K}}^{(n_K)}(y-z)}{n_K!} \cdot dF_{S_1}(z)$$

(3.10)

where $i_{K-1,j}$ denotes a busy period for a queue with $K-1$ priority classes, the $i^{th}$ class of the original $K$ classes being absent, and with a type $j$ item in line initially. (3.10) implies

$$\tilde{H}_{KI}(s) = \tilde{S}_i(s + \sum_{j=1}^{i-1} \frac{\lambda_j(l-1)}{i^{j-1}}} \tilde{S}_{i_{K-1,j}}(s) + \sum_{j=i+1}^{K} \lambda_j(l-1) \tilde{S}_{i_{K-1,j}}(s))$$

$$= \tilde{S}_i(s + \tilde{\Lambda}_{K-1}^{*}(l-1) \tilde{S}_{i_{K-1,j}}(s))$$

(3.11)

where $\tilde{\Lambda}_{K-1}^{*} = \tilde{\Lambda}_K - \lambda_i$, $\tilde{\Lambda}_{K-1} = \sum_{j \neq i} \lambda_j \tilde{S}_{i_{K-1,j}} / \tilde{\Lambda}_K - \lambda_i$. (3.9) and

(3.11) together yield
(3.12) \[
\hat{E}_{Ki}(s) = \hat{S}_1(s + \lambda_i (1 - \hat{E}_{Ki}(s))) + \bigwedge_{K-1}(1 - \hat{E}_{K-1}(s + \lambda_i (1 - \hat{E}_{Ki}(s))))
\]

\(\hat{E}_{Ki}\) is in fact the unique solution to the functional equation (3.12) subject to the regularity conditions (3.2). The proof of uniqueness can be obtained either from the proof of L. Takács for the simple queue or Theorem 3.1 with the appropriate identifications.

The moments of \(E_{Ki}\) are derivable from (3.12). In particular,

\[
E(B_{Ki}) = \frac{E(S_1)}{1 - \rho_1 - \cdots - \rho_K}
\]

(3.13)

\[
E(B_{Ki}^2) = \frac{E(S_1^2) \left[ 1 - \sum_{j \neq 1} \rho_j \right] + E(S_1) \left[ \sum_{j \neq 1} \lambda_j E(S_1^2) \right]}{\left[ 1 - \sum_{j \neq 1} \rho_j \right]^3}
\]

The distribution of the busy period for a preemptive resume priority queue is identical to the busy period distribution for the same queue with head-of-the-line discipline. The order of service is immaterial to the busy period as long as preemption does not increase the time spent in the service mechanism which is the case for a resume discipline. Hence, all the previous results for head-of-the-line discipline apply as well to the preemptive resume queue. The indifferent server queue is included as a special case of the resume discipline. A similar identification cannot be made for a preemptive repeat priority
queue since the length of time a lower priority item spends in the service mechanism is no longer equal to the service time in isolation. To date no characterization has been obtained for the busy period of a preemptive repeat priority queue.
§4. Distribution of Number of Items Serviced During A Busy Period.

The waiting time distributions are of interest both to the items in the queue and to the organization or management in charge of the queue whereas the busy period distributions are a primary concern only of the management. Another variable which is of interest to the management is the number of items which will be serviced during a busy period. This variable in effect measures the output of the service mechanism.

L. Takács in [21] characterized the distribution of the number of items serviced during a busy period for the simple queue, and this method can be generalized in a fashion analogous to §3. As before, the equilibrium condition \( 1 - \rho_1 \ldots \rho_K > 0 \) will be assumed so that all the distributions which will be discussed have total variation one. The reader can easily modify the discussion to cover the transient case.

Consider a priority queue with \( K \) priority classes and head-of-the-line discipline. Let \( f^{*}_{K_1}(j) \) be the probability that a total of \( j \) items, irrespective of class, are serviced during a busy period commencing with a single type 1 item in the queue, and let \( f^{*}_{K}(j) = \sum_{l=1}^{K} f^{*}_{K_1}(j) \) be the probability of servicing a total of \( j \) items where the class of the initial item is unspecified. \( \hat{f}^{*}_{K_1}(s) \) and \( \hat{f}^{*}_{K}(s) \) will denote the generating functions of \( \{ f^{*}_{K_1}(j) \} \) and \( \{ f^{*}_{K}(j) \} \), respectively. For a specific class let \( f^{*}_{K_1}(j) \) be the
probability of servicing \( j \) type \( i \) items in a service period which commences with a type \( i \) item in line initially. \( f_{Ki}(j) \) will denote the generating function of \( \{ f_{Ki}(j) \} \).

The determination of \( f^*_K(s) \) will be treated first. Let \( p_{K:n_1 \ldots n_K} \) be the probability that during the service period for the initial, unspecified item \( n_j \) type \( j \) items, \( j=1,\ldots,K \), arrive. Since the initial item is type \( j \) with probability \( \tau_j \),

\[
(4.1) \quad p_{K:n_1 \ldots n_K} = \int_0^\infty \left[ \frac{\lambda_k^t}{n_1! \ldots n_K!} (\tau_1)^{n_1} \ldots (\tau_K)^{n_K} \right] dF_{K}(t)
\]

and

\[
(4.2) \quad p_{K}(s_1, \ldots, s_K) = \sum_{n_1, \ldots, n_K} p_{K:n_1 \ldots n_K} s_1^{n_1} \ldots s_K^{n_K}
\]

where

\[
= \frac{\lambda_k^t}{n_1! \ldots n_K!} (\tau_1)^{n_1} \ldots (\tau_K)^{n_K} \int_0^\infty \left[ \frac{\lambda_k^t}{n_1! \ldots n_K!} (\tau_1)^{n_1} \ldots (\tau_K)^{n_K} \right] dF_{K}(t)
\]

By an argument analogous to that employed in \( \xi_3 \), the \( f^*_K(s) \) and \( f^*_{K1}(j) \) can be shown to satisfy the system of relations.
\[ f^*_K(l) = p_{K:0\ldots0} \]

\[
\begin{align*}
\sum_{n_1, \ldots, n_K} p_{K:n_1 \ldots n_K} \sum_{j_{11}, \ldots, j_{Kn_K}} f^*_K(j_{11}) \ldots f^*_K(j_{1n_1}) \ldots f^*_K(j_{K1}) \\
0 < \sum n_i \leq j-1 \\
n_1 + \ldots + n_K = j-1 \\
\ldots f^*_K(j_{Kn_K}), \ j \geq 2
\end{align*}
\]

This yields for the generating function

\[
\begin{align*}
f^*_K(s) &= \sum_{j=2}^{\infty} s^j \sum_{n_1, \ldots, n_K} p_{K:n_1 \ldots n_K} \sum_{j_{11}, \ldots, j_{Kn_K}} f^*_K(j_{11}) \ldots f^*_K(j_{K1}) \\
&0 < \sum n_i \leq j-1 \\
n_1 =j-1 \\
\sum_{j_{11}+\ldots+n_K = j-1} f^*_K(j_{Kn_K}) \\
+ s p_{K:0\ldots0} \\
&= s \sum_{n_1, \ldots, n_K} p_{K:n_1 \ldots n_K} \sum_{j=\sum n_i+1}^{\infty} s^{j-1} \sum_{j_{11}, \ldots, j_{Kn_K}} f^*_K(j_{11}) \ldots f^*_K(j_{K1}) \\
&0 < \sum n_i \\
\sum_{j_{11}+\ldots+n_K = j-1} f^*_K(j_{Kn_K}) \\
+ s p_{K:0\ldots0} \\
&(4.4)
\end{align*}
\]

\[
\sum_{n_1, \ldots, n_K} p_{K:n_1 \ldots n_K} (\tilde{f}_K(s))^{n_1} \ldots (\tilde{f}_K(s))^{n_K} + s p_{K:0\ldots0} \\
\sum_{n_1 > 0} \\
= s P(\tilde{f}_K(s), \ldots, \tilde{f}_K(s)) = s \tilde{f}_K(1 - \tilde{f}_K(s))
\]
Moreover, \( \tilde{f}_K^*(s) \) is the unique solution to the functional equation

\[
(4.5) \quad f(s) = s \tilde{S}_K^*(\Lambda_K^*(1 - f(s))) \quad |s| \leq 1
\]

subject to the regularity conditions

\[
(4.6) \quad (i) \quad f(s) \text{ analytic for } |s| \leq 1
\]

\[
(ii) \quad f(0) = 0 .
\]

The proof of uniqueness is omitted since it follows either from the simple queue proof [21] or from an argument similar to Theorem 3.1.

The moments of \( N_K^* \), the total number of items serviced during a busy period, are obtainable from (4.4). In particular,

\[
(4.7) \quad E(N_K^*) = \frac{1}{1 - \Lambda_K^*E(S_K^*)} = \frac{1}{1 - \rho_1 - \cdots - \rho_K}
\]

\[
E(N_K^{*2}) = \frac{\Lambda_K^*E(S_K^{*2})}{(1 - \Lambda_K^*E(S_K^*))^3} + \frac{2\Lambda_K^*E(S_K^*)}{(1 - \Lambda_K^*E(S_K^*))^2} + \frac{1}{1 - \Lambda_K^*E(S_K^*)}
\]

The property that the number of type 1 items serviced during a busy period is independent of the priority discipline makes it feasible to obtain a functional relation for \( \tilde{f}_{K1}^*(s) \). Let \( p_{K1:n} \) be the
probability that \( n \) type 1 items arrive during the time it takes to
service the initial type 1 item and then clear the queue of other
type items without admitting any type 1 items into service.

\[
(4.8) \quad P_{K_i:n} = \int_0^\infty e^{-\lambda_1 t} \frac{(\lambda_1 t)^n}{n!} dH_{K_i}(t)
\]

and

\[
(4.9) \quad P_{K_i}(s) = \sum_{n=0}^{\infty} P_{K_i:n} s^n = \tilde{H}_{K_i}(\lambda_1 (1-s))
\]

where \( H_{K_i}(t) \) and \( \tilde{H}_{K_i}(s) \) were defined in (3.10) and (3.11).

Since the discipline of servicing the initial type 1 item and then
clearing the queue of the other class items can be assumed to be in
effect, the \( f_{K_i}(j) \) must satisfy the relations

\[
(4.10) \quad f_{K_i}(1) = P_{K_i:0}
\]

\[
(4.10) \quad f_{K_i}(j) = \sum_{n=1}^{j-1} P_{K_i:n} \sum_{j_1 + \ldots + j_n = j-1} f_{K_i}(j_1) \ldots f_{K_i}(j_n), \quad j \geq 2.
\]

For the generating function this yields
\[ \tilde{f}_{Ki}(s) = \sum_{j=2}^{\infty} s^j \sum_{n=1}^{j-1} p_{Ki:n} \sum_{j_1 + \ldots + j_n = j-1} f_{Ki}(j_1) \ldots f_{Ki}(j_n) + sp_{Ki:0} \]
\[
= s \sum_{n=1}^{\infty} p_{Ki:n} (\tilde{f}_{Ki}(s))^n + sp_{Ki:0} 
= s \tilde{H}_{Ki}(\lambda_i^{l} (1-\tilde{f}_{Ki}(s))).
\]

The proof that \( \tilde{f}_{Ki}(s) \) is the unique solution to (4.11) subject to
the regularity conditions (4.6) is omitted.

The first two moments of \( N_{Ki} \), the number of type \( i \) items
serviced during a busy period commencing with a type \( i \) item, are
determinable from (4.11).

\[
E(N_{Ki}) = \frac{1}{1 - \sum_{j \neq i} \rho_j} \left[ 1 - \sum_{j \neq i} \rho_j \right]^{\rho_i}
\]
\[
E(N_{Ki}^2) = \frac{\lambda_i^2 E(S_i^2) \left[ 1 - \sum_{j \neq i} \rho_j \right] + \lambda_i^2 E(S_i) \left[ \sum_{j \neq i} \lambda_j E(S_j^2) \right]}{\left[ 1 - \sum_{j \neq i} \rho_j \right]^3}
\]
\[
+ \frac{2\rho_i \left[ 1 - \sum_{j \neq i} \rho_j \right]}{\left[ 1 - \sum_{j \neq i} \rho_j \right]^2} + \frac{1 - \sum_{j \neq i} \rho_j}{1 - \sum_{j \neq i} \rho_j}
\]

The distributions of the total number of items and the number of
type \( i \) items serviced when the initial arrival is of type \( j \) can
be determined by forming the appropriate convolutions of service periods
and busy periods with the distributions already determined in this
section.

As in the case of the busy period distributions the above results
for head-of-the-line discipline apply equally as well to the preemptive
resume priority queue. This also includes as a special case the
priority queue with indifferent server. The corresponding distributions
for the preemptive repeat priority queue still remain to be determined.
§5. Acknowledgements

I would like to thank Professor Samuel Karlin for suggesting this research topic and for generously contributing his help and guidance. My thanks also to Mr. J. H. Kullback for checking part of the algebraic gymnastics.
§6. Bibliography


<table>
<thead>
<tr>
<th>Address</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armed Services Technical Information Agency</td>
<td>5</td>
</tr>
<tr>
<td>Arlington Hall Station</td>
<td></td>
</tr>
<tr>
<td>Arlington 12, Virginia</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td></td>
</tr>
<tr>
<td>Office of Naval Research Branch Office</td>
<td></td>
</tr>
<tr>
<td>The John Crerar Library Building</td>
<td>1</td>
</tr>
<tr>
<td>86 East Randolph Street</td>
<td></td>
</tr>
<tr>
<td>Chicago 1, Illinois</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td></td>
</tr>
<tr>
<td>Office of Naval Research Branch Office</td>
<td></td>
</tr>
<tr>
<td>346 Broadway</td>
<td>2</td>
</tr>
<tr>
<td>New York 13, New York</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td></td>
</tr>
<tr>
<td>Office of Naval Research Branch Office</td>
<td></td>
</tr>
<tr>
<td>Navy No. 100</td>
<td></td>
</tr>
<tr>
<td>c/o Fleet Post Office</td>
<td></td>
</tr>
<tr>
<td>New York, New York</td>
<td></td>
</tr>
<tr>
<td>Commanding Officer</td>
<td></td>
</tr>
<tr>
<td>Office of Naval Research Branch Office</td>
<td></td>
</tr>
<tr>
<td>1000 Geary Street</td>
<td></td>
</tr>
<tr>
<td>San Francisco 9, California</td>
<td>1</td>
</tr>
<tr>
<td>Commanding Officer</td>
<td></td>
</tr>
<tr>
<td>Office of Naval Research Branch Office</td>
<td></td>
</tr>
<tr>
<td>1030 East Green Street</td>
<td></td>
</tr>
<tr>
<td>Pasadena 1, California</td>
<td>1</td>
</tr>
<tr>
<td>Logistics Branch</td>
<td></td>
</tr>
<tr>
<td>Room 2718, Code 436</td>
<td></td>
</tr>
<tr>
<td>Office of Naval Research Branch Office</td>
<td></td>
</tr>
<tr>
<td>Washington 25, D. C.</td>
<td>2</td>
</tr>
<tr>
<td>Commanding Officer</td>
<td></td>
</tr>
<tr>
<td>U.S. Naval Ammunition Depot Earle</td>
<td></td>
</tr>
<tr>
<td>Red Bank, New Jersey</td>
<td>1</td>
</tr>
<tr>
<td>Industrial College of the Armed Forces</td>
<td></td>
</tr>
<tr>
<td>Fort Lesley J. McNair</td>
<td></td>
</tr>
<tr>
<td>Washington 25, D. C.</td>
<td></td>
</tr>
<tr>
<td>Attn: Mr. L. L. Henkel</td>
<td>1</td>
</tr>
<tr>
<td>Director</td>
<td></td>
</tr>
<tr>
<td>National Science Foundation</td>
<td></td>
</tr>
<tr>
<td>Washington 25, D. C.</td>
<td>1</td>
</tr>
<tr>
<td>Director</td>
<td></td>
</tr>
<tr>
<td>National Security Agency</td>
<td></td>
</tr>
<tr>
<td>Attn: REMP-1</td>
<td></td>
</tr>
<tr>
<td>Fort George G. Meade, Md.</td>
<td>1</td>
</tr>
<tr>
<td>Director</td>
<td></td>
</tr>
<tr>
<td>Naval Research Laboratory</td>
<td></td>
</tr>
<tr>
<td>Washington 25, D. C.</td>
<td></td>
</tr>
<tr>
<td>Attn: Technical Information Officer</td>
<td>6</td>
</tr>
<tr>
<td>Naval War College</td>
<td></td>
</tr>
<tr>
<td>Logistics Department, Luce Hall</td>
<td></td>
</tr>
<tr>
<td>Newport, Rhode Island</td>
<td>1</td>
</tr>
<tr>
<td>Office of Technical Services</td>
<td></td>
</tr>
<tr>
<td>Department of Commerce</td>
<td></td>
</tr>
<tr>
<td>Washington 25, D. C.</td>
<td>1</td>
</tr>
<tr>
<td>Dr. Edward Barankin</td>
<td></td>
</tr>
<tr>
<td>University of California</td>
<td></td>
</tr>
<tr>
<td>Berkeley 4, California</td>
<td>1</td>
</tr>
<tr>
<td>Dr. Richard Bellman</td>
<td></td>
</tr>
<tr>
<td>The RAND Corporation</td>
<td></td>
</tr>
<tr>
<td>1700 Main Street</td>
<td></td>
</tr>
<tr>
<td>Santa Monica, California</td>
<td>1</td>
</tr>
<tr>
<td>Dean L. M. K. Boelter</td>
<td></td>
</tr>
<tr>
<td>Department of Engineering</td>
<td></td>
</tr>
<tr>
<td>University of California</td>
<td></td>
</tr>
<tr>
<td>Los Angeles 24, California</td>
<td>1</td>
</tr>
<tr>
<td>Professor S. S. Cairns, Head</td>
<td></td>
</tr>
<tr>
<td>Department of Mathematics</td>
<td></td>
</tr>
<tr>
<td>University of Illinois</td>
<td></td>
</tr>
<tr>
<td>Urbana, Illinois</td>
<td>1</td>
</tr>
<tr>
<td>Name</td>
<td>Address</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Dr. E. W. Cannon</td>
<td>Applied Mathematics Division National Bureau of Standards</td>
</tr>
<tr>
<td>Dr. A. Charnes</td>
<td>The Technological Institute Northwestern University</td>
</tr>
<tr>
<td>Dr. Randolph Church, Chairman</td>
<td>Department of Mathematics and Mechanics U.S. Naval Postgraduate School</td>
</tr>
<tr>
<td>Professor William W. Cooper</td>
<td>Graduate School of Industrial Administration Carnegie Institute of Technology</td>
</tr>
<tr>
<td>Dr. George B. Dantzig</td>
<td>The RAND Corporation 1700 Main Street Santa Monica, California</td>
</tr>
<tr>
<td>Document Center</td>
<td>Stanford Research Institute Menlo Park, California</td>
</tr>
<tr>
<td>Rear Admiral H. E. Eccles, USN (RET)</td>
<td>101 Washington Street Newport, Rhode Island</td>
</tr>
<tr>
<td>Dr. Eberhard Fels</td>
<td>3420 Lincoln Avenue Shadyside, Ohio</td>
</tr>
<tr>
<td>Professor David Gale</td>
<td>Department of Mathematics Brown University Providence, Rhode Island</td>
</tr>
<tr>
<td>Mr. Murray A. Geisler</td>
<td>The RAND Corporation 1700 Main Street Santa Monica, California</td>
</tr>
<tr>
<td>Mrs. Dorothy M. Gilford</td>
<td>Office of Naval Research Code 433 Room 2709, T-3 Building</td>
</tr>
<tr>
<td>Dr. Theodore E. Harris</td>
<td>The RAND Corporation 1700 Main Street Santa Monica, California</td>
</tr>
<tr>
<td>Dr. J. Heller</td>
<td>Navy Management Office Washington, D.C.</td>
</tr>
<tr>
<td>Dr. C. C. Holt</td>
<td>Graduate School of Industrial Engineering Carnegie Institute of Technology</td>
</tr>
<tr>
<td>Dr. Walter Jacobs</td>
<td>Hqds. USAF, DSC/Comptroller AFADA-3D, The Pentagon Washington, D.C.</td>
</tr>
<tr>
<td>Professor J. R. Jackson</td>
<td>Management Sciences Research Project University of California Los Angeles, California</td>
</tr>
<tr>
<td>Capt. W. H. Keen</td>
<td>Aircraft Design Division Bureau of Aeronautics Navy Department Washington, D.C.</td>
</tr>
<tr>
<td>Professor Harold Kuhn</td>
<td>Department of Mathematics Bryn Mawr College Bryn Mawr 6, Pennsylvania</td>
</tr>
</tbody>
</table>
Professor S. B. Littauer  
Department of Industrial Engineering  
Columbia University  
409 Engineering Building  
New York 27, New York

Mr. J. R. Simpson  
Bureau of Supplies and Accounts  
Code SS  
Arlington Annex  
U. S. Department of the Navy  
Washington 25, D. C.

Dr. William Madow  
Stanford Research Institute  
Menlo Park, California

Professor R. M. Thrall  
Department of Mathematics  
University of Michigan  
Ann Arbor, Michigan

Dr. W. H. Marlow  
The George Washington University  
Logistics Research Project  
707 22nd Street  
Washington 7, D. C.

Professor L. M. Tichvinsky  
Department of Engineering  
University of California  
Berkeley 4, California

Professor Jacob Marshack  
Cowles Commission  
Yale University  
New Haven, Connecticut

Professor James Tobin  
Cowles Foundation for Research in Economics  
Box 2125, Yale Station  
New Haven, Connecticut

Mr. Joseph Mehr  
Head, Operations Research Desk  
U. S. N. Training Device Center  
Port Washington, Long Island, N. Y.

Dr. C. B. Tompkins, Director  
Numerical Analysis Research  
University of California  
405 Hilgard Avenue  
Los Angeles 24, California

Professor Oskar Morgenstern  
Economics Research Project  
Princeton University  
9-11 Lower Pyne Building  
92 A Nassau Street  
Princeton, New Jersey

Professor A. W. Tucker  
Department of Mathematics  
Fine Hall, Box 708  
Princeton University  
Princeton, New Jersey

Professor R. R. O'Neill  
Department of Engineering  
University of California  
Los Angeles, California

Professor Jacob Wolfowitz  
Department of Mathematics  
Cornell University  
Ithaca, New York

Dr. I. Richard Savage  
School of Business, Vincent Hall  
University of Minnesota  
Minneapolis, Minnesota

Dr. Max A. Woodbury  
Department of Mathematics  
New York University  
University Heights  
New York 53, New York

Professor H. N. Shapiro  
New York University  
Institute of Mathematical Sciences  
New York, New York

Professor H. A. Simon  
Head, Department of Industrial Administration  
Carnegie Institute of Technology  
Schenley Park  
Washington 25, D. C.

Additional copies for project leader and assistants and reserve for future requirements 50